

КАЗАНСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
Факультет вычислительной математики и кибернетики

Р.З. ДАУТОВ

**МЕТОД ГАЛЕРКИНА С ВОЗМУЩЕНИЯМИ
ДЛЯ ЗАДАЧ НА СОБСТВЕННЫЕ ЗНАЧЕНИЯ**

Учебное пособие

Казань 2010

УДК 517.5

Р.З. Даутов. Метод галеркина с возмущениями для задач на собственные значения. 94 с.

В пособии излагается метод Галеркина с возмущениями для самосопряженных задач на собственные значения в вещественном гильбертовом пространстве. Рассмотрено применение этого метода для решения задачи о поверхностных собственных волнах цилиндрических диэлектрических волноводов.

Пособие рассчитано на студентов старших курсов и аспирантов, специализирующихся в области уравнений математической физики и численных методов их решения. Оно может быть полезно также научным сотрудникам, чьи интересы лежат в указанных областях.

Научный редактор:

доктор физико-математических наук М.М. Карчевский

Рецензенты:

доктор физико-математических наук М.Р. Тимербаев

кандидат физико-математических наук Е.В. Стребков

Печатается по постановлению редакционно-издательского совета факультета ВМК Казанского университета

Оглавление

Предисловие	5
ГЛАВА 1. Метод Галеркина с возмущениями для задач на собственные значения	7
§ 1. Обобщенная задача на собственные значения	7
§ 2. Метод Галеркина с возмущениями	10
2.1. Определение метода.	10
2.2. Краткий обзор литературы.	13
2.3. Оценки точности.	16
§ 3. Параметрическая задача на собственные значения	19
3.1. Прямая задача.	21
3.2. Обратная задача.	25
§ 4. Метод Галеркина с возмущениями для параметрической задачи	26
4.1. Дискретная задача.	26
4.2. Оценки точности.	28
ГЛАВА 2. Скалярная задача о поверхностных собственных волнах	30
§ 1. Постановка задачи в ограниченной области	31
1.1. Первый метод.	31
1.2. Пространство $H^{1/2}(\Gamma)$	34
1.3. Свойства функций \mathbb{K}_n	40
1.4. Метагармоническое продолжение функции.	41
1.5. Второй метод.	44
1.6. Пространства функций.	46
§ 2. Существование и свойства решений	47
2.1. Свойства форм a и b	48
2.2. Существование и свойства решений задачи (\mathcal{P})	51
2.3. Аналитичность дисперсионных кривых.	54
2.4. Множество решений задачи (\mathcal{P}_∞)	57
§ 3. Дискретная задача	60
3.1. Пространство конечных элементов.	60
3.2. Формулы численного интегрирования.	63
3.3. Дискретная задача. Свойства решений.	66
§ 4. Оценки точности	69
4.1. Оценки погрешности численного интегрирования.	70
4.2. Вспомогательные утверждения.	73
4.3. Оценки возмущений.	74
4.4. Оценки точности приближенных решений.	76

ГЛАВА 3. Результаты численных экспериментов	78
§ 1. Некоторые аспекты программной реализации	78
§ 2. Волновод кругового поперечного сечения	81
§ 3. Волновод квадратного поперечного сечения	83
§ 4. Волновод прямоугольного поперечного сечения	85
§ 5. Волновод с поперечным сечением из трех кругов	88
Литература	91

ПРЕДИСЛОВИЕ

В пособии излагается метод Галеркина с возмущениями для самосопряженных задач на собственные значения в вещественном гильбертовом пространстве. Подобные методы широко и успешно используются в практике вычислений. Мы рассмотрим применение этого метода для решения скалярной задачи о поверхностных собственных волнах цилиндрических диэлектрических волноводов.

Предполагается, что читатель знаком с основами функционального анализа, в частности, со спектральной теорией компактных (вполне непрерывных) операторов в гильбертовых пространствах. Все необходимые сведения из этих разделов можно найти, например, в книгах [1], [2].

Вкратце содержание пособия таково. В *первой главе* дается постановка задачи на собственные значения, формулируются условия на операторы задачи и определяются методы Галеркина и методы Галеркина с возмущениями. Здесь же приводится краткий обзор литературы, касающейся исследования сходимости и точности этих методов, а также излагается одна достаточно простая методика получения оценок точности. Рассматриваются также задачи на собственные значения, нелинейно зависящие от параметра.

Во *второй главе*, имеющей самостоятельный интерес, показывается, как эти абстрактные результаты могут быть применены при решении достаточно сложной конкретной задачи, а именно, скалярной задачи о поверхностных собственных волнах цилиндрических диэлектрических волноводов; она представляет собой параметрическую спектральную задачу для оператора Гельмгольца на всей плоскости. Получается эквивалентная формулировка этой задачи в ограниченной области, на основе метода конечных элементов определяется приближенный метод его решения. Результаты первой главы используются как для исследования существования решений исходной и дискретной задачи, так и для получения оценок точности предложенной численной схемы.

Третья глава посвящена описанию результатов различных вычислительных экспериментов, целью которых является демонстрация эффективности и точности рассмотренного во второй главе численного метода.

В тексте пособия принята двойная нумерация, сквозная в пределах главы. Символ $:=$ означает, что правая часть выражения определяет левую. Аналогично используется символ $=:$.

Рукопись была внимательно прочитана М.М. Карчевским. Автор с благодарностью учел его замечания.

Замеченные ошибки, опечатки, а также комментарии и пожелания, просьба направлять по адресу *rdautov@ksu.ru*.

ГЛАВА 1

МЕТОД ГАЛЕРКИНА С ВОЗМУЩЕНИЯМИ ДЛЯ ЗАДАЧ НА СОБСТВЕННЫЕ ЗНАЧЕНИЯ

В данной главе рассматриваются обобщенные задачи на собственные значения в гильбертовом пространстве и изучается метод Галеркина с возмущениями для их приближенного решения.

§ 1. Обобщенная задача на собственные значения

Пусть V — бесконечномерное вещественное гильбертово пространство с нормой $\|\cdot\|$ и скалярным произведением (\cdot, \cdot) ; A, B — самосопряженные операторы в V . Рассмотрим задачу на собственные значения: *найти такие $\lambda \in \mathbb{R}$ и ненулевые $u \in V$, что*

$$Au = \lambda Bu, \quad (1.1)$$

или, на языке симметричных билинейных форм этих операторов,

$$(P_\lambda) \quad a(u, v) = \lambda b(u, v) \quad \forall v \in V.$$

Дополнительно будем предполагать, что оператор B является компактным, размерность образа B равна бесконечности¹⁾,

$$m_A \|u\|^2 \leq a(u, u) \leq M_A \|u\|^2, \quad 0 \leq b(u, u) \leq M_B \|u\|^2, \quad u \in V. \quad (1.2)$$

Кратко условия (1.2) будем записывать в виде²⁾:

$$m_A I \leq A \leq M_A I, \quad 0 \leq B \leq M_B I.$$

Отметим, что из условий на формы a и b следует, что

$$|a(u, v)| \leq M_A \|u\| \|v\|, \quad |b(u, v)| \leq M_B \|u\| \|v\|, \quad u, v \in V.$$

¹⁾образ $B := \text{Im} B := \{f \in V : f = Bu, u \in V\}$.

²⁾ I обозначает тождественный оператор в V . Оценки снизу в (1.2) означают, что A — положительно определенный оператор, а B — неотрицательный.

Пара (λ, u) , $u \neq 0$, удовлетворяющая (P_λ) , называется собственной парой; λ — собственным числом; u — собственным элементом, соответствующим λ . Собственные элементы определяются неоднозначно: легко видеть, что, если u — собственный элемент, то cu , $c \in \mathbb{R}$, также есть собственный элемент; если u_1, u_2, \dots — собственные элементы, соответствующие λ , то их произвольная линейная комбинация $c_1u_1 + c_2u_2 + \dots$ также есть собственный элемент, соответствующий λ . В связи с этим определяется пространство $U(\lambda) := \ker(A - \lambda B)$ как множество всех собственных элементов, соответствующих собственному числу λ . Оно называется собственным подпространством, соответствующим λ , а его размерность $(\dim U(\lambda))$ — кратностью числа λ ³⁾. Учитывая сказанное, в дальнейшем задачи типа (1.1), как это принято, будем понимать как задачу нахождения пар $(\lambda, U(\lambda))$.

Нетрудно видеть, что элементы из $\ker B$ не могут быть собственными элементами, а собственные числа λ являются положительными. Ясно также, что задача на собственные значения $Bu = \mu Au$ или

$$(P_\mu) \quad b(u, v) = \mu a(u, v) \quad \forall v \in V, \quad u \neq 0,$$

имеет собственное число $\mu = 0$, которому соответствует собственное подпространство $\ker B$, причем, пара (μ, u) , $\mu > 0$, является собственной парой задачи (P_μ) тогда и только тогда, когда (λ, u) — собственная пара (P_λ) и $\lambda = 1/\mu$.

Задача (P_μ) равносильна задаче на собственные значения

$$Tu = \mu u, \quad T := A^{-1}B. \quad (1.3)$$

Оператор T , как произведение ограниченного и компактного операторов, является компактным. Кроме того, он является самосопряженным в гильбертовом пространстве V_A .¹⁾ Из теории компактных операторов непосредственно следует (см., напр., [3, с. 245])

Теорема 1.1. *Существует счетное множество положительных чисел λ^K , $K = 1, 2, \dots$, с единственной точкой накопления $+\infty$, образующих полный набор собственных чисел задачи (P_λ) . Соответствующие им собственные подпространства U^K , $K \geq 1$, — конечномерны, $V_A = \bigoplus_{K=1}^{\infty} U^K \oplus \ker B$.²⁾*

³⁾ $\ker A := \ker a := \{u \in V : a(u, u) = 0\}$.

¹⁾ Элементами V_A являются элементы V , форма a определяет скалярное произведение в нем. В силу (1.2) норма $\|\cdot\|_A := a^{1/2}(\cdot, \cdot)$ в V_A эквивалентна норме V .

Пусть $\mu^K := 1/\lambda^K$. Тогда (μ^K, U^K) , $K \geq 1$, — решения задачи (P_μ) .

Введем другую нумерацию собственных чисел λ^K , $K = 1, 2, \dots$ (нумерацию по возрастанию с учетом кратности). А именно, пронумеруем их по возрастанию,

$$0 < \lambda_1 \leq \lambda_2 \leq \lambda_3 \leq \dots,$$

указывая каждое число λ^K в этом ряду столько раз, какова его кратность. Известно следующее вариационное описание этих чисел (минимаксный принцип Куранта — Фишера):

$$\frac{1}{\lambda_i} = \mu_i = \max_{V_i \subset V} \min_{v \in V_i \setminus \{0\}} R(v), \quad i = 1, 2, \dots, \quad R(v) := \frac{b(v, v)}{a(v, v)}.$$

Здесь максимум берется по всем подпространствам V размерности i ; функционал R называется отношением Рэля.

Далее будем считать, что r_K есть кратность λ^K , а k означает такой номер, соответствующий K , что

$$\lambda^K = \lambda_k, \quad \lambda_{k-1} < \lambda_k = \lambda_{k+1} = \dots = \lambda_{k+r_K-1} < \lambda_{k+r_K}. \quad (1.4)$$

Соответственно, полагая $\mu_i = 1/\lambda_i$, $i = 1, 2, \dots$, имеем¹⁾:

$$0 < \dots \leq \mu_3 \leq \mu_2 \leq \mu_1, \\ \mu^K = \mu_k, \quad \mu_{k+r_K} < \mu_{k+r_K-1} = \dots = \mu_{k+1} = \mu_k < \mu_{k-1}. \quad (1.5)$$

В каждом пространстве U^K фиксируем некоторый ортонормированный в V_A базис так, что (span означает линейную оболочку)

$$U^K = \text{span}\{u_k, u_{k+1}, \dots, u_{k+r_K-1}\}. \quad (1.6)$$

Тогда последовательности λ_i (или μ_i), $i = 1, 2, \dots$, будет поставлена в соответствие ортонормированная в V_A последовательность собственных элементов, образующих базис в $V_A \ominus \ker B$:

$$u_1, u_2, u_3, \dots, \quad a(u_i, u_j) = \delta_{ij}, \quad b(u_i, u_j) = \mu_i \delta_{ij}, \quad i, j = 1, 2, \dots$$

Произвольный элемент $v \in V$ представим в виде

$$v = v_0 + \sum_{i=1}^{\infty} c_i u_i, \quad v_0 \in \ker B, \quad c_i = a(v, u_i).$$

²⁾ \oplus означает прямую сумму пространств. Если $V = U \oplus W$, то $U = V \ominus W = W^\perp$ есть ортогональное дополнение W .

¹⁾В (1.5) нужно опустить правое неравенство при $k = 1$, как и левое — в (1.4).

Отметим, что $a(v_0, u_i) = \lambda_i b(v_0, u_i) = 0$ для всех $i \geq 1$. Кроме того,

$$a(v, v) = c_0^2 + \sum_{i=1}^{\infty} c_i^2, \quad b(v, v) = \sum_{i=1}^{\infty} \mu_i c_i^2, \quad c_0^2 = a(v_0, v_0).$$

§ 2. Метод Галеркина с возмущениями

2.1. Определение метода.

Пусть $h > 0$ — малый параметр, $\{V_h\}_h$ — предельно плотная последовательность конечномерных подпространств V , т. е.

$$(H_1) \quad V_h \subseteq V; \quad \forall u \in V: \quad \inf_{v_h \in V_h} \|u - v_h\| \rightarrow 0 \quad \text{при } h \rightarrow 0.^1)$$

Введем симметричные билинейные формы a_h и b_h , являющиеся возмущениями (аппроксимациями) форм a и b на $V_h \times V_h$. Будем считать, что для любого $v \in V_h$ справедливы оценки

$$(H_2) \quad t_a \|v\|^2 \leq a_h(v, v) \leq M_a \|v\|^2, \quad 0 \leq b_h(v, v) \leq M_b \|v\|^2,$$

где постоянные t_a , M_a и M_b не зависят от h , $t_a > 0$.

Рассмотрим аппроксимацию задачи (P_μ) методом Галеркина с возмущениями: *найти $(\mu^h, y) \in \mathbb{R} \times V_h \setminus \{0\}$ такие, что*

$$(P_\mu^h) \quad b_h(y, v) = \mu^h a_h(y, v) \quad \forall v \in V_h.$$

Пусть h фиксировано, N_h есть размерность пространства V_h , а элементы $\varphi_1, \varphi_2, \dots, \varphi_{N_h}$ определяют в нем базис. Тогда задача (P_μ^h) сводится к алгебраической задаче на собственные значения

$$B_h x = \mu^h A_h x, \quad 0 \neq x \in \mathbb{R}^{N_h}. \quad (1.7)$$

Здесь x есть вектор коэффициентов разложения y по указанному базису, а симметричные матрицы A_h и B_h имеют следующие элементы:

$$a_{ij} := a_h(\varphi_i, \varphi_j), \quad b_{ij} := b_h(\varphi_i, \varphi_j), \quad i, j = 1, 2, \dots, N_h.$$

¹⁾в теории аппроксимации, как правило, вместо малого параметра h используется “большой” параметр n , указывающий на размерность подпространства. Параметр h общепринят в теории МКЭ и обозначает максимальный размер конечных элементов. Использование h вместо n существенно облегчит нам дальнейшие ссылки.

Из условий (H_2) следует, что A_h является положительно определенной матрицей, а B_h — неотрицательно определенной. Ясно, что задача (1.7) имеет нулевое собственное число кратности $\dim(\ker B_h)$ и, скажем, n_h положительных собственных чисел μ^{hK} , суммарной кратности равной $N_h^b := N_h - \dim(\ker B_h)$, $n_h \leq N_h^b$.²⁾ Как и в исходной задаче (P_μ) , занумеруем положительные собственные числа в убывающем порядке с учетом кратности:

$$0 < \mu_{N_h^b}^h \leq \dots \leq \mu_2^h \leq \mu_1^h.$$

Поставим им в соответствие собственные элементы

$$y_{N_h^b}, \dots, y_2, y_1, \quad a_h(y_i, y_j) = \delta_{ij}, \quad i, j = 1, 2, \dots, N_h^b.$$

Тогда любой элемент $v_h \in V_h$ представим в виде

$$v_h = y_0 + \sum_{i=1}^{N_h^b} c_i y_i, \quad y_0 \in \ker B_h, \quad c_i = a_h(v_h, y_i). \quad (1.8)$$

Отметим, что $a_h(y_0, y_i) = 0$, $1 \leq i \leq N_h^b$. Кроме того,

$$a_h(v_h, v_h) = c_0^2 + \sum_{i=1}^{N_h^b} c_i^2, \quad b_h(v_h, v_h) = \sum_{i=1}^{N_h^b} \mu_i^h c_i^2, \quad c_0^2 = a_h(y_0, y_0). \quad (1.9)$$

Аппроксимацию собственного подпространства U^K , $K \geq 1$, задачи (P_μ) , соответствующего собственному числу μ^K кратности r_K , определим следующим образом (см. (1.6)):

$$U_h^K := \text{span}(y_k, y_{k+1}, \dots, y_{k+r_K-1}),$$

а числа $\mu_{k+r_K-1}^h, \dots, \mu_{k+1}^h, \mu_k^h$ будем рассматривать как приближения собственного числа μ^K . Здесь k связано с K согласно (1.5).

Прежде чем перейти к рассмотрению вопросов, касающихся сходимости и точности описанных приближений, определим способ измерения близости собственных подпространств U^K и U_h^K посредством величины, называемой раствором. Дадим его определение.

Пусть H — гильбертово пространство с нормой $\|\cdot\|_H$ и скалярным произведением $(\cdot, \cdot)_H$, U — его подпространство. Оператор P

²⁾Далее мы наложим дополнительные ограничения на формы a_h и b_h (см. условие (H_3)). В этом случае $N_h^b \rightarrow \infty$ при $h \rightarrow 0$.

называется ортопроектором в H на U , если Pu есть такой элемент U , что

$$\|u - Pu\|_H := \inf_{v \in U} \|u - v\|_H, \quad u \in H,$$

или, эквивалентно,

$$(u - Pu, v)_H = 0 \quad \forall v \in U, \quad u \in H.$$

Ясно, что P является линейным оператором и $\|Pu\|_H \leq \|u\|_H$.

Пусть теперь U и U_h — конечномерные подпространства H одинаковой размерности. Раствор между ними определяется следующим образом (см., напр., [4, с. 195]):

$$\Theta_H(U, U_h) := \max_{u \in U, \|u\|_H=1} \|u - P_h u\|_H := \max_{y \in U_h, \|y\|_H=1} \|y - Py\|_H,$$

где P и P_h — ортопроекторы в H на U и U_h соответственно.

Имеют место равенства

$$\min_{u \in U, \|u\|_H=1} \|P_h u\|_H^2 = \min_{y \in U_h, \|y\|_H=1} \|Py\|_H^2 = 1 - \Theta_H^2(U, U_h). \quad (1.10)$$

Отметим, что из оценки $\Theta_V(U^K, U_h^K) \leq \varepsilon$ следует, что для любого $y \in U_h^K$, $\|y\| = 1$, найдется такой элемент $u \in U^K$, что $\|y - u\| \leq \varepsilon$. Ясно также, что

$$c^{-1} \Theta_V(U^K, U_h^K) \leq \Theta_{V_A}(U^K, U_h^K) \leq c \Theta_V(U^K, U_h^K)$$

при $c = (M_A/m_A)^{1/2}$. Для заданных форм d и d_h положим

$$E_d(\varphi_h) := \sup_{v_h \in V_h, \|v_h\|=1} |d(\varphi_h, v_h) - d_h(\varphi_h, v_h)|, \quad \varphi_h \in V_h.$$

Следующие величины характеризуют максимально возможные возмущения форм a и b соответственно:

$$\sup_{v \in V_h, \|v\|=1} E_a(v), \quad \sup_{v \in V_h, \|v\|=1} E_b(v).$$

Далее нам понадобится также функционал

$$E_{ab}(v) := E_a(v) + E_b(v), \quad v \in V_h.$$

2.2. Краткий обзор литературы.

В контексте предыдущих определений введем дополнительные обозначения. Во-первых отметим, что оператор $T : V \rightarrow V$ в (1.3) на языке билинейных форм определяется так, что Tf есть решение уравнения

$$a(Tf, v) = b(f, v) \quad \forall v \in V$$

при произвольно заданном $f \in V$. Аналогично введем оператор T_h :

$$T_h : V_h \rightarrow V_h, \quad a_h(T_h f_h, v) = b_h(f_h, v) \quad \forall v \in V_h, \quad \forall f_h \in V_h.$$

Тогда приближенная задача (P_μ^h) примет вид:

$$T_h y = \mu^h y. \quad (1.11)$$

Введем также оператор $T^h : V_h \rightarrow V_h$ так, что

$$a(T^h f_h, v) = b(f_h, v) \quad \forall v \in V_h$$

при любом $f_h \in V_h$. Легко видеть, что $T^h = P_h T$, где P_h — ортопроектор в V_A на V_h .

Дискретная задача: *найти* $(\mu^h, y) \in \mathbb{R} \times V_h \setminus \{0\}$ *такие, что*

$$T^h y = \mu^h y, \quad (1.12)$$

определяет метод Галеркина для задачи (1.3).

Изучению сходимости и оценкам скорости сходимости метода Галеркина и метода Галеркина с возмущениями посвящено множество работ. В первую очередь, отметим здесь результаты, представленные в монографиях [4], [1], относящиеся к 60-м годам прошлого века (см. также цитированную там литературу). В них изучается близость решений задач (1.11), (1.12) к решениям (1.3) в общей ситуации, когда оператор T — произвольный компактный оператор, действующий в банаховом пространстве.

Например, в [4, с. 257-261] доказано, что условий

$$\|T - P_h T\| \rightarrow 0, \quad \|T_h - T^h\| \rightarrow 0 \quad \text{при } h \rightarrow 0, \quad (1.13)$$

достаточно для сходимости $\mu_k^h \rightarrow \mu_k$ при $h \rightarrow 0$, $k = 1, 2, \dots$ ¹⁾ Кроме того, при достаточно малых h получены оценки, которые в наших

¹⁾Из условий (1.13) следует, что $\|T - T_h\| \rightarrow 0$ при $h \rightarrow 0$. Поэтому $N_h^b := \dim V_h - \dim(\ker B_h) = \dim(\text{Im } T_h) \rightarrow \infty$ при $h \rightarrow 0$, так как $\dim(\text{Im } T) = \infty$.

обозначениях имеют вид²⁾:

$$|\mu_k - \mu_k^h| \leq c_k (\epsilon_h^2 + \|T_h - T^h\|), \quad k = 1, 2, \dots, \quad (1.14)$$

$$\Theta_V(U^K, U_h^K) \leq c_K (\epsilon_h + \|T_h - T^h\|), \quad K = 1, 2, \dots \quad (1.15)$$

Здесь величина

$$\epsilon_h := \max_{u \in U^K, \|u\|=1} \|u - P_h u\|$$

характеризует погрешность приближения собственных элементов элементами из V_h . Оценка величины ϵ_h является стандартной задачей в теории проекционных методов (в частности, в теории метода конечных элементов). Отметим, что оценка (1.14) используется при доказательстве (1.15).

Обсудим условия (1.13). Из эквивалентности норм в V_A и в V следует, что проектор P_h ограничен в V и для любого $u \in V$

$$\|(I - P_h)u\| = \|u - P_h u\| \leq c \inf_{v_h \in V_h} \|u - v_h\| \rightarrow 0 \quad \text{при } h \rightarrow 0$$

в силу условия (H_1) . Поэтому $\|T - P_h T\| = \|(I - P_h)T\| \rightarrow 0$ при $h \rightarrow 0$ (см., напр., [4, Лемма 15.4, с. 202]). Таким образом, первое условие в (1.13) выполняется, если справедливо предположение (H_1) .

Малость величины $\|T_h - T^h\|$ зависит от малости возмущения форм a и b . Действительно, поскольку для любых $f, v \in V_h$ имеет место представление

$$\begin{aligned} a_h(T^h f - T_h f, v) &= [a(T^h f, v) - a_h(T_h f, v)] + [a_h(T^h f, v) - a(T^h f, v)] = \\ &= [b(f, v) - b_h(f, v)] + [a_h(T^h f, v) - a(T^h f, v)], \end{aligned}$$

а $T^h = P_h T$, то $|a_h(T^h f - T_h f, v)| \leq (E_b(f) + E_a(P_h T f)) \|v\|$. Полагая здесь $v = T^h f - T_h f$ и учитывая условия (H_2) , получим

$$\|T^h - T_h\| \leq c \sup_{f \in V_h, \|f\|=1} (E_b(f) + E_a(P_h T f)).$$

Таким образом, условий (H_1) , (H_2) и условия

$$(H_3) \quad \sup_{f \in V_h, \|f\|=1} (E_b(f) + E_a(P_h T f)) \rightarrow 0 \quad \text{при } h \rightarrow 0$$

²⁾здесь и далее буквой c , возможно с индексом, обозначаются различные положительные постоянные, не зависящие от h .

достаточно для сходимости $\mu_k^h \rightarrow \mu_k$ при $h \rightarrow 0$, $k = 1, 2, \dots$

Оценки (1.14), (1.15) являются неулучшаемыми для метода Галеркина (в этом случае $T_h = T^h$), но не всегда являются таковыми для метода Галеркина с возмущениями. Дело в том, что погрешность возмущений в этих оценках оценивается на самом “плохом” элементе V_h .

Уточнениям и упрощениям доказательств отмеченных выше оценок было посвящено множество работ [5], [6], [7], [8], [9], [10], [11], [12], [13], [14], [15]. Особо выделим среди них работы Дж. Осборна (J. Osborn) и его коллег. В разработанной ими методике сначала оценивается точность собственных элементов, а лишь затем — собственных чисел. Например, в [6] были получены абстрактные оценки точности, приводящие к оптимальным оценкам спектральных аппроксимаций задач на собственные значения для дифференциальных или интегральных операторов (не обязательно самосопряженных). В [12] получены аналогичные оценки, приводящие к оптимальным оценкам, в частности, в методе конечных элементов с численным интегрированием. Например, там получена оценка (ср. с (1.15))

$$\Theta_V(U^K, U_h^K) \leq c_K (\epsilon_h + \|(T_h - T^h)|_{U_{Kh}}\|), \quad (1.16)$$

где $(T_h - T^h)|_{U_{Kh}}$ представляет собой сужение оператора $T_h - T^h$ на U_{Kh} — собственное подпространство в методе Галеркина (без возмущений), определяемое аналогично U_h^K . Из (1.15) следует, что $\Theta_V(U^K, U_{Kh}) \leq c_K \epsilon_h$. Этот факт может быть использован в конкретной задаче для оптимальной оценки правой части в (1.16).

Удобные для применения в разнообразных приложениях оценки точности метода Галеркина с возмущениями для самосопряженных положительно определенных операторов A и B в гильбертовом пространстве получены в [15], [16], [17]. В этих работах доказан следующий аналог оценки (1.16):

$$\Theta_V(U^K, U_h^K) \leq c_K (\epsilon_h + \max_{u \in U^K, \|u\|=1} E_{ab}(P_h u)).$$

Ниже мы получим обобщение этих результатов на случай неотрицательно определенного оператора B .

2.3. Оценки точности.

Исследуем близость решений задач (P_μ) и (P_μ^h) . Пусть (μ^K, U^K) — решение задачи (P_μ) , r_K — кратность μ^K , $K \geq 1$. Пусть далее $\mu_{k+r_K-1}^h, \dots, \mu_{k+1}^h, \mu_k^h$ и U_h^K — их аппроксимации, определенные выше в п. 2.1. Положим

$$\epsilon_h(u) := \inf_{v_h \in V_h} \|u - v_h\| + E_{ab}(P_h u), \quad u \in V,$$

$$\Sigma_h(y) := |a(y, y) - a_h(y, y)| + |b(y, y) - b_h(y, y)|, \quad y \in V_h,$$

через P_h обозначим ортопроектор в V_A на V_h . По определению

$$a(u - P_h u, v) = 0 \quad \forall v \in V_h, \quad u \in V.$$

Из условия (1.2) следует, что

$$\|u - P_h u\| \leq c \inf_{v \in V_h} \|u - v\| \leq c \epsilon_h(u), \quad c = (M_A/m_A)^{1/2}.$$

Будем предполагать, что выполнены условия $(H_1) - (H_3)$. Как было показано выше, эти условия обеспечивают сходимость приближенных собственных чисел: $\mu_k^h \rightarrow \mu_k$ при $h \rightarrow 0$, $k = 1, 2, \dots$

Теорема 1.2. Пусть выполнены условия $(H_1) - (H_3)$. Тогда при достаточно малом h справедлива оценка

$$\Theta_V(U^K, U_h^K) \leq c_K \max_{u \in U^K, \|u\|=1} \epsilon_h(u).$$

Здесь постоянная c_K зависит от μ^K , но не зависит от h , $K \geq 1$.

ДОКАЗАТЕЛЬСТВО. Пусть $u \in U^K$ — собственный элемент. Положим $y := P_h u$ и разложим y по системе $\{y_i\}$ (см. (1.8)). Имеем

$$y = y_0 + \sum_{i=1}^{N_h^b} c_i y_i =: v_h + Q_h u + w_h,$$

$$v_h := y_0 + \sum_{i=k+r_K}^{N_h^b} c_i y_i, \quad Q_h u := \sum_{i=k}^{k+r_K-1} c_i y_i \in U_h^K, \quad w_h := \sum_{i=1}^{k-1} c_i y_i.$$

Рассмотрим величину

$$\sigma_h(y) := \sup_{\eta_h \in V_h, \|\eta_h\|=1} |b_h(y, \eta_h) - \mu_k a_h(y, \eta_h)|.$$

Так как $a(y, \eta_h) = a(P_h u, \eta_h) = a(u, \eta_h) = 1/\mu^K b(u, \eta_h)$, то

$$\begin{aligned} b_h(y, \eta_h) - \mu^K a_h(y, \eta_h) &= \\ &= [b_h(y, \eta_h) - b(y, \eta_h)] + \mu^K [a(y, \eta_h) - a_h(y, \eta_h)] + b(y - u, \eta_h). \end{aligned}$$

Отсюда при $c := \max\{1, \mu^K, M_B\}$ имеем:

$$\sigma_h(y) \leq c (E_{ab}(y) + \|u - P_h u\|) = c \epsilon_h(u).$$

Докажем, что справедливы оценки

$$\|v_h\| \leq c_K \sigma_h(y), \quad \|w_h\| \leq c_K \sigma_h(y). \quad (1.17)$$

Тогда $\|P_h u - Q_h u\| = \|v_h + w_h\| \leq c_K \sigma_h(y) \leq c_K \epsilon_h(u)$, а также

$$\begin{aligned} \Theta_V(U^K, U_h^K) &\leq \max_{u \in U^K, \|u\|=1} \|u - Q_h u\| \leq \max_{u \in U^K, \|u\|=1} (\|u - P_h u\| + \\ &\quad + \|P_h u - Q_h u\|) \leq c_K \max_{u \in U^K, \|u\|=1} \epsilon_h(u), \end{aligned}$$

то есть утверждение теоремы будет доказано.

