Московский государственный университет
им. М.В.Ломоносова.
Факультет вычислительной математики и кибернетики

**И.В. Дмитриева, В.А. Морозова, С.И. Орлик.**

# Методические материалы для подготовки к государственному экзамену по прикладной математике и информатике по дополнительной части программы специалистов

Планета Знаний
Москва
2007

УДК 512; 517; 519.6; 510.5
ББК 22.1
Д54

*Печатается по решению*
*Редакционно-издательского Совета факультета*
*Вычислительной математики и кибернетики МГУ*
*им. М.В. Ломоносова.*

Рецензенты:
профессор А.В. Гулин,
профессор Е.В. Захаров.

**И.В.Дмитриева, В.А. Морозова, С.И.Орлик.**
**Методические материалы для подготовки к государственному экзамену по прикладной математике и информатике**
Д54 **по дополнительной части программы специалистов:**
– ООО "Планета Знаний", 2007. –315с.
ISBN 978–5–903242–05–04

Пособие предназначено для студентов факультета ВМиК МГУ, обучающихся на кафедрах Математической физики, Вычислительных технологий и моделирования, Вычислительных методов, Автоматизации научных исследований, Нелинейных динамических систем и процессов управления, Общей математики, Квантовой информатики.

Пособие не является учебником, но весьма полно отражает содержание всех вопросов **дополнительной части** программы государственного экзамена.

Пособие подготовлено преподавателями кафедры Математической физики.

ISBN 978-5-903242-05-04

УДК 512; 517; 519.6; 510.5
ББК 22.1
Д54

## ВОПРОСЫ ГОСЭКЗАМЕНА

по направлению «Прикладная математика и информатика»
(дополнительная часть)

Программа утверждена Ученым Советом факультета ВМиК.

## Литература к дополнительной части программы:

1. **Ильин В.А., Садовничий В.А., Сёндов Бл.Х.** Математический анализ, т.I, II.
2. **Колмогоров А.Н., Фомин С.В.** Элементы теории функций и функционального анализа.
3. **Ильин В.А., Позняк Э.Г.** Линейная алгебра. – М.: Наука, 1998.
4. **Ильин В.А., Позняк Э.Г.** Аналитическая геометрия. – М.: Наука, 1998.
5. **Тихонов А.Н., Самарский А.А** Уравнения математической физики. – М.: Наука, 1966.
6. **Тихонов А.Н., Васильева А.Б., Свешников А.Г.** Курс обыкновенных дифференциальных уравнений. – М.: Наука, 1980.
7. **Самарский А.А., Гулин А.В.** Численные методы. – М.: Наука, 1989.
8. **Свешников А.Г., Тихонов А.Н.** Основы теории аналитических функций комплексного переменного. – М.: Наука, 1979.

## Вопрос 1.

## Необходимые условия экстремума функции нескольких переменных. Достаточные условия.

1. Экстремумом называются наибольшее или наименьшее значения функции $f(x)$, принимающей действительные значения. Если в области определения $Dom\, f$ функции $f$ существуют точки $x^*$, в которых $f$ принимает наибольшее или наименьшее значения на всей области её определения, то такие точки называются точками **абсолютного максимума** или **абсолютного минимума**. Если существует такая окрестность $O$ точки $x^*$, что для всех точек из $Dom\, f$, принадлежащих этой окрестности, $x^*$ является точкой экстремума, то говорят о **локальном экстремуме**. Различают точки **строгого и нестрогого экстремумов**; например, точка $x^*$ называется точкой нестрогого (строгого) локального минимума функции $f$, если существует такая окрестность $O$ точки $x^*$, что для всех $x \in O \cap Dom\, f$ выполнено неравенство $f(x) \geq f(x^*)$ (соответственно, $f(x) > f(x^*)$ при $x \neq x^*$).

Если функция $f$ непрерывна на компактном множестве (в конечномерном пространстве все компактные множества – это замкнутые ограниченные множества), то она в некоторых его точках принимает свои максимальное и минимальное значения. В более общем случае приходится искать $\inf\limits_{x \in Dom\, f} f(x)$ и $\sup\limits_{x \in Dom\, f} f(x)$, которые могут и не достигаться; в таких задачах требуется найти эти числа и, возможно, построить минимизирующую (максимизирующую) последовательность $x_k \in Dom\, f$.

Далее мы будем считать, что $x \in \Omega \subseteq \mathrm{R}^n$. Напомним сначала основные факты об экстремуме функции **одной** переменной; $n = 1$.

**Теорема Ферма**. Пусть функция $f(x)$ одной действительной переменной определена в некотором интервале

$a < x < b$ и во **внутренней** точке $x^*$ этого интервала принимает наименьшее (наибольшее) значение. Если существует конечная производная $f'(x^*)$, то $f'(x^*) = 0$. (Коротко: если $f(x)$ дифференцируема в точке $x^*$ и имеет в $x^*$ локальный экстремум, то $f'(x^*) = 0$)♦

**Замечания**. $f'(x^*) = 0$ геометрически означает, что в точке $(x^*, f(x^*))$ касательная к графику функции $f$ параллельна оси абсцисс. Корни уравнения $f'(x) = 0$ называют **стационарными точками** функции $f$.

В теореме существенно предположение, что $x^*$ – **внутренняя** точка интервала; без него теорема перестаёт быть верной (приведите примеры!). ♦

**Задачи**. Найдите стационарные точки следующих функций, если такие существуют; выясните, являются ли они точками экстремумов. Найдите точки экстремумов, если они существуют.

$f(x) = x^3$. $\quad f(x) = x^{\frac{2}{3}}$. $\quad f(x) = |x|$. $\quad f(x) = x^{\frac{1}{3}}$.

$$f(x) = \begin{cases} x \cdot \sin\dfrac{1}{x}, & x \neq 0, \\ 0, & x = 0. \end{cases} \qquad f(x) = \begin{cases} x^2 \cdot \sin\dfrac{1}{x}, & x \neq 0, \\ 0, & x = 0. \end{cases}$$

$$f(x) = \begin{cases} \dfrac{x}{1 + e^{\frac{1}{x}}}, & x \neq 0, \\ 0, & x = 0. \end{cases} \qquad f(x) = \begin{cases} x \cdot e^{-\frac{1}{x^2}}, & x \neq 0, \\ 0, & x = 0. \end{cases}$$ ♦

**Достаточное условие наличия или отсутствия экстремума**. Пусть функция $f(x)$ дифференцируема всюду в некоторой окрестности точки $x^*$, кроме, быть может, самой точки $x^*$, и непрерывна в $x^*$. Если в пределах этой окрестности $f'(x) < 0$ при $x < x^*$ и $f'(x) > 0$ при $x > x^*$, то в точке $x^*$ функция $f$ имеет локальный минимум (аналогично по смыслу и для максимума). Если же $f'(x)$ имеет один и тот же знак слева и справа от $x^*$, то экстремума в $x^*$ нет. ♦

Это правило позволяет ответить на вопрос об экстремуме лишь в том случае, если на интервале $a < x < b$ имеется не более конечного числа стационарных точек функции $f$ или точек, где $f$ не имеет конечной производной. Для решения задачи об экстремуме можно использовать и высшие производные функции $f$, если они существуют:

**Достаточное условие наличия или отсутствия экстремума.** Пусть $m$ – целое положительное число. Пусть в окрестности точки $x^*$ функция $f(x)$ имеет производную порядка $m$, которая непрерывна в $x^*$. Пусть $f'(x^*) = 0$, $f^{(2)}(x^*) = 0, \ldots, f^{(m-1)}(x^*) = 0$, $f^{(m)}(x^*) \neq 0$. Если $m$ – **чётное** число, то функция $f$ имеет в $x^*$ минимум при $f^{(m)}(x^*) > 0$ (максимум при $f^{(m)}(x^*) < 0$). Если же $m$ – **нечётное** число, то экстремума в $x^*$ нет. ♦

Пусть теперь функция $f(x)$ является кусочно непрерывной и кусочно гладкой **на отрезке** $a \le x \le b$. Это значит, что на $a \le x \le b$ может существовать лишь конечное число точек, в которых $f$ либо терпит разрыв первого рода, либо непрерывна, но не имеет производной. Точками экстремума функции $f$ на отрезке могут быть только те точки $x^*$, в которых

$1^0$. либо $f$ терпит разрыв;

$2^0$. либо $f$ непрерывна, но не имеет производной;

$3^0$. либо $f$ имеет равную нулю производную;

$4^0$. либо $x^* = a$ или $x^* = b$.

Такие точки называют **подозрительными на экстремум**. Чтобы найти абсолютный минимум (максимум) функции на отрезке, надо перебрать все точки локального минимума (максимума) и выбрать среди них точку с наименьшим (наибольшим) значением функции, если такое существует.

**Задачи**. Найдите точки экстремумов функций на $\mathbb{R}^1$.

$$f(x)=\begin{cases} x^2\cdot\left(1+\sin\dfrac{1}{x}\right), & x\neq 0,\\ 0, & x=0.\end{cases} \qquad f(x)=e^x+e^{-x}+2\cos x.\ \blacklozenge$$

**Задача**. Найдите точки экстремума функции $f(x)=\begin{cases} x\cdot\ln|x|, & x\neq 0,\\ 0, & x=0,\end{cases}$ на отрезке $-1\le x\le e$. ◆

**Задача**. Найдите точки экстремумов функции $f(x)=\sin^3 x+\cos^3 x$ на отрезках $0\le x\le \dfrac{3\pi}{4}$, $0\le x\le 2\pi$.

◆

**Задача**. Найдите точки экстремумов функции $f(x)=\begin{cases} \dfrac{1}{1+e^{\frac{1}{x}}}, & x\neq 0,\\ 0, & x=0.\end{cases}$ на отрезках $0\le x\le 1$, $-1\le x\le 0$, $-1\le x\le 1$, $1\le x\le 2$ и на $\mathbb{R}^1$. ◆

**Задача**. Непрерывная на отрезке $a\le x\le b$ функция $f(x)$ имеет в точке $x^*$ локальный строгий минимум; $a<x^*<b$. Означает ли это, что в некоторой окрестности точки $x^*$ функция $f$ убывает при $x<x^*$ и возрастает при $x>x^*$? Рассмотрите функцию $f(x)=\begin{cases} 2x^2+x^2\cdot\sin\dfrac{1}{x}, & x\neq 0,\\ 0, & x=0,\end{cases}$ на отрезке $-1\le x\le 1$. ◆

**Задача**. Пусть $f(x)$ дифференцируема на отрезке $a\le x\le b$ и достигает своего минимума в точке $x^*$ этого отрезка. Верно ли, что для всех точек $x$ этого отрезка выполнено неравенство $f'(x^*)\cdot(x-x^*)\ge 0$? Является ли выполнение указанного неравенства достаточным для того, чтобы в точке $x^*$

функция $f$ достигала минимума на отрезке? Рассмотрите функцию $f(x) = -\cos x$ на $0 \le x \le \frac{3\pi}{2}$. Является ли выполнение указанного неравенства достаточным для того, чтобы в $x^*$ **выпуклая** непрерывно дифференцируемая функция $f$ достигала минимума на отрезке? ♦

**Задача**. Докажите, что на отрезке $[-1,1]$ функция

$$f(x) = \begin{cases} x^4 e^{-\frac{x^2}{4}} \cdot \sin\frac{8}{x^3}, & x \ne 0, \\ 0, & x = 0, \end{cases}$$

имеет ограниченную производную $f'(x)$. Достигает ли $f'(x)$ абсолютного экстремума на $[-1,1]$?♦

2. Пусть теперь $f(x)$ – функция $n$ действительных независимых переменных, определённая в области $\Omega \subseteq \mathrm{R}^n$; $x = (x_1, \dots, x_n) \in \mathrm{R}^n$.

**Необходимое условие существования экстремума.**

Если функция $f(x)$ во **внутренней** точке $x^*$ области $\Omega$ имеет экстремум, и в этой точке существуют конечные частные производные $\frac{\partial f(x^*)}{\partial x_i}$, $i = 1, \dots, n$, то все эти производные равны нулю.

**Доказательство**. Фиксируем значения $x_i = x_i^*$, $i = 2, \dots, n$, оставляя величину $x_1$ переменной. Тогда получим функцию $f(x_1, x_2^*, \dots, x_n^*)$ от **одной** переменной $x_1$. Поскольку $x^*$ – точка экстремума функции $f$ (пусть для определённости это точка минимума), то и для указанной функции одной переменной $x_1$ точка $x_1^*$ является точкой минимума. Тогда по теореме Ферма

$\dfrac{\partial f\left(x^*\right)}{\partial x_1} = 0$. Точно так же равны нулю и остальные частные производные. ♦

**Замечание**. Если во внутренней точке $x^*$ области $\Omega$ функция $f$ имеет экстремум и **дифференцируема** в $x^*$, то в этой точке её дифференциал $df(x)\big|_{x=x^*} = 0$. Это означает, что гиперплоскость в $\mathrm{R}^{n+1}$, проведённая через точку $\left(x^*, f\left(x^*\right)\right)$, касательная к графику функции $f$ в этой точке, параллельна подпространству независимых переменных. Такую точку $x^*$ называют стационарной. ♦

**Замечание**. Доказанная теорема не является достаточным для экстремума условием. В стационарной точке $x^*$ функции $n$ переменных экстремума может не быть. Например, функция $f(x) = x_1 \cdot x_2$ имеет седловую точку $x^* = 0$. ♦

Прежде чем сформулировать достаточное условие наличия или отсутствия экстремума у функции $n$ переменных, напомним некоторые понятия из теории квадратичных форм. Пусть $\left(a_{ij}\right)_{\substack{i=1,\dots,n \\ j=1,\dots,n}}$ – **симметричная** действительная матрица, $h = \left(h_1, \dots, h_n\right)$ – действительные независимые переменные. Функция $Q(h) = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} h_i h_j$ называется квадратичной формой, определённой на линейном пространстве $\mathrm{R}^n$. Квадратичная форма $Q(h)$ называется неотрицательной, если $Q(h) \geq 0$ для всех $h \in \mathrm{R}^n$; $Q(h)$ называется неположительной, если $Q(h) \leq 0$ для всех $h \in \mathrm{R}^n$; $Q(h)$ называется **положительно определённой**, если для любого $h \in \mathrm{R}^n$ из условия $h \neq 0$ следует $Q(h) > 0$; $Q(h)$ называется **отрицательно определённой**, если для любого $h \in \mathrm{R}^n$ из условия $h \neq 0$ следует $Q(h) < 0$. Квадратичная форма $Q(h)$ является положительно определённой в том и только в том случае, если её положительный индекс инерции равен размерности $n$

пространства $\mathbb{R}^n$; $Q(h)$ является отрицательно определённой в том и только в том случае, если её отрицательный индекс инерции равен $n$. Если для некоторого $\tilde{h} \in \mathbb{R}^n$ $Q(\tilde{h}) > 0$, а для некоторого $\tilde{\tilde{h}} \in \mathbb{R}^n$ $Q\left(\tilde{\tilde{h}}\right) < 0$, то квадратичную форму называют **знакопеременной**. Знакоопределённость квадратичной формы устанавливает

**Критерий Сильвестра**. Чтобы квадратичная форма $Q(h)$ была **положительно определённой**, необходимо и достаточно, чтобы все ведущие миноры матрицы $(a_{ij})$

$$a_{11}, \quad \begin{vmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{vmatrix}, \quad \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{12} & a_{22} & a_{23} \\ a_{13} & a_{23} & a_{33} \end{vmatrix}, \ldots, \quad \begin{vmatrix} a_{11} & \ldots & a_{1n} \\ . & \ldots & . \\ a_{1n} & \ldots & a_{nn} \end{vmatrix}$$

были положительными. Чтобы $Q(h)$ была **отрицательно определённой**, необходимо и достаточно, чтобы все ведущие миноры нечётного порядка были отрицательными, четного – положительными. ♦

Если функция $f(x)$ дважды дифференцируема в точке $x^*$, то в этой точке её второй дифференциал

$$d^2 f(x)\Big|_{x=x^*} = \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{\partial^2 f(x)}{\partial x_i \partial x_j}\Bigg|_{x=x^*} \cdot dx_i dx_j$$

представляет собой квадратичную форму относительно дифференциалов $dx_1, \ldots, dx_n$ независимых переменных. Наличие или отсутствие экстремума функции $f$ в её стационарной точке $x^*$ определяется знакоопределённостью или знакопеременностью этой квадратичной формы.

**Достаточное условие наличия или отсутствия экстремума**. Пусть функция $f(x)$ определена и дважды дифференцируема в окрестности её **стационарной точки** $x^*$, а все её вторые частные производные непрерывны в $x^*$. Если в точке $x^*$ второй дифференциал $d^2 f$ является **положительно определённой**

квадратичной формой относительно $dx_1, \ldots, dx_n$, то функция $f$ имеет в $x^*$ **локальный минимум**. Если в $x^*$ второй дифференциал $d^2 f$ является **отрицательно определённой** квадратичной формой, то $f$ имеет в $x^*$ **локальный максимум**. Если $d^2 f$ в точке $x^*$ является **знакопеременной** квадратичной формой, то в $x^*$ экстремума **нет**. ♦

**Замечание**. Если в стационарной точке $x^*$ второй дифференциал является лишь неотрицательной или неположительной квадратичной формой, то никакого заключения об экстремуме функции $f$ в $x^*$ сделать нельзя. В этом случае нужно дополнительное исследование с привлечением, возможно, высших производных. Оно может оказаться очень громоздким. ♦

**Задача**. Каковы условия, достаточные для наличия или отсутствия экстремума в стационарной точке функции **двух переменных**, которым должны удовлетворять её частные производные второго порядка? ♦

**Задачи**. Исследуйте на экстремум данные функции.

$f(x) = x_1^2 + x_2^2$. $\quad f(x) = x_1^2 - x_2^2$. $\quad f(x) = x_1^4 + x_2^4$.

$f(x) = x_1^3 + x_2^3$. $\quad f(x) = x_1^3 + x_2^3 - 9x_1x_2 + 1$.

$f(x) = (x_1 - x_2)^2 + x_1^3 + x_2^3$.

$f(x) = x_1^4 + x_2^4 - 2x_1^2 + 4x_1x_2 - 2x_2^2 + 1$.

$f(x) = x_1^2 - 2x_1x_2^2 + x_2^4 - x_2^5$.

$f(x) = (x_1 + x_2 - 1) \cdot \exp\{-(x_1^2 - x_1x_2 + x_2^2)\}$. ♦

**Задача**. Найдите точки экстремумов функции $f(x) = \sin(x_1 + x_2) - \sin x_1 - \sin x_2$ на $\mathrm{R}^2$; на множестве $\{x \in R^2 \mid x_1 \geq 0, x_2 \geq 0, x_1 + x_2 \leq 2\pi\}$. ♦

**Задача**. В $\mathrm{R}^n$ даны $m$ точек. Найдите точку, сумма квадратов евклидовых расстояний от которой до этих точек минимальна. ♦

**Задача**. Может ли функция двух переменных иметь на $\mathrm{R}^2$ бесконечно много точек локального минимума и ни одной точки локального максимума? Рассмотрите функцию $f(x) = x_1 e^{x_1} - (1 + e^{x_1}) \cdot \cos x_2$. ♦

**Задача**. Имеет ли функция $f(x) = (x_1 - x_2^2)(2x_1 - x_2^2)$ локальный минимум в точке $x^* = 0$? Имеет ли эта функция локальный минимум вдоль каждой прямой, проходящей через точку $x^* = 0$? ♦

3. Пусть функция $f(x)$ определена на множестве $\Omega \subseteq \mathrm{R}^n$. Говорят, что $f$ **полунепрерывна снизу** (сверху) в точке $x_0 \in \Omega$, если для любой последовательности $\{x_k\} \in \Omega$, сходящейся к $x_0$, нижний предел $\varliminf_{k \to \infty} f(x_k) \geq f(x_0)$ (соответственно, верхний предел $\varlimsup_{k \to \infty} f(x_k) \leq f(x_0)$). Функция $f(x)$ полунепрерывна снизу (сверху) в точке $x_0 \in \Omega$ в том и только в том случае, если для любого $\varepsilon > 0$ существует $\delta > 0$ такое, что для всех $x \in \Omega$, удовлетворяющих неравенству $\|x - x_0\| < \delta$, выполнено неравенство $f(x) \geq f(x_0) - \varepsilon$ (соответственно, $f(x) \leq f(x_0) + \varepsilon$). Функцию $f$ называют полунепрерывной снизу (сверху) на множестве $\Omega$, если она полунепрерывна снизу (сверху) в каждой точке этого множества. ($\|\cdot\|$ – евклидова норма в $\mathrm{R}^n$.)

**Теорема Вейерштрасса**. Пусть $\Omega$ – **компактное** множество, а функция $f(x)$ определена, конечна и полунепрерывна снизу на $\Omega$. Тогда $\inf_{x \in \Omega} f(x) > -\infty$, множество $\left\{x^* \in \Omega \,\middle|\, f(x^*) = \inf_{x \in \Omega} f(x)\right\}$ непусто, компактно, а любая минимизирующая последовательность сходится к этому множеству. ♦

4. Функция $f(x)$, определённая на **выпуклом** множестве $\Omega \subseteq \mathrm{R}^n$, называется **выпуклой** на этом множестве, если

$$f(\alpha x + (1-\alpha)y) \le \alpha \cdot f(x) + (1-\alpha) \cdot f(y) \qquad (1)$$

при всех $x, y \in \Omega$ и при всех $\alpha$ на отрезке $0 \le \alpha \le 1$. Если в (1) равенство возможно только при $\alpha = 0$ и $\alpha = 1$, то $f$ называется **строго выпуклой** на $\Omega$.

**Теорема**. Пусть $\Omega$ – выпуклое множество, а функция $f(x)$ определена и выпукла на $\Omega$. Тогда всякая точка локального минимума функции $f$ является точкой её абсолютного минимума на $\Omega$, причём множество $\left\{x^* \in \Omega \,\middle|\, f(x^*) = \inf\limits_{x \in \Omega} f(x)\right\}$ выпукло. Если функция $f$ **строго выпукла**, то это множество содержит не более одной точки. ♦

**Теорема**. Пусть $\Omega$ – выпуклое множество, а функция $f(x)$ непрерывно дифференцируема на $\Omega$. Пусть $x^*$ – точка минимума функции $f$ на множестве $\Omega$. Тогда для всех $x \in \Omega$ выполнено неравенство

$$\left(grad\, f(x^*), x - x^*\right) \ge 0 . \qquad (2)$$

Если $x^*$ – **внутренняя** точка $\Omega$, то (2) превращается в равенство $grad\, f(x^*) = 0$. Если, кроме того, $f(x)$ **выпукла** на $\Omega$, то условие (2) является **достаточным** для того, чтобы $x^*$ была точкой минимума $f$ на $\Omega$.♦

Непрерывная выпуклая функция на выпуклом замкнутом множестве может не достигать своей точной нижней грани на этом множестве.

**Пример**. $\Omega = \left\{x \in \mathrm{R}^1 \,\middle|\, x \ge 1\right\}$, $\quad f(x) = \dfrac{1}{x} > 0$ при всех $x \in \Omega$, $\inf\limits_{x \in \Omega} f(x) = 0$.♦

Для некоторого подкласса выпуклых функций такая ситуация невозможна. Функция $f(x)$, определённая на выпуклом множестве $\Omega$, называется **сильно выпуклой на** $\Omega$, если существует постоянная $\theta > 0$ такая, что

$$f(\alpha x+(1-\alpha)y)\le \alpha\cdot f(x)+(1-\alpha)\cdot f(y)-\alpha(1-\alpha)\cdot\theta\cdot\|x-y\|^2$$

при всех $x, y\in\Omega$ и при всех $\alpha$ на отрезке $0\le\alpha\le 1$ ($\|\cdot\|$ – евклидова норма в $\mathrm{R}^n$). Сильно выпуклая на $\Omega$ функция является выпуклой и даже строго выпуклой.

**Теорема**. Пусть множество $\Omega$ выпукло и замкнуто, а функция $f(x)$ **сильно выпукла** и **полунепрерывна снизу** на $\Omega$. Тогда $\inf\limits_{x\in\Omega} f(x)>-\infty$; во множестве $\Omega$ существует, и притом единственная, точка $x^*$, для которой $f(x^*)=\inf\limits_{x\in\Omega} f(x)$; а любая минимизирующая последовательность $\{x_k\}$ сходится к $x^*$.♦

**Задача**. $f(x)=\dfrac{1}{2}(Ax,x)-(b,x)$, где $x\in\mathrm{R}^n$, $A$ – симметричная действительная $n\times n$-матрица, $A\ge 0$, $b$ – фиксированный вектор из $\mathrm{R}^n$. Найдите необходимое и достаточное условие (абсолютного) минимума функции $f$ в точке $x^*$.♦

**Задача**. $f(x)=\|Ax-b\|^2$, где $x\in\mathrm{R}^n$, $A$ – действительная $m\times n$-матрица, $b$ – фиксированный вектор из $\mathrm{R}^m$, $\|\cdot\|$ – евклидова норма в $\mathrm{R}^m$. Найдите необходимое и достаточное условие (абсолютного) минимума функции $f$ в точке $x^*$.♦

## Вопрос 2.

## Формулы Стокса, Остроградского.

1. **Формула Стокса** выражает равенство между потоком ротора векторного поля через двумерное ориентированное многообразие и циркуляцией этого векторного поля по ориентированному краю этого многообразия.

В трехмерном пространстве введём декартову прямоугольную систему координат $Oxyz$; пусть $\vec{i}$, $\vec{j}$, $\vec{k}$ – её единичные $\left(\left|\vec{i}\right| = \left|\vec{j}\right| = \left|\vec{k}\right| = 1\right)$ орты. Будем считать $\vec{i}$, $\vec{j}$, $\vec{k}$ **правой тройкой** векторов. Напомним понятие **ротора** векторного поля. Для дифференцируемого векторного поля

$$\vec{A}(x, y, z) = \{P(x, y, z), Q(x, y, z), R(x, y, z)\}$$

можно ввести дифференциальную операцию $rot\,\vec{A}$, которая в **ортонормированном базисе** $\vec{i}$, $\vec{j}$, $\vec{k}$ имеет вид

$$rot\,\vec{A} = \left(\frac{\partial R}{\partial y} - \frac{\partial Q}{\partial z}\right)\vec{i} + \left(\frac{\partial P}{\partial z} - \frac{\partial R}{\partial x}\right)\vec{j} + \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right)\vec{k}.$$

Если ввести **оператор Гамильтона**

$$\vec{\nabla} = \vec{i}\,\frac{\partial}{\partial x} + \vec{j}\,\frac{\partial}{\partial y} + \vec{k}\,\frac{\partial}{\partial z},$$

то ротор поля $\vec{A}$ имеет смысл векторного произведения: $rot\,\vec{A} = \vec{\nabla} \times \vec{A}$, которое символически можно записать в виде определителя:

$$rot\,\vec{A} = \begin{vmatrix} \vec{i} & \vec{j} & \vec{k} \\ \dfrac{\partial}{\partial x} & \dfrac{\partial}{\partial y} & \dfrac{\partial}{\partial z} \\ P & Q & R \end{vmatrix}.$$

Ориентация пространства $Oxyz$ существенна: $rot\,\vec{A}$ является **псевдовектором** (или, что то же самое, **– аксиальным вектором**); это значит, что при изменении ориентации пространства на противоположную (правой тройки $\vec{i}$, $\vec{j}$, $\vec{k}$ на левую) он преобразуется в противоположный вектор. В то же время, $rot\,\vec{A}$ инвариантен относительно преобразований декартовых прямоугольных систем координат, имеющих одну и ту же ориентацию. Символ $\vec{\nabla}$ можно рассматривать как **вектор**, потому что его компоненты $\dfrac{\partial}{\partial x}$, $\dfrac{\partial}{\partial y}$, $\dfrac{\partial}{\partial z}$ преобразуются при переходе от ортонормированного базиса $\vec{i}$, $\vec{j}$, $\vec{k}$ к другому ортонормированному базису $\vec{i}_1$, $\vec{j}_1$, $\vec{k}_1$ по тем же правилам, по которым преобразуются компоненты обычных векторов. Поэтому если

$$\vec{A} = P(x,y,z)\vec{i} + Q(x,y,z)\vec{j} + R(x,y,z)\vec{k} =$$
$$= P_1(x_1,y_1,z_1)\vec{i}_1 + Q_1(x_1,y_1,z_1)\vec{j}_1 + R_1(x_1,y_1,z_1)\vec{k}_1,$$

то для векторного произведения двух векторов имеем в одной и той же точке

$$\left(\frac{\partial R}{\partial y} - \frac{\partial Q}{\partial z}\right)\vec{i} + \left(\frac{\partial P}{\partial z} - \frac{\partial R}{\partial x}\right)\vec{j} + \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right)\vec{k} =$$
$$= \left(\frac{\partial R_1}{\partial y_1} - \frac{\partial Q_1}{\partial z_1}\right)\vec{i}_1 + \left(\frac{\partial P_1}{\partial z_1} - \frac{\partial R_1}{\partial x_1}\right)\vec{j}_1 + \left(\frac{\partial Q_1}{\partial x_1} - \frac{\partial P_1}{\partial y_1}\right)\vec{k}_1,$$

– инвариантность формы $rot\,\vec{A}$.

Мы будем далее рассматривать ограниченную полную кусочно гладкую двустороннюю связную поверхность $S$ без особых точек (и без самопересечений) с кусочно гладкой границей $\partial S$. Под окрестностью этой поверхности будем понимать любое открытое в пространстве $Oxyz$ множество, содержащее $S$. Двусторонняя гладкая поверхность имеет **непрерывное** поле нормалей к ней; если такая поверхность не имеет особых точек, то можно ввести поле $\vec{n}$ единичных $\left(|\vec{n}| = 1\right)$ нормалей к этой поверхности: $\vec{n} = \{\cos\alpha, \cos\beta, \cos\gamma\}$, где $\alpha$, $\beta$, $\gamma$ – углы, которые составляет с ортами $\vec{i}$, $\vec{j}$, $\vec{k}$ вектор $\vec{n}$ в данной точке поверхности. Выбор поля единичных нормалей задаёт **ориентацию** гладкой поверхности; если эта поверхность имеет край, то устанавливается ориентация края, согласованная с ориентацией поверхности: правило «положительного обхода контура». Кусочно гладкая поверхность $S$ называется ориентированной, если каждый из её гладких кусков ориентирован, и возникающие при этом направления обхода каждого края согласованы: вдоль каждой дуги, где два контура совпадают, направления их обхода противоположны. Под $\partial S$ будем понимать кусочно гладкий край поверхности $S$, ориентированный с помощью единичного вектора $\vec{t}$ $(|\vec{t}| = 1)$, касательного к $\partial S$ в точках его гладкости. $\partial S$ – полная граница поверхности $S$ – состоит из конечного числа кусочно гладких простых замкнутых контуров (связных компонент границы). Ориентацию края $\partial S$ считаем согласованной с ориентацией поверхности $S$: если смотреть со стороны положительного направления нормали $\vec{n}$, то поверхность $S$ при обходе $\partial S$ остаётся слева.

Пусть в некоторой окрестности поверхности $S$ задано непрерывно дифференцируемое векторное поле

$$\vec{A}(x,y,z) = \{P(x,y,z), Q(x,y,z), R(x,y,z)\}$$

(т.е. функции $P$, $Q$, $R$ непрерывны и имеют непрерывные частные производные первого порядка в этой окрестности).

Через $dS$ будем обозначать элемент площади поверхности $S$, а через $dl$ – дифференциал длины дуги кривой $\partial S$.

**Формула Стокса:**

$$\iint\limits_S \left(rot\,\vec{A}, \vec{n}\right) dS = \oint\limits_{\partial S} \left(\vec{A}, \vec{t}\right) dl. \qquad (1)$$

В формуле (1) интеграл по $\partial S$ представляет собой сумму интегралов по всем связным компонентам границы $\partial S$. Этот интеграл можно записать в виде криволинейного интеграла **второго рода**:

$$\oint\limits_{\partial S} \left(\vec{A}, \vec{t}\right) dl = \oint\limits_{\partial S} Pdx + Qdy + Rdz.$$

Интеграл по $S$ можно записать в виде поверхностного интеграла **первого рода**:

$$\iint_S \left(rot\,\vec{A},\vec{n}\right)dS =$$

$$= \iint_S \left[\left(\frac{\partial R}{\partial y} - \frac{\partial Q}{\partial z}\right)\cos\alpha + \left(\frac{\partial P}{\partial z} - \frac{\partial R}{\partial x}\right)\cos\beta + \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right)\cos\gamma\right]dS$$

или в виде поверхностного интеграла **второго рода**:

$$\iint_S \left(rot\,\vec{A},\vec{n}\right)dS =$$

$$= \iint_S \left(\frac{\partial R}{\partial y} - \frac{\partial Q}{\partial z}\right)dydz + \left(\frac{\partial P}{\partial z} - \frac{\partial R}{\partial x}\right)dzdx + \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right)dxdy.$$

**Доказательство** формулы Стокса наиболее просто в случае, когда $S$ – **односвязная** поверхность, которая однозначно проектируется на каждую из координатных плоскостей системы $Oxyz$. (Односвязность поверхности $S$ означает, что всякий простой замкнутый контур, целиком лежащий на $S$, можно стянуть в точку при помощи непрерывной деформации, не выводящей его с поверхности $S$. Иными словами, поверхность $S$ не имеет дыр, пусть даже и точечных. Или, что то же самое, граница $\partial S$ поверхности $S$ состоит из одной связной компоненты.) В этом случае формула Стокса легко выводится из **интегральной формулы Грина** (докажите!).

В общем случае поверхность $S$ можно разбить на конечное число достаточно малых гладких её частей, каждая из которых однозначно проектируется на каждую из координатных плоскостей в некоторой специально для этой части выбранной декартовой системе координат. При этом существенно, что скалярное произведение $\left(rot\,\vec{A},\vec{n}\right)$ и элемент площади поверхности $dS$ не зависят от выбора **декартовой прямоугольной** системы координат. Поэтому поверхностный интеграл в формуле (1) инвариантен относительно выбора декартовой системы координат в пространстве, т.е. при переходе от $\vec{i}$, $\vec{j}$, $\vec{k}$ к новому ортонормированному базису $\vec{i}_1$, $\vec{j}_1$, $\vec{k}_1$ он не изменит ни своего значения, ни формы. Значение и форма криволинейного интеграла в (1) также не изменяются при переходе к новой декартовой прямоугольной системе координат. Именно на этих обстоятельствах и основано доказательство формулы Стокса в общем случае, именно инвариантный характер обоих интегралов и позволяет записывать формулу Стокса в виде (1). ♦

**Задача**. Верна ли формула Стокса для плоской поверхности $S$, параллельной одной из осей координат $Oxyz$ ?♦

**Замечание**. Формулу Грина можно рассматривать как частный случай формулы Стокса. (Для какой поверхности $S$ и для какого векторного поля $\vec{A}$ ?)♦

**Задача**. Пользуясь формулой Стокса, вычислите интеграл $\oint\limits_L ydx + zdy + xdz$, $L$ – окружность $x^2 + y^2 + z^2 = a^2$, $x + y + z = 0$, пробегаемая против хода часовой стрелки, если смотреть с положительной стороны оси $Ox$. Проверьте результат без применения формулы Стокса. ♦

**Задача**. Пользуясь формулой Стокса, вычислите интеграл $\oint\limits_L (y+z)dx + (z+x)dy + (x+y)dz$, где $L$ – эллипс $x = a \cdot \sin^2 \tau$, $y = 2a \cdot \sin\tau \cdot \cos\tau$, $z = a \cdot \cos^2 \tau$, $0 \le \tau \le \pi$, пробегаемый в направлении возрастания параметра $\tau$ .♦

**Задача**. Проверьте формулу Стокса для векторного поля $\vec{A} = \{y, z, x\}$ и круга $S$ в пересечении плоскости $x + z = a$ и сферы $x^2 + y^2 + z^2 = a^2$ .♦

2. Из формулы Стокса вытекает важное следствие для **поверхностно односвязных** областей $\Omega \subseteq Oxyz$ .

Область $\Omega$ в пространстве называется поверхностно односвязной, если для любого кусочно гладкого простого замкнутого контура $L$, целиком лежащего в $\Omega$, найдётся кусочно гладкая ориентируемая поверхность (без самопересечений), которая имеет $L$ своим краем и целиком содержится в $\Omega$.

**Примеры**. Всё пространство, шар, область между двумя концентрическими сферами – поверхностно односвязные области. Тор не является поверхностно односвязной областью. ♦

**Теорема**. Пусть в поверхностно односвязной области $\Omega \subseteq Oxyz$ функции $P(x, y, z)$, $Q(x, y, z)$, $R(x, y, z)$ определены и непрерывны вместе со своими частными производными первого порядка. Тогда следующие утверждения равносильны:

$1^0$. Интеграл $\oint\limits_L Pdx + Qdy + Rdz$ по **любому** замкнутому контуру $L$, целиком лежащему в $\Omega$, равен нулю.

$2^0$. Интеграл $\oint\limits_{KN} Pdx + Qdy + Rdz$ не зависит от пути, соединяющего точки $K$ и $N$, для любых $K, N \in \Omega$.

$3^0$. Выражение $Pdx + Qdy + Rdz$ является полным дифференциалом некоторой определённой в $\Omega$ **однозначной** функции $u(x, y, z)$.

$4^0$. В области $\Omega$ выполнены тождества $\dfrac{\partial Q}{\partial x} \equiv \dfrac{\partial P}{\partial y}$, $\dfrac{\partial R}{\partial y} \equiv \dfrac{\partial Q}{\partial z}$, $\dfrac{\partial P}{\partial z} \equiv \dfrac{\partial R}{\partial x}$.◆

Векторное поле $\vec{A} = \{P, Q, R\}$, которое в поверхностно односвязной области $\Omega$ имеет $rot\,\vec{A} \equiv 0$, называется **потенциальным**: существует такая скалярная функция $u(x, y, z)$ (потенциал поля $\vec{A}$), что $\vec{A} = grad\,u$.

3. **Формула Остроградского** (Остроградского-Гаусса) выражает равенство объёмного интеграла от дивергенции векторного поля по ограниченной области полному потоку этого векторного поля через границу этой области, ориентированную в направлении её внешней нормали.

Напомним понятие **дивергенции** векторного поля. Для дифференцируемого векторного поля

$$\vec{A}(x, y, z) = \{P(x, y, z), Q(x, y, z), R(x, y, z)\}$$

можно ввести дифференциальную операцию $div\,\vec{A}$, которая в **ортонормированном базисе** $\vec{i}$, $\vec{j}$, $\vec{k}$ имеет вид

$$div\,\vec{A} = \frac{\partial P}{\partial x} + \frac{\partial Q}{\partial y} + \frac{\partial R}{\partial z}.$$

Дивергенцию поля $\vec{A}$ можно рассматривать как символическое скалярное произведение оператора Гамильтона $\vec{\nabla}$ на вектор $\vec{A}$: $div\,\vec{A} = \left(\vec{\nabla}, \vec{A}\right)$. При переходе от ортонормированного базиса $\vec{i}$, $\vec{j}$, $\vec{k}$ к другому ортонормированному базису $\vec{i}_1$, $\vec{j}_1$, $\vec{k}_1$ символ $\vec{\nabla}$ преобразуется как обычный вектор; а скалярное произведение векторов инвариантно относительно преобразований декартовых прямоугольных координат. Поэтому $div\,\vec{A}$ является инвариантом относительно таких преобразований: если

$$\vec{A} = P(x,y,z)\vec{i} + Q(x,y,z)\vec{j} + R(x,y,z)\vec{k} =$$
$$= P_1(x_1,y_1,z_1)\vec{i}_1 + Q_1(x_1,y_1,z_1)\vec{j}_1 + R_1(x_1,y_1,z_1)\vec{k}_1,$$

то в одной и той же точке имеем

$$div\,\vec{A} = \frac{\partial P}{\partial x} + \frac{\partial Q}{\partial y} + \frac{\partial R}{\partial z} = \frac{\partial P_1}{\partial x_1} + \frac{\partial Q_1}{\partial y_1} + \frac{\partial R_1}{\partial z_1}.$$

Дивергенция поля $\vec{A}$ имеет простой физический смысл: если считать, что в некоторой пространственной области $\Omega$ имеется стационарное течение жидкости, скорость которой в каждой точке $(x,y,z)$ равна $\vec{A}(x,y,z)$, то $div\,\vec{A}$ представляет собой производительность в точке $(x,y,z)$ непрерывно распределённых по $\Omega$ источников (или стоков) этой жидкости; условие $div\,\vec{A} = 0$ означает отсутствие источников или стоков.

Пусть $\Omega$ – ограниченная область в $Oxyz$, быть может, многосвязная. Пусть кусочно гладкая поверхность $S = \partial\Omega$ является её полной границей; будем считать, что $S$ ориентирована при помощи поля единичных нормалей $\vec{n}$, направленных **во внешность** области $\Omega$. Пусть функции $P(x,y,z)$, $Q(x,y,z)$, $R(x,y,z)$ непрерывны в замкнутой области $\Omega \cup S$ и имеют непрерывные частные производные первого порядка в (открытой) области $\Omega$. Пусть существуют интегралы по области $\Omega$ от **всех** частных производных первого порядка функций $P$, $Q$, $R$, вообще говоря – **несобственные**. $\vec{A} = \{P, Q, R\}$.

**Формула Остроградского**:

$$\iiint\limits_{\Omega} div\,\vec{A}\,d\Omega = \oiint\limits_{S}\left(\vec{A},\vec{n}\right)dS\,. \qquad (2)$$

В формуле (2) интеграл по $S$ представляет собой сумму интегралов по всем связным компонентам границы $\partial\Omega$ области $\Omega$. Этот интеграл можно записать в виде поверхностного интеграла **второго рода** по внешней стороне поверхности $S$:

$$\oiint\limits_{S}\left(\vec{A},\vec{n}\right)dS = \oiint\limits_{S} Pdydz + Qdzdx + Rdxdy\,,$$

либо в виде интеграла **первого рода**:

$$\oiint\limits_{S}\left(\vec{A},\vec{n}\right)dS = \oiint\limits_{S}\left(P\cdot\cos\alpha + Q\cdot\cos\beta + R\cdot\cos\gamma\right)dS\,.$$

В выбранной декартовой системе координат $Oxyz$

$$\iiint\limits_{\Omega} div\,\vec{A}\,d\Omega = \iiint\limits_{\Omega}\left(\frac{\partial P}{\partial x}+\frac{\partial Q}{\partial y}+\frac{\partial R}{\partial z}\right)dxdydz\,.$$

Формула (2) записана в инвариантном виде.

**Замечание**. Интеграл по $S$ в (2) называется **потоком** вектора $\vec{A}$ через поверхность $S$. Если интерпретировать $\vec{A}$ как скорость течения жидкости в области $\Omega$, то поток – это количество жидкости, пересекающей поверхность $S$ в единицу времени (вытекающей из $\Omega$, если $\vec{n}$ – внешняя нормаль). ♦

**Замечание**. Формулу Грина можно рассматривать как частный случай формулы Остроградского. (Для какой области $\Omega$ и для какого векторного поля $\vec{A}$ ?)♦

**Замечание**. Формула Остроградского лежит в основе вывода 1-ой, 2-ой и 3-ей формул Грина, из которых вытекают все основные свойства гармонических функций трёх переменных. ♦

**Замечание**. Формула Остроградского позволяет выразить объём $|\Omega|$ области $\Omega$, например, в виде

$$|\Omega| = \frac{1}{3}\oiint\limits_{S} xdydz + ydzdx + zdxdy\,.\blacklozenge$$

**Задача**. $S$ – кусочно гладкая поверхность, ограничивающая конечную область. Вычислите интеграл $\oiint\limits_S yzdydz + zxdzdx + xydxdy$. ♦

**Задача**. $S$ – внешняя сторона поверхности куба $0 \le x \le a$, $0 \le y \le a$, $0 \le z \le a$. Вычислите интеграл $\oiint\limits_S x^2dydz + y^2dzdx + z^2dxdy$. ♦

**Задача**. $S$ – внешняя сторона сферы $x^2 + y^2 + z^2 = a^2$. Вычислите интеграл $\oiint\limits_S x^3dydz + y^3dzdx + z^3dxdy$. ♦

**Задача**. Докажите, что для кусочно гладкой поверхности, ограничивающей конечную область, среднее значение единичной нормали к этой поверхности равно нулю. ♦

4. Из формулы Остроградского вытекает важное следствие для **пространственно односвязных** областей $\Omega \subseteq Oxyz$.

Область $\Omega$ в пространстве называется пространственно односвязной, если для любой простой (несамопересекающейся) замкнутой кусочно гладкой ориентируемой поверхности $S$, ограничивающей область $\Omega'$ и целиком лежащей в $\Omega$, область $\Omega'$ также целиком содержится в $\Omega$. Такая область $\Omega$ не содержит дыр, хотя бы и точечных.

**Примеры**. Всё пространство, шар, тор – пространственно односвязные области. Область между двумя концентрическими сферами не является пространственно односвязной. ♦

Векторное поле $\vec{A}$ называется **соленоидальным** в области $\Omega$, если поток этого поля $\iint\limits_S (\vec{A},\vec{n})dS$ через **любую** кусочно гладкую (несамопересекающуюся) поверхность $S$, расположенную в $\Omega$ и ограничивающую область $\Omega'$, равен нулю.

**Теорема**. Пусть в пространственно односвязной области $\Omega \subseteq Oxyz$ задано непрерывно дифференцируемое векторное поле интеграл по её границе: $\iint\limits_{M}\left(\frac{\partial Q}{\partial x}-\frac{\partial P}{\partial y}\right)dxdy=\oint\limits_{\partial M}Pdx+Qdy$ (контур $\partial M$ ориентирован так, что при его обходе область $M$ остаётся слева – обход против часовой стрелки);

– формула Стокса $\iint\limits_{M}\left(rot\,\vec{A},\vec{n}\right)dS=\oint\limits_{\partial M}\left(\vec{A},\vec{t}\right)dl$;

– формула Остроградского $\iiint\limits_{M}div\,\vec{A}\,d\Omega=\oiint\limits_{\partial M}\left(\vec{A},\vec{n}\right)dS$.

$\vec{A}(x, y, z)$. Чтобы поле $\vec{A}$ было соленоидальным в $\Omega$, необходимо и достаточно, чтобы в $\Omega$ выполнялось тождество $div\,\vec{A} \equiv 0$. ♦

Всякое соленоидальное поле в пространственно односвязной области можно представить в виде ротора некоторого векторного поля: если $div\,\vec{A} \equiv 0$, то существует такое поле $\vec{B}$, что $\vec{A} = rot\,\vec{B}$. При этом $\vec{B}$ называется вектор-потенциалом поля $\vec{A}$; вектор-потенциал определён не однозначно, а с точностью до произвольного слагаемого вида $grad\,u(x, y, z)$.

**Теорема Гельмгольца**. Пусть в каждой точке области $\Omega$ для векторного поля $\vec{A}$ определены $div\,\vec{A}$ и $rot\,\vec{A}$. Тогда в $\Omega$ поле $\vec{A} = \vec{C} + \vec{G}$, где поле $\vec{C}$ соленоидально, а поле $\vec{G}$ потенциально. ♦

5. Формулы (1) и (2) допускают далеко идущее обобщение на многообразия произвольной конечной размерности. (Для формулировки этого обобщения требуется ввести понятия **дифференциальной формы** и **внешнего дифференциала**.) Интеграл от внешнего дифференциала дифференциальной формы по ориентированному компактному многообразию $M$ равен интегралу от самой формы по краю $\partial M$ этого многообразия, который ориентирован согласованно с ориентацией $M$. Частными случаями этого утверждения являются известные теоремы математического анализа:

– формула Ньютона-Лейбница, выражающая значение определённого интеграла от заданной функции $f$ по отрезку через значения любой её первообразной $F$ на концах этого отрезка: $\int\limits_a^b f(x)dx = F(b) - F(a)$ ($M$ –одномерное многообразие $a \le x \le b$, краем $\partial M$ которого являются две точки $x = a$, $x = b$);

– формула Грина, связывающая двойной интеграл по ограниченной области $M$ на плоскости $Oxy$ и криволинейный

## Вопрос 3.

## Почленное интегрирование и дифференцирование функциональных рядов.

1. Пусть на множестве $X$ определена последовательность действительнозначных или комплекснозначных функций $\{u_k(x)\}$; она задаёт другую последовательность функций $f_n(x)=\sum_{k=1}^{n}u_k(x)$, которые называются **частичными суммами** функционального ряда с общим членом $u_k(x)$. Под сходимостью функционального ряда

$$\sum_{k=1}^{\infty}u_k(x) \qquad (1)$$

к сумме $u(x)$ **в том или ином смысле** на множестве $X$ (поточечной, равномерной, в среднем квадратическом и др.) понимают соответствующую сходимость его частичных сумм.

С другой стороны, любую последовательность определённых на $X$ функций $\{f_n(x)\}$ можно рассматривать как последовательность частичных сумм некоторого функционального ряда, если ввести $u_k(x)=f_k(x)-f_{k-1}(x)$ при $k>1$. Поэтому одни и те же утверждения можно формулировать и в терминах функциональных последовательностей, и в терминах функциональных рядов.

2. **Теорема 1**. Пусть функциональный ряд (1) сходится к своей сумме $u(x)$ **равномерно** на отрезке $[a,b]$. Пусть каждая функция $u_k(x)$ интегрируема по Риману на отрезке $[a,b]$. Тогда и сумма ряда $u(x)$ интегрируема по Риману на $[a,b]$, причём ряд (1) можно интегрировать на $[a,b]$ почленно, т.е. числовой ряд

$\sum_{k=1}^{\infty}\int_a^b u_k(x)dx$ сходится, и его сумма равна $\int_a^b u(x)dx$:

$$\int_a^b\left(\sum_{k=1}^{\infty}u_k(x)\right)dx=\sum_{k=1}^{\infty}\int_a^b u_k(x)dx\,.\blacklozenge$$

**Замечание**. Требование **равномерной** сходимости ряда (1) не является необходимым для возможности почленно интегрировать ряд (1). ♦

**Задача**. Сходится ли равномерно на отрезке $[0,1]$ функциональный ряд $\sum_{k=1}^{\infty}\left[\frac{kx}{1+k^2x^2}-\frac{(k-1)x}{1+(k-1)^2x^2}\right]$? Можно ли его проинтегрировать почленно на $[0,1]$?♦

**Задача**. Сходится ли равномерно на отрезке $[0,1]$ функциональный ряд $\sum_{k=1}^{\infty}2x\left[k^2e^{-k^2x^2}-(k-1)^2e^{-(k-1)^2x^2}\right]$? Можно ли его проинтегрировать почленно на $[0,1]$?♦

**Задача**. Возможно ли почленное интегрирование на отрезке $[0,1]$ функционального ряда $\sum_{k=1}^{\infty}\left[x^{\frac{1}{2k+1}}-x^{\frac{1}{2k-1}}\right]$? ♦

**Задача**. Переформулируйте теорему 1 в терминах функциональных последовательностей. ♦

**Задача**. На отрезке $[0,1]$ определены функции

$$f_n(x)=\begin{cases}2n\alpha_n x, & 0\le x\le\dfrac{1}{2n},\\ 2n\alpha_n\left(\dfrac{1}{n}-x\right), & \dfrac{1}{2n}\le x\le\dfrac{1}{n},\\ 0, & \dfrac{1}{n}\le x\le 1,\end{cases}$$

где $\{\alpha_n\}$ – заданная последовательность чисел. Какому необходимому и достаточному условию должна удовлетворять

$\{\alpha_n\}$, чтобы $\{f_n(x)\}$ равномерно сходилась на $[0,1]$? Какому необходимому и достаточному условию должна удовлетворять $\{\alpha_n\}$, чтобы последовательность $\{f_n(x)\}$ можно было почленно интегрировать на $[0,1]$?♦

**Задача**. Найдите все значения параметра $\alpha$, при которых функциональная последовательность $f_n(x) = n^\alpha xe^{-nx}$

– поточечно сходится на $[0,1]$;
– равномерно сходится на $[0,1]$;
– Может быть проинтегрирована почленно $\left(\lim\limits_{n\to\infty}\int\limits_0^1 f_n(x)dx = \int\limits_0^1\left(\lim\limits_{n\to\infty} f_n(x)\right)dx\right)$.♦

**Задача**. Сходится ли функциональная последовательность $f_n(x) = nxe^{-nx^2}$ на $[0,1]$ поточечно? Сходится ли она на $[0,1]$ равномерно? Можно ли её интегрировать почленно на $[0,1]$?♦

**Задача**. Сходится ли функциональная последовательность $f_n(x) = nx(1-x)^n$ на $[0,1]$ поточечно? Сходится ли она на $[0,1]$ равномерно? Можно ли её интегрировать почленно на $[0,1]$?♦

**Задача**. Можно ли интегрировать почленно на отрезке $[0,1]$ функциональную последовательность

$$f_n(x) = \begin{cases} 2n^2x, & 0 \le x \le \dfrac{1}{2n}, \\ n - 2n^2\left(x - \dfrac{1}{2n}\right), & \dfrac{1}{2n} \le x \le \dfrac{1}{n}, \\ 0, & \dfrac{1}{n} \le x \le 1 \end{cases} \qquad ?$$

**Задача**. Можно ли интегрировать почленно на отрезке $[0,b]$ $(0 < b \le 1)$ функциональную последовательность

$$f_n(x)=\begin{cases} 2n^3x, & 0\le x\le \dfrac{1}{2n}, \\ n^2-2n^3\left(x-\dfrac{1}{2n}\right), & \dfrac{1}{2n}\le x\le \dfrac{1}{n}, \\ 0, & \dfrac{1}{n}\le x\le 1 \end{cases} \quad ?$$

**Задача**. На отрезке $[0,1]$ определены функции

$$f_n(x)=\begin{cases} 1, \textit{ если } x - \textit{рациональное число}, \\ \textit{выражаемое несократимой дробью } \dfrac{m}{n}, \\ 0 \quad \textit{в противном случае}, \end{cases}$$

$n=1,2,\dots$ . Интегрируемы ли функции $f_n$ на $[0,1]$ по Риману? Каков поточечный предел последовательности $\{f_n(x)\}$? Сходится ли эта последовательность равномерно на $[0,1]$? Интегрируема ли по Риману её предельная функция? ♦

**Задача**. Является ли равномерно сходящейся **на полупрямой** $[0,+\infty)$ функциональная последовательность

$$f_n(x)=\begin{cases} \dfrac{1}{n}, & 0\le x\le n, \\ 0, & x>n \end{cases} \quad ?$$

Можно ли её проинтегрировать почленно на $[0,+\infty)$? ♦

**Задача**. Является ли равномерно сходящейся **на полупрямой** $[0,+\infty)$ функциональная последовательность

$$f_n(x)=\begin{cases} \dfrac{1}{n}, & 0\le x\le n^2, \\ 0, & x>n^2 \end{cases} \quad ?$$

Можно ли её проинтегрировать почленно на $[0,+\infty)$? ♦

3. **Теорема 2**. Пусть каждая функция $u_k(x)$ имеет производную на отрезке $[a,b]$. Пусть **ряд из производных** $\sum_{k=1}^{\infty} u'_k(x)$ **сходится равномерно** на $[a,b]$, а **сам ряд** $\sum_{k=1}^{\infty} u_k(x)$ **сходится хотя бы в одной точке** отрезка $[a,b]$. Тогда функциональный ряд $\sum_{k=1}^{\infty} u_k(x)$ сходится равномерно на $[a,b]$, а его сумма $u(x)$ имеет производную на $[a,b]$, которую можно вычислить, дифференцируя ряд почленно: $\left(\sum_{k=1}^{\infty} u_k(x)\right)' = \sum_{k=1}^{\infty} u'_k(x)$.

**Замечание**. Требование **равномерной** сходимости ряда $\sum_{k=1}^{\infty} u'_k(x)$ не является необходимым для возможности почленно дифференцировать ряд (1). ♦

**Задача**. Можно ли дифференцировать почленно на отрезке $[0,1]$ функциональный ряд

$$\frac{1}{2}\ln\left(1+x^2\right)+\sum_{k=2}^{\infty}\left[\frac{1}{2k}\ln\left(1+k^2x^2\right)-\frac{1}{2(k-1)}\ln\left(1+(k-1)^2x^2\right)\right]?\blacklozenge$$

**Задача**. Можно ли дифференцировать почленно на отрезке $[0,1]$ функциональный ряд $\sum_{k=1}^{\infty}\left[e^{-(k-1)^2x^2}-e^{-k^2x^2}\right]$ ? ♦

**Задача**. Можно ли дифференцировать почленно на $(-\infty,+\infty)$ функциональный ряд $\sum_{k=1}^{\infty} arctg\,\frac{x}{k^2}$ ? ♦

**Задача**. Переформулируйте теорему 2 в терминах функциональных последовательностей. ♦

**Задача**. Сходится ли равномерно на отрезке $[-1,1]$ функциональная последовательность $f_n(x)=\frac{x}{1+n^2x^2}$ ? Можно ли дифференцировать её почленно на $[-1,1]$ ? ♦

**Задача**. Сходится ли равномерно на $(-\infty,+\infty)$ функциональная последовательность $f_n(x)=\dfrac{\sin(nx)}{\sqrt{n}}$ ? Можно ли продифференцировать её почленно при каком-либо $x$ ? ♦

**Задача**. Сходится ли равномерно на $(-\infty,+\infty)$ функциональная последовательность

$f_n(x)=x^2+\dfrac{1}{n}\sin\left[n\left(x+\dfrac{\pi}{2}\right)\right]$ ? Можно ли дифференцировать её почленно на $(-\infty,+\infty)$? ♦

Из теоремы 2 вытекает

**Теорема 3**. Пусть функциональный ряд (1) сходится к своей сумме $u(x)$ **равномерно** на отрезке $[a,b]$. Пусть каждая функция $u_k(x)$ имеет первообразную на $[a,b]$. Тогда и $u(x)$ имеет первообразную на отрезке $[a,b]$, причём, если $x_0\in[a,b]$, то ряд $\sum\limits_{k=1}^{\infty}U_k(x)$ из первообразных $U_k(x)$ функций $u_k(x)$, которые удовлетворяют условиям $U_k(x_0)=0$, $\quad k=1,2,\ldots,$ сходится равномерно на $[a,b]$ к первообразной $U(x)$ суммы $u(x)$, удовлетворяющей условию $U(x_0)=0$.

4. Пусть все функции $f_n(x)$, $n=1,2,\ldots,$ и функция $f(x)$ интегрируемы по Риману на отрезке $[a,b]$. Говорят, что последовательность функций $\{f_n(x)\}$ сходится **в среднем квадратическом** к функции $f(x)$ на $[a,b]$, если $\lim\limits_{n\to\infty}\int\limits_a^b|f_n(x)-f(x)|^2\,dx=0$. Говорят, что функциональный ряд (1) сходится к функции $u(x)$ на $[a,b]$ **в среднем квадратическом**, если последовательность его частичных сумм сходится в среднем квадратическом к $u(x)$ на $[a,b]$.

**Замечание**. Из равномерной сходимости последовательности $\{f_n(x)\}$ к функции $f(x)$ на $[a,b]$ вытекает сходимость $\{f_n(x)\}$ к $f(x)$ в среднем квадратическом на $[a,b]$ (докажите!). Но сходимость в среднем квадратическом не влечёт не только равномерной сходимости, но и сходимости последовательности $\{f_n(x)\}$ хотя бы в одной точке $x$ (приведите пример!). ♦

Сходимость ряда (1) к функции $u(x)$ на $[a,b]$ в среднем квадратическом не подразумевает даже существования поточечного предела его частичных сумм, но гарантирует возможность почленно интегрировать ряд (1).

**Теорема 4**. Пусть все функции $u_k(x)$ интегрируемы по Риману на отрезке $[a,b]$ и функция $u(x)$ интегрируема по Риману на $[a,b]$. Пусть функциональный ряд (1) сходится к $u(x)$ в среднем квадратическом на $[a,b]$. Тогда ряд (1) можно интегрировать на $[a,b]$ почленно, т.е. числовой ряд $\sum_{k=1}^{\infty}\int_a^b u_k(x)dx$ сходится, и его сумма равна $\int_a^b u(x)dx$. ♦

**Пример**. Тригонометрический ряд Фурье любой кусочно непрерывной на отрезке $[-\pi,\pi]$ функции сходится к ней на этом отрезке в среднем квадратическом; поэтому тригонометрический ряд Фурье такой функции можно почленно интегрировать на $[-\pi,\pi]$. Это вытекает из замкнутости тригонометрической системы функций. ♦

**Задача**. Пусть все функции $u_k(x)$ непрерывно дифференцируемы на отрезке $[a,b]$. Пусть ряд (1) сходится на $[a,b]$, а ряд $\sum_{k=1}^{\infty}u_k'(x)$ сходится в среднем квадратическом к непрерывной функции на $[a,b]$. Докажите, что сумма ряда (1) непрерывно дифференцируема, и её производную можно найти почленным дифференцированием ряда (1). ♦

5. **Теорема 5**. Пусть функция $u(x)$ и все её производные вплоть до порядка $m \ge 1$ непрерывны на отрезке $[-\pi, \pi]$ и удовлетворяют условиям $u(-\pi) = u(\pi)$, $u'(-\pi) = u'(\pi)$, $\ldots, u^{(m)}(-\pi) = u^{(m)}(\pi)$. Кроме того, пусть $u(x)$ имеет на $[-\pi, \pi]$ кусочно непрерывную производную порядка $m+1$. Тогда тригонометрический ряд Фурье функции $u(x)$ можно $m$ раз дифференцировать почленно на $[-\pi, \pi]$.

**Доказательство** вытекает из того, что в условиях теоремы при целом $m \ge 0$ сходится числовой ряд $\sum\limits_{k=1}^{\infty} k^m \left(|\alpha_k| + |\beta_k|\right)$, где $\alpha_k$ и $\beta_k$ – коэффициенты Фурье функции $u(x)$ по тригонометрической системе функций на $[-\pi, \pi]$ (докажите этот факт!). ♦

**Задача**. Пусть функция $u(x)$ на $[0, l]$ имеет непрерывную $m$-ую производную и кусочно непрерывную $(m+1)$-ую производную, причём $u^{(i)}(0) = u^{(i)}(l) = 0$ для всех **чётных** $i$, $0 \le i \le m$. Докажите, что ряд Фурье функции $u(x)$ по системе функций $\left\{\sin\dfrac{\pi k}{l} x\right\}$, $k = 1, 2, \ldots$, на $[0, l]$ можно $m$ раз дифференцировать почленно. ♦

6. **Теорема 6**. Пусть степенной ряд ($x$ – действительная переменная)

$$\sum_{k=0}^{\infty} c_k x^k \tag{2}$$

имеет радиус сходимости $R > 0$. Если $|x| < R$, то ряд (2) можно почленно интегрировать на отрезке $[0, x]$:

$$\int_0^x \left( \sum_{k=0}^{\infty} c_k \xi^k \right) d\xi = \sum_{k=0}^{\infty} \frac{c_k}{k+1} x^{k+1} \; ; \tag{3}$$

при этом степенной ряд в правой части (3) имеет тот же радиус сходимости, что и исходный ряд (2). На интервале $(-R, R)$ ряд (2) можно дифференцировать почленно:

$$\left( \sum_{k=0}^{\infty} c_k x^k \right)' = \sum_{k=1}^{\infty} k c_k x^{k-1} \; ; \tag{4}$$

при этом степенной ряд в правой части (4) имеет тот же радиус сходимости, что и исходный ряд (2). ♦

**Замечание**. Если ряд (2) сходится в каком-либо из концов интервала сходимости $(-R, R)$, то в (3) $x$ может совпадать с этим концом. Если ряд в правой части (4) сходится в каком-либо из концов интервала $(-R, R)$, то в (4) $x$ может совпадать с этим концом.♦

**Замечание**. Внутри интервала $(-R, R)$ ряд (2) можно интегрировать и дифференцировать почленно **любое** число раз. ♦

**Замечание**. Степенной ряд ( $z$ – **комплексная** переменная) $\sum_{k=0}^{\infty} c_k z^k$ внутри **круга сходимости** $\left\{ z \,\middle|\, |z| < R \right\}$ можно почленно интегрировать и дифференцировать любое число раз. Радиус сходимости получаемых при этом степенных рядов равен радиусу сходимости исходного ряда. ♦

**Задача**. Пользуясь формулой геометрической прогрессии при $-1 < x < 1$, напишите разложение функции $\dfrac{1}{1+x}$ в ряд по степеням переменной $x$. Проинтегрировав этот ряд почленно, получите разложение в степенной ряд функции $\ln(1+x)$. Сходится ли этот ряд при $x = -1$? При $x = 1$?♦

**Задача**. Напишите разложение функции $\dfrac{1}{1+x^2}$ в ряд по степеням переменной $x$. Проинтегрировав этот ряд почленно, постройте разложение в степенной ряд функции $arctg\, x$. Какова область сходимости полученного ряда? Сходится ли он при $x=-1$? При $x=1$?♦

**Задача**. Напишите разложение функции $\dfrac{1}{\sqrt{1-x^2}}$ в ряд по степеням переменной $x$. Проинтегрировав этот ряд почленно, постройте разложение в степенной ряд функции $\arcsin x$. Какова область сходимости полученного ряда? Сходится ли он при $x=-1$? При $x=1$?♦

**Задача**. Пользуясь разложением в ряд по степеням переменной $x$ функции $\dfrac{1}{\sqrt{1+x^2}}$, постройте разложение в степенной ряд функции $\ln\left(x+\sqrt{1+x^2}\right)$. Какова его область сходимости? ♦

**Задачи**. Постройте разложения заданных функций в ряды по степеням переменной $x$ и найдите области сходимости этих степенных рядов:

$$\int_0^x e^{-\xi^2}d\xi, \quad \int_0^x \frac{\sin\xi}{\xi}d\xi, \quad \int_0^x \frac{arctg\,\xi}{\xi}d\xi, \quad \int_0^x \frac{\ln(1+\xi)}{\xi}d\xi .♦$$

**Задачи**. С помощью почленного интегрирования или почленного дифференцирования найдите суммы данных степенных рядов:

$$\sum_{k=1}^{\infty}(-1)^{k-1}k^2x^k, \qquad \sum_{k=1}^{\infty}\frac{(-1)^{k-1}x^{2k}}{k(2k-1)}, \qquad \sum_{k=1}^{\infty}k^2x^{k-1} .♦$$

7. Теоремы о почленном интегрировании и о почленном дифференцировании можно существенно обобщить, если интегрирование понимать **в смысле Лебега**, а дифференцирование допускать **почти всюду**.

Мы будем рассматривать множество $X$ с заданной на нём $\sigma$-аддитивной мерой $\mu$, определённой на $\sigma$-алгебре множеств из $X$, единицей которой является $X$; $\mu(X)<\infty$. Меру $\mu$ будем считать **полной** (т.е. из условий $\mu(Y)=0$ и $Z\subset Y$ вытекает, что множество $Z$ $\mu$-измеримо и $\mu(Z)=0$). Если же $X$ имеет бесконечную меру $\mu$, то будем предполагать его представимым в виде **счётного** объединения возрастающей последовательности множеств конечной меры (например, таким множеством является числовая прямая с линейной мерой Лебега на ней); в этом случае мера $\mu$ называется $\sigma$-конечной. Действительнозначная функция $f(x)$ на $X$ называется $\mu$-измеримой, если для всякого борелевского множества $B$ числовой прямой множество $f^{-1}(B)$ $\mu$-измеримо.

**Теорема Лебега** об ограниченной сходимости. Пусть $X$ – измеримое множество конечной или бесконечной меры $\mu$. Пусть на $X$ задана последовательность $\mu$-измеримых функций $\{f_n(x)\}$, которая сходится к функции $f(x)$ всюду на $X$ (либо **почти всюду** на $X$; либо **по мере** $\mu$ на $X$). Пусть существует такая определённая на $X$ и интегрируемая на $X$ по мере $\mu$ функция $\varphi(x)$, что при всех $n$ всюду на $X$ (либо **почти всюду** на $X$) выполнены неравенства $|f_n(x)| \leq \varphi(x)$. Тогда все функции $f_n$ и предельная функция $f$ интегрируемы на множестве $X$ по мере $\mu$ и

$$\lim_{n\to\infty}\int_X f_n(x)d\mu = \int_X f(x)d\mu \,.\blacklozenge \tag{5}$$

Важный частный случай теоремы Лебега получим, если последовательность $f_n$ равномерно ограничена: $|f_n(x)| \leq const$ при всех $n$, а множество $X$ имеет конечную меру $\mu$.

**Замечание**. Наличие мажорантной функции $\varphi$ не является необходимым для выполнения (5) (приведите пример!). ♦

**Задача**. Переформулируйте теорему Лебега в терминах функциональных рядов. ♦

**Теорема Б. Леви** о монотонной сходимости. Пусть $X$ – измеримое множество конечной или бесконечной меры $\mu$. Пусть на $X$ задана последовательность функций $\{f_n(x)\}$, для которой всюду на $X$ (либо **почти всюду** на $X$) выполнены неравенства $f_1(x) \leq f_2(x) \leq \ldots \leq f_n(x) \leq \ldots$ . Пусть все функции $f_n$

интегрируемы на $X$ по мере $\mu$, а их интегралы **ограничены в совокупности**: $\int\limits_X f_n(x)d\mu \le M = const$. Тогда почти всюду на $X$ существует конечный предел $f(x) = \lim\limits_{n\to\infty} f_n(x)$, функция $f(x)$ интегрируема на $X$ по мере $\mu$, и $\lim\limits_{n\to\infty}\int\limits_X f_n(x)d\mu = \int\limits_X f(x)d\mu$. ♦

**Замечание**. Нельзя отказаться от ограниченности в совокупности интегралов $\int\limits_X f_n(x)d\mu$: без этого предположения даже если существует конечный предел $f(x) = \lim\limits_{n\to\infty} f_n(x)$ почти всюду на $X$, функция $f(x)$ может оказаться неинтегрируемой (приведите пример!). Но если $f(x)$ интегрируема, то $\int\limits_X f_n(x)d\mu$ ограничены в совокупности (докажите!). ♦

**Задача**. Пусть все $u_k(x) \ge 0$ почти всюду на множестве $X$ и интегрируемы по мере $\mu$ на $X$. Пусть почти всюду на $X$ функциональный ряд (1) сходится к функции $u(x)$, которая интегрируема по мере $\mu$ на $X$. Докажите, что сходится числовой ряд $\sum\limits_{k=1}^{\infty}\int\limits_X u_k(x)d\mu$ и выполнено равенство $\int\limits_X u(x)d\mu = \sum\limits_{k=1}^{\infty}\int\limits_X u_k(x)d\mu$. ♦

**Задача**. Пусть $\{f_n(x)\}$ – последовательность интегрируемых на множестве $X$ конечной или бесконечной меры, почти всюду неотрицательных функций. Пусть существует такое число $M$, что для всех $n$ выполнены неравенства $\int\limits_X f_n(x)d\mu \le M$. Докажите, что функция $f(x) = \varliminf\limits_{n\to\infty} f_n(x)$ интегрируема, причём $\int\limits_X f(x)d\mu \le \varliminf\limits_{n\to\infty}\int\limits_X f_n(x)d\mu$. Обязательно ли при этом функция $\overline{\lim_{n\to\infty}} f_n(x)$ интегрируема? Обязательно ли выполнено равенство $\int\limits_X f(x)d\mu = \overline{\lim_{n\to\infty}} \int\limits_X f_n(x)d\mu$ ?♦

Напомним, что определённая на отрезке $[a,b]$ **монотонная** функция имеет почти всюду (в смысле линейной меры Лебега на прямой) на $[a,b]$ конечную производную, – теорема Лебега.

**Теорема («малая» теорема Фубини)**. Пусть все $u_k(x)$ – **монотонно неубывающие** на отрезке $[a,b]$ функции, а функциональный ряд (1) сходится всюду на $[a,b]$. Тогда почти всюду (в смысле меры Лебега) на $[a,b]$ ряд (1) допускает почленное дифференцирование: $\left(\sum\limits_{k=1}^{\infty} u_k(x)\right)' = \sum\limits_{k=1}^{\infty} u_k'(x)$.♦

## Вопрос 4.

## Формула Тейлора с остаточным членом в форме Лагранжа. Разложение элементарных функций.

1. Если функция $f(x)$ одной действительной переменной $n$ раз дифференцируема в точке $x = a$, то ей можно сопоставить многочлен

$$P_n(x,a) = \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(x-a)^k ,$$

который называется многочленом Тейлора функции $f$, и попытаться аппроксимировать функцию $f$ в некоторой окрестности точки $a$ её многочленом Тейлора.

**Теорема Тейлора**. Пусть функция $f(x)$ в окрестности точки $a$ имеет конечную производную $(n+1)$-го порядка. Пусть $x$ – любое значение аргумента функции $f$ из этой окрестности. Тогда для любого числа $p > 0$ между точками $a$ и $x$ найдётся такая точка $\xi$, что $f(x) = P_n(x,a) + R_{n+1}(x)$, где

$$R_{n+1}(x) = \left(\frac{x-a}{x-\xi}\right)^p \cdot \frac{(x-\xi)^{n+1}}{n!\,p} \cdot f^{(n+1)}(\xi). \qquad (1)$$

**Доказательство** состоит в обосновании формулы (1) для разности $R_{n+1}(x) = f(x) - P_n(x,a)$. Пусть для определённости $x > a$, и переменная $y$ изменяется на отрезке $[a,x]$. Функция $g(y) = f(x) - P_n(x,y) - \dfrac{(x-y)^p}{(x-a)^p} R_{n+1}(x)$ непрерывна и дифференцируема (по $y$) на $[a,x]$, поскольку функция $f(y)$ и все её производные до $n$-го порядка непрерывны на $[a,x]$, и на этом отрезке существует конечная $f^{(n+1)}(y)$. Кроме того, $g(a) = g(x) = 0$ (проверьте!). По теореме Ролля внутри отрезка $[a,x]$ найдётся такая точка $\xi$, что $g'(\xi) = 0$. Но на этом отрезке

$$g'(y) = -\frac{f^{(n+1)}(y)}{n!}(x-y)^n + p\cdot\frac{(x-y)^{p-1}}{(x-a)^p}R_{n+1}(x)$$

(проверьте!). Отсюда и получаем формулу (1). ♦

Остаточный член $R_{n+1}$ в (1) записан в форме Шлёмильха-Роша; в нём $\xi$ зависит от $p$. Поскольку $\xi$ лежит между $a$ и $x$, найдётся такое число $\theta$, $0<\theta<1$, что $\xi = a+\theta\cdot(x-a)$. Тогда (1) примет вид

$$R_{n+1}(x) = \frac{(x-a)^{n+1}}{n!\,p}\cdot(1-\theta)^{n-p+1}\cdot f^{(n+1)}(a+\theta\cdot(x-a)). \qquad (2)$$

Чаще всего остаточный член записывают **в форме Лагранжа**, которая получается из (2) при $p=n+1$:

$$R_{n+1}(x) = \frac{(x-a)^{n+1}}{(n+1)!}\cdot f^{(n+1)}(a+\theta\cdot(x-a)). \qquad (3)$$

Если в (2) выбрать $p=1$, то получим остаточный член **в форме Коши**:

$$R_{n+1}(x) = \frac{(x-a)^{n+1}}{n!}\cdot(1-\theta)^n\cdot f^{(n+1)}(a+\theta\cdot(x-a)). \quad (4)$$

(Значения $\theta$ в (3) и (4), вообще, различны.)

**Замечание**. При $n=0$ формула Тейлора с остаточным членом в форме Лагранжа или Коши является формулой Лагранжа: $f(x)-f(a)=(x-a)\cdot f'(a+\theta\cdot(x-a))$, $0<\theta<1$.♦

**Замечание**. Если $f^{(n+1)}$ непрерывна в окрестности точки $a$ (или $f^{(n)}$ непрерывна, а $f^{(n+1)}$ кусочно непрерывна), то формулу Тейлора можно записать с остаточным членом **в интегральной форме**:

$$R_{n+1}(x) = \frac{1}{n!}\int_a^x f^{(n+1)}(y)\cdot(x-y)^n\,dy. \qquad (5)$$

Для этого интеграл в (5) надо последовательно вычислять по частям (проведите вычисления); можно вместо этого последовательно вычислять по частям $\int_a^x f(y)\cdot\frac{d^{n+1}}{dy^{n+1}}(x-y)^n\,dy$ (проведите вычисления). Остаточный член (5), в отличие от (1) – (4), не содержит неизвестных чисел. Если $f^{(n+1)}$ непрерывна, то из (5) можно вывести, например, остаточный член в форме Лагранжа: поскольку множитель $(x-y)^n$ под знаком интеграла в (5) не меняет знака, к этому интегралу можно применить обобщённую формулу среднего значения. Остаточный член (5) удобно записывать в виде

$$R_{n+1}(x)=\frac{(x-a)^{n+1}}{n!}\cdot\int_0^1(1-t)^n\cdot f^{(n+1)}(a+(x-a)\cdot t)dt\,. \qquad (6)$$

Если $f^{(n+1)}$ непрерывна, то и $R_{n+1}(x)$ – непрерывная функция от $x$; если $f$ имеет непрерывные производные более высокого порядка $f^{(n+k)}$, $k>1$, то $R_{n+1}(x)$ можно дифференцировать $k-1$ раз при помощи дифференцирования по $x$ под знаком интеграла в (6). ♦

**Теорема** об остаточном члене в форме Пеано. Пусть функция $f(x)$ в окрестности точки $a$ имеет конечные производные до порядка $n-1$ и конечную производную порядка $n$ в самой точке $a$. Тогда $f(x)-P_n(x,a)\equiv R_{n+1}(x)=o((x-a)^n)$ при $x\to a$.♦

**Задача**. Предполагая, что функция $f$ имеет **непрерывную** производную порядка $n$ в точке $a$, докажите утверждение последней теоремы, исходя из остаточного члена $R_n(x)$ в форме Лагранжа. ♦

Значения многочлена Тейлора $P_n(x,a)$ и его производных до $n$-го порядка в точке $a$ совпадают со значениями функции $f(x)$ и её соответствующих производных в точке $a$: $f^{(k)}(a)=P_n^{(k)}(a,a)$, $k=0,1,\ldots,n$. Многочлен Тейлора является многочленом наилучшего приближения функции $f$ при $x\to a$ в том смысле, что $f(x)-P_n(x,a)=o((x-a)^n)$ при $x\to a$, и если некоторый многочлен $Q(x)$ степени, **не превышающей** $n$, обладает свойством $f(x)-Q(x)=o\left((x-a)^m\right)$ при $x \to a$, где $m \geq n$, то $Q$ совпадает с $P_n$. Если в точке $a$ **хотя бы одна** из производных $f^{(k)}(a) \neq 0$, $k=0,1,\ldots,n$, то многочлен Тейлора является главной частью формулы Тейлора.

2. Формулу Тейлора с центром в точке $a=0$ называют формулой Маклорена.

**Задачи**. Напишите разложения следующих элементарных функций по формуле Маклорена с остаточным членом $R_{n+1}(x)$ в форме Лагранжа. Укажите оценку остаточного члена на заданном промежутке изменения $x$. Стремится ли $R_{n+1}(x)$ к нулю при $n \to \infty$?

$1^0$. $e^x$ на отрезке $[-r,r]$, $r$ – любое положительное число.

$2^0$. $\sin x$ на отрезке $[-r,r]$, $r$ – любое положительное число.

