

**Р.З. Даутов, М.М. Карчевский**

**ВВЕДЕНИЕ В ТЕОРИЮ  
МЕТОДА КОНЕЧНЫХ ЭЛЕМЕНТОВ**

**ИЗДАНИЕ ВТОРОЕ, ИСПРАВЛЕННОЕ**

Допущено учебно-методическим советом по прикладной математике и информатике  
Учебно-методического объединения по классическому университетскому образованию  
в качестве учебного пособия для студентов высших учебных заведений,  
обучающихся по специальности «Прикладная математика и информатика» и по  
направлению «Прикладная математика и информатика»

Казань  
Казанский федеральный университет  
2011

УДК 519.3  
ББК 22.172  
Д 21

Печатается по постановлению  
Редакционно-издательского совета  
Казанского государственного университета

Научный редактор  
доктор физико-математических наук, профессор А.В. Лапин

Рецензенты:

доктор физико-математических наук, профессор В.Н. Паймушин,  
доктор физико-математических наук А.Б. Мазо

**Даутов Р.З., Карчевский М.М.**

**Д 21** Введение в теорию метода конечных элементов. Учебное пособие. — Казань: Казанский государственный университет им. В.И. Ульянова–Ленина, 2004. — 239 с.

**ISBN 5–9810–103–4**

Излагаются основные способы построения и исследования схем метода конечных элементов приближенного решения краевых задач для линейных эллиптических уравнений второго и четвертого порядков.

Книга рассчитана на студентов старших курсов и аспирантов, специализирующихся в области методов численного решения задач математической физики, а также на научных сотрудников, чьи интересы лежат в указанной области.

УДК 519.3  
ББК 22.172

**ISBN 5–9810–103–4**

© Даутов Р.З.  
Карчевский М.М., 2004

---

---

## Оглавление

Предисловие . . . . .	4
ГЛАВА 1. Элементарное введение в метод конечных элементов . . .	7
ГЛАВА 2. Вспомогательные результаты . . . . .	20
§ 1. Уравнения в пространстве Гильберта . . . . .	20
§ 2. Элементы теории пространств Соболева . . . . .	28
§ 3. Обобщенные решения краевых задач . . . . .	59
ГЛАВА 3. Пространства конечных элементов . . . . .	75
§ 1. Интерполяция в $R^n$ . . . . .	77
§ 2. Конечный элемент . . . . .	83
§ 3. Ассоциированные конечные элементы . . . . .	100
§ 4. Аффинно-эквивалентные конечные элементы . . . . .	107
§ 5. Криволинейные конечные элементы . . . . .	113
§ 6. Изопараметрические конечные элементы . . . . .	119
§ 7. Пространства конечных элементов . . . . .	131
§ 8. Примеры пространств конечных элементов . . . . .	138
§ 9. Аппроксимация подпространств $H^k(\Omega)$ , $k = 1, 2$ . . . . .	142
§ 10. Оценки погрешности интерполяции . . . . .	147
ГЛАВА 4. Метод конечных элементов для эллиптических уравнений	157
§ 1. Схемы с точной аппроксимацией границы для уравнений второго порядка . . . . .	157
§ 2. Схемы с точной аппроксимацией границы для уравнений четвер- того порядка . . . . .	164
§ 3. Дополнительные замечания . . . . .	167
§ 4. Численное интегрирование в методе конечных элементов. Общие построения . . . . .	176
§ 5. Схемы с численным интегрированием для многоугольных областей	180
§ 6. Схемы с численным интегрированием для областей с произвольной границей . . . . .	200
Приложение. Алгоритмические аспекты метода конечных элемен- тов . . . . .	215
§ 1. Форма данных для представления триангуляции в ЭВМ . . . . .	216
§ 2. Алгоритм сборки системы уравнений . . . . .	219
Литература . . . . .	229
Основные обозначения . . . . .	231
Предметный указатель . . . . .	235

---

---

## Предисловие ко второму изданию

Во втором издании исправлены неточности и опечатки, замеченные в первом издании.

Мы благодарны всем читателям приславшим свои отклики. Особую признательность выражаем В.Б. Андрееву, С.И. Соловьеву, М.Р. Тимербаеву, указавшим ряд неточностей и недостатков, которые мы постарались устранить.

## Предисловие к первому изданию

Книга является расширенным и переработанным изложением курсов лекций, читавшихся авторами для студентов факультета вычислительной математики и кибернетики Казанского федерального университета, и может рассматриваться как введение в современную теорию метода конечных элементов.

При отборе материала авторы стремились дать представление о наиболее устоявшихся и употребительных способах построения и исследования схем метода конечных элементов для типичных задач математической физики.

Вкратце содержание книги таково.

Первая глава носит подготовительный характер. Она может служить как быстрое элементарное введение в предмет. В ней на примерах простейших краевых задач демонстрируется основная идея построения приближенных решений по методу конечных элементов, дается его сопоставление с разностным методом.

Современная теория метода конечных элементов широко использует концепцию обобщенных решений краевых задач, причем многие из наиболее употребительных схем метода конечных элементов могут рассматриваться как варианты метода Галеркина (или метода Рунта) при специальном выборе подпространств, в которых разыскивается приближение к обобщенному решению.

В связи с этим во второй главе книги излагаются необходимые сведения из теории линейных уравнений в пространстве Гильберта, элементы теории пространств Соболева, а также основные способы

исследования обобщенной разрешимости краевых задач для эллиптических уравнений. Рассматриваются методы Рунге и Галеркина для приближенного решения этих задач.

Третья глава посвящена построению специальных конечномерных пространств, используемых в дальнейшем для определения приближенных решений. Рассмотрены пространства, основанные на лагранжевых и эрмитовых элементах, для многоугольных и криволинейных областей. В последнем случае изучаются криволинейные элементы с точной аппроксимацией границы и изопараметрические элементы, приводящие к полиномиальной аппроксимации границы области. Получены оценки точности аппроксимации функций из соболевских пространств элементами сконструированных конечноэлементных подпространств.

В четвертой главе построены и исследованы схемы метода конечных элементов для основных граничных задач для линейных эллиптических уравнений второго и четвертого порядков.

Формирование матрицы системы линейных алгебраических уравнений метода конечных элементов сводится к вычислению интегралов от выражений, содержащих коэффициенты дифференциального уравнения и базисные функции элементов. Даже для уравнений с постоянными коэффициентами далеко не всегда эти интегралы можно вычислить точно. Как правило, приходится прибегать к приближенным, квадратурным, формулам.

Выбор подходящей квадратурной формулы, сохраняющей свойства разрешимости и оценку точности исходной схемы конечных элементов — довольно сложная задача. Ее исследованию на примере задачи Дирихле для эллиптического уравнения второго порядка посвящена значительная часть четвертой главы. Рассмотрены схемы с лагранжевыми треугольными и прямоугольными элементами в областях с полигональными границами. Для областей с криволинейными границами изучены схемы с численным интегрированием, построенные на основе изопараметрических треугольных и четырехугольных элементов.

Книга заканчивается приложением, в котором кратко излагаются алгоритмические вопросы, возникающие при формировании системы линейных алгебраических уравнений метода конечных элементов.

Таким образом, в книге с достаточной степенью подробности рассмотрены лишь так называемые конформные варианты метода конечных элементов для линейных эллиптических уравнений второго и четвертого порядков. Более сложные методы, основанные на исполь-

зовании несовместных элементов, а также смешанные методы конечных элементов не рассматриваются.

Книга рассчитана на студентов старших курсов физико-математических специальностей. Предполагается, что читатель знаком с элементами функционального анализа, теории уравнений с частными производными и численных методов.

Список литературы, приведенный в конце книги, ни в коей мере не претендует на полноту. Он содержит лишь использованные нами учебные пособия и монографии по методу конечных элементов и близким разделам численного анализа и теории уравнений с частными производными.

Многие вопросы, затронутые в книге, активно обсуждались с сотрудниками кафедры вычислительной математики Казанского государственного университета. Авторы приносят им свою искреннюю благодарность.

Рукопись книги была внимательно прочитана М.Р. Тимербаевым. Авторы с благодарностью учли его замечания.

Структура книги обсуждалась с А.В. Лапиным, который инициировал написание книги и взял на себя труд по ее редактированию. Авторы выражают ему свою искреннюю признательность.

---

---

ГЛАВА 1  
Элементарное введение в метод конечных  
элементов

Продemonстрируем основные идеи метода конечных элементов на примерах простейших краевых задач.

Рассмотрим обыкновенное дифференциальное уравнение второго порядка с граничными условиями первого рода:

$$Lu \equiv -\frac{d}{dx} \left( p \frac{du}{dx} \right) + qu = f, \quad 0 < x < 1, \quad (1)$$

$$u(0) = u(1) = 0. \quad (2)$$

Обозначим через  $C^k[0, 1]$  множество функций, непрерывных и  $k$  раз непрерывно дифференцируемых на отрезке  $[0, 1]$ . Через  $C_0^k[0, 1]$  обозначим множество функций из  $C^k[0, 1]$ , равных нулю на концах отрезка. Будем предполагать, что  $p \in C^1[0, 1]$ ,  $q, f \in C[0, 1]$ , причем

$$p(x) \geq c_0 > 0, \quad q(x) \geq 0 \quad \forall x \in [0, 1], \quad c_0 = \text{const}. \quad (3)$$

Умножим уравнение (1) на произвольную функцию  $v \in C_0^1[0, 1]$ , проинтегрируем полученное равенство по отрезку  $[0, 1]$  и используем формулу интегрирования по частям для преобразования интеграла, содержащего производные функции  $u$ . В результате получим:

$$\int_0^1 (pu'v' + quv) dx = \int_0^1 f v dx \quad \forall v \in C_0^1[0, 1]. \quad (4)$$

Соотношение (4) принято называть интегральным тождеством, соответствующим краевой задаче (1), (2). Заметим, что интегральное тождество (4) в некотором смысле эквивалентно задаче (1), (2). Действительно, если  $u$  — решение задачи (1), (2), то  $u$  удовлетворяет тождеству (4). Обратно, если  $u \in C_0^2[0, 1]$  и удовлетворяет тождеству (4), то, применяя формулу интегрирования по частям, получим:

$$\int_0^1 (Lu - f)v dx = 0 \quad \forall v \in C_0^1[0, 1],$$

откуда в силу произвольности функции  $v$  вытекает, что  $Lu = f$  на отрезке  $[0, 1]$ , т. е.  $u$  — решение задачи (1), (2).

Нетрудно видеть, что тождество (4) сохраняет смысл при гораздо более слабых предположениях относительно входящих в него функций, чем те, которые были сформулированы выше. Это используется при введении понятия обобщенного решения для задачи (1), (2). Аналогично вводится понятие обобщенного решения и для уравнений с частными производными (см. §3 гл. 2).

Тождество (4) часто используют и при построении методов приближенного решения задач вида (1), (2).

Опишем, например, метод Галеркина. Введем в рассмотрение так называемые базисные или координатные функции  $\varphi_1, \varphi_2, \dots, \varphi_n \in C_0^1[0, 1]$ . Приближенное решение задачи (1), (2) будем разыскивать в виде

$$u_n(x) = \sum_{j=1}^n c_j \varphi_j(x), \quad (5)$$

определяя коэффициенты  $c_1, c_2, \dots, c_n$  из системы линейных алгебраических уравнений

$$\int_0^1 (p u_n' \varphi_i' + q u_n \varphi_i) dx = \int_0^1 f \varphi_i dx, \quad i = 1, 2, \dots, n. \quad (6)$$

Учитывая (5), запишем эту систему более подробно:

$$\sum_{j=1}^n a_{ij} c_j = b_i, \quad i = 1, 2, \dots, n,$$

или  $Ac = b$ , где  $A = (a_{ij})_{i,j=1}^n$ ,  $b = (b_1, \dots, b_n)^T$ ,

$$a_{ij} = \int_0^1 (p \varphi_j' \varphi_i' + q \varphi_j \varphi_i) dx, \quad b_i = \int_0^1 f \varphi_i dx.$$

Для вычисления участвующих здесь интегралов, обычно, применяют квадратурные формулы.

Свойства системы (6) и построенного с ее помощью приближенного решения определяются координатной системой  $\varphi_1, \varphi_2, \dots, \varphi_n$ . Если координатная система выбрана удачно, то с увеличением  $n$  точность приближенного метода улучшается. Можно показать, что если координатная система линейно независима, то матрица  $A$  системы линейных уравнений (6) положительно определена. Отсюда, в частности, вытекает однозначная разрешимость системы метода Галеркина.

Приведем примеры координатных систем метода Галеркина:

а) полиномиальная система:

$$\varphi_1(x) = x(1-x), \quad \varphi_2(x) = x(1-x)x, \quad \dots, \quad \varphi_n(x) = x(1-x)x^{n-1};$$

б) тригонометрическая система:

$$\varphi_k(x) = \sin k\pi x, \quad k = 1, 2, \dots, n.$$

При использовании указанных, или аналогичных, координатных систем матрица метода Галеркина оказывается заполненной, т. е. все ее элементы отличны от нуля. Это принципиально отличает метод Галеркина от разностного метода, при использовании которого матрица системы линейных уравнений — разреженная матрица.

В 1943 году Р. Курантом было замечено, что при специальном выборе базисных функций метод Галеркина приводит к системам линейных уравнений, по свойствам весьма близким к разностным уравнениям. В простейших случаях метод совпадает с разностным. Впоследствии (в пятидесятых годах) метод, предложенный Курантом, был переоткрыт инженерами, существенно развит и обобщен. В настоящее время этот метод, называемый методом конечных элементов (МКЭ), принадлежит к числу наиболее распространенных способов приближенного решения дифференциальных уравнений.

Опишем простейший вариант метода конечных элементов на примере задачи (1), (2).

Построим на отрезке  $[0, 1]$  сетку

$$\omega = \{0 = x_0 < x_1 < x_2 < \dots < x_N = 1\}$$

(вообще говоря, неравномерную). Ячейки сетки  $e_i = [x_{i-1}, x_i]$ ,  $i = 1, 2, \dots, N$ , назовем конечными элементами. Положим  $h_i = x_i - x_{i-1}$ ,  $h = \max_{1 \leq i \leq N} h_i$ . Обозначим приближенное решение задачи (1), (2) через  $u_h$  и будем разыскивать его как функцию, непрерывную на отрезке  $[0, 1]$ , равную нулю при  $x = 0$ ,  $x = 1$  и линейную на каждом элементе  $e_i$ ,  $i = 1, 2, \dots, N$  (см. рис. 1). Ясно, что

$$u_h(x) = \sum_{i=1}^{N-1} u_h(x_i) \varphi_i(x), \quad x \in [0, 1],$$

где  $\varphi_i$  — линейные на каждом элементе, непрерывные на отрезке  $[0, 1]$  функции, удовлетворяющие условиям (см. рис. 1):

$$\varphi_i(x_j) = \begin{cases} 1, & i = j, \\ 0, & i \neq j, \end{cases} \quad i = 0, 1, \dots, N.$$

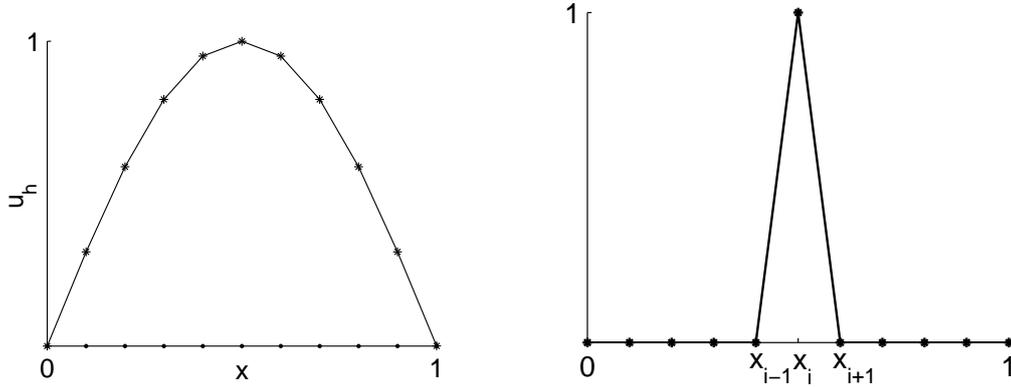


Рис. 1. Кусочно линейная функция (слева). Базисная функция  $\varphi_i$  (справа).

Функции  $\varphi_i(x)$  принято называть базисными функциями с локальным носителем, поскольку множество точек, где  $\varphi_i(x) \neq 0$  (носитель функции  $\varphi_i(x)$ ) — малый отрезок (именно, объединение элементов  $e_i$  и  $e_{i+1}$ ). Элементы  $a_{j,i}$  матрицы системы (6) в этом случае, очевидно, отличны от нуля, лишь при условии, что носители функций  $\varphi_i(x)$ ,  $\varphi_j(x)$  пересекаются, т. е., когда  $|i - j| \leq 1$ . Как следствие, матрица системы (6) трехдиагональна.

Рассмотрим теперь самый простой случай. Пусть  $p(x) \equiv 1$ ,  $q(x) \equiv 0$ , сетка равномерна, т. е.  $x_i - x_{i-1} = h$ ,  $i = 1, 2, \dots, N$ . Тогда, как нетрудно подсчитать,

$$a_{i,i-1} = -1/h, \quad a_{i,i+1} = -1/h, \quad a_{i,i} = 2/h, \quad b_i = \int_{x_{i-1}}^{x_{i+1}} f \varphi_i(x) dx,$$

$i = 1, 2, \dots, N - 1$ . Вычислим последний интеграл приближенно, полагая  $f(x) \approx f(x_i)$ ,  $x \in [x_{i-1}, x_{i+1}]$ . Тогда

$$b_i = f(x_i) \int_{x_{i-1}}^{x_{i+1}} \varphi_i(x) dx = hf(x_i),$$

и, следовательно, система (6) принимает вид

$$-\frac{u_h(x_{i-1}) - 2u_h(x_i) + u_h(x_{i+1}))}{h^2} = f(x_i), \quad i = 1, 2, \dots, N - 1.$$

Добавляя условия равенства нулю функции  $u_h$  на концах отрезка:

$$u_h(x_0) = 0, \quad u_h(x_N) = 0,$$

получим простейшую разностную схему для задачи (1), (2).

Исследуем сходимость метода конечных элементов. Предварительно получим необходимые вспомогательные результаты.

**Лемма 1.** Пусть  $u$  — непрерывная и кусочно непрерывно дифференцируемая на отрезке  $[0, 1]$  функция,  $u(0) = 0$ . Тогда

$$\|u\|_C^2 \leq \int_0^1 u'^2(x) dx,$$

где  $\|u\|_C = \max_{x \in [0,1]} |u(x)|$ .

**ДОКАЗАТЕЛЬСТВО.** По формуле Ньютона — Лейбница

$$u(x) = \int_0^x u'(\xi) d\xi \quad \forall x \in (0, 1),$$

откуда, применяя неравенство Коши — Буняковского, получим

$$u^2(x) \leq \int_0^x d\xi \int_0^x u'^2(\xi) d\xi \leq \int_0^1 u'^2(\xi) d\xi \quad \forall x \in [0, 1]. \quad \square$$

Здесь и далее символ « $\square$ » обозначает конец доказательства.

**Лемма 2.** Пусть  $u \in C_0^2[0, 1]$ ,

$$u_I(x) = \sum_{i=1}^{N-1} u(x_i) \varphi_i(x)$$

есть кусочно линейная функция, совпадающая в узлах сетки с функцией  $u$ <sup>1)</sup>. Тогда

$$\|u - u_I\|_C \leq \frac{h^2}{8} \|u''\|_C.$$

**ДОКАЗАТЕЛЬСТВО.** Пусть  $x \in (x_{i-1}, x_i)$  для некоторого  $i \in \{1, 2, \dots, N\}$ . Используя представление для остаточного члена интерполяционного полинома в форме Лагранжа, получим

$$u(x) - u_I(x) = \frac{u''(\xi)}{2} (x - x_{i-1})(x - x_i), \quad \xi \in (x_{i-1}, x_i),$$

<sup>1)</sup>Функцию  $u_I$  называют, обычно, интерполянтom функции  $u$ .

откуда вследствие очевидного неравенства  $(x - x_{i-1})(x_i - x) \leq h_i^2/4$  непосредственно вытекает утверждение леммы.  $\square$

**Теорема 1.** Пусть  $u$  — решение задачи (1), (2),  $u_h$  — приближенное решение, построенное по методу конечных элементов. Тогда

$$\|u - u_h\|_C \leq ch^2 \|u''\|_C,$$

где  $c$  — положительная постоянная, зависящая только от коэффициентов уравнения (1).

**ДОКАЗАТЕЛЬСТВО.** Обозначим через  $V_h$  линейное множество всех функций вида

$$v_h(x) = \sum_{i=1}^{N-1} v_i \varphi_i(x), \quad x \in [0, 1].$$

Покажем, что если  $u$  — решение задачи (1), (2), то

$$a(u, v_h) \equiv \int_0^1 (pu'v'_h + quv_h) dx = \int_0^1 f v_h dx \quad \forall v_h \in V_h. \quad (7)$$

Действительно, применяя формулу интегрирования по частям и учитывая, что функция  $v_h$  непрерывна и кусочно непрерывно дифференцируема на отрезке  $[0, 1]$ , а  $v_h(0) = v_h(1) = 0$ , получим, что

$$\int_0^1 pu'v'_h dx = - \int_0^1 (pu')' v_h dx,$$

откуда следует, что

$$a(u, v_h) = \int_0^1 f v_h dx.$$

Из определения функции  $u_h$  вытекает тождество

$$a(u_h, v_h) = \int_0^1 f v_h dx \quad \forall v_h \in V_h. \quad (8)$$

Используя тождества (7), (8), нетрудно получить, что

$$\begin{aligned} a(u_h - u_I, u_h - u_I) &= a(u_h, u_h - u_I) - a(u_I, u_h - u_I) = \\ &= \int_0^1 f(u_h - u_I) dx - a(u_I, u_h - u_I) = a(u - u_I, u_h - u_I). \end{aligned}$$

Поскольку функция  $(u_h - u_I)'$  постоянна на каждом элементе  $e_i$ , то, используя формулу интегрирования по частям, можно написать

$$\begin{aligned} \int_{x_{i-1}}^{x_i} p(u - u_I)'(u_h - u_I)' dx &= - \int_{x_{i-1}}^{x_i} p'(u - u_I)(u_h - u_I)' dx + \\ &+ p(u - u_I)(u_h - u_I)' \Big|_{x_{i-1}}^{x_i} = - \int_{x_{i-1}}^{x_i} p'(u - u_I)(u_h - u_I)' dx. \end{aligned}$$

Следовательно, применяя лемму 1, получим

$$\begin{aligned} a(u_h - u_I, u_h - u_I) &= \\ &= a(u - u_I, u_h - u_I) = \int_0^1 (-p'(u - u_I)(u_h - u_I)' + q(u - u_I)(u_h - u_I)) dx \leq \\ &\leq c \|u - u_I\|_C \left( \int_0^1 (u_h' - u_I')^2 dx \right)^{1/2}. \quad (9) \end{aligned}$$

Используя условия (3), напомним оценку снизу

$$a(u_h - u_I, u_h - u_I) \geq c_0 \int_0^1 (u_h' - u_I')^2 dx. \quad (10)$$

Из оценок (9), (10) и леммы 1 вытекает, очевидно, что

$$\|u_h - u_I\|_C \leq c \|u - u_I\|_C$$

и, следовательно,

$$\|u - u_h\|_C \leq \|u - u_I\|_C + \|u_h - u_I\|_C \leq c \|u - u_I\|_C.$$

Для завершения доказательства теоремы достаточно воспользоваться леммой 2.  $\square$

Таким образом, мы показали, что при  $h \rightarrow 0$  приближенное решение, построенное по методу конечных элементов, равномерно сходится к точному решению задачи (1), (2), причем со скоростью  $O(h^2)$ .

Аналогичный подход к построению приближенных решений можно использовать и для уравнений с частными производными. Продемонстрируем это на примере задачи Дирихле для уравнения Пуассона:

$$-\Delta u = f, \quad x \in \Omega, \quad (11)$$

$$u(x) = 0, \quad x \in \Gamma, \quad (12)$$

где

$$u = u(x) = u(x_1, x_2), \quad \Delta u = \frac{\partial^2 u}{\partial x_1^2} + \frac{\partial^2 u}{\partial x_2^2}.$$

Будем предполагать, что область  $\Omega$  — единичный квадрат:

$$\Omega = \{x = (x_1, x_2) : x_1, x_2 \in (0, 1)\},$$

$\Gamma$  — граница  $\Omega$ .

Метод Галеркина основан на формулировке задачи (11), (12) в виде интегрального тождества.

Умножим уравнение (11) на произвольную непрерывную, кусочно непрерывно дифференцируемую и равную нулю на  $\Gamma$  функцию  $v$  и проинтегрируем полученное равенство по области  $\Omega$ . После применения формулы интегрирования по частям будем иметь

$$\int_{\Omega} \left( \frac{\partial u}{\partial x_1} \frac{\partial v}{\partial x_1} + \frac{\partial u}{\partial x_2} \frac{\partial v}{\partial x_2} \right) dx = \int_{\Omega} f v dx.$$

Приближенное решение разыскивается в виде

$$u_N = \sum_{k=1}^N c_j \varphi_j(x),$$

где  $\varphi_j(x) = \varphi_j(x_1, x_2)$  — некоторые заданные (базисные) функции, удовлетворяющие граничному условию (12), а коэффициенты  $c_j$  подлежат определению. Их находят, решая систему линейных алгебраических уравнений:

$$\int_{\Omega} \left( \frac{\partial u_N}{\partial x_1} \frac{\partial \varphi_i}{\partial x_1} + \frac{\partial u_N}{\partial x_2} \frac{\partial \varphi_i}{\partial x_2} \right) dx = \int_{\Omega} f \varphi_i dx, \quad i = 1, 2, \dots, N.$$

Классические примеры выбора базисных функций:

$$1) \quad \varphi_{kl}(x_1, x_2) = x_1(1-x_1)x_2(1-x_2)x_1^k x_2^l, \quad k, l = 0, 1, \dots, n-1$$

(первые множители здесь обеспечивают выполнение граничных условий),

$$2) \quad \varphi_{kl}(x_1, x_2) = \sin k\pi x_1 \sin l\pi x_2, \quad k, l = 1, 2, \dots, n.$$

Отметим, что для удобства принята двойная нумерация,  $N = n^2$ .

Метод конечных элементов основан на использовании специальных базисных функций. При этом матрица системы метода Галеркина оказывается разреженной, т. е. большинство ее элементов — нули. Приведем простейший пример таких построений.

Введем на области  $\Omega$  квадратную сетку с шагом  $h$ ,  $h = 1/n$ :

$$\omega = \{x_{kl} = (kh, lh), k, l = 0, 1, \dots, n\}.$$

Каждую ячейку сетки разделим на два треугольника диагональю,

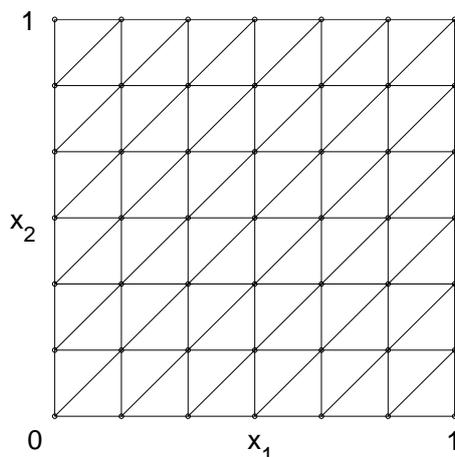


Рис. 2. Триангуляция прямоугольной области  $\Omega$ .

параллельной биссектрисе первого координатного угла. Получим разбиение области  $\Omega$  на треугольные элементы  $e_i$ ,  $i = 1, 2, \dots, 2n^2$ , т. е. триангуляцию области (см. рис. 2). Приближенное решение задачи

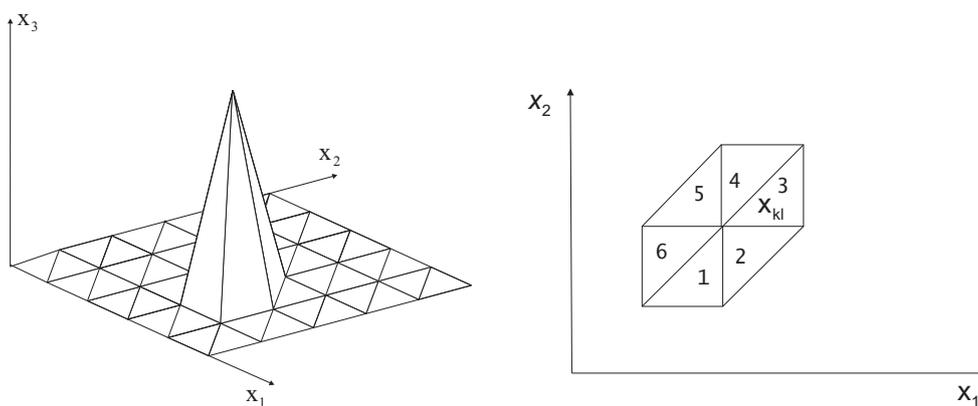


Рис. 3. Базисная функция Куранта (слева). Область  $\Omega_{kl}$ ; нумерация элементов (справа).

будем искать как функцию  $u^h$ , непрерывную в области  $\Omega$ , равную нулю на  $\Gamma$  и линейную на каждом элементе  $e_i$ . Ясно, что функцию  $u^h$

$n$	$\partial\varphi_{kl}/\partial x_1$	$\partial\varphi_{kl}/\partial x_2$
1	0	$1/h$
2	$-1/h$	$1/h$
3	$-1/h$	0
4	0	$-1/h$
5	$1/h$	$-1/h$
6	$1/h$	0

Таблица 1. Производные функции  $\varphi_{kl}$ .

можно однозначно представить в виде

$$u^h(x) = \sum_{k,l=1}^{n-1} u_{kl}^h \varphi_{kl}(x),$$

где  $u_{kl}^h = u^h(x_{kl})$ ,  $\varphi_{kl}(x)$  — функция, непрерывная на  $\Omega$ , равная нулю на  $\Gamma$ , линейная на каждом элементе и удовлетворяющая условиям:

$$\varphi_{kl}(x_{ij}) = \begin{cases} 0, & x_{ij} \neq x_{kl}, \\ 1 & x_{ij} = x_{kl}, \quad i, j = 0, 1, \dots, n. \end{cases}$$

Функция  $\varphi_{kl}$  («шпалочка Куранта»), как нетрудно проверить, тождественно равна нулю вне области  $\Omega_{kl}$ , представляющей собой объединение элементов, имеющих вершиной точку  $x_{kl}$  (см. рис. 3).

Построим уравнения для определения значений  $u_{kl}^h$ . Для этого подсчитаем коэффициенты и правую часть в системе метода Галеркина, а именно,

$$a_{klk'l'} = \int_{\Omega} \left( \frac{\partial\varphi_{kl}}{\partial x_1} \frac{\partial\varphi_{k'l'}}{\partial x_1} + \frac{\partial\varphi_{kl}}{\partial x_2} \frac{\partial\varphi_{k'l'}}{\partial x_2} \right) dx, \quad b_{kl} = \int_{\Omega} f \varphi_{kl} dx.$$

Понятно, что  $a_{klk'l'}$  может быть не нулем только в том случае, когда пересечение областей  $\Omega_{kl}$ ,  $\Omega_{k'l'}$  не пусто. Для фиксированных  $k, l$  таких областей  $\Omega_{k'l'}$  шесть, а именно:  $\Omega_{k-1,l}$ ,  $\Omega_{k-1,l-1}$ ,  $\Omega_{k,l-1}$ ,  $\Omega_{k+1,l}$ ,  $\Omega_{k+1,l+1}$ ,  $\Omega_{k,l+1}$ . При вычислении коэффициентов  $a_{klk'l'}$  потребуются значения производных функции  $\varphi_{kl}$ . Пронумеруем элементы, принадлежащие  $\Omega_{kl}$ , как показано на рис. 3. Сведем результаты очевидных вычислений в таблицу 1 ( $n$  — номер треугольника). Теперь ясно, что

$$a_{k,l,k-1,l-1} = \int_{e_6 \cup e_1} \left( \frac{\partial\varphi_{kl}}{\partial x_1} \frac{\partial\varphi_{k-1,l-1}}{\partial x_1} + \frac{\partial\varphi_{kl}}{\partial x_2} \frac{\partial\varphi_{k-1,l-1}}{\partial x_2} \right) dx = 0,$$

$$\begin{aligned} a_{k,l,k+1,l+1} &= 0, \quad a_{k,l,k,l-1} = -1, \quad a_{k,l,k-1,l} = -1, \\ a_{k,l,k,l+1} &= -1, \quad a_{k,l,k+1,l} = -1, \quad a_{k,l,k,l} = 4. \end{aligned}$$

Это означает, что уравнение с номером  $k, l$  системы метода Галеркина принимает вид

$$\frac{-u_{k-1,l}^h - u_{k+1,l}^h + 4u_{k,l}^h - u_{k,l-1}^h - u_{k,l+1}^h}{h^2} = b_{k,l}, \quad (13)$$

где

$$b_{k,l} = \frac{1}{h^2} \int_{\Omega_{kl}} f(x) \varphi_{kl}(x) dx.$$

Заметим, что при достаточно малом  $h$

$$\int_{\Omega_{kl}} f(x) \varphi_{kl}(x) dx \approx f(x_{kl}) \int_{\Omega_{kl}} \varphi_{kl}(x) dx = f(x_{kl}) h^2.$$

Таким образом, уравнение (13) приближенно представляется в виде

$$\frac{-u_{k-1,l}^h - u_{k+1,l}^h + 4u_{k,l}^h - u_{k,l-1}^h - u_{k,l+1}^h}{h^2} = f(x_{kl}).$$

Точно такое же уравнение получается при замене в уравнении (11) производных разделенными разностями:

$$\begin{aligned} \frac{\partial^2 u}{\partial x_1^2} &\approx \frac{u(x_{k-1,l}) - 2u(x_{k,l}) + u(x_{k+1,l}))}{h^2}, \\ \frac{\partial^2 u}{\partial x_2^2} &\approx \frac{u(x_{k,l-1}) - 2u(x_{k,l}) + u(x_{k,l+1}))}{h^2}. \end{aligned}$$

Записывая уравнения (13) во внутренних точках сетки, т. е. при  $k, l = 1, 2, \dots, n-1$ , и присоединяя к ним граничные условия, соответствующие (12):

$$\begin{aligned} u_{0,l}^h &= 0, \quad u_{n,l}^h = 0, \quad l = 0, 1, \dots, n, \\ u_{k,0}^h &= 0, \quad u_{k,n}^h = 0, \quad k = 0, 1, \dots, n, \end{aligned}$$

получим полную систему линейных алгебраических уравнений для отыскания приближенного решения в точках сетки.

Аналогично строят систему метода конечных уравнений для произвольной области: сначала область аппроксимируют многоугольником, затем производят триангуляцию, т. е. разбиение многоугольника на достаточно малые треугольники, далее определяют базисные

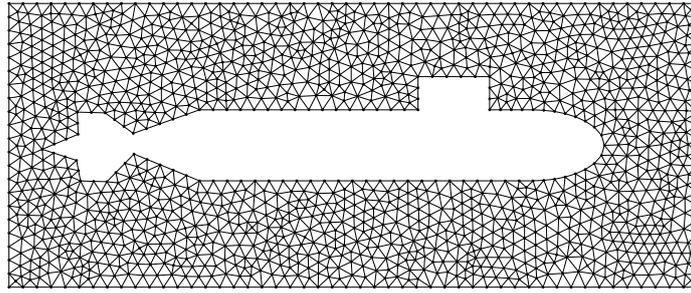


Рис. 4. Триангуляция сложной области.

функции  $\varphi_{kl}$ , при этом, конечно, области  $\Omega_{kl}$  могут быть достаточно сложными (см. рис. 4).

Описанные способы построения приближенного решения допускают очевидные обобщения. Например, в одномерном случае функцию  $u_h$  можно искать в классе функций, являющихся полиномами некоторой фиксированной степени  $m \geq 1$  на каждом элементе. Можно повысить гладкость приближенного решения, считая непрерывными и его производные вплоть до определенного порядка. Аналогичные обобщения можно применять и для уравнений с частными производными. При этом можно разбивать область не только на треугольные элементы, но и на элементы более сложной формы, например, четырехугольные и даже криволинейные.

Понятно, что указанные построения вызывают целый ряд вопросов. Прежде всего, интересно было бы оценить точность возникающих приближенных решений в зависимости от размеров и формы применяемых конечных элементов, от степени гладкости приближенного решения, вида полиномов, применяемых при построении приближенного решения, а также дифференциальных свойств точного решения исходной задачи. Интегралы, возникающие при формировании системы метода конечных элементов, вообще говоря, могут быть вычислены только приближенно при помощи некоторых квадратурных формул. Получающиеся при этом погрешности также должны быть учтены. Все эти вопросы с той или иной степенью подробности исследуются в настоящей книге применительно к краевым задачам для линейных эллиптических уравнений.

С точки зрения практического использования метода конечных элементов существенную роль играет выбор эффективного прямого или итерационного метода решения соответствующей системы линейных алгебраических уравнений, по возможности, достаточно полно учитывающего ее специфику, например, разреженную структуру матрицы системы. Сколько-нибудь подробное рассмотрение этих вопро-

сов, составляющих важный раздел современных численных методов линейной алгебры (см., например, [17], [4], [16]), заняло бы слишком много места, и мы не сочли возможным включать их в настоящую книгу.