Получим сначала первую оценку в (1.17). Пусть ρ_K характеризует отдаленность собственного числа μ^K , т. е.

$$\rho_K := \max_{\mu_i: \mu_i \neq \mu_k} |\mu_k - \mu_i|^{-1}.$$

Поскольку $\mu_k^h \rightarrow \mu_k$ при $h \rightarrow 0$, то для достаточно малых h имеем

$$\mu_k - \mu_{k+r_K}^h > c(\mu_k - \mu_{k+r_K}), \quad \mu_{k-1}^h - \mu_k > c(\mu_{k-1} - \mu_k) \quad (1.18)$$

с некоторой постоянной $c \in (0, 1)$. Так как система $\{y_i\}$ является a_h -ортогональной, то из определения v_h и (1.9) вытекает, что

$$\begin{aligned} a_h(y, v_h) &= a_h(v_h, v_h), \quad b_h(y, v_h) = b_h(v_h, v_h), \\ b_h(v_h, v_h) &\leq \mu_{k+r_K}^h a_h(v_h, v_h). \end{aligned} \quad (1.19)$$

Поэтому из (1.19), (1.18) и (H_2) имеем:

$$\begin{aligned} -b_h(y, v_h) + \mu^K a_h(y, v_h) &= -b_h(v_h, v_h) + \mu_k a_h(v_h, v_h) \geq \\ &\geq (\mu_k - \mu_{k+r_K}^h) a_h(v_h, v_h) \geq c m_a (\mu_k - \mu_{k+r_K}) \|v_h\|^2 \geq c m_a \rho_K^{-1} \|v_h\|^2. \end{aligned}$$

Отсюда следует первая оценка в (1.17).

Докажем вторую оценку в (1.17). Аналогично (1.19) имеем:

$$\begin{aligned} a_h(y, w_h) &= a_h(w_h, w_h), & b_h(y, w_h) &= b_h(w_h, w_h), \\ b_h(w_h, w_h) &\geq \mu_{k-1}^h a_h(w_h, w_h). \end{aligned}$$

Таким образом,

$$\begin{aligned} b_h(y, w_h) - \mu_k a_h(y, w_h) &= b_h(w_h, w_h) - \mu_k a_h(w_h, w_h) \geq \\ &\geq (\mu_{k-1}^h - \mu_k) a_h(w_h, w_h) \geq c m_a \rho_K^{-1} \|w_h\|^2. \quad \square \end{aligned}$$

Теорема 1.3. Пусть выполнены условия (H_1) – (H_3) и

$$\max_{u \in U^K, \|u\|=1} \epsilon_h(u) \rightarrow 0, \quad h \rightarrow 0.$$

Тогда при достаточно малом h справедливы оценки

$$|\mu^K - \mu_i^h| \leq c_K \max_{u \in U^K, \|u\|=1} (\epsilon_h^2(u) + \Sigma_h(P_h u)), \quad i = k, \dots, k + r_K - 1.$$

Здесь постоянная c_K зависит от μ^K , но не зависит от h , $K \geq 1$.

ДОКАЗАТЕЛЬСТВО. Непосредственно проверяется, что для любого $u \in U^K$ и $y \in V_h$ имеет место следующее равенство ($\mu_i = \mu^K$ при $i = k, \dots, k + r_K - 1$):

$$\begin{aligned} (\mu_i^h - \mu^K) a_h(y, y_i) &= b(y - u, y_i - u) - \mu_i a(y - u, y_i - u) + \\ &+ [b_h(y, y) - b(y, y)] - \mu_i [a_h(y, y) - a(y, y)] + \\ &+ [b_h(y, y_i - y) - b(y, y_i - y)] - \mu_i [a_h(y, y_i - y) - a(y, y_i - y)], \end{aligned}$$

где y_i — собственный элемент, соответствующий μ_i^h , $\|y_i\|_A = 1$. Тогда

$$\begin{aligned} |\mu^K - \mu_i^h| |a_h(y, y_i)| &\leq c_K (\|y - u\| \|y_i - u\| + \Sigma_h(y) + E_{ab}(y) \|y_i - y\|) \leq \\ &\leq c_K (\|y - u\| \|y_i - u\| + \Sigma_h(y) + E_{ab}(y) (\|y_i - u\| + \|y - u\|)). \end{aligned}$$

Выберем здесь $u = P^K y_i$ (P^K — ортопроектор в V_A на U^K), положим $y = P_h u$ и учтем, что

$$\|y - u\| := \|P_h u - u\| \leq c \epsilon_h(u), \quad E_{ab}(y) \leq \epsilon_h(u),$$

$$\begin{aligned} \|y_i - u\| &\leq c \|y_i - u\|_A \leq c \sup_{y \in U_h^K, \|y\|_A=1} \|y - P^K y\|_A =: \\ &=: c \Theta_{V_A}(U^K, U_h^K) \leq c \Theta_V(U^K, U_h^K) \leq c_K \max_{u \in U^K, \|u\|=1} \epsilon_h(u). \end{aligned}$$

Тогда будем иметь:

$$|\mu^K - \mu_i^h| |a_h(P_h u, y_i)| \leq c_K \max_{u \in U^K, \|u\|=1} (\epsilon_h^2(u) + \Sigma_h(P_h u)).$$

Из этой оценки следует утверждение теоремы, если

$$a_h(P_h u, y_i) \geq c > 0. \quad (1.20)$$

Докажем (1.20). Пусть $\psi_h := a_h(P_h u, y_i) - a(P_h u, y_i)$. Учтем, что

$$\begin{aligned} \|P^K y_i\|_A &\leq \|y_i\|_A := 1, \\ |\psi_h| &\leq c E_a(P_h u) \|y_i\|_A \leq c E_{ab}(P_h u) \leq c \epsilon_h(u). \end{aligned}$$

Тогда

$$\begin{aligned} a_h(P_h u, y_i) &= a(P_h u, y_i) + \psi_h = a(u, y_i) + \psi_h = a(u, P^K y_i) + \psi_h = \\ &= \|P^K y_i\|_A^2 + \psi_h \geq 1 - \Theta_{V_A}^2(U^K, U_h^K) - c \epsilon_h(u) \geq 1 - c_K \epsilon_h(u). \end{aligned}$$

Здесь мы воспользовались свойством ортопроектора (1.10). Отсюда следует (1.20) при достаточно малом h . \square

§3. Параметрическая задача на собственные значения

Пусть $\mathbb{R}_+ := \{x \in \mathbb{R} : x \geq 0\}$ и пусть задано семейство самосопряженных операторов $A(p)$ и $B(p)$, $p \in \mathbb{R}_+^1$, действующих в вещественном бесконечномерном гильбертовом пространстве V с нормой $\|\cdot\|$ и скалярным произведением (\cdot, \cdot) . При каждом заданном $p \in \mathbb{R}_+$ требуется найти такие $\lambda \in \mathbb{R}$ и ненулевые $u \in V$, что

$$(\mathcal{P}) \quad A(p)u = \lambda B(p)u,$$

или, на языке симметричных билинейных форм этих операторов,

$$a(p, u, v) = \lambda b(p, u, v) \quad \forall v \in V.$$

В данном параграфе нас будут интересовать ответы на следующие вопросы:

- (i) сколько решений (λ, u) имеет задача при заданном p и как собственные числа λ зависят от p (прямая задача)?

¹⁾или, равносильно, самосопряженные оператор-функции $p \rightarrow A(p)$ и $p \rightarrow B(p)$ на \mathbb{R}_+

(ii) сколько имеется решений (λ, u) с заданным λ при различных значениях параметра p (обратная задача)?

Ответы на эти вопросы зависят, естественно, от свойств операторов задачи. В связи с этим напомним некоторые определения. Говорят, что самосопряженная оператор-функция $p \rightarrow C(p)$, $p \in [0, \infty)$, является

(a) непрерывной на $[0, \infty)$, если $\|C(p) - C(\bar{p})\| \rightarrow 0$ при $\bar{p} \rightarrow p \in [0, \infty)$;

(b) локально липшиц-непрерывной на $(0, \infty)$, если для любого замкнутого отрезка $\omega \subset (0, \infty)$ найдется такая постоянная $L_C(\omega)$, что ¹⁾

$$\|C(p) - C(\bar{p})\| \leq L_C(\omega)|p - \bar{p}|, \quad p, \bar{p} \in \omega;$$

(c) дифференцируемой на $(0, \infty)$, если существует такая самосопряженная оператор-функция $p \rightarrow C'(p)$, определенная на $(0, \infty)$, что

$$\left\| \frac{C(p) - C(\bar{p})}{p - \bar{p}} - C'(p) \right\| \rightarrow 0 \text{ при } \bar{p} \rightarrow p \in (0, \infty);$$

(d) аналитической на $(0, \infty)$, если в некоторой окрестности произвольной точки $p_0 \in (0, \infty)$ она представляется сходящимся рядом

$$C(p) = \sum_{n=0}^{\infty} C_n(p_0) (p - p_0)^n,$$

имеющим ненулевой радиус сходимости. Здесь C_n — самосопряженные операторы;

(e) строго монотонно возрастающей (неубывающей, невозрастающей, ...), если функция $p \rightarrow c(p, u, u)$ является строго монотонно возрастающей (неубывающей, невозрастающей, ...) при каждом фиксированном $u \in V$. Здесь $c(p, u, u) := (C(p)u, u)$ — квадратичная форма оператора $C(p)$.

Напомним также, что

$$\|C(p)\| = \sup_{u \in V \setminus \{0\}} \frac{|c(p, u, u)|}{\|u\|^2} = \sup_{u \in V, \|u\|=1} |c(p, u, u)|.$$

Ясно, что дифференцируемая оператор-функция является локально липшиц-непрерывной и, тем более, — непрерывной. Далее, если

¹⁾если $L_C(\omega)$ не зависит от ω , то $C(p)$ является липшиц-непрерывной.

функция $p \rightarrow c(p, u, u)$ является дифференцируемой при каждом фиксированном $u \in V$ и

$$\left| \frac{d}{dp} c(p, u, u) \right| \leq M(p) \|u\|^2,$$

M — локально ограниченная функция¹⁾, то $C(p)$ является дифференцируемой и

$$(C'(p)u, u) = \frac{d}{dp} c(p, u, u) \quad \forall u \in V.$$

Определение 1.1. Оператор-функции со свойствами а) и б) отнесем к классу $\mathcal{L}(\mathbb{R}_+, V)$. \square

Определим класс рассматриваемых задач. Будем считать, что:

- $A_1)$ $0 \leq A(p) \leq M_A(p)I$, $0 \leq B(p) \leq M_B I$ при $p \in \mathbb{R}_+$, где M_A — локально ограниченная функция;
- $A_2)$ $r_0 := \dim(\ker A(0)) \in [1, \infty)$. При каждом $p \in \mathbb{R}_+$: $B(p)$ — компактный оператор, $\ker B(p) = \ker B(0)$, $\dim(\text{Im} B(p)) = \infty$;
- $A_3)$ $A(p) + B(p) \geq m_{AB}I$ при $p \in \mathbb{R}_+$, $m_{AB} > 0$;
- $A_4)$ отношение Рэля $R(p, u) := a(p, u, u)/b(p, u, u)$ строго монотонно возрастает по $p \in \mathbb{R}_+$ при фиксированном $u \in \tilde{V}$, где \tilde{V} есть ортогональное дополнение $\ker B(0)$ до пространства $V_{A(p)+B(p)}$;
- $A_5)$ оператор-функции $A(p)$ и $B(p)$ принадлежат классу $\mathcal{L}(\mathbb{R}_+, V)$;

Отметим, что независимость величин M_B и m_{AB} от p не является для дальнейшего существенным.

Изучим по отдельности вопросы (i) и (ii), поставленные выше.

3.1. Прямая задача.

На вопросы (i) отвечает

¹⁾ограниченная на каждом компакте из $(0, \infty)$

Теорема 1.4. При каждом $p \in \mathbb{R}_+$ существует счетное множество чисел $\lambda^K(p)$, $K \geq 1$, с единственной точкой накопления $+\infty$, образующих полный набор собственных чисел задачи (\mathcal{P}) . Соответствующие им собственные подпространства $U^K(p)$, $K \geq 1$, — конечномерны, $V_{A(p)+B(p)} = \bigoplus_{k=1}^{\infty} U^K(p) \oplus \ker B(0)$. Кроме того,

$$\lambda^1(0) := \min_{K \geq 1} \lambda^K(0) = 0, \quad \dim U^1(0) = r_0.$$

ДОКАЗАТЕЛЬСТВО. Задача (\mathcal{P}) при фиксированном p является обобщенной задачей на собственные значения. Преобразуем ее к виду

$$\tilde{A}(p)u = \tilde{\lambda}B(p)u, \quad u \in V \setminus \{0\}. \quad (1.21)$$

Здесь $\tilde{A}(p) := A(p) + B(p)$, $\tilde{\lambda} := \lambda + 1$. Нетрудно видеть, что

$$m_{AB}I \leq \tilde{A}(p) \leq \tilde{M}_A(p)I, \quad \tilde{M}_A(p) := M_A(p) + M_B.$$

При фиксированном p задача (1.21) была рассмотрена нами ранее. Существование пар $(\lambda^K(p), U^K(p))$ непосредственно следуют из теоремы 1.1. Заключительные утверждения являются очевидными, так как $\ker A(0) \cap \ker B(0) = \{0\}$ (см. A_3). \square

Следствие 1.1. Задача (\mathcal{P}) эквивалентна задаче нахождения при каждом $p \in \mathbb{R}_+$ таких чисел $\lambda \in \mathbb{R}_+$ и ненулевых $u \in \tilde{V}$, что

$$a(p, u, v) = \lambda b(p, u, v) \quad \forall v \in \tilde{V}. \quad \square$$

Рассмотрим зависимость собственных чисел от параметра p . Для этого на плоскости с осями (p, λ) при каждом $p \in \mathbb{R}_+$ отметим точки $(p, \lambda^K(p))$, $K = 1, 2, \dots$. Совокупность этих точек образует на плоскости счетное множество кривых (возможно пересекающихся, сливающихся, расщепляющихся и т. д.), которые назовем дисперсионными (см. далее теорему 1.5). Каждой точке такой кривой, например $(p, \lambda^K(p))$, соответствует собственное подпространство $U^K(p)$, размерность которого (кратность $\lambda^K(p)$) обозначим через $r_K(p)$.

Чтобы изучить свойства дисперсионных кривых, нам будет удобно ввести нумерацию чисел $\lambda^K(p)$ с учетом кратности. А именно, пронумеруем их по возрастанию,

$$0 \leq \lambda_1(p) \leq \lambda_2(p) \leq \lambda_3(p) \leq \dots, \quad \lambda_n(p) \rightarrow \infty \text{ при } n \rightarrow \infty,$$

указывая каждое число $\lambda^K(p)$ в этом ряду столько раз, какова его кратность. Ясно, что функции $\lambda_i = \lambda_i(p)$, $p \in \mathbb{R}_+$, $i \geq 1$, в совокупности полностью определяют дисперсионные кривые.

Как и в предыдущем параграфе, номеру K числа $\lambda^K(p)$ поставим в соответствие целое k по правилу: $\lambda_k(p) = \lambda^K(p)$,

$$\lambda_{k-1}(p) < \lambda_k(p) = \dots = \lambda_{k+r_K(p)-1}(p) < \lambda_{k+r_K(p)}(p), \quad (1.22)$$

а в каждом пространстве $U^K(p)$ фиксируем некоторый ортонормированный в $V_{A(p)+B(p)}$ базис так, что

$$U^K(p) = \text{span}\{u_k(p), u_{k+1}(p), \dots, u_{k+r_K(p)-1}(p)\}. \quad (1.23)$$

Тогда при каждом $p \in \mathbb{R}_+$ последовательности $\lambda_i(p)$, $i = 1, 2, \dots$, будет поставлена в соответствие (неоднозначно) последовательность собственных элементов $u_i(p)$.

Теорема 1.5. *При каждом $i \geq 1$ функция $p \rightarrow \lambda_i(p)$*
a) является неотрицательной и строго монотонно возрастает;
b) непрерывна в нуле, локально липшиц-непрерывна на $(0, \infty)$.
Кроме того, $\lambda_i(p) \searrow +0$ при $p \rightarrow +0$, $i = 1, \dots, r_0$; $\lambda_{r_0+1}(0) > 0$.

ДОКАЗАТЕЛЬСТВО. При каждом $p \in \mathbb{R}_+$ числа $\lambda_i(p) + 1$, $i \geq 1$, являются собственными числами задачи (1.21). Учитывая следствие 1.1 и используя минимаксный принцип Куранта-Фишера, имеем

$$\frac{1}{\lambda_i(p) + 1} = \max_{V_i \subset \tilde{V}} \min_{v \in V_i \setminus \{0\}} \tilde{R}(p, v), \quad (1.24)$$

где максимум берется по всем подпространствам \tilde{V} размерности i ,

$$\tilde{R}(p, v) := \frac{b(p, v, v)}{a(p, v, v) + b(p, v, v)} = \frac{1}{a(p, v, v)/b(p, v, v) + 1}.$$

Из условия A_4 заключаем, что $\tilde{R}(p, v)$ строго монотонно убывает по p при фиксированном $v \in \tilde{V}$; поэтому утверждение *a)* следует из (1.24).

Пусть $p, \bar{p} \in \mathbb{R}_+$, $a(p) = a(p, v, v)$, $b(p) = b(p, v, v)$. Имеем

$$R(p, v) - R(\bar{p}, v) = \frac{(b(p) - b(\bar{p}))a(\bar{p}) + (a(\bar{p}) - a(p))b(\bar{p})}{(a(p) + b(p))(a(\bar{p}) + b(\bar{p}))}.$$

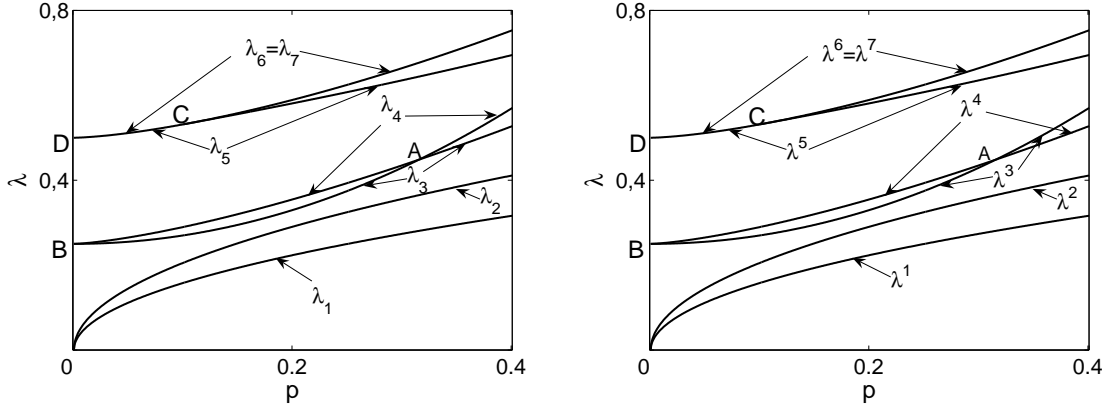


Рис. 1. На левом рисунке изображена часть плоскости (p, λ) и гладкие дисперсионные кривые, пронумерованные по возрастанию с учетом кратности. В точках 0 и B рождается по две кривых кратности 1, в D — одна кривая кратности 3; в точке A кривые пересекаются, в C — расщепляются. На правом рисунке использована другая нумерация этих же кривых (также с учетом кратности).

Отсюда легко выводится оценка $|R(p, v) - R(\bar{p}, v)| = \epsilon(p, \bar{p})$, где

$$\begin{aligned} \epsilon(p, \bar{p}) &:= c(\bar{p}) (\|A(p) - A(\bar{p})\| + \|B(p) - B(\bar{p})\|), \\ c(\bar{p}) &:= m_{AB}^{-2} (M_A(\bar{p}) + M_B). \end{aligned}$$

Таким образом, $R(p, v) - \epsilon(p, \bar{p}) \leq R(\bar{p}, v) \leq R(p, v) + \epsilon(p, \bar{p})$; поэтому из принципа Куранта-Фишера имеем

$$\lambda_i(p) - \epsilon(p, \bar{p}) \leq \lambda_i(\bar{p}) \leq \lambda_i(p) + \epsilon(p, \bar{p}),$$

или $|\lambda_i(p) - \lambda_i(\bar{p})| \leq \epsilon(p, \bar{p})$. Отсюда следуют требуемые утверждения b), поскольку при $\bar{p} = 0$ имеем (см. условие A_5):

$$\epsilon(p, 0) = c(0) (\|A(p) - A(0)\| + \|B(p) - B(0)\|) \rightarrow 0 \quad \text{при } p \rightarrow 0,$$

а для всех $p, \bar{p} \in \omega := [\omega_1, \omega_2] \subset (0, \infty)$

$$\epsilon(p, \bar{p}) \leq c_\omega (L_A(\omega) + L_B(\omega)) |p - \bar{p}|,$$

где c_ω есть верхняя грань $c(p)$ на ω .

Наконец, число $\lambda^1(0) = 0$ имеет кратность r_0 (см. теорему 1.4). Поэтому $\lambda_i(0) = 0$, $i \leq r_0$, и заключительное утверждение теоремы следует из монотонности и непрерывности функций λ_i в нуле. \square

ЗАМЕЧАНИЕ 1.1. Утверждение о локальной липшицевости функций $\lambda_i(p)$ не может быть усилено, даже если дисперсионные кривые являются бесконечно дифференцируемыми. Нетрудно видеть, что это связано с возможным пересечением кривых (см. левый рис. 1; функция $\lambda_3(p)$ не имеет производной в точке, соответствующей точке A).

3.2. Обратная задача.

Посмотрим на задачу (\mathcal{P}) с обратной точки зрения (см. вопросы (ii)). Будем считать $\lambda \geq 0$ фиксированным и искать такие $(p, u) \in \mathbb{R}_+ \times V \setminus \{0\}$, что

$$A(p)u = \lambda B(p)u.$$

Эта задача существенно сложнее, чем (\mathcal{P}) , поскольку она представляет собой нелинейную (относительно p) задачу на собственные значения. Тем не менее, благодаря монотонной зависимости отношения Рэлея от p , оказывается возможным дать конструктивный ответ на вопрос (ii). С этой целью рассмотрим “линейную” задачу

$$A(0)u = \lambda^0 B(0)u, \quad u \in V \setminus \{0\}. \quad (1.25)$$

Эта задача на собственные значения была изучена нами выше. Она имеет счетное множество неотрицательных собственных чисел λ^{0K} , которые пронумеруем по возрастанию с учетом кратности:

$$0 = \lambda_1^0 = \lambda_2^0 = \dots = \lambda_{r_0}^0 < \lambda_{r_0+1}^0 \leq \lambda_{r_0+2}^0 \leq \dots, \quad \lambda_n^0 \rightarrow \infty \text{ при } n \rightarrow \infty.$$

Числа λ_i^0 , $i \geq 1$ будем называть критическими числами, а на задачу (1.25) будем ссылаться как на задачу определения критических чисел.

Определим на полуоси \mathbb{R}_+ ступенчатую функцию, непрерывную слева, монотонно возрастающую и принимающую целые значения:

$$n(x) := \max\{i : \lambda_i^0 \leq x, i = 1, 2, \dots\}.$$

Отметим, что точками разрыва этой функции являются числа λ^{0K} , на множестве $[0, \lambda_{r_0+1}^0)$ ее значения равны r_0 .

Теорема 1.6. *При каждом $\lambda \in \mathbb{R}_+$ существует в точности $n(\lambda)$ таких неотрицательных чисел $p_i(\lambda)$ (необязательно различных) и соответствующих им элементов $u_i(\lambda)$ (одинаковым p_i соответствуют линейно-независимые u_i), что $(\lambda, u_i(\lambda))$ есть решение задачи (\mathcal{P}) при $p = p_i(\lambda)$.*

ДОКАЗАТЕЛЬСТВО. Пусть задано неотрицательное число λ , $(\lambda_i(p), u_i(p))$, $i = 1, 2, \dots$ — решения задачи (\mathcal{P}) . Поскольку функции $p \rightarrow \lambda_i(p)$ строго монотонно возрастают, то алгебраическое уравнение

$$\lambda_i(p) = \lambda$$

при фиксированном $i \geq 1$ либо имеет одно решение $p_i = p_i(\lambda)$ (если $\lambda_i(0) \leq \lambda$), либо не имеет решения. Из определения функции $n(x)$ следует, что в точности $n(\lambda)$ таких уравнений имеют решение. \square

§ 4. Метод Галеркина с возмущениями для параметрической задачи

4.1. Дискретная задача.

Пусть, как и в § 2, $\{V_h\}_h$ — предельно плотная последовательность конечномерных подпространств V , симметричные билинейные формы $a_h(p, \cdot, \cdot)$ и $b_h(p, \cdot, \cdot)$ есть возмущения (аппроксимации) форм $a(p, \cdot, \cdot)$ и $b(p, \cdot, \cdot)$ на $V_h \times V_h$ при фиксированных h и p . Рассмотрим аппроксимацию задачи (\mathcal{P}) методом Галеркина с возмущениями: *при каждом $p \in \mathbb{R}_+$ найти такие $\lambda^h \in \mathbb{R}$ и ненулевые $y \in V_h$, что*

$$(\mathcal{P}_h) \quad a_h(p, y, v) = \lambda^h b_h(p, y, v) \quad \forall v \in V_h.$$

Пусть h и p — фиксированы, N_h есть размерность пространства V_h , а элементы $\varphi_1, \varphi_2, \dots, \varphi_{N_h}$ определяют в нем базис. Тогда задача (\mathcal{P}_h) сводится к алгебраической задаче на собственные значения

$$A_h(p)x = \lambda^h B_h(p)x, \quad 0 \neq x \in \mathbb{R}^{N_h}.$$

Здесь x есть вектор коэффициентов разложения y по указанному базису, симметричные матрицы $A_h(p)$ и $B_h(p)$ имеют элементы

$$a_{ij}(p) := a_h(p, \varphi_i, \varphi_j), \quad b_{ij}(p) := b_h(p, \varphi_i, \varphi_j), \quad i, j = 1, 2, \dots, N_h.$$

Будем говорить, что вектор x соответствует элементу y , а матрица $A_h(p)$ ($B_h(p)$) порождается формой a_h (b_h).

Определение 1.2. Матричная функция $p \rightarrow C_h(p)$ является
(a) непрерывной в нуле, если

$$\|C_h(p) - C_h(0)\| \rightarrow 0 \text{ при } p \rightarrow +0;$$

(b) локально липшиц-непрерывной на $(0, \infty)$, если для любого замкнутого отрезка $\omega \subset (0, \infty)$ найдется такая постоянная $L_c(\omega)$, что

$$\|C_h(p) - C_h(\bar{p})\| \leq L_c(\omega)|p - \bar{p}|, \quad p, \bar{p} \in \omega, \quad u \in V_h.$$

Симметричные матричные функции со свойствами (a) и (b) отнесем к классу $\mathcal{L}(\mathbb{R}_+, V_h)$. \square

Наложим следующие ограничения на формы a_h и b_h :

- A_1^h) $|a_h(p, u, u)| \leq M_a(p) \|u\|_{1,\Omega}^2$, $0 \leq b_h(p, u, u) \leq M_b \|u\|_{1,\Omega}^2$, $p \in \mathbb{R}_+$,
 $u \in V_h$, где M_a — локально ограниченная функция;
- A_2^h) $\ker B_h(p) = \ker B_h(0)$ при каждом $p \in \mathbb{R}_+$;
- A_3^h) $A_h(p) + B_h(p) \geq m_{ab} I_h$ при $p \in \mathbb{R}_+$, $m_{ab} > 0$;
- A_4^h) отношение Рэля $a_h(p, u, u)/b_h(p, u, u)$ строго монотонно возрастает по p при любом фиксированном $u \in \tilde{V}_h$, где \tilde{V}_h есть ортогональное дополнение $\ker B_h(0)$ до пространства $V_{h, A_h(p)+B_h(p)}$ ¹⁾;
- A_5^h) матричные функции $A_h(p)$ и $B_h(p)$ принадлежат классу $\mathcal{L}(\mathbb{R}_+, V_h)$.

Здесь величины M_a , M_b и m_{ab} считаются не зависящими от h . Положим $N_h^b := N_h - \dim(\ker B_h(0))$.

Теорема 1.7. *При каждом $p \in \mathbb{R}_+$ задача (\mathcal{P}_h) имеет конечное число (скажем $n_h(p)$) собственных чисел $\lambda^{hK}(p)$ суммарной кратности N_h^b и соответствующих им собственных подпространств*

$$U_h^K(p), V_{h, A_h(p)+B_h(p)} = \bigoplus_{K=1}^{n_h(p)} U_h^K(p) \oplus \ker B_h(0).$$

Кроме того, если $\lambda_1^h(p) \leq \lambda_2^h(p) \leq \dots \leq \lambda_{N_h^b}^h(p)$ есть собственные числа, пронумерованные с учетом кратности, то функции $p \rightarrow \lambda_i^h(p)$, $i = 1, 2, \dots, N_h^b$, являются непрерывными в нуле, локально липшиц-непрерывными на $(0, \infty)$ и строго монотонно возрастающими.

ДОКАЗАТЕЛЬСТВО. Размерность пространства \tilde{V}_h равна N_h^b . Поскольку $\ker B_h(p) = \ker B_h(0)$ при всех $p \in \mathbb{R}_+$ и $\ker A_h(0) \cap \ker B_h(0) = \{0\}$ (см. условие A_3^h), то элементы из $\ker B_h(0)$ не могут быть собственными элементами. Преобразуем задачу (\mathcal{P}_h) к виду

$$\tilde{a}_h(p, y, v) = \tilde{\lambda}^h b_h(p, y, v), \quad v \in \tilde{V}_h,$$

¹⁾ $V_{h, A_h(p)+B_h(p)}$ есть пространство V_h с скалярным произведением $a_h(p, \cdot, \cdot) + b_h(p, \cdot, \cdot)$.

где $\tilde{a}_h := a_h + b_h$, $\tilde{\lambda}^h := \lambda^h + 1$. Обе формы \tilde{a}_h и b_h положительно определены на \tilde{V}_h . Отсюда следует существование пар $(\lambda^{hK}(p), U_h^K(p))$ и свойства пространств $U_h^K(p)$.

Доказательство остальных утверждений теоремы проводится совершенно аналогично доказательству соответствующих утверждений теоремы 1.5. \square

ЗАМЕЧАНИЕ 1.2. Сопоставим условия $A_1 - A_5$ и $A_1^h - A_5^h$. Видим, что формы a_h и b_h наследуют свойства невозмущенных форм a и b , кроме условия неотрицательности формы a . Если условия $A_1^h - A_5^h$ дополнить условием

$$A_0^h \quad A_h(p) \geq 0, \quad r_0 := \dim(\ker A_h(0)),$$

то можно утверждать также, что (см. аналогичное утверждение теоремы 1.5):

$$\lambda_i^h(p) \rightarrow 0 \text{ при } p \rightarrow +0, \quad i = 1, 2, \dots, r_0; \quad \lambda_{r_0+1}^h(0) > 0.$$

4.2. Оценки точности.