$3^0$. $\cos x$ на отрезке $[-r,r]$, $r$ – любое положительное число.

$4^0$. $\ln(1+x)$ на отрезке $[0,1]$. То же на отрезке $[-r,0]$, $r$ – произвольное число из интервала $0<r<1$ (для оценки $R_{n+1}(x)$ запишите этот остаточный член в форме Коши). Стремится ли $R_{n+1}(x)$ к нулю, если $x>1$? Проведите оценку остаточного члена, записав его в интегральной форме.

$5^0$. $(1+x)^\alpha$, $\alpha$ – действительное число. В промежутке $0 \leq x < 1$ для оценки $R_{n+1}(x)$ при достаточно больших $n$ используйте остаточный член в форме Лагранжа. В промежутке $-1<x<0$ для оценки $R_{n+1}(x)$ используйте остаточный член в форме Коши. Будет ли остаточный член стремиться к нулю при $n \to \infty$, если $|x|>1$? Каким будет $R_{n+1}(x)$, если $\alpha = n$ – натуральное число? ♦

**Задачи**. Напишите **асимптотические** (при $x \to 0$) разложения следующих элементарных функций по формуле

Маклорена с остаточным членом **в форме Пеано**: $e^x$, $\sin x$, $\cos x$, $\ln(1+x)$, $arctg\, x$, $\dfrac{1}{1-x^2}$. ♦

**Задачи**. Напишите асимптотические (при $x \to 0$) разложения по формуле Маклорена с остаточным членом в форме Пеано:

$1^0$. $tg\, x$ с точностью до $o(x^4)$.

$2^0$. $e^{\sin x}$ с точностью до $o(x^3)$.

$3^0$. $e^{tg\, x}$ с точностью до $o(x^3)$.

$4^0$. $\ln \cos x$ с точностью до $o(x^6)$.

$5^0$. $\ln\left(x + \sqrt{1+x^2}\right)$ с точностью до $o(x^5)$.

$6^0$. $\ln \dfrac{\sin x}{x}$ с точностью до $o(x^6)$.

$7^0$. $\sin(\sin x)$ с точностью до $o(x^6)$. ♦

**Задачи**. Пользуясь формулой Маклорена, найдите следующие пределы:

$$\lim_{x\to 0} \frac{x - \sin x}{x^3}; \quad \lim_{x\to 0} \frac{\cos x - e^{-\frac{x^2}{2}}}{x^3 \sin x}; \quad \lim_{x\to 0} \frac{e^x \cdot \sin x - x \cdot (1+x)}{x^3};$$

$$\lim_{x\to 0} \frac{\sin(\sin x) - x \cdot \left(1 - x^2\right)^{\frac{1}{3}}}{x^5}; \quad \lim_{x\to 0} \frac{1 - (\cos x)^{\sin x}}{x^3}. \ ♦$$

3. Формула Тейлора со всеми формами остаточного члена (1) – (6) обобщается на случай функций нескольких действительных переменных.

**Теорема**. Пусть функция $u = f(x_1, \dots, x_m)$ определена и $(n+1)$ раз дифференцируема в некотором открытом шаре с центром в точке $A\left(x_1^0, \dots, x_m^0\right)$. Тогда полное приращение этой функции в точке $A$ для любой точки из указанного шара можно представить в виде

$$f(x_1, \dots, x_m) - f\left(x_1^0, \dots, x_m^0\right) =$$

$$= du\ \Big|_A + \frac{1}{2!}d^2u\ \Big|_A + \ldots + \frac{1}{n!}d^n u\ \Big|_A + \frac{1}{(n+1)!}d^{n+1}u\ \Big|_B, \qquad (7)$$

где все дифференциалы переменных $dx_i = x_i - x_i^0$, а точка $B$ – некоторая точка указанного шара, $B = B(A)$.♦

Формулу (7) можно символически записать в виде

$$f(x_1, \ldots, x_m) - f(x_1^0, \ldots, x_m^0) =$$

$$= \sum_{k=1}^{n} \frac{1}{k!}\left[\sum_{i=1}^{m}(x_i - x_i^0)\frac{\partial}{\partial x_i}\right]^k f(x_1^0, \ldots, x_m^0) + R_{n+1},$$ где

$$R_{n+1} = \frac{1}{(n+1)!}\left[\sum_{i=1}^{m}(x_i - x_i^0)\frac{\partial}{\partial x_i}\right]^{n+1} f(x_1^0 + \theta dx_1, \ldots, x_m^0 + \theta dx_m)$$

– остаточный член в форме Лагранжа; $dx_i = x_i - x_i^0$, $0 < \theta < 1$.

**Задача**. Напишите в развернутом виде входящее в (7) выражение $du\ \Big|_A + \frac{1}{2!}d^2u\ \Big|_A + \frac{1}{3!}d^3u\ \Big|_A$ для функции двух переменных. ♦

**Задача**. Пусть функция $f(x_1, \ldots, x_m)$ непрерывна в замкнутой ограниченной области $\overline{\Omega} \subset \mathrm{R}^m$ и имеет все непрерывные частные производные $\dfrac{\partial f}{\partial x_i}$ в открытой области $\Omega$. Пусть точки $(x_1^0, \ldots, x_m^0)$ и $(x_1, \ldots, x_m)$ из $\Omega$ можно соединить отрезком прямой, целиком лежащим в $\Omega$. Напишите формулу конечных приращений Лагранжа для функции $f$.♦

**Задачи.**

$1^0$. Функцию $x^3 + y^2 + yz$ разложите по формуле Тейлора в окрестности точки $A(1,1,1)$.

$2^0$. Функцию $arctg\left(\dfrac{1+x+y}{1-x+y}\right)$ разложите по формуле Маклорена, включая члены второго порядка. ♦

**Задачи**. Разложите следующие функции по формуле Маклорена, включая члены 6-го порядка:

$\left[\cos\left(x^2+y^2\right)\right]^{\frac{1}{3}}$; $arctg\left(\dfrac{x^2+y^2}{1+xy}\right)$; $e^{x+y}\cdot arctg\left(x^2-y^2\right)$.♦

4. Формула Тейлора справедлива и для отображений нормированных пространств.

**Теорема**. Пусть $\mathsf{X}$ и $\mathsf{Y}$ – линейные нормированные пространства, $\Omega$ – открытое множество в $\mathsf{X}$, отображение $f:\ \Omega \to \mathsf{Y}$ имеет на $\Omega$ все производные вплоть до $(n+1)$-го порядка включительно. Пусть точки $a$ и $x$ множества $\Omega$ таковы, что соединяющий их отрезок $[a,x]$ целиком лежит в $\Omega$; обозначим $x-a=h$. Тогда

$$f(x)=f(a)+\frac{f'(a)}{1!}\{h\}+\ldots+\frac{f^{(n)}(a)}{n!}\underbrace{\{h,\ldots,h\}}_{n\ раз}+R_{n+1},$$

где $\left\|R_{n+1}\right\|_{\mathsf{Y}} \le \dfrac{1}{(n+1)!}\sup\limits_{z\in[a,x]}\left\|f^{(n+1)}(z)\right\|\cdot\|h\|_{\mathsf{X}}^{n+1}$.

**Замечание**. Если отображение $f$ $k$ раз дифференцируемо (по Фреше) в точке $a$, то $k$-ую производную $f^{(k)}(a)$ этого отображения можно понимать как симметричный $k$-линейный непрерывный по совокупности аргументов оператор, действующий из $\mathsf{X}^k$ в $\mathsf{Y}$; дифференциал $k$-го порядка отображения $f$ в точке $a$ – это выражение $f^{(k)}(a)\underbrace{\{h,\ldots,h\}}_{k\ раз}$, где $\underbrace{\{h,\ldots,h\}}_{k\ раз}\in \mathsf{X}^k$.♦

**Замечание**. Если отображение $f$ имеет на $\Omega$ производные вплоть до $n$-го порядка, то

$$\left\|R_{n+1}\right\|_{\mathsf{Y}} \le \frac{\|h\|_{\mathsf{X}}^n}{n!}\sup_{x\in\Omega}\left\|f^{(n)}(x)-f^{(n)}(a)\right\|.$$

Если $f^{(n)}(x)$ как функция от $x$ непрерывна в точке $a$, то $\left\|R_{n+1}\right\|_{\mathsf{Y}} = o\left(\|h\|_{\mathsf{X}}^n\right)$ при $h\to 0$.♦

## Вопрос 5.

## Ряд Лорана. Классификация изолированных особых точек.

1. Пусть даны комплексное число $a$ и последовательность комплексных чисел $\{c_n\}_{-\infty}^{+\infty}$. **Ряд Лорана** – это ряд вида

$$\sum_{n=-\infty}^{+\infty} c_n (z-a)^n, \tag{1}$$

который является обобщением степенного ряда по целым неотрицательным степеням разности $(z-a)$ ($z$ – комплексная переменная). (1) понимается как сумма двух рядов:

$$\sum_{n=0}^{+\infty} c_n (z-a)^n \tag{2}$$

– **правильной части** ряда Лорана и

$$\sum_{n=-\infty}^{-1} c_n (z-a)^n \tag{3}$$

– **главной части** ряда Лорана. Ряд (1) считается сходящимся в том и только в том случае, если сходятся его правильная и главная части. Областью сходимости степенного ряда (2) является круг $\{z \mid |z-a| < R\}$, радиус которого определяется формулой Коши-Адамара: $R = \left[\overline{\lim_{n\to+\infty}} \sqrt[n]{|c_n|}\right]^{-1}$; при этом возможны случаи $R = 0$ (тогда (2) сходится только в точке $z = a$), $0 < R < +\infty$ и $R = +\infty$. Если в (3) сделать замену переменной $\varsigma = (z-a)^{-1}$, то получим степенной ряд $\sum_{k=1}^{+\infty} c_{-k}\varsigma^k$ с кругом сходимости, радиус которого обозначим через $r^{-1}$. Тогда областью сходимости ряда (3) будет внешность круга: $\{z \mid |z-a| > r\}$. Если $r < R$, то ряд (1) сходится в кольце $D = \{z \mid r < |z-a| < R\}$. Поскольку суммы рядов (2) и (3) в их областях сходимости суть аналитические функции, сумма ряда (1) есть функция, аналитическая в кольце $D$ (если $r < R$, т.е. $D$ содержит внутренние точки). Во всех внутренних точках кольца сходимости $D$ ряд (1) сходится абсолютно, а в точках двух его граничных окружностей поведение ряда (1) может оказаться каким угодно. На любом лежащем в $D$ компактном множестве ряд (1) сходится равномерно. Ряд (1) можно дифференцировать и интегрировать в $D$ почленно.

**Задачи**. Найдите области сходимости следующих рядов Лорана:

$$\sum_{n=-\infty}^{+\infty} \frac{z^n}{2^{|n|}}; \quad \sum_{n=-\infty}^{+\infty} \frac{z^n}{3^n+1}; \quad \sum_{n=-\infty}^{+\infty} \frac{(z+1)^n}{2^{n^2}}; \quad \sum_{n=-\infty}^{+\infty} 2^n z^n. \blacklozenge$$

Любая однозначная функция $f(z)$, аналитическая внутри **круга**, единственным образом представима в этом круге сходящимся **рядом** Тейлора. Функция, аналитическая **в круговом кольце**, представима в нём рядом Лорана.

**Теорема Лорана**. Пусть $f(z)$ – однозначная аналитическая в кольце $D = \left\{z \mid r < |z-a| < R\right\}$ функция комплексной переменной $z$. Тогда $f$ единственным образом представима в $D$ сходящимся рядом Лорана. $\left(0 \le r < R \le +\infty\right)$

**Доказательство**. Пусть $z$ – внутренняя точка области $D$. Выберем кольцо $\widetilde{D} = \left\{z \mid \widetilde{r} < |z-a| < \widetilde{R}\right\}$, $r < \widetilde{r} < \widetilde{R} < R$, для которого выбранная точка $z$ также является внутренней. По формуле Коши для двусвязной области $\widetilde{D}$ имеем

$$f(z) = \frac{1}{2\pi i}\int_{\gamma_{\widetilde{R}}} \frac{f(\varsigma)}{\varsigma - z}d\varsigma - \frac{1}{2\pi i}\int_{\gamma_{\widetilde{r}}} \frac{f(\varsigma)}{\varsigma - z}d\varsigma, \tag{4}$$

где $\gamma_{\widetilde{R}}$, $\gamma_{\widetilde{r}}$ – граничные окружности кольца $\widetilde{D}$, причём их обход производится **против часовой стрелки**.

Для всех $\varsigma \in \gamma_{\tilde{R}}$ выполнено неравенство $\left|\dfrac{z-a}{\varsigma-a}\right| \le q < 1$,

поэтому на $\gamma_{\tilde{R}}$ ряд $\dfrac{1}{\varsigma - z} = \dfrac{1}{(\varsigma-a)\left(1-\dfrac{z-a}{\varsigma-a}\right)} = \sum\limits_{n=0}^{+\infty}\dfrac{(z-a)^n}{(\varsigma-a)^{n+1}}$

сходится равномерно по $\varsigma$. Умножая его на **ограниченную** функцию $\dfrac{1}{2\pi i}f(\varsigma)$ и интегрируя почленно вдоль $\gamma_{\tilde{R}}$, получим

$$\frac{1}{2\pi i}\int\limits_{\gamma_{\tilde{R}}}\frac{f(\varsigma)}{\varsigma-z}d\varsigma = \sum_{n=0}^{+\infty}c_n(z-a)^n, \quad c_n = \frac{1}{2\pi i}\int\limits_{\gamma_{\tilde{R}}}\frac{f(\varsigma)}{(\varsigma-a)^{n+1}}d\varsigma .$$

Для всех $\varsigma \in \gamma_{\tilde{r}}$ выполнено неравенство $\left|\dfrac{\varsigma-a}{z-a}\right| \le q < 1$,

поэтому на $\gamma_{\tilde{r}}$ ряд $-\dfrac{1}{\varsigma - z} = \dfrac{1}{(z-a)\left(1-\dfrac{\varsigma-a}{z-a}\right)} = \sum\limits_{n=1}^{+\infty}\dfrac{(\varsigma-a)^{n-1}}{(z-a)^{n}}$

сходится равномерно по $\varsigma$. Умножая его на **ограниченную** функцию $\dfrac{1}{2\pi i}f(\varsigma)$ и интегрируя почленно вдоль $\gamma_{\tilde{r}}$, получим

$$-\frac{1}{2\pi i}\int\limits_{\gamma_{r}}\frac{f(\varsigma)}{\varsigma-z}d\varsigma = \sum_{k=-1}^{-\infty}c_k(z-a)^k, \quad c_k = \frac{1}{2\pi i}\int\limits_{\gamma_{r}}\frac{f(\varsigma)}{(\varsigma-a)^{k+1}}d\varsigma .$$

По теореме Коши в выражениях коэффициентов $c_n$, $c_k$ окружности $\gamma_{\tilde{R}}$ и $\gamma_{\tilde{r}}$ можно заменить любой окружностью $\gamma_\rho = \left\{z \,\middle|\, |z-a| = \rho\right\}$, $r < \rho < R$. Тогда (4) примет вид

$$f(z) = \sum_{n=-\infty}^{+\infty}c_n(z-a)^n, \; c_n = \frac{1}{2\pi i}\int\limits_{\gamma_\rho}\frac{f(\varsigma)}{(\varsigma-a)^{n+1}}d\varsigma, \; n = 0, \pm 1, \ldots \quad (5)$$

Пусть наряду с разложением (5) в $D$ имеется другое разложение $f(z)=\sum_{n=-\infty}^{+\infty} b_n(z-a)^n$; внутри кольца $D$ $\sum_{n=-\infty}^{+\infty} c_n(z-a)^n \equiv \sum_{n=-\infty}^{+\infty} b_n(z-a)^n$. Оба этих ряда сходятся равномерно на окружности $\gamma_\rho=\{z \mid |z-a|=\rho\}$, $r<\rho<R$. Фиксируем целое число $m$, умножим оба ряда на $(z-a)^{-m-1}$ и проинтегрируем их почленно вдоль $\gamma_\rho$. Учтём, что $\int_{\gamma_\rho}(z-a)^{n-m-1}dz=\begin{cases} 0, & n\neq m, \\ 2\pi i, & n=m; \end{cases}$ (проведите вычисления!). Получаем $c_m=b_m$. В силу произвольности $m$ этим доказана единственность разложения (5). ♦

**Задача**. Докажите **неравенства Коши** для коэффициентов ряда Лорана: $|c_n|\leq \dfrac{M}{\rho^n}$, $n=0,\pm 1,\ldots$, где $\rho$ – радиус окружности $\gamma_\rho$ в (5), $M=\max\limits_{z\in\gamma_\rho}|f(z)|$. ♦

**Задача**. Что можно утверждать о коэффициентах разложения $\sum_{n=-\infty}^{+\infty} c_n z^n$, если оно представляет чётную функцию? Нечётную функцию? ♦

**Замечания**. **Наибольшее** кольцо, в котором сходится ряд (1), является точной (открытой) областью сходимости ряда Лорана. Вне замыкания этого кольца ряд (1) расходится. Такое кольцо может оказаться кругом с выколотым центром, внешностью круга (без бесконечно удалённой точки), или всей комплексной плоскостью без точек $z=a$ и $z=\infty$.

Напомним, что особой точкой аналитической функции $f(z)$ называется точка комплексной плоскости, являющаяся препятствием для аналитического продолжения функции $f$ вдоль какого-либо проходящего через эту точку пути. На границе круга сходимости ряда вида (2) лежит хотя бы одна особая точка суммы этого ряда. Поэтому точной областью сходимости ряда Лорана является круговое кольцо, на обеих граничных окружностях которого имеется хотя бы по одной особой точке его суммы.

Если функция $f(z)$ аналитична в круге $\{z \mid |z-a|<R\}$, то главная часть (3) её ряда Лорана равна нулю, и ряд Лорана сводится к ряду Тейлора (всюду в этом круге, включая $z=a$).

**Замечание**. Формулы (5) для коэффициентов ряда Лорана применяются редко. Обычно тейлоровские и лорановские разложения заданной функции находят косвенным образом: при помощи разложения дробей на простейшие, применением формулы суммы членов геометрической прогрессии, разложениями в ряды элементарных функций, умножением или делением степенных рядов, почленным интегрированием или дифференцированием и т.п. В силу единственности лорановского разложения все такие приёмы приведут к искомому ряду. ♦

Под лорановским разложением функции $f$ с центром **в бесконечно удалённой точке** $(a=\infty)$ понимают ряд

$$f(z)=\sum_{n=-\infty}^{+\infty} c_n z^n , \tag{6}$$

**правильная часть** которого есть ряд $\sum_{n=-\infty}^{0} c_n z^n$, а **главная часть** – ряд $\sum_{n=1}^{+\infty} c_n z^n$. Областью сходимости ряда (6) является кольцо $\{z \mid r<|z|<R\}$, $0 \le r < R \le +\infty$, а коэффициенты имеют вид

$$c_n=-\frac{1}{2\pi i}\int_{\gamma_\rho} \varsigma^{n+1} f(\varsigma)d\varsigma, \qquad n=0,\pm 1,\ldots, \qquad r<\rho<R.$$

Преобразованием $\varsigma=\dfrac{1}{z}$ разложение (6) с центром в $a=\infty$ можно свести к разложению с центром в нуле; но часто приходится строить разложение (6) функции $f$ в окрестности точки $a=\infty$ (исключая саму бесконечно удалённую точку), т.е. во внешности некоторого круга с центром в нуле: $\left\{z \mid r < |z| < +\infty\right\}$.

**Задачи**. Функцию $f(z) = \dfrac{1}{(z-1)(z-2)}$ разложите в ряд Лорана в круге $|z| < 1$; в кольце $1 < |z| < 2$; в кольце $2 < |z| < +\infty$; по степеням разности $(z - i)$; по степеням разности $\left(z - \dfrac{3}{2} - i\right)$.♦

**Задачи**. Данную функцию разложите в ряд Лорана в данной области .

$1^0$. $(z-i)^{-2}$, $1 < |z| < +\infty$.

$2^0$. $\dfrac{1}{\left(z^2-9\right)z^2}$, $1 < |z| < 2$.

$3^0$. $z^4 e^{\frac{1}{z}}$, $0 < |z| < +\infty$.

$4^0$. $\sin\dfrac{z}{z-1}$ в проколотой окрестности точки $a = 1$.♦

2. Ряд Лорана функции $f(z)$, рассматриваемый **на единичной окружности** $\left\{z \mid |z| = 1\right\}$, является тригонометрическим рядом Фурье функции $g(t) = f\left(e^{it}\right)$ на отрезке $-\pi \le t \le \pi$.

Тригонометрический ряд Фурье

$$g(t) \propto \frac{a_0}{2} + \sum_{n=1}^{+\infty} a_n \cos nt + b_n \sin nt\,, \tag{7}$$

$$a_n = \frac{1}{\pi}\int_{-\pi}^{\pi} g(t)\cos nt\,dt\,,\quad b_n = \frac{1}{\pi}\int_{-\pi}^{\pi} g(t)\sin nt\,dt\,,$$

интегрируемой на отрезке $-\pi \le t \le \pi$ функции $g(t)$ можно записать в комплексной форме. Для этого подставим $\cos nt = \dfrac{e^{int} + e^{-int}}{2}$ и $\sin nt = \dfrac{e^{int} - e^{-int}}{2i}$ в (7); тогда получим (7) в виде

$$g(t) \propto \sum_{n=-\infty}^{+\infty} c_n e^{int},$$

(8)

$$c_n = \frac{a_n - ib_n}{2} = \frac{1}{2\pi}\int_{-\pi}^{\pi} g(t) e^{-int}\, dt, \quad n = 0, 1, 2, \ldots,$$

$$c_n = \frac{a_{-n} + ib_{-n}}{2} = \frac{1}{2\pi}\int_{-\pi}^{\pi} g(t) e^{-int}\, dt, \quad n = -1, -2, \ldots.$$

Положим $z = e^{it}$ и $f(z) = f(e^{it}) = g(t)$; тогда (8) примет вид $\sum\limits_{n=-\infty}^{+\infty} c_n z^n$, $\quad c_n = \dfrac{1}{2\pi i}\displaystyle\int_{|z|=1} \frac{f(\varsigma)}{\varsigma^{n+1}} d\varsigma$. Таким образом, тригонометрический ряд Фурье функции $g$, записанный в комплексной форме, является рядом Лорана функции $f$ **на единичной окружности**. Отметим, что так построенный ряд Лорана будет иметь непустую **область** сходимости только если функция $g$ удовлетворяет существенным ограничениям.

**Задачи**. Пользуясь рядами Лорана, разложите в тригонометрические ряды Фурье следующие функции действительной переменной $t$ ($q$ – параметр).

$1^0$. $\dfrac{1 - q\cos t}{1 - 2q\cos t + q^2}$, $\quad -1 < q < 1$.

$2^0$. $\dfrac{q\sin t}{1 - 2q\cos t + q^2}$, $\quad -1 < q < 1$.

$3^0$. $\dfrac{1 - q^2}{1 - 2q\cos t + q^2}$, $\quad 0 < q < 1$.

$4^0$. $\ln\left(1 - 2q\cos t + q^2\right)$, $\quad 0 < q < 1$.

3. Пусть **однозначная** функция $f(z)$ **аналитична** в проколотой окрестности $\{z \mid 0 < |z - a| < R\}$ точки $a$ $\quad (a \neq \infty)$; в самой точке $a$ функция $f$ может быть не определена. Если существует такое конечное комплексное число $A$, что положив

$f(a)=A$, мы получим функцию $f$, аналитическую во всём круге $\{z \mid |z-a|<R\}$ (включая точку $a$), то $a$ называют **правильной** точкой функции $f$. Если же такого числа $A$ не существует, то $a$ называют **изолированной особой точкой однозначного характера** функции $f$. Различают три типа изолированных особых точек в зависимости от поведения функции $f$ в их окрестности.

Точка $a$ называется **устранимой** особой точкой функции $f$, если существует **конечный** предел $\lim\limits_{z\to a} f(z)$. (Это фактически означает, что $a$ – правильная точка, но удобно придерживаться указанной терминологии.)

Точка $a$ называется **полюсом** функции $f$, если существует $\lim\limits_{z\to a} f(z)=\infty$.

Точка $a$ называется **существенно особой** точкой функции $f$, если $\lim\limits_{z\to a} f(z)$ не существует.

В проколотой окрестности точки $a$ функцию $f$ можно разложить в ряд Лорана, вид которого зависит от характера особой точки.

**Теорема**. Точка $a$ является устранимой особой точкой однозначной аналитической функции $f(z)$ в том и только в том случае, если равна нулю главная часть лорановского разложения функции $f$ в проколотой окрестности точки $a$. ♦

«Устранимость» особой точки $a$ означает, что функцию $f$ можно переопределить (или доопределить), положив $f(a)=\lim\limits_{z\to a} f(z)$; тогда $f$ станет аналитической во всём круге $\{z \mid |z-a|<R\}$. Если аналитическая в проколотой окрестности точки $a$ функция $f$ ограничена, то $a$ – устранимая особая точка функции $f$ (докажите, используя неравенства Коши для коэффициентов ряда Лорана). В окрестности устранимой особой точки $a$ функция $f$ (после её переопределения в точке $a$) может

быть представлена в виде $f(z)=(z-a)^m h(z)$, где целое число $m \geq 0$, а функция $h$ аналитична и $h(a) \neq 0$; число $m$ является порядком нуля функции $f$ в точке $a$.

**Примеры**. В точке $a=0$ функции $\dfrac{\sin z^3}{z}$, $\dfrac{1-e^{z^4}}{z}$, $\dfrac{1-\cos z}{z}$, $\dfrac{1-\cos z}{z^2}$ имеют устранимую особую точку. Нулём какого порядка она является? ♦

**Теорема**. Точка $a$ является полюсом однозначной аналитической функции $f(z)$ в том и только в том случае, если главная часть лорановского разложения функции $f$ в проколотой окрестности точки $a$ содержит лишь конечное число ненулевых членов:

$$f(z)=\frac{c_{-m}}{(z-a)^m}+\ldots+\frac{c_{-1}}{z-a}+\sum_{n=0}^{+\infty} c_n (z-a)^n, \quad c_{-m} \neq 0, \ c_n = 0$$

при $n<-m$, $m>0$.♦

В проколотой окрестности полюса $a$ функцию $f$ можно представить в виде $f(z)=\dfrac{h(z)}{(z-a)^m}$, где функция $h$ аналитична (включая точку $a$) и $h(a) \neq 0$; число $m>0$ называется порядком полюса функции $f$ в точке $a$. Если $a$ является изолированным нулём порядка $m$ аналитической функции $w(z)$, то в некоторой окрестности точки $a$ функция $f(z)=\dfrac{1}{w(z)}$ аналитична всюду, кроме точки $a$, которая является полюсом порядка $m$ функции $f$ (докажите). Если $a$ является полюсом порядка $m$ функции $f(z)$, то в некоторой окрестности точки $a$ функция $w(z)=\dfrac{1}{f(z)}$ аналитична, причём $a$ является нулём функции $w$ порядка $m$ (докажите).

**Задачи**. Найдите все полюсы функций $tg\, z$, $ctg\, z$, $\frac{1}{\sin z^2}$, $\frac{1}{z^2+1}$, $\frac{1}{z^4+1}$, $\frac{1}{z^4+2z^2+1}$, $\frac{1}{e^{z^2}+1}$ и укажите их порядки. Имеют ли эти функции конечные особые точки, отличные от полюсов? ♦

**Теорема**. Точка $a$ является существенно особой точкой однозначной аналитической функции $f(z)$ в том и только в том случае, если главная часть лорановского разложения функции $f$ в проколотой окрестности точки $a$ содержит бесконечное число ненулевых членов. ♦

Поведение функции $f$ в окрестности существенно особой точки описывает

**Теорема Сохоцкого-Казорати-Вейерштрасса**. Для любого конечного комплексного числа $A$ и для $A=\infty$ найдётся такая последовательность точек $\{z_k\}$, сходящаяся к существенно особой точке $a$ функции $f(z)$, что $\lim\limits_{k\to\infty} f(z_k)=A$.♦

Если точка $a$ является существенно особой точкой функции $f(z)$, причём $f(z)\neq 0$ в некоторой окрестности точки $a$, то и для функции $w(z)=\frac{1}{f(z)}$ точка $a$ будет существенно особой (докажите).

Более точно поведение функции $f$ в окрестности существенно особой точки характеризует

**«Большая» теорема Пикара**. В любой (сколь угодно малой) окрестности существенно особой точки функция $f(z)$ бесконечно много раз принимает любое конечное комплексное значение $A$, за исключением, быть может, одного. ♦

**Примеры**. Точка $a=0$ является существенно особой для функций $e^{\frac{1}{z}}$ и $\sin\frac{1}{z}$. Для функции $e^{\frac{1}{z}}$ пикаровское исключённое значение $A=0$; у $\sin\frac{1}{z}$ исключённого значения нет. ♦

Бесконечно удалённая точка $a=\infty$ называется **изолированной** особой точкой однозначной аналитической функции $f(z)$, если существует такое число $r>0$, что в кольце $D=\{z \mid r<|z|<+\infty\}$ у функции $f$ нет (конечных) особых точек. В этом случае в $D$ функцию $f$ можно разложить в ряд Лорана:

$$f(z)=\sum_{n=-\infty}^{+\infty} c_n z^n . \tag{9}$$

Точка $a=\infty$ называется **устранимой** особой точкой функции $f$, если существует конечный предел $\lim\limits_{z\to\infty} f(z)$, т.е. (9) не содержит ненулевых членов с положительными степенями $z$. (В каком случае $a=\infty$ будет нулём порядка $m$ функции $f$?)

Точка $a=\infty$ называется **полюсом** функции $f$, если существует $\lim\limits_{z\to\infty} f(z)=\infty$, т.е. (9) содержит лишь конечное число ненулевых членов с положительными степенями $z$ ($c_m \neq 0$, $c_n=0$ при $n>m>0$; тогда $m$ – порядок полюса).

Точка $a=\infty$ называется **существенно особой** точкой функции $f$, если $\lim\limits_{z\to\infty} f(z)$ не существует, т.е. (9) содержит бесконечное число ненулевых членов с положительными степенями $z$.

Замена переменной $\varsigma=\frac{1}{z}$ переводит бесконечно удалённую точку в нуль, а кольцо $D$ – в проколотую окрестность нуля. Характер особой точки функции $f\left(\frac{1}{\varsigma}\right)$ при этом сохраняется.

**Примеры**. Тейлоровские разложения функций $e^z$, $\sin z$, $\cos z$ по неотрицательным степеням $z$ можно рассматривать как лорановские разложения в окрестности $a=\infty$. Эти функции имеют существенно особую точку на бесконечности. Каково поведение на бесконечности функций $e^{\frac{1}{z}}$, $\frac{1}{z^3}e^{\frac{1}{z}}$, $z^3 e^{\frac{1}{z}}$, $e^{z+\frac{1}{z}}$?♦

По теореме Лиувилля аналитическая во всей плоскости функция, которая не имеет конечных особых точек и ограничена на бесконечности, постоянна. Аналитическая функция, не имеющая конечных особых точек, называется целой. Точка $a=\infty$ является изолированной особой точкой целой функции $f$. Если это – устранимая особая точка, то $f \equiv const$. Если это – полюс, то $f$ является многочленом (докажите!). Если же $a=\infty$ – существенно особая точка целой функции $f$, то $f$ называют **целой трансцендентной**.

**Задачи**. Найдите все изолированные особые точки однозначного характера следующих функций и определите их тип. Имеются ли неизолированные особые точки?

$1^0$. $\dfrac{z}{\sin z}$. $2^0$. $z^2 \sin\dfrac{z}{z+1}$. $3^0$. $z\left(e^{\frac{1}{z}}-1\right)$. $4^0$. $\sin\left(e^{\frac{1}{z}}\right)$.

$5^0$. $\dfrac{1}{e^z-1}-\dfrac{1}{z}$. $6^0$. $\sin\dfrac{1}{z}+\dfrac{1}{z^2}$. $7^0$. $ctg\, z-\dfrac{2}{z}$. $8^0$. $e^{tg\frac{1}{z}}$.♦

4. Понятие особой точки аналитической функции возникает в задаче аналитического продолжения однозначной функции и в связи с изучением многозначных функций комплексной переменной. Пусть в проколотой окрестности $D$ точки $a$ (конечной или бесконечно удалённой) возможно аналитическое продолжение (при помощи степенных рядов) однозначной аналитической функции $f(z)$ вдоль любого пути, расположенного в $D$ и не проходящего через $a$; тогда $a$ называется **изолированной особой точкой**. Если при этом аналитическое продолжение функции $f$ вдоль всевозможных замкнутых путей в $D$ при возвращении в ту же точку $z$ не изменяет значений функции $f$, то $a$ называется особой

точкой **однозначного характера**. Такая особая точка $a$ (если она не является устранимой) может быть либо полюсом, либо существенно особой точкой. Если же при аналитическом продолжении вдоль замкнутых путей в $D$, окружающих точку $a$, функция $f$ перестаёт быть однозначной, то точка $a$ называется особой точкой **многозначного характера** или **точкой ветвления**. Это легко представить себе, разрезав кольцо $D$ вдоль некоторого его радиуса: если на обоих берегах разреза значения продолженной на $D$ функции $f$ совпадают, то для данной ветви аналитической функции $a$ – особая точка однозначного характера, если же значения $f$ на разных берегах разреза не совпадают, то $a$ – точка ветвления полной аналитической функции $F(z)$, элементом которой является $f(z)$. Полная аналитическая функция $F$, вообще говоря, не есть функция точки $z$ в обычном смысле этого слова: одному и тому же значению $z$ может отвечать конечное или счётное число значений $F$. Как однозначную аналитическую функцию $F$ можно рассматривать только на отвечающей ей римановой поверхности, которая состоит из некоторого числа расширенных комплексных плоскостей, определённым образом соединённых друг с другом.

Пусть однозначная аналитическая функция $f$ (вместе с её областью определения) является элементом полной функции $F$, $a$ – точка ветвления. Будем строить всевозможные аналитические продолжения функции $f$. Если после некоторого минимального числа $k \geq 2$ однократных обходов в $D$ точки $a$ (в одном и том же направлении) мы получим исходный элемент $f$ многозначной функции $F$, причём при $z \to a$ все определённые элементом $f$ ветви функции $F$ стремятся к одному и тому же пределу (конечному или бесконечному), то $a$ называется **алгебраической точкой ветвления**, а число $k-1$ – её порядком. Если же такие обходы точки $a$ дают всё новые и новые элементы многозначной функции $F$ или предел при $z \to a$ не существует, то $a$ называется **трансцендентной точкой ветвления**. Если ни при каком числе последовательных обходов (в одном и том же направлении) вокруг точки $a$ аналитическое продолжение некоторого элемента $f$ не приводит к этому исходному элементу, то $a$ называется **логарифмической точкой ветвления**.

**Примеры**. Точки $a=0$ и $a=\infty$ являются алгебраическими точками ветвления порядка $k-1$ функции $\sqrt[k]{z}$ ($k$ – натуральное). В каждой из этих точек соединяются вместе $k$ листов римановой поверхности данной функции. При $z\to a$ все ветви этой функции стремятся к одному и тому же пределу (докажите!). Каковы точки ветвления функции $\dfrac{1}{\sqrt[k]{z}}$ ?

Для функции $e^{\frac{1}{\sqrt[k]{z}}}$ точка $a=0$ является трансцендентной точкой ветвления (предел ветвей этой функции при $z\to a$ не существует); точка $a=\infty$ – алгебраическая точка ветвления.

Для (многозначной) функции $Ln\, z=\ln|z|+iArg\, z$ точки $a=0$ и $a=\infty$ – логарифмические точки ветвления. В каждой из этих точек соединяется бесконечное (счётное) число листов римановой поверхности. ♦

**Задачи**. Найдите точки ветвления функций $arctg\, z=\dfrac{1}{2i}\ln\dfrac{1+iz}{1-iz}$, $\sqrt{e^{z^2}+1}$, $e^{\sqrt[3]{z}}$, $\sin\sqrt[4]{z}$, $\sin\dfrac{1}{\sqrt[5]{z}}$ .♦

## Вопрос 6.

## Билинейные и квадратичные формы. Приведение их к каноническому виду. Закон инерции.

1. **Билинейные формы**.

Будем рассматривать $x$ и $y$ как независимые переменные, пробегающие все линейное пространство $\mathbf{V}$ над полем $\mathbf{P}$. Введем отображение $B:\mathbf{V}\times\mathbf{V}\to\mathbf{P}$, причем будем считать, что значения $B(x;y)$ определяются только элементами $x,y$ и не зависят от выбора базиса в $\mathbf{V}$. $B(x;y)$ называют **билинейной формой**, если при каждом фиксированном $y$ $B(x;y)$ является линейной формой (линейным функционалом) от $x$, а при каждом фиксированном $x$ – линейной формой от $y$.

**Примеры** билинейных форм.

1. Если $l_1(x)$ и $l_2(y)$ – линейные формы, определенные на пространстве $\mathbf{V}$, $x,y\in\mathbf{V}$, то $l_1(x)\cdot l_2(y)$ является билинейной формой.
2. $\mathbf{V}=C[a,b]$. Пусть фиксирована непрерывная функция двух переменных $K(s,t)\in C([a,b]\times[a,b])$. Для каждой пары непрерывных функций $x,y\in\mathbf{V}$ положим $B(x;y)=\int\limits_a^b\int\limits_a^b K(s,t)x(s)y(t)dsdt$. В частности, если $K(s,t)\equiv 1$, то $B(x;y)=\int\limits_a^b x(s)ds\cdot\int\limits_a^b y(t)dt$ – произведение линейных функционалов.
3. $\mathbf{V}=\mathbf{R}^m$, $\vec{x}=(x_1,\ldots,x_m)^T\in\mathbf{R}^m$, $\vec{y}=(y_1,\ldots,y_m)^T\in\mathbf{R}^m$. Пусть фиксирована матрица

$A = (a_{ik}) \in \mathbf{R}^{m \times m}$. Положим

$B(\vec{x};\vec{y}) = \vec{x}^T A\, \vec{y} = \sum_{i=1}^{m}\sum_{k=1}^{m} a_{ik} x_i y_k$ .◆

Если $\dim \mathbf{V} = m < \infty$ и $e = (e_1,\ldots,e_m)$ – базис пространства $\mathbf{V}$, то $x = \sum_{i=1}^{m} x_i e_i$ , $y = \sum_{k=1}^{m} y_k e_k$ и билинейная форма в базисе $e$ имеет вид

$$B(x;y) = B\left(\sum_{i=1}^{m} x_i e_i;\ \sum_{k=1}^{m} y_k e_k\right) = \sum_{i=1}^{m}\sum_{k=1}^{m} x_i y_k \cdot B(e_i;e_k) =$$

$= \sum_{i=1}^{m}\sum_{k=1}^{m} a_{ik} x_i y_k$ . Коэффициенты $a_{ik} = B(e_i;e_k)$ составляют матрицу билинейной формы в базисе $e$ .

Пусть $e$ и $f$ – два базиса в $\mathbf{V}$, связанные матрицей перехода $C:\ f = eC$. Напомним, что последнее символическое равенство означает, что каждый элемент базиса $f = (f_1,\ldots,f_m)$ следующим образом раскладывается по элементам базиса $e = (e_1,\ldots,e_m)$:

$$f_1 = c_{11}e_1 + c_{21}e_2 + \ldots + c_{m1}e_m\,,$$
$$f_2 = c_{12}e_1 + c_{22}e_2 + \ldots + c_{m2}e_m\,,$$
$$.\qquad .\qquad .\quad \ldots\quad .$$
$$f_m = c_{1m}e_1 + c_{2m}e_2 + \ldots + c_{mm}e_m\,;$$

если перечислять элементы базисов в строке, то $(f_1,\ldots,f_m) = (e_1,\ldots,e_m)\cdot C$ , где

$$C = \begin{pmatrix} c_{11} & c_{12} & \ldots & c_{1m} \\ c_{21} & c_{22} & \ldots & c_{2m} \\ . & . & \ldots & . \\ c_{m1} & c_{m2} & \ldots & c_{mm} \end{pmatrix}.$$

Из разложений $x = \sum_{i=1}^{m} x_i e_i = \sum_{i=1}^{m} \tilde{x}_i f_i$, $y = \sum_{k=1}^{m} y_k e_k = \sum_{k=1}^{m} \tilde{y}_k f_k$ получаем выражения билинейной формы в базисах $e$ и $f$:

$$B(x;y) = \sum_{i=1}^{m}\sum_{k=1}^{m} a_{ik} x_i y_k = \sum_{i=1}^{m}\sum_{k=1}^{m} \tilde{a}_{ik} \tilde{x}_i \tilde{y}_k .$$

Введем обозначения $m \times m$-матриц билинейной формы $B(x;y)$ в базисах $e$ и $f$: $A = (a_{ik})_{k=1,\ldots,m}^{i=1,\ldots,m}$, $\tilde{A} = (\tilde{a}_{ik})_{k=1,\ldots,m}^{i=1,\ldots,m}$. Как связаны матрицы $A$ и $\tilde{A}$?

**Теорема 1**. Если $f = eC$, то $\tilde{A} = C^T AC$.

**Следствие**. $rang\ A = rang\ \tilde{A}$.♦

Поскольку ранг матрицы билинейной формы не зависит от выбора базиса в пространстве $\mathbf{V}$, это число естественно назвать **рангом билинейной формы**.

Билинейные формы, определенные на $\mathbf{V}$, можно складывать и умножать на числа из поля $\mathbf{P}$; при этом будут получаться новые билинейные формы.

**Теорема 2**. Множество **всех** билинейных форм, определенных на линейном пространстве $\mathbf{V}$, является линейным пространством. Если $\dim \mathbf{V} = m < \infty$, то размерность пространства всех билинейных форм на $\mathbf{V}$ равна $m^2$.

**Задача**. Найдите базис в пространстве билинейных форм ($\dim \mathbf{V} < \infty$).♦

Рассмотрим два важных класса билинейных форм. Билинейную форму называют **симметричной**, если для любых $x, y \in \mathbf{V}$ $B(x;y) = B(y;x)$. Билинейную форму называют **кососимметричной**, если для любых $x, y \in \mathbf{V}$ $B(x;y) = -B(y;x)$.

**Пример**. В действительном евклидовом пространстве скалярное произведение является симметричной билинейной формой. ♦

В базисе $e$ коэффициенты билинейной формы $a_{ik}=B(e_i;e_k)$, поэтому матрица симметричной билинейной формы симметрична: $a_{ik}=a_{ki}$ для всех $i$ и $k$, а матрица кососимметричной билинейной формы кососимметрична: $a_{ik}=-a_{ki}$ для всех $i$ и $k$ (в частности, все $a_{ii}=0$).

**Теорема 3**. В линейном пространстве всех билинейных форм, определенных на $\mathbf{V}$, множество всех симметричных билинейных форм образует подпространство; множество всех кососимметричных билинейных форм также образует подпространство. Пространство всех билинейных форм является прямой суммой этих двух подпространств.

Доказательство. Очевидно, что если $B_1(x;y)$ и $B_2(x;y)$ – симметричные билинейные формы, то и $\alpha B_1(x;y)+\beta B_2(x;y)$ – симметричная билинейная форма, каковы бы ни были $\alpha,\beta\in\mathbf{P}$. Это значит, что симметричные билинейные формы образуют подпространство.

Если $B_1(x;y)$ и $B_2(x;y)$ – кососимметричные билинейные формы, то и $\alpha B_1(x;y)+\beta B_2(x;y)$ – кососимметричная билинейная форма при любых $\alpha,\beta\in\mathbf{P}$. Это значит, что кососимметричные билинейные формы образуют подпространство.

Ясно, что билинейная форма может быть одновременно симметричной и кососимметричной в том и только в том случае, если она нулевая, т.е. для любых $x,y\in\mathbf{V}$ $B(x;y)=0$. Следовательно, сумма двух рассматриваемых подпространств является **прямой** суммой. С другой стороны, всякую билинейную форму можно представить в виде суммы симметричной и кососимметричной билинейных форм:

$$B(x;y)=\frac{1}{2}\bigl(B(x;y)+B(y;x)\bigr)+\frac{1}{2}\bigl(B(x;y)-B(y;x)\bigr);$$

$\frac{1}{2}\bigl(B(x;y)+B(y;x)\bigr)$ – симметричная, $\frac{1}{2}\bigl(B(x;y)-B(y;x)\bigr)$ – кососимметричная. Поэтому прямая сумма двух рассматриваемых подпространств совпадает со всем пространством билинейных форм. ♦

**Замечание**. Если $\dim \mathbf{V} = m < \infty$, то размерность подпространства всех симметричных билинейных форм равна $\dfrac{m(m+1)}{2}$, а размерность подпространства всех кососимметричных билинейных форм равна $\dfrac{m(m-1)}{2}$. ♦

**Замечание** о термине «форма». Многочлен $h(x_1, \ldots, x_m)$ называют **однородным многочленом степени** $k$, если $h(\alpha x_1, \ldots, \alpha x_m) = \alpha^k \cdot h(x_1, \ldots, x_m)$. Например, выражение $x_1^2 + x_2^2 + x_3^2 + x_1 x_2 + x_1 x_3 + x_2 x_3$ является однородным многочленом степени 2 относительно переменных $x_1$, $x_2$, $x_3$. Если в пространстве $\mathbf{V}$ фиксирован базис, то линейный функционал $l(x)$ является однородным многочленом первой степени относительно координат элементов $x$ в этом базисе. Однородные многочлены степени $k$ принято называть **формами** степени $k$. При $k = 1$ получаем линейные формы, при $k = 2$ – квадратичные формы.

2. **Квадратичные формы**.

Пусть $B(x; y)$ – **симметричная** билинейная форма, определенная на линейном пространстве $\mathbf{V}$. $Q(x) = B(x; x)$ называют **квадратичной формой**, определенной на $\mathbf{V}$.

**Замечание**. Мы потребовали $B(x; y) = B(y; x)$. Это объясняется следующими соображениями. Если взять **произвольную** билинейную форму $B(x; y)$, то выражение $\frac{1}{2}\left(B(x; y) + B(y; x)\right)$ будет **симметричной** билинейной формой. Положим в нем $y = x$, тогда получим $\frac{1}{2}\left(B(x; x) + B(x; x)\right) = B(x; x)$, т.е. то же самое, как если бы в исходной форме $B(x; y)$ мы положили $y = x$. ♦

**Пример.** $\mathbf{V} = \mathbf{R}^m$, $A = \left(a_{ik}\right)_{k=1,\ldots,m}^{i=1,\ldots,m}$ – заданная действительная матрица, $\vec{x} = \left(x_1, \ldots, x_m\right)^T \in \mathbf{R}^m$, $\vec{y} = \left(y_1, \ldots, y_m\right)^T \in \mathbf{R}^m$, $B(\vec{x}; \vec{y}) = \sum_{i=1}^{m}\sum_{k=1}^{m} a_{ik} x_i y_k$. Квадратичная форма $Q(\vec{x}) = B(\vec{x}; \vec{x}) = \sum_{i=1}^{m}\sum_{k=1}^{m} a_{ik} x_i x_k$. Для изучения квадратичной формы матрицу $A$ можно считать симметричной – этого всегда можно добиться следующей группировкой слагаемых: $a_{ik} x_i x_k + a_{ki} x_k x_i = \dfrac{a_{ik} + a_{ki}}{2} x_i x_k + \dfrac{a_{ik} + a_{ki}}{2} x_k x_i$. Если в $\mathbf{R}^m$ введено скалярное произведение $(\vec{x}, \vec{y}) = x_1 y_1 + \ldots + x_m y_m$, то рассматриваемая квадратичная форма имеет вид $Q(\vec{x}) = (A\vec{x}, \vec{x}) = (\vec{x}, A\vec{x}) = \vec{x}^T A \vec{x}$, где $A = A^T$. ♦

Исходную **симметричную** билинейную форму $B(x; y)$ называют **полярной** для квадратичной формы $Q(x) = B(x; x)$. Матрицей квадратичной формы, определенной на конечномерном линейном пространстве, называют матрицу ее полярной билинейной формы.

**Теорема 4**. Полярная билинейная форма однозначно определяется своей квадратичной формой.

**Доказательство**. Пусть $Q(x) = B(x; x)$ – квадратичная форма, где симметричная билинейная форма $B(x; y)$ нам неизвестна; найдем её. По свойствам билинейной формы и в силу ее симметричности имеем

$Q(x + y) = B(x + y; x + y) = B(x; x) + B(x; y) + B(y; x) + B(y; y) = Q(x) + 2B(x; y) + Q(y)$. Отсюда

$$B(x; y) = \frac{1}{2}\left(Q(x + y) - Q(x) - Q(y)\right). ♦$$

Доказанная теорема служит еще одним обоснованием того, что при изучении квадратичных форм достаточно рассматривать только **симметричные** билинейные формы.

**Пример.** $\mathbf{V} = C[a,b]$. Положим $Q(x) = \int_a^b x^2(t)dt$ для любой непрерывной на $[a,b]$ функции $x = x(t)$. Билинейная форма, полярная к квадратичной форме $Q(x)$, имеет вид $B(x;y) = \frac{1}{2}(Q(x+y) - Q(x) - Q(y)) = \int_a^b x(t)y(t)dt$. ♦

Всюду далее мы будем рассматривать только **действительные** конечномерные линейные пространства $\mathbf{V}$, $\dim \mathbf{V} = m$. Пусть $e$ и $f$ – два базиса в $\mathbf{V}$, $f = eC$, $\vec{x}_e \in \mathbf{R}^m$ и $\vec{x}_f \in \mathbf{R}^m$ – координаты элемента $x \in \mathbf{V}$ в этих базисах: $\vec{x}_e = (x_1, \dots, x_m)^T$, $\vec{x}_f = (\tilde{x}_1, \dots, \tilde{x}_m)^T$. Тогда $\vec{x}_e = C\vec{x}_f$. Определенная на $\mathbf{V}$ квадратичная форма $Q(x)$ в базисе $e$ задается при помощи действительной симметричной $m \times m$-матрицы $A$: $Q(x) = \vec{x}_e^T A \vec{x}_e$. $Q(x)$ можно рассматривать как функцию $m$ независимых действительных переменных $x_1, \dots, x_m$ – однородный многочлен степени $2$ от этих переменных. В базисе $f$ та же квадратичная форма $Q(x)$ является функцией независимых действительных переменных $\tilde{x}_1, \dots, \tilde{x}_m$ и задается при помощи действительной симметричной матрицы $C^T AC$: $Q(x) = \vec{x}_f^T C^T AC \vec{x}_f$ (см. теорему 1). Основная задача, связанная с изучением заданной в базисе $e$ квадратичной формы $Q(x)$, состоит в нахождении такого базиса $f$ (или, что то же самое, – таких переменных $\tilde{x}_1, \dots, \tilde{x}_m$), что $Q(x)$ имеет в переменных $\tilde{x}_1, \dots, \tilde{x}_m$ наиболее простой вид.

Рассмотрим сначала случай, когда $\mathbf{V}$ – **евклидово** пространство и $e$ – **ортонормированный** базис. Всякая действительная симметричная матрица $A$ ортогонально подобна действительной диагональной матрице $D$: $A = C_0 D C_0^{-1}$, где

$$D=\begin{pmatrix}\lambda_1 & & & 0\\ & \lambda_2 & & \\ & & \dots & \\ 0 & & & \lambda_m\end{pmatrix},\ \lambda_k \text{ – собственные значения матрицы } A \text{, а } C_0$$

– ортогональная матрица, т.е. $C_0^{-1}=C_0^T$. Поэтому, если в качестве новых независимых переменных выбрать $\vec{x}_f=(\tilde{x}_1,\ldots,\tilde{x}_m)^T=C_0^{-1}\vec{x}_e$, то получим $Q(x)=\vec{x}_f^T C_0^{\,T} A C_0 \vec{x}_f=\vec{x}_f^T D\vec{x}_f=\lambda_1\tilde{x}_1^{\,2}+\lambda_2\tilde{x}_2^{\,2}+\ldots+\lambda_m\tilde{x}_m^{\,2}$.

**Определение.** Вид квадратичной формы с диагональной матрицей называют ее **каноническим** видом. ♦

Мы только что обнаружили важный факт:

**Теорема 5**. Всякая квадратичная форма, определенная на действительном евклидовом пространстве, может быть приведена к каноническому виду при помощи линейной замены независимых переменных с ортогональной матрицей преобразования этих переменных. ♦

Очевидно, что базис $f$ евклидова пространства $\mathbf{V}$, в котором квадратичная форма имеет указанный в теореме 5 канонический вид, является ортонормированным базисом, состоящим из собственных векторов матрицы квадратичной формы. Операция построения **ортонормированного** базиса, в котором квадратичная форма имеет канонический вид, называется приведением ее к **главным осям**. Вообще, привести квадратичную форму $Q(x)$ к каноническому виду – значит найти переменные $\vec{x}_f$ и выражение $Q(x)$, содержащее только квадраты этих переменных. Если приведение к каноническому виду производится при помощи ортогонального преобразования, то для записи этого канонического вида $Q(x)$ достаточно найти собственные значения матрицы $A$. Но чтобы найти переменные, в которых построен этот канонический вид, надо знать и собственные векторы матрицы $A$. Действительно, условие $A=C_0DC_0^{-1}$ означает, что $AC_0=C_0D$, или поэлементно: $\sum_{j=1}^{m}a_{ij}c_{jk}=\lambda_k c_{ik}$. Следовательно, $k$-ый столбец матрицы $C_0$

состоит из координат собственного вектора матрицы $A$, отвечающего собственному значению $\lambda_k$, $k=1,\ldots,m$. Напомним, что столбцы ортогональной матрицы $C_0$ образуют ортонормированную систему в смысле скалярного произведения $(\vec{u},\vec{v})=\sum_{i=1}^{m}u_i v_i$ в $\mathbf{R}^m$.

Полученный в теореме 5 результат является выражением следующего факта:

**Утверждение**. Для любой квадратичной формы $Q(x)$, определенной на евклидовом пространстве $\mathbf{V}$, существует единственный **самосопряженный** линейный оператор $A$ такой, что $Q(x)=(Ax,x)_{\mathbf{V}}$ для любого $x\in\mathbf{V}$.

Ортогональное преобразование переменных $\vec{x}_e$ определено только в евклидовом пространстве, и оно дает лишь один из канонических видов квадратичной формы. Канонический вид квадратичной формы не является однозначно определенным. Если в некотором базисе линейного (не обязательно евклидова) пространства $\mathbf{V}$ квадратичная форма $Q(x)$ имеет канонический вид, то переставляя элементы этого базиса (т.е. производя перенумерацию независимых переменных), снова получим $Q(x)$ в каноническом виде. Если $Q(x)=\lambda_1 x_1^2+\lambda_2 x_2^2+\ldots+\lambda_m x_m^2$, то полагая $x_1=\alpha_1\tilde{x}_1,\, x_2=\alpha_2\tilde{x}_2,\ldots,\, x_m=\alpha_m\tilde{x}_m$, где все коэффициенты $\alpha_i\neq 0$, получим другой канонический вид $Q(x)=(\lambda_1\alpha_1^2)\tilde{x}_1^{\,2}+(\lambda_2\alpha_2^2)\tilde{x}_2^{\,2}+\ldots+(\lambda_m\alpha_m^2)\tilde{x}_m^{\,2}$.

**Пример**. Пусть $Q(x)=\gamma_1 x_1^2+\gamma_2 x_2^2+\ldots+\gamma_m x_m^2$ – некоторый канонический вид квадратичной формы. Предположим для простоты, что

$$\gamma_1>0,\ldots,\gamma_p>0,\gamma_{p+1}<0,\ldots,\gamma_{p+n}<0,\gamma_{p+n+1}=0,\ldots,\gamma_m=0.$$

Выберем новые переменные

$$\widetilde{x}_i = \begin{cases} \sqrt{\gamma_i}\, x_i, & 1 \le i \le p; \\ \sqrt{-\gamma_i}\, x_i, & p+1 \le i \le p+n; \\ x_i, & p+n+1 \le i \le m. \end{cases}$$

Тогда $Q(x) = \widetilde{x}_1^{\,2} + \ldots + \widetilde{x}_p^2 - \widetilde{x}_{p+1}^2 - \ldots - \widetilde{x}_{p+n}^2$ – другой канонический вид той же квадратичной формы. Он получен из предыдущего действительным линейным невырожденным преобразованием независимых переменных.♦

**Определение**. Канонический вид квадратичной формы с коэффициентами $-1, 0, +1$ называют ее **нормальным** видом.♦

Что общего у разных канонических видов, к которым приводится одна и та же квадратичная форма? По следствию из теоремы 1 для любых двух ее канонических видов $Q(x) = \gamma_1 x_1^2 + \gamma_2 x_2^2 + \ldots + \gamma_m x_m^2 = \beta_1 \widetilde{x}_1^2 + \beta_2 \widetilde{x}_2^2 + \ldots + \beta_m \widetilde{x}_m^2$ задающие их диагональные матрицы имеют одинаковые ранги. Поэтому число ненулевых членов во всех канонических видах $Q(x)$ одинаково; оно равно рангу матрицы квадратичной формы и называется **рангом квадратичной формы**. Но этим не исчерпывается общность канонических видов одной и той же $Q(x)$.

**Теорема 6**. (Закон инерции квадратичных форм.) При любом преобразовании квадратичной формы, определенной на действительном линейном пространстве, к её каноническому виду число $p$ положительных членов и число $n$ отрицательных членов в каноническом виде будут одними и теми же.

**Доказательство**. Пусть $\dim \mathbf{V} = m < \infty$, $e, f, g$ – три базиса линейного пространства $\mathbf{V}$, $f = eC$, $g = eG$, где $C$ и $G$ – матрицы перехода между базисами. Пусть в базисе $e$ (в переменных $\vec{x}_e = (x_1, \ldots, x_m)^T$) квадратичная форма имеет вид $Q(x) = \sum_{i=1}^{m} \sum_{k=1}^{m} a_{ik} x_i x_k$. Пусть в базисе $f$ (в переменных $\vec{x}_f = (\widetilde{x}_1, \ldots, \widetilde{x}_m)^T = C^{-1} \vec{x}_e$) она имеет канонический вид $Q(x) = \alpha_1 \widetilde{x}_1^{\,2} + \ldots + \alpha_{p_1} \widetilde{x}_{p_1}^2 - \alpha_{p_1+1} \widetilde{x}_{p_1+1}^2 - \ldots - \alpha_{p_1+n_1} \widetilde{x}_{p_1+n_1}^2$, где все коэффициенты $\alpha_i > 0$. Пусть в базисе $g$ (в переменных $\vec{x}_g = \left(\tilde{\tilde{x}}_1, \ldots, \tilde{\tilde{x}}_m\right)^T = G^{-1}\vec{x}_e$) она имеет другой канонический вид $Q(x) = \beta_1\tilde{\tilde{x}}_1^2 + \ldots + \beta_{p_2}\tilde{\tilde{x}}_{p_2}^2 - \beta_{p_2+1}\tilde{\tilde{x}}_{p_2+1}^2 - \ldots - \beta_{p_2+n_2}\tilde{\tilde{x}}_{p_2+n_2}^2$, где все коэффициенты $\beta_i > 0$.

Для любого элемента $x \in \mathbf{V}$ имеем равенство

$$\alpha_1\tilde{x}_1^2 + \ldots + \alpha_{p_1}\tilde{x}_{p_1}^2 + \beta_{p_2+1}\tilde{\tilde{x}}_{p_2+1}^2 + \ldots + \beta_{p_2+n_2}\tilde{\tilde{x}}_{p_2+n_2}^2 =$$

$$= \alpha_{p_1+1}\tilde{x}_{p_1+1}^2 + \ldots + \alpha_{p_1+n_1}\tilde{x}_{p_1+n_1}^2 + \beta_1\tilde{\tilde{x}}_1^2 + \ldots + \beta_{p_2}\tilde{\tilde{x}}_{p_2}^2. \quad (1)$$

Предположим, что $p_1 < p_2$. Тогда рассмотрим те элементы $x \in \mathbf{V}$, для которых

$$\begin{gathered} \tilde{x}_1 = 0, \ldots, \tilde{x}_{p_1} = 0; \tilde{\tilde{x}}_{p_2+1} = 0, \ldots \\ \ldots, \tilde{\tilde{x}}_{p_2+n_2} = 0,\ \tilde{\tilde{x}}_{p_2+n_2+1} = 0, \ldots, \tilde{\tilde{x}}_m = 0. \end{gathered} \quad (2)$$

Поскольку $p_1 < p_2$, число условий в (2) меньше, чем $m$. При помощи линейных преобразований $C^{-1}$ и $G^{-1}$ выразим переменные $\tilde{x}_i$ и переменные $\tilde{\tilde{x}}_i$ через переменные $x_i$. Тогда (2) является однородной системой линейных алгебраических уравнений относительно координат $(x_1, \ldots, x_m)$ элементов $x$ в базисе $e$. В этой системе число уравнений меньше, чем число $m$ неизвестных. Поэтому система уравнений имеет ненулевое решение $(x_1, \ldots, x_m)$. То есть, в пространстве $\mathbf{V}$ существует элемент $x \neq 0$, для которого выполнены условия (2).

С другой стороны, если некоторый элемент $x \in \mathbf{V}$ удовлетворяет условиям (2), то в равенстве (1) левая часть равна нулю. Тогда из (1) получаем $\tilde{\tilde{x}}_1 = 0, \ldots, \tilde{\tilde{x}}_{p_2} = 0$. Итак, если элемент $x \in \mathbf{V}$ удовлетворяет (2), то в базисе $g$ **все** его координаты $\tilde{\tilde{x}}_1 = 0, \ldots, \tilde{\tilde{x}}_m = 0$, т.е. $x = 0$. – Противоречие! Следовательно, неравенство $p_1 < p_2$ невозможно.

Точно так же доказывается, что невозможно неравенство $p_2 < p_1$. Точно так же доказывается, что невозможны неравенства $n_1 < n_2$ и $n_2 < n_1$.♦

Число $p$ положительных членов в каноническом виде квадратичной формы называют ее **положительным индексом инерции**. Число $n$ отрицательных членов в каноническом виде квадратичной формы называют ее **отрицательным индексом инерции**. Пару чисел $(p,n)$ называют **сигнатурой** квадратичной формы. (Иногда сигнатурой называют число $p-n$.) Ясно, что сумма $p+n$ равна рангу квадратичной формы. Если $p+n=\dim \mathbf{V}$, то квадратичную форму называют **невырожденной**; если $p+n<\dim \mathbf{V}$, то – **вырожденной**.

Квадратичную форму $Q(x)$ называют **неотрицательной**, если $Q(x)\geq 0$ для всех $x\in \mathbf{V}$. Квадратичную форму называют **строго положительной** (или положительно определенной), если для любого $x\in \mathbf{V}$ из условия $x\neq 0$ следует $Q(x)>0$.

**Теорема 7**. Квадратичная форма является положительно определенной в том и только в том случае, если ее положительный индекс инерции $p=\dim \mathbf{V}$.

**Доказательство**. Пусть $\dim \mathbf{V}=m$, и в некотором базисе $e$ квадратичная форма $Q(x)$ приведена к нормальному виду: $Q(x)=x_1^2+\ldots+x_p^2-x_{p+1}^2-\ldots-x_{p+n}^2+0\cdot x_{p+n+1}^2+\ldots+0\cdot x_m^2$. Если $Q(x)$ является положительно определенной, то в ее нормальном виде нет отрицательных или равных нулю членов. В противном случае для некоторого элемента $y\in \mathbf{V}$, координаты которого в базисе $e$ равны, например, $\left(\underbrace{0,\ldots,0}_{p},\underbrace{1,\ldots,1}_{n},\underbrace{0,\ldots,0}_{m-p-n}\right)$, выполнено неравенство $Q(y)<0$; а для некоторого $z\in \mathbf{V}$, координаты которого в базисе $e$ равны, например,

$\left(\underbrace{0,\ldots,0}_{p},\underbrace{0,\ldots,0}_{n},\underbrace{1,\ldots,1}_{m-p-n}\right)$, выполнено равенство $Q(z)=0$, хотя $z\neq 0$. Поэтому $p=m$.

Если же в нормальном виде квадратичной формы $p=m$, т.е. $Q(x)=x_1^2+\ldots+x_m^2$, то $Q(x)\geq 0$ для любого $x\in\mathbf{V}$, причем $Q(x)=0$ только при $x=0$.♦

Важность класса положительно определенных квадратичных форм объясняется следующим обстоятельством. Пусть $Q(x)$ – **положительно определенная** квадратичная форма, а $B(x;y)$ – ее полярная билинейная форма. Тогда $B(x;y)$ обладает свойствами:

1. $B(x;y)=B(y;x)$ для любых $x,y\in\mathbf{V}$.
2. $B(x+z;y)=B(x;y)+B(z;y)$ для любых $x,y,z\in\mathbf{V}$.
3. $B(\alpha x;y)=\alpha B(x;y)$ для любых $x,y\in\mathbf{V}$, $\alpha\in\mathbf{R}$.
4. $B(x;x)\geq 0$, причем $B(x;x)=0$ только для $x=0\in\mathbf{V}$.

Тем самым $B(x;y)$ удовлетворяет всем аксиомам скалярного произведения в действительном линейном пространстве. Справедливо

**Утверждение**. Скалярное произведение в действительном линейном пространстве является билинейной формой, полярной к **положительно определенной** квадратичной форме. Любая такая билинейная форма может быть принята за скалярное произведение.♦

Квадратичную форму $Q(x)$ естественно назвать **неположительной**, если $Q(x)\leq 0$ для всех $x\in\mathbf{V}$, и **строго отрицательной** (или отрицательно определенной), если из условия $x\neq 0$ следует $Q(x)<0$. Очевидно, что $Q(x)$ является отрицательно определенной в том и только в том случае, если ее отрицательный индекс инерции $n=\dim\mathbf{V}$.

Квадратичную форму называют **знакопеременной**, если для некоторого $x\in\mathbf{V}$ $Q(x)>0$, а для некоторого $y\in\mathbf{V}$ $Q(y)<0$.

**Задача**. В терминах индексов инерции найдите необходимые и достаточные условия того, что $Q(x)$ неотрицательна; $Q(x)$ неположительна; $Q(x)$ знакопеременна. ♦

3. **Методы приведения квадратичной формы к каноническому виду.**

**Приведение квадратичной формы к главным осям** уже было рассмотрено в предыдущем разделе. Будем считать, что в $\mathbf{R}^m$ введено скалярное произведение $(\vec{x}, \vec{y}) = \sum_{i=1}^{m} x_i y_i$. Определённую на $\mathbf{R}^m$ квадратичную форму $Q(\vec{x}) = \sum_{i=1}^{m}\sum_{k=1}^{m} a_{ik} x_i x_k$ можно считать заданной в ортонормированном базисе $e_1 = (1,0,\ldots,0)^T$, $e_2 = (0,1,0,\ldots,0)^T, \ldots, e_m = (0,\ldots,0,1)^T \in \mathbf{R}^m$. Найдем корни $\lambda_1,\ldots,\lambda_m$ характеристического многочлена $\det(A - \lambda E)$; поскольку действительная матрица квадратичной формы $A = A^T$, все $m$ корней действительны. Если $\lambda_k$ является собственным значением кратности 1, то найдём отвечающий ему собственный вектор и нормируем его. Если кратность собственного значения $\lambda_k$ больше, чем 1, то найдём систему всех линейно независимых собственных векторов, отвечающих $\lambda_k$; построим по этой системе **ортонормированную** систему отвечающих $\lambda_k$ собственных векторов. Поступая так для всех $\lambda_k$, мы найдем матрицу $C_0$ требуемого ортогонального преобразования.

**Метод Лагранжа** состоит в последовательном **выделении полных квадратов**. В этом методе строится линейное невырожденное преобразование (от $\vec{x}_e$ к $\vec{x}_f$), которое не обязательно окажется ортогональным.

Прежде всего заметим, что при приведении квадратичной формы $Q(\vec{x}) = \sum_{i=1}^{m}\sum_{k=1}^{m} a_{ik} x_i x_k$ к каноническому виду можно добиться наличия в ней коэффициента $a_{kk} \neq 0$. Действительно, если $Q(\vec{x}) \not\equiv 0$ и не содержит ни одного $x_k^2$, то в $Q$ имеется хотя бы одно произведение $2a_{ik}x_ix_k$, причём $a_{ii} = 0$ и $a_{kk} = 0$. Выполним замену переменных $\begin{cases} x_i = \widetilde{x}_i + \widetilde{x}_k \\ x_k = \widetilde{x}_i - \widetilde{x}_k, \end{cases}$, а остальные переменные оставим без изменения. Тогда $2a_{ik}x_ix_k = 2a_{ik}\left(\widetilde{x}_i^2 - \widetilde{x}_k^2\right)$ войдёт в квадратичную форму (и не сократится с другими членами: $a_{ii} = 0$, $a_{kk} = 0$).

Будем считать, что $a_{11} \neq 0$. Выделим все члены, содержащие $x_1$: $a_{11}x_1^2 + 2a_{12}x_1x_2 + \cdots + 2a_{1m}x_1x_m$. Дополним эти члены до полного квадрата, т.е. запишем их в виде $\dfrac{1}{a_{11}}\left(a_{11}x_1 + \cdots + a_{1m}x_m\right)^2 - R$, где $R$ содержит только квадраты и попарные произведения членов $a_{12}x_2, \cdots, a_{1m}x_m$ и не содержит $x_1$ и $x_1^2$. Положим $\widetilde{x}_1 = a_{11}x_1 + \cdots + a_{1m}x_m$, а все остальные переменные оставим без изменения: $\widetilde{x}_i = x_i$, $i = 2, \ldots, m$. Тогда $Q = \dfrac{1}{a_{11}}\widetilde{x}_1^2 + Q_1$, где $Q_1 = \sum\limits_{i=2}^{m}\sum\limits_{k=2}^{m}\widetilde{a}_{ik}\widetilde{x}_i\widetilde{x}_k$. С квадратичной формой $Q_1$ поступим точно так же. И так далее.

**Метод Якоби** состоит в построении невырожденного линейного преобразования (от $\vec{x}_e$ к $\vec{x}_f$) с **треугольной** матрицей. Пусть $B(\vec{x};\vec{y})$ – билинейная форма, полярная к квадратичной форме $Q(\vec{x})$: $B(\vec{x};\vec{y}) = \dfrac{1}{2}\left(Q(\vec{x}+\vec{y}) - Q(\vec{x}) - Q(\vec{y})\right)$. Пусть матрица $A$ этой билинейной формы в базисе $e$ такова, что все стоящие в её левом верхнем углу миноры

$$\Delta_1 = a_{11} \neq 0\,, \Delta_2 = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} \neq 0\,,\ \ldots,\ \Delta_m = \begin{vmatrix} a_{11}\ a_{12} \ldots a_{1m} \\ \cdot \quad \cdot \quad \ldots \quad \cdot \\ a_{m1}\ a_{m2} \ldots a_{mm} \end{vmatrix} \neq 0\,.$$

Мы приведем квадратичную форму $Q$ к каноническому виду, если построим базис $f$, все элементы которого удовлетворяют условиям $B(f_i; f_k) = 0$ при $i \neq k$. Процесс построения такого базиса $f$ по исходному базису $e$ совпадает по смыслу с процессом ортогонализации линейно независимой системы элементов; в нём надо только заменить все скалярные произведения $(x, y)$ элементов $x$ и $y$ на значения билинейной формы $B(x; y)$. Элементы $f_1, \ldots, f_m$ строим последовательно по правилу ( $f = eC$ )

$$f_1 = c_{11}e_1,$$
$$f_2 = c_{12}e_1 + c_{22}e_2,$$
$$\ldots$$
$$f_m = c_{1m}e_1 + c_{2m}e_2 + \ldots + c_{mm}e_m,$$

подчинив выбор коэффициентов $c_{ik}$ требованиям $B(f_k; e_i) = 0$ при всех $i < k$, а $B(f_k; e_k) = 1$ – для всех $k = 1, \ldots, m$ В базисе $f$ квадратичная форма $Q$ будет иметь канонический вид
$$Q = \frac{1}{\Delta_1}\widetilde{x}_1^{\,2} + \frac{\Delta_1}{\Delta_2}\widetilde{x}_2^{\,2} + \ldots + \frac{\Delta_{m-1}}{\Delta_m}\widetilde{x}_m^{\,2}.$$

Метод Якоби применим и в случае, если $rang\, A = r < m$; тогда надо предполагать, что $\Delta_1 \neq 0$, $\Delta_2 \neq 0, \ldots, \Delta_r \neq 0$.

**Задача**. Приведите квадратичную форму $Q(x) = x_1x_2 + x_1x_3 + x_2x_3$ к каноническому виду методом Лагранжа; найдите её нормальный вид.

**Задача**. Считая, что та же квадратичная форма задана в ортонормированном базисе, приведите её к главным осям ортогональным преобразованием переменных.

**Задача**. Квадратичную форму $Q(x)=2x_1^2+x_2^2+x_3^2+3x_1x_2+4x_1x_3$ приведите к каноническому виду методом Якоби.

4. **Замечания**.

1. Если $\mathbf{V}$ – конечномерное действительное линейное пространство, то гиперповерхностью второго порядка называют геометрическое место точек $x\in\mathbf{V}$, удовлетворяющих уравнению $Q(x)+2l(x)+c=0$, где $Q(x)$ – определённая на $\mathbf{V}$ ненулевая квадратичная форма, $l(x)$ – линейная форма, $c=const$. Классификация таких гиперповерхностей (например, кривых в $\mathbf{R}^2$, поверхностей в $\mathbf{R}^3$) основана на приведении $Q(x)$ к каноническому виду.

2. Дифференциальные уравнения в частных производных второго порядка, линейные относительно старших производных, с действительными коэффициентами можно классифицировать и приводить к наипростейшему виду при помощи приведения к каноническому виду отвечающих им квадратичных форм.

3. В законе инерции квадратичных форм существенной является упорядоченность поля $\mathbf{R}$ действительных чисел, над которым задано линейное пространство $\mathbf{V}$. Однако и в комплексном линейном пространстве имеется класс квадратичных форм, для которых справедлив закон инерции. Это так называемые эрмитовы квадратичные формы, которые характеризуются тем, что принимают лишь действительные значения; для задания такой квадратичной формы вместо симметричной билинейной формы надо выбрать эрмитову полуторалинейную форму (в определении билинейной симметричной формы равенства $B(x;\alpha y)=\alpha B(x;y)$ и $B(x;y)=B(y;x)$ надо заменить на $B(x;\alpha y)=\overline{\alpha}B(x;y)$ и $B(x;y)=\overline{B(y;x)}$ – для любых $x,y\in\mathbf{V}$ и любого $\alpha\in\mathbf{C}$).

## Вопрос 7.

## Принцип сжимающих отображений в полных метрических пространствах. Примеры применения.

1. Пусть $M$ – произвольное множество. Говорят, что в $M$ введено расстояние между его элементами, если каждой паре элементов (точек) $x$, $y$ поставлено в соответствие действительное число $\rho(x,y)$, причем выполнены следующие аксиомы:

1. $\rho(x,y)=\rho(y,x)$ для любых $x,y\in M$;
2. $\rho(x,y)>0$, если $x\neq y$; $\rho(x,y)=0$, если $x=y$;
3. $\rho(x,y)\le\rho(x,z)+\rho(z,y)$ для любых $x,y,z\in M$.

Такая функция $\rho(x,y)$ двух переменных $x,y\in M$ называется **расстоянием** или **метрикой**. Множество $M$ с метрикой $\rho$ называется **метрическим пространством**. Наличие метрики во множестве $M$ позволяет ввести понятия предела последовательности элементов этого множества и фундаментальной последовательности его элементов. Из третьей аксиомы метрики следует, что всякая сходящаяся последовательность фундаментальна. Однако, можно привести примеры метрических пространств, в которых некоторые фундаментальные последовательности не имеют предела (приведите примеры!). Если в метрическом пространстве $M$ любая фундаментальная последовательность его элементов имеет предел (принадлежащий $M$), то $M$ называется **полным** метрическим пространством. Если пространство $M$ не полно, то его можно включить некоторым способом в полное метрическое пространство, которое называется **пополнением** пространства $M$ (это пополнение по существу единственно). Например, пространство всех действительных чисел является пополнением пространства всех рациональных чисел.

Если $M$ – линейное нормированное пространство, то $\rho(x,y)=\|x-y\|$ удовлетворяет всем аксиомам метрики. **Полное** нормированное пространство называют банаховым.

2. Если отображение $A$ действует из метрического пространства $M$ в $M$, то точка $x\in M$, для которой $Ax=x$,

называется **неподвижной** точкой отображения $A$. Многие теоретические и практические задачи сводятся к отысканию неподвижных точек некоторого отображения.

**Пример**. Систему $k$ линейных алгебраических уравнений

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ \ . \ \ . \ \ . \ \ + a_{1k}x_k = b_1, \\ a_{21}x_1 + a_{22}x_2 + \ \ . \ \ . \ \ . \ \ + a_{2k}x_k = b_2, \\ . \ . \ . \ . \ . \ . \ . \ . \ . \ . \ . \ . \ . \\ a_{k1}x_1 + a_{k2}x_2 + \ \ . \ \ . \ \ . \ \ + a_{kk}x_k = b_k \end{cases}$$

с $k$ неизвестными $x_1, \ldots, x_k$ можно записать в эквивалентной форме

$$\begin{cases} x_1 = (1 - a_{11})x_1 - a_{12}x_2 - \ \ . \ \ . \ \ . \ \ . \ \ - a_{1k}x_k + b_1, \\ x_2 = -a_{21}x_1 + (1 - a_{22})x_2 - \ \ . \ \ . \ \ . \ \ . \ \ - a_{2k}x_k + b_2, \\ . \ . \ . \ . \ . \ . \ . \ . \ . \ . \ . \ . \ . \\ x_k = -a_{k1}x_1 - a_{k2}x_2 - \ \ . \ \ . \ \ . \ \ . \ \ + (1 - a_{kk})x_k + b_k. \end{cases}$$

Полагая $c_{ij} = \delta_{ij} - a_{ij}$, где $\delta_{ij} = \begin{cases} 1 & при \ i = j, \\ 0 & при \ i \neq j, \end{cases}$ получим ту же систему в более короткой записи:

$$x_i = \sum_{j=1}^{k} c_{ij} x_j + b_i, \quad i = 1, \ldots, k.$$

Введем отображение $A$ пространства $R^k$ в $R^k$: $Ax = Cx + b$, где $x = (x_1, \ldots, x_k)^T \in R^k$, $C$ – $n \times n$-матрица с действительными элементами $c_{ij}$, $b = (b_1, \ldots, b_k)^T \in R^k$ – фиксированный вектор. Тогда решения исходной системы уравнений суть неподвижные точки отображения $A$. $\square$

Для нахождения неподвижных точек часто применяют **метод последовательных приближений**. Пусть отображение $A$ действует из некоторого замкнутого множества $V$ пространства $M$ в $V$ и непрерывно на $V$. Выберем произвольную точку $x_0 \in V$ и определим последовательность точек $x_1 = Ax_0$, $x_2 = Ax_1$, ...., $x_n = Ax_{n-1}, \ldots$. Если $x_0$ – неподвижная точка отображения $A$, то

$x_0 = x_1 = x_2 = \ldots = x_n = \ldots$; но при произвольном выборе $x_0$, как правило, $x_n \neq x_{n-1}$. Однако, по условию, если $x_{n-1} \in V$, то и $x_n \in V$, т. е. построение бесконечной последовательности точек $\{x_n\}$ возможно. Если окажется, что существует $x = \lim\limits_{n\to\infty} x_n$, то вследствие замкнутости $V$, $x \in V$. Переходя к пределу при $n \to \infty$ в равенстве $x_n = Ax_{n-1}$ и пользуясь непрерывностью отображения $A$, получим $x = Ax$. Возникают вопросы о выборе начального приближения $x_0$, о сходимости метода, о скорости сходимости. (В отношении самого уравнения $x = Ax$ прежде всего ставятся вопросы о существовании и о единственности его решения.) Ответы на эти вопросы легко дать для важного класса отображений, которые называются сжимающими.