---

---

ГЛАВА 2  
ВСПОМОГАТЕЛЬНЫЕ РЕЗУЛЬТАТЫ

§ 1. Уравнения в пространстве Гильберта

**1. Уравнение с ограниченным положительно определенным оператором.** Пусть  $H$  — вещественное пространство Гильберта со скалярным произведением  $(\cdot, \cdot)$  и нормой  $\|\cdot\| = (\cdot, \cdot)^{1/2}$ . Обозначим через  $a(\cdot, \cdot): H \times H \rightarrow R$  непрерывную билинейную форму, т. е. форму, линейную по каждому аргументу и удовлетворяющую условию

$$|a(u, v)| \leq M\|u\|\|v\| \quad \forall u, v \in H, \quad (1)$$

где  $M = \text{const} > 0$ . Через  $f: H \rightarrow R$  обозначим непрерывную линейную форму (линейный ограниченный функционал). Вследствие теоремы Рисса об общем виде линейного ограниченного функционала форма  $a(\cdot, \cdot)$  порождает линейный ограниченный оператор  $A$ , определяемый соотношением

$$(Au, v) = a(u, v) \quad \forall u, v \in H.$$

Функционалу  $f$  будем сопоставлять элемент  $f \in H$ , определяемый по теореме Рисса (единственным образом) тождеством

$$f(v) = (f, v) \quad \forall v \in H.$$

В дальнейшем будем рассматривать уравнение

$$Au = f, \quad (2)$$

где  $f \in H$  — заданный элемент. Эквивалентная формулировка: найти такой элемент  $u \in H$ , что

$$a(u, v) = f(v) \quad \forall v \in H.$$

Форма  $a(\cdot, \cdot)$  называется положительно определенной (эллиптической, коэрцитивной), если существует такая положительная постоянная  $m$ , что

$$a(u, u) \geq m\|u\|^2 \quad \forall u \in H. \quad (3)$$

Соответствующий оператор  $A$  также будем называть положительно определенным.

**Лемма 3 (Лакса — Мильграма).** Пусть оператор  $A$  ограничен и положительно определен. Тогда задача (2) при любом  $f \in H$  имеет единственное решение.

**ДОКАЗАТЕЛЬСТВО.** При любом  $\tau \neq 0$  задача (2) эквивалентна уравнению

$$u = u - \tau(Au - f).$$

Покажем, что при достаточно малом положительном  $\tau$  оператор  $Su = u - \tau(Au - f)$  — оператор сжатия. Тогда утверждение теоремы будет вытекать из теоремы Банаха о сжимающих отображениях. Пусть  $u, v$  — произвольные элементы пространства  $H$ . Имеем:

$$\begin{aligned} \|Su - Sv\|^2 &= \|u - v - \tau(Au - Av)\|^2 = \\ &= \|u - v\|^2 - 2\tau(Au - Av, u - v) + \tau^2\|Au - Av\|^2, \end{aligned}$$

откуда в силу условий (1), (3) получаем, что

$$\|Su - Sv\|^2 \leq \rho(\tau)\|u - v\|^2,$$

где  $\rho(\tau) = 1 - 2\tau m + \tau^2 M^2$ . Ясно, что  $\rho(\tau) < 1$  при  $\tau \in (0, 2m/M^2)$ , и, следовательно,  $S$  — оператор сжатия.  $\square$

**ЗАМЕЧАНИЕ 1.** Из доказательства леммы Лакса — Мильграма вытекает, что решение задачи (2) может быть построено при помощи итерационного метода

$$u^{k+1} = u^k - \tau(Au^k - f), \quad k = 0, 1, 2, \dots,$$

сходящегося, если  $\tau \in (0, 2m/M^2)$ , при любом начальном приближении  $u^0 \in H$ .

## 2. Уравнение с самосопряженным положительно определенным оператором.

**Определение 1.** Ограниченная билинейная форма называется симметричной, если

$$a(u, v) = a(v, u) \quad \forall u, v \in H.$$

Линейный ограниченный оператор  $A$ , порожденный симметричной формой, называется самосопряженным.

Ясно, что в этом случае

$$(Au, v) = (Av, u) \quad \forall u, v \in H.$$

С уравнением (2) свяжем задачу минимизации квадратичного функционала: найти такой элемент  $u \in H$ , что

$$F(u) = \min_{v \in H} F(v), \quad F(v) = \frac{1}{2}a(v, v) - f(v). \quad (4)$$

**Теорема 1.** Пусть оператор  $A$  самосопряжен и положительно определен. Тогда задачи (2), (4) эквивалентны.

**ДОКАЗАТЕЛЬСТВО.** В силу самосопряженности и положительной определенности оператора  $A$  на пространстве  $H$  можно ввести новое скалярное произведение  $(\cdot, \cdot)_a$  при помощи соотношения:  $(u, v)_a = a(u, v)$ . Скалярное произведение  $(\cdot, \cdot)_a$  называют энергетическим скалярным произведением, соответствующим оператору  $A$ . Вследствие условий (1), (3) норма, порождаемая энергетическим скалярным произведением (энергетическая норма  $\|\cdot\|_a$ ), эквивалентна<sup>1)</sup> исходной норме пространства  $H$ . Пусть  $u$  — решение задачи (2). Тогда функционал  $F$  можно представить в виде

$$F(v) = \frac{1}{2}(v, v)_a - (u, v)_a.$$

После элементарных преобразований получим

$$F(v) = F_0(v) - \frac{1}{2}\|u\|_a^2,$$

где  $F_0(v) = \frac{1}{2}\|u-v\|_a^2$ . Таким образом, задача (4) эквивалентна задаче

$$F_0(u) = \min_{v \in H} F_0(v) \quad (5)$$

и, очевидно, имеет единственное решение, совпадающее с решением задачи (2).  $\square$

**3. Метод Ритца.** Пусть  $H_n$  — замкнутое подпространство пространства  $H$ .

**Определение 2.** Элемент  $u_n \in H_n$  называется приближенным решением задачи (2) по методу Ритца, если

$$F(u_n) = \min_{v \in H_n} F(v).$$

**Теорема 2.** Пусть выполнены условия (1), (3). Тогда приближенное решение по методу Ритца существует и определяется единственным образом.

**ДОКАЗАТЕЛЬСТВО.** В силу эквивалентности задач (4), (5) приближенное решение по методу Ритца можно интерпретировать как элемент  $u_n \in H_n$  такой, что

$$\|u - u_n\|_a = \inf_{v_n \in H_n} \|u - v_n\|_a.$$

<sup>1)</sup>Нормы  $\|\cdot\|_1$  и  $\|\cdot\|_2$  называются эквивалентными на линейном нормированном пространстве  $X$ , если существуют постоянные  $c_0, c_1 > 0$  такие, что  $c_0\|u\|_1 \leq \|u\|_2 \leq c_1\|u\|_1 \forall u \in X$ .

Иными словами,  $u_n$  — элемент наилучшего приближения к  $u$  (в смысле энергетической нормы) и потому, как известно, он существует и определяется единственным образом.  $\square$

ЗАМЕЧАНИЕ 2. Если  $u_n$  — приближенное решение по методу Ритца, то, очевидно,

$$\begin{aligned} \|u - u_n\|_a &\leq \|u - v_n\|_a \quad \forall v_n \in H_n, \\ \|u - u_n\| &\leq \sqrt{M/m} \|u - v_n\| \quad \forall v_n \in H_n. \end{aligned} \quad (6)$$

Эти неравенства часто используются при получении оценок погрешности метода Ритца.

**Следствие 1.** Пусть семейство подпространств  $H_n$ ,  $n = 1, 2, \dots$  предельно полно, т. е.

$$\delta_n(v) = \inf_{v_n \in H_n} \|v - v_n\| \rightarrow 0 \quad \text{при } n \rightarrow \infty \quad \forall v \in H.$$

Тогда последовательность  $u_n$  приближений по методу Ритца сходится к  $u$ .

Доказательство сразу вытекает из неравенства

$$\|u - u_n\| \leq \sqrt{\frac{M}{m}} \delta_n(u),$$

являющегося очевидным следствием (6).  $\square$

**Следствие 2.** Пусть  $u_n$  — приближение по методу Ритца. Тогда

$$a(u_n, v_n) = f(v_n) \quad \forall v_n \in H_n. \quad (7)$$

ДОКАЗАТЕЛЬСТВО. Поскольку  $u_n$  — приближение по методу Ритца, то  $F(u_n + tv_n) \geq F(u_n)$  для любого  $v_n \in H_n$  и любого вещественного числа  $t$ . Нетрудно подсчитать, что

$$F(u_n + tv_n) = F(u_n) + t(a(u_n, v_n) - (f, v_n)) + \frac{t^2}{2}a(v, v),$$

следовательно, квадратный трехчлен  $t(a(u_n, v_n) - (f, v_n)) + \frac{t^2}{2}a(v, v)$  неотрицателен при любом  $t$ , что, очевидно, возможно лишь при  $a(u_n, v_n) - (f, v_n) = 0$ .  $\square$

**Следствие 3.** Пусть  $u_n$  — приближенное решение по методу Ритца,  $u$  — решение задачи (2). Тогда

$$(u - u_n, v_n)_a = 0 \quad \forall v_n \in H_n. \quad (8)$$

Справедливость (8) немедленно следует из (7) и очевидного тождества:  $a(u, v_n) = (f, v_n) \quad \forall v_n \in H_n$ .

В приложениях чаще всего пространство  $H_n$  конечномерно. В этом случае построение приближенного решения по методу Рунге сводится к решению системы линейных алгебраических уравнений. Действительно, пусть  $\varphi_1, \varphi_2, \dots, \varphi_n$  — какой-либо базис подпространства  $H_n$ . Представим  $u_n$  в виде  $u_n = \sum_{j=1}^n c_j \varphi_j$ , где  $c_1, c_2, \dots, c_n$  — коэффициенты, подлежащие определению. Вследствие (7)

$$a(u_n, \varphi_i) = (f, \varphi_i), \quad i = 1, \dots, n,$$

или

$$A_n c = f_n, \quad (9)$$

где

$$A_n = \{a_{ij} = a(\varphi_j, \varphi_i)\}_{i,j=1}^n, \quad c = \{c_k\}_{k=1}^n, \quad f_n = \{(f, \varphi_k)\}_{k=1}^n.$$

Матрица  $A_n$  называется матрицей Грама системы элементов  $\varphi_1, \dots, \varphi_n$ .

Матрица Грама симметрична и положительно определена. Более того, справедлива

**Лемма 4.** *Для того, чтобы система элементов  $\varphi_1, \varphi_2, \dots, \varphi_n$  была линейно независимой, необходимо и достаточно, чтобы матрица Грама была положительно определена.*

**ДОКАЗАТЕЛЬСТВО.** Для любых  $c_1, c_2, \dots, c_n$  имеем<sup>1)</sup>

$$A_n c \cdot c = \sum_{i,j=1}^n a_{ij} c_j c_i = a \left( \sum_{j=1}^n c_j \varphi_j, \sum_{i=1}^n c_i \varphi_i \right) = \left\| \sum_{i=1}^n c_i \varphi_i \right\|_a^2.$$

Если  $\varphi_1, \varphi_2, \dots, \varphi_n$  линейно независимы, то  $A_n c \cdot c \geq 0$  и обращается в нуль лишь при  $c_1 = c_2 = \dots = c_n = 0$ . Очевидно, справедливо и обратное.  $\square$

**ЗАМЕЧАНИЕ 3.** Выбор базиса в подпространстве  $H_n$  существенно сказывается на обусловленности матрицы  $A_n$  системы (9). Идеальным является случай, когда функции  $\varphi_k$ ,  $k = 1, 2, \dots, n$ , ортонормированы в смысле энергетического скалярного произведения. Тогда  $A_n$  — единичная матрица. Вообще же говоря, число обусловленности матрицы  $A_n$  может сильно возрастать с ростом  $n$ . Один из возможных способов избежать этого состоит в том, что в качестве базисных функций в методе Рунге выбирают ортонормированные собственные функции некоторого простого самосопряженного положительно определенного оператора  $B_n$ , энергетически эквивалентного оператору  $A_n$ , т. е. такого самосопряженного оператора  $B_n$ , что нормы  $\|\cdot\|_{a_n}$ ,  $\|\cdot\|_{b_n}$  эквивалентны с постоянными эквивалентности, не зависящими от  $n$  (подробнее см. в [12]).

<sup>1)</sup>  $x \cdot y = x_1 y_1 + \dots + x_n y_n$  — скалярное произведение векторов в  $R^n$ .

**4. Метод Галеркина.** Пусть теперь выполнены условия ограниченности и положительной определенности оператора  $A$ . Условия самосопряженности могут не выполняться. Определение приближения по методу Ритца для уравнения (2) при помощи тождества (7) позволяет сформулировать понятие приближенного решения и в этой, более общей, ситуации.

**Определение 3.** Будем говорить, что  $u_n \in H_n$  — приближенное решение уравнения (2) по методу Галеркина, если для любого  $v_n \in H_n$  выполнено равенство

$$a(u_n, v_n) = f(v_n). \quad (10)$$

Таким образом, метод Галеркина можно рассматривать как непосредственное обобщение метода Ритца. Понятно, что формально приближенное решение по методу Галеркина можно определить для любого уравнения в гильбертовом пространстве, не обязательно — линейного.

**Теорема 3.** Пусть выполнены условия (1), (3). Тогда приближенное решение по методу Галеркина существует и определяется единственным образом.

**ДОКАЗАТЕЛЬСТВО.** Используя теорему Рисса о представлении линейного ограниченного функционала, получим, что задача (10) эквивалентна уравнению

$$A_n u_n = f_n, \quad (11)$$

где  $A_n : H_n \rightarrow H_n$  и  $f_n \in H_n$  определяются соотношениями

$$(A_n u_n, v_n) = a(u_n, v_n) \quad \forall u_n, v_n \in H_n, \quad (f_n, v_n) = f(v_n) \quad \forall v_n \in H_n.$$

Оператор  $A_n : H_n \rightarrow H_n$  ограничен и положительно определен, поэтому существование и единственность решения уравнения (11) следуют из леммы Лакса — Мильграма.  $\square$

Точность метода Галеркина характеризует

**Лемма 5 (лемма Сеа).** Пусть  $u$  — решение уравнения (2),  $u_n$  — приближенное решение по методу Галеркина, выполнены условия (1), (3). Тогда

$$\|u - u_n\| \leq \frac{M}{m} \|u - v_n\| \quad \forall v_n \in H_n. \quad (12)$$

**ДОКАЗАТЕЛЬСТВО.** Для любого  $v_n \in H_n$  имеем

$$a(u - u_n, v_n) = (f, v_n) - (f, v_n) = 0,$$

поэтому

$$\begin{aligned} a(u - u_n, u - u_n) &= a(u - u_n, u - v_n) + a(u - u_n, v_n - u_n) = \\ &= a(u - u_n, u - v_n). \end{aligned}$$

Следовательно,

$$m\|u - u_n\|^2 \leq M\|u - u_n\|\|u - v_n\|,$$

откуда

$$\|u - u_n\| \leq \frac{M}{m}\|u - v_n\|. \quad \square$$

**ЗАМЕЧАНИЕ 4.** Поскольку  $M/m \geq 1$ , то оценка погрешности (6) метода Ритца, учитывающая самосопряженность оператора  $A$ , лучше оценки погрешности (12), полученной для метода Галеркина.

**Следствие 4.** Если семейство подпространств  $H_n$  предельно полно, то последовательность приближений по методу Галеркина сходится к решению задачи (2).

Доказательство дословно совпадает с доказательством следствия 1.

Ясно, что если подпространство  $H_n$  конечномерно, то коэффициенты разложения приближенного решения

$$u_n = \sum_{k=1}^n c_k \varphi_k$$

по базису  $\varphi_1, \varphi_2, \dots, \varphi_n \in H_n$  можно найти, решая систему линейных алгебраических уравнений вида (9). Матрица этой системы, как и в случае метода Ритца, положительно определена, но не симметрична, если оператор  $A$  исходной задачи не самосопряжен.

**5. Возмущенный метод Галеркина** При практическом использовании метода Галеркина форму  $a(\cdot, \cdot)$  и функционал  $f$  часто приходится заменять их приближенными аналогами, что, например, может быть вызвано приближенным вычислением элементов матрицы и правой части системы (9). Мы изучим возникающую при этом ситуацию в достаточно общей трактовке.

Итак, пусть наряду с задачей (2) рассматривается задача, отыскания элемента  $u_n \in H_n$  такого, что

$$a_n(u_n, v_n) = f_n(v_n) \quad \forall v_n \in H_n, \quad (13)$$

где  $a_n(\cdot, \cdot) : H_n \times H_n \rightarrow R$  — билинейная форма,  $f_n$  — линейная форма, определенная на  $H_n$ . На всем пространстве  $H$  эти формы могут быть и не определены. Метод (13) называют возмущенным методом Галеркина. Понятно, что погрешность приближенного решения теперь зависит не только от «близости» пространств  $H_n$  и  $H$ , но и от степени аппроксимации формами  $a_n(\cdot, \cdot)$ ,  $f_n(\cdot)$  форм  $a(\cdot, \cdot)$ ,  $f(\cdot)$  соответственно. Предположим выполненными следующие условия:

$$|a_n(u_n, v_n)| \leq M_n \|u_n\| \|v_n\| \quad \forall u_n, v_n \in H_n, \quad (14)$$

$$|f_n(v_n)| \leq c_n \|v_n\| \quad \forall v_n \in H_n, \quad (15)$$

$$a_n(u_n, u_n) \geq \tilde{m} \|u_n\|^2 \quad \forall u_n \in H_n. \quad (16)$$

Постоянная  $\tilde{m}$  считается не зависящей от  $n$ , в то время как  $M_n$ ,  $c_n$ , вообще говоря, могут зависеть от  $n$ . Условие (16) называют, обычно, условием равномерной положительной определенности форм  $a_n(\cdot, \cdot)$ .

Однозначная разрешимость задачи (13) при выполнении условий (14)–(16) есть очевидное следствие леммы Лакса — Мильграма.

Получим оценку погрешности возмущенного метода Галеркина.

**Лемма 6 (лемма Стренга).** Пусть выполнены условия (1), (14), (16). Тогда существует такая не зависящая от  $n$  постоянная  $C$ , что

$$\|u - u_n\| \leq C \left\{ \inf_{v_n \in H_n} \left( \|u - v_n\| + \sup_{w_n \in H_n} \frac{|a(v_n, w_n) - a_n(v_n, w_n)|}{\|w_n\|} \right) + \sup_{w_n \in H_n} \frac{|f(w_n) - f_n(w_n)|}{\|w_n\|} \right\}. \quad (17)$$

**ДОКАЗАТЕЛЬСТВО.** Заметим, что для любого элемента  $v_n \in H_n$  справедливо тождество

$$a_n(u_n - v_n, u_n - v_n) = a(u - v_n, u_n - v_n) + \{a(v_n, u_n - v_n) - a_n(v_n, u_n - v_n)\} + \{f_n(u_n - v_n) - f(u_n - v_n)\},$$

из которого, используя условия (16), (1), получим

$$\tilde{m} \|u_n - v_n\|^2 \leq \left\{ M \|u - v_n\| + \frac{|a(v_n, u_n - v_n) - a_n(v_n, u_n - v_n)|}{\|u_n - v_n\|} + \frac{|f_n(u_n - v_n) - f(u_n - v_n)|}{\|u_n - v_n\|} \right\} \|u_n - v_n\|,$$

или

$$\begin{aligned} \tilde{m}\|u_n - v_n\| \leq M\|u - v_n\| + \sup_{w_n \in H_n} \frac{|a(v_n, w_n) - a_n(v_n, w_n)|}{\|w_n\|} + \\ + \sup_{w_n \in H_n} \frac{|f_n(w_n) - f(w_n)|}{\|w_n\|}. \end{aligned}$$

Комбинируя это неравенство с неравенством

$$\|u - u_n\| \leq \|u - v_n\| + \|u_n - v_n\|,$$

непосредственно вытекающим из неравенства треугольника, и вычисляя затем нижнюю грань по  $v_n \in H_n$ , получим (17).  $\square$

## § 2. Элементы теории пространств Соболева

**1. Полилинейные формы.** Пусть  $R^n$  —  $n$ -мерное евклидово пространство,  $x = (x_1, \dots, x_n)^T$  — элементы пространства  $R^n$ ,  $e_1 = (1, 0, \dots, 0)^T$ ,  $e_2 = (0, 1, \dots, 0)^T, \dots, e_n = (0, 0, \dots, 1)^T$  — канонический базис  $R^n$ ,  $x \cdot y = x_1y_1 + \dots + x_ny_n$  — скалярное произведение векторов,  $|x| = (x \cdot x)^{1/2}$  — длина (норма) вектора<sup>1)</sup>.

Прямое произведение  $k$  экземпляров евклидова пространства  $R^n$ , т. е.  $R^n \times R^n \times \dots \times R^n$  будем обозначать через  $[R^n]^k$ .

Отображение  $A : [R^n]^k \rightarrow R$  называется полилинейной формой ( $k$ -линейной формой), если оно линейно по каждому аргументу:

$$A(\xi^1, \dots, \xi^i + \eta^i, \dots, \xi^k) = A(\xi^1, \dots, \xi^i, \dots, \xi^k) + A(\xi^1, \dots, \eta^i, \dots, \xi^k),$$

$$A(\xi^1, \dots, \lambda\xi^i, \dots, \xi^k) = \lambda A(\xi^1, \dots, \xi^i, \dots, \xi^k) \quad \forall \lambda \in R.$$

Полилинейная форма  $A(\xi^1, \dots, \xi^k)$  называется симметричной, если ее значение не меняется при любой перестановке аргументов.

При совпадении аргументов  $\xi^1 = \dots = \xi^k = \xi$  будем использовать обозначение:

$$A\xi^k = A(\xi, \dots, \xi).$$

Полагая  $\xi^i = \sum_{j=1}^n \xi_j^i e_j$ , получим

$$A(\xi^1, \dots, \xi^k) = \sum_{j_1, j_2, \dots, j_k=1}^n a^{j_1, j_2, \dots, j_k} \xi_{j_1}^1 \xi_{j_2}^2 \dots \xi_{j_k}^k, \quad (1)$$

<sup>1)</sup>  $x = (x_1, \dots, x_n)$  — вектор-строка,  $x = (x_1, \dots, x_n)^T \in R^n$  — вектор-столбец, так что  $x \cdot y = x^T y$ .

где

$$a^{j_1, j_2, \dots, j_k} = A(e_{j_1}, e_{j_2}, \dots, e_{j_k}).$$

Отметим, что число слагаемых в правой части (1) равно  $n^k$ .

Ясно, что если форма  $A$  симметрична, то ее коэффициенты  $a^{j_1, j_2, \dots, j_k}$  также симметричны (не меняются при любой перестановке индексов). Очевидно, справедливо и обратное: симметрия коэффициентов  $a^{j_1, j_2, \dots, j_k}$  обеспечивает симметрию полилинейной формы (1). Форму  $A$  можно отождествить с  $k$ -мерной таблицей ее коэффициентов.

Заметим, что для симметричной  $k$ -линейной формы выражение  $A\xi^k$ , где  $\xi = (\xi_1, \xi_2, \dots, \xi_n)^T$ , можно представить в виде

$$A\xi^k = \sum_{|\alpha|=k} a_\alpha \xi^\alpha, \quad \xi^\alpha = \xi_1^{\alpha_1} \xi_2^{\alpha_2} \dots \xi_n^{\alpha_n}.$$

Здесь  $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_n)$  — вектор с целочисленными неотрицательными компонентами (мультииндекс),  $|\alpha| = \alpha_1 + \alpha_2 + \dots + \alpha_n$  — длина мультииндекса.

Множество всех  $k$ -линейных форм  $A : [R^n]^k \rightarrow R$  образует конечномерное линейное пространство, если ввести естественным образом операции сложения двух форм и умножения формы на вещественное число. Обозначим его через  $\mathcal{L}_k(R^n, R)$ . Норма на этом пространстве, обычно, вводится при помощи соотношения:

$$\|A\| = \max_{|\xi^1|=\dots=|\xi^k|=1} |A(\xi^1, \dots, \xi^k)|.$$

Непосредственно из определения следует оценка:

$$|A(\xi^1, \dots, \xi^k)| \leq \|A\| |\xi^1| \dots |\xi^k| \quad \forall \xi^1, \xi^2, \dots, \xi^k \in R^n,$$

которой мы неоднократно будем пользоваться. В некоторых случаях удобнее пользоваться нормой

$$\|A\|_\infty = \max_{j_1, \dots, j_k} |a^{j_1, j_2, \dots, j_k}|.$$

Важно помнить, что вследствие конечномерности пространства  $\mathcal{L}_k(R^n, R)$  все нормы на нем эквивалентны.

Далее будем использовать обозначение:

$$[A] = \max_{i=1, \dots, n} |A(e^i)^k| = \max_{i=1, \dots, n} |a^{i, \dots, i}|.$$

Отметим, что операция  $[\cdot]$  определяет только полунорму на множестве  $\mathcal{L}_k(R^n, R)$  при  $k \geq 2$ .

Множество симметричных  $k$ -линейных форм также образует конечномерное линейное пространство. Норма на этом пространстве может быть определена равенством:

$$\|A\| = \max_{|\xi|=1} |A\xi^k|. \quad (2)$$

Отметим, что соотношение (2) не определяет нормы на пространстве всех  $k$ -линейных форм. Действительно, например, для любой кососимметричной формы, т. е. формы, меняющей знак при перестановке какой-либо пары аргументов,  $A\xi^k = 0 \forall \xi \in R^n$ .

В дальнейшем будут рассматриваться также полилинейные (линейные по каждому аргументу) отображения пространства  $[R^n]^k$  в  $R^n$ . Множество всех таких отображений обозначим через  $\mathcal{L}_k(R^n, R^n)$ . Понятно, что задание  $A \in \mathcal{L}_k(R^n, R^n)$  эквивалентно заданию упорядоченного набора  $n$  штук  $k$ -линейных форм  $A_1, A_2, \dots, A_n \in \mathcal{L}_k(R^n, R)$ :

$$A(\xi^1, \dots, \xi^k) = (A_1(\xi^1, \dots, \xi^k), \dots, A_n(\xi^1, \dots, \xi^k))^T, \quad \xi^i \in R^n.$$

Норму на этом пространстве можно ввести при помощи соотношения:

$$\|A\| = \max_{|\xi^1|=\dots=|\xi^k|=1} |A(\xi^1, \dots, \xi^k)|. \quad (3)$$

В некоторых случаях удобнее пользоваться нормой:

$$\|A\|_\infty = \max_{i=1, \dots, n} \|A_i\|_\infty.$$

Далее будем использовать обозначение:

$$[A] = \max_{i=1, \dots, n} |A(e_i)^k|, \quad A\xi^k = A(\xi, \dots, \xi).$$

Симметрия отображения  $A$ , очевидно, эквивалентна симметрии каждой из форм  $A_1, A_2, \dots, A_n$ .

**2. Некоторые определения и обозначения из теории дифференцируемых функций.** Открытое связное множество точек пространства  $R^n$  называется областью. Пусть  $u = u(x)$  — вещественная функция, определенная на некоторой области  $\Omega \subset R^n$ . Будем говорить, что эта функция  $l$  раз непрерывно дифференцируема на области  $\Omega$ , если у нее существуют и непрерывны на  $\Omega$  все частные производные вплоть до порядка  $l$  включительно. Как известно, в этом случае ее любая смешанная производная

$$\frac{\partial^k u}{\partial x_{i_1} \partial x_{i_2} \dots \partial x_{i_k}}, \quad i_1, i_2, \dots, i_k \in \{1, 2, \dots, n\},$$

порядка  $k \leq l$  не меняется при любой перестановке индексов  $i_1, i_2, \dots, i_k$ . В связи с этим естественно пользоваться следующим обозначением. Пусть  $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_n)$  — мультииндекс. Тогда

$$D^\alpha u = \frac{\partial^{|\alpha|} u}{\partial x_1^{\alpha_1} \partial x_2^{\alpha_2} \dots \partial x_n^{\alpha_n}}$$

обозначает производную порядка  $|\alpha|$ .

Множество всех производных функции  $u$  фиксированного порядка  $k$ , вычисленных в точке  $x \in \Omega$ , порождает симметричную  $k$ -линейную форму  $D^k u(x) \in \mathcal{L}_k(R^n, R)$ , определенную равенством:

$$D^k u(x)(h^1, h^2, \dots, h^k) = \sum_{i_1, i_2, \dots, i_k=1}^n \frac{\partial^k u(x)}{\partial x_{i_1} \partial x_{i_2} \dots \partial x_{i_k}} h_{i_1}^1 h_{i_2}^2 \dots h_{i_k}^k.$$

Нетрудно видеть, что

$$\frac{\partial^k u(x)}{\partial x_{i_1} \partial x_{i_2} \dots \partial x_{i_k}} = D^k u(x)(e_{i_1}, e_{i_2}, \dots, e_{i_k}).$$

Подчеркнем, что, несмотря на сходство обозначений,  $D^k u(x)$ , где  $k$  есть целое положительное число, и  $D^\alpha u(x)$ , где  $\alpha$  — мультииндекс длины  $n$  это — совершенно разные объекты:  $D^k u(x)$  — полилинейная форма, определяемая множеством всех частных производных порядка  $k$ , а  $D^\alpha u(x)$  — число, равное значению частной производной порядка  $|\alpha|$  функции  $u$  в точке  $x$ .

В дальнейшем также используется обозначение

$$D^k u(x) h^k = D^k u(x)(h, h, \dots, h),$$

$D^k u(x) h^k$  обычно называют дифференциалом порядка  $k$  функции  $u$  в точке  $x$ . Отметим следующие частные случаи: при  $k = 1$

$$Du(x)h \equiv D^1 u(x)h = (\nabla u(x))^T h \quad \forall h \in R^n,$$

вектор

$$\nabla u(x) = \left( \frac{\partial u(x)}{\partial x_1}, \frac{\partial u(x)}{\partial x_2}, \dots, \frac{\partial u(x)}{\partial x_n} \right)^T$$

называют градиентом функции  $u$  в точке  $x$ ; если  $k = 2$ , то

$$D^2 u(x)(h^1, h^2) = H(x)h^1 \cdot h^2 \quad \forall h^1, h^2 \in R^n.$$

Симметричная матрица

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

называется гессианом функции  $u$  в точке  $x$ .

### 3. Дифференцируемые отображения. Отображение

$$u = (u_1, u_2, \dots, u_n)^T : R^n \rightarrow R^n$$

будем называть  $k$  раз дифференцируемым в точке  $x \in R^n$ , если каждая его компонента  $u_i$  дифференцируема  $k$  раз в точке  $x$ . При этом по определению  $D^k u(x) = (D^k u_1(x), D^k u_2(x), \dots, D^k u_n(x))^T \in \mathcal{L}_k(R^n, R^n)$ ,  $D^k u_i(x) \in \mathcal{L}_k(R^n, R)$ . В частности,  $Du(x) = D^1 u(x)$  — линейное отображение, порождаемое матрицей Якоби  $\left( \frac{\partial u_i(x)}{\partial x_j} \right)_{i,j=1}^n$  системы функций  $u_1, u_2, \dots, u_n$ . В дальнейшем через  $\det(Du(x))$  обозначается определитель этой матрицы, называемый якобианом отображения  $u$ .

**4. Формула Тейлора.** Пусть функция  $u$  определена и  $m+1$  раз непрерывно дифференцируема в области  $\Omega \subset R^n$ ,  $x, x+h \in \Omega$ , причем вектор  $h$  настолько мал, что  $x+th \in \Omega \forall t \in [0, 1]$ . Тогда функция  $\varphi(t) = u(x+th)$  дифференцируема  $m+1$  раз на отрезке  $[0, 1]$ . По формуле Тейлора с остаточным членом в интегральной форме

$$\varphi(1) = \varphi(0) + \varphi'(0) + \frac{1}{2!} \varphi''(0) + \dots + \frac{1}{m!} \varphi^{(m)}(0) + \frac{1}{m!} \int_0^1 \varphi^{(m+1)}(t) (1-t)^m dt.$$

Нетрудно проверить, что

$$\varphi^{(k)}(t) = \frac{d^k}{dt^k} u(x+th) = D^k u(x+th) h^k,$$

следовательно,

$$u(x+h) = \sum_{k=0}^m \frac{1}{k!} D^k u(x) h^k + R_m, \quad (4)$$

где

$$R_m = \frac{1}{m!} \int_0^1 (1-t)^m D^{m+1} u(x+th) h^{m+1} dt.$$

Формула (4) — формула Тейлора для функций многих переменных с остаточным членом в интегральной форме.

Покажем, что если функция  $u$  непрерывно дифференцируема  $m$  раз, то

$$u(x+h) = \sum_{k=0}^m \frac{1}{k!} D^k u(x) h^k + o(|h|^m). \quad (5)$$

Действительно, используя вторую теорему о среднем и равенство (4), получим

$$u(x+h) = \sum_{k=0}^{m-1} \frac{1}{k!} D^k u(x) h^k + R_{m-1},$$

где

$$R_{m-1} = \frac{1}{m!} D^m u(x + \xi h) h^m, \quad \xi \in (0, 1),$$

есть остаточный член, записанный в форме Лагранжа. Далее,

$$|D^m u(x + \xi h) h^m - D^m u(x) h^m| \leq \|D^m u(x + \xi h) - D^m u(x)\| |h|^m = o(|h|^m)$$

в силу непрерывности  $D^m u$  в точке  $x$ .

Формулу (5) называют формулой Тейлора с остаточным членом в форме Пеано.

Таким образом, используя формулу Тейлора, мы приближаем функцию  $u$  в окрестности точки  $x_0$  полиномом

$$P_m(x) = \sum_{k=0}^m \frac{1}{k!} D^k u(x_0) (x - x_0)^k$$

степени  $m$ .

Заметим, что формулу Тейлора иногда удобно использовать для вычисления производных. А именно, если каким-либо образом получено представление

$$u(x+h) = \sum_{k=0}^m A_k h^k + o(|h|^m),$$

где  $A_k \in \mathcal{L}_k(R^n, R)$  — симметричные  $k$ -линейные формы, то  $D^k u(x) h^k = k! A_k h^k$ , и, следовательно, производные функции  $u$  порядка  $k$  в точке  $x$  легко выражаются через коэффициенты формы  $A_k$ . Пример такого применения формулы Тейлора см. далее в п. 10 настоящего параграфа.

**УПРАЖНЕНИЕ 1.** Доказать, что формула Тейлора (5) справедлива и для вектор-функций  $u: R^n \rightarrow R^n$ .

**5. Основные функциональные пространства. Простейшие неравенства.** Всюду в дальнейшем  $\Omega$  — ограниченная область евклидова пространства  $R^n$ ,  $\Gamma$  — граница  $\Omega$ . Иногда будем обозначать ее через  $\partial\Omega$ . Через  $L_p(\Omega)$ , как обычно, будем обозначать линейное вещественное пространство функций, измеримых на  $\Omega$  и суммируемых

со степенью  $p \geq 1$  в смысле Лебега. Норма на пространстве  $L_p(\Omega)$  задается равенством:

$$\|u\|_{p,\Omega} = \left( \int_{\Omega} |u|^p dx \right)^{1/p}.$$

Пространство  $L_{\infty}(\Omega)$  — линейное пространство функций с конечной нормой

$$\|u\|_{\infty,\Omega} = \operatorname{ess\,sup}_{x \in \Omega} |u(x)|,$$

где  $\operatorname{ess\,sup}_{x \in \Omega} |u(x)|$  — существенная верхняя грань функции  $u$  на  $\Omega$ , т. е.

нижняя грань всех чисел  $N$  таких, что  $|u(x)| \leq N$  почти всюду на  $\Omega$ .

Отметим, что  $\|u\|_{\infty,\Omega} = \lim_{p \rightarrow \infty} \|u\|_{p,\Omega}$ .

Через  $C(\bar{\Omega})$  будем обозначать пространство непрерывных на  $\bar{\Omega}$  вещественных функций с нормой

$$\|u\|_{C(\bar{\Omega})} = \max_{x \in \bar{\Omega}} |u(x)|.$$

Пусть  $u \in L_{p_1}(\Omega)$ ,  $v \in L_{p_2}(\Omega)$ ,  $1/p_1 + 1/p_2 = 1$ . Тогда  $uv \in L_1(\Omega)$  и выполнено неравенство Гельдера:

$$\left| \int_{\Omega} u(x)v(x)dx \right| \leq \left( \int_{\Omega} |u(x)|^{p_1} dx \right)^{1/p_1} \left( \int_{\Omega} |v(x)|^{p_2} dx \right)^{1/p_2}. \quad (6)$$

Справедливо и более общее неравенство:

$$\begin{aligned} & \left| \int_{\Omega} u_1(x)u_2(x) \cdots u_m(x)dx \right| \leq \\ & \leq \left( \int_{\Omega} |u_1(x)|^{p_1} dx \right)^{1/p_1} \left( \int_{\Omega} |u_2(x)|^{p_2} dx \right)^{1/p_2} \cdots \left( \int_{\Omega} |u_m(x)|^{p_m} dx \right)^{1/p_m}, \end{aligned} \quad (7)$$

если  $u_i \in L_{p_i}(\Omega)$ ,  $i = 1, 2, \dots, m$ ,  $1/p_1 + 1/p_2 + \dots + 1/p_m = 1$ .