Пусть $K \geq 1$, $\lambda^K(p)$ — собственное число кратности $r_K(p)$ задачи (\mathcal{P}) ; $U^K(p)$ — соответствующее ему собственное подпространство (см. (1.23), с. 23); целое k связано с K согласно (1.22), с. 23. Пусть далее $\lambda_i^h(p), i = 1, \dots, N_h^b$ — упорядоченные по возрастанию с учетом кратности собственные числа задачи (\mathcal{P}_h) , $y_i^h(p)$ — соответствующие им собственные элементы. Числа $\lambda_k^h(p), \lambda_{k+1}^h(p), \dots, \lambda_{k+r_K(p)-1}^h(p)$ будем рассматривать как приближения собственного числа $\lambda^K(p)$, а пространство

$$U_h^K(p) := \text{span}\{y_k^h(p), y_{k+1}^h(p), \dots, y_{k+r_K(p)-1}^h(p)\},$$

как аппроксимацию собственного подпространства $U^K(p)$.

Введем обозначения. Для заданных форм d и d_h положим

$$E_d(\varphi_h) := \sup_{v_h \in V_h, \|v_h\|=1} |d(p, \varphi_h, v_h) - d_h(p, \varphi_h, v_h)|, \quad \varphi_h \in V_h,$$

через P_h обозначим ортопроектор в $V_{A(p)+B(p)}$ на V_h . Пусть

$$\epsilon_h(u) := \inf_{v_h \in V_h} \|u - v_h\| + E_a(P_h u) + E_b(P_h u), \quad u \in V,$$

$$\Sigma_h(y) := |a(p, y, y) - a_h(p, y, y)| + |b(p, y, y) - b_h(p, y, y)|, \quad y \in V_h.$$

Дополнительно к условиям $A_1^h - A_5^h$ предположим, что

$$A_6^h) \quad \sup_{f \in V_h, \|f\|=1} (E_b(f) + E_b(P_h T(p)f) + E_a(P_h T(p)f)) \rightarrow 0 \quad \text{при } h \rightarrow 0,$$

где $T(p) := (A(p) + B(p))^{-1} B(p)$.

Теорема 1.8. Пусть выполнены условия (H_1) , $A_1 - A_5$, $A_1^h - A_6^h$, $\max_{u \in U^K, \|u\|=1} \epsilon_h(u) \rightarrow 0$ при $h \rightarrow 0$. Тогда для любого $p \in \mathbb{R}_+$ и $K \geq 1$ при достаточно малом h имеют место оценки

$$\begin{aligned} \Theta_V(U^K(p), U_h^K(p)) &\leq c_K \max_{u \in U^K(p), \|u\|=1} \epsilon_h(u), \\ |\lambda^K(p) - \lambda_i^h(p)| &\leq c_K \max_{u \in U^K(p), \|u\|=1} (\epsilon_h^2(u) + \Sigma_h(P_h u)), \end{aligned}$$

где постоянная c_K зависит от $\lambda^K(p)$ и p , но не зависит от h .

ДОКАЗАТЕЛЬСТВО. Воспользуемся результатами § 2. Исходную задачу (\mathcal{P}) представим в виде

$$b(p, u, v) = \mu \tilde{a}(p, u, v) \quad \forall v \in V,$$

где $\tilde{a} := a + b$, $\mu := (\lambda + 1)^{-1}$. При фиксированном p эта задача типа (P_μ) . Задачу (\mathcal{P}_h) представим в виде ее аппроксимации (P_μ^h) :

$$b_h(p, y, v) = \mu^h \tilde{a}_h(p, y, v) \quad \forall v \in V_h,$$

где $\tilde{a}_h := a_h + b_h$, $\mu^h := (\lambda^h + 1)^{-1}$. Из условий A_2^h , A_3^h и A_6^h следует, что выполнены условия (H_2) и (H_3) . Поэтому мы можем воспользоваться теоремой 1.2, из которой следует первая оценка теоремы.

Из теоремы 1.3 вытекает, что

$$|\mu^K(p) - \mu_i^h(p)| \leq c_K \max_{u \in U^K(p), \|u\|=1} (\epsilon_h^2(u) + \Sigma_h(P_h u)).$$

Отсюда следует вторая оценка, поскольку при достаточно малом h

$$|\lambda^K(p) - \lambda_i^h(p)| = |1/\mu^K(p) - 1/\mu_i^h(p)| \leq c_K |\mu^K(p) - \mu_i^h(p)|. \quad \square$$

ГЛАВА 2

СКАЛЯРНАЯ ЗАДАЧА О ПОВЕРХНОСТНЫХ СОБСТВЕННЫХ ВОЛНАХ

В данной главе изучается один численный метод решения задачи о поверхностных собственных волнах цилиндрических диэлектрических волноводов, находящихся в однородной окружающей среде, в приближении слабонаправляющего волновода. Он основан на сочетании метода точных нелокальных граничных условий и метода конечных элементов.

Рассматривается классическая модель, согласно которой волновод предполагается неограниченным и линейно изотропным, то есть относительная диэлектрическая проницаемость ε волновода не меняется вдоль оси Ox_3 и является достаточно гладкой функцией поперечных координат. Будем для определенности считать, что поперечное сечение волновода Ω_i есть односвязная ограниченная область в \mathbb{R}^2 , имеющая кусочно-гладкую границу $\gamma := \partial\Omega_i$,

$$\varepsilon := \varepsilon_\infty > 0 \text{ в } \Omega_e, \quad \varepsilon_+ := \max_{x \in \bar{\Omega}_i} \varepsilon(x) > \varepsilon_\infty, \quad \min_{x \in \bar{\Omega}_i} \varepsilon(x) > \varepsilon_\infty^1). \quad (2.1)$$

где $\Omega_e := \mathbb{R}^2 \setminus \bar{\Omega}_i$. Магнитная проницаемость волновода всюду предполагается равной магнитной проницаемости свободного пространства.

Математическая формулировка задачи сводится к нахождению пар чисел $(\beta, k) \in \Lambda$ и ненулевых вещественных $u \in H^1(\mathbb{R}^2)$, удовлетворяющих п. вс. в \mathbb{R}^2 уравнению [18]

$$-\Delta u + \beta^2 u = k^2 \varepsilon u, \quad (2.2)$$

или, в обобщенной постановке, тождеству

$$(\mathcal{P}_\infty) \quad \int_{\mathbb{R}^2} (\nabla u \cdot \nabla v + \beta^2 uv) dx = k^2 \int_{\mathbb{R}^2} \varepsilon uv dx \quad \forall v \in H^1(\mathbb{R}^2).$$

¹⁾Рассмотрение многосвязных и несвязных областей Ω_i , так же как и волноводов с размытой границей (в этом случае необходимо писать знак \geq вместо $>$ в последнем условии в (2.1)), не вносит дополнительных трудностей.

Здесь

$$\Lambda := \{(\beta, k) : \beta/\sqrt{\varepsilon_+} < k < \beta/\sqrt{\varepsilon_\infty}, \beta > 0\}.$$

Отметим, что в $\Omega_e := \mathbb{R}^2 \setminus \bar{\Omega}_i$ уравнение (2.2) имеет вид

$$-\Delta u + p^2 u = 0, \quad p := (\beta^2 - k^2 \varepsilon_\infty)^{1/2}. \quad (2.3)$$

Параметр p определяет скорость затухания решения u на бесконечности и называется поперечным волновым числом.

На задачу (2.2) можно смотреть как на параметрическую спектральную задачу, считая параметром либо β , либо k .

§ 1. Постановка задачи в ограниченной области

Введем новую пару неизвестных параметров (β, p) , предполагая, что p и k связаны равенством

$$p = (\beta^2 - k^2 \varepsilon_\infty)^{1/2}, \quad (2.4)$$

и множество

$$K := \{(\beta, p) : \beta > 0, 0 < p < \sqrt{(\varepsilon_+ - \varepsilon_\infty)/\varepsilon_+} \beta\}.$$

Легко проверяется, что формула (2.4) устанавливает взаимно-однозначное соответствие между Λ и K .¹⁾

Укажем два метода, каждый из которых позволяет свести исходную задачу в неограниченной области (на плоскости) к задаче в ограниченной области (в круге). Первый метод основывается на дифференциальном уравнении (2.2), а второй — на ее слабой форме (\mathcal{P}_∞).

1.1. Первый метод.

Рассмотрим новую задачу: найти $(\beta, p) \in K$ и ненулевые $u \in H^1(\mathbb{R}^2)$, удовлетворяющие п. вс. в \mathbb{R}^2 уравнению

$$-\Delta u + p^2 \sigma u = \beta^2 (\sigma - 1) u, \quad \sigma = \varepsilon/\varepsilon_\infty. \quad (2.5)$$

Легко заметить, что уравнение (2.5) переходит в (2.2) при условии (2.4) и наоборот. Поэтому эти задачи эквивалентны в том смысле,

¹⁾ достаточно заметить, что лучи $\{k = k_0 \beta, \beta \in \mathbb{R}_+\}$, лежащие в Λ , преобразуются в лучи $\{p = p_0 \beta, \beta \in \mathbb{R}_+\}$ из K , $p_0^2 = 1 - k_0^2 \varepsilon_\infty$.

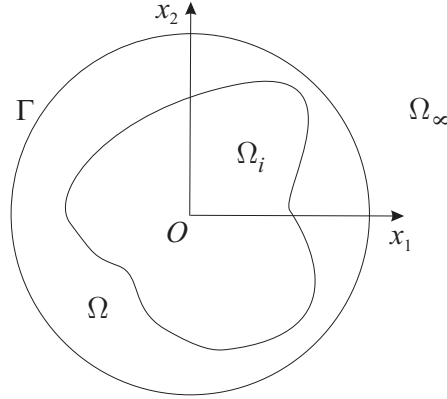


Рис. 1. Разбиение плоскости.

что между их решениями имеется простое взаимно-однозначное соответствие: (β, k, u) — решение задачи (2.2) тогда и только тогда, когда (β, p, u) — решение (2.5).

Сведем задачу (2.5) к задаче в ограниченной области, предполагая, что начало координат находится в Ω_i .

Пусть $B_{R_0} := \{x : |x| < R_0\}$ — круг минимального радиуса R_0 такой, что $\Omega_i \subseteq B_{R_0}$. Выберем R так, что $R \geq R_0$ и положим $\Omega := B_R$, $\Gamma := \partial\Omega$, $\Omega_\infty := R^2 \setminus \bar{\Omega}$ (см. рис. 1).

Решение задачи (2.5) удовлетворяет уравнению (2.3) вне Ω_i и поэтому является там гладким. Обозначим через u_p (u) его сужение на область $\bar{\Omega}_\infty$ ($\bar{\Omega}$). Тогда u_p удовлетворяет уравнению (2.3), а на Γ справедливы равенства

$$u_p = u, \quad u_{p\nu} = u_\nu.$$

Здесь ν — единичная нормаль из Ω на Γ , u_ν — производная u в направлении ν . Функция u_p , как решение внешней краевой задачи

$$-\Delta u_p + p^2 u_p = 0, \quad x \in \Omega_\infty, \quad u_p = u, \quad x \in \Gamma,$$

легко находится методом разделения переменных. Имеем

$$u_p = \sum_{n=-\infty}^{\infty} \frac{K_n(pr)}{K_n(pR)} a_n(u) e^{in\varphi}, \quad a_n(u) := \frac{1}{2\pi} \int_0^{2\pi} u|_{r=R} e^{-in\varphi} d\varphi. \quad (2.6)$$

Здесь (r, φ) — полярные координаты с началом в центре B_R , $K_n(r)$ — модифицированная функция Бесселя порядка n . Таким образом, для

определения (β, p, u) получим следующую задачу в круге, в которую параметр p нелинейно входит в нелокальное краевое условие:

$$\begin{aligned} -\Delta u + p^2 \sigma u &= \beta^2 (\sigma - 1)u, & x \in \Omega, \\ u_\nu + S_\Gamma(p)u &= 0, & x \in \Gamma. \end{aligned} \quad (2.7)$$

Здесь

$$S_\Gamma(p)u := -u_{p\nu} = \frac{1}{R} \sum_{n=-\infty}^{\infty} \mathbb{K}_n(Rp) a_n(u) e^{in\varphi}, \quad \mathbb{K}_n(r) := -r \frac{K'_n(r)}{K_n(r)}.$$

Эта задача и есть искомая задача в ограниченной области.

Далее мы исследуем разрешимость задачи (2.7), а также построим и изучим численный метод ее решения на основе метода конечных элементов. Для этого нам понадобится ее обобщенная (слабая) формулировка. Она получается стандартным способом: уравнение (2.7), предварительно умноженное на произвольно выбранную функцию $v \in V := H^1(\Omega)$, интегрируется по области Ω , а затем используется формула интегрирования по частям и учитывается краевое условие на Γ . Обобщенная формулировка задачи (2.7) имеет следующий вид:

найти $(\beta, p, u) \in K \times V \setminus \{0\}$ такие, что

$$(\mathcal{P}) \quad a(p, u, v) = \beta^2 b(u, v) \quad \forall v \in V,$$

где

$$\begin{aligned} a(p, u, v) &:= a_0(p, u, v) + s_\infty(p, u, v), \\ s_\infty(p, u, v) &:= 2\pi \sum_{n=-\infty}^{\infty} \mathbb{K}_n(pR) a_n(u) \overline{a_n(v)}, \end{aligned} \quad (2.8)$$

$$a_0(p, u, v) := \int_{\Omega} (\nabla u \cdot \nabla v + p^2 \sigma uv) dx, \quad b(u, v) := \int_{\Omega} (\sigma - 1)uv dx.$$

Здесь для комплексного числа $z = a + ib$, $\bar{z} = a - ib$ означает его комплексное сопряжение.

ЗАМЕЧАНИЕ 2.3. Формы a_0 и b , в отличие от формы s_∞ , являются традиционными. Свойства s_∞ определяются как скоростью убывания с ростом n коэффициентов Фурье $a_n(u)$, $u \in V$, так и свойствами функций $r \rightarrow \mathbb{K}_n(r)$ при произвольных n . Далее мы покажем, в частности, что $\mathbb{K}_n(r) > |n|$ при всех $r > 0$ и n и $\mathbb{K}_n(r) < |n| + r$ при $n \neq 0$. Отметим также, что в силу равенства Парсеваля

$$2\pi \sum_{n=-\infty}^{\infty} |a_n(u)|^2 = \frac{1}{R} \int_{\Gamma} u^2 d\Gamma. \quad (2.9)$$

ЗАМЕЧАНИЕ 2.4. Условие $(\beta, p) \in K$ в формулировке задачи (\mathcal{P}) не является ограничением. Действительно, множество K допускает следующее эквивалентное определение:

$$K := \{(\beta, p) : p > 0, \beta > (\varepsilon_+ / (\varepsilon_+ - \varepsilon_\infty))^{1/2} p\}.$$

Если (β, p, u) удовлетворяет тождеству (\mathcal{P}) , то нетрудно видеть, что $(1 \leq \sigma \leq \varepsilon_+ / \varepsilon_\infty)$

$$\int_{\Omega} (|\nabla u|^2 + \beta^2 u^2) dx + s_\infty(p, u, u) = \int_{\Omega} \sigma(\beta^2 - p^2) u^2 dx.$$

Отсюда следует $\beta^2 > p^2$, так как $s_\infty(p, u, u) \geq 0$, а также оценка $\beta^2 < \varepsilon_+ / \varepsilon_\infty (\beta^2 - p^2)$, равносильная оценке $\beta > (\varepsilon_+ / (\varepsilon_+ - \varepsilon_\infty))^{1/2} p$; следовательно, $(\beta, p) \in K$.

Поэтому задачу (\mathcal{P}) можно рассматривать как задачу на собственные значения относительно β^2 , а $p > 0$ — считать параметром. Задачи такого типа были рассмотрены нами в предыдущей главе.

Прежде чем перейти к второму методу вывода задачи (\mathcal{P}) , рассмотрим эквивалентные нормировки пространства $H^{1/2}(\Gamma)$ и укажем гладкие продолжения функций из $H^1(\Omega)$ с сохранением класса. На этой основе мы получим другое представление формы s_∞ .

1.2. Пространство $H^{1/2}(\Gamma)$.

Используем общепринятое обозначение $H^{1/2}(\Gamma)$ для пространства следов на Γ функций из $H^1(\mathbb{R}^2)$ (см., напр., [19, с. 55]); через $u|_\Gamma$ будем обозначать след функции u на Γ ¹⁾. Норма в $H^{1/2}(\Gamma)$ может быть определена следующими двумя эквивалентными способами²⁾:

$$\|u\|_{1/2, \Gamma} := \inf_{v \in H^1(\Omega), v|_\Gamma = u} \|v\|_{1, \Omega}, \quad |u|_{1/2, \Gamma} := \inf_{v \in H^1(\Omega_\infty), v|_\Gamma = u} \|v\|_{1, \Omega_\infty}.$$

Известно, что $H^{1/2}(\Gamma) \subset L_2(\Gamma)$, бесконечно-дифференцируемые функции образуют плотное подмножество в нем. В частности, ряд Фурье функции $u \in H^{1/2}(\Gamma)$

$$\sum_{n=-\infty}^{\infty} a_n(u) e^{in\varphi}, \quad a_n(u) = \frac{1}{2\pi} \int_0^{2\pi} u(R\varphi) e^{-in\varphi} d\varphi, \quad (2.10)$$

¹⁾ понятие следа обобщает понятие сужения. Для гладких функций эти понятия совпадают.

²⁾ две нормы $\|\cdot\|_{(1)}$ и $\|\cdot\|_{(2)}$ эквивалентны на V , если $c_1 \|u\|_{(2)} \leq \|u\|_{(1)} \leq c_2 \|u\|_{(2)} \quad \forall u \in V$, $c_1 > 0$. Числа c_1 и c_2 называют постоянными эквивалентности.

³⁾ $s = R\varphi$ есть дуговая координата на Γ

сходится к u , то есть, если u_N есть отрезок ряда Фурье,

$$u_N(s) = \sum_{n=-N}^N a_n(u) e^{in\varphi},$$

то $\|u - u_N\|_{1/2,\Gamma} \rightarrow 0$ при $N \rightarrow \infty$ (также $\|u_N\|_{1/2,\Gamma} \rightarrow \|u\|_{1/2,\Gamma}$). Отметим, что поскольку u_N вещественная функция, то

$$u_N = \bar{u}_N = \sum_{n=-N}^N \overline{a_n(u)} e^{-in\varphi}.$$

Лемма 2.1. Пусть $I_n(r)$, $K_n(r)$ — модифицированные функции Бесселя порядка n ,

$$\mathbb{I}_n(r) := r \frac{I'_n(r)}{I_n(r)}, \quad \mathbb{K}_n(r) := -r \frac{K'_n(r)}{K_n(r)}, \quad r \in \mathbb{R}_+,$$

функция $u \in H^{1/2}(\Gamma)$ представлена рядом Фурье (2.10). Тогда

$$\|u\|_{1/2,\Gamma}^2 = 2\pi \sum_{n=-\infty}^{\infty} \mathbb{I}_n(R) |a_n(u)|^2, \quad |u|_{1/2,\Gamma}^2 = 2\pi \sum_{n=-\infty}^{\infty} \mathbb{K}_n(R) |a_n(u)|^2.$$

ДОКАЗАТЕЛЬСТВО. Вычислим норму $\|u\|_{1/2,\Gamma}$. Пусть u_N — отрезок ряда Фурье функции u ,

$$\|\tilde{u}_N\|_{1,\Omega}^2 = \inf_{v \in H^1(\Omega), v|_{\Gamma} = u_N} \left\{ \|v\|_{1,\Omega}^2 := \int_{\Omega} (|\nabla v|^2 + |v|^2) dx \right\}.$$

Хорошо известно, что \tilde{u}_N является решением следующей внутренней краевой задачи¹⁾:

$$-\Delta \tilde{u}_N + \tilde{u}_N = 0, \quad x \in \Omega, \quad \tilde{u}_N|_{\Gamma} = u_N.$$

Ее решение находится методом разделения переменных:

$$\tilde{u}_N = \sum_{n=-N}^N \frac{I_n(r)}{I_n(R)} a_n(u) e^{in\varphi}.$$

¹⁾решение этой задачи является бесконечно дифференцируемым в $\bar{\Omega}$.

Таким образом, используя интегрирование по частям, имеем:

$$\begin{aligned} \|u_N\|_{1/2,\Gamma}^2 &:= \|\tilde{u}_N\|_{1,\Omega}^2 = \int_{\Omega} (-\Delta \tilde{u}_N + \tilde{u}_N) \tilde{u}_N dx + \int_{\Gamma} \frac{\partial \tilde{u}_N}{\partial r} u_N ds = \\ &= \int_{\Gamma} \frac{\partial \tilde{u}_N}{\partial r} \tilde{u}_N ds = 2\pi \sum_{n=-N}^N \mathbb{I}_n(R) |a_n(u)|^2. \end{aligned}$$

Переходя здесь к пределу при $N \rightarrow \infty$, получим первую формулу. Аналогично получается вторая формула. При этом внутренняя краевая задача заменяется на внешнюю (в области Ω_{∞}), функция Бесселя I_n — на K_n . \square

ЗАМЕЧАНИЕ 2.5. Целью данного замечания является вариационное описание функций $R \rightarrow \mathbb{I}_n(R)$ и $R \rightarrow \mathbb{K}_n(R)$, $R > 0$. Как показано выше

$$2\pi \sum_{n=-\infty}^{\infty} \mathbb{I}_n(R) |a_n(u)|^2 = \inf_{v \in H^1(B_R), v|_{\partial B_R} = u} \left\{ \int_{B_R} (|\nabla v|^2 + |v|^2) dx \right\}. \quad (2.11)$$

Выберем здесь $u = (e^{ik\varphi} + e^{-ik\varphi})/2 = \cos(k\varphi)$. Тогда левая часть в (2.11) будет равна $c_k \mathbb{I}_k(R)$, где все $c_k = \pi$, кроме $c_0 = 2\pi$ (поскольку $\mathbb{I}_{-n}(R) = \mathbb{I}_n(R)$, см. далее). В правой части сделаем замену переменных в интеграле ($r := r/R$), переходя к единичному кругу B_1 . В результате получим для любого целого k следующую формулу:

$$\mathbb{I}_k(R) = \frac{1}{c_k} \inf_{v \in H^1(B_1), v|_{\partial B_1} = \cos(k\varphi)} \left\{ \int_{B_1} (|\nabla v|^2 + R^2 |v|^2) dx \right\}.$$

Обозначая через B_1^{∞} дополнение B_1 до \mathbb{R}^2 , аналогично получим

$$\mathbb{K}_k(R) = \frac{1}{c_k} \inf_{v \in H^1(B_1^{\infty}), v|_{\partial B_1^{\infty}} = \cos(k\varphi)} \left\{ \int_{B_1^{\infty}} (|\nabla v|^2 + R^2 |v|^2) dx \right\}.$$

Отсюда следует, в частности, что функция $r \rightarrow \mathbb{K}_n(r)$ (как и $r \rightarrow \mathbb{I}_n(r)$) — строго монотонно возрастает при всех целых n и $\mathbb{K}_k(R) \leq (R/R_0)^2 \mathbb{K}_k(R_0)$ при $R \geq R_0$.

Лемма 2.2. *Для всех целых n справедливы оценки*

$$\mathbb{I}_n(r) \leq \mathbb{K}_n(r) \leq c_0^2(r) \mathbb{I}_n(r), \quad r \in \mathbb{R}_+,$$

где c_0 — непрерывная на $(0, \infty)$ монотонно убывающая функция; $c_0(r) \rightarrow +\infty$ при $r \rightarrow 0$.

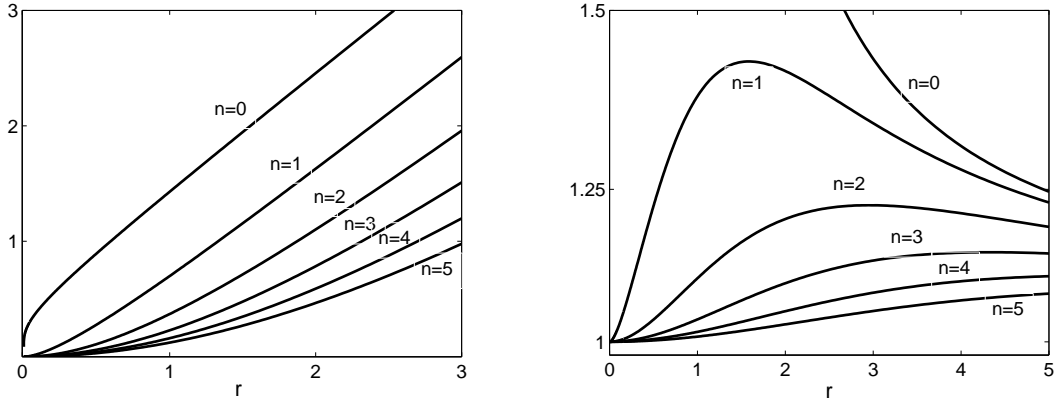


Рис. 2. На левом рисунке представлены графики функций $\mathbb{K}_n(r)$ — n для ряда значений n , на правом — функций $\mathbb{K}_n(r)/\mathbb{I}_n(r)$.

ДОКАЗАТЕЛЬСТВО. Модифицированные функции Бесселя $I_n(z)$ и $K_n(z)$ ¹⁾ являются линейно независимыми решениями уравнения

$$z^2 w'' + z w' - (z^2 + n^2)w = 0.$$

Ясно, что $K_{-n}(z) = K_n(z)$ и $\mathbb{K}_{-n}(z) = \mathbb{K}_n(z)$. Аналогично $\mathbb{I}_{-n}(z) = \mathbb{I}_n(z)$. Поэтому можно ограничиться неотрицательными n .

Учитывая асимптотические формулы при малых z ¹⁾

$$K_0(z) \approx \ln\left(\frac{1}{z}\right), \quad K_n(z) \approx \frac{(n-1)!}{2} \left(\frac{2}{z}\right)^n, \quad I_n(z) \approx \frac{1}{n!} \left(\frac{z}{2}\right)^n, \quad (2.12)$$

нетрудно получить, что $\mathbb{K}_n(0) = \mathbb{I}_n(0) = n$, $n \geq 0$. Для определителя Вронского справедлива формула

$$W(K_n, I_n) := -K'_n(z)I_n(z) + K_n(z)I'_n(z) = 1/z,$$

откуда следует, что $\mathbb{K}_n(r) + \mathbb{I}_n(r) = 1/K_n(r)I_n(r)$. Продифференцируем это тождество:

$$\mathbb{K}'_n(r) + \mathbb{I}'_n(r) = \frac{\mathbb{K}_n(r) - \mathbb{I}_n(r)}{rK_n(r)I_n(r)}.$$

Полагая здесь $a(r) := (rK_n(r)I_n(r))^{-1}$, $f(r) := 2\mathbb{K}'_n(r)$, для определения $y(r) := \mathbb{K}_n(r) - \mathbb{I}_n(r)$ получим задачу:

$$y'(r) + a(r)y(r) = f(r), \quad r > 0, \quad y(0) = 0.$$

¹⁾относительно свойств этих функций см., напр., [20, с. 247], [21, с. 374]. Далее z означает комплексную переменную, $w'(z) := dw(z)/dz$; $r \in \mathbb{R}_+$.

¹⁾ $f(z) \approx g(z)$, если $f(z) = g(z) + o(g(z))$.

Так как функции a и f являются неотрицательными (I_n и K_n — положительны, $\mathbb{K}_n(r)$ — строго монотонно возрастает по r согласно замечанию 2.5), то y также является неотрицательной, т. е. $\mathbb{I}_n(r) \leq \mathbb{K}_n(r)$.

Получим вторую оценку. Из рекуррентных соотношений

$$-K'_n(z) = K_{n-1}(z) + \frac{n}{z}K_n(z), \quad -K'_n(z) = K_{n+1}(z) - \frac{n}{z}K_n(z)$$

следуют другие формулы для \mathbb{K}_n :

$$\mathbb{K}_n(z) = n + z \frac{K_{n-1}(z)}{K_n(z)} = -n + z \frac{K_{n+1}(z)}{K_n(z)}. \quad (2.13)$$

Поскольку $0 < K_{n-1}(r) < K_n(r)$ при $n \geq 1$ и $r > 0$, то

$$n \leq \mathbb{K}_n(r) \leq n + r, \quad \mathbb{K}_n(r) \geq r - n, \quad n \geq 1, \quad r \in \mathbb{R}_+.$$

Аналогичные рассуждения приводят к оценкам

$$\mathbb{I}_n(r) = n + r \frac{I_{n+1}(r)}{I_n(r)} \geq n, \quad \mathbb{I}_n(r) = -n + r \frac{I_{n-1}(r)}{I_n(r)} \geq r - n,$$

т. к. $I_n(r) \leq I_{n-1}(r)$ при $n \geq 1$. Следовательно,

$$\mathbb{K}_n(r) \leq n + r \leq 2n + \mathbb{I}_n(r) \leq 3\mathbb{I}_n(r), \quad n = 1, 2, \dots$$

Нетрудно убедиться, что функция

$$C_0(r) := \frac{\mathbb{K}_0(r)}{\mathbb{I}_0(r)} = \frac{K_1(r)}{K_0(r)} \frac{I_0(r)}{I_1(r)}$$

строго убывает на $(0, \infty)$ (каждая дробь убывает по r) и $C_0(r) \rightarrow +\infty$ при $r \rightarrow +0$, $C_0(r) \rightarrow 1$ при $r \rightarrow +\infty$. Полагая $c_0^2(r) = \max\{C_0(r), 3\}$, получим искомую оценку $\mathbb{K}_n(r) \leq c_0^2(r) \mathbb{I}_n(r)$, $n \geq 0$, $r \in (0, \infty)$. \square

ЗАМЕЧАНИЕ 2.6. Функции $\mathbb{K}_n(r)/\mathbb{I}_n(r)$ для ряда значений n изображены на правом рисунке 2. При больших значениях r они монотонно стремятся к единице, сохраняя порядок. Из графиков следует, что $c_0^2(r) = C_0(r)$; $C_0(r) \approx 2/(\ln(1/r)r^2)$ при малых r .

Следствие 2.2. Нормы $\|\cdot\|_{1/2,\Gamma}$ и $|u|_{1/2,\Gamma}$ эквивалентны; точнее

$$\|u\|_{1/2,\Gamma} \leq |u|_{1/2,\Gamma} \leq c_0 \|u\|_{1/2,\Gamma}, \quad u \in H^{1/2}(\Gamma).$$

Здесь постоянная $c_0 := c_0(R_0)$ зависит лишь от R_0 ¹⁾.

¹⁾напомним, что Ω есть круг радиуса $R \geq R_0$, $\Gamma := \partial\Omega$.