**Определение**. Отображение $A$ метрического пространства $M$ в $M$ называется **сжимающим**, если существует такое число $\alpha$, $0 \le \alpha < 1$, что для любых двух точек $x, y \in M$ выполнено неравенство $\rho(Ax, Ay) \le \alpha \cdot \rho(x, y)$. ▯

В этом определении коэффициент сжатия $\alpha$ не зависит от $x$, $y$.

Очевидно, всякое сжимающее отображение непрерывно.

3. **Теорема 1 (принцип сжимающих отображений)**. Пусть $M$ – **полное** метрическое пространство. Если сжимающее отображение $A$ действует из $M$ в $M$, то оно имеет в $M$ неподвижную точку, и притом единственную.

**Доказательство**. Выберем произвольную точку $x_0 \in M$ и определим последовательность точек $x_1 = Ax_0$, $x_2 = Ax_1$, …., $x_n = Ax_{n-1} = A^n x_0, \ldots$ . Докажем, что эта последовательность фундаментальна. Пусть $m \ge n$. Тогда

$$\rho(x_n, x_m) = \rho(A^n x_0, A^m x_0) = \rho(A^n x_0, A^n A^{m-n} x_0) \le$$

$$\le \alpha^n \cdot \rho(x_0, x_{m-n}) \le$$

$$\le \alpha^n \cdot \left(\rho(x_0, x_1) + \rho(x_1, x_2) + \ldots + \rho(x_{m-n-1}, x_{m-n})\right) \le$$

$$\le \alpha^n \cdot \left(1+\alpha+\alpha^2+...+\alpha^{m-n-1}\right) \cdot \rho(x_0,x_1) \le \frac{\alpha^n}{1-\alpha} \cdot \rho(x_0,x_1).$$

Поскольку коэффициент сжатия $0 \le \alpha < 1$, правая часть полученного неравенства стремится к нулю при $n \to \infty$. Следовательно, $\{x_n\}$ фундаментальна.

Из полноты пространства $M$ вытекает существование в $M$ точки $x$, являющейся пределом последовательности $\{x_n\}$. Поскольку сжимающее отображение $A$ непрерывно, имеем $Ax = A\left(\lim\limits_{n\to\infty} x_n\right) = \lim\limits_{n\to\infty} Ax_n = \lim\limits_{n\to\infty} x_{n+1} = x$, т.е. $x$ – неподвижная точка отображения $A$.

Докажем единственность неподвижной точки. Пусть наряду с $x$ ещё и $y$ является неподвижной точкой отображения $A$: $Ax = x$, $Ay = y$. Тогда $\rho(x,y) = \rho(Ax,Ay) \le \alpha \cdot \rho(x,y)$. Из условия $\alpha < 1$ получаем $\rho(x,y) = 0$, т.е. $x = y$. □

**Замечание.** Если в неравенстве $\rho(x_n,x_m) \le \dfrac{\alpha^n}{1-\alpha} \cdot \rho(x_0,x_1)$ перейти к пределу при $m \to \infty$, то получим оценку погрешности $n$-го приближения: $\rho(x_n,x) \le \dfrac{\alpha^n}{1-\alpha} \cdot \rho(x_0,Ax_0)$. Эта погрешность убывает со скоростью геометрической прогрессии со знаменателем $\alpha$ и зависит от выбора начального приближения $x_0$. Для сжимающего отображения метод последовательных приближений сходится к неподвижной точке $x$ при любом начальном приближении $x_0$. □

**Замечание**. В доказанной теореме нельзя отказаться от полноты пространства $M$ (приведите пример неполного пространства $M$ и сжимающего отображения в нём, которое не имеет в $M$ неподвижной точки). Теорема перестанет быть справедливой и в случае, если от $A$ потребовать лишь выполнения условия $\rho(Ax,Ay) < \rho(x,y)$ для всех $x \ne y$. Например, если $M = R^1$,

$\rho(x,y)=|x-y|$, $Ax=\dfrac{\pi}{2}+x-arctg\,x$, то $A$ не имеет неподвижных точек, хотя

$$\rho(Ax,Ay)=|Ax-Ay|=|A'(\xi)|\cdot|x-y|=$$
$$=\left|\frac{\xi^2}{1+\xi^2}\right|\cdot|x-y|<|x-y|=\rho(x,y)$$

(мы применили формулу Лагранжа). Другой пример: $M=\{x\in R^1 \mid x\geq 1\}$, $\rho(x,y)=|x-y|$, $Ax=x+\dfrac{1}{x}$. ▯

**Теорема 2**. Пусть отображение $A$ преобразует замкнутое множество $M$ в свою компактную часть и удовлетворяет на $M$ условию $\rho(Ax,Ay)<\rho(x,y)$ для всех $x\neq y$. Тогда $A$ имеет в $M$ единственную неподвижную точку.

**Доказательство**. Пусть отображение $A$ переводит замкнутое множество $M$ в относительно компактное множество $V\subset M$. Рассмотрим $A$ на компакте $\overline{V}$ (на замыкании множества $V$). В силу непрерывности отображения $A$ непрерывен и функционал $f(z)=\rho(z,Az)$. Поэтому в некоторой точке $x\in\overline{V}$ функционал $f$ принимает наименьшее значение. Это наименьшее значение равно нулю: иначе мы имели бы

$$f(Ax)=\rho(Ax,A^2x)<\rho(x,Ax)=f(x)=\min_{z\in\overline{V}}f(z).$$

$f(x)=\rho(x,Ax)=0$ означает $Ax=x$. Если ещё $y$ была бы неподвижной точкой отображения $A$ в $M$, то мы имели бы $\rho(x,y)=\rho(Ax,Ay)<\rho(x,y)$, что невозможно. ▯

Отметим, что отображение $A$, удовлетворяющее условиям теоремы 2, не обязательно является сжимающим. Например, $M=\{x\in R^1 \mid 0\leq x\leq 1\}$, $\rho(x,y)=|x-y|$, $Ax=x-\dfrac{x^2}{2}$.

Часто удобно рассматривать отображения, которые определены и являются сжимающими не на всем пространстве $M$, а только на некотором его шаре.

**Теорема 3**. Пусть $A$ является сжимающим отображением, определённым на замкнутом шаре $\mathrm{B}=\{x\in M \mid \rho(x,y_0)\le r\}$ полного метрического пространства $M$, причём $\rho(Ay_0,y_0)\le(1-\alpha)\cdot r$ ($\alpha$ – коэффициент сжатия отображения $A$). Тогда $A$ имеет в $\mathrm{B}$ единственную неподвижную точку.

**Доказательство**. Достаточно проверить, что $A\mathrm{B}\subset\mathrm{B}$. Если $x\in\mathrm{B}$, т.е. $\rho(x,y_0)\le r$, то имеем неравенства

$$\rho(Ax,y_0)\le\rho(Ax,Ay_0)+\rho(Ay_0,y_0)\le$$
$$\le\alpha\cdot\rho(x,y_0)+(1-\alpha)\cdot r\le r,$$

т.е. $Ax\in\mathrm{B}$. Замкнутое множество полного метрического пространства само есть полное метрическое пространство. Поэтому имеем сжимающее отображение $A$, действующее из полного метрического пространства $\mathrm{B}$ в $\mathrm{B}$. Оно имеет в $\mathrm{B}$ единственную неподвижную точку. □

**Пример**. $M=R^1$, $\rho(x,y)=|x-y|$, $Ax=x^3$. Отображение $A$ имеет неподвижные точки -1, 0, 1. Оно является сжимающим на шаре $\mathrm{B}=\left\{x\in R^1 \,\middle|\, |x|\le r\right\}$, где $r<\dfrac{\sqrt{3}}{3}$. В окрестностях неподвижных точек –1 и 1 оно не является сжимающим. □

**Пример**. $M=C[0,1]$, $\rho(x,y)=\|x-y\|_{C[0,1]}$, $Ax=\int\limits_0^1 x^2(t)dt$. В шаре $\|x\|_{C[0,1]}\le\dfrac{1}{4}$ отображение $A$ является сжимающим и поэтому имеет единственную неподвижную точку $x(t)\equiv 0$. Но уравнение $Ax=x$ имеет в $C[0,1]$ второе решение $x(t)\equiv 1$. □

Таким образом, из единственности неподвижной точки отображения в некотором пространстве не следует единственность его неподвижной точки в более широком пространстве.

**Теорема 4**. Пусть $A$ – непрерывное отображение полного метрического пространства $M$ в $M$. Пусть при некотором натуральном $m > 1$ отображение $A^m$ является сжимающим. Тогда отображение $A$ имеет одну и только одну неподвижную точку.

**Доказательство**. Существует точка $x \in M$, для которой $\rho(A^m x, x) = 0$. К этой точке $x$ сходится последовательность $x_n = (A^m)^n x_0$ при любом выборе начального приближения $x_0$. Очевидны равенства $Ax_n = AA^m x_{n-1} = A^m Ax_{n-1}$. Поскольку $\{x_n\} \underset{n\to\infty}{\longrightarrow} x$, а отображения $A$ и $A^m$ непрерывны, имеем

$$\rho(Ax, A^m Ax) = 0.$$

Тогда

$$\rho(Ax, x) \le \rho(Ax, A^m Ax) + \rho(A^m Ax, A^m x) + \rho(A^m x, x) =$$
$$= \rho(A^m Ax, A^m x) \le \alpha \cdot \rho(Ax, x),$$

где $\alpha$ – коэффициент сжатия отображения $A^m$; $0 \le \alpha < 1$. Отсюда $\rho(Ax, x) = 0$. Единственность неподвижной точки отображения $A$ следует из единственности неподвижной точки отображения $A^m$, поскольку всякая точка, неподвижная относительно $A$, неподвижна и относительно $A^m$. □

Доказанная теорема допускает простое обобщение:

**Терема 5**. Пусть отображение $A$ действует из полного метрического пространства $M$ в $M$ и коммутирует с определенным на $M$ сжимающим отображением $B$ ($ABx = BAx$ для любого $x \in M$). Тогда неподвижная точка отображения $B$ является и неподвижной точкой (возможно, неединственной) отображения $A$.

**Доказательство**. Если $Bx = x$, то $BAx = ABx = Ax$, т.е. $Ax$ является неподвижной точкой отображения $B$. У сжимающего отображения $B$ неподвижная точка единственна, поэтому $Ax = x$. □

4. **Примеры**. Пусть $f$ – определенная на сегменте $[a,b]$ действительная функция, которая отображает $[a,b]$ в $[a,b]$ и удовлетворяет условию Липшица $|f(x) - f(y)| \le \alpha \cdot |x - y|$ для всех $x, y \in [a,b]$. Если $\alpha < 1$, то $f$ является сжимающим

отображением, и уравнение $f(x)=x$ имеет на $[a,b]$ единственное решение. Это решение можно найти с любой заданной точностью методом последовательных приближений. В частности, если указанная функция $f$ дифференцируема, а её производная удовлетворяет на $[a,b]$ неравенству $|f'(x)|\le \alpha <1$, то $f$ – сжимающее отображение.

Эта простая идея переносится и на случай уравнения $f(x)=x$ в банаховом пространстве. Пусть отображение $f$ определено на замкнутом **выпуклом** множестве $V$ банахова пространства $M$ и действует из $V$ в $V$. Если $f$ непрерывно дифференцируемо (по Фреше) на $V$, а его производная удовлетворяет на $V$ неравенству $\|f'(x)\|\le \alpha <1$, то уравнение $f(x)=x$ имеет в $V$ единственное решение.

Задачу решения уравнения $g(x)=0$ часто сводят к задаче о неподвижной точке некоторого отображения: уравнение $g(x)=0$ равносильно уравнению $f(x)=x$, где $f(x)=x-\lambda\cdot g(x)$.

5. **Примеры**. Будем решать систему линейных алгебраических уравнений $x=Cx+b$, рассмотренную в п. 2. Отображение $Ax\equiv Cx+b$ действует из $R^k$ в $R^k$; $x=(x_1,\ldots,x_k)^T\in R^k$.

$M=R^k$, $\rho(x,y)=\|x-y\|_1=\sum_{i=1}^{k}|x_i-y_i|$. Докажем, что отображение $A$ является сжимающим в том, и только в том случае, если элементы матрицы $C$ удовлетворяют условию

$$\max_{1\le j\le k}\sum_{i=1}^{k}|c_{ij}|\le \alpha <1. \qquad (1)$$

Пусть условие (1) выполнено. Тогда

$$\rho(Ax,Ay)=\sum_{i=1}^{k}\left|\sum_{j=1}^{k}c_{ij}(x_j-y_j)\right|\le \sum_{i=1}^{k}\sum_{j=1}^{k}|c_{ij}|\cdot|x_j-y_j|=$$

$$= \sum_{j=1}^{k} \left( \sum_{i=1}^{k} |c_{ij}| \right) \cdot |x_j - y_j| \le \left( \max_{1 \le j \le k} \sum_{i=1}^{k} |c_{ij}| \right) \cdot \rho(x, y).$$

и $A$ является сжимающим. Докажем, что (1) и необходимо. Пусть для всех $x, y \in R^k$ выполнено $\rho(Ax, Ay) \le \alpha \cdot \rho(x, y)$, где $\alpha < 1$. Выберем векторы $x = x(j_0) = (0, \ldots, 0, 1, 0, \ldots, 0)^T$ (единственная единица находится на $j_0$-ом месте) и $y = (0, \ldots, 0)^T$. Тогда $\rho(x, y) = 1$ и $\rho(Ax, Ay) = \sum_{i=1}^{k} |c_{ij_0}| \le \alpha \cdot \rho(x, y) = \alpha < 1$. Так поступим для всех $j_0 = 1, \ldots, k$; отсюда получаем (1).

$M = R^k$, $\rho(x, y) = \|x - y\|_\infty = \max_{1 \le i \le k} |x_i - y_i|$. Докажем, что отображение $A$ является сжимающим в том, и только в том случае, если элементы матрицы $C$ удовлетворяют условию

$$\max_{1 \le i \le k} \sum_{j=1}^{k} |c_{ij}| \le \alpha < 1. \tag{2}$$

Пусть условие (2) выполнено. Тогда

$$\rho(Ax, Ay) = \max_{1 \le i \le k} \left| \sum_{j=1}^{k} c_{ij} \cdot (x_j - y_j) \right| \le \max_{1 \le i \le k} \sum_{j=1}^{k} |c_{ij}| \cdot |x_j - y_j| \le$$

$$\le \left( \max_{1 \le i \le k} \sum_{j=1}^{k} |c_{ij}| \right) \cdot \left( \max_{1 \le j \le k} |x_j - y_j| \right) = \left( \max_{1 \le i \le k} \sum_{j=1}^{k} |c_{ij}| \right) \cdot \rho(x, y).$$

и $A$ является сжимающим. Докажем, что (2) и необходимо. Пусть для всех $x, y \in R^k$ выполнено $\rho(Ax, Ay) \le \alpha \cdot \rho(x, y)$, где $\alpha < 1$. Выберем векторы $x = x(i_0) = (sign\, c_{i_0 1}, sign\, c_{i_0 2}, \ldots, sign\, c_{i_0 k})^T$ и $y = (0, \ldots, 0)^T$. Тогда $\rho(x, y) = 1$ и

$$\rho(Ax, Ay) == \max_{1 \le i \le k} \left| \sum_{j=1}^{k} c_{ij} \cdot sign\, c_{i_0 j} \right| \ge \left| \sum_{j=1}^{k} c_{i_0 j} \cdot sign\, c_{i_0 j} \right| = \sum_{j=1}^{k} |c_{i_0 j}|.$$

Поэтому $\sum_{j=1}^{k} \left| c_{i_0 j} \right| \le \alpha < 1$. Так поступим для всех $i_0 = 1, \ldots, k$; отсюда получаем (2).

$$M = R^k, \qquad \rho(x, y) = \|x - y\|_2 = \left( \sum_{i=1}^{k} (x_i - y_i)^2 \right)^{\frac{1}{2}}.$$

Докажем, что условие

$$\sum_{i=1}^{k} \sum_{j=1}^{k} c_{ij}^2 < 1 \tag{3}$$

является достаточным, чтобы отображение $A$ было сжимающим.

$$\rho^2(Ax, Ay) = \sum_{i=1}^{k} \left( \sum_{j=1}^{k} c_{ij} \cdot (x_j - y_j) \right)^2 \le \left( \sum_{i=1}^{k} \sum_{j=1}^{k} c_{ij}^2 \right) \cdot \rho^2(x, y)$$

– по неравенству Коши-Буняковского (при каждом фиксированном $i$) для векторов $(c_{i1}, \ldots, c_{ik})$ и $((x_1 - y_1), \ldots, (x_k - y_k))$. Поэтому, если выполнено (3), то $A$ – сжимающее отображение. Условие (3) не является необходимым. Пусть, например, $k = 2$ и $C = \frac{\sqrt{3}}{2} \cdot \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$. Тогда $\rho(Ax, Ay) = \frac{\sqrt{3}}{2} \cdot \rho(x, y)$, но $\sum_{i=1}^{2} \sum_{j=1}^{2} c_{ij}^2 = \frac{3}{2} > 1$.

Систему уравнений $x = Cx + b$ можно решать методом последовательных приближений и в случае, когда $A$ не является сжимающим отображением. Если $\sup_{x \ne 0} \frac{\|Cx\|}{\|x\|} < 1$, то эта система имеет единственное решение, и метод последовательных приближений $x_n = Ax_{n-1}$ сходится к нему со скоростью геометрической прогрессии при любом $x_0$. Если система имеет единственное решение, то метод последовательных приближений будет сходиться к нему **при любом начальном приближении** в том, и только в том случае, если все собственные значения матрицы $C$ по модулю меньше 1.

6. **Пример**. Интегральное уравнение Фредгольма 2-го рода:

$$\lambda \cdot \int_p^q K(t,\tau)x(\tau)d\tau + b(t) = x(t), \qquad p \le t \le q. \tag{4}$$

Здесь $\lambda$ – параметр, $K(t,\tau)$ – заданная непрерывная на $[p,q]\times[p,q]$ функция, $b(t)$ – заданная непрерывная на $[p,q]$ функция, $x(t)$ – искомая функция. Пусть $N = \max\limits_{[p,q]\times[p,q]} |K(t,\tau)|$.

Отображение $Ax \equiv \lambda \cdot \int\limits_p^q K(t,\tau)x(\tau)d\tau + b(t)$ будем рассматривать в банаховом пространстве $C[p,q]$; $\|x\|_{C[p,q]} = \max\limits_{[p,q]} |x(t)|$, $\rho(x,y) = \|x-y\|_{C[p,q]}$. Очевидно неравенство $\rho(Ax,Ay) \le |\lambda| \cdot N \cdot (q-p) \cdot \rho(x,y)$. Поэтому, если $|\lambda| < \dfrac{1}{N\cdot(q-p)}$, то $A$ – сжимающее отображение.

Уравнение (4) можно решать и в других предположениях. Пусть его ядро $K \in L_2([p,q]\times[p,q])$, т.е. $\int\limits_p^q\int\limits_p^q K^2(t,\tau)dtd\tau < \infty$, $b \in L_2[p,q]$. Если $x \in L_2[p,q]$, то для почти всех $t \in [p,q]$ существует $\int\limits_p^q K(t,\tau)x(\tau)d\tau$, который является функцией из $L_2[p,q]$. Отображение $A$ можно рассматривать в гильбертовом пространстве $L_2[p,q]$.

$\rho(x,y) = \|x-y\|_{L_2[p,q]} = \left(\int\limits_p^q (x(t)-y(t))^2 dt\right)^{\frac{1}{2}}$. Из неравенства

$$\rho(Ax,Ay)=|\lambda|\cdot\left(\int_p^q\left(\int_p^q K(t,\tau)(x(\tau)-y(\tau))d\tau\right)^2 dt\right)^{\frac{1}{2}}\leq$$

$$\leq|\lambda|\cdot\left(\int_p^q\int_p^q K^2(t,\tau)dtd\tau\right)^{\frac{1}{2}}\cdot\left(\int_p^q(x(\tau)-y(\tau))^2 d\tau\right)^{\frac{1}{2}}$$

вытекает, что если $|\lambda|<\left(\int_p^q\int_p^q K^2(t,\tau)dtd\tau\right)^{-\frac{1}{2}}$, то $A$ – сжимающее отображение.

7. **Пример**. Нелинейное интегральное уравнение:

$$\lambda\cdot\int_p^q K(t,\tau;x(\tau))d\tau+b(t)=x(t),\qquad p\leq t\leq q\,.$$

Здесь $\lambda$ – параметр, заданные функции $K(t,\tau;u)$ и $b(t)$ непрерывны, причем $K$ удовлетворяет условию Липшица по переменной $u$: $|K(t,\tau;u_1)-K(t,\tau;u_2)|\leq\sigma\cdot|u_1-u_2|$.

Отображение $Ax\equiv\lambda\cdot\int_p^q K(t,\tau;x(\tau))d\tau+b(t)$ можно рассматривать как действующее из $C[p,q]$ в $C[p,q]$. Тогда выполнено неравенство $\rho(Ax,Ay)\leq|\lambda|\cdot\sigma\cdot(q-p)\cdot\rho(x,y)$.

Поэтому, если $|\lambda|<\dfrac{1}{\sigma\cdot(q-p)}$, то $A$ – сжимающее отображение.

8. **Пример**. Интегральное уравнение Вольтерра 2-го рода:

$$\lambda\cdot\int_p^t K(t,\tau)x(\tau)d\tau+b(t)=x(t),\qquad p\leq t\leq q\,.$$

Здесь $\lambda$ – параметр, $K(t,\tau)$ – заданная непрерывная при $p\leq t\leq q$, $p\leq\tau\leq t$ функция, $b(t)$ – заданная непрерывная при

$p \le t \le q$ функция. Отображение $Ax \equiv \lambda \cdot \int\limits_p^t K(t,\tau)x(\tau)d\tau + b(t)$ можно рассматривать как действующее из $C[p,q]$ в $C[p,q]$. Пусть $N = \max\limits_{\substack{p \le \tau \le t \\ p \le t \le q}} |K(t,\tau)|$. Для отображения $A$ имеем неравенство

$$|Ax(t) - Ay(t)| \le |\lambda| \cdot N \cdot \max_{p \le \tau \le q} |x(\tau) - y(\tau)| \cdot (t-p) =$$

$$= |\lambda| \cdot N \cdot (t-p) \cdot \rho(x,y) \le |\lambda| \cdot N \cdot (q-p) \cdot \rho(x,y).$$

Отсюда следует непрерывность $A$. Заметим, что

$$A^2 x(t) = \lambda \cdot \int\limits_p^t K(t,\tau_1)\left(\lambda \cdot \int\limits_p^{\tau_1} K(\tau_1,\tau_2)x(\tau_2)d\tau_2\right)d\tau_1 +$$

$$+ \lambda \cdot \int\limits_p^t K(t,\tau)b(\tau)d\tau.$$

Для отображения $A^2$ имеем неравенство

$$|A^2 x(t) - A^2 y(t)| \le$$

$$\le |\lambda|^2 \cdot \left| \int\limits_p^t \int\limits_p^{\tau_1} K(t,\tau_1)K(\tau_1,\tau_2)d\tau_2 d\tau_1 \right| \cdot \max_{p \le \tau_2 \le q} |x(\tau_2) - y(\tau_2)| \le$$

$$\le |\lambda|^2 \cdot N^2 \cdot \frac{(t-p)^2}{2} \cdot \rho(x,y).$$

Поступая так далее, найдем при $m \ge 1$ выражения для $A^m x(t)$ и по индукции получим для $A^m$ неравенства

$$|A^m x(t) - A^m y(t)| \le |\lambda|^m \cdot N^m \cdot \frac{(t-p)^m}{m!} \cdot \rho(x,y).$$

Тогда имеем оценку

$$\rho(A^m x, A^m y) = \max_{p \le t \le q} |A^m x(t) - A^m y(t)| \le$$

$$\le |\lambda|^m \cdot N^m \cdot \frac{(q-p)^m}{m!} \cdot \rho(x,y).$$

Отсюда видно, что при любом $\lambda$ число $m$ можно выбрать столь большим, что будет выполнено неравенство

$$\frac{|\lambda|^m \cdot N^m (q-p)^m}{m!} < 1 .$$

При таких $m$ отображение $A^m$ сжимающее. По теореме 4 отображение $A$ (при **любом** $\lambda$) имеет единственную неподвижную точку.

9. **Примеры**. Задача Коши для обыкновенного дифференциального уравнения 1-го порядка:

$$\frac{dx}{dt} = f(x,t), \quad x(0) = 0 .$$

Здесь заданная функция $f$ непрерывна по совокупности аргументов и удовлетворяет условию Липшица

$$|f(t,x) - f(t,y)| \le \sigma \cdot |x-y|, \quad \sigma > 0, \tag{5}$$

при $-\infty < x < +\infty$, $0 \le t \le T$. Задача Коши эквивалентна уравнению $\int\limits_0^t f(\tau, x(\tau))d\tau = x(t)$ при $0 \le t \le T$.

Отображение $Ax(t) \equiv \int\limits_0^t f(\tau, x(\tau))d\tau$ при каждом $0 < t \le T$ можно рассматривать как действующее из $C[0,t]$ в $C[0,t]$. Из (5) вытекает, что $\|Ax - Ay\|_{C[0,t]} \le \sigma \cdot t \cdot \|x-y\|_{C[0,t]}$. Поэтому при $t < \dfrac{1}{\sigma}$ отображение $A$ является сжимающим в пространстве $C[0,t]$. Значит, задача Коши имеет единственное решение, **определенное в промежутке** $0 \le t < \min\left\{T, \dfrac{1}{\sigma}\right\}$.

Докажем, что задача Коши имеет единственное решение, определённое **на всём промежутке** $0 \le t \le T$. В банаховом пространстве $C[0,T]$ наряду с нормой $\|x\|_{C[0,T]} = \max\limits_{0 \le t \le T}|x(t)|$

введём норму $\|x\|_* = \max\limits_{0\le t\le T}\left(e^{-\sigma t}\cdot|x(t)|\right)$. (Очевидно, что $e^{-\sigma T}\cdot\|x\|_{C[0,T]} \le \|x\|_* \le \|x\|_{C[0,T]}$, т.е. две нормы эквивалентны.) Из (5) получаем

$$\max_{0\le t\le T}\left|e^{-\sigma t}\cdot\int_0^t \left(f(\tau,x(\tau)) - f(\tau,y(\tau))\right)\right| \le$$

$$\le \sigma\cdot\max_{0\le t\le T}\int_0^t e^{\sigma\cdot(\tau-t)}\cdot e^{-\sigma\tau}\cdot|x(\tau)-y(\tau)|d\tau \le$$

$$\le \sigma\cdot\|x-y\|_*\cdot\max_{0\le t\le T}\int_0^t e^{\sigma\cdot(\tau-t)}d\tau = \left(1-e^{-\sigma T}\right)\cdot\|x-y\|_* .$$

Это значит, что отображение $A$ является сжимающим в норме $\|\cdot\|_*$ с коэффициентом сжатия $\alpha = 1-e^{-\sigma T} < 1$. Поэтому задача Коши имеет единственное решение, определённое при всех $0\le t\le T$.

Принцип сжимающих отображений позволяет доказать существование и единственность решения задачи Коши

$$\frac{dx}{dt} = f(t,x), \quad x(t_0) = x_0$$

и в случае, когда функция $f$, не равная тождественно нулю, задана и непрерывна в **прямоугольнике** $D = \left\{|t-t_0|\le a, |x-x_0|\le b\right\}$ и удовлетворяет в нем условию Липшица по переменной $x$ с постоянной $\sigma$. Доказательство основано на том, что для

$$0 < \theta \le \min\left\{a, \frac{b}{\max\limits_D|f(t,x)|}\right\}$$

отображение

$$Ax(t) \equiv x_0 + \int_{t_0}^t f(\tau,x(\tau))d\tau ,$$

примененное к непрерывной при $t_0\le t\le t_0+\theta$ функции $x(t)$, график которой не выходит из прямоугольника $D$, даёт непрерывную функцию, график которой также не выходит из этого прямоугольника. Отображение $A$ будет сжимающим, если выбрать $\theta < \dfrac{1}{\sigma}$. Аналогично при $t_0 - \theta \le t \le t_0$. После этого рассматривают вопрос о распространении решения на больший промежуток изменения $t$.

**Задача.** Пусть в задаче Коши

$\dfrac{dx}{dt} = f(t,x)$, $x(0)=0$, $0 \le t \le 1$, $-\infty < x < +\infty$, функция $f$ определена следующим образом:

$$f(t,x) = \begin{cases} 0, & t=0, \quad -\infty < x < +\infty, \\ 2t, & 0 < t \le 1, \; -\infty < x < 0, \\ 2t - \dfrac{4x}{t}, & 0 < t \le 1, \quad 0 \le x \le t^2, \\ -2t, & 0 < t \le 1, \quad t^2 < x < +\infty. \end{cases}$$

Является ли функция $f$ непрерывной? Ограниченной? Удовлетворяет ли $f$ условию Липшица по переменной $x$? Постройте последовательные приближения $x_{n+1}(t) = \int\limits_0^t f(\tau, x_n(\tau))d\tau$ на отрезке $0 \le t \le 1$ (начальное приближение $x_0(t) \equiv 0$); сходится ли эта последовательность? Сходятся ли к решению задачи Коши подпоследовательности $\{x_{2k}(t)\}$, $\{x_{2k-1}(t)\}$? □

**Задача.** $\dfrac{dx}{dt} = x^{\frac{1}{3}}$, $x(0)=0$, $t \ge 0$. Постройте последовательные приближения; сходятся ли они к решению задачи Коши? Является ли решением задачи Коши $x(t) = \left(\dfrac{2}{3}t\right)^{\frac{3}{2}}$? □

## Вопрос 8.

## Гильбертовы пространства. Теорема об ортогональной проекции.

1. В **действительном** линейном пространстве введем функцию двух переменных: каждой паре элементов $x$, $y$ поставим в соответствие действительное число $(x,y)$. Такая функция называется скалярным произведением, если выполнены следующие аксиомы:

1. $(x,y)=(y,x)$ для всех $x,y$;
2. $(x+y,z)=(x,z)+(y,z)$ для всех $x,y,z$;
3. $(\alpha x,y)=\alpha\cdot(x,y)$ для всех $x,y$ и любого действительного числа $\alpha$;
4. $(x,x)>0$, если $x\neq 0$; $(0,0)=0$.

В **комплексном** линейном пространстве тоже можно ввести скалярное произведение. Комплекснозначная функция $(x,y)$ двух переменных $x,y$ называется скалярным произведением, если выполнены аксиомы:

1. $(x,y)=\overline{(y,x)}$ для всех $x,y$ (черта над $(y,x)$ означает комплексное сопряжение);

2, 3, 4 – те же, что в случае действительного пространства.

Различие аксиом скалярного произведения в действительном и в комплексном пространствах приводит к следующим правилам:

в действительном пространстве

$(x,\alpha y)=\alpha\cdot(x,y)$, $\alpha$ – действительное число;

в комплексном пространстве

$(x,\alpha y)=\overline{\alpha}\cdot(x,y)$, $\alpha$ – комплексное число.

Аксиома 1 скалярного произведения в комплексном пространстве формулируется в виде $(x,y)=\overline{(y,x)}$ для того, чтобы могла выполняться аксиома 4.

Линейное пространство со скалярным произведением называется евклидовым.

Для любых двух элементов $x, y$ евклидова пространства справедливо неравенство Коши-Буняковского:

$$|(x, y)| \le \sqrt{(x, x)} \cdot \sqrt{(y, y)}.$$

Величина $\sqrt{(x, x)}$ называется евклидовой нормой $\|x\|$ элемента $x$. Евклидовым расстоянием между точками $x$ и $y$ называется $\|x - y\|$.

Норма в линейном пространстве является евклидовой в том и только в том случае, если для любых $x, y$ выполнено равенство параллелограмма:

$$\|x + y\|^2 + \|x - y\|^2 = 2\left(\|x\|^2 + \|y\|^2\right).$$

Норму называют строго выпуклой, если строго выпукла единичная сфера $\{x \mid \|x\| = 1\}$ (т.е. такая сфера не содержит отрезков: из $\|x\| = \|y\| = 1$, $x \ne y$, следует, что $\|\alpha x + (1 - \alpha)y\| < 1$ при $0 < \alpha < 1$). Строгая выпуклость нормы равносильна тому, что равенство $\|x + y\| = \|x\| + \|y\|$ для $x \ne 0$ и $y \ne 0$ возможно только если $x = \beta y$, где $\beta > 0$. Линейное пространство с такой нормой называется строго нормированным. Евклидова норма строго выпукла (докажите!).

**Утверждение**. Скалярное произведение является непрерывной функцией относительно сходимости по евклидовой норме.

**Доказательство**. Пусть при $n \to \infty$ $\|x_n - x\| \to 0$, $\|y_n - y\| \to 0$. Тогда числовые последовательности $\{\|x_n\|\}$ и $\{\|y_n\|\}$ ограничены. Обозначим через $K$ некоторую верхнюю грань последовательности $\{\|x_n\|\}$, например.

$$|(x_n, y_n) - (x, y)| \le |(x_n, y_n) - (x_n, y)| + |(x_n, y) - (x, y)| =$$

$$= |(x_n, y_n - y)| + |(x_n - x, y)| \le \|x_n\| \cdot \|y_n - y\| + \|x_n - x\| \cdot \|y\| \le$$

$$\le K \cdot \|y_n - y\| + \|y\| \cdot \|x_n - x\| \to 0 \text{ при } n \to \infty. \ \Box$$

Евклидово пространство является метрическим пространством с метрикой $\rho(x, y) = \|x - y\|$. Поэтому естественно поставить вопрос о полноте такого пространства. Если для всякой фундаментальной последовательности элементов метрического пространства найдется принадлежащий ему элемент, являющийся пределом этой последовательности, то пространство называют полным. Если метрическое пространство не полно, то его некоторым способом можно включить в полное пространство, называемое пополнением исходного; это пополнение единственно с точностью до изометрии, оставляющей неподвижными точки исходного пространства.

**Определение**. **Гильбертовым пространством** $H$ называется полное евклидово бесконечномерное пространство. ▯

Гильбертовы пространства составляют наиболее распространенный и важный для приложений класс бесконечномерных пространств. Они представляют собой естественное обобщение понятия конечномерного линейного пространства со скалярным произведением: в конечномерном пространстве свойство полноты выполнено автоматически. Гильбертово пространство является естественным обобщением трехмерного пространства евклидовой геометрии; многие геометрические понятия переносятся в гильбертово пространство.

Если евклидово пространство $H$ бесконечномерно, но не полно, то его часто называют предгильбертовым. Пополнение предгильбертова пространства будет гильбертовым.

Подмножество метрического пространства называется всюду плотным, если его замыкание совпадает со всем пространством. Метрическое пространство, в котором имеется счетное всюду плотное множество, называют **сепарабельным**. Понятие сепарабельности играет важную роль в теории гильбертовых пространств.

2. **Пример**. Через $l_2$ обозначают пространство всех действительных бесконечных числовых последовательностей $x = (x_1, x_2, \dots)$, удовлетворяющих условию $\sum_{i=1}^{\infty} x_i^2 < \infty$.

Операции сложения элементов и умножения их на действительные числа определены покомпонентно: $x+y=(x_1+y_1, x_2+y_2, \dots)$. $\alpha \cdot x=(\alpha \cdot x_1, \alpha \cdot x_2, \dots)$. Из сходимости числовых рядов $\sum_{i=1}^{\infty} x_i^2$ и $\sum_{i=1}^{\infty} y_i^2$ следует сходимость ряда $\sum_{i=1}^{\infty} x_i y_i$, поэтому в $l_2$ можно ввести скалярное произведение $(x,y)=\sum_{i=1}^{\infty} x_i y_i$ и евклидову норму $\|x\|=\left(\sum_{i=1}^{\infty} x_i^2\right)^{\frac{1}{2}}$. $l_2$ – полное сепарабельное евклидово пространство (гильбертово).

Гильбертовым пространством $l_2$ будет, разумеется, и множество всех **комплексных** числовых последовательностей, удовлетворяющих условию $\sum_{i=1}^{\infty}|x_i|^2<\infty$. В этом случае под скалярным произведением надо понимать $(x,y)=\sum_{i=1}^{\infty} x_i \overline{y}_i$.

**Пример**. В пространстве $\widetilde{L}_2[a,b]$ непрерывных на $[a,b]$ действительнозначных функций выражение $(x,y)=\int_a^b x(t)y(t)dt$ задаёт скалярное произведение. Такое евклидово пространство не является полным (приведите пример последовательности непрерывных функций, сходящейся в среднем квадратическом к разрывной функции).

Расширим (пополним) это евклидово пространство. Будем рассматривать все функции, определенные почти всюду на $[a,b]$, измеримые и такие, что существует **интеграл Лебега** $\int_a^b x^2(t)dt$. (Эквивалентные функции, т.е. различающиеся лишь на множестве меры, равной нулю, считаются совпадающими.) Такое множество функций является евклидовым пространством со скалярным произведением $(x,y)=\int_a^b x(t)y(t)dt$, которое теперь уже является **полным**. Его обозначают через $L_2[a,b]$. Это сепарабельное гильбертово пространство, которое изометрично действительному пространству $l_2$.

Более общим примером является пространство $L_2(T,\Sigma,\mu)$ действительных функций $x=x(t)$, определенных на множестве $T$ с вполне аддитивной положительной мерой $\mu$ ($\mu$ задана на $\sigma$-алгебре $\Sigma$ подмножеств множества $T$), измеримых и имеющих интегрируемый квадрат: $\int_T x^2(t)d\mu(t)<\infty$. В этом гильбертовом пространстве скалярное произведение определяется выражением $(x,y)=\int_T x(t)y(t)d\mu(t)$. Пространство сепарабельно, если мера $\mu$ имеет счетный базис.

**Пример**. Важными для математической физики примерами гильбертовых пространств являются пространства Соболева $W_2^l(T)$. Элементы такого пространства – это определенные на открытом множестве $T\subset R^k$ функции $x=x(t)=x(t_1,\ldots,t_k)$, интегрируемые с квадратом их модуля вместе со своими обобщенными производными до порядка $l$ включительно. Если за норму принять выражение $\left(\int_T \sum_{|i|\le l}\left|x^{(i)}(t)\right|^2 dt\right)^{\frac{1}{2}}$, то получим гильбертово пространство. Полным его делает то обстоятельство, что в его определении участвуют не обычные, а обобщенные производные.

**Пример**. Множество всех действительнозначных функций $x(t)$, определенных и измеримых на всей оси $-\infty<t<+\infty$, для которых существует $\lim_{T\to\infty}\frac{1}{2T}\int_{-T}^{T}x^2(t)dt<\infty$, образует действительное линейное пространство. Если ввести в нем скалярное произведение $(x,y)=\lim\limits_{T\to\infty}\dfrac{1}{2T}\int\limits_{-T}^{T}x(t)y(t)dt$, то получим гильбертово пространство. Это пространство не сепарабельно.

**Пример**. Пусть $\mathrm{H}$ – пространство всех действительнозначных функций $x=x(t)$, определенных на числовой прямой $-\infty<t<+\infty$ и имеющих не более счетного числа отличных от нуля значений: для каждой функции $x$ множество точек $t_1,t_2,\ldots$, в которых $x$ отлична от нуля, не более чем счетно. Кроме того, пусть взятая по всем таким точкам сумма $\sum\limits_{t}x^2(t)$ конечна (её можно рассматривать как обычный числовой ряд). Операции сложения и умножения на действительные числа определим в этом пространстве как сложение функций и умножение их на числа. Введем в $\mathrm{H}$ скалярное произведение $(x,y)=\sum x(t)y(t)$, где сумма берётся по множеству тех $t$, в которых $x(t)y(t)\neq 0$. Из полноты действительного пространства $l_2$ следует полнота пространства $\mathrm{H}$; $\mathrm{H}$ – гильбертово пространство. Обозначим через $x_\tau$ элемент

$$x_\tau(t)=\begin{cases}1 & при \quad t=\tau,\\ 0 & при \quad t\neq\tau.\end{cases}$$

Очевидно, что если $\tau_1\neq\tau_2$, то $\left\|x_{\tau_1}-x_{\tau_2}\right\|=\sqrt{2}$. Поскольку имеем континуум таких $x_\tau$, пространство $\mathrm{H}$ не сепарабельно.

Рассмотренное пространство $\mathrm{H}$ естественно обозначить через $l_2(-\infty,+\infty)$. Вместо числовой прямой $-\infty<t<+\infty$ можно взять произвольное бесконечное множество $T$ и описанным только что способом построить гильбертово пространство $l_2(T)$. Если $T$ имеет мощность, большую счетного множества, то получим несепарабельное пространство. Оказывается, что всякое гильбертово пространство изоморфно пространству $l_2(T)$ для некоторого соответствующим образом подобранного $T$. (Изоморфизм евклидовых пространств – это взаимно однозначное соответствие, сохраняющее определенные в этих пространствах линейные операции и скалярное произведение.)

Любые два **сепарабельных** гильбертовых пространства (над одним и тем же полем чисел) изоморфны между собой. В частности, любое такое пространство изоморфно $l_2$. Это значит, что с точностью до изоморфизма существует только одно сепарабельное гильбертово пространство, и $l_2$ можно рассматривать как его координатную реализацию. Поэтому принято говорить об абстрактном гильбертовом сепарабельном пространстве $\mathsf{H}$ или об абстрактном несепарабельном пространстве $\mathsf{H}$.

3. Элементы $x$, $y$ гильбертова пространства $\mathsf{H}$ называются **ортогональными**, если $(x, y) = 0$ (это записывают $x \perp y$).

Линейным подпространством $\mathsf{L}$ в гильбертовом пространстве $\mathsf{H}$ называется **замкнутое** линейное многообразие в $\mathsf{H}$. Если элемент $x \in \mathsf{H}$ ортогонален любому элементу $y \in \mathsf{L} \subset \mathsf{H}$, то говорят, что $x$ ортогонален подпространству $\mathsf{L}$ (это записывают $x \perp \mathsf{L}$).

**Теорема**. Пусть $\mathsf{H}$ – комплексное гильбертово пространство и $\mathsf{L}$ – линейное подпространство в $\mathsf{H}$. Для любого $x \in \mathsf{H}$ имеет место разложение $x = y + z$, где $y \in \mathsf{L}$, $z \perp \mathsf{L}$. Это разложение единственно.

**Доказательство**. Если $x \in \mathsf{L}$, то $y = x$, $z = 0$. Поэтому надо рассмотреть случай $x \notin \mathsf{L}$. Обозначим через $\sqrt{d}$ расстояние от точки $x$ до подпространства $\mathsf{L}$: $d = \inf\limits_{y \in \mathsf{L}} \|x - y\|^2$. Пусть $\{y_n\} \in \mathsf{L}$ – такая последовательность, что $d_n = \|x - y_n\|^2 \underset{n \to \infty}{\longrightarrow} d$. Пусть $h \in \mathsf{L}$ – произвольный элемент, $h \neq 0$. Тогда для любого комплексного числа $\alpha$ элемент $y_n + \alpha h \in \mathsf{L}$ и поэтому $\|x - (y_n + \alpha h)\|^2 \geq d$.

$$\|x-(y_n+\alpha h)\|^2 = ((x-y_n)-\alpha h,(x-y_n)-\alpha h) =$$
$$= \|x-y_n\|^2 - \overline{\alpha}\cdot(x-y_n,h) - \alpha\cdot(h,x-y_n) + |\alpha|^2\cdot\|h\|^2 \geq d.$$

Выберем $\alpha = \dfrac{(x-y_n,h)}{\|h\|^2}$. Тогда $\|x-y_n\|^2 - \dfrac{|(x-y_n,h)|^2}{\|h\|^2} \geq d$,

$|(x-y_n,h)|^2 \leq \|h\|^2\cdot(d_n-d)$, $|(x-y_n,h)| \leq \|h\|\cdot\sqrt{d_n-d}$. (1)

Мы выбирали $h \neq 0$, но неравенство (1) выполнено и при $h=0$. Таким образом, для любого $h \in \mathrm{L}$ имеем

$$|(y_n-y_m,h)| \leq |(y_n-x,h)| + |(y_m-x,h)| \leq$$
$$\leq \left(\sqrt{d_n-d}+\sqrt{d_m-d}\right)\cdot\|h\|.$$

В частности, для $h = y_n - y_m$ получим

$\|y_n-y_m\| \leq \sqrt{d_n-d}+\sqrt{d_m-d} \to 0$ при $n\to\infty$ и $m\to\infty$.

Это значит, что последовательность $\{y_n\}$ фундаментальна. В силу полноты пространства $\mathrm{H}$ существует предел этой последовательности $y \in \mathrm{H}$. Так как все $y_n \in \mathrm{L}$ – **замкнутому** подпространству, то и $y \in \mathrm{L}$.

Используя непрерывность скалярного произведения, перейдем к пределу при $n\to\infty$ в неравенстве (1); получим $(x-y,h)=0$. Поскольку $h$ – произвольный элемент подпространства $\mathrm{L}$, то мы доказали, что $x-y \perp \mathrm{L}$. Полагая $x-y=z$, получаем $x=y+z$, $y\in\mathrm{L}$, $z\perp\mathrm{L}$.

Осталось доказать единственность разложения элемента $x$. Пусть $x=y+z=\widetilde{y}+\widetilde{z}$, $y\in\mathrm{L}$, $\widetilde{y}\in\mathrm{L}$, $z\perp\mathrm{L}$, $\widetilde{z}\perp\mathrm{L}$. Тогда $y-\widetilde{y}=\widetilde{z}-z$ и $\|y-\widetilde{y}\|^2=(y-\widetilde{y},\widetilde{z}-z)=0$, поскольку $y-\widetilde{y}\in\mathrm{L}$, $\widetilde{z}-z\perp\mathrm{L}$. Но тогда $y=\widetilde{y}$ и $z=\widetilde{z}$. □

Такая же теорема справедлива и в действительном гильбертовом пространстве.

Элемент $y$ в разложении $x=y+z$, $y\in\mathrm{L}$, $z\perp\mathrm{L}$, называется **ортогональной проекцией** элемента $x$ на подпространство $\mathsf{L}$. Совокупность $\mathsf{L}^{\perp}$ всех элементов пространства $\mathsf{H}$, ортогональных подпространству $\mathsf{L}$, также является подпространством : все указанные элементы образуют линейное многообразие, а замкнутость $\mathsf{L}^{\perp}$ следует из непрерывности скалярного произведения. $\mathsf{L}^{\perp}$ называют **ортогональным дополнением** к $\mathsf{L}$ и говорят, что $\mathsf{H} = \mathsf{L} \oplus \mathsf{L}^{\perp}$ – ортогональная сумма подпространств $\mathsf{L}$ и $\mathsf{L}^{\perp}$. Это значит, что любой элемент $x \in \mathsf{H}$ однозначно представим в виде $x = y + z$, $y \in \mathsf{L}$, $z \in \mathsf{L}^{\perp}$, $y \perp z$ (очевидно, что $\|x\|^2 = \|y\|^2 + \|z\|^2$). $\mathsf{L}$ и $\mathsf{L}^{\perp}$ – взаимно дополнительные (до $\mathsf{H}$) ортогональные друг другу подпространства.

Часто приходится находить ортогональную проекцию элемента $x \in \mathsf{H}$ на **конечномерное** подпространство $\mathsf{L}$. Пусть $\dim \mathsf{L} = k < \infty$, и $e_1, \dots, e_k$ – базис этого конечномерного пространства (т.е. $\mathsf{L}$ является линейной оболочкой элементов $e_1, \dots, e_k$). Тогда ортогональная проекция $x$ на $\mathsf{L}$ имеет вид $y = \alpha_1 e_1 + \dots + \alpha_k e_k$. Элемент $x - y$ должен быть ортогонален каждому из элементов $e_1, \dots, e_k$ :

$$(x, e_i) - \alpha_1 (e_1, e_i) - \dots - \alpha_k (e_k, e_i) = 0 \qquad (2)$$

при всех $i = 1, \dots, k$. (2) является системой $k$ линейных алгебраических уравнений относительно $\alpha_1, \dots, \alpha_k$. Основная матрица этой системы имеет элементы $a_{ij} = (e_j, e_i)$ – матрица Грама, а её определитель $G(e_1, \dots, e_k)$ – определитель Грама элементов $e_1, \dots, e_k$. Для линейной независимости $e_1, \dots, e_k$ необходимо и достаточно, чтобы определитель Грама этих элементов был отличен от нуля. Поэтому система (2) имеет единственное решение. Нетрудно найти и расстояние $\sqrt{d} = \min_{y \in \mathsf{L}} \|x - y\|$. Оно равно $\|z\| = \|x - y\|$ при условии, что коэффициенты разложения элемента $y$ по базису $e_1, \dots, e_k$ определяются системой (2). Следовательно,

$$d=(z,z)=(z,x)=(x-y,x)=$$
$$=(x,x)-\alpha_1(e_1,x)-\ldots-\alpha_k(e_k,x)\,. \qquad (3)$$

Исключая $\alpha_1,\ldots,\alpha_k$ из системы уравнений (2), (3), получим
$$d=\frac{G(x,e_1,\ldots,e_k)}{G(e_1,\ldots,e_k)}.$$

4. Множество $M\subset \mathrm{H}$ называется **ортонормированной системой** элементов, если любые два её различные элемента ортогональны, а норма каждого её элемента равна единице. Ортонормированная система называется **полной**, если в $\mathrm{H}$ не существует ненулевого элемента, ортогонального ко всем элементам этой системы. Полные ортонормированные системы в гильбертовом пространстве $\mathrm{H}$ содержат бесконечное число элементов, поэтому возникает вопрос об их мощности. На этот вопрос имеется простой ответ в случае сепарабельного $\mathrm{H}$: пространство $\mathrm{H}$ имеет полную ортонормированную **последовательность** (счетного числа элементов) в том и только в том случае, когда оно сепарабельно. В сепарабельном пространстве $\mathrm{H}$ всякая полная ортонормированная система является замкнутой (т.е. для любого $x\in \mathrm{H}$ справедливо равенство Парсеваля) и обратно. Это означает, что в сепарабельном гильбертовом пространстве всякий элемент $x$ можно разложить в ряд Фурье по некоторому ортонормированному базису (из счетного числа элементов), а само пространство $\mathrm{H}$ представимо ортогональной суммой $\mathrm{H}=\mathrm{L}_1\oplus \mathrm{L}_2\oplus\ldots$ **счетного** числа попарно ортогональных подпространств: каждый элемент $x\in \mathrm{H}$ имеет вид $x=y_1+y_2+\ldots$, где все $y_i\in \mathrm{L}_i$, причем при заданных $\mathrm{L}_1,\mathrm{L}_2,\ldots$ такое представление единственно, и $\|x\|^2=\sum\limits_{i=1}^{\infty}\|y_i\|^2<\infty$.

## Вопрос 9.

## Теорема Рисса о представлении линейного функционала.

1. Пусть $X$ – линейное пространство над полем чисел $P$. Будем рассматривать только случаи $P = R$ и $P = C$, т.е. действительное или комплексное пространство $X$. Отображение $f$, действующее из $X$ в $P$ называют функционалом; его область определения – некоторое множество $Dom\,f$ элементов пространства $X$. Пусть $Dom\,f$ является линейным многообразием в $X$, тогда функционал $f$ называется **однородным**, если $f(\alpha x) = \alpha \cdot f(x)$ для всех $x \in Dom\,f$ и всех $\alpha \in P$; функционал $f$ называется **аддитивным**, если $f(x + y) = f(x) + f(y)$ для всех $x, y \in Dom\,f$. Аддитивный и однородный функционал называется **линейным**.

Если $X$ – нормированное пространство, то функционал $f$ называют **непрерывным в точке** $x_0 \in Dom\,f$, если для любого числа $\varepsilon > 0$ найдется такое число $\delta(\varepsilon) > 0$, что для всех $x \in Dom\,f$, удовлетворяющих неравенству $\|x - x_0\| < \delta$, выполнено неравенство $|f(x) - f(x_0)| < \varepsilon$. Аддитивный функционал, непрерывный в некоторой точке $x_0 \in Dom\,f$, будет непрерывен в каждой точке множества $Dom\,f$ (докажите!).

Линейный функционал $f$ называется **ограниченным**, если существует такое число $C$, что для всех $x \in Dom\,f$ выполнено неравенство $|f(x)| \le C \cdot \|x\|$. Линейный функционал $f$ непрерывен на $Dom\,f$ в том и только в том случае, если он ограничен на $Dom\,f$ (докажите!). Для **ограниченного** линейного функционала $\sup\limits_{\substack{x \in Dom\,f \\ \|x\| \le 1}} |f(x)| < \infty$; это число является наименьшим из чисел $C$, при которых выполнено неравенство $|f(x)| \le C \cdot \|x\|$, и называется нормой линейного функционала $f$:

$$\|f\| = \sup_{\substack{x \in Dom\, f \\ \|x\| \le 1}} |f(x)| = \sup_{\substack{x \in Dom\, f \\ \|x\| = 1}} |f(x)| = \sup_{\substack{x \in Dom\, f \\ x \ne 0}} \frac{|f(x)|}{\|x\|} = \inf_{\{C \,|\, |f(x)| \le C \|x\|\}} C .$$

Если $Dom\, f \ne \mathsf{X}$, то линейный ограниченный функционал $f$ можно продолжить на всё пространство $\mathsf{X}$ с сохранением его нормы: по теореме Хана-Банаха найдется такой линейный ограниченный функционал $g$, что $Dom\, g = \mathsf{X}$, $\|g\| = \|f\|$ и $g(x) = f(x)$ для любого $x \in Dom\, f$. Поэтому далее будем считать, что $Dom\, f = \mathsf{X}$.

Множество $\{x \in \mathsf{X} \,|\, f(x) = c\}$, где $c = const$, называется гиперплоскостью в $\mathsf{X}$. Гиперплоскость $\{x \in \mathsf{X} \,|\, f(x) = \|f\|\}$ обладает тем свойством, что для всех точек единичного шара $\{x \in \mathsf{X} \,|\, \|x\| \le 1\}$ выполнено $f(x) \le \|f\|$. Ни при каком $\varepsilon > 0$ параллельная ей гиперплоскость $\{x \in \mathsf{X} \,|\, f(x) = \|f\| - \varepsilon\}$ этим свойством уже не обладает. Поэтому гиперплоскость $\{x \in \mathsf{X} \,|\, f(x) = \|f\|\}$ можно назвать опорной к шару $\{x \in \mathsf{X} \,|\, \|x\| \le 1\}$. Эту геометрическую интерпретацию нормы линейного функционала можно выразить несколько иначе: обозначим через $d$ расстояние от точки $0 \in \mathsf{X}$ до гиперплоскости $\{x \in \mathsf{X} \,|\, f(x) = 1\}$, т.е. $d = \inf_{\{x \,|\, f(x) = 1\}} \|x\|$; из определения $\|f\|$ следует, что $\|f\| = \dfrac{1}{d}$ (докажите!).

Если $\mathsf{X}$ – **конечномерное** пространство, то всякий линейный функционал на $\mathsf{X}$ автоматически непрерывен. В случае бесконечномерного пространства $X$ из линейности функционала его непрерывность, вообще говоря, не вытекает.

Линейный функционал является частным случаем линейного оператора. Если линейный оператор действует из конечномерного пространства $X$ в конечномерное пространство $Z$, и в этих пространствах фиксированы их базисы, то оператор задаётся матрицей чисел. Поэтому линейный функционал на конечномерном пространстве $X$ можно задать набором $n$ чисел ($n = \dim X$): если $e = (e_1, \dots, e_n)$ – базис в $X$, $x = x_1 e_1 + \dots + x_n e_n$, то $f(x) = x_1 f(e_1) + \dots + x_n f(e_n)$, и функционал $f$ полностью определен числами $y_1 = f(e_1), \dots, y_n = f(e_n)$. Линейное пространство всех определённых на $X$ линейных непрерывных функционалов называется пространством $X^*$, сопряженным к $X$; если $\dim X < \infty$, то $\dim X^* = \dim X$, и мы имеем общий вид линейного функционала на $X$. В бесконечномерных пространствах построение общего вида линейных непрерывных функционалов является значительно более сложной задачей.

Различные нормы в пространстве $X$ индуцируют различные нормы в $X^*$.

**Задача**. $X = \mathbf{C}^n$, $x = (x_1, \dots, x_n)$, $f(x) = x_1 \cdot y_1 + \dots + x_n \cdot y_n$, $\|x\| = \left( \sum_{i=1}^{n} |x_i|^2 \right)^{1/2}$. Докажите, что $\|f\| = \left( \sum_{i=1}^{n} |y_i|^2 \right)^{1/2}$. □

**Задача**. То же, но $\|x\| = \left( \sum_{i=1}^{n} |x_i|^p \right)^{1/p}$, $1 < p < \infty$. Докажите, что $\|f\| = \left( \sum_{i=1}^{n} |y_i|^q \right)^{1/q}$, где $\frac{1}{p} + \frac{1}{q} = 1$. □

**Задача**. То же, но $\|x\| = \max\limits_{1 \le i \le n} |x_i|$. Докажите, что $\|f\| = \sum\limits_{i=1}^{n} |y_i|$. □

**Задача**. То же, но $\|x\| = \sum\limits_{i=1}^{n} |x_i|$. Докажите, что $\|f\| = \max\limits_{1 \le i \le n} |y_i|$. □

2. Общий вид линейного непрерывного функционала, определенного на бесконечномерном пространстве, проще всего установить, если это пространство **гильбертово**.

**Теорема Ф. Рисса**. Пусть $H$ – действительное или комплексное гильбертово пространство, $(x, y)$ – скалярное произведение его элементов. Пусть $f$ – определённый всюду на $H$ линейный непрерывный функционал. Тогда существует единственный элемент $y \in H$ такой, что при всех $x \in H$ выполнено $f(x) = (x, y)$ и, кроме того, $\|f\| = \|y\|_H$.

**Доказательство**. Обозначим через $Ker\, f = \{x \in H \mid f(x) = 0\}$ ядро функционала $f$. $Ker\, f$ является подпространством в $H$: вследствие линейности $f$ его ядро есть линейное многообразие, а вследствие непрерывности $f$ оно замкнуто. Если $Ker\, f = H$, т.е. $f(x) \equiv 0$ на $H$, то положим $y = 0 \in H$; в этом случае утверждение теоремы очевидно. Пусть $Ker\, f \ne H$. Тогда существует элемент $z \ne 0$, $z \in (Ker\, f)^{\perp}$ – ортогональному дополнению к ядру функционала. При любом $x \in H$ элемент $f(x) \cdot z - f(z) \cdot x$ принадлежит $Ker\, f$, поскольку $f(f(x) \cdot z - f(z) \cdot x) = f(x) \cdot f(z) - f(z) \cdot f(x) = 0$.

Следовательно, $z$ и $f(x) \cdot z - f(z) \cdot x$ ортогональны: $(f(x) \cdot z - f(z) \cdot x, z) = 0$. Отсюда $f(x) \cdot (z, z) - f(z) \cdot (x, z) = 0$,

$f(x) = \left( x, \overline{f(z)} \cdot \dfrac{z}{\|z\|_H^2} \right)$ (в случае действительного пространства $H$ черту комплексного сопряжения над $f(z)$ в последнем равенстве можно не принимать во внимание). Положим $y = \overline{f(z)} \cdot \dfrac{z}{\|z\|_H^2}$, тогда для любого $x \in H$ значение функционала $f$ в точке $x$ равно $f(x) = (x, y)$.

Докажем единственность полученного представления. Предположим, что для любого $x \in H$ $f(x) = (x, y_1) = (x, y_2)$. Тогда, выбирая $x = y_1 - y_2$, получим $\|y_1 - y_2\|_H^2 = 0$, т.е. $y_1 = y_2$.

Докажем, что $\|f\| = \|y\|_H$. По неравенству Коши-Буняковского $|f(x)| = |(x, y)| \le \|x\|_H \cdot \|y\|_H$, поэтому $\|f\| = \sup\limits_{x \neq 0} \dfrac{|f(x)|}{\|x\|_H} \le \|y\|_H$. При $x = y$ имеем $\dfrac{|f(y)|}{\|y\|_H} = \dfrac{\|y\|_H^2}{\|y\|_H} = \|y\|_H$, поэтому $\|f\| = \|y\|_H$. □

Через $H^*$ обозначают линейное нормированное пространство всех определённых на гильбертовом пространстве $H$ линейных непрерывных функционалов. Теорема Ф.Рисса устанавливает взаимно однозначное соответствие между $H$ и $H^*$, сохраняющее норму: если $f_1 \leftrightarrow y_1$ и $f_2 \leftrightarrow y_2$, то для любых чисел $\alpha_1$ и $\alpha_2$ будет $\alpha_1 f_1 + \alpha_2 f_2 \leftrightarrow \overline{\alpha}_1 y_1 + \overline{\alpha}_2 y_2$. (В случае действительного $H$ указанное соответствие является линейным изоморфизмом, в случае комплексного $H$ – сопряженно линейным изоморфизмом.)

**Пример**. Элементами $x$ пространства $l_2$ являются всевозможные бесконечные числовые последовательности

$x=(x_1, x_2, \dots)$, удовлетворяющие условию $\sum_{i=1}^{\infty}|x_i|^2<\infty$.

Скалярное произведение $(x,y)=\sum_{i=1}^{\infty}x_i\overline{y}_i$. $l_2$ – гильбертово пространство. Всякий линейный непрерывный функционал в $l_2$ задаётся единственной числовой последовательностью $y\in l_2$ и имеет вид $f(x)=\sum_{i=1}^{\infty}x_i\overline{y}_i$.

Любое **сепарабельное** гильбертово пространство H изоморфно пространству $l_2$, причем изоморфизм можно установить с сохранением нормы и скалярного произведения. Поэтому каждый линейный непрерывный функционал на сепарабельном гильбертовом пространстве можно интерпретировать как функционал с той же нормой в $l_2$. ▯

**Задача**. Найдите норму функционала $f(x)=x_3+x_5+x_7$, $x\in l_2$.

**Пример**. Элементами $x$ пространства $L_2[a,b]$ являются функции $x(t)$, интегрируемые по Лебегу вместе с квадратом их модуля на отрезке $[a,b]$. Скалярное произведение $(x,y)=\int_a^b x(t)\overline{y(t)}dt$. $L_2[a,b]$ – гильбертово пространство. Всякий линейный непрерывный функционал в $L_2[a,b]$ задаётся единственной функцией $y\in L_2[a,b]$ (если считать совпадающими функции, различающиеся лишь на множестве, мера которого равна нулю) и имеет вид $f(x)=\int_a^b x(t)\overline{y(t)}dt$. ▯

**Задача**. Найдите норму функционала $f(x)=\int_0^1 t\cdot x(t)dt$, $x\in L_2[0,1]$. ▯

Пусть линейный непрерывный функционал $f$ определен на (замкнутом) подпространстве $Dom\, f$ гильбертова пространства $\mathsf{H}$. В $Dom\, f$ применима теорема Ф.Рисса: по функционалу $f$ однозначно определяется такой элемент $y \in Dom\, f$, что для любого $x \in Dom\, f$ $f(x) = (x, y)$ и $\|f\| = \|y\|_{\mathsf{H}}$. Расширим функционал $f$ на всё пространство $\mathsf{H}$, положив $f(x) = (x, y)$ для любого $x \in \mathsf{H}$. Это расширение не увеличивает норму функционала, а всякое другое расширение $f$ на всё $\mathsf{H}$ увеличит его норму. Действительно, если для линейного непрерывного функционала $g$ $Dom\, g = \mathsf{H}$ и $g(x) = f(x)$ при $x \in Dom\, f$, то при некотором $u \in \mathsf{H}$ $g(x) = (x, u)$, $\|g\| = \|u\|_{\mathsf{H}}$, и $(x, y) = (x, u)$ при $x \in Dom\, f$, т.е. $u - y \perp Dom\, f$. Поскольку $y \in Dom\, f$, имеем $\|u\|_{\mathsf{H}}^2 = \|y\|_{\mathsf{H}}^2 + \|u - y\|_{\mathsf{H}}^2$. Значит, $\|g\| \ge \|f\|$, причем в случае $u \ne y$ будет $\|g\| > \|f\|$.

3. Через $C[a,b]$ обозначают пространство непрерывных на отрезке $[a,b]$ функций $x(t)$ с нормой $\|x\| = \max\limits_{a \le t \le b} |x(t)|$.

**Пример.** $f(x) = \int\limits_a^b x(t)dt$ представляет собой линейный функционал в пространстве $C[a,b]$. Этот функционал ограничен, и $\|f\| = b - a$. □

**Пример.** Пусть дана определённая на $[a,b]$ функция $y(t)$, для которой существует конечный $\int\limits_a^b |y(t)|dt$. Тогда

$f(x)=\int\limits_a^b x(t)y(t)dt$ представляет собой линейный ограниченный на $C[a,b]$ функционал, и $\|f\|=\int\limits_a^b |y(t)|dt$. □

**Пример**. Пусть фиксирована точка $t_0\in[a,b]$. Тогда $f(x)=x(t_0)$ представляет собой линейный ограниченный на $C[a,b]$ функционал, и $\|f\|=1$. □

В отличие от случая гильбертова пространства, пространство, сопряженное к $C[a,b]$, не является изоморфным пространству $C[a,b]$. Чтобы указать общий вид линейных непрерывных функционалов, определённых на $C[a,b]$, напомним сначала понятие функции с ограниченным изменением. Функция $y(t)$, заданная на $[a,b]$, называется функцией с ограниченным изменением, если существует такая постоянная $K$, что для любого разбиения отрезка $[a,b]$ точками $a=t_0<t_1<\ldots<t_n=b$ выполнено неравенство $\sum\limits_{i=1}^n |y(t_i)-y(t_{i-1})|\le K$. Величина суммы в левой части этого неравенства зависит от способа разбиения отрезка $[a,b]$. Но если множество всех таких сумм ограничено, то оно имеет конечную точную верхнюю грань, которая называется **полным изменением** функции $y$ на отрезке $[a,b]$ и обозначается $V_a^b[y]$. Всякая монотонная функция имеет ограниченное изменение. Разность двух неубывающих функций есть функция с ограниченным изменением, и всякая функция с ограниченным изменением может быть представлена в виде разности двух неубывающих функций.

Какова бы ни была функция $y(t)$ с ограниченным изменением, для любой функции $x\in C[a,b]$ существует интеграл Стилтьеса $\int\limits_a^b x(t)dy(t)$. Если $x$ рассматривать как аргумент функционала

$$f(x)=\int_a^b x(t)dy(t), \qquad (1)$$

то (1) задаёт определённый на $C[a,b]$ **линейный непрерывный** функционал $f$. Для каждого линейного непрерывного функционала $f$, определённого на $C[a,b]$, можно подобрать такую функцию $y(t)$ с ограниченным изменением , что функционал $f$ будет представим в виде (1). Эти утверждения составляют содержание следующей теоремы:

**Теорема Ф. Рисса**. Общий вид линейного непрерывного функционала $f$ на пространстве $C[a,b]$ даётся интегралом Стилтьеса (1), где $y(t)$ – функция с ограниченным изменением. $\|f\| \le V_a^b[y]$. Функцию $y(t)$ можно подобрать так, что $\|f\| = V_a^b[y]$. □

**Замечание**. По заданному на $C[a,b]$ линейному непрерывному функционалу $f$ функция $y(t)$ с ограниченным изменением определяется не единственным образом. Действительно, величина интеграла (1) для любой непрерывной функции $x(t)$ не зависит от значений $y(t)$ на конечном или счётном множестве точек и не меняется при замене $y(t)$ на $y(t)+const$. Поэтому в общем случае можно лишь утверждать, что $\|f\| \le V_a^b[y]$. Но среди всех функций $y(t)$, задающих один и тот же функционал $f$, можно выбрать такую, что $\|f\| = V_a^b[y]$.

## Вопрос 10.

## Сопряжённый оператор в гильбертовом пространстве. Вполне непрерывные операторы.

1. Пусть $X$ и $Y$ – линейные пространства над одним и тем же полем чисел $P$. Отображение $A$, действующее из $X$ в $Y$, принято называть оператором; его область определения – некоторое множество $Dom\ A$ элементов пространства $X$. Пусть $Dom\ A$ является линейным многообразием в $X$, тогда оператор $A$ называется **однородным**, если $A(\alpha x) = \alpha Ax$ для всех $x \in Dom\ A$ и всех $\alpha \in P$; оператор $A$ называется **аддитивным**, если $A(x+y) = Ax + Ay$ для всех $x, y \in Dom\ A$. Аддитивный и однородный оператор называется **линейным**.

Если $X$ и $Y$ – нормированные пространства, то оператор $A$ называется **непрерывным в точке** $x_0 \in Dom\ A$, если для любого числа $\varepsilon > 0$ найдется такое число $\delta(\varepsilon) > 0$, что для всех $x \in Dom\ A$, удовлетворяющих неравенству $\|x - x_0\|_X < \delta$, выполнено неравенство $\|Ax - Ax_0\|_Y < \varepsilon$. Аддитивный оператор, непрерывный в некоторой точке $x_0 \in Dom\ A$, будет непрерывен в каждой точке множества $Dom\ A$ (докажите!).

Оператор $A$ называют **линейным непрерывным**, если $Dom\ A = X$, $A$ линеен и непрерывен в каждой точке $x \in X$. Линейный оператор $A$ называют **ограниченным**, если $Dom\ A = X$ и существует такое число $C$, что для всех $x \in X$ выполнено неравенство $\|Ax\|_Y \le C \cdot \|x\|_X$. Образ всякого ограниченного в $X$ множества при отображении $A$ будет множеством, ограниченным в $Y$, если $A$ – ограниченный оператор (это свойство можно принять за определение ограниченного оператора). Линейный оператор является непрерывным в том и только в том случае, если он ограничен (докажите!). Для ограниченного линейного оператора

$\sup\limits_{x\neq 0}\dfrac{\|Ax\|_{Y}}{\|x\|_{X}}<\infty$; это число является наименьшим из чисел $C$, при которых выполнено неравенство $\|Ax\|_{Y}\leq C\cdot\|x\|_{X}$, и называется **нормой оператора** $A$:

$$\|A\|=\sup_{x\neq 0}\frac{\|Ax\|_{Y}}{\|x\|_{X}}=\sup_{\|x\|_{X}\leq 1}\|Ax\|_{Y}=\sup_{\|x\|_{X}=1}\|Ax\|_{Y}=\inf_{\left\{C\,\middle|\,\|Ax\|_{Y}\leq C\cdot\|x\|_{X}\right\}}C.$$

**Задача.** $X=Y=C[0,1]$, $Dom\ A$ – линейное многообразие непрерывно дифференцируемых на отрезке $[0,1]$ функций, $Ax(t)=\dfrac{dx}{dt}$. Является ли $A$ линейным непрерывным оператором? □

**Задача**. $X=Y=C[0,1]$, $Ax(t)=\int\limits_{0}^{t}x(\tau)d\tau$. Докажите, что $A$ – линейный непрерывный оператор и найдите его норму. □

2. **Сопряжённый оператор в гильбертовом пространстве**.

Пусть $H$ – гильбертово пространство, $(x,y)$ – скалярное произведение его элементов. Пусть $A$ – **линейный непрерывный** оператор, действующий в $H$. Фиксируем элемент $y\in H$ и рассмотрим функционал $f(x)=(Ax,y)$. Очевидно, что $f$ – линейный функционал, а в силу неравенств

$$|f(x)|=|(Ax,y)|\leq\|Ax\|\cdot\|y\|\leq\|A\|\cdot\|x\|\cdot\|y\|$$

функционал $f$ является ограниченным, причём $\|f\|\leq\|A\|\cdot\|y\|$. По теореме Ф. Рисса существует такой элемент $y^{*}\in H$, что для всех $x\in H$ $f(x)=(x,y^{*})$, т.е. $(Ax,y)=(x,y^{*})$. Элемент $y^{*}$ определяется функционалом $f$ однозначно, а функционал $f$ однозначно определён выбором элемента $y$. Таким образом, получаем определённый на всём $H$ оператор $A^{*}$, действующий по правилу $A^{*}y=y^{*}$, т.е. $(Ax,y)=(x,A^{*}y)$ для всех $x,y\in H$. Такой оператор $A^{*}$ называется **сопряжённым** к оператору $A$.

**Теорема**. Если $A$ – линейный непрерывный оператор, действующий в гильбертовом пространстве $\mathsf{H}$, то сопряжённый к нему оператор $A^*$ является линейным и непрерывным, причём $\|A^*\| = \|A\|$.

**Доказательство**. Выберем произвольные элементы $y_1, y_2 \in \mathsf{H}$. При всех $x \in \mathsf{H}$ выполнены соотношения $(Ax, y_1) = (x, A^* y_1)$ и $(Ax, y_2) = (x, A^* y_2)$; складывая эти соотношения, получим $(Ax, y_1 + y_2) = (x, A^* y_1 + A^* y_2)$ для всех $x$. Это означает, что $A^*(y_1 + y_2) = A^* y_1 + A^* y_2$, т.е. $A^*$ аддитивен.

Докажем однородность оператора $A^*$: при всех $x, y \in \mathsf{H}$

$$(Ax, \alpha y) = \overline{\alpha} \cdot (Ax, y) = \overline{\alpha} \cdot (x, A^* y) = (x, \alpha \cdot A^* y),$$

т.е. $A^*(\alpha y) = \alpha \cdot A^* y$ (черту комплексного сопряжения над $\alpha$ в случае действительного пространства $\mathsf{H}$ можно не учитывать).

Для функционала $f(x) = (x, y^*)$, с помощью которого был определён оператор $A^*$, имеем по теореме Ф. Рисса $\|f\| = \|y^*\|$. Поэтому $\|A^* y\| = \|y^*\| \le \|A\| \cdot \|y\|$, т.е. $A^*$ – ограниченный оператор, причём $\|A^*\| \le \|A\|$. Докажем обратное неравенство. Рассмотрим оператор $A^{**} = (A^*)^*$. Для всех $x, y \in \mathsf{H}$ имеют место равенства $(A^* x, y) = (x, A^{**} y)$ и $(A^* x, y) = \overline{(y, A^* x)} = \overline{(Ay, x)} = (x, Ay)$, поэтому $(x, A^{**} y) = (x, Ay)$, т.е. $A^{**} y = Ay$ при любом $y \in \mathsf{H}$. Это значит, что операторы $A$ и $A^{**}$ совпадают. Заменим в только что полученном неравенстве оператор $A$ на $A^*$, тогда получим $\|A\| = \|(A^*)^*\| \le \|A^*\|$. Отсюда $\|A^*\| = \|A\|$. □

Операция сопряжения линейных непрерывных операторов обладает простыми свойствами (докажите!):

$$(A + B)^* = A^* + B^*, \ (\alpha A)^* = \overline{\alpha} \cdot A^*, \ (AB)^* = B^* A^*.$$

**Задача**. $\mathsf{H}=L_2[0,1]$, $Ax(t)=\int_0^t x(\tau)d\tau$. Найдите сопряжённый оператор $A^*$. □

**Задача**. $\mathsf{H}=L_2[0,1]$, $Ax(t)=t\cdot x(t)$. Найдите сопряжённый оператор $A^*$. □

Линейный непрерывный оператор $A$, совпадающий со своим сопряжённым, называется **самосопряжённым**; он характеризуется равенством $(Ax,y)=(x,Ay)$ для всех $x,y\in\mathsf{H}$.

**Пример**. $\mathsf{H}=L_2[a,b]$ – действительное гильбертово пространство, $y=Ax$ означает, что $y(\tau)=\int_a^b K(\tau,t)x(t)dt$, где $K(\tau,t)$ – заданная непрерывная в квадрате $[a,b]\times[a,b]$ функция (или более общо: $\int_a^b\int_a^b |K(\tau,t)|^2 d\tau\, dt<\infty$). Тогда сопряжённый оператор $A^*$ также является интегральным: $y^*=A^*y$ означает, что $y^*(\tau)=\int_a^b K(t,\tau)y(t)dt$. Действительно, для любой функции $x\in L_2[a,b]$ должно быть $(Ax,y)=(x,y^*)$, т.е.

$$\int_a^b\left(\int_a^b K(\tau,t)x(t)dt\right)y(\tau)d\tau=\int_a^b x(t)y^*(t)dt,$$

$\int_a^b x(t)\left(y^*(t)-\int_a^b K(\tau,t)y(\tau)d\tau\right)dt=0$. В силу произвольности функции $x$ имеем $y^*(t)=\int_a^b K(\tau,t)y(\tau)d\tau$. Меняя ролями $t$ и $\tau$, получаем требуемое выражение для $y^*(\tau)$. Оператор $A$ будет самосопряжённым, если $K(t,\tau)=K(\tau,t)$. (В случае комплексного гильбертова пространства $L_2[a,b]$ интегральный оператор $A$ будет самосопряжённым, если $K(t,\tau)=\overline{K(\tau,t)}$.) □

**Задача**. Пусть $A$ и $B$ – самосопряжённые линейные непрерывные операторы в гильбертовом пространстве. В каком случае оператор $AB$ будет самосопряжённым? □

**Теорема**. Если $A$ – самосопряжённый линейный непрерывный оператор, то $\|A\|=\sup\limits_{\|x\|=1}|(Ax,x)|$.

**Доказательство** проведём для случая действительного гильбертова пространства $H$; для комплексного $H$ оно меняется незначительно. Обозначим $C=\sup\limits_{\|x\|=1}|(Ax,x)|$. При $\|x\|=1$ имеем $|(Ax,x)|\le\|Ax\|\cdot\|x\|\le\|A\|$, поэтому $C\le\|A\|$.

Докажем обратное неравенство. Заметим, что любой элемент $z$ в гильбертовом пространстве $H$ можно представить в виде $z=\|z\|\cdot z'$, где $\|z'\|=1$ (если $z\neq 0$, то $z'=\dfrac{z}{\|z\|}$; если $z=0$, то $z'$ – произвольный элемент, норма которого равна 1). Поэтому для любого $z\in H$

$$(Az,z)=\|z\|^2\cdot(Az',z')\le C\cdot\|z\|^2. \tag{1}$$

Учитывая, что $(Ay,x)=(Ax,y)$, имеем равенства

$$(A(x+y),x+y)=(Ax,x)+2(Ax,y)+(Ay,y),$$

$$(A(x-y),x-y)=(Ax,x)-2(Ax,y)+(Ay,y).$$

Отсюда $4\cdot(Ax,y)=(A(x+y),x+y)-(A(x-y),x-y)$. Тогда по неравенству (1)

$$|(Ax,y)|\le\frac{1}{4}C\cdot\left(\|x+y\|^2+\|x-y\|^2\right)=\frac{1}{2}C\cdot\left(\|x\|^2+\|y\|^2\right)$$ –

в силу равенства параллелограмма для евклидовой нормы. Предполагая, что $Ax\neq 0$, положим в последнем неравенстве $y=\dfrac{\|x\|}{\|Ax\|}\cdot Ax$. Тогда $\|y\|=\|x\|$, и мы получаем

$\dfrac{\|x\|}{\|Ax\|}\cdot\|Ax\|^2 \le C\cdot\|x\|^2$, т.е. $\|Ax\| \le C\cdot\|x\|$. Это же неравенство верно и в случае $Ax = 0$. Следовательно, $\|A\| \le C$. □

Из этой теоремы следует, что в случае $A = A^*$ $\|A\| = \max\left\{\left|\sup\limits_{\|x\|=1}(Ax,x)\right|, \left|\inf\limits_{\|x\|=1}(Ax,x)\right|\right\}$ (докажите!). При этом для любого $x \in \mathrm{H}$ число $(Ax,x)$ действительно и $\inf\limits_{\|x\|=1}(Ax,x)\cdot\|x\|^2 \le (Ax,x) \le \sup\limits_{\|x\|=1}(Ax,x)\cdot\|x\|^2$ (докажите!).