Неравенство (6) при  $p_1 = p_2 = 2$  называют неравенством Коши — Буняковского.

Мы будем также часто использовать аналогичные неравенства для сумм: неравенство Коши:

$$\left| \sum_{i=1}^m a_i b_i \right| \leq \left( \sum_{i=1}^m |a_i|^2 \right)^{1/2} \left( \sum_{i=1}^m |b_i|^2 \right)^{1/2},$$

неравенство Гельдера:

$$\left| \sum_{i=1}^m a_i b_i \right| \leq \left( \sum_{i=1}^m |a_i|^{p_1} \right)^{1/p_1} \left( \sum_{i=1}^m |b_i|^{p_2} \right)^{1/p_2}, \quad 1/p_1 + 1/p_2 = 1,$$

а также неравенство для сумм, соответствующее обобщенному неравенству Гельдера (7).

Часто оказывается полезной следующая формула для вычисления нормы в пространстве  $L_p$ :

$$\|u\|_{L_p} = \sup_{v \in L_q} \left( \int_{\Omega} u v dx / \|v\|_{L_q} \right), \quad 1/p + 1/q = 1. \quad (8)$$

Для ее доказательства достаточно заметить, что вследствие неравенства Гельдера

$$\int_{\Omega} u v dx / \|v\|_{L_q} \leq \|u\|_{L_p} \quad \forall v \in L_q,$$

полагая же  $v(x) = |u(x)|^{p-1} \text{sign } u(x)$ , получим

$$v \in L_q, \quad \|v\|_{L_q} = \|u\|_{L_p}^{p-1}, \quad \int_{\Omega} u v dx = \|u\|_{L_p}^p,$$

т. е.

$$\int_{\Omega} u v dx / \|v\|_{L_q} = \|u\|_{L_p}.$$

**6. Обобщенные производные.** Функция  $u$  называется финитной на  $\Omega$ , если существует компакт  $K \subset \Omega$  такой, что  $u(x) = 0$  при  $x \notin K$ . Через  $C_0^\infty(\Omega)$  будем обозначать множество бесконечно дифференцируемых финитных на  $\Omega$  функций.

**Определение 1.** Функция  $u(x) \in L_p(\Omega)$  имеет обобщенную производную  $\partial u / \partial x_i \in L_p(\Omega)$  по переменной  $x_i$ , если существует функция  $u^*(x) \in L_p(\Omega)$  такая, что

$$\int_{\Omega} u(x) \frac{\partial v}{\partial x_i} dx = - \int_{\Omega} u^*(x) v dx \quad \forall v \in C_0^\infty(\Omega).$$

По определению полагают  $\partial u(x) / \partial x_i = u^*(x)$ .

Нетрудно видеть, что обобщенная производная, если она существует, определяется единственным образом. Действительно, если  $\tilde{u}^*$  — также обобщенная производная функции  $u$  по переменной  $x_i$ , то

$$\int_{\Omega} (\tilde{u}^* - u^*)v \, dx = 0 \quad \forall v \in C_0^\infty(\Omega),$$

откуда в силу плотности  $C_0^\infty(\Omega)$  в  $L_p(\Omega)$  следует, что  $\tilde{u}^* = u^*$ . Если функция  $u$  дифференцируема в области  $\Omega$  в обычном смысле, и ее производная  $\partial u / \partial x_i \in L_p(\Omega)$ , то она совпадает с обобщенной производной.

Обобщенные производные высших порядков определяются аналогично:

$$D^\alpha u = \frac{\partial^{|\alpha|} u}{\partial x_1^{\alpha_1} \partial x_2^{\alpha_2} \dots \partial x_n^{\alpha_n}} \in L_p(\Omega), \quad |\alpha| = \alpha_1 + \alpha_2 + \dots + \alpha_n,$$

$$\int_{\Omega} u(x) D^\alpha v(x) \, dx = (-1)^{|\alpha|} \int_{\Omega} D^\alpha u(x) v(x) \, dx \quad \forall v \in C_0^\infty(\Omega).$$

Будем считать, что  $D^0 u = u$ .

Приведем примеры, иллюстрирующие различия между понятиями классической и обобщенной производной.

**ПРИМЕР 1.** Пусть функция  $u$  определена на интервале  $(-1, 1)$  соотношением:

$$u(x) = \begin{cases} -x, & x \leq 0, \\ 1 + x, & x > 0. \end{cases}$$

Ясно, что  $u \in L_p(-1, 1)$  при любом  $p \geq 1$ . Функция  $u$  дифференцируема на интервале  $(-1, 1)$  всюду, кроме точки  $x = 0$ . Ее производная

$$u'(x) = \begin{cases} -1, & x < 0, \\ 1, & x > 0. \end{cases}$$

принадлежит  $L_p(-1, 1)$ , но не является обобщенной производной функции  $u$ . Действительно, в противном случае должно было бы выполняться тождество:

$$\int_{-1}^1 uv' \, dx = - \int_{-1}^1 u'v \, dx \quad \forall v \in C_0^\infty(-1, 1). \quad (9)$$

Используя формулу интегрирования по частям, получим:

$$\int_{-1}^1 uv' dx = \int_{-1}^0 uv' dx + \int_0^1 uv' dx = - \int_{-1}^1 u'v dx - v(0),$$

и поскольку, вообще говоря,  $v(0) \neq 0$ , тождество (9) не выполняется, т. е. функция  $u'$  не является обобщенной производной функции  $u$ .

Этот пример наводит на мысль, что функция, имеющая обобщенную производную, должна быть непрерывной. В одномерном случае это действительно так. Более того, всякая функция, имеющая обобщенную производную из  $L_p(0, 1)$ ,  $p \geq 1$ , абсолютно непрерывна на отрезке  $[0, 1]$ . Можно показать, что и в случае произвольного числа измерений функция, непрерывная и кусочно непрерывно дифференцируемая на некоторой области, имеет все обобщенные производные первого порядка (подробнее см. ниже, с. 44).

Однако, как показывает следующий пример, при  $n \geq 2$  функция, имеющая обобщенные производные, не обязательно является ограниченной.

ПРИМЕР 2. Пусть  $\Omega = \{x \in R^2: x_1^2 + x_2^2 \leq R < 1\}$ ,  $u(x) = \ln \ln \frac{1}{r}$ , где  $r = (x_1^2 + x_2^2)^{1/2}$ . Функция  $u$  и ее первые производные

$$\frac{\partial u}{\partial x_i} = -\frac{x_i}{r^2 \ln r},$$

неограниченно возрастают при  $r \rightarrow 0$ . Легко проверить, что  $u, \partial u / \partial x_i \in L_2(\Omega)$ . Покажем, что  $\partial u / \partial x_i$ ,  $i = 1, 2$ , — обобщенные производные функции  $u$ . Для любой функции  $v \in C_0^\infty(\Omega)$  можно написать

$$\int_{\Omega} \frac{\partial v}{\partial x_i} \ln \ln \frac{1}{r} dx = \lim_{\varepsilon \rightarrow 0} \int_{\Omega_\varepsilon} \frac{\partial v}{\partial x_i} \ln \ln \frac{1}{r} dx,$$

где  $\Omega_\varepsilon = \{x \in \Omega : r > \varepsilon\}$ . Используя формулу интегрирования по частям, получим:

$$\int_{\Omega_\varepsilon} \frac{\partial v}{\partial x_i} \ln \ln \frac{1}{r} dx = - \int_{\Omega_\varepsilon} v \frac{\partial u}{\partial x_i} dx + \int_{|r|=\varepsilon} v \ln \ln \frac{1}{r} \cos(\nu, x_i) dx,$$

где  $\nu$  — нормаль, внешняя к  $\Omega_\varepsilon$ . Нетрудно убедиться, что

$$\lim_{\varepsilon \rightarrow 0} \int_{|r|=\varepsilon} v \ln \ln \frac{1}{r} \cos(\nu, x_i) dx = 0,$$

и, следовательно,

$$\int_{\Omega} u \frac{\partial v}{\partial x_i} dx = - \int_{\Omega} v \frac{\partial u}{\partial x_i} dx \quad \forall v \in C_0^\infty(\Omega),$$

т. е.  $\partial u / \partial x_i$  — обобщенная производная функции  $u$ .

**7. Пространства Соболева.** Множество функций из  $L_p(\Omega)$ , имеющих обобщенные производные  $D^\alpha u \in L_p(\Omega)$  для всех  $\alpha$ ,  $|\alpha| \leq s$  ( $s \geq 1$ ) образует линейное пространство. Вводя на этом пространстве норму

$$\|u\|_{s,p,\Omega} = \left( \int_{\Omega} \sum_{|\alpha| \leq s} |D^\alpha u|^p dx \right)^{1/p} \quad (10)$$

при конечных  $p \geq 1$ , и

$$\|u\|_{s,\infty,\Omega} = \max_{|\alpha| \leq s} \operatorname{ess\,sup}_{x \in \Omega} |D^\alpha u(x)| \quad (11)$$

при  $p = \infty$ , получим линейное нормированное пространство (пространство Соболева), обозначаемое через  $W_p^s(\Omega)$ . Известно, что  $W_p^s(\Omega)$  — полное пространство.

Далее будут использоваться следующие обозначения для полунорм:

$$|u|_{s,p,\Omega} = \left( \int_{\Omega} \sum_{|\alpha|=s} |D^\alpha u|^p dx \right)^{1/p}, \quad [u]_{s,p,\Omega} = \left( \int_{\Omega} \sum_{i=1}^n \left| \frac{\partial^s u}{\partial x_i^s} \right|^p \right)^{1/p},$$

$$|u|_{s,\infty,\Omega} = \max_{|\alpha|=s} \operatorname{ess\,sup}_{x \in \Omega} |D^\alpha u(x)|, \quad [u]_{s,\infty,\Omega} = \max_{i=1,\dots,n} \operatorname{ess\,sup}_{x \in \Omega} \left| \frac{\partial^s u}{\partial x_i^s}(x) \right|.$$

Описание свойств пространств Соболева связано с сужением класса рассматриваемых областей. Приведем соответствующие определения.

**Определение 2.** Заданная на некотором множестве  $D$  пространства  $R^n$  вещественная функция  $u$  удовлетворяет условию Липшица с показателем  $\lambda$ , если существуют постоянные  $c > 0$  и  $\lambda \in (0, 1]$  такие, что

$$|u(x) - u(x')| \leq c|x - x'|^\lambda \quad \forall x, x' \in D.$$

**Определение 3.** Ограниченная область  $\Omega$  евклидова пространства  $R^n$  и ее граница  $\Gamma$  принадлежат классу  $C^{k,\lambda}$ , если для любой точки  $x \in \Gamma$  найдется ее окрестность  $\Delta(x) \subset R^n$  и декартова система

координат  $(\xi_1, \xi_2, \xi_3, \dots, \xi_n)$ , такие, что в  $\Delta(x)$  уравнение  $\Gamma$  может быть записано в виде  $\xi_n = \varphi(\xi_1, \xi_2, \xi_3, \dots, \xi_{n-1})$  с функцией  $\varphi$  класса  $C^{k,\lambda}$ , т. е. функцией  $k$  раз дифференцируемой, все  $k$ -е производные которой удовлетворяют условию Липшица с показателем  $\lambda$ .

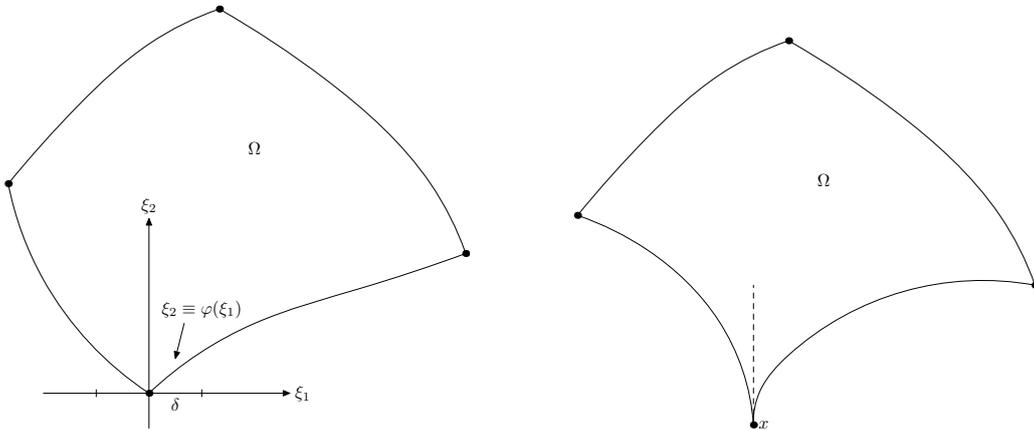


Рис. 1. Пример областей с липшицевой (слева) и не липшицевой границей (справа). В точке  $x$  имеется «нулевой» угол.

Будем предполагать, что все встречающиеся в дальнейшем области как минимум принадлежат классу  $C^{0,1}$ , иначе говоря, являются липшицевыми.

Для таких областей можно дать другое, эквивалентное, определение пространства  $W_p^s(\Omega)$ .

Пусть  $C^\infty(\bar{\Omega})$  — линейное пространство, получаемое сужением на  $\bar{\Omega} = \Omega \cup \Gamma$  множества бесконечно дифференцируемых на  $R^n$  функций. Пространство  $W_p^s(\Omega)$  определяется как замыкание  $C^\infty(\bar{\Omega})$  по норме (10).

Чаще всего в книге используется пространство Соболева  $W_2^s(\Omega)$ . Это пространство становится пространством Гильберта, если ввести на нем скалярное произведение

$$(u, v)_{s,\Omega} = \int_{\Omega} \sum_{|\alpha| \leq s} D^\alpha u D^\alpha v \, dx.$$

Пространство  $W_2^s(\Omega)$  будем обозначать, обычно, через  $H^s(\Omega)$ , а норму функции  $u \in H^s(\Omega)$  через  $\|u\|_{s,\Omega}$ .

Функции, принадлежащие  $W_p^s(\Omega)$ , при достаточно больших  $s$  и  $p$  обладают определенной гладкостью. Дадим в связи с этим

**Определение 4.** Пусть  $X$  и  $Y$  линейные нормированные пространства с нормами  $\|\cdot\|_X, \|\cdot\|_Y$ . Говорят, что  $X$  непрерывно вкладывается

вадается в  $Y$ , если каждый элемент  $u$ , принадлежащий  $X$ , принадлежит  $Y$ , и существует постоянная  $c$  такая, что

$$\|u\|_Y \leq c \|u\|_X \quad \forall u \in X.$$

Иными словами, оператор, сопоставляющий каждому элементу пространства  $X$  тот же элемент как элемент пространства  $Y$  (оператор вложения) — линейный ограниченный оператор.

**Определение 5.** Линейное нормированное пространство  $X$  вполне непрерывно вкладывается в линейное нормированное пространство  $Y$ , если оператор вложения  $X$  в  $Y$  вполне непрерывен<sup>1)</sup>.

**Теорема 1.** Пусть  $sp < n$ . Тогда пространство  $W_p^s(\Omega)$  непрерывно вкладывается в  $L_q(\Omega)$  при любом  $q \leq np/(n - sp)$ . Вложение  $W_p^s(\Omega)$  в  $L_q$  вполне непрерывно при  $q < np/(n - sp)$ . Если  $sp = n$ , то  $W_p^s(\Omega)$  вполне непрерывно вкладывается в  $L_q(\Omega)$  при любом  $q \in [1, \infty)$ .

**Теорема 2.** Вложение  $W_p^{s+1}(\Omega)$  в  $W_p^s(\Omega)$  вполне непрерывно при любых  $s \geq 0$ ,  $p \in [1, \infty]$ .

ЗАМЕЧАНИЕ 1. Теорема 1, в частности, показывает, что функции из  $W_2^1(\Omega)$  при  $n = 2$  суммируемы по  $\Omega$  с любой степенью  $q < \infty$ .

**Теорема 3.** Пусть  $(s - r)p > n$ . Тогда пространство  $W_p^s(\Omega)$  вполне непрерывно вкладывается в пространство  $C^r(\bar{\Omega})$ .

ЗАМЕЧАНИЕ 2. Нужно иметь в виду, что по определению функция  $u$  из  $W_p^s(\Omega)$  — элемент  $L_p(\Omega)$ , т. е. под  $u$  фактически понимается класс эквивалентности. Иными словами, функции, совпадающие почти всюду на  $\Omega$ , неразличимы. Таким образом, например, принадлежность функции  $u$  из  $W_p^s(\Omega)$  пространству  $C(\bar{\Omega})$  означает, что соответствующий класс эквивалентности содержит функцию, непрерывную на  $\bar{\Omega}$ .

Для областей класса  $C^{s,\lambda}$  функции из  $W_p^s(\Omega)$  допускают продолжение на более широкую область с сохранением нормы. Точнее, справедлива

**Теорема 4.** Пусть  $u \in W_p^s(\Omega)$ ,  $\Omega'$  — область, содержащая  $\bar{\Omega}$ . Тогда существует функция  $\tilde{u} \in \overset{\circ}{W}_p^s(\Omega')$  такая, что  $\tilde{u}(x) = u(x)$  при  $x \in \Omega$  и

$$\|\tilde{u}\|_{s,p,\Omega'} \leq c \|u\|_{s,p,\Omega},$$

где  $c$  — постоянная, зависящая от областей  $\Omega$  и  $\Omega'$ , но не зависящая от  $u$ .

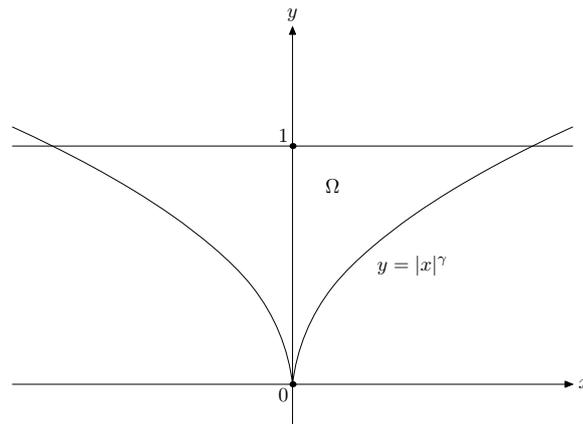


Рис. 2. Область из контрпримера.

Приведем пример, показывающий, что утверждение теоремы 3 может не выполняться, если область  $\Omega$  не принадлежит классу  $C^{0,1}$ .

Пусть  $\Omega$  — двумерная область, ограниченная кривой  $y = |x|^\gamma$  ( $\gamma < 1$ ) и прямой  $y = 1$  (рис. 2). В начале координат граница области имеет нулевой угол и потому, как нетрудно проверить, не принадлежит классу  $C^{0,1}$ . Рассмотрим функцию  $u(x, y) = y^\beta$ . Эта функция неограниченно возрастает, когда  $y \rightarrow 0$  при любом  $\beta < 0$ . Однако она имеет обобщенные производные первого порядка в области  $\Omega$ :

$$\frac{\partial u}{\partial x} = 0, \quad \frac{\partial u}{\partial y} = \beta y^{\beta-1},$$

причем

$$\int_{\Omega} \left| \frac{\partial u}{\partial y} \right|^{2+\varepsilon} d\Omega = \beta^{2+\varepsilon} \int_0^1 \int_{-y^{1/\gamma}}^{y^{1/\gamma}} y^{(\beta-1)(2+\varepsilon)} dx dy = 2\beta^{2+\varepsilon} \int_0^1 y^{(\beta-1)(2+\varepsilon)+1/\gamma} dy.$$

Ясно, что существуют такие  $\varepsilon > 0$ ,  $\beta < 0$ , что  $(\beta-1)(2+\varepsilon)+1/\gamma > -1$  при любом  $\gamma < 1$ . Это означает, что  $u \in W_{2+\varepsilon}^1(\Omega)$ , и, следовательно, теорема 3 для рассматриваемой области неверна.

Ввиду особой важности для теории МКЭ теоремы 3 дадим элементарное доказательство непрерывности вложения  $W_p^s(\Omega)$  в  $C(\bar{\Omega})$ , опирающееся, правда, на теорему 4 о продолжении. При этом будем считать, что  $\Omega$  — область класса  $C^{s,\lambda}$ .

Обозначим через  $\Omega'$  область, содержащую  $\bar{\Omega}$ . Рассмотрим сначала некоторую функцию  $u(x) \in C_0^\infty(\Omega')$  и продолжим ее нулем на все

<sup>1)</sup>т. е. всякое множество, ограниченное в  $X$ , компактно в пространстве  $Y$ , или, что эквивалентно, из последовательности, ограниченной в  $X$ , можно выделить подпоследовательность, сходящуюся в  $Y$ .

пространство  $R^n$ . Пусть  $S(x_0, R)$  — шар радиуса  $R$  с центром в точке  $x_0$ ,  $S(x_0, R) \supseteq \Omega'$ . Введем сферическую систему координат с началом в точке  $x_0$ . Имеем

$$u(x_0) = - \int_0^R u'_r(r) dr,$$

где  $r$  — радиальная координата. Интегрируя по частям  $s - 1$  раз, получим

$$u(x_0) = b \int_0^R r^{s-1} \frac{\partial^s u}{\partial r^s} dr,$$

где  $b$  — постоянная, зависящая лишь от  $s$ . Проинтегрируем это равенство по угловым переменным. Будем иметь:

$$u(x_0) = b_1 \int_{S(x_0, R)} r^{s-n} \frac{\partial^s u}{\partial r^s} dv,$$

где  $dv = r^{n-1} dr d\sigma$  — элемент объема в сферических координатах,  $b_1$  — постоянная, зависящая лишь от  $n$  и  $s$ . Используя неравенство Гёльдера, получим:

$$|u(x_0)| \leq b_1 \left( \int_{S(x_0, r)} r^{p'(s-n)} dv \right)^{1/p'} \left( \int_{S(x_0, r)} \left| \frac{\partial^s u}{\partial r^s} \right|^p dv \right)^{1/p},$$

где  $1/p + 1/p' = 1$ . Очевидно, первый интеграл справа конечен и зависит лишь от  $n, s, p$  и  $R$  при  $p'(s-n) + n - 1 > -1$ , т. е. при  $ps > n$ . Возвращаясь к декартовым координатам, получим, что второй интеграл не превосходит  $\|u\|_{s,p,\Omega'}$ . Вследствие произвольности  $x_0 \in \Omega$  отсюда вытекает, что

$$\|u\|_{C(\bar{\Omega})} \leq c \|u\|_{s,p,\Omega'} \quad (12)$$

с постоянной  $c$ , не зависящей от  $u$ .

Пусть теперь  $u \in W_p^s(\Omega)$  — произвольная функция,  $\tilde{u}$  — продолжение  $u$ , такое, что  $\|\tilde{u}\|_{s,p,\Omega'} \leq c \|u\|_{s,p,\Omega}$ . Рассмотрим последовательность функций  $u_m \in C_0^\infty(\Omega')$  таких, что  $\|u_m - \tilde{u}\|_{s,p,\Omega'} \rightarrow 0$  при  $m \rightarrow \infty$ . Последовательность  $u_m$  фундаментальна в  $W_p^s(\Omega')$  и в силу неравенства (12), справедливого для функций из  $C_0^\infty(\Omega')$ , фундаментальна и в  $C(\bar{\Omega})$ , поэтому последовательность  $u_m$  сходится в смысле пространства  $C(\bar{\Omega})$ , а ее предел  $u(x)$  — непрерывная в  $\bar{\Omega}$  функция.

Для каждого члена последовательности  $u_m$  выполнено неравенство  $\|u_m\|_{C(\bar{\Omega})} \leq c \|u_m\|_{s,p,\Omega}$ . Это значит, что оно справедливо и для предельной функции  $\tilde{u}$ . Для завершения доказательства теоремы достаточно заметить, что  $\|\tilde{u}\|_{s,p,\Omega} \leq c \|u\|_{s,p,\Omega}$ .  $\square$

**8. Следы функций из соболевских пространств.** Как уже говорилось (см. замечание 2), функции из  $W_p^s(\Omega)$ , совпадающие почти всюду на  $\Omega$ , неразличимы. Вследствие этого непосредственно говорить о значениях (следах) функций из  $W_p^s(\Omega)$  на многообразиях размерности, меньшей чем  $n$ , и, в частности, на границе области, не имеет смысла. Тем не менее, при определенных условиях на  $p$  и  $s$  можно дать корректное определение понятия следа, обобщающее обычное понятие сужения непрерывной на  $\bar{\Omega}$  функции на некоторую часть  $\bar{\Omega}$ . Приведем в общих чертах соответствующие рассуждения. Справедлива

**Теорема 5.** Пусть  $u \in C^\infty(\Omega)$ . Обозначим через  $\Gamma_m$  сечение области  $\Omega$  гиперплоскостью размерности  $m < n$ . Пусть  $m > n - sp > 0$ ,  $q = mp/(n - sp)$ . Тогда

$$\|u\|_{L_q(\Gamma_m)} \leq c \|u\|_{s,p,\Omega}, \quad (13)$$

где  $c$  — постоянная, не зависящая от выбора  $u$ .

Пусть теперь  $u$  — произвольная функция из  $W_p^s$ ,  $u_k$  — последовательность функций из  $C^\infty(\bar{\Omega})$ , сходящаяся к  $u$ . Тогда на основании неравенства (13) получаем, что последовательность  $u_k$  фундаментальна в  $L_q(\Gamma_m)$ . Предельная функция  $u \in L_q(\Gamma_m)$  этой последовательности называется следом функции  $u \in W_p^s(\Omega)$  на  $\Gamma_m$ .

Можно показать, что такое определение следа корректно в том смысле, что:

- 1) след функции  $u \in W_p^s(\Omega)$  не зависит от выбора последовательности  $u_k$ , сходящейся к  $u$ ;
- 2) след непрерывно меняется (в смысле нормы  $L_q(\Gamma_m)$ ) при непрерывном движении сечения области  $\bar{\Omega}$ ;
- 3) для непрерывной функции из  $W_p^s(\Omega)$  след совпадает с ее сужением на  $\Gamma_m$ .

Понятно, что для следа функции  $u$  справедлива оценка (13). Более того, имеет место

**Теорема 6.** Пусть  $\Gamma_m$  — сечение области  $\bar{\Omega}$  гиперплоскостью размерности  $m < n$ ,  $m > n - sp \geq 0$ . Тогда:

- 1) если  $n > sp$ , то пространство  $W_p^s(\Omega)$  непрерывно вкладывается в пространство  $L_q(\Gamma_m)$ , где  $q = mp/(n - sp)$ ; пространство  $W_p^s(\Omega)$  компактно вкладывается в пространство  $L_{q^*}(\Gamma_m) \forall q^* < q$ ;

2) если  $n = sp$ , то пространство  $W_p^s(\Omega)$  компактно вкладывается в  $L_q(\Gamma_m) \forall q < \infty$ .

Утверждение теоремы справедливо и в том случае, когда  $\Gamma_m$  — кусок достаточно гладкой поверхности<sup>1)</sup> размерности  $m$ , например,  $\Gamma_m$  может быть частью границы области  $\Omega$ .

Пусть  $g \in L_q(\Gamma)$ . Будем говорить, что  $u(x) = g(x)$  на  $\Gamma$ , если след функции  $u$  на  $\Gamma$  совпадает с  $g$ . В частности, если  $g = 0$ , будем говорить, что функция  $u$  обращается в нуль на  $\Gamma$ .

ЗАМЕЧАНИЕ 3. Не следует думать, что множество следов на  $\Gamma$  функций из  $W_p^s(\Omega)$  заполняет все пространство  $L_q(\Gamma)$ . Более того, не всякая непрерывная на  $\Gamma$  функция, даже при сколь угодно гладкой  $\Gamma$ , может быть продолжена в  $\Omega$  так, чтобы она принадлежала  $W_p^s(\Omega)$ .

**Теорема 7.** Если  $\Gamma$  — липшицева кусочно непрерывно дифференцируемая поверхность, то для любой функции  $u \in W_p^s(\Omega)$  на  $\Gamma$  определены следы всех производных  $D^\alpha u$  при  $|\alpha| < s$ .

ДОКАЗАТЕЛЬСТВО. Ясно, что если  $|\alpha| < s$ , то  $D^\alpha u \in W_p^1(\Omega)$ , а при  $m = n - 1$ ,  $s = 1$  условие  $m > n - sp$  выполнено при любом  $p > 1$ .  $\square$

Отметим, что для функции  $u \in W_p^s(\Omega)$  справедлива формула интегрирования по частям:

$$\int_{\Omega} u \frac{\partial v}{\partial x_i} dx = - \int_{\Omega} \frac{\partial u}{\partial x_i} v dx + \int_{\Gamma} uv \cos(\widehat{\nu}, \widehat{x}_i) dx \quad \forall v \in C^1(\overline{\Omega}). \quad (14)$$

УПРАЖНЕНИЕ 2. Показать, что формула (14) справедлива для любых функций  $u, v \in W_2^1(\Omega)$ .

Используя понятие следа, дадим часто используемый в теории МКЭ критерий принадлежности функции пространству  $W_p^s$ .

**Теорема 8.** Пусть  $\overline{\Omega} = \overline{\Omega}_1 \cup \dots \cup \overline{\Omega}_m$ , где  $\Omega_i$  — подобласть  $\Omega$ ,  $i = 1, 2, \dots, m$ ,  $\Omega_i \cap \Omega_j = \emptyset$  при  $i \neq j$ ; каждая из областей  $\Omega_i$  имеет кусочно гладкую границу. Пусть далее  $u_i$  — сужение на  $\Omega_i$  функции  $u$ , определенной на области  $\Omega$ ,  $u_i \in W_p^s(\Omega_i)$ . Предположим, что для любых двух областей  $\Omega_i, \Omega_j$ , имеющих общий кусок  $\Gamma_{ij}$  границы ненулевой  $(n-1)$ -меры, следы на  $\Gamma_{ij}$  функций  $u_i, u_j$  и всех их производных вплоть до порядка  $s - 1$  совпадают. Тогда  $u \in W_p^s(\Omega)$ .

ДОКАЗАТЕЛЬСТВО. Ограничимся простейшим случаем, когда  $s = 1$ ,  $m = 2$ , предлагая читателю рассмотрение общей ситуации в качестве упражнения. Ясно, что  $\partial u / \partial x_k \in L_p(\Omega)$ ,  $k = 1, 2, \dots, n$ . Покажем, что эти функции — обобщенные производные функции  $u$ . Для

<sup>1)</sup> Достаточно предположить, что  $\Gamma_m \in C^s$

произвольной функции  $v \in C_0^\infty$ , используя формулу интегрирования по частям (14), получим:

$$\begin{aligned} \int_{\Omega} \frac{\partial u}{\partial x_i} v dx &= \int_{\Omega_1} \frac{\partial u_1}{\partial x_i} v dx + \int_{\Omega_2} \frac{\partial u_2}{\partial x_i} v dx = - \int_{\Omega} \frac{\partial v}{\partial x_i} u dx + \\ &+ \int_{\Gamma_{12}} u_1 v \cos(\widehat{\nu_1 x_i}) dx + \int_{\Gamma_{12}} u_2 v \cos(\widehat{\nu_2 x_i}) dx. \end{aligned} \quad (15)$$

Вследствие непрерывности функции  $v$  и равенства следов  $u_1, u_2$  на  $\Gamma_{12}$  интегралы по  $\Gamma_{12}$  равны по модулю и противоположны по знаку. Следовательно, равенство (15) совпадает с равенством, определяющим обобщенную производную функции  $u$ .  $\square$

Замыкание по норме (10) линейного множества функций  $C_0^\infty(\Omega)$  образует замкнутое подпространство пространства  $W_p^s(\Omega)$ , обозначаемое через  $\overset{\circ}{W}_p^s(\Omega)$ . При  $p = 2$  будем также использовать обозначение  $H_0^s(\Omega)$ .

Из определения следа и теоремы 7 вытекает, что если  $\Gamma$  — липшицева кусочно непрерывно дифференцируемая поверхность, то любая функция  $u \in \overset{\circ}{W}_p^s(\Omega)$  обращается в нуль на  $\Gamma$  вместе со всеми своими производными до порядка  $s - 1$  включительно.

### 9. Эквивалентные нормировки соболевских пространств.

На пространстве  $W_p^s(\Omega)$  часто бывает удобно вводить норму не при помощи соотношения (10) (или (11)), а некоторыми другими, эквивалентными, способами<sup>1)</sup>. Прежде всего заметим, что поскольку

$$\left( \sum_{|\alpha|=m} |D^\alpha u(x)|^p \right)^{1/p}, \quad p \geq 1,$$

есть норма симметричной полилинейной формы  $D^m u(x)$ , то норму на пространстве  $W_p^s(\Omega)$  можно определять эквивалентным образом при помощи равенства

$$\|u\|_{s,p,\Omega}^p = \int_{\Omega} \sum_{k=0}^s \|D^k u(x)\|^p dx,$$

где  $\|\cdot\|$  — любая норма полилинейной формы  $D^k u(x)$ ,  $k = 1, 2, \dots, s$ .

Значительно менее очевиден следующий результат о том, что в выражении для нормы пространства  $W_p^s$  можно не использовать смешанные производные функции  $u$ .

<sup>1)</sup>См. сноску на с. 22.

**Теорема 9.** Норма  $\|u\|'_{s,p,\Omega} = \|u\|_{0,p,\Omega} + [u]_{s,p,\Omega}$  эквивалентна на  $W_p^s(\Omega)$  норме, определенной равенством (10) (или (11)).

Весьма общий прием построения норм, эквивалентных исходной норме, дает следующая теорема С.Л. Соболева.

**Теорема 10.** Пусть  $f_0, f_1, \dots, f_k$  — линейные непрерывные функционалы на пространстве  $W_p^s(\Omega)$ ,  $s \geq 1$ ,  $p \in [1, \infty]$ , обладающие тем свойством, что из равенств

$$|u|_{r,p,\Omega} = 0, \quad r = l, l+1, \dots, s, \quad f_0(u) = f_1(u) = \dots = f_k(u) = 0$$

вытекает, что  $u = 0$ . Тогда выражение

$$\|u\|'_{s,p,\Omega} = |u|_{l,p,\Omega} + |u|_{l+1,p,\Omega} + \dots + |u|_{s,p,\Omega} + |f_0(u)| + \dots + |f_k(u)|$$

определяет норму на пространстве  $W_p^s(\Omega)$ , эквивалентную исходной.

**ДОКАЗАТЕЛЬСТВО.** Нетрудно проверить, выражение  $\|u\|'_{s,p}$  действительно определяет норму на пространстве  $W_p^s(\Omega)$ . Неравенство  $\|u\|'_{s,p} \leq c_1 \|u\|_{s,p} \quad \forall u \in W_p^s(\Omega)$ ,  $c_1 = \text{const}$ , — следствие непрерывности функционалов  $f_i$ ,  $i = 0, 1, \dots, k$ . Остается доказать справедливость обратного неравенства  $\|u\|_{s,p} \leq c_2 \|u\|'_{s,p} \quad \forall u \in W_p^s(\Omega)$ ,  $c_2 = \text{const}$ . Предположим противное. Тогда найдется такая последовательность  $\{u_m\}$ , что

$$\|u_m\|_{s,p} = 1, \quad m \geq 1, \tag{16}$$

и в то же время

$$\|u_m\|'_{s,p} \rightarrow 0. \tag{17}$$

По теореме 2 пространство  $W_p^s$  компактно вкладывается в пространство  $W_p^{s-1}$  при любых  $1 \leq p \leq \infty$ ,  $s \geq 1$ . Поэтому из последовательности  $\{u_m\}$  можно извлечь подпоследовательность  $\{u_{m_k}\}$ , фундаментальную в пространстве  $W_p^{s-1}$ , т. е. такую, что

$$\|u_{m_k} - u_{m'_k}\|_{s-1,p} \rightarrow 0 \quad \text{при } m_k, m'_k \rightarrow \infty. \tag{18}$$

Кроме того, из (17) следует, что  $|u_{m_k}|_{s,p} \rightarrow 0$ , поэтому  $|u_{m_k} - u_{m'_k}|_{s,p} \rightarrow 0$  при  $m_k, m'_k \rightarrow \infty$ . Вместе с (18) это означает, что последовательность  $\{u_{m_k}\}$  фундаментальна в  $W_p^s(\Omega)$  и потому имеет предел:  $u_{m_k} \rightarrow u$  в  $W_p^s(\Omega)$ . На основании (17), используя непрерывность функционалов  $f_i$  и полунорм  $|\cdot|_{j,p}$ , получим:

$$|u|_{j,p} = 0, \quad j = l, l+1, \dots, s, \quad f_j(u) = 0, \quad j = 0, 1, \dots, k,$$

и, значит,  $u = 0$ . С другой стороны, из (16) вытекает, что

$$\lim_{m_k \rightarrow \infty} \|u_{m_k}\|_{s,p} = \|u\|_{s,p} = 1.$$

Полученное противоречие и доказывает теорему.  $\square$

Аналогично теореме 10 с использованием теоремы 9 доказывается

**Теорема 11.** Пусть  $f_0, f_1, \dots, f_k$  — линейные непрерывные функционалы на  $W_p^s(\Omega)$ , обладающие тем свойством, что из равенств

$$[u]_{r,p,\Omega} = 0, \quad r = l, l+1, \dots, s, \quad f_0(u) = f_1(u) = \dots = f_k(u) = 0$$

вытекает, что  $u = 0$ . Тогда выражение

$$\|u\|'_{s,p,\Omega} = [u]_{l,p,\Omega} + [u]_{l+1,p,\Omega} + \dots + [u]_{s,p,\Omega} + |f_0(u)| + \dots + |f_k(u)|$$

определяет норму на пространстве  $W_p^s(\Omega)$ , эквивалентную исходной.