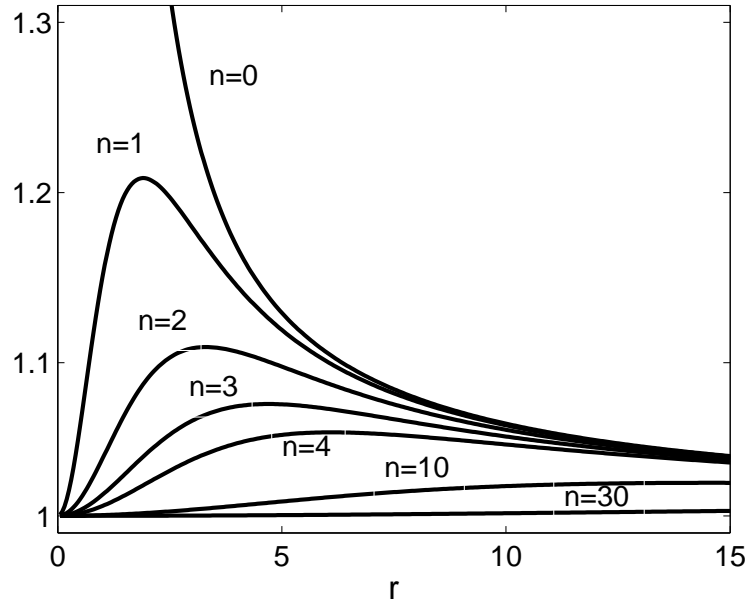


Рис. 3. Графики функций $(n^2 + r^2)^{1/2}/\mathbb{I}_n(r)$ для ряда значений n .

ДОКАЗАТЕЛЬСТВО. Требуемые оценки непосредственно вытекают из лемм 2.1, 2.2, если учесть, что $c_0(R) \leq c_0(R_0)$. \square

Следствие 2.3. Для любой функции $u \in H^1(\Omega)$

$$|u|_{1/2,\Gamma} \leq c_0 \|u\|_{1,\Omega}, \quad c_0 := c_0(R_0).$$

ДОКАЗАТЕЛЬСТВО. Из определения следует, что $\|u\|_{1/2,\Gamma} \leq \|u\|_{1,\Omega}$. Поэтому $|u|_{1/2,\Gamma} \leq c_0 \|u\|_{1/2,\Gamma} \leq c_0 \|u\|_{1,\Omega}$. \square

ЗАМЕЧАНИЕ 2.7. Норма $\|\cdot\|_{1/2,\Gamma}$, введенная выше, имеет два важных свойства, не зависящих от радиуса R области Ω :

$$\|u\|_{1/2,\Gamma} \leq \|u\|_{1,\Omega}, \quad u \in H^1(\Omega), \quad \|u\|_{1/2,\Gamma} \leq \|u\|_{1,\Omega_\infty}, \quad u \in H^1(\Omega_\infty).$$

Вторая оценка следует из неравенств $\|u\|_{1/2,\Gamma} \leq |u|_{1/2,\Gamma} \leq \|u\|_{1,\Omega_\infty}$. Учитывая следующие оценки, на доказательстве которых мы не будем останавливаться (см. рис. 3):

$$1 \leq (n^2 + r^2)^{1/2}/\mathbb{I}_n(r) \leq C^2(r), \quad r \in \mathbb{R}_+, \quad n = 0, \pm 1, \pm 2, \dots$$

придем к третьей эквивалентной нормировке пространства $H^{1/2}(\Gamma)$:

$$\|u\|_{1/2,\Gamma}^2 := \sum_{n=-\infty}^{\infty} (n^2 + R^2)^{1/2} |a_n(u)|^2. \quad (2.14)$$

Имеем $\|u\|_{1/2,\Gamma} \leq \|u\|_{1/2,\Gamma} \leq C(r) \|u\|_{1/2,\Gamma}$. Здесь $C^2(r) := r/\mathbb{I}_0(r)$ — строго монотонно убывающая функция (см. кривую $n = 0$ на рис. 3), $C^2(r) \approx 2/r$ при малых r . Ясно, что можно принять $R = 1$ в (2.14) (при этом изменятся лишь постоянные эквивалентности). Норма (2.14) наиболее часто используется для нормировки пространства $H^{1/2}(\Gamma)$.

1.3. Свойства функций \mathbb{K}_n .

Учитывая важность для нас функций $r \rightarrow \mathbb{K}_n(r)$, зафиксируем их свойства, отмеченные в ходе доказательства леммы 2.2, а также укажем некоторые новые.

Лемма 2.3. *При каждом целом n функция $z \rightarrow \mathbb{K}_n(z)$ является аналитической¹⁾ в правой полуплоскости $\mathbb{C}_+ := \{r + iy : r > 0\}$ и непрерывной в $\overline{\mathbb{C}_+} := \{r + iy : r \geq 0\}$. Кроме того*

$$\mathbb{K}_n(\bar{z}) = \overline{\mathbb{K}_n(z)}, \quad \mathbb{K}_n = \mathbb{K}_{-n}, \quad \mathbb{K}_n(0) = |n|; \quad (2.15)$$

$$|\tilde{\mathbb{K}}_n(z)/z| \leq 1, \quad |\mathbb{K}'_n(z)| \leq 2(|n| + |z|), \quad z \in \mathbb{C}_+, \quad n \neq 0, \quad (2.16)$$

где $\tilde{\mathbb{K}}_n(z) = \mathbb{K}_n(z) - \mathbb{K}_n(0)$. При малых z справедливы формулы

$$\tilde{\mathbb{K}}_0(z) \approx \ln^{-1}\left(\frac{1}{z}\right), \quad \tilde{\mathbb{K}}_{\pm 1}(z) = \frac{z^2}{\tilde{\mathbb{K}}_0(z)}, \quad \tilde{\mathbb{K}}_n(z) \approx \frac{z^2}{2(|n| - 1)}. \quad (2.17)$$

При $r \in \mathbb{R}_+$ имеем: $|n| \leq \mathbb{K}_n(r) \leq |n| + r$, $0 \leq \mathbb{K}'_n(r) \leq 2|n|$, если $n \neq 0$; $\mathbb{K}_0(0) = 0$, $\mathbb{K}'_0(r) > 0$ при $r > 0$.

ДОКАЗАТЕЛЬСТВО. Известно, что при целых n модифицированная функция Бесселя K_n является аналитической на z -плоскости с разрезом вдоль отрицательной вещественной оси, причем в $\overline{\mathbb{C}_+}$ она не имеет нулей (см., напр., [21, с. 377]). Поскольку

$$\mathbb{K}_n(z) := -zK'_n(z)/K_n(z) = \mathbb{K}_n(0) + zK_{n-1}(z)/K_n(z),$$

то отсюда следует, что \mathbb{K}_n является аналитической в \mathbb{C}_+ и непрерывной в $\overline{\mathbb{C}_+} \setminus \{0\}$. Непрерывность в нуле следует из асимптотик (2.17), которые легко вывести из асимптотик в нуле функций K_n (см. (2.12)). Равенства (2.15), как и оценки $|n| \leq \mathbb{K}_n(r) \leq |n| + r$ при $n \neq 0$, были установлены в лемме 2.2.

Докажем первую оценку в (2.16). Функция $\phi(z) := \tilde{\mathbb{K}}_n(z)/z$ равна $K_{n-1}(z)/K_n(z)$ (см. (2.13)) и является аналитической в \mathbb{C}_+ . Поэтому максимум ее модуля достигается либо на мнимой оси, либо в бесконечно удаленной точке. Поскольку $\mathbb{K}_n(\bar{z}) = \overline{\mathbb{K}_n(z)}$, то достаточно ограничиться рассмотрением мнимой полуоси $\{0 + iy, y \geq 0\}$. В бесконечно удаленной точке $|K_{n-1}(z)/K_n(z)| = 1$, поскольку при $|z| \rightarrow \infty$

$$K_n(z) = (2/\pi z)^{1/2} e^{-z} (1 + O(z^{-1})).$$

¹⁾ дифференцируемой в смысле комплексного анализа

Пусть $z = iy$, $y \geq 0$. По определению $K_n(z)$ имеем в этом случае:

$$K_{n-1}(z)/K_n(z) = i \overline{H_{n-1}^{(1)}(y)/H_n^{(1)}(y)},$$

где $H_n^{(1)}(y) = J_n(y) + iY_n(y)$ — функция Ханкеля порядка n . Поэтому

$$|\phi(z)| := |\tilde{\mathbb{K}}_n(z)/z| = \frac{J_{n-1}^2(y) + Y_{n-1}^2(y)}{J_n^2(y) + Y_n^2(y)} \leq 1,$$

так как согласно формуле Николсона [22, с. 444]

$$J_n^2(y) + Y_n^2(y) = \frac{8}{\pi^2 y} \int_0^\infty K_0(2y \operatorname{sh}(t)) \operatorname{ch}(2nt) dt \nearrow n.$$

Докажем вторую оценку в (2.16). Дифференцируя функцию $\mathbb{K}_n := -zK'_n/K_n$ и используя равенство $(zK'_n)' = (z^2 + n^2)K_n/z$, легко вычислить, что

$$\mathbb{K}'_n(z) = (\mathbb{K}_n^2(z) - z^2 - n^2)/z.$$

Отсюда при вещественных $z = r$ получим $\mathbb{K}'_n(r) \leq 2|n|$ при $n \neq 0$, поскольку $\mathbb{K}_n(r) \leq |n| + r$. При $z \in \mathbb{C}_+$ и $n \neq 0$ имеем $\mathbb{K}_n(z) =: |n| + z\phi(z)$, где $|\phi(z)| \leq 1$. Поэтому

$$\mathbb{K}'_n(z) = 2|n|\phi(z) + (\phi^2(z) - 1)z.$$

Отсюда следует требуемая оценка: $|\mathbb{K}'_n(z)| \leq 2(|n| + |z|)$. \square

Графики функций $r \rightarrow \mathbb{K}_n(r) - \mathbb{K}_n(0)$ для ряда значений n приведены на рис. 2. Отметим качественное различие свойств кривых при $n = 0$ и $n \neq 0$. Оно, в основном, объясняется формулами (2.17).

1.4. Метагармоническое продолжение функции.

Положим

$$H_0^1(\Omega_\infty) := \{v \in H^1(\Omega_\infty) : v|_\Gamma = 0\},$$

произвольно фиксируем вещественное p , $p > 0$, и определим функции

$$1_x := \min\{1, x^2\}, \quad 1^x := \max\{1, x^2\}, \quad x \in \mathbb{R}_+.$$

Отметим, что $1_x \rightarrow 0$ при $x \rightarrow 0$.

Определение 2.3. Функцию u_p назовем метагармонической в области $D \subset \mathbb{R}^2$, если $-\Delta u_p + p^2 u_p = 0$ в D .

Определение 2.4. Функцию $u_p \in H^1(\mathbb{R}^2)$ назовем метагармоническим продолжением функции $u \in H^1(\Omega)$, если u_p является метагармонической в области Ω_∞ и $u_p|_\Omega = u$.

Корректность этого определения вытекает из следующей леммы.

Лемма 2.4. *Метагармоническое продолжение u_p произвольной функции $u \in H^1(\Omega)$ существует, единственно и*

$$\|u_p\|_{1,\mathbb{R}^2} \leq c(p)\|u\|_{1,\Omega},$$

где c — локально ограниченная функция на $(0, \infty)$. Кроме того, симметричная неотрицательная форма

$$s_\infty(p, u, v) := \int_{\Omega_\infty} (\nabla u_p \cdot \nabla v_p + p^2 u_p v_p) dx, \quad u, v \in H^1(\Omega),$$

совпадает с формой (2.8) и является ограниченной на $H^1(\Omega)$:

$$0 \leq s_\infty(p, u, u) \leq M_s(p)\|u\|_{1,\Omega}^2, \quad M_s(p) := c_0^2(R_0)1^p. \quad (2.18)$$

ДОКАЗАТЕЛЬСТВО. Определим на $H^1(\Omega_\infty)$ функционал

$$|u|_{1,p}^2 := \int_{\Omega_\infty} (|\nabla u|^2 + p^2 u^2) dx.$$

По определению $-\Delta u_p + p^2 u_p = 0$ в Ω_∞ , $u_p|_\Gamma = u$. Поэтому $u_p|_{\Omega_\infty}$ является решением задачи минимизации

$$|u_p|_{1,p}^2 = \min_{v \in H^1(\Omega_\infty): v|_\Gamma = u} |v|_{1,p}^2.$$

Хорошо известно, что решение этой задачи существует и единственно, поскольку квадратичный функционал $u \rightarrow |u|_{1,p}^2$ является ограниченным и коэрцитивным:

$$1_p \|u\|_{1,\Omega_\infty}^2 \leq |u|_{1,p}^2 \leq 1^p \|u\|_{1,\Omega_\infty}^2.$$

Функционал $|u_p|_{1,p}^2$ нетрудно вычислить в терминах коэффициентов Фурье функции $u|_{\Gamma}$. Рассуждая также как при доказательстве леммы 2.1, получим

$$s_{\infty}(p, u, u) := |u_p|_{1,p}^2 = 2\pi \sum_{n=-\infty}^{\infty} \mathbb{K}_n(pR) |a_n(u)|^2 \quad \forall u \in V.$$

Положим здесь $u := u + tv$, $t \in \mathbb{R}$, $v \in V$, и вычислим производную по t от обеих частей при $t = 0$. В результате придем к формуле (2.8).

Принимая во внимание следствие 2.3, теперь имеем:

$$\begin{aligned} s_{\infty}(p, u, u) := |u_p|_{1,p}^2 &\leq 1^p \min_{v \in H^1(\Omega_{\infty}): v|_{\Gamma}=u} \|v\|_{1,\Omega_{\infty}}^2 =: \\ &=: 1^p |u|_{1/2,\Gamma}^2 \leq c_0^2(R_0) 1^p \|u\|_{1,\Omega}^2. \end{aligned}$$

Следовательно,

$$\begin{aligned} \|u_p\|_{1,\mathbb{R}^2}^2 &= \|u\|_{1,\Omega}^2 + \|u_p\|_{1,\Omega_{\infty}}^2 \leq \|u\|_{1,\Omega}^2 + s_{\infty}(p, u, u)/1_p \leq \\ &\leq (1 + M_s(p)/1_p) \|u\|_{1,\Omega}^2 =: c^2(p) \|u\|_{1,\Omega}^2. \quad \square \end{aligned}$$

Следствие 2.4. *Справедливо тождество*

$$\int_{\Omega_{\infty}} (\nabla u_p \cdot \nabla v + p^2 u_p v) dx = 0 \quad \forall v \in H_0^1(\Omega_{\infty}).$$

Лемма 2.5. *Пусть ограниченная функция u является гармонической или метагармонической вне замкнутого круга B_{R_0} , $a_n(u|_{\partial B_{R_0}})$ обозначает коэффициент Фурье функции $u|_{\partial B_{R_0}} \in L_2(\partial B_{R_0})$. Тогда при всех целых n и $R \geq R_0$ справедливы оценки*

$$|a_n(u|_{\partial B_R})| \leq (R_0/R)^{|n|} |a_n(u|_{\partial B_{R_0}})|.$$

ДОКАЗАТЕЛЬСТВО. Рассмотрим случай метагармонической функции u . Воспользуемся представлением (2.6). Оно справедливо в $\mathbb{R}^2 \setminus B_R$ при любом $R \geq R_0$. Поэтому вне B_{R_0} имеем:

$$u = \sum_{n=-\infty}^{\infty} \frac{K_n(pr)}{K_n(pR_0)} a_n(u|_{\partial B_{R_0}}) e^{in\varphi}. \quad (2.19)$$

Пусть $\varrho_n(p) := K_n(Rp)/K_n(pR_0)$. Имеем $\varrho_n(p) = \varrho_{-n}(p) > 0$. Используя асимптотики (2.12) функций Бесселя, нетрудно вычислить, что $\varrho_n(0) = (R_0/R)^{|n|}$. Учтем, что $\varrho'_n(p) = (\mathbb{K}_n(R_0p) - \mathbb{K}_n(Rp))\varrho_n(p)/p \leq 0$, поскольку функция $r \rightarrow \mathbb{K}_n(r)$ — возрастающая. Поэтому $0 < \varrho_n(p) \leq \varrho_n(0) = (R_0/R)^{|n|}$ при всех n .

Рассмотрим теперь равенство (2.19) в точках $r = R$ и заметим, что ряд, в силу равенства Парсеваля (см. (2.9)), равномерно сходится на $[0, 2\pi)$. Умножим обе его части на $e^{-ik\varphi}/2\pi$ и проинтегрируем по $\varphi \in [0, 2\pi)$. В результате получим $a_k(u|_{\partial B_R}) = \varrho_k(p)a_k(u|_{\partial B_{R_0}})$ при всех k . Отсюда следует искомая оценка: $|a_k(u|_{\partial B_R})| \leq (R_0/R)^{|k|}|a_k(u|_{\partial B_{R_0}})|$.

Случай гармонической функции u рассматривается аналогично. В этом случае в формуле (2.19) множитель $K_n(pr)/K_n(pR_0)$ достаточно заменить на $(R_0/r)^{|n|}$. \square

1.5. Второй метод.

Приведем прямое доказательство эквивалентности задач (\mathcal{P}_∞) и (\mathcal{P}) .

Теорема 2.9. *Задачи (\mathcal{P}_∞) и (\mathcal{P}) эквивалентны в следующем смысле. Если (β, k, u) — решение задачи (\mathcal{P}_∞) , то $(\beta, p, u|_\Omega)$ — решение задачи (\mathcal{P}) при $p = (\beta^2 - \varepsilon_\infty k^2)^{1/2}$. Обратно: если (β, p, u) — решение задачи (\mathcal{P}) , то (β, k, u_p) — решение задачи (\mathcal{P}_∞) , где $k = ((\beta^2 - p^2)/\varepsilon_\infty)^{1/2}$, u_p — метагармоническое продолжение u .*

ДОКАЗАТЕЛЬСТВО. Пусть (β, k, u) удовлетворяет тождеству (\mathcal{P}_∞) . Определим число p согласно (2.4). Тогда $(\beta, p) \in K$, поскольку $(\beta, k) \in \Lambda$. Запишем тождество (\mathcal{P}_∞) в следующем виде:

$$\int_{\Omega} (\nabla u \cdot \nabla v + \beta^2 uv) dx + \int_{\Omega_\infty} (\nabla u \cdot \nabla v + p^2 uv) dx = k^2 \int_{\Omega} \varepsilon uv dx. \quad (2.20)$$

Здесь v произвольная функция из $H^1(\mathbb{R}^2)$. Заметим, что если принять $v = 0$ в области $\bar{\Omega}$, то получим тождество

$$\int_{\Omega_\infty} (\nabla u \cdot \nabla v + p^2 uv) dx = 0 \quad \forall v \in H_0^1(\Omega_\infty).$$

Следовательно, u является метагармонической в Ω_∞ (см. следствие 2.4) и $u = u_p$. Ограничимся в (2.20) только такими $v \in H^1(\mathbb{R}^2)$, что

$v = v_p$, и учтем определение формы s_∞ . Тогда получим, что

$$\int_{\Omega} (\nabla u \cdot \nabla v + \beta^2 uv) dx + s_\infty(p, u, v) = k^2 \int_{\Omega} \varepsilon uv dx \quad \forall v \in H^1(\Omega). \quad (2.21)$$

Так как $p^2 = \beta^2 - \varepsilon_\infty k^2$, то

$$k^2 \int_{\Omega} \varepsilon uv dx = -p^2 \int_{\Omega} \sigma uv dx + \beta^2 \int_{\Omega} \sigma uv dx.$$

Поэтому тождество (2.21) преобразуется к виду

$$\int_{\Omega} (\nabla u \cdot \nabla v + p^2 \sigma uv) dx + s_\infty(p, u, v) = \beta^2 \int_{\Omega} (\sigma - 1) uv dx.$$

Следовательно, $(\beta, p, u|_{\Omega})$ есть решение задачи (\mathcal{P}) .

Докажем обратное утверждение. Пусть (β, p, u) — решение (\mathcal{P}) . Определим $k = ((\beta^2 - p^2)/\varepsilon_\infty)^{1/2}$. Так как $(\beta, p) \in K$, то $(\beta, k) \in \Lambda$. Пусть u_p и w_p — метагармонические продолжения u и произвольного $w \in V$ соответственно. Учитывая определение формы s_∞ , имеем

$$\begin{aligned} \int_{\Omega} (\nabla u_p \cdot \nabla w + \beta^2 u_p w) dx + \int_{\Omega_\infty} (\nabla u_p \cdot \nabla w_p + p^2 u_p w_p) dx = \\ = k^2 \int_{\Omega} \varepsilon u_p w dx. \end{aligned} \quad (2.22)$$

Пусть v — произвольная функция из $H^1(\mathbb{R}^2)$ такая, что $v|_{\Omega} = w$. В Ω_∞ представим ее в виде $v = w_p + \eta$. Ясно, что $\eta \in H_0^1(\Omega_\infty)$ и

$$\int_{\Omega_\infty} (\nabla u_p \cdot \nabla \eta + p^2 u_p \eta) dx = 0.$$

Учтем это равенство в (2.22). Тогда получим

$$\int_{\Omega} (\nabla u_p \cdot \nabla v + \beta^2 u_p v) dx + \int_{\Omega_\infty} (\nabla u_p \cdot \nabla v + p^2 u_p v) dx = k^2 \int_{\Omega} \varepsilon u_p v dx.$$

С учетом равенства $p^2 = \beta^2 - \varepsilon_\infty k^2$ оно преобразуется к виду

$$\int_{\mathbb{R}^2} (\nabla u_p \cdot \nabla v + \beta^2 u_p v) dx = k^2 \int_{\mathbb{R}^2} \varepsilon u_p v dx.$$

Отсюда следует, что (β, k, u_p) — решение задачи (\mathcal{P}_∞) . \square

ЗАМЕЧАНИЕ 2.8. По определению задача (\mathcal{P}) зависит от выбора области Ω , точнее, зависит от параметра R — радиуса Ω . Поэтому правильнее было бы писать (\mathcal{P}_R) , Ω_R , Γ_R и т.д. Чтобы не перегружать обозначения, мы эти зависимости от R будем подразумевать и не будем указывать явно. Это соглашение дополнительно оправдывается тем, что в тройке (β, p, u) , определяющей решение задачи (\mathcal{P}) , от R зависит только u , причем тривиальным образом: u есть сужение на область Ω функции, определенной в \mathbb{R}^2 , которая не зависит от R . С теоретической точки зрения выбор R не важен при условии, что $\Omega \supseteq \Omega_i$. С практической точки зрения величина R имеет важное значение и его выбор будет влиять на эффективность численных методов, которые будут обсуждаться далее.

ЗАМЕЧАНИЕ 2.9. По-видимому, D. Givoli был первым, кто предложил использовать метод разделения переменных для получения точного нелокального краевого условия на фиктивной границе при решении краевых задач в неограниченных областях [23] [24]. Оператор S_Γ был им назван DtN оператором (Dirichlet-to-Neumann), поскольку он преобразует, фактически, краевые условия Дирихле на фиктивной границе в условия Неймана. Мы выбрали конечную (расчетную) область Ω в форме круга исходя из соображений простоты изложения и проводимых выкладок. Для рассматриваемых нами задач оператор S_Γ нетрудно определить и для эллиптической области Ω (см., напр., [25]). Такой выбор оправдан на практике, если волновод Ω_i является сильно вытянутым в одном направлении.

1.6. Пространства функций.

Далее нам понадобятся следующие пространства функций, определенных на области $D \subseteq \mathbb{R}^2$. Все интегралы ниже понимаются в смысле Лебега.

$C^k(D)$ — пространство k -раз непрерывно дифференцируемых на D функций, $k \geq 0$; $C(D) = C^0(D)$;

$L_p(D)$ — пространство измеримых по Лебегу функций, для которых конечна норма

$$\|u\|_{0,p,D} = \left(\int_D |u|^p dx \right)^{1/p}, \quad 1 \leq p < \infty;$$

$L_\infty(D)$ — пространство измеримых по Лебегу существенно ограниченных функций, с нормой

$$\|u\|_{0,\infty,D} = \inf \{ M : |u| \leq M \text{ п.в. в } \Omega \};$$

$W_p^k(D)$ — пространство Соболева, $1 \leq p \leq \infty$, k — целое число,

$$W_p^k(D) = \{ u : D^\alpha u \in L_p(D), |\alpha| \leq k \}.$$

Здесь D^α обозначает обобщенную производную порядка $\alpha = (\alpha_1, \alpha_2)$,

$$D^\alpha = \frac{\partial^{|\alpha|}}{\partial x_1^{\alpha_1} \partial x_2^{\alpha_2}}, \quad |\alpha| = \alpha_1 + \alpha_2, \quad \alpha_1, \alpha_2 \geq 0, \quad D^{(0,0)}u = u.$$

При $1 \leq p < \infty$ пространство $W_p^k(D)$ оснащается нормой

$$\|u\|_{k,p,D} = \left(\sum_{|\alpha| \leq k} \int_D |D^\alpha u|^p dx \right)^{1/p},$$

для полунорм используются обозначения

$$|u|_{s,p,D} = \left(\sum_{|\alpha|=s} \int_D |D^\alpha u|^p dx \right)^{1/p}, \quad 0 \leq s \leq k.$$

При $p = 2$ используются сокращения: $H^k(D) = W_2^k(D)$,

$$\|u\|_{k,D} = \|u\|_{k,2,D}, \quad |u|_{s,D} = |u|_{s,2,D}.$$

Определим, наконец, пространство $W_\infty^k(D)$ — пространство $(k-1)$ -раз непрерывно-дифференцируемых в \bar{D} функций, производные $k-1$ порядка которых удовлетворяют в \bar{D} условию Липшица. Норма в нем

$$\|u\|_{k,\infty,D} = \sum_{|\alpha| \leq k} |D^\alpha u|_{0,\infty,D}.$$

Пространства Соболева подробно описаны в книгах [26], [27].

§ 2. Существование и свойства решений

Задачу (\mathcal{P}) (см. замечание 2.4) далее будем рассматривать как параметрическую задачу на собственные значения: *при каждом заданном $p > 0$ найти такие $(\beta, u) \in \mathbb{R}_+ \times V \setminus \{0\}$, что*

$$(\mathcal{P}) \quad a(p, u, v) = \beta^2 b(u, v) \quad \forall v \in V.$$

Формы a и b зависят от функции $\sigma := \varepsilon/\varepsilon_\infty$. В данном параграфе будут использованы лишь следующие свойства этой функции:

$$1 \leq \sigma \leq \sigma_+ := \varepsilon_+/\varepsilon_\infty \quad \text{в } \Omega; \quad \sigma|_{\Omega \setminus \Omega_i} = 1; \quad C(\bar{\Omega}_i) \ni \sigma|_{\bar{\Omega}_i} \geq 1 + \sigma_0.$$

Здесь $\sigma_0 > 0$ ¹⁾. Предположения о гладкости σ в Ω_i , а также границы $\partial\Omega_i$, понадобятся нам в следующих параграфах.

Укажем необходимые в дальнейшем свойства форм a и b . Отметим, что пространство $V := H^1(\Omega)$ является гильбертовым. Скалярное произведение в нем определяется равенством

$$(u, v) := \int_{\Omega} (\nabla u \cdot \nabla v + uv) dx, \quad u, v \in V.$$

2.1. Свойства форм a и b .

Наиболее просто устроена форма b ,

$$b(u, v) := \int_{\Omega} (\sigma - 1)uv dx.$$

Лемма 2.6. *Форма $b(\cdot, \cdot)$ является симметричной,*

$$0 \leq b(u, u) \leq M_B \|u\|_{1,\Omega}^2, \quad u \in V.$$

Порождаемый ею оператор B является самосопряженным, компактным и неотрицательным¹⁾, $\ker B = \{u \in V : u = 0 \text{ на } \Omega_i\}$.

ДОКАЗАТЕЛЬСТВО. Поскольку $\sigma = 1$ вне Ω_i , то

$$0 \leq b(u, u) = \int_{\Omega_i} (\sigma - 1)u^2 dx \leq M_B \|u\|_{0,\Omega}^2 \leq M_B \|u\|_{1,\Omega}^2, \quad (2.23)$$

где $M_B := \sigma_+ - 1$. Из симметричности и ограниченности формы b следует самосопряженность B ; компактность B есть следствие первой оценки сверху в (2.23) и компактности вложения $H^1(\Omega) \subset L_2(\Omega)$. Отметим, что ядро и образ B являются бесконечномерными. \square

Рассмотрим форму a . Она была определена нами при $p > 0$ следующим образом:

$$\begin{aligned} a(p, u, v) &:= a_0(p, u, v) + s_{\infty}(p, u, v) := \\ &:= \int_{\Omega} (\nabla u \cdot \nabla v + p^2 \sigma uv) dx + 2\pi \sum_{n=-\infty}^{\infty} \mathbb{K}_n(pR) a_n(u) \overline{a_n(v)}. \end{aligned} \quad (2.24)$$

¹⁾ см. соответствующее замечание на с. 30.

¹⁾ оператор B в V определяется по теореме Рисса равенством $(Bu, v) = b(u, v)$, $u, v \in V$.

Далее без дополнительных ссылок будем использовать свойства функций \mathbb{K}_n , указанные в лемме 2.3. Поскольку $\mathbb{K}_n(pR) \rightarrow |n|$ при $p \rightarrow 0$ при всех n , то по определению примем

$$a(0, u, v) := \int_{\Omega} \nabla u \cdot \nabla v \, dx + 2\pi \sum_{n=-\infty}^{\infty} |n| a_n(u) \overline{a_n(v)}. \quad (2.25)$$

В силу этого, далее будем считать, что при всех $p \in \mathbb{R}_+$ форма a имеет определение (2.24).

Лемма 2.7. *При каждом $p \in \mathbb{R}_+$ форма a является симметричной, неотрицательной и ограниченной,*

$$m_A(p) \|u\|_{1,\Omega}^2 \leq a(p, u, u) \leq M_A(p) \|u\|_{1,\Omega}^2, \quad u \in V, \quad (2.26)$$

где $m_A(p) := 1_p$, $M_A(p) := (\sigma_+ + c_0^2(R_0)) 1^p$.

Если $(A(p)u, v) = a(p, u, v)$ для любых $u, v \in V$, то $A(p)$ является самосопряженным и неотрицательным оператором; $\ker A(0) = \{u : u = \text{const в } \Omega\}$; $A(p) + B \geq m_{AB}I$, где постоянная $m_{AB} > 0$ зависит от Ω и ε .

ДОКАЗАТЕЛЬСТВО. Симметричность и неотрицательность формы a являются очевидными. Напомним, что при $p > 0$ (см. (2.18))

$$0 \leq s_{\infty}(p, u, u) \leq M_s(p) \|u\|_{1,\Omega}^2, \quad M_s(p) := c_0^2(R_0) 1^p. \quad (2.27)$$

Ясно, что эти оценки справедливы и при $p = 0$, т. к. $\mathbb{K}_n(0) \leq \mathbb{K}_n(Rp)$. Учитывая их, из (2.24) легко вывести оценки (2.26) (напомним, что $1 \leq \sigma \leq \sigma_+$). Из них следует, что $A(p)$ является самосопряженным и неотрицательным. Отметим, что $m_A(0) = 0$, $m_A(p) > 0$ при $p > 0$.