Самосопряжённый оператор называется **положительным**, если $(Ax,x) \ge 0$ для любого $x \in \mathrm{H}$; это символически записывают в виде неравенства $A \ge 0$.

**Пример**. Пусть $\mathrm{L}$ – подпространство (замкнутое) гильбертова пространства $\mathrm{H}$ и $P$ – оператор ортогонального проектирования на $\mathrm{L}$. Тогда $P^2 = P = P^* \ge 0$. $P \le E$ ($E$ – тождественный оператор), $\|P\| = 1$. □

**Задача**. Верно ли, что для любых двух самосопряжённых операторов $A$ и $B$ либо $A \ge B$, либо $B \ge A$ **?** □

Пусть $A = A^* \ge 0$. Положительный самосопряжённый оператор $B$ называется квадратным корнем из оператора $A$, если $B^2 = A$.

Обозначение: $B = A^{\frac{1}{2}}$. Квадратный корень из $A \ge 0$ существует, единственный, он перестановочен с любым оператором, перестановочным с $A$.

Подпространство $\mathrm{L}$ гильбертова пространства $\mathrm{H}$ называется **инвариантным** относительно оператора $A$, если из $x \in \mathrm{L}$ следует $Ax \in \mathrm{L}$. Если подпространство $\mathrm{L}$ инвариантно относительно линейного непрерывного оператора $A$, то его ортогональное дополнение $\mathrm{L}^\perp$ инвариантно относительно $A^*$ (докажите!). В частности, если $A$ – **самосопряжённый** оператор, то ортогональное дополнение к любому его инвариантному подпространству само инвариантно относительно $A$.

Собственные векторы самосопряжённого оператора, отвечающие **различным** его собственным значениям, ортогональны (докажите!). Спектр самосопряжённого оператора лежит на действительной оси комплексной плоскости.

Понятие сопряжённого оператора можно ввести и в том случае, когда оператор $A$ является линейным **неограниченным**. Пусть $Dom\ A$ – область определения оператора $A$. Снова будем рассматривать скалярные произведения $(Ax, y)$, где $x$ пробегает $Dom\ A$. Теперь уже нельзя утверждать, что при любом элементе $y$ такое скалярное произведение как функция от $x$ представимо в виде $(x, y^*)$. Однако существуют пары элементов $y$ и $y^*$, для которых $(Ax, y) = (x, y^*)$ при любом $x \in Dom\ A$; например, $y = y^* = 0$. Это ещё не позволяет ввести оператор $A^*$, сопряжённый к $A$: необходимо, чтобы элемент $y^*$ **однозначно** определялся элементом $y$. Последнее требование будет выполнено в том и только в том случае, если множество $Dom\ A$ плотно в $\mathsf{H}$ (т.е. замыкание множества $Dom\ A$ совпадает с $\mathsf{H}$). В этом случае оператор $A$ имеет сопряжённый оператор $A^*$, его областью определения $Dom\ A^*$ является множество всех тех $y$, для которых существуют $y^*$, удовлетворяющие при любом $x \in Dom\ A$ равенству $(Ax, y) = (x, y^*)$, и $A^* y = y^*$.

3. Множество $S$ метрического пространства $\mathsf{X}$ называется **компактным**, если из любого покрытия его открытыми множествами можно выделить конечное подпокрытие.

Множество $S$ метрического пространства $\mathsf{X}$ называется **счётно-компактным**, если из любой бесконечной последовательности его элементов можно выделить подпоследовательность, сходящуюся к элементу множества $S$.

Пусть $S$ и $N$ – множества метрического пространства $\mathsf{X}$. $N$ называется $\varepsilon$ – сетью для $S$, если для любого элемента $x \in S$ найдётся такой элемент $y \in N$, что расстояние $\rho(x, y) < \varepsilon$.

Множество $S$ называется **вполне ограниченным**, если для любого $\varepsilon > 0$ для него найдётся $\varepsilon$ – сеть, состоящая из конечного числа точек.

**Теорема**. Пусть $S$ – **замкнутое** множество в **полном** метрическом пространстве $X$. Тогда утверждения «$S$ компактно», «$S$ счётно-компактно», «$S$ вполне ограничено» эквивалентны. □

Множество $S$ называется **относительно компактным**, если его замыкание в $X$ компактно.

Для доказательства (относительной) компактности некоторого множества надо либо построить для него конечную $\varepsilon$ – сеть при любом $\varepsilon > 0$, либо использовать критерий компактности множеств в конкретном пространстве $X$.

**Задача**. $X = l_2$, $S = \left\{ x \in l_2 \middle| \ |x_n| \le \dfrac{1}{2^n}, n = 1,2,... \right\}$. При каждом $\varepsilon > 0$ постройте для $S$ конечную $\varepsilon$ – сеть элементов из $l_2$. □

**Теорема Арцела**. Семейство функций в пространстве $C[a,b]$ относительно компактно в том и только в том случае, если оно равномерно ограничено и равностепенно непрерывно. □

В **конечномерном** линейном нормированном пространстве $X$ компактными являются все замкнутые ограниченные множества, и только они. В **бесконечномерном** пространстве это уже неверно. В бесконечномерном банаховом (т.е. полном нормированном) пространстве $X$ шар $\left\{ x \in X \middle| \ \|x\| \le 1 \right\}$ не компактен.

**Пример**. $X = l_2$, $S = \left\{ x \in l_2 \middle| \ \|x\| \le 1 \right\}$. Множеству $S$ принадлежит бесконечная последовательность элементов $x_1 = (1,0,...)$, $x_2 = (0,1,0...)$, …., $x_n = (0,...,0,1,0,...)$, из которой невозможно выделить сходящуюся подпоследовательность, поскольку $\|x_i - x_j\| = \sqrt{2}$ при $i \ne j$. □

Пусть $X$ – линейное нормированное пространство. Последовательность $\{x_n\} \in X$ называется **слабо сходящейся** к элементу $x_0 \in X$, если для любого линейного непрерывного функционала $f(x)$, определённого на $X$, числовая последовательность $\{f(x_n)\}$ сходится к $f(x_0)$. Из слабой сходимости последовательности $\{x_n\}$ не следует, вообще говоря, её сходимость по норме (сильная сходимость); можно лишь утверждать, что слабо сходящаяся последовательность ограничена по норме.

4. **Вполне непрерывные операторы**. Оператор $A$, действующий из банахова пространства $X$ в банахово пространство $Y$, называется **компактным,** если он преобразует всякое множество, ограниченное в $X$, во множество, относительно компактное в $Y$. Если при этом $A$ непрерывен, то он называется **вполне непрерывным**. В случае, когда оператор $A$ **линеен**, эти понятия совпадают.

**Задача**. Докажите, что линейный вполне непрерывный оператор ограничен. ▯

**Задача**. Пусть линейный непрерывный оператор действует из банахова пространства $X$ в **конечномерное** нормированное пространство $Y$. Докажите, что он вполне непрерывен. ▯

**Задача**. Оператор $A$ действует из $l_2$ в $l_2$ по правилу: если $x = (x_1, x_2, \ldots) \in l_2$, то $Ax = (0, x_1, x_2, \ldots)$. Является ли он вполне непрерывным? ▯

**Задача**. Оператор $A$ действует из $l_2$ в $l_2$ по правилу: если $x = (x_1, x_2, \ldots) \in l_2$, то $Ax = \left(\frac{x_1}{2}, \frac{x_2}{4}, \ldots, \frac{x_n}{2^n}, \ldots\right)$. Является ли он вполне непрерывным? ▯

**Задача**. Оператор $A$ действует из $C[0,1]$ в $C[0,1]$ по правилу: если функция $x(t)$ непрерывна на $[0,1]$, то $Ax(t) = \int_0^t x(\tau)d\tau$. Является ли он вполне непрерывным? ▯

**Пример**. Пусть $K(\tau,t)$ – заданная непрерывная в квадрате $[a,b]\times[a,b]$ функция, $y=Ax$ означает $y(\tau)=\int_a^b K(\tau,t)x(t)\,dt$. Такой оператор $A$ можно считать действующим из $C[a,b]$ в $C[a,b]$. В этой паре пространств он вполне непрерывен; доказательство этого факта основано на теореме Арцела. Можно доказать, что $A$ будет вполне непрерывным и как оператор из $L_2[a,b]$ в $L_2[a,b]$. □

**Теорема**. Пусть $A$ – линейный вполне непрерывный оператор, отображающий бесконечномерное банахово пространство $\mathsf{X}$ в себя, а $B$ – линейный непрерывный оператор, действующий в том же пространстве. Тогда $AB$ и $BA$ – вполне непрерывные операторы.

**Доказательство**. Оператор $B$ преобразует произвольное ограниченное множество $S\subset\mathsf{X}$ в ограниченное множество $B(S)$, а оператор $A$ преобразует $B(S)$ в относительно компактное множество $A(B(S))$. Следовательно, оператор $AB$ переводит любое ограниченное множество в относительно компактное, т.е. он вполне непрерывен. Аналогично доказывается, что $BA$ вполне непрерывен. □

**Следствие**. Действующий в **бесконечномерном** банаховом пространстве линейный вполне непрерывный оператор $A$ не может иметь непрерывного обратного оператора $A^{-1}$.

**Доказательство**. Тождественный оператор в бесконечномерном пространстве не является вполне непрерывным, поскольку шар $\{x\in\mathsf{X}\mid \|x\|\le 1\}$ не компактен. □

Этот простой факт приводит к тому, что многие практически важные задачи, которые формулируются в виде уравнения первого рода $Ax=y$ относительно $x$ со вполне непрерывным оператором $A$, оказываются **некорректно поставленными**.

Всякий линейный непрерывный оператор, действующий из $\mathsf{X}$ в $\mathsf{Y}$, переводит последовательность $\{x_n\}\subset\mathsf{X}$, сходящуюся по норме $\|\cdot\|_{\mathsf{X}}$, в последовательность $\{Ax_n\}$, сходящуюся по норме

$\|\cdot\|_Y$. Линейный вполне непрерывный оператор обладает более сильным свойством: всякую **слабо** сходящуюся в $X$ последовательность он переводит в последовательность, сходящуюся **сильно** (т.е. по норме $\|\cdot\|_Y$).

**Теорема**. Пусть $A$ – линейный непрерывный оператор, действующий в гильбертовом пространстве $H$. $A$ вполне непрерывен тогда и только тогда, когда $A^*$ вполне непрерывен.

**Доказательство**. Поскольку $A = (A^*)^*$, достаточно доказать, что $A^*$ вполне непрерывен, если этим свойством обладает $A$.

Докажем сначала, что если $A$ – линейный непрерывный на $H$ оператор, и если оператор $A^*A$ вполне непрерывен, то и оператор $A$ вполне непрерывен. Пусть $S$ – какое-нибудь бесконечное ограниченное по норме множество точек $x \in H$: $\|x\| < C$ для всех $x \in S$. Тогда любая последовательность $\{x_n\} \subset S$ переводится оператором $A^*A$ в последовательность $\{A^*Ax_n\} \subset H$, из которой можно выделить фундаментальную (по норме) подпоследовательность $\{A^*Ax_{n_i}\}$. Поскольку

$$\left\|Ax_{n_i} - Ax_{n_j}\right\|^2 = \left(A\left(x_{n_i} - x_{n_j}\right), A\left(x_{n_i} - x_{n_j}\right)\right) =$$

$$= \left(A^*A\left(x_{n_i} - x_{n_j}\right), x_{n_i} - x_{n_j}\right) \le \left\|A^*Ax_{n_i} - A^*Ax_{n_j}\right\| \cdot \left\|x_{n_i} - x_{n_j}\right\|,$$

$\lim\limits_{i,j\to\infty} \left\|A^*Ax_{n_i} - A^*Ax_{n_j}\right\| = 0$, $\left\|x_{n_i} - x_{n_j}\right\| \le \left\|x_{n_i}\right\| + \left\|x_{n_j}\right\| < 2C$, то $\lim\limits_{i,j\to\infty} \left\|Ax_{n_i} - Ax_{n_j}\right\| = 0$. В силу полноты пространства $H$ последовательность $\{Ax_{n_i}\}$ сходится. Это значит, что оператор $A$ вполне непрерывен.

Если оператор $A$ вполне непрерывен, то вполне непрерывен и оператор $AA^*$ (для непрерывного $A$ оператор $A^*$ непрерывен). Но $AA^* = (A^*)^* A^*$, и для доказательства теоремы остаётся применить только что доказанное утверждение. □

## Вопрос 11.

## Теоремы Фредгольма.

1. Пусть $X$ – **конечномерное** евклидово или унитарное пространство, $T$ – действующий в $X$ линейный оператор, $T^*$ – сопряженный оператор. Введём подпространства $Ker\,T = \{x \in X \,|Tx = 0\}$, $Ran\,T = \{y \in X \,|\exists x,\, y = Tx\}$ и $Ker\,T^*$, $Ran\,T^*$. Справедливо

**Утверждение 1**. $Ker\,T = (Ran\,T^*)^{\perp}$, $Ker\,T^* = (Ran\,T)^{\perp}$.

**Доказательство**. Пусть $x \in Ker\,T$, т.е. $Tx = 0$. Тогда для любого элемента $y \in X$ имеем $(x, T^*y) = (Tx, y) = (0, y) = 0$. В силу произвольности $x$ и $y$ это означает, что $Ran\,T^* \perp Ker\,T$. Вспоминая, что $(T^*)^* = T$, точно так же получаем $Ran\,T \perp Ker\,T^*$. Воспользуемся известными соотношениями

$$\dim Ran\,T + \dim Ker\,T = \dim X,$$

$$\dim Ran\,T^* + \dim Ker\,T^* = \dim X.$$

Число $\dim Ran\,T$ называется рангом оператора $T$, а число $\dim Ran\,T^*$ – рангом $T^*$. Ранги операторов $T$ и $T^*$ совпадают, поэтому

$$\dim Ran\,T + \dim Ker\,T = \dim X,$$

$$\dim Ran\,T^* + \dim Ker\,T^* = \dim X. \ \Box$$

Из утверждения 1 немедленно вытекает

**Теорема Фредгольма**. Для разрешимости уравнения $T\varphi = f$ необходимо и достаточно, чтобы его правая часть $f$ была ортогональна ко всем решениям $\psi_0$ сопряженного однородного уравнения $T^*\psi_0 = 0$. $\Box$

**Альтернатива Фредгольма**. Либо неоднородное уравнение $T\varphi = f$ имеет, и притом единственное, решение при любой правой части $f$, либо сопряженное однородное уравнение $T^*\psi_0 = 0$ имеет ненулевое решение.

**Доказательство**. Пусть ранг оператора $T$ и ранг оператора $T^*$ равны $r$. Возможны два случая: $r = \dim X$, или $r < \dim X$. В первом случае область значений оператора $T$ совпадает со всем пространством: $Ran\,T = X$. Поэтому уравнение $T\varphi = f$ имеет решение при любой правой части $f$, оператор $T$ невырожденный. В этом же случае $\dim Ker\,T^* = 0$. Поэтому $Ker\,T^*$ не имеет ненулевых векторов, т.е. сопряженное однородное уравнение $T^*\psi_0 = 0$ не имеет ненулевых решений.

В случае $r < \dim X$ область значений оператора $T$ не совпадает с $X$, и неоднородное уравнение $T\varphi = f$ может иметь решение **не при любой** правой части $f$. В этом же случае $Ker\,T^*$ состоит не только из нулевого элемента, т.е. уравнение $T^*\psi_0 = 0$ имеет ненулевые решения. ▯

С точки зрения изучения систем линейных алгебраических уравнений эти теоремы означают следующее. Для разрешимости неоднородной системы с **квадратной** матрицей необходимо и достаточно, чтобы её правая часть была ортогональна ко всем решениям сопряжённой однородной системы. Либо неоднородная система с **квадратной** матрицей имеет, и притом единственное, решение при любой правой части, либо сопряженная однородная система имеет ненулевое решение.

2. В теории Фредгольма изучаются интегральные уравнения

$$\varphi(t) - \mu \cdot \int_a^b K(t,\tau)\varphi(\tau)d\tau = f(t), \tag{1}$$

$$\varphi_0(t) - \mu \cdot \int_a^b K(t,\tau)\varphi_0(\tau)d\tau = 0, \tag{2}$$

$$\psi(t) - \mu \cdot \int_a^b K(\tau,t)\psi(\tau)d\tau = g(t), \tag{3}$$

$$\psi_0(t)-\mu\cdot\int_a^b K(\tau,t)\psi_0(\tau)d\tau=0; \qquad (4)$$

здесь $a\le t\le b$, $\mu$ – действительный параметр, $K$, $f$, $g$ – заданные действительные функции, $\varphi$, $\varphi_0$, $\psi$, $\psi_0$ – искомые функции. Если параметр $\mu$, заданные и искомые функции комплексны, то вместо (3) и (4) рассматривают уравнения

$$\psi(t)-\overline{\mu}\cdot\int_a^b \overline{K(\tau,t)}\psi(\tau)d\tau=g(t),$$

$$\psi_0(t)-\overline{\mu}\cdot\int_a^b \overline{K(\tau,t)}\psi_0(\tau)d\tau=0;$$

мы будем далее считать, что $\mu$ и все функции действительны.

Интегральный оператор $(A\varphi)(t)\equiv\int_a^b K(t,\tau)\varphi(\tau)d\tau$ обычно называют оператором Фредгольма, если он действует в некотором функциональном пространстве $X$ и **вполне непрерывен** в $X$. Уравнение (1) называют интегральным уравнением 2-го рода. Предполагается, что правая часть $f$ и искомая функция $\varphi$ принадлежат одному и тому же функциональному пространству $X$; это требует определённых предположений о ядре $K(t,\tau)$ интегрального оператора $A$. Значения параметра $\mu$, при которых однородное интегральное уравнение (2) имеет в $X$ ненулевые решения, называют **характеристическими числами** оператора $A$, а сами эти решения – его собственными функциями. Характеристические числа и собственные значения оператора $A$ взаимно обратны, а их собственные функции совпадают.

Теоремы Фредгольма – это сформулированные ниже утверждения, которые справедливы при некоторых предположениях о ядре $K(t,\tau)$ и при определенном выборе функционального пространства $X$, содержащего $f$ и искомые функции $\varphi$, $\psi_0$.

I. Неоднородное уравнение (1) разрешимо при тех и только тех правых частях $f$, которые ортогональны каждому решению союзного (сопряженного) однородного уравнения (4):

$\int\limits_a^b f(t)\psi_0(t)dt = 0$ $\quad$ $(\int\limits_a^b f(t)\overline{\psi_0(t)}dt = 0$ в случае комплекснозначных функций). ▯

II. (Альтернатива Фредгольма.) Либо неоднородное уравнение (1) при любой правой части $f$ имеет одно и только одно решение, либо однородное уравнение (2) имеет ненулевое решение. ▯

III. При фиксированном значении параметра $\mu$ однородные уравнения (2) и (4) имеют одно и то же, притом конечное, число линейно независимых решений. ▯

IV. Множество характеристических чисел оператора $A$ не более чем счетно и может иметь единственную предельную точку лишь на бесконечности. ▯

Часто под альтернативой Фредгольма понимают вытекающее из теорем Фредгольма следующее утверждение: либо уравнение (1) и союзное (сопряженное) с ним уравнение (3) имеют единственные решения $\varphi$, $\psi$, каковы бы ни были заданные функции $f$, $g$, либо соответствующие однородные уравнения (2) и (4) имеют ненулевые решения, причем число линейно независимых решений конечно и одинаково для обоих уравнений. Альтернатива Фредгольма означает, что если $\mu$ не является характеристическим числом оператора $A$, то оба уравнения (1) и (3) однозначно разрешимы **при любых правых частях**, а если $\mu$ – характеристическое число, то оба однородных уравнения (2) и (4) имеют **одинаковое конечное** число линейно независимых решений.

**Пример**. $K(t,\tau) = \sum\limits_{i=1}^{n} \alpha_i(t)\beta_i(t)$, где $\alpha_1, \ldots, \alpha_n$ – непрерывные линейно независимые на $[a,b]$ функции, $\beta_1, \ldots, \beta_n$ – непрерывные линейно независимые на $[a,b]$ функции; такое ядро $K$ называется вырожденным. Уравнения (1) – (4) с вырожденными ядрами сводятся к системам линейных алгебраических уравнений.

Пусть в уравнении (1) с вырожденным ядром функция $f$ и искомая функция $\varphi$ непрерывны на $[a,b]$. Уравнение (1) имеет вид

$$\varphi(t) - \mu \cdot \sum_{i=1}^{n} \alpha_i(t) \cdot \int_a^b \beta_i(\tau)\varphi(\tau)d\tau = f(t).$$

Если ввести обозначения $p_i = \int_a^b \beta_i(\tau)\varphi(\tau)d\tau$, то получим $\varphi(t) = \mu \cdot \sum_{i=1}^{n} p_i \cdot \alpha_i(t) + f(t)$, т.е. решение интегрального уравнения с вырожденным ядром сводится к нахождению чисел $p_i$, $i = 1, \ldots, n$. Подставим в (1) последнее выражение функции $\varphi$ и воспользуемся линейной независимостью функций $\alpha_1, \ldots, \alpha_n$; тогда получим при $i = 1, \ldots, n$ уравнения

$$p_i - \int_a^b \beta_i(\tau) \cdot \left( f(\tau) + \mu \cdot \sum_{j=1}^{n} p_j \cdot \alpha_j(\tau) \right) d\tau = 0 .$$

Вводя обозначения $f_i = \int_a^b \beta_i(\tau) f(\tau)d\tau$ и $\theta_{ij} = \int_a^b \beta_i(\tau)\alpha_j(\tau)d\tau$, будем иметь для определения неизвестных $p_1, \ldots, p_n$ систему линейных алгебраических уравнений

$$p_i - \mu \cdot \sum_{j=1}^{n} \theta_{ij} p_j = f_i , \quad i = 1, \ldots, n , \tag{5}$$

которая эквивалентна уравнению (1) с вырожденным ядром. Её определитель

$$D(\mu) = \begin{vmatrix} 1 - \mu \cdot \theta_{11} & -\mu \cdot \theta_{12} & \ldots & -\mu \cdot \theta_{1n} \\ -\mu \cdot \theta_{21} & 1 - \mu \cdot \theta_{22} & \ldots & -\mu \cdot \theta_{2n} \\ . & . & \ldots & . \\ -\mu \cdot \theta_{n1} & -\mu \cdot \theta_{n2} & \ldots & 1 - \mu \cdot \theta_{nn} \end{vmatrix}$$

является многочленом от $\mu$ степени не выше $n$. $D(\mu)$ не равен тождественно нулю, поскольку $D(0) = 1$. Поэтому имеется не более $n$ различных значений $\mu$, при которых $D(\mu) = 0$. При этих $\mu$ система уравнений (5) (а вместе с ней и интегральное уравнение (1)) либо неразрешима, либо имеет бесконечно много решений. При остальных $\mu$ система (5) (и уравнение (1)) имеет единственное решение. Отсюда следуют утверждения теорем Фредгольма для уравнений (1) – (4) с вырожденными ядрами (докажите!). □

**Задача**. $\varphi(t)-\mu\cdot\int\limits_0^{\pi}\sin(t+\tau)\varphi(\tau)d\tau=f(t)$.

Найдите характеристические числа и собственные функции интегрального оператора. Что означают теоремы Фредгольма в данном случае?

**Пример**. Ядро $K(t,\tau)$ непрерывно, $f$, $g$ и искомые функции принадлежат пространству $C[a,b]$. Тогда справедливы теоремы Фредгольма. □

**Пример**. Ядро $K(t,\tau)$ удовлетворяет условию $\int\limits_a^b\int\limits_a^b K^2(t,\tau)dt\,d\tau<\infty$, $f$, $g$ и искомые функции принадлежат пространству $L_2[a,b]$. Тогда справедливы теоремы Фредгольма. □

В качестве области интегрирования в уравнениях (1) – (4) вместо отрезка $[a,b]$ можно рассматривать некоторое ограниченное или неограниченное измеримое множество в пространстве любого конечного числа измерений.

К интегральным уравнениям 2-го рода сводятся разнообразные теоретические и практические задачи: краевые задачи Штурма-Лиувилля о собственных значениях дифференциальных операторов, задачи Дирихле и Неймана для уравнения Лапласа и др. В наиболее простых случаях оказываются справедливыми теоремы Фредгольма. Однако, многие важные задачи приводят к нефредгольмовым интегральным уравнениям.

**Пример** интегрального уравнения, для которого не выполнены теоремы Фредгольма:

$$\varphi(t)=\mu\cdot\int\limits_0^{+\infty}\sin(t\cdot\tau)\varphi(\tau)d\tau.$$

При $\mu_{1,2} = \pm\sqrt{\dfrac{2}{\pi}}$ имеется **бесконечное** число линейно независимых решений: при этих значениях $\mu$ уравнению удовлетворяют соответственно функции $\varphi_{1,2} = \sqrt{\dfrac{\pi}{2}}e^{-ct} \pm \dfrac{t}{t^2+c^2}$ **при любом** $c > 0$. (Отметим, что заменой $t = tg\,\xi$, $\tau = tg\,\zeta$ уравнение можно свести к уравнению с конечным интервалом интегрирования.) ▯

3. Пусть $B$ – вполне непрерывный линейный оператор, действующий в гильбертовом пространстве $\mathrm{H}$. Тогда и сопряженный к нему оператор $B^*$ вполне непрерывен. Докажем некоторые свойства операторов $T = E - B$ ($E$ – тождественный оператор) и $T^* = E - B^*$. Будем рассматривать множества $Ker\,T = \{x \in \mathrm{H}\,|\,Tx = 0\}$ – совокупность всех нулей оператора $T$, $Ran\,T = \{y \in \mathrm{H}\,|\,\exists x,\, y = Tx\}$ – область значений оператора $T$ и $Ker\,T^*$, $Ran\,T^*$. Совокупность всех нулей любого линейного непрерывного оператора является (замкнутым) линейным подпространством. Область значений линейного непрерывного оператора есть линейное многообразие, но не обязательно замкнутое. Однако для операторов $T$ и $T^*$ замкнутость их областей значений имеет место:

**Утверждение 2**. Множества $Ran\,T$ и $Ran\,T^*$ замкнуты и, следовательно, являются подпространствами в $\mathrm{H}$. (В доказательстве этого утверждения существенно используется полная непрерывность оператора $B$.)

**Утверждение 3**. $Ker\,T = \left(Ran\,T^*\right)^{\perp}$, $Ker\,T^* = \left(Ran\,T\right)^{\perp}$.

**Доказательство**. Подпространства $Ker\,T$ и $Ran\,T^*$ замкнуты. Кроме того, для любого $x \in Ker\,T$ имеем $\left(x, T^*y\right) = \left(Tx, y\right) = 0$ при всех $y \in \mathrm{H}$, т.е. $Ker\,T \perp Ran\,T^*$. Если $x \perp Ran\,T^*$, то для любого $y \in \mathrm{H}$ получим $\left(Tx, y\right) = \left(x, T^*y\right) = 0$, т.е. $x \in Ker\,T$. Поэтому в $H$ нет ни одного ненулевого элемента, ортогонального одновременно и к $Ker\,T$, и к $Ran\,T^*$. Следовательно, пространство $H$ является ортогональной суммой подпространств $Ker\,T$ и $Ran\,T^*$.

Заменяя в доказанном утверждении $T$ на $T^*$, получаем разложение $H$ в ортогональную сумму подпространств $Ker\,T^*$ и $Ran\,T$. ▯

При каждом натуральном $k$ положим $H_k = Ran\left(T^k\right)$; в частности, $H_1 = Ran\,T$. $T\,H_k = H_{k+1}$, имеют место включения $H \supset H_1 \supset H_2 \supset \ldots$. По утверждению 2 все подпространства $H_k$ замкнуты (докажите!).

**Утверждение 4**. Существует такой номер $k_0$, что $H_{k+1} = H_k$ при всех $k \geq k_0$.

**Доказательство**. Пусть указанного $k_0$ не существует, т.е. все подпространства $H_k$ различны. Тогда можно построить такую ортонормированную последовательность $\{x_k\}$, что $x_k \in H_k$ и $x_k \perp H_{k+1}$. Пусть $m > k$, тогда $Bx_m - Bx_k = (x_m - Tx_m + Tx_k) - x_k$, $(x_m - Tx_m + Tx_k) \in H_{k+1}$. Поэтому $\|Bx_m - Bx_k\| \geq \|x_k\| = 1$. Следовательно, из последовательности $\{Bx_k\}$ выбрать сходящуюся подпоследовательность невозможно. Это противоречит полной непрерывности оператора $B$. ▯

**Утверждение 5**. Если $Ker\,T = 0$, то $Ran\,T = H$; если $Ker\,T^* = 0$, то $Ran\,T^* = H$. Если $Ran\,T = H$, то $Ker\,T = 0$; если $Ran\,T^* = H$, то $Ker\,T^* = 0$.

**Доказательство**. Если $Ker\,T = 0$, то оператор $T$ отображает $H$ на $Ran\,T$ взаимно однозначно. При этом, если $Ran\,T \neq H$, то в цепочке включений $H \supset H_1 \supset H_2 \supset \ldots$ все подпространства $\mathrm{H}_k$ различны (докажите!), что невозможно по утверждению 4. Поэтому $Ran\,T = \mathrm{H}$. Аналогично, $Ran\,T^* = \mathrm{H}$, если $Ker\,T^* = 0$. Но тогда $Ran\,T^* = \mathrm{H}$ и $Ker\,T = 0$. □

Из утверждений 2 – 5 вытекают теоремы Фредгольма для уравнений

$$T\varphi = f \quad (6); \qquad T\varphi_0 = 0 \quad (7);$$

$$T^*\psi = g \quad (8); \qquad T^*\psi_0 = 0 \quad (9)$$

в гильбертовом пространстве $\mathrm{H}$.

I. Для разрешимости уравнения (6) необходимо и достаточно, чтобы его правая часть $f$ была ортогональна ко всем решениям $\psi_0$ уравнения (9).

**Доказательство** следует из утверждения 3. □

II. (Альтернатива Фредгольма.) Либо неоднородное уравнение (6) имеет, и притом единственное, решение при любой правой части $f \in \mathrm{H}$, либо сопряженное однородное уравнение (9) имеет ненулевое решение. (Другая, эквивалентная формулировка: … либо уравнение (7) имеет ненулевое решение.)

**Доказательство** следует из утверждений 3 и 5. □

III. Однородные уравнения (7) и (9) имеют одно и то же, притом **конечное**, число линейно независимых решений.

**Доказательство**. Подпространство $Ker\,T$ конечномерно. Действительно, в противном случае в нём нашлась бы бесконечная ортонормированная последовательность $\{x_k\}$, для которой $Bx_k = x_k$. Следовательно, при $k \neq j$ будем иметь $\|Bx_k - Bx_j\| = \|x_k - x_j\| = \sqrt{2}$. Но тогда из последовательности $\{Bx_k\}$ невозможно выбрать сходящуюся подпоследовательность, что противоречит полной непрерывности оператора $B$. Аналогично и $Ker\,T^*$ конечномерно.

Пусть $\dim Ker\,T = m$, $\dim Ker\,T^* = n$. Предположим, что $m < n$. Выберем ортонормированный базис $\{\varphi_1, \ldots, \varphi_m\}$ в $Ker\,T$ и ортонормированный базис $\{\psi_1, \ldots, \psi_n\}$ в $Ker\,T^*$.

Построим оператор $S$ по правилу $Sx = Tx + \sum_{i=1}^{m} (x, \varphi_i) \cdot \psi_i$.

Оператор $S$ отличается от оператора $T$ на конечномерный оператор, поэтому для $S$ выполнены утверждения 2 – 5. Докажем, что уравнение $Sx = 0$ имеет только тривиальное решение. Пусть $Sx_0 = 0$. По утверждению 3 все элементы $\psi_i$ ортогональны всем элементам вида $Tx$, поэтому $Tx_0 = 0$ и $(x_0, \varphi_i) = 0$ при всех $i = 1, \dots, m$. Следовательно, $x_0 = 0$, т.е. $Ker\, S = 0$. Тогда по утверждению 5 $Ran\, S = H$, откуда следует, что для $\psi_{m+1}$ найдется такой элемент $y$, что $Sy = Ty + \sum_{i=1}^{m} (y, \varphi_i) \cdot \psi_i = \psi_{m+1}$. Умножим последнее равенство скалярно на $\psi_{m+1}$ и вспомним, что $Ty \in Ran\, T$, $Ran\, T \perp Ker\, T^*$. Тогда получим противоречие: 0=1. Следовательно, $m \geq n$.

Меняя ролями $T$ и $T^*$, точно так же докажем, что $n \geq m$. Поэтому $m = n$. □

**Замечание**. В любом банаховом пространстве однородное уравнение $B\varphi_0 = \varphi_0$, где $B$ – вполне непрерывный оператор, может иметь лишь конечное число линейно независимых решений (докажите!). Это и означает, что подпространство $Ker(E - B)$ конечномерно. □

В качестве оператора $B$ можно взять оператор $\mu \cdot A$, где $\mu$ – параметр, $A$ – вполне непрерывный оператор. Тогда уравнения (6) – (9) будут иметь вид $\varphi - \mu \cdot A\varphi = f$, $\varphi_0 = \mu \cdot A\varphi_0$, $\psi - \overline{\mu} \cdot A^*\psi = g$, $\psi_0 = \overline{\mu} \cdot A^*\psi_0$. Полагая $\lambda = \frac{1}{\mu}$, получим $\lambda \cdot \varphi - A\varphi = \widetilde{f}$, $A\varphi_0 = \lambda\varphi_0$, $\overline{\lambda} \cdot \psi - A^*\psi = \widetilde{g}$, $A^*\psi_0 = \overline{\lambda} \cdot \psi_0$ при $\lambda \neq 0$. Спектр вполне непрерывного оператора $A$ в H состоит из конечного или счетного числа отличных от нуля собственных значений конечной кратности и точки $\lambda = 0$ (точка $\lambda = 0$

принадлежит спектру оператора $A$, поскольку $A^{-1}$ не может быть ограниченным), причем только $\lambda = 0$ может быть предельной точкой для последовательности собственных значений. Это и составляет утверждение теоремы IV.

Теоремы Фредгольма распространяются на случай, когда оператор $T$ действует в банаховом пространстве $X$; тогда $T^*$ действует в сопряженном пространстве $X^*$. Теоремы Фредгольма справедливы не только для оператора $T = E - B$, где $B$ – вполне непрерывный оператор; для выполнения теорем Фредгольма требуется, чтобы линейный оператор $T$ был нормально разрешим, и чтобы $Ker\,T$ и $Ker\,T^*$ были конечномерны, причем $\dim Ker\,T = \dim Ker\,T^*$.

## Вопрос 12.

## Теорема Гильберта-Шмидта.

1. Линейный оператор $A$, действующий в **конечномерном** евклидовом или унитарном пространстве $X$, является самосопряженным в том и только в том случае, если в $X$ существует ортонормированный базис, в котором матрица оператора $A$ действительна и диагональна. Это означает, что все собственные значения такого оператора действительны, и в $X$ имеется ортонормированный базис, целиком состоящий из собственных векторов оператора $A$. Теорема Гильберта-Шмидта распространяет этот факт на **вполне непрерывные самосопряженные** операторы в **гильбертовом** пространстве.

Прежде чем сформулировать теорему Гильберта-Шмидта, напомним основные спектральные свойства вполне непрерывных операторов и самосопряженных ограниченных операторов.

**Теорема 1**. Пусть $A$ – линейный вполне непрерывный оператор, действующий в банаховом пространстве. Если его собственное значение $\lambda \neq 0$, то отвечающее этому $\lambda$ собственное подпространство конечномерно. Для любого $\delta > 0$ вне круга $|\lambda| \leq \delta$ комплексной плоскости (вне интервала $|\lambda| \leq \delta$ в случае действительного пространства $X$) может содержаться лишь конечное число собственных значений оператора $A$. ▯

Из теоремы 1 вытекает, что все собственные значения вполне непрерывного оператора можно перенумеровать в порядке невозрастания их модулей: $|\lambda_1| \geq |\lambda_2| \geq \dots$. Будем при этом повторять каждое собственное значение столько раз, какова его геометрическая кратность, и в том же порядке перечислять отвечающие этим $\lambda_k$ собственные векторы: $\psi_1, \psi_2, \dots$.

**Теорема 2**. Если $A$ – линейный вполне непрерывный оператор, действующий в банаховом пространстве $X$, то его спектр состоит из конечного или счетного множества точек, единственной предельной точкой которого может быть $\lambda = 0$. Если $X$ бесконечномерно, то $\lambda = 0$ принадлежит спектру оператора $A$.

Если $\lambda \neq 0$ не является собственным значением оператора $A$, то это число $\lambda$ – регулярная точка оператора $A$. □

**Теорема 3**. Пусть $A$ – самосопряженный линейный ограниченный оператор, действующий в гильбертовом пространстве $\mathrm{H}$. Все собственные значения такого оператора действительны, а собственные векторы, отвечающие различным собственным значениям, ортогональны. Весь спектр такого оператора лежит на отрезке $[m, M]$ действительной оси, где $m = \inf\limits_{\|x\|=1}(Ax, x)$, $M = \sup\limits_{\|x\|=1}(Ax, x)$, причем числа $m$ и $M$ являются точками спектра. □

Отметим, что из теоремы 3 не следует существование собственных значений самосопряженного оператора.

**Пример**. В пространстве $L_2[0,1]$ оператор $A = A^*$ действует по правилу $Ax(t) = t \cdot x(t)$, $0 \le t \le 1$. Все точки отрезка $[0,1]$ являются точками спектра, $m = 0$, $M = 1$. Оператор $A$ не имеет собственных значений. □

Из теоремы 3 вытекает лишь, что самосопряженный оператор имеет непустой спектр.

Спектральные свойства оператора, действующего в гильбертовом пространстве, можно существенно уточнить, если он одновременно является вполне непрерывным и самосопряженным.

**Теорема 4**. Пусть $A$ – вполне непрерывный самосопряженный линейный оператор, действующий в гильбертовом пространстве $\mathrm{H}$. Если $A \neq 0$, то $A$ имеет по крайней мере одно собственное значение $\lambda \neq 0$. Если $M = \sup\limits_{\|x\|=1}(Ax, x) \neq 0$, то число $M$ является наибольшим собственным значением оператора $A$; если $m = \inf\limits_{\|x\|=1}(Ax, x) \neq 0$, то число $m$ является наименьшим собственным значением оператора $A$. $\|A\| = |\lambda_1|$, где $\lambda_1$ – наибольшее по модулю собственное значение. □

Если отличный от нулевого вполне непрерывный оператор не является самосопряженным в $\mathrm{H}$, то он может и не иметь ни одного собственного вектора.

**Пример**. В пространстве $l_2$ оператор $A$ действует по правилу $Ax = (0, x_1, x_2/2, x_3/3, \ldots)$ для любого $x = (x_1, x_2, x_3, \ldots) \in l_2$. Оператор $A$ вполне непрерывен, его спектр состоит только из точки непрерывного спектра $\lambda = 0$. □

**Пример**. Интегральный оператор Вольтерра $Ax(t) = \int_a^t K(t,\tau)x(\tau)d\tau$ с непрерывным ядром $K(t,\tau)$ действует в пространстве $L_2[a,b]$. Его спектр состоит из единственной точки $\lambda = 0$. Если $K(t,\tau)$ не равно тождественно нулю, то $\lambda = 0$ не является собственным значением. □

Введём обозначения: $Ker\ A = \{x \in \mathsf{H} \,|\, Ax = 0\}$ – совокупность всех нулей оператора $A$. $Ran\ A = \{y \in \mathsf{H} \,|\, \exists x,\ y = Ax\}$ – область значений оператора $A$ (для отличного от конечномерного линейного вполне непрерывного оператора $A$ в банаховом пространстве $\mathsf{X}$ областью значений будет **незамкнутое** в $\mathsf{X}$ линейное многообразие).

**Теорема Гильберта-Шмидта**. Для любого вполне непрерывного самосопряженного линейного оператора $A \neq 0$, действующего в гильбертовом пространстве $\mathsf{H}$, существует ортонормированная система $\{\psi_k\}$ собственных векторов, отвечающих отличным от нуля собственным значениям $\{\lambda_k\}$ оператора $A$, такая, что для любого $x \in \mathsf{H}$ имеется единственное представление в виде $x = \sum_k c_k \psi_k + x_0$, где $x_0 \in Ker\ A$ (число слагаемых в сумме конечно или счетно в зависимости от оператора $A$). При этом $Ax = \sum_k \lambda_k c_k \psi_k$, и если система $\{\psi_k\}$ бесконечна, то $\lim_{k\to\infty} \lambda_k = 0$. В многообразии $Ran\ A$ система элементов $\{\psi_k\}$ является замкнутой, т.е. при любом $x \in \mathsf{H}$ для $Ax$ выполнено равенство Парсеваля. □

Указанное в теореме разложение элемента $Ax$ есть разложение в сходящийся по $\|\cdot\|_H$ ряд Фурье по ортонормированной системе $\{\psi_k\}$; коэффициенты Фурье в этом разложении равны $(Ax,\psi_k)=(x,A\psi_k)=(x,\lambda_k\psi_k)=\lambda_k\cdot(x,\psi_k)$. $c_k=(x,\psi_k)$.

Теорема Гильберта-Шмидта означает, что если H – **сепарабельное** гильбертово пространство, то для всякого вполне непрерывного самосопряженного оператора $A$ в H существует счетный ортонормированный базис, целиком состоящий из собственных векторов этого оператора. Для построения такого базиса в случае $Ker\ A\neq\{0\}$ достаточно дополнить систему собственных векторов $\{\psi_k\}$ произвольным (не более чем счётным) ортонормированным базисом подпространства $Ker\ A$, что возможно в силу сепарабельности $Ker\ A$; векторы базиса в $Ker\ A$ будут собственными векторами оператора $A$, отвечающими нулевому собственному значению.

**Задача**. Пусть действующий в гильбертовом пространстве H вполне непрерывный самосопряженный оператор обратим. Докажите, что в H существует ортонормированный базис из его собственных векторов, отвечающих отличным от нуля собственным значениям. Какой части спектра принадлежит в этом случае $\lambda=0$? □

2. Пусть $A$ – вполне непрерывный самосопряженный линейный оператор, действующий в гильбертовом пространстве H, $f$ – фиксированный элемент из H, $\mu$ – числовой параметр. Уравнение $\varphi-\mu\cdot A\varphi=f$ можно исследовать не только на основе общих теорем Фредгольма, но и при помощи теоремы Гильберта-Шмидта. Представим искомое решение $\varphi$ и заданный элемент $f$ в виде

$$\varphi=\sum_k c_k\psi_k+\varphi_0,\quad \varphi_0\in Ker\ A,\quad c_k=(\varphi,\psi_k),$$

$$f=\sum_k d_k\psi_k+f_0,\quad f_0\in Ker\ A,\quad d_k=(f,\psi_k).$$

Тогда получим уравнение

$$\varphi_0 + \sum_k c_k \psi_k - \mu \cdot \sum_k \lambda_k c_k \psi_k = f_0 + \sum_k d_k \psi_k ,$$

из которого следует $\varphi_0 = f_0$, $c_k \cdot (1 - \mu \cdot \lambda_k) = d_k$.

Если **при всех** $k$ выполнены неравенства $1 - \mu \cdot \lambda_k \neq 0$, то $c_k = \dfrac{d_k}{1 - \mu \cdot \lambda_k}$, $\sum_k |c_k|^2 \leq \dfrac{1}{\min_k |1 - \mu \cdot \lambda_k|^2} \sum_k |d_k|^2 < \infty$, и в $H$ существует единственное решение $\varphi$ исходного уравнения:

$$\varphi = \sum_k \frac{d_k}{1 - \mu \cdot \lambda_k} \cdot \psi_k + f_0 .$$

Выразим его через $f$, $\mu$, $\{\lambda_k\}$ и $\{\psi_k\}$:

$$\varphi = f + \mu \cdot A\varphi = f + \mu \sum_k \frac{d_k \cdot A\psi_k}{1 - \mu \cdot \lambda_k} + \mu \cdot Af_0 =$$

$$= f + \sum_k \frac{\mu \cdot \lambda_k}{1 - \mu \cdot \lambda_k} \cdot d_k \cdot \psi_k .$$

Если же $\dfrac{1}{\mu}$ совпадает с одним из собственных значений оператора $A$ (это собственное значение может иметь конечную кратность), то решение $\varphi$ существует лишь в том случае, если соответствующие коэффициенты $d_k = 0$, т.е. если элемент $f$ ортогонален всем собственным векторам, отвечающим собственному значению $\dfrac{1}{\mu}$. При этом соответствующие коэффициенты $c_k$ можно выбрать произвольно, поэтому решение $\varphi$ не единственно: если некоторый элемент $\varphi$ является решением, а $\varphi_{\frac{1}{\mu}}$ – отвечающий $\dfrac{1}{\mu}$ собственный вектор, то и $\varphi + \varphi_{\frac{1}{\mu}}$ является решением.

3. Интегральный оператор $Ax = \int\limits_a^b K(t,\tau)x(\tau)d\tau$ определяется ядром $K(t,\tau)$. Если ядро непрерывно при $a \le t \le b$ и $a \le \tau \le b$, или, по крайней мере, удовлетворяет условию $\int\limits_a^b\int\limits_a^b |K(t,\tau)|^2 dt\, d\tau < \infty$, то оператор $A$ действует из пространства $L_2[a,b]$ в $L_2[a,b]$ и вполне непрерывен.

Ядро оператора называется симметричным, если $K(t,\tau) \equiv \overline{K(\tau,t)}$ (в случае действительного ядра $K(t,\tau) \equiv K(\tau,t)$). По теореме Фубини имеем для любых функций $x, y \in L_2[a,b]$

$$(Ax,y)_{L_2} = \int\limits_a^b \left( \int\limits_a^b K(t,\tau)x(\tau)d\tau \right) \overline{y(t)}dt =$$

$$= \int\limits_a^b \left( \int\limits_a^b K(t,\tau)\overline{y(t)}dt \right) x(\tau)d\tau =$$

$$= \int\limits_a^b x(\tau) \cdot \overline{\left( \int\limits_a^b \overline{K(t,\tau)} y(t)dt \right)} d\tau = \left(x, A^* y\right)_{L_2},$$

где $A^* y = \int\limits_a^b \overline{K(t,\tau)} y(t)dt$. Если ядро $K(t,\tau)$ симметрично, то

$A^* y = \int\limits_a^b \overline{K(t,\tau)} y(t)dt = \int\limits_a^b K(\tau,t)y(t)dt = Ay$, т.е. $A = A^*$. Везде далее будем предполагать, что ядро оператора $A$ симметрично.

Интегральный оператор $A$ с симметричным ядром имеет **конечное** число ненулевых собственных значений в том и только в том случае, если ядро $K(t,\tau)$ **вырожденное**. Будем далее предполагать, что ядро оператора $A$ не является вырожденным, и, следовательно, имеется бесконечная последовательность ненулевых собственных значений $\lambda_k \underset{k\to\infty}{\to} 0$, которой отвечает в $L_2[a,b]$ ортонормированная система $\{\psi_k(t)\}$ собственных функций оператора $A$.

Говорят, что функция $y(t)$ **истокообразно представима** через ядро $K(t,\tau)$, если существует такая функция $x \in L_2[a,b]$, что $y(t) = \int_a^b K(t,\tau)x(\tau)d\tau$ при $a \le t \le b$, т.е. $y \in Ran\, A$.

Истокообразно представимая функция $y(t)$ может быть разложена на отрезке $a \le t \le b$ в ряд Фурье по ортонормированной системе $\{\psi_k(t)\}$, который сходится к $y(t)$ в **среднем квадратическом** (т.е. по норме $\|\cdot\|_{L_2}$). Однако для рассматриваемого интегрального оператора $A$ этот результат может быть усилен: ряд Фурье такой функции $y(t)$ по системе $\{\psi_k(t)\}$ сходится к $y(t)$ **абсолютно** и **равномерно**. Этот факт играет основную роль в теории интегральных уравнений с симметричным ядром.

**Утверждение**. Пусть функция $K(t,\tau)$ непрерывна по совокупности аргументов при $a \le t \le b$ и $a \le \tau \le b$. Тогда для любой функции $x \in L_2[a,b]$ функция $y(t) = \int_a^b K(t,\tau)x(\tau)d\tau$ **непрерывна** на отрезке $a \le t \le b$ и все собственные функции $\psi_k(t)$ **непрерывны** на $a \le t \le b$. □

**Теорема Гильберта-Шмидта**. Пусть симметричное ядро $K(t,\tau)$ не равно тождественно нулю, причем функция $K(t,\tau)$ непрерывна по совокупности аргументов. Если непрерывная функция $y(t)$ истокообразно представима через ядро $K(t,\tau)$, то она может быть разложена в ряд по ортонормированной системе собственных функций $\{\psi_k(t)\}$ интегрального оператора

$\int\limits_a^b K(t,\tau)x(\tau)d\tau$, который сходится на отрезке $a \le t \le b$ к функции $y(t)$ **абсолютно** и **равномерно**.

**Доказательство** проведем для случая **действительного** пространства $L_2[a,b]$: тогда ядро $K(t,\tau) = K(\tau,t)$, функции $x(t)$, $y(t)$ и все собственные функции $\psi_k(t)$ действительны.

Абсолютную и равномерную сходимость ряда

$$\sum_{k=1}^{\infty} \lambda_k \cdot c_k \cdot \psi_k(t), \tag{1}$$

где $c_k = \int\limits_a^b x(\tau)\psi_k(\tau)d\tau$, докажем по критерию Коши. Фиксируем произвольное $t$ на отрезке $a \le t \le b$. По определению собственных функций $\lambda_k \cdot \psi_k(t) = \int\limits_a^b K(t,\tau)\psi_k(\tau)d\tau$, т.е. при **фиксированном** $t$ числа $\lambda_k \cdot \psi_k(t)$ суть коэффициенты Фурье по ортонормированной системе $\{\psi_k(\tau)\}$ ядра $K(t,\tau)$ как функции от аргумента $\tau$. По неравенству Бесселя $\sum\limits_{k=1}^{\infty} c_k^2 \le \|x\|_{L_2}^2$ и

$$\sum_{k=1}^{\infty} \lambda_k^2 \cdot \psi_k^2(t) \le \int\limits_a^b |K(t,\tau)|^2 d\tau \le N^2 \cdot (b-a),$$

где $N = \max|K(t,\tau)|$. При любых натуральных $n$ и $p$ имеем неравенства

$$\left|\sum_{k=n}^{n+p} \lambda_k \cdot c_k \cdot \psi_k(t)\right| \le \sum_{k=n}^{n+p} |c_k| \cdot \left(|\lambda_k| \cdot |\psi_k(t)|\right) \le$$

$$\leq \left(\sum_{k=n}^{n+p} c_k^2\right)^{1/2} \cdot \left(\sum_{k=n}^{n+p} \lambda_k^2 \cdot \psi_k^2(t)\right)^{1/2} \leq \left(\sum_{k=n}^{n+p} c_k^2\right)^{1/2} \cdot N \cdot \sqrt{b-a},$$

из которых в силу сходимости ряда $\sum_{k=1}^{\infty} c_k^2$ следует абсолютная и равномерная сходимость ряда (1).

Из **непрерывности** всех функций $\psi_k(t)$ и равномерной сходимости ряда (1) следует, что его сумма непрерывна. Тогда и функция $h(t) = y(t) - \sum_{k=1}^{\infty} \lambda_k \cdot c_k \cdot \psi_k(t)$ непрерывна как разность непрерывных функций.

$$\|h\|_{L_2}^2 = (h, y) - \sum_{k=1}^{\infty} \lambda_k \cdot c_k \cdot (h, \psi_k),$$

$$(h, \psi_i) = (y, \psi_i) - \sum_{k=1}^{\infty} \lambda_k \cdot c_k \cdot (\psi_k, \psi_i) = \lambda_i c_i - \lambda_i c_i = 0$$

для всех $i$. Поэтому $\|h\|_{L_2}^2 = (h, y) = (h, Ax) = (Ah, x)$. Но равенства $(h, \psi_i) = 0$ для всех $i$ означают, что $h \in Ker\, A$, т.е. $\|h\|_{L_2}^2 = 0$. Таким образом, ряд (1) сходится именно к функции $y(t)$. □

**Замечание**. В условии теоремы Гильберта-Шмидта не предполагаются ни полнота системы функций $\{\psi_k(t)\}$, ни сходимость ряда Фурье по $\{\psi_k(t)\}$ функции $x(t)$.

**Замечание**. Теорема Гильберта-Шмидта справедлива не только для **непрерывного** ядра $K(t,\tau)$, но и в случае, если $\int_a^b \int_a^b |K(t,\tau)|^2 dt\, d\tau < \infty$, а функция $\int_a^b |K(t,\tau)|^2 d\tau$ ограничена на отрезке $a \leq t \leq b$. Эти требования выполнены для ядра $K(t,\tau) = \dfrac{G(t,\tau)}{|t-\tau|^{\alpha}}$, если $G(t,\tau) = \overline{G(\tau,t)}$ – непрерывная функция,

$0 \le \alpha < \dfrac{1}{2}$; для такого ядра функция $y(t) = \int\limits_a^b K(t,\tau)x(\tau)d\tau$ будет непрерывной при любой функции $x \in L_2[a,b]$ и все $\psi_k(t)$ также непрерывны. ▯

Из теоремы Гильберта-Шмидта вытекает важное

**Следствие**. Пусть ядро $K(t,\tau)$ удовлетворяет условиям теоремы Гильберта-Шмидта. Если функция $f(t)$ непрерывна при $a \le t \le b$, а число $\dfrac{1}{\mu}$ не является собственным значением интегрального оператора $A$, то уравнение

$$\varphi(t) - \mu \cdot \int\limits_a^b K(t,\tau)\varphi(\tau)d\tau = f(t), \qquad a \le t \le b,$$

имеет одно и только одно непрерывное решение

$$\varphi(t) = f(t) + \sum_k \frac{\mu \cdot \lambda_k}{1 - \mu \cdot \lambda_k} \cdot d_k \cdot \psi_k(t),$$

где $d_k = \int\limits_a^b f(\tau)\overline{\psi_k(t)}d\tau$, причем определяющий его ряд сходится **абсолютно** и **равномерно** на отрезке $a \le t \le b$.

## Вопрос 13.

## Функция Грина первой краевой задачи для обыкновенного дифференциального уравнения второго порядка. Условия существования решения краевой задачи.

1. Функция Грина – это функция, связанная с интегральным представлением решений краевых задач для дифференциальных уравнений. Функция Грина краевой задачи для линейного дифференциального уравнения является ядром интегрального оператора, обратного к дифференциальному оператору, который порожден данным дифференциальным уравнением и однородными краевыми условиями. Она позволяет найти решения неоднородного уравнения, удовлетворяющие однородным краевым условиям. Нахождение функции Грина сводит изучение свойств дифференциального оператора к исследованию свойств соответствующего интегрального оператора.

Определение и способ построения функции Грина рассмотрим на примере обыкновенного линейного дифференциального уравнения второго порядка в самосопряженной форме. Пусть на отрезке $0 \le x \le l$ заданы **положительная непрерывно дифференцируемая** функция $p(x)$ и **непрерывные** функции $q(x)$ и $f(x)$. Все вводимые функции считаем действительными. Построим для функции $u \in C^2[0,l]$ дифференциальное выражение

$$L[u] \equiv \frac{d}{dx}\left(p(x)\frac{du}{dx}\right) - q(x)u(x), \qquad 0 \le x \le l.$$

Задача нахождения функции $u(x)$, удовлетворяющей уравнению

$$L[u] = f(x), \qquad 0 \le x \le l, \tag{1}$$

и краевым условиям

$$u(0) = u_0, \qquad u(l) = u_l, \tag{2}$$

где $u_0$, $u_l$ – заданные числа, называется **первой** краевой задачей (т.е. с краевыми условиями первого рода).

Краевая задача может не иметь решений, а в случае её разрешимости возможны как единственность так и неединственность решений.

**Пример.** $u'' + u = 0$, $u(0) = 0$, $u(l) = u_l$.

Общее решение уравнения имеет вид $u(x) = C_1 \cos x + C_2 \sin x$. Из краевого условия при $x = 0$ вытекает $C_1 = 0$. Если $l = \pi n$, где $n$ – целое число ($n \neq 0$), а $u_l \neq 0$, то краевая задача не имеет решений. Если же $l = \pi n$, а $u_l = 0$, то решений бесконечно много: $u(x) = C_2 \sin x$ является решением при любой постоянной $C_2$. Наконец, если $l \neq \pi n$, то имеется единственное решение $u(x) = \dfrac{u_l}{\sin l} \sin x$. □

Уравнение второго порядка

$$u'' + a_1(x)u' + a_0(x)u = g(x), \qquad 0 \le x \le l, \tag{3}$$

где $a_0(x)$, $a_1(x)$, $g(x)$ – **непрерывные** при $0 \le x \le l$ функции, можно свести к уравнению вида (1). Для этого умножим (3) на функцию $p(x) = \exp\left\{ \int\limits_0^x a_1(\xi) d\xi \right\} > 0$, тогда и получим уравнение (1), в котором $q(x) = -p(x) \cdot a_0(x)$, $f(x) = p(x) \cdot g(x)$.

Задачу (1), (2) можно свести к аналогичной краевой задаче с **однородными** краевыми условиями. Если $U(x)$ – такая достаточно гладкая функция, что $U(0) = u_0$, $U(l) = u_l$, то для функции $v(x) = u(x) - U(x)$ получим

$$L[v] = h(x), \quad v(0) = 0, \quad v(l) = 0, \quad h = f - L[U].$$

Например, можно выбрать $U(x) = \dfrac{u_l - u_0}{l} x + u_0$. Далее будем рассматривать только краевую задачу

$$L[u] = f(x) \text{ при } 0 \le x \le l, \quad p \in C^1[0,l], \tag{4}$$

$$p(x) > 0 \text{ при } 0 \le x \le l, \quad q \in C[0,l], \quad f \in C[0,l],$$

$$u(0) = 0, \quad u(l) = 0. \tag{5}$$

Решением задачи (4), (5) называется функция $u(x)$, имеющая непрерывную вторую производную на отрезке $0 \le x \le l$, удовлетворяющая уравнению (4) при $0 \le x \le l$ и граничным условиям (5). $u \in C^2[0,l]$.

При определении функции Грина задачи (4), (5) необходимо различать следующие два случая (определения функции Грина в этих случаях различны).

**Случай I:** однородная краевая задача

$$L[u]=0, \qquad 0 \le x \le l, \tag{6}$$

$$u(0)=0, \qquad u(l)=0 \tag{7}$$

имеет только тривиальное решение $u(x) \equiv 0$. Это означает, что $\lambda = 0$ не является собственным значением задачи

$$-L[u]=\lambda u, \qquad 0 \le x \le l, \tag{8}$$

$$u(0)=0, \qquad u(l)=0. \tag{9}$$

Если решение задачи (4), (5) существует, то оно единственно.

**Пример**. $u''=f(x)$, $0 \le x \le l$, $u(0)=0$, $u(l)=0$. □

**Случай II:** однородная задача (6), (7) имеет нетривиальное решение. Число $\lambda = 0$ является собственным значением задачи (8), (9). Решение задачи (4), (5) не может быть единственным.

**Пример**. $u''+u=f(x)$, $0 \le x \le \pi$, $u(0)=0$, $u(\pi)=0$. □

**Замечание**. В случае II задача (8), (9) имеет (с точностью до произвольного ненулевого постоянного множителя) ровно одну собственную функцию. Действительно, предположим, что имеются две линейно независимые собственные функции $\varphi_0(x)$ и $\psi_0(x)$. Обе они являются решениями линейного однородного уравнения (8) с непрерывными на отрезке $0 \le x \le l$ коэффициентами. Поэтому их определитель Вронского не может обратиться в нуль ни в одной точке отрезка $0 \le x \le l$. Это противоречит краевым условиям (9), из которых следует обращение в нуль определителя Вронского при $x=0$ и $x=l$. □

**Пример** двух линейно независимых собственных функций, отвечающих собственному значению $\lambda = 0$. Функции $\varphi_0(x)=\sin x$ и $\psi_0(x)=\sin 2x$ принадлежат $C^2[0,\pi]$ и удовлетворяют задаче $u'' - 3ctg\,x \cdot u' + \left(1 + 3ctg^2 x\right) \cdot u = 0$, $0 < x < \pi$, $u(0) = 0$, $u(l) = 0$. Коэффициенты уравнения имеют **особенности** на концах отрезка, поэтому уравнение задаём лишь на открытом интервале. □

2. Пусть $u(x)$ и $v(x)$ удовлетворяют уравнениям $L[u] = f(x)$ и $L[v] = g(x)$. Тогда справедливо тождество Лагранжа:

$$\frac{d}{dx}\left[p(x)\left(v\frac{du}{dx} - u\frac{dv}{dx}\right)\right] = f(x)v(x) - g(x)u(x).$$

Его интегральная форма называется формулой Грина:

$$\int_0^l \left(vL[u] - uL[v]\right)dx = \left[p(x)\left(v\frac{du}{dx} - u\frac{dv}{dx}\right)\right]_{x=0}^{x=l} =$$

$$= \int_0^l \left(f(x)v(x) - g(x)u(x)\right)dx. \qquad (10)$$

Из тождества Лагранжа вытекает, что если $u_1(x)$ и $u_2(x)$ – линейно независимые решения однородного уравнения $L[u] = 0$, то для определителя Вронского $W[u_1, u_2]$ этих решений выполнено тождество $p(x) \cdot W[u_1, u_2](x) \equiv const$, $0 \le x \le l$.

3. **Функция Грина в случае I.**

**Определение**. Функцией Грина задачи (4), (5) называется определённая при $0 \le x \le l$, $0 \le \xi \le l$ функция $G(x,\xi)$, которая удовлетворяет следующим условиям.

$1^0$. $G(x,\xi)$ непрерывна при $0 \le x \le l$, $0 \le \xi \le l$.

$2^0$. Для каждого фиксированного $\xi$, $0 < \xi < l$, функция $G(x,\xi)$ имеет равномерно непрерывные первую и вторую производные по переменной $x$ в каждом из полусегментов $0 \le x < \xi$ и $\xi < x \le l$, причем производная первого порядка имеет в точке $x = \xi$ разрыв первого рода с величиной скачка

$$\left.\frac{d}{dx}G(x,\xi)\right|_{x=\xi+0}-\left.\frac{d}{dx}G(x,\xi)\right|_{x=\xi-0}=\frac{1}{p(\xi)}.$$

$3^0$. В каждом из полусегментов $0\le x<\xi$ и $\xi<x\le l$ функция $G(x,\xi)$ по переменной $x$ удовлетворяет однородному уравнению $L[G]=0$ и однородным краевым условиям $G(0,\xi)=0$, $G(l,\xi)=0$. ▯

**Теорема 1**. Если однородная краевая задача (6), (7) имеет только тривиальное решение, то функция Грина существует и единственна.

**Доказательство**. Построим два линейно независимых решения однородного уравнения (6), каждое из которых удовлетворяет **только одному** из двух однородных граничных условий (7). Этим требованиям удовлетворяют функции $u_1(x)$, $u_2(x)$, являющиеся решениями задач

$$L[u_1]=0,\quad 0\le x\le l,\qquad\qquad L[u_2]=0,\quad 0\le x\le l,$$
$$u_1(0)=0,\quad u_1'(0)=1,\qquad\qquad u_2(l)=0,\quad u_2'(l)=1.$$

Функция $u_1(x)$ не равна тождественно нулю и не может удовлетворять условию $u(l)=0$, поскольку задача (6), (7) имеет лишь тривиальное решение. Аналогично $u_2(x)$ не равна тождественно нулю и не может удовлетворять условию $u(0)=0$.

Функцию $G(x,\xi)$ можно искать в виде

$$G(x,\xi)=\begin{cases}C_1u_1(x), & 0\le x\le\xi,\\ C_2u_2(x), & \xi\le x\le l.\end{cases}$$

Постоянные $C_1$, $C_2$ однозначно определяются из условий непрерывности $G(x,\xi)$ и скачка её производной при $x=\xi$:

$$\begin{aligned}&C_2u_2(\xi)-C_1u_1(\xi)=0,\\ &C_2u_2'(\xi)-C_1u_1'(\xi)=\frac{1}{p(\xi)}.\end{aligned}\qquad(11)$$

Определитель основной матрицы системы линейных алгебраических уравнений (11) является определителем Вронского функций $u_1$, $u_2$ в точке $\xi$; он отличен от нуля. □

$$G(x,\xi)=\frac{1}{p(\xi)\cdot W[u_1,u_2](\xi)}\begin{cases}u_2(\xi)u_1(x), & 0\le x\le\xi,\\ u_1(\xi)u_2(x), & \xi\le x\le l,\end{cases}$$

где $W[u_1,u_2](\xi)$ – значение определителя Вронского функций $u_1$, $u_2$ в точке $\xi$. Отсюда видно, что $G(x,\xi)=G(\xi,x)$.

**Замечание**. В качестве $u_1$ и $u_2$ можно взять **произвольные** нетривиальные решения уравнения $L[u]=0$, удовлетворяющие соответственно условиям $u_1(0)=0$ и $u_2(l)=0$. При этом произведение $p(\xi)\cdot W[u_1,u_2](\xi)$ **не зависит** от $\xi$. □

**Теорема 2**. Если однородная краевая задача (6), (7) имеет только тривиальное решение, то решение неоднородной задачи (4), (5) существует для любой непрерывной при $0\le x\le l$ функции $f$ и выражается через функцию Грина в виде

$$u(x)=\int_0^l G(x,\xi)f(\xi)d\xi. \qquad (12)$$

**Доказательство** проводится проверкой того, что функция

$$u(x)=\int_0^x G(x,\xi)f(\xi)d\xi+\int_x^l G(x,\xi)f(\xi)d\xi$$

удовлетворяет задаче (4), (5) (найдите для этого $\frac{du}{dx}$ и $\frac{d}{dx}\left(p(x)\frac{du}{dx}\right)$, учтите условие скачка производной функции Грина при $\xi=x$). □

**Замечание**. При каждом $\xi$ функция Грина удовлетворяет по переменной $x$ в смысле обобщённых функций уравнению $L[G(x,\xi)]=\delta(x-\xi)$, в правой части которого содержится $\delta$ – функция. $G(x,\xi)$ есть возмущение, порождаемое точечным источником единичной интенсивности, находящимся в точке $\xi$. Если считать, что $f$ описывает распределение точечных источников, то (12) показывает, что для построения решения задачи (4), (5) надо просуммировать вклады от всех таких источников возмущений. □

**Задача**. $u'' + u = f(x)$, $0 \le x \le \dfrac{\pi}{2}$, $u(0) = 0$, $u\left(\dfrac{\pi}{2}\right) = 0$.

Найдите функцию Грина. □

4. **Функция Грина в случае II**.

Если однородная краевая задача (6), (7) имеет (единственное с точностью до ненулевого постоянного множителя) нетривиальное решение $\varphi_0(x)$, то функция Грина в смысле предыдущего определения не существует: краевым условиям $G(0,\xi) = 0$, $G(l,\xi) = 0$ может удовлетворять только решение $const \cdot \varphi_0(x)$ уравнения (6), которое не имеет разрыва первой производной. Рассматриваемые в теореме 1 функции $u_1$, $u_2$ будут удовлетворять обоим краевым условиям (7), они линейно зависимы; поэтому их определитель Вронского равен нулю, и система линейных алгебраических уравнений (11) неразрешима.

**Теорема 3**. Если задача (4), (5) разрешима, то $f$ и $\varphi_0$ ортогональны: $\int_0^l f(x)\varphi_0(x)dx = 0$.

**Доказательство**. Применим формулу Грина (10) к решению $u(x)$ задачи (4), (5) и $\varphi_0(x)$:

$$\int_0^l \varphi_0(x) f(x)dx = \int_0^l (\varphi_0(x)L[u] - u(x)L[\varphi_0])dx =$$

$$= \left[p(x)\cdot(\varphi_0(x)u'(x) - u(x)\varphi_0'(x))\right]\Big|_{x=0}^{x=l} = 0 . \square$$

Если задача (4), (5) имеет решение $u(x)$, то и любая функция $u(x) + const \cdot \varphi_0(x)$ является её решением. Предполагая разрешимость задачи (4), (5), выделим из множества всех её решений **единственное** решение. Для этого подчиним его дополнительному условию: будем искать функцию $u(x)$, **ортогональную** функции $\varphi_0(x)$, т.е. $\int_0^l u(x)\varphi_0(x)dx = 0$.

**Теорема 4**. Только тривиальное решение однородной задачи (6), (7) ортогонально к $\varphi_0$. Если неоднородная задача (4), (5) разрешима, то ортогональное к $\varphi_0$ её решение единственно. □

В рассматриваемом случае II вместо функции Грина вводят **обобщённую** функцию Грина.

**Определение**. Обобщённой функцией Грина задачи (4), (5) называется определённая при $0 \le x \le l$, $0 \le \xi \le l$ функция $\widetilde{G}(x,\xi)$, которая удовлетворяет следующим условиям.

$1^0$, $2^0$ – те же, что в определении обычной функции Грина.

$3^0$. По переменной $x$ при $x \ne \xi$ функция $\widetilde{G}(x,\xi)$ удовлетворяет неоднородному уравнению $L\left[\widetilde{G}(x,\xi)\right] = -\varphi_0(\xi)\varphi_0(x)$, где $\varphi_0$ – **нормированная** собственная функция $\left(\int_0^l \varphi_0^2(x)dx = 1\right)$, и краевым условиям $\widetilde{G}(0,\xi) = 0$, $\widetilde{G}(l,\xi) = 0$.

$4^0$. По переменной $x$ функция $\widetilde{G}(x,\xi)$ ортогональна собственной функции $\varphi_0(x)$, $\int_0^l \widetilde{G}(x,\xi)\varphi_0(x)dx = 0$. □

**Теорема 5**. Решение $u(x)$ задачи (4), (5), ортогональное собственной функции $\varphi_0(x)$, существует и единственно в том и только в том случае, если правая часть $f(x)$ уравнения (4) ортогональна к $\varphi_0(x)$. Это решение выражается через обобщённую функцию Грина в виде

$$u(x)=\int_0^l \widetilde{G}(x,\xi)f(\xi)d\xi . \square$$

Для построения $\widetilde{G}(x,\xi)$ наряду с **нормированной** собственной функцией $\varphi_0(x)$ выберем решение $v(x)$ уравнения $L[v]=0$, линейно независимое с $\varphi_0(x)$ и удовлетворяющее условию $p(x)\cdot W[v,\varphi_0](x)\equiv 1$. Кроме того, найдём какое-либо частное решение $w(x)$ уравнения $L[w]=-\varphi_0(\xi)\varphi_0(x)$. Тогда обобщённую функцию Грина можно искать в виде

$$\widetilde{G}(x,\xi)=w(x)+\begin{cases} C_1 v(x)+C_3\varphi_0(x), & 0\le x\le \xi,\\ C_2 v(x)+C_4\varphi_0(x), & \xi\le x\le l.\end{cases}$$

Постоянные $C_1$, $C_2$, $C_3$, $C_4$ однозначно определяются из требований, предъявляемых к функции $\widetilde{G}(x,\xi)$.

**Задача**. $u''+u=f(x)$, $0\le x\le \pi$, $u(0)=0$, $u(\pi)=0$. Найдите обобщённую функцию Грина. $\square$

## Вопрос 14.

## Задача Штурма-Лиувилля и свойства её решений. Теорема Стеклова.

1. Пусть на отрезке $0 \le x \le l$ заданы **положительная непрерывно дифференцируемая** функция $p(x)$ и **непрерывная** действительная функция $q(x)$. Построим для функции $u \in C^2[0,l]$ самосопряженное дифференциальное выражение

$$L[u] = \frac{d}{dx}\left(p(x)\frac{du}{dx}\right) - q(x)u(x), \qquad 0 \le x \le l.$$

Обозначим через $\lambda$ комплексный параметр. Введём **положительную непрерывную** на отрезке $0 \le x \le l$ функцию $\rho(x)$ (она играет роль весовой функции в определении скалярного произведения функций $f$ и $g$: $(f,g) = \int\limits_0^l f(x)g(x)\rho(x)dx$).

Фиксируем действительные числа $\alpha$ и $\beta$.

Задача Штурма-Лиувилля состоит в нахождении нетривиальных решений уравнения

$$-L[u] = \lambda\rho(x)u(x), \qquad 0 \le x \le l, \tag{1}$$

удовлетворяющих краевым условиям

$$\begin{aligned} &u(0)\cos\alpha - p(0)u'(0)\sin\alpha = 0, \\ &u(l)\cos\beta + p(l)u'(l)\sin\beta = 0. \end{aligned} \tag{2}$$

$p \in C^1[0,l]$, $p(x) > 0$ при $0 \le x \le l$, $q \in C[0,l]$, $\rho \in C[0,l]$, $\rho(x) > 0$ при $0 \le x \le l$.

Значения параметра $\lambda$, при которых существует нетривиальное решение задачи (1), (2), называются собственными значениями задачи Штурма-Лиувилля, а сами нетривиальные решения $u(x,\lambda)$ – её собственными функциями. Собственная функция определена с точностью до ненулевого постоянного множителя.

**Задачи**. Найдите все собственные значения и отвечающие им собственные функции для уравнения $-u'' = \lambda u$, $0 \le x \le l$, при следующих краевых условиях: а) $u(0)=0$, $u(l)=0$; б) $u'(0)=0$, $u(l)=0$; в) $u(0)=0$, $u'(l)=0$; г) $u'(0)=0$, $u'(l)=0$; д) $u(0)=0$, $u'(l)+H\cdot u(l)=0$, $H=const>0$; е) $u'(0)-h\cdot u(0)=0$, $h=const>0$, $u'(l)=0$. □

**Замечание**. Если первоначально задача на собственные значения была поставлена на отрезке $a\le x\le b$, то заменой независимой переменной $z=\dfrac{x-a}{b-a}\cdot l$ её легко преобразовать в задачу на отрезке $0\le z\le l$. Часто выбирают $l=\pi$. □

**Замечание**. Если первоначально задача на собственные значения поставлена на отрезке $a\le x\le b$, на котором функция $p(x)$ имеет непрерывную первую производную, а функция $p(x)\cdot\rho(x)$ – непрерывную вторую производную, то с помощью подстановок $z=\dfrac{1}{c}\int\limits_a^x\left(\dfrac{\rho(\xi)}{p(\xi)}\right)^{\frac{1}{2}}d\xi$, $c=\dfrac{1}{\pi}\int\limits_a^b\left(\dfrac{\rho(\xi)}{p(\xi)}\right)^{\frac{1}{2}}d\xi$, $U=[\rho(x)p(x)]^{\frac{1}{4}}\cdot u(x)$, $\mu=c\cdot\lambda$ уравнение $-\dfrac{d}{dx}\left(p(x)\dfrac{du}{dx}\right)+q(x)u(x)=\lambda\rho(x)u(x)$, $a\le x\le b$, приводится к виду $-\dfrac{d^2U}{dz^2}+Q(z)U(z)=\mu U(z)$, $0\le z\le\pi$, где $Q(z)=\dfrac{\theta''(z)}{\theta(z)}-c^2\cdot\dfrac{q(x(z))}{\rho(x(z))}$, $\theta(z)=[\rho(x(z))p(x(z))]^{\frac{1}{4}}$. Это преобразование, называемое преобразованием Лиувилля, позволяет при изучении свойств задачи (1), (2) считать, что $p(x)\equiv1$ и $\rho(x)\equiv1$. Однако во многих прикладных задачах важно сохранить первоначальную постановку (1), (2). □

**Свойства решений** задачи Штурма-Лиувилля сформулируем для случая

$$-u''(x)+q(x)u(x)=\lambda u(x),\quad 0\le x\le\pi, \tag{3}$$

$$\begin{aligned} &u(0)\cos\alpha - u'(0)\sin\alpha = 0, \\ &u(l)\cos\beta + u'(l)\sin\beta = 0. \end{aligned} \tag{4}$$

$1^0$. Все собственные значения задачи (3), (4) действительны и образуют бесконечную неограниченно возрастающую последовательность: $\lambda_0 < \lambda_1 < \lambda_2 < \ldots$.

$2^0$. Каждому собственному значению $\lambda_n$ отвечает единственная (с точностью до произвольного ненулевого постоянного множителя) собственная функция $u(x, \lambda_n)$. (Все собственные функции можно выбрать действительными.)

$3^0$. Собственные функции $u(x, \lambda_n)$, $u(x, \lambda_m)$, отвечающие различным собственным значениям, ортогональны: $\int\limits_0^\pi u(x, \lambda_n) u(x, \lambda_m) dx = 0$ при $n \neq m$.

$4^0$. При всех $n = 0, 1, 2, \ldots$ собственная функция $u(x, \lambda_n)$ имеет в интервале $0 < x < \pi$ в точности $n$ нулей.

$5^0$. Если комплексное число $\lambda$ отлично от всех собственных значений, то при $0 \le x \le \pi$, $0 \le \xi \le \pi$ определена такая непрерывная функция $G(x, \xi; \lambda) = \overline{G(\xi, x; \bar{\lambda})}$, что для любой непрерывной на отрезке $0 \le x \le \pi$ функции $f(x)$ уравнение

$$v'' + [\lambda - q(x)]v(x) = f(x), \qquad 0 \le x \le \pi, \tag{5}$$

имеет единственное решение, удовлетворяющее граничным условиям (4); это решение имеет вид

$$v(x) = \int\limits_0^\pi G(x, \xi; \lambda) f(\xi) d\xi .$$

При действительных $\lambda$ функция $G(x, \xi; \lambda)$ действительна.

$6^0$. Если для некоторого $n$ $\lambda = \lambda_n$, а функция $f(x)$ непрерывна на отрезке $0 \le x \le \pi$, то задача (5), (4) имеет решение в том и только в том случае, если $\int\limits_0^\pi u(x, \lambda_n) f(x) dx = 0$. При этом если $v(x)$ является решением задачи (5), (4), то и функция $v(x)+const\cdot u(x,\lambda_n)$ будет её решением; все решения задачи (5), (4) имеют такой вид.