Приведем примеры, иллюстрирующие применение теорем об эквивалентных нормировках.

**1) Неравенство Фридрихса.** Используя теорему 6 при  $p = 2$ ,  $s = 1$ , получим, что для  $u \in W_2^1(\Omega)$  определен след на границе области  $\Omega$ , принадлежащий  $L_2(\Gamma)$ , причем  $\|u\|_{L_2(\Gamma)} \leq c \|u\|_1$ , поэтому можно определить линейный ограниченный на пространстве  $W_2^1(\Omega)$  функционал

$$f_0(u) = \int_{\Gamma} u \, dx.$$

Ясно, что если для  $u \in W_2^1(\Omega)$  выполнены равенства

$$|u|_1^2 = \int_{\Omega} |\nabla u|^2 \, dx = 0, \tag{19}$$

$$f_0(u) = 0, \tag{20}$$

то  $u(x) \equiv \text{const}$  вследствие (19), а в силу (20) имеем  $u(x) \equiv 0$ . Поэтому на основании теоремы 10 норма

$$\|u\|'_1 = |u|_1 + \left| \int_{\Gamma} u \, dx \right|$$

эквивалентна исходной норме

$$\|u\|_1 = (|u|_1^2 + |u|_0^2)^{1/2}$$

на пространстве  $W_2^1(\Omega)$ . Отсюда вытекает, что

$$\|u\|_{L_2(\Omega)}^2 \leq \|u\|_1^2 \leq c \left( \int_{\Omega} |\nabla u|^2 dx + \left| \int_{\Gamma} u dx \right|^2 \right).$$

В частности, для  $u \in \mathring{W}_2^1(\Omega)$  имеем:

$$\|u\|_{L_2(\Omega)}^2 \leq c \int_{\Omega} |\nabla u|^2 dx. \quad (21)$$

Неравенство (21) называется неравенством Фридрихса. Из него вытекает, что полунорма  $|\cdot|_1$  — норма на пространстве  $\mathring{W}_2^1(\Omega)$ , эквивалентная  $\|\cdot\|_1$ . Действительно, как следствие (21) немедленно получаем:

$$\|u\|_1^2 = \int_{\Omega} (|\nabla u|^2 + u^2) dx \leq c \int_{\Omega} |\nabla u|^2 dx = c|u|_1^2.$$

Оценка  $|u|_1 \leq \|u\|_1$  очевидна.

**ЗАМЕЧАНИЕ 4.** Неравенство Фридрихса (21) доказано нами как следствие теоремы об эквивалентных нормировках и общей теоремы вложения 10. При этом использовано предположение о том, что граница области  $\Omega$  липшицева. На самом деле неравенство Фридрихса справедливо для произвольной ограниченной области. Докажем это при  $n = 2$ . Случай пространства произвольной размерности рассматривается аналогично. Пусть область  $\Omega \subset R^2$  и ограничена. Можно считать, что  $\Omega \subset \Omega'$ , где  $\Omega'$  — квадрат со стороной  $a$ :  $\Omega' = \{x \in R^2 : 0 \leq x_1, x_2 \leq a\}$ . Для функции  $u \in C_0^\infty(\Omega)$  определим функцию  $\tilde{u} \in C_0^\infty(\Omega')$  соотношением

$$\tilde{u}(x) = \begin{cases} u(x), & x \in \Omega, \\ 0, & x \notin \Omega. \end{cases}$$

Пусть  $x \in \Omega'$ . Используя формулу Ньютона — Лейбница, можно написать:

$$\tilde{u}(x_1, x_2) = \int_0^{x_1} \frac{\partial \tilde{u}(\xi_1, x_2)}{\partial \xi_1} d\xi_1,$$

откуда в силу неравенства Коши — Буняковского

$$\tilde{u}^2(x_1, x_2) \leq a \int_0^a \left( \frac{\partial \tilde{u}(\xi_1, x_2)}{\partial \xi_1} \right)^2 d\xi_1.$$

Интегрируя последнее неравенство по  $\Omega'$ , получим:

$$\int_{\Omega'} \tilde{u}^2 dx \leq a^2 \int_{\Omega'} \left( \frac{\partial \tilde{u}}{\partial x_1} \right)^2 dx \leq a^2 \int_{\Omega'} |\nabla \tilde{u}|^2 dx,$$

или, так как  $\tilde{u}(x) \equiv 0$  вне  $\Omega$ ,

$$\int_{\Omega} u^2 dx \leq a^2 \int_{\Omega} |\nabla u|^2 dx \quad \forall u \in C_0^\infty. \quad (22)$$

Пусть теперь  $u \in \mathring{W}_2^1(\Omega)$ . По определению пространства  $\mathring{W}_2^1(\Omega)$  существует последовательность  $u_m \in C_0^\infty(\Omega)$  такая, что  $\|u - u_m\|_1 \rightarrow 0$  при  $m \rightarrow \infty$ . Ясно, что при этом

$$\int_{\Omega} u_m^2 dx \rightarrow \int_{\Omega} u^2 dx, \quad \int_{\Omega} |\nabla u_m|^2 dx \rightarrow \int_{\Omega} |\nabla u|^2 dx.$$

Записывая неравенство (22) для функций  $u_m$  и переходя к пределу при  $m \rightarrow \infty$ , получим, что (22) выполнено для любой функции  $u \in \mathring{W}_2^1(\Omega)$ .

**2) Неравенство Пуанкаре.** Пусть, как и в предыдущем примере,  $p = 2$ ,  $s = 1$ , а функционал  $f_0(u)$  определен на  $W_2^1(\Omega)$  соотношением

$$f_0(u) = \int_{\Omega} u dx.$$

Вследствие неравенства Коши — Буняковского

$$|f_0(u)| \leq c \left( \int_{\Omega} u^2 dx \right)^{1/2} \leq c \|u\|_1,$$

т. е.  $f_0(u)$  — линейный ограниченный функционал. Ясно, что, если

$$f_0(u) = 0, \quad |u|_1 = 0,$$

то  $u(x) \equiv 0$ , и по теореме 10 норма

$$\|u\|_1' = |u|_1 + \left| \int_{\Omega} u dx \right|$$

эквивалента норме пространства  $W_2^1(\Omega)$ . Отсюда вытекает, что

$$\int_{\Omega} u^2 dx \leq c \left\{ \int_{\Omega} |\nabla u|^2 dx + \left| \int_{\Omega} u dx \right|^2 \right\}.$$

Это неравенство называют неравенством Пуанкаре.

**3) Эквивалентные нормировки пространства  $H_0^2(\Omega)$ .** Если  $\Omega$  — ограниченная область с липшицевой границей, то нормы

$$\|u\|_{2,\Omega}^{(1)} = \left( \int_{\Omega} \sum_{|\alpha|=2} (D^\alpha u)^2 dx \right)^{1/2}, \quad \|u\|_{2,\Omega}^{(2)} = \left( \int_{\Omega} \sum_{i=1}^n \left( \frac{\partial^2 u}{\partial x_i^2} \right)^2 \right)^{1/2} dx,$$

$$\|u\|_{2,\Omega}^{(3)} = \left( \int_{\Omega} (\Delta u)^2 \right)^{1/2} dx,$$

определенные на  $H_0^2(\Omega)$ , эквивалентны норме пространства  $H^2(\Omega)$ .

Напомним, что  $H_0^2(\Omega)$  — подпространство пространства  $H^2(\Omega)$ , получающееся замыканием множества  $C_0^\infty(\Omega)$  по норме  $H^2(\Omega)$ . Понятно поэтому, что соответствующие неравенства эквивалентности достаточно установить для функций из  $C_0^\infty(\Omega)$ . Сделаем это сначала для нормы  $\|\cdot\|_{2,\Omega}^{(1)}$ . Неравенство  $\|u\|_{2,\Omega}^{(1)} \leq \|u\|_{2,\Omega}$  очевидно. Используя неравенство Фридрихса, получим

$$\int_{\Omega} u^2 dx \leq c \int_{\Omega} |\nabla u|^2 dx,$$

$$\int_{\Omega} \left( \frac{\partial u}{\partial x_i} \right)^2 dx \leq c \int_{\Omega} \sum_{j=1}^n \left( \frac{\partial^2 u}{\partial x_i \partial x_j} \right)^2 dx, \quad i = 1, \dots, n,$$

следовательно, обратное неравенство  $\|u\|_{2,\Omega} \leq c \|u\|_{2,\Omega}^{(1)}$  также доказано.

Сравним нормы  $\|u\|_{2,\Omega}^{(1)}$  и  $\|u\|_{2,\Omega}^{(2)}$ . Ясно, что  $\|u\|_{2,\Omega}^{(2)} \leq \|u\|_{2,\Omega}^{(1)}$ . Для того, чтобы оценить  $\|u\|_{2,\Omega}^{(1)}$  сверху, рассмотрим выражение

$$\int_{\Omega} \frac{\partial^2 u}{\partial x_i \partial x_j} \frac{\partial^2 u}{\partial x_i \partial x_j} dx$$

при  $i \neq j$  и преобразуем его при помощи формулы интегрирования по частям, перебрасывая производную по  $x_j$  на второй множитель, а затем производную по  $x_i$  — на первый. В результате получим:

$$\int_{\Omega} \frac{\partial^2 u}{\partial x_i \partial x_j} \frac{\partial^2 u}{\partial x_i \partial x_j} dx = \int_{\Omega} \frac{\partial^2 u}{\partial x_i^2} \frac{\partial^2 u}{\partial x_j^2} dx \leq \frac{1}{2} \int_{\Omega} \left( \left( \frac{\partial^2 u}{\partial x_i^2} \right)^2 + \left( \frac{\partial^2 u}{\partial x_j^2} \right)^2 \right) dx.$$

Эквивалентность норм  $\|u\|_{2,\Omega}^{(1)}$ ,  $\|u\|_{2,\Omega}^{(2)}$  доказана. Доказательство эквивалентности норм  $\|u\|_{2,\Omega}^{(1)}$ ,  $\|u\|_{2,\Omega}^{(3)}$  оставляем читателю в качестве упражнения.

**4) Леммы типа Брамбла — Гильберта.** В этом примере с помощью теорем 10, 11 мы докажем следующие часто используемые в теории МКЭ результаты.

**Лемма 1.** Пусть  $f$  — линейный ограниченный функционал на пространстве  $W_p^s(\Omega)$ , обращающийся в нуль на любом полиноме степени  $l \leq s - 1$  по совокупности переменных, т. е.  $f(P_l) = 0$ ,

если  $P_l(x) = \sum_{|\alpha| \leq l} c_\alpha x^\alpha$ , где  $c_\alpha$  — постоянные,  $x^\alpha = x_1^{\alpha_1} \cdots x_n^{\alpha_n}$ . Тогда существует такая постоянная  $c$ , зависящая только от области  $\Omega$  и функционала  $f$ , что

$$|f(u)| \leq c (|u|_{l+1,p} + |u|_{l+2,p} + \dots + |u|_{s,p}) \quad \forall u \in W_p^s(\Omega).$$

ДОКАЗАТЕЛЬСТВО. Введем в рассмотрение норму

$$\|u\|_s^* = |u|_{l+1,p} + |u|_{l+2,p} + \dots + |u|_{s,p} + \sum_{|\alpha| \leq l} \left| \int_{\Omega} D^\alpha u \, dx \right|$$

на пространстве  $W_p^s(\Omega)$ . Покажем, что она эквивалентна исходной норме. Для этого достаточно проверить, что линейные ограниченные функционалы

$$f_\alpha(u) = \int_{\Omega} D^\alpha u \, dx$$

удовлетворяют условиям теоремы 10. Пусть

$$|u|_{l+1,p} = |u|_{l+2,p} = \dots = |u|_{s,p} = 0, \quad (23)$$

$$\int_{\Omega} D^\alpha u \, dx = 0, \quad |\alpha| \leq l. \quad (24)$$

Из равенств (23) вытекает, что  $D^\alpha u = 0$  для всех  $\alpha$  таких, что  $|\alpha| = l+1$ . Это означает, что  $u(x)$  — полином степени  $l$  по совокупности переменных:  $u(x) = \sum_{|\alpha| \leq l} c_\alpha x^\alpha$ . Равенства (24) представляют собой систему ли-

нейных алгебраических уравнений относительно коэффициентов этого полинома. Покажем, что она имеет только тривиальное решение. Действительно, рассмотрим мультииндекс  $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_n)$  такой, что  $|\alpha| = l$ . Очевидно,  $D^\alpha u = c_\alpha \alpha_1! \alpha_2! \cdots \alpha_n!$ , и условие

$$\int_{\Omega} D^\alpha u \, dx = 0$$

дает, что  $c_\alpha = 0$ . Таким образом, все старшие коэффициенты полинома  $u(x)$ , т. е. коэффициенты  $c_\alpha$  при  $|\alpha| = l$ , равны нулю. Рассуждая аналогично, установим, что все коэффициенты полинома  $u(x)$  равны нулю, а это означает, что условия теоремы 10 выполнены, т. е. норма  $\|\cdot\|_s^*$  эквивалентна норме пространства  $W_p^s(\Omega)$ . Ясно теперь, что

для любой функции  $u \in W_p^s(\Omega)$  и любого полинома  $P_l(x)$  степени  $l$  по совокупности переменных

$$|f(u)| = |f(u + P_l)| \leq \|f\| \|u + P_l\|_{s,p} \leq c \|f\| \|u + P_l\|_{s,p}^*.$$

Полином  $P_l$  можно выбрать так, что

$$\int_{\Omega} D^{\alpha} (u + P_l) dx = 0, \quad |\alpha| \leq l.$$

Действительно, для этого достаточно решить систему линейных алгебраических уравнений

$$\int_{\Omega} D^{\alpha} P_l dx = - \int_{\Omega} D^{\alpha} u dx, \quad |\alpha| \leq l \quad (25)$$

относительно коэффициентов полинома  $P_l$ . Система (25) имеет решение, и притом единственное, поскольку соответствующая ей однородная система, как показано выше, имеет только тривиальное решение. При указанном выборе полинома  $P_l$  вследствие определения нормы  $\|\cdot\|_s^*$  имеем:

$$|f(u)| \leq c \|f\| \|u + P_l\|_{s,p}^* \leq c (|u|_{l+1,p} + |u|_{l+2,p} + \dots + |u|_{s,p}). \quad \square$$

Аналогично лемме 1 с использованием теоремы 11 об эквивалентных нормировках доказывается

**Лемма 2.** Пусть  $f$  — линейный ограниченный функционал на пространстве  $W_p^s(\Omega)$ , обращающийся в нуль на любом полиноме степени  $l \leq s - 1$  по каждой переменной, т. е.

$$f(P_l) = 0, \quad \text{если } P_l(x) = \sum_{\alpha_1, \dots, \alpha_n=0}^l c_{\alpha} x^{\alpha}, \quad c_{\alpha} - \text{постоянные.}$$

Тогда существует такая постоянная  $c$ , зависящая только от области  $\Omega$  и функционала  $f$ , что

$$|f(u)| \leq c ([u]_{l+1,p} + [u]_{l+2,p} + \dots + [u]_{s,p}) \quad \forall u \in W_p^s(\Omega).$$

**10. Замена переменных.** При исследовании метода конечных элементов важную роль играют формулы замены независимых переменных при вычислении производных и оценки, связанные с заменой независимых переменных в выражениях для норм на соболевских пространствах.

Пусть  $e, \hat{e}$  — ограниченные области пространства  $R^n$  с липшицевыми границами. Отображение  $x_e: \hat{e} \rightarrow e$  будем называть  $C^l$ -диффеоморфизмом, если  $x_e$  — непрерывное взаимно однозначное отображение  $\hat{e}$  на  $e$ , непрерывно дифференцируемое  $l$  раз на  $\hat{e}$ , обратное отображение  $x_e^{-1}$  непрерывно дифференцируемо  $l$  раз на  $e^{-1}$ .

Пусть  $l, m$  — целые неотрицательные числа. Положим

$$I(l, m) = \{i = (i_1, \dots, i_m) : i_1 + \dots + i_m = l, i_1 + 2i_2 + \dots + mi_m = m\},$$

где  $i_1, \dots, i_m$  — целые неотрицательные числа. В силу симметрии полилинейной формы, порожденной всеми производными функции  $u$  некоторого фиксированного порядка (см. с. 31), при  $i_1 + i_2 + \dots + i_m = l$  естественно использовать обозначение

$$D^l u(x)(h_1^{i_1}, \dots, h_m^{i_m}) = D^l u(x)(h_1, \dots, h_1, \dots, h_m, \dots, h_m);$$

в правой части равенства  $h_k$  повторяется  $i_k$  раз, или отсутствует, если  $i_k = 0, k = 1, \dots, m$ .

**Лемма 3.** Пусть  $x = x_e(\hat{x})$  есть  $C^m$ -диффеоморфизм области  $\hat{e} \subset R^n$  на область  $e \subset R^n, m \geq 1, u \in C^m(\bar{e}), \hat{u}(\hat{x}) = u(x_e(\hat{x}))$ . Тогда<sup>2)</sup>

$$D^m \hat{u}(\hat{x}) \hat{\xi}^m = m! \sum_{l=0}^m \sum_{i \in I(l, m)} \frac{1}{i_1! \dots i_m!} \frac{1}{(1!)^{i_1} \dots (m!)^{i_m}} \times \\ \times D^l u(x) \left( (Dx_e(\hat{x}) \hat{\xi})^{i_1}, \dots, (D^m x_e(\hat{x}) \hat{\xi}^m)^{i_m} \right), \quad (26)$$

$$\|D^m \hat{u}(\hat{x})\| \leq c \sum_{l=0}^m \|D^l u(x)\| \sum_{i \in I(l, m)} \|Dx_e(\hat{x})\|^{i_1} \dots \|D^m x_e(\hat{x})\|^{i_m}, \quad (27)$$

$$[D^m \hat{u}(\hat{x})] \leq c \sum_{l=0}^m \|D^l u(x)\| \sum_{i \in I(l, m)} [Dx_e(\hat{x})]^{i_1} \dots [D^m x_e(\hat{x})]^{i_m}, \quad (28)$$

если  $u \in C^m(\bar{e})$ . Здесь постоянная  $c$  зависит лишь от  $m$ .

**ДОКАЗАТЕЛЬСТВО.** Пусть для достаточно малого  $\hat{\xi}$  отрезок  $[\hat{x}, \hat{x} + \hat{\xi}]$  принадлежит  $\hat{e}$  и  $x = x_e(\hat{x})$ . Воспользуемся формулой Тей-

<sup>1)</sup>Если значение  $l$  явно не указано, отображение будем называть диффеоморфизмом.

<sup>2)</sup>Сумма  $\sum_{i \in I(l, m)}$  считается нулевой, если множество  $I(l, m)$  является пустым.

лора (5), с. 32:

$$\begin{aligned}\hat{u}(\hat{x} + \hat{\xi}) &= u(x_e(\hat{x} + \hat{\xi})) = \sum_{l=0}^m \frac{1}{l!} D^l u(x) (x_e(\hat{x} + \hat{\xi}) - x)^l + o(|\hat{\xi}|^m) = \\ &= \sum_{l=0}^m \frac{1}{l!} D^l u(x) (\Delta x_e)^l + o(|\hat{\xi}|^m).\end{aligned}\quad (29)$$

Здесь  $\Delta x_e = x_e(\hat{x} + \hat{\xi}) - x = x_e(\hat{x} + \hat{\xi}) - x_e(\hat{x})$ . При выводе (29) мы учли, что  $|x_e(\hat{x} + \hat{\xi}) - x| = o(|\hat{\xi}|)$ . В свою очередь,

$$\Delta x_e = \sum_{k=1}^m \frac{1}{k!} D^k x_e(\hat{x}) \hat{\xi}^k + o(|\hat{\xi}|^m).$$

Таким образом,

$$\begin{aligned}\hat{u}(\hat{x} + \hat{\xi}) &= \sum_{l=0}^m \frac{1}{l!} D^l u(x) \left( \sum_{k=1}^m \frac{1}{k!} D^k x_e(\hat{x}) \hat{\xi}^k + o(|\hat{\xi}|^m), \right. \\ &\quad \left. \dots, \sum_{k=1}^m \frac{1}{k!} D^k x_e(\hat{x}) \hat{\xi}^k + o(|\hat{\xi}|^m) \right) + o(|\hat{\xi}|^m).\end{aligned}$$

Раскроем скобки, используя полилинейность формы  $D^l u(x)$ . Полезно обратить внимание, что соответствующие преобразования формально следуют биномиальной формуле

$$(a_1 + a_2 + \dots + a_m)^l = \sum_{i_1+i_2+\dots+i_m=l} \frac{l!}{i_1!i_2!\dots i_m!} a_1^{i_1} a_2^{i_2} \dots a_m^{i_m}.$$

В итоге получим

$$\begin{aligned}\hat{u}(\hat{x} + \hat{\xi}) &= \sum_{l=0}^m \sum_{i_1+i_2+\dots+i_m=l} \frac{1}{i_1!i_2!\dots i_m!} \times \\ &\quad \times D^l u(x) \left( \left( \frac{1}{1!} D x_e(\hat{x}) \hat{\xi} \right)^{i_1}, \dots, \left( \frac{1}{m!} D^m x_e(\hat{x}) \hat{\xi}^m \right)^{i_m} \right) + o(|\hat{\xi}|^m).\end{aligned}\quad (30)$$

С другой стороны

$$\hat{u}(\hat{x} + \hat{\xi}) = \sum_{p=0}^m \frac{1}{p!} D^p \hat{u}(\hat{x}) \hat{\xi}^p + o(|\hat{\xi}|^m).$$

Сравнивая эту формулу с (30), получим искомое представление (26). Из этого представления, очевидно, вытекает неравенство

$$\begin{aligned} |D^m \hat{u}(\hat{x}) \hat{\xi}^m| &\leq c \sum_{l=0}^m \sum_{i \in I(l,m)} \|D^l u(x)\| |Dx_e(\hat{x}) \hat{\xi}|^{i_1} \dots |D^m x_e(\hat{x}) \hat{\xi}^m|^{i_m} \leq \\ &\leq c |\xi|^m \sum_{l=0}^m \|D^l u(x)\| \sum_{i \in I(l,m)} \|Dx_e(\hat{x})\|^{i_1} \dots \|D^m x_e(\hat{x})\|^{i_m} \quad (31) \end{aligned}$$

с постоянной  $c$ , зависящей только от  $m$ . Отсюда следует (27). Теперь, выбирая в первом неравенстве (31)  $\hat{\xi} = e_j$ , будем иметь

$$\|D^m \hat{u}(\hat{x}) e_j^m\| \leq c \sum_{l=0}^m \|D^l u(x)\| \sum_{i \in I(l,m)} |Dx_e(\hat{x}) e_j|^{i_1} \dots |D^m x_e(\hat{x}) e_j^m|^{i_m},$$

откуда после перехода к максимуму по всем  $j$  получим (28).  $\square$

**Следствие 1.** Лемма 3 справедлива и для вектор-функций  $u(x) = (u_1(x), \dots, u_n(x))^T$  таких, что  $u_i \in C^m(\bar{e})$ ,  $i = 1, \dots, n$ .

Для того, чтобы убедиться в этом, достаточно записать равенства (26) для каждой компоненты  $u$  и воспользоваться определением  $m$ -линейной формы  $D^m \hat{u} \in \mathcal{L}_m(R^n, R^n)$ .

**Лемма 4.** Пусть  $x = x_e(\hat{x})$  есть  $C^m$ -диффеоморфизм области  $\hat{e} \subset R^n$  на область  $e \subset R^n$ ,  $\hat{u}(\hat{x}) = u(x_e(\hat{x}))$ ,  $m \geq 0$ . Тогда

$$\begin{aligned} |\hat{u}|_{m,p,\hat{e}}^p &\leq c_1 \max_{x \in e} |\det(Dx_e^{-1}(x))| \sum_{l=0}^m |u|_{l,p,e}^p \times \\ &\times \sum_{i \in I(l,m)} \max_{\hat{x} \in \hat{e}} (\|Dx_e(\hat{x})\|^{p i_1} \dots \|D^m x_e(\hat{x})\|^{p i_m}) \quad (32) \end{aligned}$$

для всех  $u \in W_p^m(e)$ ,  $p \in [1, \infty)$ . Если  $u \in W_\infty^m(e)$ , то

$$|\hat{u}|_{m,\infty,\hat{e}} \leq c_2 \sum_{l=0}^m |u|_{l,\infty,e} \sum_{i \in I(l,m)} \max_{\hat{x} \in \hat{e}} (\|Dx_e(\hat{x})\|^{i_1} \dots \|D^m x_e(\hat{x})\|^{i_m}). \quad (33)$$

Здесь  $c_1 = c_1(m, n, p)$ ,  $c_2 = c_2(m, n)$ .

**ДОКАЗАТЕЛЬСТВО.** В силу плотности вложения  $C^\infty(\bar{e}) \subset W_p^m(e)$  достаточно считать, что  $u \in C^\infty(\bar{e})$ . Применяя неравенство Гельдера

для оценки правой части (27), будем иметь, что

$$\|D^m \hat{u}(\hat{x})\|^p \leq c \sum_{l=0}^m \|D^l u(x)\|^p \sum_{i \in I(l,m)} \|Dx_e(\hat{x})\|^{p^{i_1}} \dots \|D^m x_e(\hat{x})\|^{p^{i_m}}$$

с постоянной  $c = c(m, p)$ . После интегрирования этого неравенства по переменной  $\hat{x} \in \hat{e}$  получим

$$\begin{aligned} \int_{\hat{e}} \|D^m \hat{u}(\hat{x})\|^p d\hat{x} &\leq c \sum_{l=0}^m \int_{\hat{e}} \|D^l u(x)\|^p d\hat{x} \times \\ &\times \sum_{i \in I(l,m)} \max_{\hat{x} \in \hat{e}} (\|Dx_e(\hat{x})\|^{p^{i_1}} \dots \|D^m x_e(\hat{x})\|^{p^{i_m}}). \end{aligned}$$

Отсюда следует оценка (32), поскольку  $d\hat{x} = |\det(Dx_e^{-1}(x))| dx$ , а полунормы  $|u|_{m,p,e}$  и

$$\left( \int_e \|D^m u(x)\|^p dx \right)^{1/p}$$

эквивалентны (см. с. 45). Оценка (33) следует из (28).  $\square$

**Следствие 2.** *Поскольку отображения  $x_e$  и  $x_e^{-1}$  равноправно фигурируют в формулировке леммы 4, то*

$$\begin{aligned} |u|_{m,p,e}^p &\leq c \max_{\hat{x} \in \hat{e}} |\det(Dx_e(\hat{x}))| \sum_{l=0}^m |\hat{u}|_{l,p,\hat{e}}^p \times \\ &\times \sum_{i \in I(l,m)} \max_{x \in e} (\|Dx_e^{-1}(x)\|^{p^{i_1}} \dots \|D^m x_e^{-1}(x)\|^{p^{i_m}}), \quad (34) \end{aligned}$$

$$|u|_{m,\infty,e} \leq c \sum_{l=0}^m |\hat{u}|_{l,\infty,\hat{e}} \sum_{i \in I(l,m)} \max_{x \in e} (\|Dx_e^{-1}(x)\|^{i_1} \dots \|D^m x_e^{-1}(x)\|^{i_m}).$$

**Следствие 3.** *Если в условиях леммы 3 преобразование  $x_e$  — аффинное, т. е.  $x_e(\hat{x}) = B_e \hat{x} + b_e$ , где  $B_e$  — заданная матрица,  $b_e$  — заданный вектор, то при  $m \geq 0$ ,  $p \in [1, \infty]$*

$$\begin{aligned} D^m \hat{u}(\hat{x})(\hat{\xi}, \dots, \hat{\xi}) &= D^m u(x)(B_e \hat{\xi}, \dots, B_e \hat{\xi}), \\ |\hat{u}|_{m,p,\hat{e}} &\leq c |\det B_e|^{-1/p} \|B_e\|^m |u|_{m,p,e}, \\ |u|_{m,p,e} &\leq c |\det B_e|^{1/p} \|B_e^{-1}\|^m |\hat{u}|_{m,p,\hat{e}}. \end{aligned}$$

ДОКАЗАТЕЛЬСТВО. В формулах (26), (32), (34) остается лишь одно слагаемое с  $l = i_1 = m$ , поскольку для аффинных отображений

$$Dx_e(\hat{x}) \equiv B_e, \quad D^k x_e(\hat{x}) \equiv 0, \quad D^k x_e^{-1}(\hat{x}) \equiv 0 \quad \forall k \geq 2. \quad \square$$

**Лемма 5.** Пусть  $x = x_e(\hat{x})$  есть  $C^m$ -диффеоморфизм области  $\hat{e}$  на область  $e \in R^n$ ,  $u \in W_p^m(e)$ ,  $m \geq 0$ ,  $1 \leq p \leq \infty$ ,  $\hat{u}(\hat{x}) = u(x_e(\hat{x}))$ . Тогда

$$[\hat{u}]_{m,p,\hat{e}}^p \leq c_1 \max_{x \in e} |\det(Dx_e^{-1}(x))| \sum_{l=0}^m |u|_{l,p,e}^p \times \\ \times \sum_{i \in I(l,m)} \max_{\hat{x} \in \hat{e}} ([Dx_e(\hat{x})]^{p^{i_1}} \dots [D^m x_e(\hat{x})]^{p^{i_m}}),$$

$$[\hat{u}]_{m,\infty,\hat{e}} \leq c_2 \sum_{l=0}^m |u|_{l,\infty,e} \sum_{i \in I(l,m)} \max_{\hat{x} \in \hat{e}} ([Dx_e(\hat{x})]^{i_1} \dots [D^m x_e(\hat{x})]^{i_m}).$$

Здесь  $c_1 = c_1(n, m, p)$ ,  $c_2 = c_2(n, m)$ .

ДОКАЗАТЕЛЬСТВО. Рассуждения полностью аналогичны доказательству леммы 4. При этом надо учесть, что

$$[\hat{u}]_{m,p,\hat{e}}^p = \int_{\hat{e}} \sum_{i=1}^n |D^m \hat{u}(\hat{x}) e_i^m|^p d\hat{x} \leq n \int_{\hat{e}} [D^m \hat{u}(\hat{x})]^p d\hat{x},$$

$$[\hat{u}]_{m,\infty,\hat{e}} = \text{ess sup}_{\hat{x} \in \hat{e}} [D^m \hat{u}(\hat{x})]. \quad \square$$

**Лемма 6.** Пусть  $x = x_e(\hat{x})$  есть  $C^m$ -диффеоморфизм области  $\hat{e}$  на область  $e \in R^n$ . Тогда

$$Dx_e^{-1}(x) = (Dx_e(\hat{x}))^{-1},$$

и при  $m \geq 2$

$$D^m x_e^{-1}(x) \eta^m = -m! \sum_{l=1}^{m-1} \sum_{i \in I(l,m)} \frac{1}{i_1! \dots i_m!} \frac{1}{(1!)^{i_1} \dots (m!)^{i_m}} \times \\ \times D^l x_e^{-1}(x) \left( \eta^{i_1}, (D^2 x_e(\hat{x}) (Dx_e^{-1}(x) \eta)^2)^{i_2}, \dots, (D^m x_e(\hat{x}) (Dx_e^{-1}(x) \eta)^m)^{i_m} \right),$$

$$\begin{aligned} \max_{x \in e} \|D^m x_e^{-1}(x)\| &\leq c \sum_{l=1}^{m-1} \max_{\hat{x} \in \hat{e}} \|D^l x_e^{-1}(\hat{x})\| \sum_{i \in I(l,m)} \max_{x \in e} \|Dx_e^{-1}(x)\|^{m-i_1} \times \\ &\quad \times \max_{\hat{x} \in \hat{e}} (\|D^2 x_e(\hat{x})\|^{i_2} \dots \|D^m x_e(\hat{x})\|^{i_m}), \end{aligned}$$

где постоянная  $c$  зависит только от  $m$ .

**ДОКАЗАТЕЛЬСТВО.** Воспользуемся следствием 1 и выберем в формуле (26)  $u(x) = x_e^{-1}(x)$ . Очевидно, что  $\hat{u}(\hat{x}) = \hat{x}$ , а  $I(0, m) = \emptyset$ ,  $I(m, m) = (m, 0, \dots, 0)$  при  $m \geq 1$ . При  $m = 1$  получим

$$D\hat{u}(\hat{x}) = D\hat{x} = E, \quad \xi = Du(x)Dx_e(\hat{x})\xi \quad \forall \xi \in R^n$$

( $E$  — единичная матрица). Отсюда следует первое утверждение леммы. При  $m \geq 2$  имеем

$$\begin{aligned} 0 &= m! \sum_{l=1}^m \sum_{i \in I(l,m)} \frac{1}{i_1! \dots i_m!} \frac{1}{(1!)^{i_1} \dots (m!)^{i_m}} \times \\ &\quad \times D^l x_e^{-1}(x) \left( (Dx_e(\hat{x})\hat{\xi})^{i_1}, \dots, (D^m x_e(\hat{x})\hat{\xi}^m)^{i_m} \right). \end{aligned}$$

Из этого равенства вытекает второе утверждение леммы, если положить, что  $\eta = Dx_e(\hat{x})\hat{\xi}$ . Заключительное утверждение леммы есть следствие второго.  $\square$

### Задачи.

1. Пусть  $x = x_e(\hat{x})$  есть отображение области  $\hat{e} \subset R^n$  на область  $e \subset R^n$  и  $\hat{u}(\hat{x}) = u(x_e(\hat{x}))$ . Доказать, что

$$\begin{aligned} D^2 \hat{u}(\hat{x}) \xi^2 &= D^2 u(x) (Dx_e(\hat{x}) \xi)^2 + Du(x) (D^2 x_e(\hat{x}) \xi^2), \\ D^3 \hat{u}(\hat{x}) \xi^3 &= D^3 u(x) (Dx_e(\hat{x}) \xi)^3 + \\ &+ 3D^2 u(x) (D^2 x_e(\hat{x}) \xi^2, Dx_e(\hat{x}) \xi) + Du(x) (D^3 x_e(\hat{x}) \xi^3). \end{aligned}$$

2. Доказать, что

$$D^2 x_e^{-1}(x) \eta^2 = -Dx_e^{-1}(x) \left( D^2 x_e(\hat{x}) (Dx_e^{-1}(x) \eta)^2 \right),$$

$$\max_{x \in e} \|D^2 x_e^{-1}(x)\| \leq \max_{\hat{x} \in \hat{e}} \|D^2 x_e(\hat{x})\| \max_{x \in e} \|Dx_e^{-1}(x)\|^3.$$

### § 3. Обобщенные решения краевых задач

**1. Введение.** Понятие обобщенного решения естественным образом возникает при изучении вариационных постановок задач математической физики. Рассмотрим в качестве примера задачу о равновесии мембраны, жестко закрепленной по контуру и находящейся под действием внешней силы  $f$ . Эта задача может быть сформулирована как задача об отыскании функции  $u = u(x)$ , доставляющей минимальное значение интегралу (потенциальной энергии системы мембрана — внешние силы)

$$F(u) = \frac{1}{2} \int_{\Omega} T |\nabla u|^2 dx - \int_{\Omega} f u dx \quad (1)$$

на множестве функций, равных нулю на границе  $\Gamma$  области  $\Omega \subset R^2$ . Здесь  $T = T(x)$  — заданная непрерывно дифференцируемая функция — натяжение мембраны.

Уравнение равновесия мембраны получается как необходимое условие минимума (уравнение Эйлера) функционала (1) и имеет вид

$$-\operatorname{div}(T \nabla u) = f, \quad x \in \Omega. \quad (2)$$

Присоединяя к уравнению (2) условия жесткого закрепления

$$u(x) = 0, \quad x \in \Gamma, \quad (3)$$

получим граничную задачу Дирихле для отыскания функции  $u$ .

Под классическим решением этой задачи понимают функцию  $u$ , непрерывную в области  $\bar{\Omega}$ , дважды непрерывно дифференцируемую в открытой области  $\Omega$  и удовлетворяющую уравнению (2) и граничному условию (3).

Заметим, что функционал (1) определен на гораздо более широком множестве функций. Именно, он, очевидно, имеет смысл для любой функции  $u \in H^1(\Omega)$ . Учтем граничные условия (3), полагая, что  $u \in H_0^1(\Omega)$ . Таким образом, возникает задача об отыскании минимума функционала  $F$  на пространстве  $H_0^1(\Omega)$ :

$$F(u) = \min_{v \in H_0^1} F(v). \quad (4)$$

Решение этой задачи естественно назвать обобщенным решением задачи Дирихле (2), (3).

В дальнейшем будет показано, что задача (4) имеет единственное решение при любой функции  $f \in L_2(\Omega)$ .