Утверждение о ядре $A(0)$ становится очевидным, если учесть, что при $u = \text{const в } \Omega$, имеем $u|_{\Gamma} = \text{const}$; $a_n(u) = 0$ при $n \neq 0$ (коэффициент $a_0(u) \neq 0$ отсутствует в определении $a(0, u, u)$). Наконец,

$$a(p, u, u) + b(u, u) \geq \int_{\Omega} |\nabla u|^2 \, dx + \int_{\Omega} (\sigma - 1) u^2 \, dx =: |u|_{1,\Omega}^2 + \rho^2(u). \quad (2.28)$$

Отметим, что ρ есть непрерывная полунорма на $L_2(\Omega)$ и она не обращается в нуль на постоянных в Ω функциях. Хорошо известно, что в этом случае справедлива оценка (см., напр., [28, с. 25])

$$\int_{\Omega} u^2 \, dx \leq c_{\Omega} (|u|_{1,\Omega}^2 + \rho^2(u)),$$

где постоянная c_Ω зависит от области Ω . Прибавляя к обеим частям этой оценки $|u|_{1,\Omega}^2$, получим: $|u|_{1,\Omega}^2 + \rho^2(u) \geq m_{AB} \|u\|_{1,\Omega}^2$, а из (2.28) — заключительное утверждение леммы. \square

Лемма 2.8. *Функция $p \rightarrow a(p, u, u)$ является дифференцируемой на $(0, \infty)$, непрерывной в нуле и строго монотонно возрастающей при каждом фиксированном $u \in V$.*

ДОКАЗАТЕЛЬСТВО. Пусть фиксировано $u \in V$. Убедимся, что ряд, определяющий $s_\infty(p, u, u)$, можно почленно дифференцировать по p на любом отрезке $\omega := [\omega_1, \omega_2]$ из $(0, \infty)$. Для этого достаточно убедиться, что формальная производная ряда

$$s'(p) := 2\pi R \sum_{n=-\infty}^{\infty} \mathbb{K}'_n(pR) |a_n(u)|^2$$

сходится равномерно в ω . Напомним, что все функции $r \rightarrow \mathbb{K}_n(r)$ являются строго монотонно возрастающими, непрерывными в нуле и аналитическими на $(0, \infty)$ ¹⁾, $0 < \mathbb{K}'_n(r) \leq |n| + r$ при $r > 0$ и $n \neq 0$. Пусть $\mu_\omega := \max\{\mathbb{K}'_0(pR), pR : p \in \omega\}$. Нетрудно видеть, что ряд $s'(p)$ мажорируется сходящимся числовым рядом

$$\begin{aligned} 2\pi R \sum_{n=-\infty}^{\infty} (|n| + \mu_\omega) |a_n(u)|^2 &= R s_\infty(0, u, u) + \mu_\omega 2\pi R \sum_{n=-\infty}^{\infty} |a_n(u)|^2 = \\ &= R s_\infty(0, u, u) + \mu_\omega \int_{\Gamma} u^2 d\Gamma \leq D_\omega \|u\|_{1,\Omega}^2. \end{aligned}$$

Здесь мы воспользовались оценкой (2.27) при $p = 0$, равенством Парсеваля (см. (2.9)) и непрерывностью вложения $H^1(\Omega) \subset L_2(\Gamma)$. Следовательно, ряд $s'(p)$ сходится равномерно на ω (признак Вейерштрасса). Поскольку $s'(p) \geq 0$, то при $p \in \omega$ и $u \neq 0$ имеем:

$$0 < \frac{d}{dp} a(p, u, u) = 2p \int_{\Omega} \sigma u^2 dx + s'(p) \leq D_A(p) \|u\|_{1,\Omega}^2,$$

где $D_A(p) := 2p\sigma_+ + D_\omega$. Отсюда следует также, что $a(p, u, u)$ строго монотонно возрастает по p при фиксированном u .

¹⁾вещественная функция аналитична в области ω , если в каждой точке ω она раскладывается в равномерно сходящийся ряд Тейлора.

Пусть теперь $\epsilon(p) := \max\{\mathbb{K}_0(pR)/R, p\}$. Поскольку при всех n имеем $0 \leq \mathbb{K}_n(pR) - |n| \leq R \epsilon(p)$, то

$$\begin{aligned} 0 \leq a(p, u, u) - a(0, u, u) &= p^2 \int_{\Omega} \sigma u^2 dx + 2\pi \sum_{n=-\infty}^{\infty} (\mathbb{K}_n(pR) - |n|) |a_n(u)|^2 \\ &\leq p^2 \sigma_+ \|u\|_{0,\Omega}^2 + \epsilon(p) 2\pi R \sum_{n=-\infty}^{\infty} |a_n(u)|^2 = \\ &= p^2 \sigma_+ \|u\|_{0,\Omega}^2 + \epsilon(p) \|u\|_{0,\Gamma}^2 \leq (p^2 \sigma_+ + c_{\Omega} \epsilon(p)) \|u\|_{1,\Omega}^2, \end{aligned} \quad (2.29)$$

где c_{Ω} — постоянная вложения $H^1(\Omega) \subset L_2(\Gamma)$. Отсюда следует, что a является непрерывной в нуле, так как $\epsilon(p) \rightarrow 0$ при $p \rightarrow 0$. \square

Следствие 2.5. *Оператор $A_1(p) := A(p) - A(0)$ является неотрицательным и компактным при $p > 0$, $\|A_1(p)\| \rightarrow 0$ при $p \rightarrow 0$.*

ДОКАЗАТЕЛЬСТВО. Так как $(A_1(p)u, u) := a(p, u, u) - a(0, u, u)$, то из (2.29) имеем:

$$0 \leq (A_1(p)u, u) \leq p^2 \sigma_+ \|u\|_{0,\Omega}^2 + \epsilon(p) \|u\|_{0,\Gamma}^2.$$

Отсюда следует компактность оператора $A_1(p)$, так как вложения $H^1(\Omega) \subset L_2(\Omega)$ и $H^1(\Omega) \subset L_2(\Gamma)$ являются компактными. Ясно, что

$$\|A_1(p)\| = \sup_{u \in V, u \neq 0} \frac{|(A_1(p)u, u)|}{\|u\|_{1,\Omega}^2} \leq p^2 \sigma_+ + c_{\Omega} \epsilon(p) \rightarrow 0, \quad p \rightarrow 0.$$

2.2. Существование и свойства решений задачи (P).

Напомним операторную формулировку задачи: при каждом заданном $p > 0$ найти такие $(\beta, u) \in \mathbb{R}_+ \times V \setminus \{0\}$, что $A(p)u = \beta^2 Bu$. Рассмотрим также задачу

$$A(0)u = \beta^2 Bu, \quad (2.30)$$

собственные числа которой определяют критические числа.

Задачи подобного вида были изучены нами в главе 1 (см. § 3, с. 19), где спектральный параметр β^2 обозначался через λ , а параметр

p пробегал \mathbb{R}_+ ¹⁾. Были доказаны две теоремы (1.4 и 1.5), посвященные соответственно существованию решений и свойствам дисперсионных кривых. Чтобы воспользоваться этими результатами, необходимо проверить пять условий (см. условия $A_1 - A_5$ на с. 21), которым должны удовлетворять операторы $A(p)$ и B при $p \in \mathbb{R}_+$ (предыдущий пункт, фактически, и был посвящен проверке этих условий).

Лемма 2.9. *Оператор-функция $p \rightarrow A(p)$ и оператор B удовлетворяют условиям $A_1 - A_5$.*

ДОКАЗАТЕЛЬСТВО. Напомним условия $A_1 - A_5$ (опуская части, касающиеся зависимости B от p), а также укажем леммы предыдущего пункта, посвященные их проверке.

В леммах 2.6 и 2.7 было доказано, что

$$m_A(p) \leq A(p) \leq M_A(p)I, \quad p \in \mathbb{R}_+, \quad 0 \leq B \leq M_B I, \\ r_0 := \dim(\ker A(0)) = 1, \quad \dim(\operatorname{Im} B(p)) = \infty,$$

где $m_A(0) = 0$, $m_A(p) > 0$ при $p > 0$. Эти оценки означают, что выполнены как условия A_1 , так и условия A_2 .

В лемме 2.7 при всех p из \mathbb{R}_+ была установлена оценка $A(p) + B(p) \geq m_{AB}I$, определяющая условие A_3 . Там же было доказано, что функция $p \rightarrow a(p, u, u)$ строго монотонно возрастает при любом $u \in V$, поэтому отношения Рэлея $R(p, u) := a(p, u, u)/b(u, u)$ строго монотонно возрастает по p на \tilde{V} ¹⁾ (условие A_4).

Непрерывность в нуле функции $p \rightarrow a(p, u, u)$ была доказана в лемме 2.8. Там же была установлена ее дифференцируемость на $(0, \infty)$. Такие функции локально-липиц непрерывны, поскольку

$$|a(p, u, u) - a(\bar{p}, u, u)| = \left| \int_{\bar{p}}^p \frac{d}{d\rho} a(\rho, u, u) d\rho \right| \leq \\ \leq \left| \int_{\bar{p}}^p L_\omega \|u\|_{1,\Omega}^2 d\rho \right| \leq L_\omega |p - \bar{p}| \|u\|_{1,\Omega}^2,$$

где $p, \bar{p} \in \omega := [\omega_1, \omega_2] \subset (0, \infty)$. Таким образом функция $p \rightarrow A(p)$, как и B , принадлежат классу $\mathcal{L}(\mathbb{R}_+, V)$ (условия A_5). \square

¹⁾рассматриваемая задача несколько проще, поскольку оператор B не зависит от p .

¹⁾ \tilde{V} есть ортогональное дополнение $\ker B$ до пространства $V_{A(p)+B}$.

Теорема 2.10. При каждом $p \in \mathbb{R}_+$ существует счетное множество чисел $\beta^K(p)$, $K \geq 1$, с единственной точкой накопления $+\infty$, квадраты которых образуют полный набор собственных чисел задачи (P) при $p > 0$ и задачи (2.30) при $p = 0$. Соответствующие им собственные подпространства $U^K(p)$ — конечномерны, $V_{A(p)+B} = \bigoplus_{k=1}^{\infty} U^K(p) \oplus \ker B$.²⁾

Кроме того, если $\{\beta_i(p)\}_{i=1}^{\infty}$ есть числа $\beta^K(p)$, $K = 1, 2, \dots$, пронумерованные по возрастанию с учетом кратности, то

a) функции $p \rightarrow \beta_i(p)$ и $p \rightarrow \beta_i^2(p) - p^2$, $i \geq 1$, являются строго монотонно возрастающими, непрерывными в нуле и локально липшиц-непрерывными на $(0, \infty)$;

b) $\beta_1(p) \searrow +0$ при $p \rightarrow +0$, $\beta_2(0) > 0$;

c) если $k_0 := (\varepsilon_+ / (\varepsilon_+ - \varepsilon_{\infty}))^{1/2}$, то $\frac{\beta_i(p)}{p} > k_0$, $\lim_{p \rightarrow \infty} \frac{\beta_i(p)}{p} = k_0$.

ДОКАЗАТЕЛЬСТВО. Из теорем 1.4 и 1.5 главы 1, которыми мы можем воспользоваться благодаря предыдущей лемме, следуют все утверждения теоремы, кроме утверждений о монотонности функций $p \rightarrow \beta_i^2(p) - p^2$ и c).

Пусть $p > 0$. Используем минимаксный принцип Куранта-Фишера. Имеем:

$$\beta_i^2(p) = \min_{V_i \subset \tilde{V}} \max_{v \in V_i \setminus \{0\}} R(p, v), \quad R(p, v) := \frac{a(p, v, v)}{b(v, v)},$$

Поэтому

$$\beta_i^2(p) - p^2 = \min_{V_i \subset \tilde{V}} \max_{v \in V_i \setminus \{0\}} \tilde{R}(p, v), \quad \tilde{R}(p, v) := R(p, v) - p^2. \quad (2.31)$$

Простые преобразования приводят к равенству

$$\tilde{R}(p, v) = \frac{1}{b(v, v)} \left(\int_{\Omega} (|\nabla v|^2 + p^2 v^2) dx + s_{\infty}(p, v, v) \right).$$

Теперь из (2.31) следует, что функция $p \rightarrow \beta_i^2(p) - p^2$ является неотрицательной и строго монотонно возрастающей.

²⁾кратность $\beta^K(p)$ определим как кратность $\beta^{K^2}(p) := (\beta^K(p))^2$. Соответствующие им собственные элементы будем называть также собственными функциями.

Докажем второе утверждение в с), поскольку первое следует из принадлежности $(\beta_i(p), p)$ множеству K . Обозначим через B_r круг малого радиуса r с центром в такой точке $x_+ \in \bar{\Omega}_i$, что

$$\varepsilon_+ = \varepsilon(x_+) = \max_{x \in \bar{\Omega}_i} \varepsilon(x), \quad \delta_r := \max_{x \in B_r} \frac{\varepsilon_+ - \varepsilon(x)}{\varepsilon_\infty},$$

и пусть V_r есть множество функций из V , равных нулю вне B_r . Отметим, что $s_\infty(p, u, u) = 0$ на V_r ; $\sigma(x) := \varepsilon(x)/\varepsilon_\infty \leq \varepsilon_+/\varepsilon_\infty$ в B_r ,

$$\sigma(x) - 1 = \frac{\varepsilon_+}{\varepsilon_\infty} - 1 - \frac{\varepsilon_+ - \varepsilon(x)}{\varepsilon_\infty} \geq \frac{\varepsilon_+ - \varepsilon_\infty - \varepsilon_\infty \delta_r}{\varepsilon_\infty} =: \frac{d_r}{\varepsilon_\infty}.$$

Нетрудно видеть, что для любого $v \in V_r$ справедлива оценка

$$R(p, v) \leq \frac{\varepsilon_\infty}{d_r} \left(\int_{B_r} |\nabla v|^2 dx \right) \left(\int_{B_r} v^2 dx \right)^{-1} + \frac{\varepsilon_+ p^2}{d_r}.$$

Обозначим через (λ_r^i, u^i) — собственные пары оператора Лапласа в круге B_r при краевых условиях Дирихле (u^i продолжим нулем вне B_r). Поскольку $\lambda_r^i = r^{-2} \lambda_1^i$, то

$$\beta_i^2(p) = \min_{V_i \subset \tilde{V}} \max_{v \in V_i \setminus \{0\}} R(p, v) \leq \frac{\varepsilon_\infty \lambda_1^i}{r^2 d_r} + \frac{\varepsilon_+ p^2}{d_r}.$$

Следовательно,

$$k_0^2 \leq \frac{\beta_i^2(p)}{p^2} \leq \frac{\varepsilon_\infty \lambda_1^i}{p^2 r^2 d_r} + \frac{\varepsilon_+}{d_r}.$$

Переходя здесь к пределу по $r \rightarrow 0$, $p \rightarrow \infty$ так, чтобы $pr \rightarrow \infty$, получим второе утверждение с), так как $\delta_r \rightarrow 0$, $\varepsilon_+/d_r \rightarrow k_0^2$. \square

2.3. Аналитичность дисперсионных кривых.

Согласно следствию 2.5 оператор $A(p)$ допускает представление в виде суммы самосопряженных неотрицательных операторов, $A(p) =: A(0) + A_1(p)$, причем операторы $A_1(p)$ являются компактными и $\|A_1(p)\| \rightarrow 0$ при $p \rightarrow 0$. Оператор $A(0)$ порождается формой $a(0, \cdot, \cdot)$ (см. (2.25)),

$$(A_1(p)u, v) := p^2 \int_{\Omega} \sigma uv dx + 2\pi \sum_{n=-\infty}^{\infty} \tilde{\mathbb{K}}_n(pR) a_n(u) \overline{a_n(v)}, \quad (2.32)$$

где $\tilde{\mathbb{K}}_n(z) = \mathbb{K}_n(z) - \mathbb{K}_n(0)$. Задачу (\mathcal{P}) преобразуем к виду

$$\tilde{A}(p)u := A_0u + A_1(p)u = \lambda Bu, \quad A_0 := A(0) + B, \quad \lambda := \beta^2 + 1. \quad (2.33)$$

Далее покажем, что оператор $A_1(p)$, рассматриваемый как возмущение положительно определенного оператора A_0 , аналитически зависит от p . Тогда, привлекая аналитическую теорию возмущений линейных операторов, получим следующее утверждение.

Теорема 2.11. *Все собственные числа задачи (\mathcal{P}) представляются аналитическими на $(0, \infty)$ функциями. Точнее, существует последовательность скалярных функций $\lambda^i(p)$ и последовательность вектор-функций $u^i(p)$, $i = 1, 2, \dots$, аналитических на $(0, \infty)$ и таких, что для каждой точки $p \in (0, \infty)$ последовательность $\lambda^i(p)$ представляет полный набор собственных чисел задачи (\mathcal{P}) , а последовательность $u^i(p)$ образует полное в $V/\ker B(0)$ семейство собственных элементов.*

ДОКАЗАТЕЛЬСТВО. Рассмотрим эквивалентную формулировку (2.33) задачи (\mathcal{P}) . Задачи такого типа при более общих предположениях относительно операторов задачи изучались в монографии [29, с. 520]. Поскольку оператор $\tilde{A}(p)$ положительно определен при всех p , то определен самосопряженный и компактный оператор $T(p) := \tilde{A}^{-1/2}(p)B\tilde{A}^{-1/2}(p)$. Он также аналитически зависит от p на $(0, \infty)$ [29, с. 523, 524]. Ясно, что

$$T(p)v = \mu(p)v, \quad \lambda(p) =: \mu^{-1}(p) - 1, \quad u =: \tilde{A}^{-1/2}(p)v.$$

Нас интересуют только ненулевые собственные числа семейства $T(p)$. Так как $\ker T(p) = \ker B(0)$ при всех p , то они могут быть получены как решения задачи

$$\tilde{T}(p)v = \mu(p)v, \quad v \in \tilde{V} \setminus \{0\}, \quad \tilde{V} := V/\ker B(0), \quad (2.34)$$

где $\tilde{T}(p)$ есть сужение $T(p)$ на \tilde{V} . Известно, что все собственные числа и соответствующие им собственные элементы этой задачи представляются аналитическими на $(0, \infty)$ функциями ([29, см. теорему 3.9 и замечание 3.11 на с. 490]), причем собственные элементы образуют базис в \tilde{V} . Отсюда следуют утверждения теоремы, если учесть связь между исходной задачей и (2.34). \square

ЗАМЕЧАНИЕ 2.10. Разницу между двумя нумерациями $\{\lambda_i(p)\}$ и $\{\lambda^i(p)\}$ дисперсионных кривых демонстрирует рис. 1, с. 24.

ЗАМЕЧАНИЕ 2.11. Согласно теореме 2.10 аналитические функции $\lambda^i(p)$ определены по непрерывности при $p = 0$. Аналогичное утверждение, на доказательстве которого мы не останавливаемся, справедливо и для собственных элементов $u^i(p)$.

Докажем аналитичность семейства операторов $A_1(p)$. Для этого достаточно доказать следующее утверждение.

Лемма 2.10. *Функция $p \rightarrow \sum_{n=-\infty}^{\infty} \tilde{\mathbb{K}}_n(pR) a_n(u) \overline{a_n(v)}$ при фиксированных $u, v \in V$ является аналитической на $(0, \infty)$. Радиус сходимости его ряда Тейлора в точке $p_0 > 0$ равен p_0 .*

ДОКАЗАТЕЛЬСТВО. Пусть фиксированы $u, v \in V$, $c_n := a_n(u) \overline{a_n(v)}$, $\phi_n(z) := \tilde{\mathbb{K}}_n(z)/z$ ($z \in \mathbb{C}_+ := \{p + iy : p > 0\}$). Определим ряды

$$s_0 := \sum_{n=-\infty}^{\infty} |c_n|, \quad s(z) := \sum_{n=-\infty}^{\infty} \tilde{\mathbb{K}}_n(Rz) c_n.$$

Из равенства Парсеваля (см. (2.9), с. 33) и вложения $V \subset L_2(\Gamma)$ следует, что числовой ряд s_0 сходится. Действительно,

$$s_0^2 \leq \sum_{n=-\infty}^{\infty} |a_n(u)|^2 \sum_{n=-\infty}^{\infty} |a_n(v)|^2 \leq c \|u\|_{1,\Omega}^2 \|v\|_{1,\Omega}^2.$$

Ряд $s(z)$ преобразуется к виду

$$s(z) = c_0 \mathbb{K}_0(Rz) + Rz \sum_{n=-\infty, n \neq 0}^{\infty} c_n \phi_n(Rz) =: c_0 \mathbb{K}_0(Rz) + Rz s_1(z).$$

Функциональный ряд s_1 мажорируется рядом s_0 , так как $|\phi_n(z)| \leq 1$ в \mathbb{C}_+ при $n \neq 0$. Поэтому s_1 сходится равномерно в \mathbb{C}_+ . Поскольку при каждом n функции $\phi_n(z)$, а также $\mathbb{K}_0(z)$, являются аналитическими в односвязной области \mathbb{C}_+ , то по теореме Вейерштрасса $s_1(z)$, а следовательно и $s(z)$, являются аналитическими в \mathbb{C}_+ . Отметим, что поскольку аналитичность s теряется в нуле, то радиус сходимости ряда Тейлора

$$s(z) = \sum_{k=0}^{\infty} s^{(k)}(z_0, u, v) (z - z_0)^k, \quad s^{(k)}(z_0, u, v) := \frac{s^{(k)}(z_0)}{k!},$$

при вещественном положительном $z_0 = p_0$, равен в точности p_0 . Коэффициенты этого ряда представляют собой симметричные, вещественные и ограниченные билинейные формы на V и могут быть получены разложением в ряд Тейлора в точке z_0 функций $\tilde{\mathbb{K}}_n(Rz)$ в определении $s(z)$ и группировкой слагаемых. \square

2.4. Множество решений задачи (\mathcal{P}_∞) .

Напомним, что задача (\mathcal{P}_∞) является обобщенной формулировкой задачи нахождения пар чисел $(\beta, k) \in \Lambda$ и ненулевых вещественных $u \in H^1(\mathbb{R}^2)$, удовлетворяющих п.вс. в \mathbb{R}^2 уравнению

$$-\Delta u + \beta^2 u = k^2 \varepsilon u,$$

где продольное волновое число k пропорционально частоте электромагнитных колебаний ω , $k^2 := \varepsilon_0 \mu_0 \omega^2$.

Между множествами решений этой задачи и задачи (\mathcal{P}) имеется взаимно-однозначная связь, определяемая теоремой 2.9: если $(\beta_i(p), p, u_i(p))$ есть решение задачи (\mathcal{P}) , $i \geq 1$, то $(\beta_i(p), k_i(p), u_{ip}(p))$ — решение задачи (\mathcal{P}_∞) при $k_i(p) = (\beta_i^2(p) - p^2)/\varepsilon_\infty)^{1/2}$ и наоборот¹⁾. Таким образом, из теоремы 2.10, посвященной описанию множества решений задачи (\mathcal{P}) , нетрудно получить информацию о множестве и свойствах решений задачи (\mathcal{P}_∞) . Дадим, предварительно, несколько пояснений.

В задаче (\mathcal{P}_∞) будем считать волновое число k (или ω) параметром, а неизвестными (β, u) и будем интересоваться зависимостью решений от k . Пусть $(\beta_i(p), u_i(p))$, $i \geq 1$, $p \in (0, \infty)$, есть решения задачи (\mathcal{P}) и $k_i(p) = (\beta_i^2(p) - p^2)/\varepsilon_\infty)^{1/2}$. Тогда решениями задачи (\mathcal{P}_∞) являются тройки $(\beta_i(p), k_i(p), u_{ip}(p))$, $i \geq 1$, $p \in (0, \infty)$. В плоскости (k, β) кривые

$$k = k_i(p), \quad \beta = \beta_i(p), \quad p \in (0, \infty), \quad i = 1, 2, \dots$$

определяют дисперсионные кривые задачи (\mathcal{P}_∞) , параметризованные поперечным волновым числом p . Согласно утверждению а) теоремы 2.10, функции $p \rightarrow k_i(p)$, $p \rightarrow \beta_i(p)$ локально удовлетворяют условию Липшица, непрерывны в нуле и строго монотонно возрастают.

¹⁾здесь $u_{ip}(p)$ есть метагармоническое продолжение $u_i(p)|_\Omega$.

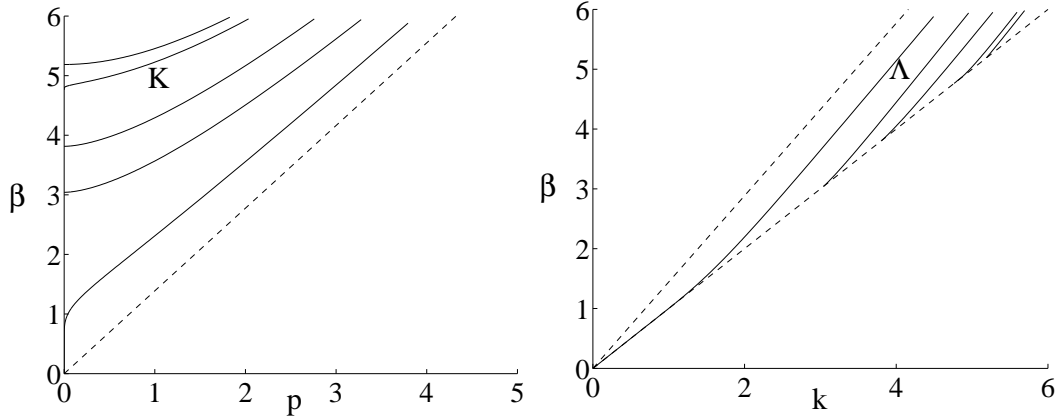


Рис. 4. Дисперсионные кривые для волновода прямоугольного поперечного сечения с размерами 1.5×1 ; $\varepsilon(x) = 2.0$, $x \in \Omega_i$; $\varepsilon_\infty = 1$. На левом рисунке пунктирной линией изображена граница области K , на правом — граница Λ .

При $p \rightarrow 0$ имеем $k_i(p) \rightarrow k_i^0 = \beta_i^0 / \varepsilon_\infty^{1/2}$, $\beta_i(p) \rightarrow \beta_i^0$. Таким образом, i -тая дисперсионная кривая начинается в критической точке (k_i^0, β_i^0) (число $\omega_i^0 := k_i^0 / (\varepsilon_0 \mu_0)^{1/2}$ называется критической частотой; см. рис. 4, где представлены первые дисперсионные кривые задачи (\mathcal{P}) (слева) и (\mathcal{P}_∞) (справа) для однородного волновода прямоугольного поперечного сечения, полученные численно). Ясно, что дисперсионные кривые допускают параметризацию

$$\beta_i = \beta_i(k), \quad k \in (k_i^0, \infty), \quad i = 1, 2, \dots^1$$

Также ясно, что числа β_i^{02} являются решениями следующей задачи: найти $(\beta^2, u) \in \mathbb{R}_+ \times V \setminus \{0\}$, удовлетворяющие тождеству

$$(\mathcal{P}^0) : \quad a(0, u, v) = \beta^2 b(u, v) \quad \forall v \in V.$$

Напомним, что

$$a(0, u, v) := \int_{\Omega} \nabla u \cdot \nabla v \, dx + 2\pi \sum_{n=-\infty}^{\infty} |n| a_n(u) \overline{a_n(v)}.$$

Задачу (\mathcal{P}^0) назовем задачей определения критических чисел. Согласно теореме 2.10

$$0 = \beta_1^0 < \beta_2^0 \leq \beta_3^0 \leq \dots, \quad \beta_i^0 \rightarrow \infty, \quad i \rightarrow \infty.$$

¹⁾впрочем, как и параметризацию $k_i = k_i(\beta)$, $\beta \in (\beta_i^0, \infty)$, $i = 1, 2, \dots$

Определим ступенчатую функцию

$$n(k) := \max\{i : k_i^0 < k, i = 1, 2, \dots\}.$$

Ясно, что $n(k)$ в точности равно числу решений $(\beta_i(k), u_i(k))$ задачи (\mathcal{P}_∞) при фиксированном $k > 0$. Из приведенных выше рассуждений и теоремы 2.10 непосредственно следует

Теорема 2.12. *При каждом $k > 0$ существует конечное число собственных чисел $\beta^K(k)$, $1 \leq K \leq m(k)$, суммарной кратности $n(k)$, и соответствующих им собственных подпространств $U^K(k)$ таких, что $(\beta^K(k), k, u)$ есть решение задачи (\mathcal{P}_∞) при любом $u \in U^K(k)$. Кроме того, если*

$$0 < \beta_1(k) \leq \beta_2(k) \leq \dots \leq \beta_{n(k)}(k),$$

есть нумерация чисел $\{\beta^K(k)\}_{K=1}^{m(k)}$ с учетом кратности, то

- a) $n(k) \rightarrow \infty$ при $k \rightarrow \infty$; $\beta_1(k) \rightarrow 0$ при $k \rightarrow +0$;
- b) функции $k \rightarrow \beta_i(k)$, $k \geq k_i^0$, — строго монотонно возрастают и локально липшицевы, $i = 1, \dots, n(k)$;
- c) $\varepsilon_\infty^{1/2} k < \beta_i(k) < \varepsilon_+^{1/2} k$, $k > k_i^0$, $\lim_{k \rightarrow \infty} (\beta_i(k)/k) = \varepsilon_+^{1/2}$, $i = 1, \dots, n(k)$.

При каждом $k > 0$, как следует из теоремы, существует по крайней мере одно решение задачи (\mathcal{P}_∞) . Число решений неограниченно возрастает с ростом k .

ЗАМЕЧАНИЕ 2.12. Обратим внимание, что качественное поведение основной дисперсионной кривой $p \rightarrow \beta_1(p)$ на рис. 4 отличается от поведения других кривых. В частности, она вогнута и экспоненциально приближается к оси ординат при малых p .

Проанализируем поведение $\beta_1(p)$ при малых p . Оператор $A(p)$ имеет вид $A(p) = A(0) + A_1(p)$, где компактный оператор $A_1(p)$ определяется формулой (2.32). При малых p функции $\tilde{\mathbb{K}}_n(Rp)$ имеют асимптотику (2.17). Пусть $\varepsilon := p^2 \ln(1/p)$, $\varepsilon_0 := \mathbb{K}_0(pR) \approx 1/\ln(1/p)$. Определим операторы

$$(A_0 u, v) := a(0, u, v) + \varepsilon_0 2\pi a_0(u) \overline{a_0(v)},$$

$$(E(p)u, v) := p^2 \int_{\Omega} \sigma uv dx + 2\pi \sum_{n=-\infty, n \neq 0}^{\infty} \tilde{\mathbb{K}}_n(pR) a_n(u) \overline{a_n(v)}.$$

Тогда задача (\mathcal{P}) примет вид

$$A_0 u + E(p)u = \beta^2(p)Bu.$$

Используя равенство Парсеваля легко показывается, что $\|E(p)\| \leq c\varepsilon$. Поэтому при малых p задача на собственные значения

$$A_0 u = \lambda Bu,$$

является хорошим приближением задачи (\mathcal{P}) , поскольку $0 \leq \beta^2(p) - \lambda \leq c\epsilon$. Эта задача зависит от малого параметра ϵ_0 . Пусть (λ_0, u_0) его собственные пары при нулевом значении ϵ_0 . Используя теорию асимптотических разложений собственных пар (см., напр., [29, теорема 2.6, с. 551]) имеем разложение $\lambda = \lambda_0 + \lambda_1\epsilon_0 + O(\epsilon_0^2)$. Для простого собственного числа $\lambda_1 = (A_0 u_0, u_0)/(B u_0, u_0)$. Окончательно получим:

$$\beta^2(p) = \lambda_0 + \lambda_1\epsilon_0 + O(\epsilon_0^2) + O(\epsilon) = \lambda_0 + \lambda_1\mathbb{K}_0(pR) + O(\mathbb{K}_0^2(pR)).$$

Рассмотрим первую собственную пару $(\beta_1^2(p), u_1(p))$. В этом случае $\lambda_0 = 0$, $u_0 = c$ — постоянная в Ω функция (он принадлежит ядру A_0 при $\epsilon_0 = 0$). Поэтому

$$\beta_1^2(p) = \lambda_1 \mathbb{K}_0(Rp) + O(\mathbb{K}_0^2(pR)), \quad \lambda_1 := 2\pi / \int_{\Omega} (\sigma - 1) dx.$$

Таким образом, при малых p поведение функции $p \rightarrow \beta_1(p)$ совпадает с поведением функции $\mathbb{K}_0^{1/2}(Rp)$ (график функции $\mathbb{K}_0(p)$ см. на левом рис. 2, с. 37).