$7^0$. Пусть $\{\varphi_n\}$ – ортонормированная последовательность собственных функций задачи (3), (4): $\varphi_n(x)=u(x,\lambda_n)$, $\int\limits_0^\pi \varphi_n^2(x)dx=1$. Система $\{\varphi_n\}$ является замкнутой в пространстве $L_2[0,\pi]$, т.е. для любой действительной функции $f\in L_2[0,\pi]$ имеет место равенство Парсеваля $\int\limits_0^\pi f^2(x)dx=\sum\limits_{n=0}^{\infty} f_n^2$, где $f_n=(f,\varphi_n)=\int\limits_0^\pi f(x)\varphi_n(x)dx$; ряд $\sum\limits_{n=0}^{\infty} f_n\cdot\varphi_n(x)$ сходится на $0\le x\le\pi$ к функции $f(x)$ в среднем квадратическом (т.е. по норме $\|\cdot\|_{L_2}$).

$8^0$. **(Теорема Стеклова)** Если функция $f\in C^2[0,\pi]$ и удовлетворяет граничным условиям (4), то ряд $\sum\limits_{n=0}^{\infty} f_n\cdot\varphi_n(x)$ сходится на $0\le x\le\pi$ к функции $f(x)$ **абсолютно** и **равномерно**. Ряд $\sum\limits_{n=0}^{\infty} f_n\cdot\varphi_n'(x)$ сходится к функции $f'(x)$ на $0\le x\le\pi$ **абсолютно** и **равномерно**.

Совершенно аналогичными свойствами обладает и задача (1), (2); нужно только учитывать весовую функцию $\rho(x)$ при вычислении всех скалярных произведений и в записи решений через функцию Грина.

2. Отметим важный случай задачи (1), (2). Будем считать для простоты $\rho(x)\equiv 1$.

Пусть непрерывная функция $q(x) \geq 0$ при $0 \leq x \leq l$, а граничные условия записаны в виде

$$h_1 u(0) - h_2 u'(0) = 0, \qquad H_1 u(l) + H_2 u'(l) = 0, \tag{6}$$

где $h_1 \geq 0$, $h_2 \geq 0$, $h_1 + h_2 > 0$, $H_1 \geq 0$, $H_2 \geq 0$, $H_1 + H_2 > 0$.

Если $p(x) > 0$, $q(x) \geq 0$, то отличное от постоянного решение уравнения $L[u] = 0$ не может достигать положительного максимального значения или отрицательного минимального значения во внутренних точках отрезка $0 \leq x \leq l$ (докажите!); отсюда следует, что уравнение $-L[u] = f(x)$ не может иметь двух различных решений, удовлетворяющих граничным условиям (6). Число $\lambda = 0$ может быть собственным значением задачи (1), (6) только если $q(x) \equiv 0$ и $h_1 = H_1 = 0$.

Если $\lambda = 0$ не является собственным значением, то существует функция Грина $G(x, \xi; 0)$, и для любой непрерывной функции $f(x)$ уравнение $-L[u] = \lambda u(x) + f(x)$ имеет единственное удовлетворяющее (6) решение

$$u(x) = \lambda \int_0^l G(x, \xi; 0) \cdot u(\xi) d\xi + \int_0^l G(x, \xi; 0) f(\xi) d\xi .$$

При $f(x) \equiv 0$ это означает, что задача Штурма-Лиувилля (1), (6) эквивалентна задаче на собственные значения для однородного интегрального уравнения с невырожденным симметричным ядром:

$$u(x) = \lambda \int_0^l G(x, \xi; 0) \cdot u(\xi) d\xi \tag{7}$$

(в предположении, что $\lambda = 0$ не является собственным значением (1), (6)). От последнего предположения можно избавиться, если заметить, что $\mu = 0$ не является собственным значением задачи $-L[u] + u(x) = \mu \rho(x) u(x)$ с условиями (6), т.е. заменить $q(x)$ на $q(x) + 1$ и положить $\mu = \lambda + 1$.

Основные свойства задачи Штурма-Лиувилля вытекают из (7). В частности, теорема Стеклова следует из теоремы Гильберта-Шмидта.

3. **Задача**. Докажите свойство $3^0$ задачи (3), (4), пользуясь формулой Грина. Пользуясь $3^0$, докажите, что все собственные значения задачи (3), (4) действительны. ▯

Свойства $1^0$ и $4^0$ задачи (3), (4) составляют содержание теоремы об осцилляции Штурма. Она вытекает из следующей фундаментальной теоремы.

**Теорема Штурма**. Пусть даны два уравнения $u'' + \alpha(x)u = 0$, $v'' + \beta(x)v = 0$, $a \le x \le b$. Если $\alpha(x) < \beta(x)$ на всём отрезке $a \le x \le b$, то между каждыми двумя нулями любого нетривиального решения первого уравнения заключён по крайней мере один нуль каждого решения второго уравнения. (Докажите!) .▯

## Вопрос 15.

## Зависимость решений дифференциальных уравнений от параметров и начальных данных.

1. Многие математические модели реальных процессов формулируются в виде задач для дифференциальных уравнений. Определяющие эти модели величины, которые входят в уравнения, в начальные или краевые условия, могут быть известны лишь в пределах конечной точности. Эти величины часто берутся из эксперимента или вычисляются приближенно. Математические модели содержат параметры, характеризующие изучаемую физическую систему (величину массы, заряда, и др.). Возникает имеющий принципиальное значение вопрос о характере зависимости решения поставленной задачи от этих параметров; в первую очередь – зависит ли решение от параметров непрерывно: если даже незначительные изменения параметра могут привести к сильному изменению решения, то это решение нельзя использовать для моделирования изучаемого реального процесса.

В ряде случаев параметры вводят в модель искусственно и распоряжаются их значениями произвольно. Это может быть вызвано невозможностью точно сформулировать описание явления на ранней стадии его изучения или сложностью построения модели в замкнутой форме. Часто параметры модели служат для «подгонки» результатов расчётов к экспериментальным данным в процессе их интерпретации. Параметры вводят и при конструировании вычислительных алгоритмов для решения поставленной задачи, а также для обоснования их применимости. Важным классом приближённых методов, основанных на введении параметров, являются асимптотические методы, при помощи которых строятся формулы, описывающие качественное поведение решения.

Мы рассмотрим некоторые вопросы, связанные с изучением зависимости решения **задачи Коши для обыкновенного дифференциального уравнения** от параметра $\mu$ и начальных данных:

$$\frac{du}{dt} = f(t,u,\mu), \quad u(t_0) = u_0, \quad u, u_0 \in \mathrm{R}^n, \quad \mu \in \mathrm{R}^m. \quad (1)$$

**Пример**. Если в задаче Коши $\frac{du}{dt} = f(t,u)$, $u(t_0) = u_0$ изменять $t_0$ и $u_0$, то как будет изменяться её решение? □

**Пример**. Если в задаче Коши $\frac{du}{dt} = f(t,u,\mu)$, $u(t_0) = u_0$ функция $f$ непрерывно зависит от параметра $\mu \in R^1$, то будет ли её решение зависеть от $\mu$ непрерывно? Можно ли решение дифференцировать по параметру? Можно ли в окрестности значения $\mu_0$ представить решение формулой Тейлора по степеням $(\mu - \mu_0)$? Можно ли разложить решение в ряд Тейлора по степеням $(\mu - \mu_0)$? □

**Пример**. Как ведёт себя зависящее от параметра $\mu \in R^1$ решение задачи Коши $\frac{du}{dt} = f(t,u,\mu)$, $u(t_0) = u_0$, если функция $f$ имеет особенность при $\mu = \mu_0$? Такой вопрос возникает при исследовании задачи $\frac{du}{dt} = \frac{1}{\mu} \cdot f(t,u)$, $u(t_0) = u_0$, если $\mu$ изменяется в окрестности точки $\mu_0 = 0$; в этом случае уравнение $\mu \cdot \frac{du}{dt} = f(t,u)$ имеет **малый параметр** при производной.

Задачу Коши для уравнения $n$-го порядка

$$\mu \cdot U^{(n)}(t) = F\left(t, U, U', \ldots, U^{(n-1)}\right),\quad U \in R^1,\quad \mu \in R^1,$$

с **малым параметром при старшей производной** можно переформулировать в виде задачи Коши для системы уравнений 1-го порядка в нормальной форме:

$$\mu \cdot \frac{du_1}{dt} = f_1(t, u_1, \ldots, u_n), \qquad u_1 \equiv U^{(n-1)},$$

$$\frac{du_2}{dt} = u_1, \qquad u_2 \equiv U^{(n-2)},$$

$$\cdots$$

$$\frac{du_n}{dt} = u_{n-1}. \qquad\qquad u_n \equiv U.$$

Если в этом уравнении, или в этой системе уравнений, положить $\mu = 0$ и вместо исходной задачи решить полученную **вырожденную** (укороченную) задачу, то как будут связаны решения исходной и вырожденной задач? □

Исследование зависимости решения задачи Коши **от начальных данных** можно свести к изучению зависимости решения **от параметров в правой части уравнения**. Обозначим через $u(t;t_0,u_0)$ решение задачи Коши $\dfrac{du}{dt} = f(t,u)$, $\quad u(t_0) = u_0$. Введём новую функцию $v = u(t;t_0,u_0) - u_0$ и новую независимую переменную $\tau = t - t_0$; тогда $v = v(\tau) = u(\tau + t_0;t_0,u_0) - u_0$. Получаем задачу $\dfrac{dv}{d\tau} = f(\tau + t_0, v + u_0)$, $\quad v(0) = 0$, в которой $\mu = (t_0, u_0)$ – параметры.

Обратно: зависимость от параметра $\mu \in \mathrm{R}^m$ можно рассматривать как частный вид зависимости от начальных данных. Для этого задачу (1) следует заменить задачей Коши $\dfrac{du}{dt} = f(t,u,\mu)$, $\quad \dfrac{d\mu}{dt} = 0$, $\quad u(t_0) = u_0$, $\quad \mu(t_0) = \mu_0$ для вектор-функции $(u,\mu) \in \mathrm{R}^{n+m}$, в которой уже нет внешнего параметра; здесь в качестве $t_0$, $u_0$, $\mu_0$ можно взять любые допустимые значения. Тогда можно изучать зависимость решения $(u,\mu)$ от начальных данных $t_0$, $u_0$, $\mu_0$.

**Замечание**. Решения задачи (1) будем изучать на **конечном отрезке** изменения независимой переменной $t$. Не будем исследовать решения в бесконечном интервале изменения $t$, даже если такие решения существуют. □

2.1. **Теорема 1** (о непрерывной зависимости решения задачи Коши от параметров и начальных данных).

Пусть $D_{t,u}$ – область пространства $\mathrm{R}^{1+n}$, в которой изменяются независимые переменные $t$ и $u$, а $D_{\mu} = \left\{ \mu \in \mathrm{R}^m \middle| \; \|\mu - \mu_0\|_{\mathrm{R}^m} < c = const \right\}$.

Пусть функция $f(t,u,\mu)$ определена в области $D = D_{t,u} \times D_{\mu}$, непрерывна в $D$ по совокупности переменных $t$, $u$, $\mu$ и удовлетворяет условию Липшица по переменной $u$ равномерно в $D$ (т.е. постоянная Липшица не зависит от $t$ и $\mu$).

Пусть $w(t)$ – решение уравнения

$$\frac{du}{dt} = f(t,u,\mu) \tag{2}$$

на отрезке $a \le t \le b$ при $\mu = \mu_0$.

Можно найти такое число $\delta > 0$, что для всех точек $(t_0, u_0, \mu)$ области $G(\delta, a, b) \subset \mathrm{R}^{1+n+m}$, которая определяется неравенствами $a < t_0 < b$, $\|u_0 - w(t_0)\|_{R^n} + \|\mu - \mu_0\|_{R^m} < \delta$, **существует единственное** решение $u(t)$ уравнения (2) на отрезке $a \le t \le b$, удовлетворяющее начальному условию $u(t_0) = u_0$. $u = u(t; t_0, u_0, \mu)$. Функция $u$ **непрерывна по совокупности переменных** $t$, $t_0$, $u_0$, $\mu$ при $a < t < b$ и $(t_0, u_0, \mu) \in G(\delta, a, b)$. □

**Замечание**. Для справедливости теоремы 1 условие Липшица не является необходимым; в предположениях теоремы важна единственность решения: достаточно потребовать, чтобы функция $f$ была непрерывна и ограничена в $D$, а при $\mu = \mu_0$ задача (1) имела **единственное** решение. В этих условиях при стремлении начальных данных $\widetilde{t}_0$, $\widetilde{u}_0$ к значениям $t_0$, $u_0$, а параметра $\mu$ к значению $\mu_0$, решение $u(t; \widetilde{t}_0, \widetilde{u}_0, \mu)$ будет стремиться к решению $u(t; t_0, u_0, \mu)$ **равномерно** на отрезке $a \le t \le b$. □

**Замечание**. Для **непродолжаемых** решений задачи (1) в аналогичных предположениях имеет место непрерывная зависимость от параметров и начальных данных: непродолжаемое решение $u = u(t; t_0, u_0, \mu)$ определено на некотором **открытом** множестве пространства переменных $t$, $t_0$, $u_0$, $\mu$ и непрерывно на нём. □

**Теорема 2**. Пусть выполнены условия теоремы 1 и, кроме того, в области $D$ функция $f(t, u, \mu)$ имеет непрерывные частные производные по переменным $u_1, \ldots, u_n, \mu_1, \ldots, \mu_m$. Тогда определённое в теореме 1 решение $u = u(t; t_0, u_0, \mu)$ непрерывно дифференцируемо при $a < t < b$ и $(t_0, u_0, \mu)$ из области $G(\delta, a, b)$. □

**Замечание**. Если функция $f$ имеет непрерывные производные по $u_1, \ldots, u_n, \mu_1, \ldots, \mu_m$ вплоть до порядка $p \geq 1$, то и решение задачи (1) имеет непрерывные по совокупности переменных $t$, $\mu$ производные по параметрам до $p$ -го порядка. □

**Замечание**. Из существования непрерывных частных производных функции $f$ по $u_1, \ldots, u_n, \mu_1, \ldots, \mu_m$ следует существование смешанных частных производных $\dfrac{\partial^2 u_i(t)}{\partial t \partial t_0}$, $\dfrac{\partial^2 u_i(t)}{\partial t \partial u_{0j}}$, $\dfrac{\partial^2 u_i(t)}{\partial t \partial \mu_k}$, $i, j = 1, \ldots, n$, $k = 1, \ldots, m$, которые непрерывны и не зависят от порядка дифференцирования. □

**Замечание**. Скалярная функция $\Phi(t, u)$, непрерывная вместе со всеми своими частными производными первого порядка и отличная от постоянной, называется **первым интегралом** системы

$$\frac{du}{dt} = f(t, u), \qquad u \in \mathrm{R}^n, \tag{3}$$

если при подстановке в $\Phi$ произвольного решения $u = u(t)$ системы (3) мы получаем постоянную относительно $t$ величину, т.е. функция $\Phi(t, u(t))$ зависит только от выбора решения $u(t)$, но не от переменной $t$. Знание первого интеграла позволяет понизить порядок системы (3) на единицу, а отыскание $n$ функционально независимых первых интегралов равносильно отысканию общего решения в неявном виде. Вопросы о существовании первых интегралов и о числе независимых первых интегралов сводятся к изучению системы (нелинейных) уравнений $u = u(t; t_0, u_0)$ относительно $u_0$. Существование у функции $u$ **непрерывных производных по начальным данным** позволяет для этой цели ввести соответствующий якобиан. □

2.2. В некоторых случаях нужно найти производные решения $u(t;\mu)$ задачи (1) по параметрам $\mu_1, \ldots, \mu_m$ при **фиксированном** значении $\mu^*$. Оказывается, что для этого нет необходимости искать решение $u(t;\mu)$ при переменном $\mu$, а затем дифференцировать его по $\mu_1, \ldots, \mu_m$. Вместо этого можно решить некоторую систему линейных дифференциальных уравнений. Для нахождения производных решения $u(t;u_0)$ по начальным значениям $u_{01}, \ldots, u_{0n}$ при **фиксированном** значении $u_0^*$ можно поступить аналогичным образом. Рассмотрим для простоты случай одного параметра $(m=1)$, но будем считать, что $u_0 = u_0(\mu)$.

**Теорема 3**. Пусть в задаче

$$\frac{du}{dt} = f(t,u,\mu), \;\; u(t_0) = u_0(\mu), \;\; u, u_0 \in \mathrm{R}^n, \;\; \mu \in \mathrm{R}^1, \tag{4}$$

функции $f$ и $u_0$ непрерывны и имеют непрерывные производные по $u_1, \ldots, u_n, \mu$. Тогда решение задачи (4) имеет непрерывную производную по параметру $\mu$.

Производные
$$\left.\frac{\partial u_i}{\partial \mu}\right|_{\mu=\mu^*} \equiv v_i(t), \qquad i = 1, \ldots, n,$$

удовлетворяют линейной системе уравнений

$$\frac{dv_i}{dt} = \sum_{j=1}^{n} \frac{\partial f_i}{\partial u_j} \cdot v_j + \frac{\partial f_i}{\partial \mu}, \qquad i = 1, \ldots, n. \tag{5}$$

и начальным условиям $v_i(t_0)=\frac{du_{0i}}{d\mu}$, $i=1,\ldots,n$, причём производные $\frac{\partial f_i}{\partial u_j}$ и $\frac{\partial f_i}{\partial \mu}$ в (5) берутся **в силу решения задачи** (4) **при** $\mu=\mu^*$ (т.е. вдоль её интегральной траектории $u=u(t;\mu^*)$).□

Система (5) называется **системой уравнений в вариациях** (по параметрам) для задачи (4) при $\mu=\mu^*$. Построение системы (5) основано на идее **линеаризации** (4) в окрестности выбранного решения $u(t;\mu^*)$.

В частном случае, когда $f=f(t,u)$, $u_{0k}(\mu)=\mu$, $u_{0i}(\mu)=const$ при $i\neq k$, мы получим из теоремы 3, что для задачи $\frac{du}{dt}=f(t,u)$, $u(t_0)=u_0$ производные $\left.\frac{\partial u_i}{\partial u_{0k}}\right|_{u=u_0^*}\equiv v_i(t)$ по начальному значению $u_{0k}$ существуют и удовлетворяют системе уравнений

$$\frac{dv_i}{dt}=\sum_{j=1}^{n}\frac{\partial f_i}{\partial u_j}\cdot v_j\,,\qquad i=1,\ldots,n\,, \tag{6}$$

и начальным условиям $v_i(t_0)=0$ при $i\neq k$, $v_k(t_0)=1$, причём производные $\frac{\partial f_i}{\partial u_j}$ в (6) берутся **в силу решения исходной задачи при** $u_0=u_0^*$. Задачу (6) можно получить формальным дифференцированием исходной задачи по параметру $\mu$.

**Задача.** $\frac{du}{dt}=u+\mu\cdot(t+u^2)$, $u(0)=1$. Найдите $\left.\frac{\partial u}{\partial \mu}\right|_{\mu=0}$.□

**Задача.** $\frac{du}{dt}=u+u^2+tu^3$, $u(2)=u_0$. Найдите $\left.\frac{\partial u}{\partial u_0}\right|_{u_0=0}$.□

**Задача.** $\dfrac{du}{dt}=\sin(t\cdot u)$, $u(t_0)=u_0$. Найдите $\left.\dfrac{\partial u}{\partial t_0}\right|_{\substack{t_0=0\\u_0=0}}$, $\left.\dfrac{\partial u}{\partial u_0}\right|_{\substack{t_0=0\\u_0=0}}$. ▯

2.3. **Теорема 4**. Пусть в области $D$ переменных $t$, $u$, $\mu$ ($\mu\in\mathbb{R}^1$) функция $f(t,u,\mu)$ имеет непрерывные и равномерно ограниченные частные производные по $u_1,\dots,u_n,\mu$ вплоть до порядка $p+1$. Тогда существует отрезок $t_0\le t\le T$, на котором для решения задачи (1) **в окрестности значения параметра** $\mu_0=0$ справедливо асимптотическое представление

$$u(t;\mu)=u(t;0)+\mu\cdot\frac{\partial u(t;0)}{\partial\mu}+\dots+\frac{\mu^p}{p!}\cdot\frac{\partial^p u(t;0)}{\partial\mu^p}+R_{p+1}(t;\mu),$$

где остаточный член $R_{p+1}(t;\mu)$ при $\mu\to 0$ стремится к нулю **равномерно** на отрезке $t_0\le t\le T$, причём $R_{p+1}(t;\mu)=O\left(\mu^{p+1}\right)$ при $\mu\to 0$. ▯

Для нахождения коэффициентов разложения $\dfrac{\partial^k u(t;0)}{\partial\mu^k}$ нет необходимости искать решение $u(t;\mu)$ при переменном $\mu$, а затем вычислять его производные по $\mu$ при $\mu=0$ (именно этого и позволяет избежать теорема 4!). Перепишем указанное в теореме тейлоровское разложение в виде $u(t;\mu)=v_0(t)+\mu\cdot v_1(t)+\dots+\dfrac{\mu^p}{p!}\cdot v_p(t)+R_{p+1}$ и подставим его в (1), разложив функцию $f(t,u,\mu)$ по степеням переменной $\mu$. Приравняем коэффициенты при одинаковых степенях $\mu$. Тогда для определения $v_0,\dots,v_p$ получим систему дифференциальных уравнений, из которой **последовательно** найдём все нужные нам коэффициенты разложения.

**Задача**. $\dfrac{du}{dt} = 4\mu \cdot t - u^2$, $u(1) = 1$. Найдите 2 члена разложения решения $u(t;\mu)$ в окрестности значения $\mu_0 = 0$. □

**Задача**. $\dfrac{du}{dt} = 6 \cdot \dfrac{\mu}{t} - u^2$, $u(1) = 1 + 3\mu$. Найдите 3 члена разложения решения $u(t;\mu)$ по степеням параметра $\mu$. □

**Замечание**. В теореме 4 ничего не говорится о поведении остаточного члена при $p \to \infty$ в случае, если $f$ бесконечно дифференцируема. □

**Теорема Пуанкаре** (об аналитической зависимости решения задачи Коши от параметра). Пусть функция $f(t,u,\mu)$ в области изменения переменных $t$, $u$ и в окрестности значения параметра $\mu_0$ ($\mu \in \mathsf{R}^1$) непрерывна по $t$ и аналитически зависит от $\mu$ и $u$. Тогда решение задачи (1) аналитически зависит от параметра $\mu$ в окрестности значения $\mu_0$. □

Тейлоровские разложения искомого решения по параметру важны с практической точки зрения и приводят к многочисленным теоретическим результатам. Эти разложения называются **асимптотиками**. Асимптотики нужны уже по той причине, что эффективное представление решений через элементарные функции возможно лишь для отдельных, специальных видов уравнений. Асимптотические методы, в частности – метод малого параметра, являются одним из способов построения приближённых решений и качественного изучения их свойств. Обоснованием асимптотик занимается **теория возмущений**. Будем считать, что нас интересует поведение решения задачи (1) в окрестности значения параметра $\mu_0 = 0$ ($\mu \in \mathsf{R}^1$). Задачу $\dfrac{du}{dt} = f(t,u,0)$, $u(t_0) = u_0$ считаем невозмущенной, а задачу $\dfrac{du}{dt} = f(t,u,\mu)$, $u(t_0) = u_0$ – возмущенной. Теория возмущений изучает асимптотики для $u(t;\mu)$ при $\mu \to 0$. Выше рассматривался случай **регулярных**

**возмущений**: $f(t,u,\mu)$ обладала достаточной гладкостью по $u$ и $\mu$, удовлетворяла условиям теоремы существования и единственности решения задачи (1), причём при $\mu \to 0$ эти условия не нарушались. Но если при $\mu = 0$ функция $f$ не является непрерывной по $\mu$ или условия существования и единственности решения задачи (1) не выполнены, то построить указанную в теореме 4 асимптотику невозможно; такой случай называется **сингулярным возмущением**. Свойства сингулярно возмущённых задач коренным образом отличаются от регулярно возмущённых.

3. **Теорема Тихонова** (о малом параметре при старшей производной).

**Пример.** $\mu \cdot \dfrac{du}{dt} = \alpha \cdot u + \beta$, $\qquad u(0) = u_0$, $u \in \mathrm{R}^1$. $\mu \in \mathrm{R}^1$, $\alpha$, $\beta$, $u_0$ – постоянные, $\alpha \neq 0$. Точное решение задачи Коши: $u(t;\mu) = \left(u_0 + \dfrac{\beta}{\alpha}\right) e^{\frac{\alpha t}{\mu}} - \dfrac{\beta}{\alpha}$.

Полагая $\mu = 0$, получим вырожденное (не дифференциальное) уравнение $\alpha \cdot u + \beta = 0$. Его решение $\bar{u} \equiv -\dfrac{\beta}{\alpha}$, вообще говоря, не удовлетворяет условию $u(0) = u_0$ (черта над $u$ указывает на выбор значения $\mu = 0$). Если $\alpha < 0$, то $u(t;\mu) \underset{\mu \to +0}{\to} \bar{u}$ на **открытом** множестве $t > 0$; если $\alpha > 0$, то $u(t;\mu) \underset{\mu \to -0}{\to} \bar{u}$ на $t > 0$. При $u_0 \neq -\dfrac{\beta}{\alpha}$ в обоих случаях предел не является равномерным по $t$. Если $\mu \to 0$ **произвольным** образом, то $u(t;\mu)$ не имеет предела ни при $\alpha > 0$, ни при $\alpha < 0$.
▯

**Пример.** $\mu \cdot U''(t) + U'(t) = 0$, $U(0) = U_0$, $U'(0) = U_1$, $U \in \mathrm{R}^1$, $\mu \in \mathrm{R}^1$. Дифференциальное уравнение имеет общее решение $\quad U(t;\mu) = C_1(\mu) e^{-\frac{t}{\mu}} + C_2(\mu)$, где $C_1$, $C_2$ –

произвольные функции параметра $\mu$. При $t \neq 0$ это решение не является непрерывным по $\mu$ в точке $\mu = 0$, за исключением случая, когда $C_1(\mu) \equiv 0$, а $C_2(\mu)$ непрерывна. Предполагая, что $U_0$ и $U_1$ не зависят от $\mu$, получим единственное решение задачи Коши $U(t;\mu) = \mu U_1\left(1 - e^{-\frac{t}{\mu}}\right) + U_0$, которое не является непрерывной функцией параметра $\mu$ при $\mu = 0$, за исключением точки $t = 0$ или случая $U_1 = 0$. Тем не менее при $t \geq 0$ существует **односторонний** предел: $U(t;\mu) \underset{\mu \to +0}{\longrightarrow} U_0$.

Полагая $\mu = 0$, получим вырожденное дифференциальное уравнение $U'(t) = 0$ **меньшего порядка**, чем исходное. Его решение $\overline{U} \equiv U_0$ удовлетворяет первому из начальных условий (а начальное условие, задающее значение производной, при выборе $\mu = 0$, вообще говоря, теряется). $U(t;\mu) \underset{\mu \to +0}{\longrightarrow} \overline{U}$ при $t \geq 0$ ($U(t;\mu) \underset{\mu \to -0}{\longrightarrow} \overline{U}$ при $t \leq 0$), но если $t$ и $\mu$ имеют противоположные знаки, а $U_1 \neq 0$, то при $\mu \to 0$ предел не существует. ▯

Задачи с малым параметром при старшей производной удобно записывать в виде системы уравнений в нормальной форме. Будем далее рассматривать системы **двух** уравнений и введём обозначения $z = u_1$, $y = u_2$:

$$
\begin{aligned}
&\mu \cdot \frac{dz}{dt} = g(t,z,y), && z = z(t;\mu), \\
&\frac{dy}{dt} = f(t,z,y), && y = y(t;\mu), \\
&z(0) = z_0, \quad y(0) = y_0, && \mu > 0.
\end{aligned}
\tag{7}
$$

Полагая $\mu = 0$, получим вырожденную систему

$$0 = g(t, z, y),$$
$$\frac{dy}{dt} = f(t, z, y);$$

при этом принципиальное отличие от случая регулярного возмущения состоит в том, что вырожденная система имеет более низкий порядок, чем исходная, и её решение, вообще, не может удовлетворить обоим начальным условиям.

Первое из уравнений вырожденной системы в силу нелинейности определяет в некоторой области $\Omega$ изменения переменных $t$, $y$, вообще говоря, несколько решений (корней) $z = \varphi_k(t, y)$. Будем предполагать, что они изолированы друг от друга. Всякий такой корень можно представить себе в виде поверхности в пространстве $Otyz$. На каждой такой поверхности функция $g$ обращается в нуль, а между ними сохраняет знак. Выделим корень $z = \varphi_k(t, y)$, обладающий тем свойством, что при переходе через него возрастающей переменной $z$ функция $g$ меняет знак с "+" на "-": $\left.\frac{\partial g}{\partial z}\right|_{z=\varphi_k(t,y)} < 0$. Такой корень называют **устойчивым**. Область

$$\left\{(t, y, z) \middle| \ (t, y) \in \Omega, \varphi_{k-1}(t, y) < z < \varphi_{k+1}(t, y)\right\}$$

между двумя соседними с $\varphi_k$ корнями называют **областью влияния** корня $\varphi_k$. Пусть начальная точка $P(t = 0, y = y_0, z = z_0)$ принадлежит области влияния корня $\varphi_k$. Построим поле $\vec{T}$ векторов, касательных к интегральным кривым системы (7): $\vec{T} = \{dt, dy, dz\}$. Каждый вектор $\vec{T}$ пропорционален вектору $\left\{1, f, \frac{1}{\mu} \cdot g\right\}$. При малых $\mu > 0$ $z$ – компонента этого векторного поля значительно превосходит остальные компоненты, и вне малых окрестностей корней вектор $\vec{T}$ почти параллелен оси $Oz$. Поэтому кривая, начинающаяся в точке $P$, идёт почти параллельно $Oz$, пока не попадёт в малую окрестность корня $\varphi_k$.

Попав в эту окрестность, интегральная кривая уже не выйдет из неё, коль скоро её проекция на $Oty$ остаётся в $\Omega$. В этой малой окрестности корня $\varphi_k$ задача сводится к задаче с **регулярным** возмущением (регулярное возмущение содержится и в начальных данных!). Из теории регулярных возмущений следует, что $\lim_{\mu \to 0} y(t;\mu) = \bar{y}(t)$ на отрезке $0 \le t \le T$, где $\bar{y}(t)$ – решение задачи $\frac{d\bar{y}}{dt} = f(t, \varphi_k(t, \bar{y}), \bar{y})$, $\bar{y}(0) = y_0$ (от $T$ требуется только, чтобы на $0 \le t \le T$ точка $(t, \bar{y})$ не покидала $\Omega$). Что же до $z$ – компоненты решения, то $\lim_{\mu \to 0} z(t;\mu) = \bar{z}(t) \equiv \varphi_k(t, \bar{y}(t))$ в **полуоткрытом** промежутке $0 < t \le T$, поскольку, вообще говоря, $z_0 \ne \varphi_k(0, y_0)$. Первый предельный переход выполнен **равномерно** на отрезке $0 \le t \le T$, а второй равномерным не является. В окрестности $t = 0$ имеется область, в которой $z$ – компонента решения задачи (7) сильно отличается от $z$ – компоненты решения вырожденной задачи, т.е. от $\varphi_k(t, \bar{y}(t))$. Эта область называется **пограничным слоем**.

Описанные два предельных перехода составляют утверждение **теоремы Тихонова**. Они справедливы при следующих предположениях.

$1^0$. $g$ и $f$ непрерывны вместе с частными производными по $z$ и $y$.

$2^0$. Корни $z = \varphi_k(t, y)$ изолированы, а функции $\varphi_k(t, y)$ непрерывны вместе с их производными по $y$.

$3^0$. Выбранный корень является устойчивым: $\frac{\partial g}{\partial z}(t, \varphi_k(t, \bar{y}(t)), \bar{y}(t)) < 0$.

$4^0$. Решение $\bar{y}(t)$ вырожденной задачи определено на отрезке $0 \le t \le T$; $(t, \bar{y})$ принадлежит ограниченной области $\Omega$.

$5^0$. Начальная точка $(0, y_0, z_0)$ принадлежит области влияния выбранного корня.

**Замечание**. В задаче (7) предполагалось, что $\mu > 0$. При $\mu < 0$ та же теорема справедлива при $t \le 0$. Если же $\mu \to 0$ произвольным образом, то решение не будет стремиться к какому-либо пределу. □

**Замечание**. Понятие устойчивого корня связано с асимптотической устойчивостью по Ляпунову. □

**Замечание**. Пограничными слоями в течениях жидкостей с малой вязкостью являются узкие области вблизи определённых частей границы, где скорость течения быстро меняется от нуля на границе до значений скорости практически невязкого течения. Математически такие течения описываются дифференциальными уравнениями, в которых производные наивысшего порядка умножаются на коэффициент вязкости. □

**Задача**. $\mu \cdot \dfrac{dz}{dt} = g(z)$, $z(0) = z_0$. Пусть $z = \varphi \equiv const$ является устойчивым корнем уравнения $g(z) = 0$. Изобразите на плоскости $Otz$ поведение интегральных кривых. □

**Задача**. $\mu \cdot U''(t) + \gamma \cdot U'(t) = F(t, U)$, $U(0) = U_0$, $U'(0) = U_1$, $\gamma \ne 0$, $U \in \mathrm{R}^1$, $\mu > 0$; функции $F$ и $\dfrac{\partial F}{\partial U}$ непрерывны по совокупности переменных $t$, $U$, причём выполнено условие согласования $U_1 = \dfrac{1}{\gamma} \cdot F(0, U_0)$.

Переформулируйте поставленную задачу в виде задачи Коши для системы двух дифференциальных уравнений в нормальной форме на отрезке $0 \le t \le T$. Постройте вырожденную систему. Проверьте, выполняются ли условия теоремы Тихонова (рассмотрите случаи $\gamma > 0$ и $\gamma < 0$). Что можно сказать о вырожденной системе, если условие согласования не выполнено? □

Теорема Тихонова даёт для решения задачи (7) асимптотическое представление

$$y(t;\mu) = \bar{y}(t) + \varepsilon_1(t;\mu), \quad z(t;\mu) = \bar{z}(t) + \varepsilon_2(t;\mu).$$

Изучение остаточных членов $\varepsilon_1$, $\varepsilon_2$ и уточнение асимптотики производится в **теории пограничного слоя**.

## Вопрос 16.

## Постановка вариационных задач. Необходимые условия экстремума.

1. Функционалом называется отображение $J$ некоторого множества $V$ во множество действительных или комплексных чисел. В вариационном исчислении изучают действительнозначные функционалы, определённые на множествах функций, кривых, поверхностей.

**Пример**. График непрерывно дифференцируемой на отрезке $x_0 \le x \le x_1$ функции $y = y(x)$ является плоской спрямляемой кривой. Длина её дуги $J = \int\limits_{x_0}^{x_1} \sqrt{1 + (y'(x))^2}\,dx$. Если рассматривать всевозможные такие кривые, то $J = J[y]$ – функционал; его аргумент – функция $y(x)$. Среди всех указанных кривых можно выделить, например, кривые, соединяющие фиксированные точки $A(x_0, y_0)$, $B(x_1, y_1)$, и поставить вопрос о выборе среди них кривой наименьшей длины. Можно не требовать, чтобы соединяющая $A$ и $B$ кривая была графиком некоторой функции, но была бы задана параметрическими уравнениями $x = \alpha(t)$, $y = \beta(t)$, $t_0 \le t \le t_1$; длина дуги такой кривой равна $\int\limits_{t_0}^{t_1} \sqrt{(\alpha'(t))^2 + (\beta'(t))^2}\,dt$ и определяется выбором кривой. Точки $A$ и $B$ можно не фиксировать, а разрешить им двигаться по некоторым заданным кривым. Во всех подобных случаях можно интересоваться кривой наименьшей длины при некоторых специальных ограничениях. ▯

**Пример**. Если в области $\Omega$ на плоскости $Oxy$ определена непрерывно дифференцируемая функция $z = z(x, y)$, то задана поверхность, площадь которой $J = \iint\limits_{\Omega} \sqrt{1 + (z'_x)^2 + (z'_y)^2}\,dx\,dy$.

**Пример** нелокального функционала:

$$J[y]=\frac{\int\limits_{x_0}^{x_1} x\cdot\sqrt{1+(y'(x))^2}\,dx}{\int\limits_{x_0}^{x_1}\sqrt{1+(y'(x))^2}\,dx}$$ – абсцисса центра тяжести материальной кривой $y=y(x)$, $x_0\le x\le x_1$.

Одна из основных задач вариационного исчисления – задача Лагранжа – состоит в нахождении экстремума функционала $J[y]=\int\limits_{x_0}^{x_1}F(x,y,y')dx$, $y\in\mathrm{R}^n$, $x\in\mathrm{R}^1$, при наличии дифференциальных ограничений типа равенств: $\varphi(x,y,y')=0$, $\varphi\in\mathrm{R}^m$, $m<n$, и $s$ граничных условий $(s\le 2n+2)$. Это задача на условный экстремум. Если в вариационной задаче присутствуют только граничные условия, а дополнительные ограничения не накладываются, то говорят о безусловном экстремуме.

Вариационная задача, в которой концы кривой, доставляющей экстремум, фиксированы, называется **задачей с закреплёнными концами**. Например, в задаче

$$\int\limits_{x_0}^{x_1}F(x,y,y')dx \;\rightarrow\; extr,\;\; y\in\mathrm{R}^1,\;\; x\in\mathrm{R}^1, \tag{1}$$

могут быть заданы начальная и конечная точки $y(x_0)=y_0$, $y(x_1)=y_1$, через которые должна проходить искомая кривая $y(x)$. Оказывается, что кривая, доставляющая экстремум в (1), удовлетворяет краевой задаче для некоторого дифференциального уравнения – **уравнения Эйлера** – с указанными краевыми условиями первого рода. Эта краевая задача может иметь единственное или неединственное решение, а может быть неразрешимой.

**Замечание**. Задача Лагранжа может быть сформулирована по-другому: в подлежащий минимизации функционал и в ограничения может входить **управление**.

Далее будем рассматривать только вариационные задачи о безусловном экстремуме (т.е. без ограничений), без управления.

необходимые условия слабого экстремума являются и необходимыми условиями сильного экстремума.

Можно ввести понятие экстремума, выбирая в качестве уклонения функций $\|y - y_*\|_{C^l}$, – в смысле близости $l$-го порядка. Вообще, для введения понятия экстремума требуется сначала договориться о том, какие кривые $y(x)$ считаются допустимыми, и что понимать под окрестностью кривой $y_*(x)$. Для этого удобно рассматривать кривые $y(x)$ как элементы некоторого линейного нормированного пространства.

3. Пусть $\mathsf{V}$ – линейное нормированное пространство. Предположим, что функционал $J$ определён в некоторой окрестности точки $y_* \in \mathsf{V}$.

Функционал $J[y]$ называется дифференцируемым в точке $y_*$ в смысле Фреше, если существует такой **линейный непрерывный** функционал $L$, что для любого элемента $h \in \mathsf{V}$ $J[y_* + h] - J[y_*] = L[h] + \omega[h]$, где $|\omega[h]| = o(\|h\|_{\mathsf{V}})$ при $h \to 0$. Функционал $L$ называется производной Фреше функционала $J$ в точке $y_*$ и обозначается $J'(y_*)$. Главная линейная часть $J'(y_*)[h]$ приращения дифференцируемого функционала $J$, отвечающего приращению аргумента $h$, называется дифференциалом Фреше функционала $J$ в точке $y_*$. Очевидно, что из дифференцируемости по Фреше функционала в точке $y_*$ следует его непрерывность в этой точке.

Если **для всех** $h \in \mathsf{V}$ существует предел $\lim\limits_{\alpha \to 0} \dfrac{J(y_* + \alpha h) - J(y_*)}{\alpha}$, то этот предел называется (первой) вариацией функционала $J$ в точке $y_*$; он обозначается $\delta J(y_*)[h]$. Очевидно, что $\delta J(y_*)[h] = \left(\dfrac{\partial}{\partial \alpha} J(y_* + \alpha h)\right)\Bigg|_{\alpha=0}$. При фиксированном $y_*$ первая вариация является, вобще говоря, нелинейным функционалом от $h$.

**Пример**. $\mathsf{V}=\mathsf{R}^2$, $y=(y_1,y_2)\in\mathsf{R}^2$,

$$J[y]=\begin{cases}\dfrac{y_1^3-3y_1\cdot y_2^2}{y_1^2+y_2^2}, & y\neq 0,\\ 0, & y=0.\end{cases}$$

Функционал $J$ (функция двух переменных) имеет в точке $y_*=0$ нелинейную первую вариацию (докажите!). ▯

Если функционал $J$ имеет в точке $y_*$ первую вариацию вида $\delta J(y_*)[h]=L[h]$, где $L$ – **линейный непрерывный** функционал, то говорят, что $J$ дифференцируем в точке $y_*$ в смысле Гато. При этом функционал $L$ называется производной Гато функционала $J$ в точке $y_*$ и обозначается $J'(y_*)$, а сама первая вариация называется дифференциалом Гато в точке $y_*$ **по направлению** $h$: $\delta J(y_*)[h]=J'(y_*)[h]$.

**Замечание**. Понятие вариации функционала не зависит от нормы в пространстве $\mathsf{V}$, оно определяется только множеством всех элементов $\mathsf{V}$. ▯

**Замечание**. Под вариацией элемента $y_*$ понимают элемент $h$ в выражении $y_*+\alpha h$. Если $y_*=y_*(x)$ – некоторая функция, то выберем какую-нибудь близкую к ней (в смысле близости по норме $\|\cdot\|_{\mathsf{V}}$) допустимую функцию $y(x)$ и включим их в однопараметрическое семейство $y_*(x)+\alpha\cdot(y(x)-y_*(x))$. Разность $y(x)-y_*(x)$ называется вариацией функции $y_*$. Тем самым, в вариационном исчислении применяют вариации аргумента функционала **по направлениям**. Однако, используют и другие виды вариаций аргумента, например, **игольчатые** вариации. ▯

Из дифференцируемости $J$ в точке $y_*$ по Фреше следует дифференцируемость $J$ в точке $y_*$ по Гато и совпадение производных Гато и Фреше в $y_*$. Однако из дифференцируемости

$J$ по Гато не вытекает ни дифференцируемость $J$ по Фреше, ни даже непрерывность $J$.

**Пример**. $\mathsf{V} = \mathsf{R}^2$, $y = (y_1, y_2) \in \mathsf{R}^2$,

$$J[y] = \begin{cases} 1, \ y_2 = y_1^2, \ y \neq 0, \\ 0, \ y_2 \neq y_1^2 \ \text{или} \ y = 0. \end{cases}$$

Функционал $J$ (функция двух переменных) разрывен в точке $y_* = 0$. Но $J$ дифференцируем в точке $y_* = 0$ по Гато и его дифференциал Гато равен нулю (докажите!). ▯

**Задача**. Найдите производную Фреше функционала $J[y] = \int_{x_0}^{x_1} F(x, y, y')dx$, определённого на пространстве $\overset{0}{C^1}[x_0, x_1]$ непрерывно дифференцируемых на отрезке $x_0 \le x \le x_1$ функций $y(x) \in \mathsf{R}^1$, которые обращаются в нуль на концах этого отрезка. ▯

**Теорема**. Пусть точка $y_* \in \mathsf{V}$ является точкой локального экстремума функционала $J$ и в этой точке существует вариация $\delta J(y_*)[h]$. Тогда для любого $h \in \mathsf{V}$ $\delta J(y_*)[h] = 0$.

**Доказательство**. При $h = 0$ утверждение очевидно. Пусть $h \neq 0$. Докажем теорему в случае, когда $y_*$ – точка минимума. Найдётся такое $r > 0$, что для всех $y \in \mathsf{V}$, удовлетворяющих неравенству $\|y - y_*\|_{\mathsf{V}} < r$, выполнено $J[y] \ge J[y_*]$. Пусть $|\alpha| < \dfrac{r}{\|h\|_{\mathsf{V}}}$, тогда $\|(y_* + \alpha h) - y_*\|_{\mathsf{V}} = |\alpha| \cdot \|h\|_{\mathsf{V}} < r$; поэтому $J[y_* + \alpha h] \ge J[y_*]$. Следовательно, при $\alpha = 0$ функция $J[y_* + \alpha h]$ действительной переменной $\alpha$ имеет локальный минимум. В силу того, что $J$ имеет вариацию в точке $y_*$, эта функция дифференцируема при $\alpha = 0$. Отсюда $\delta J(y_*)[h] = 0$. ▯

Точка $y_* \in \mathsf{V}$ называется **стационарной точкой** функционала $J$, если его вариация определена в этой точке и равна нулю **при любом** $h \in \mathsf{V}$. Необходимые условия экстремума функционала – это условия, которым должна удовлетворять кривая (функция, поверхность и т.д.), являющаяся его стационарной точкой.

4. Опишем вывод необходимого условия экстремума в **задаче** (1) **с закреплёнными концами**. Будем исследовать эту задачу на **слабый** экстремум, т.е. в пространстве $C^1[x_0, x_1]$. Предположим, что функция $F(x, y, y')$ непрерывно дифференцируема в некоторой области пространства $\mathsf{R}^3$.

**Замечание**. Рассматриваемый функционал $J$ непрерывен на $C^1[x_0, x_1]$. Но он может не оказаться непрерывным на $C[x_0, x_1]$ (приведите пример!). □

**Теорема**. Чтобы функция $y_*(x)$ доставляла слабый локальный экстремум в задаче (1), необходимо, чтобы она удовлетворяла уравнению Эйлера:

$$F_y(x, y_*(x), y_*'(x)) - \frac{d}{dx} F_{y'}(x, y_*(x), y_*'(x)) = 0\,. \;\square$$

**Вывод уравнения Эйлера** основан на одной из двух лемм, которые называют основными леммами вариационного исчисления.

**Лемма Лагранжа**. Пусть функция $v(x)$ непрерывна на отрезке $x_0 \le x \le x_1$. Предположим, что для любой непрерывно дифференцируемой функции $h(x)$, обращающейся в нуль на концах отрезка $x_0 \le x \le x_1$, выполнено равенство $\int_{x_0}^{x_1} v(x) h(x) dx = 0$. Тогда $v(x) \equiv 0$ на этом отрезке. □

**Лемма Дюбуа-Реймона** . Пусть функция $w(x)$ непрерывна на отрезке $x_0 \le x \le x_1$. Предположим, что для любой непрерывной функции $g(x)$, удовлетворяющей условию $\int_{x_0}^{x_1} g(x) dx = 0$, выполнено равенство $\int_{x_0}^{x_1} w(x) g(x) dx = 0$. Тогда $w(x) \equiv const$ на этом отрезке. □

Первый этап вывода уравнения Эйлера состоит в доказательстве того, что функционал $J$ в задаче (1) обладает первой вариацией (в любой точке $y_*$, лишь бы $F$ была определена)

$$\delta J(y_*)[h] = \int_{x_0}^{x_1} \left( q(x)h(x) + p(x)h'(x) \right) dx ,$$

где $q(x) = F_y(x, y_*(x), y_*'(x))$, $p(x) = F_{y'}(x, y_*(x), y_*'(x))$.

Здесь $h$ – произвольная функция, принадлежащая $\overset{0}{C^1}[x_0, x_1]$, т. е. непрерывно дифференцируемая и обращающаяся в нуль на концах отрезка. Необходимое условие экстремума состоит в том, что $\delta J(y_*)[h] = 0$.

Второй этап состоит в преобразовании выражения для первой вариации (**для указанных** $h$) интегрированием по частям. Это можно сделать двумя способами. Следуя Лагранжу, интегрируют по частям второе слагаемое; при этом надо дополнительно предполагать, что функция $p(x)$ непрерывно дифференцируема. Тогда получим

$$\delta J(y_*)[h] = \int_{x_0}^{x_1} v(x)h(x)dx , \quad v(x) = \left( F_y - \frac{d}{dx} F_{y'} \right)\bigg|_{y_*(x)} . \tag{3}$$

Следуя Дюбуа-Реймону, интегрируют по частям первое слагаемое. В этом случае

$$\delta J(y_*)[h] = \int_{x_0}^{x_1} w(x)h'(x)dx , \quad w(x) = \int_{x}^{x_1} \left( F_y \big|_{y_*(\xi)} \right) d\xi + F_{y'} \big|_{y_*(x)} . \tag{4}$$

Сопоставив (3) с леммой Лагранжа, получаем уравнение Эйлера. Сопоставив (4) с леммой Дюбуа-Реймона, получаем

$$\int_{x}^{x_1} q(\xi)d\xi + p(x) \equiv const . \tag{5}$$

Первое слагаемое в (5) можно продифференцировать по $x$, а следовательно, и $p(x)$ можно продифференцировать по $x$. Это и даёт уравнение Эйлера. □

Интегральные кривые уравнения Эйлера называются **экстремалями** вариационной задачи (1). Экстремали – это подозрительные на экстремум кривые.

**Задача**. Найдите экстремали в задаче $\int_0^{\frac{\pi}{2}} \left[(y')^2 - y^2\right]dx \rightarrow extr, \; y(0)=0, \; y\left(\frac{\pi}{2}\right)=1.$ □

**Задача**. Найдите экстремали в задаче $\int_0^1 \left[(y')^2 + 12xy\right]dx \rightarrow extr, \; y(0)=0, \; y(1)=1.$ □

**Задача**. Какой вид имеет уравнение Эйлера и какие заключения можно сделать об экстремалях в следующих случаях задачи (1)? $F = F(x,y)$. $F = F(y')$. $F = F(x,y')$. $F = F(y,y')$.□

**Замечание**. Уравнение Эйлера для функционала $J$ задачи (1) является уравнением второго порядка. Но может оказаться, что кривая, на которой $J$ достигает экстремума, **не является дважды дифференцируемой**. Пример: $J[y] = \int_{-1}^{1} y^2 \cdot (1-y')^2\,dx$, $y(-1)=0$, $y(1)=1$. Минимальное значение $J$ равно нулю и достигается на функции $y_*(x) = \begin{cases} 0, & x \le 0, \\ x, & x > 0. \end{cases}$ Эта функция удовлетворяет уравнению Эйлера (проверьте!), хотя она не принадлежит $C^1[-1,1]$. Такая экстремаль называется **ломаной экстремалью**. □

В вариационной **задаче с подвижными концами** концам кривой, доставляющей экстремум, разрешено перемещаться вдоль заданных многообразий. В этом случае произвольные постоянные, от которых зависит решение уравнения Эйлера, можно определить из граничных условий и **условий трансверсальности**, которые являются необходимым требованием для обращения в нуль первой вариации функционала. Например, если в задаче

$$\int_{x_0}^{x_1} F(x,y,y')dx \rightarrow extr, \qquad y \in \mathrm{R}^1,$$

левый и правый концы экстремали могут смещаться вдоль заданных линий: $y_0 = \Phi(x_0)$ и $y_1 = \Psi(x_1)$, то

$$\left[F(x,y,y') + (\Phi'(x) - y'(x)) \cdot F_{y'}(x,y,y')\right]\Big|_{x=x_0} = 0,$$

$$\left[F(x,y,y') + (\Psi'(x) - y'(x)) \cdot F_{y'}(x,y,y')\right]\Big|_{x=x_1} = 0.$$

Условия трансверсальности устанавливают зависимости между угловыми коэффициентами $y'$ и $\Phi'$, $y'$ и $\Psi'$ в граничных точках.

**Задача**. Что означают условия трансверсальности для функционала $\int\limits_{x_0}^{x_1} F(x,y) \cdot \sqrt{1 + (y'(x))^2}\, dx$ в задаче с подвижными вдоль кривых $y_0 = \Phi(x_0)$, $y_1 = \Psi(x_1)$ концами? ▯

Если на один из концов экстремали (например, на правый) не наложено никаких ограничений, то $F(x,y,y')\Big|_{x=x_1} = 0$ и $F_{y'}(x,y,y')\Big|_{x=x_1} = 0$.

**Задача**. По какой плоской кривой материальная точка должна скатываться вниз из положения $A(x_0, y_0)$ под действием только силы тяжести, чтобы в кратчайшее время достигнуть вертикальной прямой $x = x_1$ в плоскости $Oxy$? ▯

**Замечание**. Кроме необходимых условий экстремума, вытекающих из равенства нулю **первой вариации** функционала, имеются необходимые условия, которые следуют из требования неотрицательности **второй вариации** функционала в точке его **минимума**. Сформулируем одно из таких условий – условие Лежандра: чтобы кривая $y_*(x)$ доставляла **минимум** функционалу в задаче (2) с закреплёнными концами, необходимо, чтобы во всех точках кривой $y_*(x)$ выполнялось неравенство $F_{y'y'}(x, y_*(x), y_*'(x)) \geq 0$, $x_0 \leq x \leq x_1$. Как и уравнение Эйлера, условие Лежандра является необходимым условием **слабого** экстремума. При нарушении условия Лежандра вторая вариация функционала не сохраняет свой знак, и кривая $y(x)$ не доставляет экстремум. ▯

**Замечание**. Уравнения Эйлера можно получить и в вариационных задачах для функционалов другого вида, например,

$$J[y]=\int_{x_0}^{x_1} F(x,y,y')dx, \qquad y\in \mathrm{R}^n;$$

$$J[y]=\int_{x_0}^{x_1} F(x,y,y',\ldots,y^{(l)})dx, \qquad y\in \mathrm{R}^1;$$

$$J[z]=\iint_{\Omega} F(x,y,z,z'_x,z'_y)dxdy, \qquad z=z(x,y)\in \mathrm{R}^1; \text{ и др.}$$

Кривые и поверхности удобно задавать в параметрической форме. В этом случае значения подлежащего минимизации функционала должны определяться самой кривой (или поверхностью), а не способом её параметризации. Вариационная задача для такого функционала также приводит к уравнению Эйлера, которое надо рассматривать совместно с условиями, характеризующими конкретный выбор параметров.

**Пример**. Экстремали функционала

$J[y]=\int_{x_0}^{x_1} F(x,y_1,y_2,y'_1,y'_2)dx$, $\quad y=(y_1,y_2)\in \mathrm{R}^2$, в задаче с закреплёнными концами удовлетворяют системе уравнений Эйлера $F_{y_i}-\dfrac{d}{dx}F_{y'_i}=0$, $i=1,2$.

Найдите экстремали функционала

$\int_0^{\frac{\pi}{2}}\left[(y'_1)^2+(y'_2)^2+2y_1y_2\right]dx$, $\quad y_1(0)=0$, $\; y_2(0)=0$, $\; y_1\left(\dfrac{\pi}{2}\right)=1$, $y_2\left(\dfrac{\pi}{2}\right)=-1$. ▯

**Пример**. Экстремали функционала $J[y]=\int\limits_{x_0}^{x_1}F(x,y,y',y'')dx$, $y\in \mathrm{R}^1$, в задаче с закреплёнными концами удовлетворяют уравнению Эйлера $F_y-\dfrac{d}{dx}F_{y'}+\dfrac{d^2}{dx^2}F_{y''}=0$.

Найдите экстремали функционала $\int\limits_0^{\frac{\pi}{2}}\left[(y'')^2-y^2+x^2\right]dx$, $y(0)=1$, $y'(0)=0$, $y\left(\dfrac{\pi}{2}\right)=0$, $y'\left(\dfrac{\pi}{2}\right)=-1$.□

**Пример**. Экстремали функционала $J[z]=\iint\limits_{\Omega}F(x,y,z,z'_x,z'_y)dxdy$, $z=z(x,y)\in \mathrm{R}^1$, в задаче с закреплённым краем удовлетворяют уравнению Эйлера $F_z-\dfrac{\partial}{\partial x}F_{z'_x}-\dfrac{\partial}{\partial y}F_{z'_y}=0$.

Найдите уравнение экстремалей функционала $\iint\limits_{\Omega}\left((z'_x)^2+(z'_y)^2-2zf(x,y)\right)dxdy$; на границе области $\Omega$ функция $z$ задана.□

# Вопрос 17.

## Вариационные задачи на условный экстремум. Метод множителей Лагранжа.

1. **Условный экстремум функции конечного числа независимых переменных.**

$$J(y) \equiv J(y_1, \ldots, y_n) \to extr, \quad y = (y_1, \ldots, y_n) \in \mathrm{R}^n, \quad (1)$$

$$\varphi_i(y_1, \ldots, y_n) = 0, \quad i = 1, \ldots, m, \quad m < n. \quad (2)$$

Говорят, что в точке $y^* = (y_1^*, \ldots, y_n^*) \in \mathrm{R}^n$, **удовлетворяющей уравнениям связи** (2), функция $J(y)$ имеет условный минимум, если неравенство $J(y) \geq J(y^*)$ выполняется в $\mathrm{R}^n$ в некоторой окрестности точки $y^*$ для всех точек этой окрестности, **удовлетворяющих** (2). Будем предполагать, что заданные в (1), (2) функции $J$ и $\varphi_i$, $i = 1, \ldots, m$, имеют в окрестности точки $y^*$ непрерывные частные производные по всем аргументам. Пусть, кроме того, в точке $y^*$ ранг матрицы частных производных $\left(\dfrac{\partial \varphi_i}{\partial y_j}\right)_{\substack{i=1,\ldots,m \\ j=1,\ldots,n}}$ равен $m$; для определённости предположим, что определитель $\dfrac{D(\varphi_1, \ldots, \varphi_m)}{D(y_{n-m+1}, \ldots, y_n)} \neq 0$ в точке $y^*$. Тогда в достаточно малой окрестности точки $y^*$ систему уравнений (2) можно разрешить относительно $y_{n-m+1}, \ldots, y_n$: $y_{n-m+1} = \psi_1(y_1, \ldots, y_{n-m}), \ldots, y_n = \psi_m(y_1, \ldots, y_{n-m})$. И тем самым вопрос об условном экстремуме функции $J(y_1, \ldots, y_n)$ в точке $y^*$ сводится к вопросу о безусловном экстремуме сложной функции

$$J(y_1, \ldots, y_{n-m}, \psi_1(y_1, \ldots, y_{n-m}), \ldots, \psi_m(y_1, \ldots, y_{n-m})) \quad (3)$$

от $n-m$ переменных в точке $z^* = (y_1^*, \ldots, y_{n-m}^*) \in \mathrm{R}^{n-m}$.

Для вывода необходимого условия экстремума в задаче (1), (2) не обязательно искать функции $\psi_i$, $i = 1, \ldots, m$, заданные неявно уравнениями (2). Если в точке $y^*$ функция $J(y)$ имеет условный экстремум (то же самое: сложная функция (3) в точке $z^*$ имеет безусловный экстремум), то дифференциал функции (3) в точке $z^*$ равен нулю тождественно относительно дифференциалов **независимых** переменных $dy_1, \ldots, dy_{n-m}$. В силу инвариантности формы первого дифференциала это условие можно записать в виде

$$\sum_{j=1}^{n} \frac{\partial J}{\partial y_j} dy_j = 0, \tag{4}$$

где $dy_1, \ldots, dy_{n-m}$ – дифференциалы независимых переменных, $dy_{n-m+1}, \ldots, dy_n$ – дифференциалы функций $\psi_1, \ldots, \psi_m$ в точке $z^*$, а все производные вычислены в точке $y^*$. Из (4) не вытекают равенства нулю всех производных $\dfrac{\partial J}{\partial y_j}$, поскольку не все дифференциалы $dy_j$ являются дифференциалами независимых переменных. Чтобы исключить дифференциалы зависимых переменных, продифференцируем полным образом уравнения связи (2), понимая под $y_{n-m+1}, \ldots, y_n$ функции $\psi_1, \ldots, \psi_m$:

$$\sum_{j=1}^{n} \frac{\partial \varphi_i}{\partial y_j} dy_j = 0, \quad i = 1, \ldots, m; \tag{5}$$

здесь все производные вычислены в точке $y^*$. Поскольку в точке $y^*$ определитель $\dfrac{D(\varphi_1, \ldots, \varphi_m)}{D(y_{n-m+1}, \ldots, y_n)} \neq 0$, дифференциалы $dy_{n-m+1}, \ldots, dy_n$ в этой точке можно из (5) линейно выразить через $dy_1, \ldots, dy_{n-m}$. Подставив эти выражения в (4), мы приведём (4) к виду

$$A_1 dy_1 + \ldots + A_{n-m} dy_{n-m} = 0. \tag{6}$$

В (6) фигурируют только дифференциалы **независимых** переменных. Поэтому в точке $y^*$

$$A_1 = 0, \quad . \quad . \quad . \quad . \quad , \quad A_{n-m} = 0. \tag{7}$$

Соотношения (7) вместе с уравнениями связи (2) составляют систему $n$ уравнений, которым должна удовлетворять подозрительная на экстремум точка $\left(y_1^*, \ldots, y_n^*\right)$.

Можно устранить неравноправие переменных в проведённых выше рассуждениях. Для этого умножим равенства (5), соответственно, на «неопределённые» пока множители $\lambda_i$, $i = 1, \ldots, m$, и результаты почленно сложим с (4):

$$\sum_{j=1}^{n}\left(\frac{\partial J}{\partial y_j} + \lambda_1 \cdot \frac{\partial \varphi_1}{\partial y_j} + \ldots + \lambda_m \cdot \frac{\partial \varphi_m}{\partial y_j}\right) dy_j = 0; \tag{8}$$

здесь снова $dy_1, \ldots, dy_{n-m}$ – дифференциалы независимых переменных, $dy_{n-m+1}, \ldots, dy_n$ – дифференциалы неявно заданных уравнениями (2) функций, а все производные вычислены в $y^*$. Выберем такие значения $\lambda_i^*$, $i = 1, \ldots, m$, чтобы все коэффициенты при дифференциалах $dy_{n-m+1}, \ldots, dy_n$ **зависимых** переменных обратились в нуль:

$$\frac{\partial J}{\partial y_j} + \lambda_1^* \cdot \frac{\partial \varphi_1}{\partial y_j} + \ldots + \lambda_m^* \cdot \frac{\partial \varphi_m}{\partial y_j} = 0, \quad j = n - m + 1, \ldots, n. \tag{9}$$

Это возможно, поскольку $\dfrac{D(\varphi_1, \ldots, \varphi_m)}{D(y_{n-m+1}, \ldots, y_n)} \neq 0$ в точке $y^*$. При всех $\lambda_i = \lambda_i^*$ (8) принимает вид

$$\sum_{j=1}^{n-m}\left(\frac{\partial J}{\partial y_j} + \lambda_1^* \cdot \frac{\partial \varphi_1}{\partial y_j} + \ldots + \lambda_m^* \cdot \frac{\partial \varphi_m}{\partial y_j}\right) dy_j = 0,$$

где все $dy_1, \ldots, dy_{n-m}$ – дифференциалы **независимых** переменных. Поэтому

$$\frac{\partial J}{\partial y_j} + \lambda_1^* \cdot \frac{\partial \varphi_1}{\partial y_j} + \ldots + \lambda_m^* \cdot \frac{\partial \varphi_m}{\partial y_j} = 0, \quad j = 1, \ldots, n - m. \tag{10}$$

Для определения $n$ неизвестных $y_1^*, \ldots, y_n^*$ и $m$ множителей $\lambda_1^*, \ldots, \lambda_m^*$ имеем $n+m$ уравнений (2), (9), (10). Эти уравнения можно записать единообразно, если ввести **функцию Лагранжа** $\mathrm{L}(y,\lambda) = J(y) + \sum_{i=1}^{m} \lambda_i \cdot \varphi_i(y)$, $\lambda = (\lambda_1, \ldots, \lambda_m)$; тогда (2), (9), (10) означают

$$\left.\frac{\partial \mathrm{L}}{\partial y_j}\right|_{\substack{y=y^* \\ \lambda=\lambda^*}} = 0, \qquad j = 1, \ldots, n, \tag{11}$$

$$\left.\frac{\partial \mathrm{L}}{\partial \lambda_i}\right|_{\substack{y=y^* \\ \lambda=\lambda^*}} \equiv \varphi_i(y^*) = 0, \qquad i = 1, \ldots, m. \tag{12}$$

$y^*$ является стационарной точкой функции $\mathrm{L}(y,\lambda^*)$ по переменной $y$.

**Замечание**. Если оптимизируемая функция $J(y)$ квадратична, а уравнения связи (2) линейны, то система (11), (12) оказывается линейной. В общем случае получаемая с помощью функции Лагранжа система необходимых условий экстремума (11), (12) нелинейна. Помимо вычислительных трудностей при её решении возникает проблема нахождения **всех** стационарных точек $(y^*, \lambda^*)$. □

**Замечание**. Функция Лагранжа применяется и для решения задач на условный экстремум с ограничениями типа **неравенств**. □

**Задачи**.

$1^0$. $J(y) = y_1 \cdot y_2 \to extr$, $\qquad y_1 + y_2 = 1$.

$2^0$. $J(y) = y_1 \cdot y_2 \to extr$, $\qquad y_1^2 + y_2^2 = 1$.

$3^0$. $J(y) = 2y_1 + 3y_2 \to extr$, $\qquad y_1^2 + y_2^2 = 1$.

$4^0$. $J(y) = y_1^2 + y_2^2 \to extr$, $\qquad 2y_1 + 3y_2 = 1$.

$5^0$. Найдите кратчайшее расстояние от точки $(a,b,c)$ до плоскости $Ay_1 + By_2 + Cy_3 + D = 0$.

$6^0$. $J(y) = y_1 + y_2 + y_3 + y_4 \to extr$, $\qquad y_1 y_2 y_3 y_4 = 1$.

$7^0$. Найдите наименьшее и наибольшее значения квадратичной формы $J(y)=\sum_{i=1}^{n}\sum_{k=1}^{n}a_{ik}\,y_i y_k$ (все $a_{ik}=a_{ki}$) при условии $\sum_{i=1}^{n}y_i^2=1$.

**Пример** задачи, в которой метод Лагранжа не приводит к решению: $J(y)=(y_1+1)^2+y_2^2\to\min$, $y_2^2=y_1^3$. ($J(y)$ есть квадрат расстояния от точки $y_1=-1$, $y_2=0$ до точки на кривой $y_2^2=y_1^3$.) ▯

**Замечание**. В некоторых случаях функцию Лагранжа следует строить в виде

$$\mathrm{L}(y,\lambda)=\lambda_0\cdot J(y)+\sum_{i=1}^{m}\lambda_i\cdot\varphi_i(y).$$

$\lambda_0^*$ выбирают равным 0 или 1 в зависимости от рангов матриц

$$P=\begin{pmatrix}\dfrac{\partial\varphi_1}{\partial y_1} & \cdots & \dfrac{\partial\varphi_1}{\partial y_n}\\ \cdot & \cdots & \cdot\\ \dfrac{\partial\varphi_m}{\partial y_1} & \cdots & \dfrac{\partial\varphi_m}{\partial y_n}\end{pmatrix}\quad\text{и}\quad \widetilde{P}=\begin{pmatrix}\dfrac{\partial\varphi_1}{\partial y_1} & \cdots & \dfrac{\partial\varphi_1}{\partial y_n}\\ \cdot & \cdots & \cdot\\ \dfrac{\partial\varphi_m}{\partial y_1} & \cdots & \dfrac{\partial\varphi_m}{\partial y_n}\\ \dfrac{\partial J}{\partial y_1} & \cdots & \dfrac{\partial J}{\partial y_n}\end{pmatrix}$$

в точке $y^*$. Если $rang\,\widetilde{P}(y^*)=rang\,P(y^*)$, то $\lambda_0^*=1$; если же $rang\,\widetilde{P}(y^*)>rang\,P(y^*)$, то необходимо положить $\lambda_0^*=0$. В случае $rang\,\widetilde{P}(y^*)=rang\,P(y^*)=m$ все множители $\lambda_i^*$ определены однозначно. В случаях $rang\,\widetilde{P}(y^*)>rang\,P(y^*)$ или $rang\,\widetilde{P}(y^*)=rang\,P(y^*)<m$ множители $\lambda_i^*$ определяются не единственным образом. ▯

2. **Изопериметрическая задача вариационного исчисления**.

$$J[y] \equiv \int_{x_0}^{x_1} F(x, y, y')dx \to extr\,, \quad y(x_0) = y_0\,, \quad y(x_1) = y_1\,, \quad (13)$$

$$\Phi[y] \equiv \int_{x_0}^{x_1} G(x, y, y')dx = c = const\,. \quad (14)$$

**Замечание**. Название «изопериметрическая» происходит от следующей задачи вариационного исчисления: среди всех замкнутых кривых на плоскости с заданным периметром найти кривую, которая ограничивает наибольшую площадь. В этой задаче оптимизируемым функционалом $J$ является площадь, а сохраняющий постоянное значение функционал $\Phi$ – длина кривой; такую задачу формулируют в параметрической форме, а не в виде (13), (14). ▯

Будем предполагать, что заданные в (13), (14) функционалы $J[y]$ и $\Phi[y]$ определены в пространстве $C^1[x_0, x_1]$ и непрерывно дифференцируемы (по Фреше) в $C^1[x_0, x_1]$ (можно потребовать, например, чтобы функции $F$ и $G$ были дважды непрерывно дифференцируемы при $x_0 \le x \le x_1$ и при любых значениях $y$ и $y'$). $y \in \mathrm{R}^1$, $c \in \mathrm{R}^1$.

**Теорема**. Пусть $y^* \in C^1[x_0, x_1]$, $y^*(x_0) = y_0$, $y^*(x_1) = y_1$, – точка экстремума функционала $J[y]$ при условии $\Phi[y] = c$, причём $y^*$ **не является стационарной точкой функционала** $\Phi$ (т.е. вариация $\delta\Phi(y^*)[h]$ не равна нулю тождественно по $h$). Тогда существует такое число $\lambda^*$, что $y^*$ будет стационарной точкой функционала $\mathrm{L}[y, \lambda] \equiv \int_{x_0}^{x_1} (F + \lambda G)dx$ при $\lambda = \lambda^*$. ▯

Эта теорема означает, что точки условного экстремума в задаче (13), (14) следует искать либо среди экстремалей функционала

которые являются экстремалями функционала $J$ или экстремалями $\Phi$. Например, задача о максимуме площади, ограниченной замкнутой кривой заданной длины, и задача о минимуме длины замкнутой кривой, ограничивающей заданную площадь, имеют общие экстремали. □

3. **Задача Лагранжа** состоит в нахождении экстремума функционала

$$J[y] \equiv \int_{x_0}^{x_1} F(x, y, y')dx \to extr, \qquad y \in \mathrm{R}^n, \tag{15}$$

при $m$ ограничениях типа равенств

$$\varphi_i(x, y, y') = 0, \qquad i = 1, \ldots, m, \quad m < n, \tag{16}$$

и выполнении $s$ граничных условий на концах отрезка $x_0 \le x \le x_1$, $s \le 2n + 2$. Допускаются **неразделённые** граничные условия вида

$$\theta_k(x_0, y(x_0), x_1, y(x_1)) = 0, \qquad k = 1, \ldots, s,$$

причём в некоторых задачах $x_0$ и $x_1$ подлежат определению. Уравнения связей (16) могут содержать, а могут не содержать производную (дифференциальные уравнения или конечные уравнения). Решения задачи (15), (16) будем понимать в смысле слабого экстремума.

**Замечание**. Изопериметрическая задача (13), (14) сводится к задаче Лагранжа при помощи введения новой переменной $z$, удовлетворяющей краевой задаче $z'_x = G(x, y, y')$, $z(x_0) = 0$, $z(x_1) = c$. (Постройте отвечающую задаче (13), (14) задачу Лагранжа.) □

Необходимые условия оптимальности в задаче Лагранжа на условный экстремум оказывается удобным записывать как необходимые условия стационарности для функционала Лагранжа, построенного с помощью множителей Лагранжа. Например, в задаче Лагранжа **с закреплёнными концами** необходимые условия экстремума получаются как необходимые условия безусловного экстремума функционала $\mathrm{L}[y, \lambda] \equiv \int_{x_0}^{x_1} Q(x, y, y', \lambda)dx$,

$y(x_0)=y_0\in\mathrm{R}^n$, $y(x_1)=y_1\in\mathrm{R}^n$. (Граничные условия не должны противоречить уравнениям связей (16)). Отметим, что в задаче с закреплёнными концами $x_0$ и $x_1$ фиксированы, задача решается на не изменяющемся отрезке $x_0\le x\le x_1$.

Если в задаче Лагранжа $s<2n+2$, то один или оба конца экстремали не закреплены, а могут смещаться вдоль заданных граничными условиями (многомерных) поверхностей; при этом $x_0$ и $x_1$ варьируются. В вариационных задачах **с подвижными концами** необходимым условием оптимальности, кроме уравнения Эйлера, является **условие трансверсальности**, с помощью которого (наряду с граничными условиями) определяются произвольные постоянные, входящие в решение уравнения Эйлера. Условие трансверсальности является необходимым условием обращения в нуль первой вариации функционала. Поскольку в задаче с подвижными концами $x_0$ и $x_1$ не заданы, теперь необходимо определить $2n+2$ произвольные постоянные (включая $x_0$ и $x_1$). С помощью условия трансверсальности получают ровно столько соотношений, позволяющих определить эти произвольные постоянные.

Дифференциальные уравнения связей (16) называются **неголономными**, если их невозможно проинтегрировать по $x$, т.е. заменить на конечные связи. Если же уравнения связей являются конечными, т.е. имеют вид

$$\varphi_i(x,y)=0,\qquad i=1,\ldots,m,\qquad m<n, \tag{21}$$

то их называют **голономными**. Если уравнения (21) независимы, т.е. $rang\left(\dfrac{\partial\varphi_i}{\partial y_j}\right)_{\substack{i=1,\ldots,m\\ j=1,\ldots,n}}=m$, то для получения необходимых условий экстремума в задаче (15), (21) можно ввести функционал Лагранжа и искать решения среди его экстремалей. Рассмотрим частный случай $n=2$, $m=1$, т.е. будем искать экстремум функционала

$$J[y]\equiv\int_{x_0}^{x_1}F(x,y_1,y_2,y_1',y_2')dx \tag{22}$$

**Замечание**. Правило множителей Лагранжа состоит в том, что необходимое условие локального экстремума в задаче на условный экстремум при некотором выборе множителей Лагранжа совпадает с необходимым условием безусловного экстремума по $y$ функционала Лагранжа. Это правило требует обоснования. Для «гладких задач» обоснование можно провести при помощи теоремы Люстерника. ▯

## Вопрос 18.

## Классификация уравнений в частных производных второго порядка. Приведение к каноническому виду.

1. Дифференциальным уравнением с частными производными второго порядка называется уравнение, содержащее хотя бы одну производную второго порядка от неизвестной функции $u(x_1,...,x_n)$ и не содержащее производных более высокого порядка.

Производные функции $u$ будем записывать в виде $u_{x_i} = \dfrac{\partial u}{\partial x_i}$,

$u_{x_i x_j} = \dfrac{\partial^2 u}{\partial x_i \partial x_j}$; $x_1,\ldots, x_n$ – действительные переменные.

$$\Phi\left(x_1,...,x_n,u,u_{x_1},...,u_{x_n},u_{x_1x_1},u_{x_1x_2},...,u_{x_nx_n}\right)=0 \qquad (1)$$

– общий вид такого уравнения; здесь предполагается, что заданная действительнозначная функция $\Phi$ имеет производные $\dfrac{\partial \Phi}{\partial u_{x_i x_j}}$, причем $\displaystyle\sum_{i=1}^{n}\sum_{j=1}^{n}\left(\frac{\partial \Phi}{\partial u_{x_i x_j}}\right)^2 \neq 0$. Уравнение (1), вообще говоря, нелинейно относительно функции $u$ и её производных. Если (1) имеет вид

$$\sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}\cdot u_{x_i x_j} + F\left(x_1,...,x_n,u,u_{x_1},...,u_{x_n}\right)=0, \qquad (2)$$

где $a_{ij} = a_{ij}\left(x_1,...,x_n\right)$ – заданные действительнозначные функции, то уравнение называется линейным относительно старших производных. Именно уравнения вида (2) мы и будем классифицировать.

2. Опишем подробно классификацию уравнений (2) в случае двух независимых переменных; будем обозначать их через $x$ и $y$:

$$\begin{aligned}&a_{11}(x,y)\cdot u_{xx} + 2\cdot a_{12}(x,y)\cdot u_{xy} + a_{22}(x,y)\cdot u_{yy}\\ &+F\left(x,y,u,u_x,u_y\right)=0,\end{aligned} \qquad (3)$$

$u = u(x, y)$. Предположим, что коэффициенты при старших производных непрерывны и нигде не обращаются в нуль одновременно. Выполним некоторую невырожденную замену независимых переменных $\begin{cases} \xi = \varphi(x, y) \\ \eta = \psi(x, y) \end{cases}$. Тогда искомая функция примет вид $u = u_1(\xi(x, y), \eta(x, y))$. Выразим частные производные $u_x$, $u_y$, $u_{xx}$, $u_{xy}$, $u_{yy}$ в новых переменных:

$$u_x = u_\xi \xi_x + u_\eta \eta_x,$$

$$u_y = u_\xi \xi_y + u_\eta \eta_y,$$

$$u_{xx} = \left(u_{\xi\xi}\xi_x^2 + 2u_{\xi\eta}\xi_x\eta_x + u_{\eta\eta}\eta_x^2\right) + u_\xi \xi_{xx} + u_\eta \eta_{xx},$$

$$u_{xy} = \left(u_{\xi\xi}\xi_x\xi_y + u_{\xi\eta}\left(\xi_x\eta_y + \xi_y\eta_x\right) + u_{\eta\eta}\eta_x\eta_y\right) + u_\xi \xi_{xy} + u_\eta \eta_{xy},$$

$$u_{yy} = \left(u_{\xi\xi}\xi_y^2 + 2u_{\xi\eta}\xi_y\eta_y + u_{\eta\eta}\eta_{yy}^2\right) + u_\xi \xi_{yy} + u_\eta \eta_{yy}.$$

Найдём зависимость переменных $x$, $y$ от новых независимых переменных $\xi$, $\eta$ и подставим найденные выражения для функции $u$ и её производных в уравнение (3). Тогда (3) будет иметь вид

$$\overline{a}_{11}(\xi, \eta) \cdot u_{\xi\xi} + 2 \cdot \overline{a}_{12}(\xi, \eta) \cdot u_{\xi\eta} + \overline{a}_{22}(\xi, \eta) \cdot u_{\eta\eta} +$$

$$+ \overline{F}(\xi, \eta, u, u_\xi, u_\eta) = 0,$$

где коэффициенты при старших производных

$$\overline{a}_{11} = a_{11}\xi_x^2 + 2a_{12}\xi_x\xi_y + a_{22}\xi_y^2,$$

$$\overline{a}_{12} = a_{11}\xi_x\eta_x + a_{12}\left(\xi_x\eta_y + \xi_y\eta_x\right) + a_{22}\xi_y\eta_y,$$

$$\overline{a}_{22} = a_{11}\eta_x^2 + 2a_{12}\eta_x\eta_y + a_{22}\eta_y^2$$

и функция $\overline{F}$ выражены в переменных $\xi$, $\eta$.

