ПОЛНЫЙ КУРС ЛЕКЦИЙ
по матем. анализу, 2 семестр, 2020 г.

Лекция 1 (07.02.20)
Интегральное исчисление. Неопределенный интеграл

Первообразная

Напомним определение промежутка.

Промежутком называется любое множество на $\mathbb{R}$, содержащее вместе с каждой парой точек и все точки, лежащие между ними.

В прошлом семестре было установлено, что промежутками являются числовая прямая $\mathbb{R}$, лучи $(a, +\infty)$, $[a, +\infty)$, $(-\infty, b)$, $(-\infty, b]$, отрезки $[a, b]$, интервалы (a, b), полуотрезки $[a, b)$ и $(a, b]$ с $a, b \in \mathbb{R}$, пустое множество $\varnothing$; приведенные множества исчерпывают совокупность промежутков.

Определение 1. Если действительнозначная функция f определена на промежутке $I \subset \mathbb{R}$, то ее **точной первообразной** (**примитивной**) на I является такая действительнозначная функция F, которая всюду на I имеет конечную производную (по I, т.е. правую в левом конце I, если он входит в I, и левую в правом конце I, если он входит в I) и $F' = f$ всюду на I.

Замечание. В приведенном определении можно вместо действительнозначных функций рассматривать комплекснозначные.

Определение 2. Если действительнозначная функция f определена на промежутке $I \subset \mathbb{R}$ за исключением конечного множества точек, то ее **обобщенной первообразной** (**обобщенной примитивной**) на I является такая действительнозначная функция F, которая непрерывна на I и дифференцируема всюду на I, за исключением конечного множества точек, и $F' = f$ всюду на I, кроме конечного множества точек.

Очевидно, всякая точная первообразная f на I является и обобщенной первообразной f на I.

Теорема 1 (описание класса первообразных). *Если F — обобщенная (точная) первообразная функции f на промежутке I, то $F+C$, где C — постоянная, также обобщенная (точная) первообразная f на I; если F_1 и F_2 — обобщенные первообразные функции f на промежутке I, то их разность $F_1 - F_2$ постоянна на I (то есть $F_1 - F_2 = C$ — постоянной на I).*

▼ Так как $(F + C)' = F' + C' = F' = f$ на I всюду, кроме конечного числа точек (всюду, если F — точная первообразная), то $F + C$ — обобщенная (точная) первообразная f на I. Если F_1 и F_2 — обобщенные первообразные f на I, то $F_1 - F_2$ непрерывна на I и $F_1 - F_2 = F_1' - F_2' = 0$ всюду на I, кроме конечного множества точек (в которых $F_1 - F_2$ непрерывна). Точки этого множества разбивают I на конечное число промежутков, на каждом из которых по теореме Лагранжа разность $F_1 - F_2$ постоянна. А так как в точках указанного множества разность $F_1 - F_2$ непрерывна, то $F_1 - F_2$ постоянна на I. ▲

Неопределенный интеграл и его свойства

Определение 3. Произвольная обобщенная первообразная функции f на промежутке I называется **неопределенным интегралом** f на I и обозначается $\int f\,dx$, где $\int$ — интеграл (знак интеграла), $f\,dx$ — подынтегральное выражение, f — подынтегральная функция, dx — дифференциал.

К сожалению, в этом стандартном обозначении не фигурирует промежуток I, что иногда приводит к недоразумениям.

Часто под неопределенным интегралом понимают совокупность всех первообразных функции, но мы будем пользоваться приведенным определением.

Определение 3. Интегрирование функции f на промежутке I — операция нахождения неопределенного интеграла f на I (обратная к операции дифференцирования).

С другим употреблением этого термина встретимся позже в разделе "Определенные интегралы".

Свойства неопределенного интеграла.

1. Для любой интегрируемой на промежутке I функции f

$$\left(\int f\,dx\right)' = f \quad \textit{или} \quad d\left(\int f\,dx\right) = f\,dx$$

всюду на I, за исключением конечного множества точек (а для неопределенного интеграла, являющегося точной первообразной, всюду);

2. если функция F непрерывна на промежутке I и дифференцируема на I всюду, кроме конечного множества точек, то

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

где C — некоторая постоянная;

3. если функция f интегрируема на промежутке I, то для любой константы $k \in \mathbb{R}$ функция kf интегрируема на I и

$$\int kf\,dx = k\int f\,dx + C,$$

где C — некоторая постоянная; при $k \neq 0$ оба интеграла одновременно существуют или несуществуют;

4. если функции f и g интегрируемы на промежутке I, то и $f \pm g$ интегрируема на I и

$$\int f \pm g\,dx = \int f\,dx \pm \int g\,dx + C,$$

где C — некоторая постоянная; если существуют любые два из трех написанных интегралов, то существует и третий и имеет место написанное равенство.

5. Формула интегрирования по частям.

Если функции u и v непрерывны на промежутке I и дифференцируемы на нем всюду, кроме конечного множества точек, то uv' и vu' одновременно интегрируемы или неинтегрируемы на I и в случае интегрируеимости

$$\int uv'dx = u\cdot v - \int vu'dx + C \quad \textit{или} \quad \int u\,dv = u\cdot v - \int v\,du + C,$$

где C — некоторая постоянная.

6. Формула замены переменной.

Если функция $f(x)$ интегрируема на промежутке I и $F(x)$ — ее неопределенный интеграл на I, а функция $\varphi(t)$ непрерывна на промежутке J, дифференцируема на нем всюду, кроме конечного множества точек, и лишь в конечном множестве точек J принимает значения, в которых $F(x)$ не дифференцируема, $\varphi(J) \subset I$, то функция $f(\varphi(t)) \cdot \varphi'(t)$ интегрируема на J и

$$\int f(\varphi(t)) \cdot \varphi'(t)\, dt \equiv \int f(\varphi(t))\, d\varphi(t) = F(\varphi(t)) + C,$$

где C — некоторая постоянная.

7. Формула интегрирования обратной функции.

Если функция f всюду на промежутке I имеет конечную не равную нулю производную и F — неопределенный интеграл f на I, то обратная к функции f на I функция f^{-1} интегрируема на промежутке $\varphi(I)$ и

$$\int f^{-1}(y)\, dy = y f^{-1}(y) - F\left(f^{-1}(y)\right) + C,$$

где C — некоторая постоянная.

▼ Проверка всех перечисленных свойств и равенств производится в соответствии с определением дифференцирования на указанных промежутках. ▲

Лекция 2 (11.02.20)
Таблица основных неопределенных интегралов
Интегрирование в элементарных функциях

Таблица основных неопределенных интегралов

$$1. \quad \boxed{\int x^\alpha dx = \frac{x^{\alpha+1}}{\alpha+1} + C}$$

на $\mathbb{R}$ для целых $\alpha \geqslant 0$ и $\alpha = \frac{1}{n}$, n — нечетное натуральное число; на $(-\infty, 0)$ и на $(0, +\infty)$ для целых $\alpha < -1$; на $[0, +\infty)$ для $\alpha \geqslant 0$; на $(0, +\infty)$ для $\alpha < 0$, $\alpha \neq -1$; *особо выделим, что во всех случаях $\alpha \neq -1$.*

$$2. \quad \boxed{\int \frac{1}{x}\, dx = \ln|x| + C}$$

на $(-\infty, 0)$ и на $(0, +\infty)$.

$$3. \quad \boxed{\int a^x dx = \frac{a^x}{\ln a} + C}$$

на $\mathbb{R}$, где $a > 0$ и $a \neq 1$.

$$4. \quad \boxed{\int \sin x\, dx = -\cos x + C}$$

на $\mathbb{R}$.

$$5. \quad \boxed{\int \cos x\,dx = \sin x + C}$$

на $\mathbb{R}$.

$$6. \quad \boxed{\int \frac{1}{\sin^2 x}\,dx = \int 1 + \operatorname{ctg}^2 x\,dx = -\operatorname{ctg} x + C}$$

на $(k\pi, (k+1)\pi)$ для $k \in \mathbb{Z}$.

$$7. \quad \boxed{\int \frac{1}{\cos^2 x}\,dx = \int 1 + \operatorname{tg}^2 x\,dx = \operatorname{tg} x + C}$$

на $\big((k-\frac{1}{2})\pi, (k+\frac{1}{2})\pi\big)$ для $k \in \mathbb{Z}$.

$$8. \quad \boxed{\int \frac{1}{\sqrt{a^2 - x^2}}\,dx = \arcsin\frac{x}{a} + C = -\arccos\frac{x}{a} + \tilde{C}}$$

на $[-a, a]$, где $a > 0$.

$$9. \quad \boxed{\int \frac{1}{a^2 + x^2}\,dx = \frac{1}{a}\operatorname{arctg}\frac{x}{a} + C = -\frac{1}{a}\operatorname{arcctg}\frac{x}{a} + \tilde{C}}$$

на $\mathbb{R}$, a — любое, не равное нулю.

$$10. \quad \boxed{\int \operatorname{sh} x\,dx = \operatorname{ch} x + C}$$

на $\mathbb{R}$.

$$11. \quad \boxed{\int \operatorname{ch} x\,dx = \operatorname{sh} x + C}$$

на $\mathbb{R}$.

$$12. \quad \boxed{\int \frac{1}{\operatorname{sh}^2 x}\,dx = \int \operatorname{cth}^2 x - 1\,dx = -\operatorname{cth} x + C}$$

на $(-\infty, 0)$ и на $(0, +\infty)$.

$$13. \quad \boxed{\int \frac{1}{\operatorname{ch}^2 x}\,dx = \int 1 - \operatorname{th}^2 x\,dx = \operatorname{th} x + C}$$

на $\mathbb{R}$.

$$14. \quad \boxed{\begin{gathered}\int \frac{1}{\sqrt{x^2 \pm a^2}}\,dx = \\ = \ln\left|x + \sqrt{x^2 \pm a^2}\right| + C = -\ln\left|x - \sqrt{x^2 \pm a^2}\right| + \tilde{C}\end{gathered}}$$

на $\mathbb{R}$ для знака “$+$” и на $(-\infty, -|a|)$ и $(|a|, +\infty)$ для знака “$-$”, где $a \neq 0$.

$$15. \quad \boxed{\int \frac{1}{a^2 - x^2}\,dx = \frac{1}{2a}\ln\left|\frac{x+a}{x-a}\right| + C}$$

на $(-\infty, -|a|)$, на $(-|a|, |a|)$ и на $(|a|, +\infty)$, где $a \neq 0$.

Написанные формулы сразу следуют из формул дифференцирования простейших элементарных, гиперболических и обратных гиперболических функций. Эти формулы легко также проверить непосредственным дифференцированием.

В связи с введением понятия неопределенного интеграла и операции интегрирования возникают вопросы. Какие функции имеют неопределенный интеграл или, иначе, какие функции можно интегрировать? И как найти неопределенный интеграл, если он существует? Хотя бы в случае элементарных функций.

Частичный ответ на первый вопрос будет получен позднее, будет установлено достаточное условие интегрируемости: всякая непрерывная на промежутке функция имеет точную первообразную, т.е. неопределенный интеграл. Что касается интегрирования элементарных функций, то, в отличие от дифференцирования, эта операция выводит за пределы элементарных функций (существуют элементарные функции интегралы от которых неэлементарны, например, e^{-x^2}, $\frac{\sin x}{x}$, $\frac{\cos x}{x}$, $\frac{1}{\ln x}$), а в тех случаях, когда не выводит, то часто нет общего алгоритма нахождения неопределенного интеграла, хотя для некоторых классов элементарных функций такие алгоритмы существуют.

Интегрирование рациональных дробей

Простейший класс интегрируемых функций — многочлены. Более сложный класс интегрируемых функций — рациональные функции, т.е. функции вида $\frac{P(x)}{Q(x)}$, где $P(x)$ и $Q(x)$ — многочлены.

Определение 1. Простейшими (**элементарными**) рациональными дробями называются дроби

$$\frac{A}{(x-a)^n} \quad \text{и} \quad \frac{Ax+B}{(x^2+px+q)^n},$$

где a, p, q, A, B — действительные числа, n — натуральное число и квадратичный трехчлен x^2+px+q не имеет действительных корней.

Покажем, как интегрируются простейшие дроби.

$$\int \frac{A}{x-a}\,dx = A\ln|x-a| + C,$$

$$\int \frac{A}{(x-a)^n}\,dx = A\frac{(x-a)^{-n+1}}{-n+1} + C = \frac{-A}{(n-1)(x-a)^{n-1}} + C$$

при $n > 1$.

Далее, $x^2+px+q = \left(x+\frac{p}{2}\right)^2 + \left(q-\frac{p^2}{4}\right)$, где $q-\frac{p^2}{4} > 0$, так как трехчлен не имеет действительных корней. Вводя обозначения $t = x+\frac{p}{2}$, $\lambda^2 = q-\frac{p^2}{4}$, $\alpha = A$ и $\beta = B-A\frac{p}{2}$ получаем, что при $n=1$ интеграл от второй простейшей дроби сводится к интегралу

$$\int \frac{\alpha t+\beta}{t^2+\lambda^2}\,dt = \frac{\alpha}{2}\int \frac{d(t^2+\lambda^2)}{t^2+\lambda^2} + \beta\int \frac{1}{t^2+\lambda^2}\,dt = \frac{\alpha}{2}\ln\left(t^2+\lambda^2\right) + \frac{\beta}{\lambda}\operatorname{arctg}\frac{t}{\lambda} + C,$$

т.е.

$$\int \frac{Ax+B}{x^2+px+q}\,dx = \frac{A}{2}\ln(x^2+px+q) + \frac{B-A\frac{p}{2}}{\sqrt{q-\frac{p^2}{4}}}\operatorname{arctg}\frac{x+\frac{p}{2}}{\sqrt{q-\frac{p^2}{4}}} + C.$$

Аналогично, заменой $t = x + \frac{p}{2}$ преобразуем интеграл $\int \frac{Ax+B}{(x^2+px+q)^n}\,dx$ с $n > 1$ в интеграл $\int \frac{\alpha t+\beta}{(t^2+\lambda^2)^n}\,dt$, где $\lambda^2 = q - \frac{p^2}{4} > 0$, $\alpha = A$, $\beta = B - A\frac{p}{2}$ (как и раньше). Тогда

$$\int \frac{\alpha t+\beta}{(t^2+\lambda^2)^n}\,dt = \frac{\alpha}{2}\int \frac{d(t^2+\lambda^2)}{(t^2+\lambda^2)^n} + \beta\int \frac{dt}{(t^2+\lambda^2)^n} =$$
$$= \frac{-\alpha}{2(n-1)(t^2+\lambda^2)^{n-1}} + \beta\int \frac{dt}{(t^2+\lambda^2)^n},$$

натуральное $n > 1$. Проинтегрируем последний интеграл по частям, обозначая его для краткости I_n, $n \geqslant 1$.

$$I_n = \frac{t}{(t^2+\lambda^2)^n} + n\int \frac{2t^2\,dt}{(t^2+\lambda^2)^{n+1}} = \frac{t}{(t^2+\lambda^2)^n} +$$
$$+2n\int \frac{dt}{(t^2+\lambda^2)^n} - 2n\lambda^2\int \frac{dt}{(t^2+\lambda^2)^{n+1}}.$$

Значит,

$$I_n = \frac{t}{(t^2+\lambda^2)^n} + 2nI_n - 2n\lambda^2 I_{n+1},$$

откуда

$$I_{n+1} = \frac{2n-1}{2n\lambda^2}I_n + \frac{t}{2n\lambda^2(t^2+\lambda^2)^n},$$
$$I_1 = \frac{1}{\lambda}\mathrm{arctg}\frac{t}{\lambda} + C,$$

или

$$I_n = \frac{2n-3}{2(n-1)\lambda^2}I_{n-1} + \frac{t}{2(n-1)\lambda^2(t^2+\lambda^2)^{n-1}},\ n > 1;$$
$$I_1 = \frac{1}{\lambda}\mathrm{arctg}\frac{t}{\lambda} + C.$$

Таким образом, интеграл I_n можно последовательно сводить к интегралам $I_{n-1},\ I_{n-2},\ldots,\ I_1$, последний их которых берется (т.е. вычисляется) непосредственно.

В итоге установлено, что все простейшие дроби можно проинтегрировать за конечное число указанных шагов и их первообразные — элементарные функции.

Определение 2. Рациональную дробь называют **правильной**, если степень многочлена в числителе строго меньше степени многочлена в знаменателе.

Из курса алгебры известно, что любую правильную рациональную дробь можно разложить в сумму простейших дробей. Точнее, верен следующий результат.

Утверждение (о разложении правильных дробей). *Если знаменатель правильной рациональной дроби с действительными коэффициентами* $\frac{P(x)}{Q(x)}$ *разложен в произведение*

$$Q(x) = \alpha\prod_i (x-a_i)^{\nu_i}\prod_j \left(x^2+p_jx+q_j\right)^{\mu_j},$$

где α *— действительное число,* a_i *— различные действительные числа,* (p_j, q_j) *— различные пары действительных чисел, причём квадратичные трехчлены* $x^2+p_jx+q_j$

не имеют действительных корней, ν_i, μ_j — натуральные числа, то тогда правильная дробь $\frac{P(x)}{Q(x)}$ представляется единственным образом как сумма простейших дробей

$$\frac{P(x)}{Q(x)} = \sum_i \sum_{k=1}^{\nu_i} \frac{A_k^i}{(x-a_i)^k} + \sum_j \sum_{k=1}^{\mu_j} \frac{B_k^j x + C_k^j}{(x^2+p_j x+q_j)^k},$$

где A_k^i, B_k^j, C_k^j — действительные числа.

Теорема 1 (об интегрировании рациональных дробей). *Интеграл от любой рациональной дроби с действительными коэффициентами сводится к интегрированию многочлена и простейших дробей, он выражается через рациональные функции, а также функции* ln *и* arctg *и, следовательно, является элементарной функцией.*

▼ Если рациональная дробь $\frac{P(x)}{Q(x)}$ не является правильной, т.е. степень числителя не меньше степени знаменателя, то разделив $P(x)$ на $Q(x)$ с остатком мы от дроби $\frac{P(x)}{Q(x)}$ перейдем к выражению $\mathcal{R}(x) + \frac{\mathcal{P}(x)}{Q(x)} = \frac{P(x)}{Q(x)}$, где $\mathcal{R}(x)$ — частное, многочлен с действительными коэффициентами, а $\mathcal{P}(x)$ — остаток от деления, многочлен с действительными коэффициентами степень которого меньше степени $Q(x)$. Интеграл от многочлена является многочленом степени на единицу больше, а интеграл от правильной рациональной дроби $\frac{\mathcal{P}(x)}{Q(x)}$ сводится к интегрированию простейших дробей и, значит, интеграл от любой рациональной функции выражается через рациональные функции, ln и arctg. ▲

Методы интегрирования

Нам известно, как интегрировать многочлены и простейшие дроби, поэтому рассмотрим сейчас методы разложения правильной дроби в сумму простейших.

Метод неопределенных коэффициентов.

Для правильной дроби $\frac{\mathcal{P}(x)}{Q(x)}$ пишется вышеприведенное разложение, в котором коэффициенты A_k^i, B_k^j и C_k^j считаются неизвестными. После этого простейшие дроби приводятся к общему знаменателю $Q(x)$ и складываются, а получившийся в числителе многочлен приравнивается к многочлену $P(x)$. Их коэффициенты должны совпадать и в результате получаем систему линейных уравнений с неизвестными A_k^i, B_k^j и C_k^j, которая имеет единственное решение (в силу единственности разложения в сумму простейших дробей). Найдя его, получим искомое разложение $\frac{\mathcal{P}(x)}{Q(x)}$ в сумму простейших дробей. Дополнительно заметим, что вместо приравнивания коэффициентов можно приравнивать значения получившегося многочлена значениям многочлена $P(x)$ в некоторых точках.

Метод вычеркивания

Разложив знаменатель правильной дроби $\frac{\mathcal{P}(x)}{Q(x)}$ в произведение

$$Q(x) = \alpha \prod_i (x-a_i)^{\nu_i} \prod_j \left(x^2+p_j x+q_j\right)^{\mu_j},$$

домножим дробь на $(x-a_l)^{\nu_l}$ — один из сомножителей в разложении $Q(x)$. Тогда, в соответствии с написанным ранее, имеет место равенство

$$\frac{P(x)}{\alpha\prod_{i\neq l}(x-a_i)^{\nu_i}\prod_j(x^2+p_jx+q_j)^{\mu_j}}=\sum_{k=1}^{\nu_l}A_k^l(x-a_l)^{\nu_l-k}+$$
$$+(x-a_l)^{\nu_l}\left(\sum_{i\neq l}\sum_{k=1}^{\nu_i}\frac{A_k^i}{(x-a_i)^k}+\sum_j\sum_{k=1}^{\mu_j}\frac{B_k^jx+C_k^j}{(x^2+p_jx+q_j)^k}\right),$$

полагая в нем $x=a_l$, получаем

$$\frac{P(a_l)}{\alpha\prod_{i\neq l}(a_l-a_i)^{\nu_i}\prod_j(a_l^2+p_ja_l+q_j)^{\mu_j}}=A_{\nu_l}^l,$$

т.е. вычеркивая в правильной дроби $\frac{\mathcal{P}(x)}{Q(x)}$ в разложении $Q(x)$ один из сомножителей $(x-a_l)^{\nu_l}$ и подставляя в оставшееся выражение $x=a_l$ получим $A_{\nu_l}^l$. Таким образом можно установить все $A_{\nu_i}^i$. Остальные коэффициенты обычно находят методом неопределенных коэффициентов.

Существуют и другие методы интегрирования, например, метод Остроградского. Изложение метода Остроградского имеется в учебнике Ильина и Позднякa, но мы его касаться не будем.

Лекция 3 (14.02.20) Интегрирование некоторых иррациональных функций и тригонометрических выражений

Интегрирование некоторых иррациональных функций

Определение 3. Многочленом двух переменных x и y называется конечная сумма членов вида ax^ky^l, где $a\in\mathbb{R}$, k и l из $\mathbb{Z}^+$.

Определение 4. Рациональной функцией (дробью) двух переменных x и y называется дробь, числитель и знаменатель которой — многочлены двух переменных.

Всюду в этой лекции далее $R(x,y)$ обозначает рациональную функцию двух переменных x и y.

Дробно-линейная иррациональность

1. $\int R\left(x,\sqrt[m]{\frac{ax+b}{cx+d}}\right)dx$ — **дробно-линейная иррациональность**, где a,b,c,d из $\mathbb{R}$, натуральное $m>1$, определитель $\begin{vmatrix}a&b\\c&d\end{vmatrix}\neq0$. Замена переменной $\boxed{t=\sqrt[m]{\dfrac{ax+b}{cx+d}}}$, $x=\frac{b-dt^m}{ct^m-a}$ сводит этот интеграл к интегралу от рациональной дроби.

Квадратичная иррациональность

2. $\int R\left(x, \sqrt{a^2x+bx+c}\right)\,dx$ — **квадратичная иррациональность**, a, b, c из $\mathbb{R}$. Сводится к интегралам от рациональных функций с помощью подстановок Эйлера.
1) $a>0$, $\boxed{\sqrt{a^2x+bx+c}=\pm t\pm x\sqrt{a}}$ (комбинация знаков произвольна). Так как при этом $a^2x+bx+c=t^2\pm 2\sqrt{a}xt+ax^2$, то $x=\frac{t^2-c}{b\mp 2t\sqrt{a}}$.
2) $c>0$, $\boxed{\sqrt{a^2x+bx+c}=\pm xt\pm\sqrt{c}}$ (комбинация знаков произвольна). Так как при этом $a^2x+bx+c=x^2t^2\pm 2\sqrt{c}xt+c$, то $x=\frac{b\mp 2t\sqrt{c}}{t^2-a}$.
3) $a^2x+bx+c=a(x-x_1)(x-x_2)$, где x_1 и x_2 — действительные числа, причем $x_1\neq x_2$, $\boxed{\sqrt{a^2x+bx+c}=\pm(x-x_1)t}$, $x=\frac{x_1t^2-ax_2}{t^2-a}$.

Заметим, что если у квадратичного трехчлена нет корней, то он или всюду отрицателен и подынтегральное выражение всюду не имеет смысла или всюду положителен, а тогда и $c>0$ (значение трехчлена при $x=0$) и $a>0$ (знак трехчлена совпадает со знаком a при достаточно большом $|x|$), т.е. при этом применимы и 1) и 2) постановки Эйлера. Если же корни трехчлена совпадают, $x_1=x_2$, то $\sqrt{a^2x+bx+c}=\sqrt{a}|x-x_1|$ и интеграл сразу является интегралом от рациональной функции (возможно, от различных функций при $x\geqslant x_1$ и $x\leqslant x_1$). Значит, подстановки Эйлера позволяют всегда проинтегрировать квадратичную иррациональность.

Укажем и некоторые другие использующиеся при интегрировании квадратичных иррациональностей приемы. Так при вычислении
$\int R\left(x, \sqrt{\lambda^2-x^2}\right)\,dx$ используют замены $x=\lambda\sin t$ или $x=\lambda\cos t$ или $x=\lambda\,\mathrm{th}\,t$;
$\int R\left(x, \sqrt{\lambda^2+x^2}\right)\,dx$ используют замены $x=\lambda\,\mathrm{tg}\,t$ или $x=\lambda\,\mathrm{ctg}\,t$ или $x=\lambda\,\mathrm{sh}\,t$;
$\int R\left(x, \sqrt{x^2-\lambda^2}\right)\,dx$ используют замены $x=\lambda\sec t$ или $x=\lambda\,\mathrm{cosec}\,t$ или $x=\lambda\,\mathrm{ch}\,t$.

Дифференциальный бином (или биномиальный дифференциал)

3. $\int x^p(a+bx^q)^r dx$ — интеграл от **дифференциального бинома** (или **биномиального дифференциала**), где p, q и r — рациональные, а a и b — действительные числа. Заменой $t=x^q$ сводится к интегралу $\frac{1}{q}\int(a+bt)^r t^{\frac{p+1}{q}-1}dt$, который в трех случаях сводится к интегрированию рациональных функций (в остальных случаях, как показал П.Л.Чебышёв, он не берется в элементарных функциях).
1) r — целое, замена $u=\boxed{\sqrt[N]{t}}$, где N — знаменатель дроби $\frac{p+1}{q}$;
2) $\frac{p+1}{q}$ — целое, является интегралом от дробно-линейной иррациональности, заменой $\boxed{u=a+bt}$ может быть сведен к случаю 1);
3) $\frac{p+1}{q}+r$ — целое, как интеграл $\frac{1}{q}\int\left(\frac{a+bt}{t}\right)^r t^{\frac{p+1}{q}+r-1}dt$ является интегралом от дробно-линейной иррациональности.

Рациональное тригонометрическое выражение

4. $\int R(\sin x,\cos x)\,dx$ — интеграл от **рационального тригонометрического выражения**. Заменой $\boxed{t=\mathrm{tg}\frac{x}{2}}$ сводится к интегралу $\int R\left(\frac{2t}{1+t^2},\frac{1-t^2}{1+t^2}\right)\frac{2dt}{1+t^2}$ — интегралу от рациональной функции. Но в ряде случаев выгоднее другие замены. Лучше
1) если $R(u,v)=-R(-u,v)$, замена $\boxed{t=\cos x}$;
2) если $R(u,v)=-R(u,-v)$, замена $\boxed{t=\sin x}$;
3) если $R(u,v)=R(-u,-v)$, замена $\boxed{t=\mathrm{tg}\,x}$ или $\boxed{t=\mathrm{ctg}\,x}$.

Отметим, что всегда можно обойтись такими заменами, так как $R(u,v) = \frac{R(u,v)-R(-u,v)}{2} + \frac{R(u,v)-R(u,-v)}{2} + \frac{R(-u,v)+R(u,-v)}{2}$, где к каждому слагаемому применима соответственно замена 1), 2) или 3).

Пример. $\int \frac{dx}{\sin x \cos^3 x}$. При замене $t = \operatorname{tg}\frac{t}{2}$ он сводится к $\int \frac{(1+t^2)^3}{t(1-t^2)^3}\,dt$. При замене $t = \cos x$ он сводится к $-\int \frac{dt}{t^3(1-t^2)}$. При замене $t = \sin x$ он сводится к $\int \frac{dt}{t(1-t^2)^2}$. При замене $t = \operatorname{tg} x$ он сводится к $\int \frac{1+t^2}{t}\,dt$, при замене $t = \operatorname{ctg} x$ он своится к $-\int \frac{1+t^2}{t^3}\,dt$. Это показывает, что использование различных замен может приводить к интегралам весьма разной сложности. Возможны и другие приемы интегрирования, например $\int \frac{dx}{\sin x \cos^3 x} = \int \frac{(\sin^2 + \cos^2 x)\,dx}{\sin x \cos^3 x} = \int \frac{\sin x\,dx}{\cos^3 x} + \int \frac{dx}{\sin x \cos x} + C = \int \frac{-d\cos x}{\cos^3 x} + \int \frac{(\sin^2 + \cos^2 x)\,dx}{\sin x \cos x} + C = \frac{1}{2\cos^2 x} - \ln|\cos x| + \ln|\sin x| + C$.

Рациональная функция от экспоненты

5. $\int R(a^x)\,dx$ — интеграл с рациональной функцией от экспоненты. Сводится к интегралу от рациональной функции заменой $\boxed{t = a^x}$. Иногда также используют замены с различными гиперболическими функциями.

Квазимногочлены

6. $\int P(x)a^x\,dx$, $\int P(x)\cos x\,dx$, $\int P(x)\sin x\,dx$, где $P(x)$ — многочлен, — интегралы с квазимногочленами, вычисляются интегрированием по частям, приводящим к интегралам того же вида, но с многочленом меньшей степени. Тем же приемом интегрирования по частям вычисляются интегралы вида
$\int P(x)\arcsin x\,dx$, $\int P(x)\arccos x\,dx$, $\int P(x)\operatorname{arctg} x\,dx$, $\int P(x)\operatorname{arcctg} x\,dx$, $\int P(x)\ln x\,dx$, $\int P(a^x)\sin \beta x\,dx$, $\int P(a^x)\cos \beta x\,dx$.

Интегрирование $\sin^m x \cdot \cos^n x$

7. $\int \sin^m x \cdot \cos^n x\,dx$, где m и n — целые или рациональные числа. С помощью замены $\boxed{t = \sin x}$ или $\boxed{t = \cos x}$ сводится к интегралу от дифференциального бинома. Применяются и другие замены, а также интегрирование по частям или формулы преобразования произведений тригонометрических функций в суммы.

Полезно также знать наиболее часто встречающиеся неберущиеся в элементарных функциях интегралы.
$\int \frac{\sin x}{x}\,dx$ — интегральный синус;
$\int \frac{\cos x}{x}\,dx$ — интегральный косинус;
$\int \frac{1}{\ln x}\,dx$ — интегральный логарифм;
$\int e^{-x^2} dx$ — интеграл Пуассона;
$\int \sin(x^2)\,dx$, $\int \cos(x^2)\,dx$ — интегралы Френеля;
$\int R\left(x, \sqrt{P(x)}\right)\,dx$, где $P(x)$ — многочлен третьей или четвертой степени, — эллиптические интегралы, в общем случае не выражаются через элементарные функции.

Лекция 4 (18.02.20)
Определенные интегралы.
Интегралы Римана и Курцвейля–Хенстока

Определения

Определение 1. Два отрезка $[a,b]$ и $[c,d]$ **не перекрываются**, если они не пересекаются или их пересечение — точка, которая является концом каждого из них.

Определение 2. Системой неперекрывающихся отрезков будем называть систему отрезков, в которой любые два отрезка не перекрываются.

Определение 3. Рассмотрим отрезок $[a,b] \subset \mathbb{R}$. **Разбиение** T отрезка $[a,b]$ — любой конечный набор неперекрывающихся отрезков $\{\Delta_i\}_{i=1}^n$, объединение которых равно $[a,b]$.

Порядок нумерации отрезков разбиения неважен, но обычно их нумеруют в порядке расположения на оси следующим образом: $\Delta_i = [a_{i-1}, a_i]$, где $a = a_0 < a_1 < a_2 < \cdots < a_n = b$.

Так как система концов отрезков разбиения $\{a_i\}_{i=0}^n$ позволяет однозначно восстановить отрезки разбиения, то иногда *разбиением* называют систему концов отрезков разбиения, но мы эту терминологию не будем использовать.

Определение 4. Длина отрезка $[a,b]$ равна $b-a$ и будем обозначаться как $|[a,b]|$, причем в дальнейшем длину пустого или одноточечного множества считаем равной 0.

Определение 5. Если $T = \{\Delta_i\}_{i=1}^n$ — разбиение отрезка $[a,b]$, а $\boldsymbol{\xi} = \{\xi_i\}_{i=1}^n$ — набор точек $\xi_i \in \Delta_i$, возможно, повторяющихся, то **разбиением с отмеченными точками** или, кратко, **отмеченным разбиением** $\mathbb{T}$ называют множество пар $\{(\Delta_i, \xi_i)\}_{i=1}^n$. Каждую пару (Δ_i, ξ_i) будем называть элементом отмеченного разбиения. Точки ξ_i будем называть **отмеченными точками** или **метками**.

Все рассматриваемые дальше функции считаем действительнозначными, хотя при желании можно считать их и комплекснозначными. Те места, где различие между действительным и комплексным случаями существенно, будем отмечать особо.

Определение 6. Пусть на отрезке $[a,b]$ определена функция f. **Интегральной суммой** или **суммой Римана** функции f на отрезке $[a,b]$, соответствующей отмеченному разбиению (подразбиению) $\mathbb{T} = \{(\Delta_i, \xi_i)\}_{i=1}^n$, называют сумму

$$\mathfrak{S}(f, \mathbb{T}) = \sum_{\mathbb{T}} f(\xi_i)|\Delta_i| = \sum_{i=1}^n f(\xi_i)|\Delta_i|.$$

Интеграл Римана

Определение 7. Функция f **интегрируема** на отрезке $[a,b]$ (по отрезку $[a,b]$) **в смысле Римана** и ее интеграл равен числу $\boldsymbol{I}$, если f определена на $[a,b]$ и для любого $\varepsilon > 0$ существует такое число $\delta > 0$, что для любого отмеченного разбиения $\mathbb{T} = \{(\Delta_i, \xi_i)\}$ с $|\Delta_i| < \delta$ для всех i верна оценка $|\mathfrak{S}(f,\mathbb{T}) - \boldsymbol{I}| = \left|\sum_{\mathbb{T}} f(\xi_i)|\Delta_i| - \boldsymbol{I}\right| < \varepsilon$, т.е.

$$\forall \varepsilon > 0 \quad \exists \text{число } \delta > 0 \quad \forall \mathbb{T} = \{(\Delta_i, \xi_i)\}, \quad |\Delta_i| < \delta \text{ для всех } i,$$

$$|\mathfrak{S}(f,\mathbb{T}) - \boldsymbol{I}| = \left|\sum_{\mathbb{T}} f(\xi_i)|\Delta_i| - \boldsymbol{I}\right| < \varepsilon.$$

Число $\boldsymbol{I}$ называют **определенным интегралом Римана** функции f по отрезку $[a,b]$ (на отрезке $[a,b]$).

Определенный интеграл Римана функции f на отрезке $[a,b]$ (по отрезку $[a,b]$) обозначают как $\int\limits_a^b f\,dx$ или $\int\limits_{[a,b]} f\,dx$, а если хотят подчеркнуть, что это интеграл Римана (в смысле Римана), то как $(\mathcal{R})\int\limits_a^b f\,dx$ или $(\mathcal{R})\int\limits_{[a,b]} f\,dx$.

В обозначении $\int\limits_a^b f\,dx$ знак $\int$ — интеграл, a — нижний предел, b — верхний предел, f — подынтегральная функция, $f\,dx$ — подынтегральное выражение, dx — дифференциал.

Определение 8. Масштабом на множестве E называется любая действительнозначная строго положительная функция на E.

Интеграл Курцвейля–Хенстока

Определение 9. Функция f интегрируема на отрезке $[a,b]$ (по отрезку $[a,b]$) **в смысле Курцвейля–Хенстока** и ее интеграл равен числу $\boldsymbol{I}$, если f определена на $[a,b]$ и для любого $\varepsilon>0$ существует такой масштаб δ на $[a,b]$, что для любого отмеченного разбиения $\mathbb{T}=\{(\Delta_i,\xi_i)\}$ с $|\Delta_i|<\delta(\xi_i)$ для всех i верна оценка $|\mathfrak{S}(f,\mathbb{T})-\boldsymbol{I}|=\left|\sum\limits_{\mathbb{T}} f(\xi_i)|\Delta_i|-\boldsymbol{I}\right|<\varepsilon$, т.е.

$$\forall\varepsilon>0\quad \exists\text{масштаб } \delta \text{ на } [a,b]\quad \forall\mathbb{T}=\{(\Delta_i,\xi_i)\},\quad |\Delta_i|<\delta(\xi_i) \text{ для всех } i,$$
$$|\mathfrak{S}(f,\mathbb{T})-\boldsymbol{I}|=\left|\sum_{\mathbb{T}} f(\xi_i)|\Delta_i|-\boldsymbol{I}\right|<\varepsilon.$$

Число $\boldsymbol{I}$ называют **определённым интегралом Курцвейля–Хенстока** функции f на отрезке $[a,b]$ (по отрезку $[a,b]$).

Определенный интеграл Курцвейля–Хенстока от функции f на отрезке $[a,b]$ (по отрезку $[a,b]$) обозначают как $\int\limits_a^b f\,dx$ или $\int\limits_{[a,b]} f\,dx$, а если хотят подчеркнуть, что это интеграл Курцвейля–Хенстока (в смысле Курцвейля–Хенстока), то как $(\mathcal{H})\int\limits_a^b f\,dx$ или $(\mathcal{H})\int\limits_{[a,b]} f\,dx$.

Интегралы как пределы по базе

Заметим, что оба определения интеграла фактически определяют интеграл как предел интегральных сумм Римана по базе.

Действительно, пусть $M=\{\mathbb{T}\}$ — множество отмеченных разбиений отрезка $[a,b]$. На M при помощи функции f определена функция $\sum\limits_{\mathbb{T}} f(\xi_i)|\Delta_i|$, сопоставляющая каждому отмеченному разбиению $\mathbb{T}=\{(\Delta_i,\xi_i)\}$ интегральную сумму Римана $\sum\limits_i f(\xi_i)|\Delta_i|$.

Для интеграла Римана соответствующая база $\mathfrak{B}_{\mathcal{R}}$ состоит из множеств $\boldsymbol{B}_\delta = \{\mathbb{T} \in M: \ |\Delta_i| < \delta$ для всех $i\}$, где число $\delta > 0$.

Для интеграла Курцвейля–Хенстока соответствующая база $\mathfrak{B}_{\mathcal{H}}$ состоит из множеств $\boldsymbol{B}_\delta^{\mathcal{H}} = \{\mathbb{T} \in M: \ |\Delta_i| < \delta(\xi_i)$ для всех $i\}$, где δ — масштаб на $[a, b]$.

По определению база $\mathfrak{B}$ должна быть непустой и должна содержать только непустые множества, кроме того, для любых двух элементов $B_1, B_2 \in \mathfrak{B}$ должен существовать такой элемент $B \in \mathfrak{B}$, что $B \subset B_1 \cap B_2$.

Для базы $\mathfrak{B}_{\mathcal{R}}$ очевидно, что все множества $\boldsymbol{B}_\delta$ непусты. Для базы $\mathfrak{B}_{\mathcal{H}}$ непустоту элементов необходимо проверить.

Определение 10. Будем говорить, что отмеченное разбиение $\mathbb{T} = \{(\Delta_i, \xi_i)\}_{i=1}^n$ **мельче** числа $\delta > 0$, если $|\Delta_i| < \delta$ для всех i, $1 \leqslant i \leqslant n$.

Определение 11. Скажем, что отмеченное разбиение $\mathbb{T} = \{(\Delta_i, \xi_i)\}_{i=1}^n$ **согласовано** с масштабом δ на $[a, b]$ (или что $\mathbb{T}$ является δ-**разбиением)**, если для любого i, $1 \leqslant i \leqslant n$, $|\Delta_i| < \delta(\xi_i)$.

Основная лемма. *Для любого заданного на* $[a, b]$ *масштаба* δ *существует согласованное с этим масштабом* δ *разбиение* $\mathbb{T}$ *отрезка* $[a, b]$.

▼ Предположим, что такого разбиения отрезка $[a, b]$ нет. Обозначим отрезкок $[a, b]$ как I_1 и разобъем его на две половины — два неперекрывающихся отрезка равной длины. Тогда хотя бы для одной из половин также не существует требуемого согласованного с масштабом δ отмеченного разбиения $\mathbb{T}$ (ведь если бы для каждой половины такие отмеченные разбиения существовали, то их объединение было бы искомым отмеченным разбиением отрезка I_1). Обозначим эту половину как отрезок I_2 и разобъем ее на две половины. Хотя бы одна из них не имеет требуемого согласованного с масштабом δ отмеченного разбиения. Обозначим ее I_3 и опять разобъем пополам .... Продолжая такое построение до бесконечности, получим последовательность вложенных отрезков с длинами, стремящимися к нулю (ведь при каждом шаге получается отрезок вдвое меньшей длины), где каждый из отрезков не имеет согласованного с масштабом δ отмеченного разбиения $\mathbb{T}$ (с $\xi_i \in \Delta_i$ для всех i). По принципу полноты Кантора существует точка ξ, принадлежащая всем построенным отрезкам I_k. Так как $\delta(\xi) > 0$, то найдется такой отрезок I_m, что $|I_m| < \delta(\xi)$. Но тогда отмеченное разбиение, состоящее из единственной пары (I_m, ξ), будет требуемым разбиением отрезка I_m, ведь $\xi \in I_m$ и $|I_m| < \delta(\xi)$. Это противоречит построению отрезков I_k и тем самым доказывает ложность предположения, что требуемого отмеченного разбиения отрезка $[a, b]$ нет. ▲

Теперь удостоверимся, что обе базы обладают необходимым свойством

$$\forall B_1, B_2 \in \mathfrak{B} \ \exists B \in \mathfrak{B}: \ B \subset B_1 \cap B_2.$$

▼ Действительно,

$$\boldsymbol{B}_{\delta_1} \cap \boldsymbol{B}_{\delta_2} = \boldsymbol{B}_\delta, \text{ где } \delta = \min\{\delta_1, \delta_2\};$$
$$\boldsymbol{B}_{\delta_1}^{\mathcal{H}} \cap \boldsymbol{B}_{\delta_2}^{\mathcal{H}} = B_\delta^{\mathcal{H}}, \text{ где } \delta(x) = \min\{\delta_1(x), \delta_2(x)\}. \blacktriangle$$

Итак, проверено, что $\mathfrak{B}_{\mathcal{R}}$ и $\mathfrak{B}_{\mathcal{H}}$ — базы в множестве отмеченных разбиений. По определениям интегралов Римана и Курцвейля–Хенстока

$$(\mathcal{R})\int\limits_a^b f\,dx = \lim_{\mathfrak{B}_\mathcal{R}} \mathfrak{S}(f,\mathbb{T}) = \lim_{\mathfrak{B}_\mathcal{R}} \sum_{\mathbb{T}} f(\xi_i)|\Delta_i|,$$

$$(\mathcal{H})\int\limits_a^b f\,dx = \lim_{\mathfrak{B}_\mathcal{H}} \mathfrak{S}(f,\mathbb{T}) = \lim_{\mathfrak{B}_\mathcal{H}} \sum_{\mathbb{T}} f(\xi_i)|\Delta_i|.$$

Простейшие свойства

Из свойств предела по базе сразу следуют некоторые простейшие свойства обоих интегралов римановского типа.

Свойство 1 (взаимоотношение интегралов). *Если функция f на $[a,b]$ интегрируема в смысле Римана и число $\boldsymbol{I}$ — ее интеграл, то f на $[a,b]$ интегрируема в смысле Курцвейля–Хенстока и ее интеграл то же самое число $\boldsymbol{I}$.*

▼ Действительно, так как $\boldsymbol{B}_\delta \supset \boldsymbol{B}^{\mathcal{H}}_{\delta(x)}$ при $0 < \delta(x) < \frac{\delta}{2}$ на $[a,b]$ и $\boldsymbol{B}^{\mathcal{M}}_{\delta(x)} \supset \boldsymbol{B}^{\mathcal{H}}_{\delta(x)}$, то по теореме о пределах по разным базам это свойство верно. ▲

Замечание. Если в определении интеграла Курцвейля-Хенстока ограничиться постоянными масштабами, то, как легко видеть, получится определение интеграла Римана.

Следующий пример показывает, что существует функции, которые не интегрируемы по Риману, но интегрируемы по Курцвейлю-Хенстоку, т.е., что интеграл Курцвейля-Хенстока существенно более общий, чем интеграл Римана.

Лекция 5 (21.02.20)
Свойства интегралов
Римана и Курцвейля–Хенстока

Пример. Функция Дирихле

$$\mathcal{D}(x) = \begin{cases} 1, & \text{если } x \text{ — рационально,} \\ 0, & \text{если } x \text{ — иррационально,} \end{cases}$$

на любом отрезке $[a,b]$, $b > a$, неинтегрируема по Риману, но интегрируема по Курцвейлю-Хенстоку.

▼ Действительно, $\mathcal{D}(x)$ неинтегрируема по Риману на любом отрезке $[a,b]$, $b > a$, так как для любого разбиения T можно выбрать все точки $\xi_i \in \Delta_i$ рациональными и тогда $\mathfrak{S}(f,\mathbb{T}) = \sum\limits_i |\Delta_i| = b - a$, а можно выбрать все точки $\xi_i \in \Delta_i$ иррациональными и тогда $\mathfrak{S}(f,\mathbb{T}) = 0$. Нетрудно убедиться, что функция Дирихле $\mathcal{D}(x)$ на любом отрезке $[a,b]$ интегрируема по Курцвейлю–Хенстоку, и интеграл от нее равен нулю. Для этого достаточно произвольному $\varepsilon > 0$ поставить в соответствие масштаб δ определенный следующим образом. Пусть $\{r_k\}_{k=1}^\infty$ — последовательность всех рациональных точек отрезка $[a,b]$. Положим

$$\delta(x) = \begin{cases} 1, & \text{если } x \text{ иррационально,} \\ \varepsilon \cdot 2^{-k-1}, & \text{если } x = r_k. \end{cases}$$

Тогда при любом согласованном с этим масштабом δ отмеченном разбиении $\mathbb{T} = \{(\Delta_i, \xi_i)\}$ справедлива оценка

$$0 \leqslant \sum_{\mathbb{T}} \mathcal{D}(\xi_i)|\Delta_i| = \sum_{k=1}^{\infty} \sum_{\xi_i = r_k} |\Delta_i| < \sum_{k=1}^{\infty} 2\varepsilon \cdot 2^{-k-1} = \varepsilon$$

и, поэтому, $(\mathcal{H}) \int\limits_a^b \mathcal{D}(x)\, dx = 0$. ▲

Отметим, что существуют более сложные примеры ограниченных функций, которые неинтегрируемы по Курцвейлю–Хенстоку.

Свойство 2 (единственность интегралов). *Если интеграл Римана или Курцвейля–Хенстока от функции f на отрезке $[a, b]$ существует, то он единственен.*

▼ Это непосредственное следствие теоремы о единственности предела по базе. ▲

Свойство 3 (линейность по функциям). *Если функция f интегрируема на отрезке $[a, b]$ в смысле Римана или Курцвейля–Хенстока, а c — число, то cf интегрируема на $[a, b]$ в том же смысле и $\int\limits_a^b cf\, dx = c\int\limits_a^b f\, dx$. Если функции f и g интегрируемы на $[a, b]$ в смысле Римана или Курцвейля–Хенстока, то и $f \pm g$ интегрируема на $[a, b]$ в том же смысле и $\int\limits_a^b (f \pm g)\, dx = \int\limits_a^b f\, dx \pm \int\limits_a^b g\, dx$.*

▼ Действительно, функции cf на множестве отмеченных разбиений соответствует функция $\mathfrak{S}(cf, \mathbb{T}) = c\mathfrak{S}(f, \mathbb{T})$ и если $\mathfrak{B}$ — любая из двух указанных баз, то по теореме о пределе произведения существует $\int\limits_a^b cf\, dx = \lim\limits_{\mathfrak{B}} \mathfrak{S}(cf, \mathbb{T}) = \lim\limits_{\mathfrak{B}} c\mathfrak{S}(f, \mathbb{T}) = c \lim\limits_{\mathfrak{B}} \mathfrak{S}(f, \mathbb{T}) = c\int\limits_a^b f\, dx$.