§ 3. Дискретная задача

Для приближенного решения задачи (\mathcal{P}) используем метод Галеркина с возмущениями. Абстрактная оценка точности этого метода была получена в первой главе (см. теорему 1.8 на с. 29). Для дискретизации задачи используем метод конечных элементов (МКЭ) с численным интегрированием, основанный на конформной аппроксимации V_h пространства Соболева $V := H^1(\Omega)$. Наряду с численным интегрированием, дополнительное возмущение в метод Галеркина вносит усечение бесконечной суммы, входящей в определение формы a .

3.1. Пространство конечных элементов.

Опишем кратко конструкцию пространства конечных элементов V_h (подробнее см. в [30, с. 47], либо [31], [32], либо [33, с. 114]). Пусть $\{\mathcal{T}_h\}_h$ — семейство точных регулярных триангуляций круга Ω , зависящее от малого параметра $h > 0$. А именно, \mathcal{T}_h есть такая совокупность треугольных конечных элементов e , что¹⁾

$$\bar{\Omega}_h := \bigcup_{e \in \mathcal{T}_h} e = \bar{\Omega}, \quad h := \max_{e \in \mathcal{T}_h} \text{diam}(e).$$

¹⁾общими у элементов могут быть либо сторона, либо вершина; e — замкнутое множество.

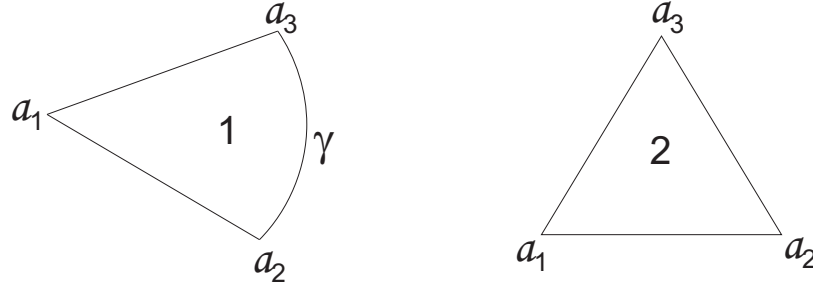


Рис. 5. Приграничные (1) и внутренние (2) треугольники, γ есть часть $\partial\Omega$ или $\partial\Omega_i$.

Допускается использование элементов двух типов — элементов с тремя прямолинейными сторонами (прямолинейные элементы) и элементов с двумя прямолинейными и одной криволинейной стороной (криволинейные элементы). С практической точки зрения естественно использовать криволинейные элементы только в качестве приграничных, то есть когда две вершины треугольника e лежат на $\partial\Omega$ (окружности), или на $\partial\Omega_i$ (границе волновода) (см. рис. 5). Для упрощения изложения будем считать, что $\partial\Omega_i$ является липшицевой, кусочно-гладкой кривой класса C^∞ .²⁾

Рассмотрим элемент $e \in \mathcal{T}_h$ с вершинами a_1 , a_2 и a_3 (либо криволинейный, либо прямолинейный). Будем считать, что две его стороны a_1a_2 и a_1a_3 — прямолинейные, а третья лежит на кривой ℓ , параметризованной дуговой координатой s так, что $a_2 = \chi(s_2)$, $a_3 = \chi(s_3)$ ($l_e = s_3 - s_2$ — длина дуги a_2a_3)¹⁾.

Пусть \hat{e} — базисный треугольник в плоскости $(\hat{x}_1\hat{x}_2)$ с вершинами $\hat{a}_1 = (0, 0)$, $\hat{a}_2 = (1, 0)$, $\hat{a}_3 = (0, 1)$. Определим отображение

$$x = x_e(\hat{x}) := a_1 + B_e \hat{x} + \hat{x}_1 \frac{\Phi(\hat{x}_2) - \Phi_I(\hat{x}_2)}{1 - \hat{x}_2}.$$

Здесь $\Phi(t) := \chi(s_2 + t(s_3 - s_2))$; $\Phi_I(t) := a_2 + t(a_3 - a_2)$, $0 \leq t \leq 1$, — параметрическое представление хорды a_2a_3 ;

$$B_e := \begin{pmatrix} (a_2 - a_1)_1 & (a_3 - a_1)_1 \\ (a_2 - a_1)_2 & (a_3 - a_1)_2 \end{pmatrix}.$$

Отображение $x = x_e(\hat{x})$ задает преобразование \hat{e} на e , сохраняющее ориентацию, причем $a_i = x_e(\hat{a}_i)$. Известно, что это взаимно однознач-

²⁾Напомним, что $\Omega = B_R \supset \Omega_i$, Ω_i — поперечное сечение волновода.

¹⁾ ℓ является либо кривой линией, либо — прямой, $\ell = \{x \in \mathbb{R}^2 : x = \chi(s), s \in [0, L]\}$.

ное диффеоморфное отображение, если h — достаточно мало. Точнее, если

$$c^{-1}h^2 \leq \det B_e \leq ch^2, \quad \|B_e\| \leq ch, \quad \|B_e^{-1}\| \leq ch, \quad (2.35)$$

то

$$c^{-1}h^2 \leq \det(Dx_e(\hat{x})) \leq ch^2 \quad \forall \hat{x} \in \hat{e}, \quad (2.36)$$

$$\|\det(Dx_e)\|_{j,\infty,\hat{e}} \leq ch^{j+2}, \quad \|x_e\|_{j,\infty,\hat{e}} \leq ch^j, \quad j = 1, 2, \dots, \quad (2.37)$$

где Dx_e — матрица Якоби отображения x_e . Здесь и далее буква c (возможно с индексом) используется для обозначения различных постоянных, не зависящих от h .

Обозначим через $\hat{x} = x_e^{-1}(x)$ обратное отображение к x_e , через \hat{P}_m — множество полиномов степени не выше m на \hat{e} , $m \geq 0$,

$$\hat{P}_m := \left\{ \sum_{0 \leq \alpha_1 + \alpha_2 \leq m} c_\alpha \hat{x}_1^{\alpha_1} \hat{x}_2^{\alpha_2}, c_\alpha \in \mathbb{R} \right\}, \quad \dim \hat{P}_m =: \hat{M} := \frac{(m+1)(m+2)}{2}.$$

Введем следующее пространство функций на e :

$$P_m^e := \{p : p(x) = \hat{p}(x_e^{-1}(x)), \hat{p} \in \hat{P}_m, x \in e\}, \quad m \geq 1.$$

Если $\hat{\varphi}_i(\hat{x})$, $i = 1, \dots, \hat{M}$, — базис Лагранжа в \hat{P}_m , $m \geq 1$, связанный с узлами интерполяции $\hat{\omega} := \{\hat{a}_i, i = 1, \dots, \hat{M}\}$ ¹⁾, то базис Лагранжа в P_m^e составляют функции $\varphi_i^e(x) = \hat{\varphi}_i(x_e^{-1}(x))$, $i = 1, \dots, \hat{M}$. Они связаны с узлами интерполяции $\omega^e := \{a_i^e := x_e(\hat{a}_i), \hat{a}_i \in \hat{\omega}\}$ на элементе e .

ЗАМЕЧАНИЕ 2.13. Если ℓ — отрезок прямой (то есть все стороны e прямолинейные), то $x_e(\hat{x}) = a_1 + B_e \hat{x}$ — аффинное отображение, а P_m^e — пространство полиномов степени не выше m . Заметим также, что для криволинейных элементов e функции из P_m^e не являются полиномами. Тем не менее, сужение произвольной функции из P_m^e на любую сторону e (в том числе криволинейную) является полиномом степени не выше m относительно дуговой координаты этой стороны (см., напр., [33, с. 118]).

Триангуляция \mathcal{T}_h называется регулярной, если выполнены условия (2.36) и (2.37) для любого элемента $e \in \mathcal{T}_h$. Далее будем предполагать, что длины сторон элементов из \mathcal{T}_h имеют порядок h , а углы треугольников отделены от нуля и π равномерно по h . Этими условиями

¹⁾то есть $\hat{\varphi}_i \in \hat{P}_m$, $\hat{\varphi}_i(\hat{a}_j) = \delta_{ij}$ для $i, j = 1, \dots, \hat{M}$. Предполагается, что на каждой стороне \hat{e} выбрано $m+1$ равномерно расположенных узла интерполяции, включая вершины элемента.

достаточно для выполнения (2.35), то есть для обеспечения регулярности триангуляции.

Аппроксимация V_h пространства V определяется теперь следующим образом:

$$V_h := \{v_h \in C(\bar{\Omega}) : v_h|_e \in P_m^e \forall e \in \mathcal{T}_h\}.$$

Это пространство конечномерно и базис Лагранжа в нем определяется обычным способом: если $\omega_h := \{a_i, i = 1, \dots, N_h\}$ — сетка узлов на $\bar{\Omega}$, образованная различными точками из $\bigcup\{\omega^e, e \in \mathcal{T}_h\}$, то узлу $a_l \in \omega_h$ ставится в соответствие базисная функция $\varphi_l \in V_h$ такая, что $\varphi_l(a_j) = \delta_{lj}$, $l, j = 1, \dots, N_h$. По определению имеем:

$$u_h(x) = \sum_{i=1}^{N_h} c_i \varphi_i(x), \quad x \in \bar{\Omega}, \quad c_i = u_h(a_i), \quad u_h \in V_h.$$

Введем обозначения: $\hat{u}(\hat{x}) := u(x_e(\hat{x}))$ — образ функции $u(x)$, определенной на e , при преобразовании координат $x = x_e(\hat{x})$, $\hat{x} \in \hat{e}$;

$$|u|_{k,h}^2 := \sum_{e \in \mathcal{T}_h} |u|_{k,e}^2, \quad \|u\|_{k,h}^2 := \sum_{e \in \mathcal{T}_h} \|u\|_{k,e}^2, \quad \|u\|_{k,\infty,h} := \max_{e \in \mathcal{T}_h} \|u\|_{k,\infty,e}.$$

Регулярность триангуляции \mathcal{T}_h обеспечивает справедливость следующих оценок:

$$|\hat{u}|_{s,p,\hat{e}} \leq c h^{s-2/p} \|u\|_{s,p,e}, \quad |u|_{s,p,e} \leq c h^{2/p-s} |\hat{u}|_{s,p,\hat{e}}, \quad (2.38)$$

$$\|v_h\|_{k,h} \leq c h^{l-k} \|v_h\|_{l,h}, \quad 0 \leq l < k, \quad v_h \in V_h, \quad (2.39)$$

$$\inf_{v_h \in V_h} \|u - v_h\|_{1,\Omega} \rightarrow 0 \quad \text{при } h \rightarrow 0, \quad u \in V, \quad (2.40)$$

$$\inf_{v_h \in V_h} \|u - v_h\|_{1,\Omega} \leq c h^k \|u\|_{k+1,h}, \quad k = 0, \dots, m. \quad (2.41)$$

Здесь $s \geq 0$, $p \in [1, \infty]$; $\|\cdot\|_{s,p,e}$, $|\cdot|_{s,p,e}$ — норма и полунорма в $W_p^s(e)$. Оценка (2.39) называется в теории МКЭ обратным неравенством.

3.2. Формулы численного интегрирования.

Для приближенного вычисления интегралов по области Ω используем составные квадратурные формулы. Пусть $m \geq 1$ то же, что и в определении V_h , а \hat{S} — некоторая заданная квадратурная формула

на элементе \hat{e} с положительными коэффициентами, точная на полиномах из \hat{P}_{2m-1}^1):

$$\int_{\hat{e}} \varphi(\hat{x}) d\hat{x} \approx \sum_{i=1}^Q \hat{c}_i \varphi(\hat{b}_i) =: \hat{S}(\varphi), \quad \hat{c}_i > 0.$$

Примеры таких квадратур хорошо известны в литературе (см., напр., [33, с. 192]). Например, при $m = 1$ подходящей является формула с одним узлом (центр тяжести \hat{e}) или с тремя узлами (вершины \hat{e}).

Пусть $J_e(\hat{x}) := Dx_e(\hat{x})$ — матрица Якоби преобразования x_e , $|J_e(\hat{x})| := \det J_e(\hat{x}) > 0$ — его якобиан. Тогда следующие формулы определяют искомую составную квадратуру $S_h(\cdot)$:

$$\begin{aligned} \int_{\Omega} \varphi(x) dx &= \sum_{e \in \mathcal{T}_h} \int_e \varphi(x) dx = \sum_{e \in \mathcal{T}_h} \int_{\hat{e}} |J_e(\hat{x})| \varphi(x_e(\hat{x})) d\hat{x} \approx \\ &\approx \sum_{e \in \mathcal{T}_h} \sum_{i=1}^Q \hat{c}_i |J_e(\hat{b}_i)| \varphi(x_e(\hat{b}_i)) =: \sum_{e \in \mathcal{T}_h} S_e(\varphi) =: S_h(\varphi). \end{aligned}$$

Определим дискретные аналоги норм в $H^1(\Omega)$ и $L_2(\Omega)$:

$$\|u_h\|_{1, V_h}^2 := S_h(|\nabla u_h|^2 + u_h^2), \quad \|u_h\|_{0, V_h}^2 := S_h(u_h^2).$$

Теорема 2.13. Пусть триангуляция \mathcal{T}_h является регулярной, а квадратурная формула \hat{S} точна на полиномах из \hat{P}_{2m-2} и имеет положительные коэффициенты. Тогда имеют место оценки:

$$c^{-1} \|u_h\|_{1, \Omega} \leq \|u_h\|_{1, V_h} \leq c \|u_h\|_{1, \Omega}, \quad u_h \in V_h, \quad (2.42)$$

$$\|u_h\|_{0, V_h} \leq c \|u_h\|_{0, \Omega}, \quad u_h \in V_h. \quad (2.43)$$

ДОКАЗАТЕЛЬСТВО. Пусть $e \in \mathcal{T}_h$, $\hat{P}_m \ni \hat{u}(\hat{x}) := u_h(x_e(\hat{x}))$ — образ функции $u_h = u_h|_e$ при преобразовании координат $x = x_e(\hat{x})$. При $J_e^{-T} := (J_e^{-1})^T$ имеем $|\nabla u_h(x)| = |J_e^{-T}(\hat{x}) \hat{\nabla} \hat{u}(\hat{x})|$.¹⁾ Из условий регулярности триангуляции вытекают следующие равномерные по $\hat{x} \in \hat{e}$ оценки:

$$c^{-1} h^2 \leq |J_e(\hat{x})| \leq c h^2, \quad \|J_e(\hat{x})\| \leq c h, \quad \|J_e^{-1}(\hat{x})\| \leq c h^{-1}. \quad (2.44)$$

¹⁾Использование таких квадратур в МКЭ при решении краевых задач для эллиптических уравнений второго порядка приводит к оптимальным оценкам точности в нормах H^1 и L_2 .

¹⁾ $\hat{\nabla} = (\partial/\partial \hat{x}_1, \partial/\partial \hat{x}_2)^T$.

Здесь $\|A\|$ означает спектральную (вторую) норму матрицы A ($|Ay| \leq \|A\| |y|$, $y \in \mathbb{R}^2$; $\|A\| = \|A^T\|$). Из условий на \widehat{S} следует, что

$$\widehat{I}(\hat{u}) := \int_{\hat{e}} \left(|\widehat{\nabla} \hat{u}|^2 + \hat{u}^2 \right) d\hat{x} \sim \widehat{S} \left(|\widehat{\nabla} \hat{u}|^2 + \hat{u}^2 \right) =: \hat{s}(\hat{u}) \text{ на } \widehat{P}_m,^1$$

поскольку, как нетрудно видеть, как $\widehat{I}^{1/2}$, так и $\hat{s}^{1/2}$ есть нормы на \widehat{P}_m . Докажем, что

$$I(u_h) := \int_e (h^2 |\nabla u_h|^2 + u_h^2) dx \sim S_e (h^2 |\nabla u_h|^2 + u_h^2) =: S(u_h) \quad (2.45)$$

на P_m^e . Используя оценку $|y| \leq \|A\| |A^{-1}y|$, $y \in \mathbb{R}^2$, а также оценки (2.44), имеем:

$$\begin{aligned} I(u_h) &= \int_{\hat{e}} |J_e| \left(h^2 |J_e^{-T} \widehat{\nabla} \hat{u}|^2 + \hat{u}^2 \right) d\hat{x} \leq c h^2 \widehat{I}(\hat{u}) \leq c h^2 \hat{s}(\hat{u}) \leq \\ &\leq c \frac{h^2}{\min |J_e|} S_e (\|J^T\|^2 |J^{-T} \nabla u_h|^2 + u_h^2) \leq c S(u_h). \end{aligned}$$

Такие же рассуждения приводят к обратному неравенству $S(u_h) \leq c I(u_h)$, что доказывает (2.45). Аналогично доказывается, что также

$$\int_e |\nabla u_h|^2 dx \sim S_e (|\nabla u_h|^2) \text{ на } P_m^e. \quad (2.46)$$

Из (2.45) и (2.46) имеем:

$$\int_e ((1 + h^2) |\nabla u_h|^2 + u_h^2) dx \sim S_e ((1 + h^2) |\nabla u_h|^2 + u_h^2).$$

Отсюда, после суммирования по всем $e \in \mathcal{T}_h$, следуют оценки (2.42).

Функционал $\widehat{S}^{1/2}(\hat{u}^2)$ определяет полунорму на \widehat{P}_m , а $\left(\int_{\hat{e}} \hat{u}^2 d\hat{x} \right)^{1/2}$ — норму. Поэтому, как и ранее, имеем:

$$S_e(u_h^2) \leq c h^2 \widehat{S}(\hat{u}^2) \leq c h^2 \int_{\hat{e}} \hat{u}^2 d\hat{x} \leq c \frac{h^2}{\min |J_e|} \int_e u_h^2 dx \leq c \int_e u_h^2 dx.$$

¹⁾ $f(u) \sim g(u)$ на U , если $c_1 f(u) \leq g(u) \leq c_2 g(u)$ для $\forall u \in U$, c_1, c_2 не зависят от h и u , $c_1 > 0$.

Суммируя эти оценки по всем $e \in \mathcal{T}_h$, приходим к (2.43). \square

Следствие 2.6. $S_h(|\nabla u_h|^2) \sim |u_h|_{1,\Omega}^2$,

$$S_e(|\nabla u_h|^2 + u_h^2) \sim \int_e (|\nabla u_h|^2 + u_h^2) dx.$$

3.3. Дискретная задача. Свойства решений.

Определим билинейные формы a_{0h} , b_h , s_N , являющиеся аппроксимациями на $V_h \times V_h$ форм a_0 , b и s_∞ соответственно. Для этого заменим интегралы по области Ω в определении форм a_0 и b на составную квадратуру S_h , бесконечный ряд в определении s_∞ — конечным. В итоге получим:

$$a_{0h}(p, u, v) := S_h(\nabla u \cdot \nabla v + p^2 \sigma uv), \quad \sigma := \varepsilon/\varepsilon_\infty,$$

$$b_h(u, v) := S_h((\sigma - 1) uv),$$

$$s_N(p, u, v) := 2\pi \sum_{n=-N}^N \mathbb{K}_n(pR) a_n(u) \overline{a_n(v)}.$$

Здесь $u, v \in V_h$, $N \geq 0$. Предполагается, что коэффициенты Фурье функций u и v вычисляются точно. Последнее нетрудно выполнить аналитически, так как u и v на Γ являются кусочно-полиномиальными (степени m) функциями относительно угловой координаты φ .

Отметим, что при аппроксимации формы s_∞ возникает дополнительный параметр N . Далее будем считать, что $N = N(h)$, поэтому введем обозначение

$$a_h(p, u, v) := a_{0h}(p, u, v) + s_N(p, u, v).$$

Рассмотрим конечномерную задачу: *при каждом $p \in \mathbb{R}_+$ найти числа $\beta^h \in \mathbb{R}_+$ и ненулевые $u_h \in V_h$ такие, что*¹⁾

$$(\mathcal{P}_h) \quad a_h(p, u_h, v) = \beta^{h2} b_h(u_h, v) \quad \forall v \in V_h.$$

¹⁾числа β^h , как и β^{h2} , мы далее будем называть собственными числами.

При $p > 0$ эта задача является дискретным аналогом задачи (\mathcal{P}) , при $p = 0$ — задачи (\mathcal{P}^0) для определения критических чисел.

Пусть $u_h = \sum_{i=1}^{N_h} y_i \varphi_i$, $y_i = u_h(a_i)$; $y := (y_1, y_2, \dots, y_{N_h})^T$ — вектор узловых параметров u_h . Введем матрицы:

$$A_h(p) := \{a_h(p, \varphi_j, \varphi_i)\}_{i,j=1}^{N_h}, \quad B_h := \{b_h(\varphi_j, \varphi_i)\}_{i,j=1}^{N_h}.$$

Тогда задача (\mathcal{P}_h) может быть записана как параметрическая обобщенная алгебраическая задача на собственные значения:

$$A_h(p)y = \beta^{h2} B_h y, \quad y \in \mathbb{R}^{N_h} \setminus \{0\}.$$

Задачи подобного вида были изучены нами в главе 1 (см. § 4, с. 26), где спектральный параметр β^{h2} обозначался через λ^h . Чтобы воспользоваться результатами теоремы 1.7, посвященной существованию решений задачи и свойствам дисперсионных кривых, необходимо предварительно убедиться, что семейства матриц $A_h(p)$ и B_h удовлетворяют условиям $A_1^h - A_5^h$ на с. 27.

Лемма 2.11. Семейства матриц $A_h(p)$, $p \in \mathbb{R}_+$ и B_h удовлетворяют условиям $A_1^h - A_5^h$.

ДОКАЗАТЕЛЬСТВО. Учтем, что $\mathbb{K}_n(r) \geq 0$ при $r \in \mathbb{R}_+$. Привлекая лемму 2.4, имеем:

$$0 \leq s_N(p, u_h, u_h) \leq s_\infty(p, u_h, u_h) \leq M_s(p) \|u_h\|_{1,\Omega}^2.$$

Учитывая (2.42), также имеем ($1_p := \min\{1, p^2\}$, $1^p := \max\{1, p^2\}$):

$$a_{0h}(p, u_h, u_h) \geq 1_p S_h(|\nabla u_h|^2 + u_h^2) \geq 1_p c^{-2} \|u_h\|_{1,\Omega}^2.$$

Аналогично доказывается неравенство $a_{0h}(p, u_h, u_h) \leq 1^p c^2 \|u_h\|_{1,\Omega}^2$. Из этих оценок следует, что

$$m_a(p) \|u_h\|_{1,\Omega}^2 \leq a_h(p, u_h, u_h) \leq M_a(p) \|u_h\|_{1,\Omega}^2, \quad (2.47)$$

где $m_a(p) := 1_p c^{-2}$, $M_a(p) := 1^p c^2 + M_s(p)$, $p \in \mathbb{R}_+$. Оценки

$$0 \leq b_h(u_h, u_h) \leq M_b \|u_h\|_{1,\Omega}^2, \quad M_b := (\sigma_+ - 1)c^2, \quad (2.48)$$

являются очевидными. Из (2.47) и (2.48) следует, что выполнены условия A_1^h . Далее, по определению

$$a_h(0, u_h, u_h) = S_h(|\nabla u_h|^2) + 2\pi \sum_{n=-N}^N |n| |a_n(u_h)|^2 \geq 0.$$

Учитывая (2.46), имеем $\ker A_h(0) = \{u_h : u_h = \text{const в } \Omega\}$. Поэтому $r_0 = 1$ и выполнены условия A_2^h (см. замечание 1.2).

Поскольку $\sigma|_{\tilde{\Omega}_i} \geq 1 + \sigma_0$, то, учитывая (2.46) и следствие 2.6, имеем при $c_0 := \min\{1, \sigma_0\}$:

$$\begin{aligned} a_h(p, u_h, u_h) + b_h(u_h, u_h) &\geq S_h(|\nabla u_h|^2) + \sigma_0 \sum_{e \in \Omega_i} S_e(u_h^2) \geq \\ &\geq c_0 \left(\sum_{e \in \Omega \setminus \Omega_i} S_e(|\nabla u_h|^2) + \sum_{e \in \Omega_i} S_e(|\nabla u_h|^2 + u_h^2) \right) \geq \\ &\geq c \left(\int_{\Omega} |\nabla u_h|^2 dx + \int_{\Omega_i} u_h^2 dx \right) \geq m_{ab} \|u_h\|_{1, \Omega}^2. \end{aligned}$$

Отметим, что постоянная m_{ab} зависит от Ω .

Перейдем к условию A_4^h . Имеем ($1 \geq \sigma/\sigma_+$):

$$b_h(u, u) = \sum_{e \in \Omega_i} S_e((\sigma - 1)u^2) \geq \sigma_0 \sum_{e \in \Omega_i} S_e(u^2) \geq \sigma_0/\sigma_+ \sum_{e \in \Omega_i} S_e(\sigma u^2).$$

С другой стороны, очевидно, $b_h(u, u) \leq \sum_{e \in \Omega_i} S_e(\sigma u^2) =: S(u)$. Поэтому,

если $u \in \tilde{V}_h^1$, то $S(u) \neq 0$. Отсюда легко следует, что отношение Рэля $R_h(p, u) := a_h(p, u, u)/b_h(u, u)$ строго монотонно возрастает по p при $u \in \tilde{V}_h$ (условие A_4^h).

Проверим условие A_5^h . Пусть $u \in V_h$. Непрерывность в нуле функции $p \rightarrow a_h(p, u, u)$ доказывается также, как и для формы a (см. лемму 2.8). Учитывая, что $\mathbb{K}'_n(r) > 0$ при $r > 0$, имеем:

$$\begin{aligned} 0 \leq \frac{d}{dp} a_h(p, u, u) &= 2p S_h(\sigma u^2) + \frac{d}{dp} s_N(p, u, u) \leq cp \|u\|_{0, V_h}^2 + \\ &+ \frac{d}{dp} s_\infty(p, u, u) \leq cp \|u\|_{0, \Omega}^2 + \tilde{M}_s(p) \|u\|_{1, \Omega}^2 \leq \tilde{M}_a(p) \|u\|_{1, \Omega}^2. \end{aligned}$$

¹⁾ \tilde{V}_h есть ортогональное дополнение $\ker B_h$ до пространства $V_{h, A_h(p)+B_h}$.

Таким образом, $A_h(p) \in \mathcal{L}(\mathbb{R}_+, V_h)$. \square

Пусть $N_h^b := \dim(\text{Im } B_h)$. Из теоремы 1.7, с. 27 (см. также замечание 1.2 к нему), непосредственно следует

Теорема 2.14. *При каждом $p \in \mathbb{R}_+$ задача (\mathcal{P}_h) имеет конечное число (скажем $n_h(p)$) собственных чисел $\beta^{hK}(p)$ суммарной кратности N_h^b и соответствующих им собственных подпространств*

$$U_h^K(p), \quad K = 1, 2, \dots, n_h(p), \quad V_{h, A_h(p)+B_h} = \bigoplus_{K=1}^{n_h(p)} U_h^K(p) \oplus \ker B_h.$$

Кроме того, если $0 \leq \beta_1^h(p) \leq \beta_2^h(p) \leq \dots \leq \beta_{N_h^b}^h(p)$ есть собственные числа, пронумерованные с учетом кратности, то:

функции $p \rightarrow \beta_i^h(p)$, $i = 1, 2, \dots, N_h^b$, являются непрерывными в нуле, локально липшиц-непрерывными на $(0, \infty)$ и строго монотонно возрастающими; $\beta_1^h(p) \rightarrow 0$ при $p \rightarrow +0$; $\beta_2^h(0) > 0$.

§ 4. Оценки точности

При каждом фиксированном $p \in \mathbb{R}_+$ задача (\mathcal{P}_h) является схемой Галеркина с возмущениями для задачи (\mathcal{P}) . Задачи такого типа были нами изучены в главе 1.

Пусть, $\beta_i^2(p)$ и $\beta_i^{h2}(p)$, $i = 1, 2, \dots$ — упорядоченные по возрастанию с учетом кратности собственные числа задачи (\mathcal{P}) и (\mathcal{P}_h) соответственно, $u_i(p)$ и $u_i^h(p)$ — соответствующие им собственные функции. Пусть далее $\beta^K(p)$ имеет кратность $r_K = r_K(p)$, $K \geq 1$,

$$\begin{aligned} \beta^K(p) &= \beta_k(p), \quad \beta_{k-1}(p) < \beta_k(p) = \dots = \beta_{k+r_K-1}(p) < \beta_{k+r_K}(p), \\ U^K(p) &:= \text{span}\{u_k(p), \dots, u_{k+r_K-1}(p)\}, \\ U_h^K(p) &:= \text{span}\{u_k^h(p), \dots, u_{k+r_K-1}^h(p)\}. \end{aligned}$$

Целью данного параграфа является оценка зависимости величин $|\beta^K(p) - \beta_i^h(p)|$, $k \leq i \leq k + r_K - 1$, и раствора $\Theta_V(U^K(p), U_h^K(p))$ от параметров h и N дискретной задачи (\mathcal{P}_h) .

Теорема 1.8, с. 29, приводит к следующим оценкам этих величин:

$$\Theta_V(U^K(p), U_h^K(p)) \leq c \max_{u \in U^K(p), \|u\|_{1,\Omega}=1} \epsilon_h(u), \quad (2.49)$$

$$|\beta^{K2}(p) - \beta_i^{h2}(p)| \leq c \max_{u \in U^K(p), \|u\|_{1,\Omega}=1} (\epsilon_h^2(u) + \Sigma_h(P_h u)). \quad (2.50)$$

Здесь $c = c(K, p)$, $P_h := P_h(p)$ — ортопроектор в $V_{A(p)+B}$ на V_h ,

$$\epsilon_h(u) := \inf_{v_h \in V_h} \|u - v_h\|_{1,\Omega} + E_a(P_h u) + E_b(P_h u), \quad u \in V,$$

$$\Sigma_h(y) := |a(p, y, y) - a_h(p, y, y)| + |b(y, y) - b_h(y, y)|, \quad y \in V_h,$$

для заданных форм d и d_h

$$E_d(\varphi_h) := \sup_{v_h \in V_h, \|v_h\|=1} |d(\varphi_h, v_h) - d_h(\varphi_h, v_h)|, \quad \varphi_h \in V_h.$$

Для того, чтобы воспользоваться оценками (2.49), (2.50), необходимо предварительно убедиться в выполнении условий (H_1) , A_6^h . Фактически, необходимо проверить лишь условие¹⁾

$$A_6^h) \quad \sup_{f \in V_h, \|f\|=1} (E_b(f) + E_b(P_h T(p)f) + E_a(P_h T(p)f)) \rightarrow 0 \quad \text{при } h \rightarrow 0,$$

поскольку (H_1) выполняется очевидным образом (см. (2.40)).