Подберём такую замену независимых переменных, при которой уравнение (3) будет приведено к наиболее простому виду. Под этим подразумевается, что уравнение (3) с **переменными** коэффициентами при старших производных мы постараемся преобразовать в уравнение с **постоянными** коэффициентами при старших производных. При этом нас не будет интересовать, станет ли $\overline{F}$ более сложным выражением, чем $F$. Зафиксируем точку

$(x, y)$ и поставим в соответствие уравнению (3) квадратичную форму

$$Q = a_{11}(x, y) \cdot q_1^2 + 2 \cdot a_{12}(x, y) \cdot q_1 \cdot q_2 + a_{22}(x, y) \cdot q_2^2$$

от независимых переменных $q_1, q_2$. Симметричную матрицу коэффициентов квадратичной формы $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$, $a_{12} = a_{21}$, приведём невырожденным линейным преобразованием $B$ к диагональному виду. Это значит, что после преобразования переменных $\vec{q} = B\vec{p}$, $\vec{q} = \begin{pmatrix} q_1 \\ q_2 \end{pmatrix}$, $\vec{p} = \begin{pmatrix} p_1 \\ p_2 \end{pmatrix}$, квадратичная форма примет канонический вид $Q = \alpha_1 \cdot p_1^2 + \alpha_2 \cdot p_2^2$ в новых независимых переменных $p_1, p_2$. Можно выбрать такое преобразование $B$, что коэффициенты $\alpha_1$ и $\alpha_2$ будут равны по модулю единице или нулю (такой вид квадратичной формы называется нормальным). Закон инерции квадратичных форм утверждает, что число положительных и число отрицательных членов в каноническом виде не зависит от способа приведения. Это позволяет провести классификацию уравнений вида (3) и подобрать требуемую замену переменных $\begin{cases} \xi = \varphi(x, y) \\ \eta = \psi(x, y) \end{cases}$.

Уравнение (3) называется уравнением параболического типа в точке $(x, y)$, если один из коэффициентов $\alpha_1$, $\alpha_2$ в каноническом виде квадратичной формы $Q$ равен нулю. Тогда другой коэффициент можно выбрать по модулю равным единице. Для уравнения параболического типа в точке $(x, y)$ выражение $D(x, y) = a_{12}^2 - a_{11} \cdot a_{22} = 0$.

Уравнение (3) называется уравнением эллиптического типа в точке $(x, y)$, если коэффициенты $\alpha_1$ и $\alpha_2$ в каноническом виде квадратичной формы $Q$ имеют одинаковые знаки. Их можно выбрать одновременно равными либо +1, либо $-1$. Для уравнения эллиптического типа в точке $(x, y)$ $D(x, y) = a_{12}^2 - a_{11} \cdot a_{22} < 0$.

Уравнение (3) называется уравнением гиперболического типа в точке $(x, y)$, если $\alpha_1$ и $\alpha_2$ в каноническом виде квадратичной формы $Q$ имеют разные знаки. Их можно выбрать равными $+1$ и $-1$. Для уравнения гиперболического типа в точке $(x, y)$ $D(x, y) = a_{12}^2 - a_{11} \cdot a_{22} > 0$.

**Теорема.** Пусть новые независимые переменные $\xi$, $\eta$ выбраны так, что $\xi_x = b_{11}$, $\xi_y = b_{21}$, $\eta_x = b_{12}$, $\eta_y = b_{22}$, где $b_{ij}$ – элементы матрицы $B$, с помощью которой квадратичная форма $Q$ была приведена к каноническому виду. Тогда коэффициенты при вторых производных функции $u(x, y)$ в уравнении (3) будут преобразованы при замене $\begin{cases} \xi = \varphi(x, y) \\ \eta = \psi(x, y) \end{cases}$ так же, как коэффициенты квадратичной формы $Q$ при замене $\vec{p} = B^{-1}\vec{q}$.

**Доказательство.** Если в переменных $q_1, q_2$ квадратичная форма $Q$ имеет матрицу коэффициентов $A$, а в переменных $p_1, p_2$ – матрицу коэффициентов $\overline{A}$, причём $\vec{q} = B\vec{p}$, то $\overline{A} = B^T AB$. ▯

3. Совершенно аналогично проводится классификация уравнений вида (2) с произвольным числом $n$ независимых переменных. Зафиксируем точку $(x_1, \ldots, x_n)$ и поставим в соответствие уравнению (2) квадратичную форму

$$Q = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij}(x_1, \ldots, x_n) \cdot q_i \cdot q_j \tag{4}$$

от независимых переменных $q_1, \ldots, q_n$. Симметричную матрицу $A$ коэффициентов квадратичной формы приведем невырожденным линейным преобразованием к диагональному виду. Согласно закону инерции число положительных и число отрицательных членов в каноническом виде квадратичной формы $Q$ не зависят от способа приведения. Это и позволяет провести классификацию уравнений (2). Если все коэффициенты квадратичной формы $Q$ в её каноническом виде отличны от нуля и одного знака, то уравнение (2) называется эллиптическим в точке $(x_1,...,x_n)$. Если один из этих коэффициентов отрицателен, а все остальные положительны (или наоборот), то уравнение (2) называется гиперболическим в точке $(x_1,...,x_n)$. В случае, когда $l$ $(1<l<n-1)$ коэффициентов положительны, а $n-l$ отрицательны, уравнение (2) называется ультрагиперболическим в точке $(x_1,...,x_n)$. Если же хотя бы один из этих коэффициентов (но не все) равен нулю, то уравнение (2) называется параболическим в точке $(x_1,...,x_n)$. Когда в разных частях области изменения независимых переменных уравнение (2) принадлежит к различным типам, говорят, что оно является уравнением смешанного типа.

4. Отметим, что в предположении непрерывности всех производных $\dfrac{\partial \Phi}{\partial u_{x_i x_j}}$ можно провести классификацию уравнений общего вида (1). Для этого надо фиксировать некоторое решение $u^*(x_1,...,x_n)$ этого уравнения и привести к каноническому виду квадратичную форму (4) с коэффициентами $a_{ij}(x_1,...,x_n)=\left.\dfrac{\partial \Phi}{\partial u_{x_i x_j}}\right|_{u=u^*(x_1,...,x_n)}$. Тогда уравнение (1) назовём эллиптическим, гиперболическим и т.д. в точке $(x_1,...,x_n)$ **для выбранного решения** $u^*$.

5. Приведение к каноническому виду квадратичной формы $Q$, отвечающей уравнению (2), позволяет записать уравнение в наиболее простой – канонической – форме в каждой фиксированной точке $(x_1,...,x_n)$. Но оно не даёт способа нахождения замены независимых переменных сразу в некоторой области их изменения. Для уравнений с двумя независимыми переменными (и только для них) такой способ существует. В его основе лежит понятие характеристик уравнения (3).

**Определение.** Пусть кривая $\gamma$ на плоскости $x$, $y$ задана уравнением $\omega(x, y) = 0$, где $\omega$ – непрерывно дифференцируемая функция, причём во всех точках этой кривой $\omega_x^2 + \omega_y^2 \neq 0$. Кривая $\gamma$ называется характеристикой уравнения (3), если функция $\omega(x, y)$ удовлетворяет характеристическому уравнению

$$a_{11}(x, y) \cdot \omega_x^2 + 2 \cdot a_{12}(x, y) \cdot \omega_x \cdot \omega_y + a_{22}(x, y) \cdot \omega_y^2 = 0. \quad (5)$$

В курсе обыкновенных дифференциальных уравнений доказана следующая

**Теорема.** Функция $\omega(x, y)$ является частным решением уравнения (5) тогда и только тогда, когда равенство $\omega(x, y) = const$ представляет собой общий интеграл обыкновенного дифференциального уравнения

$$a_{11} \cdot (dy)^2 - 2 \cdot a_{12} \cdot dx \cdot dy + a_{22} \cdot (dx)^2 = 0. \quad (6)$$

Уравнение (6) называется дифференциальным уравнением характеристик. Знание характеристик уравнения (3) позволяет привести его к канонической форме сразу для всех точек $x$, $y$ в области, в которой (3) имеет один и тот же тип. При этом единая для указанной области замена переменных $\begin{cases} \xi = \varphi(x, y) \\ \eta = \psi(x, y) \end{cases}$ может быть построена в явном виде для уравнений каждого из трёх типов. Предполагая, что $a_{11}(x, y) \neq 0$, найдём из (6)

$$\frac{dy}{dx} = \frac{a_{12} \pm \sqrt{a_{12}^2 - a_{11} \cdot a_{22}}}{a_{11}}. \quad (7)$$

**Гиперболический тип.** $D(x, y) = a_{12}^2 - a_{11} \cdot a_{22} > 0$. Тогда правые части обоих уравнений (7) действительны и различны. Поэтому имеем два независимых интеграла уравнения (6): $\varphi(x, y) = const$, $\psi(x, y) = const$. Полагая $\begin{cases} \xi = \varphi(x, y) \\ \eta = \psi(x, y) \end{cases}$, получим $\overline{a}_{11} = 0$, $\overline{a}_{22} = 0$, $\overline{a}_{12} \neq 0$ (см. выражения $\overline{a}_{ij}$ в переменных $\xi$, $\eta$). Разделив уравнение в переменных $\xi$ и $\eta$ на $\overline{a}_{12}$, получим каноническую форму уравнения гиперболического типа (первую каноническую форму):

$$u_{\xi\eta} = G_1\left(\xi, \eta, u, u_{\xi}, u_{\eta}\right).$$

Если выполнить замену переменных $\begin{cases} \tilde{\xi} = \dfrac{1}{2}(\xi + \eta) \\ \tilde{\eta} = \dfrac{1}{2}(\xi - \eta) \end{cases}$, то получим вторую каноническую форму:

$$u_{\tilde{\xi}\tilde{\xi}} - u_{\tilde{\eta}\tilde{\eta}} = G_2\left(\tilde{\xi}, \tilde{\eta}, u, u_{\tilde{\xi}}, u_{\tilde{\eta}}\right).$$

**Эллиптический тип.** $D(x, y) < 0$. Имеем комплексный интеграл $\varphi(x, y) = const$ уравнения (7) ($\varphi$ – комплекснозначная функция) и комплексно сопряженный с ним. Выберем действительные независимые переменные $\begin{cases} \tilde{\xi} = \operatorname{Re}\varphi(x, y) \\ \tilde{\eta} = \operatorname{Im}\varphi(x, y) \end{cases}$, $\begin{cases} \tilde{\xi} = \dfrac{1}{2}(\xi + \eta) \\ \tilde{\eta} = \dfrac{1}{2}(\xi - \eta) \end{cases}$.

Тогда получим каноническую форму уравнения эллиптического типа:

$$u_{\tilde{\xi}\tilde{\xi}} + u_{\tilde{\eta}\tilde{\eta}} = E\left(\tilde{\xi}, \tilde{\eta}, u, u_{\tilde{\xi}}, u_{\tilde{\eta}}\right).$$

**Параболический тип.** $D(x, y) = 0$. Это значит, что $0 \le a_{12}^2 = a_{11} \cdot a_{22} = |a_{11}| \cdot |a_{22}|$, поэтому $sign\, a_{11} = sign\, a_{22}$. Тогда $\overline{a}_{11} = \left(\sqrt{|a_{11}|} \cdot \xi_x + \sqrt{|a_{22}|} \cdot \xi_y\right)^2 \cdot sign\, a_{11}$. Выбирая $\xi = \varphi(x, y)$, где $\varphi(x, y) = const$ – единственный интеграл уравнения (7), получим $\overline{a}_{11} = 0$. Но тогда автоматически $\overline{a}_{12} = 0$, поскольку $\overline{a}_{12} = \left(\sqrt{|a_{11}|} \cdot \xi_x + \sqrt{|a_{22}|} \cdot \xi_y\right) \cdot \left(\sqrt{|a_{11}|} \cdot \eta_x + \sqrt{|a_{22}|} \cdot \eta_y\right) \cdot sign\, a_{11}$.

Поэтому в качестве $\eta(x, y)$ можно выбрать любую функцию, независимую с $\varphi(x, y)$. В переменных $\xi$, $\eta$ получим каноническую форму уравнения параболического типа:

$$u_{\eta\eta} = P\left(\xi, \eta, u, u_{\xi}, u_{\eta}\right).$$

6. **Задача.** Найдите тип уравнений $u_{tt} = a^2 u_{xx}$, $u_t = a^2 u_{xx}$, $u_{xx} + u_{yy} = 0$. Какие из этих уравнений имеют действительные семейства характеристик? Найдите все такие семейства.

**Задача.** Приведите к канонической форме уравнение $u_{xx} + x \cdot u_{yy} = 0$ в каждой области плоскости $x$, $y$, в которой оно сохраняет тип.

**Решение.** Запишем дифференциальное уравнение характеристик: $(dy)^2 + x \cdot (dx)^2 = 0$. Если $x < 0$, то исходное уравнение гиперболического типа. В этом случае $dy = \pm\sqrt{-x} \cdot dx$, и мы имеем два независимых действительных интеграла уравнения характеристик: $y + \dfrac{2}{3} \cdot (-x)^{3/2} = const$, $y - \dfrac{2}{3} \cdot (-x)^{3/2} = const$.

Выполните замену переменных $\begin{cases} \xi = y + \dfrac{2}{3} \cdot (-x)^{3/2} \\ \eta = y - \dfrac{2}{3} \cdot (-x)^{3/2} \end{cases}$. Для этого

выразите $u_x, u_{xx}, u_y, u_{yy}$ в переменных $\xi$, $\eta$ и подставьте $u_{xx}, u_{yy}$ в исходное уравнение. Вы должны получить $u_{\xi\eta} = \dfrac{u_\xi - u_\eta}{8 \cdot (-x)^{3/2}}$. В силу выполняемой замены переменных $(-x)^{3/2} = \dfrac{3}{4} \cdot (\xi - \eta)$.

Окончательно: $u_{\xi\eta} = \dfrac{u_\xi - u_\eta}{6 \cdot (\xi - \eta)}$ — первая каноническая форма уравнения в области $\xi > \eta$. Можно вместо $\xi$, $\eta$ ввести другие

переменные: $\begin{cases} \tilde{\xi} = y \\ \tilde{\eta} = \dfrac{2}{3} \cdot (-x)^{3/2} \end{cases}$. Снова выразите $u_x, u_{xx}, u_y, u_{yy}$ в

переменных $\tilde{\xi}, \tilde{\eta}$ и из исходного уравнения получите

$u_{\tilde\xi\tilde\xi} - u_{\tilde\eta\tilde\eta} = \dfrac{u_{\tilde\eta}}{3\tilde\eta}$ — его вторую каноническую форму. Зная, что $\begin{cases} \tilde\xi = \dfrac{1}{2}\cdot(\xi+\eta) \\ \tilde\eta = \dfrac{1}{2}\cdot(\xi-\eta) \end{cases}$, проверьте полученные ответы, переходя от первой канонической формы ко второй и наоборот.

Если $x > 0$, то исходное уравнение эллиптического типа. В этом случае получаем два комплексно сопряженных интеграла уравнения характеристик: $y + \dfrac{2}{3} i \cdot x^{3/2} = const, \quad y - \dfrac{2}{3} i \cdot x^{3/2} = const.$ Выберем новые действительные независимые переменные: $\begin{cases} \tilde\xi = y \\ \tilde\eta = \dfrac{2}{3}\cdot x^{3/2} \end{cases}$. Выразите $u_x, u_{xx}, u_y, u_{yy}$ в новых переменных и получите каноническую форму исходного уравнения $u_{\tilde\xi\tilde\xi} + u_{\tilde\eta\tilde\eta} + \dfrac{u_{\tilde\eta}}{3\tilde\eta} = 0$ в области $\tilde\eta > 0$.

Если $x = 0$, то формально исходное уравнение параболического типа. Но имеет ли смысл при $x = 0$ указывать его тип как тип уравнения в частных производных?

Замечание. Рассмотренное уравнение называется уравнением Трикоми. В газовой динамике оно описывает дозвуковое движение в области гиперболичности и сверхзвуковое движение в области эллиптичности. Уравнение Трикоми – уравнение смешанного типа.

**Задача.** $u_{xx} + 2u_{xy} - 2u_{xz} + 2u_{yy} + 6u_{zz} = 0$ – уравнение с постоянными коэффициентами при старших производных, поэтому в каждой точке $(x, y, z)$ оно имеет один и тот же тип. Приводя отвечающую этому уравнению квадратичную форму к каноническому виду, запишите уравнение в канонической форме. Ответ: $u_{\xi\xi} + u_{\eta\eta} + u_{\varsigma\varsigma} = 0$.

**Задача.** Пусть все коэффициенты линейного уравнения

$$a_{11} \cdot u_{xx} + 2 \cdot a_{12} \cdot u_{xy} + a_{22} \cdot u_{yy} + b \cdot u +$$
$$+ c_1 \cdot u_x + c_2 \cdot u_y + f(x, y) = 0$$

постоянны. После приведения к канонической форме выполните замену неизвестной функции: $u(x, y) = e^{\alpha x + \beta y} \cdot v(x, y)$, где $\alpha$ и $\beta$ – параметры. Как надо выбрать $\alpha$ и $\beta$, чтобы в канонической форме уравнения для новой функции $v$ исчезли первые производные (для уравнений гиперболического или эллиптического типа )? Как надо выбрать $\alpha$ и $\beta$, чтобы в канонической форме уравнения исчезла одна из первых производных и сама искомая функция (для уравнений параболического типа)?

## Вопрос 19.

## Первая краевая задача для уравнения колебаний струны. Интеграл энергии и единственность решения первой краевой задачи.

1. Первая смешанная (начально-краевая) задача для уравнения колебаний состоит в определении функции $u(x,t)$, удовлетворяющей в открытой области $Q=\{0<x<l\}\times\{0<t<\infty\}$ уравнению

$$\frac{\partial^2 u(x,t)}{\partial t^2}=a^2\,\frac{\partial^2 u(x,t)}{\partial x^2}+f(x,t), \tag{1}$$

краевым условиям первого рода

$$u(0,t)=\mu_1(t),\quad u(l,t)=\mu_2(t) \tag{2}$$

при всех $0\le t<\infty$ и начальным условиям

$$u(x,0)=\varphi(x),\quad \frac{\partial u(x,0)}{\partial t}=\psi(x) \tag{3}$$

при $0\le x\le l$. Такую задачу коротко называют первой краевой задачей. (Возможны и другие краевые условия, например, линейные краевые условия второго или третьего рода.)

**Определение.** Классическим решением задачи (1), (2), (3) называется функция $u(x,t)$, непрерывная в замкнутой области $\overline{Q}=\{0\le x\le l\}\times\{0\le t<\infty\}$, имеющая непрерывные производные второго порядка в открытой области $Q$, удовлетворяющая в $Q$ уравнению колебаний (1), на сегменте $0\le x\le l$ – начальным условиям (3), а на его концах – краевым условиям (2). $u(x,t)\in C(\overline{Q})\cap C^2(Q)$.

**Замечание.** Для существования классического решения необходимо согласование начальных и краевых условий:

$$\mu_1(0)=\varphi(0),\quad \mu_2(0)=\varphi(l),\quad \frac{\partial\mu_1(0)}{\partial t}=\psi(0),\quad \frac{\partial\mu_2(0)}{\partial t}=\psi(l).$$

Часто встречаются задачи, решения которых не могут удовлетворять требованиям предъявляемым к **классическим** решениям. Если не выполняется согласование начальных и краевых условий или искомая функция $u(x,t)$ не может обладать требуемой гладкостью, то понятие решения задачи надо рассматривать в более широком, **обобщенном** смысле. □

В силу линейности задачи (1), (2), (3) её решение $u(x,t)$ можно представить в виде суммы решений трёх задач: $u(x,t)=u_1(x,t)+u_2(x,t)+u_3(x,t)$, где $u_1(x,t)$ – решение задачи с однородным уравнением, однородными краевыми условиями и неоднородными начальными условиями, $u_2(x,t)$ – решение задачи с неоднородным уравнением и однородными краевыми и начальными условиями, $u_3(x,t)$ – решение задачи с однородным уравнением, однородными начальными условиями и неоднородными краевыми условиями. При этом задачу для $u_3(x,t)$ можно свести к первым двум задачам, например, заменой искомой функции $u_3(x,t)=v(x,t)+\frac{x}{l}\cdot(\mu_2(t)-\mu_1(t))+\mu_1(t)$.

2. **Теорема единственности решения первой краевой задачи.** Задача (1), (2), (3) может иметь только одно классическое решение.

**Доказательство.** Предположим, что существуют два различных решения $u_1(x,t)$ и $u_2(x,t)$ задачи (1), (2), (3). Функция $w(x,t)=u_1(x,t)-u_2(x,t)$ удовлетворяет задаче

$$w_{tt}(x,t)=a^2 w_{xx}(x,t),\quad 0<x<l,\quad 0<t<\infty,$$

$$w(0,t)=0,\quad w(l,t)=0,\quad 0\le t<\infty,$$

$$w(x,0)=0,\quad w_t(x,0)=0,\quad 0\le x\le l.$$

(Через $w_t$, $w_{tt}$, $w_{xx}$ обозначены частные производные функции $w(x,t)$.)

Введём функционал

$$E(t)=\frac{1}{2}\int_0^l\left(w_t^2(x,t)+a^2 w_x^2(x,t)\right)dx.$$

Из определения $E(t)$ и из начальных условий следует, что $E(t) \geq 0$, $E(0) = 0$.

$$\frac{dE}{dt} = \int_0^l \left( w_t(x,t) w_{tt}(x,t) + a^2 w_x(x,t) w_{xt}(x,t) \right) dx.$$

Интеграл $a^2 \int_0^l w_x(x,t) w_{xt}(x,t) dx$ вычислим по частям:

$$a^2 \int_0^l w_x(x,t) w_{xt}(x,t) dx =$$

$$= a^2 \left( w_x(x,t) w_t(x,t) \right)\Big|_{x=0}^{x=l} - a^2 \int_0^l w_{xx}(x,t) w_t(x,t) dx.$$

Тогда получим

$$\frac{dE}{dt} = \int_0^l \left( w_t \left( w_{tt} - a^2 w_{xx} \right) \right) dx + a^2 \left( w_x(x,t) w_t(x,t) \right)\Big|_{x=0}^{x=l}.$$

Так как функция $w(x,t)$ удовлетворяет однородному уравнению колебаний, то

$$\frac{dE}{dt} = a^2 \left( w_x(x,t) w_t(x,t) \right)\Big|_{x=0}^{x=l}.$$

$w(0,t) = w(l,t) = 0$ во все моменты времени $t \geq 0$. Поэтому $w_t(0,t) = w_t(l,t) = 0$. Отсюда $\dfrac{dE}{dt} = 0$, $E(t) = const$. Но $E(0) = 0$, поэтому $E(t) \equiv 0$. Следовательно, $w_t(x,t) = 0$ и $w_x(x,t) = 0$, $0 < x < l$, $t > 0$. Поэтому $w(x,t) = const$. Так как $w(x,0) = 0$, то $w(x,t) \equiv 0$. □

3. Уравнение (1) можно интерпретировать как уравнение малых поперечных колебаний гибкой однородной струны под действием внешней силы. В этом случае величина $u(x,t)$ является поперечным смещением в момент времени $t$ точки струны с координатой $x$. Уравнение (1) справедливо в пределах действия закона Гука и в предположениях, что все возникающие в струне

напряжения направлены по касательным к профилю струны, и что сумма проекций всех сил на ось $Ox$ в каждой точке равна нулю. Для однородной струны её линейная плотность массы $\rho = const$.

Пусть концы струны жестко закреплены, а внешние силы отсутствуют, т.е. уравнение (1) однородно. В момент времени $t$ струна имеет кинетическую энергию $\frac{\rho}{2}\int_0^l (u_t(x,t))^2\,dx$ и потенциальную энергию $\frac{\rho}{2}\int_0^l a^2\cdot(u_x(x,t))^2\,dx$. Сумма кинетической и потенциальной энергий механической системы в некоторый момент времени называется интегралом энергии. Как видно из доказательства теоремы, в рассматриваемом случае интеграл энергии не зависит от времени: полная энергия колеблющейся системы **при отсутствии внешних возмущений** не меняется со временем – закон сохранения энергии. В момент времени $t = 0$ потенциальная энергия струны определяется её профилем $\varphi(x)$, а кинетическая энергия – скоростями $\psi(x)$.

4. Решение задачи

$$u_{tt}(x,t) = a^2 u_{xx}(x,t), \quad 0 < x < l, \quad 0 < t < \infty,$$

$$u(0,t) = 0, \quad u(l,t) = 0, \quad 0 \le t < \infty,$$

$$u(x,0) = \varphi(x), \quad u_t(x,0) = \psi(x), \quad 0 \le x \le l,$$

можно построить в виде ряда

$$u(x,t) = \sum_{n=1}^{\infty} u_n(x,t), \tag{4}$$

где $u_n(x,t) = \left(A_n \cos\left(\frac{\pi n a}{l}t\right) + B_n \sin\left(\frac{\pi n a}{l}t\right)\right)\cdot \sin\left(\frac{\pi n}{l}x\right)$,

а коэффициенты $A_n$ и $B_n$ определяются из начальных условий. (4) представляет движение струны как сумму **независимых** движений

$$u_n(x,t) = \sqrt{A_n^2 + B_n^2}\cdot\cos(\omega_n\cdot t - \delta_n)\cdot\sin\left(\frac{\pi n}{l}x\right),$$

## Вопрос 20.

## Принцип максимума для уравнения теплопроводности. Единственность решения первой краевой задачи и задачи Коши.

1. Уравнением теплопроводности называется уравнение с частными производными

$$\frac{\partial u}{\partial t} = a^2 \Delta_M u \tag{1}$$

относительно функции $u = u(M,t)$, где $\Delta_M u$ – оператор Лапласа по переменным $x$, $y$, $z$: если $x$, $y$, $z$ – координаты точки $M$ в пространстве $R^3$ в **декартовой** системе координат, то $\Delta_M u = \dfrac{\partial^2 u}{\partial x^2} + \dfrac{\partial^2 u}{\partial y^2} + \dfrac{\partial^2 u}{\partial z^2}$. В дальнейшем входящие в уравнение теплопроводности частные производные будем обозначать через $u_t$, $u_{xx}$, $u_{yy}$, $u_{zz}$. (1) является уравнением параболического типа.

Процесс распространения тепла в пространстве можно описать температурой $u(M,t)$, зависящей от точек $M(x,y,z)$ и времени $t$. Если температура зависит от точек $M$, то возникают потоки тепла, направленные от областей с высокой температурой к областям с меньшей температурой. Эти потоки подчинены закону Фурье, из которого в предположении однородности среды и следует уравнение (1). Если в среде действуют источники и стоки тепла, то получим неоднородное уравнение теплопроводности

$$u_t = a^2 \Delta_M u + f(x,y,z,t). \tag{2}$$

Если среда неоднородна, то температура $u(M,t)$ удовлетворяет уравнению

$$C\rho \cdot u_t = div_M \left(k(M) \cdot grad_M u\right) + F(M,t),$$

где $C$ – удельная теплоёмкость, $\rho$ – объёмная плотность массы, $k$ – коэффициент теплопроводности. В случае $C = const$, $\rho = const$, $k = const$ получаем (2), где $a^2 = \dfrac{k}{C\rho}$.

2. Для изучения основных свойств решений уравнения теплопроводности введём область изменения независимых переменных в четырёхмерном пространстве $R^3 \times [0, T]$. Будем рассматривать в пространстве $R^3$ ограниченную область $\Omega$ с границей $S$. Назовём открытым цилиндром в $R^3 \times [0, T]$ область $Q_T$ вида $Q_T = \Omega \times (0, T] = \{(M, t) \mid M \in \Omega, t \in (0, T]\}$. замкнутым цилиндром – область $\overline{Q}_T = \overline{\Omega} \times [0, T] = \{(M, t) \mid M \in \overline{\Omega}, t \in [0, T]\}$, где $\overline{\Omega} = \Omega \cup S$. Если $T = \infty$, то полагаем $Q = Q_T$.

Классическим решением уравнения (1) в $Q_T$ называется функция $u(M, t)$, имеющая в $Q_T$ непрерывные частные производные $u_t$, $u_{xx}$, $u_{yy}$, $u_{zz}$ и удовлетворяющая (1); $M = M(x, y, z)$. $u(M, t) \in C^{2,1}(Q_T)$.

Следующая теорема является основной, из неё вытекают важнейшие следствия, описывающие свойства решений уравнения теплопроводности.

**Теорема (принцип максимума).** Классическое решение уравнения теплопроводности $u_t = a^2 \cdot \Delta u$, $(M, t) \in Q_T$, непрерывное в замкнутом цилиндре $\overline{Q}_T$, внутри этого цилиндра не может принимать значения, большие, чем значения при $t = 0$ или на границе $S$ области $\Omega$.

**Доказательство.**

Введём обозначение $A = \max\left\{ \underset{M \in \overline{\Omega}}{u(M, 0)};\ \underset{P \in S;\, t \in [0, T]}{u(P, t)} \right\}$. Надо доказать, что $u(M, t) \leq A$ для всех точек $(M, t) \in \overline{Q}_T$. Это утверждение будем доказывать от противного. Пусть в точке $(M_0, t_0) \in Q_T$ функция $u(M, t)$ достигает своего максимального значения, большего $A$, т.е. $u(M_0, t_0) = A + \varepsilon$, $\varepsilon > 0$.

Рассмотрим вспомогательную функцию

$v(M,t)=u(M,t)+\alpha(t_0-t)$, $\alpha>0$. Очевидно, $v(M_0,t_0)=u(M_0,t_0)=A+\varepsilon$. Теперь оценим максимальное значение $v(M,t)$ на границе $S$ области $\overline{\Omega}$ или в начальный момент времени: $\max\left\{\underset{M\in\overline{\Omega}}{v(M,0)};\quad \underset{P\in S;\ t\in[0,T]}{v(P,t)}\right\}\le A+\alpha T<A+\dfrac{\varepsilon}{2}$, если $\alpha<\dfrac{\varepsilon}{2T}$. Таким образом, максимальное значение функции $v(M,t)$ на границе цилиндра меньше, чем некоторое её значение внутри. Следовательно, существует точка $(M_1,t_1)$ внутри цилиндра, в которой функция $v(M,t)$ должна достигать своего максимума: $(M_1,t_1)\in Q_T$, $v(M_1,t_1)\ge v(M_0,t_0)=A+\varepsilon$. Так как $(M_1,t_1)$ – точка максимума, то должны выполняться условия для первых производных: $grad_M v(M_1,t_1)=0$, $\left.\dfrac{\partial v}{\partial t}\right|_{\substack{t=t_1\\ M=M_1}}\ge 0$ ($\left.\dfrac{\partial v}{\partial t}\right|_{\substack{t=t_1\\ M=M_1}}=0$, если $t\ne T$, или $\left.\dfrac{\partial v}{\partial t}\right|_{\substack{t=T\\ M=M_1}}\ge 0$) и для вторых производных: $\Delta v(M_1,t_1)\le 0$. Посмотрим, что это даёт для функции $u(M,t)$:

$$u(M,t)=v(M,t)-\alpha(t_0-t),\quad \alpha>0;$$

$$grad_M u(M_1,t_1)=grad_M v(M_1,t_1)=0;$$

$$\left.\frac{\partial u}{\partial t}\right|_{\substack{M=M_1\\ t=t_1}}=\left.\frac{\partial v}{\partial t}\right|_{\substack{M=M_1\\ t=t_1}}+\alpha\ge\alpha>0;$$

$$\Delta u(M_1,t_1)=\Delta v(M_1,t_1)\le 0.$$

Таким образом, в точке $(M_1,t_1)$, лежащей внутри области $Q_T$, $\Delta u\le 0$, $u_t>0$, т.е. не выполняется уравнение теплопроводности. Пришли к противоречию. □

**Замечание.** Принцип максимума является выражением того очевидного факта, что тепло перемещается от мест с большей температурой к местам с меньшей температурой, т.е. «растекается». При отсутствии источников и стоков тепла это и приводит к доказанному только что утверждению.

3. Для однородного уравнения теплопроводности справедлив и принцип минимума:

**Теорема.** Классическое решение уравнения теплопроводности $u_t = a^2 \cdot \Delta u$, $(M,t) \in Q_T$, непрерывное в замкнутом цилиндре $\overline{Q}_T$, внутри этого цилиндра не может принимать значения, меньшие, чем значения при $t = 0$ или на границе $S$ области $\Omega$.

**Доказательство.** Функция $u_1(M,t) = -u(M,t)$ – тоже решение уравнения теплопроводности. Максимальное значение для функции $u(M,t)$ является минимальным для функции $u_1(M,t)$. Следовательно, справедливость этой теоремы следует из предыдущей. ▯

**Следствие.** Из доказанных теорем следует принцип экстремума: значения функции $u(M,t)$ для всех точек $(M,t) \in \overline{Q}_T$ лежат между максимальным и минимальным значениями функции на границе, т.е.

$$\min\left\{\underset{M\in\overline{\Omega}}{u(M,0)};\ \underset{P\in S;\ t\in[0,T]}{u(P,t)}\right\} \le u(M,t) \le \max\left\{\underset{M\in\overline{\Omega}}{u(M,0)};\ \underset{P\in S;\ t\in[0,T]}{u(P,t)}\right\}.$$

**Замечание.** Функция $u(M,t) = const$ является решением уравнения $u_t = a^2 \cdot \Delta u$ и не противоречит принципам максимума и минимума.

**Теорема сравнения 1.** Пусть функции $u_i(M,t)$, $i = 1, 2$, удовлетворяют однородному уравнению теплопроводности $u_t = a^2 \cdot \Delta u$, непрерывны в $\overline{Q}_T$ и удовлетворяют условиям $u_1(M,0) \ge u_2(M,0)$, $M \in \overline{\Omega}$, и $u_1(P,t) \ge u_2(P,t)$, $P \in S$, $t \in [0,T]$. Тогда $u_1(M,t) \ge u_2(M,t)$ во всех точках замкнутого цилиндра $\overline{Q}_T$.

**Доказательство.** Введём функцию $v(M,t)=u_1(M,t)-u_2(M,t)$. Пусть функция $v(M,t)$ не равна тождественно нулю, в противном случае утверждение теоремы очевидно. Так как уравнение $u_t=a^2\cdot\Delta u$ линейно и однородно, то линейная комбинация решений тоже является решением этого уравнения. Следовательно, функция $v(M,t)$ является решением уравнения $v_t=a^2\cdot\Delta v$, $v(M,t)\in C(\overline{Q}_T)$. Т.е. функция $v(M,t)$ – классическое решение уравнения $v_t=a^2\cdot\Delta v$, и выполнены неравенства $v(M,0)\ge 0$, $M\in\overline{\Omega}$, и $v(P,t)\ge 0$, $P\in S$, $t\in[0,T]$. Для функции $v(M,t)$ выполняется принцип минимума. Следовательно, $v(M,t)\ge 0$, $(M,t)\in\overline{Q}_T$, и $u_1(M,t)\ge u_2(M,t)$, $(M,t)\in\overline{Q}_T$. ▯

**Теорема сравнения 2.** Пусть функции $u_i(M,t)$, $i=1,2$, удовлетворяют однородному уравнению теплопроводности $u_t=a^2\cdot\Delta u$, непрерывны в $\overline{Q}_T$ и удовлетворяют условиям $|u_1(M,0)-u_2(M,0)|\le\varepsilon$, $M\in\overline{\Omega}$, и $|u_1(P,t)-u_2(P,t)|\le\varepsilon$, $P\in S$, $t\in[0,T]$. Тогда $|u_1(M,t)-u_2(M,t)|\le\varepsilon$ во всех точках замкнутого цилиндра $\overline{Q}_T$.

**Доказательство.** Рассмотрим три функции: $v_1(M,t)=-\varepsilon$, $v_3(M,t)=\varepsilon$ и $v_2(M,t)=u_1(M,t)-u_2(M,t)$. Все функции $v_i(M,t)$, $i=1,2,3$, удовлетворяют уравнению $v_t=a^2\cdot\Delta v$, принадлежат классу непрерывных функций $C(\overline{Q}_T)$ и удовлетворяют условиям $v_1(M,0)\le v_2(M,0)\le v_3(M,0)$, $M\in\overline{\Omega}$, и $v_1(P,t)\le v_2(P,t)\le v_3(P,t)$, $P\in S$, $t\in[0,t]$. Из принципов максимума и минимума следует $v_1(M,t)\le v_2(M,t)\le v_3(M,t)$, $(M,t)\in\overline{Q}_T$. ▯

**Замечание.** Принципы максимума и минимума имеют место и для более общего уравнения $C\rho u_t = div\big(k(M)\cdot grad\,u\big) - qu\,,\quad q \ge 0.$ □

4. Для построения математической модели распространения тепла в ограниченном теле $\overline{\Omega}$ необходимо к уравнению (2) добавить дополнительные условия. Такими условиями являются, например, начальное условие, определяющее температуру во всех точках тела в начальный момент времени, и граничное условие. Так как уравнение содержит только первую производную по времени, то достаточно лишь одного начального условия.

Если граница $S$ области $\Omega$ поддерживается при заданной температуре, то

$$u(P,t) = \mu(P,t),\quad P \in S\,,\quad t \ge 0\,,$$

где $\mu(P,t)$ – заданная функция. Это условие называется граничным условием I рода или условием Дирихле.

Математическая модель будет разумно сформулирована только в том случае, если указанные дополнительные условия выделяют **единственное** решение уравнения (2). (Кроме того, требуется, разумеется, чтобы это решение существовало и непрерывно зависело от определяющих эту модель функций.)

**Первая** начально-краевая задача для уравнения теплопроводности состоит в нахождении функции $u(M,t)$, удовлетворяющей следующим требованиям:

$$u_t = a^2\Delta u + f(M,t),\quad (M,t) \in Q\,, \tag{2}$$

$$u(M,0) = \varphi(M),\quad M \in \overline{\Omega}\,, \tag{3}$$

$$u(P,t) = \mu(P,t),\quad P \in S\,,\quad t \in [0,\infty); \tag{4}$$

$$Q = Q_{T=\infty} = \Omega \times [0,\infty).$$

**Определение.** Классическим решением первой начально-краевой задачи (2), (3), (4) называется функция $u(M,t)$, $u(M,t) \in C(\overline{Q}) \cap C^{2,1}(Q)$, **непрерывная** в цилиндре $\overline{Q}$, имеющая непрерывные производные первого порядка по $t$ и второго порядка по координатам точки $M$ в цилиндре $Q$, удовлетворяющая в $Q$ уравнению (2), начальному условию (3) и краевому условию (4).

**Замечание.** Для существования классического решения задачи (2), (3), (4) необходимо требовать согласования условий (3) и (4): $\varphi(P)=\mu(P,0)$ при $P\in S$. Часто возникают задачи, решения которых не могут удовлетворять требованиям, предъявляемым к **классическим** решениям. Например, может не выполняться согласование начального и краевого условий. Такие решения надо понимать в некотором **обобщенном** смысле.

**Теорема единственности решения.** Задача (2), (3), (4) может иметь только одно классическое решение.

**Доказательство.** Допустим, что существуют две функции $u_i(M,t)$, $i=1,2$, являющиеся классическими решениями задачи (2), (3), (4). Пусть $T>0$. Функции $u_i(M,t)\in C(\overline{Q}_T)\cap C^{2,1}(Q_T)$, $i=1,2$. Тогда функция $v(M,t)=u_1(M,t)-u_2(M,t)$ является классическим решением задачи

$$v_t=a^2\Delta v,\quad (M,t)\in Q_T,$$

$$v(M,0)=0,\quad M\in\overline{\Omega},$$

$$v(P,t)=0,\quad P\in S,\quad t\in[0,T],\quad v\in C(\overline{Q}_T).$$

Для функции $v(M,t)$ справедливы принципы максимума и минимума. ( Для функций $u_i(M,t)$, $i=1,2$, принципы максимума и минимума не применимы, так как они удовлетворяют неоднородному уравнению). Следовательно, $v(M,t)$ удовлетворяет принципу экстремума: $0\le v(M,t)\le 0$, т.е. $v(M,t)\equiv 0$. □

Из принципа максимума (при любом $T>0$) следует также

**Теорема об устойчивости решения.** Классическое решение задачи (2), (3), (4) устойчиво по начальным и граничным условиям.

**Доказательство.** Пусть функции $u_i(M,t)$, $i=1,2$, являются решениями задач

$$u_t=a^2\Delta u+f(M,t),\quad (M,t)\in Q_T,$$

$$u(M,0)=\varphi_i(M),\quad M\in\overline{\Omega},$$

$$u(P,t)=\mu_i(P,t),\quad P\in S,\quad t\in[0,T],\quad i=1,2.$$

Предположим, что $\left|\varphi_1(M)-\varphi_2(M)\right|\le\varepsilon$, $\quad M\in\overline{\Omega}$, и $\left|\mu_1(P,t)-\mu_2(P,t)\right|\le\varepsilon$, $\quad P\in S$, $\quad t\in[0,T]$. Надо доказать, что $\left|u_1(M,t)-u_2(M,t)\right|\le\varepsilon$ для всех $(M,t)\in\overline{Q}_T$. Рассмотрим функцию $v(M,t)=u_1(M,t)-u_2(M,t)$, которая является решением задачи

$$v_t=a^2\Delta v,\quad (M,t)\in Q_T,$$

$$v(M,0)=\varphi_1(M)-\varphi_2(M),\quad M\in\overline{\Omega},$$

$$v(P,t)=\mu_1(P,t)-\mu_2(P,t),\quad P\in S,\quad t\in[0,T].$$

Так как функция $v(M,t)$ в граничных и начальных точках удовлетворяет условиям

$\left|v(M,0)\right|\le\varepsilon$, $\quad M\in\overline{\Omega}$, и $\quad\left|v(P,t)\right|\le\varepsilon$, $\quad P\in S$, $\quad t\in[0,T]$, то из теоремы сравнения 2 следует , что $\left|v(M,t)\right|\le\varepsilon$, $\quad(M,t)\in\overline{Q}_T$. ▯

5. **Пример.** $\Omega=\{0<x<l\}\subset\mathrm{R}^1$, $u=u(x,t)$.

$$u_t=a^2u_{xx},\qquad 0<x<l,\qquad 0<t\le T,\tag{5}$$

$$u(x,0)=\varphi(x),\qquad 0\le x\le l,\tag{6}$$

$$u(0,t)=0,\quad u(l,t)=0,\quad 0\le t\le T.\tag{7}$$

Эта задача описывает распространение тепла в однородном стержне с теплоизолированной боковой поверхностью, если в стержне нет источников тепла. Начальная температура стержня задана непрерывной функцией $\varphi(x)$, а на его концах поддерживается нулевая температура. $T$ – любое положительное число, допускается $T=+\infty$. Очевидно, классическое решение задачи возможно только при $\varphi(0)=0$ и $\varphi(l)=0$. Если функция $\varphi(x)$ непрерывно дифференцируема на сегменте $0\le x\le l$ и обращается в нуль на его концах, то нетрудно доказать, что сумма ряда

$$u(x,t)=\sum_{n=1}^{\infty}c_n\cdot\exp\left\{-\left(\frac{\pi na}{l}\right)^2t\right\}\cdot\sin\frac{\pi n}{l}x\tag{8}$$

с коэффициентами $c_n = \dfrac{2}{l}\displaystyle\int_0^l \varphi(x)\sin\dfrac{\pi n}{l}x\,dx$ является классическим решением поставленной задачи.

Из принципа максимума следует, что задача (5), (6), (7) **корректно поставлена:** её решение существует, оно **единственно** и **устойчиво** к малым возмущениям начального и краевых условий.

Отметим, что ряд в (8) при $t<0$ может вовсе не иметь смысла. По этой причине первая начально-краевая задача для уравнения (1) не рассматривается при $t<0$.

6. Пусть теперь $\Omega = \mathrm{R}^3$, $Q = \{(M,t) \mid M \in \mathrm{R}^3, t>0\}$, $\overline{Q} = \{(M,t) \mid M \in \mathrm{R}^3, t \ge 0\}$. Задача Коши состоит в нахождении температуры $u(M,t)$ всюду в $\mathrm{R}^3$ при $t>0$, если всюду в $\mathrm{R}^3$ известна начальная температура:

$$u_t = a^2\Delta u\,,\quad M\in \mathrm{R}^3,\ 0<t<\infty, \tag{9}$$

$$u(M,0) = \varphi(M),\quad M\in \mathrm{R}^3. \tag{10}$$

**Определение.** Классическим решением задачи Коши для уравнения теплопроводности называется функция $u(M,t)\in C(\overline{Q})\cap C^{2,1}(Q)$, удовлетворяющая уравнению (9) в $Q$ и начальному условию (10) при $t=0$ всюду в $\mathrm{R}^3$. □

Нетрудно проверить, что если $\varphi(M)$ – непрерывная и **ограниченная** функция, то функция

$$u(M,t) = \left(\frac{1}{2a\sqrt{\pi t}}\right)^3 \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} e^{-\frac{(x-\xi)^2+(y-\eta)^2+(z-\varsigma)^2}{4a^2t}}\cdot\varphi(\xi,\eta,\varsigma)d\xi d\eta d\varsigma,$$

$M = M(x,y,z)$, является классическим решением задачи Коши, **ограниченным** в $\overline{Q}$. Возникает вопрос о единственности решения этой задачи. Справедлива

**Теорема единственности.** Задача (9), (10) может иметь только одно классическое решение, **ограниченное** в области $\overline{Q}$.

**Замечание.** В этой теореме существенным является требование ограниченности искомой функции во всей области $\overline{Q}$.

Оно как раз и является условием, достаточным для единственности решения задачи (9), (10). ▯

**Замечание.** Непосредственно воспользоваться принципом максимума для доказательства этой теоремы нельзя, так как $\Omega = \mathrm{R}^3$ – неограниченная область. ▯

Ради простоты обозначений доказательство теоремы единственности проведём для случая, когда $\varphi(M)$ не зависит от $y$ и $z$ (а, следовательно, и $u(M,t)$ не зависит от $y$ и $z$ ). То есть, докажем теорему единственности решения задачи Коши на прямой: $Q = \{(x,t) \mid x \in \mathrm{R}^1, t > 0\}$, $\overline{Q} = \{(x,t) \mid x \in \mathrm{R}^1, t \geq 0\}$,

$$u_t = a^2 u_{xx}, \quad (x,t) \in Q, \tag{11}$$

$$u(x,0) = \varphi(x), \quad x \in \mathrm{R}^1, \tag{12}$$

и существует $K = const$, что $|u(x,t)| \leq K$, $(x,t) \in \overline{Q}$.

**Доказательство теоремы единственности.** Предположим, что существуют два ограниченных решения $u_i(x,t)$, $i = 1, 2$, которые удовлетворяют задаче (11), (12). Введём функцию $v(x,t) = u_1(x,t) - u_2(x,t)$, не равную тождественно нулю. В силу линейности задачи (11) , (12) функция $v(x,t)$ будет удовлетворять однородной задаче Коши:

$$v_t = a^2 v_{xx}, \quad (x,t) \in Q, \tag{13}$$

$$v(x,0) = 0, \quad x \in \mathrm{R}^1. \tag{14}$$

Условие ограниченности для функций $u_1(x,t)$ и $u_2(x,t)$ даёт условие ограниченности для функции $v(x,t)$:

$|v(x,t)| = |u_1(x,t) - u_2(x,t)| \leq |u_1(x,t)| + |u_2(x,t)| \leq 2K$, где $|u_1(x,t)| \leq K$, $|u_2(x,t)| \leq K$. Таким образом, функция $v(x,t)$ является решением задачи (13) ,(14) и ограничена в области $\overline{Q}$. Покажем, что $v(x,t) \equiv 0$, $(x,t) \in Q$. Выберем в полуплоскости $Q$ линии $|x| = L$ и $t = T$ и будем рассматривать ограниченную область $Q_L$:

$Q_L = [-L, L]\times(0, T]$, $\overline{Q}_L = [-L, L]\times[0, T]$. Введём вспомогательную функцию $w(x,t) = \dfrac{4K}{L^2}\left(\dfrac{x^2}{2} + a^2 t\right)$.

Функция $w(x,t)$ удовлетворяет уравнению $w_t = a^2 w_{xx}$. Положим $t = 0$, тогда $w(x,0) = \dfrac{2Kx^2}{L^2} \geq |v(x,0)| = 0$. Пусть $|x| = L$, тогда $w(\pm L, t) = 2K + \dfrac{4Ka^2 t}{L^2} \geq 2K \geq |v(\pm L, t)|$. Так как область $Q_L$ ограничена, внутри этой области функции $v(x,t)$ и $w(x,t)$ удовлетворяют однородному уравнению теплопроводности, а на границе выполняются условия $|v(x,0)| \leq w(x,0)$ и $|v(\pm L, t)| \leq w(\pm L, t)$, то к функциям $v(x,t)$ и $w(x,t)$ можно применить следствие из принципа максимума: $|v(x,t)| \leq w(x,t)$, $(x,t)\in \overline{Q}_L$, или $-\dfrac{4K}{L^2}\left(\dfrac{x^2}{2} + a^2 t\right) \leq v(x,t) \leq \dfrac{4K}{L^2}\left(\dfrac{x^2}{2} + a^2 t\right)$. Зафиксируем точку $(x,t)\in \overline{Q}_L$ и перейдём к пределу при $L \to \infty$, получим $\lim\limits_{L\to\infty} v(x,t) = 0$. В силу независимости $v(x,t)$ от $L$ и произвольности выбора точки $(x,t)$ получим, что всюду в области $\overline{Q}$ $v(x,t) \equiv 0$. Следовательно, $u_1(x,t) \equiv u_2(x,t)$ в $Q$, и решение задачи единственно. □

**Замечание.** Можно значительно ослабить ограничение на рост решения при $|x| \to \infty$ (при $M \to \infty$ в $\mathrm{R}^3$), отказавшись от ограниченности решения задачи Коши. Для этого введём класс $U$ функций, которые всюду в $\mathrm{R}^1$ и при всех $t \geq 0$ удовлетворяют неравенствам $|u(x,t)| < K\cdot\exp(N\cdot x^2)$ с некоторыми положительными постоянными $K$ и $N$ ($K$ и $N$ определяются функцией $u(x,t)$). В $U$ справедлива теорема единственности.

## Вопрос 21.

## Постановка внешних и внутренних краевых задач для уравнения Лапласа. Условие разрешимости внутренней задачи Неймана.

1. Дифференциальный оператор $\Delta u \equiv div\, grad\, u$ называется оператором Лапласа; $\Delta u = 0$ – уравнение Лапласа. $\Delta u = 0$ – уравнение эллиптического типа. Будем рассматривать удовлетворяющие уравнению Лапласа действительные функции $u = u(M)$ точки $M$ в пространстве $\mathrm{R}^3$ или на плоскости $\mathrm{R}^2$. Дважды непрерывно дифференцируемая в области $\Omega$ функция $u(M)$, удовлетворяющая в $\Omega$ уравнению Лапласа, называется **гармонической** в этой области функцией.

В декартовой системе координат в $\mathrm{R}^3$ уравнение Лапласа имеет вид

$$\Delta_{(x,y,z)} u(x,y,z) = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2} = 0,$$

в сферической ( $x = r\sin\theta\cos\varphi$ , $y = r\sin\theta\sin\varphi$ , $z = r\cos\theta$ ) –

$$\Delta_{(r,\theta,\varphi)} u(r,\theta,\varphi) = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial u}{\partial r}\right) + \frac{1}{r^2\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial u}{\partial\theta}\right) +$$

$$+\frac{1}{r^2\sin^2\theta}\frac{\partial^2 u}{\partial\varphi^2} = 0,$$

в цилиндрической $(x = r\cos\varphi,\quad y = r\sin\varphi,\quad z = z)$ –

$$\Delta_{(r,\varphi,z)} u(r,\varphi,z) = \frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial u}{\partial r}\right) + \frac{1}{r^2}\frac{\partial^2 u}{\partial\varphi^2} + \frac{\partial^2 u}{\partial z^2} = 0.$$

В плоском случае уравнение Лапласа имеет вид $\Delta_{(x,y)} u(x,y) = \dfrac{\partial^2 u}{\partial x^2} + \dfrac{\partial^2 u}{\partial y^2} = 0$ в декартовой системе координат или $\Delta_{(r,\varphi)}u(r,\varphi)=\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial u}{\partial r}\right)+\frac{1}{r^2}\frac{\partial^2 u}{\partial\varphi^2}=0$ – в полярной $(x=r\cos\varphi,\quad y=r\sin\varphi)$.

Уравнение Лапласа описывает разнообразные физические явления и процессы:

- стационарное тепловое поле в однородном теле или в однородной пластине в отсутствие источников тепла;
- потенциальное течение несжимаемой жидкости;
- потенциал стационарного электрического тока;
- потенциал статического поля электрических зарядов;
- и др.

Однородность уравнения Лапласа выражает отсутствие источников поля в рассматриваемой области; если такие источники имеются, то вместо уравнения Лапласа получим уравнение Пуассона $\Delta u=-f$.

Для однозначного определения искомой функции $u(M)$ в области $\Omega$ требуются дополнительные условия. Чаще всего на границе $S$ области $\Omega$ ставятся **краевые условия** одного из следующих типов:

I. $u(P)=\mu(P),\quad P\in S$, (первая краевая задача)

II. $\left.\frac{\partial u}{\partial n}\right|_{P\in S}=\nu(P)$, (вторая краевая задача)

III. $\alpha\frac{\partial u(P)}{\partial n}+\beta u(P)=\chi(P),\quad P\in S$, (третья краевая задача), где $\alpha+\beta\neq 0$, $\alpha\geq 0$, $\beta\geq 0$, $\mu$, $\nu$, $\alpha$, $\beta$, $\chi$ – заданные функции, $\frac{\partial u}{\partial n}$ – производная функции $u(M)$ по внешней нормали к границе $S$.

Первую краевую задачу называют задачей Дирихле, а вторую – задачей Неймана. Изучая именно третью краевую задачу (а не первую и не вторую), будем предполагать, что $\alpha(P)\neq 0$ всюду на $S$. Тогда краевое условие $\alpha\frac{\partial u}{\partial n}+\beta u=\chi$ можно записать в виде $\frac{\partial u}{\partial n}+hu=\eta$, где $h=\frac{\beta}{\alpha}$, $\eta=\frac{\chi}{\alpha}$.

Если ищется решение задачи в ограниченной области, то задачу называют внутренней, если вне ограниченной области, то – внешней краевой задачей.

Решения всех краевых задач будем подразумевать **классическими**; это накладывает определенные требования на постановку задачи. Следует отметить, что многие практические задачи не удовлетворяют таким требованиям; это приводит к необходимости обобщения понятия решения.

**Замечание.** В случае двух независимых действительных переменных гармонические функции очевидным образом связаны с аналитическими функциями одной комплексной переменной. Например, если $f(z)=u(x,y)+iv(x,y)$, $z=x+iy$, – аналитическая функция, то $u$ и $v$ – гармонические функции, связанные условиями Коши-Римана. Это обстоятельство, разумеется, проявляет себя в краевых задачах для гармонических функций двух переменных.

2. **Внутренняя задача Дирихле.**

$$\Delta u(M)=0\,,\quad M\in\Omega\,,$$

$$u(P)=\mu(P),\quad P\in S\,.$$

Пусть $\Omega$ – ограниченная область (в $\mathsf{R}^3$ или в $\mathsf{R}^2$) с достаточно гладкой границей $S$ ($S$ – замкнутая поверхность в $\mathsf{R}^3$ или $S$ – замкнутая кривая в $\mathsf{R}^2$).

Найти функцию $u(M)$, которая определена и непрерывна в замкнутой области $\overline{\Omega}=\Omega\cup S$, удовлетворяет внутри области $\Omega$ уравнению Лапласа $\Delta u=0$ и принимает на границе $S$ заданные значения: $u(P)=\mu(P)$, $P\in S$.

Таким образом, мы ищем функцию $u(M)$, гармоничную внутри $\Omega$. Требование гармоничности функции $u(M)$ на границе излишне. Условие непрерывности функции $u(M)$ в замкнутой области необходимо для единственности решения задачи. Если это условие отбросить, то единственность решения задачи нарушится, так как любая функция вида $u(M)=\begin{cases} const\,, & M\in\Omega\,,\\ \mu(M), & M\in S\end{cases}$ будет решением задачи. Поэтому требуем $\mu\in C(S)$. В этом случае функция $u(M)$ называется классическим решением задачи. $u(M) \in C(\overline{\Omega}) \cap C^2(\Omega)$, $\mu \in C(S)$.

**Теорема единственности решения внутренней задачи Дирихле.** Задача Дирихле не может иметь более одного классического решения.

**Доказательство.** Допустим, что существуют два решения задачи Дирихле $u_1(M)$ и $u_2(M)$. Введём функцию $v(M) = u_1(M) - u_2(M)$. $v(M)$ будет удовлетворять условиям

$$v \in C(\overline{\Omega}); \quad \Delta v(M) = 0, \quad M \in \Omega; \quad v(P)\big|_{P \in S} = 0.$$

Функция $v(M)$ является решением задачи Дирихле с однородными граничными условиями. По теореме Вейерштрасса любая непрерывная функция в замкнутой ограниченной области достигает своего максимального и минимального значений. Если $v(M) > 0$ хотя бы в одной точке области $\Omega$, то она достигает своего максимального значения внутри области $\Omega$, но это невозможно в силу принципа максимума, так как $v(M)\big|_{M \in S} = 0$. Аналогично доказывается, что $v(M)$ не может быть меньше нуля внутри $\Omega$. Следовательно, $v(M) \equiv 0$. ▯

Из принципа максимума следует также **устойчивость** решения внутренней задачи Дирихле. Рассмотрим две задачи Дирихле:

$$\Delta u_i(M) = 0, \quad M \in \Omega,$$

$$u_i(P)\big|_{P \in S} = \mu_i(P), \quad i = 1, 2.$$

Введём расстояния между функциями:

$$\rho(\mu_1, \mu_2) = \|\mu_1 - \mu_2\|_{C(S)} = \max_{P \in S} |\mu_1(P) - \mu_2(P)|,$$

$$\rho(u_1, u_2) = \|u_1 - u_2\|_{C(\overline{\Omega})} = \max_{M \in \overline{\Omega}} |u_1(M) - u_2(M)|.$$

**Определение.** Решение задачи Дирихле называется устойчивым, если для любого $\varepsilon > 0$ существует $\delta(\varepsilon) > 0$ такое, что из неравенства $\rho(\mu_1, \mu_2) < \delta(\varepsilon)$ следует $\rho(u_1, u_2) < \varepsilon$.

**Теорема.** Для внутренней задачи Дирихле, если $\rho(\mu_1, \mu_2) \le \varepsilon$, то $\rho(u_1, u_2) \le \varepsilon$. ▯

**Замечание.** Требование достаточной гладкости границы $S$ связано с желанием пользоваться формулами Грина и представлять решение задачи Дирихле с помощью производной по нормали к $S$ от функции Грина этой краевой задачи. Выражение «$S$ – достаточно гладкая граница области $\Omega$» как раз и означает, что для неё справедлива первая формула Грина.

Всюду далее мы считаем, что $S$ – достаточно гладкая граница области $\Omega$.

3. **Внутренняя задача Неймана.**

$$\Delta u(M) = 0, \quad M \in \Omega, \tag{1}$$

$$\frac{\partial u(P)}{\partial n} = \nu(P), \quad P \in S. \tag{2}$$

Найти функцию $u(M)$, которая определена, непрерывна и непрерывно дифференцируема в замкнутой области $\overline{\Omega}$ $\left(\Omega \subset \mathrm{R}^3\right.$ или $\left.\Omega \subset \mathrm{R}^2\right)$, удовлетворяет внутри области $\Omega$ уравнению Лапласа $\Delta u = 0$, а её нормальная производная принимает на границе заданные значения $\left.\frac{\partial u(P)}{\partial n}\right|_{P \in S} = \nu(P), \quad P \in S$.

$u(M) \in C^1(\overline{\Omega}) \cap C^2(\Omega), \quad \nu \in C(S)$.

Очевидно, если $u(M)$ – решение задачи Неймана, то $v(M) = u(M) + C$, где $C = const$, – тоже решение той же задачи Неймана. Это легко проверить, если подставить функцию $v(M)$ в задачу (1), (2).

**Теорема.** Решение задачи Неймана определяется с точностью до произвольной постоянной, т.е. если $u_1(M)$ и $u_2(M)$ – решения одной и той же задачи Неймана, то $u_1(M) - u_2(M) = const$.

**Доказательство.** Здесь нельзя воспользоваться принципом максимума, так как неизвестно чему равны значения функции $u(M)$ на границе. Поэтому используем формулы Грина. Допустим, что существуют две функции $u_1(M)$ и $u_2(M)$, являющиеся решениями задачи Неймана (1), (2). Рассмотрим функцию $v(M)=u_1(M)-u_2(M)$, которая не равна тождественно нулю. $v(M)\in C^1(\overline{\Omega})\cap C^2(\Omega)$. Функция $v(M)$ удовлетворяет задаче

$$\Delta v(M)=0,\quad M\in\Omega,$$

$$\frac{\partial v(P)}{\partial n}=0,\quad P\in S.$$

Очевидно, решением этой задачи является $v(M)=const$. Докажем, что других решений нет. Применим первую формулу Грина к функции $v(M)$:

$$\iiint_{\Omega} v\Delta v d\tau=\iint_{S} v\frac{\partial v}{\partial n}d\sigma-\iiint_{\Omega}(grad\, v, grad\, v)d\tau .$$

Учитывая, что $\Delta v(M)=0,\quad M\in S$, и $\left.\frac{\partial v(P)}{\partial n}\right|_{P\in S}=0$, получим

$\iiint_{\Omega}(grad\, v, grad\, v)d\tau=0$, т.е.

$(grad\, v, grad\, v)=\left(\frac{\partial v}{\partial x}\right)^2+\left(\frac{\partial v}{\partial y}\right)^2+\left(\frac{\partial v}{\partial z}\right)^2=0$. Так как сумма неотрицательных слагаемых равна нулю, то $\frac{\partial v}{\partial x}=\frac{\partial v}{\partial y}=\frac{\partial v}{\partial z}=0$ всюду в области $\Omega$. Отсюда получаем $v(M)=const$.▯

**Свойство гармонической функции.** Если функция $u(M)$ гармонична в ограниченной области $\Omega$ с достаточно гладкой границей $S$, то $\iint_{S}\frac{\partial u(P)}{\partial n}d\sigma=0$.

Доказательство. Положим в первой формуле Грина $v(M)\equiv 1$, тогда получим

$\iint_{S} v\frac{\partial u}{\partial n}d\sigma=\iiint_{\Omega} v\Delta u\, d\tau-\iiint_{\Omega}(grad\, v, grad\, u)d\tau=0$, так как $\Delta u=0$ и $grad\, v=0$.▯

Из доказанного свойства вытекает **необходимое условие существования решения** внутренней задачи Неймана для уравнения Лапласа:

решение задачи $\Delta u = 0\,,\quad M \in \Omega,\quad \left.\dfrac{\partial u}{\partial n}\right|_{P\in S} = \nu(P)$

может существовать только если $\iint\limits_S \nu(P)d\sigma = 0$. Условие разрешимости аналогичной задачи для уравнения Пуассона

$$\Delta u = -f(M),\quad M \in \Omega,$$

$$\left.\frac{\partial u}{\partial n}\right|_{P\in S} = \nu(P)$$

выглядит так: $\iiint\limits_\Omega f(M)d\tau + \iint\limits_S \nu(P)d\sigma = 0$. Это условие также вытекает из первой формулы Грина, если положить $v(M) \equiv 1$. Физический смысл последнего условия состоит в том, что стационарный поток тепла (несжимаемой жидкости, напряженности электрического поля и т.п.) через замкнутую поверхность $S$ равен суммарной величине всех источников (зарядов и т.п.), находящихся внутри $S$.

Совершенно аналогичное необходимое условие разрешимости задачи Неймана для уравнения Лапласа имеет место и для плоской области $\Omega$, ограниченной кривой $S$: $\int\limits_S \nu(P)\,dl = 0$.

4. **Внутренняя третья краевая задача.**

$$\Delta u(M) = 0,\quad M \in \Omega\,,$$

$$\alpha(P)\frac{\partial u(P)}{\partial n} + \beta(P)u(P) = \chi(P),\ P \in S\,,\ \alpha(P) \ge 0\,,\ \beta(P) \ge 0\,,$$

$$\alpha(P) + \beta(P) > 0\,.$$

Пусть коэффициент $\alpha(P)$ не равен нулю всюду на $S$. Тогда разделим граничное условие на $\alpha(P)$ и получим задачу

$$\Delta u(M) = 0, \quad M \in \Omega, \tag{3}$$

$$\frac{\partial u(P)}{\partial n} + h(P)u(P) = \eta(P), \quad P \in S, \tag{4}$$

где функция $h(P) = \dfrac{\beta(P)}{\alpha(P)}$ не равна тождественно нулю на $S$ (иначе получим вторую краевую задачу, а не третью), а $\eta(P) = \dfrac{\chi(P)}{\alpha(P)}$.

Найти функцию $u(M)$, которая определена, непрерывна и непрерывно дифференцируема в замкнутой области $\overline{\Omega}$ ($\Omega \subset \mathrm{R}^3$ или $\Omega \subset \mathrm{R}^2$), удовлетворяет внутри области $\Omega$ уравнению Лапласа $\Delta u = 0$, на границе удовлетворяет условию (4). $u(M) \in C^1(\overline{\Omega}) \cap C^2(\Omega)$, $\eta(P) \in C(S)$, $h(P) \in C(S)$.

**Теорема единственности решения третьей внутренней краевой задачи.** Если функция $h(P)$ неотрицательна всюду на $S$, не равна тождественно нулю и непрерывна на $S$, то третья краевая задача не может иметь более одного решения.
Доказательство этой теоремы основано на применении первой формулы Грина и существенно использует условие $h(P) \geq 0$.

Условие $h(P) \geq 0, \quad P \in S$, существенно для единственности решения задачи (3), (4).

**Пример.** Если $h(P) < 0, \quad P \in S$, то можно построить задачу, которая имеет более одного решения. Рассмотрим круг радиуса $a$ (область $\Omega \subset \mathrm{R}^2$ с границей $S$). Надо найти функцию $u(M)$, удовлетворяющую условиям

$$\Delta u(M) = 0, \quad M \in \Omega,$$

$$\frac{\partial u(P)}{\partial n} + h(P)u(P) = 0, \quad P \in S.$$

Возьмём две функции $u_1 \equiv 0$ и $u_2 = x$ и подберём $h(P)$ таким образом, чтобы обе они были решениями этой задачи. Запишем задачу в полярной системе координат. Тогда получим

$x = r\cdot\cos\varphi$, $u_2 = r\cdot\cos\varphi$ и $\dfrac{\partial u_2}{\partial n} = \dfrac{\partial u_2}{\partial r} = \cos\varphi$ (нормаль $n$ к границе $S$ совпадает с направлением радиуса). Если $h = -\dfrac{1}{a}$, то условие $\dfrac{\partial u(P)}{\partial n} + h(P)u(P) = 0$, $P \in S$, превратится в тождество как для функции $u_1$, так и для функции $u_2$. Следовательно, из-за выбора $h < 0$ мы построили два решения третьей краевой задачи. □

5. **Внешние краевые задачи** ставятся по-разному в $\mathsf{R}^3$ и в $\mathsf{R}^2$. Это связано с различием поведения на бесконечности фундаментальных решений уравнения Лапласа в трехмерном и в двумерном случаях.

Пусть $M_0(x_0, y_0, z_0)$ – некоторая фиксированная точка в $\mathsf{R}^3$. Найдём решение уравнения Лапласа, зависящее только от $R_{MM_0} = \sqrt{(x-x_0)^2 + (y-y_0)^2 + (z-z_0)^2}$. Введём сферическую систему координат $(r,\theta,\varphi)$ с центром в точке $M_0(r_0,\theta_0,\varphi_0)$. Такое решение будет обладать сферической симметрией, т.е. $\dfrac{\partial u}{\partial\theta} = 0$, $\dfrac{\partial u}{\partial\varphi} = 0$. Следовательно, будет выполняться уравнение $\dfrac{1}{r^2}\dfrac{d}{dr}\left(r^2\dfrac{du}{dr}\right) = 0$. Общее решение этого уравнения имеет вид $u(r) = \dfrac{C_1}{r} + C_2$, где $r = R_{MM_0}$. Решение $u(M, M_0) = \dfrac{1}{R_{MM_0}}$ называется **фундаментальным решением** уравнения Лапласа в пространстве. Это гармоническая функция переменной $M$, определённая всюду в $\mathsf{R}^3$, кроме точки $M = M_0$.

В двумерном случае введём полярную систему координат $(r,\varphi)$ с центром в точке $M_0(x_0, y_0)$. Найдём решение уравнения Лапласа, зависящее только от $R_{MM_0} = \sqrt{(x-x_0)^2 + (y-y_0)^2}$. В этом случае $\frac{\partial u}{\partial \varphi}=0$, и будет выполняться уравнение $\frac{1}{r}\frac{d}{dr}\left(r\frac{du}{dr}\right)=0$. Его общее решение имеет вид $u(r)=C_1\ln\frac{1}{r}+C_2$, где $r=R_{MM_0}$. Функция $u(M,M_0)=\ln\frac{1}{R_{MM_0}}$ называется **фундаментальным решением** уравнения Лапласа на плоскости.

Рассмотрим сначала случай трех переменных. Пусть $r=\sqrt{x^2+y^2+z^2}$.

**Определение.** Функция $u(M)$, $M\in R^3$, называется регулярной на бесконечности, если при всех достаточно больших $r$ выполнены неравенства

$$|u|\le\frac{A}{r},\quad \left|\frac{\partial u}{\partial x}\right|\le\frac{A}{r^2},\quad \left|\frac{\partial u}{\partial y}\right|\le\frac{A}{r^2},\quad \left|\frac{\partial u}{\partial z}\right|\le\frac{A}{r^2},\quad A=const.\square$$

Рассмотрим некоторую ограниченную область $\Omega\subset R^3$, границей которой является замкнутая поверхность $S$. Обозначим через $\Omega'$ область, являющуюся внешней к $S$. $R^3=\Omega\cup S\cup\Omega'$.

**Теорема.** Пусть функции $u(M)$ и $v(M)$ регулярны на бесконечности и $u,v\in C^2(\Omega')\cap C^1(\overline{\Omega'})$. Тогда для $u(M)$ и $v(M)$ имеет место первая формула Грина:

$$\iiint\limits_{\Omega'}u\Delta v d\tau=\iint\limits_{S}u\frac{\partial v}{\partial n}d\sigma-\iiint\limits_{\Omega'}(grad\,u,grad\,v)d\tau. \quad (5)\ \square$$

Поменяв местами функции $u$ и $v$ в соотношении (5) и вычтя из одного соотношения другое, получим вторую формулу Грина:

$$\iiint\limits_{\Omega'}(u\Delta v-v\Delta u)d\tau=\iint\limits_{S}\left(u\frac{\partial v}{\partial n}-v\frac{\partial u}{\partial n}\right)d\sigma.$$

Учитывая, что фундаментальное решение $\dfrac{1}{R_{MM_0}}$ уравнения Лапласа регулярно на бесконечности, получим третью (основную) формулу Грина для функции $u(M)$ в области $\Omega'$. При этом нормаль к поверхности $S$ должна быть внешней по отношению к $\Omega'$, т.е. направлена внутрь $\Omega$.

$$u(M_0)=\frac{1}{4\pi}\iint_S\left(\frac{1}{R_{M_0P}}\frac{\partial u(P)}{\partial n_P}-u(P)\frac{\partial}{\partial n_P}\left(\frac{1}{R_{M_0P}}\right)\right)d\sigma_P-$$
$$-\frac{1}{4\pi}\iiint_{\Omega'}\frac{\Delta u(M)}{R_{M_0M}}d\tau_M\ .$$

**Теорема.** Гармоническая в области $\Omega'\subset \mathrm{R}^3$ функция, равномерно стремящаяся к нулю на бесконечности, является регулярной на бесконечности.
Доказательство этой теоремы основано на преобразовании Кельвина гармонической функции $u(M)$.

В случае двух переменных будем рассматривать аналогичные по смыслу $\Omega$, $S$, $\Omega'$ в $\mathrm{R}^2$. Пусть $r=\sqrt{x^2+y^2}$.

**Определение.** Функция двух переменных $u(x,y)$ называется регулярной на бесконечности, если она имеет конечный предел на бесконечности. ▯

**Теорема.** Для **гармонической** в $\Omega'$ функции $u(x,y)$, регулярной на бесконечности, при достаточно больших значениях $r$ справедливы **н**еравенства

$$\left|\frac{\partial u}{\partial x}\right|\le\frac{A}{r^2},\quad\left|\frac{\partial u}{\partial y}\right|\le\frac{A}{r^2},\qquad A=const\ .▯$$

**Теорема.** Для функций двух переменных, **гармонических** в $\Omega'$ и регулярных на бесконечности, справедливы формулы Грина. ▯

Чтобы внешняя краевая задача для уравнения Лапласа имела единственное решение, требуют его регулярности на бесконечности.

6. **Внешняя задача Дирихле** в $\Omega' \subset \mathrm{R}^3$.

Найти функцию $u(M)$, удовлетворяющую уравнению Лапласа $\Delta u = 0$ в неограниченной области $\Omega'$, непрерывную в замкнутой области $\overline{\Omega'} = \Omega' \cup S$, принимающую на границе заданные значения $u(P) = \mu(P)$, $P \in S$, и **равномерно стремящуюся к нулю на бесконечности**. (Здесь $\Omega' = \mathrm{R}^3 \setminus \overline{\Omega}$, $\Omega$ – ограниченная область с границей $S$.) $u(M) \in C(\overline{\Omega'}) \cap C^2(\Omega')$.

**Замечание.** Условие равномерного стремления к нулю функции $u(M)$ на бесконечности важно для единственности решения задачи.

**Пример.** Рассмотрим задачу $\Delta u(r) = 0$, $r > a$, $u(r = a) = 1$. Функция $u(r)$ вне шара радиуса $a$ удовлетворяет уравнению Лапласа, а на границе принимает заданное значение $u(r = a) = 1$.

Тогда функции $u_1(r) \equiv 1$ и $u_2(r) = \dfrac{a}{r}$ удовлетворяют и уравнению Лапласа вне шара и граничному условию. Получили два решения поставленной задачи. Если учесть условие на бесконечности, то функция $u_1(r) = 1$ не подходит. Из двух решений $u_1(r) = 1$ и $u_2(r) = \dfrac{a}{r}$ можно построить целое семейство решений $u(r) = \alpha u_1(r) + \beta u_2(r)$, $\alpha + \beta = 1$, которое будет удовлетворять поставленной задаче. Условие $\lim\limits_{r \to \infty} u = 0$ позволяет выделить единственное решение внешней задачи Дирихле. ▯

**Теорема.** Внешняя задача Дирихле в пространстве $\mathrm{R}^3$ может иметь только одно решение.

Доказательство этой теоремы основано на применении принципа максимума в области, ограниченной поверхностью $S$ и сферой $S_r$ достаточно большого радиуса $r$.

7. **Внешняя задача Неймана** в $\Omega' \subset \mathrm{R}^3$.

Найти функцию $u(M)$, удовлетворяющую уравнению Лапласа $\Delta u = 0$ в неограниченной области $\Omega'$, удовлетворяющую условию $\dfrac{\partial u(P)}{\partial n}=\nu(P),\quad P\in S$, и **равномерно стремящуюся к нулю на бесконечности.** (Здесь $\Omega'=\mathrm{R}^3\setminus\overline{\Omega}$, $\Omega$ – ограниченная область с границей $S$.) $u(M)\in C^1(\overline{\Omega'})\cap C^2(\Omega')$, $\nu(P)\in C(S)$.

Требование регулярности искомой функции на бесконечности гарантирует единственность решения поставленной задачи:

**Теорема.** Внешняя задача Неймана в $\mathrm{R}^3$ имеет единственное классическое решение.

Доказательство этой теоремы основано на применении первой формулы Грина в неограниченной области $\Omega'$.

**Замечание.** Решение внешней задачи Неймана, регулярное на бесконечности, существует только при выполнении условия $\iint\limits_S \nu(P)d\sigma=0$. Для доказательства необходимости этого условия применим вторую формулу Грина в области, ограниченной $S$ и сферой $S_r$ достаточно большого радиуса $r$, к функциям $u(M)$ – решению задачи Неймана в $\Omega'$ – и $\nu(M)\equiv 1$. Переходя к пределу при $r\to\infty$ и пользуясь оценкой $\left|grad\,u(M)\right|=O\left(\dfrac{1}{r^2}\right)$ для регулярной на бесконечности гармонической функции $u(M)$, мы и получим условие разрешимости внешней задачи Неймана.

8. **Внешняя третья краевая задача** в $\Omega'\subset\mathrm{R}^3$.

Найти функцию $u(M)$, удовлетворяющую уравнению Лапласа $\Delta u=0$ в неограниченной области $\Omega'$, удовлетворяющую условию $\dfrac{\partial u(P)}{\partial n}+h(P)u(P)=\eta(P),\quad P\in S$, $\eta(P)\in C(S)$, $h(P)\geq 0$, $h(P)\in C(S)$ ($h(P)$ не равна тождественно нулю), и **равномерно стремящуюся к нулю на бесконечности.** (Здесь $\Omega'=\mathrm{R}^3\setminus\overline{\Omega}$, $\Omega$ – ограниченная область с границей $S$.) $u(M)\in C^1(\overline{\Omega'})\cap C^2(\Omega')$.

Требование регулярности искомой функции на бесконечности позволяет применить формулы Грина в неограниченной области $\Omega'$ и доказать единственность решения поставленной задачи:

**Теорема.** Внешняя третья краевая задача имеет единственное классическое решение, если $h(P) \geq 0$ и не равна тождественно нулю на границе области.

9. **Внешняя задача Дирихле** в $\Omega' \subset \mathsf{R}^2$.

Для уравнения Лапласа **на плоскости** требование равномерного стремления решения к нулю на бесконечности является жестким, такого решения может не существовать.

**Пример.** В полярных координатах $(r, \varphi)$ на плоскости рассмотрим задачу Дирихле вне круга радиуса $a \neq 1$:

$$\Delta u(r) = 0, \quad r > a,$$

$$u(r = a) = 1.$$

Можно построить два решения этой задачи: $u_1(r) \equiv 1$ и $u_2(r) = \dfrac{\ln r}{\ln a}$. Других линейно независимых решений у этой задачи нет, общее решение уравнения $\Delta u(r) = 0$ имеет вид $u(r) = C_1 + C_2 \ln r$. Если потребовать, чтобы функция $u(r)$ равномерно стремилась к нулю на бесконечности, то ни одно из этих решений не подходит, задача не имеет решений. Условие равномерного стремления к нулю на бесконечности надо заменить требованием существования конечного предела решения на бесконечности. Тогда подходит решение $u_1(r) \equiv 1$.□

Во внешней задаче Дирихле на плоскости требуется найти функцию $u(M)$, удовлетворяющую уравнению Лапласа $\Delta u = 0$ в неограниченной области $\Omega'$, непрерывную в замкнутой области $\overline{\Omega'} = \Omega' \cup S$, принимающую на границе заданные значения $u(P) = \mu(P)$, $P \in S$, и **ограниченную на бесконечности.** (Здесь $\Omega' = \mathsf{R}^2 \setminus \overline{\Omega}$, $\Omega$ – ограниченная область с границей $S$.) $u(M) \in C(\overline{\Omega'}) \cap C^2(\Omega')$, $\mu(P) \in C(S)$.

**Теорема.** Внешняя задача Дирихле на плоскости может иметь не более одного классического решения, регулярного на бесконечности.

10. **Внешняя задача Неймана** в $\Omega' \subset \mathrm{R}^2$.

Найти функцию $u(M)$, удовлетворяющую уравнению Лапласа $\Delta u = 0$ в неограниченной области $\Omega'$, удовлетворяющую условию $\dfrac{\partial u(P)}{\partial n} = \nu(P), \quad P \in S$, и **регулярную на бесконечности.** (Здесь $\Omega' = \mathrm{R}^2 \setminus \overline{\Omega}$, $\Omega$ – ограниченная область с границей $S$.) $u(M) \in C^1(\overline{\Omega'}) \cap C^2(\Omega')$, $\nu(P) \in C(S)$.