Аналогично, функции $f \pm g$ на множестве отмеченных разбиений соответствует функция $\mathfrak{S}(f \pm g, \mathbb{T}) = \mathfrak{S}(f, \mathbb{T}) \pm \mathfrak{S}(g, \mathbb{T})$ и если $\mathfrak{B}$ — любая из двух указанных баз, то по теореме о пределе суммы-разности существует $\int\limits_a^b f \pm g\, dx = \lim\limits_{\mathfrak{B}} \mathfrak{S}(f \pm g, \mathbb{T}) = \lim\limits_{\mathfrak{B}} (\mathfrak{S}(f, \mathbb{T}) \pm \mathfrak{S}(g, \mathbb{T})) = \lim\limits_{\mathfrak{B}} \mathfrak{S}(f, \mathbb{T}) \pm \lim\limits_{\mathfrak{B}} \mathfrak{S}(g, \mathbb{T}) = \int\limits_a^b f\, dx \pm \int\limits_a^b g\, dx$. ▲

Свойство 4 (сохранение неравенств). *Если функция f интегрируема на $[a, b]$ в любом из двух смыслов, функция g интегрируема на $[a, b]$ в любом из двух смыслов и $f(x) \leqslant g(x)$ на $[a, b]$, то для их интегралов (возможно, в разных смыслах) справедливо неравенство $\int\limits_a^b f\, dx \leqslant \int\limits_a^b g\, dx$.*

▼ Действительно, в силу свойства 1 о взаимоотношении интегралов, можно ограничиться случаем интегрируемости f и g в смысле Курцвейля–Хенстока. Если $f(x) \leqslant g(x)$ на $[a, b]$, то на множестве отмеченных разбиений $\sum\limits_{\mathbb{T}} f(\xi_i)|\Delta_i| \leqslant \sum\limits_{\mathbb{T}} g(\xi_i)|\Delta_i|$ и, значит, в силу теоремы о переходе к пределу в неравенствах (для пределов по базе) $\int\limits_a^b f\, dx = \lim\limits_{\mathfrak{B}_{\mathcal{H}}} \sum\limits_{\mathbb{T}} f(\xi_i)|\Delta_i| \leqslant \lim\limits_{\mathfrak{B}_{\mathcal{H}}} \sum\limits_{\mathbb{T}} g(\xi_i)|\Delta_i| = \int\limits_a^b g\, dx$. ▲

Критерии Коши существования интегралов Римана и Курцвейля–Хенстока

Теперь вспомним критерий Коши существования предела по базе.

Критерий Коши. *Конечный предел функции f по базе $\mathfrak{B}$ существует тогда и только тогда, когда для функции f выполняется условие: $\forall\varepsilon > 0\ \exists B \in \mathfrak{B}\ \forall x, x' \in B:\ |f(x) - f(x')| < \varepsilon$.*

Используя критерий Коши получаем два следующих критерия интегрируемости.

Критерий Коши $\mathcal{R}$-интегрируемости. *Функция f, определенная на отрезке $[a, b]$, интегрируема на $[a, b]$ в смысле Римана тогда и только тогда, когда для любого $\varepsilon > 0$ найдется такое $\delta > 0$, что для любых разбиений $\mathbb{T}$, $\mathbb{T}'$ из $\boldsymbol{B}_\delta$ выполняется неравенство $|\mathfrak{S}(f, \mathbb{T}) - \mathfrak{S}(f, \mathbb{T}')| < \varepsilon$, т.е.*

$$\forall\varepsilon > 0\ \exists\delta > 0\ \forall\mathbb{T}, \mathbb{T}' \in \boldsymbol{B}_\delta :$$
$$|\mathfrak{S}(f, \mathbb{T}) - \mathfrak{S}(f, \mathbb{T}')| < \varepsilon.$$

Критерий Коши $\mathcal{H}$-интегрируемости. *Функция f, определенная на отрезке $[a, b]$, интегрируема на $[a, b]$ в смысле Курцвейля–Хенстока тогда и только тогда, когда для любого $\varepsilon > 0$ найдется такой масштаб δ, что для любых разбиений Хенстока $\mathbb{T}$, $\mathbb{T}'$ из $\boldsymbol{B}_\delta^{\mathcal{H}}$ выполняется неравенство $|\mathfrak{S}(f, \mathbb{T}) - \mathfrak{S}(f, \mathbb{T}')| < \varepsilon$, т.е.*

$$\forall\varepsilon > 0\ \exists \textit{масштаб } \delta \textit{ на } [a, b]\ \forall\mathbb{T}, \mathbb{T}' \in \boldsymbol{B}_\delta^{\mathcal{H}} :$$
$$|\mathfrak{S}(f, \mathbb{T}) - \mathfrak{S}(f, \mathbb{T}')| < \varepsilon.$$

Приведенные критерии позволяют доказать следующее свойство.

Свойство 5 (интегрируемость на подотрезках). *Если функция f интегрируема на отрезке $[a, b]$ в смысле Римана или Курцвейля–Хенстока, то она интегрируема в том же смысле и на любом отрезке $[\overline{a}, \overline{b}] \subset [a, b]$.*

▼ Аналогично введенным для отрезка $[a, b]$ классам отмеченных разбиений $\boldsymbol{B}_\delta$ и $\boldsymbol{B}_\delta^{\mathcal{H}}$ введем для отрезка $[\overline{a}, \overline{b}]$ такие же классы отмеченных разбиений $\overline{\boldsymbol{B}}_\delta$ и $\overline{\boldsymbol{B}}_\delta^{\mathcal{H}}$. Пусть для заданного $\varepsilon > 0$ найдено такое подходящее число $\delta > 0$ (масштаб δ на $[a, b]$), что вышенаписанное условие критерия Коши выполнено. Покажем, что при тех же ε и δ условие критерия Коши выполнено на $[\overline{a}, \overline{b}]$. Рассмотрим два произвольных разбиения $\overline{\mathbb{T}}$ и $\overline{\mathbb{T}}'$ из $\overline{\boldsymbol{B}}_\delta$ (из $\overline{\boldsymbol{B}}_\delta^{\mathcal{H}}$). Если $a \neq \overline{a}$, то дополним $\overline{\mathbb{T}}$ и $\overline{\mathbb{T}}'$ одним и тем же отмеченным разбиением отрезка $[a, \overline{a}]$ мельче δ (согласованным с масштабом δ на $[a, \overline{a}]$); если $\overline{b} \neq b$, то еще дополним $\overline{\mathbb{T}}$ и $\overline{\mathbb{T}}'$ одним и тем же отмеченным разбиением отрезка $[\overline{b}, b]$ мельче δ (согласованным с масштабом δ на $[\overline{b}, b]$). В результате из отмеченных разбиений $\overline{\mathbb{T}}$ и $\overline{\mathbb{T}}'$ отрезка $[\overline{a}, \overline{b}]$ получим разбиения отрезка $[a, b]$ $\mathbb{T}$ и $\mathbb{T}'$ из $\boldsymbol{B}_\delta$ (из $\boldsymbol{B}_\delta^{\mathcal{H}}$). Так как дополнялись разбиения $\overline{\mathbb{T}}$ и $\overline{\mathbb{T}}'$ одинаковым образом, то разность интегральных сумм $\mathfrak{S}(f, \overline{\mathbb{T}}) - \mathfrak{S}(f, \overline{\mathbb{T}'})$ на отрезке $[\overline{a}, \overline{b}]$ равна разности интегральных сумм $\mathfrak{S}(f, \mathbb{T}) - \mathfrak{S}(f, \mathbb{T}')$ на отрезке $[a, b]$ и, значит, (в силу критерия Коши) меньше ε по абсолютной величине. Следовательно, условие критерия Коши интегрируемости на отрезке $[\overline{a}, \overline{b}]$ выполнено и f интегрируема на $[\overline{a}, \overline{b}]$ в смысле Римана (Курцвейля–Хенстока). ▲

В заключение раздела отметим, что интеграл Римана — широко известный классический интеграл. Что же касается интеграла Курцвейля–Хенстока, то он эквивалентен узкому интегралу Данжуа и интегралу Перрона и фактически также является новым определением известного специалистам интеграла.

Лекция 6 (25.02.20)
Ограниченнсть интегрируемых
по Риману функций.
Формула Ньютона–Лейбница

Ограниченность интегрируемых по Риману функций

Теорема 1 (необходимое усл. $\mathcal{R}$-интегрируемости). *Необходимым условием интегрируемости по Риману функции f на отрезке $[a,b]$ является ограниченность f на $[a,b]$.*

▼ Покажем, что неограниченная на отрезке $[a,b]$ функция f неинтегрируема по Риману на $[a,b]$.

Если функция f неограничена на отрезке $[a,b]$, то тогда для любого разбиения $T=\{\Delta_i\}_{i=1}^n$ функция f неограничена хотя бы на одном из отрезков разбиения Δ_j (ведь если $|f(x)|\leqslant C_i$ на Δ_i, то $|f(x)|\leqslant \max\limits_{1\leqslant i\leqslant n} C_i$ на $[a,b]$). Значит, $|f(\xi_j)|$, где $\xi_j\in\Delta_j$, может принимать сколь угодно большие значения. Так как $|\Delta_j|>0$, то тоже самое можно сказать и о $|f(\xi_j)|\cdot|\Delta_j|$. Зафиксировав каким-либо образом $\xi_i\in\Delta_i$ при $i\neq j$ и меняя $\xi_j\in\Delta_j$ видим, что интегральная сумма

$$\mathfrak{S}(f,\mathbb{T})=\sum_i f(\xi_i)|\Delta_i| = f(\xi_j)|\Delta_j| + \sum_{i\neq j} f(\xi_i)|\Delta_i|$$

также не является ограниченной. По свойству предела по базе из интегрируемости f следует ограниченность $\mathfrak{S}(f,\mathbb{T})$ на некотором элементе $\boldsymbol{B}_\delta$ базы $\mathfrak{B}_{\mathcal{R}}$, что находится в противоречии с тем, что при изменении одной из отмеченных точек $\xi_j\in\Delta_j$ интегральная сумма $\mathfrak{S}(f,\mathbb{T})$ не является ограниченной величиной. ▲

Что ограниченности функции недостаточно для ее интегрируемости по Риману, видно на примере функции Дирихле.

А сейчас установим еще одно свойство введенных интегралов.

Аддитивность интегралов по отрезкам

Теорема 2 (аддитивность по отрезкам). *Пусть $a<b<c$ и функция f интегрируема на отрезках $[a,b]$ и $[b,c]$ в смысле Римана или Курцвейля–Хенстока. Тогда f интегрируема на $[a,c]$ в том же смысле и*

$$\int\limits_a^c f\,dx = \int\limits_a^b f\,dx + \int\limits_b^c f\,dx.$$

▼ Обозначим для краткости $I^1=\int\limits_a^b f\,dx$, $I^2=\int\limits_b^c f\,dx$. Рассмотрим сначала случай интеграла Римана. Возьмем произвольное $\varepsilon>0$ и найдем такое $\delta_1>0$, что для любого разбиения $\mathbb{T}^1$ отрезка $[a,b]$ мельче δ_1 верно неравенство

$$\left|\mathfrak{S}(f,\mathbb{T}^1)-I^1\right|<\varepsilon;$$

потом найдем такое $\delta_2>0$, что для любого разбиения $\mathbb{T}^2$ отрезка $[b,c]$ мельче δ_2 верно неравенство

$$\left|\mathfrak{S}(f,\mathbb{T}^2)-I^2\right|<\varepsilon.$$

Возьмем такое $\delta > 0$, что $\delta < \delta_1$, $\delta < \delta_2$ и $\sup\limits_{[a,c]}|f| \cdot \delta < \varepsilon$ (конечность $\sup\limits_{[a,b]}|f|$ и $\sup\limits_{[b,c]}|f|$, а значит, и $\sup\limits_{[a,c]}|f|$, следует из теоремы 1) и пусть $\mathbb{T}$ — произвольное разбиение отрезка $[a,c]$ мельче δ. Если b не является внутренней точкой ни одного отрезка Δ_i разбиения $\mathbb{T}$ (то есть это конец одного отрезка разбиения и начало какого-то другого), то положим $\mathbb{T}^1 = \{(\Delta_i, \xi_i) \in \mathbb{T} : \ \Delta_i \subset [a,b]\}$, $\mathbb{T}^2 = \{(\Delta_i, \xi_i) \in \mathbb{T} : \ \Delta_i \subset [b,c]\}$, это разбиения $[a,b]$ и $[b,c]$ соответственно и $|\mathfrak{S}(f,\mathbb{T}^1) - I^1| < \varepsilon$, $|\mathfrak{S}(f,\mathbb{T}^2) - I^2| < \varepsilon$. А так как $\mathfrak{S}(f,\mathbb{T}) = \mathfrak{S}(f,\mathbb{T}^1) + \mathfrak{S}(f,\mathbb{T}^2)$, то

$$\left|\mathfrak{S}(f,\mathbb{T}) - (I^1 + I^2)\right| \leqslant \left|\mathfrak{S}(f,\mathbb{T}^1) - I^1\right| + \left|\mathfrak{S}(f,\mathbb{T}^2) - I^2\right| < 2\varepsilon.$$

Если же b является внутренней точкой какого-то отрезка $\Delta_j = [a_{j-1}, a_j]$ разбиения $\mathbb{T}$, то перейдем от разбиения $\mathbb{T}$ к разбиению $\overline{\mathbb{T}}$, заменив пару $([a_{j-1}, a_j], \xi_j)$ двумя парами $([a_{j-1}, b], b)$ и $([b, a_j], b)$. При этом

$$\left|\mathfrak{S}(f,\mathbb{T}) - \mathfrak{S}(f,\overline{\mathbb{T}})\right| = |f(\xi_i)(a_j - a_{j-1}) - f(b)(b - a_{j-1}) - f(b)(a_j - b)| \leqslant$$
$$\leqslant \sup_{[a,b]}|f|\left((a_j - a_{j-1}) + (b - a_{j-1}) + (a_j - b)\right) = 2\sup_{[a,b]}|f|(a_j - a_{j-1}) < 2\sup_{[a,b]}|f| \cdot \delta < 2\varepsilon.$$

А разбиение $\overline{\mathbb{T}}$ входит в ранее рассмотренный случай и, значит, $|\mathfrak{S}(f,\overline{\mathbb{T}}) - (I^1 + I^2)| < 2\varepsilon$. В итоге всегда справедлива оценка $|\mathfrak{S}(f,\mathbb{T}) - (I^1 + I^2)| < 4\varepsilon$.

Итак, для любого $\varepsilon > 0$ существует такое $\delta > 0$, что для любого отмеченного разбиения $\mathbb{T}$ отрезка $[a,b]$ мельче δ верно неравенство $|\mathfrak{S}(f,\mathbb{T}) - (I^1 + I^2)| < 4\varepsilon$. Следовательно, f интегрируема на $[a,c]$ по Риману и $\int\limits_a^c f\,dx = I^1 + I^2$.

Теперь рассмотрим случай интеграла Курцвейля–Хенстока. Возьмем произвольное $\varepsilon > 0$ и сначала найдем такой масштаб δ_1 на $[a,b]$, что для любого согласованного с ним разбиения $\mathbb{T}^1$ отрезка $[a,b]$ верно неравенство

$$\left|\mathfrak{S}(f,\mathbb{T}^1) - I_1\right| < \varepsilon,$$

а затем найдем такой масштаб δ_2 на $[b,c]$, что для любого согласованного с ним разбиения $\mathbb{T}^2$ отрезка $[b,c]$ верно неравенство

$$\left|\mathfrak{S}(f,\mathbb{T}^2) - I_2\right| < \varepsilon.$$

Положим

$$\delta(x) = \begin{cases} \min\{\delta_1(x), b - x\} & \text{при } x \in [a,b), \\ \min\{\delta_1(x), \delta_2(x)\} & \text{при } x = b, \\ \min\{\delta_2(x), x - b\} & \text{при } x \in (b,c]. \end{cases}$$

Если $\xi_i \in [a,b)$, то $|\Delta_i| < b - \xi_i$ и $b \notin \Delta_i$, а если $\xi_i \in (b,c]$, то $|\Delta_i| < \xi_i - c$ и $b \notin \Delta_i$. Тогда для любого согласованного с δ разбиения $\mathbb{T}$ отрезка $[a,c]$, $\mathbb{T} = \{(\Delta_i, \xi_i)\}_{i=1}^n$, имеем, что если $\xi_i \neq b$, то $b \notin \Delta_i$. Поэтому среди пар (Δ_i, ξ_i) есть пара или две с отмеченной точкой $\xi_j = b$. В случае присутствия в разбиении $\mathbb{T}$ двух таких отрезков точка b является их общим концом. Тогда $\mathbb{T}^1 = \{(\Delta_i, \xi_i) \in \mathbb{T} : \ \Delta_i \subset [a,b]\}$ — разбиение отрезка $[a,b]$, согласованное с $\delta(x) \leqslant \delta_1(x)$ и $|\mathfrak{S}(f,\mathbb{T}^1) - I^1| < \varepsilon$, а $\mathbb{T}^2 = \{(\Delta_i, \xi_i) \in \mathbb{T} : \Delta_i \subset [b,c]\}$ — разбиение отрезка $[b,c]$, согласованное с $\delta(x) \leqslant \delta_2(x)$ и $|\mathfrak{S}(f,\mathbb{T}^2) - I^2| < \varepsilon$. А так как $\mathfrak{S}(f,\mathbb{T}) = \mathfrak{S}(f,\mathbb{T}^1) + \mathfrak{S}(f,\mathbb{T}^2)$, то

$$\left|\mathfrak{S}(f,\mathbb{T}) - (I^1 + I^2)\right| \leqslant \left|\mathfrak{S}(f,\mathbb{T}^1) - I^1\right| + \left|\mathfrak{S}(f,\mathbb{T}^2) - I^2\right| < 2\varepsilon.$$

Если же только один отрезок $\Delta_j = [a_{j-1}, a_j]$ из разбиения $\mathbb{T}$ содержит точку b, $b \in (a_{j-1}, a_j)$, $\xi_j = b$, где $1 \leqslant j \leqslant n$, то перейдем от разбиения $\mathbb{T}$ к разбиению $\overline{\mathbb{T}}$ заменив пару (Δ_j, b) на две пары $([a_{j-1}, b], b)$ и $([b, a_j], b)$. Очевидно,

$$\mathfrak{S}(f, \mathbb{T}) = \mathfrak{S}(f, \overline{\mathbb{T}}),$$

а $\overline{\mathbb{T}}$ — согласованное с масштабом δ разбиение отрезка $[a, c]$ уже рассмотренного типа и

$$\left|\mathfrak{S}(f, \overline{\mathbb{T}}) - (I^1 + I^2)\right| < 2\varepsilon.$$

В итоге получаем, что для любого $\varepsilon > 0$ существует такой масштаб δ на $[a, c]$, что для любого согласованного с ним разбиения $\mathbb{T}$ отрезка $[a, c]$ имеем: $|\mathfrak{S}(f, \mathbb{T}) - (I^1 + I^2)| < 2\varepsilon$. Следовательно, f интегрируема на $[a, c]$ в смысле Хенстока и $\int\limits_a^c f\,dx = I^1 + I^2$. ▲

Формула Ньютона–Лейбница

Теорема 3 (формула Ньютона–Лейбница). *Если функция f определена на отрезке $[a, b]$, F непрерывна на $[a, b]$ и $F' = f$ всюду на интервале (a, b), то f интегрируема на $[a, b]$ в смысле Курцвейля–Хенстока и верна формула Ньютона–Лейбница*

$$(\mathcal{H}) \int\limits_a^b f\,dx = F(b) - F(a).$$

▼ Возьмем произвольное $\varepsilon > 0$ и поставим ему в соответствие масштаб δ на $[a, b]$ следующим образом. Используя непрерывность F в точке a найдем такое $\delta(a) > 0$, что

$$|f(a)|\delta(a) < \frac{\varepsilon}{8} \quad \text{и} \quad \forall x,\ a < x < a + \delta(a),\ |F(x) - F(a)| < \frac{\varepsilon}{8}.$$

Используя непрерывность F в точке b найдем такие $\delta(b) > 0$, что

$$|f(b)|\delta(b) < \frac{\varepsilon}{8} \quad \text{и} \quad \forall x,\ b - \delta(b) < x < b,\ |F(x) - F(b)| < \frac{\varepsilon}{8}.$$

А для любой точки $x \in (a, b)$ подберем такое $\delta(x) > 0$, что для любого Δx,

$$\forall \Delta x,\ |\Delta x| < \delta(x),\ x + \Delta x \in [a, b]:$$
$$|F(x + \Delta x) - F(x) - f(x) \cdot \Delta x| < \frac{\varepsilon|\Delta x|}{2(b - a)}$$

(ведь $F(x + \Delta x) - F(x) = F'(x) \cdot \Delta x + o(\Delta x)$).

Возьмем любое согласованное с построенным масштабом δ на $[a, b]$ разбиение $\mathbb{T} = \{(\Delta_i, \xi_i)\}_{i=1}^n$, $\Delta_i = [a_{i-1}, a_i]$.

Тогдв если $\xi_1 = a$, то имеем неравенства

$$|f(a)| \cdot |\Delta_1| < \frac{\varepsilon}{8} \quad \text{и} \quad |F(a) - F(a_2)| < \frac{\varepsilon}{8},$$

из которых следует неравенство

$$|f(\xi_1) \cdot |\Delta_1| - (F(a_2) - F(a_1))| < 2\frac{\varepsilon}{8} = \frac{\varepsilon}{4}.$$

Аналогично, если $\xi_n = b$, то имеем неравенства

$$|f(b)| \cdot |\Delta_n| < \frac{\varepsilon}{8} \quad \text{и} \quad |F(b) - F(a_{n-1})| < \frac{\varepsilon}{8},$$

из которых следует неравенство

$$|f(\xi_n) \cdot |\Delta_n| - (F(a_n) - F(a_{n-1}))| < 2\frac{\varepsilon}{8} = \frac{\varepsilon}{4}.$$

Если $a < \xi_i < b$, $1 \leqslant i \leqslant n$, то $F'(\xi_i) = f(\xi_i)$ и $|F(a_i) - F(\xi_i) - f(\xi_i)(a_i - \xi_i)| < \frac{\varepsilon(a_i - \xi_i)}{2(b-a)}$, $|F(\xi_i) - F(a_{i-1}) - f(\xi_i)(\xi_i - a_{i-1})| < \frac{\varepsilon(\xi_i - a_{i-1})}{2(b-a)}$, откуда имеем неравенство $|F(a_i) - F(a_{i-1}) - f(\xi_i)(a_i - a_{i-1})| < \frac{\varepsilon(a_i - a_{i-1})}{2(b-a)}$, то есть неравенство

$$|f(\xi_i) \cdot |\Delta_i| - (F(a_i) - F(a_{i-1}))| < \frac{\varepsilon(a_i - a_{i-1})}{2(b-a)}. \tag{**}$$

Из неравенств (*) и (**) получаем, что

$$|\mathfrak{S}(f, \mathbb{T}) - (F(b) - F(a))| = \left| \sum_{i=1}^{n} f(\xi_i) \cdot |\Delta_i| - \sum_{i=1}^{n} (F(a_i) - F(a_{i-1})) \right| <$$
$$< \frac{\varepsilon}{4} + \sum_{a < \xi_i < b} |f(\xi_i) \cdot |\Delta_i| - (F(a_i) - F(a_{i-1}))| + \frac{\varepsilon}{4} \leqslant$$
$$\leqslant \frac{\varepsilon}{4} + \sum_{i=1}^{n} \frac{\varepsilon \cdot |\Delta_i|}{2(b-a)} + \frac{\varepsilon}{4} = \frac{\varepsilon}{4} + \frac{\varepsilon}{2} + \frac{\varepsilon}{4} = \varepsilon.$$

Значит, функция f интегрируема на отрезке $[a, b]$ в смысле Курцвейля–Хенстока и $\int\limits_a^b f\,dx = F(b) - F(a)$. ▲

Следствие 1 (формула Ньютона–Лейбница). *Если функция f интегрируема на отрезке $[a, b]$ в смысле Римана или Курцвейля–Хенстока, а F — ее точная первообразная на $[a, b]$, то*

$$\int\limits_a^b f\,dx = F(b) - F(a).$$

▼ Это непосредственное следствие теоремы 3. ▲

Следствие 2 (формула Ньютона–Лейбница). *Если функция f интегрируема на отрезке $[a, b]$ в смысле Римана или Курцвейля–Хенстока, а F — ее обобщеннвя первообразная на $[a, b]$, то*

$$\int\limits_a^b f\,dx = F(b) - F(a).$$

▼ Это непосредственное следствие теоремы 3 и аддитивности интегралов по отрезкам. ▲

Замечание. Исаак Ньютон определял интеграл как приращение первообразной, поэтому иногда такой интеграл называют интегралом Ньютона. Теорема 3 показывает, что интеграл Курцвейля–Хенстока более общий (сильнее), чем интеграл Ньютона, а следствие теоремы показывает, что интеграл Ньютона не противоречит интегралам Римана и Курцвейля–Хенстока.

Напомним, что промежутком называется любое подмножество $\mathbb{R}$, содержащее вместе с каждой парой точек и все точки, лежащие между ними.

Промежутками являются пустое множество $\varnothing$, одноточечное множество, интервалы (a, b), полуотрезки $[a, b)$ и $(a, b]$, отрезки $[a, b]$, лучи (a, ∞), $(-\infty, b)$, $[a, \infty)$, $(-\infty, b]$, числовая прямая $\mathbb{R}$. Перечисленными множествами исчерпываются все промежутки.

Следствие 3. *Если функции F_1 и F_2 непрерывны на промежутке I и конечные производные F_1' и F_2' равны всюду на I за исключением конечного множества (в точках которого не являются дифференцируемыми F_1 или F_2 или $F_1' \neq F_2'$), то разность $F_1 - F_2$ постоянна на I.*

▼ Действительно, возьмем произвольную функцию f на I совпадающую с F_1' всюду, где F_1' существует и конечна. Зафиксируем точку $a \in I$. Тогда по теореме 2 для $b > a$, $b \in I$,

$$F_1(b) - F_1(a) = (\mathcal{H}) \int\limits_a^b f\, dx = F_2(b) - F_2(a).$$

Значит, для любых точек $x_1, x_2 \in I$ имеем $F_1(x_2) - F_1(x_1) = F_2(x_2) - F_2(x_1)$, то есть $F_1(x_2) - F_2(x_2) = F_1(x_1) - F_2(x_1) = \mathrm{const}$ — постоянной. ▲

Лекция 7 (28.02.20)
 Верхняя мера множств.
 Множества меры ноль.
 Интегрируемость по Риману
 ограниченных непрерывных почти всюду функций

Верхняя мера множества

Определение 1. Верхней (внешней) мерой Лебега или просто **верхней (внешней) мерой** множества $E \subset \mathbb{R}$ называется величина

$$\mu^* E = \inf_{E \subset \bigcup\limits_i l_i} \sum_i |l_i|$$

— точная нижняя грань сумм длин интервалов l_i, $i \in \mathbb{N}$, покрывающих E (множество интервалов не более чем счетно).

Очевидно, что $0 \leqslant \mu^* E \leqslant +\infty$.

Замечание. *В определении вместо интервалов могут фигурировать отрезки — получится эквивалентное определение.*

▼ Действительно, если система интервалов $\{l_i\}$ такова, что $E \subset \bigcup\limits_i l_i$, то замкнув интервалы, то есть добавив к ним концы, получим такую систему отрезков $\{\bar{l}_i\}$, что $E \subset \bigcup\limits_i \bar{l}_i$. Так как $|\bar{l}_i| = |l_i|$, то при замене в определении интервалов на отрезки внешняя мера множества не может увеличиться. А если система отрезков $\{h_i\}$ такова, что $E \subset \bigcup\limits_i h_i$, то взяв любое $\varepsilon > 0$ и подобрав для каждого отрезка h_i такой интервал l_i, что $h_i \subset l_i$ и $|l_i| \leqslant |h_i| + \varepsilon 2^{-i}$, получим систему интервалов $\{l_i\}$, $E \subset \bigcup\limits_i l_i$ и

$\sum\limits_i |l_i| \leqslant \sum\limits_i |h_i| + \varepsilon$. В силу произвольности $\varepsilon > 0$ при переходе от определения с отрезками к определению с интервалами внешняя мера множества увеличиться не может. Значит, определение с интервалами эквивалентно определению с отрезками. ▲

Внешняя мера обладает следующим свойством.

Свойство внешней меры. *Если $E \subset \bigcup\limits_j E^j$, где $\{E^j\}$, $j \in \mathbb{N}$, — не более чем счетная система множеств, то $\mu^* E \leqslant \sum\limits_j \mu^* E^j$.*

▼ Возьмем произвольное $\varepsilon > 0$ и для каждого множества E^j построим такую не более чем счетную систему интервалов $\{l_i^j\}_i$ покрывающую E^j, что $\sum\limits_i \left|l_i^j\right| \leqslant \mu^* E^j + 2^{-j}\varepsilon$. Тогда не более чем счетная система интервалов $\{l_i^j\}_{j,i}$ покрывает $\bigcup\limits_i E^j$, а значит и E, и $\sum\limits_{j,i} \left|l_i^j\right| = \sum\limits_j \sum\limits_i \left|l_i^j\right| \leqslant \sum\limits_j (\mu^* E^j + 2^{-j}\varepsilon) = \sum\limits_j \mu^* E^j + \sum\limits_j 2^{-j}\varepsilon \leqslant \sum\limits_j \mu^* E^j + \varepsilon$. Значит, $\mu^* E \leqslant \sum\limits_j \mu^* E^j + \varepsilon$. В силу произвольности $\varepsilon > 0$ свойство доказано. ▲

Определение 2. Множество $E \subset \mathbb{R}$ **меры нуль по Лебегу**, если $\mu^* E = 0$.

Из определения вытекает следующее свойство.

Свойство 1 (множеств меры нуль). *Множество $E \subset \mathbb{R}$ является множеством меры нуль по Лебегу тогда и только тогда, когда для любого $\varepsilon > 0$ найдётся не более чем счетная система интервалов $\{l_i\}$, $\bigcup\limits_i l_i \supset E$, $\sum\limits_i |l_i| < \varepsilon$.*

Следствиями свойств внешней меры являются следующие свойства множеств меры нуль по Лебегу.

Свойство 2 (множеств меры нуль). *Подмножество множества меры нуль по Лебегу — множество меры нуль по Лебегу.*

Свойство 3 (множеств меры нуль). *Не более чем счетное объединение множеств меры нуль по Лебегу — множество меры нуль по Лебегу.*

Следствие. *Не более чем счетное множество имеет меру нуль по Лебегу.*

Это сразу следует из предыдущего свойства и того факта, что одноточечное множество имеет меру нуль по Лебегу.

Определение 3. Если какое-либо свойство выполняется для всех точек множества E, кроме подмножества E нулевой меры по Лебегу, то говорят, что оно выполняется **почти всюду** на E.

Колебание и его свойства

Определение 4. Если функция f определена на множестве E, то ее **колебанием** (**осцилляцией**) на E называется величина

$$\operatorname*{osc}_E f = \sup_{x,y \in E} |f(x) - f(y)|.$$

Очевидно $0 \leqslant \operatorname*{osc}\limits_E f \leqslant +\infty$.

Лемма 1. *Если функция f на множестве E действительнозначна, то*

$$\operatorname*{osc}_E f = \sup_E f - \inf_E f.$$

▼ Если точки $x, y \in E$, то значения функции $f(x), f(y) \in [\inf\limits_E f, \sup\limits_E f] \subset \mathbb{R}$ и $|f(x) - f(y)| \leqslant \sup\limits_E f - \inf\limits_E f$; значит $\operatorname*{osc}\limits_E f \leqslant \sup\limits_E f - \inf\limits_E f$. Если $\sup\limits_E f = +\infty$ или $\inf\limits_E f = -\infty$, то для любого $y \in E$ $\sup\limits_{x\in E} |f(x) - f(y)| = +\infty$ и, значит, $\operatorname*{osc}\limits_E f = +\infty = \sup\limits_E f - \inf\limits_E f$. Если $\sup\limits_E f$ и $\inf\limits_E f$ конечны, то для любого $\varepsilon > 0$ найдутся такие точки $x, y \in E$, что $f(x) > \sup\limits_E f - \varepsilon$, $f(y) < \inf\limits_E f + \varepsilon$; тогда $f(x) - f(y) > \sup\limits_E f - \inf\limits_E f - 2\varepsilon$ и, в силу произвольности $\varepsilon > 0$, $\operatorname*{osc}\limits_E f \geqslant \sup\limits_E f - \inf\limits_E f$. Значит, $\operatorname*{osc}\limits_E f = \sup\limits_E f - \inf\limits_E f$. ▲

В дальнейшем нам понадобится следующая лемма.

Лемма 2 (о разбиении подотрезка). *Пусть $T = \{\Delta_i\}_{i=1}^n$ — разбиение отрезка $[a, b]$. Тогда для любого отрезка $I = [c, d] \subset [a, b]$ невырожденные отрезки $I \cap \Delta_i$ образуют разбиение отрезка I и верно равенство*

$$|I| = d - c = \sum_{i=1}^{n} |I \cap \Delta_i|.$$

▼ Можно считать, что отрезки $\Delta_i = [a_{i-1}, a_i]$ занумерованы в порядке их расположения на $\mathbb{R}$, $a = a_0 < a_1 < \dots < a_n = b$, и $a_{k-1} \leqslant c < a_k$, $a_{m-1} < d \leqslant a_m$, $k \leqslant m$. Тогда отрезки $[c, a_k]$, Δ_i, $k < i < m$, и $[a_{m-1}, d]$ образуют разбиение отрезка $I = [c, d]$ и $d - c = d - a_{m-1} + \sum\limits_{i=k+1}^{m-1} (a_i - a_{i-1}) + a_k - c = |[c, d] \cap \Delta_m| + \sum\limits_{i=k+1}^{m-1} |[c, d] \cap \Delta_i| + |[c, d] \cap \Delta_k| = \sum\limits_{i=1}^{n} |[c, d] \cap \Delta_i|$. ▼

Теорема 1. *Если функция f на отрезке $[a, b]$ ограничена и непрерывна почти всюду, то f интегрируема на $[a, b]$ в смысле Римана.*

▼ Возьмем произвольное $\varepsilon > 0$ и выберем (в соответствии со свойством множеств меры нуль) не более чем счетную систему интервалов $\{l_i\}$, которые покрывают все точки разрыва f на $[a, b]$ и $\sum\limits_i |l_i| < \varepsilon$. Для каждой точки x непрерывности f на $[a, b]$ найдем такое $\delta(x) > 0$, что $f\big(B_{3\delta(x)}(x) \cap [a, b]\big) \subset B_\varepsilon(f(x))$ и, значит, $\operatorname*{osc}\limits_{B_{3\delta(x)}(x)\cap[a,b]} f \leqslant 2\varepsilon$. Система интервалов $\{l_i, B_\delta(x)\}$ покрывает $[a, b]$, выделим из нее конечное подпокрытие $l_{i_1}, \dots, l_{i_p}, B_{\delta(x_1)}(x_1), \dots, B_{\delta(x_q)}(x_q)$. Выберем такое $\delta > 0$, что $\delta < \min\limits_{1\leqslant k\leqslant q} \delta(x_k)$ и $2p\delta < \varepsilon$. Пусть $\mathbb{T}$, где $\mathbb{T} = \{(\Delta_i, \xi_i)\}_{i=1}^n$, и $\mathbb{T}'$, где $\mathbb{T}' = \{(\Delta'_j, \xi'_j)\}_{j=1}^{n'}$, — два произвольных отмеченных разбиения отрезка $[a, b]$ из $\boldsymbol{B}_\delta$, то есть с $\xi_i \in \Delta_i$ и $|\Delta_i| < \delta$ для всех i, с $\xi'_j \in \Delta'_j$ и $|\Delta'_i| < \delta$ для всех j (с $\Delta_i \subset B_\delta(\xi_i)$ для всех i, с $\Delta'_j \subset B_\delta(\xi'_j)$ для всех j). Оценим разность

$$\mathfrak{S}(f, \mathbb{T}) - \mathfrak{S}(f, \mathbb{T}') = \sum_{i=1}^{n} f(\xi_i)|\Delta_i| - \sum_{j=1}^{n'} f(\xi'_j)|\Delta'_j| =$$

$$= \sum_{i=1}^{n}\sum_{j=1}^{n'} f(\xi_i)\left|\Delta_i \cap \Delta'_j\right| - \sum_{j=1}^{n'}\sum_{i=1}^{n} f(\xi'_i)\left|\Delta_i \cap \Delta'_j\right| = \sum_{i=1}^{n}\sum_{j=1}^{n'} \big(f(\xi_i) - f(\xi'_j)\big)\left|\Delta_i \cap \Delta'_j\right|.$$

Если $\xi_i \in l_{i_k} = (\alpha, \beta)$, $1 \leqslant k \leqslant p$, то $\Delta_i \subset B_\delta(\xi_i) \subset (\alpha - \delta, \beta + \delta)$ — интервалу l_{i_k}, увеличенному на δ в обе стороны. Поэтому

$$\sum_{\xi_i \in \bigcup\limits_{k=1}^{p} l_{i_k}} |\Delta_i| \leqslant \sum_{k=1}^{p} (|l_{i_k}| + 2\delta) \leqslant \sum_i |l_i| + 2p\delta < 2\varepsilon.$$

Значит, используя лемму о разбиении подотрезка, имеем оценку

$$\sum_{\xi_i \in \bigcup\limits_{k=1}^{p} l_{i_k}} \sum_{j=1}^{n'} \left| f(\xi_i) - f(\xi'_j) \right| \cdot \left| \Delta_i \cap \Delta'_j \right| \leqslant$$

$$\leqslant 2 \sup_{[a,b]} |f| \cdot \sum_{\xi_i \in \bigcup\limits_{k=1}^{p} l_{i_k}} |\Delta_i| < 4 \sup_{[a,b]} |f| \cdot \varepsilon.$$

Если $\xi_i \in B_{\delta(x_k)}(x_k)$, то $\Delta_i \subset B_\delta(\xi_i) \subset B_{2\delta(x_k)}(x_k)$; так как $\Delta'_j \subset B_\delta(\xi'_j)$, то если $\Delta_i \cap \Delta'_j \neq \varnothing$, то $\xi'_j \in B_{3\delta(x_k)}(x_k)$ и, значит, $\left| f(\xi_i) - f(\xi'_j) \right| \leqslant \operatorname*{osc}\limits_{B_{3\delta(x)}(x) \cap [a,b]} f \leqslant 2\varepsilon$. Отсюда имеем оценку

$$\sum_{\xi_i \in \bigcup\limits_{k=1}^{q} B_{\delta(x_k)}(x_k)} \sum_{j=1}^{n'} \left| f(\xi_i) - f(\xi'_j) \right| \cdot \left| \Delta_i \cap \Delta'_j \right| \leqslant$$

$$\leqslant \sum_{i=1}^{n} \sum_{j=1}^{n'} 2\varepsilon \cdot \left| \Delta_i \cap \Delta'_j \right| = 2\varepsilon \sum_{i=1}^{n} |\Delta_i| = 2\varepsilon(b-a)$$

(согласно лемме о разбиении подотрезка).

Значит,

$$|\mathfrak{S}(f, \mathbb{T}) - \mathfrak{S}(f, \mathbb{T}')| < 4 \sup_{[a,b]} |f| \cdot \varepsilon + 2(b-a)\varepsilon =$$

$$= \left(4 \sup_{[a,b]} |f| + 2(b-a) \right) \varepsilon = C(f, [a,b]) \cdot \varepsilon,$$

то есть выполнен критерий Коши интегрируемости f на $[a,b]$ по Риману. ▲

Следствие 1. *Непрерывная на отрезке $[a,b]$ функция f интегрируема на $[a,b]$ в смысле Римана.*

▼ Это следует из доказанной теоремы. ▲

Следствие 2. *Если функция f ограничена и имеет не более чем счетное множество точек разрыва на $[a,b]$, то f интегрируема на $[a,b]$ в смысле Римана.*

▼ Это также следует из доказанной теоремы. ▲

Следствие 3. *Если функция f монотонна на $[a,b]$, то f интегрируема на $[a,b]$ в смысле Римана.*

▼ Действительно, f принимает наибольшее и наименьшее значения в концах отрезка и, значит, ограничена, множество точек разрыва f на $[a,b]$ не более чем счетно по теореме о разрывах монотонной функции. ▲

Лекция 8 (03.03.20)
Критерий интегрируемости Лебега. Дополнительные свойства интеграла Римана

Колебание и его свойства

Определение 1. Если функция f определена на множестве E, то ее **колебанием** (**осцилляцией**) на E называется величина

$$\operatorname*{osc}_E f = \sup_{x,y\in E} |f(x) - f(y)|.$$

Очевидно $0 \leqslant \operatorname*{osc}_E f \leqslant +\infty$.

Лемма 1. *Если функция f на множестве E действительнозначна, то*

$$\operatorname*{osc}_E f = \sup_E f - \inf_E f.$$

▼ Если точки $x, y \in E$, то значения функции $f(x), f(y) \in [\inf_E f, \sup_E f] \subset \mathbb{R}$ и $|f(x) - f(y)| \leqslant \sup_E f - \inf_E f$; значит $\operatorname*{osc}_E f \leqslant \sup_E f - \inf_E f$. Если $\sup_E f = +\infty$ или $\inf_E f = -\infty$, то для любого $y \in E$ $\sup_{x\in E} |f(x) - f(y)| = +\infty$ и, значит, $\operatorname*{osc}_E f = +\infty = \sup_E f - \inf_E f$. Если $\sup_E f$ и $\inf_E f$ конечны, то для любого $\varepsilon > 0$ найдутся такие точки $x, y \in E$, что $f(x) > \sup_E f - \varepsilon$, $f(y) < \inf_E f + \varepsilon$; тогда $f(x) - f(y) > \sup_E f - \inf_E f - 2\varepsilon$ и, в силу произвольности $\varepsilon > 0$, $\operatorname*{osc}_E f \geqslant \sup_E f - \inf_E f$. Значит, $\operatorname*{osc}_E f = \sup_E f - \inf_E f$. ▲

Лемма 1. *Если действительнозначная функция f определена на $[a, b]$ и $\mathbb{T} = \{\Delta_i\}_{i=1}^n$ — разбиение $[a, b]$, то*

$$\operatorname*{osc}_{\{\boldsymbol{\xi}:\, \xi_i \in \Delta_i\}} \mathfrak{S}(f, \mathbb{T}) = \sum_{i=1}^n \operatorname*{osc}_{\Delta_i} f \cdot |\Delta_i|.$$

▼ $$\operatorname*{osc}_{\{\boldsymbol{\xi}:\, \xi_i \in \Delta_i\}} \mathfrak{S}(f, \mathbb{T}) = \sup_{\{\xi, \xi' :\, \xi_i, \xi'_i \in \Delta_i,\}} \left| \sum_{i=1}^n (f(\xi_i) - f(\xi'_i)) \cdot |\Delta_i| \right| \leqslant$$
$$\leqslant \sup_{\{\xi, \xi' :\, \xi_i, \xi'_i \in \Delta_i,\}} \sum_{i=1}^n |f(\xi_i) - f(\xi'_i)| \cdot |\Delta_i| \leqslant$$
$$\leqslant \sum_{i=1}^n \sup_{\xi_i, \xi'_i \in \Delta_i} |f(\xi_i) - f(\xi'_i)| \cdot |\Delta_i| = \sum_{i=1}^n \operatorname*{osc}_{\Delta_i} f \cdot |\Delta_i|.$$

Теперь докажем противоположное неравенство. Если для некоторого j, $1 \leqslant j \leqslant n$, $\operatorname*{osc}_{\Delta_j} f = +\infty$, то

$$\operatorname*{osc}_{\{\boldsymbol{\xi}:\, \xi_i \in \Delta_i\}} \mathfrak{S}(f, \mathbb{T}) = \sup_{\{\xi, \xi' :\, \xi_i, \xi'_i \in \Delta_i,\}} \sum_{i=1}^n |f(\xi_i) - f(\xi'_i)| \cdot |\Delta_i| \geqslant$$
$$\geqslant \sup_{\xi_j, \xi'_j \in \Delta_j} |f(\xi_j) - f(\xi'_j)| \cdot |\Delta_j| = \operatorname*{osc}_{\Delta_j} f \cdot |\Delta_j| = +\infty.$$

Если для всех i $\operatorname*{osc}_{\Delta_i} f < +\infty$, то для любого $\varepsilon > 0$ и любого i, $1 \leqslant i \leqslant n$, найдутся такие $\xi_i, \xi_i' \in \Delta_i$, что $f(\xi_i) - f(\xi_i') > \operatorname*{osc}_{\Delta_i} f - \varepsilon$. Но тогда

$$\sum_{i=1}^{n} (f(\xi_i) - f(\xi_i')) \cdot |\Delta_i| > \sum_{i=1}^{n} \operatorname*{osc}_{\Delta_i} f \cdot |\Delta_i| -$$

$$-\varepsilon \sum_{i=1}^{n} |\Delta_i| = \sum_{i=1}^{n} \operatorname*{osc}_{\Delta_i} f \cdot |\Delta_i| - \varepsilon(b-a)$$

и, в силу произвольности $\varepsilon > 0$,

$$\operatorname*{osc}_{\{\boldsymbol{\xi}\colon \xi_i \in \Delta_i\}} \mathfrak{S}(f, \mathbb{T}) \geqslant \sum_{i=1}^{n} \operatorname*{osc}_{\Delta_i} f \cdot |\Delta_i|. \blacktriangle$$

О непрерывности почти всюду интегрируемых по Риману функций

Теперь докажем еще одну лемму.

Лемма 1. *Если действительнозначная функция f интегрируема по Риману на отрезке $[a, b]$, то для любого $\varepsilon > 0$ найдется такое разбиение $T = \{\Delta_i\}_{i=1}^n$ отрезка $[a, b]$, что $\sum\limits_{i=1}^{n} \operatorname*{osc}_{\Delta_i} f \cdot |\Delta_i| < \varepsilon$.*

$\blacktriangledown$ Для данного $\varepsilon > 0$ найдем такое $\delta > 0$, что для любого разбиения $\mathbb{T} = \{(\Delta_i, \xi_i)\}_{i=0}^n$ мельче δ выполнено неравенство $|\mathfrak{S}(f, \mathbb{T}) - I| < \frac{\varepsilon}{3}$, где $I = \int\limits_a^b f\,dx$. Взяв любое разбиение $T = \{\Delta_i\}_{i=1}^n$ мельче δ будем иметь в силу написанного неравенства $\operatorname*{osc}_{\{\boldsymbol{\xi}\colon \xi_i \in \Delta_i\}} \mathfrak{S}(f, \mathbb{T}) \leqslant \frac{2\varepsilon}{3} < \varepsilon$. Отсюда и из равенства предыдущей леммы $\operatorname*{osc}_{\{\boldsymbol{\xi}\colon \xi_i \in \Delta_i\}} \mathfrak{S}(f, \mathbb{T}) = \sum\limits_{i=1}^{n} \operatorname*{osc}_{\Delta_i} f \cdot |\Delta_i|$ получаем доказываемое утверждение. $\blacktriangle$

Теорема 1. *Если функция f интегрируема по Риману на отрезке $[a, b]$, то f на $[a, b]$ ограничена и непрерывна почти всюду.*

$\blacktriangledown$ Ранее было доказано, что необходимым условием интегрируемости по Риману функции f на отрезке $[a, b]$ является ограниченность f на $[a, b]$. Поэтому остается только доказать, что f непрерывна почти всюду на $[a, b]$. Рассмотрим случай действительнозначной функции f. Возьмем произвольное $\varepsilon > 0$. Для каждого $j \in \mathbb{N}$ найдем такое разбиение $T^j = \{\Delta_i^j\}_{i=1}^{n_j}$, что

$$\sum_{i=1}^{n_j} \operatorname*{osc}_{\Delta_i^j} f \cdot |\Delta_i^j| < \varepsilon 2^{-2j}.$$

Для каждого $j \in \mathbb{N}$ из отрезков Δ_i^j разбиения T^j выберем те, для которых $\operatorname*{osc}_{\Delta_i^j} f \geqslant 2^{-j}$. Сумма их длин строго меньше $\varepsilon 2^{-j}$, так как $\sum\limits_{i\,\text{выбр.}} \operatorname*{osc}_{\Delta_i^j} f \cdot |\Delta_i^j| \geqslant 2^{-j} \sum\limits_{i\,\text{выбр.}} |\Delta_i^j|$ и, значит, $\sum\limits_{i\,\text{выбр.}} |\Delta_i^j| < \varepsilon 2^{-j}$. Если x не принадлежит ни одному из выбранных отрезков, то существует такая $B_\delta(x)$, что для любого $t \in B_\delta(x) \cap [a, b]$ верно неравенство $|f(t) -$

$f(x)| < 2^{-j}$. Все выбранные для $j = 1, 2, \dots$ отрезки образуют систему отрезков, сумма длин которых строго меньше $\sum\limits_{j=1}^{\infty} \varepsilon 2^{-j} = \varepsilon$. А если x не принадлежит выбранной системе, то для любого $\varepsilon > 0$ найдется такое $j \in \mathbb{N}$, что $2^{-j} < \varepsilon$. Далее, найдется такая $B_\delta(x)$, что для любого $t \in B_\delta(x) \cap [a, b]$ верно неравенство $|f(t) - f(x)| < 2^{-j} < \varepsilon$. Значит, x — точка непрерывности f на $[a, b]$. Все точки разрыва принадлежат системе выбранных отрезков, сумма длин которых строго меньше ε, следовательно f непрерывна на $[a, b]$ почти всюду.