Обобщенной постановке задачи (2), (3) можно придать и другую, эквивалентную, зачастую более удобную, форму. Заметим, если  $u$  — обобщенное решение задачи (2), (3), то для любой функции  $\eta \in H_0^1$  и любого вещественного  $t$  имеем

$$F(u + t\eta) \geq F(u).$$

Иными словами, функция вещественного переменного

$$\varphi(t) = F(u + t\eta)$$

достигает минимального значения при  $t = 0$ . Нетрудно видеть, что  $\varphi$  — дифференцируемая функция, следовательно,  $\varphi'(0) = 0$ . Легко подсчитать, что

$$\varphi'(0) = \int_{\Omega} T \nabla u \cdot \nabla \eta \, dx - \int_{\Omega} f \eta \, dx,$$

поэтому, если  $u$  — решение задачи (4), то

$$\int_{\Omega} T \nabla u \cdot \nabla \eta \, dx = \int_{\Omega} f \eta \, dx \quad \forall \eta \in H_0^1. \quad (5)$$

Теперь можно сформулировать определение, эквивалентное введенному ранее определению обобщенного решения задачи (2), (3). Именно, будем говорить, что функция  $u \in H_0^1(\Omega)$  — обобщенное решение задачи (2), (3), если выполнено тождество (5).

Заметим, что к тождеству (5) можно придти и другим путем, удобным с точки зрения распространения понятия обобщенного решения на более сложные задачи. А именно, считая, что  $u$  — классическое решение задачи (2), (3), умножим уравнение (2) на некоторую непрерывно дифференцируемую на  $\Omega$  и равную нулю на  $\Gamma$  функцию  $\eta$  и проинтегрируем полученное равенство по области  $\Omega$ :

$$- \int_{\Omega} \eta \operatorname{div}(T \nabla u) \, dx = \int_{\Omega} f \eta \, dx. \quad (6)$$

Преобразуем интеграл в левой части равенства (6) при помощи формулы интегрирования по частям, учитывая при этом граничное условие для функции  $\eta$ . В результате придем к тождеству вида (5).

Таким образом, если  $u$  — классическое решение задачи (2), (3), для которого существует интеграл

$$\int_{\Omega} T |\nabla u|^2 \, dx$$

(называемый интегралом Дирихле), то оно удовлетворяет интегральному тождеству (5). В этом смысле классическое решение задачи (2), (3) является и обобщенным решением.

Пусть теперь функция  $u$ , обобщенное решение задачи (2), (3), дважды непрерывно дифференцируема в  $\Omega$ . Покажем, что тогда  $u$  есть классическое решение задачи (2), (3). Действительно, выбирая в качестве функции  $\eta$  в (5) непрерывную и непрерывно дифференцируемую функцию, равную нулю на  $\Gamma$ , при помощи формулы интегрирования по частям получим

$$\int_{\Omega} (-\operatorname{div}(T\nabla u) - f) \eta \, dx = 0. \quad (7)$$

Отсюда следует, что  $u$  — решение уравнения (2). Если предположить противное, то найдется точка  $x \in \Omega$ , такая, что

$$(\operatorname{div}(T\nabla u) + f)(x) \neq 0.$$

Не ограничивая общности, можно считать, что

$$(\operatorname{div}(T\nabla u) + f)(x) > 0.$$

В силу непрерывности функции  $\operatorname{div}(T\nabla u) + f$  это неравенство сохраняется и для некоторой окрестности точки  $x$ , лежащей внутри  $\Omega$ . Выберем теперь функцию  $\eta$  так, чтобы она была положительной внутри указанной окрестности и тождественно равной нулю вне этой окрестности. Тогда интеграл в левой части равенства (7) будет положительным, что по доказанному невозможно ни при какой функции  $\eta$ , следовательно,  $\operatorname{div}(T\nabla u) + f = 0$  всюду в области  $\Omega$ , т. е.  $u$  — классическое решение задачи (2), (3).

Приведенные рассуждения показывают, что данное здесь определение обобщенного решения естественно обобщает понятие классического решения.

**2. Обобщенные решения краевых задач для эллиптических уравнений второго порядка.** Пусть  $\Omega \subset R^n$  — ограниченная область с липшицевой границей  $\Gamma$ . Будем искать функцию

$$u \in C^2(\Omega) \cap C(\bar{\Omega}), \quad (8)$$

удовлетворяющую уравнению

$$-\sum_{i,j=1}^n \frac{\partial}{\partial x_i} \left( a_{ij} \frac{\partial u}{\partial x_j} \right) + \sum_{i=1}^n a_i \frac{\partial u}{\partial x_i} + a_0 u = f, \quad x \in \Omega, \quad (9)$$

и граничному условию

$$u(x) = 0, \quad x \in \Gamma. \quad (10)$$

Задача (9), (10) называется первой краевой задачей, или задачей Дирихле.

Относительно коэффициентов и правой части уравнения (9) будем предполагать выполненными следующие условия:

$$a_{ij} \in C^1(\Omega), \quad i, j = 1, \dots, n, \quad a_i \in C(\Omega), \quad i = 0, \dots, n, \quad f \in C(\Omega). \quad (11)$$

Вводя в рассмотрение матрицу  $A(x) = \{a_{ij}\}_{i,j=1}^n$  и вектор-строку  $a = (a_1, \dots, a_n)$ , уравнению (9) можно придать более компактный вид

$$-\operatorname{div}(A\nabla u) + a\nabla u + a_0u = f, \quad x \in \Omega.$$

В дальнейшем предполагается выполненным условие эллиптичности. Это означает, что матрица  $A$  равномерно по  $x \in \bar{\Omega}$  положительно определена:

$$A(x)\xi \cdot \xi \geq c_0 |\xi|^2 \quad \forall \xi \in R^n, \quad \forall x \in \bar{\Omega}, \quad c_0 = \operatorname{const} > 0. \quad (12)$$

При ограниченности всех коэффициентов  $a_{ij}$  в  $\bar{\Omega}$  справедливо, очевидно, и обратное неравенство

$$A(x)\xi \cdot \xi \leq c_1 |\xi|^2 \quad \forall \xi \in R^n, \quad \forall x \in \bar{\Omega}.$$

Сформулируем определение обобщенного решения задачи (9), (10). Рассуждая по аналогии с пунктом 1, умножим уравнение на функцию  $\eta$ , удовлетворяющую условию (10) и условиям гладкости (8), и преобразуем интеграл, содержащий старшие производные, при помощи формулы интегрирования по частям. Получим

$$\begin{aligned} \int_{\Omega} (A\nabla u \cdot \nabla \eta + a\nabla u \eta + a_0u\eta) dx = \\ = \int_{\Omega} f\eta dx + \int_{\Gamma} (A\nabla u \cdot \nu) \eta dx, \end{aligned} \quad (13)$$

где  $\nu$  — единичный вектор внешней нормали к  $\Gamma$ . Второе слагаемое справа, очевидно, обращается в нуль за счет граничного условия для функции  $\eta$ . Таким образом, если функция  $u$  — решение задачи (9), (10), то для любой функции  $\eta$ , удовлетворяющей условиям гладкости (8) и граничному условию (10),

$$\int_{\Omega} (A\nabla u \cdot \nabla \eta + a\nabla u \eta + a_0u\eta) dx = \int_{\Omega} f\eta dx. \quad (14)$$

Равенство (14) называется интегральным тождеством, соответствующим задаче (9), (10). Его естественно положить в основу определения обобщенного решения. Понятно, что (14) сохраняет смысл для любых функций  $u, \eta \in H^1(\Omega)$ . Условия на коэффициенты и правую часть уравнения также могут быть ослаблены. Именно, достаточно считать, что

$$a_{ij}, a_i \in L_\infty(\Omega), \quad f \in L_2(\Omega). \quad (15)$$

**Определение 1.** Функция  $u \in H_0^1$  называется обобщенным решением задачи (9), (10), если для любой функции  $\eta \in H_0^1$  выполнено интегральное тождество (14).

Точно так же, как в пункте 1, показывается, что, если обобщенное решение существует и выполнены условия гладкости (8)–(11), то оно будет классическим решением задачи (9), (10).

Исследуем существование и единственность обобщенного решения задачи (9), (10). Определим линейную форму  $f : H_0^1(\Omega) \rightarrow R$  и билинейную форму  $a(\cdot, \cdot) : H_0^1(\Omega) \times H_0^1(\Omega) \rightarrow R$  при помощи соотношений

$$f(v) = \int_{\Omega} f v dx \quad \forall v \in H_0^1(\Omega),$$

$$a(u, \eta) = \int_{\Omega} (A \nabla u \cdot \nabla \eta + a \nabla u \eta + a_0 u \eta) dx \quad \forall u, v \in H_0^1(\Omega). \quad (16)$$

**Лемма 1.** Пусть выполнены условия (15). Тогда форма  $a(\cdot, \cdot)$  непрерывна.

**ДОКАЗАТЕЛЬСТВО.** Для любых  $u, v \in H_0^1(\Omega)$  имеем

$$|a(u, v)| \leq \int_{\Omega} (c_1 |\nabla u| |\nabla v| + c_2 |\nabla u| |v| + c_3 |u| |v|) dx.$$

Используем неравенство Коши — Буняковского, а затем — Фридрихса. Получим

$$|a(u, v)| \leq c_1 |u|_1 |v|_1 + c_2 |u|_1 |v|_0 + c_3 |u|_0 |v|_0 \leq c |u|_1 |v|_1,$$

где  $c = c_1 + c_2 c_\Omega^{1/2} + c_3 c_\Omega$ . Здесь  $c_\Omega$  — постоянная из неравенства Фридрихса,

$$c_2 = \operatorname{ess\,sup}_{x \in \Omega} |a(x)|, \quad c_3 = \operatorname{ess\,sup}_{x \in \Omega} |a_0(x)|. \quad \square$$

Аналогично при помощи неравенств Коши — Буняковского и Фридрихса доказывается, что для любой функции  $f \in L_2(\Omega)$  линейная форма  $f$  непрерывна.

Для формулировки условий положительной определенности формы  $a(\cdot, \cdot)$  представим функцию  $a_0$  в виде  $a_0(x) = a_0^+(x) + a_0^-(x)$ , где  $a_0^+(x) = (a_0(x) + |a_0(x)|) / 2 \geq 0$  — положительная часть  $a_0$ ,  $a_0^-(x) = (a_0(x) - |a_0(x)|) / 2 \leq 0$  — отрицательная часть  $a_0$ .

**Лемма 2.** Пусть выполнено условие (12),

$$|a_0^-| \leq c_4, \quad c_0 - c_2 c_\Omega^{1/2} - c_4 c_\Omega > 0. \quad (17)$$

Тогда форма  $a(\cdot, \cdot)$  положительно определена.

**ДОКАЗАТЕЛЬСТВО.** Полагая в равенстве (16)  $\eta = u$ , получим

$$\begin{aligned} a(u, u) &\geq \int_{\Omega} \left( c_0 |\nabla u|^2 - c_2 |\nabla u| |u| - c_4 |u|^2 \right) dx \geq \\ &\geq \left( c_0 - c_2 c_\Omega^{1/2} - c_4 c_\Omega \right) |u|_1^2. \quad \square \end{aligned}$$

Из лемм 1, 2 и леммы Лакса — Мильграма непосредственно вытекает

**Теорема 1.** Пусть выполнены условия (12), (15), (17). Тогда задача (9), (10) имеет единственное обобщенное решение при любой правой части.

**ЗАМЕЧАНИЕ 1.** Мы доказали однозначную обобщенную разрешимость задачи (9), (10) при условии, что младшие члены уравнения достаточно малы. Привлекая теорию Фредгольма разрешимости уравнений с вполне непрерывными операторами, можно указать значительно более общие критерии существования решения этой задачи.

Особо остановимся на важном для приложений случае, когда форма  $a(\cdot, \cdot)$ , соответствующая задаче (9), (10), симметрична. Как показывает соотношение (16), для этого достаточно потребовать выполнения условий

$$A = A^T, \quad a = 0.$$

Если при этом

$$c_0 - c_4 c_\Omega > 0,$$

то форма  $a(\cdot, \cdot)$  положительно определена, и, как показано в п. 2, задача отыскания обобщенного решения задачи (9), (10) эквивалентна задаче минимизации

$$F(u) = \min_{u \in H_0^1(\Omega)} F(v),$$

где

$$F(u) = \frac{1}{2} \int_{\Omega} (A \nabla u \cdot \nabla u + a_0 u^2) dx - \int_{\Omega} f u dx. \quad (18)$$

Функционал (18) принято называть энергетическим функционалом.

Аналогично задаче (9)–(10) формулируется и исследуется неоднородная задача Дирихле. Пусть функция  $g$  определена на  $\Gamma$  и допускает продолжение в  $H^1(\Omega)$  в том смысле, что существует функция  $v \in H^1(\Omega)$  такая, что след ее на  $\Gamma$  совпадает с  $g$ .

**Определение 2.** Функция  $u \in H^1(\Omega)$  — обобщенное решение неоднородной задачи Дирихле, если  $u(x) = g(x)$ ,  $x \in \Gamma$ , и

$$\int_{\Omega} (A \nabla u \cdot \nabla \eta + a \nabla u \cdot \eta + a_0 u \eta) dx = \int_{\Omega} f \eta dx \quad \forall \eta \in H_0^1(\Omega).$$

Возможна постановка и других краевых условий для уравнения (9).

Будем считать, что область  $\Omega$  принадлежит классу  $C^1$  и рассмотрим так называемую третью краевую задачу. Под классическим решением этой задачи понимается функция  $u \in C^1(\overline{\Omega}) \cap C^2(\Omega)$ , удовлетворяющая уравнению (9) и граничному условию

$$A \nabla u \cdot \nu + \sigma u = g, \quad x \in \Gamma. \quad (19)$$

Здесь  $\sigma = \sigma(x)$ ,  $g = g(x)$  — заданные функции,  $\nu$ , как и выше, — единичный вектор внешней нормали к  $\Gamma$ . Предполагаются выполненными условия гладкости (11), функции  $\sigma$ ,  $g$  считаются непрерывными.

**Замечание 2.** Выражение  $A \nabla u \cdot \nu$  называется конормальной производной функции  $u$ . В простейшем случае, когда  $A = E$  — единичная матрица, т. е. старшие производные уравнения (9) образуют оператор Лапласа

$$\Delta u = \frac{\partial^2 u}{\partial x_1^2} + \frac{\partial^2 u}{\partial x_2^2} + \cdots + \frac{\partial^2 u}{\partial x_n^2},$$

это выражение совпадает с нормальной производной функции  $u$ .

Введем понятие обобщенного решения задачи (9), (19). С этой целью рассмотрим соотношение (13), которому удовлетворяет решение уравнения (9) при любой функции  $\eta$ , и учтем граничное условие (19). Соотношение (13) примет вид

$$\int_{\Omega} (A\nabla u \cdot \nabla \eta + a\nabla u \eta + a_0 u \eta) dx + \int_{\Gamma} \sigma u \eta dx = \\ = \int_{\Omega} f \eta dx + \int_{\Gamma} g \eta dx. \quad (20)$$

Ясно, что оно имеет смысл при любых  $u, \eta \in H^1(\Omega)$ , поскольку из теоремы 6 при  $p = 2$ ,  $s = 1$  следует, что  $H^1(\Omega)$  непрерывно вкладывается в  $L_2(\Gamma)$ . Условия на исходные данные также можно ослабить. Именно, будем считать, что выполнены условия (15), и, кроме того,  $\sigma \in L_{\infty}(\Gamma)$ ,  $g \in L_2(\Gamma)$ .

**Определение 3.** Обобщенным решением задачи (9), (19) назовем функцию  $u \in H^1(\Omega)$ , удовлетворяющую интегральному тождеству (20) при любой функции  $\eta \in H^1(\Omega)$ .

Аналогично случаю первой краевой задачи можно показать, что при выполнении соответствующих условий гладкости обобщенное решение задачи удовлетворяет уравнению (9) и граничному условию (19) в классическом смысле.

**Замечание 3.** Следует обратить внимание на принципиальную разницу между граничными условиями первого и третьего рода. Граничные условия третьего рода вытекают как следствие из интегрального тождества (20), в то время как граничные условия Дирихле накладываются на искомое решение априори. В связи с этим принята такая терминология: граничные условия Дирихле называют главными граничными условиями, граничные условия третьего рода — естественными.

При исследовании обобщенной разрешимости задачи (9), (19) будем предполагать выполненными условия (12), (17). Кроме того предположим, что  $a_0(x) \geq m_0 = \text{const} \geq 0$ ,  $\sigma(x) \geq m_1 = \text{const} \geq 0$ , причем  $m_0 + m_1 > 0$ .

**Теорема 2.** При сформулированных выше условиях задача (9), (19) имеет единственное обобщенное решение, если постоянная  $c_2$  достаточно мала.

**ДОКАЗАТЕЛЬСТВО.** Заметим прежде всего, что вследствие теоремы 6

$$\left| \int_{\Gamma} \sigma u \eta dx \right| \leq c_{1\Omega} \|u\|_1 \|\eta\|_1,$$

где  $c_{1\Omega}$  — постоянная, зависящая от области  $\Omega$  и от  $\sigma$ . Отсюда выте-

кает, что билинейная форма  $a(\cdot, \cdot) : H^1(\Omega) \times H^1(\Omega) \rightarrow R$ ,

$$a(u, v) = \int_{\Omega} (A \nabla u \cdot \nabla v + a \nabla u \cdot v + a_0 u v) dx + \int_{\Gamma} \sigma u v dx,$$

непрерывна. Функционал

$$f(\eta) = \int_{\Omega} f \eta dx + \int_{\Gamma} g \eta dx$$

есть линейный ограниченный функционал на  $H^1(\Omega)$ . Таким образом, задача отыскания обобщенного решения задачи (9), (19) эквивалентна задаче отыскания функции  $u \in H^1(\Omega)$  такой, что

$$a(u, \eta) = f(\eta) \quad \forall \eta \in H^1(\Omega),$$

и для доказательства ее однозначной разрешимости осталось проверить, что форма  $a(\cdot, \cdot)$  положительно определена. Будем считать, что  $m_1 > 0$ . Случай, когда функция  $a_0$  строго положительна, рассматривается вполне аналогично. Рассуждая, как в случае первой краевой задачи, получим

$$a(u, u) \geq c_0 |u|_{1,\Omega}^2 - c_2 |u|_{1,\Omega} |u|_{0,\Omega} + m_1 |u|_{0,\Gamma}^2. \quad (21)$$

Нетрудно проверить, что норма

$$\|u\|_1^* = \left( |u|_{1,\Omega}^2 + \int_{\Gamma} u^2 dx \right)^{1/2}$$

эквивалентна исходной норме на пространстве  $H^1(\Omega)$ , следовательно,  $|u|_{0,\Omega} \leq c_{\Omega} \|u\|_1^*$  с постоянной  $c_{\Omega}$ , зависящей лишь от области  $\Omega$ . Поэтому, усиливая неравенство (21), получим  $a(u, u) \geq (c - c_2 c_{\Omega}) \|u\|_1^{*2}$ , где  $c = \max(c_0, m_1) > 0$ , причем  $c - c_2 c_{\Omega} > 0$ , если постоянная  $c_2$  достаточно мала.  $\square$

**3. Краевые задачи для систем уравнений теории упругости.** В качестве важного для приложений примера систем линейных уравнений с частными производными, исследование разрешимости которых проходит во многом по той же схеме, что и для одного эллиптического уравнения, приведем классическую систему уравнений линейной теории упругости.

Пусть  $\Omega \subset R^3$  — область, занимаемая упругим телом, отнесенная к декартовой системе координат  $x_1, x_2, x_3$ . Разыскивается вектор-функция  $u = (u_1, u_2, u_3)^T$  — вектор смещений — как решение системы

уравнений равновесия упругого тела

$$\sum_{j=1}^3 \frac{\partial \sigma_{ij}}{\partial x_j} + f_i = 0, \quad i = 1, 2, 3, \quad x \in \Omega. \quad (22)$$

Здесь  $\sigma_{ij} = \sigma_{ji}$ ,  $i, j = 1, 2, 3$ , — компоненты тензора напряжений,  $f = (f_1, f_2, f_3)^T$  — вектор внешних объемных сил, действующих на упругое тело. Компоненты тензора напряжений связаны законом Гука с компонентами тензора деформаций

$$\sigma_{ij} = \sigma_{ij}(u) = \sum_{k,l=1}^3 a_{ijkl} \varepsilon_{kl}, \quad (23)$$

где  $a_{ijkl} = a_{ijkl}(x)$  — заданные функции, определяющие упругие свойства тела,  $\varepsilon_{kl}$  — компоненты тензора деформаций — вычисляются через компоненты вектора смещений по формулам

$$\varepsilon_{ij} = \varepsilon_{ij}(u) = \frac{1}{2} \left( \frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} \right), \quad i, j = 1, 2, 3. \quad (24)$$

Таким образом, система (22)–(24) есть система трех линейных уравнений с частными производными относительно компонент вектора смещений  $u$ .

К системе (22)–(24) нужно добавить граничные условия. Мы рассмотрим условия двух типов.

1) Граничные условия Дирихле (граничные условия первого рода):

$$u(x) = 0, \quad x \in \Gamma. \quad (25)$$

Эти условия означают, что граница тела жестко закреплена, т. е. ее точки при деформации тела не смещаются.

2) Граничные условия второго рода:

$$\sum_{j=1}^3 \sigma_{ij} \cos(\nu, x_j) = g_i, \quad x \in \Gamma, \quad (26)$$

где  $g_i(x)$ ,  $x \in \Gamma$  — заданные функции. Условие (26) означает, что на  $\Gamma$  заданы поверхностные силы с плотностью  $g = (g_1, g_2, g_3)^T$  (вектор напряжений).

В дальнейшем будет рассматриваться тот случай, когда граница области разбита на две части:  $\Gamma = \Gamma_u \cup \Gamma_\sigma$ , на  $\Gamma_u$  задано условие (25),

а на  $\Gamma_\sigma$  — условие (26). Возникающая при этом задача называется смешанной граничной задачей для системы уравнений теории упругости.

Относительно исходных данных задачи предположим, что

$$f \in [C(\Omega)]^3, \quad g \in [C(\Gamma)]^3, \quad a_{ijkl} \in C^1(\Omega),$$

коэффициенты  $a_{ijkl}$  удовлетворяют условиям симметрии

$$a_{ijkl} = a_{jikl} = a_{klij}, \quad i, j, k, l = 1, 2, 3,$$

порождаемая ими квадратичная форма равномерно положительно определена:

$$\sum_{i,j,k,l=1}^3 a_{ijkl} \xi_{ij} \xi_{kl} \geq c_0 \sum_{i,j=1}^3 \xi_{ij}^2, \quad c_0 = \text{const} > 0, \quad (27)$$

при любых вещественных  $\xi_{ij}$  таких, что  $\xi_{ij} = \xi_{ji}$ ,  $i, j = 1, 2, 3$ ,  $x \in \bar{\Omega}$ .

ЗАМЕЧАНИЕ 4. В простейшем случае, когда упругое тело однородно и изотропно, все коэффициенты в законе Гука определяются лишь двумя постоянными, и соотношения (23) принимают вид

$$\sigma_{ij} = \lambda \left( \sum_{i=1}^3 \varepsilon_{ii} \right) \delta_{ij} + 2\mu \varepsilon_{ij}, \quad \delta_{ij} = \begin{cases} 1, & i = j, \\ 0, & i \neq j, \end{cases}$$

где  $\lambda$ ,  $\mu$  — положительные числа (постоянные Ламе). Соответствующая квадратичная форма

$$\sum_{i,j=1}^3 \sigma_{ij} \varepsilon_{ij} = \lambda \left( \sum_{i=1}^3 \varepsilon_{ii} \right)^2 + 2\mu \sum_{i,j=1}^3 \varepsilon_{ij}^2,$$

очевидно, положительно определена,  $c_0 = 2\mu$ .

Для построения интегрального тождества, соответствующего смешанной граничной задаче, умножим  $i$ -тое уравнение системы (22) на функцию  $\eta_i$ , равную нулю на  $\Gamma_u$ , просуммируем полученные равенства, проинтегрируем по области  $\Omega$  и преобразуем интегралы, содержащие производные от  $\sigma_{ij}$ , при помощи формул интегрирования по частям с учетом граничных условий на  $\Gamma_u$  и  $\Gamma_\sigma$  и симметрии тензора напряжений. Получим

$$\int_{\Omega} \sum_{i,j=1}^3 \sigma_{ij}(u) \varepsilon_{ij}(\eta) dx = \int_{\Omega} f \cdot \eta dx + \int_{\Gamma_\sigma} g \cdot \eta dx. \quad (28)$$

Будем рассматривать пространство

$$V = \left\{ u = (u_1, u_2, u_3)^T \in [H^1(\Omega)]^3 : \text{след } u_i \text{ на } \Gamma_u \text{ равен нулю,} \right. \\ \left. i = 1, 2, 3 \right\}$$

как гильбертово пространство с нормой  $|u|_1 = (|u_1|_1^2 + |u_2|_1^2 + |u_3|_1^2)^{1/2}$ .

**Определение 4.** Функция  $u \in V$  — обобщенное решение задачи смешанной граничной задачи линейной теории упругости, если она удовлетворяет тождеству (28) при любой функции  $\eta \in V$ .

**Теорема 3.** Обобщенное решение смешанной граничной задачи линейной теории упругости при  $\text{mes } \Gamma_u > 0$  существует и единственно.

ДОКАЗАТЕЛЬСТВО. Определим билинейную форму  $a : V \times V \rightarrow R$ :

$$a(u, \eta) = \int_{\Omega} \sum_{i,j=1}^3 \sigma_{ij}(u) \varepsilon_{ij}(\eta) dx$$

и линейный функционал  $f : V \rightarrow R$ :

$$f(\eta) = \int_{\Omega} f \cdot \eta dx + \int_{\Gamma_{\sigma}} g \cdot \eta dx.$$

Ясно, что форма  $a(\cdot, \cdot)$  симметрична и ограничена, функционал  $f$  непрерывен. Форма  $a(\cdot, \cdot)$  положительно определена. Мы установим этот факт в частном случае, когда  $\Gamma_u = \Gamma$ , т. е., когда для системы уравнений (22)–(24) рассматривается задача Дирихле. В силу (27) имеем

$$a(u, u) \geq c_0 \int_{\Omega} \sum_{i,j=1}^3 \varepsilon_{ij}^2 dx.$$

Покажем, что

$$\int_{\Omega} \sum_{i,j=1}^3 \varepsilon_{ij}^2 dx \geq c |u|_1^2. \quad (29)$$

Тогда форма  $a(\cdot, \cdot)$  положительно определена. Неравенство (29) называется неравенством Корна. Понятно, что его достаточно проверить лишь для функций  $u \in [C_0^{\infty}(\Omega)]^3$ . Для функций из  $V$  оно затем может быть установлено предельным переходом. Для любого  $\delta \in (0, 1)$  имеем

$$I \equiv \int_{\Omega} \sum_{i,j=1}^3 \varepsilon_{ij}^2 dx \geq \sum_{i=1}^3 \int_{\Omega} \left( \frac{\partial u_i}{\partial x_i} \right)^2 dx +$$

$$\begin{aligned}
& + \frac{\delta}{2} \int_{\Omega} \left( \left( \frac{\partial u_1}{\partial x_2} + \frac{\partial u_2}{\partial x_1} \right)^2 + \left( \frac{\partial u_1}{\partial x_3} + \frac{\partial u_3}{\partial x_1} \right)^2 + \left( \frac{\partial u_2}{\partial x_3} + \frac{\partial u_3}{\partial x_2} \right)^2 \right) dx = \\
& = \int_{\Omega} \left( \sum_{i=1}^3 \left( \frac{\partial u_i}{\partial x_i} \right)^2 + \frac{\delta}{2} \sum_{i,j=1, i \neq j}^3 \left( \frac{\partial u_i}{\partial x_j} \right)^2 + \delta \sum_{i,j=1, i \neq j}^3 \frac{\partial u_i}{\partial x_j} \frac{\partial u_j}{\partial x_i} \right) dx.
\end{aligned}$$

Используя дважды формулу интегрирования по частям, получим

$$\int_{\Omega} \frac{\partial u_i}{\partial x_j} \frac{\partial u_j}{\partial x_i} dx = \int_{\Omega} \frac{\partial u_i}{\partial x_i} \frac{\partial u_j}{\partial x_j} dx \geq -\frac{1}{2} \int_{\Omega} \left( \left( \frac{\partial u_i}{\partial x_i} \right)^2 + \left( \frac{\partial u_j}{\partial x_j} \right)^2 \right) dx, \quad i \neq j,$$

следовательно,

$$I \geq \min(\delta/2, 1 - \delta) \sum_{i=1}^3 |u_i|_1^2,$$

т. е. неравенство (29) доказано.  $\square$

Если  $\Gamma_u \neq \Gamma$ ,  $\text{mes } \Gamma_u > 0$ , неравенство (29) также справедливо, но доказательство его далеко не элементарно (см., например, [22]).

ЗАМЕЧАНИЕ 5. Смешанная граничная задача линейной теории упругости эквивалентна задаче минимизации

$$F(u) = \min_{v \in V} F(v),$$

где

$$F(v) = \frac{1}{2} \int_{\Omega} \sum_{i,j,k,l=1}^3 a_{ijkl} \varepsilon_{ij}(v) \varepsilon_{kl}(v) dx - \int_{\Omega} f \cdot v dx - \int_{\Gamma_{\sigma}} g \cdot v dx.$$

Функционал  $F$  интерпретируется в механике как потенциальная энергия системы упругое тело — внешние силы.

**4. Двумерные задачи теории упругости.** При некоторых условиях деформацию упругого тела можно характеризовать двумерным вектором смещений  $u = (u_1, u_2)$ , где  $u_1, u_2$  — функции двух декартовых координат,  $u_i(x) = u_i(x_1, x_2)$ ,  $i = 1, 2$ . Уравнения равновесия при этом записываются в виде

$$\sum_{j=1}^2 \frac{\partial \sigma_{ij}}{\partial x_j} + f_i = 0, \quad i, j = 1, 2, \quad x \in \Omega, \quad (30)$$

где

$$\sigma_{ij} = \sigma_{ij}(u) = \sum_{k,l=1}^2 a_{ijkl} \varepsilon_{kl}, \quad (31)$$

$$\varepsilon_{ij} = \varepsilon_{ij}(u) = \frac{1}{2} \left( \frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} \right), \quad i, j = 1, 2, \quad (32)$$

а  $\Omega$  — ограниченная двумерная область с границей  $\Gamma$ . К уравнениям (30)–(32) присоединяют граничные условия, аналогичные (25)–(26). Пусть  $\Gamma = \Gamma_u \cup \Gamma_\sigma$ . Положим

$$u(x) = 0, \quad x \in \Gamma_u, \quad (33)$$

$$\sum_{j=1}^2 \sigma_{ij} \cos(\nu, x_j) = g_i, \quad i = 1, 2, \quad x \in \Gamma_\sigma. \quad (34)$$

Задача (30)–(34) называется смешанной граничной задачей плоской теории упругости.

Исследование обобщенной разрешимости этой задачи при условии, что коэффициенты  $a_{ijkl}$  удовлетворяют условиям симметрии и положительной определенности, проводится точно так же, как и для задачи (22)–(26).

**5. Регулярность решений эллиптических уравнений.** В предыдущих пунктах мы исследовали разрешимость краевых задач для эллиптических уравнений при минимальных требованиях к гладкости исходных данных. Если гладкость исходных данных увеличивается, то и свойства решений краевых задач соответственно улучшаются. В дальнейшем будет использован такой результат.

**Теорема 4.** Пусть  $u$  есть обобщенное решение задачи Дирихле (9)–(10). Предположим, что для целого  $k \geq 0$  и для допустимых  $i, j = 0, 1, \dots, n$  выполнены условия:

$$a_{ij} \in W_\infty^{k+1}(\Omega), \quad a_i \in W_\infty^k(\Omega), \quad f \in W_p^k(\Omega), \quad \Gamma \in W_\infty^{k+2}.$$

Тогда  $u \in W_p^{k+2}(\Omega)$ , причем

$$\|u\|_{k+2,p,\Omega} \leq c \|f\|_{k,p,\Omega},$$

где постоянная  $c$  зависит лишь от коэффициентов уравнения и области  $\Omega$ .

**6. Уравнения четвертого порядка.** Пусть  $\Omega$  — ограниченная двумерная область с границей  $\Gamma$  класса  $C^1$ . Требуется найти

функцию  $u \in C^4(\Omega) \cap C^1(\overline{\Omega})$ , удовлетворяющую уравнению

$$\begin{aligned} \sum_{i,j,k,l=1}^2 \frac{\partial^2}{\partial x_i \partial x_j} \left( a_{ijkl} \frac{\partial^2 u}{\partial x_k \partial x_l} \right) - \sum_{i,j=1}^2 \frac{\partial}{\partial x_i} a_{ij} \frac{\partial u}{\partial x_j} + \\ + \sum_{i=1}^2 a_i \frac{\partial u}{\partial x_i} + a_0 u = 0, \quad x \in \Omega, \end{aligned} \quad (35)$$

и граничным условиям

$$u(x) = 0, \quad \frac{\partial u}{\partial \nu} = 0, \quad x \in \Gamma. \quad (36)$$

Прежде, чем сформулировать понятие обобщенного решения задачи (35), (36), заметим, что, если функция  $u \in C^1(\overline{\Omega})$  и удовлетворяет условиям (36), то

$$\frac{\partial u}{\partial x_1} = \frac{\partial u}{\partial x_2} = 0, \quad x \in \Gamma.$$

Действительно, при выполнении первого условия (36) имеем  $\frac{\partial u}{\partial \tau} = 0$ , где  $\tau$  — единичный вектор касательной к  $\Gamma$ , а

$$\frac{\partial u}{\partial x_i} = \frac{\partial u}{\partial \nu} \cos(\nu, x_i) + \frac{\partial u}{\partial \tau} \cos(\nu, x_i), \quad i = 1, 2. \quad (37)$$

Умножим теперь уравнение (35) на функцию  $\eta \in C^4(\Omega) \cap C^1(\overline{\Omega})$ , удовлетворяющую граничным условиям (36). После интегрирования по частям (дважды — в слагаемых, содержащих четвертые производные, и один раз в слагаемых, содержащих вторые производные) с учетом граничных условий (36) для функции  $\eta$  получим

$$\begin{aligned} \int_{\Omega} \left( \sum_{i,j,k,l=1}^2 a_{ijkl} \frac{\partial^2 u}{\partial x_k \partial x_l} \frac{\partial^2 \eta}{\partial x_i \partial x_j} + \sum_{i,j=1}^2 a_{ij} \frac{\partial u}{\partial x_j} \frac{\partial \eta}{\partial x_i} + \right. \\ \left. + \sum_{i=1}^2 a_i \frac{\partial u}{\partial x_i} \eta + a_0 u \eta \right) dx = \int_{\Omega} f \eta dx. \end{aligned} \quad (38)$$

Теперь естественно дать

**Определение 5.** Функция  $u \in H_0^2(\Omega)$  называется обобщенным решением задачи (35)–(36), если для любой функции  $\eta \in H_0^2(\Omega)$  она удовлетворяет интегральному тождеству (38).

Если предположить выполненными условия:  $f \in L_2(\Omega)$ ,

$$\sum_{i,j,k,l=1}^2 a_{ijkl} \xi_{ij} \xi_{kl} \geq c_0 \sum_{i,j=1}^2 \xi_{ij}^2, \quad c_0 > 0, \quad (39)$$

$$\sum_{i,j=1}^2 a_{ij} \xi_i \xi_j \geq 0, \quad (40)$$

а коэффициенты  $a_i$ ,  $i = 1, 2, \dots, n$ , и  $a_0^-$  считать достаточно малыми, то с использованием результатов об эквивалентных нормировках пространства  $H_0^2(\Omega)$  (см. с. 49) существование и единственность обобщенного решения устанавливается точно так же, как для уравнений второго порядка.

Простейшим примером эллиптического уравнения четвертого порядка является бигармоническое уравнение

$$\Delta^2 u \equiv \frac{\partial^2 u}{\partial x_1^2} + 2 \frac{\partial^2 u}{\partial x_1 \partial x_2} + \frac{\partial^2 u}{\partial x_2^2} = f,$$

описывающее, например, малые прогибы тонкой упругой пластинки, находящейся под действием поперечной силы  $f$ . Для бигармонического уравнения условия (39), (40), очевидно, выполнены.

---

---

ГЛАВА 3  
Пространства конечных элементов

Пусть  $\Omega$  — область в  $R^n$  с липшицевой границей  $\Gamma$ ,  $V = W_p^k(\Omega)$ ,  $k \geq 1$ ,  $p \geq 1$ , — пространство Соболева. Рассмотрим вопрос о построении конечномерных аппроксимаций  $V$  пространствами  $V_h$  — пространствами конечных элементов. Неотрицательная величина  $h \in R$  в обозначении  $V_h$  служит характеристикой точности аппроксимации и является малым параметром, причем  $N_h = \dim V_h \rightarrow \infty$  при  $h \rightarrow 0$ .

Отличительной особенностью пространств конечных элементов является наличие в них естественно конструируемого базиса из функций с малыми носителями<sup>1)</sup>, точнее такого базиса  $\{\varphi_i\}_{i=1}^{N_h}$  в  $V_h$ , что

$$\text{mes}(\text{supp } \varphi_i) = O(h^n), \quad i = 1, 2, \dots, N_h.$$

Построение различных пространств  $V_h$  в методе конечных элементов осуществляется по единой схеме и включает следующие шаги.

1) Выбирается семейство  $\mathcal{T}_h$  замкнутых областей  $e$ , обычно одной и той же формы, называемых конечными элементами (в двумерном случае это треугольники или четырехугольники, возможно с криволинейными сторонами, в трехмерном — тетраэдры, параллелепипеды или призмы). Семейство  $\mathcal{T}_h$  выбирается так, что любые два различных элемента  $e_1, e_2 \in \mathcal{T}_h$  не имеют общих внутренних точек, и область

$$\Omega_h = \bigcup_{e \in \mathcal{T}_h} e$$

аппроксимирует  $\Omega$ . Иногда может выполняться равенство  $\Omega_h = \bar{\Omega}$ .