Прежде чем приступить к анализу правых частей в (2.49) и (2.50), оценим два типа возмущений в формах a и b , а именно возмущений от использования квадратурных формул и от усечения ряда.

До конца главы будем предполагать, что $K \geq 1$ и $p \in \mathbb{R}_+$ являются фиксированными, $\varepsilon|\Omega_i \subset W_\infty^{2m}(\Omega_i)$, а собственные функции u из $U^K(p)$ обладают следующей гладкостью:

$$u|_{\Omega_i} \subset H^{m+1}(\Omega_i), \quad u|_{\Omega \setminus \Omega_i} \subset H^{m+1}(\Omega \setminus \Omega_i), \quad (2.51)$$

где m то же число, что и в определении пространства V_h . Будем использовать дополнительные обозначения

$$|u|_j := |u|_{j,\Omega_i} + |u|_{j,\Omega \setminus \Omega_i}, \quad \|u\|_j := \|u\|_{j,\Omega_i} + \|u\|_{j,\Omega \setminus \Omega_i}, \quad j \geq 0.$$

4.1. Оценки погрешности численного интегрирования.

Введем обозначения функционалов погрешности квадратурных формул:

$$\widehat{E}(\hat{f}) := \int_{\hat{e}} \hat{f}(\hat{x}) d\hat{x} - \hat{S}(\hat{f}), \quad E_e(f) := \int_e f(x) dx - S_e(f),$$

$$E_h(f) := \sum_{e \in \mathcal{T}_h} E_e(f).$$

¹⁾здесь $T(p) := (A(p) + B)^{-1}B$, $A(p)$ и B — операторы задачи (P) .

Лемма 2.12. Пусть $\widehat{E}(\hat{f}) = 0$ для любого полинома $\hat{f} \in \widehat{P}_{2m-1}$, $p|_e \in W_\infty^{2m}(e)$ для $\forall e \in \mathcal{T}_h$. Пусть также E_h обозначает любую из величин $E_h(puv)$, $E_h\left(pu\frac{\partial v}{\partial x_i}\right)$ или $E_h\left(p\frac{\partial u}{\partial x_j}\frac{\partial v}{\partial x_i}\right)$, $i, j = 1, 2$. Тогда

$$|E_h| \leq c \|p\|_{2m, \infty, h} h^{2m} \|u\|_{m, h} \|v\|_{m, h} \quad \forall u, v \in V_h.$$

ДОКАЗАТЕЛЬСТВО. Обоснование оценок подобного типа проводится по стандартной в МКЭ схеме: в функционале погрешности осуществляется переход на базисный элемент, используется лемма Брамбла — Гильберта и совершается обратный переход на исходный элемент (см., напр, [34, с. 178], [33, с. 176]). Имеет место оценка:

$$|E_e(puv)| = |\widehat{E}(\hat{f})| \leq \|\hat{f}\|_{0, \infty, \hat{e}} \leq \|\hat{f}\|_{2m, \infty, \hat{e}},$$

где $\hat{f} := \hat{\sigma}\hat{u}\hat{v}$, $\hat{\sigma} := |J_e(\hat{x})|\hat{p}(\hat{x})$, $\hat{p}(\hat{x}) := p(x_e(\hat{x}))$, $|J_e(\hat{x})|$ — якобиан преобразования x_e . Так как $\widehat{E}(\widehat{P}_{2m-1}) = 0$, то из леммы Брамбла — Гильберта [35] следует, что

$$|E_e(puv)| \leq c |\hat{f}|_{2m, \infty, \hat{e}}. \quad (2.52)$$

Поскольку $\hat{u}, \hat{v} \in \widehat{P}_m$, а полунормы $|\cdot|_{i, \infty, \hat{e}}$ и $|\cdot|_{i, \hat{e}}$ эквивалентны на \widehat{P}_m , то пользуясь правилом Лейбница, получим:

$$\begin{aligned} |\hat{f}|_{2m, \infty, \hat{e}} &\leq c \sum_{i, j=0}^m |\hat{\sigma}|_{2m-i-j, \infty, \hat{e}} |\hat{u}|_{i, \infty, \hat{e}} |\hat{v}|_{j, \infty, \hat{e}} \leq \\ &\leq c \sum_{i, j=0}^m |\hat{\sigma}|_{2m-i-j, \infty, \hat{e}} |\hat{u}|_{i, \hat{e}} |\hat{v}|_{j, \hat{e}}. \end{aligned} \quad (2.53)$$

Из условий регулярности триангуляции (см. оценки (2.36), (2.37), (2.38) на с. 63) и правила Лейбница следует:

$$\begin{aligned} |\hat{\sigma}|_{s, \infty, \hat{e}} &\leq c \sum_{l=0}^s |J_e|_{s-l, \infty, \hat{e}} |\hat{p}|_{l, \infty, \hat{e}} \leq \\ &\leq c \sum_{l=0}^s h^{s-l+2} h^l \|p\|_{l, \infty, e} = c h^{s+2} \|p\|_{s, \infty, e}. \end{aligned}$$

Таким образом, из (2.53) будем иметь:

$$\begin{aligned} |\hat{f}|_{2m,\infty,\hat{e}} &\leq c \sum_{i,j=0}^m h^{2m-i-j+2} \|p\|_{2m-i-j,\infty,e} h^{i-1} \|u\|_{i,e} h^{j-1} \|v\|_{j,e} \leq \\ &\leq c h^{2m} \|p\|_{2m,\infty,e} \|u\|_{m,e} \|v\|_{m,e}. \end{aligned}$$

Комбинируя последнюю оценку и (2.52), имеем требуемое:

$$\begin{aligned} |E_h(puv)| &= \left| \sum_{e \in \mathcal{T}_h} E_e(puv) \right| \leq c h^{2m} \sum_{e \in \mathcal{T}_h} \|p\|_{2m,\infty,e} \|u\|_{m,e} \|v\|_{m,e} \leq \\ &\leq c h^{2m} \|p\|_{2m,\infty,h} \sum_{e \in \mathcal{T}_h} \|u\|_{m,e} \|v\|_{m,e} \leq c h^{2m} \|p\|_{2m,\infty,h} \|u\|_{m,h} \|v\|_{m,h}. \end{aligned}$$

Приведенные выше рассуждения полностью повторяются и в двух оставшихся случаях. Например,

$$E_h := E_h \left(p \frac{\partial u}{\partial x_j} \frac{\partial v}{\partial x_i} \right) = \sum_{e \in \mathcal{T}_h} I_e, \quad I_e := E_e \left(p \frac{\partial u}{\partial x_j} \frac{\partial v}{\partial x_i} \right).$$

Поскольку $\nabla u = J_e^T \hat{\nabla} \hat{u}$, J_e — матрица Якоби преобразования x_e , то I_e представляет собой сумму слагаемых вида $\hat{E}(\hat{f})$,

$$\hat{f} := \hat{\sigma} \frac{\partial \hat{u}}{\partial \hat{x}_j} \frac{\partial \hat{v}}{\partial \hat{x}_i}, \quad \hat{\sigma} := \frac{1}{|J_e(\hat{x})|} \frac{\partial x_{ei}(\hat{x})}{\partial \hat{x}_k} \frac{\partial x_{ej}(\hat{x})}{\partial \hat{x}_l} \hat{p},$$

причем $|\hat{E}(\hat{f})| \leq c |\hat{f}|_{2m,\infty,\hat{e}}$. Как и ранее, имеем: $|\hat{\sigma}|_{s,\infty,\hat{e}} \leq c h^s \|p\|_{s,\infty,e}$,

$$\begin{aligned} |\hat{E}(\hat{f})| &\leq c |\hat{f}|_{2m,\infty,\hat{e}} \leq c \sum_{i,j=0}^{m-1} |\hat{\sigma}|_{2m-i-j,\infty,\hat{e}} |\hat{u}|_{i+1,\hat{e}} |\hat{v}|_{j+1,\hat{e}} \leq \\ &\leq c h^{2m} \|p\|_{2m,\infty,e} \|u\|_{m,e} \|v\|_{m,e}. \end{aligned}$$

Отсюда, после суммирования по $e \in \mathcal{T}_h$, следует оценка E_h . \square

Следствие 2.7. Пусть $\varepsilon|_{\Omega_i} \in W_\infty^{2m}(\Omega_i)$, $u, v \in V_h$, $p \in \mathbb{R}_+$. Тогда

$$\begin{aligned} |a_0(p, u, v) - a_{0h}(p, u, v)| &\leq c h^{2m} \|u\|_{m,h} \|v\|_{m,h}, \\ |b(u, v) - b_h(u, v)| &\leq c h^{2m} \|u\|_{m,h} \|v\|_{m,h}. \end{aligned}$$

4.2. Вспомогательные утверждения.

Следующие две леммы понадобятся нам при оценке возмущений метода Галеркина.

Лемма 2.13. Пусть $u \in U^K(p)$, $k = 1, 2, \dots, m$. Тогда

$$\|u - P_h u\|_{1,\Omega} \leq c h^k \|u\|_{k+1}, \quad \|P_h u\|_{k,h} \leq c \|u\|_k. \quad (2.54)$$

ДОКАЗАТЕЛЬСТВО. Из эквивалентности норм в $V_{A(p)+B}$ и в V следует, что проектор P_h ограничен в V и

$$\|u - P_h u\|_{1,\Omega} \leq c \inf_{v_h \in V_h} \|u - v_h\|_{1,\Omega} \quad \forall u \in V.$$

Первая оценка леммы следует из этой оценки и (2.41), так как $\|u\|_{k,h} = \|u\|_k$. Из ограниченности P_h в V вытекает справедливость второй оценки в (2.54) при $k = 1$, поскольку $\|P_h u\|_{1,h} = \|P_h u\|_{1,\Omega}$.

Пусть $k \geq 2$, π_h есть такой проектор V на V_h , что

$$\|u - \pi_h u\|_{1,\Omega} \leq c h^{k-1} \|u\|_k, \quad \|u - \pi_h u\|_{k,h} \leq c \|u\|_k.$$

Например, в качестве π_h можно выбрать оператор V_h -интерполяции¹⁾. Используя обратное неравенство (2.39), имеем:

$$\begin{aligned} \|\pi_h u - P_h u\|_{k,h} &\leq c h^{1-k} \|\pi_h u - P_h u\|_{1,\Omega} \leq \\ &\leq c h^{1-k} (\|u - \pi_h u\|_{1,\Omega} + \|u - P_h u\|_{1,\Omega}) \leq c \|u\|_k. \end{aligned}$$

Следовательно,

$$\|P_h u\|_{k,h} \leq \|u\|_{k,h} + \|u - \pi_h u\|_{k,h} + \|\pi_h u - P_h u\|_{k,h} \leq c \|u\|_k. \quad \square$$

Обозначим через $\mathcal{A}(\Omega)$ множество функций из $H^1(\Omega)$, гармонических или метагармонических в $\Omega \setminus \bar{\Omega}_i$. Отметим, что $U^K(p) \subset \mathcal{A}(\Omega)$.

Лемма 2.14. Пусть $p \in \mathbb{R}_+$, $S_N(u) := 2\pi \sum_{|n| > N} \mathbb{K}_n(Rp) |a_n(u)|^2$, $u \in \mathcal{A}(\Omega)$. Тогда²⁾

$$S_N(P_h u) \leq c (\|u - P_h u\|_{1,\Omega}^2 + (R_0/R)^{2N} \|u\|_{1,\Omega}^2).$$

¹⁾ $\pi_h u \in V_h$, $\pi_h u(a_i) = u(a_i)$ в каждой точке сетки a_i .

²⁾ напомним, что $R \geq R_0$, $R_0 := \min\{r : \Omega_i \subseteq B_r\}$

ДОКАЗАТЕЛЬСТВО. Имеем: $S_N(P_h u) \leq 2(S_N(u - P_h u) + S_N(u))$, $\mathbb{K}_n(Rp) \leq 1^p \mathbb{K}_n(R)$ (см. замечание 2.5). Поэтому (см. следствие 2.3):

$$\begin{aligned} S_N(u - P_h u) &:= 2\pi \sum_{|n|>N} \mathbb{K}_n(Rp) |a_n(u - P_h u)|^2 \leq \\ &\leq 1^p 2\pi \sum_{|n|>N} \mathbb{K}_n(R) |a_n(u - P_h u)|^2 \leq 1^p |u - P_h u|_{1/2,\Gamma}^2 \leq c \|u - P_h u\|_{1,\Omega}^2. \end{aligned}$$

Из леммы 2.5 следует, что

$$|a_n(u)| \leq (R_0/R)^{|n|} |a_n(u_0)|, \quad u_0 = u|_{\partial B_{R_0}}.$$

Кроме того $\mathbb{K}_n(Rp) \leq (Rp/R_0)^2 \mathbb{K}_n(R_0)$. Поэтому

$$\begin{aligned} S_N(u) &\leq p^2 (R_0/R)^{2N} 2\pi \sum_{|n|>N} \mathbb{K}_n(R_0) |a_n(u_0)|^2 \leq \\ &\leq p^2 (R_0/R)^{2N} |u|_{1/2,\partial B_{R_0}}^2 \leq c (R_0/R)^{2N} \|u\|_{1,B_{R_0}}^2 \leq c (R_0/R)^{2N} \|u\|_{1,\Omega}^2. \end{aligned}$$

Теперь из оценок $S_N(u - P_h u)$ и $S_N(u)$ следует требуемое утверждение.

4.3. Оценки возмущений.

Оценим возмущения форм, вызванные численным интегрированием и усечением ряда. Положим

$$\epsilon_{hm} := h^m + (R_0/R)^N, \quad \epsilon_{h0} := \|(I - P_h)T(p)\| + h^2 + (R_0/R)^N.$$

Теорема 2.15. Пусть $p \in \mathbb{R}_+$, $u_h \in V_h$, $u \in U^K(p)$. Тогда

$$E_a(P_h T(p)u_h) \leq c \epsilon_{h0} \|u_h\|_{1,\Omega}, \quad (2.55)$$

$$E_a(P_h u) \leq c \epsilon_{hm} \|u\|_{m+1}, \quad (2.56)$$

$$|a(p, P_h u, P_h u) - a_h(p, P_h u, P_h u)| \leq c \epsilon_{hm}^2 \|u\|_{m+1}^2. \quad (2.57)$$

ДОКАЗАТЕЛЬСТВО. Имеем для любых $f_h, v_h \in V_h$ ¹⁾:

$$(a - a_h)(p, f_h, v_h) = (a_0 - a_{0h})(p, f_h, v_h) + 2\pi \sum_{|n|>N} \mathbb{K}_n(Rp) a_n(f_h) \overline{a_n(v_h)}.$$

¹⁾ $(a - a_h)(p, u, v) = a(p, u, v) - a_h(p, u, v)$.

Воспользуемся следствием 2.7. Тогда

$$|(a - a_h)(p, f_h, v_h)| \leq c h^{2m} \|f_h\|_{m,h} \|v_h\|_{m,h} + 2\pi \sum_{|n|>N} \mathbb{K}_n(Rp) |a_n(f_h)| |a_n(v_h)|. \quad (2.58)$$

Выберем здесь $f_h = P_h v$, $v = T(p)u_h$ и учтем, что $v \in \mathcal{A}(\Omega)$,

$$2\pi \sum_{|n|>N} \mathbb{K}_n(Rp) |a_n(v_h)|^2 \leq c \|v_h\|_{1,\Omega}^2,$$

$\|v_h\|_{m,h} \leq c h^{1-m} \|v_h\|_{1,\Omega}$ (обратное неравенство). Используя неравенство Коши — Буняковского и лемму 2.14, получим:

$$\begin{aligned} |(a - a_h)(p, P_h v, v_h)| &\leq c h^2 \|P_h v\|_{1,\Omega} \|v_h\|_{1,\Omega} + c S_N^{1/2}(P_h v) \|v_h\|_{1,\Omega} \leq \\ &\leq c (h^2 \|v\|_{1,\Omega} + \|v - P_h v\|_{1,\Omega} + (R_0/R)^N \|v\|_{1,\Omega}) \|v_h\|_{1,\Omega}. \end{aligned}$$

Отметим, что $T(p)$ компактный оператор при каждом $p \in \mathbb{R}_+$, $\|v\|_{1,\Omega} \leq c \|u_h\|_{1,\Omega}$, $\|v - P_h v\|_{1,\Omega} \leq \|(I - P_h)T(p)\| \|u_h\|_{1,\Omega}$. Отсюда следует первая оценка (2.55).

Выберем теперь в (2.58) $f_h = P_h u$ и учтем оценки (2.54). Получим:

$$\begin{aligned} |(a - a_h)(p, P_h u, v_h)| &\leq c h^{m+1} \|P_h u\|_{m,h} \|v_h\|_{1,\Omega} + c S_N^{1/2}(P_h u) \|v_h\|_{1,\Omega} \leq \\ &\leq c \epsilon_{hm} \|u\|_{m+1} \|v_h\|_{1,\Omega}. \end{aligned}$$

Из этой оценки следует (2.56). Наконец, выбирая в (2.58) функции $f_h = v_h = P_h u$, из (2.54) и леммы 2.14 получим оценку (2.57). \square

Теорема 2.16. Пусть $u_h \in V_h$, $u \in U^K(p)$. Тогда

$$E_b(u_h) \leq c h^2 \|u_h\|_{1,\Omega}, \quad (2.59)$$

$$E_b(P_h u) \leq c h^{m+1} \|u\|_m, \quad (2.60)$$

$$|b(P_h u, P_h u) - b_h(P_h u, P_h u)| \leq c h^{2m} \|u\|_m^2. \quad (2.61)$$

ДОКАЗАТЕЛЬСТВО. Согласно следствию 2.7

$$|b(u_h, v_h) - b_h(u_h, v_h)| \leq c h^{2m} \|u_h\|_{m,h} \|v_h\|_{m,h}, \quad u_h, v_h \in V_h.$$

Дважды применяя здесь обратное неравенство, получим (2.59). Выбирая $u_h = P_h u$, используя (2.54) и обратное неравенство, будем иметь (2.60). Полагая $u_h = v_h = P_h u$ и учитывая (2.54), получим (2.61). \square

4.4. Оценки точности приближенных решений.

Используя полученные выше оценки величин, входящих в условие A_6^h и неравенства (2.49), (2.50), оценим точность приближенного метода.

Теорема 2.17. Пусть $p \in \mathbb{R}_+$, $\varepsilon|_{\Omega_i} \in W_\infty^{2m}(\Omega_i)$, выполнены условия гладкости (2.51) и

$$N \geq c_0 \ln(1/h), \quad c_0 = \frac{m}{\ln(R/R_0)}.$$

Тогда при достаточно малом h имеют место оценки:

$$\Theta_V(U^K(p), U_h^K(p)) \leq ch^m, \quad |\beta^K(p) - \beta_i^h(p)| \leq ch^{2m},$$

где $K \geq 1$, $i = k, \dots, k + r_K - 1$, $c = c(K, p)$.

ДОКАЗАТЕЛЬСТВО. Воспользуемся оценками (2.55), (2.59). Пусть $u_h \in V_h$, $v_h = P_h T(p)u_h$. Тогда

$$E_b(v_h) \leq ch^2 \|v_h\|_{1,\Omega} \leq ch^2 \|u_h\|_{1,\Omega}.$$

Легко видеть, что $(R_0/R)^N \leq h^m$. Поэтому

$$\begin{aligned} E_b(u_h) + E_b(P_h T(p)u_h) + E_a(P_h T(p)u_h) &\leq c(h^2 + \epsilon_{h0}) \|u_h\|_{1,\Omega} \leq \\ &\leq c(h + \|(I - P_h)T(p)\|) \|u_h\|_{1,\Omega}. \end{aligned}$$

Отсюда следует справедливость условия A_6^h , поскольку, в силу условия (H_1) и компактности оператора $T(p)$, имеем $\|(I - P_h)T(p)\| \rightarrow 0$ при $h \rightarrow 0$ (см., напр., [4, Лемма 15.4, с. 202]).

Обратимся к оценкам (2.49), (2.50). Из оценок (2.56) и (2.60) получаем, что $E_a(P_h u) + E_b(P_h u) \leq ch^m \|u\|_{m+1}$. Следовательно, с учетом (2.41), имеем $\epsilon_h(u) \leq ch^m \|u\|_{m+1}$. Оценка $\Sigma_h(P_h u) \leq ch^{2m} \|u\|_{m+1}^2$ вытекает непосредственно из (2.57) и (2.61). Таким образом, из оценок (2.49), (2.50) имеем:

$$\Theta_V(U^K(p), U_h^K(p)) \leq ch^m, \quad |\beta^{K^2}(p) - \beta_i^{h^2}(p)| \leq ch^{2m}, \quad (2.62)$$

поскольку, в силу конечномерности пространства $U^K(p)$,

$$\max_{u \in U^K(p), \|u\|=1} \|u\|_{m+1} \leq c, \quad c = c(K, p).$$

Ясно, что из (2.62) следует также, что $|\beta^K(p) - \beta_i^h(p)| \leq c h^{2m}$. \square

ЗАМЕЧАНИЕ 2.14. В случае волновода кругового поперечного сечения ($\Omega_i = B_{R_0}$) естественно выбрать $R = R_0$. Однако этот выбор исключается в формулировке теоремы (иначе $c_0 = \infty$). Преодолеть это ограничение нетрудно, если уточнить лемму 2.14. В этом случае

$$S_N(P_h u) \leq c (\|u - P_h u\|_{1,\Omega}^2 + (N+1)^{-2m} \|u\|_{1,\Omega}^2).$$

Для доказательства этого утверждения достаточно m -раз проинтегрировать по частям в выражении $a_n(u)$ при оценке $S_N(u)$. Далее, повторяя доказательство теоремы 2.17, нетрудно вывести, что ее результаты остаются справедливыми при условии $N = O(1/h)$.

ГЛАВА 3
РЕЗУЛЬТАТЫ ЧИСЛЕННЫХ
ЭКСПЕРИМЕНТОВ

Эта глава посвящена описанию результатов различных вычислительных экспериментов, целью которых является практическая оценка точности рассмотренного в предыдущей главе метода. В вычислениях была использована простейшая схема МКЭ ($m = 1$ в определении пространства конечных элементов V_h) и квадратурная формула с одним узлом в центре тяжести конечного элемента.

§ 1. Некоторые аспекты программной реализации

Напомним формулировку дискретной задачи: *при каждом $p \in \mathbb{R}_+$ найти $\beta^h \in \mathbb{R}_+$ и $y \in \mathbb{R}^{N_h} \setminus \{0\}$ такие, что*

$$A_h(p)y = \beta^{h2} B_h y, \quad (3.1)$$

где $A_h(p) = \{a_h(p, \varphi_j, \varphi_i)\}_{i,j=1}^{N_h}$, $B_h = \{b_h(\varphi_j, \varphi_i)\}_{i,j=1}^{N_h}$, функции $\{\varphi_i\}_{i=1}^{N_h}$ образуют базис Лагранжа в V_h ,

$$a_h(p, u, v) = S_h(\nabla u \cdot \nabla v + p^2 \sigma uv) + 2\pi \sum_{n=-N}^N \mathbb{K}_n(pR) a_n(u) \overline{a_n(v)},$$

$$b_h(u, v) = S_h((\sigma - 1) uv).$$

Представим матрицу $A_h(p)$ в виде суммы матриц $A_{0h}(p)$ и $S_h^N(p)$, где

$$S_h^N(p) = \{s_{ij}\}_{i,j=1}^{N_h}, \quad s_{ij} = 2\pi \sum_{n=-N}^N \mathbb{K}_n(pR) a_n(\varphi_j) \overline{a_n(\varphi_i)}.$$

Отметим, что матрицы $A_{0h}(p)$ и B_h являются обычными для метода конечных элементов. Их вычисление осуществляется известным алгоритмом сборки, который сводится к однотипным поэлементным вычислениям. Укажем способ вычисления матрицы $S_h^N(p)$.

Пусть в результате триангуляции области Ω ее граница Γ разбита точками ϕ_i на n_Γ равных частей длины $h = 2\pi/n_\Gamma$. Будем считать также, что граничные точки последовательно имеют номера от 1 до n_Γ , внутренние — от $n_\Gamma + 1$ до N_h . Поскольку φ_i при $n_\Gamma + 1 \leq i \leq N_h$ равны нулю на Γ , то матрица $S_h^N(p)$ имеет следующий вид:

$$S_h^N(p) = \begin{bmatrix} S(p) & 0 \\ 0 & 0 \end{bmatrix}, \quad S(p) = \{s_{kl}\}_{k,l=1}^{n_\Gamma}.$$

По определению пространства V_h функции φ_k , $k = 1, 2, \dots, n_\Gamma$, равны нулю на Γ вне отрезка $[\phi_{k-1}, \phi_{k+1}]$ и

$$\varphi_k|_\Gamma = \begin{cases} (\varphi - \phi_{k-1})/h, & \varphi \in [\phi_{k-1}, \phi_k], \\ (\phi_{k+1} - \varphi)/h, & \varphi \in [\phi_k, \phi_{k+1}]. \end{cases}$$

Простые вычисления приводят к следующей формуле:

$$a_n(\varphi_k) := \frac{1}{2\pi} \int_0^{2\pi} \varphi_k|_\Gamma e^{-in\varphi} d\varphi = d_n e^{-in\phi_k}, \quad d_n = \frac{h}{2\pi} \left(\frac{\sin(hn/2)}{hn/2} \right)^2.$$

Учтем, что $a_n(\varphi_k) = \overline{a_{-n}(\varphi_k)}$. Получим¹⁾:

$$s_{kl} = 4\pi \sum'_{n=0}^N d_n^2 \mathbb{K}_n(pR) (\cos(n\phi_l) \cos(n\phi_k) + \sin(n\phi_l) \sin(n\phi_k)).$$

Введем в рассмотрение прямоугольные матрицы Q_c и Q_s размера $n_\Gamma \times (N+1)$ с элементами $\cos(n\phi_l)$ и $\sin(n\phi_l)$ соответственно, а также диагональную матрицу

$$D(p) = 4\pi \operatorname{diag}(0.5d_0^2 \mathbb{K}_0(pR), d_1^2 \mathbb{K}_1(pR), \dots, d_N^2 \mathbb{K}_N(pR)).$$

Тогда, очевидно, справедлива формула

$$S(p) = Q_c D(p) Q_c^T + Q_s D(p) Q_s^T.$$

Матрицы Q_c и Q_s вычисляются один раз и при новом значении p переычисляется только матрица $D(p)$. Следовательно, для вычисления $S(p)$ при данном p требуется порядка $O(n_\Gamma N^2)$ арифметических

¹⁾ в сумме $\sum' a_n$ слагаемое a_0 умножается на 0.5

операций. Выразим это число в терминах h . В силу регулярности триангуляции $n_\Gamma = O(h^{-1})$; из оценок точности метода следует, что достаточно выбрать $N = O(\ln(1/h))$. Таким образом, для вычисления $S(p)$ при данном p требуется порядка $O(h^{-1} \ln^2(1/h))$ арифметических операций, тогда как для вычисления матриц $A_{0h}(p)$ и B_h требуется порядка $O(h^{-2})$ операций.

Далее мы рассчитаем конкретные волноводы, характеризующихся различной формой поперечного сечения Ω_i и постоянной проницаемостью ε ¹⁾. Для каждого волновода мы приводим графики дисперсионных кривых, а также исследуем точность вычисления собственных чисел β в зависимости от расчетных параметров h и N .

Опишем алгоритм определения дисперсионных кривых. Напомним, что исходная задача состоит в определении кривых $\beta = \beta(k)$ и соответствующих собственных функций $u = u(k)$, зависящих от волнового числа k (k пропорционально частоте электромагнитных колебаний ω). Предположим, что нас интересуют все дисперсионные кривые и соответствующие им собственные волны в интервале частот $[0, \omega_0]$ (т. е. при $k \in [0, k_0]$). Мы знаем, что дисперсионные кривые лежат в области

$$\Lambda_0 := \{(\beta, k) : \varepsilon_\infty^{1/2} k \leq \beta \leq \varepsilon_+^{1/2} k, k \in [0, k_0]\}.$$

Используя предлагаемый нами метод, эта задача решается следующим образом. Мы переформулируем задачу в терминах параметров (β, p) , вместо (β, k) , где $p = (\beta^2 - k^2 \varepsilon_\infty)^{1/2}$. Области Λ_0 взаимнооднозначно соответствует область K_0 параметров (β, p) ,

$$K_0 := \{(\beta, p) : p \in [0, p_0], \gamma p \leq \beta \leq (p^2 + k_0 \varepsilon_+)^{1/2}\},$$

где $\gamma := (\varepsilon_+ / (\varepsilon_+ - \varepsilon_\infty))^{1/2}$, $p_0 := (\varepsilon_+ - \varepsilon_\infty)^{1/2} k_0$.

На отрезке $[0, p_0]$ введем равномерную сетку узлов и в каждой точке сетки p_j найдем все собственные числа $\beta_i^{h^2}$ дискретной задачи (3.1) из отрезка $[\gamma^2 p_j^2, p_j^2 + k_0 \varepsilon_+]$ и соответствующие им собственные функции. В результате интерполяции (например, кусочно-линейной) получим приближенные дисперсионные кривые $\beta = \beta_i^h(p)$, лежащие в области K_0 . Положим $k_i^{h^2}(p) := (\beta_i^{h^2}(p) - p^2) / \varepsilon_\infty$. Найденные кривые

$$\beta = \beta_i^h(p), \quad k = k_i^h(p), \quad p \in [0, p_0],$$

¹⁾Будем считать, что единицы размерностей выбраны так, что $\varepsilon_\infty = 1$, $\mu_0 = 1$.

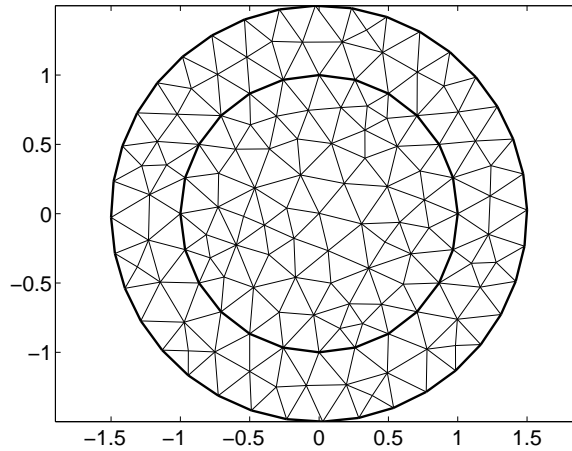


Рис. 1. Триангуляция области Ω для кругового волновода, $R = 1.5$, $N_h = 146$.

образуют в совокупности приближения ко всем искомым дисперсионным кривым точности $O(h^{2m})$. Отметим, что в результате решения дискретных задач на собственные значения (мы использовали метод Ланцоша) определяются приближения лишь к сужениям собственных функций на круг Ω . При необходимости они могут быть метагармонически продолжены вне Ω .

§ 2. Волновод кругового поперечного сечения

Рассмотрим однородный волновод с круговым поперечным сечением радиуса 1 и проницаемости $\varepsilon = 2$. Этот пример интересен тем, что для него известно “точное” решение задачи.