**Теорема.** Классическое решение внешней задачи Неймана на плоскости, ограниченное и регулярное на бесконечности, определяется с точностью до постоянного слагаемого.

Доказательство этой теоремы основано на применении первой формулы Грина к регулярной на бесконечности гармонической функции в неограниченной области $\Omega'$.

**Замечание.** Для существования решения внешней задачи Неймана, регулярного на бесконечности, необходимо выполнение условия $\int\limits_S \nu(P)\,dl = 0$.

11. **Внешняя третья краевая задача** в $\Omega' \subset \mathrm{R}^2$.

Найти функцию $u(M)$, удовлетворяющую уравнению Лапласа $\Delta u = 0$ в неограниченной области $\Omega'$, удовлетворяющую условию $\dfrac{\partial u(P)}{\partial n} + h(P)u(P) = \eta(P), \quad P \in S$, $\eta(P) \in C(S)$, $h(P) \geq 0$, $h(P) \in C(S)$ ($h(P)$ не равна тождественно нулю), и **регулярную на бесконечности.** (Здесь $\Omega' = \mathrm{R}^2 \setminus \overline{\Omega}$, $\Omega$ – ограниченная область с границей $S$.) $u(M) \in C^1(\overline{\Omega'}) \cap C^2(\Omega')$.

**Теорема.** Если $h(P) \geq 0$ и $h(P)$ не равна тождественно нулю, то внешняя третья краевая задача на плоскости может иметь не более одного классического решения, регулярного на бесконечности.

## Вопрос 22.

## Формулы Грина. Функция Грина для внутренней задачи Дирихле.

1. **Первая формула Грина** в области $\Omega \subset \mathrm{R}^3$.

Пусть $\Omega$ – ограниченная область в $\mathrm{R}^3$, границей которой является замкнутая поверхность $S$; $\overline{\Omega} = \Omega \cup S$. Будем считать всюду далее, что $S$ – достаточно гладкая поверхность, к которой применима теорема Остроградского-Гаусса.

**Теорема.** Пусть в области $\Omega$ заданы две функции $u(M)$ и $v(M)$: $u, v \in C^1(\overline{\Omega}) \cap C^2(\Omega)$. Тогда в $\Omega$ справедлива первая формула Грина:

$$\iiint\limits_{\Omega} v\Delta u\, d\tau = \iint\limits_{S} v\frac{\partial u}{\partial n} d\sigma - \iiint\limits_{\Omega} (grad\, u, grad\, v) d\tau . \qquad (1)$$

**Доказательство.** Первая формула Грина получается из теоремы Остроградского-Гаусса

$$\iiint\limits_{\Omega} div\, \vec{A}\, d\tau = \iint\limits_{S} (\vec{A}, \vec{n}) d\sigma$$

($\vec{n}$ – единичная внешняя нормаль к $S$), если специальным образом выбрать векторное поле $\vec{A}$. Возьмем в качестве этого векторного поля $\vec{A} = v(M) \cdot grad\, u(M)$.

$div(v \cdot grad\, u) = v \cdot div(grad\, u) + (grad\, v, grad\, u) =$
$= v \cdot \Delta u + (grad\, v, grad\, u), \quad \Delta u = div(grad\, u)$. Отсюда
$v \cdot \Delta u = div(v \cdot grad\, u) - (grad\, v, grad\, u)$. Теперь, проинтегрировав по $\Omega$, получим формулу (1). □

2. **Вторая формула Грина** в $\Omega \subset \mathrm{R}^3$.

Если поменять местами функции $u(M)$ и $v(M)$, то

$$\iiint\limits_{\Omega} u\Delta v\, d\tau = \iint\limits_{S} u\frac{\partial v}{\partial n} d\sigma - \iiint\limits_{\Omega} (grad\, u, grad\, v) d\tau . \qquad (2)$$

Вычтем из формулы (1) формулу (2), тогда получим **вторую формулу Грина**:

$$\iiint_{\Omega}(v\Delta u-u\Delta v)d\tau=\iint_{S}\left(v\frac{\partial u}{\partial n}-u\frac{\partial v}{\partial n}\right)d\sigma\,. \tag{3}$$

Вторая формула Грина симметрична относительно функций $u(M)$ и $v(M)$.

### 3. Третья (основная) формула Грина в $\Omega\subset\mathrm{R}^3$.

Пусть $u(M)$ – произвольная функция, дважды непрерывно дифференцируемая в ограниченной области $\Omega$ и непрерывная вместе с производными первого порядка в замкнутой области $\overline{\Omega}$: $u\in C^1(\overline{\Omega})\cap C^2(\Omega)$. Фиксируем точку $M_0\in\Omega$ и выберем функцию $v(M,M_0)=\dfrac{1}{R_{MM_0}}$, где $R_{MM_0}$ – расстояние от $M_0$ до $M$; $v(M,M_0)$ является фундаментальным решением уравнения Лапласа в $\mathrm{R}^3$. Применим вторую формулу Грина к функциям $u(M)$ и $v(M,M_0)$. Функция $v(M,M_0)$ имеет особенность, когда точка $M$ совпадает с точкой $M_0$, поэтому применить вторую формулу Грина во всей области $\Omega$ нельзя.

Окружим точку $M_0$ шаром $K_\varepsilon^{M_0}$ радиуса $\varepsilon$ с центром $M_0$, ограниченным сферой $\Sigma_\varepsilon^{M_0}$. Применим к функциям $u(M)$ и $v(M,M_0)$ вторую формулу Грина в области $\Omega\setminus K_\varepsilon^{M_0}$:

$$\iiint_{\Omega\setminus K_\varepsilon^{M_0}}(u\Delta v-v\Delta u)d\tau=\iint_{S}\left(u\frac{\partial v}{\partial n}-v\frac{\partial u}{\partial n}\right)d\sigma+$$

$$+\iint_{\Sigma_\varepsilon^{M_0}}\left(u\frac{\partial v}{\partial n}-v\frac{\partial u}{\partial n}\right)d\sigma\,. \tag{4}$$

В интеграле $\iint\limits_S \left( u\dfrac{\partial v}{\partial n} - v\dfrac{\partial u}{\partial n} \right) d\sigma$ $\vec{n}$ – внешняя нормаль к поверхности $S$, а в интеграле $\iint\limits_{\Sigma_\varepsilon^{M_0}} \left( u\dfrac{\partial v}{\partial n} - v\dfrac{\partial u}{\partial n} \right) d\sigma$ $\vec{n}$ – внутренняя нормаль к поверхности сферы $\Sigma_\varepsilon^{M_0}$. Учитывая, что $v(P)\Big|_{P\in\Sigma_\varepsilon^{M_0}} = \dfrac{1}{\varepsilon}$, $\dfrac{\partial v(P)}{\partial n}\Bigg|_{P\in\Sigma_\varepsilon^{M_0}} = -\dfrac{\partial}{\partial r}\left(\dfrac{1}{r}\right)\Bigg|_{r=\varepsilon} = \dfrac{1}{\varepsilon^2}$, вычислим интеграл по сфере:

$$\iint\limits_{\Sigma_\varepsilon^{M_0}} \left( u(P)\frac{\partial v(P,M_0)}{\partial n_P} - v(P,M_0)\frac{\partial u(P)}{\partial n_P} \right) d\sigma_P =$$

$$= \iint\limits_{\Sigma_\varepsilon^{M_0}} \left( u(P)\frac{1}{\varepsilon^2} - \frac{1}{\varepsilon}\frac{\partial u(P)}{\partial n} \right) d\sigma_P = *.$$

Воспользуемся теоремой о среднем:

$$* = \left( \frac{1}{\varepsilon^2}u(P^*) - \frac{1}{\varepsilon}\frac{\partial u(P^{**})}{\partial n} \right) 4\pi\varepsilon^2 =$$

$$= 4\pi\, u(P^*) - 4\pi\varepsilon\frac{\partial u(P^{**})}{\partial n} = **,$$

где точки $P^*$ и $P^{**}$ принадлежат сфере $\Sigma_\varepsilon^{M_0}$. Перейдем к пределу при $\varepsilon \to 0$, тогда сфера $\Sigma_\varepsilon^{M_0}$ и шар $K_\varepsilon^{M_0}$ будут стремиться к точке $M_0$, значения $u(P^*)$ и $\dfrac{\partial u(P^{**})}{\partial n}$ будут стремиться к $u(M_0)$ и $\dfrac{\partial u(M_0)}{\partial n}$; в силу ограниченности $\dfrac{\partial u}{\partial n}$ и непрерывности функции $u$ получаем $** = 4\pi\, u(M_0)$. Таким образом, формула (4) примет вид

$$\iiint\limits_\Omega \frac{1}{R_{MM_0}} \Delta u \, d\tau_M = 4\pi\, u(M_0) +$$

$$+\iint_S\left(u(P)\frac{\partial}{\partial n_P}\left(\frac{1}{R_{PM_0}}\right)-\frac{1}{R_{PM_0}}\frac{\partial u(P)}{\partial n_P}\right)d\sigma_P$$

или

$$u(M_0)=\frac{1}{4\pi}\iint_S\left(\frac{1}{R_{PM_0}}\frac{\partial u(P)}{\partial n_P}-u(P)\frac{\partial}{\partial n_P}\left(\frac{1}{R_{PM_0}}\right)\right)d\sigma_P-$$

$$-\frac{1}{4\pi}\iiint_\Omega\frac{\Delta u(M)}{R_{MM_0}}d\tau_M\ , \tag{5}$$

где $M_0$ – **внутренняя** точка области $\Omega$. Формула (5) называется **третьей формулой Грина** или интегральным представлением функции $u(M)$.

**Замечание.** Если $M_0\in\mathrm{R}^3\setminus\overline{\Omega}$, то точки $M$ и $M_0$ не могут совпасть. Тогда правая часть формулы (5) равна нулю. Если $M_0$ принадлежит гладкой границе области $\overline{\Omega}$, то вырезать эту точку можно сферическим куполом, в пределе это будет полусфера, а её поверхность будет равна $2\pi\varepsilon^2$. Тогда формулу (5) можно переписать так:

$$\frac{1}{4\pi}\iint_S\left(\frac{1}{R_{PM_0}}\frac{\partial u(P)}{\partial n_P}-u(P)\frac{\partial}{\partial n_P}\left(\frac{1}{R_{PM_0}}\right)\right)d\sigma_P-$$

$$-\frac{1}{4\pi}\iiint_\Omega\frac{\Delta u(M)}{R_{MM_0}}d\tau_M=\begin{cases}u(M_0), & M_0\in\Omega,\\ 0\quad, & M_0\in\mathrm{R}^3\setminus\overline{\Omega},\\ \dfrac{u(M_0)}{2}, & M_0\in S,\\ S-\textit{гладкая в точке } M_0\,.\end{cases}$$

4. **Формулы Грина** в области $\Omega\subset\mathrm{R}^2$.

Пусть теперь $\Omega$ – ограниченная область на плоскости, граница $S$ которой является достаточно гладкой кривой. В **интегральной формуле Грина**

$$\iint\limits_{\Omega}\left(\frac{\partial G}{\partial x}+\frac{\partial H}{\partial y}\right)dxdy=\oint\limits_{S}Gdy-Hdx$$

выберем $G=v(x,y)\cdot\dfrac{\partial u(x,y)}{\partial x}$, $H=v(x,y)\cdot\dfrac{\partial u(x,y)}{\partial y}$, где $u,v\in C^1(\overline{\Omega})\cap C^2(\Omega)$.

Тогда получим **первую формулу Грина**

$$\iint\limits_{\Omega}v\Delta u dxdy=\oint\limits_{S}v\frac{\partial u}{\partial n}dl-\iint\limits_{\Omega}(grad u, grad v)dxdy. \qquad (6)$$

Меняя ролями $u$ и $v$ в (6), немедленно получаем **вторую формулу Грина**

$$\iint\limits_{\Omega}(v\Delta u-u\Delta v)dxdy=\oint\limits_{S}\left(v\frac{\partial u}{\partial n}-u\frac{\partial v}{\partial n}\right)dl. \qquad (7)$$

Точно так же, как в случае трёх независимых переменных, выводим из (7) **третью (основную) формулу Грина:**

$$\frac{1}{2\pi}\oint\limits_{S}\left(\ln\frac{1}{R_{PM_0}}\cdot\frac{\partial u(P)}{\partial n_P}-u(P)\cdot\frac{\partial}{\partial n_P}\left(\ln\frac{1}{R_{PM_0}}\right)\right)dl_P-$$

$$-\frac{1}{2\pi}\iint\limits_{\Omega}\Delta u(M)\cdot\ln\frac{1}{R_{MM_0}}\cdot dxdy=\begin{cases}u(M_0), & M_0\in\Omega,\\ 0, & M_0\in \mathrm{R}^2\setminus\overline{\Omega},\\ \dfrac{u(M_0)}{2}, & M_0\in S,\\ S-\textit{гладкая в точке } M_0;\end{cases}$$

здесь $M=M(x,y)$, $M_0=M_0(x_0,y_0)$ – фиксированная точка, $R_{MM_0}$ – расстояние между $M_0$ и $M$, $\ln\dfrac{1}{R_{MM_0}}$ – фундаментальное решение уравнения Лапласа на плоскости, $u\in C^1(\overline{\Omega})\cap C^2(\Omega)$.

5. **Следствия из формул Грина.**

Из формул Грина вытекают основные свойства гармонических функций.

$1^0$. Если функция $u(M)$ гармонична в ограниченной области $\Omega \subset \mathrm{R}^3$ с достаточно гладкой границей $S$, то $\iint\limits_S \frac{\partial u}{\partial n} d\sigma = 0$. $\left( \oint\limits_S \frac{\partial u}{\partial n} dl = 0 \right.$ в случае $\left. \Omega \subset \mathrm{R}^2. \right)$

$2^0$. Формула среднего значения.

Если $u(M)$ – гармоническая функция в области $\Omega \subset \mathrm{R}^3$, то для любой точки $M \in \Omega$ имеет место представление $u(M) = \frac{1}{4\pi a^2} \iint\limits_{\Sigma_a^M} u(P) d\sigma_P$, где $\Sigma_a^M$ – сфера радиуса $a$ с центром в точке $M$, целиком лежащая в $\Omega$, т.е. $\Sigma_a^M \subset \Omega$.

В случае гармонической функции двух переменных $u(M) = \frac{1}{2\pi a} \int\limits_{L_a^M} u(P) dl_P$, где $L_a^M$ – окружность радиуса $a$ с центром в точке $M$, целиком лежащая в области гармоничности функции $u(M)$.

$3^0$. Гармоническая функция бесконечно дифференцируема.

$4^0$. Принцип максимума (минимума) гармонической функции:

**Теорема.** Пусть $\Omega$ – ограниченная область с границей $S$. Пусть функция $u(M)$ гармонична в $\Omega$ и непрерывна в $\overline{\Omega}$. Тогда она достигает своего максимального и минимального значений на границе области $\overline{\Omega}$, т.е. $\max\limits_{M \in \overline{\Omega}} u(M) = \max\limits_{M \in S} u(M)$, $\min\limits_{M \in \overline{\Omega}} u(M) = \min\limits_{M \in S} u(M)$.

6. **Функция Грина для внутренней задачи Дирихле** в области $\Omega \subset \mathrm{R}^3$.

Пусть $\Omega$ – ограниченная область в $\mathrm{R}^3$ с границей $S$. Внутренняя (т.е. в **ограниченной** области $\Omega$) задача Дирихле для уравнения Пуассона состоит в нахождении функции $u(M)$, удовлетворяющей уравнению

$$\Delta u(M) = -f(M)\ ,\ M \in \Omega\ ,$$

и краевому условию первого рода

$$u(P) = \mu(P),\ \ P \in S\ .$$

Решение этой задачи будем предполагать классическим: $u \in C(\overline{\Omega}) \cap C^2(\Omega)$. (Мы не требуем $u \in C^1(\overline{\Omega})$, поэтому для применения формул Грина нужны дополнительные предположения о поверхности $S$, которые гарантируют выполнение формул Грина для функций $u$, $v$ из класса $C(\overline{\Omega}) \cap C^2(\Omega)$, имеющих правильные нормальные производные на $S$ и $\Delta u, \Delta v \in L_2(\Omega)$.)
Запишем интегральное представление решения во внутренней точке $M_0$ области $\Omega$:

$$u(M_0) = \frac{1}{4\pi}\iint\limits_S \left\{\frac{\partial u(P)}{\partial n_P}\frac{1}{R_{M_0P}} - u(P)\frac{\partial}{\partial n_P}\left(\frac{1}{R_{M_0P}}\right)\right\} d\sigma_P -$$

$$-\frac{1}{4\pi}\iiint\limits_\Omega \frac{\Delta u(M)}{R_{M_0M}} d\tau_M\ . \tag{9}$$

Пусть функция $v(M)$ гармонична в области $\Omega$ и непрерывна вместе с первыми производными в замкнутой области $\overline{\Omega}$. Применим к функциям $u(M)$ и $v(M)$ вторую формулу Грина:

$$0 = \iint\limits_S \left\{v(P)\frac{\partial u(P)}{\partial n_P} - u(P)\frac{\partial v(P)}{\partial n_P}\right\} d\sigma_P -$$

$$-\iiint\limits_\Omega v(M)\Delta u(M) d\tau_M\ . \tag{10}$$

Сложим формулы (9) и (10), получим

$$u(M_0) = \iint\limits_S \left\{\frac{\partial u(P)}{\partial n_P} G(M_0, P) - u(P)\frac{\partial G(M_0, P)}{\partial n_P}\right\} d\sigma_P -$$

$$-\iiint\limits_\Omega G(M_0, M)\Delta u(M) d\tau_M\ ,$$

где $G(M_0, M) = \dfrac{1}{4\pi R_{M_0 M}} + v(M)$.

Если потребовать, чтобы выполнялось условие $G(M_0, P) = 0, \quad P \in S$, то

$$u(M_0) = -\iint_S u(P) \frac{\partial G(M_0, P)}{\partial n_P} d\sigma_P -$$

$$- \iiint_\Omega \Delta u(M) G(M_0, M) d\tau_M . \tag{11}$$

**Определение.** Функция $G(M_0, M)$ называется функцией Грина внутренней задачи Дирихле для оператора Лапласа в $\Omega \subset \mathrm{R}^3$, если

1. $G(M_0, M) = \dfrac{1}{4\pi R_{M_0 M}} + v(M)$, где функция $v(M)$ гармонична всюду в области $\Omega$;
2. $G(M_0, P) = 0, \quad P \in S$. □

Из определения следует, что функция Грина $G(M_0, M)$ с точностью до гармонической функции $v$ является фундаментальным решением уравнения Лапласа. Второе требование в определении продиктовано типом граничных условий. Если функция Грина $G(M_0, M)$ существует, то решение задачи Дирихле для уравнения Пуассона находится в явном виде по формуле

$$u(M_0) = -\iint_S \mu(P) \frac{\partial G(M_0, P)}{\partial n_P} d\sigma_P +$$

$$+ \iiint_\Omega f(M) G(M_0, M) d\tau_M . \tag{12}$$

**Замечание.** Формула (12) содержит производную $\dfrac{\partial G(M_0, M)}{\partial n_M}$, существование которой не следует из определения.

Формула (12) даёт классическое решение задачи Дирихле при выполнении условий $\mu \in C(S)$ и $f \in C^1(\overline{\Omega})$.

Для построения функции Грина $G(M_0, M)$ необходимо найти функцию $v(M) = v(M, M_0)$ ($M_0$ – параметр), удовлетворяющую задаче

$$\Delta v(M) = 0, \quad M \in \Omega, \tag{13}$$

$$v(P) = -\frac{1}{4\pi R_{M_0 P}}, \quad P \in S. \tag{14}$$

Для достаточно широкого класса поверхностей (так называемых поверхностей Ляпунова) задача (13), (14) разрешима, поэтому функция Грина существует. ▯

7. **Свойства функции Грина.**

**Теорема.** Если $M, M_0 \in \Omega$, то $G(M_0, M) > 0$.

**Доказательство.** Точку $M_0$, в которой функция Грина имеет особенность, вырежем шаром $K_\varepsilon$ с центром $M_0$, ограниченным сферой $\Sigma_\varepsilon$. Из представления функции Грина следует, что на поверхности $\Sigma_\varepsilon$ и внутри неё $G(M_0, M) > 0$ и $G(M_0, P) = 0$, $P \in S$. Так как функция Грина $G(M_0, M)$ в области, заключенной между поверхностями $\Sigma_\varepsilon$ и $S$, удовлетворяет уравнению Лапласа, то в силу принципа максимума $G(M_0, M) > 0$ всюду в области, заключенной между поверхностями $\Sigma_\varepsilon$ и $S$. В силу произвольности выбора $\varepsilon$ получаем $G(M_0, M) > 0$ всюду в области $\Omega$. ▯

**Замечание.** Так как функция $v(M)$ удовлетворяет задаче (13), (14), то из принципа максимума следует, что $v(M) < 0$ всюду в области $\Omega$. Следовательно, $G(M_0, M) = \dfrac{1}{4\pi R_{M_0 M}} + v(M) < \dfrac{1}{4\pi R_{M_0 M}}$.

Отсюда получаем границы изменения значений функции $G(M_0, M)$:

$$0 \le G(M_0, M) < \frac{1}{4\pi R_{M_0 M}}, \quad M_0 \ne M,$$

причём равенство нулю учитывает, что $G(M_0, P) = 0, \quad P \in S$. □

**Теорема.** Функция Грина симметрична относительно точек $M_0$ и $M$: $G(M_0, M) = G(M, M_0)$. (Это свойство называется принципом взаимности.)

**Доказательство.** Пусть $M_1$ и $M_2$ – некоторые фиксированные точки области $\Omega$. Построим сферы $\Sigma_\varepsilon^{M_1}$ и $\Sigma_\varepsilon^{M_2}$ радиусов $\varepsilon$ с центрами в точках $M_1$ и $M_2$. Введём функции $w_1(M) = G(M, M_1)$ и $w_2(M) = G(M, M_2)$ и применим к ним вторую формулу Грина в области $\Omega \setminus \left(K_\varepsilon^{M_1} \cup K_\varepsilon^{M_2}\right)$, где $K_\varepsilon^{M_1}$ и $K_\varepsilon^{M_2}$ – шары радиусов $\varepsilon$ с центрами в точках $M_1$ и $M_2$:

$$\iiint\limits_{\Omega \setminus \left(K_\varepsilon^{M_1} \cup K_\varepsilon^{M_2}\right)} \left(w_1(M)\Delta w_2(M) - w_2(M)\Delta w_1(M)\right) d\tau_M =$$

$$= \iint\limits_S \left( w_1(P)\frac{\partial w_2(P)}{\partial n_P} - w_2(P)\frac{\partial w_1(P)}{\partial n_P} \right) d\sigma_P +$$

$$+ \iint\limits_{\Sigma_\varepsilon^{M_1}} \left( w_1(P)\frac{\partial w_2(P)}{\partial n_P} - w_2(P)\frac{\partial w_1(P)}{\partial n_P} \right) d\sigma_P +$$

$$+ \iint\limits_{\Sigma_\varepsilon^{M_2}} \left( w_1(P)\frac{\partial w_2(P)}{\partial n_P} - w_2(P)\frac{\partial w_1(P)}{\partial n_P} \right) d\sigma_P .$$

Учитывая, что функции $w_1(M)$ и $w_2(M)$ удовлетворяют уравнению Лапласа в $\Omega \setminus M_1$ и в $\Omega \setminus M_2$ соответственно, а на границе $S$ области $\Omega$ принимают нулевые значения, получим

$$\iint\limits_{\Sigma_\varepsilon^{M_1}} \left( w_1(P)\frac{\partial w_2(P)}{\partial n_P} - w_2(P)\frac{\partial w_1(P)}{\partial n_P} \right) d\sigma_P +$$

$$+ \iint\limits_{\Sigma_\varepsilon^{M_2}} \left( w_1(P)\frac{\partial w_2(P)}{\partial n_P} - w_2(P)\frac{\partial w_1(P)}{\partial n_P} \right) d\sigma_P = 0,$$

где нормали $\vec{n}_P$ к поверхностям $\Sigma_\varepsilon^{M_1}$ и $\Sigma_\varepsilon^{M_2}$ направлены внутрь шаров $K_\varepsilon^{M_1}$ и $K_\varepsilon^{M_2}$.

$$\iint\limits_{\Sigma_\varepsilon^{M_1}}\left(G(P,M_1)\frac{\partial G(P,M_2)}{\partial n_P}-G(P,M_2)\frac{\partial G(P,M_1)}{\partial n_P}\right)d\sigma_P+ \tag{15}$$

$$+\iint\limits_{\Sigma_\varepsilon^{M_2}}\left(G(P,M_1)\frac{\partial G(P,M_2)}{\partial n_P}-G(P,M_2)\frac{\partial G(P,M_1)}{\partial n_P}\right)d\sigma_P=0.$$

Функции $w_1(M)$ и $w_2(M)$ внутри шаров $K_\varepsilon^{M_2}$ и $K_\varepsilon^{M_1}$ соответственно удовлетворяют уравнению Лапласа. Напишем интегральное представление функции $w_1(M)$ в точке $M_2$ и функции $w_2(M)$ в точке $M_1$:

$$w_1(M_2)=\iint\limits_{\Sigma_\varepsilon^{M_2}}\left(w_1(P)\frac{\partial G(P,M_2)}{\partial n_P}-G(P,M_2)\frac{\partial w_1(P)}{\partial n_P}\right)d\sigma_P=$$

$$=\iint\limits_{\Sigma_\varepsilon^{M_2}}\left(G(P,M_1)\frac{\partial G(P,M_2)}{\partial n_P}-G(P,M_2)\frac{\partial G(P,M_1)}{\partial n_P}\right)d\sigma_P.$$

$$w_2(M_1)=\iint\limits_{\Sigma_\varepsilon^{M_1}}\left(w_2(P)\frac{\partial G(P,M_1)}{\partial n_P}-G(P,M_1)\frac{\partial w_2(P)}{\partial n_P}\right)d\sigma_P=$$

$$=\iint\limits_{\Sigma_\varepsilon^{M_1}}\left(G(P,M_2)\frac{\partial G(P,M_1)}{\partial n_P}-G(P,M_1)\frac{\partial G(P,M_2)}{\partial n_P}\right)d\sigma_P.$$

Сравнивая, полученные выражения для $w_1(M_2)$ и $w_2(M_1)$ с формулой (15), получим $w_1(M_2)-w_2(M_1)=0$ или $G(M_1,M_2)=G(M_2,M_1)$. ▯

8. **Построение функции Грина** методом зеркальных изображений.

Функция Грина задачи Дирихле $G(M,M_0)$ допускает различные физические интерпретации. Мы примем электростатическую интерпретацию. Пусть поверхность $S$, ограничивающая область $\Omega \subset R^3$, сделана из идеального проводника и заземлена. Поместим в точке $M_0$ внутри $\Omega$ электрический заряд величины $\dfrac{1}{4\pi}$. Этот заряд индуцирует распределение зарядов на $S$. Поэтому потенциал электростатического поля в области $\Omega$ равен сумме потенциала $\dfrac{1}{4\pi R_{MM_0}}$ поля точечного заряда и потенциала $v(M, M_0)$ поля индуцированных зарядов. Эта сумма и равна $G(M, M_0)$.

Если мысленно убрать индуцированные на $S$ заряды, то для сохранения прежнего потенциала $G$ в области $\Omega$ придется разместить некоторые точечные заряды вне поверхности $S$, которые в $\Omega$ создадут поле с потенциалом $v$. Эти заряды являются зеркальными относительно $S$ изображениями заряда, помещенного в точку $M_0$. Такой прием позволяет для областей простой геометрической формы найти $v(M, M_0)$ и построить функцию Грина.

9. **Функция Грина для внутренней задачи Дирихле** в области $\Omega \subset R^2$.

Отличие случая двух независимых переменных от только что рассмотренного состоит в характере особенности фундаментального решения уравнения Лапласа при $M = M_0$:

**Определение.** Функция $G(M_0, M)$ называется функцией Грина внутренней задачи Дирихле для оператора Лапласа в $\Omega \subset R^2$, если

1. $G(M_0, M) = \dfrac{1}{2\pi} \ln\left(\dfrac{1}{R_{M_0M}}\right) + v(M)$, где функция $v(M)$ гармонична всюду в области $\Omega$;
2. $G(M_0, P) = 0$, $P \in S$. □

В случае двух независимых переменных для построения функции Грина в **односвязной** области $\Omega$ можно использовать конформные отображения:

**Теорема.** Пусть точке $M_0(x_0, y_0)$ отвечает точка комплексной $z$ - плоскости $z_0 = x_0 + iy_0$, а точке $M(x, y)$ – точка $z = x + iy$. Если функция $w = \varphi(z; z_0)$ конформно отображает односвязную область $\Omega$ на $\{w \mid |w| < 1\}$ так, что произвольная фиксированная точка $z_0 \in \Omega$ переводится в центр единичного круга, то $G(M, M_0) = \dfrac{1}{2\pi} \ln \dfrac{1}{|\varphi(z; z_0)|}$.

## Вопрос 23.

## Примеры и канонический вид одношаговых итерационных методов решения систем линейных алгебраических уравнений.

**1. Канонический вид одношаговых итерационных методов.** Дана система линейных алгебраических уравнений

$$Ay = f \qquad (1)$$

с квадратной невырожденной матрицей $A = \{a_{ij}\}$ порядка $m$, где $y = (y_1, y_2, \ldots, y_m)^T$, $f = (f_1, f_2, \ldots, f_m)^T$. **Итерационные методы** состоят в том, что решение $y$ системы (1) находится как предел при $n \to \infty$ последовательных приближений $y_n$, где $y_n = (y_1^n, y_2^n, \ldots, y_m^n)^T$ и $n$ - номер итерации. **Одношаговым** итерационным методом называется такой итерационный метод, в котором для нахождения $y_{n+1}$ используется только одна предыдущая итерация $y_n$.

**Замечание.** Для решения систем уравнений можно применять и **многошаговые** итерационные методы, в которых $y_{n+1}$ определяется через значения $y_j$ на двух и более предыдущих итерациях, т. е. $y_{n+1} = f(y_n, y_{n-1}, \ldots, y_{n-k})$. При этом значения $y_n, y_{n-1}, \ldots, y_{n-k}$ приходится хранить в памяти компьютера. ▯

Будем рассматривать точное решение $y$ исходной системы и приближения $y_n$ как элементы некоторого линейного конечномерного пространства с нормой $\|y_n\|$. **Нормой матрицы** $A$, подчиненной данной норме вектора, называется число $\|A\| = \sup\limits_{x \neq 0} \frac{\|Ax\|}{\|x\|}$. Говорят, что итерационный метод **сходится**, если $\|z_n\| \to 0$ при $n \to \infty$, где $z_n = y_n - y$ - **погрешность** метода.

**Каноническим видом** одношагового итерационного метода решения системы (1) называется его запись в виде

$$B_{n+1}\frac{y_{n+1}-y_n}{\tau_{n+1}}+Ay_n=f,\ n=0,1,\ldots,n_0. \qquad (2)$$

Здесь $B_{n+1}$ - матрица, задающая тот или иной итерационный метод, $\tau_{n+1}$ - итерационный параметр. Предполагается, что существуют матрицы $B_n^{-1}, n=1,2,\ldots,n_0-1$ и задано начальное приближение $y_0$. Тогда для вычисления $y_{n+1}$ по известным $y_n$ и $f$ достаточно решить систему уравнений $B_{n+1}y_{n+1}=F_n$, где $F_n=(B_{n+1}-\tau_{n+1}A)y_n+\tau_{n+1}f$.

**Замечание.** Матрицу $B_{n+1}$ надо выбирать так, чтобы система $B_{n+1}y_{n+1}=F_n$ решалась легче, чем исходная система $Ay=f$.□

Итерационный метод называют **явным (неявным)**, если $B_n=E$ ($B_n\neq E$), где $E$ - единичная матрица. Итерационный метод называется **стационарным**, если $B_{n+1}=B$ и $\tau_{n+1}=\tau$ не зависят от номера итерации, и **нестационарным** – в противоположном случае.

**2. Метод простой итерации.** Простейшим итерационным методом решения систем линейных алгебраических уравнений является метод простой итерации

$$\frac{y_{n+1}-y_n}{\tau}+Ay_n=f\,, \qquad (3)$$

который относится к явным стационарным методам. Перепишем формулу (3) в виде

$$y_{n+1}=Sy_n+\tau f,$$

где $S=E-\tau A$ - **матрица перехода**.

**Теорема 1** (достаточное условие сходимости). Если $\|S\|<1$, то метод простой итерации сходится при любом начальном приближении.

**Доказательство.** Погрешность метода $z_n = y_n - y$ удовлетворяет уравнению $z_{n+1} = S z_n$, откуда получаем равенство $z_n = S^n z_0$. Следовательно, $\|z_n\| \le \|S\|^n \|z_0\| \to 0$ при $n \to \infty$.▯

**Теорема 2** (критерий сходимости). Пусть система (1) имеет единственное решение. Метод простой итерации сходится к решению системы (1) при любом начальном приближении тогда и только тогда, когда все собственные значения матрицы $S$ по модулю меньше 1. ▯

**3. Метод Якоби.** Представим матрицу $A$ в виде суммы трех матриц

$$A = A_1 + D + A_2,$$

где $D = diag[a_{11}, a_{22}, \ldots, a_{mm}]$ – диагональная матрица с той же главной диагональю, что и матрица $A$, матрица $A_1$ - нижняя треугольная и матрица $A_2$ - верхняя треугольная с нулевыми главными диагоналями. Тогда исходная система примет вид

$$y = -D^{-1} A_1 y - D^{-1} A_2 y + D^{-1} f$$

или в покоординатной записи

$$y_i = -\sum_{j=1}^{i-1} \frac{a_{ij}}{a_{ii}} y_j - \sum_{j=i+1}^{m} \frac{a_{ij}}{a_{ii}} y_j + \frac{f_i}{a_{ii}}, \ i = 1,2,\ldots,m. \qquad (4)$$

Здесь предполагается, что все $a_{ii}$ отличны от нуля и что значение суммы равно нулю, если верхний предел суммирования меньше нижнего.

Исходя из записи системы в виде (4) итерации в методе Якоби определяются следующим образом

$$y_i^{n+1} = -\sum_{j=1}^{i-1} \frac{a_{ij}}{a_{ii}} y_j^n - \sum_{j=i+1}^{m} \frac{a_{ij}}{a_{ii}} y_j^n + \frac{f_i}{a_{ii}}, \quad i = 1,2,\ldots,m, \ \ n = 0,1,\ldots,n_0$$

Начальные значения $y_i^0, i = 1,2,\ldots,m$ задаются произвольно. Окончание итераций определяется либо заданием максимального числа итераций $n_0$, либо условием $\max\limits_{1 \le i \le m} \left| y_i^{n+1} - y_i^n \right| < \varepsilon$, где $\varepsilon > 0$ - заданное число.

Метод Якоби можно записать в векторном виде:

$$y_{n+1} = -D^{-1}A_1 y_n - D^{-1}A_2 y_n + D^{-1} f$$

или $Dy_{n+1} + (A_1 + A_2)y_n = f$. Учитывая, что $A_1 + A_2 = A - D$, получим канонический вид (2) метода Якоби

$$D(y_{n+1} - y_n) + Ay_n = f,$$

т. е. $B_{n+1} = D, \quad \tau_{n+1} = 1$.

**Замечание.** Метод Якоби можно рассматривать как некоторый метод простой итерации $y_{n+1} = S y_n + D^{-1} f$ с матрицей $S = -D^{-1}(A_1 + A_2)$. Поэтому для исследования его сходимости можно применять теоремы 1 и 2. □

**Пример.** Рассмотрим разностную аппроксимацию уравнения Пуассона в единичном квадрате на квадратной сетке:

$$\frac{y_{i-1\,j} - 2y_{ij} + y_{i+1\,j}}{h^2} + \frac{y_{i\,j-1} - 2y_{ij} + y_{i\,j+1}}{h^2} = -f_{ij}, \tag{5}$$

$$i, j = 1, 2, \ldots, N-1, \quad hN = 1, \quad y_{ij}\Big|_{\gamma_h} = \mu_{ij},$$

где $f_{ij}$, $\mu_{ij}$ - заданные функции, $\gamma_h$ - граница сеточной области.

Реализация метода Якоби для разностной задачи (5) имеет вид

$$\frac{y^n_{i-1\,j} - 2y^{n+1}_{ij} + y^n_{i+1\,j}}{h^2} + \frac{y^n_{i\,j-1} - 2y^{n+1}_{ij} + y^n_{i\,j+1}}{h^2} = -f_{ij}, \tag{6}$$

$$i, j = 1, 2, \ldots, N-1, \quad hN = 1, \quad y^{n+1}_{ij}\Big|_{\gamma_h} = \mu_{ij},$$

причем решение на новой итерации находится явным образом:

$$y^{n+1}_{ij} = \frac{1}{4}\left(y^n_{i-1\,j} + y^n_{i+1\,j} + y^n_{i\,j-1} + y^n_{i\,j+1}\right) + \frac{h^2}{4} f_{ij}. \ \square$$

**4. Метод Зейделя.** Метод Зейделя строится таким образом, чтобы в уравнении с номером $i$ неизвестные, имеющие индекс больший, чем $i$, вычислялись бы по значениям на $n$-ой итерации, а остальные неизвестные – по значениям на новой $n+1$-ой итерации.

В координатной форме метод Зейделя записывается так:

$$y^{n+1}_i = -\sum_{j=1}^{i-1} \frac{a_{ij}}{a_{ii}} y^{n+1}_j - \sum_{j=i+1}^{m} \frac{a_{ij}}{a_{ii}} y^n_j + \frac{f_i}{a_{ii}},$$

$$i = 1,2,\ldots,m,\ \ n = 0,1,\ldots,n_0.$$

Для понимания того, как реализуется данный метод, выпишем подробнее первые два уравнения:

$$y_1^{n+1} = -\sum_{j=2}^{m} \frac{a_{1j}}{a_{11}} y_j^n + \frac{f_1}{a_{11}},$$

$$y_2^{n+1} = -\frac{a_{21}}{a_{22}} y_1^{n+1} - \sum_{j=3}^{m} \frac{a_{2j}}{a_{22}} y_j^n + \frac{f_2}{a_{22}}.$$

Зная $f_1$ и компоненты $y_j^n, j = 2,3,\ldots,m$ с предыдущей итерации, из первого уравнения находят $y_1^{n+1}$. Далее, во второе уравнение подставляют известные значения $y_j^n, j = 3,\ldots,m$ и только что найденное значение $y_1^{n+1}$ и вычисляют $y_2^{n+1}$. Таким образом, компоненты вектора $y_{n+1}$ находятся последовательно, начиная с $i = 1$.

В матричной форме метод Зейделя имеет вид

$$y_{n+1} = -D^{-1}A_1 y_{n+1} - D^{-1}A_2 y_n + D^{-1} f.$$

После элементарных преобразований отсюда получим канонический вид (2) данного метода

$$(D + A_1)(y_{n+1} - y_n) + Ay_n = f.$$

**Замечание.** Можно записать метод Зейделя в виде

$$y_{n+1} = -(D + A_1)^{-1} A_2 y_n + (D + A_1)^{-1} f.$$

Обозначая $S = -(D + A_1)^{-1} A_2$, получим, что метод Зейделя эквивалентен некоторому методу простой итерации. Поэтому для его сходимости необходимо и достаточно, чтобы все собственные значения матрицы $S = -(D + A_1)^{-1} A_2$ по модулю были меньше 1.

**Пример.** Рассмотрим реализацию метода Зейделя для разностной задачи (5):

$$\frac{y_{i-1j}^{n+1} - 2y_{ij}^{n+1} + y_{i+1j}^{n}}{h^2} + \frac{y_{ij-1}^{n+1} - 2y_{ij}^{n+1} + y_{ij+1}^{n}}{h^2} = -f_{ij}, \quad (7)$$

$$i,j = 1,2,\ldots,N-1, \quad hN = 1, \quad y_{ij}^{n+1}\Big|_{\gamma_h} = \mu_{ij}.$$

Несмотря на то, что метод Зейделя является неявным, найти значения $y_{ij}^{n+1}$ на новой итерации несложно. Зная граничные значения $y_{01}^{n+1} = \mu_{01}$ и $y_{10}^{n+1} = \mu_{10}$ из уравнения (7) сначала находят $y_{11}^{n+1}$. Теперь можно найти $y_{12}^{n+1}$ и т. д. Таким образом, неизвестные вычисляются в следующем порядке изменения индексов: $(1,1),\ (1,2),\ldots,(1,N-1),(2,1),(2,2),\ldots,(2,N-1),\ldots,(N-1,1),$ $(N-1,2),\ldots,(N-1,N-1)$. ▯

**Пример.** В отличие от метода Якоби, итерации в методе Зейделя существенно зависят от порядка уравнений, причем при перестановке уравнений метод может стать расходящимся. Исследуем сходимость метода Зейделя для системы с матрицей

$$A = \begin{pmatrix} a & 1 \\ 1 & a \end{pmatrix}, \quad |a| > 1.$$

Представим матрицу $A$ в виде суммы трех матриц $A = A_1 + D + A_2$, где

$$A_1 = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \quad D = \begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix}, \quad A_2 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix},$$

и найдем матрицу $S = -(D + A_1)^{-1} A_2 = \dfrac{1}{a^2}\begin{pmatrix} 0 & -a \\ 0 & 1 \end{pmatrix}$.

Собственными значениями этой матрицы будут числа $\lambda_1 = 0$ и $\lambda_2 = 1/(a^2) < 1$. Следовательно, условия теоремы 2 выполнены, и метод Зейделя сходится.

Переставляя в матрице $A$ строки, получим матрицу

$$A' = \begin{pmatrix} 1 & a \\ a & 1 \end{pmatrix}, \quad |a| > 1.$$

Нетрудно найти собственные значения соответствующей матрицы перехода: $\lambda_1' = 0$ и $\lambda_2' = a^2 > 1$, т. е. метод Зейделя будет расходиться. ▯

**5. Задача.** Для системы $\begin{cases} y_1 + y_2 + 2y_3 = -1, \\ 2y_1 - y_2 + 2y_3 = 3, \\ 4y_1 + y_2 + 4y_3 = -3 \end{cases}$ записать метод Зейделя и найти $y_n = (y_1^n, y_2^n, y_3^n)^T$, $n = 1,2,3$, если начальное приближение равно нулю. ▯

**Задача.** Проведя исключение вспомогательного приближения $\overline{y}_n$, привести итерационный метод

$$(E + \omega_1 P)\overline{y}_n = (E - \omega_1 Q)y_n + \omega_1 f,$$
$$(E + \omega_2 Q)y_{n+1} = (E - \omega_2 P)\overline{y}_n + \omega_2 f$$

к каноническому виду (2). Здесь $\omega_1, \omega_2$ – параметры и $Q, P$ – некоторые матрицы. ▯

**Задача.** При каких $\alpha, \beta$ сходится метод простой итерации $y_{n+1} = Sy_n + \tau f$, где $S = \begin{pmatrix} \alpha & \beta & 0 \\ \beta & \alpha & \beta \\ 0 & \beta & \alpha \end{pmatrix}$. ▯

## Вопрос 24.

## Теорема о сходимости итерационного метода для систем с симметрической положительно определенной матрицей.

**1.** Рассмотрим систему линейных алгебраических уравнений $Ay = f$ с квадратной невырожденной действительной матрицей $A = \{a_{ij}\}$ порядка $m$, где $y = (y_1, y_2, \dots, y_m)^T$, $f = (f_1, f_2, \dots, f_m)^T$. Пусть в пространстве $R^m$ введены скалярное произведение $(x, y) = \sum_{i=1}^{m} x_i y_i$ и норма $\|x\| = \left( \sum_{i=1}^{m} x_i^2 \right)^{1/2}$.

Для решения системы выберем одношаговый **стационарный** итерационный метод, записанный в **каноническом виде**

$$B \frac{y_{n+1} - y_n}{\tau} + Ay_n = f, \quad n = 0,1,\dots, \; y_0 \text{ задан.} \qquad (1)$$

Здесь $n$ - номер итерации и $y_n$ – это вектор, полученный в результате $n$-ой итерации.

**Погрешностью** метода на $n$-ой итерации называется вектор $z_n = y_n - y$. Говорят, что итерационный метод **сходится**, если $\|z_n\| \to 0$ при $n \to \infty$.

**Теорема Самарского.** Пусть $A$ - симметрическая положительно определенная матрица, $\tau > 0$, и пусть выполнено неравенство

$$B - 0.5\tau A > 0. \qquad (2)$$

Тогда итерационный метод сходится.

**Доказательство.** Погрешность $z_n$ удовлетворяет однородному уравнению

$$B \frac{z_{n+1} - z_n}{\tau} + Az_n = 0, \quad n = 0,1,\dots, \qquad (3)$$

поэтому $z_{n+1} = \left(E - \tau B^{-1} A\right) z_n$, $\quad Az_{n+1} = \left(A - \tau AB^{-1} A\right) z_n$. Тогда

$$(Az_{n+1}, z_{n+1}) = (Az_n, z_n) - \tau(AB^{-1}Az_n, z_n) - \tau(Az_n, B^{-1}Az_n) + \\ + \tau^2\left(AB^{-1}Az_n, B^{-1}Az_n\right).$$

Так как матрица $A$ – симметрическая, то

$$(AB^{-1}Az_n, z_n) = (B^{-1}Az_n, A^T z_n) = (B^{-1}Az_n, Az_n) = \\ = (Az_n, B^{-1}Az_n).$$

Поэтому

$$(Az_{n+1}, z_{n+1}) = (Az_n, z_n) - 2\tau(Az_n, B^{-1}Az_n) + \\ + \tau^2(AB^{-1}Az_n, B^{-1}Az_n) = (Az_n, z_n) - \\ - 2\tau((B - 0.5\tau A)B^{-1}Az_n, B^{-1}Az_n).$$

Учитывая условие (2), получаем неравенство $(Az_{n+1}, z_{n+1}) \le (Az_n, z_n)$. Следовательно, последовательность $(Az_n, z_n)$ монотонна и ограничена снизу нулем. Поэтому существует $\lim\limits_{n\to\infty}(Az_n, z_n) = P$.

По условию теоремы матрица $B - 0.5\tau A$ является положительно определенной. Значит, существует константа $\delta > 0$ такая, что $((B - 0.5\tau A)B^{-1}Az_n, B^{-1}Az_n) \ge \delta \left\| B^{-1}Az_n \right\|^2$.

Поэтому $(Az_{n+1}, z_{n+1}) \le (Az_n, z_n) - 2\delta\tau \left\| B^{-1}Az_n \right\|^2$ или $(Az_{n+1}, z_{n+1}) - (Az_n, z_n) + 2\delta\tau \left\| B^{-1}Az_n \right\|^2 \le 0$. Перейдем в полученном неравенстве к пределу при $n \to \infty$ и учтем, что $\lim\limits_{n\to\infty}(Az_n, z_n) = P$. Тогда $\lim\limits_{n\to\infty}\| w_n \| = 0$, где $w_n = B^{-1}Az_n$.

Далее, из положительной определенности матрицы $A$ следует ее обратимость, поэтому $z_n = A^{-1}Bw_n$ и $\| z_n \| \le \| A^{-1}B \| \| w_n \|$. Таким образом, $\lim\limits_{n\to\infty}\| z_n \| = \lim\limits_{n\to\infty}\| w_n \| = 0$. Следовательно, итерационный метод сходится. ▯

**Следствие 1.** Пусть $A$ - симметрическая положительно определенная матрица с **диагональным преобладанием**, то есть

$$a_{ii} > \sum_{j \neq i} \left| a_{ij} \right|, \, i = 1, 2, \ldots, m. \qquad (4)$$

Тогда метод Якоби $D(y_{n+1} - y_n) + Ay_n = f$, где $D = diag\left[a_{11}, a_{22}, \ldots, a_{mm}\right]$ сходится.

**Доказательство.** Метод Якоби имеет канонический вид (1), где $\tau = 1, B = D$. Условие сходимости (2) принимает вид $A < 2D$. Покажем, что это неравенство следует из условия (4). Применяя неравенство $y_i y_j \le \dfrac{y_i^2 + y_j^2}{2}$ к каждому слагаемому положительно определенной квадратичной формы $(Ay, y) = \sum\limits_{i,j=1}^{m} a_{ij} y_i y_j$, получим оценку

$$(Ay, y) \le \frac{1}{2} \sum_{i,j=1}^{m} |\, a_{ij} \,|\, y_i^2 + \frac{1}{2} \sum_{i,j=1}^{m} |\, a_{ij} \,|\, y_j^2 =$$

$$= \frac{1}{2} \sum_{i,j=1}^{m} |\, a_{ij} \,|\, y_i^2 + \frac{1}{2} \sum_{i,j=1}^{m} |\, a_{ji} \,|\, y_i^2 .$$

Так как $A$ - симметрическая и положительно определенная матрица, то $a_{ij} = a_{ji}, a_{ii} > 0, i, j = 1, 2, \ldots, m.$ Следовательно,

$$(Ay, y) \le \frac{1}{2} \sum_{i,j=1}^{m} |\, a_{ij} \,|\, y_i^2 + \frac{1}{2} \sum_{i,j=1}^{m} |\, a_{ij} \,|\, y_i^2 = \sum_{i,j=1}^{m} |\, a_{ij} \,|\, y_i^2 =$$

$$= \sum_{i=1}^{m} y_i^2 \left( \sum_{j \neq i} |a_{ij}| + a_{ii} \right).$$

Заметим, что условие диагонального преобладания (4) можно записать в виде $a_{ii} + \sum\limits_{j \neq i} \left| a_{ij} \right| < 2a_{ii}, \, i = 1, 2, \ldots, m.$

Тогда $(Ay, y) < 2 \sum\limits_{i=1}^{m} a_{ii} y_i^2 = 2(Dy, y)$ или, в матричной записи, $A < 2D$. ▯

**Следствие 2.** Пусть $A$ - симметрическая положительно определенная матрица. Тогда метод верхней релаксации $(D+\omega A_1)\dfrac{y_{n+1}-y_n}{\omega}+Ay_n=f$ сходится при условии $0<\omega<2$. В частности, метод Зейделя ($\omega=1$) сходится.

**Доказательство.** При использовании метода верхней релаксации исходную матрицу $A$ представляют в виде суммы трех матриц $A=A_1+D+A_2$, где $D=diag[a_{11},a_{22},\ldots,a_{mm}]$ – диагональная матрица с той же главной диагональю, что и матрица $A$, матрица $A_1$ - нижняя треугольная и матрица $A_2$ - верхняя треугольная с нулевыми главными диагоналями. Вследствие симметричности исходной матрицы $A_2^T=A_1$, поэтому

$$(Ay,y)=(A_1y,y)+(Dy,y)+(A_2y,y)=2(A_1y,y)+(Dy,y)\ .$$

Метод верхней релаксации можно записать в каноническом виде (1) с $B=D+\omega A_1,\tau=\omega$. Тогда условия сходимости $\tau>0$ и $B-0.5\tau A>0$ примут вид $\omega>0$ и

$$(By,y)-0.5\omega(Ay,y)=((D+\omega A_1)y,y)-0.5\omega(2(A_1y,y)+(Dy,y))=(1-0.5\omega)(Dy,y)>0.$$

Видно, что это условие будет выполняться при $\omega<2$.□

**Следствие 3.** Пусть $A$ - симметрическая положительно определенная матрица. Тогда метод простой итерации $\dfrac{y_{n+1}-y_n}{\tau}+Ay_n=f$ сходится тогда и только тогда, когда

$$\tau<2/\lambda_{\max}\,, \tag{5}$$

где $\lambda_{\max}$ – максимальное собственное значение матрицы $A$.

**Доказательство.** Докажем сначала **достаточность.** Согласно условию (2) метод сходится при условии

$$E-0.5\tau A>0. \tag{6}$$

Пусть $\lambda_i,\ i=1,2,\ldots m,$ – собственные значения матрицы $A$, расположенные в порядке возрастания. Условие (6) эквивалентно тому, что все собственные значения матрицы $E-0.5\tau A$ положительны. Поэтому достаточно потребовать положительности минимального собственного значения этой матрицы, равного $1-0.5\tau\lambda_m$. Таким образом, метод простой итерации сходится, если выполнено условие (5).

**Необходимость.** Покажем, что, если условие (5) нарушено, то найдется начальное приближение $y_0$, при котором $\|y_n - y\| \not\to 0$ при $n \to \infty$.

Возьмем в качестве начального приближения вектор $y_0 = y + \mu$, где $y$ – точное решение системы $Ay = f$, а $\mu$ - собственный вектор матрицы $A$, отвечающий собственному значению $\lambda_{\max} = \lambda_m$, т. е. $A\mu = \lambda_m \mu$. При таком выборе начального приближения имеем $z_0 = y_0 - y = \mu$. Из уравнения (3) при $B = E$ получим

$$z_n = (E - \tau A)^n z_0 = (E - \tau A)^n \mu$$

и, следовательно, $z_n = (1 - \tau\lambda_m)^n \mu$, $\|z_n\| = |1 - \tau\lambda_m|^n \|\mu\|$.

Если $\tau = 2\lambda_m^{-1}$, то $\|z_n\| = \|\mu\| \not\to 0$ при $n \to \infty$. Если же $\tau > 2\lambda_m^{-1}$, то $|1 - \tau\lambda_m| > 1$ и $\|z_n\| \to \infty$ при $n \to \infty$. Таким образом, условие (5) необходимо для сходимости метода простой итерации. ▯

**2.** Будем решать систему $Ay = f$ с помощью явного **нестационарного** итерационного метода

$$\frac{y_{n+1} - y_n}{\tau_{n+1}} + Ay_n = f, \; n = 0,1,\ldots \qquad (7)$$

**Оптимальным** итерационным методом назовем такой метод (7), в котором итерационные параметры $\tau_1, \tau_2, \ldots, \tau_m$ выбраны так, что норма погрешности $\|y_m - y\|$ на $m$-ой итерации минимальна.

**Теорема.** Пусть матрица $A$ является симметричной положительно определенной матрицей, $\lambda_{\min}(A) > 0$ и $\lambda_{\max}(A) > 0$ - ее наименьшее и наибольшее собственные значения. Пусть задано число итераций $m$. Среди методов вида (7) оптимальным будет метод, для которого

**Задача.** Выяснить условие сходимости метода простой итерации для системы с матрицей

$$A = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix}. \square$$

**Задача.** Дана линейная система

$$\begin{pmatrix} 5 & 4 \\ 4 & 5 \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = \begin{pmatrix} 6 \\ 3 \end{pmatrix}$$

и начальное приближение $y_0 = (0, 1)^T$. Найти $y_3$ при решении этой системы

1) методом простой итерации с параметром $\tau = 0.2$,
2) методом с чебышевским набором параметров для $m = 3$,
3) методом Зейделя. $\square$

**Задача.** Написать формулы для вычисления компонент итерационного приближения $y_{n+1}$ по методу верхней релаксации, примененного к решению системы

$$\begin{pmatrix} a & b & 0 \\ b & a & c \\ 0 & c & a \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix} = \begin{pmatrix} f_1 \\ f_2 \\ f_3 \end{pmatrix}, \quad a > b\sqrt{2}, \quad b > c > 0. \square$$

## Вопрос 25.

## Интерполяционная формула Лагранжа и оценка ее погрешности.

**1.** Пусть на отрезке $a \le x \le b$ заданы точки $x_k$, $k = 0,1,\ldots,n$, которые называются **узлами интерполирования**. В этих точках известны значения функции $f(x)$. Задача **интерполирования алгебраическими многочленами** состоит в том, чтобы построить такой многочлен $L_n(x) = a_0 + a_1 x + \ldots + a_n x^n$ степени $n$, что его значения в заданных точках $x_k$, $k = 0,1,\ldots,n$, совпадают со значениями функции $f(x)$ в этих точках.

Данная задача имеет единственное решение для любой непрерывной функции $f(x)$, так как для отыскания коэффициентов $a_0, a_1, \ldots, a_n$ нужно решить систему уравнений

$$a_0 + a_1 x_i + a_2 x_i^2 + \ldots + a_n x_i^n = f(x_i),\ i = 0,1,\ldots,n$$

с определителем Вандермонда, не равным нулю при условии, что среди точек $x_i$, $i = 0,1,\ldots,n$, нет совпадающих.

**Интерполяционным многочленом** для функции $f(x)$, построенным по узлам $x_i$, $i = 0,1,\ldots,n$, называется многочлен $L_n(x)$, удовлетворяющий условиям $L_n(x_i) = f(x_i),\ i = 0,1,\ldots,n$.

**2. Интерполяционная формула Лагранжа.** Для построения интерполяционного многочлена в форме Лагранжа представим искомый многочлен в виде линейной комбинации

$$L_n(x) = \sum_{k=0}^{n} c_k(x) f(x_k)$$

и найдем явное выражение для коэффициентов $c_k(x)$. Так как значения $L_n(x)$ в заданных точках $x_k$, $k = 0,1,\ldots,n$, должны совпадать со значениями функции $f(x)$ в этих точках, то получаем условия

$$\sum_{k=0}^{n} c_k(x_i) f(x_k) = f(x_i), \quad i = 0,1,\dots n.$$

Эти соотношения будут выполнены, если

$$c_k(x_i) = \begin{cases} 0, & i \neq k, \\ 1, & i = k, i = 0,1,\dots,n, \end{cases}$$

то есть каждая из функций $c_k(x)$ имеет не менее $n$ нулей на $[a,b]$.

Ищем $c_k(x)$ в виде многочленов степени $n$:

$$c_k(x) = \lambda_k (x - x_0)(x - x_1)\dots(x - x_{k-1})(x - x_{k+1})\dots(x - x_n)$$

с неопределенными коэффициентами $\lambda_k$. При этом условия $c_k(x_i) = 0, \quad i \neq k$, выполнены. Из условия $c_k(x_k) = 1$ получаем

$$\lambda_k^{-1} = (x_k - x_0)(x_k - x_1)\dots(x_k - x_{k-1})(x_k - x_{k+1})\dots(x_k - x_n).$$

Следовательно, коэффициенты $c_k(x)$ находятся по формулам

$$c_k(x) = \frac{\prod\limits_{j \neq k}(x - x_j)}{\prod\limits_{j \neq k}(x_k - x_j)}. \tag{1}$$

Можно записать коэффициенты $c_k(x)$ в другом виде. Рассмотрим многочлен $\omega(x)$ степени $n+1$:

$$\omega(x) = (x - x_0)(x - x_1)\dots(x - x_{k-1})(x - x_k)(x - x_{k+1})\dots(x - x_n)$$

и вычислим его производную в точке $x_k$:

$$\omega'(x_k) = (x_k - x_0)\dots(x_k - x_{k-1})(x_k - x_{k+1})\dots(x_k - x_n).$$

Отсюда получим $\qquad c_k(x) = \dfrac{\omega(x)}{(x - x_k)\omega'(x_k)}.$

Итак, **интерполяционный многочлен Лагранжа** имеет вид

$$L_n(x) = \sum_{k=0}^{n} \frac{\prod\limits_{j \neq k}(x - x_j)}{\prod\limits_{j \neq k}(x_k - x_j)} f(x_k)$$

или $$L_n(x)=\sum_{k=0}^{n}\frac{\omega(x)}{(x-x_k)\omega'(x_k)}f(x_k).$$

**Замечание.** Интерполяционный многочлен Лагранжа применяется при выводе формул численного дифференцирования и интегрирования (например, при выводе квадратурных формул трапеций и Симпсона). Интерполирование используется также при сгущении таблиц, когда вычисление значений $f(x)$ трудоемко, и при построении сеточных аналогов задач математической физики. □

**Замечание.** При увеличении числа узлов интерполяции приходится перестраивать весь многочлен Лагранжа заново, поэтому формулой Лагранжа удобно пользоваться при интерполировании нескольких функций с фиксированным количеством узлов интерполяции. □

**Пример.** Построим интерполяционный многочлен по заданным точкам: $x_0=0, f(x_0)=1$; $x_1=1, f(x_1)=2$; $x_2=2, f(x_2)=4$. Поскольку значения $f(x)$ заданы в трех узлах, степень интерполяционного многочлена $n=2$. Найдем коэффициенты $c_k(x)$ по формуле (1):

$$c_0(x)=\frac{(x-1)(x-2)}{(0-1)(0-2)}=\frac{1}{2}(x-1)(x-2),$$

$$c_1(x)=-x(x-2),\ c_2(x)=\frac{1}{2}x(x-1).$$

Тогда

$$L_2(x)=\frac{1}{2}(x-1)(x-2)f(0)-x(x-2)f(1)+\frac{1}{2}x(x-1)f(2)=$$

$$=\frac{1}{2}(x-1)(x-2)-2x(x-2)+2x(x-1)=\frac{x^2}{2}+\frac{x}{2}+1. \ □$$

**Примеры.**

**1)** Интерполяционный многочлен можно записать в другой форме - **форме Ньютона**:

$$L_n(x)=f(x_0)+(x-x_0)f(x_0,x_1)+(x-x_0)(x-x_1)f(x_0,x_1,x_2)+$$
$$+\ldots+(x-x_0)(x-x_1)\ldots(x-x_{n-1})f(x_0,x_1,\ldots,x_n),$$

где $$f(x_0,x_1,\ldots,x_n)=\sum_{i=0}^{n}\frac{f(x_i)}{\prod\limits_{\substack{j=0,\\ j\neq i}}^{n}(x_i-x_j)}.$$

Данная формула является разностным аналогом формулы Тейлора.

**2) Интерполяционный многочлен Эрмита.** Пусть в узлах интерполирования $x_k, k=0,1,\ldots,m$, среди которых нет совпадающих, известны не только значения функции $f(x_k)$, но и ее производных $f^{(i)}(x_k), i=1,2,\ldots,N_k-1$, до порядка $N_k-1$ включительно. Числа $N_k$ при этом называют **кратностью узла** $x_k$. Тогда на всей совокупности узлов известно $N_0+N_1+\ldots+N_m$ величин, поэтому можно ставить вопрос о построении многочлена $H_n(x)$ степени $n=N_0+N_1+\ldots+N_m-1$, удовлетворяющего требованиям

$$H_n^{(i)}(x_k)=f^{(i)}(x_k), i=0,1,\ldots,N_k-1,\ \ k=0,1,\ldots,m.$$

Такой многочлен называется интерполяционным многочленом Эрмита для функции $f(x)$. Рассмотренный выше многочлен Лагранжа является частным случаем данного многочлена при условии, что все узлы простые: $N_k=1,\ \ k=0,1,\ldots,m$. □

**3. Погрешностью интерполирования** или **остаточным членом интерполяционной формулы** называется функция $r_n(x)=f(x)-L_n(x)$. В узлах интерполирования $x_k, k=0,1,\ldots,n$ погрешность равна нулю, поэтому будем получать оценку погрешности в точке $x\in[a,b]$, не являющейся узлом интерполирования.

**Теорема.** Пусть функция $f(x)$ имеет $n+1$ непрерывную производную на отрезке $[a,b]$. Тогда для погрешности интерполирования верна оценка

$$|r_n(x)|\leq\frac{M_{n+1}}{(n+1)!}|\omega(x)|, \qquad (2)$$

где $M_{n+1} = \sup\limits_{x\in[a,b]} |f^{(n+1)}(x)|$, $\omega(x) = \prod\limits_{k=0}^{n}(x - x_k)$.

**Доказательство.** Введем вспомогательную функцию $\varphi(y) = f(y) - L_n(y) - K\omega(y)$, где $K$ – некоторая константа, $y \in [a,b]$. Зафиксируем точку $x \in [a,b]$, не являющуюся узлом интерполирования и выберем постоянную $K$ из условия $\varphi(x) = 0$. Тогда получим $K = \dfrac{f(x) - L_n(x)}{\omega(x)}$. Заметим, что функция $\varphi(y)$ равна нулю в точках $x, x_k, k = 0,1,\ldots,n$, среди которых нет совпадающих. Следовательно, $\varphi(y)$ имеет не менее $n+2$ нулей на этом отрезке. Поэтому по теореме Ролля производная $\varphi'(y)$ имеет не менее, чем $n+1$ нуль на $[a,b]$, $\varphi''(y)$ - не менее $n$ нулей и т. д., функция $\varphi^{(n+1)}(y)$ по крайней мере один раз обращается в нуль на $[a,b]$. Следовательно, существует точка $\xi \in [a,b]$ такая, что $\varphi^{(n+1)}(\xi) = 0$. Тогда $\varphi^{(n+1)}(\xi) = f^{(n+1)}(\xi) - (n+1)!K = 0$ и, следовательно, $f^{(n+1)}(\xi) = \dfrac{f(x) - L_n(x)}{\omega(x)}(n+1)!$. Отсюда находим, что $r_n(x) = f(x) - L_n(x) = \dfrac{f^{(n+1)}(\xi)}{(n+1)!}\omega(x)$. Оценивая последнее равенство по модулю, получим искомую оценку (2). ▯

**Замечание.** Если $f(x)$ – алгебраический многочлен степени $n$, то интерполирование, проведенное по любым точкам $x_k, k = 0,1,\ldots,n$, осуществляется точно, то есть $L_n(x) \equiv f(x)$. ▯

**Замечание.** Важную роль в оценке (2) играет полином $\omega(x)$. На отрезке $[x_0, x_n]$ он имеет $n+1$ нуль, а его значения между этими нулями сравнительно невелики. В этом случае множитель $|\omega(x)|$ не обесценивает оценку (2). Если же точка $x$ выходит за пределы отрезка $[x_0, x_n]$ и удаляется от точки $x_0$ влево или от точки $x_n$ вправо, то оценка (2) будет резко ухудшаться за счет быстрого роста функции $|\omega(x)|$.

**4. Задача.** Написать интерполяционный многочлен второй степени для функции $y = \sin x$ по ее значениям в трех точках $x_0 = 0,\ x_1 = \dfrac{\pi}{6},\ x_2 = \dfrac{\pi}{2}$. Получить оценку погрешности в точке $x = \dfrac{\pi}{4}$.

**Задача.** Определить, с какой точностью можно вычислить по формуле Лагранжа $f\left(\frac{1}{2}\right)$, если известны значения $f(-1)$, $f(0)$, $f(1)$, где $f(x) = x^3 + x^2 - 2x$. Сравнить с оценкой по формуле (2).

**Задача.** Функция $f(x) = \dfrac{1}{A^2 - x}$ приближается на $[-4,-1]$ многочленом Лагранжа по узлам $x_i = -4,-3,-2,-1$. При каких значениях $A$ оценка погрешности в точке $x = -\dfrac{3}{2}$ не превосходит $\dfrac{15}{16} \cdot 10^{-5}$?

## Вопрос 26.

## Метод прогонки решения разностных уравнений.

**1.** Дана система линейных разностных уравнений

$$a_j y_{j-1} - c_j y_j + b_j y_{j+1} = -f_j, \quad j = 1,2,\ldots,N-1, \qquad (1)$$

$$y_0 = k_1 y_1 + \mu_1, \; y_N = k_2 y_{N-1} + \mu_2. \qquad (2)$$

Эта система относится к системам с **трехдиагональной** матрицей, т. е. с матрицей, все элементы которой, не лежащие на главной и двух побочных диагоналях, равны нулю. Для численного решения таких систем применяется **метод прогонки**, который является вариантом метода Гаусса. Метод прогонки основан на предположении, что искомые величины $y_j, y_{j+1}$ связаны рекуррентным соотношением

$$y_j = \alpha_{j+1} y_{j+1} + \beta_{j+1}, \quad j = 0,1,\ldots,N-1. \qquad (3)$$

Здесь коэффициенты $\alpha_{j+1}, \beta_{j+1}$ (так называемые **прогоночные коэффициенты**) подлежат определению, исходя из условий (1),(2).

Выведем расчетные формулы данного метода. Для этого с помощью равенства (3) выразим $y_{j-1}$ через $y_{j+1}$:

$$y_{j-1} = \alpha_j y_j + \beta_j = \alpha_j(\alpha_{j+1} y_{j+1} + \beta_{j+1}) + \beta_j =$$
$$= \alpha_j \alpha_{j+1} y_{j+1} + (\alpha_j \beta_{j+1} + \beta_j), \; j = 1,2,\ldots,N-1.$$

Подставляя полученные выражения для $y_{j-1}$ и $y_j$ в исходное уравнение (1), получим

$$(\alpha_{j+1}(a_j\alpha_j - c_j) + b_j) y_{j+1} + (\beta_{j+1}(a_j\alpha_j - c_j) + a_j\beta_j + f_j) = 0.$$

Это равенство будет выполнено, если потребовать, чтобы при $j = 1,2,\ldots,N-1$ имели место соотношения

$$\alpha_{j+1}(a_j\alpha_j - c_j) + b_j = 0, \; \beta_{j+1}(a_j\alpha_j - c_j) + a_j\beta_j + f_j = 0$$

или

$$\alpha_{j+1} = \frac{b_j}{c_j - a_j\alpha_j}, \; \beta_{j+1} = \frac{a_j\beta_j + f_j}{c_j - a_j\alpha_j}, \; j = 1,2,\ldots,N-1. \qquad (4)$$

Для решения нелинейных разностных уравнений (4) необходимо задать начальные значения $\alpha_1, \beta_1$. Первое из условий

(2) и соотношение (3) при $j=0$ должны быть эквивалентны. Для этого достаточно положить

$$\alpha_1 = k_1,\ \beta_1 = \mu_1. \qquad (5)$$

Теперь можно вычислить по формулам (4) остальные коэффициенты $\alpha_{j+1}, \beta_{j+1},\ j=1,2,\dots,N-1$. Этот этап вычислений называется **прямой прогонкой**.

Следующий этап (**обратная прогонка**) – это нахождение решения системы (1) по рекуррентной формуле (3). Счет начинается с $j=N-1$. Для начала счета требуется задать значение $y_N$, которое определяется из уравнений $y_{N-1}=\alpha_N y_N+\beta_N$, $y_N = k_2 y_{N-1}+\mu_2$ и равно

$$y_N = \frac{k_2\beta_N+\mu_2}{1-k_2\alpha_N}.$$

Далее находятся $y_j$ по формулам

$$y_j = \alpha_{j+1}y_{j+1}+\beta_{j+1}, \quad j=N-1, N-2,\dots,0.$$

**Замечание.** Изложенный выше метод называется методом **правой прогонки**. Существуют и другие варианты метода прогонки.

**1) Метод левой прогонки.** В этом случае решение системы (1), (2) находится по формулам

$$y_{j+1} = \xi_{j+1}y_j + \eta_{j+1}, \quad j=0,1,\dots,N-1,$$

$$\xi_j = \frac{a_j}{c_j - b_j\xi_{j+1}},\ \eta_j = \frac{b_j\eta_{j+1}+f_j}{c_j-b_j\xi_{j+1}},\ j=N-1, N-2,\dots,2,1,$$

$$\xi_N = k_2,\ \eta_N=\mu_2,\quad y_0 = \frac{k_1\eta_1+\mu_1}{1-k_1\xi_1}.$$

**2) Метод встречных прогонок** – это комбинация правой и левой прогонок. Пусть $j=j_0,\ 0<j_0<N$, – некоторый внутренний узел. Тогда в области $0\le j\le j_0+1$ вычисляются по формулам правой прогонки коэффициенты $\alpha_j, \beta_j$, а в области $j_0\le j\le N$ по формулам левой прогонки находятся $\xi_j, \eta_j$. Теперь при $j=j_0$ из формул $y_{j_0}=\alpha_{j_0+1}y_{j_0+1}+\beta_{j_0+1}$,

$y_{j_0+1} = \xi_{j_0+1} y_{j_0} + \eta_{j_0+1}$ получаем $y_{j_0} = \dfrac{\beta_{j_0+1} + \alpha_{j_0+1}\eta_{j_0+1}}{1 - \alpha_{j_0+1}\xi_{j_0+1}}$. Зная $y_{j_0}$, можно найти все остальные $y_j$: $y_j = \alpha_{j+1} y_{j+1} + \beta_{j+1}$ при $j < j_0$ и $y_{j+1} = \xi_{j+1} y_j + \eta_{j+1}$ при $j > j_0$. ▯

**2.** Метод прогонки называется **устойчивым**, если $|\alpha_j| \le 1,\ j = 1,2,\ldots,N$.

**Замечание.** Неравенства $|\alpha_j| \le 1,\ \ j = 1,\ldots,N$ обеспечивают устойчивость счета по рекуррентным формулам (3), т. е. погрешность, внесенная на каком-либо шаге вычислений, не будет возрастать при переходе к следующим шагам. ▯

**Теорема.** Пусть коэффициенты исходной системы (1), (2) удовлетворяют условиям

$$a_j \ne 0, b_j \ne 0, |c_j| \ge |a_j| + |b_j|,\ j = 1,2,\ldots,N-1, \qquad (6)$$

$$|k_1| \le 1,\ |k_2| < 1. \qquad (7)$$

Тогда метод прогонки применим и устойчив.

**Доказательство.** По индукции докажем, что $|\alpha_j| \le 1,\ \ j = 1,\ldots,N$. Из формул (5), (7) видно, что $|\alpha_1| = |k_1| \le 1$. Пусть $|\alpha_j| \le 1$ для некоторого $j$; докажем, что $|\alpha_{j+1}| \le 1$. Так как $\left|c_j - \alpha_j a_j\right| \ge \left||c_j| - |\alpha_j|\,|a_j|\right| \ge \left||c_j| - |a_j|\right|$, то, учитывая (6), получим $\left|c_j - \alpha_j a_j\right| \ge \left|b_j\right| > 0$. Следовательно, знаменатели в равенствах (4) не обращаются в нуль и $|\alpha_{j+1}| = \dfrac{|b_j|}{|c_j - a_j\alpha_j|} \le 1$, $j = 1,\ldots,N-1$. Поэтому прогонка устойчива.

Далее, из условия $|k_2| < 1$ и только что доказанного неравенства $|\alpha_N| \le 1$ имеем $|1 - k_2\alpha_N| \ge 1 - |k_2|\,|\alpha_N| \ge$ $\ge 1 - |k_2| > 0$. Значит, не обращается в нуль знаменатель в выражении для $y_N$. ▯

**Замечание.** Теорема будет верна, если условия (6), (7) заменить на соотношения $a_j \neq 0, b_j \neq 0, |c_j| > |a_j| + |b_j|$, $j = 1,2,\ldots,N-1$, $|k_1| \leq 1$, $|k_2| \leq 1$. □

**Пример.** Рассмотрим первую краевую задачу для уравнения теплопроводности

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2}, \quad 0 < x < l, \quad t > 0, \tag{8}$$
$$u(x,0) = u_0(x), \quad 0 \leq x \leq l,$$
$$u(0,t) = \nu_1(t), \quad u(l,t) = \nu_2(t), \quad t \geq 0.$$

На плоскости $(x,t)$ введем сетку $\omega_{h,\tau} = \{(x_j, t_n), x_j = jh,$ $j = 0,1,\ldots,N,\ Nh = l,\ t_n = n\tau, n = 0,1,\ldots\}$. Аппроксимируем задачу (8) разностной схемой

$$y^n_{t,j} = \sigma y^{n+1}_{\bar{x}x,j} + (1-\sigma) y^n_{\bar{x}x,j}, \tag{9}$$
$$j = 1,2,\ldots,N-1, \quad n = 0,1,\ldots,$$
$$y^n_0 = \nu_1(t_n),\ y^n_N = \nu_2(t_n),\ n = 0,1,\ldots,\ y^0_j = u_0(x_j),\ j = 0,1,\ldots,N.$$

Здесь $\sigma$ - действительный параметр, $y^n_j = y(x_j, t_n)$,

$$y^n_{t,j} = \frac{y^{n+1}_j - y^n_j}{\tau}, \quad y^n_{\bar{x}x,j} = \frac{y^n_{j+1} - 2y^n_j + y^n_{j-1}}{h^2}.$$

Считая $y^n_j$ известным, сведем схему (9) к виду (1), (2):

$$\frac{1}{h^2} y^{n+1}_{j-1} - \left(\frac{1}{\sigma\tau} + \frac{2}{h^2}\right) y^{n+1}_j + \frac{1}{h^2} y^{n+1}_{j+1} = -\varphi^n_j,$$
$$j = 1,2,\ldots,N-1, \quad n = 0,1,\ldots,$$
$$y^{n+1}_0 = \nu_1(t_{n+1}),\ y^{n+1}_N = \nu_2(t_{n+1}),\ n = 0,1,\ldots,$$

где $\varphi^n_j = \dfrac{1}{\sigma\tau} y^n_j + \left(\dfrac{1}{\sigma} - 1\right) y^n_{\bar{x}x,j}$, если $\sigma \neq 0$. Отсюда видно, что

$$a_j = b_j = \frac{1}{h^2}, \quad c_j = \frac{1}{\sigma\tau} + \frac{2}{h^2},\ f_j = \varphi^n_j,\ j = 1,2,\ldots,N-1,$$

$k_1 = k_2 = 0, \quad \mu_1 = \nu_1(t_{n+1}), \ \mu_2 = \nu_2(t_{n+1})$. Найдем условия, при которых построенную систему можно будет решать методом прогонки. Из теоремы следует, что должно быть выполнено условие

$$\left| \frac{1}{\sigma\tau} + \frac{2}{h^2} \right| \geq \frac{2}{h^2}.$$

Решая это неравенство, найдем достаточное условие применимости и устойчивости прогонки

$$\sigma \geq -\frac{h^2}{4\tau}. \ \square$$

**Замечание.** Рассмотренный выше метод прогонки предназначен для однопроцессорного компьютера. Однако в настоящее время все более распространенными и доступными становятся многопроцессорные вычислительные системы. Многие вычислительные алгоритмы могут быть распараллелены, т. е. оказывается возможным выделить в них фрагменты, исполняемые одновременно на различных процессорах, за счет чего увеличивается скорость решения требуемой задачи. Наиболее эффективными являются алгоритмы, учитывающие архитектуру ЭВМ.