Следствие (критерий Лебега $\mathcal{R}$-интегрируемости). *Функция f интегрируема на отрезке $[a, b]$ в смысле Римана тогда и только тогда, когда на $[a, b]$ f ограничена и непрерывна почти всюду.*

▼ Это следует из доказанной теоремы 2 и теоремы 1. ▲

Из критерия Лебега интегрируемости по Риману легко следует ряд свойств интеграла Римана. Например, очевидно, что если f интегрируема на $[a, b]$, то она интегрируема по Риману и на любом $[\overline{a}, \overline{b}] \subset [a, b]$, или, если f интегрируема по Риману на $[a, b]$ и на $[b, c]$, то она интегрируема по Риману и на $[a, c]$. Но эти свойства были установлены нами ранее. Сейчас мы установим ряд свойств интеграла Римана, которые ранее не устанавливались и которыми в большинстве своем не обладает интеграл Курцвейля–Хенстока.

Обозначения. Будем обозначать класс интегрируемых по Риману на отрезке $[a, b]$ функций $\mathcal{R}[a, b]$, класс интегрируемых по Курцвейлю–Хенстоку на отрезке $[a, b]$ функций $\mathcal{H}[a, b]$.

Свойство 1. *Если функции f и g из $\mathcal{R}[a, b]$, то $f \cdot g \in \mathcal{R}[a, b]$.*

Свойство 2. *Если функция $f \in \mathcal{R}[a, b]$, функция φ непрерывна и ограничена на множестве $f([a, b])$, то $\varphi(f) \in \mathcal{R}[a, b]$.*

Свойство 3. *Если функцию $f \in \mathcal{R}[a, b]$ изменить в конечном числе точек, то полученная после такого изменения функция $\tilde{f} \in \mathcal{R}[a, b]$ и интеграл от нее совпадает с интегралом от начальной функции.*

Свойство 4. *Если $f \in \mathcal{R}[a, b]$, $a < b$, то $|f| \in \mathcal{R}[a, b]$ и*

$$\left| \int\limits_a^b f\, dx \right| \leqslant \int\limits_a^b |f|\, dx.$$

Свойство 5. *Если $f \in \mathcal{R}[a, b]$, $a < b$, $f(x) \geqslant 0$ на $[a, b]$ и в точке x_0 непрерывности f по $[a, b]$ значение $f(x_0) > 0$, то верно неравенство $\int\limits_a^b f\, dx > 0$.*

▼ Докажем перечисленные свойства.

1. Если f и g ограничены и непрерывны почти всюду на $[a, b]$, то и $f \cdot g$ ограничена и непрерывна почти всюду на $[a, b]$, а значит, интегрируема по Риману на $[a, b]$.

2. Из ограниченности φ на $f([a, b])$ следует ограниченность $\varphi(f)$ на $[a, b]$, а множество точек разрыва $\varphi(f)$ на $[a, b]$ — подмножество множества точек разрыва f на $[a, b]$. Значит, по критерию Лебега $\varphi(f) \in \mathcal{R}[a, b]$.

3. Разность $\tilde{f} - f$ отлична от нуля в конечном числе точек, значит $\tilde{f} - f \in \mathcal{R}[a, b]$ и $\int\limits_a^b \tilde{f} - f\, dx = 0$. Используя линейность по функциям интеграла Римана получаем,

что $\tilde f = (\tilde f - f) + f \in \mathcal{R}[a,b]$ и $\int\limits_a^b \tilde f\,dx = \int\limits_a^b (\tilde f - f)\,dx + \int\limits_a^b f\,dx = \int\limits_a^b f\,dx$.

4. Так как множество точек разрыва $|f|$ на $[a,b]$ — подмножество множества точек разрыва f на $[a,b]$, то $|f| \in \mathcal{R}[a,b]$. Так как для любого отмеченного разбиения $\mathbb{T} = \{(\Delta_i, \xi_i)\}$ верно неравенство

$$|\mathfrak{S}(f,\mathbb{T})| = \left|\sum_i f(\xi_i)|\Delta_i|\right| \leqslant \sum_i |f(\xi_i)| \cdot |\Delta_i| = \mathfrak{S}(|f|,\mathbb{T}),$$

то, в силу сохранения неравенств при переходе к пределу по базе, имеем

$$\left|\int\limits_a^b f\,dx\right| \leqslant \int\limits_a^b |f|\,dx.$$

5. Так как f непрерывна в $x_0 \in [a,b]$, то найдётся $B_\delta(x_0)$, такая что для любого $x \in B_\delta(x_0) \cap [a,b]$: $f(x) \geqslant \frac{f(x_0)}{2} > 0$. Положим

$$g(x) = \begin{cases} \frac{f(x_0)}{2} & \text{на } B_\delta(x_0) \cap [a,b], \\ 0 & \text{в остальных точках.} \end{cases}$$

Тогда $g(x) \leqslant f(x)$ на $[a,b]$ и $\int\limits_a^b g\,dx = \frac{f(x_0)}{2} \cdot |B_\delta(x_0) \cap [a,b]| > 0$, что, в силу сохранения неравенств при интегрировании, влечет свойство 5. ▲

Лекция 9 (06.03.20)
Измеримые функции
Интегрирование по Курцвейлю-Хенстоку ограниченных измеримых функций

Содержание этой прочитанной лекции не включено в программу экзамена.

Измеримые функции

Дадим два эквивалентных определения измеримости функции на отрезке.

Определение 1. Функцию f называют **измеримой** на отрезке $[a,b]$, если f определена почти всюду на $[a,b]$ и для любого $\varepsilon > 0$ найдется такая функция $g \in \mathrm{C}[a,b]$ (пространству непрерывных на $[a,b]$ функций), что $\mu^*\{x \in [a,b] : f(x) \neq g(x)\} < \varepsilon$.

Определение 2. Функцию f называют **измеримой** на отрезке $[a,b]$, если f определена почти всюду на $[a,b]$ и для любого $\varepsilon > 0$ найдется такое множество $E \subset [a,b]$, $\mu^* E < \varepsilon$, что f непрерывна на $[a,b] \setminus E$ (относительно $[a,b] \setminus E$).

Для доказательства эквивалентности этих определений нам понадобится ряд свойств.

Напомним, что множество $G \subset \mathbb{R}$ называют открытым, если любая точка $x \in G$ имеет окрестность $B_\delta(x) \subset G$.

Лемма 2. *Любое открытое множество $G \subset \mathbb{R}$ является объединением не более чем счетного набора попарно непересекающихся интервалов (α_i, β_i), ограниченных или неограниченных, концы которых не принадлежат G.*

▼ Очевидно утверждение верно, хотя и бессодержательно, для пустого множества $\varnothing$ — оно является объединением пустого набора интервалов. Если $x \in G$, то пусть $\alpha = \inf\{t \in \mathbb{R} : [t, x] \subset G\}$, $\beta = \sup\{y \in \mathbb{R} : [x, y] \subset G\}$. Если $\alpha < z < x$, то найдется такое t, $\alpha \leqslant t < z$, что $[t, x] \subset G$, но тогда $z \in [t, x] \subset G$. Если $x < z < \beta$, то найдется такое y, $z < y \leqslant \beta$, что $[x, y] \subset G$, но тогда $z \in [x, y] \subset G$. Следовательно $(\alpha, \beta) \subset G$. Если $\alpha \in G$, то, так как G — открытое множество, есть такое $t < \alpha$, что $[t, \alpha]$, а значит, и $[t, x]$, входит в G, что противоречит определению α. Значит, $\alpha \notin G$. Аналогично и $\beta \notin G$. Так как $(\alpha, \beta) \subset G$, $\alpha, \beta \notin G$, то построенные для различных точек x интервалы или совпадают или не пересекаются. Таким образом можно построить набор попарно непересекающихся интервалов, ограниченных или неограниченных, концы которых не принадлежат G, а их объединение совпадает с G. Но на прямой $\mathbb{R}$ любой набор попарно непересекающихся интервалов не более чем счетен — ведь каждому такому интервалу можно сопоставить лежащее в нем рациональное число, тогда такой набор будет эквивалентен некоторому подмножеству счетного множества рациональных чисел и, следовательно, сам не более чем счетен. ▲

Определение 3. Если G непустое открытое подмножество $\mathbb{R}$, то интервалы $(\alpha_i, \beta_i) \subset G$, $\alpha_i, \beta_i \notin G$, объединение которых совпадает с G, называют **составляющими** интервалами G.

Напомним, что множество $F \subset \mathbb{R}$ называют замкнутым, если $\mathbb{R} \setminus F$ открытое множество.

Определение 4. Если F замкнутое собственное подмножество $\mathbb{R}$ (т.е. $F \subset \mathbb{R}$, $F \neq \mathbb{R}$), то составляющие интервалы множества $\mathbb{R} \setminus F$ называют **смежными** или **дополнительными** интервалами F.

Лемма 1. *Пусть на непустом замкнутом множестве $F \subset \mathbb{R}$ определена непрерывная на нем функция f. Если $F \neq \mathbb{R}$, то доопределив f на каждом конечном смежном интервале (α_i, β_i) множества F линейным образом так, что доопределенная f линейна на $[\alpha_i, \beta_i]$, а на любом бесконечном смежном интервале (если они есть) постоянной, совпадающей со значением f в конце такого интервала, получим непрерывную функцию на $\mathbb{R}$.*

▼ Доопределенная функция линейна, а значит, непрерывна, на всех смежных интервалах F. Поэтому надо только доказать, что f сохранила непрерывность на F. Если $x \in F$ и $x = \alpha_i$ — левому концу некоторого смежного интервала F, то, в силу линейности f на $[\alpha_i, \beta_i]$, получаем, что предел справа $\lim\limits_{t \to x+0} f(t) = f(x)$. Если $x \in F$ и не является левым концом никакого смежного интервала F, то для любого $\gamma > 0$ верно, что $(x, x+\gamma) \cap F \neq \varnothing$. Так как f непрерывна на F, то для любого $\varepsilon > 0$ найдется такое $\delta > 0$, что для любого $t \in B_\delta(x) \cap F = (x-\delta, x+\delta) \cap F$ верно неравенство $|f(t) - f(x)| < \varepsilon$. Возьмем такое λ, $0 < \lambda < \delta$, что $x + \lambda \in F$. Тогда, в силу доопределения функции f, имеем, что $f[x, x+\lambda] \subset B_\varepsilon(f(x))$ и, значит, предел справа $\lim\limits_{t \to x+0} f(t) = f(x)$. Аналогично устанавливается, что в любом случае предел слева $\lim\limits_{t \to x-0} f(t) = f(x)$. Значит, $\lim\limits_{t \to x} f(t) = f(x)$, доопределенная функция непрерывна в точке x. ▲

Теорема 1. *Определения 1 и 2 измеримости функции на отрезке эквивалентны.*

▼ Из определения измеримости 1 следует определение измеримости 2. Действительно, пусть f измерима по определению 1. Тогда возьмем произвольное $\varepsilon > 0$ и найдем такую $g \in \mathrm{C}[a, b]$, что $\mu^*\{x \in [a, b] : f(x) \neq g(x)\} < \varepsilon$. Тогда, положив $E = \{x \in [a, b] : f(x) \neq g(x)\}$, видим, что $\mu^* E < \varepsilon$ и f, совпадающая на $[a, b] \setminus E$ с

$g \in \mathrm{C}[a,b]$, непрерывна на $[a,b] \setminus E$.

Теперь покажем, что из определения измеримости 2 следует определение измеримости 1. Возьмем произвольное $\varepsilon > 0$ и найдем такое множество E, $\mu^* E < \varepsilon$, что f непрерывна на $[a,b] \setminus E$. По определению верхней меры существует покрывающая E система интервалов $\{l_i\}$, $\sum\limits_i |l_i| < \varepsilon$. Пусть $F = [a,b] \setminus \bigcup\limits_i l_i$. Это замкнутое множество (как пересечение замкнутых множеств $[a,b]$ и $\mathbb{R} \setminus \bigcup\limits_i l_i$). Если F пусто, то $[a,b]$ покрыто системой интервалов $\{l_i\}$, для любой $g \in \mathrm{C}[a,b]$ $\{x \in [a,b] : f(x) \neq g(x)\} \subset [a,b] \subset \bigcup\limits_i l_i$ и, следовательно, $\mu^*\{x \in [a,b] : f(x) \neq g(x)\} < \sum\limits_i |l_i| < \varepsilon$. Если F непусто, то воспользуемся леммой 1 и построим такую $g \in \mathrm{C}[a,b]$, что $g = f$ на F. Тогда $\{x \in [a,b] : f(x) \neq g(x)\} \subset ([a,b] \setminus F)$ и, значит, $\mu^*\{x \in [a,b] : f(x) \neq g(x)\} < \mu^*\left([a,b] \setminus F\right) \leqslant \sum\limits_i |l_i| < \varepsilon$. ▲

Класс измеримых функций широко используется в математике. Его свойства будут подробнее рассмотрены далее. Пока только отметим, что к измеримым на отрезке $[a,b]$ функциям f относятся непрерывные на $[a,b]$ почти всюду функции, ведь в качестве множества E из определения 2 можно брать множество всех точек, в которых функция f не является непрерывной на $[a,b]$.

Интегрируемость по Курцвейлю-Хенстоку ограниченных измеримых функций

Перейдем к теореме об интегрируемости ограниченных измеримых функций. Сначала очевидное утверждение.

Лемма 2. *Для любой постоянной $C \geqslant 0$ функция*

$$\max\{C, |x|\} = \frac{C + |x| + |C - |x||}{2}$$

непрерывна на $\mathbb{R}$ (и на $\mathbb{C}$).

Теорема 2. *Если функция f измерима и ограничена на $[a,b]$, то она интегрируема в смысле Курцвейля-Хенстока на $[a,b]$.*

▼ Пусть $C > 0$ и $|f(x)| \leqslant C$ на $[a,b]$. Возьмем произвольное $\varepsilon > 0$ и, воспользовавшись определением измеримости 1, найдем такую функцию $g_0 \in \mathrm{C}[a,b]$, что $\mu^*\{x \in [a,b] : f(x) \neq g_0(x)\} < \varepsilon$. Положим

$$g(x) = \frac{C g_0(x)}{\max\{C, |g_0(x)|\}} = \begin{cases} g_0(x), & \text{если } |g(x)| \leqslant C, \\ C\frac{g_0(x)}{|g_0(x)|}, & \text{если } |g(x)| > C. \end{cases}$$

Это непрерывная на $[a,b]$ функция (как частное двух непрерывных функций со знаменателем отличным от 0), $|g(x)| \leqslant C$ и $\{x \in [a,b] : f(x) \neq g(x)\} = \{x \in [a,b] : f(x) \neq g_0(x)\}$, $\mu^*\{x \in [a,b] : f(x) \neq g(x)\} < \varepsilon$. Обозначим множество $\{x \in [a,b] : f(x) \neq g(x)\}$ как E. Непрерывная функция g интегрируема по Риману на $[a,b]$. Поэтому существует такое $\gamma > 0$, что для любого отмеченного разбиения отрезка $[a,b]$ $\mathbb{T} = \{(\Delta_i, \xi_i)\}_{i=1}^n$ с отмеченными точками $\xi_i \in \Delta_i$ и $|\Delta_i| < \gamma$ для всех i имеем неравенство $|\mathfrak{S}(g, \mathbb{T}) - I| < \varepsilon$, где I - интеграл от g по $[a,b]$. Теперь найдем такую покрывающую множество E не более чем счетную систему интервалов l_i, что $\sum\limits_i |l_i| < \varepsilon$. Определим масштаб δ на $[a,b]$ следующим образом: если в точке x имеем равенство $f(x) = g(x)$, то $\delta(x) = \gamma$; если в точке x $f(x) \neq g(x)$, то есть $x \in E$, то найдем интервал $l_i \ni x$ и возьмем $\delta(x) \leqslant \gamma$ таким, что $B_{\delta(x)}(x) \subset l_i$. Пусть $\mathbb{T}$ и $\mathbb{T}'$ — два произвольных согласованных с масштабом δ разбиения. Тогда

$$|\mathfrak{S}(f,\mathbb{T}) - \mathfrak{S}(g,\mathbb{T})| = \left|\sum_{\xi_i \in E} (f(\xi_i) - g(\xi_i))|\Delta_i|\right| \leqslant 2C \sum_j \sum_{\xi_i \in l_j} |\Delta_i| \leqslant 2C \sum_j |l_j| < 2C\varepsilon.$$

Аналогично,

$$|\mathfrak{S}(f,\mathbb{T}') - \mathfrak{S}(g,\mathbb{T}')| < 2C\varepsilon.$$

Следовательно,

$$\begin{gathered}|\mathfrak{S}(f,\mathbb{T}) - \mathfrak{S}(f,\mathbb{T}')| \leqslant |\mathfrak{S}(f,\mathbb{T}) - \mathfrak{S}(g,\mathbb{T})| + |\mathfrak{S}(g,\mathbb{T}) - I| + \\ + |I - \mathfrak{S}(g,\mathbb{T}')| + |\mathfrak{S}(g,\mathbb{T}') - \mathfrak{S}(f,\mathbb{T}')| < (4C+2)\varepsilon.\end{gathered}$$

Выполнен критерий Коши интегрируемости по Курцвейлю-Хенстоку. ▲

Лекция 10 (10.03.20). О функциях, равных нулю почти всюду. Интегралы Римана и Курцвейля-Хенстока с переменным верхним пределом

Интегрируемость по Курцвейлю-Хенстоку функций, равных нулю почти всюду

Содержание этого прочитанного раздела не включено в программу экзамена.

Теперь докажем одну теорему для интеграла Курцвейля-Хенстока.

Теорема 3. *Если функция f на $[a,b]$ определена и равна нулю почти всюду, то $f \in \mathcal{H}[a,b]$ и $\int\limits_a^b f\,dx = 0$.*

▼ Пусть $E^j = \{x \in [a,b] : j-1 < |f(x)| \leqslant j\}$, $j \in \mathbb{N}$. Возьмем произвольное $\varepsilon > 0$. Так как множество E^j меры нуль по Лебегу, то для каждого $j \in \mathbb{N}$ найдется такая система интервалов $\{l_i^j\}_i$ покрывающая E^j, что $\sum\limits_i |l_i^j| < \frac{1}{j}2^{-j}\varepsilon$. Определим теперь масштаб δ на $[a,b]$ следующим образом: если $f(x) = 0$, то положим $\delta(x) = 1$; если $f(x) \neq 0$, то найдем сначала $E_j \ni x$, а потом найдем какое-либо $l_i^j \ni x$ и возьмем $\delta(x)$ таким, что $B_{\delta(x)}(x) \subset l_i^j$. Таким образом мы определим масштаб δ на $[a,b]$. Теперь возьмем любое разбиение $\mathbb{T}$ согласованное с масштабом δ. Тогда

$$\begin{gathered}|\mathfrak{S}(f,\mathbb{T})| = \left|\sum_k f(\xi_k)|\Delta_k|\right| \leqslant \sum_j \sum_{\xi_k \in E_j} |f(\xi_k)| \cdot |\Delta_k| \leqslant \\ \leqslant \sum_j j \sum_{\Delta_k \subset \cup_i l_i^j} |\Delta_k| \leqslant \sum_j j \sum_i \left|l_i^j\right| < \sum_{j=1}^{\infty} j \cdot \frac{1}{j} 2^{-j}\varepsilon = \varepsilon.\end{gathered}$$

Значит, f интегрируема по Курцвейлю-Хенстоку на $[a,b]$ и интеграл от нее нуль. ▲

Следствие. *Если функция f на $[a,b]$ интегрируема в смысле Курцвейля–Хенстока, то любая определенная и равная f почти всюду на $[a,b]$ функция g интегрируема на $[a,b]$ в том же смысле и $\int\limits_a^b g\,dx = \int\limits_a^b f\,dx$.*

▼ Это сразу следует из аддитивности интеграл Курцвейля-Хенстока по функциям и того, то $g-f$ интегрируема по Курцвейлю-Хенстоку и интеграл от нее по $[a,b]$ равен нулю. ▲

Теперь естественно ввести более широкое определение интегрируемости в смысле Курцвейля-Хенстока.

Определение. Будем определенную почти всюду на отрезке $[a,b]$ функцию f называть **интегрируемой в смысле Курцвейля–Хенстока** на $[a,b]$, если при некотором (а значит, и при любом) доопределении на $[a,b]$ она будет интегрируема в этом смысле; интегралом по $[a,b]$ от f будем считать интеграл по $[a,b]$ от доопределенной функции.

Определения

Определение 1. При изучении всех интегралов на отрезках прямой удобно считать, что всегда по определению

$$\int\limits_a^a f\,dx = 0;\ \int\limits_b^a f\,dx = -\int\limits_a^b f\,dx \text{ при } a < b.$$

Используя это определение, можно сформулировать свойство аддитивности по отрезкам интегралов Римана и Курцвейля–Хенстока в следующем более общем виде.

Утверждение. *Если для некоторых $a, b, c \in \mathbb{R}$ из трех интегралов $\int\limits_a^b f\,dx$, $\int\limits_b^c f\,dx$ и $\int\limits_a^c f\,dx$ (в смысле Римана или Курцвейля–Хенстока) два существуют, то существует и третий (в том же смысле) и выполняется равенство*

$$\int\limits_a^b f\,dx + \int\limits_b^c f\,dx = \int\limits_a^c f\,dx.$$

▼ Проверка этого утверждения проста и сводится к рассмотрению нескольких случаев различного порядка расположения точек a, b, c на $\mathbb{R}$. ▲

Теперь напомним одно ранее установленное свойство изучаемых интегралов — если f интегрируема на $[a,b]$ в смысле Римана или Курцвейля–Хенстока, то f интегрируема в том же смысле на любом отрезке $[\overline{a}, \overline{b}] \subset [a,b]$. Поэтому можно ввести следующее определение.

Определение 2. Если функция f интегрируема на $[a,b]$ в смысле Римана или Курцвейля–Хенстока, то, зафиксировав точку $x_0 \in [a.b]$ и постоянную $C \in \mathbb{R}$, функцию

$$F(x) = \int\limits_{x_0}^{x} f\,dt + C$$

на $[a,b]$ называют **неопределённым интегралом** или **интегралом с переменным верхним пределом**, соответственно, Римана или Курцвейля–Хенстока.

Два неопределенных интеграла

$$F_1(x) = \int\limits_{x_1}^{x} f\,dt + C_1 \quad \text{и} \quad F_2(x) = \int\limits_{x_2}^{x} f\,dt + C_2,$$

где $x_1, x_2 \in [a,b]$, отличаются на постоянную,

$$F_1(x) - F_2(x) = \left(\int\limits_{x_1}^{x} - \int\limits_{x_2}^{x}\right) f\,dt + C_1 - C_2 = \int\limits_{x_1}^{x_2} f\,dt + (C_1 - C_2).$$

Простейшие свойства неопределенных интегралов

Прежде, чем начать изучение неопределенных интегралов, введем еще одно определение.

Определение 3. Функция f на множестве $E \subset \mathbb{R}$ принадлежит **классу Липшица** (иногда говорят, **классу Гёльдера**), если f определена на E и существует такая постоянная $C \geqslant 0$, что для любых x_1 и x_2 из E выполняется неравенство $|f(x_1) - f(x_2)| \leqslant C|x_1 - x_2|$.

Обозначение. Если функция f на множестве $E \subset \mathbb{R}$ принадлежит классу Липшица, то пишут: $f \in \mathrm{Lip}(E)$.

Если $f \in \mathrm{Lip}(E)$, то f равномерно непрерывна на E.

Теорема 1. *Пусть функция f интегрируема на отрезке $[a,b]$ в смысле Римана или Курцвейля–Хенстока. Если f ограничена на $[a,b]$ (или, для интеграла Курцвейля–Хенстока, совпадает с ограниченной функцией почти всюду на $[a,b]$), то неопределенный интеграл $F(x) = \int\limits_{x_0}^{x} f\,dt + C \in \mathrm{Lip}([a,b])$. Если x — точка непрерывности f на $[a,b]$, то неопределенный интеграл $F(x)$ имеет производную (по $[a.b]$) в точке x и $F'(x) = f(x)$.*

▼ Если функция f ограничена на отрезке $[a,b]$, $[\overline{a},\overline{b}] \subset [a,b]$ и $\overline{\mathbb{T}} = \{(\Delta_i.\xi_i)\}$ — отмеченное разбиение отрезка $[\overline{a},\overline{b}]$, то $\left|\mathfrak{S}(f,\overline{\mathbb{T}})\right| = |\sum_i f(\xi_i)|\Delta_i|| \leqslant \sup_{[a,b]} |f| \sum_i |\Delta_i| = \sup\limits_{[a,b]} |f| \cdot (\overline{b} - \overline{a})$. При переходе к пределу по базе нестрогие неравенства сохраняются, значит, для любого $[\overline{a},\overline{b}] \subset [a,b]$ верно неравенство $\left|\int\limits_{\overline{a}}^{\overline{b}} f\,dt\right| \leqslant \sup\limits_{[a,b]} |f| \cdot (\overline{b} - \overline{a})$. Если точки $x, y \in [a,b]$, то $|F(y) - F(x)| = \left|\int\limits_{x}^{y} f\,dt\right| \leqslant \sup\limits_{[a,b]} |f| \cdot |y - x|$ и, значит, $F \in \mathrm{Lip}([a,b])$.

Если функция f непрерывна в точке $x \in [a,b]$ по $[a,b]$, то для $y \in [a,b]$, $y \neq x$, $\left|\frac{F(y)-F(x)}{y-x} - f(x)\right| = \left|\frac{1}{y-x}\int\limits_{x}^{y} f(t)\,dt - \frac{1}{y-x}\int\limits_{x}^{y} f(x)\,dt\right| = \frac{1}{|y-x|}\left|\int\limits_{x}^{y} f(t) - f(x)\,dt\right| \leqslant$ $\leqslant \frac{1}{|y-x|} \sup\limits_{t \in [x,y]} |f(t) - f(x)| \cdot |y - x| \leqslant \operatorname*{osc}\limits_{[x,y]} f = o(1)$ при $y \to x$; через $[x,y]$ в данном случае обозначается отрезок с концами в точках x и y при любых соотношениях порядка между ними. Тем самым доказано, что $F'(x) = f(x)$. ▲

Следствие 1. *У непрерывной на промежутке функции f имеется точная первообразная (функция, производная которой в каждой точке промежутка равна f).*

▼ Действительно, если взять интеграл с переменным верхним пределом (в любом из двух смыслов) $F(x) = \int\limits_{x_0}^{x} f\,dt$, где x_0 из данного промежутка, то по предыдущей теореме $F' = f$ всюду на промежутке ▲

Следствие 2. *У ограниченной на промежутке имеющей конечное число точек разрыва функции f имеется обобщенная первообразная.*

▼ Действительно, интеграл с переменным верхним пределом $F(x) = \int\limits_{x_0}^{x} f\,dt$, где x_0 из данного промежутка, будет обобщенной первообразной, ведь по предыдущей теореме это непрерывная функция на промежутке и $F' = f$ всюду на промежутке, кроме конечного числа точек. ▲

Замечание. Условие ограниченности f на промежутке можно заменить условием ограниченности f на любом отрезке из промежутка.

Следствие 3. *Если $f \in \mathcal{R}[a,b]$, F — неопределённый интеграл f, то $F'(x) = f(x)$ почти всюду на $[a,b]$.*

▼ Действительно, равенство $F'(x) = f(x)$ выполняется в каждой точке x, в которой f непрерывна, а по критерию Лебега f непрерывна почти всюду на $[a,b]$. ▲

Изложенные свойства неопределенного интеграла в основном характеризуют неопределенный интеграл Римана. Интегрируемые по Курцвейлю–Хенстоку функции могут быть неограниченными и не иметь точек непрерывности, тем не менее их неопределенные интегралы непрерывны, дифференцируемы почти всюду и их производные почти всюду равны подынтегральной функции.

Лекция 11 (13.03.20)
Интегралы Стилтьеса

Определения

В этой лекции будут рассмотрены обобщения ранее введенных интегралов. Идея обобщения принадлежит Стилтьесу.

Все рассматриваемые дальше функции считаем действительнозначными, хотя можно считать их и комплекснозначными. Те места, где различие между действительным и комплексным случаями существенно, будем отмечать особо.

Обозначение 1. Если дан отрезок $\Delta = [a,b]$ и определенная в его концах функция g, то приращение $g(b) - g(a)$ функции g на отрезке Δ будем обозначать $g(\Delta)$.

Определение 1. Пусть на отрезке $[a,b]$ определены функции f и g. **Интегральной суммой Римана–Стилтьеса** или просто **суммой Римана–Стилтьеса** функции f по функции g на отрезке $[a,b]$, соответствующей отмеченному разбиению $\mathbb{T} = \{(\Delta_i, \xi_i)\}_{i=1}^{n}$, называют сумму

$$\mathfrak{S}(f\,dg, \mathbb{T}) = \sum_{\mathbb{T}} f(\xi_i) g(\Delta_i) = \sum_{i=1}^{n} f(\xi_i) g(\Delta_i),$$

где $g(\Delta_i)$ — приращение функции g на отрезке Δ_i.

Интеграл Римана–Стилтьеса

Определение 2. Функция f **интегрируема по функции** g на отрезке $[a,b]$ **в смысле Римана–Стилтьеса** и ее интеграл равен числу $\boldsymbol{I}$, если f и g определены на $[a,b]$ и для любого $\varepsilon > 0$ существует такое число $\delta > 0$, что для любого разбиения $\mathbb{T} = \{(\Delta_i, \xi_i)\}_{i=1}^n$ мельче δ верно неравенство $|\mathfrak{S}(f\,dg, \mathbb{T}) - \boldsymbol{I}| < \varepsilon$, т.е.

$$\forall \varepsilon > 0\ \exists \text{число } \delta > 0\ \forall \mathbb{T},\ \xi_i \in \Delta_i \text{ и } |\Delta_i| < \delta \text{ для всех } i :$$
$$|\mathfrak{S}(f\,dg, \mathbb{T}) - \boldsymbol{I}| < \varepsilon.$$

Число $\boldsymbol{I}$ называют **определенным интегралом Римана–Стилтьеса** от функции f по функции g на отрезке $[a,b]$ (по отрезку $[a,b]$).

Обозначение 2. Определенный интеграл Римана–Стилтьеса от функции f по функции g на отрезке $[a,b]$ (по отрезку $[a,b]$) обозначают как $\int\limits_a^b f\,dg$ или $\int\limits_{[a,b]} f\,dg$, а если хотят подчеркнуть, что это интеграл Римана–Стилтьеса (в смысле Римана–Стилтьеса), то как $(\mathcal{R}-\mathcal{S})\int\limits_a^b f\,dg$ или $(\mathcal{R}-\mathcal{S})\int\limits_{[a,b]} f\,dg$.

Интеграл Курцвейля–Хенстока–Стилтьеса

Определение 3. Функция f **интегрируема по функции** g на отрезке $[a,b]$ (по отрезку $[a,b]$) **в смысле Курцвейля–Хенстока–Стилтьеса** и ее интеграл равен числу $\boldsymbol{I}$, если f и g определены на $[a,b]$ и для любого $\varepsilon > 0$ существует такой масштаб δ на $[a,b]$, что для любого согласованного с масштабом δ разбиения $\mathbb{T} = \{(\Delta_i, \xi_i)\}_{i=1}^n$ верно неравенство $|\mathfrak{S}(f\,dg, \mathbb{T}) - \boldsymbol{I}| < \varepsilon$, т.е.

$$\forall \varepsilon > 0\ \exists \text{масштаб } \delta \text{ на } [a,b]\ \forall \mathbb{T},\ \xi_i \in \Delta_i \text{ и } |\Delta_i| < \delta(\xi_i) \text{ для всех } i :$$
$$|\mathfrak{S}(f\,dg, \mathbb{T}) - \boldsymbol{I}| < \varepsilon.$$

Число $\boldsymbol{I}$ называют **определенным интегралом Курцвейля–Хенстока–Стилтьеса** от функции f по отрезку $[a,b]$ (на отрезке $[a,b]$).

Обозначение 3. Определенный интеграл Курцвейля–Хенстока–Стилтьеса от функции f по функции g на отрезке $[a,b]$ (по отрезку $[a,b]$) обозначают как $\int\limits_a^b f\,dg$ или $\int\limits_{[a,b]} f\,dg$, а если хотят подчеркнуть, что это интеграл Курцвейля–Хенстока–Стилтьеса (в смысле Курцвейля–Хенстока–Стилтьеса), то как $(\mathcal{H}-\mathcal{S})\int\limits_a^b f\,dg$ или $(\mathcal{H}-\mathcal{S})\int\limits_{[a,b]} f\,dg$.

Интегралы Стилтьеса как пределы по базе

Заметим, что оба определения интеграла фактически определяют интеграл как предел интегральных сумм Римана по базе.

Действительно, пусть $M = \{\mathbb{T}\}$ — множество отмеченных разбиений отрезка $[a,b]$. На M по функциям f и g определена функция $\mathfrak{S}(f\,dg, \mathbb{T})$, сопоставляющая каждому отмеченному разбиению $\mathbb{T}$ интегральную сумму Римана–Стилтьеса $\sum\limits_i f(\xi_i)g(\Delta_i)$. По определениям интегралов Римана–Стилтьеса и Курцвейля–Хенстока–Стилтьеса

$$(\mathcal{R}-\mathcal{S})\int\limits_a^b f\,dg = \lim_{\mathfrak{B}_\mathcal{R}} \mathfrak{S}(f\,dg, \mathbb{T}),$$

$$(\mathcal{H}-\mathcal{S})\int\limits_a^b f\,dg = \lim_{\mathfrak{B}_\mathcal{H}} \mathfrak{S}(f\,dg, \mathbb{T}).$$

Из свойств предела по базе сразу следуют некоторые простейшие свойства обоих обобщенных интегралов Римана.

Простейшие свойства интегралов Стилтьеса

Свойство 1 (взаимоотношение интегралов). *Если функция f интегрируема по функции g на $[a,b]$ в смысле Римана–Стилтьеса и $\boldsymbol{I}$ — ее интеграл, то f интегрируема по g на $[a,b]$ в смысле Курцвейля–Хенстока–Стилтьеса и ее интеграл то же число $\boldsymbol{I}$.*

▼ Действительно, так как $\mathfrak{B}_\mathcal{R} \subset \mathfrak{B}_\mathcal{H}$, то по теореме о пределах по разным базам свойство верно. ▲

Свойство 2 (единственность интегралов). *Если интеграл Римана–Стилтьеса или Курцвейля–Хенстока–Стилтьеса от f по g на $[a,b]$ существует, то он единственен.*

▼ Это непосредственное следствие теоремы о единственности предела по базе. ▲

Свойство 3 (линейность по функциям). *Если функция f интегрируема по функции g на $[a,b]$ в смысле Римана–Стилтьеса или Курцвейля–Хенстока–Стилтьеса, а c — число (действительное или комплексное), то cf интегрируема по g на $[a,b]$ в том же смысле и $\int\limits_a^b cf\,dg = c\int\limits_a^b f\,dg$; также f интегрируема по cg на $[a,b]$ в том же смысле и $\int\limits_a^b f\,d(cg) = c\int\limits_a^b f\,dg$.*

Если функции f_1 и f_2 интегрируемы по функции g на $[a,b]$ в смысле Римана–Стилтьеса или Курцвейля–Хенстока–Стилтьеса, то функция $f_1 \pm f_2$ интегрируема по функции g на $[a,b]$ в том же смысле и $\int\limits_a^b f_1 \pm f_2\,dg = \int\limits_a^b f_1\,dg \pm \int\limits_a^b f_2\,dg$; если функция f интегрируема по функциям g_1 и g_2 на $[a,b]$ в смысле Римана–Стилтьеса или Курцвейля–Хенстока–Стилтьеса, то функция f интегрируема по функции $g_1 \pm g_2$ на $[a,b]$ в том же смысле и $\int\limits_a^b f\,d(g_1 \pm g_2) = \int\limits_a^b f\,dg_1 \pm \int\limits_a^b f\,dg_2$.

▼ Действительно, функциям cf и g на множестве отмеченных разбиений $\mathbb{T}$ соответствует функция $\mathfrak{S}(cf\,dg, \mathbb{T}) = c\mathfrak{S}(f\,dg, \mathbb{T})$ и если $\mathfrak{B}$ — любая из двух баз, то по теореме о пределе произведения существует $\int\limits_a^b cf\,dg = \lim\limits_{\mathfrak{B}} \mathfrak{S}(cf\,dg, \mathbb{T}) = \lim\limits_{\mathfrak{B}} c\mathfrak{S}(f\,dg, \mathbb{T}) = c\lim\limits_{\mathfrak{B}} \mathfrak{S}(f\,dg, \mathbb{T}) = c\int\limits_a^b f\,dg$.

Аналогично, функциям f и cg на множестве отмеченных разбиений $\mathbb{T}$ соответствует функция $\mathfrak{S}(f\,dcg, \mathbb{T}) = c\mathfrak{S}(f\,dg, \mathbb{T})$ и если $\mathfrak{B}$ — любая из двух баз, то теореме о пределе суммы-разности существует $\int\limits_a^b f\,d(cg) = \lim\limits_{\mathfrak{B}} \mathfrak{S}(f\,dcg, \mathbb{T}) = \lim\limits_{\mathfrak{B}} c\mathfrak{S}(f\,dg, \mathbb{T}) =$ $c \lim\limits_{\mathfrak{B}} \mathfrak{S}(f\,dg, \mathbb{T}) = c \int\limits_a^b f\,dg.$

Функциям $f_1 \pm f_2$ и g на множестве отмеченных разбиений соответствует функция $\mathfrak{S}((f_1 \pm f_2)dg, \mathbb{T}) = \mathfrak{S}(f_1 dg, \mathbb{T}) \pm \mathfrak{S}(f_2 dg, \mathbb{T})$ и если $\mathfrak{B}$ — любая из двух указанных баз, то по теореме о пределе суммы-разности существует $\int\limits_a^b f_1 \pm f_2\,dg = \lim\limits_{\mathfrak{B}} \mathfrak{S}((f_1 \pm f_2)dg, \mathbb{T}) =$ $\lim\limits_{\mathfrak{B}} (\mathfrak{S}(f_1 dg, \mathbb{T}) \pm \mathfrak{S}(f_2 dg, \mathbb{T})) = \lim\limits_{\mathfrak{B}} \mathfrak{S}(f_1 dg, \mathbb{T}) \pm \lim\limits_{\mathfrak{B}} \mathfrak{S}(f_2 dg, \mathbb{T}) = \int\limits_a^b f_1\,dg \pm \int\limits_a^b f_2\,dg.$

Аналогично, функциям f и $g_1 \pm g_2$ на множестве отмеченных разбиений соответствует функция $\mathfrak{S}(f\,d(g_1 \pm g_2), \mathbb{T}) = \mathfrak{S}(f\,dg_1, \mathbb{T}) \pm \mathfrak{S}(f\,dg_2, \mathbb{T})$ и если $\mathfrak{B}$ — любая из двух указанных баз, то по теореме о пределе суммы-разности существует $\int\limits_a^b f\,d(g_1 \pm g_2) = \lim\limits_{\mathfrak{B}} \mathfrak{S}(f\,d(g_1 \pm g_2), \mathbb{T}) = \lim\limits_{\mathfrak{B}} (\mathfrak{S}(f\,dg_1, \mathbb{T}) \pm \mathfrak{S}(f\,dg_2, \mathbb{T})) =$ $\lim\limits_{\mathfrak{B}} \mathfrak{S}(f\,dg_1, \mathbb{T}) \pm \lim\limits_{\mathfrak{B}} \mathfrak{S}(f\,dg_2, \mathbb{T}) = \int\limits_a^b f\,dg_1 \pm \int\limits_a^b f\,dg_2.$ ▼

Свойство 4 (сохранение неравенств). *Если функция f интегрируема по функции g на $[a,b]$ в любом из двух смыслов, функция h интегрируема по функции g на $[a,b]$ в любом из двух смыслов и $f(x) \leqslant h(x)$ на $[a,b]$, а функция g неубывает на $[a,b]$, то для их интегралов (возможно, в разных смыслах) справедливо неравенство $\int\limits_a^b f\,dg \leqslant \int\limits_a^b h\,dg$.*

▼ Действительно, в силу свойства взаимоотношения интегралов можно ограничиться случаем интегрируемости f по g и h по g в смысле Курцвейля–Хенстока–Стилтьеса. Если $f(x) \leqslant h(x)$ на $[a,b]$, а функция g неубывает на $[a,b]$, то на множестве отмеченных разбиений

$$\mathfrak{S}(f\,dg, \mathbb{T}) = \sum_i f(\xi_i) g(\Delta_i) \leqslant \sum_i h(\xi_i) g(\Delta_i) = \mathfrak{S}(h\,dg, \mathbb{T})$$

и, значит, в силу теоремы о переходе к пределу в неравенствах (для пределов по базе)

$$\int_a^b f\,dg = \lim_{\mathfrak{B}_{\mathcal{H}}} \mathfrak{S}(f\,dg, \mathbb{T}) \leqslant \lim_{\mathfrak{B}_{\mathcal{H}}} \mathfrak{S}(h\,dg, \mathbb{T}) = \int_a^b h\,dg. \blacktriangle$$

Критерии Коши существования интегралов Стилтьеса

Теперь вспомним критерий Коши существования предела по базе.

Конечный предел функции f по базе $\mathfrak{B}$ существует тогда и только тогда, когда для функции f выполняется условие: для любого $\varepsilon > 0$ существует такой элемент $B \in \mathfrak{B}$, что для любых $x, x' \in B$ верно неравенство $|f(x) - f(x')| < \varepsilon$.

Используя критерий Коши получаем два следующих критерия интегрируемости.

Критерий Коши интегрируемости по Риману–Стилтьесу. *Если функции f и g определены на отрезке $[a,b]$, то функция f интегрируема по функции g на $[a,b]$*

в смысле Римана–Стилтьеса тогда и только тогда, когда для любого $\varepsilon > 0$ существует такое число $\delta > 0$, что для любых разбиений $\mathbb{T}, \mathbb{T}' \in \boldsymbol{B}_\delta$ верно неравенство $|\mathfrak{S}(f\,dg, \mathbb{T}) - \mathfrak{S}(f\,dg, \mathbb{T}')| < \varepsilon$, т.е.

$$\forall \varepsilon > 0\ \exists \delta > 0\ \forall \mathbb{T}, \mathbb{T}' \in \boldsymbol{B}_\delta :$$
$$|\mathfrak{S}(f\,dg, \mathbb{T}) - \mathfrak{S}(f\,dg, \mathbb{T}') < \varepsilon.$$

Критерий Коши интегрируемости по Курцвейлю–Хенстоку–Стилтьесу. *Если функции f и g определены на отрезке $[a, b]$, то функция f интегрируема по функции g на $[a, b]$ в смысле Курцвейля–Хенстока–Стилтьеса тогда и только тогда, когда для любого $\varepsilon > 0$ существует такой масштаб δ, что для любых разбиений Хенстока $\mathbb{T}, \mathbb{T}' \in \boldsymbol{B}_\delta$ верно неравенство $|\mathfrak{S}(f\,dg, \mathbb{T}) - \mathfrak{S}(f\,dg, \mathbb{T}')| < \varepsilon$, т.е.*

$$\forall \varepsilon > 0\ \exists \text{масштаб } \delta \text{ на } [a, b]\ \forall \mathbb{T}, \mathbb{T}' \in \boldsymbol{B}_\delta :$$
$$|\mathfrak{S}(f\,dg, \mathbb{T}) - \mathfrak{S}(f\,dg, \mathbb{T}')| < \varepsilon.$$

Приведенные критерии Коши интегрируемости позволяют доказать следующее свойство.

Свойство 5 (интегрируемость на подотрезках). *Если функция f интегрируема по функции g на отрезке $[a, b]$ в смысле Римана–Стилтьеса или Курцвейля–Хенстока–Стилтьеса, то она интегрируема в том же смысле и на любом отрезке $[\overline{a}, \overline{b}] \subset [a, b]$.*

▼ Аналогично введенным для отрезка $[a, b]$ классам отмеченных разбиений $\boldsymbol{B}_\delta$, $\boldsymbol{B}_\delta^{\mathcal{M}}$ и $\boldsymbol{B}_\delta^{\mathcal{H}}$ введем для отрезка $[\overline{a}, \overline{b}]$ такие же классы отмеченных разбиений $\overline{\boldsymbol{B}}_\delta$, $\overline{\boldsymbol{B}}_\delta^{\mathcal{M}}$ и $\overline{\boldsymbol{B}}_\delta^{\mathcal{H}}$. Пусть для заданного $\varepsilon > 0$ найдено такое число $\delta > 0$ (масштаб δ на $[a, b]$), что вышенаписанное условие критерия Коши выполнено. Покажем, что при тех же ε и δ (масштабе δ) условие критерия Коши выполнено на $[\overline{a}, \overline{b}]$. Рассмотрим два произвольных разбиения$\overline{\mathbb{T}}$ и $\overline{\mathbb{T}}'$ из $\overline{\boldsymbol{B}}_\delta$ (из $\overline{\boldsymbol{B}}_\delta^{\mathcal{M}}$, из $\overline{\boldsymbol{B}}_\delta^{\mathcal{H}}$). Если $a \neq \overline{a}$, то дополним их одним и тем же разбиением отрезка $[a, \overline{a}]$ мельче δ (согласованным с масштабом δ на$[a, \overline{a}]$); если $\overline{b} \neq b$, то дополним их одним и тем же разбиением отрезка $[\overline{b}, b]$ мельче δ (согласованным с масштабом δ на $[\overline{b}, b]$). В результате из отмеченных разбиений $\overline{\mathbb{T}}$ и $\overline{\mathbb{T}}'$ отрезка $[\overline{a}, \overline{b}]$ получим разбиения отрезка $[a, b]$ $\mathbb{T}$ и $\mathbb{T}'$ из $\boldsymbol{B}_\delta$ (из $\boldsymbol{B}_\delta^{\mathcal{M}}$, из $\boldsymbol{B}_\delta^{\mathcal{H}}$). Так как дополнялись разбиения $\overline{\mathbb{T}}$ и $\overline{\mathbb{T}}'$ одинаковым образом, то разность интегральных сумм $\mathfrak{S}(f\,dg, \overline{\mathbb{T}}) - \mathfrak{S}(f\,dg, \overline{\mathbb{T}}')$ на отрезке $[\overline{a}, \overline{b}]$ равна разности интегральных сумм $\mathfrak{S}(f\,dg, \mathbb{T}) - \mathfrak{S}(f\,dg, \mathbb{T}')$ на отрезке $[a, b]$ и, значит, (в силу критерия Коши) меньше ε по абсолютной величине. Следовательно, условие критерия Коши на отрезке $[\overline{a}, \overline{b}]$ выполнено и f интегрируема на $[\overline{a}, \overline{b}]$ в смысле Римана–Стилтьеса (Курцвейля–Хенстока–Стилтьеса). ▲

В заключение отметим, что интеграл Римана–Стилтьеса широко известен. Интеграл Курцвейля–Хенстока–Стилтьеса известен мало.

Аддитивность по отрезкам

Свойством аддитивности по отрезкам интеграл Римана–Стилтьеса не обладает, что показывает следующий пример.

Пример. Функция

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

интегрируема по функции

$$g(x) = \begin{cases} 0, & \text{если } x \in [-1, 0), \\ 1, & \text{если } x \in [0, 1] \end{cases}$$

на отрезках $[-1, 0]$ и $[0, 1]$ в смысле Римана–Стилтьеса, но не интегрируема на отрезке $[-1, 1]$ в том же смысле.

▼ Поскольку $f(x) = 0$ на $[-1, 0]$, то для любого отмеченного разбиения отрезка $[-1, 0]$ интегральная сумма функции f по функции g равна 0, предел интегральных сумм по базе Римана $\mathfrak{B}_{\mathcal{R}}$ равен 0, т.е. $(\mathcal{R} - \mathcal{S}) \int\limits_{-1}^{0} f\,dg = 0$.

Поскольку $g(x)$ постоянна на $[0, 1]$, то на любом отрезке $\Delta \subset [0, 1]$ приращение функции g будет равно 0, значит для любого отмеченного разбиения отрезка $[0, 1]$ интегральная сумма функции f по функции g равна 0, предел интегральных сумм по базе Римана $\mathfrak{B}_{\mathcal{R}}$ равен 0, т.е. $(\mathcal{R} - \mathcal{S}) \int\limits_{0}^{1} f\,dg = 0$.

Теперь рассмотрим любое такое отмеченное разбиение отрезка $[-1, 1]$, что точка 0 лежит внутри одного из отрезков разбиения Δ_k. Тогда $g(\Delta_k) = 1$, а при $i \neq k$ $g(\Delta_i) = 0$. Если отмеченная точка $\xi_k \leqslant 0$, то $f(\xi_k) = 0$ и интегральная сумма функции f по функции g равна 0, а если отмеченная точка $\xi_k > 0$, то $f(\xi_k) = 1$ и интегральная сумма функции f по функции g равна 1. На любом элементе $\boldsymbol{B}_\delta \in \mathfrak{B}_{\mathcal{R}}$ интегральные суммы принимают значения 0 и 1, критерий Коши интегрируемости в смысле Римана–Стилтьеса не выполняется. ▲

Лекция 12 (17.03.20)
Интегралы Стилтьеса.
Функции ограниченной вариации

Интеграл Курцвейля–Хенстока–Стилтьеса обладает свойством аддитивности по отрезкам.