Радиус R области Ω был выбран равным 1.5. На рис. 2 изображены первые семь дисперсионных кривых $\beta = \beta(p)$ задачи (\mathcal{P}) , вычисленные на сетке с 2493 узлами и числом Фурье-гармоник $N = 10$. Сплошными линиями обозначено точное решение, точками — приближенное, штриховой линией изображена прямая $\beta = k_0 p$, определяющая границу области K . Все дисперсионные кривые лежат выше этой прямой.

На рис. 3 сплошными линиями изображены первые семь дисперсионных кривых $\beta = \beta(k)$ исходной задачи (\mathcal{P}_∞) . Штриховыми линиями представлены прямые $\beta = k\sqrt{\varepsilon_+}$ и $\beta = k\sqrt{\varepsilon_\infty}$, определяющие границы области Λ .

Представим теперь результаты численного исследования зависи-

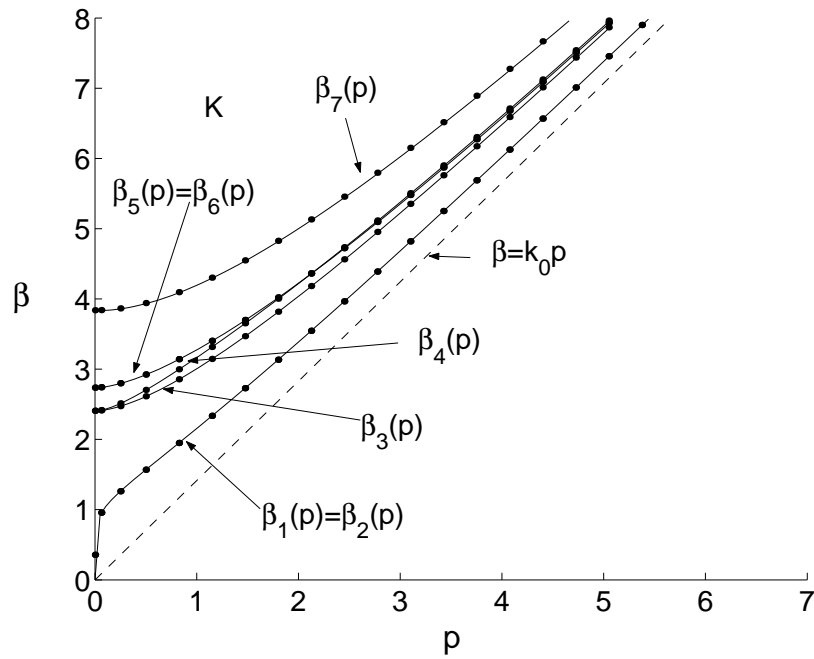


Рис. 2. Первые семь дисперсионных кривых $\beta = \beta(p)$ для кругового волновода.

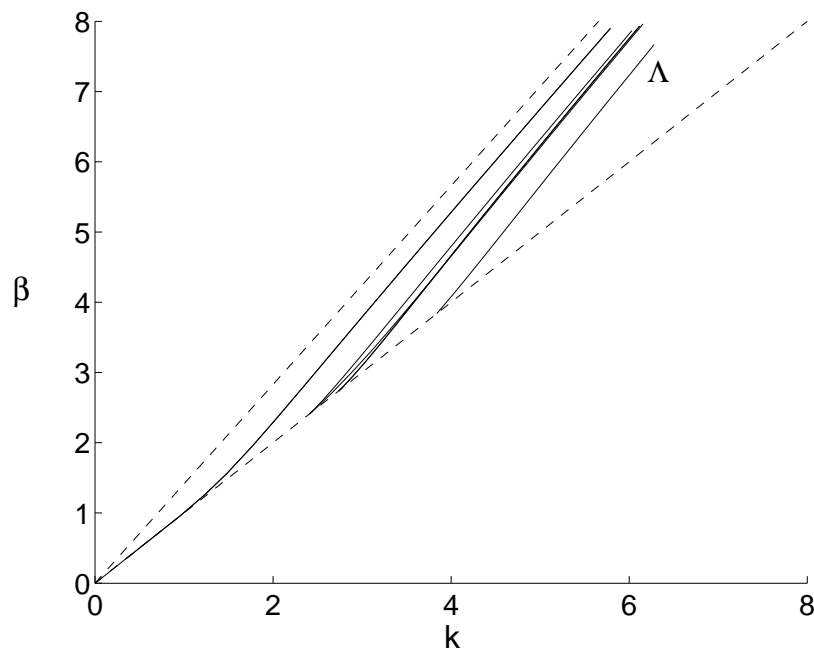


Рис. 3. Первые семь дисперсионных кривых $\beta = \beta(k)$ для кругового волновода, $\beta_1(k) = \beta_2(k)$, $\beta_5(k) = \beta_6(k)$.

мости точности метода от параметров N_h — общего числа точек сетки и N — числа Фурье-гармоник. Для фиксированного параметра $p = 1$ разыскивались собственные значения β^h задачи (\mathcal{P}_h) , а затем сравнивались с точными. Результаты вычислений представлены в таблице 4.1 для β_4^h . Из этой таблицы видно, что достаточно выбрать $N = 1$ или $N = 2$ и при этом $|\beta_4 - \beta_4^h|/|\beta_4| \approx 0.7h^2$.

$N \setminus N_h(n_\Gamma)$	45(16)	330(52)	1125(92)	2881(152)
1	0.64	0.748	0.631	0.668
3	0.641	0.748	0.631	0.668
5	0.641	0.748	0.631	0.668
7	0.641	0.748	0.631	0.668
15	0.642	0.748	0.631	0.668

Таблица 4.1: зависимость $e = h^{-2}|\beta_4 - \beta_4^h|/|\beta_4|$ при $p = 1$ от h и N .

§ 3. Волновод квадратного поперечного сечения

Рассмотрим однородный волновод с поперечным сечением в виде единичного квадрата и проницаемости $\varepsilon = 2.08$, для которого известны результаты физических экспериментов [36]. Круг Ω расположим так, чтобы центр круга совпадал с центром квадрата Ω_i , радиус R круга выберем равным 1.5. Пример триангуляции области Ω приведен на рис. 4.

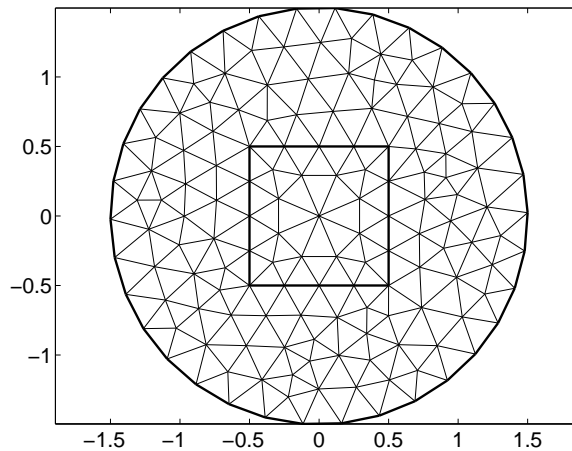


Рис. 4. Триангуляция области для волновода квадратного поперечного сечения, $R = 1.5$, $N_h = 151$.

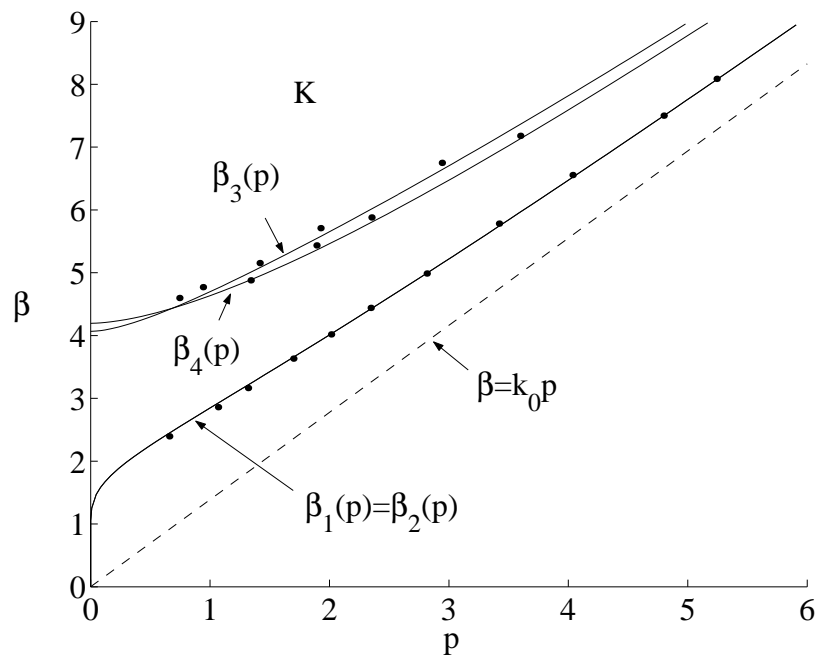


Рис. 5. Первые четыре дисперсионные кривые $\beta = \beta(p)$ для квадратного волновода.

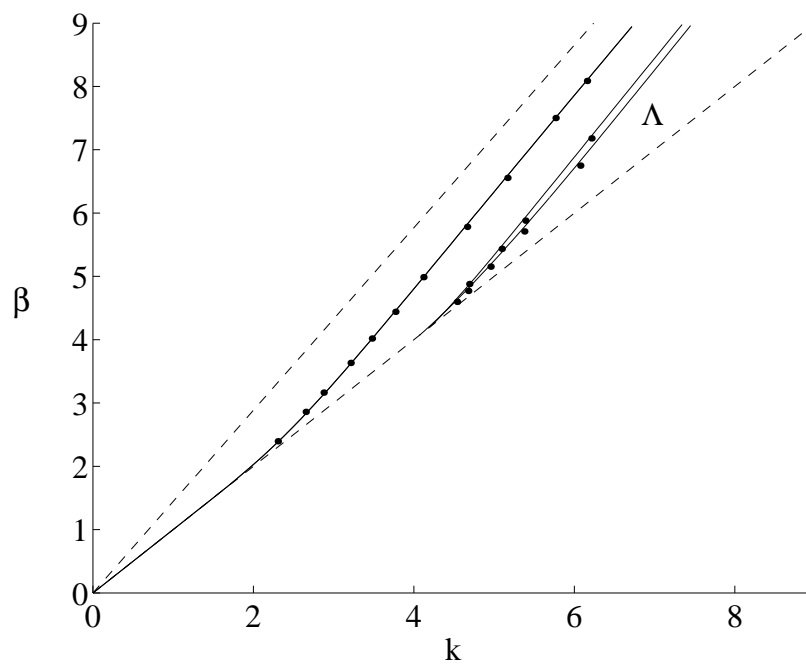


Рис. 6. Первые четыре дисперсионные кривые $\beta = \beta(k)$ для квадратного волновода, $\beta_1(k) = \beta_2(k)$.

На рис. 5 сплошными линиями показаны первые четыре дисперсионные кривые $\beta = \beta(p)$ задачи (\mathcal{P}) , построенные на сетке с общим числом узлов $N_h = 2500$ и вычисленные при количестве Фурье-гармоник $N = 10$. Нижняя кривая является кратной $\beta_1(p) = \beta_2(p)$, две другие пересекаются. Экспериментальные данные обозначены точками. График показывает хорошее соответствие между полученными приближенными решениями и экспериментальными данными. На рис. 6 изображены соответствующие им дисперсионные кривые $\beta = \beta(k)$ задачи (\mathcal{P}_∞) .

Представим теперь результаты численного исследования зависимости точности метода от параметров N_h и N при $p = 1$. В отличие от кругового волновода в данном случае не известно точное решение. Поэтому за “точное решение задачи” принималось приближенное решение, полученное на сетке с числом узлов $N_h = 6000$ ($n_\Gamma = 212$). Результаты вычислений представлены в таблице 4.2 для третьего собственного значения β_3 .

$N \mid N_h(n_\Gamma)$	31(16)	341(50)	1012(92)
1	2.26	1.6	1.61
3	2.27	1.61	1.64
5	2.27	1.61	1.64
7	2.27	1.61	1.64
15	2.27	1.61	1.64

Таблица 4.2: зависимость $e = h^{-2}|\beta_3 - \beta_3^h|/|\beta_3|$ при $p = 1$ от h и N .

Из этой таблицы можно заключить, что достаточно выбрать N равное 3, и при этом $|\beta_4 - \beta_4^h|/|\beta_4| \approx 1.6h^2$.

§ 4. Волновод прямоугольного поперечного сечения

Приведем еще один пример волновода, для которого известны экспериментальные данные [36]. Рассмотрим однородный волновод, поперечное сечение которого представляет собой прямоугольник размера 1.5×1 с $\varepsilon = 2.08$. Круг Ω расположим так, чтобы центр круга совпадал с центром прямоугольника Ω_i , радиус R круга выберем равным 1.5. Пример триангуляции области Ω приведен на рис. 7.

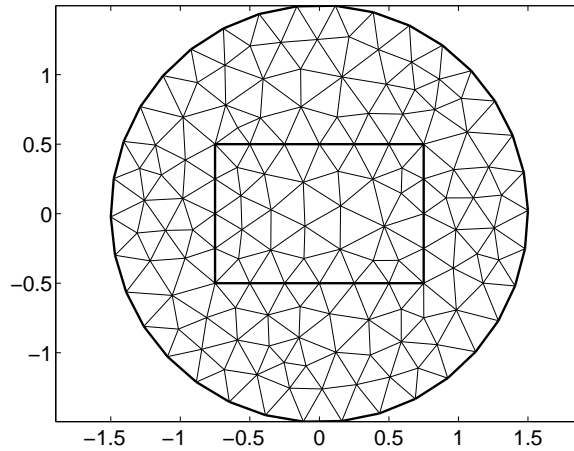


Рис. 7. Триангуляция области для прямоугольного волновода, $R = 1.5$, $N_h = 148$.

На рис. 8 сплошными линиями показаны первые четыре дисперсионные кривые $\beta = \beta(p)$ задачи (\mathcal{P}) для прямоугольного волновода, построенные на сетке с общим числом узлов $N_h = 2179$ и вычисленные при количестве Фурье-гармоник $N = 10$. Экспериментальные данные обозначены точками. График показывает хорошее соответствие между полученными приближенными решениями и экспериментальными данными. На рис. 9 изображены соответствующие им дисперсионные кривые $\beta = \beta(k)$ задачи (\mathcal{P}_∞) .

Представим теперь результаты численного исследования зависимости точности метода от параметров N_h и N . Так как в данном случае точное решение не известно, то схема исследования такая же, как в случае квадратного волновода: за “точное решение задачи” принималось приближенное решение, полученное на сетке с общим числом узлов $N_h = 6015$ ($n_\Gamma = 212$). Результаты вычислений представлены в таблице 4.3 для третьего собственного значения β_3 .

N $N_h(n_\Gamma)$	40(17)	304(50)	1016(92)
1	1.36	0.987	0.327
3	1.39	1.19	1.04
5	1.39	1.19	1.05
7	1.39	1.19	1.05
15	1.39	1.19	1.05

Таблица 4.3: зависимость $e = h^{-2}|\beta_3 - \beta_3^h|/|\beta_3|$ при $p = 1$ от h и N .

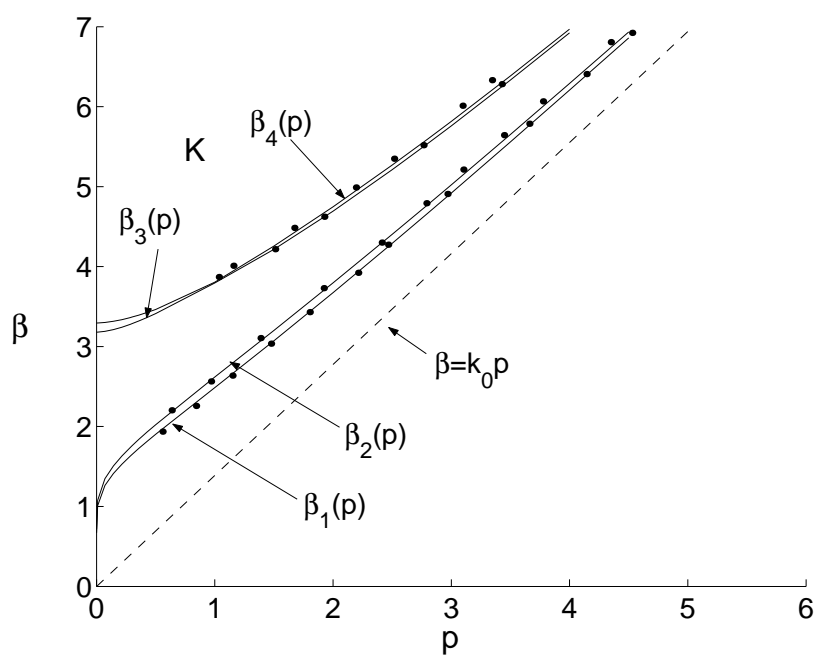


Рис. 8. Первые четыре дисперсионные кривые для прямоугольного волновода.

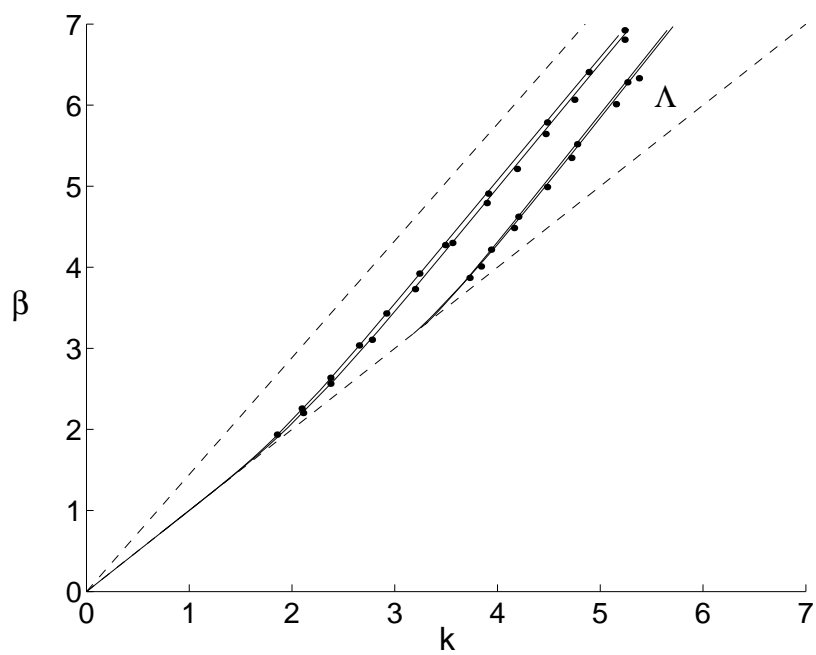


Рис. 9. Первые четыре дисперсионные кривые для прямоугольного волновода.

Из этой таблицы можно заключить, что достаточно выбрать число Фурье-гармоник $N = 5$, и при этом $|\beta_3 - \beta_3^h|/|\beta_3| \approx 1.1h^2$.

§ 5. Волновод с поперечным сечением из трех кругов

Рассмотрим однородный волновод, состоящей из трех касающихся друг друга кругов радиуса 0.4. В этом случае не известно ни точное решение, ни экспериментальные данные. Радиус R круга Ω был выбран равным 1.5, центр области Ω_i совпадает с центром Ω , $\varepsilon = 2$. Грубая триангуляция области приведена на рис. 10.

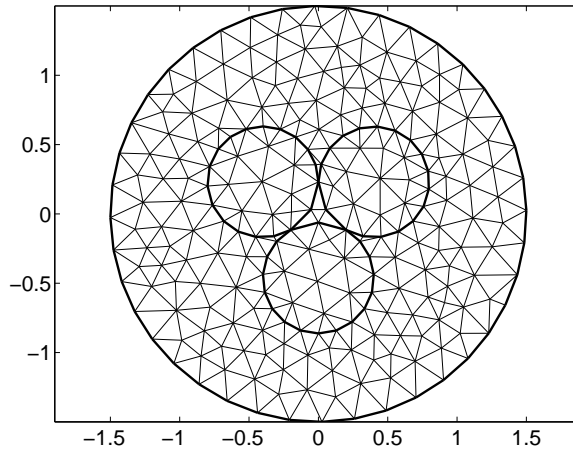


Рис. 10. Триангуляция области поперечного сечения волновода, состоящего из трех касающихся друг друга кругов радиуса 0.4, $R = 1.5$, $N_h=243$.

На рис. 11 изображены первые шесть дисперсионных кривых $\beta = \beta(p)$ задачи (\mathcal{P}) , вычисленные на сетке с общим числом узлов $N_h = 2226$ при количестве Фурье-гармоник $N = 10$. Первая и пятая кривые, изображенные на рисунке, являются кратными: $\beta_1(p) = \beta_2(p)$, $\beta_5(p) = \beta_6(p)$. На рис. 12 изображены дисперсионные кривые $\beta = \beta(k)$ задачи (\mathcal{P}_∞) .

Представим результаты численного исследования зависимости точности метода от параметров N_h и N . Так как в данном случае точное решение не известно, то за “точное решение задачи” принималось приближенное решение, полученное на сетке с числом узлов $N_h = 6006$ ($n_\Gamma = 216$). Результаты вычислений представлены в таблице 4.4 для четвертого собственного значения β_4 .

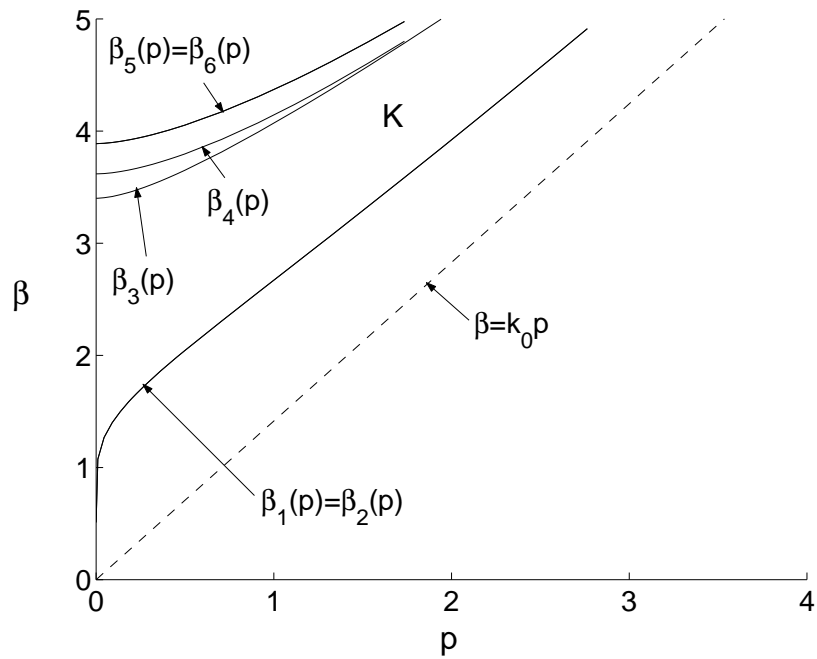


Рис. 11. Первые шесть дисперсионных кривых для волновода с поперечным сечением, состоящим из трех касающихся друг друга кругов.

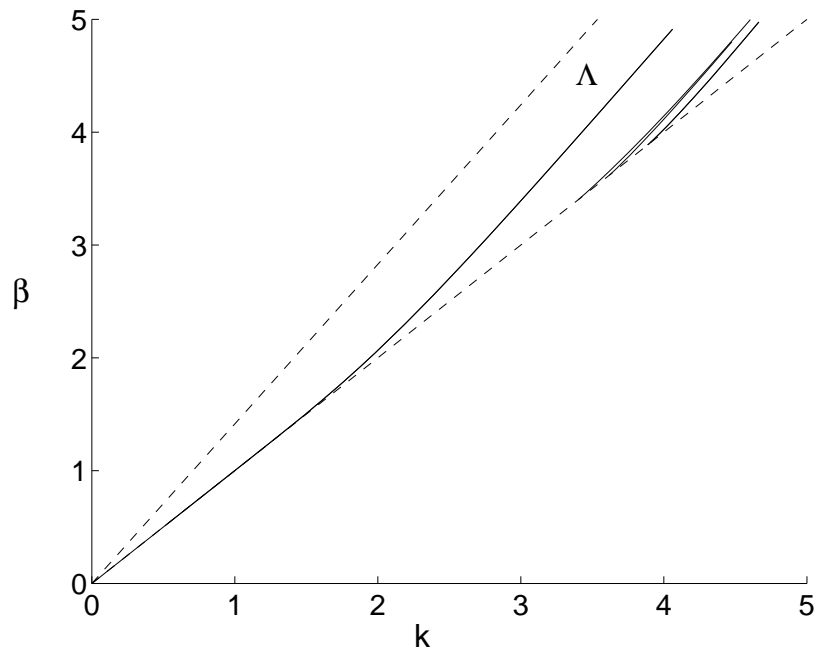


Рис. 12. Первые шесть дисперсионных кривых для волновода с поперечным сечением, состоящим из трех касающихся друг друга кругов.

N $N_h(n_\Gamma)$	78(16)	335(50)	1093(90)
1	0.5	23.3	92.5
3	0.619	1.67	1.56
5	0.62	1.67	1.57
7	0.62	1.67	1.57
15	0.62	1.67	1.57

Таблица 4.4: зависимость $e = h^{-2}|\beta_4 - \beta_4^h|/|\beta_4|$ при $p = 1$ от h и N .

Из этой таблицы можно заключить, что достаточно выбрать число Фурье-гармоник $N = 5$, и при этом $|\beta_4 - \beta_4^h|/|\beta_4| \approx 1.6h^2$.

Литература

1. Канторович Л.В., Акилов Г.П. Функциональный анализ в нормированных пространствах – М.: Наука, 1977. – 744 с.
2. Колмогоров А.Н., Фомин С.В. Элементы теории функций и функционального анализа в нормированных пространствах – М.: Наука, 1976. – 544 с.
3. Рисс Ф. Секефальви-Надь Б. Лекции по функциональному анализу – М.: Мир, 1979. – 587 с.
4. Красносельский М. А., Вайникко Г.М., Забрейко П.П., Рунцицкий Я.Б., Стеценко В.Я. Приближенное решение операторных уравнений – М.: Наука, 1969. – 455 с.
5. Bramble J. Osborn J. Rates of the convergence estimates for nonselfadjoint eigenvalue approximations // Math. Comp. – 1973. – V. 27. – № 123. – P. 525 – 549.
6. Osborn J.E. Spectral approximations for compact operators // Math. Comp. – 1975. – V. 29. – № 131. – P. 712 – 725.
7. Vainikko G. Über die Konvergenz und Divergenz von Näherungsmethoden bei Eigenwertproblemen // Math. Nachr. – 1977. – V. 78. – P. 145 – 164.
8. Descloux J., Nassif N., Rappaz J. On spectral approximation. Part 1. The problem of convergence // R.A.I.R.O. Numer. Anal. – 1978. – V. 12. – P. 97 – 112.
9. Descloux J., Nassif N., Rappaz J. On spectral approximation. Part 2. Error estimates for the Galerkin method // R.A.I.R.O. Numer. Anal. – 1978. – V. 12. – P. 113 – 119.
10. Chatelin F. Spectral Approximation of Linear Operators – New York: Academic Press, 1983.

11. Babushka I., Osborn J. Finite element Galerkin approximation of the eigenvalues and eigenvectors of selfadjoint problems // *Math. Comp.* – 1989. – V. 52. – P. 275 – 297.
12. Banerjee U., Osborn J. Estimation of the effect of numerical integration in finite element eigenvalue approximation // *Numer. Math.* – 1990. – V. 56. P. 735 – 762.
13. Lebaud M.P. Error estimates in an isoparametric finite element eigenvalue problem // *Math. Comp.* – 1994. – V. 63. – № 207. – P. 19 – 40.
14. Vanmaele M., Zenishek A. The combined effect of numerical integration and approximation of the boundary in the finite element method for eigenvalue problems // *Numer. Math.* – 1995. – V. 71. – P. 253 – 273.
15. Соловьев С.И. Погрешность метода Бубнова-Галеркина с возмущениями для симметричных спектральных задач с нелинейным вхождением спектрального параметра // *Ж. вычисл. матем. и матем. физ.* – 1992. – Т. 32. – № 5. – С. 675 – 691.
16. Соловьев С.И. Аппроксимация симметричных спектральных задач с нелинейным вхождением параметра // *Изв. Вузов. Математика.* – 1993. – № 10. – С. 60 – 68.
17. Соловьев С.И. Оценки погрешности метода конечных элементов для симметричных спектральных задач с нелинейным вхождением параметра / С.И. Соловьев // *Изв. Вузов. Математика.* – 1994. – № 9. – С. 70 – 77.
18. Снайдер А., Лав Дж. Теория оптических волноводов – М.: Радио и связь, 1987. – 656 с.
19. Лионс Ж.-Л., Мадженес Э. Неоднородные граничные задачи и их приложения – М.: Мир, 1971. – 371 с.
20. Янке Е., Эмде Ф., Леш Ф. Специальные функции – М.: Наука, 1968. – 344 с.

21. Абрамовиц Е. Справочник по специальным функциям / Е. Абрамовиц, Стиган. – М.: Мир, 1968. – 344 с.
22. Ватсон Дж. Н. Теория бесселевых функций – М.: Изд-во иностр. лит., 1949.
23. Givoli D., Keller J.B. Exact non-reflecting boundary conditions // J. Comput. Phys. – 1989. – V. 82. – P. 172 – 192.
24. Givoli D. Nonreflecting boundary conditions (review article) // J. Comput. Phys. – 1991. – V. 94. – P. 1 – 29.
25. Ben-Porat G., Givoli D. Solution of unbounded domain problems using elliptic artificial boundaries // Comm. Num. Meth. in Eng. – 1995. – V. 11. – P. 735 – 741.
26. Adams R.A. Sobolev spaces – New York, San Francisco, London: Academic Press, 1975. – 269 с.
27. Павлова М.Ф., Тимербаев М.Р. Пространства Соболева. – Казань, Изд-во КГУ, 2002, 120 с.
28. Темам Р. Математические задачи теории пластичности – М.: Наука, 1991. – 288 с.
29. Като Т. Теория возмущений линейных операторов / Т. Като. – М.: Мир, 1972. – 740 с.
30. Корнеев В.Г. Схемы метода конечных элементов высоких порядков точности – Ленинград.: Из-во Ленинградского университета, 1977. – 206 с.
31. Zlamal M. Curved elements in finite element method. I. // SIAM J. Numer. Anal. – 1973. – V. 10. – № 1. – P. 229 – 240.
32. Zlamal M. Curved elements in finite element method. II. // SIAM J. Numer. Anal. – 1974. – V. 11. – № 2. – P. 347 – 362.
33. Даутов Р.З., Карчевский М.М. Введение в теорию метода конечных элементов – Казань: Изд-во КГУ, 2004. – 239 с.
34. Сьярле Ф. Метод конечных элементов для эллиптических задач – М.: Мир, 1980. – 512 с.

-
35. Bramble J., Hilbert S. Bounds for the class of linear functionals with applications to Hermite interpolation // Numer. Math. – 1971. – V. 16. – № 4. – P. 362 – 369.
 36. Карпенко В.А., Столяров Ю.Д., Холмогоров В.Ф. Теоретические и экспериментальные исследования прямоугольного диэлектрического волновода // Радиотехника и электроника. – 1980. – Т. 25. – № 1. – С. 51 – 57.