Независимо от того, выполняется равенство  $\Omega_h = \bar{\Omega}$ , или нет, и независимо от формы элементов  $e$  говорят, что проведена триангуляция  $\mathcal{T}_h$  области  $\Omega$ . Величина  $h$  обычно определяется как максимальный из диаметров элементов, принадлежащих  $\mathcal{T}_h$ . Пример триангуляции области приведен на рисунке 1.

2) С каждым элементом  $e \in \mathcal{T}_h$  связывается тройка  $(e, P_e, \Sigma_e)$ , также называемая конечным элементом. Здесь  $P_e \subset C^k(e)$  — некоторое конечномерное пространство функций, определенных на элементе  $e$ , с

---

<sup>1)</sup>  $\text{supp } \varphi = \{x \in R^n : \varphi(x) \neq 0\}$  — носитель функции  $\varphi$ .

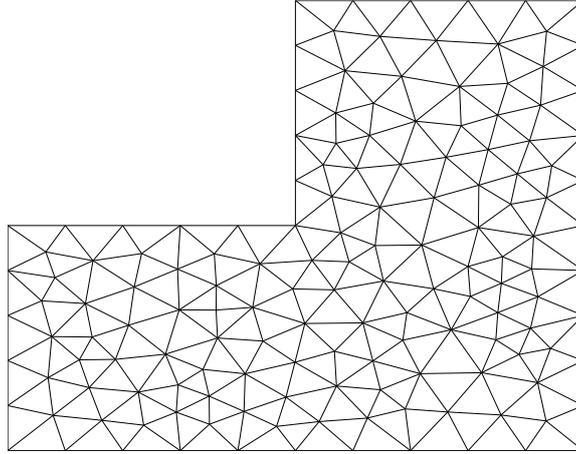


Рис. 1. Пример триангуляции  $L$ -образной области  $\Omega$ ,  $\Omega_h = \Omega$ .

базисом  $\{\varphi_1^e, \dots, \varphi_{m_e}^e\}$ ,  $\Sigma_e = \{\phi_1^e, \dots, \phi_{m_e}^e\}$  — такое множество линейных функционалов над  $P_e$ , что произвольный элемент  $p \in P_e$  может быть однозначно представлен в виде

$$p(x) = \sum_{i=1}^{m_e} \phi_i^e(p) \varphi_i^e(x). \quad (1)$$

Выражение в правой части равенства (1) по форме напоминает интерполяционный полином и часто в конкретных ситуациях таковым и является. Множество  $\mathcal{E}_h = \{(e, P_e, \Sigma_e), e \in \mathcal{T}_h\}$  называют семейством конечных элементов на области  $\Omega$ .

3) Наконец, конструируется конечномерное пространство  $V_h$  функций, определенных на  $\Omega_h$ , таких, что  $V_h \subset C^{k-1}(\overline{\Omega}_h)$ , и сужение на каждый элемент  $e \in \mathcal{T}_h$  любой функции из  $V_h$  принадлежит  $P_e$ .

Здесь важно подчеркнуть, что в силу теоремы 8, с. 44,

$$V_h \subset W_p^k(\Omega_h). \quad (2)$$

Это позволяет трактовать в дальнейшем метод конечных элементов как специальный вариант метода Галеркина.

Важно также отметить, что надлежащий выбор семейства конечных элементов  $\mathcal{E}_h$  не только автоматически обеспечивает нетривиальность пространства  $V_h$ , но позволяет также легко указать в нем базисные функции с малыми носителями. А именно, они конструируются так, что сужение каждой базисной функции  $\varphi_i(x)$  из  $V_h$  на элемент  $e \in \mathcal{T}_h$  либо оказывается тождественно равным нулю, либо совпадает с одной из функций  $\varphi_k^e(x)$ .

В описанной схеме построения пространств конечных элементов основным является понятие конечного элемента  $(e, P_e, \Sigma_e)$ . Можно

сказать, что *конечный элемент порождает пространство конечных элементов* при помощи формальных процедур, единых для всех типов элементов.

Сказанное выше определяет структуру настоящей главы.

Поскольку в основе понятия «конечный элемент» лежит понятие интерполяции, то сначала, в §1, мы кратко рассмотрим вопросы, связанные с интерполяцией в пространстве  $R^n$ .

Затем, в §2, дадим определение базисного конечного элемента, ключевого понятия для всего дальнейшего изложения, и приведем различные примеры базисных элементов. Характерной особенностью базисного элемента является его простая форма и фиксированные размеры.

Важность понятия базисного элемента определяется тем, что:

1) пользуясь им можно достаточно просто строить другие конечные элементы, более сложной конфигурации (так называемые ассоциируемые элементы); эти вопросы рассматриваются в §3–§6;

2) оно удобно при теоретическом исследовании погрешности аппроксимации пространства  $V$  пространствами  $V_h$ ; фактически, исследование погрешности достаточно провести на базисном элементе (см. ниже §10);

3) введение базисного элемента позволяет автоматизировать процедуру построения системы алгебраических уравнений метода конечных элементов.

**ЗАМЕЧАНИЕ 6.** Как уже было сказано, в настоящей главе мы рассматриваем только такие конечномерные подпространства, для которых выполнено условие (2). Методы конечных элементов, основанные на использовании таких подпространств, называются конформными. Применяются также неконформные методы конечных элементов, для которых условие (2) нарушается.

## § 1. Интерполяция в $R^n$

**1. Постановка задачи.** Пусть  $\Delta$  — некоторое множество в пространстве  $R^n$ ,  $P$  — конечномерное пространство вещественных функций переменной  $x \in \Delta$ ,  $\Sigma = \{\phi_i\}_{i=1}^N$  — заданное множество линейных функционалов на пространстве  $P$ .

Рассмотрим задачу интерполяции в пространстве  $R^n$ , т. е. задачу построения функции  $p \in P$ , удовлетворяющей условиям

$$\phi_i(p) = p_i, \quad i = 1, 2, \dots, N, \quad (1)$$

где  $p_1, p_2, \dots, p_N$  — заданные числа.

Важнейшим примером такой задачи является задача полиномиальной лагранжевой интерполяции. Дадим ее постановку. Пусть  $A$

есть некоторое множество мультииндексов, т. е. векторов  $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_n)$ , компоненты которых — целые неотрицательные числа. Множество  $P_A$  полиномов вида

$$P(x) = \sum_{\alpha \in A} c_\alpha x^\alpha, \quad x^\alpha = x_1^{\alpha_1} \dots x_n^{\alpha_n},$$

с действительными коэффициентами  $c_\alpha$  образует линейное пространство. Мономы  $x^\alpha$ ,  $\alpha \in A$ , — базис пространства  $P_A$ ,  $\dim P_A = \text{card } A^1$ . Приведем используемые на протяжении всей книги примеры таких пространств:

1) пространство

$$P_m = P_m(x) = \left\{ p : p = \sum_{|\alpha| \leq m} c_\alpha x^\alpha \right\}, \quad |\alpha| = \alpha_1 + \alpha_2 + \dots + \alpha_n,$$

полиномов степени не выше  $m$  по совокупности переменных  $x_1, \dots, x_n$ ;

2) пространство

$$Q_m = Q_m(x) = \left\{ p : p = \sum_{\alpha_1, \dots, \alpha_n=0}^m c_\alpha x^\alpha \right\}$$

полиномов степени не выше  $m$  по каждой переменной.

Индукцией по размерности пространства  $R^n$  нетрудно показать, что

$$\dim P_m = \frac{(m+n)!}{m!n!}, \quad \dim Q_m = (m+1)^n.$$

Пусть далее  $\omega = \{a_i\}_{i=1}^N$  — заданное множество попарно различных точек из  $\Delta$ . Задача полиномиальной лагранжевой интерполяции заключается в нахождении такого полинома  $p \in P_A$ , который принимает заданные значения  $p_1, p_2, \dots, p_N$  в точках  $a_1, a_2, \dots, a_N$ , т. е. удовлетворяет условиям

$$p(a_i) = p_i, \quad i = 1, \dots, N.$$

Таким образом, задача полиномиальной лагранжевой интерполяции является частным случаем задачи (1) при

$$P = P_A, \quad \phi_i(p) = p(a_i), \quad i = 1, 2, \dots, N.$$

Другим важным примером задачи (1) является задача эрмитовой интерполяции, отличающаяся от лагранжевой интерполяции лишь

<sup>1)</sup>  $\text{card } A$  — мощность множества  $A$ . Если множество  $A$  конечно,  $\text{card } A$  — число элементов  $A$ .

набором функционалов  $\{\phi_i(p)\}$ . Кроме функционалов, соответствующих значениям функции в некоторых точках, в этот набор при эрмитовой интерполяции входят также функционалы, соответствующие значениям производных функции по фиксированным направлениям в определенных точках. Примерами линейных функционалов  $\phi_i$  при эрмитовой интерполяции могут служить функционалы вида

$$\phi_i : p \rightarrow p(a_i), \quad p \rightarrow Dp(a_i)\nu_k, \quad p \rightarrow D^2p(a_i)(\nu_k, \mu_l), \dots,$$

где  $a_i$  — некоторая точка из  $\Delta$ ,  $\nu_k, \mu_l, \dots$  — заданные векторы в  $R^n$ .

**ЗАМЕЧАНИЕ 1.** Примерами применяемых при построении и исследовании методов конечных элементов линейных функционалов, отличных от приведенных выше, являются функционалы

$$\phi_i : p \rightarrow \frac{1}{|\Delta_i|} \int_{\Delta_i} p(x) dx, \quad p \rightarrow \sum_{|\alpha| \leq k} c_\alpha D^\alpha p(a_i),$$

где  $\Delta_i$  — некоторые подмножества множества  $\Delta$ ,  $c_\alpha \in R$ ,  $a_i \in \Delta$ .

**2. Разрешимость задачи интерполяции.** Как известно, в многомерном случае задача полиномиальной интерполяции не всегда имеет решение, даже если число условий интерполяции совпадает с числом неизвестных коэффициентов полинома. В связи с этим возникает общий вопрос о таком выборе множества линейных функционалов  $\Sigma$  для заданного пространства функций  $P$ , чтобы задача интерполяции имела решение и притом единственное.

**Теорема 1.** Пусть  $\dim P = \text{card } \Sigma$ . Тогда следующие утверждения, касающиеся задачи (1), эквивалентны:

- 1) для любых  $p_1, \dots, p_N$  задача имеет по крайней мере одно решение;
- 2) для любых  $p_1, \dots, p_N$  задача имеет единственное решение;
- 3) для  $p_1 = p_2 = \dots = p_N = 0$  задача имеет лишь тривиальное решение.

**ДОКАЗАТЕЛЬСТВО.** Пусть  $\varphi_i$ ,  $i = 1, 2, \dots, N$  — базис в пространстве  $P$ ,  $p = \sum_{j=1}^N c_j \varphi_j$ . Задача (1) эквивалентна системе линейных алгебраических уравнений

$$\sum_{j=1}^N \phi_i(\varphi_j) c_j = p_i, \quad i = 1, \dots, N,$$

для которой утверждения 1)–3) равносильны.  $\square$

**Определение 1.** Будем говорить, что пространство  $P$  и множество  $\Sigma$  унисольвентны, если задача интерполяции (1) имеет единственное решение для любого набора действительных чисел  $\{p_i\}_{i=1}^N$ .

В этом случае иногда также будем говорить, что  $\Sigma$  есть  $P$ -унисольвентное множество, или  $(\Sigma, P)$  — унисольвентная пара.

Примеры.

1) В одномерном случае ( $n = 1$ ) пространство  $P_m$  полиномов степени не выше  $m$  и множество функционалов

$$\phi_i : p \rightarrow p(a_i), \quad i = 0, 1, \dots, m,$$

соответствующих любым  $m + 1$  различным точкам  $a_i$  унисольвентны.

2) При  $n = 1$  пространство полиномов  $P_{2m-1}$  и множество функционалов

$$p(a_i), p'(a_i) \quad i = 0, 1, \dots, m - 1,$$

соответствующих любым  $m$  различным точкам  $a_i$  унисольвентны.

3) При  $n = 2$  пространство полиномов  $P_1 = \{c_0 + c_1x_1 + c_2x_2\}$  первой степени и множество функционалов

$$\phi_i : p \rightarrow p(a_i), \quad a_i \in R^2, \quad i = 0, 1, 2,$$

унисольвентны тогда и только тогда, когда точки  $a_0, a_1, a_2$  различны и не лежат на одной прямой.

4) Пусть  $n = 2, a_0, a_1, a_2, a_3$  — любые точки, лежащие на гиперболе  $x_1x_2 = 1$ . Тогда пространство билинейных функций

$$Q_1 = \{c_0 + c_1x_1 + c_2x_2 + c_{12}x_1x_2\}$$

и множество функционалов

$$\phi_i : p \rightarrow p(a_i), \quad i = 0, 1, 2, 3$$

не унисольвентны.

Следующее утверждение очевидным образом вытекает из теоремы 1.

**Теорема 2.** Для того, чтобы пара  $(\Sigma, P)$  была унисольвентной, необходимо и достаточно выполнения следующих условий:

- 1)  $\dim P = \text{card } \Sigma = N$ ;
- 2) единственная функция из  $P$ , удовлетворяющая условиям

$$\phi_i(p) = 0, \quad i = 1, \dots, N,$$

есть тождественный нуль.

Проверка унисольвентности при помощи этой теоремы, как правило, не является конструктивной. Достаточно общий метод построения унисольвентных пар  $(\Sigma, P)$  основан на использовании базиса Лагранжа.

**Определение 2.** Множество элементов  $\{\varphi_i\}_{i=1}^N \subset P$  называется базисом Лагранжа, соответствующим множеству функционалов  $\Sigma$ , если

$$\phi_j(\varphi_i) = \delta_{ij}, \quad i, j = 1, 2, \dots, N,$$

где  $\delta_{ij}$  — символ Кронекера, т. е.

$$\delta_{ij} = \begin{cases} 0, & i \neq j \\ 1, & i = j. \end{cases}$$

**Теорема 3.** Для того, чтобы пара  $(\Sigma, P)$  была унисольвентной, необходимо и достаточно выполнения следующих условий:

- 1)  $\dim P = \text{card } \Sigma = N$ ;
- 2) существует базис Лагранжа  $\{\varphi_i\}_{i=1}^N$ , соответствующий множеству  $\Sigma$ .

При выполнении условий 1), 2) функция  $p = \sum_{i=1}^N \phi_i(p)\varphi_i$  является решением задачи интерполяции.

**ДОКАЗАТЕЛЬСТВО.** Необходимость. Пусть  $(\Sigma, P)$  — унисольвентная пара. Тогда для любого  $i = 1, 2, \dots, N$  существует единственная функция  $\varphi_i \in P$  такая, что

$$\phi_j(\varphi_i) = \delta_{ij}, \quad j = 1, 2, \dots, N,$$

т. е. существует базис Лагранжа  $\{\varphi_i\}_{i=1}^N$ . Докажем, что  $\dim P = N$ . Для этого достаточно доказать, что множество  $\{\varphi_i\}_{i=1}^N$  образует базис в  $P$ . Прежде всего, нетрудно видеть, что эта система является линейно независимой. Действительно, если

$$\sum_{i=1}^N c_i \varphi_i(x) = 0, \quad x \in \Delta,$$

то применяя функционал  $\phi_j \in \Sigma$  к обеим частям этого равенства, получим  $c_j = 0$  для всех  $j = 1, 2, \dots, N$ . Пусть теперь  $p \in P$  — произвольная функция,  $p_i = \phi_i(p)$ . Так как

$$\phi_j\left(\sum_{i=1}^N p_i \varphi_i\right) = p_j, \quad j = 1, 2, \dots, N,$$

то вследствие унисольвентности  $(\Sigma, P)$  получим, что  $p = \sum_{i=1}^N p_i \varphi_i$ , т. е., действительно, система  $\{\varphi_i\}_{i=1}^N$  образует базис в  $P$ .

Достаточность. Пусть выполнены условия 1), 2). Как и выше, легко проверить, что функции  $\{\varphi_i\}_{i=1}^N$  образуют линейно независимую систему в пространстве  $P$ . Поскольку  $\dim P = N$ , то они образуют базис в  $P$ . Тогда любая функция  $p \in P$  однозначно представляется в виде

$$p(x) = \sum_{i=1}^N p_i \varphi_i(x).$$

Применяя теперь функционалы  $\phi_j \in \Sigma$  к обеим частям этого равенства, получим, что существует единственная функция  $p \in P$  такая, что  $\phi_j(p) = p_j$ ,  $j = 1, 2, \dots, N$ .  $\square$

Какой теоремой, 2, или 3, следует воспользоваться в конкретной ситуации для проверки унисольвентности заданной пары  $\Sigma$  и  $P$ , зависит от того, насколько трудно построить (или описать) базис Лагранжа.

**ЗАМЕЧАНИЕ 2.** С алгебраической точки зрения для однозначной разрешимости задачи интерполяции (1) необходимо и достаточно, чтобы множество  $\Sigma = \{\phi_1, \dots, \phi_N\}$  было базисом в пространстве линейных функционалов над конечномерным пространством  $P$ . Базис Лагранжа  $\varphi_1, \dots, \varphi_N \in P$  есть базис биортогональный  $\Sigma$ .

**3. Оператор интерполяции.** Знание базиса Лагранжа позволяет просто решить задачу интерполяции произвольной достаточно гладкой функции  $u$ .

**Определение 3.** Пусть  $(\Sigma, P)$  — унисольвентная пара,  $W$  — такое линейное пространство функций на  $\Delta$ , что на нем определены функционалы из  $\Sigma$ . Функция  $u_I \in P$  называется  $P$ -интерполянтом функции  $u \in W$ , если  $\phi_i(u_I) = \phi_i(u)$ ,  $i = 1, 2, \dots, N$ . Линейный оператор  $\pi : W \rightarrow P$ , ставящий в соответствие функции ее  $P$ -интерполянт, называется оператором  $P$ -интерполяции (или просто оператором интерполяции).

Очевидно, что имеет место формула

$$\pi u = \sum_{i=1}^N \phi_i(u) \varphi_i(x),$$

и в силу единственности  $P$ -интерполянта  $\pi p = p \quad \forall p \in P$ . Таким образом,  $\pi^2 = \pi$  и, следовательно,  $\pi$  является проектором. Это свойство оператора интерполяции будет неоднократно использоваться в дальнейшем.

## § 2. Конечный элемент

**1. Описание конечного элемента.** Учитывая материал предыдущего параграфа, дадим следующее

**Определение 1.** Назовем  $m$ -мерным конечным элементом в  $R^n$ ,  $m \leq n$ , тройку  $(e, P_e, \Sigma_e)$  такую, что:

- 1)  $e \subset R^n$  — диффеоморфный образ замкнутого многогранника в  $R^m$ ;
- 2)  $P_e \subset C^k(e)$ ,  $k \geq 1$ , — конечномерное пространство вещественных функций на  $e$ ;
- 3)  $\Sigma_e = \{\phi_i^e, i = 1, 2, \dots, m_e\}$  —  $P_e$ -унисольвентное множество линейных функционалов.

Если  $m = n$ , то тройку  $(e, P_e, \Sigma_e)$  будем называть конечным элементом в  $R^n$ .

Под  $k$ -мерной гранью конечного элемента будем понимать образ  $k$ -мерной грани соответствующего многогранника, под вершиной — образ вершины многогранника (0-мерной грани). Линейные функционалы  $\phi_i^e \in \Sigma_e$  называются степенями свободы конечного элемента, а элементы базиса Лагранжа  $\{\varphi_i^e(x), i = 1, 2, \dots, m_e\}$  — базисными функциями конечного элемента. Совокупность степеней свободы конечного элемента удобно представлять в виде вектора-столбца, а совокупность базисных функций — в виде вектора-строки. Введем следующие обозначения:

$$\varphi^e(x) = (\varphi_1^e(x), \dots, \varphi_{m_e}^e(x)), \quad \phi^e(\cdot) = (\phi_1^e(\cdot), \dots, \phi_{m_e}^e(\cdot))^T.$$

Тогда произвольная функция  $p \in P_e$  однозначно представима в виде

$$p(x) = \varphi^e(x)\phi^e(p).$$

**Определение 2.** Назовем  $m$ -мерный конечный элемент  $(e, P_e, \Sigma_e)$  в  $R^n$  лагранжевым, если все его степени свободы имеют вид

$$\phi_i^e : p \rightarrow p(a_i^e), \quad a_i^e \in e.$$

Если степени свободы конечного элемента есть только функционалы вида

$$\phi_i : p \rightarrow p(a_i^e), \quad p \rightarrow Dp(a_i^e)\nu_k^e, \quad p \rightarrow D^2p(a_i^e)(\nu_k^e, \mu_l^e), \dots,$$

где  $\nu_k^e, \mu_l^e, \dots$  — фиксированные векторы в  $R^n$ , т. е. в качестве степеней свободы используются только значения функции и ее производных, то

конечный элемент называется эрмитовым. Точки  $a_i^e$  называют узлами интерполяции, а степень свободы конечного элемента, связанную с узлом интерполяции, — узловым параметром. Множество всех узлов интерполяции элемента  $e$  будем обозначать через  $\omega^e$ .

Далее мы ограничимся изучением только лагранжевых и эрмитовых конечных элементов.

Данные нами определения, фактически, уточняют задачу интерполяции. Главное уточнение связано с геометрической формой конечного элемента, т. е. с видом множества  $e^1$ ). Как следует из определения, одномерный элемент в  $R^1$  является конечным отрезком в  $R^1$ , одномерный элемент  $e$  в  $R^n$  — это отрезок кривой в  $R^n$ . Концевые точки отрезка — вершины элемента. Двумерный элемент в  $R^2$  (элемент в  $R^2$ ) имеет форму многоугольника (с прямолинейными или криволинейными сторонами). Обычно это треугольник или четырехугольник. Двумерный элемент  $e$  в  $R^3$  — часть поверхности многоугольной формы. Трехмерные элементы (элементы в  $R^3$ ) — это многогранники в  $R^3$  (возможно с криволинейными гранями). Простейшие и наиболее часто используемые формы — это тетраэдр, параллелепипед или призма.

**ЗАМЕЧАНИЕ 1.** Выше мы указали наиболее широко распространенные в практических вычислениях формы элементов в  $R^n$ ,  $n = 2, 3$  (см. также начало данной главы). Широкое употребление простейших многогранников связано с тем, что произвольную область в  $R^2$  или  $R^3$  можно представить (или достаточно точно аппроксимировать) объединением, без наложения, таких элементов.

Следующее определение выделяет класс конечных элементов, который служит основой построения конечноэлементных аппроксимаций пространств  $W_p^k$ ,  $k \geq 1$ .

**Определение 3.** Отнесем  $m$ -мерный конечный элемент в  $R^n$  к классу  $C^k$ ,  $k = 0, 1, \dots$ , если  $P_e \subset C^{k+1}(e)$  и, кроме того, значения произвольной функции из  $P_e$  и всех ее частных производных до порядка  $k$  включительно на любой  $(m - 1)$ -мерной грани элемента  $e$  однозначно определяются заданием значений всех узловых параметров во всех узлах, принадлежащих этой грани.

Из этого определения следует, что с каждой  $(m - 1)$ -мерной гранью конечного элемента должно быть связано достаточное количество линейных функционалов из  $\Sigma_e$ . В частности, в каждой вершине  $a_i$  одномерного лагранжева элемента класса  $C^0$  должен быть определен функционал  $\phi : p \rightarrow p(a_i)$ , а для эрмитова элемента клас-

<sup>1)</sup>Множество  $e$  часто само называется конечным элементом. Из контекста всегда понятно, о каком элементе идет речь: о множестве  $e$  или тройке  $(e, P_e, \Sigma_e)$ .

са  $C^k$  должен быть определен по крайней мере  $k+1$  функционал вида  $\phi : p \rightarrow p^{(\alpha)}(a_i)$ ,  $\alpha = 0, 1, \dots, k$ .

Пусть  $(e, P_e, \Sigma_e)$  — произвольный  $m$ -мерный элемент класса  $C^k$  в  $R^n$ , а  $g$  — любая из его  $(m-1)$ -мерных граней. Обозначим через  $\Sigma_e|_g$  множество узловых параметров из  $\Sigma_e$ , связанных с гранью  $g$ , и пусть  $P_e|_g$  есть конечномерное пространство, состоящее из сужений функций из  $P_e$  на грань  $g$ , т. е.  $P_e|_g = \{p|_g : g \rightarrow R, p \in P_e\}$ .

**Лемма 1.** *Для того, чтобы  $m$ -мерный конечный элемент  $(e, P_e, \Sigma_e)$  в  $R^n$  принадлежал к классу  $C^k$ ,  $k \geq 0$ , необходимо и достаточно, чтобы для любой  $(m-1)$ -мерной грани  $g$  элемента  $e$  из условий*

$$\phi_j(p) = 0, \quad \phi_j \in \Sigma_e|_g, \quad p \in P_e, \quad (1)$$

следовало, что

$$D^\alpha p \equiv 0 \quad \text{на } g, \quad |\alpha| \leq k. \quad (2)$$

**ДОКАЗАТЕЛЬСТВО.** На элементе  $e$  функция  $p \in P_e$  имеет вид

$$p(x) = \sum_{j=1}^{m_e} \phi_j(p) \varphi_j(x),$$

т. е.

$$D^\alpha p(x) = \sum_{j=1}^{m_e} \phi_j(p) D^\alpha \varphi_j(x), \quad |\alpha| \leq k,$$

и соответственно

$$D^\alpha p(x)|_g = \sum_{j=1}^{m_e} \phi_j(p) D^\alpha \varphi_j(x)|_g, \quad |\alpha| \leq k. \quad (3)$$

Из определения элемента класса  $C^k$  следует, что в этой сумме должны остаться только слагаемые, соответствующие функционалам из  $\Sigma_e|_g$ . Таким образом, должны быть выполнены равенства

$$D^\alpha p(x)|_g = \sum_{\phi_j \in \Sigma_e|_g} \phi_j(p) D^\alpha \varphi_j(x)|_g, \quad |\alpha| \leq k. \quad (4)$$

Отсюда видно, что из условий (1) следует справедливость (2). Обратное, если из условий (1) следует (2), то из представления (3) вытекает, что имеет место (4), т. е. элемент  $(e, P_e, \Sigma_e)$  принадлежит классу  $C^k$ .  $\square$

**Следствие 1.** *Для того, чтобы  $m$ -мерный конечный элемент в  $R^n$  принадлежал к классу  $C^k$ ,  $k \geq 0$ , необходимо и достаточно,*

чтобы произвольная базисная функция Лагранжа  $\varphi_i(x)$ , соответствующая узловому параметру, определенному в узле  $a_i$ , обращалась в нуль вместе со своими производными до порядка  $k$  включительно вдоль всякой  $(m-1)$ -мерной грани, не содержащей узел  $a_i$ .

**ДОКАЗАТЕЛЬСТВО.** Достаточно заметить, что по определению базисных функций Лагранжа функция  $p = \varphi_i(x)$  удовлетворяет условиям (1), если узел  $a_i$  не принадлежит грани  $g$ .  $\square$

Рассмотрим подробнее лагранжевы элементы класса  $C^0$ . В этом случае

$$\Sigma_e|_g = \{\phi_i : \phi_i(p) = p(a_i), a_i \in \omega^g = \omega^e \cap g\},$$

где  $\omega^e$  — множество узлов интерполяции элемента  $e$ . Поскольку  $p(a_i) = p|_g(a_i)$ , то функционалы из  $\Sigma_e|_g$  можно считать определенными и на элементах из  $P_e|_g$ . Вследствие формулы (4) при  $\alpha=0$  имеем следующее представление произвольного элемента  $p \in P_e|_g$ :

$$p(x) = \sum_{a_i \in \omega^g} p(a_i) \varphi_i|_g(x).$$

Поскольку  $\varphi_i|_g(a_j) = \delta_{ij}$  для любых номеров  $i, j$  таких, что  $a_i, a_j \in \omega^g$ , то функции  $\{\varphi_i|_g(x), a_i \in \omega^g\}$  образуют базис Лагранжа в  $P_e|_g$ , а тройка  $(g, P_e|_g, \Sigma_e|_g)$  образует конечный элемент. Назовем этот элемент сужением лагранжева элемента  $(e, P_e, \Sigma_e)$  на грань  $g$ .

**Теорема 1.** Пусть  $(e, P_e, \Sigma_e)$  —  $m$ -мерный лагранжев конечный элемент класса  $C^0$  в  $R^n$ ,  $n \geq m \geq 2$ . Тогда его сужение  $(g, P_e|_g, \Sigma_e|_g)$  на произвольную  $(m-1)$ -мерную грань  $g$  элемента  $e$  представляет собой  $(m-1)$ -мерный лагранжев элемент класса  $C^0$  в  $R^n$ .

**ДОКАЗАТЕЛЬСТВО.** Пусть  $g$  и  $\tau$  любые две смежные  $(m-1)$ -мерные грани  $e$  с общей  $(m-2)$ -мерной гранью  $\gamma$ . Обозначим через  $\omega^\gamma$  множество (возможно пустое) тех узлов из  $\omega^e$ , которые принадлежат этой грани. Фиксируем произвольно функцию  $p$  из  $P_e$  так, чтобы она была равна нулю на  $\gamma$ . Это можно сделать, например, выбрав равными нулю узловые параметры функции  $p$  на грани  $\tau$ . Пусть  $q = p|_g$  — сужение  $p$  на грань  $g$ . Тогда

$$q = \sum_{a_i \in \omega^g \setminus \omega^\gamma} p(a_i) \varphi_i^g(x),$$

где  $\varphi_i^g = \varphi_i|_g$  — базисные функции элемента  $(g, P_e|_g, \Sigma_e|_g)$ . По построению имеем для любого  $x \in \gamma$ :

$$q|_\gamma(x) = \sum_{a_i \in \omega^g \setminus \omega^\gamma} p(a_i) \varphi_i^g|_\gamma(x) = 0. \quad (5)$$

Поскольку числа  $p(a_i)$  при  $a_i \in \omega^g \setminus \omega^\gamma$  могут быть выбраны произвольно, то из (5) вытекает, что функции  $\varphi_i^g|_\gamma$  тождественно равны нулю на  $\gamma$ , если  $a_i \in \omega^g \setminus \omega^\gamma$ . Пользуясь следствием 1, отсюда заключаем, что элемент  $(g, P_e|_g, \Sigma_e|_g)$  принадлежит классу  $C^0$ .  $\square$

**Следствие 2.** *В каждой вершине лагранжева элемента класса  $C^0$  должен быть определен узловой параметр.*

**ДОКАЗАТЕЛЬСТВО.** Применяя теорему 1 рекурсивно, получим, что каждая одномерная грань лагранжева элемента класса  $C^0$  является также лагранжевым элементом класса  $C^0$ , поэтому в его вершинах должны быть определены функционалы вида  $\phi : p \rightarrow p(a_i)$ .  $\square$

**ЗАМЕЧАНИЕ 2.** Из следствия 2 вытекает ограничение на лагранжевы конечные элементы  $(e, P_e, \Sigma_e)$ , имеющих форму многогранников:  $\dim P_e \geq N$ , где  $N$  — число вершин  $e$ . Аналогичное ограничение можно получить и для эрмитовых элементов. Например, можно показать, что для элемента в  $R^2$  класса  $C^k$  треугольной формы в каждой вершине элемента должны быть определены функционалы  $p \rightarrow D^\alpha p(a_i)$  для всех  $|\alpha| \leq 2k$ . Отметим, что узловые параметры, соответствующие производным, для эрмитовых элементов определяются не однозначно даже при фиксированных узлах интерполяции. В связи с этим далее, в п. 3 §4, мы введем понятие равных эрмитовых элементов.

Для краткости и ясности обозначений условимся в дальнейшем при определении множества  $\Sigma_e$  указывать лишь результат действия функционала на элемент, обозначаемый через  $p$ . Так, например, в случае лагранжева конечного элемента будем писать

$$\Sigma_e = \{p(a_i^e), i = 1, 2, \dots, m_e\}.$$

**2. Базисные конечные элементы.** Опишем сначала геометрические свойства используемых в дальнейшем базисных элементов.

**Определение 4.** Если  $\hat{e} = [0, 1]$ , то конечный элемент  $(\hat{e}, P_{\hat{e}}, \Sigma_{\hat{e}})$  назовем одномерным базисным конечным элементом. Если  $\hat{e}$  — единичный симплекс в  $R^n$ , т. е.

$$\hat{e} = \{\hat{x} \in R^n : \hat{x}_i \geq 0, i = 1, 2, \dots, n, \sum_{i=1}^n \hat{x}_i \leq 1\},$$

или  $\hat{e} = [0, 1]^n$  — единичный  $n$ -мерный куб, то конечный элемент  $(\hat{e}, P_{\hat{e}}, \Sigma_{\hat{e}})$  будем называть соответственно симплицальным (треугольным при  $n = 2$ , тетраэдральным при  $n = 3$ ) или прямоугольным базисным конечным элементом в  $R^n$ .

Базисный конечный элемент<sup>1)</sup> будем обозначать через  $(\hat{e}, \hat{P}, \hat{\Sigma})$ .

<sup>1)</sup>Прилагательное «базисный» означает, что элемент  $e$  имеет фиксированный (раз и навсегда определенный) размер и играет определяющую роль при построении других конечных элементов.

Перейдем к рассмотрению примеров базисных элементов. В описаниях элементов следует обратить внимание на показатель  $m$  во включениях типа  $\hat{P} \supseteq P_m$ . Величина  $m$ , как будет показано в дальнейшем, определяет аппроксимационные свойства конечного элемента.

Для сокращения записей только в этом параграфе будем считать, что базисный конечный элемент задан в координатах  $x$  и обозначать через  $\omega = \{a_i\}$ ,  $\phi_i$ ,  $\varphi_i(x)$ ,  $\pi$  его узлы интерполяции, степени свободы, базисные функции и оператор интерполяции соответственно. Начиная со следующего параграфа, где мы введем в рассмотрение и другие конечные элементы, для указанных выше величин будем использовать обозначения:  $\hat{x}$ ,  $\hat{\omega} = \{\hat{a}_i\}$ ,  $\hat{\phi}_i$ ,  $\hat{\varphi}_i(\hat{x})$ ,  $\hat{\pi}$ . Иными словами, в этом параграфе будем опускать значок «крышка» в обозначениях базисного элемента и относящихся к нему величин.

### 1. Одномерный лагранжев элемент степени $m$ .

Пусть  $e = [0, 1]$  — единичный отрезок на вещественной оси  $x$ ,  $\omega = \{a_i\}_{i=0}^m = \{0 = a_0 < a_1 < \dots < a_m = 1\}$  — множество, состоящее из  $m + 1$  различных точек элемента  $e$ , включающее его граничные точки,  $\Sigma = \{p(a_i)\}_{i=0}^m$ ,  $P = P_m$  — пространство полиномов переменной  $x$  степени не выше  $m$ .

Как уже отмечалось,  $\Sigma$  и  $P$  — унисольвентная пара. Базис Лагранжа в пространстве  $P$  образуют полиномы

$$\varphi_i(x) = \prod_{j=0, j \neq i}^m \frac{x - a_j}{a_i - a_j} = \frac{\omega(x)}{(x - a_i)\omega'(a_i)}, \quad i = 0, 1, \dots, m, \quad (6)$$

где  $\omega(x) = (x - a_0)(x - a_1) \dots (x - a_m)$ .

Итак, тройка  $(e, \Sigma, P)$  образует базисный лагранжев конечный элемент ( $m_e = m + 1$ ).

Поскольку в вершинах  $a_0 = 0$  и  $a_m = 1$  заданы узловые параметры, соответствующие значениям функции в этих вершинах, то мы определили элемент класса  $C^0$ .

Точки  $a_i$ ,  $i = 1, 2, \dots, m - 1$ , выбираются из каких-либо дополнительных соображений. При небольших значениях  $m$  (например,  $m < 10$ ) их, обычно, располагают равномерно:  $a_i = i/m$ .

## 2. Лагранжев прямоугольный элемент степени $m$ .

Пусть  $e = [0, 1] \times [0, 1]$  — единичный квадрат на плоскости  $x_1, x_2$ ,  $\{a_i\}_{i=0}^m$  — множество  $m+1$  различных точек отрезка  $[0, 1]$ , симметрично расположенных относительно середины отрезка<sup>1)</sup>, включающее его граничные точки;  $\omega = \{a_{ij} = (a_i, a_j)\}_{i,j=0}^m$  (см. рис. 2, соответствующий случаю  $m = 4$ ),  $\Sigma = \{p(a_{ij})\}_{i,j=0}^m$ . Множество  $\Sigma$  унисольвентно с пространством  $P = Q_m$  полиномов степени не выше  $m$  по каждой переменной. Действительно,  $\dim Q_m = (m+1)^2$  совпадает с числом функционалов в  $\Sigma$ . Базис Лагранжа, очевидно, образуют полиномы

$$\varphi_{ij}(x) = \varphi_i(x_1)\varphi_j(x_2), \quad i, j = 0, 1, \dots, m,$$

где функции  $\varphi_k$  определяются по формуле (6). В силу теоремы 3, с. 81, тройка  $(e, P, \Sigma)$  образует базисный лагранжев конечный элемент ( $m_e = (m+1)^2$ ). Отметим, что  $P \supset P_m$ .