Теорема 1 (аддитивность по отрезкам). *Пусть $a < b < c$ и функция f интегрируема по функции g на отрезках $[a, b]$ и $[b, c]$ в смысле Курцвейля–Хенстока–Стилтьеса. Тогда f интегрируема по функции g на $[a, c]$ в том же смысле и*

$$\int\limits_a^c f\,dg = \int\limits_a^b f\,dg + \int\limits_b^c f\,dg.$$

▼ Обозначим для краткости $I_1 = \int\limits_a^b f\,dg$, $I_2 = \int\limits_b^c f\,dg$. Возьмем произвольное $\varepsilon > 0$ и сначала найдем такой масштаб δ_1 на $[a, b]$, что для любого согласованного с ним разбиения $\mathbb{T}^1$ отрезка $[a, b]$ верно неравенство

$$|\mathfrak{S}(f\,dg, \mathbb{T}^1) - I_1| < \varepsilon;$$

а затем найдем такой масштаб δ_2 на $[b, c]$, что для любого согласованного с ним разбиения $\mathbb{T}^2$ отрезка $[a, b]$ верно неравенство

$$|\mathfrak{S}(f\,dg, \mathbb{T}^2) - I_2| < \varepsilon.$$

Положим

$$\delta(x) = \begin{cases} \min\{\delta_1(x), b - x\} & \text{при } x \in [a, b), \\ \min\{\delta_1(x), \delta_2(x)\} & \text{при } x = b, \\ \min\{\delta_2(x), x - b\} & \text{при } x \in (b, c]. \end{cases}$$

Тогда в любом согласованном с $\delta(x)$ разбиении $\mathbb{T}$ отрезка $[a, c]$, $\mathbb{T} = \{(\Delta_i, \xi_i)\}_{i=1}^{n}$, присутствуют один или два отрезка разбиения $\Delta_j \ni b$ с отмеченной точкой $\xi_j = b$, так как для любого отрезка разбиения Δ_i с $\xi_i \neq b$ из определения масштаба δ следует, что $b \notin B_{\delta(\xi_i)}(\xi_i)$, а значит, $b \notin \Delta_i$. В случае присутствия в разбиении $\mathbb{T}$ двух содержащих точку b отрезков разбиения она является их общим концом, а также отмеченной точкой для обоих отрезков. Тогда $\mathbb{T}^1 = \{(\Delta_i, \xi_i) \in \mathbb{T} : \ \Delta_i \subset [a, b]\}$ — разбиение отрезка $[a, b]$, согласованное с $\delta(x) \leqslant \delta_1(x)$ и $|\mathfrak{S}(f\,dg, \mathbb{T}^1) - I_1| < \varepsilon$, а $\mathbb{T}^2 = \{(\Delta_i, \xi_i) \in \mathbb{T} : \ \Delta_i \subset [b, c]\}$ — разбиение отрезка $[b, c]$, согласованное с $\delta(x) \leqslant \delta_2(x)$ и $|\mathfrak{S}(f\,dg, \mathbb{T}^2) - I_2| < \varepsilon$. А так как $\mathfrak{S}(f\,dg, \mathbb{T}) = (f\,dg, \mathbb{T}^1) + \mathfrak{S}(f\,dg, \mathbb{T}^2)$, то $|\mathfrak{S}(f\,dg, \mathbb{T}) - (I_1 + I_2)| \leqslant |\mathfrak{S}(f\,dg, \mathbb{T}^1) - I_1| + |\mathfrak{S}(f\,dg, \mathbb{T}^2) - I_2| < 2\varepsilon$. Если же только один отрезок $\Delta_j = [a_{j-1}, a_j]$ разбиения T содержит точку b, $b \in (a_{j-1}, a_j)$, $\xi_j = b$, где $1 \leqslant j \leqslant n$, то перейдем от разбиения $\mathbb{T}$ к разбиению $\overline{\mathbb{T}} = (\mathbb{T} \setminus \{(\Delta_j, \xi_j)\}) \cup \{([a_{j-1}, b], b), ([b, a_j], b)\}$. Очевидно

$$\mathfrak{S}(f\,dg, \mathbb{T}) = \mathfrak{S}(f\,dg, \overline{\mathbb{T}}),$$

а $\overline{\mathbb{T}}$ — согласованное с масштабом δ разбиение отрезка $[a, c]$ уже рассмотренного типа и

$$|\mathfrak{S}(f\,dg, \overline{\mathbb{T}}) - (I_1 + I_2)| < 2\varepsilon.$$

В итоге получаем, что для любого $\varepsilon > 0$ существует такой масштаб δ на $[a, c]$, что для любого отмеченного разбиения $\mathbb{T}$ отрезка $[a, c]$ верно неравенство $|\mathfrak{S}(f\,dg, \mathbb{T}) - (I_1 + I_2)| < 2\varepsilon$. Следовательно, f интегрируема на $[a, c]$ в смысле Курцвейля–Хенстока–Стилтьеса и $\int\limits_a^c f\,dx = I_1 + I_2$. ▲

Для интеграла Римана–Стилтьеса свойство аддитивности верно при дополнительном требовании существования интегралов по всем трем отрезкам.

Теорема 2 (аддитивность по отрезкам). *Пусть $a < b < c$ и функция f интегрируема по функции g на отрезках $[a, b]$, $[b, c]$ и $[a, c]$ в смысле Римана–Стилтьеса. Тогда*

$$\int\limits_a^c f\,dg = \int\limits_a^b f\,dg + \int\limits_b^c f\,dg.$$

▼ По свойству взаимосвязи интегралов, если существует интеграл Римана–Стилтьеса, то существует равный ему интеграл Курцвейля–Хенстока–Стилтьеса. А последний аддитивен по отрезкам, поэтому

$$(\mathcal{R}-\mathcal{S})\int\limits_a^c f\,dg = (\mathcal{H}-\mathcal{S})\int\limits_a^c f\,dg = (\mathcal{H}-\mathcal{S})\int\limits_a^b f\,dg +$$

$$+(\mathcal{H}-\mathcal{S})\int\limits_b^c f\,dg = (\mathcal{R}-\mathcal{S})\int\limits_a^b f\,dg + (\mathcal{R}-\mathcal{S})\int\limits_b^c f\,dg. \blacktriangle$$

В заключение отметим (без доказательства), что если потребовать чуть большую интегрируемость в смысле Римана–Стилтьеса, интегрируемость f по g на $[a,d]$ и $[b,c]$, $a<b<d<c$, то f будет интегрируема по g в смысле Римана–Стилтьеса на $[a,c]$.

Определения вариации

Вопрос интегрируемости конкретной функции f по конкретной функции g в каком-либо смысле может оказаться весьма сложным и зависит от особенностей каждой из функций. Обычно ищут такую пару классов функций, что каждая функция первого класса интегрируема по каждой функции второго класса. Одну такую пару укажем в данном разделе.

Определение 1. Пусть E — подмножество $\mathbb{R}$, φ — определенная на E функция. Через $D=\{a_i\}_{i=0}^n$ будем обозначать упорядоченный в порядке возрастания конечный набор точек из E, $I_i=[a_{i-1},a_i]$, $i=1,\ldots,n$. Точная верхняя грань сумм

$$\sum_{i=1}^n |\varphi(I_i)| = \sum_{i=1}^n |\varphi(a_i)-\varphi(a_{i-1})|,$$

взятая по всем конечным упорядоченным наборам D точек множества E, называется **вариацией** функции φ на множестве E.

Обозначение 1. Вариация функции φ на множестве E обозначается $\operatorname*{Var}_E \varphi$.

Определение 2. Если $\operatorname*{Var}_E \varphi < \infty$, то φ называют **функцией ограниченной вариации** (с **ограниченным изменением**, VB-**функцией**) на множестве E.

Обозначение 2. Пространство функций ограниченной вариации на множестве E обозначается $\mathrm{VB}(E)$.

Часто встречается другое определение вариации эквивалентное определению 1. Приведем его.

Определение 3. Точная верхняя грань сумм $\sum\limits_i |\varphi(I_i)|$, взятая по всем не более чем счетным наборам неперекрывающихся отрезков $\{I_i\}$ с концами из множества E, называется **вариацией** функции φ на множестве E и обозначается также $\operatorname*{Var}_E \varphi$.

Теорема 3. *Два приведенных определения вариации* $\operatorname*{Var}_E \varphi$ *эквивалентны.*

$\blacktriangledown$ Чтобы обозначать в доказательстве, о каком определении идет речь, будем добавлять к вариации по определению 1 индекс 1, а к вариации по определению 3 индекс 2 т.е. будем писать соответственно $\operatorname*{Var}_E^1 \varphi$ или $\operatorname*{Var}_E^2 \varphi$.

Если $D=\{a_i\}_{i=0}^n$ — упорядоченный в порядке возрастания конечный набор точек из E, то отрезки $I_i=[a_{i-1},a_i]$, $i=1,\ldots,n$, образуют конечную систему неперекрывающихся отрезков. Значит, $\sup\limits_D \sum\limits_{i=1}^n |\varphi(I_i)| = \operatorname*{Var}_E^1 \varphi \leqslant \operatorname*{Var}_E^2 \varphi$.

Покажем теперь, что $\operatorname*{Var}_E^2\varphi \leqslant \operatorname*{Var}_E^1\varphi$. Действительно, в определении 3 можно ограничиться конечными наборами неперекрывающихся отрезков, т.к. бесконечная сумма $\sum\limits_i |\varphi(I_i)|$ — предел конечных сумм (по определению). Если $\{I_k\}$ — конечный набор неперекрывающихся отрезков с концами из множества E, то, обозначив через a_i, $i = 0, 1, \dots, n$, их концы, занумерованные в порядке возрастания, получим, что среди отрезков $[a_{i-1}, a_i]$, $i =, 1, \dots, n$, содержатся все отрезки I_k и, значит, $\sum\limits_k |\varphi(I_k)| \leqslant \sum\limits_{i=1}^n |\varphi([a_{i-1}, a_i])|$, откуда следует, что $\operatorname*{Var}_E^2\varphi \leqslant \operatorname*{Var}_E^1\varphi$, что и требовалось.

В итоге имеем: $\operatorname*{Var}_E^1\varphi = \operatorname*{Var}_E^2\varphi$. ▲

Свойства функций ограниченной вариации

Свойство 1. *Если φ определена на E и $H \subset E$, то $\operatorname*{Var}_H \varphi \leqslant \operatorname*{Var}_E \varphi$.*

▼ Это непосредственное следствие любого определения вариации. ▲

Свойство 2. *Если φ — VB-функция на E, $c \in \mathbb{R}(\in \mathbb{C}$, то $c\varphi$ — VB-функция на E и $\operatorname*{Var}_E c\varphi = |c| \operatorname*{Var}_E \varphi$. Если φ и ψ— VB-функции на E, то $\varphi \pm \psi$ — VB-функция на E и $\operatorname*{Var}_E(\varphi \pm \psi) \leqslant \operatorname*{Var}_E \varphi + \operatorname*{Var}_E \psi$.*

▼ Для любого отрезка I $|c\varphi(I)| = |c| \cdot |\varphi(I)|$, поэтому $\operatorname*{Var}_E c\varphi = |c| \operatorname*{Var}_E \varphi$.

Для любого отрезка I $|(\varphi \pm \psi)(I)| \leqslant |\varphi(I)| + |\psi(I)|$, поэтому $\operatorname*{Var}_E(\varphi \pm \psi) \leqslant \operatorname*{Var}_E \varphi + \operatorname*{Var}_E \psi$. ▲

Свойство 3. *Если φ — VB-функция на E, то φ ограничена на E.*

▼ Действительно, если точка $x_0 \in E$, то для любой точки $x \in E$ верно неравенство $|\varphi(x)| \leqslant |\varphi(x) - \varphi(x_0)| + |\varphi(x_0)| \leqslant \operatorname*{Var}_E \varphi + |\varphi(x_0)|$, значит, φ ограничена на E. ▲

Свойство 4. *Если φ — VB-функция на $[a, b]$ и $[b, c]$, то φ — VB-функция на $[a, c]$ и $\operatorname*{Var}_{[a,c]} \varphi = \operatorname*{Var}_{[a,b]} \varphi + \operatorname*{Var}_{[b,c]} \varphi$.*

▼ Пусть $\{a_i\}_{i=0}^n$ — упорядоченный в порядке возрастания конечный набор точек из $[a, b]$, а $\{b_j\}_{j=0}^m$ — упорядоченный в порядке возрастания конечный набор точек из $[b, c]$. Объединив их, получим упорядоченный в порядке возрастания конечный набор точек из $[a, c]$ и

$$\sum_{i=1}^n |\varphi(a_i) - \varphi(a_{i-1})| + \sum_{j=1}^m |\varphi(b_j) - \varphi(b_{j-1})| \leqslant$$
$$\leqslant \sum_{i=1}^n |\varphi(a_i) - \varphi(a_{i-1})| + |\varphi(b_0) - \varphi(a_n)| + \sum_{j=1}^m |\varphi(b_j) - \varphi(b_{j-1})|,$$

значит,

$$\sum_{i=1}^n |\varphi(a_i) - \varphi(a_{i-1})| + \sum_{j=1}^m |\varphi(b_j) - \varphi(b_{j-1})| \leqslant \operatorname*{Var}_{[a,c]} \varphi,$$

и, следовательно,

$$\operatorname*{Var}_{[a,b]} \varphi + \operatorname*{Var}_{[b,c]} \varphi \leqslant \operatorname*{Var}_{[a,c]} \varphi.$$

Если $\{c_i\}_{i=0}^n$ — упорядоченный в порядке возрастания конечный набор точек из $[a, c]$, содержащий точку c, $c = c_j$, $0 < j < n$, то $\{c_i\}_{i=0}^j$ — упорядоченный в порядке возрастания конечный набор точек из $[a, b]$, а $\{c_i\}_{i=j}^j$ — упорядоченный в порядке возрастания конечный набор точек из $[b,c]$. Так как $\sum\limits_{i=1}^{n}|\varphi(c_i)-\varphi(c_{i-1})| = \left(\sum\limits_{i=1}^{j}+\sum\limits_{i=j+1}^{n}\right)|\varphi(c_i)-\varphi(c_{i-1})|$, то

$$\sum_{i=1}^{n}|\varphi(c_i)-\varphi(c_{i-1})| \leqslant \operatorname*{Var}_{[a,b]}\varphi + \operatorname*{Var}_{[b,c]}\varphi.$$

Поскольку к любому набору точек из $[a,c]$ можно добавить точку c и точку, которая меньше c, а также точку, которая больше c, только увеличив при этом сумму модулей приращений φ по парам соседних точек, то

$$\operatorname*{Var}_{[a,c]}\varphi \leqslant \operatorname*{Var}_{[a,b]}\varphi + \operatorname*{Var}_{[b,c]}\varphi.$$

В итоге имеем равенство $\operatorname*{Var}\limits_{[a,c]}\varphi = \operatorname*{Var}\limits_{[a,b]}\varphi + \operatorname*{Var}\limits_{[b,c]}\varphi$. ▲

Свойство 5. *Если φ и ψ — VB-функции на E, то их произведение $\varphi\psi$ — VB-функция на E и $\operatorname*{Var}\limits_{E}\varphi\psi \leqslant \sup\limits_{E}|\varphi|\cdot\operatorname*{Var}\limits_{E}\psi + \sup\limits_{E}|\psi|\cdot\operatorname*{Var}\limits_{E}\varphi$.*

▼ Для любого отрезка $I=[a,b]$ с концами $a,b\in E$ верна оценка $|(\varphi\psi)(I)| = |\varphi(b)\psi(b)-\varphi(a)\psi(a)| = |\varphi(b)(\psi(b)-\psi(a))+\psi(a)(\varphi(b)-\varphi(a))| \leqslant |\varphi(b)|\cdot|\psi(I)|+|\psi(a)|\cdot|\varphi(I)| \leqslant \sup\limits_{E}|\varphi|\cdot|\psi(I)|+\sup\limits_{E}|\psi|\cdot|\varphi(I)|$, поэтому $\operatorname*{Var}\limits_{E}\varphi\psi \leqslant \sup\limits_{E}|\varphi|\cdot\operatorname*{Var}\limits_{E}\psi+\sup\limits_{E}|\psi|\cdot\operatorname*{Var}\limits_{E}\varphi$. ▲

Свойство 6. *Если φ — VB-функция на E и $m=\inf\limits_{E}|\varphi|>0$, то $\frac{1}{\varphi}$ — VB-функция на E и $\operatorname*{Var}\limits_{E}\frac{1}{\varphi} \leqslant \frac{1}{m^2}\operatorname*{Var}\limits_{E}\varphi$.*

▼ Для любого отрезка $I=[a,b]$ с концами $a,b\in E$ верна оценка $\left|\frac{1}{\varphi}(I)\right| = \left|\frac{1}{\varphi(b)}-\frac{1}{\varphi(a)}\right| = \left|\frac{\varphi(a)-\varphi(b)}{\varphi(a)\varphi(b)}\right| \leqslant \frac{|\varphi(I)|}{m^2}$, поэтому $\operatorname*{Var}\limits_{E}\frac{1}{\varphi} \leqslant \frac{1}{m^2}\operatorname*{Var}\limits_{E}\varphi$. ▲

Свойство 7. *Если φ — VB-функция на E, $\psi\in\operatorname{Lip}(\varphi(E))$, то $\psi(\varphi)$ — VB-функция на E и $\operatorname*{Var}\limits_{E}\psi(\varphi) \leqslant C\operatorname*{Var}\limits_{E}\varphi$, где C — постоянная из определения класса Липшица.*

▼ Для любого отрезка $I=[a,b]$ с концами $a,b\in E$ верна оценка $|\psi(\varphi)(I)| = |\psi(\varphi(b))-\psi(\varphi(a)) \leqslant C|\varphi(b)-\varphi(a)| = C|\varphi(I)|$, поэтому $\operatorname*{Var}\limits_{E}\psi(\varphi) \leqslant C\operatorname*{Var}\limits_{E}\varphi$. ▲

Лекция 13 (20.03.20) **Функции ограниченной вариации. Интегрирование непрерывных функций по функциям ограниченной вариации. Интегрирование по частям для интеграла Римана–Стилтьеса**

Функции ограниченной вариации — разность монотонных функций

Обозначение. Пусть φ определена на $[a,b]$, $a\leqslant b$, тогда полагаем $\operatorname*{Var}\limits_{a}^{b}\varphi = \operatorname*{Var}\limits_{[a,b]}\varphi$, $\operatorname*{Var}\limits_{b}^{a}\varphi = -\operatorname*{Var}\limits_{[a,b]}\varphi$.

Тогда из свойства 4 следует, что при любом расположении точек a, b, c на $\mathbb{R}$, если функция φ определена на $[\min\{a,b,c\}, \max\{a,b,c\}]$, то

$$\operatorname*{Var}_a^c \varphi = \operatorname*{Var}_a^b \varphi + \operatorname*{Var}_b^c \varphi.$$

Теорема 1. *Если φ — VB-функция на промежутке I, точка $x_0 \in I$, то $\operatorname*{Var}_{x_0}^x \varphi$ — неубывающая функция на I, а если φ еще и действительнозначна, то неубывающими являются также функции $\operatorname*{Var}_{x_0}^x \varphi + \varphi(x)$ и $\operatorname*{Var}_{x_0}^x \varphi - \varphi(x)$.*

▼ Из свойства аддитивности вариации по отрезкам функций ограниченной вариации следует, что для любых x_1, x_2 из I

$$\operatorname*{Var}_{x_0}^{x_2} \varphi - \operatorname*{Var}_{x_0}^{x_1} \varphi = \operatorname*{Var}_{x_1}^{x_2} \varphi,$$

а при $x_2 \geqslant x_1$ $\operatorname*{Var}_{x_1}^{x_2} \varphi \geqslant |\varphi(x_2) - \varphi(x_1)|$. Значит, $\operatorname*{Var}_{x_0}^x \varphi$ — неубывающая функция на I и, при действительнозначности φ, $\operatorname*{Var}_{x_0}^x \varphi + \varphi(x)$ и $\operatorname*{Var}_{x_0}^x \varphi - \varphi(x)$ — неубывающие функции на I. ▲

Теорема 2 (Жордана). *Действительнозначная функция φ является VB-функцией на отрезке I тогда и только тогда, когда φ является разностью двух неубывающих на I функций, причем эти функции можно выбрать такими, что сумма их вариаций на любом отрезке $J \subset I$ будет равна вариации φ на J.*

▼ Пусть $I = [a, b]$. Ограниченная неубывающая функция ψ на I — VB-функция на I и

$$\operatorname*{Var}_I \psi = \psi(b) - \psi(a),$$

поскольку для любого набора точек $D = \{a_i\}_{i=0}^n \subset I$, $a \leqslant a_0 < a_1 < \dots < a_n \leqslant b$, имеем $\sum\limits_{i=1}^n |\psi(a_i) - \psi(a_{i-1}| = \psi(a_n) - \psi(a_0) \leqslant \psi(b) - \psi(a)$ и, значит, $\operatorname*{Var}_I \psi \leqslant \psi(b) - \psi(a)$, а с другой стороны, $\operatorname*{Var}_I \psi \geqslant \psi(b) - \psi(a)$. Поэтому, если φ — разность двух неубывающих на I функций, то по свойству 2 φ — VB-функция на I.

Если действительнозначная функция φ является VB-функцией на отрезке $I \neq \varnothing$, то, взяв $x_0 \in I$, укажем три представления φ в виде разности ограниченных неубывающих на I функций

$$\varphi(x) = \left(\operatorname*{Var}_{x_0}^x \varphi + \varphi(x)\right) - \operatorname*{Var}_{x_0}^x \varphi,$$

$$\varphi(x) = \operatorname*{Var}_{x_0}^x \varphi - \left(\operatorname*{Var}_{x_0}^x \varphi - \varphi(x)\right),$$

$$\varphi(x) = \frac{1}{2}\left(\operatorname*{Var}_{x_0}^x \varphi + \varphi(x)\right) - \frac{1}{2}\left(\operatorname*{Var}_{x_0}^x \varphi - \varphi(x)\right).$$

Покажем, что в последнем случае сумма вариацией функций разности на любом отрезке $J \subset I$ равна вариации φ на J. Действительно, если $J = [c, d]$, то имеем равенство

$$\operatorname*{Var}_c^d \frac{1}{2}\left(\operatorname*{Var}_{x_0}^x \varphi \pm \varphi(x)\right) = \frac{1}{2}\left(\operatorname*{Var}_{x_0}^d \varphi \pm \varphi(d)\right) - \frac{1}{2}\left(\operatorname*{Var}_{x_0}^c \varphi \pm \varphi(c)\right) = \frac{1}{2}\left(\operatorname*{Var}_c^d \varphi \pm (\varphi(d) - \varphi(c))\right).$$

Из него следует, что

$$\operatorname*{Var}_c^d \frac{1}{2}\left(\operatorname*{Var}_{x_0}^x \varphi + \varphi(x)\right) + \operatorname*{Var}_c^d \frac{1}{2}\left(\operatorname*{Var}_{x_0}^x \varphi - \varphi(x)\right) = \operatorname*{Var}_c^d \varphi,$$

сумма вариацией функций разности на J равна вариации φ на J. ▲

Интегрируемость непрерывных функций по функциям ограниченной вариации

Лемма. *Пусть $T = \{I_i\}_{i=1}^n$ — разбиение отрезка $[a,b]$. Тогда для любого отрезка $I = [c,d] \subset [a,b]$ и любой определенной на нем функции g верно равенство*

$$g(I) = g(d) - g(c) = \sum_{i=1}^n g(I \cap I_i),$$

где $g(I \cap I_i)$ — приращение функции g на отрезке $I \cap I_i$, причем приращение на пустом или одноточечном множестве равно нулю.

▼ Можно считать, что отрезки I_i занумерованы в порядке их расположения на $\mathbb{R}$. Пусть $I_i = [a_{i-1}, a_i]$, $a = a_0 < a_1 < \dots < a_n = b$, $c \in I_k$, $d \in I_l$, то есть $a_{k-1} \leqslant c \leqslant a_k < \dots < a_{l-1} \leqslant d \leqslant a_l$. Тогда $g(I) = g(d) - g(c) = g(d) - g(a_{l-1}) + \sum\limits_{i=k+1}^{l-1} (g(a_i) - g(a_{i-1})) + g(a_k) - g(c) = g([c,d] \cap I_l) + \sum\limits_{i=k+1}^{l-1} g([c,d] \cap I_i) + g([c,d] \cap I_k) = \sum\limits_{i=1}^{n} g(I \cap I_i)$. ▲

Теорема 3. *Если $f \in \mathrm{C}[a,b]$ (— пространству непрерывных на $[a,b]$ функций), а $g \in \mathrm{VB}[a,b]$, то f интегрируема по g на $[a,b]$ в смысле Римана–Стилтьеса.*

▼ Так как функция f равномерно непрерывна на $[a,b]$, то для любого $\varepsilon > 0$ найдётся такое $\delta > 0$, что если $x, y \in [a,b]$ и $|x-y| < 2\delta$, то $|f(x) - f(y)| < \varepsilon$. Пусть $\mathbb{T} = \{(\Delta_i, \xi_i)\}$ и $\mathbb{T}' = \{(\Delta'_j, \xi'_j)\}$ отмеченные разбиения отрезка $[a,b]$ мельче δ, тогда $\mathfrak{S}(f\,dg, \mathbb{T}) - \mathfrak{S}(f\,dg, \mathbb{T}') = \sum\limits_i f(\xi_i) g(\Delta_i) - \sum\limits_j f(\xi'_j) g(\Delta'_j) = \sum\limits_i \sum\limits_j \left(f(\xi_i) - f(\xi'_j)\right) g(\Delta_i \cap \Delta'_j)$, где $g(\Delta_i \cap \Delta'_j) = 0$, если $\Delta_i \cap \Delta'_j$ пусто или одноточечно, а если $\Delta_i \cap \Delta'_j \neq \varnothing$, то имеем $|f(\xi_i) - f(\xi'_j)| < \varepsilon$. Поэтому $\left|\mathfrak{S}(f\,dg, \mathbb{T}) - \mathfrak{S}(f\,dg, \mathbb{T}')\right| = \left|\sum\limits_i \sum\limits_j \left(f(\xi_i) - f(\xi'_j)\right) g(\Delta_i \cap \Delta'_j)\right| \leqslant \varepsilon \sum\limits_i \sum\limits_j \left|g(\Delta_i \cap \Delta'_j)\right| \leqslant \varepsilon \operatorname*{Var}_{[a,b]} g$.

По критерию Коши интегрируемости по Риману–Стилтьесу существует $(\mathcal{R} - \mathcal{S}) \int\limits_a^b f\,dg$. ▲

Для интеграла Римана–Стилтьеса в теореме нельзя расширить ни один из классов с сохранением другого. Достаточно просто проверяется невозможность расширения класса функций f: если функция g ограниченной вариации на $[a,b]$ разрывна в точке

$c \in [a, b]$ (относительно $[a, b]$), то необходимым условием интегрируемости f по g на $[a, b]$ является непрерывность f в точке c (для простоты можно ограничиться функциями $g_c(x) = \operatorname{sign}(x - c)$, где $\operatorname{sign} x = 1$ при $x > 0$, $= 0$ при $x = 0$ и $= -1$ при $x < 0$). Невозможность расширения класса функций g требует более тонких рассуждений.

Добавим к свойствам интегралов еще одно свойство.

Свойство 8. *Если f интегрируема по g на $[a, b]$ в любом из двух смыслов, $g \in \mathrm{VB}[a, b]$, то верна оценка*

$$\left|\int\limits_a^b f\,dg\right| \leqslant \sup_{[a,b]} |f| \cdot \operatorname*{Var}_{[a.b]} g.$$

▼ Действительно, для любого отмеченного разбиения $\mathbb{T} = \{(\Delta_i, \xi_i)\}$ имеем

$$|\mathfrak{S}(f\,dg, \mathbb{T})| = \left|\sum_i f(\xi_i) g(\Delta_i)\right| \leqslant$$

$$\leqslant \sup_{[a,b]} |f| \cdot \sum_i |g(\Delta_i)| \leqslant \sup_{[a,b]} |f| \cdot \operatorname*{Var}_{[a.b]} g.$$

По теореме о переходе к пределу в неравенствах (для пределов по базе) оценка верна. ▲

Теорема 1. *Если функция g интегрируема по функции f на $[a, b]$ в смысле Римана–Стилтьеса, то и f интегрируема по g на $[a, b]$ в том же смысле и*

$$(\mathcal{R} - \mathcal{S}) \int\limits_a^b f\,dg = f \cdot g\big|_a^b - (\mathcal{R} - \mathcal{S}) \int\limits_a^b g\,df =$$

$$= f(b)g(b) - f(a)g(a) - (\mathcal{R} - \mathcal{S}) \int\limits_a^b g\,df.$$

▼ Пусть $\mathbb{T} = \{(\Delta_i, \xi_i)\}$ — разбиение отрезка $[a, b]$, $\xi_i \in \Delta_i = [a_{i-1}, a_i]$, $i = 1, \ldots, n$, $a = a_0 < a_1 < \cdots < a_n = b$. Тогда $\mathfrak{S}(f\,dg, \mathbb{T}) = \sum\limits_{i=1}^n f(\xi_i)\big(g(a_i) - g(a_{i-1})\big) = \sum\limits_{i=1}^n f(\xi_i) g(a_i - \sum\limits_{i=1}^n f(\xi_i) g(a_{i-1}) = \sum\limits_{i=1}^n f(\xi_i) g(a_i) - \sum\limits_{i=0}^{n-1} f(\xi_{i+1}) g(a_i) = f(\xi_{n+1}) g(a_n) - f(\xi_0) g(a_0) - \sum\limits_{i=0}^n g(a_i)\big(f(\xi_{i+1}) - f(\xi_i)\big)$, где $\xi_0 = a$, $\xi_{n+1} = b$, $\xi_i \leqslant a_i \leqslant \xi_{i+1}$. Тогда $\tilde{\mathbb{T}} = \{([\xi_i, \xi_{i+1}], a_i)\}_{i=0}^n$ — отмеченное разбиение отрезка $[a, b]$ (пары с вырожденными отрезками $[\xi_i, \xi_{i+1}]$ можно удалить из разбиения $\tilde{\mathbb{T}}$) и выполняется равенство

$$\mathfrak{S}(f\,dg, \mathbb{T}) = f(\xi_{n+1}) g(a_n) - f(\xi_0) g(a_0) - \mathfrak{S}(g\,df, \tilde{\mathbb{T}}). \qquad (*)$$

Поскольку $a = \xi_0 = a_0 \leqslant \xi_1 \leqslant a_2 \leqslant \ldots \xi_n \leqslant a_n = \xi_{n+1} = b$, то из неравенств $|\Delta_i| = a_i - a_{i-1} < \delta$ для $i = 1, \ldots, n$ следуют неравенства $|\xi_{i+1} - \xi_i| < 2\delta$ для $i = 0, \ldots, n$. Значит, если в равенстве (*) выражение справа имеет предел по базе $\mathfrak{B}_{\mathcal{R}}$, то и выражение слева имеет предел по той же базе и

$$\int_a^b f\,dg = f(b)g(b) - f(a)g(a) - \int_a^b g\,df. \blacktriangle$$

Следствие. *Если $f \in \mathrm{VB}[a,b]$, а $g \in \mathrm{C}[a,b]$, то f интегрируема по g на $[a,b]$ в смысле Римана–Стилтьеса.*

$\blacktriangledown$ Действительно, по теореме об интегрируемости непрерывной функции по функции ограниченной вариации g интегрируема по f на $[a,b]$ в смысле Римана–Стилтьеса, а тогда и f интегрируема по g на $[a,b]$ в смысле Римана–Стилтьеса. $\blacktriangle$

Примечание. Ситуация с интегрированием по частям для интеграла Курцвейля–Хенстока более сложная. Существуют примеры, когда в формуле интегрирования по частям один из интегралов существует, а другой не существует, или оба существуют, но формула не верна. Для интеграла Курцвейля-Хенстока формула интегрирования по частям верна при дополнительных условиях.

Лекция 14 (24.03.20)
Сведение интеграла Римана–Стилтьеса к интегралу Римана. Замена переменной в интегралах

Сведение интеграла Римана–Стилтьеса к интегралу Римана

Теорема 1. *Если f ограничена на $[a,b]$, g непрерывна на $[a,b]$ и fg интегрируема по Риману на $[a,b]$, а функция G — неопределенный интеграл от функции g, то f интегрируема по G на $[a,b]$ в смысле Римана–Стилтьеса и*

$$(\mathcal{R}-\mathcal{S})\int_a^b f\,dG = (\mathcal{R})\int_a^b f\cdot g\,dx.$$

$\blacktriangledown$ Функция g непрерывна, а значит, равномерно непрерывна на $[a,b]$. Для любого $\varepsilon > 0$ найдется такое $\delta > 0$, что если $x, t \in [a,b]$ и $|x-t| < \delta$, то $|g(x) - g(t)| < \varepsilon$. Если отрезок $\Delta \subset [a,b]$ и $|\Delta| < \delta$, то для любой точки $\xi \in \Delta$

$$|G(\Delta) - g(\xi)|\Delta|| = \left|\int_\Delta (g(t) - g(\xi))dt\right| \leqslant \varepsilon|\Delta|.$$

Тогда для любого отмеченного разбиения $\mathbb{T} = \{(\Delta_i, \xi_i)\}$ мельче δ имеем

$$|\mathfrak{S}(fdG, \mathbb{T}) - \mathfrak{S}(f\cdot g, \mathbb{T})| \leqslant \sum_i |f(\xi_i)|\cdot|G(\Delta_i) - g(\xi_i)|\Delta_i|| \leqslant \sup_{[a,b]}|f|(b-a)\varepsilon.$$

Следовательно, $\mathfrak{S}(fdG, \mathbb{T}) - \mathfrak{S}(f\cdot g, \mathbb{T}) = o(1)$ по базе Римана $\mathfrak{B}_\mathcal{R}$. Так как

$$\int_a^b fg dx = \lim_{\mathfrak{B}_\mathcal{R}} \mathfrak{S}(f\cdot g, \mathbb{T}) = \lim_{\mathfrak{B}_\mathcal{R}} \mathfrak{S}(fdG, \mathbb{T}),$$

то

$$\int\limits_a^b f\,dG = \lim_{\mathfrak{B}_{\mathcal{R}}} \mathfrak{S}(fdG, \mathbb{T}) = \int\limits_a^b fgdx.$$

▲

Интегрирование по частям для интеграла Римана

Теорема 2. *Если функции u и v интегрируемы на $[a,b]$ в смысле Римана, U и V — их неопределенные интегралы, то $u \cdot V$ и $v \cdot U$ интегрируемы на $[a,b]$ в смысле Римана и*

$$(\mathcal{R}) \int\limits_a^b u \cdot V\,dx = U \cdot V\Big|_a^b - (\mathcal{R}) \int\limits_a^b v \cdot U\,dx =$$

$$= U(b)V(b) - U(a)V(a) - (\mathcal{R}) \int\limits_a^b v \cdot U\,dx.$$

▼ По предыдущей теореме $(\mathcal{R}) \int\limits_a^b u \cdot V\,dx = (\mathcal{R} - \mathcal{S}) \int\limits_a^b V\,dU$, $(\mathcal{R}) \int\limits_a^b v \cdot U\,dx = (\mathcal{R} - \mathcal{S}) \int\limits_a^b U\,dV$,а по теореме об интегрировании по частям для интеграла Римана–Стилтьеса $(\mathcal{R} - \mathcal{S}) \int\limits_a^b V\,dU = U(b)V(b) - U(a)V(a) - (\mathcal{R} - \mathcal{S}) \int\limits_a^b U\,dV$. ▲

Замена переменной в интегралах

Теорема 3 (формула замены переменной). *Если $f \in \mathcal{R}[a,b]$ $(\in \mathcal{H}[a,b])$, φ — строго возрастающая непрерывно дифференцируемая функция на $[\alpha,\beta]$, $\varphi(\alpha) = a$, $\varphi(\beta) = b$, то $f(\varphi) \cdot \varphi' \in \mathcal{R}[\alpha,\beta]$ $(\in \mathcal{H}[\alpha,\beta])$ и*

$$\int\limits_\alpha^\beta f(\varphi) \cdot \varphi'\,dt = \int\limits_a^b f\,dx.$$

▼ Возьмем любое $\varepsilon > 0$. Функция φ' непрерывна, а значит, равномерно непрерывна на $[\alpha,\beta]$, поэтому найдется такое $\delta_1 > 0$, что для любых $t, s \in [\alpha,\beta]$ из неравенства $|t - s| < \delta_1$ следует неравенство $\sup\limits_{[a,b]} |f| \cdot |\varphi'(t) - \varphi'(s)| < \varepsilon$ (для любого $t \in [\alpha,\beta]$ найдется такое $\delta_1(t) > 0$, что для любого $s \in [\alpha,\beta]$, $|t - s| < \delta_1(t)$, верно неравенство $|f(\varphi(t))(\varphi'(t) - \varphi'(s))| < \varepsilon$).

Пусть $\boldsymbol{I} = \int\limits_a^b f\,dx$. Найдется такое число $\delta_2 > 0$ (масштаб δ_2 на $[a,b]$), что для любого разбиения $\mathbb{T}_2$ отрезка $[a,b]$ мельче δ_2 (согласованного с масштабом δ_2) выполняется неравенство

$$|\mathfrak{S}(f, \mathbb{T}_2) - \boldsymbol{I}| < \varepsilon.$$

Так как функция φ непрерывна, а значит, и равномерно непрерывна на $[\alpha, \beta]$, то найдется такое число $\delta < \delta_1$ (такой масштаб $\delta < \delta_1$ на $[\alpha, \beta]$), что для любых $t, s \in [\alpha, \beta]$ из неравенства $|s - t| < \delta$ следует неравенство $|\varphi(s) - \varphi(t)| < \delta_2$ (из неравенства $|t - s| < \delta(t)$ следует неравенство $|\varphi(t) - \varphi(s)| < \delta_2(\varphi(t))$.

Возьмем любое разбиение $\mathbb{T} = \{([\alpha_{i-1}, \alpha_i], \xi_i)\}$ отрезка $[\alpha, \beta]$ мельче δ (согласованное с масштабом δ). Используя формулу Лагранжа найдем такие $\zeta_i \in [\alpha_{i-1}, \alpha_i]$, что $\varphi(\alpha_i) - \varphi(\alpha_{i-1}) = \varphi'(\zeta_i)(\alpha_i - \alpha_{i-1})$. Тогда по неравенствам выше $\Big|\mathfrak{S}(f(\varphi)\varphi', \mathbb{T}) - \boldsymbol{I}\Big| = \Big|\sum\limits_i f(\varphi(\xi_i))\varphi'(\xi_i)|\alpha_i - \alpha_{i-1}| - \boldsymbol{I}\Big| \leqslant \Big|\sum\limits_i f(\varphi(\xi_i))(\varphi'(\xi_i) - \varphi'(\zeta_i))|\alpha_i - \alpha_{i-1}|\Big| + \Big|\sum\limits_i f(\varphi(\xi_i))\varphi'(\zeta_i)|\alpha_i - \alpha_{i-1}| - \boldsymbol{I}\Big| < \sum_i \varepsilon|\alpha_i - \alpha_{i-1}| + \Big|\sum\limits_i f(\varphi(\xi_i))|\varphi(\alpha_i) - \varphi(\alpha_{i-1})| - \boldsymbol{I}\Big| < (\beta - \alpha + 1)\varepsilon$, откуда и следует утверждение теоремы. ▲

Следствие 1 (формула замены переменной). *Если* $f \in \mathcal{R}[a, b]$ *(* $\in \mathcal{H}[a, b]$*),* φ *— строго убывающая непрерывно дифференцируемая функция на* $[\alpha, \beta]$*,* $\varphi(\alpha) = b$*,* $\varphi(\beta) = a$*, то* $f(\varphi) \cdot \varphi' \in \mathcal{R}[\alpha, \beta]$ *(*$\in \mathcal{H}[\alpha, \beta]$*) и*

$$\int\limits_\alpha^\beta f(\varphi) \cdot -\varphi' \, dt = \int\limits_a^b f \, dx.$$

▼ Действительно, рассматривая функцию $f(-x)$ на отрезке $[-b, -a]$ видим, что каждому отмеченному разбиению $\mathbb{T} = \{([a_{i-1}, a_i], \xi_i)\}$ отрезка $[a, b]$ соответствует разбиение $\mathbb{T}^- = \{([-a_i, -a_{i-1}], -\xi_i)\}$ отрезка $[-b, -a]$ и $\mathfrak{S}(f(x), \mathbb{T}) = \mathfrak{S}(f(-x), \mathbb{T}^-)$, поэтому $\int\limits_{-b}^{-a} f(-x)\, dx$ и $\int\limits_a^b f(x)\, dx$ одновременно существуют и равны или не существуют, а по теореме $\int\limits_\alpha^\beta f(-(-\varphi(t))) \cdot -(\varphi'(t))\, dt = \int\limits_{-b}^{-a} f(-x)\, dx$. ▲

Для функций, имеющих точную первообразную, легко получить другую теорему о замене переменной.

Теорема 4 (формула замены переменной). *Если* $f \in \mathcal{H}[a, b]$*) и имеет точную первообразную* F *на* $[a, b]$*,* φ *— дифференцируемая функция на* $[\alpha, \beta]$*,* $\varphi(\alpha) = a$*,* $\varphi(\beta) = b$*,* $\varphi([\alpha, \beta]) = [a, b]$*, то* $f(\varphi) \cdot \varphi' \in \mathcal{H}[\alpha, \beta]$*) и*

$$\int\limits_\alpha^\beta f(\varphi) \cdot \varphi' \, dt = \int\limits_a^b f \, dx = F(b) - F(a).$$

▼ По следствию из формулы Ньютона–Лейбница $\int\limits_a^b f(x)\, dx = F(b) - F(a)$.

Функция $F(\varphi(t))$ — точная первообразная функции $f(\varphi(t))\varphi'(t)$ на $[\alpha, \beta]$, по тому же следствию $\int\limits_\alpha^\beta f(\varphi(t)) \cdot \varphi'(t)\, dt = F(\varphi(\beta)) - F(\varphi(\alpha)) = F(b) - F(a)$. ▲

Формула Тейлора с остаточным членом в интегральной форме

Теорема 5 (формула Тейлора с остаточным членом в интегральной форме). *Если функция* f $n + 1$ *раз дифференцируема на отрезке с концами* x_0 *и* x*, а* f^{n+1}

интегрируема по Риману на нём, то

$$f(x)=\sum_{k=0}^{n}\frac{f^{(k)}(x_0)}{k!}(x-x_0)^k+\frac{1}{n!}(\mathcal{R})\int\limits_{x_0}^{x}(x-t)^n f^{(n+1)}(t)\,dt,$$

(то есть остаток $r_n(x)=\frac{1}{n!}(\mathcal{R})\int\limits_{x_0}^{x}(x-t)^n f^{(n+1)}(t)\,dt,$*).*

▼ Проведем его индукцией по n. При $n=0$ $\;f(x)=f(x_0)+\int\limits_{x_0}^{x}f'(t)\,dt=f(x_0)+f(x)-f(x_0)$, что верно.

Предположим, что утверждение верно при $n=m$. Докажем его для $n=m+1$. Так как остаточный член $r_{m+1}(x)=\frac{1}{n!}(\mathcal{H})\int\limits_{x_0}^{x}(x-t)^n f^{(n+1)}(t)\,dt=\frac{1}{n!}(\mathcal{R}-\mathcal{S})\int\limits_{x_0}^{x}(x-t)^n\,df^{(n)}(t)=\frac{1}{n!}(x-t)^n f^{(n)}(t)\Big|_{x_0}^{x}-\frac{1}{n!}(\mathcal{R}-\mathcal{S})\int\limits_{x_0}^{x}f^{(n)}(t)\,d(x-t)^n=-\frac{1}{n!}f^{(n)}(x_0)(x-x_0)^n+\frac{1}{m!}(\mathcal{R})\int\limits_{x_0}^{x}(x-t)^m f^{(m+1)}(t)\,dt=-\frac{1}{n!}f^{(n)}(x_0)(x-x_0)^n+r_m(x)$, согласно формулам сведения интеграла Римана к интегралу Римана-Стилтьеса, а его к интегралу Римана, то

$$r_m(x)=\frac{f^{(m+1)}(x_0)}{(m+1)!}(x-x_0)^{m+1}+r_{m+1}(x).$$

В итоге получаем, что $f(x)=\sum\limits_{k=0}^{m}\frac{f^{(k)}(x_0)}{k!}(x-x_0)^k+\frac{1}{m!}(\mathcal{R})\int\limits_{x_0}^{x}(x-t)^m f^{(m+1)}(t)\,dt=\sum\limits_{k=0}^{m+1}\frac{f^{(k)}(x_0)}{k!}(x-x_0)^k+\frac{1}{(m+1)!}(\mathcal{R})\int\limits_{x_0}^{x}(x-t)^{m+1}f^{(m+2)}(t)\,dt$. Значит, утверждение теоремы верно для всех $n\in\mathbb{Z}^+$. ▲

Лекция 15 (27.03.20)
Первая и вторая теоремы о среднем

Первая теорема о среднем

Теорема 1 (для интеграла Римана). *Если функция f ограничена на $[a,b]$, функции fg и g интегрируемы на $[a,b]$ в смысле Римана, g сохраняет знак на $[a,b]$ (т.е. $g(x)\geqslant 0$ на $[a,b]$ или $g(x)\leqslant 0$ на $[a,b]$), то существует λ, $\inf\limits_{[a,b]}f\leqslant\lambda\leqslant\sup\limits_{[a,b]}f$, такое что*

$$\int\limits_a^b fg\,dx=\lambda\int\limits_a^b g\,dx.$$

▼ Рассмотрим случай $\int\limits_a^b g\,dx>0$. Тогда по свойству сохранения неравенств

$$\inf_{[a,b]}f\cdot\int\limits_a^b g\,dx\leqslant\int\limits_a^b fg\,dx\leqslant\sup_{[a,b]}f\cdot\int\limits_a^b g\,dx$$

и значит,

$$\inf_{[a,b]} f \leqslant \lambda = \frac{\int\limits_a^b fg\,dx}{\int\limits_a^b g\,dx} \leqslant \sup_{[a,b]} f.$$

Если $\int\limits_a^b g\,dx < 0$, то переходя от функции g к функции $-g$ сводим этот случай к предыдущему.

Если же $\int\limits_a^b g\,dx = 0$, то в силу сохранения знака $g = 0$ почти всюду на $[a,b]$, $\int\limits_a^b fg\,dx = 0$ и λ можно брать любым. ▲

Следствие 1. *Если $f \in \mathrm{C}[a,b]$, то существует $\theta \in [a,b]$, такое что $f(\theta) = \lambda$, т.е. в этом случае $\int\limits_a^b fg\,dx = f(\theta)\int\limits_a^b g\,dx$.*

Замечание. В частном случае $g(x) \equiv 1$ теорему 1 и следствие из нее также именуют иногда первой теоремой о среднем.

Для интегралов Стилтьеса имеют место аналогичные результаты.

Теорема 4 (для интеграла Римана-Стилтьеса). *Если функция f ограничена на $[a,b]$ и интегрируема по монотонной функции G на $[a,b]$ в смысле Римана-Стилтьеса, то существует λ, $\inf\limits_{[a,b]} f \leqslant \lambda \leqslant \sup\limits_{[a,b]} f$, такое что*

$$\int_a^b f\,dG = \lambda\int_a^b dG = \lambda(G(b) - G(a)).$$

▼ Рассмотрим случай $\int\limits_a^b dG = G(b) - G(a) > 0$. Тогда в силу сохранения неравенств

$$\inf_{[a,b]} f \cdot \int_a^b dG \leqslant \int_a^b f\,dG \leqslant \sup_{[a,b]} f \cdot \int_a^b dG$$

и значит,

$$\inf_{[a,b]} f \leqslant \lambda = \frac{\int\limits_a^b f\,dG}{G(b) - G(a)} \leqslant \sup_{[a,b]} f.$$

Если $G(b) - G(a) < 0$, то, переходя от функции G к функции $-G$, сводим этот случай к предыдущему.

Если $G(b) - G(a) = 0$, то G постоянна, интеграл от любой функции по G равен нулю и λ можно брать любым. ▲

Следствие 2. *Если $f \in \mathrm{C}[a,b]$, то существует $\theta \in [a,b]$, такое что $f(\theta) = \lambda$, т.е. в этом случае $\int\limits_a^b f\,dG = f(\theta)\int\limits_a^b dG = f(\theta)(G(b) - G(a))$.*

Вторая теорема о среднем

Теорема 5 (для интеграла Римана). *Если функция* $f \in \mathcal{R}[a,b]$, *а* g — *монотонная функция на* $[a,b]$, *то существует такое* $\xi \in [a,b]$, *что*

$$(\mathcal{R}) \int\limits_a^b fg\,dx = g(a) \int\limits_a^\xi f\,dx + g(b) \int\limits_\xi^b f\,dx.$$

Если впридачу g *неотрицательна и невозрастает, то существует такое* $\zeta \in [a,b]$, *что*

$$(\mathcal{R}) \int\limits_a^b fg\,dx = g(a) \int\limits_a^\zeta f\,dx,$$

а если g *неотрицательна и неубывает, то существует такое* $\zeta \in [a,b]$, *что*

$$(\mathcal{R}) \int\limits_a^b fg\,dx = g(b) \int\limits_\zeta^b f\,dx.$$

Для интеграла Римана–Стилтьеса имеет место аналогичный результат.