Покажем, например, как можно распараллелить метод прогонки. Если имеется вычислительная машина с $M$ **независимыми** процессорами, то целесообразно разбить систему (1), (2) на $M$ групп уравнений. Проведем разбиение следующим образом. Зададим целые числа $n_0, n_1, \ldots, n_M$ такие, что $0 = n_0 < n_1 < \ldots < n_{M-1} < n_M = N$. К $k$-ой группе отнесем уравнения системы (1), (2), имеющие номера $n_{k-1}, n_{k-1}+1, \ldots, n_k$. В каждой группе независимо и одновременно осуществим прямую прогонку. В результате из исходной системы (1), (2) выделится система уравнений с трехдиагональной матрицей, связывающая значения неизвестных только в граничных точках групп, т. е. значения $y_{n_1}, y_{n_2}, \ldots, y_{n_{M-1}}$. Для решения полученной системы применим обычный метод прогонки, состоящий из прямого и обратного хода. Тем самым будут найдены значения неизвестных в граничных точках групп. Далее, используя найденные граничные значения, в каждой группе можно одновременно вычислить все остальные значения $y_j$. $\square$

**3. Задача.** Разностная задача

$$\frac{y_{j-1}-2y_j+y_{j+1}}{h^2}-d_j\frac{y_{j-1}+4y_j+y_{j+1}}{6}=-f_j,\ j=1,2,\ldots,N-1,$$

$$y_0=y_N=0,\quad 0\le d_j\le Q,\quad Q=const$$

решается методом прогонки. Доказать, что прогонка устойчива при условии $h^2Q<6$. ▯

**Задача.** Найти решение системы методом прогонки

$$y_{j-1}-2y_j+y_{j+1}=-2,\quad j=1,2,3,4,$$

$$y_0=0,\quad y_5=0.5y_4+1.$$ ▯

**Задача.** Найти решение системы

$$2y_{j-1}-3y_j+y_{j+1}=0,\quad j=1,2,3,4,$$

$$y_0=0,\quad y_5=2,$$

1) пользуясь формулами метода правой прогонки,
2) пользуясь формулами метода левой прогонки. ▯

**Задача.** Доказать, что при решении краевой задачи

$$y_{j-1}-2y_j+y_{j+1}=-f_j,\quad j=1,2,\ldots,N-1,$$

$$y_0=a,\quad y_N=b$$

методом правой прогонки коэффициенты $\alpha_{j+1}$ вычисляются по формуле $\alpha_{j+1}=\dfrac{j}{j+1}$. ▯

## Вопрос 27.

## Основные понятия теории разностных схем: аппроксимация, устойчивость, сходимость.

**1.** Пусть дана исходная **дифференциальная** задача

$$Lu(x) = f(x), \quad x \in G, \tag{1}$$

где $G$ - область в $R^m$, $f(x)$ - заданная функция, $L$ - **линейный** дифференциальный оператор. Предполагаем, что дополнительные условия (например, начальные и граничные) учтены оператором $L$ и правой частью $f(x)$.

Введем **сетку** $G_h$ - конечное множество точек, принадлежащих $G$, плотность распределения которых характеризуется параметром $h$ – **шагом сетки**. В общем случае параметр $h$ является вектором, его длина $|h|$. Обычно сетка $G_h$ выбирается так, что при $|h| \to 0$ множество $G_h$ стремится заполнить всю область $\overline{G}$. Функция, определенная в точках сетки $G_h$, называется **сеточной функцией.**

**Примеры.**

**1) Равномерной сеткой на отрезке** $[a,b]$ называется множество точек $G_h = \{x_i = a + ih, i = 0,1,\ldots,N\}$, где $h = (b-a)/N$ - шаг сетки.

**2) Неравномерная сетка на отрезке** $[a,b]$.

Возьмем произвольные точки $a = x_0 < x_1 < x_2 < \ldots < x_{N-1} < x_N = b$ и введем шаги сетки в виде $x_i - x_{i-1} = h_i$, $\sum_{i=1}^{N} h_i = b - a$. Образованная таким образом сетка называется неравномерной. Шаг сетки в этом случае зависит от номера узла, т. е. является сеточной функцией.

**3) Равномерная сетка в прямоугольнике** $D = \{0 \le x \le 1, 0 \le t \le T\}$. Разобьем отрезки $[0,1]$ оси $x$ и $[0,T]$ оси $t$ соответственно на $N_1$ и $N_2$ частей и обозначим

$h = 1/N_1, \tau = T/N_2$. Тогда сеткой будет множество точек $G_{h\tau} = \{(x_i, t_j)\}$, где $x_i = ih$, $t_j = j\tau$, $i = 0,1,\ldots,N_1$, $j = 0,1,\ldots,N_2$. □

В уравнении (1) заменим дифференциальный оператор $L$ на сетке $G_h$ разностным оператором $L_h$, функции $f(x)$, $u(x)$ - соответственно сеточными функциями $\varphi_h(x), y_h(x)$. В результате получим систему разностных уравнений

$$L_h\, y_h(x) = \varphi_h(x), \quad x \in G_h, \tag{2}$$

которая называется **разностной схемой** или **разностной задачей**. Меняя параметр $h$, получаем множество решений $y_h(x)$, зависящих от $h$. Таким образом, следует рассматривать **семейство разностных задач (2)**, соответствующих различным значениям параметра $h$. В дальнейшем предполагаем, что $L_h$ – линейный оператор.

**2.** Пусть решение задачи (1) $u(x) \in V_0$, где $V_0$ - линейное пространство с введенной в нем нормой $\|\cdot\|_0$, а сеточные функции $y_h(x), \varphi_h(x)$ являются элементами линейного пространства $V_h$ (**пространства сеточных функций**) с нормой $\|\cdot\|_h$. Основной интерес для теории приближенных методов представляет оценка близости $y_h(x)$ к $u(x)$. Но эти функции являются элементами разных пространств, и для их сравнения необходимо каким- либо образом отобразить пространство $V_h$ на $V_0$ или наоборот. В первом случае сеточная функция $y_h(x)$ доопределяется (например, при помощи линейной интерполяции) во всех остальных точках области $G$. В результате получаем функцию $\widetilde{y}$ непрерывного аргумента. Близость $y_h(x)$ к $u(x)$ характеризуется числом $\|\widetilde{y} - u\|_0$.

Во втором случае каждой функции $u(x) \in V_0$ поставим в соответствие сеточную функцию $u_h(x)$ с помощью некоторого линейного **оператора проектирования** $P_h : V_0 \to V_h$, т. е. $u_h(x) = P_h u(x)$. Имея сеточную функцию $u_h(x)$, образуем разность $y_h - u_h$, являющуюся элементом пространства $V_h$. Тогда близость $y_h(x)$ к $u(x)$ характеризуется числом $\| y_h - u_h \|_h$.

В дальнейшем будем сравнивать функции $y_h(x)$ и $u(x)$ только в пространстве сеточных функций $V_h$.

**Примеры.**

**1)** Пусть $V_0$ – пространство непрерывных функций на $[0,1]$ и $G_h$ – равномерная сетка с шагом $h$. Тогда оператор проектирования можно определить следующим образом: $(P_h u)(x_i) = u(x_i),\ i = 0,1,\dots,N.$

**2)** Если $V_0$ – пространство функций, интегрируемых на $[0,1]$ и $G_h$ – та же сетка, то в качестве оператора проектирования можно взять оператор осреднения

$$(P_h u)(x_i) = \frac{1}{h} \int\limits_{x_i - 0.5h}^{x_i + 0.5h} u(x)\,dx,\ \ i = 1,2,\dots,N-1,$$

$$(P_h u)(x_0) = \frac{1}{0.5h} \int\limits_{0}^{0.5h} u(x)\,dx, \quad (P_h u)(x_N) = \frac{1}{0.5h} \int\limits_{1-0.5h}^{1} u(x)\,dx. \ \square$$

Будем требовать, чтобы нормы в $V_h$ были **согласованы** с нормой в исходном пространстве $V_0$, т. е. для любой $u \in V_0$ выполнено условие $\lim\limits_{|h| \to 0} \| P_h u \|_h = \| u \|_0$. Это условие обеспечивает единственность предела сеточных функций при $|h| \to 0$. Действительно, если для $u, v \in V_0$

$$\lim_{|h| \to 0} \| y_h - P_h u \|_h = 0, \quad \lim_{|h| \to 0} \| y_h - P_h v \|_h = 0,$$

то

$$\| P_h (u - v) \|_h = \| P_h u - P_h v \|_h = \| (P_h u - y_h) + (y_h - P_h v) \|_h \le$$
$$\le \| P_h u - y_h \|_h + \| y_h - P_h v \|_h.$$

Тогда $\left\| u-v \right\|_0 = \lim\limits_{|h|\to 0} \left\| P_h(u-v) \right\|_h = 0$, т. е. $u = v$.

**3.** Пусть $u(x)$ - решение исходной задачи (1) и $y_h(x)$ – решение разностной задачи (2) при некотором фиксированном $h$.

**Определение.** Сеточная функция $z_h(x) = y_h(x) - P_h u(x),\ x \in G_h$ называется **погрешностью разностной схемы** (2).

После подстановки $y_h(x) = P_h u(x) + z_h(x)$ в уравнение (2) получим, что погрешность $z_h(x)$ удовлетворяет уравнению

$$L_h\, z_h(x) = \psi_h(x),\ \ x \in G_h, \tag{3}$$

где

$\psi_h(x) = \varphi_h(x) - L_h\,(P_h u(x)) \equiv \varphi_h(x) - L_h\, u_h(x)$ - сеточная функция, называемая **погрешностью аппроксимации разностной задачи** (2) **<u>на решении</u> исходной дифференциальной задачи** (1).

Преобразуем выражение для $\psi_h(x)$. Спроектируем уравнение (1) на сетку: $P_h L u(x) = P_h f(x)$ или $(Lu)_h(x) = f_h(x)$. Тогда
$$\psi_h(x) = \varphi_h(x) - L_h u_h(x) + \\ + (Lu)_h(x) - f_h(x) = [(Lu)_h(x) - L_h u_h(x)] + \\ + [\varphi_h(x) - f_h(x)],$$

т. е.

$$\psi_h(x) = \psi_{h,1}(x) + \psi_{h,2}(x),$$

где

$$\psi_{h,1}(x) = (Lu)_h(x) - L_h\, u_h(x),$$
$$\psi_{h,2}(x) = \varphi_h(x) - f_h(x).$$

**Определение.** Функции $\psi_{h,1}(x)$ и $\psi_{h,2}(x)$ называются, соответственно, **погрешностью аппроксимации дифференциального оператора** $L$ **разностным оператором** $L_h$ **и погрешностью аппроксимации правой части.**

**Определение.** Разностная задача (2) **аппроксимирует** исходную задачу (1), если $\left\| \psi_h \right\|_h \to 0$ при $|h| \to 0$. Разностная схема имеет $k$ **-й порядок аппроксимации**, если существуют постоянные $k > 0,\ M_1 > 0$, не зависящие от $h$ и такие, что

$$\left\| \psi_h \right\|_h \le M_1 |h|^k .$$

Аналогично определяются погрешность аппроксимации дифференциального оператора и погрешность аппроксимации правой части.

**Замечание.** Погрешность аппроксимации на решении представляется в виде суммы погрешностей аппроксимации дифференциального оператора и правой части. Но порядок погрешности аппроксимации на решении может оказаться выше, чем порядок погрешности аппроксимации оператора и правой части в отдельности. ▯

Рассмотрим пример, иллюстрирующий данное замечание.

**Пример.** На равномерной сетке с шагом $h$ заменим дифференциальное уравнение

$$u''(x) = -f(x)$$

разностным уравнением

$$y_{\bar{x}x,i} = -\varphi_i, \quad \varphi_i = f(x_i) + \frac{h^2}{12} f''(x_i),$$

где $y_{\bar{x}x,i} = \dfrac{y_{i+1} - 2y_i + y_{i-1}}{h^2}$.

Пользуясь разложением функции в ряд Тейлора и обозначая $u(x_i) = u_i$, получим

$$u_{i\pm1} = u_i \pm h u_i' + \frac{h^2}{2} u_i'' \pm \frac{h^3}{3!} u_i''' + \frac{h^4}{4!} u_i^{(4)} \pm \frac{h^5}{5!} u_i^{(5)} + O(h^6).$$

Поэтому $\qquad u_{\bar{x}x,i} = u_i'' + \dfrac{h^2}{12} u_i^{(4)} + O(h^4)$.

Следовательно, дифференциальный оператор и правая часть имеют второй порядок аппроксимации $u_i'' - u_{\bar{x}x,i} = O(h^2)$,

$f_i - \varphi_i = O(h^2)$. Найдем порядок погрешности аппроксимации на решении:

$$\psi_i = [u_i'' + f_i] - [u_{\bar{x}x,i} + \varphi_i] = u_i'' + f_i - u_i'' -$$

$$-\frac{h^2}{12}u_i^{(4)} - f_i - \frac{h^2}{12}f_i'' + O(h^4).$$

Далее, из уравнения $u''(x) = -f(x)$ находим, что $u_i^{(4)} = -f_i''$. Поэтому $\psi_i = O(h^4)$.▯

**Пример.** Покажем, что порядок погрешности аппроксимации может зависеть от выбора сетки. Действительно, на равномерной сетке $u_i'' - u_{\bar{x}x,i} = O(h^2)$. Возьмем теперь неравномерную сетку с шагом $h_i$ и заменим $u''(x)$ разностным отношением

$$L_h u_i = \frac{1}{\hbar_i}\left(\frac{u_{i+1} - u_i}{h_{i+1}} - \frac{u_i - u_{i-1}}{h_i}\right), \quad \hbar_i = 0{,}5(h_i + h_{i+1}).$$

Вычислим погрешность аппроксимации в точке $x_i$. Для этого разложим достаточно гладкую функцию $u(x)$ в окрестности данной точки:

$$u_{i+1} = u_i + h_{i+1}u_i' + \frac{h_{i+1}^2}{2}u_i'' + \frac{h_{i+1}^3}{3!}u_i''' + O(h_{i+1}^4),$$

$$u_{i-1} = u_i - h_i u_i' + \frac{h_i^2}{2}u_i'' - \frac{h_i^3}{3!}u_i''' + O(h_i^4).$$

Тогда получим, что

$$\frac{u_{i+1} - u_i}{h_{i+1}} = u_i' + \frac{h_{i+1}}{2}u_i'' + \frac{h_{i+1}^2}{3!}u_i''' + O(h_{i+1}^3),$$

$$\frac{u_i - u_{i-1}}{h_i} = u_i' - \frac{h_i}{2}u_i'' + \frac{h_i^2}{3!}u_i''' + O(h_i^3).$$

Следовательно, $L_h u_i = u_i'' + \dfrac{h_{i+1}^2 - h_i^2}{6\hbar_i}u_i''' + O(\hbar_i^2)$ и

$$\psi_i = L_h u_i - u_i'' = \frac{h_{i+1} - h_i}{3} u_i''' + O(\hbar_i) = O(\hbar_i),$$

т. е. разностный оператор имеет первый порядок аппроксимации относительно $\hbar_i$ на произвольной неравномерной сетке. ▯

**Определение.** Разностная схема (2) называется **корректной**, если

1) ее решение существует и единственно при любых правых частях $\varphi_h \in V_h$.

2) существует постоянная $M_2 > 0$, не зависящая от $h$ и такая, что при любых $\varphi_h \in V_h$ выполняется оценка

$$\| y_h \|_h \le M_2 \| \varphi_h \|_h . \qquad (4)$$

Свойство 2) называется **устойчивостью** разностной схемы, а оценка (4) – **априорной** оценкой.

**Замечание.** Свойство 1) эквивалентно существованию оператора $L_h^{-1}$, а свойство 2) – равномерной по $h$ ограниченности оператора $L_h^{-1}$. ▯

**Определение.** Решение разностной задачи (2) **сходится** к решению дифференциальной задачи (1), если $\| y_h - P_h u \|_h \to 0$ при $|h| \to 0$. Разностная схема имеет $k$**-й порядок точности**, если существуют постоянные $k > 0, M_3 > 0,$ не зависящие от $h$ и такие, что

$$\| y_h - P_h u \|_h \le M_3 |h|^k .$$

**Теорема.** Пусть дифференциальная задача (1) поставлена корректно, разностная схема (2) является корректной и аппроксимирует исходную задачу (1). Тогда решение разностной задачи (2) сходится к решению исходной задачи (1), причем порядок точности совпадает с порядком аппроксимации.

**Доказательство.** Из корректности разностной схемы (2) следует оценка (4). Так как уравнение (3) для погрешности схемы $z_h$ имеет ту же структуру, что и разностная задача (2), то выполняется оценка

$$\| z_h \|_h \le M_2 \| \psi_h \|_h . \qquad (5)$$

Константа $M_2$ не зависит от $h$, поэтому при $\| \psi_h \|_h \to 0$ норма погрешности $z_h$ стремится к нулю, т. е. схема сходится. Далее, если $\| \psi_h \|_h \le M_1 | h |^k$, то из (5) следует оценка $\| z_h \|_h \le M_1 M_2 | h |^k$. Тем самым, разностная схема имеет $k$-й порядок точности. ▯

**Замечание.** Доказанная выше теорема позволяет разделить изучение сходимости на два отдельных этапа: доказательство аппроксимации и доказательство устойчивости. ▯

**4. Задача.** Дифференциальная задача

$$u''(x) = 0, \quad 0 < x < 1, \quad u'(0) = \mu_1, \quad u(1) = 0$$

заменяется разностной схемой

$$y_{\bar{x}x,i} = 0, \quad i = 1,2,\ldots, N-1, \; Nh = 1,$$

$$\frac{y_1 - y_0}{h} = \mu_1, \quad y_N = 0.$$

Какой порядок относительно шага $h \to 0$ имеет погрешность аппроксимации граничного условия

а) на решении исходной дифференциальной задачи,

б) на произвольной достаточно гладкой функции? ▯

**Задача.** При каком соотношении $h$ и $\tau$ разностное уравнение $\dfrac{y_i^{n+1} - y_i^n}{\tau} = \dfrac{y_{i-1}^n - 2y_i^n + y_{i+1}^n}{h^2}$ имеет порядок аппроксимации $O(\tau^2 + h^4)$ на решении дифференциального уравнения $\dfrac{\partial u}{\partial t} = \dfrac{\partial^2 u}{\partial x^2}$ ? ▯

## Вопрос 28.

## Разностная аппроксимация задачи Дирихле для уравнения Пуассона: постановка разностной задачи, оценка погрешности.

**1. Постановка разностной задачи**. Пусть $G=\{0<x_1<l_1,\, 0<x_2<l_2\}$ – прямоугольник с границей $\Gamma$. Задача Дирихле для уравнения Пуассона формулируется следующим образом: найти непрерывную в $\overline{G}=G\cup\Gamma$ функцию $u(x_1,x_2)$, удовлетворяющую уравнению

$$\frac{\partial^2 u}{\partial x_1^2}+\frac{\partial^2 u}{\partial x_2^2}=-f(x_1,x_2),\quad (x_1,x_2)\in G \tag{1}$$

и граничному условию

$$u(x_1,x_2)=\mu(x_1,x_2),\quad (x_1,x_2)\in\Gamma. \tag{2}$$

Здесь функции $f(x_1,x_2),\,\mu(x_1,x_2)$ заданы и таковы, что решение задачи (1), (2) существует, единственно и является достаточно гладкой функцией.

Введем в $\overline{G}$ прямоугольную сетку $G_h$ с шагами $h_1$ по направлению $x_1$ и $h_2$ – по направлению $x_2$: $h_1=l_1/N_1,\ h_2=l_2/N_2$, где $N_1, N_2$ – целые числа. Обозначим $x_1^i=ih_1,\ x_2^j=jh_2$. Сетка $G_h$ состоит из узлов $x_{ij}=(x_1^i,x_2^j),\ i=0,1,\dots N_1,\ j=0,1,\dots,N_2$, которые подразделяются на множество **внутренних точек**

$$\omega=\{x_{ij}\mid i=1,2,\dots N_1-1,\ j=1,2,\dots,N_2-1\}$$

и множество **граничных точек** сетки

$$\gamma=\{x_{0j},\,x_{N_1j}\}_{j=1}^{N_2-1}\bigcup\{x_{i0},\,x_{iN_2}\}_{i=1}^{N_1-1}.$$

**Замечание.** Угловые точки $(0,0),(l_1,0),(0,l_2),(l_1,l_2)$ не будут участвовать в аппроксимации исходной задачи, поэтому их не относят ни к внутренним, ни к граничным точкам. ▯

Далее, для сеточных функций и разностных отношений введем обозначения $y_{ij} = y(x_{ij})$ и

$$y_{\bar{x}_1 x_1, ij} = \frac{y_{i+1,j} - 2y_{ij} + y_{i-1,j}}{h_1^2},$$

$$y_{\bar{x}_2 x_2, ij} = \frac{y_{i,j+1} - 2y_{ij} + y_{i,j-1}}{h_2^2}.$$

Сопоставим задаче Дирихле (1), (2) следующую разностную схему

$$y_{\bar{x}_1 x_1, ij} + y_{\bar{x}_2 x_2, ij} = -f_{ij}, \tag{3}$$

$$i = 1,2,\ldots,N_1 - 1,\ j = 1,2,\ldots,N_2 - 1,$$

$$y_{i0} = \mu(x_1^i, 0),\ y_{iN_2} = \mu(x_1^i, l_2),\ i = 1,2,\ldots,N_1 - 1, \tag{4}$$

$$y_{0j} = \mu(0, x_2^j),\ y_{N_1 j} = \mu(l_1, x_2^j),\ j = 1,2,\ldots,N_2 - 1.$$

Построенная разностная схема имеет второй порядок погрешности аппроксимации по $h_1$ и по $h_2$ и представляет собой систему линейных алгебраических уравнений относительно $y_{ij}$, состоящую из $(N_1 - 1)(N_2 - 1)$ уравнений и стольких же неизвестных.

**2. Каноническая форма записи разностной схемы.** Каждой точке $x \in G_h$ поставим в соответствие один и только один **шаблон** $Ш(x)$ - любое подмножество $G_h$, содержащее данную точку $x$. **Окрестностью** точки $x$ назовем множество $Ш'(x) = Ш(x) \setminus \{x\}$, причем $Ш'(x)$ может быть и пустым множеством. Пусть заданы вещественные функции $A(x), B(x,\xi), F(x)$, где $x, \xi \in G_h$.

**Канонической формой записи разностной схемы** называется ее запись в виде

$$A(x)y(x) = \sum_{\xi \in Ш'(x)} B(x,\xi) y(\xi) + F(x), \quad x \in G_h, \tag{5}$$

где $y(x)$ - искомая сеточная функция.

Получим каноническую форму разностной схемы (3), (4). Для этого запишем уравнение (3) в виде, разрешенном относительно $y_{ij}$:

$$\left(\frac{2}{h_1^2}+\frac{2}{h_2^2}\right)y_{ij}=\frac{y_{i+1,j}+y_{i-1,j}}{h_1^2}+\frac{y_{i,j+1}+y_{i,j-1}}{h_2^2}+f_{ij}. \qquad (6)$$

Обозначим через $x$ точку $x_{ij}$ - центральную точку шаблона $Ш(x)$, который для данного уравнения состоит из пяти точек $x_{ij}$, $x_{i\pm1,j}$, $x_{i,j\pm1}$. Тогда $Ш'(x)$ - это четыре точки $x_{i\pm1,j}$, $x_{i,j\pm1}$. Определим коэффициенты $A(x), B(x,\xi)$ и правую часть $F(x)$ следующим образом:

$$A(x)=\frac{2}{h_1^2}+\frac{2}{h_2^2},\ \ B(x,x_{i\pm1,j})=\frac{1}{h_1^2},$$

$$B(x,x_{i,j\pm1})=\frac{1}{h_2^2},\quad F(x)=f_{ij}.$$

В этих обозначениях уравнение (6) будет записано в канонической форме (5). Но уравнение (6) определено только во внутренних точках сетки, т.е. при $x\in\omega$. Поэтому к нему нужно добавить граничные условия (4). Будем считать $Ш'(x)$ пустым множеством при $x\in\gamma$ и обозначим $F(x)=\mu(x)$. Тогда условия (4) можно записать в виде

$$A(x)y(x)=F(x),\quad x\in\gamma,\ \text{где}\ A(x)=1.$$

Таким образом, разностную схему (3),(4) можно представить в виде системы уравнений (5), причем для любой точки $x\in G_h$ выполняются **условия положительности коэффициентов**

$$A(x)>0,\ B(x,\xi)>0\ \text{для всех}\ \xi\in Ш'(x), \qquad (7)$$

$$D(x)=A(x)-\sum_{\xi\in Ш'(x)}B(x,\xi)\ge 0.$$

Заметим, что для задачи (3), (4) $D(x)=0$ для $x\in\omega$ и $D(x)=1$ для $x\in\gamma$.

Разностные схемы, удовлетворяющие при всех $x\in G_h$ условиям (7), называются **монотонными разностными схемами**. Приведем примеры монотонных и немонотонных схем.

**Пример.** Для дифференциальной задачи

$$u''(x) = -f(x), \quad 0 < x < 1, \quad u'(0) = 0, u(1) = 0$$

построим разностную схему второго порядка аппроксимации

$$\frac{y_{i+1} - 2y_i + y_{i-1}}{h^2} = -f_i, \quad i = 1,2,\ldots, N-1,$$

$$\frac{y_1 - y_0}{h} = -\frac{h}{2} f_0, \quad y_N = 0.$$

Запишем схему в канонической форме

$$y_i = 0{,}5(y_{i+1} + y_{i-1}) + 0{,}5h^2 f_i, \quad i = 1,2,\ldots, N-1,$$

$$y_0 = y_1 + 0{,}5h^2 f_0, \quad y_N = 0.$$

Сетка $G_h$ состоит из узлов $x_i = ih, i = 0,1,\ldots, N$ и имеет одну граничную точку $x_N$. Окрестность $Ш'(x_0)$ узла $x_0$ состоит из одного узла $x_1$. Окрестность $Ш'(x_i)$ узла $x_i$ при $i = 1,2,\ldots, N-1$ состоит из двух узлов $x_{i-1}, x_{i+1}$. Окрестность $Ш'(x_N)$ узла $x_N$ является пустым множеством. Свойства положительности коэффициентов выполнены, т. к. $A(x_0) = B(x_0, x_1) = 1$, $D(x_0) = 0$, $A(x_i) = 1, B(x_i, x_{i\pm1}) = 0{,}5$, $D(x_i) = 0$, $i = 1,2,\ldots, N-1$. $A(x_N) = D(x_N) = 1$. Таким образом, рассмотренная схема является монотонной.

**Пример.** Рассмотрим краевую задачу для уравнения первого порядка

$$\frac{\partial u}{\partial t} + \frac{\partial u}{\partial x} = 0, \quad x > 0, t > 0,$$

$$u(x,0) = u_0(x), \quad x > 0, \quad u(0,t) = \mu_1(t), \ t > 0.$$

В квадранте $x > 0, t > 0$ введем сетку с шагом $h$ по $x$ и шагом $\tau$ по $t$ и обозначим $x_i = ih, i = 0,1,\ldots,$ $t_n = n\tau, n = 0,1,\ldots, y_i^n = y(x_i, t_n)$. Одна из схем для данной задачи имеет вид

$$\frac{y_i^{n+1} - y_i^n}{\tau} + \frac{y_{i+1}^n - y_i^n}{h} = 0, \quad i = 1,2,\ldots, \quad n = 0,1\ldots, \qquad (8)$$

$y_i^0 = u_0(x_i), i = 0,1,\ldots, \quad y_0^n = \mu_1(t_n), n = 1,2,\ldots$

Записывая уравнение (8) в виде, разрешенном относительно $y_i^{n+1}$, получим

$$y_i^{n+1} = (\alpha+1)y_i^n - \alpha\, y_{i+1}^n, \text{ где } \alpha = \tau/h.$$

Обозначим через $x$ точку $(x_i, t_{n+1})$, а через $\xi_i$ точку $(x_i, t_n)$. Тогда $A(x) = 1,\ B(x,\xi_i) = \alpha + 1, \quad B(x,\xi_{i+1}) = -\alpha < 0$, т. е. условия положительности коэффициентов (7) не выполнены, и схема (8) является немонотонной. □

**3. Принцип максимума.** Сетка $G_h$ называется **связной сеткой**, если для любых двух ее узлов $x_0, x_0'$, таких что по крайней мере один из узлов имеет непустую окрестность, существует множество узлов $x_i \in G_h$ $i = 1,2,\ldots,m$, такое что

$$x_1 \in Ш'(x_0), x_2 \in Ш'(x_1), \ldots, x_0' \in Ш'(x_m),$$

т. е. каждый последующий узел принадлежит окрестности предыдущего. Аналогично определяется понятие связности любого подмножества сетки $G_h$.

**Замечание.** Смысл требования связности состоит в том, чтобы от любого узла $x_0 \in G_h$ можно было перейти к любому другому узлу $x_0' \in G_h$, пользуясь только заданными шаблонами. □

Определим сеточный оператор $L$ формулами

$$Ly(x) = A(x)y(x) - \sum_{\xi \in Ш'(x)} B(x,\xi)y(\xi), \quad x \in G_h.$$

Тогда задачу (5) можно записать в виде

$$Ly(x) = F(x), \quad x \in G_h. \tag{9}$$

Наряду с сеткой $G_h$ будем рассматривать какое-либо ее подмножество $\omega_h$ и обозначим $\overline{\omega}_h = \bigcup_{x \in \omega_h} Ш(x)$.

**Теорема (принцип максимума).** Пусть сетка $G_h$ и ее подмножество $\omega_h$ являются связными, причем $\overline{\omega}_h \subseteq G_h$. Пусть в

$\omega_h$ выполнены условия положительности коэффициентов (7). Тогда, если функция $y(x)$, заданная на $G_h$, не является постоянной на $\overline{\omega}_h$ и

$$Ly(x) \le 0 \text{ при всех } x \in \omega_h$$

( либо $Ly(x) \ge 0$ при всех $x \in \omega_h$), то $y(x)$ не может принимать наибольшего положительного (соответственно наименьшего отрицательного) значения на $\omega_h$ среди всех ее значений на $\overline{\omega}_h$. □

**Замечание.** Принцип максимума является одним из наиболее эффективных методов исследования существования и единственности разностных задач, т. к. условия положительности коэффициентов проверяются достаточно просто, в том числе и для схем с переменными коэффициентами. Однако принцип максимума дает лишь достаточные условия корректности, т. е. разностная задача может быть корректно поставленной в определенном смысле, но не удовлетворять условиям принципа максимума. □

**Следствие 1.** Если при всех $x \in G_h$

1) выполнены условия положительности коэффициентов (7),

2) $Ly(x) \le 0$ ($Ly(x) \ge 0$) и найдется хотя бы один узел $x_0 \in G_h$, в котором

$$D(x_0) > 0, \tag{10}$$

то $y(x) \le 0$ ( $y(x) \ge 0$) для всех $x \in G_h$.

**Доказательство.** Пусть $Ly(x) \le 0$. Если $y(x) \equiv const$ при $x \in G_h$, то в точке $x_0 \in G_h$, в которой $D(x_0) > 0$ получим

$$Ly(x_0) = A(x_0)y(x_0) - \sum_{\xi \in Ш'(x_0)} B(x_0, \xi)y(\xi) =$$

$$= D(x_0)y(x_0) + \sum_{\xi \in Ш'(x_0)} B(x_0, \xi)(y(x_0) - y(\xi)) =$$

$$= D(x_0)y(x_0) \le 0.$$

Следовательно, $y(x) \equiv y(x_0) \le 0$.

Для функции $y(x) \ne const$ проведем доказательство от противного. Предположим, что $y(x) > 0$ хотя бы в одной точке

$x \in G_h$. Тогда в $G_h$ существует положительный максимум функции $y(x)$, что противоречит принципу максимума. Поэтому $y(x) \le 0$.

Аналогично рассматривается случай $Ly(x) \ge 0$.▯

**Замечание.** Если сетка $G_h$ содержит хотя бы одну точку $x_0 \in G_h$ такую, что $Ш'(x_0)$ - пустое множество, то следствие 1 будет верно, т.к.

$$D(x_0) = A(x_0) - \sum_{\xi \in Ш'(x)} B(x_0, \xi) = A(x_0) > 0. \text{▯}$$

**Следствие 2.** Пусть коэффициенты оператора $L$ удовлетворяют условиям (7) при каждом $x \in G_h$ и условию (10). Тогда задача (5) имеет единственное решение.

**Доказательство.** Задача (5) представляет собой систему линейных алгебраических уравнений, в которой число уравнений равно числу неизвестных. Поэтому достаточно показать, что однородное уравнение $Ly(x) = 0$, $x \in G_h$ имеет только тривиальное решение $y(x) \equiv 0$. Для однородного уравнения одновременно выполнены условия $Ly(x) \le 0$ и $Ly(x) \ge 0$. Поэтому по следствию 1 получаем, что в каждой точке $x \in G_h$ выполняются неравенства $y(x) \le 0$ и $y(x) \ge 0$, т. е. $y(x) \equiv 0$ на $G_h$.▯

**Замечание.** Коэффициенты разностной задачи (3), (4) удовлетворяют условиям следствия 2 ($D(x) = 1$ при $x \in \gamma$), поэтому эта задача имеет **единственное решение**.▯

Рассмотрим вспомогательную задачу

$$LY(x) = \overline{F}(x), \ x \in G_h, \tag{11}$$

которая отличается от задачи (9) только правой частью.

**Теорема сравнения.** Пусть при всех $x \in G_h$ выполнены условия положительности коэффициентов (7) и условие (10). Тогда, если

$$|F(x)| \le \overline{F}(x) \text{ для всех } x \in G_h,$$

то $|y(x)| \le Y(x)$ для всех $x \in G_h$.

**Доказательство.** Для функций $v(x) = Y(x) + y(x)$ и $w(x) = Y(x) - y(x)$ имеем $Lv(x) = \overline{F}(x) + F(x) \geq 0$, $Lw(x) = \overline{F}(x) - F(x) \geq 0$. Тогда согласно следствию 1 выполнены неравенства $v(x) \geq 0,\ \ w(x) \geq 0$, т. е. $-Y(x) \leq y(x) \leq Y(x)$.▯

**Замечание.** Функция $Y(x)$ называется **мажорантной функцией** для решения задачи (9). Для получения оценки решения $y(x)$ обычно строят вспомогательную задачу (11) так, чтобы можно было легко найти ее решение, и затем применяют теорему сравнения. ▯

Теорема сравнения позволяет доказать устойчивость решения по граничным условиям для первой краевой задачи

$$Ly(x) = 0,\quad x \in \omega,\ y(x) = \mu(x),\quad x \in \gamma. \tag{12}$$

**Следствие 3 (устойчивость по граничным условиям).** Пусть при всех $x \in \omega$ выполнены условия положительности коэффициентов (7). Тогда для решения задачи (12) справедлива оценка

$$\max_{x \in \omega} |\, y(x) | \leq \max_{x \in \gamma} |\, \mu(x) |. \tag{13}$$

**Доказательство.** Введем вспомогательную задачу

$$LY(x) = 0,\ x \in \omega,\ Y(x) = \alpha,\ x \in \gamma, \tag{14}$$

где $\alpha = \max\limits_{x \in \gamma} |\, \mu(x) |$. Все условия теоремы сравнения выполнены, поэтому $\left| y(x) \right| \leq Y(x)$.

Обозначим $v(x) = \alpha - Y(x)$. Тогда $Lv(x) = L\alpha - LY(x) = L\alpha = D(x)\alpha \geq 0$ для $x \in \omega$ и $v(x) = 0$ при $x \in \gamma$. Согласно следствию 1 имеем $v(x) \geq 0$, т. е. $Y(x) \leq \alpha$. Тем самым, при всех $x \in G_h$ справедливо неравенство $\left| y(x) \right| \leq Y(x) \leq \alpha$, откуда получаем оценку (13). ▯

**4. Оценка погрешности разностной задачи Дирихле.** Запишем задачу (3), (4) в виде (9) и представим ее решение $y(x)$ как сумму $y(x) = \widetilde{y}(x) + \overline{y}(x)$, где $\widetilde{y}(x)$ - решение однородного уравнения с неоднородным граничным условием

$$L\widetilde{y}(x)=0,\quad x\in\omega\,.\,\widetilde{y}(x)=\mu(x),\quad x\in\gamma \tag{15}$$

и $\overline{y}(x)$ - решение неоднородного уравнения с нулевым граничным условием

$$L\overline{y}(x)=F(x),\quad x\in\omega\,,\,\overline{y}(x)=0,\quad x\in\gamma\,. \tag{16}$$

Заметим, что для разностной задачи (3), (4) выполнены все условия принципа максимума, поэтому к задаче (15) можно применить следствие 3, т. е. верна оценка

$$\|\widetilde{y}\|_{C(G_h)}\le\|\mu\|_{C(\gamma)}, \tag{17}$$

где

$$\|\widetilde{y}\|_{C(G_h)}=\max_{x\in G_h}|\widetilde{y}(x)|,\quad \|\mu\|_{C(\gamma)}=\max_{x\in\gamma}|\mu(x)|.$$

Оценим решение задачи (16). Введем мажорантную функцию

$$Y(x)=K\left(l_1^2+l_2^2-x_1^2-x_2^2\right), \tag{18}$$

где $K>0$ - пока произвольная постоянная. Заметим, что $Y(x)\ge 0$ при всех $x\in G_h$ и $LY(x)=-Y_{\bar{x}_1x_1}-Y_{\bar{x}_2x_2}=-4K$, т. е. можно считать, что функция $Y(x)$ является решением краевой задачи

$$LY(x)=\overline{F}(x),\quad x\in\omega\,,Y(x)=\overline{\mu}(x),\quad x\in\gamma\,, \tag{19}$$

где $\overline{F}(x)=4K$ и $\overline{\mu}(x)\ge 0$ - значение функции (18) при $x\in\gamma$. Положим $K={}^{1}\!/_{4}\|F\|_{C(\omega)}$. Тогда для задач (16) и (19) будут выполнены все условия теоремы сравнения, откуда получаем оценку

$$\|\overline{y}\|_{C(G_h)}\le\max_{x\in G_h}Y(x)\le K(l_1^2+l_2^2)=\frac{l_1^2+l_2^2}{4}\|F\|_{C(\omega)}. \tag{20}$$

Из неравенства треугольника и оценок (17), (20) следует **оценка решения задачи Дирихле**

$$\|y\|_{C(G_h)}\le\frac{l_1^2+l_2^2}{4}\|f\|_{C(\omega)}+\|\mu\|_{C(\gamma)}. \tag{21}$$

Данная оценка выражает собой **устойчивость** разностной схемы (3), (4) по правой части и по граничным условиям.

Найдем оценку погрешности схемы. Обозначим $z_{ij}=y_{ij}-u(x_1^i,x_2^j)$, где $y_{ij}$ - решение разностной задачи (3), (4) и $u(x_1,x_2)$ - решение исходной дифференциальной задачи (1), (2).

Подставляя $y_{ij} = z_{ij} + u(x_1^i, x_2^j)$ в уравнения (3), (4) получим, что погрешность $z_{ij}$ удовлетворяет системе уравнений

$$z_{\bar{x}_1 x_1, ij} + z_{\bar{x}_2 x_2, ij} = -\psi_{ij}, \; x_{ij} \in \omega, \quad z_{ij} = 0, \quad x_{ij} \in \gamma, \tag{22}$$

где $\psi_{ij} = u_{\bar{x}_1 x_1, ij} + u_{\bar{x}_2 x_2, ij} + f_{ij}$ - погрешность аппроксимации разностной схемы на решении задачи (1), (2). Пусть четвертые производные решения $u(x_1, x_2)$ ограничены. Тогда существует постоянная $M_1$, не зависящая от $h_1$ и $h_2$ и такая, что

$$\|\psi\|_{C(\omega)} \le M_1(h_1^2 + h_2^2). \tag{23}$$

Заметим, что задача (22) отличается от разностной схемы (3), (4) только правыми частями в основном уравнении и в граничном условии. Поэтому для решения задачи (22) справедлива оценка, аналогичная (21), т. е. оценка

$$\|z\|_{C(G_h)} \le \frac{l_1^2 + l_2^2}{4} \|\psi\|_{C(\omega)}. \tag{24}$$

Отсюда и из (23) получаем неравенство

$$\|z\|_{C(G_h)} \le M_2(h_1^2 + h_2^2), \; M_2 = \frac{l_1^2 + l_2^2}{4} M_1,$$

откуда следует, что **схема** (3), (4) **сходится и имеет второй порядок точности**.

**5. Задача.** Изобразить шаблон, выписать каноническую форму и проверить условия положительности коэффициентов для разностной схемы

$$\frac{y_i^{n+1} - y_i^n}{\tau} = \frac{y_{i+1}^{n+1} - 2y_i^{n+1} + y_{i-1}^{n+1}}{h^2} + \varphi_i^n,$$

$$i = 1,2,\dots N-1, \; n = 0,1,\dots,K-1,$$

$$y_0^{n+1} = \mu_1(t_{n+1}), \; y_N^{n+1} = \mu_2(t_{n+1}), \; n = 0,1,\dots,K-1,$$

$$y_i^0 = u_0(x_i), \; i = 0,1,\dots,N. \; \square$$

**Задача.** Показать, что для уравнения

$$\frac{y_i^{n+1} - 2y_i^n + y_i^{n-1}}{\tau^2} = \frac{y_{i+1}^n - 2y_i^n + y_{i-1}^n}{h^2}$$

принцип максимума не выполняется. $\square$

## Вопрос 29.

## Двуслойные разностные схемы для уравнения теплопроводности: построение, исследование погрешности аппроксимации.

**1.** Пусть дана первая краевая задача для уравнения теплопроводности с постоянными коэффициентами

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} + f(x,t), \quad 0 < x < 1, 0 < t \le T \tag{1}$$

$$u(x,0) = u_0(x), \quad 0 \le x \le 1, \tag{2}$$

$$u(0,t) = \mu_1(t), \quad u(1,t) = \mu_2(t), \quad 0 \le t \le T, \tag{3}$$

где $u_0(x), \mu_1(t), \mu_2(t)$ - заданные функции. В дальнейшем будем предполагать, что решение $u(x,t)$ обладает необходимым числом производных по $x$ и по $t$.

Для построения разностной схемы введем **пространственно-временную сетку** $\omega_{h,\tau} = \omega_h \times \omega_\tau$, где $\omega_h = \{x_i = ih, i = 0,1,\dots,N,\ Nh = 1\}$, $\omega_\tau = \{t_n = n\tau, n = 0,1,\dots,K,\ K\tau = T\}$. Точки $(x_i,t_n), i = 0,1,\dots,N,\ n = 0,1,\dots,K$, образуют узлы этой сетки. **Слоем** будем называть множество всех узлов сетки, имеющих одну и ту же временную координату. Например, $n$-ым слоем называется множество узлов $(x_0,t_n),(x_1,t_n),\dots,(x_N,t_n)$.

Для функции $y(x,t)$, определенной на сетке $\omega_{h,\tau}$, обозначим

$$y_i^n = y(x_i,t_n), \qquad y_{t,i}^n = \frac{y_i^{n+1} - y_i^n}{\tau}, \quad y_{\bar{x}x,i}^n = \frac{y_{i+1}^n - 2y_i^n + y_{i-1}^n}{h^2}.$$

Чтобы аппроксимировать уравнение (1) в точке $(x_i,t_n)$, заменим производную $\partial u/\partial t$ разностным отношением $y_{t,i}^n$, а производную $\partial^2 u / \partial x^2$ – второй разностной производной $y^n_{\bar{x}x,i}$. Правую часть заменим приближенно сеточной функцией $\varphi^n_i$.

**Замечание.** В качестве $\varphi^n_i$ можно взять одно из следующих выражений: $f(x_i, t_n)$, $\dfrac{1}{h}\displaystyle\int\limits_{x_{i-1/2}}^{x_{i+1/2}} f(x, t_n)dx$, $\dfrac{1}{h\tau}\displaystyle\int\limits_{t_n}^{t_{n+1}} dt \int\limits_{x_{i-1/2}}^{x_{i+1/2}} f(x,t)dx$.□

Начальное и граничные условия (2), (3) аппроксимируем точно. Далее, зададим произвольный действительный параметр $\sigma$ и определим **однопараметрическое семейство схем с весами** следующим образом:

$$y_{t,i} = \sigma y^{n+1}_{\bar{x}x,i} + (1-\sigma) y^n_{\bar{x}x,i} + \varphi^n_i, \tag{4}$$

$$i = 1,2,\ldots,N-1, \quad n = 0,1,\ldots,K-1,$$

$$y^0_i = u_0(x_i), \quad i = 0,1,\ldots,N,$$

$$y^{n+1}_0 = \mu_1(t_{n+1}), \quad y^{n+1}_N = \mu_2(t_{n+1}), \quad n = 0,1,\ldots,K-1.$$

Схема (4) содержит значения искомой функции $y$ на двух слоях и поэтому называется **двуслойной схемой**.

При $\sigma = 0$ получим из (4) **явную схему**. Находить решение такой системы нужно по слоям. Решение на нулевом слое задано начальными условиями $y^0_i = u_0(x_i)$, $i = 0,1,\ldots,N$. Если решение $y^n_i$, $i = 0,1,\ldots,N$ на слое $n$ уже найдено, то решение $y^{n+1}_i$ на слое $n+1$ находится по явной формуле

$$y^{n+1}_i = y^n_i + \tau(y^n_{\bar{x}x,i} + \varphi^n_i), \quad i = 1,2,\ldots,N-1,$$

а значения $y^{n+1}_0 = \mu_1(t_{n+1})$, $y^{n+1}_N = \mu_2(t_{n+1})$ доопределяются из граничных условий.

При $\sigma = 1$ схема (4) называется **чисто неявной схемой (схемой с опережением)**. Как и в случае явной схемы, решение системы находится по слоям, начиная с $n = 1$. Однако теперь для нахождения $y^{n+1}_i$ по известным $y^n_i$ требуется решить систему уравнений

$$\gamma\, y^{n+1}_{i+1} - (1+2\gamma)\, y^{n+1}_i + \gamma\, y^{n+1}_{i-1} = -F^n_i, \quad i = 1,2,\ldots,N-1,$$

$$y_0^{n+1} = \mu_1(t_{n+1}), \quad y_N^{n+1} = \mu_2(t_{n+1}), \quad n = 0,1,\ldots,K-1,$$

где $\gamma = \tau/h^2$, $F_i^n = y_i^n + \tau\varphi_i^n$, $\varphi_i^n = f(x_i, t_{n+1}) + O(\tau + h^2)$.

Эту систему можно решать методом прогонки, т. к. условия устойчивости прогонки выполнены.

При $\sigma = 1/2$ получим из (4) **шеститочечную симметричную схему**, для решения которой можно также использовать метод прогонки (докажите этот факт самостоятельно).

**Замечание.** Разностную схему (4) можно построить другим способом, например, **интегро-интерполяционным методом**. В этом случае исходят из дифференциального уравнения (1) как уравнения баланса тепла. Входящие в это уравнение производные заменяют приближенными разностными выражениями. В результате получают так называемую **консервативную** разностную схему, т. е. схему, выражающую на сетке законы сохранения. При написании разностных схем следует добиваться того, чтобы эти схемы были консервативными, иначе можно получить расходящуюся схему. □

**Пример.** Приведем пример, поясняющий это замечание. Рассмотрим дифференциальную задачу для стационарного уравнения теплопроводности с переменными коэффициентами:

$$\frac{\partial}{\partial x}\left(k(x)\frac{\partial u}{\partial x}\right) = 0, \quad 0 < x < 1, \quad u(0) = 1, \quad u(1) = 0. \quad (5)$$

На сетке $\omega_h$ заменим задачу (5) разностной схемой

$$k_i \frac{y_{i+1} - 2y_i + y_{i-1}}{h^2} + \frac{k_{i+1} - k_{i-1}}{2h} \cdot \frac{y_{i+1} - y_{i-1}}{2h} = 0, \quad (6)$$

$$k_i = k(x_i), \quad i = 1,2,\ldots,N-1, \; y_0 = 1, \quad y_N = 0.$$

Покажем, что схема (6) расходится в классе кусочно-постоянных коэффициентов

$$k(x) = \begin{cases} k_1, & 0 < x < \xi, \\ k_2, & \xi < x < 1, \end{cases}$$

где $\xi = x_n + \theta h$, $x_n = nh$, $0 < \theta < 1$.

Так как $k(x)$ имеет разрыв первого рода в точке $x = \xi$, так что $[k] = k(\xi + 0) - k(\xi - 0) \neq 0$, то в этой точке ставятся условия сопряжения $[u]=0, \left[k\dfrac{\partial u}{\partial x}\right]=0$. Точным решением задачи (5), удовлетворяющим условиям сопряжения, будет функция

$$u(x)=\begin{cases}1-\alpha_0 x, & 0\le x\le\xi,\\ \beta_0(1-x), & \xi\le x\le 1,\end{cases}$$

где $\alpha_0=\left(\gamma+(1-\gamma)\xi\right)^{-1}$, $\beta_0=\gamma\alpha_0$, $\gamma=k_1/k_2$, а решением разностной задачи (6) - функция

$$y(x_i)=\begin{cases}1-\alpha x_i, & 0\le x\le x_n,\\ \beta(1-x_i), & x_{n+1}\le x\le 1,\end{cases} \tag{7}$$

где $\alpha=\dfrac{1}{\mu+(1-\mu)\xi+h(\lambda-\theta-(1-\theta)\mu)}$, $\quad \mu=\dfrac{3+\gamma}{5-\gamma}\lambda$,

$\lambda=\dfrac{5\gamma-1}{3\gamma+1}$, $\beta=\mu\alpha$. Предельный переход при $h\to 0$ дает

$\lim\limits_{h\to 0}\alpha=\overline{\alpha}_0$, $\lim\limits_{h\to 0}\beta=\overline{\beta}_0$, $\overline{\alpha}_0=\left(\mu+(1-\mu)\xi\right)^{-1}$, $\overline{\beta}_0=\mu\overline{\alpha}_0$.

Функции (7) доопределим на всем отрезке при помощи линейной интерполяции и получим функцию $\widetilde{y}(x,h), x\in[0,1]$, совпадающую с $y(x_i)$ в узлах $x_i$. Найдем предел $\widetilde{y}(x,h)$ при $h\to 0$:

$$\overline{u}(x)=\lim_{h\to 0}\widetilde{y}(x,h)=\begin{cases}1-\overline{\alpha}_0 x, & 0\le x\le\xi,\\ \overline{\beta}_0(1-x), & \xi\le x\le 1.\end{cases}$$

Сравнивая предельную функцию $\overline{u}(x)$ с точным решением $u(x)$, видим, что $\overline{u}(x)=u(x)$ при $\overline{\alpha}_0=\alpha_0$, $\overline{\beta}_0=\beta_0$, а это возможно лишь при $\gamma=1$ или $k_1=k_2$.

Таким образом, решение разностной задачи (6) при $h\to 0$ стремится к функции $\overline{u}(x)$, которая в случае $k_1\ne k_2$ отлична от точного решения задачи (5). Следовательно, схема (6) расходится.

Выясним физический смысл функции $\overline{u}(x)$. Эта функция является решением задачи (5), удовлетворяющим при $x=\xi$

условиям $\left[\bar{u}\right]=0,\left[k\dfrac{\partial\bar{u}}{\partial x}\right]=-\bar{\alpha}_0(\mu-\gamma)k_2=q$, где $q$ – мощность сосредоточенного источника (стока) тепла в точке $x=\xi$. Величина $q$ меняется в широких пределах в зависимости от $\gamma$. Поэтому физическая причина расходимости схемы (6) в том, что она нарушает баланс тепла, приводя к появлению дополнительного источника (при $q<0$) или стока (при $q>0$) тепла в точке $x=\xi$. □

**2.** Исследуем погрешность аппроксимации схемы (4) на решении $u(x_i,t_n)$ исходной задачи (1) – (3). Подставляя функцию $y_i^n=z_i^n+u(x_i,t_n)$ в разностную задачу (4), получим следующую систему уравнений для погрешности схемы $z_i^n$:

$$z_{t,i}^n=\sigma z_{\bar{x}x,i}^{n+1}+(1-\sigma)z_{\bar{x}x,i}^n+\psi_i^n,$$

$$i=1,2,\ldots,N-1,\quad n=0,1,\ldots,K-1,$$

$$z_i^0=0,\quad i=0,1,\ldots,N,\; z_0^{n+1}=z_N^{n+1}=0,\quad n=0,1,\ldots,K-1,$$

где $\psi_i^n=\sigma u_{\bar{x}x,i}^{n+1}+(1-\sigma)u_{\bar{x}x,i}^n-u_{t,i}^n+\varphi_i^n$ - погрешность аппроксимации схемы (4) на решении задачи (1) – (3). Обозначим $\dot{u}=\partial u/\partial t$, $u''=\partial^2u/\partial x^2$, $t_{n+1/2}=t_n+0{,}5\tau$ и учтем следующие разложения по формуле Тейлора: $u_{t,i}^n=\dot{u}(x_i,t_{n+1/2})+O(\tau^2)$, $u_{\bar{x}x,i}=u''(x_i)+\dfrac{h^2}{12}u^{(4)}(x_i)+O(h^4)$. Тогда получим

$$\psi_i^n=\sigma\left(u''(x_i,t_{n+1})+\frac{h^2}{12}u^{(4)}(x_i,t_{n+1})\right)+(1-\sigma)\Big(u''(x_i,t_n)+$$

$$+\frac{h^2}{12}u^{(4)}(x_i,t_n)\Big)-\dot{u}(x_i,t_{n+1/2})+\varphi_i^n+O(\tau^2+h^4).$$

Отсюда, проводя разложение в точке $(x_i, t_{n+1/2})$ и обозначая $u = u(x_i, t_{n+1/2})$, будем иметь

$$\psi_i^n = \sigma\left(u'' + \frac{\tau}{2}\dot{u}'' + \frac{h^2}{12}u^{(4)}\right) + (1-\sigma)\left(u'' - \frac{\tau}{2}\dot{u}'' + \frac{h^2}{12}u^{(4)}\right) -$$

$$- \dot{u} + \varphi_i^n + O(\tau^2 + h^4) = (u'' - \dot{u} + \varphi_i^n) + (\sigma - 0{,}5)\tau\,\dot{u}'' +$$

$$+ \frac{h^2}{12}u^{(4)} + O(\tau^2 + h^4).$$

Дифференцируя уравнение (1) дважды по переменной $x$, получим $u^{(4)} = \dot{u}'' - f''$. Следовательно

$$\psi_i^n = \left((\sigma - 0{,}5)\tau + \frac{h^2}{12}\right)\dot{u}'' + \varphi_i^n - f(x_i, t_{n+1/2}) - \tag{8}$$

$$- \frac{h^2}{12} f''(x_i, t_{n+1/2}) + O(\tau^2 + h^4).$$

Из формулы (8) можно сделать следующие выводы. Если

$$\sigma = \sigma_* = \frac{1}{2} - \frac{h^2}{12\tau}$$

и

$$\varphi_i^n = f(x_i, t_{n+1/2}) + \frac{h^2}{12} f''(x_i, t_{n+1/2}) + O(\tau^2 + h^4),$$

то схема (4) имеет второй порядок аппроксимации по $\tau$ и четвертый – по $h$. Такая схема называется **схемой повышенного порядка аппроксимации.**

Если $\sigma = 0.5$, $\varphi_i^n = f(x_i, t_{n+1/2}) + O(\tau^2 + h^2)$, то схема (4) имеет второй порядок аппроксимации по $\tau$ и по $h$. При остальных значениях $\sigma$ и при $\varphi_i^n = f(x_i, t_{n+1}) + O(\tau + h^2)$ схема (4) имеет первый порядок аппроксимации по $\tau$ и второй - по $h$. В частности, явная схема $(\sigma = 0)$ и схема с опережением $(\sigma = 1)$ имеют аппроксимацию $O(\tau + h^2)$.

**3. Задача.** Определить порядок аппроксимации разностного уравнения

$$\frac{1}{12}\frac{y_{i+1}^{n+1}-y_{i+1}^{n}}{\tau}+\frac{5}{6}\frac{y_{i}^{n+1}-y_{i}^{n}}{\tau}+\frac{1}{12}\frac{y_{i-1}^{n+1}-y_{i-1}^{n}}{\tau}=$$

$$=\frac{1}{2}\left(\frac{y_{i-1}^{n+1}-2y_{i}^{n+1}+y_{i+1}^{n+1}}{h^2}+\frac{y_{i-1}^{n}-2y_{i}^{n}+y_{i+1}^{n}}{h^2}\right)$$

на решении дифференциального уравнения $\dfrac{\partial u}{\partial t}=\dfrac{\partial^2 u}{\partial x^2}$. □

**Задача.** Найти для $x_i = 0.2\,i,\quad i = 0,1,2,3$ решение уравнения

$$\frac{\partial u}{\partial t}=\frac{\partial^2 u}{\partial x^2},\quad 0<x<0.6,\quad 0<t\le 0.08,$$

удовлетворяющее условиям

$$u(x,0)=x^2/2,\quad 0\le x\le 0.6,$$

$$u(0,t)=t,\quad u(0.6,t)=0.18+t,\quad 0\le t\le 0.08$$

по явной разностной схеме, полагая $h=0.2,\ \tau=0.04$. □

## Вопрос 30.

## Исследование устойчивости по начальным данным схемы с весами для уравнения теплопроводности.

**1.** Будем рассматривать первую краевую задачу для уравнения теплопроводности

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2}, \quad 0 < x < l, 0 < t \le T,$$
$$u(x,0) = u_0(x), \quad 0 \le x \le l, \qquad (1)$$
$$u(0,t) = u(l,t) = 0, \quad 0 \le t \le T.$$

Введем сетку $\omega_{h,\tau} = \omega_h \times \omega_\tau$, где $\omega_h = \{x_i = ih,$ $i = 0,1,\ldots,N,\ Nh = l\}$, $\omega_\tau = \{t_n = n\tau,\ n = 0,1,\ldots,K,\ K\tau = T\}$ и обозначим

$$y_i^n = y(x_i, t_n), \quad y_{t,i}^n = \frac{y_i^{n+1} - y_i^n}{\tau}, \quad y_{\bar{x}x,i}^n = \frac{y_{i+1}^n - 2y_i^n + y_{i-1}^n}{h^2}.$$

Дифференциальную задачу (1) заменим на сетке $\omega_{h,\tau}$ **схемой с весами**

$$y_{t,i}^n = \sigma y_{\bar{x}x,i}^{n+1} + (1-\sigma) y_{\bar{x}x,i}^n, \qquad (2)$$
$$i = 1,2,\ldots,N-1, \quad n = 0,1,\ldots,K-1,$$
$$y_i^0 = u_0(x_i), \quad i = 0,1,\ldots,N, \qquad (3)$$
$$y_0^n = y_N^n = 0, \quad n = 0,1,\ldots,K-1,$$

где $\sigma$ - некоторый вещественный параметр.

**2.** Рассмотрим пространство $H$ функций $y$, заданных на сетке $\omega_h$ и обращающихся в нуль на границе: $y_0 = y_N = 0$. Введем **оператор второй разностной производной** формулами

$$(Ay)_i = -y_{\bar{x}x,i}, \quad i = 1,2,\ldots,N-1,\ y_0 = y_N = 0. \qquad (4)$$

Решим задачу на собственные значения для этого оператора, т. е. найдем такие числа $\lambda$ (**собственные значения**), для которых уравнение

$$Ay = \lambda y \tag{5}$$

имеет нетривиальные решения (**собственные функции**), и найдем собственные функции.

**Замечание.** Уравнение (5) для оператора (4) представляет собой алгебраическую задачу на собственные значения для матрицы

$$A = \frac{1}{h^2}\begin{bmatrix} 2 & -1 & 0 & 0 & \dots & 0 & 0 & 0 \\ -1 & 2 & -1 & 0 & \dots & 0 & 0 & 0 \\ 0 & -1 & 2 & -1 & \dots & 0 & 0 & 0 \\ . & . & . & . & \dots & . & . & . \\ 0 & 0 & 0 & 0 & \dots & -1 & 2 & -1 \\ 0 & 0 & 0 & 0 & \dots & 0 & -1 & 2 \end{bmatrix},$$

которая называется **матрицей разностного оператора** (4). Матрица $A$ является трехдиагональной симметричной матрицей порядка $N-1$. Известно, что у такой матрицы существуют $N-1$ действительных собственных значений и столько же линейно независимых собственных векторов. ▯

Запишем уравнение (5) в виде

$$-y_{\bar{x}x,\,i} = \lambda y_i, \quad i = 1,2,\dots,N-1, \; y_0 = y_N = 0$$

или подробнее:

$$y_{i+1} + y_{i-1} = (2 - h^2\lambda)y_i, \quad i = 1,2,\dots,N-1, \quad y_0 = y_N = 0. \tag{6}$$

Будем искать собственные функции задачи (6) в виде

$$y_k(x_i) = \sin\frac{\pi k x_i}{l}, \quad x_i = ih, \quad k = 1,2,\cdots \tag{7}$$

Граничные условия $y_0 = y_N = 0$ при этом выполнены. Подставляя (7) в уравнение (6), получим

$$\sin\frac{\pi k(x_i + h)}{l} + \sin\frac{\pi k(x_i - h)}{l} = (2 - h^2\lambda)\sin\frac{\pi k x_i}{l}$$

или

$$2\sin\frac{\pi k x_i}{l}\cos\frac{\pi k h}{l} = (2 - h^2\lambda)\sin\frac{\pi k x_i}{l}.$$

Следовательно, функция (7) будет собственной функцией оператора (4), если $2\cos\dfrac{\pi k h}{l} = 2 - h^2\lambda$, т. е.

$$\lambda = \lambda_k = \frac{2}{h^2}\left(1 - \cos\frac{\pi k h}{l}\right) = \frac{4}{h^2}\sin^2\frac{\pi k h}{2l}. \qquad (8)$$

При $k = 1,2,\ldots,N-1$ получаем $N-1$ различных действительных собственных значений и отвечающих им собственных функций (7).

**Замечание.** Из формулы (8) видно, что $0 < \lambda_1 < \lambda_2 < \ldots < \lambda_{N-1} < \dfrac{4}{h^2}$. ▯

Введем в пространстве $H$ скалярное произведение $(y,v) = \sum_{i=1}^{N-1} y_i v_i h$ и норму $\|y\| = \left(\sum_{i=1}^{N-1} h y_i^2\right)^{1/2}$. Используя разностные формулы Грина, покажем, что оператор (4) будет самосопряженным в $H$. Действительно, т. к. $v_0 = v_N = 0$, то

$$(Ay,v) = -\sum_{i=1}^{N-1} y_{\bar{x}x,i} v_i h = -\sum_{i=1}^{N-1} y_{x,i} v_i + \sum_{i=1}^{N-1} y_{\bar{x},i} v_i = \sum_{i=1}^{N-1} y_{\bar{x},i} v_i -$$

$$-\sum_{i=2}^{N} y_{\bar{x},i} v_{i-1} = \sum_{i=1}^{N} y_{\bar{x},i} v_{\bar{x},i} h.$$

Здесь обозначено $y_{\bar{x},i} = \dfrac{y_i - y_{i-1}}{h}$, $y_{x,i} = \dfrac{y_{i+1} - y_i}{h}$.

Меняя $y$ и $v$ местами, получим

$$(Av,y) = \sum_{i=1}^{N} v_{\bar{x},i} y_{\bar{x},i} h = (v, Ay),$$

откуда видно, что оператор (4) является **самосопряженным**.

Следствием самосопряженности данного оператора является ортогональность его собственных функций, отвечающих различным собственным значениям, т. е. система собственных функций (7) образует ортогональный базис в пространстве $H$.

Вычислим квадрат нормы собственной функции $y_k(x_i)$. По определению имеем

$$\|y_k\|^2 = \sum_{i=1}^{N-1} h \sin^2 \frac{\pi k x_i}{l} = \sum_{i=1}^{N} h \sin^2 \frac{\pi k x_i}{l}.$$

Используя формулу

$$\sin^2 \frac{\pi k x_i}{l} = \frac{1}{2}\left(1 - \cos \frac{2\pi k x_i}{l}\right)$$

и учитывая тождество

$$\sin \frac{2\pi k (x_i + 0.5h)}{l} - \sin \frac{2\pi k (x_{i-1} + 0.5h)}{l} = 2 \sin \frac{\pi k h}{l} \cos \frac{2\pi k x_i}{l},$$

получим

$$\|y_k\| = \frac{hN}{2} - \frac{1}{4\sin(\pi k h / l)}\left(\sin \frac{2\pi k (x_N + 0.5h)}{l} - \sin \frac{\pi k h}{l}\right) = \frac{l}{2}.$$

Таким образом, система собственных функций

$$\mu_k(x_i) = \sqrt{\frac{2}{l}} \sin \frac{\pi k x_i}{l}, \quad k = 1,2,\ldots,N-1,\ i = 1,2,\ldots,N-1 \quad (9)$$

образует **ортонормированный базис** в $H$.

**3.** Используя оператор второй разностной производной (4), запишем схему (2), (3) в операторном виде

$$\frac{y^{n+1} - y^n}{\tau} + \sigma A y^{n+1} + (1-\sigma) A y^n = 0, \quad y^0 = u^0, \qquad (10)$$

где $y^n = \left(y_1^n, y_2^n, \ldots, y_{N-1}^n\right)^T$, $u^0 = \left(u_0(x_1), u_0(x_2), \ldots, u_0(x_{N-1})\right)^T$, или в эквивалентном виде

$$(E + \sigma \tau A) y^{n+1} = (E - (1-\sigma)\tau A) y^n, \quad y^0 = u^0,$$

где $E$ - единичный оператор. Уравнение (10) имеет решение, если оператор $B = E + \sigma \tau A$ имеет обратный. В дальнейшем предполагаем, что выполнены неравенства

$$1 + \sigma \tau \lambda_k > 0, \quad k = 1,2,\ldots,N-1, \qquad (11)$$

гарантирующие существование $\left(E + \sigma \tau A\right)^{-1}$. Здесь $\lambda_k > 0$ - собственные значения оператора (4).

Будем искать решение задачи (2), (3) в виде разложения по ортонормированному базису (9) из собственных функций оператора (4). При каждом $n$ решение $y_i^n = y(x_i, t_n)$ можно представить в виде

$$y_i^n = \sum_{k=1}^{N-1} c_k(t_n)\mu_k(x_i), \qquad (12)$$

где $c_k(t_n)$ – коэффициенты Фурье функции $y_i^n$. Подставим (12) в уравнение (2) и учтем, что $(\mu_k(x_i))_{\bar{x}x,\,i} = -\lambda_k\mu_k(x_i)$. Тогда получим

$$\sum_{k=1}^{N-1}\mu_k(x_i)\left(\frac{c_k(t_{n+1}) - c_k(t_n)}{\tau} + \sigma\lambda_k c_k(t_{n+1}) + (1-\sigma)\lambda_k c_k(t_n)\right) = 0 .$$

В силу линейной независимости функций $\mu_k(x_i)$ отсюда следует, что

$$\frac{c_k(t_{n+1}) - c_k(t_n)}{\tau} + \sigma\lambda_k c_k(t_{n+1}) + (1-\sigma)\lambda_k c_k(t_n) = 0 ,$$
$$k = 1,2,\ldots,N-1, \quad n = 0,1,\ldots,K-1.$$

Из этого уравнения получаем

$$c_k(t_{n+1}) = q_k c_k(t_n), \;\; q_k = \frac{1-(1-\sigma)\tau\lambda_k}{1+\sigma\tau\lambda_k},$$

причем выражение, стоящее в знаменателе, положительно в силу условия (11).

Таким образом, решение $y_i^{n+1}$ задачи (2), (3) можно записать в виде

$$y_i^{n+1} = \sum_{k=1}^{N-1} q_k c_k(t_n)\mu_k(x_i) . \qquad (13)$$

Оценим норму функции $y^{n+1}$. Из разложения (13) и ортонормированности базиса $\mu_k(x_i)$ получаем

$$\left\|y^{n+1}\right\|^2 = \sum_{i=1}^{N-1}(y_i^{n+1})^2 h = \sum_{k=1}^{N-1} q_k^2 (c_k(t_n))^2 .$$

Поэтому

$$\left\| y^{n+1} \right\| \le \left( \sum_{k=1}^{N-1} (c_k(t_n))^2 \right)^{1/2} \max_{1\le k\le N-1} |q_k| = \left\| y^n \right\| \max_{1\le k\le N-1} |q_k|.$$

Потребуем, чтобы выполнялось условие $|q_k| \le 1,\ k = 1,2,\ldots,N-1$, которое эквивалентно условию

$$\sigma \ge \frac{1}{2} - \frac{1}{\tau \lambda_{N-1}}, \tag{14}$$

где $\lambda_{N-1} = \dfrac{4}{h^2}\cos^2\dfrac{\pi h}{2l}$ - наибольшее собственное значение оператора (4).

**Замечание.** Из условия (14) при любом $k = 1,2,\ldots,N-1$ следует неравенство

$$1 + \sigma\tau\lambda_k \ge \frac{\tau\lambda_k}{2} > 0,$$

т. е. неравенство (11). ▯

Условие (14) будет выполнено, если потребовать

$$\sigma \ge \frac{1}{2} - \frac{h^2}{4\tau}. \tag{15}$$

Итак, если выполнено (15), то справедлива оценка

$$\left\| y^{n+1} \right\| \le \left\| y^n \right\| \le \left\| y^{n-1} \right\| \le \ldots \le \left\| y^0 \right\|,$$

которая означает, что схема (2), (3) **устойчива по начальным данным** в норме

$$\left\| y \right\| = \left( \sum_{i=1}^{N-1} h y_i^2 \right)^{1/2}.$$

**Замечание.** При $\sigma = 0$ (явная схема) из условия (15) получаем условие устойчивости в виде

$$\frac{\tau}{h^2} \le \frac{1}{2},$$

а при всех $\sigma \ge 1/2$ схемы (2), (3) будут **абсолютно устойчивыми,** т. е. устойчивыми при любых шагах $h$ и $\tau$. ▯

**Примеры.**

**1)** Очень многие двуслойные схемы, аппроксимирующие уравнение теплопроводности на шеститочечном шаблоне, сводятся к схеме с весами (2) и, следовательно, нет необходимости отдельно исследовать их устойчивость. Достаточно проверить выполнения условия (15) для данной схемы. Рассмотрим один из таких примеров. Схема

$$2\gamma\, y_i^{n+1} = (\gamma - 0.5)\left(y_{i+1}^{n+1} + y_{i-1}^{n+1}\right) + 0.5\left(y_{i+1}^{n} + y_{i-1}^{n}\right),$$

$$i = 1,2,\ldots, N-1, \quad n = 0,1,\ldots, \quad \gamma = \tau/h^2,$$

$$y_i^0 = u_0(x_i), \quad i = 0,1,\ldots, N,$$

$$y_0^n = y_N^n = 0, \quad n = 0,1,\ldots$$

приводится к виду

$$y_{t,i}^n = \left(1 - \frac{1}{2\gamma}\right) y_{\bar{x}x,i}^{n+1} + \frac{1}{2\gamma}\, y_{\bar{x}x,i}^{n},$$

что соответствует схеме с весами (2) при $\sigma = \left(1 - \dfrac{1}{2\gamma}\right)$. Сравнивая с (15), видим, что рассматриваемая схема устойчива при условии $\gamma \ge 0.5$. Это довольно необычное условие устойчивости, так как чаще всего встречаются схемы, устойчивые при условии ограниченности величины $\gamma$ сверху.

**2)** Приведем пример абсолютно неустойчивой схемы, аппроксимирующей уравнение теплопроводности:

$$\frac{1}{2}\left(\frac{y_{i-1}^{n+1} - y_{i-1}^{n}}{\tau} + \frac{y_{i+1}^{n+1} - y_{i+1}^{n}}{\tau}\right) = \frac{y_{i+1}^{n} - 2y_i^n + y_{i-1}^{n}}{h^2},$$

$$i = 1,2,\ldots, N-1, \quad n = 0,1,\ldots$$

Эта двуслойная схема является частным случаем схемы с весами (2) при $\sigma = -\dfrac{h^2}{2\tau}$ и, следовательно, согласно (14), условие ее устойчивости имеет вид

$$-\frac{h^2}{2\tau} \ge \frac{1}{2} - \frac{h^2}{4\tau\cos^2(\pi h/2l)}$$

или

$$1+\frac{1}{\gamma}\frac{2\cos^2(\pi h/2l)-1}{2\cos^2(\pi h/2l)}\le 0, \qquad \gamma=\tau/h^2 \ .$$

Видно, что при всех достаточно малых $h$ это неравенство не выполняется, т.е. схема неустойчива. 

**4. Задача.** На равномерной сетке с шагом $h$ по $x$ и шагом $\tau$ по $t$ построить явную разностную схему для задачи

$$u_t=a^2u_{xx}, \quad 0<x<1, \quad t>0,$$

$$u(0,t)=u(1,t)=0, \ t\ge 0, \quad u(x,0)=u_0(x), \ 0\le x\le 1$$

и исследовать ее устойчивость при

$$h=0.05\,, \tau=0.00009\,, a^2=15\,.$$

**Задача.** Является ли устойчивой по начальным данным схема

$$y_i^{n+1}=\frac{1}{3}\left(y_{i+1}^n+y_i^n+y_{i-1}^n\right), \quad i=1,2,\ldots,N-1,$$

$$y_0^n=y_N^n=0 \ ?$$

**Задача.** Исследовать устойчивость схемы с весами для нестационарного уравнения Шредингера

$$i\,y_{t,j}^n=\sigma\, y_{\bar{x}x,j}^{n+1}+(1-\sigma)\,y_{\bar{x}x,j}^n, \quad \sigma=\sigma_0+i\sigma_1,$$

$$j=1,2,\ldots,N-1, \quad hN=1, \quad n=1,2,\ldots,$$

$$y_j^0=u_0(x_j), \quad y_0^{n+1}=y_N^{n+1}=0,$$

где $i$ - мнимая единица. 

## От авторов

При написании данного пособия мы пользовались в основном следующей литературой (не вся она является обязательной для подготовки к государственному экзамену):

## Литература.


1. **Ильин В.А., Позняк Э.Г.** Основы математического анализа, части I, II. – М.: Наука, 1971, 1973.
2. **Никольский С.М.** Курс математического анализа, т.I, II. – М.: Наука, 1990, 1991.
3. **Фихтенгольц Г.М.** Курс дифференциального и интегрального исчисления, т.I, II, III. – М.: Наука, 1969, 1970.
4. **Смирнов В.И.** Курс высшей математики , т. I-V. – М.: Наука, 1974-1981.
5. **Васильев Ф.П.** Численные методы решения экстремальных задач. – М.: Наука, 1980.
6. **Гелбаум Б., Олмстед Дж.** Контрпримеры в анализе. –М.:Мир, 1967.
7. **Натансон И.П.** Теория функций вещественной переменной. – М.: Наука, 1974.
8. **Свешников А.Г., Тихонов А.Н.** Теория функций комплексной переменной. – М.: Наука, 1974.
9. **Лаврентьев М.А., Шабат Б.В.** Методы теории функций комплексного переменного. – М.: Наука, 1973.
10. **Бицадзе А.В** Основы теории аналитических функций комплексного переменного. – М.: Наука, 1972.
11. **Маркушевич А.И.** Краткий курс теории аналитических функций. – М.: Наука, 1978.
12. **Евграфов М.А.** Аналитические функции. – М.: Наука, 1991.
13. **Ильин В.А., Позняк Э.Г.** Линейная алгебра. – М.: Наука, 1974.
14. **Ильин В.А., Ким Г.Д.** Линейная алгебра и аналитическая геометрия.– М.: МГУ, 2002.
15. **Воеводин В.В.** Линейная алгебра. – М.: Наука, 1974.
16. **Гельфанд И.М.** Лекции по линейной алгебре. – М.: Наука, 1971.
17. **Колмогоров А.Н., Фомин С.В.** Элементы теории функций и функционального анализа. – М.: Наука, 1972.
18. **Люстерник Л.А., Соболев В.И.** Краткий курс функционального анализа. – М.: Высшая школа, 1982.
19. **Канторович Л.В., Акилов Г.П.** Функциональный анализ. – М.: Наука, 1977.
20. **Треногин В.А**. Функциональный анализ. – М.: Наука, 1980.

21. **Вулих Б.З.** Введение в функциональный анализ. – М.: Физматлит, 1958.
22. **Ахиезер Н.И., Глазман И.М.** Теория линейных операторов в гильбертовом пространстве. – М.: Наука, 1966.
23. **Наймарк М.А.** Линейные дифференциальные операторы. – М.: Наука, 1969.
24. **Красносельский М.А., Вайникко Г.М., Забрейко П.П., Рутицкий Я.Б., Стеценко В.Я.** Приближённое решение операторных уравнений – М.: Наука, 1969.
25. **Петровский И.Г.** Лекции по теории интегральных уравнений. – М.: Наука, 1965.
26. **Васильева А.Б., Тихонов Н.А.** Интегральные уравнения. – М.: МГУ, 1989.
27. **Михлин С.Г.** Интегральные уравнения. – М.-Л.: ОГИЗ, 1949.
28. **Ловитт У.В.** Линейные интегральные уравнения. М.:Тех.-теор. лит., 1957.
29. **Краснов М.Л.** Интегральные уравнения. – М.: Наука, 1975.
30. **Тихонов А.Н., Васильева А.Б., Свешников А.Г.** Дифференциальные уравнения. – М.: Наука, 1980.
31. **Понтрягин Л.С.** Обыкновенные дифференциальные уравнения. – М.: Наука, 1982.
32. **Петровский И.Г.** Лекции по теории обыкновенных дифференциальных уравнений. – М.: МГУ, 1984.
33. **Эльсгольц Л.Э.** Дифференциальные уравнения и вариационное исчисление. – М.: Наука, 1969.
34. **Карташёв А.П., Рождественский Б.Л.** Обыкновенные дифференциальные уравнения и основы вариационного исчисления. – М.: Наука, 1980.
35. **Коддингтон Э.А., Левинсон Н.** Теория обыкновенных дифференциальных уравнений. – М.: Изд-во иностранной литературы, 1958.
36. **Хартман Ф.** Обыкновенные дифференциальные уравнения. – М.: Мир, 1970.
37. **Левитан Б.М., Саргсян И.С.** Введение в спектральную теорию. Самосопряжённые обыкновенные дифференциальные операторы. – М.: Наука, 1970.
38. **Гельфанд И.М., Фомин С.В.** Вариационное исчисление. – М.: Физматлит, 1961.
39. **Буслаев В.С.** Вариационное исчисление. – Л.: Изд-во Ленинградского ун-та, 1980.
40. **Иоффе А.Д., Тихомиров В.М.** Теория экстремальных задач. – М.: Наука, 1974.

41. **Тихонов А.Н., Самарский А.А.** Уравнения математической физики. – М.: Наука, 1972.
42. **Владимиров В.С.** Уравнения математической физики. – М.: Наука, 1971.
43. **Соболев С.Л.** Уравнения математической физики. – М.: Наука, 1966.
44. **Бицадзе А.В.** Уравнения математической физики. – М.: Наука, 1976.
45. **Бахвалов Н.С.** Численные методы. – М.: Наука, 1973.
46. **Бахвалов Н.С., Жидков Н.П., Кобельков Г.М.** Численные методы. – М.: Наука, 1987.
47. **Самарский А.А.** Введение в численные методы. – М.: Наука, 1987.
48. **Самарский А.А., Гулин А.В.** Численные методы. – М.: Наука, 1989.
49. **Самарский А.А., Гулин А.В.** Численные методы математической физики. – М.: Научный мир, 2000.
50. **Самарский А.А.** Теория разностных схем. – М.: Наука, 1977.
51. **Самарский А.А., Николаев Е.С.** Методы решения сеточных уравнений. – М.: Наука, 1978.
52. **Костомаров Д.П., Фаворский А.П.** Вводные лекции по численным методам. – М.: Логос, 2004.
53. **Крылов В.И., Бобков В.В., Монастырный П.И.** Вычислительные методы, т. I, II. – М.: Наука, 1976, 1977.

Рекомендуем указанные книги для подготовки к государственному экзамену по прикладной математике и информатике и для последующей профессиональной деятельности.

Учебно-методическое издание

**И.В.Дмитриева, В.А. Морозова, С.И.Орлик.**

**Методические материалы для подготовки к государственному экзамену по прикладной математике и информатике по дополнительной части программы специалистов**

ООО "Планета Знаний"

Формат 60×84/16. Усл. печ. л. 18,37. Тираж 300. Зак. 2111.

Отпечатано с оригинал-макета
в Производственной фирме «Полиграф-Книга».
160001, г. Вологда, ул. Челюскинцев, 3.