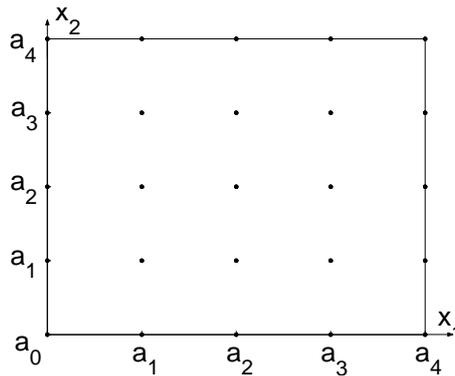


Рис. 2. Лагранжев прямоугольный элемент с 25 степенями свободы,  $P = Q_4$ .

ЗАМЕЧАНИЕ 3. Построенный элемент принадлежит классу  $C^0$ . Чтобы убедиться в этом, достаточно проверить, что базисная функция  $\varphi_{ij}$  тождественно равна нулю на всех сторонах элемента  $e$ , кроме той стороны, которой принадлежит точка  $a_{ij}$ , и воспользоваться следствием 1, с. 85.

ЗАМЕЧАНИЕ 4. Мы использовали нумерацию узлов интерполяции и связанных с ними узловых параметров и базисных функций при помощи двух индексов только для наглядности описания конечного элемента. Введем во многих случаях более удобную нумерацию с помощью одного индекса. Пронумеруем все узлы  $\omega$  произвольным образом от 1 до  $m_e$  и пусть  $a_i$  есть  $i$ -тый узел в этой нумерации. Положим  $\varphi_i(x) = \varphi_{kl}(x)$ , если  $a_i = a_{kl}$ . Тогда каждая функция  $p \in P$  будет иметь представление

$$p(x) = \sum_{i=1}^{m_e} p(a_i) \varphi_i(x) \equiv \varphi^e(x) \phi^e(p).$$

<sup>1)</sup>Такое расположение узлов выбрано из удобства построения в дальнейшем пространства конечных элементов  $V_h$ .

При  $m = 1$  определенный в этом примере конечный элемент будем называть билинейным конечным элементом. Укажем формулы для базисных функций в этом случае, считая, что узлы интерполяции пронумерованы против часовой стрелки, первый узел — начало координат:  $\varphi_1(x) = (1 - x_1)(1 - x_2)$ ,  $\varphi_2(x) = x_1(1 - x_2)$ ,  $\varphi_3(x) = x_1x_2$ ,  $\varphi_4(x) = (1 - x_1)x_2$ .

### 3. Лагранжев треугольный конечный элемент степени $m$ .

Пусть  $e$  — единичный симплекс (треугольник) в  $R^2$ ,  $\{a_i = i/m\}_{i=0}^m$  есть равномерная сетка на отрезке  $[0, 1]$ ,

$$\omega = \{a_{ij} = (a_i, a_j), \quad 0 \leq i + j \leq m\}$$

(см. рис. 3, соответствующий случаю  $m = 6$ ),

$$\Sigma = \{p(a_{ij}), \quad 0 \leq i + j \leq m\}.$$

Докажем, что  $\Sigma$  унисольвентно с пространством  $P = P_m$  полиномов степени не выше  $m$  по совокупности переменных  $x_1, x_2$ . Очевидно, что размерность  $P_m$ , равная  $(m + 1)(m + 2)/2$ , совпадает с числом элементов  $\Sigma$ . Построим базис Лагранжа. Пусть  $a_{kl}$  — некоторая точка  $\omega$ ,  $x_3 = 1 - x_1 - x_2$ . Тогда на прямых

$$\begin{aligned} x_1 &= a_0, & x_1 &= a_1, & \dots, & x_1 &= a_{k-1}, \\ x_2 &= a_0, & x_2 &= a_1, & \dots, & x_2 &= a_{l-1}, \\ x_3 &= a_0, & x_3 &= a_1, & \dots, & x_3 &= a_{m-(k+l)-1}, \end{aligned}$$

лежат все узлы  $\omega$ , кроме  $a_{kl}$  (см. рис. 3, где эти прямые выделены). Отсюда следует, что полином

$$\begin{aligned} p_{kl}(x) &= c(x_1 - a_0) \dots (x_1 - a_{k-1})(x_2 - a_0) \dots (x_2 - a_{l-1}) \dots \\ &\quad \dots (x_3 - a_0) \dots (x_3 - a_{m-(k+l)-1}), \end{aligned}$$

принадлежащий  $P_m$ , обращается в нуль во всех точках  $\omega$ , кроме точки  $a_{kl}$ . Остается нормировать  $p_{kl}(x)$ , чтобы получить базисную функцию, соответствующую точке  $a_{kl}$ . Итак, базис Лагранжа образуют полиномы

$$\varphi_{kl}(x) = \frac{p_{kl}(x)}{p_{kl}(a_{kl})}, \quad 0 \leq k + l \leq m,$$

а тройка  $(e, \Sigma, P)$  — базисный лагранжев конечный элемент класса  $C^0$ . Последнее утверждение обосновывается так же, как и в замечании 3. Замечание 4 также сохраняет силу.

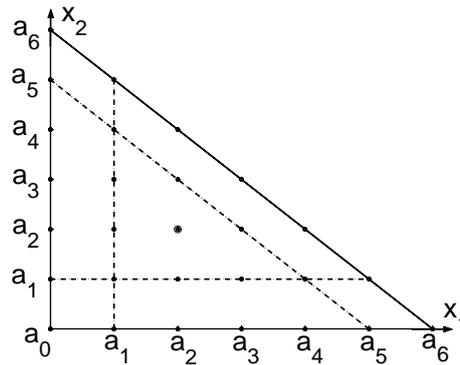


Рис. 3. Лагранжев элемент с 28 степенями свободы,  $P = P_6$ . Выделенные линии содержат все узлы кроме одного.

При  $m = 1$  определенный здесь конечный элемент будем называть линейным конечным элементом. Формулы для его базисных функций, если узлы интерполяции пронумерованы против часовой стрелки, начиная с точки  $x = (0, 0)$ , имеют вид:

$$\varphi_1(x) = 1 - x_1 - x_2, \quad \varphi_2(x) = x_1, \quad \varphi_3(x) = x_2.$$

#### 4. Максимально неполный прямоугольный элемент степени $m$ .

Пусть  $e = [0, 1] \times [0, 1]$ ,  $\{a_i\}_{i=0}^m$ ,  $m \geq 2$  — множество  $m + 1$  различных точек отрезка  $[0, 1]$ , симметрично расположенное относительно середины отрезка и включающее его граничные точки;  $\omega = \{a_{ij} = (a_i, a_j) : a_{ij} \in \partial e\}$  (см. рис. 4, соответствующий случаю  $m = 4$ ),  $\Sigma = \{p(a_{ij}) : a_{ij} \in \omega\}$ . Докажем, что  $\Sigma$  унисольвентно с

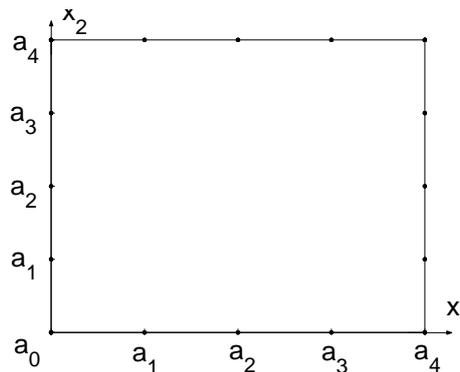


Рис. 4. Максимально неполный лагранжев прямоугольный элемент с 16 степенями свободы,  $P = Q_4^{\min}$ . Все узлы лежат на гранях элемента.

пространством полиномов  $P = Q_m^{\min}$ , которое определим следующим

образом:

$$Q_m^{\min} = \left\{ p : p(x) = \sum_{\alpha_1=0}^1 \sum_{\alpha_2=0}^m c_\alpha x_1^{\alpha_1} x_2^{\alpha_2} + \sum_{\alpha_1=0}^m \sum_{\alpha_2=0}^1 c_\alpha x_1^{\alpha_1} x_2^{\alpha_2}, c_\alpha \in R^1 \right\}.$$

Размерность пространства  $Q_m^{\min}$  равна  $4m$  и совпадает с числом элементов  $\Sigma$ . Построим базис Лагранжа. Введем функции:

$$\begin{aligned} \psi_{00}(x) &= (1 - x_1)(1 - x_2), & \psi_{m0}(x) &= x_1(1 - x_2), \\ \psi_{mm}(x) &= x_1 x_2, & \psi_{0m}(x) &= (1 - x_1)x_2. \end{aligned}$$

Пусть, далее, функции  $\varphi_k$  определяются по формуле (6). Покажем, что функции

$$\begin{aligned} \varphi_{k0}(x) &= \varphi_k(x_1)(1 - x_2), & \varphi_{km}(x) &= \varphi_k(x_1)x_2, & k &= 1, \dots, m-1, \\ \varphi_{0k}(x) &= (1 - x_1)\varphi_k(x_2), & \varphi_{mk}(x) &= x_1\varphi_k(x_2), & k &= 1, \dots, m-1, \\ \varphi_{st}(x) &= \psi_{st}(x) - \sum_{k=1}^{m-1} \psi_{st}(a_{ks})\varphi_{ks}(x) - \sum_{k=1}^{m-1} \psi_{st}(a_{tk})\varphi_{tk}(x), & s, t &= 0, m, \end{aligned}$$

образуют базис Лагранжа в  $Q_m^{\min}$ . Действительно, каждая из функций  $\varphi_{ij}$  принадлежит  $Q_m^{\min}$  и обращается в нуль во всех узлах  $\omega$ , кроме узла  $a_{ij}$ , где она равна единице. Принадлежность построенного элемента классу  $C^0$  обосновывается так же, как и в замечании 3. Замечание 4 сохраняет силу.

**ЗАМЕЧАНИЕ 5.** Использованный способ построения базисных функций легко обобщается и на случай, когда на каждой стороне элемента узлы интерполяции, не совпадающие с угловыми точками  $e$ , выбираются независимо (это также верно и для элементов, рассмотренных в примере 2). Отметим, что элемент, указанный далее в задаче 4, с. 99, находит применение при решении плоских задач теории упругости.

Для построенного элемента  $(e, \Sigma, P)$

$$P = Q_m^{\min} \supset P_m \tag{7}$$

для  $m = 1, 2, 3$ , но при  $m \geq 4$  включение (7) не выполняется. Это хорошо видно из таблицы 1, показывающей степени тех мономов  $x^\alpha = x_1^{\alpha_1} x_2^{\alpha_2}$ , которые входят в определение пространства  $Q_m^{\min}$ ,  $m = 6$ , (они помечены символом  $\oplus$ ) и степени мономов, которые имеются в  $P_5$ , но не входят в  $Q_m^{\min}$  (они помечены символом  $\ominus$ ).

Из таблицы 1 также видно, что в пространстве  $Q_m^{\min}$  не хватает ровно  $\dim P_{m-4} = (m-3)(m-2)/2$  мономов для выполнения условия (7) при  $m \geq 4$ .

Таблица 1

$\alpha_2$																			
↑																			
6	⊕	⊕																	
5	⊕	⊕																	
4	⊕	⊕																	
3	⊕	⊕	⊖																
2	⊕	⊕	⊖	⊖															
1	⊕	⊕	⊕	⊕	⊕	⊕	⊕	⊕	⊕	⊕	⊕	⊕	⊕	⊕	⊕	⊕	⊕	⊕	⊕
0	⊕	⊕	⊕	⊕	⊕	⊕	⊕	⊕	⊕	⊕	⊕	⊕	⊕	⊕	⊕	⊕	⊕	⊕	⊕
	0	1	2	3	4	5	6												→ $\alpha_1$

**5.  $(m, s)$ -неполные прямоугольные элементы.**

Пусть  $m \geq 4$ ,  $(e, \Sigma, P)$  — максимально неполный прямоугольный элемент степени  $m$ , определенный выше,  $\omega$  — множество его узлов интерполяции,  $\{\varphi_i, a_i \in \omega\}$  — базис Лагранжа в  $P$ ,  $s \in [4, m]$  — заданное целое число.

Построим прямоугольный базисный конечный элемент  $(e, P_m^s, \Sigma_m^s)$ , для которого выполнены условия

$$Q_m^{\min} \subset P_m^s, \quad P_s \subset P_m^s \not\subset P_{s+1}. \tag{8}$$

Такой элемент назовем  $(m, s)$ -неполным прямоугольным элементом.

Пусть  $\omega_s = \{a_i \in e\}$  — множество внутренних точек  $e$  таких, что узловые параметры  $\Sigma_s = \{p(a_i), a_i \in \omega_s\}$  образуют  $T_s$ -унисольвентное множество, где  $T_s$  — пространство полиномов, удовлетворяющих условию  $P_{s-4} \subseteq T_s \subseteq Q_{s-4}$ . Примеры 2–4 показывают, что такие множества  $\omega_s$  существуют. Пусть далее  $\{\psi_i, a_i \in \omega_s\}$  — базис Лагранжа в  $T_s$ .

Определим пространство полиномов

$$P_m^s = \left\{ p : p(x) = p_1(x) + x_1 x_2 (1 - x_1)(1 - x_2) p_2(x), \right. \\ \left. p_1(x) \in Q_m^{\min}, p_2(x) \in T_s \right\},$$

а также множества узлов и узловых параметров  $\omega_m^s = \omega \cup \omega_s, \Sigma_m^s = \Sigma \cup \Sigma_s$ . По определению пространства  $P_m^s$  условия (8) выполнены (см. таблицу 1, соответствующую случаю  $m = 6, s = 5, T_s = P_1$ ). По построению также выполнено условие  $\dim P_m^s = \text{card } \Sigma_m^s$ .

Осталось указать базис Лагранжа  $\{\varphi_i^s, a_i \in \omega_m^s\}$  в  $P_m^s$ .

Пусть  $w(x) = x_1 x_2 (1 - x_1)(1 - x_2)$ . Легко проверить, что

$$\varphi_i^s(x) = \begin{cases} \varphi_i(x) - \sum_{a_k \in \omega_s} \frac{w(x)}{w(a_k)} \varphi_i(a_k) \psi_k(x), & \text{если } a_i \in \omega, \\ \frac{w(x)}{w(a_i)} \psi_i(x), & \text{если } a_i \in \omega_s. \end{cases}$$

Конкретный  $(m, s)$ -неполный прямоугольный элемент определяется, таким образом, выбором граничных узлов интерполяции  $\omega$ , пространства полиномов  $T_s$  и соответствующего ему множества внутренних узлов интерполяции  $\omega_s$ . Для выбора множеств  $\omega$  и  $\omega_s$  используются дополнительные соображения, например, можно потребовать некоторой симметрии расположения узлов.

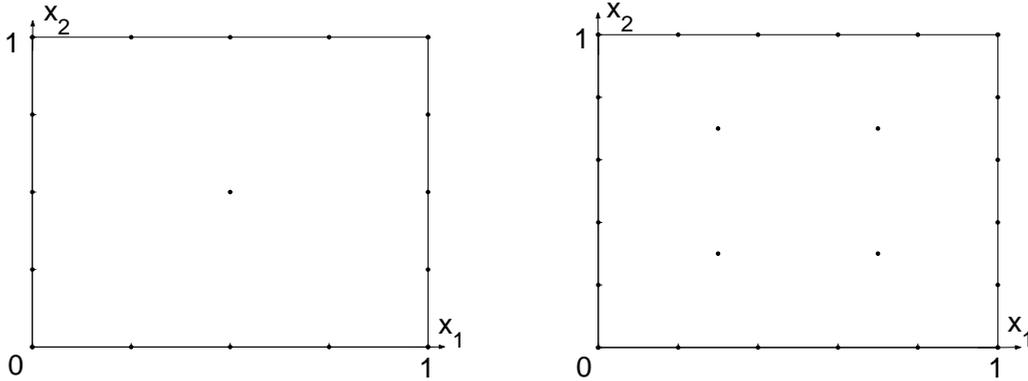


Рис. 5. Примеры неполных прямоугольных элементов типа  $(4, 4)$  и  $(5, 5)$ .

На рис. 5 приведены примеры расположения узлов интерполяции для  $(m, m)$ -неполных элементов при  $m = 4$  и  $m = 5$  соответственно. В последнем случае один внутренний узел можно исключить. Он добавлен из соображений симметрии. Принадлежность определенных здесь элементов классу  $C^0$  обосновывается так же, как и в замечании 3.

### 6. Одномерный эрмитов кубический элемент класса $C^1$ .

Пусть  $e = [0, 1]$ ,  $P = P_3$ ,  $\omega = \{a_0 = 0, a_1 = 1\}$ ,

$$\Sigma = \{\phi_{i0}(p) = p(a_i), \phi_{i1}(p) = Dp(a_i)(a_{i+1} - a_i), i = 0, 1\}, \quad a_2 = a_0.$$

Отметим, что  $\phi_{i1}(p) = (-1)^i p'(a_i)$ .

Тройка  $(e, \Sigma, P)$  образует конечный элемент ( $m_e = 4$ ), его базис Лагранжа составляют полиномы

$$\begin{aligned} \varphi_{00}(x) &= (1-x)^2(1+2x), & \varphi_{10}(x) &= x^2(3-2x), \\ \varphi_{01}(x) &= (1-x)^2x, & \varphi_{11}(x) &= x^2(1-x). \end{aligned}$$

Элемент принадлежит классу  $C^1$ , поскольку в граничных узлах элемента  $a_0 = 0$  и  $a_1 = 1$  заданы узловые параметры, соответствующие значению функции и ее первой производной.

### 7. Эрмитов бикубический элемент класса $C^1$ .

Пусть  $e = [0, 1] \times [0, 1]$ ,  $P = Q_3$ . В множество  $\omega$  включим все вершины  $e$ :  $\omega = \{a_1 = (0, 0), a_2 = (1, 0), a_3 = (1, 1), a_4 = (0, 1)\}$ . С каждым узлом  $a_i \in \omega$  свяжем функционалы

$$\begin{aligned} \phi_{i0}(p) &= p(a_i), \quad \phi_{i3}(p) = D^2 p(a_i)(a_{i-1} - a_i, a_{i+1} - a_i), \\ \phi_{ik}(p) &= \begin{cases} Dp(a_i)(a_{i-1} - a_i), & \text{если } (a_{i-1} - a_i) \parallel Ox_k, \\ Dp(a_i)(a_{i+1} - a_i), & \text{если } (a_{i+1} - a_i) \parallel Ox_k, \end{cases} \end{aligned}$$

где  $k = 1, 2$ ,  $a_0 = a_4$  (все индексы справа от знаков равенства берутся по модулю 4), и положим  $\Sigma = \{\phi_{ik}, i = 1, \dots, 4, k = 0, \dots, 3\}$ . Очевидно,  $\text{card } \Sigma = \dim P = 16$ . Базис Лагранжа строится так же, как и для лагранжева прямоугольного элемента степени  $m$ , в виде произведений базисных функций одномерных элементов. Укажем, например, базисные функции, связанные с узлом  $a_1$ . Базисные функции, связанные с остальными узлами, строятся аналогично. Пусть

$$\begin{aligned} \psi_0(t) &= (1 - t)^2(1 + 2t), \quad \psi_0^1(t) = (1 - t)^2 t, \\ \psi_1(t) &= t^2(3 - 2t), \quad \psi_1^1(t) = t^2(1 - t). \end{aligned}$$

Тогда

$$\begin{aligned} \varphi_{10}(x) &= \psi_0(x_1)\psi_0(x_2), \quad \varphi_{11}(x) = \psi_0^1(x_1)\psi_0(x_2), \\ \varphi_{12}(x) &= \psi_0(x_1)\psi_0^1(x_2), \quad \varphi_{13}(x) = \psi_0^1(x_1)\psi_0^1(x_2). \end{aligned}$$

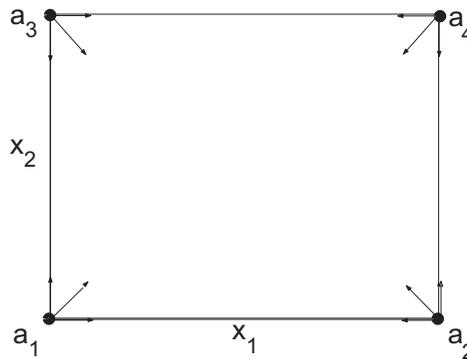


Рис. 6. Эрмитов бикубический элемент,  $P = Q_3$ .

Рисунок 6 иллюстрирует построенный элемент. Здесь черный кружок указывает, что в точке задано значение функции, вертикальная

или горизонтальная стрелка — задана первая производная по направлению, указанному стрелкой, наклонная стрелка — заданы вторые смешанные производные.

Докажем, что этот элемент принадлежит классу  $C^1$ . Заметим, что по определению  $\phi_{i3}(p) = \pm \partial_{12} p(a_i)$ <sup>1)</sup> и с точностью до знака  $\phi_{ik}(p)$  совпадает с  $\partial_k p(a_i)$  при  $k = 1, 2, i = 1, \dots, 4$ . Фиксируем теперь некоторую сторону элемента, например,  $a_1 a_2$ . Нетрудно проверить, что базисные функции  $\varphi_{1k}(x)$ ,  $k = 0, \dots, 3$ , и все их первые производные обращаются в нуль на всех сторонах  $e$ , кроме сторон, содержащих  $a_1$ . Аналогично устанавливается, что базисные функции и их первые производные, соответствующие вершинам  $a_3$  и  $a_4$ , равны нулю на стороне  $a_1 a_2$ . Таким образом, в силу следствия 1, с. 85, элемент принадлежит классу  $C^1$ .

### 8. Эрмитов треугольный кубический элемент класса $C^0$ .

Пусть  $e$  — единичный треугольник в  $R^2$ ,  $P = P_3$ . В множество  $\omega$  включим вершины  $e$  и его центр масс:

$$\omega = \{a_1 = (0, 0), a_2 = (1, 0), a_3 = (0, 1), a_4 = (a_1 + a_2 + a_3)/3\}.$$

Положим

$$\Sigma = \{\phi_{i0}(p) = p(a_i), \phi_{i1}(p) = Dp(a_i)(a_{i-1} - a_i), \\ \phi_{i2}(p) = Dp(a_i)(a_{i+1} - a_i), \quad i = 1, 2, 3, \phi_{40}(p) = p(a_4)\}, \quad a_0 = a_3$$

(индексы справа от знака равенства берутся по модулю 3). Имеем

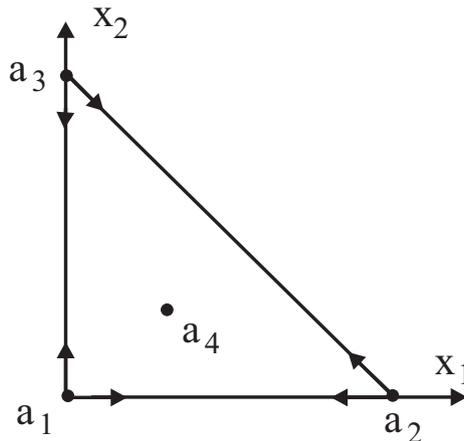


Рис. 7. Эрмитов кубический элемент класса  $C^0$ ,  $P = P_3$ .

<sup>1)</sup>  $\partial_i p = \partial p / \partial x_i$ ,  $\partial_0 p = p$ ,  $\partial_{ij} p = \partial^2 p / \partial x_i \partial x_j$ .

$\dim P = \text{card } \Sigma = 10$ . Для доказательства унисольвентности пары  $(\Sigma, P)$  воспользуемся теоремой 2, с. 80. Пусть все узловые параметры функции  $p \in P$  равны нулю. Покажем, что тогда  $p$  есть тождественный нуль. Для этого обозначим через  $s$  абсциссу стороны  $[a_i, a_{i+1}]$ , и пусть  $q = q(s)$  — след функции  $p$  на этой стороне. Тогда  $q'(a_i) = Dp(a_i)(a_{i+1} - a_i) = 0$ ,  $q'(a_{i+1}) = -Dp(a_{i+1})(a_i - a_{i+1}) = 0$ . Кроме того,  $q(a_i) = q(a_{i+1}) = 0$ . Нетрудно видеть, что  $q$  является полиномом третьей степени от  $s$ , поэтому  $q \equiv 0$ . Иными словами,  $p(x) \equiv 0$  на стороне  $[a_i, a_{i+1}]$ . Это рассуждение верно для любой стороны  $e$ . Таким образом, полином  $p$  обязан иметь вид  $p(x) = cx_1x_2x_3$ , где  $x_3 = 1 - x_1 - x_2$  (см. задачу 1, с. 99). Отсюда, используя условие  $p(a_4) = 0$ , находим, что  $c = 0$ , т. е.  $p(x) \equiv 0$  на  $e$ .

Мы показали, в частности, что если узловые параметры функции из  $P$  на некоторой стороне элемента равны нулю, то функция есть тождественный нуль на этой стороне. Отсюда в силу следствия 1, с. 85, вытекает, что рассматриваемый элемент принадлежит классу  $C^0$ .

Построенный элемент имеет 10 степеней свободы (см. рис. 7). Его базисные функции нетрудно выписать. Например,

$$\begin{aligned} \varphi_{10}(x) &= -2x_3^3 + 3x_3^2 - 7x_1x_2x_3, & \varphi_{11}(x) &= x_1x_3(2x_3 + x_1 - 1), \\ \varphi_{12}(x) &= x_2x_3(2x_3 + x_2 - 1), & \varphi_{40}(x) &= 27x_1x_2x_3, \end{aligned}$$

где  $x_3 = 1 - x_1 - x_2$ , и т. д. Однако в следующем примере выписывать базис в явном виде было бы слишком утомительно.

### 9. Конечный элемент Белла (класса $C^1$ ).

Пусть  $e$  — единичный треугольник в  $R^2$ ,  $\omega$  — множество вершин  $e$ ,

$$\omega = \{a_1 = (0, 0), a_2 = (1, 0), a_3 = (0, 1)\}.$$

Определим множество функционалов

$$\begin{aligned} \Sigma &= \{\phi_{i0}(p) = p(a_i), \phi_{i1}(p) = Dp(a_i)(a_{i-1} - a_i), \\ \phi_{i2}(p) &= Dp(a_i)(a_{i+1} - a_i), \phi_{i,j+2}(p) = D^2p(a_i)(a_{j+1} - a_j)^2, i, j = 1, 2, 3\}, \end{aligned}$$

т. е. в каждой вершине  $e$  заданы значение функции, две первые производные и три вторые производные по соответствующим направлениям (в определении  $\Sigma$  индексы справа от знака равенства берутся по модулю 3,  $a_0 = a_3$ ).

Пространство  $P$  определим следующим образом:

$$P = \{p \in P_5 : \partial_\nu p \in P_3 \text{ на каждой стороне } e\},$$

где  $\nu$  — единичный вектор внешней нормали к  $\partial e$ ,  $\partial_\nu p$  — производная функции  $p$  по направлению  $\nu$ ,  $\partial_\nu p$  рассматривается здесь как функция абсциссы вдоль соответствующей стороны треугольника  $e$ .

Из данного определения сразу следует, что  $P_4 \subset P$ .

Из определения пространства  $P$  вытекает также, что для любого полинома

$$p(x) = \sum_{|\alpha| \leq 5} a_\alpha x_1^{\alpha_1} x_2^{\alpha_2} \in P$$

справедливы равенства

$$a_{1,4} = a_{4,1} = 0, \quad 2(a_{2,3} + a_{3,2}) + 5(a_{0,5} + a_{5,0}) = 0,$$

откуда получаем, что  $\dim P = 18 = \text{card } \Sigma$ .

Для доказательства унисольвентности пары  $(\Sigma, P)$  воспользуемся теоремой 2, с. 80, и будем рассуждать аналогично тому, как это делалось в предыдущем примере. Пусть все узловые параметры функции  $p$  равны нулю. Покажем, что тогда  $p$  есть тождественный нуль. Обозначим через  $s$  абсциссу вдоль стороны  $[a_i, a_{i+1}]$ , и пусть  $q = q(s)$  — след функции  $p$  на этой стороне. Имеем  $q(a_i) = 0$ ,  $q'(a_i) = Dp(a_i)(a_{i+1} - a_i) = 0$ ,  $q''(a_i) = D^2p(a_i)(a_{i+1} - a_i)^2 = 0$ . Аналогично  $q(a_{i+1}) = q'(a_{i+1}) = q''(a_{i+1}) = 0$ . Поскольку  $q$  является полиномом от  $s$  не выше пятой степени, отсюда следует, что  $q \equiv 0$ , т. е.  $p(x) \equiv 0$  на стороне  $[a_i, a_{i+1}]$ . Пусть теперь  $q(s) = \partial_\nu p(s)$  на той же стороне  $[a_i, a_{i+1}]$ . В каждой точке  $a_i$  равны нулю две первых производных функции  $p$  по двум линейно независимым направлениям и три вторых производных функции  $p$  по трем попарно линейно независимым направлениям (см. задачу 2, с. 99). Отсюда следует, что все первые и все вторые производные функции  $p$  в каждой точке  $a_i$  равны нулю и  $q(a_i) = q'(a_i) = q(a_{i+1}) = q'(a_{i+1}) = 0$ . Поскольку  $q \in P_3$ , то  $q \equiv 0$  и  $\partial_\nu p \equiv 0$  на рассматриваемой стороне. Таким образом,  $p = \partial_\nu p = 0$  на любой стороне  $e$ , поэтому  $p(x) = cx_1^2 x_2^2 x_3^2$  (см. задачу 1, с. 99). Так как  $p \in P_5$ , то  $p(x) \equiv 0$  на  $e$ .

Выше мы показали, в частности, что если узловые параметры функции из  $P$  на некоторой стороне элемента равны нулю, то сама функция и ее нормальная производная есть тождественные нули на этой стороне. Отсюда в силу следствия 1, с. 85, вытекает, что рассматриваемый элемент принадлежит классу  $C^1$ .

Иллюстрация элемента  $(e, P, \Sigma)$  дана на рис. 8, где окружность означает, что в точке заданы все вторые производные функции.

ЗАМЕЧАНИЕ 6. Можно дать эквивалентное определение пространства  $P$ :

$$P = \{p \in P_5 : \chi_{ij}(\partial_\nu p) = 0, \quad 1 \leq i < j \leq 3\},$$

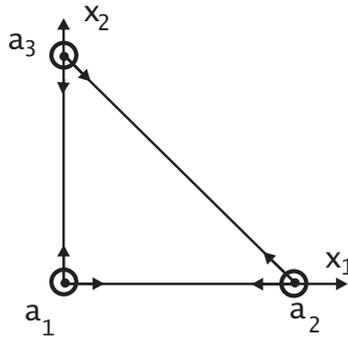


Рис. 8. Конечный элемент Белла класса  $C^1$ .

где  $\chi_{ij}(v) = 4(v(a_i) + v(a_j)) - 8v(a_{ij}) - (Dv(a_j)(a_j - a_i) - Dv(a_i)(a_j - a_i))$ ,  $a_{ij} = (a_i + a_j)/2$  (подчеркнем, что  $\partial_\nu p$  здесь понимается как функция абсциссы вдоль соответствующей стороны треугольника  $e$ ). Действительно, если разложить слагаемые в первой и третьей скобках в определении  $\chi_{ij}(v)$  в ряд Тейлора в окрестности точки  $a_{ij}$ , то  $\chi_{ij}(v) = -1/24d^4v(s)/ds^4$ , где  $s$  — абсцисса вдоль стороны  $[a_i, a_j]$ . Поэтому, если  $\partial_\nu p \in P_3$  на стороне  $[a_i, a_j]$ , то  $\chi_{ij}(\partial_\nu p) = 0$ . Обратно, если  $p \in P_5$ , то  $\partial_\nu p \in P_4$  на стороне  $[a_i, a_j]$  и из условия  $\chi_{ij}(\partial_\nu p) = 0$  получаем, что  $\partial_\nu p \in P_3$  на  $[a_i, a_j]$ .

Отметим, что условие  $\chi_{ij}(\partial_\nu p) = 0$  означает, что нормальная производная функции из  $P$  в центральной точке стороны является линейной комбинацией производных первого и второго порядка в граничных точках этой стороны треугольника.

ЗАМЕЧАНИЕ 7. Во всех рассмотренных примерах эрмитовых элементов обращает на себя внимание способ определения узловых параметров, соответствующих производным. Он удобен в дальнейшем при определении эрмитовых элементов, отличных от базисных. Совершенно ясно, что выбранный способ не является единственным. Важно только то, что производные определенного порядка образуют в некотором смысле полный набор. В связи с этим настоятельно рекомендуем решить задачу 8.

### Задачи.

1. Показать, что если полином  $p = p(x_1, x_2) \in P_m$  обращается в нуль на прямой  $l$ , определяемой уравнением  $x_3 \equiv ax_1 + bx_2 + c = 0$ , то  $p = x_3q$ ,  $q \in P_{m-1}$ . Если дополнительно  $Dp(x)\nu \equiv 0$  на  $l$ , где  $\nu$  — нормаль к  $l$ , то  $p = x_3^2q$ ,  $q \in P_{m-2}$ .

2. Пусть  $\mu, \nu, \tau \in R^2$  — попарно линейно независимы,  $a \in R^2$ . Доказать, что

i) если  $Du(a)\mu = Du(a)\nu = 0$ , то  $Du(a)l = 0 \forall l \in R^2$ ,

ii) если  $D^2u(a)(\mu, \mu) = D^2u(a)(\nu, \nu) = D^2u(a)(\tau, \tau) = 0$ , то  $D^2u(a)(l, n) = 0 \forall l, n \in R^2$ .

3. Показать, что тройка  $e = [0, 1]$ ,  $P = P_2$ ,  $\Sigma = \{p(0), \int_0^1 p dx, p(1)\}$  определяет конечный элемент класса  $C^0$ . Получить явный вид базисных функций.

4. Показать, что тройка  $e = [0, 1] \times [0, 1]$ ,  $P = Q_2^{min}$ ,  $\Sigma = \{p(a), a \in \omega\}$ , где

$$\omega = \{a_1 = (0, 0), a_2 = (1, 0), a_3 = (1, 1), a_4 = (0, 1), a_5 = (0.25, 0), a_6 = (1, 0.5), a_7 = (0.5, 1), a_8 = (0, 0.25)\},$$

определяет конечный элемент класса  $C^0$ . Получить явный вид базисных функций.

5. Показать, что следующая тройка определяет конечный элемент:  $e$  — единичный треугольник в  $R^2$ ,  $P = P_2$ ,  $\Sigma = \{p(a_i), Dp(a_i)(a_{i+1} - a_i), i = 1, 2, 3\}$ , где  $a_1 = (0, 0)$ ,  $a_2 = (1, 0)$ ,  $a_3 = (0, 1)$ . Какого он класса?

6. Показать, что следующая тройка определяет конечный элемент:  $e$  — единичный треугольник в  $R^2$ ,  $P = P_4$ ,  $\Sigma = \{p(a_i), D^2p(a_i)(a_{j+1} - a_j)^2, i, j = 1, 2, 3, p(a_{ij}), 1 \leq i < j \leq 3\}$ , где  $a_1 = (0, 0)$ ,  $a_2 = (1, 0)$ ,  $a_3 = (0, 1)$ ,  $a_{ij} = (a_i + a_j)/2$ . Какого он класса?

7. Пусть  $\Sigma_b$  — множество узловых параметров элемента Белла. Показать, что следующая тройка определяет конечный элемент класса  $C^0$  (но не  $C^1$ ):  $e$  — единичный треугольник в  $R^2$ ,  $P = P_5$ ,

$$\Sigma = \Sigma_b \cup \{Dp(a_{ij})(a_k - a_{ij}), 1 \leq i < j \leq 3, k = 1, 2, 3, k \neq i, k \neq j\},$$

где  $a_1 = (0, 0)$ ,  $a_2 = (1, 0)$ ,  $a_3 = (0, 1)$ ,  $a_{ij} = (a_i + a_j)/2$ .

8. Пусть  $e$  — единичный треугольник в  $R^2$ ,  $P = P_3$ ,

$$\Sigma = \{p(a_i), \partial_j p(a_i), i = 1, 2, 3, j = 1, 2, p(a_4)\},$$

где  $a_1 = (0, 0)$ ,  $a_2 = (1, 0)$ ,  $a_3 = (0, 1)$ ,  $a_4 = (a_1 + a_2 + a_3)/3$ .

а) Показать, что эта тройка определяет конечный элемент.

б) Показать, что в случае бикубического эрмитова элемента можно было бы положить  $\Sigma = \{p(a_i), \partial_j p(a_i), \partial_{12} p(a_i), i = 1, \dots, 4, j = 1, 2\}$ .

с) Как аналогично определить элемент Белла?

### § 3. Ассоциированные конечные элементы

#### 1. Определение и свойства ассоциированных элементов.

Предположим, что задан базисный конечный элемент  $(\hat{e}, \hat{P}, \hat{\Sigma})$  в  $R^n$  и  $\hat{\omega} = \{\hat{a}_i\}$ ,  $\hat{\phi}_i$ ,  $\hat{\varphi}_i(\hat{x})$ ,  $\hat{x} \in \hat{e}$ ,  $\hat{\pi}$  — соответственно узлы интерполяции, узловые параметры, базисные функции и оператор интерполяции.