Теорема 6 (для интеграла Римана-Стилтьеса). *Если функция* $F \in \mathrm{C}[a,b]$, *а* g — *монотонная функция на* $[a,b]$, *то существует такое* $\xi \in [a,b]$, *что*

$$(\mathcal{R}-\mathcal{S}) \int\limits_a^b g\,dF = g(a) \int\limits_a^\xi dF + g(b) \int\limits_\xi^b dF =$$
$$= g(a)(F(\xi) - F(a)) + g(b)(F(b) - F(\xi)).$$

Если впридачу g *неотрицательна и невозрастает, то существует такое* $\zeta \in [a,b]$, *что*

$$(\mathcal{R}-\mathcal{S}) \int\limits_a^b g\,dF = g(a) \int\limits_a^\zeta dF = g(a)(F(\zeta) - F(a)),$$

а если g *неотрицательна и неубывает, то существует такое* $\zeta \in [a,b]$, *что*

$$(\mathcal{R}-\mathcal{S}) \int\limits_a^b g\,dF = g(b) \int\limits_\zeta^b dF = g(b)(F(b) - F(\zeta)).$$

▼ Докажем эти две теоремы.

Если $F(x)$ — неопределенный интеграл f, то по теореме о сведении интеграла Римана к интегралу Римана–Стилтьеса

$$(\mathcal{R}) \int\limits_a^b fg\,dx = (\mathcal{R}-\mathcal{S}) \int\limits_a^b g\,dF$$

и утверждения теоремы 5 следуют из соответствующих утверждений теоремы 6, к доказательству которой приступим.

Так как g монотонна, а F непрерывна на $[a.b]$, то g интегрируема по F на $[a.b]$ в смысле Римана–Стилтьеса. По формуле интегрирования по частям

$$\int\limits_a^b g\,dF = g(b)F(b) - g(a)F(a) - \int\limits_a^b F\,dg.$$

Если g неубывает на $[a,b]$, то по свойству сохранения неравенств

$$\min_{[a,b]} F \cdot \int\limits_a^b dg \leqslant \int\limits_a^b F\,dg \leqslant \max_{[a,b]} F \cdot \int\limits_a^b dg$$

и значит, в силу непрерывности F существует такое $\xi \in [a,b]$, что

$$\int\limits_a^b F\,dg = F(\xi)\int\limits_a^b dg = F(\xi)(g(b) - g(a)).$$

Отсюда получаем равенство

$$\int\limits_a^b g\,dF = g(b)F(b) - g(a)F(a) - \int\limits_a^b F\,dg =$$
$$= g(a)(F(\xi) - F(a)) + g(b)(F(b) - F(\xi)),$$

которое и является утверждением теоремы.

А если g невозрастает, то перейдем от g к $-g$ и сведем этот случай к предыдущему.

Если g неотрицательна и не возрастает на $[a,b]$, то $g(b) \geqslant 0$ и

$$\min_{[a,b]} F \cdot g(b) \leqslant g(b)F(b) \leqslant \max_{[a,b]} F \cdot g(b)$$

и по свойству сохранения неравенств

$$\min_{[a,b]} F \cdot (g(a) - g(b)) \leqslant \int\limits_a^b F\,d(-g) \leqslant \max_{[a,b]} F \cdot (g(a) - g(b)),$$

ведь $\int\limits_a^b d(-g) = g(a) - g(b)$. Складывая написанные неравенства, получим

$$\min_{[a,b]} F \cdot g(a) \leqslant g(b)F(b) - \int\limits_a^b F\,dg \leqslant \max_{[a,b]} F \cdot g(a).$$

Из последнего неравенства и из непрерывности F следует, что существует такое $\zeta \in [a,b]$, что

$$g(a)F(\zeta) = g(b)F(b) - \int\limits_a^b F\,dg,$$

т.е.
$$\int\limits_a^b g\,dF = g(b)F(b) - g(a)F(a) - \int\limits_a^b F\,dg == g(a)(F(\zeta) - F(a)).$$

Случай, когда g неотрицательна и неубывает рассматривается аналогично. Он также может быть сведет к предыдущему случаю рассмотрением функций $f(-x)$ и $g(-x)$ на отрезке $[-b, -a]$. ▲

Лекция 16 (07.04.20)
Несобственные интегралы

Несобственные интегралы. Определения и простейшие свойства

Определение 1. Если функция f определена на полуотрезке $[a, b)$, где $a \in \mathbb{R}$, $b \in \overline{\mathbb{R}}$, $a < b$ (т.е. a и b из $\mathbb{R}$ и $a < b$ или $a \in \mathbb{R}$, $b = +\infty$), интегрируема в одном из двух смыслов (по Риману или по Курцвейлю–Хенстоку) на любом $[a, b']$ при $b' \in (a, b)$ и существует
$$\lim_{b' \to b-0} \int\limits_a^{b'} f\,dx,$$
то этот предел называют **несобственным интегралом** от f на (по) $[a, b)$ (или на (по) $[a, b]$ для $b \neq +\infty$) в соответствующем смысле и обозначают
$$\int\limits_a^b f\,dx \text{ или } \int\limits_{[a,b)} f\,dx.$$

Аналогично, если функция f определена на полуотрезке $(a, b]$, где $a \in \overline{\mathbb{R}}$, $b \in \mathbb{R}$, $a < b$, интегрируема в одном из двух смыслов (по Риману или по Курцвейлю–Хенстоку) на любом $[a', b]$ при $a' \in (a, b)$ и существует
$$\lim_{a' \to a+0} \int\limits_{a'}^{b} f\,dx,$$
то этот предел называют **несобственным интегралом** от f на (по) $(a, b]$ (или на (по) $[a, b]$ для $a \neq -\infty$) в соответствующем смысле и обозначают
$$\int\limits_a^b f\,dx \text{ или } \int\limits_{(a,b]} f\,dx.$$

В первом случае говорят про несобственный интеграл **с особенностью** в b, а во втором случае про несобственный интеграл **с особенностью** в a.

О несобственном интеграле Римана

Теорема 1. *Если функция f интегрируема в несобственном смысле на $[a, b)$ (с особенностью в точке b) по Риману и ограничена на $[a, b)$, $b \neq +\infty$, то при любом ее доопределении в точке b она будет интегрируема по Риману на $[a, b]$ и интеграл Римана по $[a, b]$ будет совпадать с несобственным интегралом по $[a, b)$.*

▼ Возьмем последовательность точек $b'_n \in [a, b)$, $b'_n \to b$ при $n \to \infty$. Из условия интегрируемости f на $[a, b'_n]$ и критерия интегрируемости Лебега получаем, что f непрерывна п.в. (почти всюду) на $[a, b'_n]$, а значит, по свойству „не более чем счетное объединение множеств меры нуль — множество меры нуль, и на $\bigcup\limits_{n=1}^{\infty} [a, b'_n] = [a, b)$. При любом дооопределении в точке b функция f останется ограниченной и непрерывной п.в. на $[a, b)$, а значит, и на $[a, b]$. Но тогда по критерию Лебега f будет интегрируема по Риману на $[a, b]$. Из свойства непрерывности интеграла Римана с переменным верхним пределом следует совпадение интеграла Римана с несобственным интегралом. ▲

Аналогичная теорема справедлива и для интегралов с особенностью в точке a на полуотрезке $(a, b]$. То же самое можно говорить и о всех дальнейших утверждениях про несобственные интегралы, которые будут формулироваться только для интегралов с особенностью в точке b на полуотрезке $[a, b)$, но их очевидные аналоги имеют место и для несобственных интегралов с особенностью в точке a на полуотрезке $(a, b]$.

Стоит также отметить, что в случае конечной точки b по доказанной теореме представляют интерес для несобственного интегрирования только функции неограниченные на $[a, b)$, а значит, и на $[b', b)$ для любого $b' \in (a, b)$.

Критерий Коши несобственной интегрируемости

Всюду дальше считаем, что функция f определена на полуотрезке $[a, b)$, где $a \in \mathbb{R}$, $b \in \overline{\mathbb{R}}$, $a < b$, и интегрируема на любом $[a, b']$ при $b' \in (a, b)$ по Риману.

Определение 2. Функция f удовлетворяет **условию Коши несобственной интегрируемости** на $[a, b)$ (с особенностью в b), если f определена на $[a, b)$, интегрируема на любом $[a, b']$ при $b' \in (a, b)$ по Риману и

$$\forall \varepsilon > 0\ \exists B_\delta(b)\ \forall b', b'' \in B_\delta(b) \cap [a, b): \quad \left| \int\limits_{b'}^{b''} f\, dx \right| < \varepsilon$$

или, что эквивалентно,

$$\forall \varepsilon > 0\ \exists \tilde{b} \in [a, b)\ \forall b', b'' \in (\tilde{b}, b): \quad \left| \int\limits_{b'}^{b''} f\, dx \right| < \varepsilon.$$

Эквивалентность очевидна, если отождествить $\tilde{b}$ с левым концом промежутка $B_\delta(b) \cap [a, b)$.

Теорема 2. *Для интегрируемости функции f в несобственном смысле на $[a, b)$ (с особенностью в b) по Риману необходимо и достаточно выполнения условия Коши несобственной интегрируемости для f на $[a, b)$ для интеграла Римана.*

▼ Пусть $F(x) = \int\limits_a^x f\,dx$ — неопределенный интеграл. Так как $\int\limits_{b'}^{b''} f\,dx = F(b'') - F(b')$, то приведенный критерий Коши — переформулировка критерия Коши существования предела функции F (в точке b по множеству $[a, b)$,

$$\lim_{b' \to b-0} F(b') = \lim_{b' \to b-0} \int\limits_a^{b'} f\,dt.$$

Значит, утверждение теоремы верно. ▲

Примеры

а) Интеграл с особенностью в точке $+\infty$

$$\int\limits_0^{+\infty} e^{-x} dx = \lim_{b' \to +\infty} \int\limits_0^{b'} e^{-x} dx = \lim_{b' \to +\infty} \left(1 - e^{-b'}\right) = 1.$$

б) Интеграл с особенностью в точке $+\infty$

$$\int\limits_1^{+\infty} x^\alpha dx = \lim_{b' \to +\infty} \int\limits_1^{b'} x^\alpha dx = \begin{cases} \lim\limits_{b' \to +\infty} \left(\frac{(b')^{\alpha+1}}{\alpha+1} - \frac{1}{\alpha+1}\right), & \alpha \neq -1, \\ \lim\limits_{b' \to +\infty} \ln b', & \alpha = -1. \end{cases}$$

Легко видеть, что при $\alpha + 1 < 0$ предел существует и равен $-\frac{1}{\alpha+1}$, при $\alpha + 1 \geqslant 0$ предел не существует.

в) Интеграл с особенностью в точке 0

$$\int\limits_0^1 x^\alpha dx = \lim_{a' \to +0} \int\limits_{a'}^1 x^\alpha dx = \begin{cases} \lim\limits_{a' \to +0} \left(\frac{1}{\alpha+1} - \frac{(a')^{\alpha+1}}{\alpha+1}\right), & \alpha \neq -1, \\ \lim\limits_{a' \to +0} -\ln a', & \alpha = -1. \end{cases}$$

▼ Легко видеть, что при $\alpha + 1 > 0$ предел существует и равен $\frac{1}{\alpha+1}$, при $\alpha + 1 \leqslant 0$ предел не существует. Отметим, что при $\alpha \geqslant 0$ $x^\alpha \in \mathcal{R}[0, 1]$. ▲

Еще отметим, что пользуясь теоремой о замене переменной в интеграле можно любой несобственный интеграл (даже по бесконечному полуотрезку) сводить к несобственному интегралу по $[0, 1)$. Для случая конечного полуотрезка $[a, b)$ подходит замена $\varphi(t) = a + (b - a)t$, а если $b = \infty$, то замена $\varphi(t) = a - 1 + \frac{1}{1-t}$.

Определение 3. Если конечный несобственный интеграл существует, то говорят, что он **сходится**, если не существует, то говорят, что он **не сходится** или **расходится**.

Абсолютная сходимость несобственного интеграла

Определение 4. Несобственный интеграл от f на $[a, b)$ **сходится абсолютно**, если f интегрируема на любом $[a, b'] \subset [a, b]$ в соответствующем смысле и сходится интеграл от абсолютной величины подынтегральной функции $|f|$ на $[a, b)$.

Теорема 3. *Абсолютно сходящийся интеграл сходится (т.е. если несобственный интеграл от f на $[a,b)$ сходится абсолютно, то он сходится).*

▼ Если интеграл абсолютно сходится, то функция $|f|$ удовлетворяет условию Коши несобственной интегрируемости на $[a,b)$, а тогда в силу неравенства

$$\left|\int\limits_{b'}^{b''} f\,dx\right| \leqslant \int\limits_{b'}^{b''} |f|\,dx$$

при $a \leqslant b' \leqslant b'' < b$ и функция f удовлетворяет условию Коши несобственной интегрируемости на $[a,b)$ и, значит, несобственный интеграл от f на $[a,b)$ сходится. ▲

Определение 5. Если несобственный интеграл от f на $[a,b)$ сходится, но не сходится абсолютно, то говорят, что он сходится **условно**.

Признаки сравнения

Теорема 4. *Пусть функции f и g определены на $[a,b)$, интегрируемы по Риману на любом $[a,b'] \subset [a,b)$ и $0 \leqslant f(x) \leqslant g(x)$ на $[a,b)$. Тогда из сходимости несобственного интеграла $\int\limits_{[a,b)} g\,dx$ Римана следует сходимость несобственного интеграла $\int\limits_{[a,b)} f\,dx$ (в соответствующем смысле) и неравенство*

$$0 \leqslant \int\limits_{[a,b)} f\,dx \leqslant \int\limits_{[a,b)} g\,dx,$$

а из расходимости интеграла $\int\limits_{[a,b)} f\,dx$ следует расходимость интеграла $\int\limits_{[a,b)} g\,dx$.

▼ Если интеграл $\int\limits_{[a,b)} g\,dx$ сходится, то для функции g выполнено условие Коши несобственной интегрируемости на $[a,b)$, а в силу неравенства

$$0 \leqslant \int\limits_{b'}^{b''} f\,dx \leqslant \int\limits_{b'}^{b''} g\,dx,$$

при $a \leqslant b' \leqslant b'' < b$ и функция f удовлетворяет условию Коши несобственной интегрируемости на $[a,b)$ и, значит, несобственный интеграл от f на $[a,b)$ сходится. Полагая в написанном неравенстве $b' = a$ и устремляя b'' к b получаем, что

$$0 \leqslant \int\limits_{[a,b)} f\,dx \leqslant \int\limits_{[a,b)} g\,dx.$$

Из доказанного видно, что если интеграл $\int\limits_{[a,b)} f\,dx$ расходится, то интеграл $\int\limits_{[a,b)} g\,dx$ сходится не может. ▲

Замечание 1. *В теореме достаточно выполнения неравенства $0 \leqslant f(x) \leqslant g(x)$ на $[a',b)$ для некоторого $a' \in (a,b)$.*

▼ Действительно, в силу равенства

$$\int\limits_a^{b'} f\,dx = \int\limits_a^{a'} f\,dx + \int\limits_{a'}^{b'} f\,dx$$

при $b' \to b-0$ пределы первого и последнего из интегралов одновременно существуют или не существуют. Значит, f интегрируема на $[a,b)$ тогда и только тогда, когда интегрируема на $[a',b)$. Отсюда и из аналогичного утверждения для функции g следует справедливость первой части замечания. ▲

Теорема 5. *Пусть функции f и g определены и строго положительны на $[a,b)$, интегрируемы по Риману на любом $[a,b'] \subset [a,b)$ и*

$$0 < C_1 \leqslant \frac{f(x)}{g(x)} \leqslant C_2 < \infty$$

на $[a,b)$. Тогда несобственные интегралы от этих функций одновременно сходятся или расходятся.

▼ Так как $f(x) \leqslant C_2 g(x)$ и $g(x) \leqslant \frac{1}{C_1} f(x)$ на $[a,b)$, то по предыдущей теореме несобственные интегралы от функций f и g одновременно сходятся или расходятся на $[a,b)$. ▲

Замечание 2. *Как и в предыдущей теореме, достаточно выполнения неравенства $0 < C_1 \leqslant \frac{f(x)}{g(x)} \leqslant C_2 < \infty$ на $[a',b)$ для некоторого $a' \in (a,b)$.*

Признаки Абеля и Дирихле

Теорема 6 (признак Абеля). *Если существует несобственный интеграл Римана от функции f на $[a,b)$, а φ — VB-функция на $[a,b)$, то сходится несобственный интеграл Римана от функции $f\varphi$ на $[a,b)$.*

Теорема 7 (признак Дирихле). *Если интегралы Римана $\int\limits_a^x f\,dt$ существуют и ограничены (как функции от x) на $[a,b)$, а φ — VB-функция на $[a,b)$ и $\lim\limits_{x\to b-0} \varphi(x) = 0$, то сходится несобственный интеграл Римана от функции $f\varphi$ на $[a,b)$.*

▼ Для обеих теорем проведем доказательство одновременно. Достаточно доказать, что существует $\lim\limits_{b'\to b-0} (\mathcal{R}) \int\limits_a^{b'} f\varphi\,dt$. Пусть $a \leqslant b' \leqslant b'' < b$,

$$(\mathcal{R}) \int\limits_{b'}^{b''} f\varphi\,dx = (\mathcal{R}-\mathcal{S}) \int\limits_{b'}^{b''} \varphi\,dF = F(b'')\varphi(b'') - (\mathcal{R}-\mathcal{S}) \int\limits_{b'}^{b''} F\,d\varphi,$$

где F такая первообразная f, что $F(b') = 0$.

В признаке Абеля f интегрируема на $[a,b)$, а значит, удовлетворяет признаку Коши несобственной интегрируемости, для любого $\varepsilon > 0$ найдется такое $\tilde{b} \in [a,b)$, что для любых $b', b'' \in (\tilde{b}, b)$ выполняется неравенство $\left|\int\limits_{b'}^{b''} f\,dx\right| < \varepsilon$, а значит, $F(b'') < \varepsilon$ и $F(x) < \varepsilon$ на $[b',b'']$. Получаем оценку

$$\left|(\mathcal{R}) \int\limits_{b'}^{b''} f\varphi\,dx\right| < \varepsilon \sup_{[a,b)} |\varphi| + \varepsilon \operatorname*{Var}_{[a,b)} \varphi = \left(\sup_{[a,b)} |\varphi| + \operatorname*{Var}_{[a,b)} \varphi\right) \varepsilon.$$

По критерию Коши несобственный интеграл от функции $f\varphi$ по $[a,b)$ существует.

По определению вариации для любого $\varepsilon > 0$ существует такой упорядоченный в порядке возрастания конечный набор точек $D = \{a_i\}_{i=0}^n$ из $[a,b)$, что $\sum\limits_{i=1}^n |\varphi(a_i) - \varphi(a_{i-1})| > \operatorname*{Var}\limits_{[a,b)} \varphi - \varepsilon$, тогда $\operatorname*{Var}\limits_{[a,a_n]} > \operatorname*{Var}\limits_{[a,b)} \varphi - \varepsilon$, а по аддитивности вариации по отрезкам для любого $b'' \in [a_n, b)$ $\operatorname*{Var}\limits_{[a_n,b'']} \varphi < \varepsilon$. По условию $\lim\limits_{x\to b-0} \varphi(x) = 0$. Следовательно, для любого $\varepsilon > 0$ найдется такое $\tilde{b} \in [a,b)$, что для любого $b'' \in (\tilde{b}, b)$ выполняется неравенство $\varphi(b'') < \varepsilon$ и что для любых $b', b'' \in (\tilde{b}, b)$ выполняется неравенство $\operatorname*{Var}\limits_{[b',b'']} \varphi < \varepsilon$. Пользуясь ограниченностью вариации по подмножеству и тем, что $F(x) = \int\limits_{b'}^{x} f\,dt = \left(\int\limits_a^x - \int\limits_a^{b'}\right) f\,dt$ и, значит, $|F(x)| \leqslant 2 \sup\limits_{x\in[a,b)} \left|\int\limits_a^x f\,dt\right|$, получаем оценку

$$\left|(\mathcal{R}) \int\limits_{b'}^{b''} f\varphi\,dx\right| \leqslant |F(b'')|\varepsilon + \sup_{[b',b'']} |F|\varepsilon \leqslant 4 \sup_{x\in[a,b)} \left|\int\limits_a^x f\,dt\right| \varepsilon,$$

По критерию Коши несобственный интеграл от функции $f\varphi$ по $[a,b)$ существует. ▲

МНОГОМЕРНЫЕ ПРОСТРАНСТВА И ФУНКЦИИ МНОГИХ ПЕРЕМЕННЫХ

Лекция 17 (10.04.20) Метрические и нормированные пространства. Пространство $\mathbb{R}^n$

Метрическое пространство

Определение 1. Метрическим пространством называется пара (M, ρ), где M — множество, а ρ — **метрика** или **расстояние**, функция из $M \times M$ в $\mathbb{R}$ (т.е. функция пары точек из M со значениями в множестве действительных чисел) со следующими свойствами:

1)$\forall x, y \in M: \quad \rho(x,y) \geqslant 0$ и $\rho(x,y) = 0 \Longleftrightarrow x = y$;

2) $\forall x, y \in M: \quad \rho(x,y) = \rho(y,x)$ — симметричность;

3) $\forall x, y, z \in M: \quad \rho(x,z) \leqslant \rho(x,y) + \rho(y,z)$ — неравенство треугольника.

Примеры. 1) M — любое множество, $\rho(x,y) = \begin{cases} 0, & \text{если } x = y, \\ 1, & \text{если } x \neq y, \end{cases}$ — тривиальная метрика;

2) $M = \mathbb{R}$, $\rho(x,y) = |x-y|$;

3) $M = \mathbb{R}^2$, $\rho((x_1,x_2),(y_1,y_2)) = \sqrt{(x_1-y_1)^2 + (x_2-y_2)^2}$;

4) $M = \mathbb{R}^2$, $\rho((x_1,x_2),(y_1,y_2)) = |x_1-y_1| + |x_2-y_2|$.

В дальнейшем в соответствии с традицией часто будем называть метрическим пространством само множество M, подразумевая при этом наличие связанной с M метрики ρ.

Отметим такой очевидный факт: подмножество метрического пространства — метрическое пространство с той же метрикой (т.е. если (M, ρ) — метрическое пространство, $M' \subset M$, то (M', ρ) также метрическое пространство).

В дальнейшем нам понадобится лемма.

Лемма (обобщение неравенства треугольника). *Для любых точек x_k, $k = 1, \ldots, n$, метрического пространства*

$$\rho(x_1, x_n) \leqslant \sum_{k=1}^{n-1} \rho\left(x_k, x_{k+1}\right).$$

▼ Для $n = 1$ имеем верное неравенство $0 \leqslant 0$, для $n = 2$ также имеем верное неравенство $\rho\left(x_1, x_2\right) \leqslant \rho\left(x_1, x_2\right)$, для $n = 3$ имеем верное неравенство треугольника $\rho\left(x_1, x_3\right) \leqslant \rho\left(x_1, x_2\right) + \rho\left(x_2, x_3\right)$. Теперь проведем индукцию по n. Предположим, что неравенство верно для $n = m \geqslant 3$ и докажем, что оно верно для $n = m + 1$. В самом деле, $\rho\left(x_1, x_n\right) \leqslant \rho\left(x_1, x_m\right) + \rho\left(x_m, x_{m+1}\right) \leqslant \sum\limits_{k=1}^{m-1} \rho\left(x_k, x_{k+1}\right) + \rho\left(x_m, x_{m+1}\right) = \sum\limits_{k=1}^{m} \rho\left(x_k, x_{k+1}\right)$. ▲

Нормированное пространство

Определение 2. Нормированным пространством называется пара $(N, \|\cdot\|)$, где N — линейное пространство над полем $\mathbb{R}$ действительных или $\mathbb{C}$ комплексных чисел, а $\|\cdot\|$ — **норма** или **длина** вектора, функция из N в $\mathbb{R}$ (т.е. функция точки из N со значениями в множестве действительных чисел) со следующими свойствами:

1)$\forall x, \in N: \ \|x\| \geqslant 0$ и $\|x\| = 0 \Longleftrightarrow x = 0$;

2) $\forall x \in N$ $\forall \alpha$ из $\mathbb{R}$ или $\mathbb{C}$ (в зависимости от того, над каким полем N является линейным пространством): $\|\alpha x\| = |\alpha| \cdot \|x\|$;

3) $\forall x, y \in N : \|x + y\| \leqslant \|x\| + \|y\|$ — неравенство треугольника.

Утверждение. *Всякое нормированное пространство является метрическим пространством с метрикой $\rho(x, y) = \|x - y\|$.*

▼ Действительно, свойство 1) метрики сразу следует из свойства 1) нормы. Из свойства 2) нормы следует выполнение свойства 2) метрики: $\rho(x, y) = \|x - y\| = \|-1 \cdot (y - x)\| = |-1| \cdot \|y - x\| = \|y - x\| = \rho(y, x)$. Из свойства 3) нормы следует выполнение свойства 3) метрики: $\rho(x, z) = \|x - z\| = \|(x - y) + (y - z)\| \leqslant \|x - y\| + \|y - z\| = \rho(x, y) + \rho(y, z)$. ▲

Пространство $\mathbb{R}^n$

Определение 1. Пространство $\mathbb{R}^n$ — множество упорядоченных наборов из n действительных чисел $\vec{x} = (x_1, x_2, \ldots, x_n)$, $x_k \in \mathbb{R}$, $k = 1, \ldots, n$. Его можно рассматривать как векторное пространство над полем действительных чисел $\mathbb{R}$ с определенными следующим образом операциями сложения векторов и умножения вектора на число: для любых $\vec{x}, \vec{y} \in \mathbb{R}^n$

$$\vec{x} + \vec{y} = (x_1, \ldots, x_n) + (y_1, \ldots, y_n) = (x_1 + y_1, \ldots, x_n + y_n),$$

для любого $\vec{x} \in \mathbb{R}^n$, для любого $\alpha \in \mathbb{R}$

$$\alpha \cdot \vec{x} = \alpha(x_1, \ldots, x_n) = (\alpha x_1, \ldots, \alpha x_n).$$

В этом пространстве вектора

$$\vec{e}_1 = (1, 0, 0, \ldots, 0),\ \vec{e}_2 = (0, 1, 0, \ldots, 0),\ \ldots,\ \vec{e}_n = (0, \ldots, 0, 1)$$

образуют стандартный базис, а числа $x_1, \ldots, x_n$ называют **координатами** вектора $\vec{x} = (x_1, \ldots, x_n)$.

Теперь докажем одно полезное неравенство.

Теорема 1 (неравенство Коши–Буняковского). *Для любых действительных чисел a_k и b_k, $k = 1, \ldots, n$, справедливо неравенство*

$$\left|\sum_{k=1}^{n} a_k b_k\right| \leqslant \sqrt{\sum_{k=1}^{n} a_k^2}\sqrt{\sum_{k=1}^{n} b_k^2},$$

причем равенство возможно лишь в случае, если один из наборов $a_1, \ldots, a_n$, $b_1, \ldots, b_n$ пропорционален другому (т.е. получается из него умножением всех членов на одно и то же число).

▼ Это неравенство является частным случаем неравенства Гёльдера из первого семестра. Здесь дадим его независимое доказательство.

Если все a_k равны нулю, то написанное неравенство превращается в равенство, при этом $a_k = 0 \cdot b_k$, $k = 1, \ldots, n$. Если не все a_k равны нулю, то рассмотрим квадратный относительно x трехчлен

$$\sum_{k=1}^{n} (a_k x - b_k)^2 = \left(\sum_{k=1}^{n} a_k^2\right) x^2 - \left(2\sum_{k=1}^{n} a_k b_k\right) x + \left(\sum_{k=1}^{n} b_k^2\right).$$

Очевидно, он не имеет двух различных корней (ведь квадратный трехчлен неотрицателен), а единственный корень x_0 существует лишь в случае, если $a_k x_0 - b_k = 0$, $k = 1, \ldots, n$, что означает пропорциональность чисел b_k числам a_k. Значит, дискриминант уравнения

$$D = 4\left(\sum_{k=1}^{n} a_k b_k\right)^2 - 4\left(\sum_{k=1}^{n} a_k^2\right) \cdot \left(\sum_{k=1}^{n} b_k^2\right) \leqslant 0,$$

причем равенство имеет место лишь в случае пропорциональности чисел b_k числам a_k. Но последнее — лишь несколько иначе написанное утверждение теоремы, которая тем самым доказана. ▼

Теорема 2. *Пространство $\mathbb{R}^n$ — нормированное пространство с нормой*

$$\|\vec{x}\| = \|(x_1, \ldots, x_n)\| = \sqrt{\sum_{k=1}^{n} x_k^2}.$$

▼ Очевидно выполнение всех свойств нормы, кроме неравенства треугольника. поэтому остается только доказать, что

$$\|\vec{x} + \vec{y}\| = \sqrt{\sum_{k=1}^{n} (x_k + y_k)^2} \leqslant \sqrt{\sum_{k=1}^{n} x_k^2} + \sqrt{\sum_{k=1}^{n} y_k^2} = \|\vec{x}\| + \|\vec{y}\|.$$

Так как обе части неравенства неотрицательны, то возводя их в квадрат получаем эквивалентное неравенство

$$\sum_{k=1}^{n}(x_k+y_k)^2=\sum_{k=1}^{n}x_k^2+2\sum_{k=1}^{n}x_ky_k+\sum_{k=1}^{n}y_k^2\leqslant$$
$$\leqslant\sum_{k=1}^{n}x_k^2+2\sqrt{\sum_{k=1}^{n}x_k^2}\sqrt{\sum_{k=1}^{n}y_k^2}+\sum_{k=1}^{n}y_k^2,$$

которое после уничтожения одинаковых в обеих частях неравенства членов становится неравенством

$$2\sum_{k=1}^{n}x_ky_k\leqslant 2\sqrt{\sum_{k=1}^{n}x_k^2}\sqrt{\sum_{k=1}^{n}y_k^2},$$

которое после сокращения на 2 отличается от доказанного в предыдущей теореме неравенства Коши-Буняковского лишь отсутствием модуля в левой части. Значит, неравенство треугольника выполняется. ▲

Из теоремы следует, что все дальнейшие результаты, относящиеся к метрическим и нормированным пространствам, справедливы и для $\mathbb{R}^n$.

Классификация точек

Вернемся к метрическому пространству (M,ρ).

Определение 2. Открытым шаром радиуса $r>0$ с центром в точке x называется множество

$$B_r(x)=\{y\in M:\ \rho(x,y)<r\}.$$

Определение 3. Замкнутым шаром радиуса $r\geqslant 0$ с центром в точке x называется множество

$$\overline{B}_r(x)=\{y\in M:\ \rho(x,y)\leqslant r\}.$$

Определение 4. Под **ε-окрестностью** (иногда в дальнейшем называемой просто **окрестностью**) точки $x\in M$, где $\varepsilon>0$, понимается открытый шар $B_r(x)$.

Определение 5. Под **проколотой ε-окрестностью** точки $x\in M$, где $\varepsilon>0$, понимается множество $B'_\varepsilon(x)=B_\varepsilon(x)\setminus\{x\}=\{y\in M:\ 0<\rho(x,y)<r\}$.

Определение 6. Точка x метрического пространства M называется **внутренней** точкой множества $E\subset M$, если существует $B_\varepsilon(x)\subset E$.

Определение 7. Точка x метрического пространства M называется **внешней** точкой множества $E\subset M$, если существует $B_\varepsilon(x)\subset M\setminus E$ — дополнению множества E.

Определение 8. Точка x метрического пространства M называется **граничной** точкой множества $E\subset M$, если для любой ε-окрестности x $B_\varepsilon(x)\cap E\neq\varnothing$ и $B_\varepsilon(x)\cap(M\setminus E)\neq\varnothing$.

По отношению к множеству E метрического пространства любая точка является точкой одного вида из этих трех.

Определение 9. Внутренностью множества E метрического пространства называется множество внутренних точек E.

Определение 10. Внешностью множества E метрического пространства называется множество внешних точек E.

Определение 11. Границей множества E метрического пространства называется множество граничных точек E.

Открытые и замкнутые множества

Определение 12. Множество E метрического пространства называется **открытым**, если оно совпадает со своей внутренностью (т.е., если все его точки внутренние).

Определение 13. Множество E метрического пространства называется **замкнутым**, если $M \setminus E$ открытое множество.

Множества $\varnothing$ и M открыты и замкнуты одновременно. Другие множества из M могут быть открытыми, могут быть замкнутыми (в том числе одновременно и открытыми и замкнутыми), могут быть и не открытыми и не замкнутымы.

Теорема 3. *Открытый шар — открытое множество, замкнутый шар — замкнутое множество.*

▼ Пусть $y \in B_r(x)$, т.е. $\rho(x,y) < r$. Найдем такое $\varepsilon > 0$, что $\rho(x,y) + \varepsilon < r$ (можно взять $\varepsilon = \frac{1}{2}(r - \rho(x,y))$). Покажем, что $B_\varepsilon(y) \subset B_r(x)$ и, значит, все точки $B_r(x)$ внутренние. Если $z \in B_\varepsilon(y)$, то по неравенству треугольника $\rho(x,z) \leqslant \rho(x,y) + \rho(y,z) < \rho(x,y) + \varepsilon < r$ и, значит, $z \in B_r(x)$, следовательно, $B_\varepsilon(y) \subset B_r(x)$. Доказано, что открытый шар — открытое множество.

Пусть $y \in M$ и $y \notin \overline{B}_r(x)$, т.е. $\rho(x,y) > r$. Найдем такое $\varepsilon > 0$, что $\rho(x,y) - \varepsilon > r$ (можно взять $\varepsilon = \frac{1}{2}(\rho(x,y) - r)$). Покажем, что $B_\varepsilon(y) \cap \overline{B}_r(x) = \varnothing$ и, значит, все точки $M \setminus \overline{B}_r(x)$ внутренние. Если $z \in B_\varepsilon(y)$, то по неравенству треугольника $\rho(x,z) \geqslant \rho(x,y) - \rho(y,z) > \rho(x,y) - \varepsilon > r$ и, значит, $z \notin \overline{B}_r(x)$, следовательно, $B_\varepsilon(y) \cap \overline{B}_r(x) = \varnothing$. Доказано, что $M \setminus E$ — открытое множество, т.е., что замкнутый шар — замкнутое множество. ▲

Лекция 18 (14.04.20)
Открытые и замкнутые множества.
Компакты в метрических пространствах

Открытые и замкнутые множества

Теорема 1. *Любые объединения и конечные пересечения открытых множеств — открытые множества. Любые пересечения и конечные объединения замкнутых множеств — замкнутые множества.*

▼ Сначала докажем, что любое объединение открытых множеств является открытым множеством. Пусть G_λ, $\lambda \in \Lambda$, — открытые множества. Проверим, что $\bigcup\limits_{\lambda\in\Lambda} G_\lambda$ — открытое множество, т.е., что любая точка $x \in \bigcup\limits_{\lambda\in\Lambda} G_\lambda$ — внутренняя точка $\bigcup\limits_{\lambda\in\Lambda} G_\lambda$. Действительно, если $x \in \bigcup\limits_{\lambda\in\Lambda} G_\lambda$, то существует такое $\lambda' \in \Lambda$, что $x \in G_{\lambda'}$. Так как $G_{\lambda'}$ — открытое множество, то существует $B_\varepsilon(x) \subset G_{\lambda'} \subset \bigcup\limits_{\lambda\in\Lambda} G_\lambda$ и, значит, x — внутренняя точка $\bigcup\limits_{\lambda\in\Lambda} G_\lambda$.

Теперь докажем, что конечное пересечение открытых множеств является открытым множеством. Пусть $G_1, \ldots, G_n$ — открытые множества. Проверим, что их пересечение $\bigcap\limits_{k=1}^{n} G_k$ — открытое множество, т.е., что любая точка $x \in \bigcap\limits_{k=1}^{n} G_k$ — внутренняя точка $\bigcap\limits_{k=1}^{n} G_k$. Действительно, если $x \in \bigcap\limits_{k=1}^{n} G_k$, то $x \in G_k$, $k = 1, \ldots, n$. Так как G_k, $k = 1, \ldots, n$, — открытые множества, то существуют $B_{\varepsilon_k}(x) \subset G_k$, $k = 1, \ldots, n$. Возьмем $\varepsilon = \min\limits_{1 \leqslant k \leqslant n} \varepsilon_k$, тогда $B_\varepsilon(x) \subset B_{\varepsilon_k}(x) \subset G_k$ для $k = 1, \ldots, n$ и, следовательно, $B_\varepsilon(x) \subset \bigcap\limits_{k=1}^{n} G_k$, т.е. x — внутренняя точка $\bigcap\limits_{k=1}^{n} G_k$.

Теперь покажем, что любое пересечение замкнутых множеств является замкнутым множеством. Пусть F_λ, $\lambda \in \Lambda$, — замкнутые множества. Тогда $M \setminus F_\lambda$ открытые множества и по законам Моргана и уже доказанной части теоремы $M \setminus \bigcap\limits_{\lambda \in \Lambda} F_\lambda = \bigcup\limits_{\lambda \in \Lambda} (M \setminus F_\lambda)$ — открытое множество, а значит, $\bigcap\limits_{\lambda \in \Lambda} F_\lambda$ — замкнутое множество.

Покажем, что конечное объединение замкнутых множеств является замкнутым множетсвом. Пусть $F_1, \ldots, F_n$ — замкнутые множества. Тогда $M \setminus F_k$ открытые множества и по законам Моргана и уже доказанной части теоремы $M \setminus \bigcup\limits_{k=1}^{n} F_k = \bigcap\limits_{k=1}^{n} (M \setminus F_k)$ — открытое множество, а значит, $\bigcup\limits_{k=1}^{n} F_k$ — замкнутое множество. ▲

В предыдущей лекции нами было введено понятие ε-окрестности. В дальнейшем будем под окрестностью точки всегда понимать некоторую ε-окрестность точки (хотя часто окрестностью точки называют любое содержащее эту точку открытое множество).

Определение 1. Точка x называется **предельной точкой** множества E метрического пространства, если в любой ее проколотой ε-окрестности $B'_\varepsilon(x)$ найдется точка множества E (т.е. $B'_\varepsilon(x) \cap E \neq \varnothing$).

Теорема 2. *Точка x является предельной точкой множества E метрического пространства тогда и только тогда, когда в любой ее ε-окрестности $B_\varepsilon(x)$ содержится бесконечно много точек множества E (т.е. $B_\varepsilon(x) \cap E$ — бесконечное множество).*

▼ Если в любой ε-окрестности x $B_\varepsilon(x)$ бесконечно много точек E, то они есть и в проколотой ε-окрестности $B'_\varepsilon(x)$ и, значит, x — предельная точка E.

Если найдется ε-окрестность точки x $B_\varepsilon(x)$ в которой содержится только конечное множество точек E, то в случае, если отличных от x точек множества E в $B_\varepsilon(x)$ нет, то $B'_\varepsilon(x) \cap E = \varnothing$ и, значит, x не является предельной точкой E, а если отличные от x точки множества E в $B_\varepsilon(x)$ есть, то возьмем в качестве $\delta > 0$ расстояние до точки x от ближайшей из них к x. Тогда в проколотой δ-окрестности точки x $B'_\delta(x)$ не будет точек E и, значит, x не является предельной точкой E и в этом случае. ▲

Определение 2. Точка x называется точкой **прикосновения** множества E метрического пространства, если в любой ее ε-окрестности $B_\varepsilon(x)$ найдется точка множества E (т.е. $B_\varepsilon(x) \cap E \neq \varnothing$).

Определение 3. Точка x называется **изолированной** точкой множества E метрического пространства, если существует такая ε-окрестность x $B_\varepsilon(x)$, пересечение которой с множеством E содержит одну точку x (т.е. $B_\varepsilon(x) \cap E = \{x\}$).

Всякая точка прикосновения внутренняя или граничная, а также предельная или изолированная.

Теорема 3. *Следующие утверждения эквивалентны:*

1) множество E замкнуто;

2) множество E содержит все свои граничные точки;

3) множество E содержит все свои точки прикосновения;

4) множество E содержит все свои предельные точки.

▼ Проведем его по схеме: 1)⇒2)⇒3)⇒4)⇒1).

Сначала покажем, что из 1) следует 2).

Если E замкнуто, то его дополнение $M \setminus E$ открыто и, значит, все точки $M \setminus E$ внутренние точки $M \setminus E$ и внешние точки E. Если $M \setminus E$ — внешность E, то E содержит все свои граничные точки.

Теперь докажем, что из 2) следует 3).

Действительно, любая точка прикосновения E — внутренняя или граничная точка E. Все внутренние точки E принадлежат E. Если E содержит все свои граничные точки, то E содержит все свои точки прикосновения.

Теперь покажем, что из 3) следует 4).

Так как любая предельная точка E является точкой прикосновения E, то если E содержит все свои точки прикосновения, то E содержит все свои предельные точки.

И, наконец, докажем, что из 4) следует 1).

Пусть $x \in M \setminus E$, значит, x не является предельной точкой E. Тогда существует такая $B'_\varepsilon(x)$, что $B'_\varepsilon(x) \cap E = \varnothing$, а значит, и $B_\varepsilon(x) \cap E = \varnothing$, т.е. $B_\varepsilon(x) \subset M \setminus E$. Итак, $M \setminus E$ — открытое множество, а E — замкнутое множество. ▲

Определение и простейшие свойства компактов

Напомним, что система множеств $\{U_\lambda\}$ образует покрытие множества E, если $\bigcup\limits_\lambda U_\lambda \supset E$.

В случае, когда все множества системы открытые, покрытие называют открытым.

Определение 4. Множество K метрического пространства называется **компактом** (или **компактным множеством**), если из любой покрывающей его системы открытых множеств можно выделить также покрывающую K конечную подсистему множеств (из любого открытого покрытия K можно выделить конечное подпокрытие).

Определение 5. Множество E метрического пространства называется **ограниченным**, если оно заключается (включается) в некоторый шар.

Теорема 4. *Любой компакт — ограниченное и замкнутое множество.*

▼ Первое следует из того, что взяв любую точку x метрического пространства M из покрытия компакта K (а фактически, всего M) открытыми множествами $\{B_n(x)\}_{n\in\mathbb{N}}$ можно выделить конечное подпокрытие K, а т.к. $B_1(x) \subset B_2(x) \subset B_3(x) \subset \ldots$, то найдется $B_n(x) \supset K$.

Второе следует из того, что если $x \notin K$, то из покрытия компакта K (а фактически, $M \setminus \{x\}$) открытыми множествами $\left\{M \setminus \overline{B}_{\frac{1}{n}}(x)\right\}_{n\in\mathbb{N}}$ можно выделить конечное подпокрытие K, а т.к. $M \setminus \overline{B}_1(x) \subset M \setminus \overline{B}_{\frac{1}{2}}(x) \subset M \setminus \overline{B}_{\frac{1}{3}}(x) \subset \ldots$, то найдется $M \setminus \overline{B}_{\frac{1}{n}}(x) \supset K$, значит, $B_{\frac{1}{n}}(x) \subset \overline{B}_{\frac{1}{n}}(x) \subset M \setminus K$, $M \setminus K$ — открытое, а K — замкнутое множество. ▲

Теорема 5. *Любое замкнутое подмножество компакта — компакт.*

▼ Пусть K компакт в метрическом пространстве M и F его замкнутое подмножество. Покажем, что F также компакт. Пусть система открытых множеств $\{G_\alpha\}$ покрывает F. Система $\{G_\alpha\}$ вместе с открытым множеством $M \setminus F$ покрывает M, а значит, и K. Выделим из этого покрытия конечное подпокрытие K. Если $M \setminus F$ входит в это подпокрытие, то удалив из него $M \setminus F$ получим искомое конечное подпокрытие F множествами системы $\{G_\alpha\}$. Если $M \setminus F$ не входит в выделенное конечное подпокрытие, то это подпокрытие является искомым конечным подпокрытием F множествами системы $\{G_\alpha\}$. Значит, F — компакт. ▲

Критерий компактности в $\mathbb{R}^n$

Теорема 6. *Любой n-мерный брус (параллелепипед)*

$$\prod_{k=1}^{n}[a_k, b_k] = \{\vec{x} = (x_1, \dots, x_n): \ a_k \leqslant x_k \leqslant b_k, \ k = 1, \dots, n\}$$

— компакт в $\mathbb{R}^n$.

▼ Предположим обратное, брус $\Pi^1 = \prod\limits_{k=1}^{n}[a_k^1, b_k^1] = \prod\limits_{k=1}^{n}[a_k, b_k]$ покрыт системой открытых множеств $\{G_\alpha\}$ и из нее нельзя выбрать конечной подсистемы, также покрывающей Π^1. Тогда разделив этот брус пополам по каждому ребру $[a_k^1, b_k^1]$, $k = 1, \dots, n$, получим 2^n брусов $\prod\limits_{k=1}^{n}[\alpha_k, \beta_k]$, где $[\alpha_k, \beta_k]$ — правая или левая половина отрезка $[a_k^1, b_k^1]$, т.е. или отрезок $\left[a_k^1, \frac{a_k^1+b_k^1}{2}\right]$ или отрезок $\left[\frac{a_k^1+b_k^1}{2}, b_k^1\right]$. Хотя бы для одного из этих 2^n брусов также нельзя выбрать из системы $\{G_\alpha\}$ конечной покрывающей его подсистемы. Обозначим такой брус $\Pi^2 = \prod\limits_{k=1}^{n}[a_k^2, b_k^2]$. Разделим его пополам по каждому ребру $[a_k^2, b_k^2]$, $k = 1, \dots, n$, и получим 2^n брусов, хотя бы для одного из которых также нельзя выбрать из системы $\{G_\alpha\}$ конечной покрывающей его подсистемы. Обозначим такой брус $\Pi^3 = \prod\limits_{k=1}^{n}[a_k^3, b_k^3]$. Разделим его пополам по каждому ребру .... Продолжая далее рассуждения, получим последовательность вложенных брусов

$$\Pi^1 \supset \Pi^2 \supset \Pi^3 \supset \dots \supset \Pi^m \supset \dots, \quad \Pi^m = \prod_{k=1}^{n}[a_k^m, b_k^m],$$

каждый из которых нельзя покрыть конечной подсистемой системы $\{G_\alpha\}$. Длина k-го ребра $[a_k^m, b_k^m]$ бруса Π^m равна $2^{1-m}(b_k^1 - a_k^1)$ и стремится к 0 при возрастании m. Так как

$$[a_k^1, b_k^1] \supset [a_k^2, b_k^2] \supset [a_k^3, b_k^3] \supset \dots \supset [a_k^m, b_k^m] \supset \dots,$$

то по принципу вложенных отрезков Кантора существует единственная точка c_k принадлежащая всем $[a_k^m, b_k^m]$, $m \in \mathbb{N}$. Так как $\vec{c} = (c_1, \dots, c_n) \in \Pi^1$, то $\vec{c} = (c_1, \dots, c_n) \in \bigcup\limits_{\alpha} G_\alpha$ и, значит, найдется такое α', что $\vec{c} = (c_1, \dots, c_n) \in G_{\alpha'}$. Множество $G_{\alpha'}$ — открытое, следовательно, найдется $B_\varepsilon(\vec{c}) \subset G_{\alpha'}$. Найдем такое m, что

$\sqrt{n}2^{1-m} \cdot \max\limits_{1\leqslant k\leqslant n}(b_k^1 - a_k^1) < \varepsilon$. Тогда для любой точки $\vec{x} \in \Pi^m$

$$\rho(\vec{x},\vec{c}) = \|\vec{x}-\vec{c}\| = \sqrt{\sum_{k=1}^{n}(x_k - c_k)^2} \leqslant \sqrt{\sum_{k=1}^{n}(b_k^m - a_k^m)^2} \leqslant$$
$$\leqslant \sqrt{n}\max_{1\leqslant k\leqslant n}(b_k^m - a_k^m) = \sqrt{n}2^{1-m}\cdot\max_{1\leqslant k\leqslant n}(b_k^1 - a_k^1) < \varepsilon$$

и, значит, $\Pi^m \subset B_\varepsilon(\vec{c}) \subset G_{\alpha'}$, что противоречит невозможности покрытия Π^m конечной подсистемой системы $\{G_\alpha\}$. Полученное противоречие доказывает теорему. ▲

Теорема 7. *Множество в $\mathbb{R}^n$ компактно тогда и только тогда, когда оно ограничено и замкнуто.*

▼ Если множество компактно, то по теореме 2 оно ограничено и замкнуто. Остается только доказать, что если множество $K \subset \mathbb{R}^n$ ограничено и замкнуто, то оно компактно. Так как K ограничено, то существует шар $\overline{B}_r(\vec{x}) \supset K$, брус $\Pi = \prod\limits_{k=1}^{n}[x_k - r, x_k + r] \supset \overline{B}_r(\vec{x}) \supset K$ ($\vec{x} = (x_1,\ldots,x_n)$). По предыдущей теореме Π компакт, K его замкнутое подмножество и, значит, по теореме 3 также компакт. ▲

Существование предельной точки у компакта

Теорема 6. *Любое бесконечное подмножество компакта имеет хотя бы одну предельную точку принадлежащую этому компакту.*

▼ Предположим обратное. Пусть K — компакт, A — его бесконечное подмножество, не имеющее предельных точек, принадлежащих K. Тогда для любого $x \in K$ найдется такая $B'_{\varepsilon_x}(x)$, что $A \cap B'_{\varepsilon_x}(x) = \varnothing$, а значит, $A \cap B_{\varepsilon_x}(x)$ пусто или равно $\{x\}$. Система открытых множеств $\{B_{\varepsilon_x}(x)\}_{x\in K}$ покрывает компакт K, но из нее нельзя выделить конечного подпокрытия K, ведь иначе получили бы, что A конечно. Полученное противоречие доказывает теорему. ▲

Следствие. *В $\mathbb{R}^n$ любое ограниченное бесконечное множество имеет хотя бы одну предельную точку.*

▼ В самом деле, ограниченное множество принадлежит какому-то шару $\overline{B}_r(x)$, компакту по критерию компактности, и, значит, если оно бесконечно, то обязательно имеет хотя бы одну предельную точку. ▲

Лекция 19 (17.04.20)
Предел последовательности.
Полные метрические пространства

Предел последовательности. Определения

Определение 1. Последовательностью в метрическом пространстве M (точнее, (M,ρ)) называется отображение множества натуральных чисел $\mathbb{N}$ в M: $n \to a_n \in M$.

Определение 2. Подпоследовательностью последовательности $\{a_n\}_{n\in\mathbb{N}}$ будем называть отображение $k \to a_{n_k}$, где $\{n_k\}$ — строго возрастающая последовательность натуральных чисел (т.е. подпоследовательность — это последовательность $\{a_{n_k}\}_{k\in\mathbb{N}}$).

Определение 3. Элемент a метрического пространства M называется **пределом последовательности** $\{a_n\}_{n\in\mathbb{N}}$ элементов M (или $\{a_n\}_{n=1}^{\infty}$ или $\{a_n\}$ или просто a_n), если

$$\rho(a_n, a) \longrightarrow 0 \text{ при } n \to \infty.$$

Пишут $\lim\limits_{n\to\infty} a_n = a$ или $\lim a_n = a$ или $a_n \xrightarrow[n\to\infty]{} a$ или $a_n \to a$.

Данное определение можно было записать иначе:

$$\forall \varepsilon > 0\ \exists N_\varepsilon \in \mathbb{R}\ \forall n > N_\varepsilon : \quad \rho(a_n, a) < \varepsilon$$

или

$$\forall B_\varepsilon(a)\ \exists N_\varepsilon \in \mathbb{R}\ \forall n > N_\varepsilon : \quad a_n \in B_\varepsilon(a).$$

Определение 4. Последовательность называют **сходящейся**, если она имеет предел (т.е. найдется a, удовлетворяющее предыдущему определению).

Определение 5. Последовательность называют **расходящейся**, если она не является сходящейся.

Свойства предела

Свойство 1 (о подпоследовательности). *Если последовательность имеет предел, то любая ее подпоследовательность имеет тот же предел.*

▼ Если числовая последовательность $\rho(a_n, a) \to 0$ при $n \to \infty$, то по свойству сходимости подпоследовательности (для числовых последовательностей) и ее подпоследовательность $\rho(a_{n_k}, a) \to 0$ при $n \to \infty$. ▲

Замечание. И эту и последующие теоремы можно доказывать так же, как они были доказана в в предыдущем семестре, но мы будем сокращать доказательства, используя уже известные свойства числовых последовательностей.

Свойство 2 (о единственности предела). *Если последовательность имеет предел, то он единственен.*

▼ Если $\rho(a_n, a) \xrightarrow[n\to\infty]{} 0$ и $\rho(a_n, b) \xrightarrow[n\to\infty]{} 0$, то переходя к пределу (при $n \to \infty$) в неравенстве $\rho(a, b) \leqslant \rho(a_n, a) + \rho(a_n, b)$ получаем, что $\rho(a, b) \leqslant 0$, т.е. $\rho(a, b) = 0$ и, значит, $a = b$. ▲

Определение 6. Последовательность $\{a_n\}$ называют **ограниченной**, если множество значений последовательности $\{x \in M : \ x = a_n,\ n \in \mathbb{N}\}$ ограничено (т.е. лежит в некотором шаре).

Свойство 3 (об ограниченности). *Если последовательность сходится, то она ограничена.*

▼ Если $a = \lim\limits_{n\to\infty} a_n$, то $\rho(a_n, a) \to 0$ и, значит, числовая последовательность $\{\rho(a_n, a)\}$ ограничена, т.е. существует $r \geqslant \rho(a_n, a)$ при $n \in \mathbb{N}$, тогда $a_n \in \overline{B}_r(a)$ при всех $n \in \mathbb{N}$. ▲

Свойство 4 (об отделимости). *Если последовательность $\{a_n\}$ сходится и $b \neq \lim a_n$, то $\exists \varepsilon > 0\ \exists N \in \mathbb{R}\ \forall n > N :\ a_n \notin B_\varepsilon(b)$.*

▼ Пусть $a = \lim a_n$, возьмем $\varepsilon = \frac{1}{2}\rho(a, b)$, а N таким, что при $n > N\ \ a_n \in B_\varepsilon(a)$. Так как $B_\varepsilon(a) \cap B_\varepsilon(b) = \varnothing$, то окрестность $B_\varepsilon(b)$ и найденное N — искомые, т.е. при $n > N\ \ a_n \notin B_\varepsilon(b)$. ▲

В метрическом пространстве вообще говоря не определены операции сложения, вычитания, умножения и деления элементов, нет отношения порядка, поэтому основанные на этих операциях теоремы о числовых последовательностях не имеют аналогов в общих метрических пространствах. В нормированных пространствах некоторые аналоги таких теорем есть.

Бесконечно малые последовательности

Определение 7. Последовательность $\{a_n\}$ в нормированном пространстве называют **бесконечно малой**, если $\lim\limits_{n\to\infty} a_n = 0$ — нулю линейного пространства или, что по определению предела то же самое, $\lim\limits_{n\to\infty} \|a_n\| = 0$.

Обозначение 1. Если последовательность $\{a_n\}$ бесконечно малая, то пишут $a_n = o(1)$.

Обозначение 2. Если пишут $a_n = O(1)$, то это означает, что последовательность $\{a_n\}$ ограничена.

Теорема 1. *В нормированном пространстве $a_n \xrightarrow[n\to\infty]{} a$ тогда и только тогда, когда $a_n = a + o(1)$ (т.е., когда $a_n - a = o(1)$).*

▼ Утверждение, что $a_n \xrightarrow[n\to\infty]{} a$ означает, что $\rho(a_n, a) = \|a_n - a\| \xrightarrow[n\to\infty]{} 0$, но это же означает, что $a_n - a = o(1)$. ▲

Теорема 2. *Если в нормированном пространстве $a_n = o(1)$ и $b_n = o(1)$, то $a_n \pm b_n = o(1)$. Если $a_n = o(1)$, а α_n — ограниченная числовая последовательность, то $\alpha_n a_n = o(1)$. Если $a_n = O(1)$, а α_n — бесконечно малая числовая последовательность, то $\alpha_n a_n = o(1)$.*

▼ Так как $a_n = o(1)$ и $b_n = o(1)$, то $\|a_n\| = o(1)$ и $\|b_n\| = o(1)$, а так как $\|a_n \pm b_n\| \leqslant \|a_n\| + \|b_n\| = o(1) + o(1) = o(1)$, то $\|a_n \pm b_n\| \xrightarrow[n\to\infty]{} 0$ и, значит, $a_n \pm b_n = o(1)$. В двух оставшихся утверждениях из числовых последовательностей $|\alpha_n, \|a_n\|$ одна ограничена, а другая бесконечно малая и, значит, по теореме о сумме бесконечно малых числовых последовательностей $\|\alpha_n a_n\| = |\alpha_n| \cdot \|a_n\| = o(1) \cdot O(1) = o(1)$, т.е. $\alpha_n a_n = o(1)$. ▲

Теорема 3. *Если в нормированном пространстве $a_n \xrightarrow[n\to\infty]{} a$ и $b_n \xrightarrow[n\to\infty]{} b$, то $a_n \pm b_n \xrightarrow[n\to\infty]{} a \pm b$. А если $a_n \xrightarrow[n\to\infty]{} a$, числовая последовательность $\alpha_n \xrightarrow[n\to\infty]{} \alpha$, то $\alpha_n a_n \xrightarrow[n\to\infty]{} \alpha a$.*

▼ Так как $a_n - a = o(1)$ и $b_n - b = o(1)$, то $(a_n \pm b_n) - (a \pm b) = o(1) \pm o(1) = o(1)$ по предыдущей теореме и, значит, $a_n \pm b_n \xrightarrow[n\to\infty]{} a \pm b$.

Далее, так как $a_n - a = o(1)$ и $\alpha_n - \alpha = o(1)$, то $\alpha_n a_n - \alpha a = \alpha_n(a_n - a) + (\alpha_n - \alpha)a = O(1) \cdot o(1) + o(1) \cdot O(1) = o(1)$ по предыдущей теореме и, значит, $\alpha_n a_n \xrightarrow[n\to\infty]{} \alpha a$. ▲

Замечание. Мы не касаемся деления $\frac{a_n}{\alpha_n}$, т.к. оно сводится к умножению $\frac{1}{\alpha_n} \cdot a_n$. Теперь перейдем к пространствам $\mathbb{R}^n$.

Теорема 4. *В $\mathbb{R}^n$ $\vec{x}^m \xrightarrow[m\to\infty]{} \vec{x}$ тогда и только тогда, когда для любого натурального k, $1 \leqslant k \leqslant n$, $x_k^m \xrightarrow[m\to\infty]{} x_k$ (т.е. в $\mathbb{R}^n$ последовательность сходится тогда и только тогда, когда она сходится покоординатно).*

▼ Необходимость. Если $\vec{x}^m \xrightarrow[m\to\infty]{} \vec{x}$, то $\|\vec{x}^m - \vec{x}\| \xrightarrow[m\to\infty]{} 0$. А так как для любого k, $1 \leqslant k \leqslant n$, $|x_k^m - x_k| \leqslant \sqrt{\sum\limits_{k=1}^{n} |x_k^m - x_k|^2} = \|\vec{x}^m - \vec{x}\|$, то $|x_k^m - x_k| \xrightarrow[m\to\infty]{} 0$ и, значит,

$x_k^m \xrightarrow[m\to\infty]{} x_k$.

Достаточность. Если для любого k, $1 \leqslant k \leqslant n$, $x_k^m \xrightarrow[m\to\infty]{} x_k$, то $(x_k^m - x_k)^2 = o(1)$ для всех k, $1 \leqslant k \leqslant n$, и значит, $\|\vec{x}^m - \vec{x}\|^2 = \sum\limits_{k=1}^{n}(x_k^m - x_k)^2 = \sum\limits_{k=1}^{n} o(1) = o(1)$, т.е. $\|\vec{x}^m - \vec{x}\| \xrightarrow[m\to\infty]{} 0$, а $\vec{x}^m \xrightarrow[m\to\infty]{} \vec{x}$. ▲

Полные метрические пространства

Вернемся к метрическим пространствам.

Определение 8. Последовательность $\{a_n\}$ элементов метрического пространства называется **фундаментальной** или **последовательностью Коши**, если

$$\forall \varepsilon > 0\ \exists N\ \forall i > N\ \forall j > N : \quad \rho(a_i, a_j) < \varepsilon.$$

Утверждение 1. *В метрическом пространстве любая сходящаяся последовательность является последовательностью Коши.*

▼ Если $a_n \xrightarrow[n\to\infty]{} a$, то $\forall \varepsilon > 0\ \exists N\ \forall n > N :\ \rho(a_n, a) < \frac{\varepsilon}{2}$. Но тогда $\forall i > N\ \forall j > N :$ $\rho(a_i, a_j) \leqslant \rho(a_i, a) + \rho(a, a_j) < \varepsilon$. ▲

Как мы уже знаем, в $\mathbb{R}$ верно и обратное утверждение, всякая последовательность Коши сходится. В произвольном метрическом пространстве последовательность Коши не обязана быть сходящейся, но особый интерес вызывают пространства, в которых любая последовательность Коши сходится.

Определение 9. Метрическое пространство называется **полным**, если в нем любая последовательность Коши сходится.

Определение 10. Полное нормированное пространство (полное, как метрическое с метрикой $\rho(x, y) = \|x - y\|$) называется **банаховым**.

Теорема 5. *Пространство $\mathbb{R}^n$ является банаховым.*

▼ Если $\{\vec{x}_m\}_{m=1}^{\infty}$ — последовательность Коши в $\mathbb{R}^n$, то в силу неравенства $|x_k^i - x_k^j| \leqslant \|\vec{x}^i - \vec{x}^j\|$ она является последовательностью Коши покоординатно, по критерию Коши в $\mathbb{R}$ она сходится покоординатно, а тогда по теореме о покоординатной сходимости она сходится в $\mathbb{R}^n$. ▲

Замечание. Укажем (без доказательства), что любое метрическое пространство можно считать подмножеством полного метрического пространства с той же метрикой (т.е. для любого данного метрического пространства найдется полное метрическое пространство, подмножеством которого является данное пространство и метрика на нем совпадает с метрикой этого полного пространства).

Теорема 6 (Больцано-Вейерштрасса). *Из любой последовательности точек компакта (в метрическом пространстве) можно выбрать сходящуюся к точке компакта подпоследовательность.*

▼ Если множество значений последовательности конечно, то хотя бы одно значение встречается в последовательности бесконечно много раз и тогда члены последовательности, равные ему, образуют постоянную, а значит, сходящуюся к точке компакта подпоследовательность.

Если множество значений последовательности $\{a_n\}$ бесконечно, то теореме о существовании предельной точки на компакте оно имеет предельную точку принадлежащую компакту. Обозначим ее через a. Найдём такое n_1, что $a_{n_1} \in B_1(a)$. Если уже выбраны $n_1, n_2, \ldots, n_m$ так, что $n_1 < n_2 < \cdots < n_m$ и $a_{n_k} \in B_{\frac{1}{k}}(a)$, $k = 1, \ldots, m$, то

n_{m+1} выбираем таким, что $n_{m+1} > n_m$ и $a_{m+1} \in B_{\frac{1}{m+1}}(a)$. Так будет построена подпоследовательность $\{a_{n_k}\}_{k=1}^{\infty}$, которая сходится к a. Действительно, $\forall \varepsilon > 0\ \exists N > \frac{1}{\varepsilon}$, но тогда $\forall k > N:\ a_{n_k} \in B_{\frac{1}{k}}(a) \subset B_{\frac{1}{N}}(a) \subset B_{\varepsilon}(a)$ и, значит, $a_{n_k} \to a \in K$ при $k \to \infty$. ▲

Следствие 1. *В $\mathbb{R}^n$ из любой ограниченной последовательности можно выбрать сходящуюся подпоследовательность.*

▼ Действительно, члены ограниченной последовательности принадлежат некоторому замкнутому шару, который по критерию компактности является компактом. ▲

В метрических пространствах существует аналог принципа вложенных отрезков Кантора. Приведем его формулировку.

Теорема 7. *Метрическое пространстве полно тогда и только тогда, когда в нем любая последовательность вложенных замкнутых шаров $\overline{B}_{r_1}(x_1) \supset \overline{B}_{r_2}(x_2) \supset \overline{B}_{r_3}(x_3) \supset \ldots$ с радиусами r_n, стремящимися к нулю, имеет общую точку.*

В приведенной теореме условие, что радиусы r_n стремятся к нулю, существенно.

Теперь перейдем к нормированным пространствам и отметим, что в полном нормированном пространстве (в отличие от метрического) любая последовательность вложенных замкнутых шаров имеет общую точку.

Лекция 20 (21.04.20)
Предел функции в метрическом пространстве

Определения

Пусть есть два метрических пространства (M, ρ_M) и (N, ρ_N) и пусть функции — отображения из M в N (т.е. отображения M или подмножества M в N).

Определение 1 (предел функции по Коши). Элемент b метрического пространства N называется **пределом функции** $f(x)$ при x стремящемся к $a \in M$ по множеству $A \subset M$, если a — предельная точка A, f определена в некоторой проколотой Δ-окрестности точки a на множестве A (т.е. на $B'_{\Delta}(a) \cap A$ для некоторго $\Delta > 0$) и

$$\forall \varepsilon > 0\ \exists \delta > 0\ \forall x \in B'_{\delta}(a) \cap A: \quad f(x) \in B_{\varepsilon}(b)$$

или, что эквивалентно,

$$\forall B_{\varepsilon}(b)\ \exists B'_{\delta}(a)\ \forall x \in B'_{\delta}(a) \cap A: \quad f(x) \in B_{\varepsilon}(b).$$

или, что эквивалентно,

$$\forall \varepsilon > 0\ \exists \delta > 0\ \forall x \in A,\ 0 < \rho_M(x, a) < \delta: \quad \rho_N(f(x), b) < \varepsilon.$$

Обозначение 1. Пишут: $\lim\limits_{A \ni x \to a} f(x) = b$ или $f(x) \xrightarrow[A \ni x \to a]{} b$ и говорят: **предел функции** f **в точке** a **по множеству** A **равен** b или $f(x)$ **стремится к** b **при стремлении** x **к** a **по множеству** A.

Определение 2 (предел функции по Гейне). Элемент b метрического пространства N называется **пределом функции** $f(x)$ при x стремящемся к $a \in M$ по множеству $A \subset M$, если a — предельная точка A, f определена в некоторой проколотой Δ-окрестности точки a на множестве A (т.е. на $B'_{\Delta}(a) \cap A$ для некоторго $\Delta > 0$) и для любой последовательности аргументов x_n из $A \setminus \{a\}$, стремящейся к a, последовательность значений $f(x_m)$ стремится к b, то есть

$$\forall \{x_n\} \in \Big\{ \{t_n\} : \ t_n \in A \setminus \{a\} \text{ и } \lim_{n\to\infty} t_n = a \Big\} : \ \lim_{n\to\infty} f(x_n) = b$$

или, эквивалентно,

$$\forall x_n \in A \setminus \{a\} : \quad (x_n \xrightarrow[n\to\infty]{} a \Rightarrow f(x_n) \xrightarrow[n\to\infty]{} b).$$

Иначе говоря, предел функции f (из M в N) в точке $a \in M$ по множеству A это такой элемент $b \in N$, который является пределом любой последовательности значений функции, если последовательность аргументов стремится к a по множеству $A \setminus \{a\}$.

Как и в случае предыдущего определения Коши, пишут: $\lim\limits_{A \ni x \to a} f(x) = b$ или $f(x) \xrightarrow[A \ni x \to a]{} b$.

Замечание. Отметим, что если рассматривать $A \cup \{a\}$ как метрическое пространство (с той же метрикой, что и в M), то вышеприведенные определения пределов по A и те же определения пределов по всему метрическому пространству $A \cup \{a\}$ очевидно равносильны. Поэтому вместо предела по подмножеству можно расссматривать предел по метрическому пространству, что упрощает формулировки и доказательства.

Теорема 1. *Определения предела функции по Коши и по Гейне эквивалентны.*

▼ В соответствии с замечанием выше считаем, что в определениях пределов $A = M$. Покажем, что оба определения одновременно выполняются или одновременно не выполняются. Сначала докажем, что если выполняется определение по Коши, то выполняется и определение по Гейне.

Итак, пусть выполняется определение по Коши, a — предельная точка M, функция f определена в некоторой проколотой Δ-окрестности $B'_\Delta(a)$ точки a и

$$\forall B_\varepsilon(b) \quad \exists B'_\delta(a)\ \forall x \in B'_\delta(a) : \quad f(x) \in B_\varepsilon(b). \tag{*}$$

Проверим, что если $x_n \neq a$, $n \in \mathbb{N}$, и $x_n \to a$, то $f(x_n) \to b$. Так как $x_n \to a$, то $\forall B_\delta(a)\ \exists N \in \mathbb{R}\ \forall n > N : \quad x_n \in B_\delta(a)$. Так как $x_n \neq a$, то, значит, $\forall n > N\ x_n \in B'_\delta(a) \cap A$ и, следовательно, по (*) $\forall n > N\ f(x_n) \in B_\varepsilon(b)$. Получаем:

$$\forall \varepsilon > 0\ \exists N \in \mathbb{R}\ \forall n > N : \quad f(x_n) \in B_\varepsilon(b),$$

то есть $\lim\limits_{n\to\infty} f(x_n) = b$. Значит, выполняется определение по Гейне.

Теперь докажем, что если не выполняется определение по Коши, то не выполняется и определение по Гейне. В начале обоих определений формулируются некоторые одинаковые требования, естественно, что их невыполнение в одном определении означает их невыполнение в другом. Поэтому пусть a — предельная точка M, функция f определена в некоторой проколотой Δ-окрестности $B'_\Delta(a)$ точки a, но определение по Коши не выполняется, то есть

$$\exists B_\varepsilon(b)\ \forall B'_\delta(a)\ \exists x \in B'_\delta(a) : \quad f(x) \notin B_\varepsilon(b).$$

Возьмем последовательность $\delta_n = \frac{1}{n}$ и найдем соответствующую последовательность точек $x_n \in B'_{\delta_n}(a)$, $f(x_n) \notin B_\varepsilon(b)$. Точки $x_n \neq a$, $x_n \to a$ (т.к. $\delta_n \to 0$), но $f(x_n) \notin B_\varepsilon(b)$, а значит, b не является пределом последовательности $f(x_n)$. Определение по Гейне не выполняется. Теорема доказана. ▲

Так как определения предела функции по Коши и по Гейне эквиваленты, то в дальнейшем мы будем говорить просто, что b является пределом f в точке a или что f имеет предел в точке a равный b.

Определение 3. Функция f **имеет предел** в точке a, если существует такое $b \in N$, что $\lim\limits_{x\to a} f(x) = b$.

Свойства предела

Свойство 1 (о пределе по подмножеству). *Если* $\lim\limits_{A\ni x\to a} f(x) = b$*, а* $A' \subset A$ *и* a *— предельная точка* A'*, то* $\lim\limits_{A'\ni x\to a} f(x) = b$*.*

▼ Если $x_n \in A' \setminus \{a\}$ и $x_n \to a$, то $f(x_n) \to b$, т.е. по Гейне $\lim\limits_{A'\ni x\to a} f(x) = b$. ▲

Свойство 2 (о единственности предела). *Если* $\lim\limits_{A\ni x\to a} f(x)$ *существует, то он единственен.*

▼ Возьмем любую последовательность точек $x_n \neq a$, $x_n \to a$. Тогда $f(x_n) \to \lim\limits_{x\to a} f(x)$ и, значит, в силу единственности предела для последовательностей, $\lim\limits_{x\to a} f(x)$ единственен. ▲

Свойство 3 (об ограниченности). *Если существует предел* $\lim\limits_{x\to a} f(x)$*, то существует такая проколотая* δ*-окрестность точки* a*, на которой функция* f *ограничена (т.е.* $f\left(B'_\delta(a)\right)$ *— ограниченное множество).*

▼ Пусть $b = \lim\limits_{x\to a} f(x)$, тогда $\exists B'_\delta(a)\ \forall x \in B'_\delta(a):\ f(x) \in B_1(b)$, т.е. $f\left(B'_\delta(a)\right) \subset B_1(b)$, значит, $f\left(B'_\delta(a)\right)$ — ограниченное множество. ▲

Свойство 4 (об отделимости). *Если* $\lim\limits_{x\to a} f(x) = b \neq c$*, то* $\exists B_\varepsilon(c)\ \exists B_\delta(a) :$ $f\left(B'_\delta(a)\right) \cap B_\varepsilon(c) = \varnothing$ *(т.е.* $\forall x \in B_\delta(a):\ f(x) \notin B_\varepsilon(c)$*).*

▼ Возьмем $\varepsilon = \frac{1}{2}\rho(b,c)$. Ясно, что $B_\varepsilon(b) \cap B_\varepsilon(c) = \varnothing$. Затем найдем такую $B'_\delta(a)$, что $\forall x \in B'_\delta(a):\ f(x) \in B_\varepsilon(b)$, т.е. $f\left(B'_\delta(a)\right) \subset B_\varepsilon(b)$. Но тогда $f\left(B'_\delta(a)\right) \cap B_\varepsilon(c) = \varnothing$. ▲

Так как в метрическом пространстве вообще говоря нет операций сложения, вычитания, умножения, деления и нет отношения порядка, то ряд теорем о пределах функций не имеют аналогов в общих метрических пространствах. Приведем некоторые аналоги таких теорем для нормированных пространств.

Определение 4. Функцию f из метрического пространства M в нормированное пространство N называют **бесконечно малой** в точке $a \in M$ (при $x \to a$), если $\lim\limits_{x\to a} f(x) = 0$ — нулю линейного пространства N или, что по определению предела то же самое, $\lim\limits_{x\to a} \|f(x)\| = 0$.

Обозначение 2. Про бесконечно малую функцию пишут $f(x) = o(1)$ при $x \to a$.

Обозначение 3. Если пишут $f(x) = O(1)$ при $x \to a$, то это означает, что f **ограничена** в некоторой проколотой окрестности точки a.

Обозначение 4. Пусть f и g функции из метрического пространства M в нормированное пространство N. Функция $f(x) = o(g(x))$ при $x \to a$, если функции f и g определены в некоторой проколотой окрестности точки a и существует такая бесконечно малая функция $h(x)$ (из M в $\mathbb{R}$ (или в $\mathbb{C}$)) при $x \to a$, что $f(x) = h(x)g(x)$ на некоторой $B'_\delta(a)$, т.е. $f(x) = o(1)g(x)$ на $B'_\delta(a)$ (при $x \to a$).

Обозначение 5. Пусть f и g функции из метрического пространства M в нормированное пространство N, $f(x) = O(g(x))$ при $x \to a$, если f и g определены в некоторой проколотой окрестности точки a и существует такая функция $h(x)$ (из M в $\mathbb{R}$ (или в $\mathbb{C}$)) ограниченная при $x \to a$, что $f(x) = h(x)g(x)$ на некоторой $B'_\delta(a)$, т.е. $f(x) = O(1)g(x)$ на $B'_\delta(a)$ (при $x \to a$).

Теорема 2. $\lim\limits_{x\to a} f(x) = b \in \mathbb{R} \Leftrightarrow f(x) = b + o(1)$ *при* $x \to a$.

▼ Возьмем любую последовательность точек $x_n \neq a$, $x_n \to a$. По аналогичной теореме для последовательностей утверждения $\lim\limits_{n\to\infty} f(x_n) = b$ и $f(x_n) = b + o(1)$ (т.е $\lim\limits_{n\to\infty}(f(x_n) - b) = 0$) эквивалентны, значит, эквивалентны и утверждения $\lim\limits_{x\to a} f(x) = b$ и $f(x) = b + o(1)$ при $x \to a$. ▲

Теорема 3. *Если f и g функции из метрического пространства M в нормированное пространство N, бесконечно малые при $x \to a$, то $f(x) \pm g(x) = o(1)$ при $x \to a$. Если из двух функций, f из метрического пространства M в нормированное пространство N и α из метрического пространства M в $\mathbb{R}$ (или в $\mathbb{C}$, если N линейное пространство над $\mathbb{C}$), одна $o(1)$, а другая $O(1)$ при $x \to a$, то их произведение $\alpha(x)f(x) = o(1)$ при $x \to a$.*

▼ Возьмем любую последовательность точек $x_n \neq a$, $x_n \to a$. Тогда $(f \pm g)(x_n) = f(x_n) \pm g(x_n) = o(1) \pm o(1) = o(1)$ по свойствам бесконечно малых последовательностей и $\alpha(x_n)f(x_n) = o(1) \cdot O(1) = o(1)$ по свойствам бесконечно малых последовательностей. Это значит (по определению Гейне), что $f(x) \pm g(x) = o(1)$ при $x \to a$ и $\alpha(x)f(x) = o(1)$ при $x \to a$. ▲

Теорема 4. *Если f и g функции из метрического пространства M в нормированное пространство N и $\lim\limits_{x\to a} f(x) = b$, $\lim\limits_{x\to a} g(x) = c$, то $\lim\limits_{x\to a}(f(x) \pm g(x) = b \pm c$. Если f — функция из метрического пространства M в нормированное пространство N и $\lim\limits_{x\to a} f(x) = b$, а α — функция из метрического пространства M в $\mathbb{R}$ (или в $\mathbb{C}$, если N линейное пространство над $\mathbb{C}$), $\lim\limits_{x\to a} \alpha(x) = \beta$, то $\lim\limits_{x\to a} \alpha(x)f(x) = \beta b$.*

▼ Возьмем любую последовательность точек $x_n \neq a$, $x_n \to a$. Тогда $(f \pm g)(x_n) = f(x_n) \pm g(x_n) \to b \pm c$ по свойствам пределов последовательностей и $(\alpha f)(x_n) = \alpha(x_n)f(x_n) \to \beta b$ по по свойствам пределов последовательностей. ▲

Теперь коснемся пространства $\mathbb{R}^n$.

Теорема 5. *Если $\vec{f}$ функция из метрического пространства M в $\mathbb{R}^n$, то $\lim\limits_{x\to a} \vec{f}(x) = \vec{b}$ тогда и только тогда, когда для любого натурального k, $1 \leqslant k \leqslant n$, $\lim\limits_{x\to a} f_k(x) = b_k$.*

▼ Это следует из теоремы о покоординатной сходимости последовательностей и определения предела по Гейне.

Определение 5. Функция f из метрического пространства M в метрическое пространство N **удовлетворяет условию Коши** в точке a, если a — предельная точка M, f определена в некоторой проколотой окрестности точки a и

$$\forall \varepsilon > 0\ \exists B'_\delta(a)\ \forall x, x' \in B'_\delta(a): \quad \rho_N(f(x) - f(x')) < \varepsilon.$$

Теорема 6. *Пусть f — функция из метрического пространства M в метрическое пространство N. Если предел функции f в точке a существует, то f удовлетворяет условию Коши в точке a. Если f удовлетворяет условию Коши в точке a и N полное метрическое пространство, то f имеет предел в точке a.*

▼ Проверим сначала первое утверждение. Пусть $\lim\limits_{x\to a} f(x) = b$, тогда

$$\forall\varepsilon > 0\ \exists B'_\delta(a)\ \forall x \in B'_\delta(a): \quad f(x) \in B_{\frac{\varepsilon}{2}}(b).$$

Но если $x \in B'_\delta(a)$ и $x' \in B'_\delta(a)$, то

$$\rho_N(f(x), f(x')) \leqslant \rho_N(f(x), b) + \rho_N(b, f(x')) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

и, следовательно, f удовлетворяет условию Коши в точке a.

Докажем теперь достаточность условия Коши в случае, если N полное метрическое пространство. Итак, a предельная точка M и

$$\forall\varepsilon > 0\ \exists B'_\delta(a)\ \forall x, x' \in B'_\delta(a): \quad \rho_N(f(x) - f(x')) < \varepsilon.$$

Возьмем любую последовательность точек $x_n \neq a$, $x_n \to a$. Тогда для любого $\delta > 0\ \exists K\ \forall n > K: \quad x_n \in B'_\delta(a)$. Значит, $\forall n, m > K: \quad \rho_N(f(x_n), f(x_m)) < \varepsilon$. Таким образом, последовательность $f(x_n)$ является последовательностью Коши и, следовательно, в силу полноты N, сходится. Покажем теперь, что предел последовательности $f(x_n)$ один и тот же для всех последовательностей $x_n \neq a$, $x_n \to a$. В самом деле, если рассмотрим две последовательности $x_n \neq a$, $x_n \to a$, и $y_n \neq a$, $y_n \to a$. Последовательность

$$x_1, y_1, x_2, y_2, x_3, y_3, \ldots, x_n, y_n, \ldots$$

сходится к a. По доказанному последовательность

$$f(x_1), f(y_1), f(x_2), f(y_2), f(x_3), f(y_3), \ldots, f(x_n), f(y_n), \ldots$$

имеет предел и, значит, по свойству сходимости подпоследовательности,

$$\lim_{n\to\infty} f(x_n) = \lim_{n\to\infty} f(y_n).$$

По определению Гейне существует $\lim\limits_{x\to a} f(x)$. ▲

Лекция 21 (24.04.20)
Непрерывные функции

Непрерывность функции в точке

Пусть, как и раньше, (M, ρ_M) и (N, ρ_n) — метрические пространства и функции — отображения из M (части M) в N.

Определение 1 (непрерывности по Коши). Функция f **непрерывна в точке** $a \in M$, если f определена в некоторой окрестности точки a и

$$\forall B_\varepsilon(f(a))\ \exists B_\delta(a)\ \forall x \in B_\delta(a): \quad f(x) \in B_\varepsilon(f(a))$$

или, эквивалентно,

$$\forall\varepsilon > 0\ \exists\delta > 0\ \forall x,\ \rho_M(x, a) < \delta: \quad \rho_N(f(x), f(a)) < \varepsilon.$$

Замечание 1. Если a изолированная точка пространства M, то условие непрерывности функции в точке a всегда выполняется при существовании $f(a)$. Если a

— предельная точка M, то условие непрерывности эквивалентно утверждению, что $\lim\limits_{x\to a} f(x) = f(a)$.

Определение 2. Если f определена в некоторой окрестности точки a, но не является непрерывной в ней, то говорят, что f **разрывна в точке** a.

Замечание 2. В отличие от одномерного случая тут мы требуем, чтобы f была определена в точке a, так как других случаев точек разрыва мы рассматривать не будем.

Определение 3. Функция f **непрерывна в точке** $a \in M$ **по множеству** $E \subset M$, если f непрерывна в точке a на $E \cup \{a\}$ как метрическом пространстве — подмножестве M.

Определение 4. Функция f **разрывна в точке** $a \in M$ **по множеству** $E \subset M$, если f разрывна в точке a на $E \cup \{a\}$ как метрическом пространстве — подмножестве M.

Определение 5 (непрерывности по Гейне). Функция f **непрерывна в точке** $a \in M$, если f определена в некоторой окрестности точки a и для любой последовательности аргументов $x_n \to a$ последовательность значений $f(x_n) \to f(a)$.

Замечание 3. Обратим внимание, что фигурирующее в определении предела по Гейне условие $x_n \neq a$ здесь отсутствует.

Теорема 1. *Определения непрерывности по Коши и по Гейне эквивалентны.*

▼ Покажем сначала, что если выполняется определение по Коши, то выполняется и определение по Гейне. Итак, f определена в некоторой окрестности точки a и

$$\forall \varepsilon > 0\ \exists \delta > 0\ \forall x,\ \rho_M(x,a) < \delta : \quad \rho(f(x), f(a)) < \varepsilon.$$

Возьмем последовательность $x_n \to a$, тогда $\forall \delta > 0\ \exists N\ \forall n > N : \quad \rho(x_n, a) < \delta$. Но тогда в силу написанного выше $\forall n > N : \quad \rho(f(x_n), f(a)) < \varepsilon$ и, значит, $f(x_n) \to f(a)$. Определение по Гейне выполняется.

Покажем теперь, что если выполняется определение по Гейне, то выполняется и определение по Коши. В силу замечания к определению по Коши при этом можно ограничиться случаем, когда a — предельная точка M. Так как для любой последовательности $x_n \neq a$, $x_n \to a$, последовательность $f(x_n) \to f(a)$), то $\lim\limits_{x\to a} f(x) = f(a)$, а значит, f непрерывна в точке a и по Коши. ▲

Теорема 2. *Пусть (M, ρ_M), (N, ρ_N), (L, ρ_L) — метрические пространства, функция f из M в N непрерывна в точке $a \in M$, а функция g из N в L непрерывна в точке $f(a) \in N$. Тогда функция $g(f)$ из M в L непрерывна в точке $a \in M$.*

▼ Возьмем последовательность $x_n \to a$, тогда $f(x_n) \to f(a)$ в силу непрерывности f в точке a, а $g(f(x_n)) \to g(f(a))$ в силу непрерывности g в точке $f(a)$. Значит, по определению Гейне $g(f)$ непрерывна в точке a. ▲

Для совершения операций сложения, вычитания, умножения и деления функций необходимо, чтобы эти операции были определены в пространстве значений функций. Поэтому нижеследующая теорема сформулирована только для случая, когда N — нормированное пространство.

Теорема 3. *Пусть M — метрическое, а N — нормированное пространство. Если f и g, функции из M в N, непрерывны в точке $a \in M$, то и $f \pm g$ непрерывна в точке $a \in M$. Если функция f из M в N и функция α из M в $\mathbb{R}$ (или в $\mathbb{C}$, если N линейное пространство над $\mathbb{C}$), непрерывны в точке $a \in M$, то αf (из M в N) непрерывна в точке $a \in M$, а если $\alpha(a) \neq 0$, то и $\frac{f}{\alpha}$ непрерывна в точке $a \in M$.*

▼ Возьмем любую последовательность $x_n \to a$ в M. Тогда $f(x_n) \to f(a)$ и $g(x_n) \to g(a)$, а значит по теореме о пределе суммы последовательностей и $(f \pm g)(x_n) = f(x_n) \pm g(x_n) \to f(a) \pm g(a) = (f \pm g)(a)$. Так как $f(x_n) \to f(a)$ и $\alpha(x_n) \to \alpha(a)$, то по теореме о произведении предела последовательностей и $(\alpha f)(x_n) = \alpha(x_n) f(x_n) \to \alpha(a) f(a) = (\alpha f)(a)$, а если $\alpha(a) \neq 0$, то и $\left(\frac{f}{\alpha}\right)(x_n) = \frac{f(x_n)}{\alpha(x_n)} \to \frac{f(a)}{\alpha(a)} = \left(\frac{f}{\alpha}\right)(a)$. ▲

Непрерывность функции на метрическом пространстве

Определение 6. Функция f, отображающая метрическое пространство M в метрическое пространство N, **непрерывна на** M, если f непрерывна в каждой точке M.

Определение 7. Функция f из метрического пространства M в метрическое пространство N **непрерывна на множестве** $E \subset M$, если f непрерывна на E как метрическом пространстве — подмножестве M.

Теорема 4. *Функция f, отображающая метрическое пространство M в метрическое пространство N, непрерывна на M тогда и только тогда, когда для любого открытого множества G из N его прообраз $f^{-1}(G) = \{x \in M : \ f(x) \in G\}$ — открытое множество в M.*

▼ Необходимость. Пусть f непрерывна на M и G — открытое подмножество N. Покажем, что $f^{-1}(G)$ — открытое подмножество M.

Пусть $x \in f^{-1}(G)$, т.е. $f(x) \in G$. Поскольку G — открытое множество, то найдется $B_\varepsilon(f(x)) \subset G$. Функция f непрерывна в точке x, поэтому найдется такая $B_\delta(x)$, что $f(B_\delta(x)) \subset B_\varepsilon(f(x)) \subset G$, т.е. $B_\delta(x) \subset f^{-1}(G)$ и, значит, $f^{-1}(G)$ — открытое множество.

Достаточность. Пусть для любого открытого множества $G \subset N$ $f^{-1}(G)$ — открытое множество в M. Покажем, что f непрерывна на M.

Пусть $x \in M$. Возьмем любую $B_\varepsilon(f(x))$, $f^{-1}\left(B_\varepsilon(f(x))\right)$ — открытое множество, x — его точка. Значит, найдется $B_\delta(x) \subset f^{-1}\left(B_\varepsilon(f(x))\right)$, т.е. найдется такое $B_\delta(x)$, что $f\left(B_\delta(x)\right) \subset B_\varepsilon(f(x))$. В силу произвольности $B_\varepsilon(f(x))$ получаем, что f непрерывна в точке x, а так как x — любая точка M, то f непрерывна на M. ▲

Непрерывные функции на компакте

Лемма 1. *Пусть (M, ρ) — метрическое пространство, $E \subset M$. Множество $G \subset E$ является открытым в E как метрическом пространстве с метрикой ρ тогда и только тогда, когда найдется такое открытое в M множество $\tilde{G}$, что $G = \tilde{G} \cap E$ (т.е. открытые множества в E — сужения открытых множеств в M).*

▼ Достаточность. Если $\tilde{G}$ — открытое множество в M, то $G = \tilde{G} \cap E$ — открытое множество в E, ведь если ε-окрестность точки x в M $B_\varepsilon(x) = \{y \in M : \ \rho(x, y) < \varepsilon\} \subset \tilde{G}$, то ε-окрестность точки x в E $B_\varepsilon(x) \cap E = \{y \in E : \ \rho(x, y) < \varepsilon\} \subset \tilde{G} \cap E = G$.

Необходимость. Если G — открытое множество в E, то для любой точки $x \in G$ существует в M такая ε-окрестность $B_\varepsilon(x)$, что ее пересечение с E, являющееся ε-окрестностью в E, принадлежит G, т.е. $B_\varepsilon(x) \cap E \subset G$. Объединение таких окрестностей

$$\tilde{G} = \bigcup_{B_\varepsilon(x)\cap E\subset G} B_\varepsilon(x) \supset G,$$

$$\tilde{G}\cap E = \bigcup_{B_\varepsilon(x)\cap E\subset G} B_\varepsilon(x)\cap E \subset G,$$

значит, $\tilde{G}\cap E = G$ и $\tilde{G}$ — открытое множество в M (как объединение открытых множеств). ▲

Лемма 2. *Если K — компакт в метрическом пространстве M, то K компакт в K как метрическом пространстве — подпространстве M.*

▼ Пусть $\{G_\lambda\}$ — открытые множества в K покрывающие K. Тогда по предыдущей лемме существуют такие открытые множества в M $\{\tilde{G}_\lambda\}$, что для каждого λ верно равенство $\tilde{G}_\lambda\cap K = G_\lambda$. Так как эти множества покрывают компакт K, то можно выделить конечное подпокрытие K $\{\tilde{G}_{\lambda_k}\}_{k=1}^n$. А тогда $\{G_{\lambda_k}\}_{k=1}^n = \{\tilde{G}_{\lambda_k}\cap K\}_{k=1}^n$ — конечное подпокрытие K в K как метрическом пространстве. Значит, K — компакт в K как метрическом пространстве. ▲

Образ компакта — компакт

Лемма 3. *Если функция f из метрического пространства M в метрическое пространство N непрерывна на M, K — компакт в M, то $f(K)$ — компакт в N.*

▼ Рассмотрим любое покрытие образа компакта $f(K)$ открытыми множествами G_λ. Тогда по теореме 4 о критерии непрерывности отображения $f^{-1}(G_\lambda)$ тоже открытые множества и они покрывают компакт K. Выделим конечное подпокрытие K $f^{-1}(G_{\lambda_1}),\ldots,f^{-1}(G_{\lambda_n})$. Тогда множества $G_{\lambda_1},\ldots,G_{\lambda_n}$ образуют конечное покрытие $f(K)$ и, значит, $f(K)$ — компакт в N. ▲

Теорема 5. *Если f — функция из метрического пространства M в метрическое пространство N, которая определена и непрерывна на компакте K из M, то $f(K)$ — компакт в N.*

▼ По лемме 2 K — компакт в K как метрическом пространстве — подпространстве M, а тогда по лемме 3 $f(K)$ — компакт в N. ▲

Теоремы Вейерштрасса

Теорема 6 (Вейерштрасса). *Если функция f из метрического пространства M в метрическое пространство N непрерывна на компакте K, то f ограничена на K.*

▼ По предыдущей теореме $f(K)$ — компакт, а значит ограниченное множество. ▲

Теорема 7 (Вейерштрасса). *Если f непрерывная действительнозначная функция на компакте K в метрическом пространстве, то f принимает на K наибольшее и наименьшее значения (достигает своего $\sup\limits_K f$ и $\inf\limits_K f$).*

▼ По предыдущей теореме существуют конечные $\sup\limits_K f$ и $\inf\limits_K f$. Это не могут быть внешние точки $f(K)$ (иначе это не точные грани), а значит, в силу замкнутости K (как компакта) это точки $f(K)$, т.е. найдутся $x_s\in K$ и $x_i\in K$, что $f(x_s) = \sup\limits_K f$ и $f(x_i) = \inf\limits_K f$. ▲

Замечание 4. Две последние теоремы можно было доказать аналогично тому, как это было сделано в одномерном случае основываясь на том, что по теореме

Больцано–Вейерштрасса из любой последовательности точек компакта можно выделить сходящуюся подпоследовательность.

Лекция 21 (28.04.20)
 Равномерная непрерывность. Связные множества и кривые

Равномерная непрерывность

Пусть, как и раньше, (M, ρ_M) и (N, ρ_n) — метрические пространства и функции — отображения из M (части M) в N.

Определение 1. Функция f, отображающая M в N, является **равномерно непрерывной** на M, если

$$\forall \varepsilon > 0 \; \exists \delta > 0 \; \forall x \in M \; \forall x' \in B_\delta(x) : \quad f(x') \in B_\varepsilon(f(x))$$

или, эквивалентно,

$$\forall \varepsilon > 0 \; \exists \delta > 0 \; \forall x, x' \in M :$$
$$\rho_M(x, x') < \delta \Rightarrow \rho_N(f(x), f(x')) < \varepsilon.$$

Определение 2. Функция f из M в N **равномерно непрерывна** на $E \subset M$, если f равномерно непрерывна на E как метрическом пространстве — подпространстве M.

Теорема 1 (Кантора). *Если функция f (из M в N) непрерывна на компакте $K \subset M$, то f равномерно непрерывна на K.*

▼ Предположим обратное, f непрерывна, но не является равномерно непрерывной на K, тогда

$$\exists \varepsilon > 0 \; \forall \delta > 0 \; \exists x, x' \in K :$$
$$\rho_M(x, x') < \delta \text{ и } \rho_N(f(x), f(x')) \geqslant \varepsilon.$$

Возьмем последовательность $\delta_n = \frac{1}{n}$ и для каждого δ_n найдем такие точки x_n и x'_n из K, что $\rho_M(x_n, x'_n) < \delta_n$ и $\rho_N(f(x_n), f(x'_n)) \geqslant \varepsilon$. Воспользуемся теоремой Больцано-Вейерштрасса и выберем сходящуюся подпоследовательность $x_{n_k} \xrightarrow[k\to\infty]{} x_0 \in K$. Так как $\rho\left(x'_{n_k}, x_0\right) \leqslant \rho\left(x'_{n_k}, x_{n_k}\right) + \rho\left(x_{n_k}, x_0\right) < \delta_{n_k} + \rho\left(x_{n_k}, x_0\right)$, то $x'_{n_k} \xrightarrow[k\to\infty]{} x_0 \in K$. В силу непрерывности f в точке x_0 $\lim\limits_{k\to\infty} f(x_{n_k}) = \lim\limits_{k\to\infty} f(x'_{n_k}) = f(x_0)$, значит, $\rho_N\left(f(x_{n_k}), f(x'_{n_k})\right) \xrightarrow[k\to\infty]{} 0$, что противоречит условию $\rho_N\left(f(x_{n_k}), f(x'_{n_k})\right) \geqslant \varepsilon > 0$. Следовательно, f равномерно непрерывна на K. ▲

Следствие 1. *Если действительнозначная функция f непрерывна на замкнутом ограниченном множестве $K \subset \mathbb{R}^n$, то f на K ограничена, принимает наибольшее и наименьшее значения и равномерно непрерывна.*

▼ Действительно, по критерию компактности в $\mathbb{R}^n$ K — компакт в $\mathbb{R}^n$ и, значит, по теоремам Вейерштрасса и по предыдущей теореме следствие справедливо. ▲

Связные множества

Определение 3. Метрическое пространство M называется **связным**, если его нельзя разбить на два непустых непересекающихся открытых множества или, эквивалентно, если его нельзя разбить на два непустых непересекающихся замкнутых множества (ведь если такие множества существуют, то они одновременно открыты и замкнуты) или, эквивалентно, если в метрическом пространстве не существует одновременно открытых и замкнутых подмножеств, кроме пустого множества $\varnothing$ и всего пространства M.

Определение 4. Метрическое пространство M называется **несвязным**, если его можно разбить на два непустых непересекающихся открытых множества или, эквивалентно, если его можно разбить на два непустых непересекающихся замкнутых множества (ведь если такие множества существуют, то они одновременно открыты и замкнуты) или, эквивалентно, если в метрическом пространстве существует одновременно открытое и замкнутое подмножество, отличное от пустого множества $\varnothing$ и всего пространства M.

Определение 5. Множество в метрическом пространстве называется **связным**, если оно связно как метрическое пространство (с той же метрикой).

Определение 6. Множество в метрическом пространстве называется **несвязным**, если оно несвязно как метрическое пространство (с той же метрикой).

Критерий связности

Теорема 2. *Метрическое пространство (множество в метрическом пространстве) несвязно тогда и только тогда, когда на нем существует непрерывная действительнозначная функция, принимающая ровно два значения.*

▼ Из определений следует, что достаточно рассмотреть только случай метрического пространства.

Необходимость. Покажем, что если метрическое пространство M несвязно, то на нем можно задать непрерывную функцию, принимающую ровно два значения $a, b \in \mathbb{R}$, $a \neq b$. Пусть $M = G_1 \cup G_2$, где G_1 и G_2 непустые непересекающиеся открытые множества. Определим функцию f на M следующим образом:

$$f(x) = \begin{cases} a, & \text{если } x \in G_1, \\ b, & \text{если } x \in G_2. \end{cases}$$

Поскольку прообраз любого множества из $\mathbb{R}$ или $\varnothing$ или G_1 или G_2 или $G_1 \cup G_2 = M$, то по критерию непрерывности функции на метрическом пространстве f — непрерывная функция на M.

Достаточность. Покажем теперь, что если на метрическом пространстве M можно задать непрерывную двузначную действительнозначную функцию, то M несвязно. Пусть функция f непрерывна на M и принимает ровно два значения a и b из $\mathbb{R}$. Найдем такое $\varepsilon > 0$, что $B_\varepsilon(a) \cap B_\varepsilon(a) = \varnothing$ (например, $\varepsilon = \frac{|b-a|}{2}$) и положим $G_1 = f^{-1}\left(B_\varepsilon(a)\right) = f^{-1}(\{a\})$ и $G_2 = f^{-1}\left(B_\varepsilon(b)\right) = f^{-1}(\{b\})$. По критерию непрерывности функции на метрическом пространстве это открытые множества. Ясно, что они непустые, непересекающиеся и $G_1 \cup G_2 = M$. ▲

Следствие 2. *Метрическое пространство (множество в метрическом пространстве) связно тогда и только тогда, когда на нем не существует непрерывной действительнозначной функции, принимающей ровно два значения.*

Теорема 3. *Множество E в метрическом пространстве M несвязно тогда и только тогда, когда в M существуют два непересекающихся открытых множества, которые покрывают E и каждое из которых пересекается с E.*

▼ Достаточность. Если такие открытые множества G_1 и G_2 в M существуют, то $G_1 \cap E$ и $G_2 \cap E$ образуют разбиение E как метрического пространства на два непустых непересекающихся открытых множества, т.е. E несвязно.

Необходимость. Если E несвязно, т.е. существуют множества $\overline{G}_1$ и $\overline{G}_2$, образующие разбиение E как метрического пространства на два непустых непересекающихся открытых множества, то положим

$$G_1 = \{x \in M : \inf_{t \in \overline{G}_1} \rho(x,t) < \inf_{t \in \overline{G}_2} \rho(x,t)\},$$
$$G_2 = \{x \in M : \inf_{t \in \overline{G}_1} \rho(x,t) > \inf_{t \in \overline{G}_2} \rho(x,t)\}.$$

Легко видеть, что $G_1 \supset \overline{G}_1$, $G_2 \supset \overline{G}_2$ и $G_1 \cap G_2 = \varnothing$. Остается показать, что это открытые множества.

По неравенству треугольника для любых точек x, y, t

$$\rho(x,t) \leqslant \rho(x,y) + \rho(y,t),$$

значит,

$$\inf_{t \in \overline{G}_1} \rho(x,t) \leqslant \rho(x,y) + \rho(y,t)$$

при любом $t \in \overline{G}_1$, откуда

$$\inf_{t \in \overline{G}_1} \rho(x,t) \leqslant \rho(x,y) + \inf_{t \in \overline{G}_1} \rho(y,t).$$

Поменяв местами x и y, получим неравенство

$$\inf_{t \in \overline{G}_1} \rho(y,t) \leqslant \rho(x,y) + \inf_{t \in \overline{G}_1} \rho(x,t).$$

Из двух последних неравенств следует, что

$$\left| \inf_{t \in \overline{G}_1} \rho(x,t) - \inf_{t \in \overline{G}_1} \rho(y,t) \right| \leqslant \rho(x,y)$$

и, значит, функция $\varphi(x) = \inf\limits_{t \in \overline{G}_1} \rho(x,t)$ непрерывна на M. Аналогично, непрерывна на M функция $\psi(x) = \inf\limits_{t \in \overline{G}_2} \rho(x,t)$. Значит, непрерывна разность этих функций $h(x) = \varphi(x) - \psi(x)$. А тогда

$$G_1 = h^{-1}((0,+\infty)) \text{ и } G_2 = h^{-1}((-\infty,0))$$

— прообразы открытых множеств, а значит, открытые множества. ▲

Теорема о промежуточных значениях

Напомним, что промежутком называется любое подмножество $\mathbb{R}$, содержащее вместе с каждой парой точек и все точки, лежащие между ними.

Теорема 4. *Подмножество $\mathbb{R}$ связно тогда и только тогда, когда это промежуток.*

▼ Достаточность. Если непрерывная на промежутке действительнозначная функция принимает в двух его точках разные значения, то между этими точками она принимает по теореме Больцано–Коши все промежуточные значения. Значит, по следствию критерия связности, промежуток — связное множество.

Необходимость. Если подмножество $\mathbb{R}$ вместе с парой точек не содержит лежащую между ними точку c, то функция

$$f(x) = \begin{cases} 0 & \text{при } x \leqslant c, \\ 1 & \text{при } x > c, \end{cases}$$

принимает ровно два значения на этом подмножестве и непрерывна на нем (она непрерывна во всех точках $\mathbb{R}$, кроме c), значит, по критерию несвязности, это несвязное множество. ▲

Теорема 5. *Образ связного множества при его непрерывном отображении в метрическое пространство связен.*

▼ Если образ $\varphi(D)$ связного множества D несвязен, то на $\varphi(D)$ существует непрерывная действительнозначная двузначная функция f. Но тогда $f(\varphi)$ будет непрерывной действительнозначной двузначной функцией на D, что по противоречит связности D. ▲

Теорема 6 (Больцано–Коши). *Действительнозначная непрерывная на связном множестве функция принимает все промежуточные значения (т.е., если она принимает значения a и b, $a < b$, то она принимает все значения отрезка $[a, b]$.*

По теореме образ связного множества связен, по теореме 4 это промежуток, который вместе с каждой парой точек содержит и все лежащие между ними точки. ▲

Кривые и линейно связные множества

Определение 7. **Непрерывной кривой** или просто **кривой** в метрическом пространстве будем называть непрерывный образ промежутка, т.е. образ промежутка при непрерывном отображении его в метрическое пространство.

Теорема 8. *Непрерывная кривая — связное множество.*

▼ Эта теорема непосредственно следует из теорем о связности промежутка и о том, что образ связного множества — связное множество. ▲

Определение 8. Множество в метрическом пространстве называется **линейно связным**, если любые две его точки можно соединить непрерывной кривой, принадлежащей этому множеству (т.е. через любые две его точки можно провести непрерывную кривую, принадлежащую этому множеству).

Теорема 9. *Линейно связное множество — связно.*

▼ Предположим, что это несвязное множество. Тогда на нем существует непрерывная действительнозначная двузначная функция f. Соединим пару точек, в которых эта функция принимает различные значения, непрерывной кривой, принадлежащей множеству, и получим непрерывную кривую, на которой непрерывная действительнозначная функция принимает ровно два значения, что невозможно по предыдущей теореме. ▲

Определение 9. Отрезком в нормированном пространстве N с концами в точках x и y называется множество $\{t \in N : \ t = (1-\alpha)x + \alpha y, \text{ где } 0 \leqslant \alpha \leqslant 1\}$. Обозначается отрезок $[x, y]$.

Отрезок — непрерывная кривая (ведь t как функция от α непрерывна).

Определение 4. Ломаной в нормированном пространстве называется кривая, состоящая из конечной последовательности отрезков $[a_{k-1}, a_k]$, $k = 1, \ldots, n$.

Это непрерывная кривая, что является очевидны, если отрезок $[a_{k-1}, a_k]$ рассматривать как непрерывное отображение отрезка $[k-1, k] \subset \mathbb{R}$ по формуле $\alpha \longrightarrow (k-\alpha)a_{k-1} + (\alpha - k + 1)a_k$, а всю ломаную как отображение отрезка $[0, n]$ складывающееся из приведенных отображений.

Теорема 10. *В нормированном пространстве открытое множество связно тогда и только тогда, когда любые две его точки можно соединить принадлежащей этому множеству ломаной.*

▼ Достаточность очевидна из предыдущей теоремы, ведь открытое множество в этом случае линейно связно.

Докажем необходимость. Пусть G — открытое связное множество в нормированном пространстве N, x_0 — произвольная точка G и $G_0 = \{x \in G : \ x \text{ и } x_0$ соединяются принадлежащей G ломаной$\}$. Очевидно, $x_0 \in G_0$. Если $x \in G_0$, то существует $B_\varepsilon(x) \subset G$, для любой точки $y \in B_\varepsilon(x)$ отрезок $[x, y] \subset B_\varepsilon(x) \subset G$ (ведь если $0 \leqslant \alpha \leqslant 1$, то $\|(1-\alpha)x + \alpha y - x\| = \|\alpha(y - x)\| = \alpha\|y - x\| < \alpha\varepsilon$), значит, все точки $B_\varepsilon(x)$ соединяются с x_0 ломаной и $B_\varepsilon(x) \subset G_0$. Следовательно, G_0 открытое множество в N, а значит, и в G как метрическом пространстве — подпространстве N. Покажем, что $G \setminus G_0$ открытое множество. Если $x \in G \setminus G_0$, то существует $B_\varepsilon(x) \subset G$, для любой точки $y \in B_\varepsilon(x)$ отрезок $[x, y] \subset B_\varepsilon(x) \subset G$, значит, $B_\varepsilon(x) \cap G_0 = \varnothing$, т.е. $B_\varepsilon(x) \subset G \setminus G_0$. В силу связности G или G_0 или $G \setminus G_0$ пусто, но $x_0 \in G_0$, значит, $G \setminus G_0 = \varnothing$, т.е. $G = G_0$. Теорема доказана. ▲

Лекция 22 (05.05.20)
Дифференцируемость функции в точке. Достаточное условие дифференцируемости. Геометрический смысл дифференцируемости

Линейные отображения

Определение 1. Отображение (функция) f из линейного пространства L_1 в линейное пространство L_2 называется **линейным**, если из существования $f(a)$ следует, что для любого $\alpha \in \mathbb{R}$ (или $\alpha \in \mathbb{C}$, если L_1 и L_2 — линейные пространства над полем комплексных чисел $\mathbb{C}$) существует $f(\alpha a) = \alpha f(a)$ и из существования $f(a)$ и $f(b)$ следует существование $f(a + b) = f(a) + f(b)$.

Теорема 1. *Определенная на $\mathbb{R}^n$ действительнозначная функция φ линейна тогда и только тогда, когда существуют $\alpha_k \in \mathbb{R}$, $k = 1, \ldots, n$, что для любого $\vec{x} \in \mathbb{R}^n$ функция $\varphi(\vec{x}) = \varphi(x_1, \ldots, x_n) = \sum\limits_{k=1}^{n} \alpha_k x_k$. Если φ линейна, то она непрерывна и существует такое $C \geqslant 0$, что для любого $\vec{x} \in \mathbb{R}^n$ верна оценка $|\varphi(\vec{x})| \leqslant C\|\vec{x}\|$.*

▼ Если $\varphi(\vec{x}) = \varphi(x_1, \ldots, x_n) = \sum\limits_{k=1}^{n} \alpha_k x_k$, то, очевидно, φ — линейная функция. Если φ — линейная функция, то положим $\alpha_k = \varphi(\vec{e}_k)$ (где $\vec{e}_k$ — вектор, k-ая координата которого 1, а остальные 0), $k = 1, \ldots, n$. Тогда

$$\varphi(\vec{x}) = \varphi\left(\sum_{k=1}^{n} x_k \vec{e}_k\right) = \sum_{k=1}^{n} x_k \varphi(\vec{e}_k) = \sum_{k=1}^{n} \alpha_k x_k.$$

Ясно, что φ непрерывна и по неравенству Коши–Буняковского

$$|\varphi(\vec{x})| = \left|\sum_{k=1}^{n} \alpha_k x_k\right| \leqslant \sqrt{\sum_{k=1}^{n} \alpha_k^2} \cdot \sqrt{\sum_{k=1}^{n} x_k^2} = \sqrt{\sum_{k=1}^{n} \alpha_k^2}\|\vec{x}\|$$

и, значит, заключительное утверждение теоремы верно с $C = \sqrt{\sum\limits_{k=1}^{n} \alpha_k^2}$. ▲

Дифференцируемость функции в точке

Определение 2. Функция f из пространства $\mathbb{R}^n$ в $\mathbb{R}$ называется **дифференцируемой в точке** $x_0 \in \mathbb{R}^n$, если f определена в окрестности этой точки и ее приращение $\Delta f = f(x_0 + \Delta x) - f(x_0) = f(x) - f(x_0)$ представляется в виде

$$\Delta f = f(x_0 + \Delta x) - f(x_0) = L(\Delta x) + o(\|\Delta x\|),$$

где L — линейная функция, отображающая $\mathbb{R}^n$ в $\mathbb{R}$, $o(\|\Delta x\|) = o(1) \cdot \|\Delta x\|$, $o(1) \to 0$ при $\Delta x \to 0$ (т.е. при $\|\Delta x\| \to 0$), а приращение аргумента Δx таково, что f определена в точке $x_0 + \Delta x$ (что выполняется при достаточно малых Δx).

Определение 3. Линейная функция L называется (**полным**) **дифференциалом** (или **производным отображением** или **полной производной**) функции f в точке x_0 и обозначается $\boldsymbol{d}f(\Delta x)\big|_{x_0} = L(\Delta x)$. Будут также использоваться обозначения $\boldsymbol{d}f(x_0, \Delta x)$, $\boldsymbol{d}f(\Delta x)$ (если понятно, в какой точке происходит дифференцирование) и $\boldsymbol{d}f(x_0)$.

Теорема 2. *Если функция f из пространства $\mathbb{R}^n$ в $\mathbb{R}$ дифференцируема в точке, то она в ней непрерывна.*

▼ Возьмем любую последовательность $\Delta x_k \to 0$. Из непрерывности линейной функции L следует, что $L(\Delta x_k) \to 0$. Функция $f(x_0 + \Delta x_k)$ определена для всех достаточно малых Δx_k, т.е. начиная с некоторого номера, и приращение функции

$$\Delta f = f(x_0 + \Delta x_k) - f(x_0) = L(\Delta x_k) + o(\|\Delta x_k\|) \xrightarrow[k \to \infty]{} 0,$$

значит (по определению предела по Гейне), $\lim\limits_{\Delta x \to 0} f(x_0 + \Delta x) = f(x_0)$ и, следовательно, f непрерывна в точке x_0. ▲

Определение 4. Частной производной (первого порядка) функции f (из $\mathbb{R}^n$ в $\mathbb{R}$) по переменной x_k в точке $\vec{x}^0$ называется предел

$$\lim_{t \to 0} \frac{f(\vec{x}^0 + t\vec{e}_k) - f(\vec{x}^0)}{t} = \lim_{t \to 0} \frac{f(x_1^0, \ldots, x_{k-1}^0, x_k^0 + t, x_{k+1}^0, \ldots, x_n^0) - f(x_1^0, \ldots, x_n^0)}{t}.$$

Обозначение. Частную производную функции f по переменной x_k в точке $\vec{x}^0$ обозначают

$$f'_{x_k}(\vec{x}^0) \text{ или } f'_{x_k}\big|_{\vec{x}=\vec{x}^0} \text{ или } \frac{\partial f(\vec{x}^0)}{\partial x_k} \text{ или } \frac{\partial f(\vec{x})}{\partial x_k}\Big|_{\vec{x}=\vec{x}^0}.$$

Фактически это обычная производная в ситуации, когда функция f рассматривается как функция переменной x_k при фиксировании остальных переменных.

Теорема 3. *Если функция f дифференцируема в точке $\vec{x}^0$, то она имеет в этой точке все частные производные $\frac{\partial f(\vec{x}^0)}{\partial x_k}$, $k = 1, \ldots, n$, и ее дифференциал*

$$\boldsymbol{d}f\left(\overrightarrow{\Delta x}\right)\big|_{\vec{x}^0} = L\left(\overrightarrow{\Delta x}\right) = \sum_{k=1}^{n} \frac{\partial f(\vec{x}^0)}{\partial x_k} \Delta x_k.$$

▼ Дифференциал — линейная функция приращения аргумента, $\boldsymbol{d}f(\overrightarrow{\Delta x}) = L(\overrightarrow{\Delta x}) = \sum\limits_{k=1}^{n} \alpha_k \Delta x_k$ (по теореме 1), значит, $\Delta f = \sum\limits_{k=1}^{n} \alpha_k \Delta x_k + o\left(\|\overrightarrow{\Delta x}\|\right)$. Используя последнее равенство получаем, что существует

$$\lim_{t\to 0} \frac{f(\vec{x}^0 + t\vec{e}_k) - f(\vec{x}^0)}{t} = \lim_{t\to 0} \frac{\alpha_k t + o(|t|)}{t} = \lim_{t\to 0} (\alpha_k + o(1)) = \alpha_k,$$

то есть существует $\frac{\partial f(\vec{x}^0)}{\partial x_k} = \alpha_k$, $k = 1, \ldots, n$, и

$$\boldsymbol{d}f(\vec{x}^0) = \sum_{k=1}^{n} \alpha_k \Delta x_k = \sum_{k=1}^{n} \frac{\partial f(\vec{x}^0)}{\partial x_k} \Delta x_k. \blacktriangle$$

Замечание 1. Если дифференциал $\boldsymbol{d}f\left(\overrightarrow{\Delta x}\right)\big|_{\vec{x}^0}$, являющийся при фиксированном $\vec{x}^0$ функцией от $\overrightarrow{\Delta x}$, рассматривать как функцию от $\vec{x}$, $\overrightarrow{\Delta x} = \vec{x} - \vec{x}^0$, то его частная производная по переменной x_k в любой точке равна $\frac{\partial f(\vec{x}^0)}{\partial x_k}$.

Отметим, что существование частных производных в точке не гарантирует даже непрерывности, тем более дифференцируемости, функции. Например, функция двух переменных

$$f(x, y) = \operatorname{sign} xy = \begin{cases} 0, & \text{если } xy = 0, \\ 1, & \text{если } xy > 0, \\ -1, & \text{если } xy < 0, \end{cases}$$

имеет в точке $(0, 0)$ равные нулю частные производные $\frac{\partial f(0,0)}{\partial x}$ и $\frac{\partial f(0,0)}{\partial y}$ и разрывна в $(0, 0)$.

Достаточное условие дифференцируемости

Определение 5. Функция f (из $\mathbb{R}^n$ в $\mathbb{R}$) имеет в точке $\vec{x}^0$ **непрерывную частную производную** $\frac{\partial f}{\partial x_k}$, если эта частная производная существует в окрестности точки $\vec{x}^0$ и непрерывна в ней.

Теорема 4. *Если функция f имеет непрерывные частные производные (первого порядка) в точке $\vec{x}^0$ $\frac{\partial f(\vec{x})}{\partial x_k}$, $k = 1, \ldots, n$, то f дифференцируема в этой точке.*

▼ Проведем индукцию по n. При $n = 1$ утверждение верно (по теореме для одной переменной), причем достаточно просто существования производной в точке). Предположим, что оно верно при $n = m$ и докажем его при $n = m + 1$. Имеем равенство

$$\Delta f = f(\vec{x}) - f(\vec{x}^0) = f(x_1, \ldots, x_m, x_{m+1}) - f(x_1, \ldots, x_m, x_{m+1}^0) +$$
$$+ f(x_1, \ldots, x_m, x_{m+1}^0) - f(x_1^0, \ldots, x_m^0, x_{m+1}^0).$$

По теореме Лагранжа

$$f(x_1,\ldots,x_m,x_{m+1})-f(x_1,\ldots,x_m,x_{m+1}^0)=$$
$$=\frac{\partial}{\partial x_{m+1}}f(x_1,\ldots,x_m,x_{m+1}^0+\theta\Delta x_{m+1})\Delta x_{m+1},$$

где $0<\theta<1$. Так как частная производная $\frac{\partial f}{\partial x_{m+1}}$ непрерывна в точке $\vec{x}^0$, то

$$\frac{\partial}{\partial x_{m+1}}f(x_1,\ldots,x_m,x_{m+1}^0+\theta\Delta x_{m+1})=\frac{\partial f(\vec{x}^0)}{\partial x_{m+1}}+o(1),$$

где $o(1)\to 0$ при $\overrightarrow{\Delta x}\to 0$.

Рассматривая функцию f как функцию только первых m переменных при фиксированной $m+1$-ой переменной, по индукционному предположению имеем:

$$f(x_1^0+\Delta x_1,\ldots,x_m^0+\Delta x_m,x_{m+1}^0)-f(x_1^0,\ldots,x_m^0,x_{m+1}^0)=$$
$$=\sum_{k=1}^{m}\frac{\partial f(\vec{x}^0)}{\partial x_k}\Delta x_k+o\left(\|(\overrightarrow{\Delta x})_m\|\right),$$

где $(\overrightarrow{\Delta x})_m=(\Delta x_1,\ldots,\Delta x_m,0)$.

В итоге имеем:

$$\Delta f=f(\vec{x}^0+\overrightarrow{\Delta x})-f(\vec{x}^0)=\frac{\partial f(\vec{x}^0)}{\partial x_{m+1}}\Delta x_{m+1}+o(1)\cdot\Delta x_{m+1}+$$
$$+\sum_{k=1}^{m}\frac{\partial f(\vec{x}^0)}{\partial x_k}\Delta x_k+o\left(\|(\overrightarrow{\Delta x})_m\|\right)=\sum_{k=1}^{m+1}\frac{\partial f(\vec{x}^0)}{\partial x_k}\Delta x_k+o\left(\|\overrightarrow{\Delta x}\|\right).\blacktriangle$$

Замечание 1. В теореме можно отказаться от непрерывности частной производной $\frac{\partial f}{\partial x_1}$, достаточно ее существования в точке $\vec{x}^0$. Ведь при $n=1$ такое утверждение верно, а далее из верности этого предположения для $n=m$ следуя доказательству теоремы получаем его верность для $n=m+1$. Так как порядок переменных несущественен, то, значит, в формулировке теоремы можно отказаться от требования непрерывности любой одной частной производной $\frac{\partial f}{\partial x_k}$, достаточно ее существования в точке $\vec{x}^0$.

Геометрический смысл дифференцируемости

Рассмотрим поверхность S в пространстве $\mathbb{R}^{n+1}$, являющуюся графиком функции f, $x_{n+1}=f(x_1,\ldots,n)$.

Определение 6. Плоскость (гиперплоскость) Π, проходящая через точку поверхности S $\vec{x}^0\in\mathbb{R}^{n+1}$, $x_{n+1}^0=f(x_1^0,\ldots,x_n^0)$, и определяемая уравнением

$$x_{n+1}-x_{n+1}^0=\sum_{k=1}^{n}\alpha_k(x_k-x_k^0),$$

называется **касательной плоскостью** к поверхности S в точке $\vec{x}^0\in\mathbb{R}^{n+1}$, если угол между двумя прямыми, прямой из плоскости, проходящей через точки $\left(x_1^0,\ldots,x_n^0,x_{n+1}^0\right)$ и $\left(x_1,\ldots,x_n,x_{n+1}^0+\sum_{k=1}^n\alpha_k(x_k-x_k^0)\right)$, и секущей, проходящей через точки $\left(x_1^0, \ldots, x_n^0, x_{n+1}^0\right)$ и $(x_1, \ldots, x_n, f(x_1, \ldots, x_n))$, стремится к нулю при $(x_1, \ldots, x_n) \to (x_1^0, \ldots, x_n^0)$.

Теорема 5. *Функция f, определенная в окрестности точки $\vec{x}^0 \in \mathbb{R}^n$, дифференцируема в этой точке тогда и только тогда, когда определяемая ею поверхность $S \subset \mathbb{R}^{n+1}$ имеет в точке $(x_1^0, \ldots, x_n^0, f(x_1^0, \ldots, x_n^0))$ касательную плоскость. При этом последняя определяется уравнением*

$$x_{n+1} - f(\vec{x}^0) = \sum_{k=1}^{n} \frac{\partial f(\vec{x}^0)}{\partial x_k}(x_k - x_k^0).$$

▼ Пусть α — угол между двумя вышеуказанными прямыми, прямой из плоскости и секущей, а β и $\beta+\alpha$ — соответственно углы между ними и прямой, проходящей через точки $(x_1^0, \ldots, x_n^0, x_{n+1}^0)$ и $(x_1, \ldots, x_n, x_{n+1}^0)$. Угол β не может приближаться сколь угодно близко к $\pm\frac{\pi}{2}$, поскольку

$$|\operatorname{tg}\beta| = \left|\frac{x_{n+1} - x_{n+1}^0}{\|\overrightarrow{\Delta x}\|}\right| = \left|\frac{\sum_{k=1}^n \alpha_k \Delta x_k}{\|\overrightarrow{\Delta x}\|}\right|,$$

где $\overrightarrow{\Delta x} = (\Delta x_1, \ldots, \Delta x_n) \in \mathbb{R}^n$, $\Delta x_k = x_k - x_k^0$, $k = 1, \ldots, n$. По теореме предыдущей лекции о линейной функции на n-мерном пространстве существует такая постоянная C, что $|\operatorname{tg}\beta| < C$. А так как

$$|\operatorname{tg}(\alpha+\beta) - \operatorname{tg}\beta| \cdot \left\|\overrightarrow{\Delta x}\right\| = \left|\frac{\sin\alpha}{\cos(\alpha+\beta)\cdot\cos\beta}\right| \cdot \left\|\overrightarrow{\Delta x}\right\| = \left|f(\vec{x}) - f(\vec{x}^0) - \sum_{k=1}^{n} \alpha_k \Delta x_k\right|,$$

где последнее выражение — расстояние между точками $(x_1, \ldots, x_n, f(\vec{x}^0) + \sum\limits_{k=1}^{n} \alpha_k \Delta x_k)$ и $(x_1, \ldots, x_n, f(\vec{x}))$, то угол α, $-\frac{\pi}{2} < \alpha < \frac{\pi}{2}$, стремится к нулю тогда и только тогда, когда указанное расстояние $\left|f(\vec{x}) - f(\vec{x}^0) - \sum\limits_{k=1}^{n} \alpha_k \Delta x_k\right|$ является $o(\|\overrightarrow{\Delta x}\|)$, что эквивалентно дифференцируемости f в точке $\vec{x}^0$. При этом $\sum\limits_{k=1}^{n} \alpha_k \Delta x_k)$ является дифференциалом функции f и, значит, $\alpha_k = \frac{\partial f(\vec{x}^0)}{\partial x_k}$ по предыдущей лекции. ▲

Лекция 23 (08.05.20)
Производная по направлению.
Градиент функции.
Правила дифференцирования

Производная по направлению

Определение 1. Производной по направлению $\vec{w} = (w_1, \ldots, w_n)$, $\|\vec{w}\| = 1$, функции f в точке $\vec{x}^0$ называется

$$\frac{\partial f(\vec{x}^0)}{\partial \vec{w}} = \lim_{t \to +0} \frac{f(\vec{x}^0 + t\vec{w}) - f(\vec{x}^0)}{t}.$$

Отметим, что если существует $\frac{\partial f(\vec{x}^0)}{\partial x_k}$, то производная по направлению $\vec{e}_k$ в точке $\vec{x}^0$ — это $\frac{\partial f(\vec{x}^0)}{\partial x_k}$, а по направлению $-\vec{e}_k$ — это $-\frac{\partial f(\vec{x}^0)}{\partial x_k}$. Производная по направлению $\vec{w}$ функции f в точке $\vec{x}^0$ — это правая производная в точке 0 функции $f(\vec{x}^0 + t\vec{w})$ как функции одной переменной t. Левая производная этой функции отличается от производной по направлению $-\vec{w}$ знаком.

Теорема 1. *Если функция f дифференцируема в точке $\vec{x}^0$, то f имеет в этой точке производную по любому направлению $\vec{w}$ и при этом*

$$\frac{\partial f(\vec{x}^0)}{\partial \vec{w}} = \sum_{k=1}^{n} \frac{\partial f(\vec{x}^0)}{\partial x_k} w_k.$$

▼ Поскольку f дифференцируема в точке $\vec{x}^0$, то

$$f(\vec{x}^0 + t\vec{w}) - f(\vec{x}^0) = \sum_{k=1}^{n} \frac{\partial f(\vec{x}^0)}{\partial x_k} \cdot t w_k + o(t\|\vec{w}\|)$$

и, значит,

$$\frac{\partial f(\vec{x}^0)}{\partial \vec{w}} = \lim_{t\to+0} \frac{f(\vec{x}^0 + t\vec{w}) - f(\vec{x}^0)}{t} = \sum_{k=1}^{n} \frac{\partial f(\vec{x}^0)}{\partial x_k} w_k. \blacktriangle$$

Отметим, что существование производных по любому направлению в точке не гарантирует даже непрерывности, тем более дифференцируемости функции в этой точке.

Например, функция двух переменных

$$f(x,y) = \begin{cases} 1, & \text{если } x^2 + \left(y - \frac{1}{2}\right)^2 = \frac{1}{4},\ x \neq 0 \text{ и } y \neq 0, \\ 0 & \text{в остальных точках} \end{cases}$$

имеет в точке $(0,0)$ производную по любому направлению равную нулю, но разрывна в ней.

Градиент функции

Определение 2. Вектор $\left(\frac{\partial f(\vec{x}^0)}{\partial x_1}, \ldots, \frac{\partial f(\vec{x}^0)}{\partial x_n}\right)$ называется **градиентом функции** f в точке $\vec{x}^0$ и обозначается $\operatorname{grad} f(\vec{x}^0)$.

Определение 3. Если $\vec{x}$ и $\vec{y}$ — вектора из $\mathbb{R}^n$, то их **скалярным произведением** называется величина $(\vec{x}, \vec{y}) = \sum\limits_{k=1}^{n} x_k y_k$.

В введенных обозначениях производная по направлению

$$\frac{\partial f(\vec{x}^0)}{\partial \vec{w}} = \left(\operatorname{grad} f(\vec{x}^0), \vec{w}\right).$$

Теорема 2. *Если функция f дифференцируема в точке $\vec{x}^0$, то длина $\operatorname{grad} f(\vec{x}^0)$ равна максимальной величине производной по направлению в точке $\vec{x}^0$ и, если $\operatorname{grad} f(\vec{x}^0) \neq 0$, то он направлен в ту же сторону, что и единственный вектор направления $\vec{w}$, вдоль которого производная $\frac{\partial f(\vec{x}^0)}{\partial \vec{w}}$ максимальна.*

▼ Из неравенства Коши-Буняковского следует, что

$$\frac{\partial f(\vec{x}^0)}{\partial \vec{w}} \leqslant \|\operatorname{grad} f(\vec{x}^0)\| \cdot \|\vec{w}\| = \|\operatorname{grad} f(\vec{x}^0)\|.$$

Если $\operatorname{grad} f(\vec{x}^0) \neq 0$, то возьмем $\vec{w} = \frac{\operatorname{grad} f(\vec{x}^0)}{\|\operatorname{grad} f(\vec{x}^0)\|}$ и тогда

$$\frac{\partial f(\vec{x}^0)}{\partial \vec{w}} = \left(\operatorname{grad} f(\vec{x}^0), \frac{\operatorname{grad} f(\vec{x}^0)}{\|\operatorname{grad} f(\vec{x}^0)\|}\right) = \|\operatorname{grad} f(\vec{x}^0)\|.$$

По тому же неравенству Коши-Буняковского равенство может достигаться лишь на векторе, пропорциональном $\operatorname{grad} f(\vec{x}^0)$, а так как вектор $-\frac{\operatorname{grad} f(\vec{x}^0)}{\|\operatorname{grad} f(\vec{x}^0)\|}$ не годится (по этому направлению $\frac{\partial f(\vec{x}^0)}{\partial \vec{w}}$ минимальна), то такое направление единственно. ▲

Правила дифференцирования

Теорема 3. *Если функции f и g дифференцируемы в точке $\vec{x}^0$, то дифференцируемы в ней и функции $f \pm g$, $f \cdot g$, а если $g(\vec{x}^0) \neq 0$, то и $\frac{f}{g}$ и при этом*

$$\begin{gathered}
\boldsymbol{d}(f \pm g)(\vec{x}^0) = \boldsymbol{d}f(\vec{x}^0) \pm \boldsymbol{d}g(\vec{x}^0), \\
\boldsymbol{d}(f \cdot g)(\vec{x}^0) = f(\vec{x}^0)\boldsymbol{d}g(\vec{x}^0) + g(\vec{x}^0)\boldsymbol{d}f(\vec{x}^0), \\
\boldsymbol{d}\left(\frac{f}{g}\right)(\vec{x}^0) = \frac{g(\vec{x}^0)\boldsymbol{d}f(\vec{x}^0) - f(\vec{x}^0)\boldsymbol{d}g(\vec{x}^0)}{g^2(\vec{x}^0)}.
\end{gathered}$$

▼ Поскольку

$$\begin{gathered}
\Delta(f \pm g) = \Delta f \pm \Delta g = \boldsymbol{d}f(\vec{x}^0) + o(\|\overrightarrow{\Delta x}\|) + \\
+\boldsymbol{d}g(\vec{x}^0) + o(\|\overrightarrow{\Delta x}\|) = \boldsymbol{d}f(\vec{x}^0) \pm \boldsymbol{d}g(\vec{x}^0) + o(\|\overrightarrow{\Delta x}\|),
\end{gathered}$$

то $f \pm g$ дифференцируема в точке $\vec{x}^0$ и $\boldsymbol{d}(f \pm g)(\vec{x}^0) = \boldsymbol{d}f(\vec{x}^0) \pm \boldsymbol{d}g(\vec{x}^0)$.

Аналогично, учитывая непрерывность (в силу дифференцируемости) функции g в точке $\vec{x}^0$, получаем

$$\begin{gathered}
\Delta(f \cdot g) = f(\vec{x}^0 + \overrightarrow{\Delta x})g(\vec{x}^0 + \overrightarrow{\Delta x}) - f(\vec{x}^0)g(\vec{x}^0) = g(\vec{x}^0 + \overrightarrow{\Delta x})\big(f(\vec{x}^0 + \overrightarrow{\Delta x}) - f(\vec{x}^0)) + \\
+f(\vec{x}^0)(g(\vec{x}^0 + \overrightarrow{\Delta x}) - g(\vec{x}^0)) = (g(\vec{x}^0) + o(1))\Delta f + f(\vec{x}^0)\Delta g = \\
= \big(g(\vec{x}^0) + o(1)\big)\left(\boldsymbol{d}f(\vec{x}^0) + o(\|\overrightarrow{\Delta x}\|)\right) + f(\vec{x}^0)\left(\boldsymbol{d}g(\vec{x}^0) + o(\|\overrightarrow{\Delta x}\|)\right) = g(\vec{x}^0)\boldsymbol{d}f(\vec{x}^0) + \\
+f(\vec{x}^0)\boldsymbol{d}g(\vec{x}^0) + g(\vec{x}^0) \cdot o(\|\overrightarrow{\Delta x}\|) + o(1) \cdot \boldsymbol{d}f(\vec{x}^0) + o(1) \cdot o(\|\overrightarrow{\Delta x}\|) + f(\vec{x}^0) \cdot o(\|\overrightarrow{\Delta x}\|),
\end{gathered}$$

где последние четыре слагаемых — $o(\|\overrightarrow{\Delta x}\|)$. Для всех, кроме $o(1) \cdot \boldsymbol{d}f(\vec{x}^0)$, это очевидно. По предыдущей лекции $\boldsymbol{d}f(\vec{x}^0) = O(\|\overrightarrow{\Delta x}\|)$, а так как $o(1) \cdot O(\|\overrightarrow{\Delta x}\|) = o(\|\overrightarrow{\Delta x}\|)$, то и $o(1) \cdot \boldsymbol{d}f(\vec{x}^0) = o(\|\overrightarrow{\Delta x}\|)$. Значит, $f \cdot g$ дифференцируема в точке $\vec{x}^0$ и для ее дифференциала верно равенство $\boldsymbol{d}(f \cdot g)(\vec{x}^0) = f(\vec{x}^0)\boldsymbol{d}g(\vec{x}^0) + g(\vec{x}^0)\boldsymbol{d}f(\vec{x}^0)$.

Точно также устанавливается дифференцируемость частного.

$$\Delta\left(\frac{f}{g}\right)=\frac{f(\vec{x}^0+\overrightarrow{\Delta x})}{g(\vec{x}^0+\overrightarrow{\Delta x})}-\frac{f(\vec{x}^0)}{g(\vec{x}^0)}=\frac{f(\vec{x}^0+\overrightarrow{\Delta x})g(\vec{x}^0)-f(\vec{x}^0)g(\vec{x}^0+\overrightarrow{\Delta x})}{g(\vec{x}^0+\overrightarrow{\Delta x})g(\vec{x}^0)}=$$

$$=\frac{g(\vec{x}^0)\Delta f-f(\vec{x}^0)\Delta g}{g(\vec{x}^0+\overrightarrow{\Delta x})g(\vec{x}^0)}=\left(\frac{1}{g(\vec{x}^0)}+o(1)\right)\frac{1}{g(\vec{x}^0)}\cdot\Big(g(\vec{x}^0)\boldsymbol{d}f(\vec{x}^0)+g(\vec{x}^0)\cdot o(\|\overrightarrow{\Delta x}\|)-$$

$$-f(\vec{x}^0)\boldsymbol{d}g(\vec{x}^0)-f(\vec{x}^0)\cdot o(\|\overrightarrow{\Delta x}\|)\Big)=\frac{g(\vec{x}^0)\boldsymbol{d}f(\vec{x}^0)-f(\vec{x}^0)\boldsymbol{d}g(\vec{x}^0)}{g^2(\vec{x}^0)}+\frac{o(\|\overrightarrow{\Delta x}\|)}{g(\vec{x}^0)}-$$

$$-\frac{f(\vec{x}^0)}{g^2(\vec{x}^0)}\cdot o(\|\overrightarrow{\Delta x}\|)+o(1)\cdot\boldsymbol{d}f(\vec{x}^0)+o(1)\cdot o(\|\overrightarrow{\Delta x}\|)-o(1)\cdot\frac{f(\vec{x}^0)}{g(\vec{x}^0)}\cdot\boldsymbol{d}g(\vec{x}^0)-$$

$$-o(1)\cdot\frac{f(\vec{x}^0)}{g(\vec{x}^0)}\cdot o(\|\overrightarrow{\Delta x}\|),$$

где, учитывая, что по предыдущей лекции $\boldsymbol{d}f(\vec{x}^0)=O(\|\overrightarrow{\Delta x}\|)$ и $\boldsymbol{d}g(\vec{x}^0)=O(\|\overrightarrow{\Delta x}\|)$, получаем, что шесть последних слагаемых — $o(\|\overrightarrow{\Delta x}\|)$. Значит, $\frac{f}{g}$ дифференцируема в точке $\vec{x}^0$ и $\boldsymbol{d}(\frac{f}{g})(\vec{x}^0)=\frac{g(\vec{x}^0)\boldsymbol{d}f(\vec{x}^0)-f(\vec{x}^0)\boldsymbol{d}g(\vec{x}^0)}{g^2(\vec{x}^0)}$. ▲

Дифференцируемость сложной функции

Теорема 4. *Если функция f дифференцируема в точке $\vec{x}^0\in\mathbb{R}^n$, а функции $x_k(\vec{t})$, $\vec{t}\in\mathbb{R}^m$, $k=1,\dots,n$, дифференцируемы в точке $\vec{t}^0\in\mathbb{R}^m$, причем $x_k(\vec{t}^0)=x_k^0$, $k=1,\dots,n$, то тогда функция $f\left(\vec{x}(\vec{t})\right)=f\left(x_1(\vec{t}),\dots,x_n(\vec{t})\right)$ дифференцируема в точке $\vec{t}^0$ и*

$$\boldsymbol{d}f\left(\vec{x}(\vec{t}^0)\right)=\sum_{i=1}^{n}\frac{\partial f(\vec{x}^0)}{\partial x_i}dx_i(\vec{t}^0)=\sum_{i=1}^{n}\frac{\partial f(\vec{x}^0)}{\partial x_i}\sum_{j=1}^{m}\frac{\partial x_i(\vec{t}^0)}{\partial t_j}\Delta t_j.$$

▼ Из дифференцируемости $f(\vec{x})$ в точке $\vec{x}^0$ следует, что

$$\Delta f=\sum_{i=1}^{n}\frac{\partial f(\vec{x}^0)}{\partial x_i}\Delta x_i+o(\|\overrightarrow{\Delta x}\|)=\sum_{i=1}^{n}\frac{\partial f(\vec{x}^0)}{\partial x_i}\left(\sum_{j=1}^{m}\frac{\partial x_i(\vec{t}^0)}{\partial t_j}\Delta t_j+o(\|\overrightarrow{\Delta t}\|)\right)+o(\|\overrightarrow{\Delta x}\|),$$

и

$$\Delta x_i=\sum_{j=1}^{m}\frac{\partial x_i(\vec{t}^0)}{\partial t_j}\Delta t_j+o(\|\overrightarrow{\Delta t}\|)=O(\|\overrightarrow{\Delta t}\|),$$

а значит, и $\|\overrightarrow{\Delta x}\|=O(\|\overrightarrow{\Delta t}\|)$, откуда следует, что $o(\|\overrightarrow{\Delta x}\|)=o(1)\cdot\frac{\|\overrightarrow{\Delta x}\|}{\|\overrightarrow{\Delta t}\|}\cdot\|\overrightarrow{\Delta t}\|=o(1)\cdot\|\overrightarrow{\Delta t}\|=o(\|\overrightarrow{\Delta t}\|)$. Следовательно,

$$\Delta f=\sum_{i=1}^{n}\frac{\partial f(\vec{x}^0)}{\partial x_i}\sum_{j=1}^{m}\frac{\partial x_i(\vec{t}^0)}{\partial t_j}\Delta t_j+o(\|\overrightarrow{\Delta t}\|),$$

где двойная сумма — линейная функция от $\overrightarrow{\Delta t}$ и, значит, дифференциал функции f в точке $\vec{t}^0$. ▲

Следствие 1. *Если функция f дифференцируема в точке $\vec{x}^0 \in \mathbb{R}^n$, а функции $x_k(t)$, $t \in \mathbb{R}$, $k = 1, \dots, n$, дифференцируемы в точке $t^0 \in \mathbb{R}$, причем $x_k(t^0) = x_k^0$, $k = 1, \dots, n$, то тогда функция одной переменной $f\left(\vec{x}(t)\right) = f\left(x_1(t), \dots, x_n(t)\right)$ дифференцируема в точке t^0 и*

$$\frac{df\left(\vec{x}(t^0)\right)}{dt} = \sum_{i=1}^{n} \frac{\partial f(\vec{x}^0)}{\partial x_i} \frac{dx_i(t^0)}{dt}.$$

▼ Следствие — частный случай предыдущей теоремы. ▲

Следствие 2. *Если функция f дифференцируема в точке $\vec{x}^0 \in \mathbb{R}^n$, а функции $x_i(\vec{t})$, $\vec{t} \in \mathbb{R}^m$, $i = 1, \dots, n$, имеют частные производные $\frac{\partial x_i(\vec{t}^0)}{\partial t_j}$, $i = 1, \dots, n$, в точке $\vec{t}^0 \in \mathbb{R}^m$, причем $x_i(\vec{t}^0) = x_i^0$, $i = 1, \dots, n$, то тогда функция $f\left(\vec{x}(\vec{t})\right) = f\left(x_1(\vec{t}), \dots, x_n(\vec{t})\right)$ имеет в точке $\vec{t}^0$ частную производную $\frac{\partial f\left(\vec{x}(\vec{t}^0)\right)}{\partial t_j}$ и*

$$\frac{\partial f\left(\vec{x}(\vec{t}^0)\right)}{\partial t_j} = \sum_{i=1}^{n} \frac{\partial f(\vec{x}^0)}{\partial x_i} \frac{\partial x_i(\vec{t}^0)}{\partial t_j}.$$

▼ Это следствие вытекает из предыдущего. ▲

Инвариантность первого дифференциала

Следствие 3. *При предположениях последней теоремы вычисление дифференциала функции $f(\vec{x}(\vec{t}))$ прямым способом*

$$\boldsymbol{d}f = \sum_{j=1}^{m} \frac{\partial f(\vec{x}(\vec{t}))}{\partial t_j}\Big|_{\vec{t}^0} \Delta t_j,$$

где $\frac{\partial f(\vec{x}(\vec{t}))}{\partial t_j}\Big|_{\vec{t}^0} = \sum\limits_{i=1}^{n} \frac{\partial f(\vec{x}^0)}{\partial x_i} \frac{\partial x_i(\vec{t}^0)}{\partial t_j}$, или последовательным способом

$$\boldsymbol{d}f = \sum_{i=1}^{n} \frac{\partial f(\vec{x}^0)}{\partial x_i} dx_i = \sum_{i=1}^{n} \frac{\partial f(\vec{x}^0)}{\partial x_i} \left(\sum_{j=1}^{m} \frac{\partial x_i(\vec{t}^0)}{\partial t_j} \Delta t_j \right)$$

(ведь $\boldsymbol{d}x_i = \sum\limits_{j=1}^{m} \frac{\partial x_i(\vec{t}^0)}{\partial t_j} \Delta t_j$, если $\boldsymbol{d}x_i$ рассматривать как дифференциал x_i, а не как приращение x_i) приводит к одинаковому результату.

Приведенное свойство называется **инвариантностью первого дифференциала**.

Лекция 24 (12.05.20)
Производные высших порядков. Теоремы о равенстве смешанных производных

Частные производные второго порядка

Определение 1. Если частная производная $\frac{\partial}{\partial x_i} f(\vec{x})$ существует в окрестности точки $\vec{x}^0$ и для нее как функции существует частная производная по x_j в точке $\vec{x}^0$

$$\frac{\partial}{\partial x_j}\left(\frac{\partial}{\partial x_i} f(\vec{x})\right)\Big|_{\vec{x}^0} = \frac{\partial}{\partial x_j}\left(\frac{\partial}{\partial x_i} f(\vec{x}^0)\right),$$

то эту производную называют **частной производной второго порядка** функции f по переменным x_i и x_j в точке $\vec{x}^0$.

Обозначение 1. Частную производную второго порядка функции f по переменным x_i и x_j в точке $\vec{x}^0$ обозначают

$$\frac{\partial^2}{\partial x_j \partial x_i} f(\vec{x})\Big|_{\vec{x}=\vec{x}^0} \text{ или } \frac{\partial^2 f(\vec{x})}{\partial x_j \partial x_i}\Big|_{\vec{x}=\vec{x}^0} \text{ или } \frac{\partial^2 f(\vec{x}^0)}{\partial x_j \partial x_i} \text{ или } f''_{x_i,x_j}(\vec{x}^0).$$

В двух первых обозначениях вместо “$\vec{x}=\vec{x}^0$” пишут также “$\vec{x}^0$”.

Определение 2. Если x_j и x_i одна и та же переменная, т.е. $j = i$, то частную производную $\frac{\partial^2 f(\vec{x}^0)}{\partial x_i \partial x_i}$ называют **чистой** частной производной и обозначают $\frac{\partial^2 f(\vec{x}^0)}{(\partial x_i)^2}$ или, по традиции опуская скобки, $\frac{\partial^2 f(\vec{x}^0)}{\partial x_i^2}$.

Определение 3. Если x_j и x_i разные переменные, т.е. $j \neq i$, то частную производную $\frac{\partial^2 f(\vec{x}^0)}{\partial x_j \partial x_i}$ называют **смешанной** частной производной.

Частные производные высших порядков

Аналогично частной производной второго порядка определяются частные производные любого натурального порядка.

Определение 4. Если частная производная порядка m ($m \in \mathbb{N}$) функции f по переменным $x_{i_1}, \ldots, x_{i_m}$ $\frac{\partial^m f(\vec{x})}{\partial x_{i_m} \ldots \partial x_{i_1}}$ существует в окрестности точки $\vec{x}^0$ и для нее как функции существует частная производная по переменной $x_{i_{m+1}}$ в точке $\vec{x}^0$

$$\frac{\partial}{\partial x_{i_{m+1}}}\left(\frac{\partial^m f(\vec{x})}{\partial x_{i_m} \ldots \partial x_{i_1}}\right)\Bigg|_{\vec{x}=\vec{x}^0},$$

то эту производную называют **частной производной порядка** $m+1$ функции f по переменным $x_{i_1}, \ldots, x_{i_m}, x_{i_{m+1}}$ в точке $\vec{x}^0$.

Обозначение 2. Частную производную порядка $m+1$ функции f по переменным $x_{i_1}, \ldots, x_{i_m}, x_{i_{m+1}}$ в точке $\vec{x}^0$ обозначают

$$\frac{\partial^{m+1}}{\partial x_{i_{m+1}} \partial x_{i_m} \ldots \partial x_{i_1}} f(\vec{x})\Bigg|_{\vec{x}=\vec{x}^0} \text{ или } \frac{\partial^{m+1} f(\vec{x})}{\partial x_{i_{m+1}} \partial x_{i_m} \ldots \partial x_{i_1}}\Bigg|_{\vec{x}^0} \text{ или}$$
$$\frac{\partial^{m+1} f(\vec{x}^0)}{\partial x_{i_{m+1}} \partial x_{i_m} \ldots \partial x_{i_1}} \text{ или } f^{(m+1)}_{x_{i_1},\ldots,x_{i_m},x_{i_{m+1}}}.$$

Определение 5. Если $x_{i_{m+1}}, x_{i_m}, \ldots, x_{i_1}$ одна и та же переменная, т.е. $i_{m+1} = i_m = \cdots = i_1$, то такую частную производную называют **чистой** и обозначают $\frac{\partial^{m+1} f(\vec{x}^0)}{(\partial x_{i_1})^{m+1}}$ или, по традиции опуская скобки, $\frac{\partial^{m+1} f(\vec{x}^0)}{\partial x_{i_1}^{m+1}}$.

Определение 6. Если среди переменных $x_{i_{m+1}}, x_{i_m}, \ldots, x_{i_1}$ встречаются различные, то частную производную $\frac{\partial^{m+1} f(\vec{x}^0)}{\partial x_{i_{m+1}} \partial x_{i_m} \ldots \partial x_{i_1}}$ называют **смешанной** частной производной.

Теоремы о равенстве смешанных производных

Теорема 1 (Шварца). *Пусть функция двух переменных $f(x,y)$ имеет в точке (x_0,y_0) непрерывные смешанные производные $\frac{\partial^2 f}{\partial y\partial x}$ и $\frac{\partial^2 f}{\partial x\partial y}$. Тогда они равны,*

$$\frac{\partial^2}{\partial x\partial y}f(x_0,y_0)=\frac{\partial^2}{\partial y\partial x}f(x_0,y_0).$$

▼ Будем брать Δx и Δy настолько малыми, что точка $(x_0+\Delta x, y_0+\Delta y)$ лежит в окрестности точки (x_0,y_0), где существуют первые частные производные и смешанные производные второго порядка. Рассмотрим выражение

$$\Delta^2 f = f(x_0+\Delta x, y_0+\Delta y) - -f(x_0, y_0+\Delta y) - f(x_0+\Delta x, y_0) + f(x_0,y_0),$$

называемое второй разностью. Его можно рассматривать как разность по переменной x от разности по переменной y,

$$\begin{gathered}\Delta^2 f = [f(x_0+\Delta x, y_0+\Delta y) - f(x_0+\Delta x, y_0)] - [f(x_0, y_0+\Delta y) - f(x_0,y_0)] = \\ = \frac{\partial}{\partial x}\left[f(x_0+\theta\Delta x, y_0+\Delta y) - f(x_0+\theta\Delta x, y_0)\right]\Delta x = \\ = \left[\frac{\partial}{\partial x}f(x_0+\theta\Delta x, y_0+\Delta y) - \frac{\partial}{\partial x}f(x_0+\theta\Delta x, y_0)\right]\Delta x = \Delta_y f'_x(x_0+\theta\Delta x, y)\Delta x,\end{gathered}$$

где $0<\theta<1$ по теореме Лагранжа. Еще раз применяя теорему Лагранжа, имеем

$$\begin{gathered}\Delta^2 f = \Delta_y f'_x(x_0+\theta\Delta x, y)\Delta x = \left[\frac{\partial}{\partial x}f(x_0+\theta\Delta x, y_0+\Delta y) - \frac{\partial}{\partial x}f(x_0+\theta\Delta x, y_0)\right]\Delta x = \\ = \frac{\partial}{\partial y}\left(\frac{\partial}{\partial x}f(x_0+\theta\Delta x, y_0+\vartheta\Delta y)\right)\cdot\Delta x\cdot\Delta y = \frac{\partial^2}{\partial y\partial x}f(x_0+\theta\Delta x, y_0+\vartheta\Delta y)\cdot\Delta x\cdot\Delta y,\end{gathered}$$

где $0<\vartheta<1$, $0<\theta<1$.

В силу непрерывности смешанной производной $\frac{\partial^2 f}{\partial y\partial x}$ в точке (x_0,y_0) получаем, что при $\Delta x\to 0$, $\Delta y\to 0$ так, что ни одно из приращений не обращается в 0,

$$\frac{\Delta^2 f}{\Delta y\,\Delta x}\xrightarrow[\Delta x\to 0\Delta y\to 0]{}\frac{\partial^2 f}{\partial y\,\partial x}f(x_0,y_0).$$

Аналогично, при $\Delta x\to 0$, $\Delta y\to 0$ так, что ни одно из приращений не обращается в 0,

$$\frac{\Delta^2 f}{\Delta y\,\Delta x}\xrightarrow[\Delta x\to 0\Delta y\to 0]{}\frac{\partial^2 f}{\partial x\,\partial y}f(x_0,y_0),$$

что и доказывает теорему. ▲

В теореме Шварца можно заменить требование существования и непрерывности смешанной производной $\frac{\partial^2 f}{\partial x\partial y}$ (или $\frac{\partial^2 f}{\partial y\partial x}$) на требование существования непрерывной частной производной первого порядка $\frac{\partial f}{\partial y}$ (соответственно,$\frac{\partial f}{\partial y}$) в окрестности точки (x_0,y_0).

Следствие 4. *Если функция f (из $\mathbb{R}^n$ в $\mathbb{R}$) имеет в точке $\vec{x}^0$ непрерывные смешанные частные производные $\frac{\partial^2}{\partial x_j\partial x_i}f$ и $\frac{\partial^2}{\partial x_i\partial x_j}f$, то эти частные производные равны.*

▼ Действительно, полагая $x_k = x_k^0$ при $k \neq i, j$, т.е. фиксируя все переменные, кроме i-ой и j-ой, сводим данное утверждение к утверждению теоремы Шварца. ▲

Следствие 1. *Если у функции f (из $\mathbb{R}^n$ в $\mathbb{R}$) любые смешанные частные производные порядка m непрерывны в точке $\vec{x}^0$, а производные порядков меньше m непрерывны в ее окрестности, то смешанные частные производные до порядка m включительно по одинаковым наборам переменных (с учетом их кратности, но без учета их порядка) равны.*

▼ Действительно, при $m = 2$ это предыдущее следствие. А если утверждение верно при $m = r$, то оно верно и при $m = r + 1$. Ведь по предположению (индукционному) все смешанные частные производные до порядка r включительно по одинаковым наборам переменных равны в некоторой окрестности точки $\vec{x}^0$ (т.е. первые r (или менее) дифференцирований можно совершать в некоторой окрестности точки $\vec{x}^0$ в любом порядке), а по следствию 1 можно менять порядок r-ого и $r+1$-ого дифференцирования в точке $\vec{x}^0$. Значит, все $r + 1$ дифференцирований можно совершать в точке $\vec{x}^0$ в любом порядке — результат не меняется. ▲

Теорема 2 (Юнга). *Если функция двух переменных $f(x, y)$ дважды дифференцируема в точке (x_0, y_0) (т.е. дифференцируемы частные производные первого порядка), то в этой точке смешанные производные второго порядка равны,*

$$\frac{\partial^2}{\partial x \partial y} f(x_0, y_0) = \frac{\partial^2}{\partial y \partial x} f(x_0, y_0).$$

▼ Как и в доказательстве теоремы Шварца рассмотрим выражение

$$\Delta^2 f = [f(x_0 + \Delta x, y_0 + \Delta y) - f(x_0 + \Delta x, y_0)] - [f(x_0, y_0 + \Delta y) - f(x_0, y_0)] =$$
$$= \frac{\partial}{\partial x} [f(x_0 + \theta\Delta x, y_0 + \Delta y) - f(x_0 + \theta\Delta x, y_0)] \cdot \Delta x =$$
$$= \left[\frac{\partial}{\partial x} f(x_0 + \theta\Delta x, y_0 + \Delta y) - \frac{\partial}{\partial x} f(x_0 + \theta\Delta x, y_0) \right] \Delta x = \Delta_y f'_x(x_0 + \theta\Delta x, y)\Delta x,$$

где $0 < \theta < 1$ (по теореме Лагранжа). В силу дифференцируемости частной производной f'_x

$$\Delta^2 f = \Delta_y f'_x(x_0 + \theta\Delta x, y)\Delta x = [f'_x(x_0 + \theta\Delta x, y_0 + \Delta y) - f'_x(x_0 + \theta\Delta x, y_0)] \Delta x =$$
$$= f'_x(x_0, y_0) \cdot \Delta x + f''_{x,x} \cdot \theta(\Delta x)^2 + f''_{x,y}(x_0, y_0) \cdot \Delta x \cdot \Delta y + o(\|(\theta\Delta x, \Delta y)\|) \cdot \Delta x -$$
$$- f'_x(x_0, y_0) \cdot \Delta x - f''_{x,x} \cdot \theta(\Delta x)^2 - o(|\theta\Delta x|) \cdot \Delta x = f''_{x,y}(x_0, y_0) \cdot \Delta x \cdot \Delta y +$$
$$+ o(\|(\Delta x, \Delta y)\|) \cdot \Delta x$$

и, значит,

$$\lim_{\Delta y = \Delta x \to 0} \frac{\Delta^2 f}{\Delta y \cdot \Delta x} = f''_{x,y}(x_0, y_0).$$

Аналогично,

$$\lim_{\Delta y = \Delta x \to 0} \frac{\Delta^2 f}{\Delta x \cdot \Delta y} = f''_{y,x}(x_0, y_0),$$

что и доказывает теорему. ▲

Следствие 2. *Если функция f из $\mathbb{R}^n$ в $\mathbb{R}$ дважды дифференцируема в точке $\vec{x}^0$, то для любых переменных x_i и x_j*

$$\frac{\partial^2}{\partial x_i \partial x_j} f(\vec{x}^0) = \frac{\partial^2}{\partial x_j \partial x_i} f(\vec{x}^0).$$

▼ Действительно, полагая $x_k = x_k^0$ при $k \neq i, j$, т.е. фиксируя все переменные, кроме i-ой и j-ой, сводим данное утверждение к утверждению теоремы. ▲

Следствие 3. *Если функция f из $\mathbb{R}^n$ в $\mathbb{R}$ m раз дифференцируема в точке $\vec{x}^0$, то смешанные производные до порядка m включительно по одинаковым наборам переменных (с учетом их кратности, но без учета их порядка) равны.*

▼ Действительно, при $m = 2$ это предыдущее следствие. А если утверждение верно при $m = r$, то оно верно и при $m = r + 1$. Ведь все частные производные до порядка r включительно непрерывны в некоторой окрестности точки $\vec{x}^0$ и, значит, совпадают при дифференцировании по одинаковым наборам переменных, а по предыдущему следствию можно менять порядок r-ого и $r + 1$-ого дифференцирования. Значит, все $r + 1$ дифференцирований можно совершать в любом порядке — результат получается одинаковый. ▲

Лекция 25 (15.05.20)
 Дифференциалы высших порядков.
 Формула Тейлора и локальные экстремумы

Кратная дифференцируемость функции

Зная определение дифференцируемости функции в точке дадим теперь по индукции определение кратной дифференцируемости функции в точке.

Определение 1. Функция f называется $m + 1$ **раз дифференцируемой в точке** $\vec{x}^0$, где $m \in \mathbb{N}$, если сама функция и все ее частные производные порядка k, $1 \leqslant k < m$, дифференцируемы в некоторой окрестности точки $\vec{x}^0$, а все ее частные производные порядка m дифференцируемы в точке $\vec{x}^0$.

Утверждение. Из теоремы о достаточном условии дифференцируемости следует, что для того, чтобы функция f была m раз дифференцируемой в точке $\vec{x}^0$ достаточно, чтобы все ее частные производные порядка k, $1 \leqslant k < m$, существовали и были непрерывны в некоторой окрестности точки $\vec{x}^0$, а все частные производные порядка m существовали и были непрерывны в точке $\vec{x}^0$.

Определение 2. Функция f называется m **раз непрерывно дифференцируемой** в точке $\vec{x}^0$, если f дифференцируема m раз в некоторой окрестности точки $\vec{x}^0$ и все ее частные производные порядка m непрерывны в точке $\vec{x}^0$.

Дифференциал второго порядка

Определение 3. Если первый дифференциал $\boldsymbol{d}f(\vec{x}, \overrightarrow{\Delta x})$ как функция от $\vec{x}$ при любом фиксированном $\overrightarrow{\Delta x}$ дифференцируем в точке $\vec{x}^0$, то выражение, являющееся дифференциалом от первого дифференциала при таком же приращении $\overrightarrow{\Delta x}$, что и фиксированное в первом дифференциале, называется **вторым дифференциалом**. Если f дважды дифференцируема в точке $\vec{x}^0$, то второй дифференциал

$$\boldsymbol{d}^2 f(\overrightarrow{\Delta x})\big|_{\vec{x}^0} = \boldsymbol{d}^2 f(\vec{x}^0, \overrightarrow{\Delta x}) = \sum_{j=1}^{n} \boldsymbol{d}\left(\frac{\partial f(\vec{x}^0)}{\partial x_j}\right)\big|_{\vec{x}^0} \Delta x_j =$$
$$= \sum_{j=1}^{n}\sum_{i=1}^{n} \frac{\partial^2 f(\vec{x}^0)}{\partial x_i \partial x_j} \Delta x_j \Delta x_i.$$

Вместо Δx_i и Δx_j часто используют обозначения dx_i и dx_j.

Замечание 1. Иногда под вторым дифференциалом понимается не квадратичная, а билинейная форма переменных $\overrightarrow{\Delta_1 x}$ и $\overrightarrow{\Delta_2 x}$

$$\boldsymbol{d}^2 f(\overrightarrow{\Delta_1 x}, \overrightarrow{\Delta_2 x})\big|_{\vec{x}^0} = \boldsymbol{d}^2 f(\vec{x}^0, \overrightarrow{\Delta_1 x}, \overrightarrow{\Delta_2 x}) =$$
$$= \sum_{j=1}^{n}\sum_{i=1}^{n} \frac{\partial^2 f(\vec{x}^0)}{\partial x_i \partial x_j} \Delta_1 x_j \Delta_2 x_i.$$

Но мы таким понятием второго дифференциала пользоваться не будем.

Дифференциалы высших порядков

Аналогично дифференциалу второго порядка определяются дифференциалы любого порядка.

Определение 4. Если дифференциал m-ого порядка $\boldsymbol{d}^m f(\vec{x}, \overrightarrow{\Delta x})$ как функция от $\vec{x}$ при любом фиксированном $\overrightarrow{\Delta x}$ дифференцируем в точке $\vec{x}^0$, то выражение, являющееся дифференциалом от m-ого дифференциала, взятое при таком же приращении $\overrightarrow{\Delta x}$, что и зафиксированное в m-ом дифференциале, называется **$m+1$-ым дифференциалом** функции f в точке $\vec{x}^0$. Если функция f $m+1$ раз дифференцируема в точке $\vec{x}^0$, то $m+1$-ый дифференциал

$$\boldsymbol{d}^{m+1} f(\overrightarrow{\Delta x})\big|_{\vec{x}^0} = \boldsymbol{d}^{m+1} f(\vec{x}^0, \overrightarrow{\Delta x}) =$$
$$= \sum_{i_{m+1}=1}^{n}\sum_{i_m=1}^{n}\cdots\sum_{i_1=1}^{n} \frac{\partial^{m+1} f(\vec{x}^0)}{\partial x_{i_{m+1}} \partial x_{i_m} \dots \partial x_{i_1}} \Delta x_{i_1} \dots \Delta x_{i_{m+1}}.$$

Замечание 2. Иногда под $m+1$-ым дифференциалом понимается не квадратичная, а полилинейная форма переменных $\overrightarrow{\Delta_1 x}$, $\overrightarrow{\Delta_2 x}, \dots, \overrightarrow{\Delta_{m+1} x}$

$$\boldsymbol{d}^{m+1} f\left(\overrightarrow{\Delta_1 x}, \overrightarrow{\Delta_2 x}, \dots, \overrightarrow{\Delta_{m+1} x}\right)\big|_{\vec{x}^0} =$$
$$= \sum_{i_{m+1}=1}^{n}\sum_{i_m=1}^{n}\cdots\sum_{i_1=1}^{n} \frac{\partial^{m+1} f(\vec{x}^0)}{\partial x_{i_{m+1}} \partial x_{i_m} \dots \partial x_{i_1}} \Delta_1 x_{i_1} \dots \Delta_{m+1} x_{i_{m+1}}.$$

Но мы таким понятием высших дифференциалов пользоваться не будем.

Формула Тейлора

Определение 5. Для функции f из $\mathbb{R}^n$ в $\mathbb{R}$ дифференцируемой в точке $\vec{x}^0$ не менее m раз равенство

$$f(\vec{x}) = f(\vec{x}^0 + \overrightarrow{\Delta x}) = \sum_{k=0}^{m} \frac{1}{k!} \boldsymbol{d}^k f(\overrightarrow{\Delta x})\big|_{\vec{x}^0} + r_m(\overrightarrow{\Delta x})\big|_{\vec{x}^0},$$

где $d^0 f(\overrightarrow{\Delta x})\big|_{\vec{x}^0} = f(\vec{x}^0)$, называется **формулой Тейлора** в точке $\vec{x}^0$ функции f; $r_m(\overrightarrow{\Delta x})\big|_{\vec{x}^0}$ — **остаточный член** формулы Тейлора. В случае $\vec{x}^0 = \vec{0}$ написанное равенство иногда называют **формулой Маклорена.**

Теорема 1 (остаточный член в форме Лагранжа). *Если функция f из $\mathbb{R}^n$ в $\mathbb{R}$ m раз непрерывно дифференцируема на отрезке $[\vec{x}^0, \vec{x}]$ и $m+1$ раз дифференцируема на интервале $(\vec{x}^0, \vec{x})$ (получающимся из отрезка отбрасыванием его концов), то найдется такое θ, $0 < \theta < 1$, что*

$$f(\vec{x}) = f(\vec{x}^0 + \overrightarrow{\Delta x}) = \sum_{k=0}^{m} \frac{1}{k!} \boldsymbol{d}^k f(\overrightarrow{\Delta x})\big|_{\vec{x}^0} + \frac{1}{(m+1)!} \boldsymbol{d}^{m+1} f(\overrightarrow{\Delta x})\big|_{\vec{x}^0 + \theta\overrightarrow{\Delta x}}$$

(т.е. $r_m(\vec{x}) = \frac{1}{(m+1)!} \boldsymbol{d}^{m+1} f(\overrightarrow{\Delta x})\big|_{\vec{x}^0 + \theta\overrightarrow{\Delta x}}$).

▼ Рассмотрим при фиксированном $\overrightarrow{\Delta x} = \vec{x} - \vec{x}^0$ функцию одного переменного t, $0 \leqslant t \leqslant 1$, $g(t) = f(\vec{x}^0 + t\overrightarrow{\Delta x})$. Тогда

$$\begin{aligned} g(0) =& f(\vec{x}^0), \\ g'(0) =& \boldsymbol{d} f(\overrightarrow{\Delta x})\big|_{\vec{x}^0} = \sum_{i=1}^{n} \frac{\partial f(\vec{x}^0)}{\partial x_i} \overrightarrow{\Delta x_i} \end{aligned}$$

(по следствию теоремы о дифференцировании сложной функции)

$$\begin{aligned} g''(0) =& \boldsymbol{d}^2 f(\overrightarrow{\Delta x})\big|_{\vec{x}^0} = \sum_{j=1}^{n} \sum_{i=1}^{n} \frac{\partial^2 f(\vec{x}^0)}{\partial x_j\, \partial x_i} \Delta x_i \Delta x_j, \\ \vdots \quad & \quad \vdots \\ g^{(m)}(0) =& \boldsymbol{d}^m f(\overrightarrow{\Delta x})\big|_{\vec{x}^0}, \\ g^{(m+1)}(t) =& \boldsymbol{d}^{m+1} f(\overrightarrow{\Delta x})\big|_{\vec{x}^0 + t\overrightarrow{\Delta x}}. \end{aligned}$$

По формуле Тейлора с остаточным членом в форме Лагранжа

$$f(\vec{x}^0 + t\overrightarrow{\Delta x}) = \sum_{k=1}^{m} \frac{1}{k!} g^{(k)}(0) t^k + \frac{1}{(m+1)!} g^{(m+1)}(\theta t) t^{m+1} =$$

$$= \sum_{k=1}^{m} \frac{1}{k!} \boldsymbol{d}^k f(\overrightarrow{\Delta x})\big|_{\vec{x}^0} \cdot t^k + \frac{1}{(m+1)!} \boldsymbol{d}^{m+1} f(\overrightarrow{\Delta x})\big|_{\vec{x}^0 + \theta t\overrightarrow{\Delta x}} \cdot t^{m+1}.$$

Теперь, полагая $t = 1$, получаем искомую формулу. ▲

Следствие 1 (остаточный член в форме Пеано). *Если функция f из $\mathbb{R}^n$ в $\mathbb{R}$ m раз непрерывно дифференцируема в точке $\vec{x}^0$, то*

$$f(\vec{x}) = f(\vec{x}^0 + \overrightarrow{\Delta x}) = \sum_{k=0}^{m} \frac{1}{k!} \boldsymbol{d}^k f(\overrightarrow{\Delta x})\big|_{\vec{x}^0} + o(\|\overrightarrow{\Delta x}\|^m)$$

(т.е. $r_m(\vec{x}) = o(\|\overrightarrow{\Delta x}\|^m)$*).*

▼ При достаточно малом $\overrightarrow{\Delta x}$ по теореме 1

$$f(\vec{x}) = f(\vec{x}^{\,0} + \overrightarrow{\Delta x}) = \sum_{k=0}^{m-1} \frac{1}{k!}\boldsymbol{d}^k f(\overrightarrow{\Delta x})\big|_{\vec{x}^{\,0}} + \frac{1}{m!}\boldsymbol{d}^m f(\overrightarrow{\Delta x})\big|_{\vec{x}^{\,0}+\theta\overrightarrow{\Delta x}}.$$

Дифференциал

$$\boldsymbol{d}^m f(\overrightarrow{\Delta x})\big|_{\vec{x}^{\,0}+\theta\overrightarrow{\Delta x}}$$

является суммой чденов

$$\frac{\partial^m f(\vec{x}^{\,0} + \theta\overrightarrow{\Delta x})}{\partial x_{i_m}\dots\partial x_{i_1}}\Delta x_{i_1}\dots\Delta x_{i_m}.$$

Так как f m раз непрерывно дифференцируема в точке $\vec{x}^{\,0}$, то

$$\frac{\partial^m f(\vec{x}^{\,0} + \theta\overrightarrow{\Delta x})}{\partial x_{i_m}\dots\partial x_{i_1}} = \frac{\partial^m f(\vec{x}^{\,0})}{\partial x_{i_m}\dots\partial x_{i_1}} + o(1).$$

Значит,

$$\boldsymbol{d}^m f(\overrightarrow{\Delta x})\big|_{\vec{x}^{\,0}+\theta\overrightarrow{\Delta x}} = \boldsymbol{d}^m f(\overrightarrow{\Delta x})\big|_{\vec{x}^{\,0}} + o(\|\overrightarrow{\Delta x}\|^m)$$

и верность следствия установлена. ▲

Замечание 1. Формула Тейлора с остаточным членом в форме Пеано верна просто при дифференцируемости функции f m раз в точке $\vec{x}^{\,0}$, но доказать это сложнее.

Локальные экстремумы функций многих переменных

Определение 6. Функция f из $\mathbb{R}^n$ в $\mathbb{R}$ имеет в точке $\vec{x}^{\,0}$ **локальный максимум** (**локальный минимум**), если f определена в некоторой окрестности точки $\vec{x}^{\,0}$ и

$$\exists B'_\delta(\vec{x}^{\,0})\ \forall \vec{x} \in B'_\delta(\vec{x}^{\,0}):\ f(\vec{x}) \leqslant f(\vec{x}^{\,0}) \quad \big(f(\vec{x}) \geqslant f(\vec{x}^{\,0})\big)\,.$$

При этом говорят, что $\vec{x}^{\,0}$ — **точка локального максимума** (**локального минимума**), а величина $f(\vec{x}^{\,0})$ — **локальный максимум** (**локальный минимум**) функции f.

Если в приведенном определении заменить нестрогое неравенство строгим, то получится определение **строгого локального максимума** (**строгого локального минимума**). При этом говорят, что $\vec{x}^{\,0}$ — **точка строгого локального максимума** (**строгого локального минимума**), а величина $f(\vec{x}^{\,0})$ — **строгий локальный максимум** (**строгий локальный минимум**) функции f.

Определение 7. Термин **локальный экстремум** означает локальный максимум или локальный минимум, а термин **строгий локальный экстремум** — строгий локальный максимум или строгий локальный минимум.

Условия локальных экстремумов

Теорема 3 (необходимое условие локального экстремума). *Если функция f из $\mathbb{R}^n$ в $\mathbb{R}$ имеет в точке $\vec{x}^0$ локальный экстремум, то все частные производные $\frac{\partial f(\vec{x}^0)}{\partial x_k}$, которые существуют, равны нулю:* $\frac{\partial f(\vec{x}^0)}{\partial x_k} = 0$.

▼ Полагая $x_i = x_i^0$ при $i \neq k$, т.е. фиксируя все переменные, кроме k-ой, сводим утверждение теоремы к теореме Ферма утверждающей, что если функция одного переменного дифференцируема в точке локального экстремума, то ее производная в этой точке равна нулю. ▼

Следствие 2. *Если функция f из $\mathbb{R}^n$ в $\mathbb{R}$ имеет в точке $\vec{x}^0$ локальный экстремум и дифференцируема в ней, то для любого $\overrightarrow{\Delta x}$ имеем $\boldsymbol{d}f(\overrightarrow{\Delta x})\big|_{\vec{x}^0} = 0$ и $grad\, f(\vec{x}^0) = \vec{0}$.*

Определение 8. Точка $\vec{x} \in \mathbb{R}^n$, в которой обращаются в нуль все частные производные функции f, называется **стационарной точкой** функции f.

Лекция 26 (19.05.20)
Условия локального экстремума. Неявные функции

Для дифференцируемых функций поиск точек локальных экстремумов обычно сводится к определению стационарных точек и исследованию, какие из них являются точками локальных экстремумов.

Теорема 1 (об условиях локального экстремума). *Пусть функция f из $\mathbb{R}^n$ в $\mathbb{R}$ m раз непрерывно дифференцируема в точке $\vec{x}^0$ и $\boldsymbol{d}^m f(\overrightarrow{\Delta x})\big|_{\vec{x}^0}$ первый отличный от нуля дифференциал функции f в точке $\vec{x}^0$.*

Если $\boldsymbol{d}^m f(\overrightarrow{\Delta x})\big|_{\vec{x}^0}$ положительно определен (т.е. $\boldsymbol{d}^m f(\overrightarrow{\Delta x})\big|_{\vec{x}^0} > 0$ при $\overrightarrow{\Delta x} \neq 0$), то f имеет в точке $\vec{x}^0$ строгий локальный минимум; если f имеет в $\vec{x}^0$ локальный минимум, то $\boldsymbol{d}^m f(\overrightarrow{\Delta x})\big|_{\vec{x}^0}$ определен неотрицательно (т.е. $\boldsymbol{d}^m f(\overrightarrow{\Delta x})\big|_{\vec{x}^0} \geqslant 0$ при любом $\overrightarrow{\Delta x}$).

Если $\boldsymbol{d}^m f(\overrightarrow{\Delta x})\big|_{\vec{x}^0}$ отрицательно определен (т.е. $\boldsymbol{d}^m f(\overrightarrow{\Delta x})\big|_{\vec{x}^0} < 0$ при $\overrightarrow{\Delta x} \neq 0$), то f имеет в точке $\vec{x}^0$ строгий локальный максимум; если f имеет в $\vec{x}^0$ локальный максимум, то $\boldsymbol{d}^m f(\overrightarrow{\Delta x})\big|_{\vec{x}^0}$ определен неположительно (т.е. $\boldsymbol{d}^m f(\overrightarrow{\Delta x})\big|_{\vec{x}^0} \leqslant 0$ при любом $\overrightarrow{\Delta x}$).

Если $\boldsymbol{d}^m f(\overrightarrow{\Delta x})\big|_{\vec{x}^0}$ принимает значения разных знаков (как строго больше нуля, так и строго меньше), то f не имеет в точке $\vec{x}^0$ локального экстремума.

▼ Если $\boldsymbol{d}^m f(\overrightarrow{\Delta x})\big|_{\vec{x}^0} > 0$ при $\overrightarrow{\Delta x} \neq 0$), то, так как единичная сфера в $\mathbb{R}^n$ компакт, значение

$$\eta = \min_{\|\overrightarrow{\Delta x}\|=1} \frac{1}{m!}\boldsymbol{d}^m f(\overrightarrow{\Delta x})\big|_{\vec{x}^0}$$

достигается в некоторой точке сферы (по 2-ой теореме Вейерштрасса) и $\eta > 0$. По формуле Тейлора с остаточным членом в форме Пеано

$$\Delta f = f(\vec{x}^0 + \overrightarrow{\Delta x}) - f(\vec{x}^0) = \frac{1}{m!}\boldsymbol{d}^m f(\overrightarrow{\Delta x})\big|_{\vec{x}^0} + o(1) \cdot \|\overrightarrow{\Delta x}\|^m,$$

а

$$\boldsymbol{d}^m f(\overrightarrow{\Delta x})\big|_{\vec{x}^0} = \boldsymbol{d}^m f\left(\frac{\overrightarrow{\Delta x}}{\|\overrightarrow{\Delta x}\|}\right)\big|_{\vec{x}^0} \cdot \left\|\overrightarrow{\Delta x}\right\|^m$$

(это очевидно для каждого из членов $\frac{\partial^m f(\vec{x}^{\,0})}{\partial x_{i_m}\ldots\partial x_{i_1}} \cdot \Delta x_{i_1} \ldots \Delta x_{i_m}$ дифференциала), поэтому

$$\Delta f \geqslant (\eta + o(1)) \cdot \|\overrightarrow{\Delta x}\|^m.$$

Найдем такую проколотую δ-окрестность $B'_\delta(\vec{x}^{\,0})$, что на ней $|o(1)| < \eta$. Тогда на $B'_\delta(\vec{x}^{\,0})$

$$\Delta f \geqslant (\eta + o(1)) \cdot \|\overrightarrow{\Delta x}\|^m > 0$$

и, значит, в точке $\vec{x}^{\,0}$ строгий локальный минимум.

Если в точке $\vec{x}^{\,0}$ локальный минимум функции f, то в некоторой окрестности $\vec{x}^{\,0}$

$$\Delta f = \frac{1}{m!}\boldsymbol{d}^m f(\overrightarrow{\Delta x})\big|_{\vec{x}^{\,0}} + o(1) \cdot \|\overrightarrow{\Delta x}\|^m \geqslant 0.$$

Значит,

$$\lim_{\alpha\to+0} \alpha^{-m} \left(\frac{1}{m!}\boldsymbol{d}^m f(\alpha\overrightarrow{\Delta x})\big|_{\vec{x}^{\,0}} + o(1) \cdot \|\alpha\overrightarrow{\Delta x}\|^m \right) =$$
$$= \frac{1}{m!}\boldsymbol{d}^m f(\overrightarrow{\Delta x})\big|_{\vec{x}^{\,0}} + \lim_{\alpha\to+0} o(1) \cdot \|\overrightarrow{\Delta x}\|^m = \frac{1}{m!}\boldsymbol{d}^m f(\overrightarrow{\Delta x})\big|_{\vec{x}^{\,0}} \geqslant 0.$$

для любого $\overrightarrow{\Delta x}$, т.е. дифференциал $\boldsymbol{d}^m f(\overrightarrow{\Delta x})\big|_{\vec{x}^{\,0}}$ определен неотрицательно.

Другие случаи, когда дифференциал $\boldsymbol{d}^m f(\overrightarrow{\Delta x})\big|_{\vec{x}^{\,0}}$ отрицательно определен или когда $\vec{x}^{\,0}$ — точка локального максимума функции f, рассматриваются аналогично. Они могут быть также сведены к уже изученным случаям рассмотрением функции $-f$.

И, наконец, если дифференциал $\boldsymbol{d}^m f(\overrightarrow{\Delta x})\big|_{\vec{x}^{\,0}}$ принимает значения разных знаков, то, в соответствии с предыдущими пунктами теоремы, не выполняется необходимое условие как локального минимума, так и локального максимума. ▲

Замечание. Отметим, что дифференциал нечетного порядка или нулевой или принимает значения разных знаков, так как $\boldsymbol{d}^m f(\overrightarrow{\Delta x}) = -\boldsymbol{d}^m f(-\overrightarrow{\Delta x})$ при нечетном m.

На практике наиболее часто встречается случай $m = 2$. В этом случае определение характера стационарной точки приводит к изучению квадратичной формы

$$\boldsymbol{d}^m f(\overrightarrow{\Delta x})\big|_{\vec{x}^{\,0}} = \sum_{j=1}^{n}\sum_{i=1}^{n} \frac{\partial^2 f(\vec{x}^{\,0})}{\partial x_j \partial x_i} \Delta x_i\, \Delta x_j.$$

Для анализа квадратичной формы применяется известный в алгебре критерий Сильверста.

Утверждение. *Квадратичная форма* $\sum\limits_{j=1}^{n}\sum\limits_{i=1}^{n} a_{j,i} t_i t_j$ *с симметричной матрицей* $(a_{j,i})_{j,i=1}^n$ *(т.е.* $a_{j,i} = a_{i,j}$*) положительно (отрицательно) определена тогда и только тогда, когда все главные миноры матрицы*

$$|a_{j,i}|_{j,i=1}^k > 0,\ k = 1, \ldots, n,$$
$$\left((-1)^k\, |a_{j,i}|_{j,i=1}^k > 0,\ k = 1, \ldots, n, \right).$$

Теорема о неявной функции

Определение 1. Если переменная y, являющаяся функцией аргументов $x_1, x_2, \dots, x_n$, задается посредством функционального уравнения

$$F(y, x_1, \dots, x_n) = 0,$$

то говорят, что y как функция $x_1, x_2, \dots, x_n$ **задана неявно** или что y — **неявная функция**.

Естественно, возникают вопросы: при каких условиях функциональное уравнение $F(y, x_1, \dots, x_n) = F(y, \vec{x}) = 0$, $\vec{x} \in \mathbb{R}^n$, однозначно определяет функцию $y(\vec{x}) = y(x_1, \dots, x_n)$; при каких условиях $y(\vec{x})$ непрерывна и при каких дифференцируема. Ответы (окончательности которых не будем касаться) содержаться в следующей теореме.

Теорема 2 (о неявной функции). *Пусть $F(y, \vec{x})$ функция $1+n$ переменного y, $\vec{x} = (x_1, \dots, x_n)$, непрерывна в некоторой окрестности точки $(y^0, \vec{x}^0) \in \mathbb{R}^{1+n}$ и имеет в этой окрестности непрерывную частную производную F'_y. Тогда если $F(y^0, \vec{x}^0) = 0$, а $F'_y(y^0, \vec{x}^0) \neq 0$, то*

$$\exists \varepsilon_0 > 0 \; \forall \varepsilon \in (0, \varepsilon_0) \; \exists B_\delta(\vec{x}^0) \subset \mathbb{R}^n \; \forall x \in B_\delta(\vec{x}^0) \; \exists! \, y \in B_\varepsilon(y^0) : \; F(y, \vec{x}) = 0,$$

т.е. для любого достаточно малого числа $\varepsilon > 0$ найдется такая окрестность $B_\delta(\vec{x}^0)$ точки $\vec{x}^0$, что в пределах этой окрестности существует единственная функция $y(\vec{x})$, удовлетворяющая условию $y(\vec{x}) \in B_\varepsilon(y^0)$ и являющаяся решением уравнения $F(y, \vec{x}) = 0$, причем эта функция $y(\vec{x})$ непрерывна в некоторой окрестности точки $\vec{x}^0$.

Если дополнительно потребовать дифференцируемость $F(y, \vec{x})$ в точке $(y^0, \vec{x}^0)$ (в окрестности точки $(y^0, \vec{x}^0)$), то функция $y(\vec{x})$ будет дифференцируема в точке $\vec{x}^0$ (в некоторой окрестности точки $\vec{x}^0$) и при этом

$$\boldsymbol{d}y(\overrightarrow{\Delta x})\big|_{\vec{x}} = \sum_{k=1}^{n} -\frac{F'_{x_k}(y, \vec{x})}{F'_y(y, \vec{x})} \Delta x_k,$$

где $y = y(\vec{x})$, в точке $\vec{x} = \vec{x}^0$ (в некоторой окрестности точки $\vec{x}^0$).

▼ Выберем число $\varepsilon_0 > 0$ таким, что в $2\varepsilon_0$-окрестности точки $(y^0, \vec{x}^0)$ функция $F(y, \vec{x})$ непрерывна, а частная производная $F'_y(y, \vec{x})$ непрерывна и сохраняет знак (строго положительна, если $F'_y(y^0, \vec{x}^0) > 0$; строго отрицательна, если $F'_y(y^0, \vec{x}^0) > 0$). Так как $F(y, \vec{x}^0)$ как функция одной переменной y имеет производную одного знака на отрезке $[-\varepsilon_0, \varepsilon_0]$, то $F(y, \vec{x}^0)$ — строго монотонная функция от y на $[-\varepsilon_0, \varepsilon_0]$, $F(y^0, \vec{x}^0) = 0$, значит, $F(y^0 + \varepsilon, \vec{x}^0) \cdot F(y^0 - \varepsilon, \vec{x}^0) < 0$ для любого $\varepsilon \in (0, \varepsilon_0)$. Пользуясь непрерывностью $F(y, \vec{x}^0)$ найдем такую окрестность точки $\vec{x}^0$ $B_\delta(\vec{x}^0)$, $0 < \delta < \varepsilon_0$, что для $\vec{x} \in B_\delta(\vec{x}^0)$

$$F(y^0 + \varepsilon, \vec{x}) \cdot F(y^0 - \varepsilon, \vec{x}) < 0.$$

При этом

$$\overline{B}_\varepsilon(y^0) \times B_\delta(\vec{x}^0) = \{(y, \vec{x}) : \; y \in \overline{B}_\varepsilon(y^0), \; \vec{x} \in B_\delta(\vec{x}^0)\} \subset$$
$$\subset \overline{B}_{\varepsilon_0}(y^0) \times B_{\varepsilon_0}(\vec{x}^0) = \{(y, \vec{x}) : \; y \in \overline{B}_{\varepsilon_0}(y^0), \; \vec{x} \in B_{\varepsilon_0}(\vec{x}^0)\} \subset B_{2\varepsilon_0}\big((y^0, \vec{x}^0)\big)$$

(в силу неравенства треугольника). Теперь по теореме Больцано–Коши о промежуточном значении и в силу строгой монотонности $F(y,\vec{x})$ как функции одной переменной y при $\vec{x}\in B_\delta(\vec{x}^0)\subset B_{\varepsilon_0}(\vec{x})$ получаем, что

$$\forall x\in B_\delta(\vec{x}^0)\ \exists!\, y\in B_\varepsilon(y^0):\ F(y,\vec{x})=0.$$

Тем самым показано, что в пределах $B_\delta(\vec{x}^0)$ существует единственная функция $y(\vec{x})$, удовлетворяющая условию $y(\vec{x})\in B_\varepsilon(y^0)$, которая является решением уравнения $F(y,\vec{x})=0$. Так как при этом $y\left(B_\delta(\vec{x}^0)\right)\subset B_\varepsilon(y^0)$, то, значит, $y(\vec{x})$ непрерывна в точке $\vec{x}^0$, $y(\vec{x}^0)=y^0$.

У любой точки $(y(\vec{x}),\vec{x})$, где $\vec{x}\in B_\delta(\vec{x}^0)\subset B_{\varepsilon_0}(\vec{x}^0)$, $y(\vec{x})\in B_\varepsilon(y^0)\subset B_{\varepsilon_0}(\vec{x}^0)$, и значит, $(y(\vec{x}),\vec{x})\in B_{2\varepsilon_0}\left((y^0,\vec{x}^0)\right)$, есть окрестность, в которой $F(y,\vec{x})$ непрерывна и имеет непрерывную частную производную F'_y, $F(y(\vec{x}),\vec{x})=0$, $F'_y(y(\vec{x}),\vec{x})\neq 0$. Значит, для точки $(y(\vec{x}),\vec{x})$ выполняются те же условия, что и для точки $y^0,\vec{x}^0)$, поэтому используя уже доказанную непрерывность $y(\vec{x})$ в точке $\vec{x}^0$ делаем вывод, что $y(\vec{x})$ непрерывна в любой точке $\vec{x}\in B_\delta(\vec{x}^0)$.

Теперь перейдем к дифференцируемости. Если $F(y,\vec{x})$ дифференцируема в точке $(y^0,\vec{x}^0)$, то

$$\Delta F=F'_y(y^0,\vec{x}^0)\Delta y+\sum_{k=1}^n F'_{x_k}(y^0,\vec{x}^0)\Delta x_k+o(1)(|\Delta y|+\|\overrightarrow{\Delta x}\|),$$

причем фигурирующее тут $o(1)$ не больше того, которое присутствует в обычной записи последнего члена в виде $o(1)\cdot\|(\Delta y,\overrightarrow{\Delta x})\|$, так как $|\Delta y|+\|\overrightarrow{\Delta x}\|=\|(\Delta y,0)\|+\|(0,\overrightarrow{\Delta x})\|\geqslant\|(\Delta y,\overrightarrow{\Delta x})\|$ (по неравенству треугольника для норм), где $\|\overrightarrow{\Delta x}\|$ — норма в $\mathbb{R}^n$, а три последующие нормы — в $\mathbb{R}^{1+n}$.

Если взять $y=y(\vec{x})$, то $\Delta F=F(y(\vec{x}),\vec{x})-F(y^0,\vec{x}^0)=0-0=0$ и, значит,

$$F'_y(y^0,\vec{x}^0)\cdot\Delta y+\sum_{k=1}^n F'_{x_k}(y^0,\vec{x}^0)\cdot\Delta x_k+o(1)\cdot\Big(|\Delta y|+\|\overrightarrow{\Delta x}\|\Big)=0.$$

Как уже установлено, если $\overrightarrow{\Delta x}\to 0$, то $\Delta y=y(\vec{x})-y^0\to 0$ и, значит, $o(1)\to 0$ при $\overrightarrow{\Delta x}\to 0$. Будем брать приращение $\overrightarrow{\Delta x}$ столь малым, что $|o(1)|\leqslant\frac{1}{2}\left|F'_y(y^0,\vec{x}^0)\right|$. Тогда

$$\Delta y=\frac{-\sum_{k=1}^n F'_{x_k}(y^0,\vec{x}^0)\cdot\Delta x_k+o(1)\cdot\|\overrightarrow{\Delta x}\|}{F'_y(y^0,\vec{x}^0)+0(1)}=\sum_{k=1}^n O(1)\cdot\Delta x_k+o(1)\cdot\|\overrightarrow{\Delta x}\|=O(1)\cdot\|\overrightarrow{\Delta x}\|.$$

Следовательно, в предшествующем равенстве член

$$o(1)\cdot\Big(|\Delta y|+\|\overrightarrow{\Delta x}\|\Big)=o(1)\cdot\|\overrightarrow{\Delta x}\|,$$

где последнее $o(1)\to 0$ при $\overrightarrow{\Delta x}\to o$, т.е.

$$F'_y(y^0,\vec{x}^0)\cdot\Delta y+\sum_{k=1}^n F'_{x_k}(y^0,\vec{x}^0)\cdot\Delta x_k+o(1)\cdot\|\overrightarrow{\Delta x}\|=0.$$

Отсюда получаем, что

$$\Delta y = \sum_{k=1}^{n} -\frac{F'_{x_k}(y^0, \vec{x}^{\,0})}{F'_y(y^0, \vec{x}^{\,0})} \cdot \Delta x_k + o(\|\overrightarrow{\Delta x}\|,$$

т.е. $y(\vec{x})$ дифференцируема в точке $\vec{x}^{\,0}$ и

$$\boldsymbol{d}y(\overrightarrow{\Delta x})\big|_{\vec{x}^{\,0}} = \sum_{k=1}^{n} -\frac{F'_{x_k}(y^0, \vec{x}^{\,0})}{F'_y(y^0, \vec{x}^{\,0})} \cdot \Delta x_k.$$

Если $F(y, \vec{x})$ дифференцируема в окрестности точки $(y^0, \vec{x}^{\,0})$, то для каждой точки $(y(\vec{x}), \vec{x})$ из некоторой окрестности точки $(y^0, \vec{x}^{\,0})$ выполняются те же условия, что и для точки $(y^0, \vec{x}^{\,0})$, и, значит, $y(\vec{x})$ при это будет дифференцируема в точке $\vec{x}$. Отсюда и из непрерывности $y(\vec{x})$ следует, что существует такая окрестность точки $\vec{x}^{\,0}$, на которой $y(\vec{x})$ дифференцируема. ▲

Следствие 1. *При предположениях теоремы с дополнительным требованием дифференцируемости $F(y, \vec{x})$ в точке $(y^0, \vec{x}^{\,0})$ (в окрестности точки $(y^0, \vec{x}^{\,0})$) неявная функция $y(\vec{x})$ имеет в точке $\vec{x}^{\,0}$ (в некоторой окрестности точки $\vec{x}^{\,0}$) частные производные*

$$\frac{\partial y(\vec{x})}{\partial x_k} = -\frac{F'_{x_k}(y, \vec{x})}{F'_y(y, \vec{x})}$$

в точке $\vec{x} = \vec{x}^{\,0}$, $y = y(\vec{x}) = y(\vec{x}^{\,0})$ (в окрестности точки $\vec{x}^{\,0}$, $y = y(\vec{x})$).

▼ Следствие сразу следует из вида дифференциала $\boldsymbol{d}y(\overrightarrow{\Delta x})$. ▲

Следствие 2. *Если $F(y, \vec{x})$ удовлетворяет предположениям теоремы и непрерывно дифференцируема в точке $(y^0, \vec{x}^{\,0})$ (в окрестности точки $(y^0, \vec{x}^{\,0})$), то неявная функция $y(\vec{x})$ непрерывно дифференцируема в точке $\vec{x}^{\,0}$ (в некоторой окрестности точки $\vec{x}^{\,0}$).*

▼ Действительно, по следствию 1 частная производная $\frac{\partial y(\vec{x})}{\partial x_k}$ непрерывна в точке $\vec{x}^{\,0}$ (в окрестности точки $\vec{x}^{\,0}$) как отношение двух непрерывных функций. ▲

ДОБАВЛЕНИЕ

Теорема о неявных функциях

Приведём для ознакомления только формулировку теоремы о неявных функциях. Доказывать её не будем.

Теорема 3 (о неявных функциях). *Пусть m функций $F_i(\vec{y},\vec{x})$, $i=1,\ldots,m$, от $m+n$ переменных $(\vec{y},\vec{x})$, $\vec{y}\in\mathbb{R}^m$, $\vec{x}\in\mathbb{R}^n$, непрерывны в некоторой окрестности точки $(\vec{y}^0,\vec{x}^0)\in\mathbb{R}^{m+n}$ и имеют в этой окрестности непрерывные частные производные $\frac{\partial F_i(\vec{y},\vec{x})}{\partial y_j}$, $i=1,\ldots,m$, $j=1,\ldots,m$. Тогда если в точке $(\vec{y}^0,\vec{x}^0)$ все функции обращаются в нуль, $F_i(\vec{y}^0,\vec{x}^0)=0$, $i=1,\ldots,m$, а определитель, называемый* ***определителем Якоби*** *или* ***якобианом***,

$$\det\left(\frac{\partial F_i(\vec{y}^0,\vec{x}^0)}{\partial y_j}\right)_{\substack{1\leqslant i\leqslant m\\ 1\leqslant j\leqslant m}}\neq 0,$$

то

$$\exists\,\varepsilon_j^0>0,\ j=1,\ldots,m,\ \forall\varepsilon_j\in(0,\varepsilon_j^0),\ j=1,\ldots,m,\ \exists B_\delta(\vec{x}^0)\subset\mathbb{R}^n\ \forall\vec{x}\in B_\delta(\vec{x}^0)$$
$$\exists!\,\vec{y}=(y_1,\ldots,y_m)\in\prod_{j=1}^m B_{\varepsilon_j}(y_j^0):\ F_i(\vec{y},\vec{x})=F_i(y_1,\ldots,y_m,\vec{x})=0,\ i=1,\ldots,m,$$

то есть для любых достаточно малых чисел $\varepsilon_j>0$, $j=1,\ldots,m$, найдется такая окрестность $B_\delta(\vec{x}^0)$ точки $\vec{x}^0$, что в пределах этой окрестности существуют единственные m функций $y_j(\vec{x})$, $j=1,\ldots,m$, удовлетворяющие условиям $y_j(\vec{x})\in B_{\varepsilon_j}(y_j^0)$, $j=1,\ldots,m$, и являющиеся решением системы уравнений $F_i(\vec{y},\vec{x})=F_i(y_1,\ldots,y_m,\vec{x})=0$, $i=1,\ldots,m$, причем эти функции непрерывны в некоторой окрестности точки $\vec{x}^0$.

Если дополнительно потребовать дифференцируемость $F_i(\vec{y},\vec{x})$, $i=1,\ldots,m$, в точке $(\vec{y}^0,\vec{x}^0)$ (в некоторой ее окрестности), то функции $y_j(\vec{x})$, $j=1,\ldots,m$, будут дифференцируемы в точке $\vec{x}^0$ (в некоторой ее окрестности). Причем, если $F_i(\vec{y},\vec{x})$, $i=1,\ldots,m$, непрерывно дифференцируемы в точке $(\vec{y}^0,\vec{x}^0)$ (в некоторой ее окрестности), то функции $y_j(\vec{x})$, $j=1,\ldots,m$, будут непрерывно дифференцируемы в точке $\vec{x}^0$ (в некоторой ее окрестности).