Достаточно общий способ построения конечных элементов  $(e, P_e, \Sigma_e)$  в  $R^n$  таких, что замкнутое множество  $e$  диффеоморфно  $\hat{e}$ , дает следующая

**Теорема 1.** Пусть задан базисный конечный элемент  $(\hat{e}, \hat{P}, \hat{\Sigma})$  в  $R^n$  и  $x = x_e(\hat{x})$  — достаточно гладкое отображение множества  $\hat{e}$  на заданное замкнутое множество  $e \in R^n$ , имеющее достаточно гладкое обратное  $\hat{x} = x_e^{-1}(x)$ . Положим

$$P_e = \{p : p(x) = \hat{p}(x_e^{-1}(x)), \hat{p} \in \hat{P}\},$$

$$\Sigma_e = \{\phi_i^e : \phi_i^e(p) = \hat{\phi}_i(\hat{p}) \forall p \in P_e, \hat{\phi}_i \in \hat{\Sigma}, \hat{p}(\hat{x}) = p(x_e(\hat{x}))\}. \quad (1)$$

Тогда тройка  $(e, P_e, \Sigma_e)$  образует конечный элемент в  $R^n$  <sup>1)</sup>.

<sup>1)</sup>Отображение  $x = x_e(\hat{x})$  должно быть настолько гладким, чтобы равенства (1), определяющие функционалы  $\Sigma_e$ , были корректны.

**ДОКАЗАТЕЛЬСТВО.** Покажем сначала, что  $\dim P_e = \text{card } \Sigma_e$ . Для этого достаточно убедиться в том, что  $\dim P_e = \dim \hat{P}$ . Последнее вытекает из того, что линейное отображение  $A : \hat{P} \rightarrow P_e$ , порожденное равенством  $p(x) = \hat{p}(x_e^{-1}(x))$ , невырождено, так как для диффеоморфного отображения  $|\det(Dx_e(\hat{x}))| \geq c > 0$ ,  $\hat{x} \in \hat{e}$ , и, следовательно,

$$\|A\hat{p}\|_{L_1(e)} = \int_e |p(x)| dx = \int_{\hat{e}} |\det(Dx_e(\hat{x}))| |\hat{p}| d\hat{x} \geq c \int_{\hat{e}} |\hat{p}| d\hat{x} = c \|\hat{p}\|_{L_1(\hat{e})}$$

для любого  $\hat{p} \in \hat{P}$ . Заметим, наконец, что базис Лагранжа в  $P_e$  образуют функции

$$\varphi_i^e(x) = \hat{\varphi}_i(x_e^{-1}(x)), \quad i = 1, \dots, N_e,$$

т. е. образы базисных функций базисного элемента. В самом деле, поскольку  $\varphi_i^e(x_e(\hat{x})) = \hat{\varphi}_i(\hat{x})$ , то по определению  $\phi_j^e(\varphi_i^e) = \hat{\phi}_j(\hat{\varphi}_i) = \delta_{ij}$ .  $\square$

**Определение 1.** Элемент  $(e, P_e, \Sigma_e)$ , построенный согласно теореме 1, назовем конечным элементом, ассоциированным с базисным элементом  $(\hat{e}, \hat{P}, \hat{\Sigma})$ .

Пусть  $\hat{\omega} = \{\hat{a}_i\}$  — множество узлов интерполяции, по которым определяются узловые параметры из  $\hat{\Sigma}$ . Преобразование  $x = x_e(\hat{x})$  однозначно определяет на элементе  $e$  множество  $\omega^e = \{a_i^e = x_e(\hat{a}_i), \hat{a}_i \in \hat{\omega}\}$ , которое будем называть множеством узлов интерполяции элемента  $e$ . Под вершинами и гранями элемента  $e$  будем понимать образы вершин и граней элемента  $\hat{e}$  при преобразовании  $x = x_e(\hat{x})$ .

**ЗАМЕЧАНИЕ 1.** Взаимно однозначное соответствие между пространствами функций  $P_e$  и  $\hat{P}$ , осуществляемое обратимым преобразованием  $x = x_e(\hat{x})$ , можно рассматривать как преобразование координат. Поэтому для произвольной функции  $v = v(x)$ , заданной на элементе  $e$ , для сокращения записи будем обозначать через  $\hat{v} = \hat{v}(\hat{x})$  ее образ при замене координат  $x = x_e(\hat{x})$ :  $\hat{v}(\hat{x}) = v(x_e(\hat{x})) = v(x)$ .

Из практических соображений ясно, что необходимо стремиться выбирать наиболее простые отображения  $x = x_e(\hat{x})$  из возможных. В этой связи дадим

**Определение 2.** Пусть  $(e, P_e, \Sigma_e)$  — конечный элемент, ассоциированный с базисным  $(\hat{e}, \hat{P}, \hat{\Sigma})$  посредством преобразования  $x = x_e(\hat{x})$ . Он называется:

- 1) аффинно-эквивалентным, если отображение  $x = x_e(\hat{x})$  — аффинное, т. е.  $x_e(\hat{x}) = B_e \hat{x} + b_e$ , где  $B_e$  — невырожденная  $(n \times n)$ -матрица,  $b_e$  — вектор в  $R^n$ ;

- 2) изопараметрическим, если  $x_e(\hat{x}) = (x_1(\hat{x}), \dots, x_n(\hat{x}))^T$ ,  $x_i(\hat{x}) \in \hat{P}$ ,  $i = 1, 2, \dots, n$ , т. е.  $x_e(\hat{x}) \in [\hat{P}]^n$ ;
- 3) криволинейным, если он не является аффинно-эквивалентным или изопараметрическим.

**ЗАМЕЧАНИЕ 2.** Название «аффинно-эквивалентный» мотивируется тем, что если  $(e, P_e, \Sigma_e)$  и  $(\tau, P_\tau, \Sigma_\tau)$  — два конечных элемента, аффинно-эквивалентных базисному  $(\hat{e}, \hat{P}, \hat{\Sigma})$ , то они аффинно-эквивалентны между собой, т. е.  $(\tau, P_\tau, \Sigma_\tau)$  является ассоциированным с элементом  $(e, P_e, \Sigma_e)$  посредством аффинного отображения и наоборот. Приставка «изо» в слове изопараметрический подчеркивает тот факт, что в определении отображения  $x = x_e(\hat{x})$  (а, следовательно, в определении  $P_e$  и  $\Sigma_e$ ) использовано само пространство  $\hat{P}$  базисного элемента  $(\hat{e}, \hat{P}, \hat{\Sigma})$ . Отметим также, что в некоторых случаях изопараметрический конечный элемент может быть одновременно и аффинно-эквивалентным.

**ЗАМЕЧАНИЕ 3.** Если  $\hat{P}$  является пространством полиномов, то пространство  $P_e$  будет полиномиальным только для аффинно-эквивалентных элементов. Отметим также, что с практической точки зрения (для построения системы алгебраических уравнений МКЭ) нам не понадобится явный вид базисных функций в  $P_e$  (так же, как и явный вид отображения  $\hat{x} = x_e^{-1}(x)$ ). Все, что необходимо для реализации МКЭ — это базисные функции базисного элемента, отображение  $x = x_e(\hat{x})$  и матрица Якоби  $Dx_e(\hat{x})$ .

Далее мы рассмотрим конкретные примеры ассоциированных элементов более подробно, при этом изучим элементы класса  $C^0$  всех трех отмеченных типов и ограничимся только аффинно-эквивалентными эрмитовыми элементами класса  $C^1$ .

Предварительно выведем из теоремы 1 три следствия.

**Следствие 1.** *Конечный элемент  $(e, P_e, \Sigma_e)$ , ассоциированный с лагранжевым базисным элементом  $(\hat{e}, \hat{P}, \hat{\Sigma})$  класса  $C^0$ , является также лагранжевым элементом класса  $C^0$  и*

$$\Sigma_e = \{p(a_i^e), a_i^e \in \omega^e = x_e(\hat{\omega})\},$$

где  $\hat{\omega}$  — множество узлов интерполяции базисного элемента.

**ДОКАЗАТЕЛЬСТВО.** Чтобы получить явное выражение для  $\phi_i^e \in \Sigma_e$ , мы должны осуществить преобразование координат  $\hat{x} \rightarrow x$  в выражении для  $\hat{\phi}_i \in \hat{\Sigma}$ . Полученный результат и будет определять  $\phi_i^e$ .

Пусть  $p \in P_e$ ,  $\hat{p} = \hat{p}(\hat{x}) = p(x_e(\hat{x})) \in \hat{P}$  — образ  $p$ . Поскольку  $\hat{\phi}_i(\hat{p}) = \hat{p}(\hat{a}_i)$ , то из (1) имеем

$$\phi_i^e(p) = \hat{p}(\hat{a}_i) = p(x_e(\hat{a}_i)) = p(a_i^e).$$

Пусть, далее,  $e'$  — некоторая грань  $e$ ,  $e' = x_e(\hat{e}')$ . Тогда

$$p(x)|_{e'} = \hat{p}(\hat{x})|_{\hat{e}'} = \sum_{\hat{a}_i \in \hat{e}'} \hat{p}(\hat{a}_i) \hat{\phi}_i(\hat{x}) = \sum_{a_i^e \in e'} p(a_i^e) \phi_i^e(x),$$

откуда вытекает, что  $(e, P_e, \Sigma_e)$  — элемент класса  $C^0$ .  $\square$

**Следствие 2.** Конечный элемент  $(e, P_e, \Sigma_e)$ , аффинно-эквивалентный эрмитову базисному элементу  $(\hat{e}, \hat{P}, \hat{\Sigma})$  класса  $C^1$ , является также эрмитовым элементом класса  $C^1$ . Если  $x = B_e \hat{x} + b_e$  есть соответствующее преобразование координат и

$$\hat{\Sigma} = \{\hat{p}(\hat{a}_i), D\hat{p}(\hat{a}_j)\hat{\nu}_k, D^2\hat{p}(\hat{a}_j)(\hat{\nu}_k, \hat{\mu}_l), \dots\},$$

то

$$\Sigma_e = \{p(a_i^e), Dp(a_j^e)B_e\hat{\nu}_k, D^2p(a_j^e)(B_e\hat{\nu}_k, B_e\hat{\mu}_l), \dots\},$$

где  $a_i^e = B_e\hat{a}_i + b_e$ . В частности, если

$$\hat{\Sigma} = \{\hat{p}(\hat{a}_i), D\hat{p}(\hat{a}_j)(\hat{a}_k - \hat{a}_j), D^2\hat{p}(\hat{a}_j)(\hat{a}_k - \hat{a}_j, \hat{a}_l - \hat{a}_j), \dots\},$$

то способ определения узловых параметров сохраняется<sup>1)</sup>:

$$\Sigma_e = \{p(a_i^e), Dp(a_j^e)(a_k^e - a_j^e), D^2p(a_j^e)(a_k^e - a_j^e, a_l^e - a_j^e), \dots\}. \quad (2)$$

**ДОКАЗАТЕЛЬСТВО.** Пусть  $x_e(\hat{x}) = B_e\hat{x} + b_e$ . Используем правило замены переменных в производных при аффинном преобразовании (см. формулу (3), с. 56). Получим:

$$D\hat{p}(\hat{a}_j)\hat{\nu}_k = Dp(a_j^e)B_e\hat{\nu}_k, \quad D^2\hat{p}(\hat{a}_j)(\hat{\nu}_k, \hat{\mu}_l) = D^2p(a_j^e)(B_e\hat{\nu}_k, B_e\hat{\mu}_l), \dots$$

Поскольку при  $\hat{\nu}_k = \hat{a}_k - \hat{a}_j$  имеем  $B_e\hat{\nu}_k = B_e(\hat{a}_k - \hat{a}_j) = a_k^e - a_j^e$ , то соотношение (2) также выполнено. Проверку того, что  $(e, P_e, \Sigma_e)$  — элемент класса  $C^1$ , оставляем читателю.  $\square$

**Следствие 3.** Пусть  $\hat{\pi}$  — оператор интерполяции базисного элемента,  $\pi_e$  — оператор интерполяции ассоциированного элемента,  $u(x)$  — достаточно гладкая функция, определенная на элементе  $e$ ,  $\hat{u}(\hat{x}) = u(x_e(\hat{x}))$ . Тогда  $(\widehat{\pi_e u}) = \hat{\pi}\hat{u}$ .

**ДОКАЗАТЕЛЬСТВО.** Действительно,

$$(\widehat{\pi_e u}) = \sum_{i=1}^{m_e} \phi_i^e(u) \widehat{\varphi_i^e(x)} = \sum_{i=1}^{m_e} \hat{\phi}_i(\hat{u}) \varphi_i^e(x_e(\hat{x})) = \sum_{i=1}^{m_e} \hat{\phi}_i(\hat{u}) \hat{\varphi}_i(\hat{x}) = \hat{\pi}\hat{u}. \quad \square$$

<sup>1)</sup>Именно по этой причине в рассмотренных выше примерах эрмитовых элементов узловые параметры были выбраны подобным образом. По поводу неоднозначности выбора узловых параметров эрмитовых элементов см. далее определение на с. 111.

## 2. Регулярность семейства ассоциированных элементов.

Пусть  $(\hat{e}, \hat{P}, \hat{\Sigma})$  — базисный конечный элемент в  $R^n$ ,  $h > 0$  — параметр,  $\mathcal{T}_h = \{e\}_h$  — семейство замкнутых областей в  $R^n$  таких, что для любого элемента  $e \in \mathcal{T}_h$ :

- i)  $e$  — диффеоморфный образ  $\hat{e}$ :  $e = x_e(\hat{e})$ ;
- ii)  $\text{diam } e \leq h$ .

С каждой областью  $e$  из  $\mathcal{T}_h$  свяжем конечный элемент  $(e, P_e, \Sigma_e)$ , ассоциированный посредством преобразования  $x = x_e(\hat{x})$  с базисным элементом. Таким образом, получим семейство элементов

$$\mathcal{E}_h = \{(e, P_e, \Sigma_e), e \in \mathcal{T}_h\}.$$

В дальнейшем чаще всего будем рассматривать два случая (см. определение 4, с. 87).

1) Базисный элемент — симплекс в пространстве  $R^n$ . Соответствующее семейство  $\mathcal{T}_h$  будем называть семейством симплицальных элементов.

2) Базисный элемент — единичный  $n$ -мерный куб. Соответствующее семейство  $\mathcal{T}_h$  будем называть семейством прямоугольных элементов.

В зависимости от формы базисного элемента отображение  $x_e : \hat{e} \rightarrow e$  мы подчиним некоторым условиям, что позволит установить оценки (в терминах параметра  $h$ ) для обратного отображения  $x_e^{-1} : e \rightarrow \hat{e}$ . Эти общие результаты будут использованы в дальнейшем при получении оценок погрешности аппроксимации соболевских пространств пространствами конечных элементов.

**Определение 3.** Семейство элементов  $\mathcal{T}_h$  и, соответственно, семейство  $\mathcal{E}_h$  ассоциированных конечных элементов в  $R^n$  назовем  $k$ -регулярным,  $k \geq 1$ , если для любого  $e \in \mathcal{T}_h$  справедлива оценка

$$c_1 h^n \leq \det(Dx_e(\hat{x})) \leq c_2 h^n \quad \forall \hat{x} \in \hat{e}, \quad (3)$$

и, кроме того<sup>1)</sup>,

$$\max_{\hat{x} \in \hat{e}} \|D^l x_e(\hat{x})\| \leq c_3 h^l, \quad l = 1, 2, \dots, k, \quad (4)$$

если  $\mathcal{T}_h$  — семейство симплицальных элементов, и

$$\max_{\hat{x} \in \hat{e}} \|D^l x_e(\hat{x})\| \leq c_3 h, \quad \max_{\hat{x} \in \hat{e}} [D^l x_e(\hat{x})] \leq c_3 h^l, \quad l = 1, 2, \dots, k, \quad (5)$$

<sup>1)</sup> Норма  $\|\cdot\|$  и полунорма  $[\cdot]$  для полилинейных отображений  $[R^n]^m \rightarrow R^n$  введены на с. 30.

если  $\mathcal{T}_h$  — семейство прямоугольных элементов. Здесь  $c_1, c_2, c_3$  — постоянные, не зависящие от  $h$ <sup>1)</sup>. Семейство  $\mathcal{T}_h$  ( $\mathcal{E}_h$ ) назовем регулярным, если оно  $k$ -регулярно для любого  $k \geq 1$ .

**Определение 4.** Семейство прямоугольных ассоциированных конечных элементов  $\mathcal{E}_h$  в  $R^n$  назовем сильно  $k$ -регулярным,  $k \geq 1$ , если для любого  $e \in \mathcal{T}_h$  справедливы оценки (3), (4).

ЗАМЕЧАНИЕ 4. Нетрудно видеть, что условия 1-регулярности симплицальных и прямоугольных элементов совпадают, а из 1-регулярности аффинно-эквивалентного семейства элементов следует его регулярность. Действительно, если  $x_e = B_e \hat{x} + b_e$ , то

$$Dx_e(\hat{x}) = B_e, \quad D^m x_e(\hat{x}) = 0, \quad m \geq 2,$$

а нормы  $[Dx_e(\hat{x})]$  и  $\|Dx_e(\hat{x})\| = \|B_e\|$  — эквивалентны, где  $\|B_e\|$  — спектральная норма матрицы  $B_e$ , т. е. норма, подчиненная евклидовой норме вектора.

ЗАМЕЧАНИЕ 5. По определению условие  $k$ -регулярности семейства симплицальных элементов идентично условию сильной  $k$ -регулярности семейства прямоугольных элементов. Более слабое условие  $k$ -регулярности семейства прямоугольных элементов вводится для того, чтобы обеспечить большее разнообразие допустимых геометрических форм элементов из  $\mathcal{T}_h$ .

Для отображения  $x = x_e(\hat{x}) = (x_1(\hat{x}), \dots, x_n(\hat{x}))^T$  положим

$$|x_e|_{l, \infty, \hat{e}} = \max_{\hat{x} \in \hat{e}} \max_{j=1, \dots, n} \max_{|\alpha|=l} |D^\alpha x_j(\hat{x})|,$$

$$[x_e]_{l, \infty, \hat{e}} = \max_{\hat{x} \in \hat{e}} \max_{i, j=1, \dots, n} \left| \frac{\partial^l x_j(\hat{x})}{\partial \hat{x}_i^l} \right|.$$

Учитывая замечание об эквивалентных нормировках конечномерного пространства  $\mathcal{L}_l(R^n, R^n)$  (см. с. 30), нетрудно убедиться в существовании постоянных  $c_1, c_2$ , зависящих только от  $n$  и  $l$ , таких, что

$$c_1 |x_e|_{l, \infty, \hat{e}} \leq \max_{\hat{x} \in \hat{e}} \|D^l x_e(\hat{x})\| \leq c_2 |x_e|_{l, \infty, \hat{e}},$$

$$c_1 [x_e]_{l, \infty, \hat{e}} \leq \max_{\hat{x} \in \hat{e}} [D^l x_e(\hat{x})] \leq c_2 [x_e]_{l, \infty, \hat{e}}.$$

Отсюда следует, что условия (4), (5) можно эквивалентным образом сформулировать в терминах норм  $|x_e|_{l, \infty, \hat{e}}$  и  $[x_e]_{l, \infty, \hat{e}}$ . Далее мы неоднократно будем это использовать.

**Теорема 2.** Пусть семейство  $\mathcal{E}_h$  ассоциированных конечных элементов в  $R^n$  является  $k$ -регулярным,  $k \geq 1$ . Тогда

$$c_1 h^{-n} \leq \det(Dx_e^{-1}(x)) \leq c_2 h^{-n} \quad \forall x \in e, \quad (6)$$

<sup>1)</sup>Условимся всюду в дальнейшем обозначать через  $c, c_1, c_2, \dots$  постоянные, не зависящие от  $h$ ; один и тот же символ может обозначать разные постоянные.

и, кроме того,

$$|\det(Dx_e)|_{l,\infty,\hat{e}} \leq ch^{n+l}, \quad l = 1, 2, \dots, k, \quad (7)$$

$$\max_{x \in e} \|D^l x_e^{-1}(x)\| \leq ch^{-l}, \quad l = 1, 2, \dots, k, \quad (8)$$

в случае, когда  $\mathcal{T}_h$  — семейство симплицальных элементов, или

$$\max_{x \in e} \|D^l x_e^{-1}(x)\| \leq ch^{-l}, \quad l = 1, 2, \dots, k, \quad (9)$$

если  $\mathcal{T}_h$  — семейство прямоугольных элементов.

ДОКАЗАТЕЛЬСТВО. Оценки (6) непосредственно вытекают из оценок (3), поскольку (см. также лемму 6, с. 57)

$$Dx_e^{-1}(x)Dx_e(\hat{x}) = E,$$

где  $E$  — единичная матрица, и следовательно,

$$\det(Dx_e^{-1}(x)) \det(Dx_e(\hat{x})) = 1.$$

Оценки (7) очевидны при  $n = 1$ . Докажем их при  $n = 2$ , оставляя случай  $n > 2$  в качестве упражнения. Пусть  $x_e(\hat{x}) = (x_1(\hat{x}), x_2(\hat{x}))^T$ . Тогда

$$J_e = \det(Dx_e(\hat{x})) = \frac{\partial x_1(\hat{x})}{\partial \hat{x}_1} \frac{\partial x_2(\hat{x})}{\partial \hat{x}_2} - \frac{\partial x_2(\hat{x})}{\partial \hat{x}_1} \frac{\partial x_1(\hat{x})}{\partial \hat{x}_2}.$$

В силу условия (4) каждое слагаемое в  $J_e$  оценивается сверху величиной  $ch^2$ . Поскольку  $\partial J_e / \partial \hat{x}_\alpha$  есть сумма четырех слагаемых вида

$$\frac{\partial^2 x_i(\hat{x})}{\partial \hat{x}_\alpha \partial \hat{x}_j} \frac{\partial x_k(\hat{x})}{\partial \hat{x}_l},$$

каждое из которых имеет порядок  $h^3$ , то отсюда следует искомая оценка при  $l = 1$ . Повторяя рассуждения, получим (7).

Оценки (8) докажем индукцией по  $l$ . Из равенства  $Dx_e^{-1}(x) = (Dx_e(\hat{x}))^{-1}$  следует, что элементы матрицы  $Dx_e^{-1}(x)$  имеют вид  $\frac{\partial x_\alpha(\hat{x})}{\partial \hat{x}_\beta} / J_e$  и оцениваются сверху величиной  $|x_e|_{1,\infty,\hat{e}} / \min_{\hat{x} \in \hat{e}} J_e \leq ch^{-1}$ .

Отсюда следует оценка (8) при  $l = 1$ . Пусть (8) имеет место при  $l = 1, 2, \dots, s-1$ . Тогда из леммы 6, с. 57, имеем

$$\begin{aligned} \max_{x \in e} \|D^s x_e^{-1}(x)\| &\leq c \sum_{k=1}^{s-1} \max_{\hat{x} \in \hat{e}} \|D^k x_e^{-1}(\hat{x})\| \sum_{i \in I(k,s)} \max_{x \in e} \|Dx_e^{-1}(x)\|^{s-i_1} \times \\ &\quad \times \max_{\hat{x} \in \hat{e}} (\|D^2 x_e(\hat{x})\|^{i_2} \dots \|D^s x_e(\hat{x})\|^{i_s}). \quad (10) \end{aligned}$$

Предположения индукции и условия регулярности позволяют оценить сверху правую часть в (10) величиной

$$c \sum_{k=1}^{s-1} h^{-1} \sum_{i \in I(k,s)} h^{i_1 - s + 2i_2 + \dots + si_s} \leq c h^{-1},$$

что доказывает оценку (8). В случае прямоугольных элементов рассуждения аналогичны. Правая часть в (10) в этом случае оценивается величиной

$$c \sum_{k=1}^{s-1} h^{-k} \sum_{i \in I(k,s)} h^{i_1 - s + i_2 + \dots + i_s} \leq c \sum_{k=1}^{s-1} h^{-k} \sum_{i \in I(k,s)} h^{k-s} \leq c h^{-s},$$

откуда следует (9).  $\square$

УПРАЖНЕНИЕ 3. Доказать (7) при  $n > 2$ .

В следующих трех параграфах мы рассмотрим конкретные примеры ассоциированных элементов и укажем условия их  $k$ -регулярности.

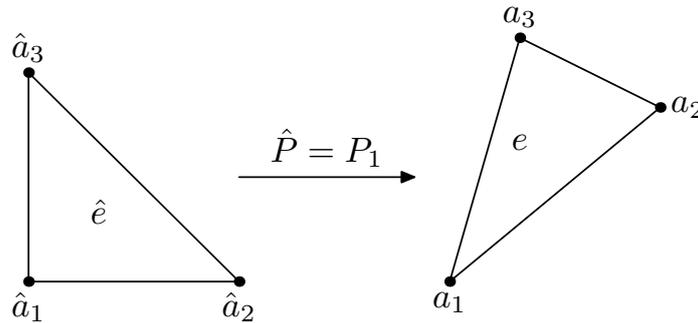


Рис. 9. Аффинно-эквивалентный элемент с 3 степенями свободы,  $P = P_1$ .

### § 4. Аффинно-эквивалентные конечные элементы

**1. Возможные геометрические формы элементов.** Пусть  $(e, P_e, \Sigma_e)$  — произвольный конечный элемент в  $R^n$ , аффинно-эквивалентный базисному элементу  $(\hat{e}, \hat{P}, \hat{\Sigma})$ . Какое разнообразие геометрических форм  $e$  при этом допускается? Очевидно, возможны только элементы следующих типов.

1) В одномерном случае  $\hat{e} = [0, 1]$  и, следовательно,  $e$  может быть только замкнутым интервалом в  $R^1$ . Пусть  $e = [a, b]$ . Отображение  $x = x_e(\hat{x})$  и обратное к нему определяются формулами

$$x = x_e(\hat{x}) = a + h_e \hat{x}, \quad h_e = b - a, \quad \hat{x} = x_e^{-1}(x) = (x - a)/h_e.$$

2) В двумерном случае (см. рис. 9), когда  $\hat{e}$  — единичный треугольник,  $e$  может быть только треугольником в  $R^2$ , так как прямые при аффинном преобразовании переходят в прямые. Пусть  $a_1^e, a_2^e, a_3^e$  — вершины  $e$ , перечисленные против часовой стрелки. Тогда отображение  $x = x_e(\hat{x})$  определяется по правилу

$$x = x_e(\hat{x}) = a_1^e + \hat{x}_1(a_2^e - a_1^e) + \hat{x}_2(a_3^e - a_1^e). \quad (1)$$

Действительно, это отображение является аффинным и вершины  $\hat{e}$  преобразуются в вершины  $e$ , причем  $(0, 0) \rightarrow a_1^e, (1, 0) \rightarrow a_2^e, (0, 1) \rightarrow a_3^e$ , т. е. отображение сохраняет ориентацию базисного элемента. Пусть

$$B_e = \begin{pmatrix} (a_2^e - a_1^e)_1 & (a_3^e - a_1^e)_1 \\ (a_2^e - a_1^e)_2 & (a_3^e - a_1^e)_2 \end{pmatrix}, \quad b_e = \begin{pmatrix} (a_1^e)_1 \\ (a_1^e)_2 \end{pmatrix},$$

тогда  $x_e(\hat{x}) = B_e \hat{x} + b_e$ . Якобиан отображения  $x = x_e(\hat{x})$  равен  $\det(B_e)$ , и, следовательно, равен  $2|e|$  — удвоенной площади элемента  $e$ , поскольку

$$|e| = \int_e dx = \int_{\hat{e}} \det(B_e) d\hat{x} = \det(B_e) \int_{\hat{e}} d\hat{x} = |\hat{e}| \det(B_e). \quad (2)$$

Итак, отображение  $x = x_e(\hat{x})$  полностью определяется координатами вершин элемента  $e$  и является взаимно однозначным. Обратное отображение имеет вид  $\hat{x} = x_e^{-1}(x) = B_e^{-1}(x - b_e)$ .

3) В случае, когда  $\hat{e}$  — единичный квадрат,  $e$  может быть только параллелограммом в  $R^2$ , так как параллельность прямых при невырожденном аффинном отображении сохраняется (см. рис. 10). Если  $a_1^e, a_2^e, a_3^e$  — любые три вершины параллелограмма, перечисленные против часовой стрелки, то отображение  $x = x_e(\hat{x})$  определяется равенством (1). Отметим важный частный случай. Если  $x_e(\hat{x}) = B_e \hat{x} + b_e$ ,

$$B_e = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} h_1^e & 0 \\ 0 & h_2^e \end{pmatrix}, \quad b_e = \begin{pmatrix} a_1^e \\ a_2^e \end{pmatrix},$$

то множество  $e = x_e(\hat{e})$  является прямоугольником в  $R^2$  с длинами сторон, равными  $h_1^e$  и  $h_2^e$ . Он получается следующим образом: сначала  $\hat{e}$  сжимается вдоль осей с коэффициентами сжатия  $h_1^e$  и  $h_2^e$ , затем полученный прямоугольник поворачивается на угол  $\theta$  против часовой стрелки (вершина  $(0, 0)$  остается неподвижной) и переносится так, что точка  $(0, 0)$  переходит в  $b_e$ . В частности, мы получим прямоугольник со сторонами, параллельными осям координат, если  $\theta = 0$ .

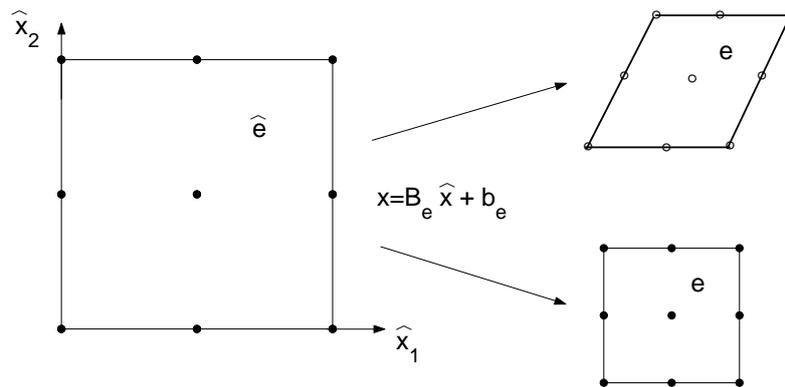


Рис. 10. Возможные формы аффинно-эквивалентных прямоугольных элементов.

Итак, мы выяснили, какие геометрические формы могут иметь аффинно-эквивалентные конечные элементы в  $R^n$  при  $n = 1, 2$  и в каждом случае определили вид преобразования  $x = x_e(\hat{x})$ . Выбирая теперь элемент  $e$  конкретной формы (из указанных выше), мы по нему однозначно определяем отображение  $x = x_e(\hat{x})$  и конечный элемент  $(e, P_e, \Sigma_e)$  так, как это описано в предыдущем параграфе. А именно, определяем множества  $\omega^e$ ,  $P_e$ ,  $\Sigma_e$ , базисные функции Лагранжа  $\varphi_i^e$  и оператор интерполяции  $\pi_e$ .

**2. Регулярность семейства аффинно-эквивалентных элементов.** Пусть  $(\hat{e}, \hat{P}, \hat{\Sigma})$  — базисный элемент в  $R^n$  и  $\mathcal{E}_h = \{(e, P_e, \Sigma_e), e \in \mathcal{T}_h\}$  — семейство аффинно-эквивалентных ему элементов максимального диаметра  $h$ . Каждый элемент  $e$  из  $\mathcal{T}_h$  является образом  $\hat{e}$  при аффинном преобразовании координат  $x = x_e(\hat{x}) = B_e \hat{x} + b_e$ .

Отметим, что если аффинно-эквивалентный конечный элемент — треугольник и  $\hat{P} = \hat{P}_s$  — множество всех полиномов степени не выше  $s$  по совокупности переменных  $\hat{x}$ , то  $P_e = P_s$  — множество всех полиномов степени не выше  $s$  по совокупности переменных  $x$ . Если конечный элемент — параллелограмм и  $\hat{P} \supset \hat{P}_s$ , то  $P_e \supset P_s$ , при этом максимальная степень мономов в  $P_e$ , вообще говоря, будет выше, чем в  $\hat{P}$ .

Сформулируем условия на геометрические характеристики элементов семейства  $\mathcal{T}_h$ , обеспечивающие его регулярность.

**Лемма 1.** Пусть  $\mathcal{T}_h$  — определенное выше семейство аффинно-эквивалентных элементов,  $\rho_e$  — максимальный радиус шара, содержащегося в элементе  $e$ . Пусть, далее, существует постоянная  $c$ , не зависящая от  $h$ , такая, что  $\rho_e \geq ch$  для любого элемента  $e \in \mathcal{T}_h$ . Тогда семейство  $\mathcal{T}_h$  регулярно.

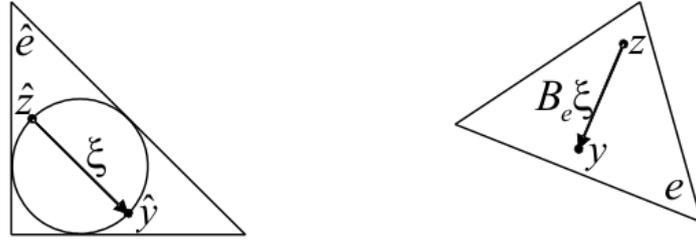


Рис. 11. К доказательству леммы 1.

ДОКАЗАТЕЛЬСТВО. Убедимся предварительно в справедливости оценок <sup>1)</sup>

$$\|B_e\| \leq h_e/\hat{\rho}, \quad \|B_e^{-1}\| \leq \hat{h}/\rho_e, \quad (3)$$

где  $h_e$ ,  $\hat{h}$  — диаметры  $e$  и  $\hat{e}$  соответственно, а  $\hat{\rho}$  — максимальный диаметр шара, содержащегося в  $\hat{e}$ . Отметим, что  $\hat{h}$ ,  $\hat{\rho}$  — постоянные, зависящие только от  $n$ . Ясно, что

$$\|B_e\| = \max_{\xi \neq 0} \frac{|B_e \xi|}{|\xi|} = \max \left\{ \frac{|B_e \xi|}{|\xi|}; \xi \in R^n, |\xi| = \hat{\rho} \right\}.$$

Каждый вектор  $\xi \in R^n$  длины  $\hat{\rho}$  может быть представлен в виде  $\xi = \hat{y} - \hat{z}$ ,  $\hat{y}, \hat{z} \in \hat{e}$ . Поэтому  $B_e \xi = (B_e \hat{y} + b_e) - (B_e \hat{z} + b_e) = y - z$ , где  $y, z \in e$ . Это означает, что  $|B_e \xi| \leq h_e$ , откуда следует первое неравенство (3). Вторая оценка доказывается аналогично.

Поскольку  $h_e \leq h$ ,  $\rho_e \geq ch$ , то из (3) вытекают, оценки:

$$\|B_e\| \leq c_1 h, \quad \|B_e^{-1}\| \leq c_2 h^{-1}, \quad (4)$$

где  $c_1 = 1/\hat{\rho}$ ,  $c_2 = 1/c$ .

Далее, согласно равенству (2), с. 108, имеем  $\det(B_e) = |e|/|\hat{e}|$ . Учитывая, что элемент  $e$  содержит шар радиуса  $\rho_e \geq ch$  и может быть вписан в  $n$ -мерный куб со стороной  $h$ , получим

$$c_3 h^n \leq \det(B_e) \leq c_4 h^n \quad (5)$$

с постоянными, не зависящими от  $h$ . Из оценок (4), (5) и замечания 4, с. 105, следует утверждение леммы.  $\square$

Рассмотрим отдельно случай  $n = 2$ , когда  $\mathcal{T}_h$  — семейство треугольников. Обозначим через  $h_i^e$  и  $\theta_i^e$ ,  $i = 1, 2, 3$ , — стороны и углы треугольника  $e$ , занумерованные против часовой стрелки, причем углы  $\theta_1^e$  и  $\theta_2^e$  будем считать прилегающими к стороне  $h_1^e$ .

<sup>1)</sup>Здесь и далее под нормой матрицы понимаем норму, согласованную с евклидовой нормой векторов.

**Следствие 1.** Если для любого элемента  $e \in \mathcal{T}_h$

$$ch \leq h_i^e \leq h, \quad \theta_i^e \geq \theta_0, \quad i = 1, 2, 3, \quad (6)$$

где постоянные  $c$  и  $\theta_0$  не зависят от  $h$ , то семейство  $\mathcal{T}_h$  регулярно.

**ДОКАЗАТЕЛЬСТВО.** Достаточно проверить, что из (6) следует, что  $\rho_e \geq ch$ . Для каждого  $i = 1, 2, 3$  имеют место равенства

$$\rho_e(h_1^e + h_2^e + h_3^e) = 2|e| = h_i^e h_{i+1}^e \sin \theta_{i+1}^e, \quad (7)$$

где индекс, превышающий 3, берется по модулю 3. Если условия (6) выполнены, то из (7) вытекает, что  $\rho_e \geq c_1 h$  с постоянной  $c_1 = (c^2/3) \sin \theta_0$ .  $\square$

**ЗАМЕЧАНИЕ 1.** Из формулы 1, очевидно, вытекает, что если семейство элементов  $\mathcal{T}_h$  регулярно, то  $\theta_i^e \leq \pi - \theta_0$ ,  $i = 1, 2, 3$ , для любого  $e \in \mathcal{T}_h$ .

**3. Равные эрмитовы элементы.** Как отмечалось ранее, в случае эрмитовых элементов имеется некоторый произвол в определении узловых параметров базисного элемента, соответствующих производным (см., например, задачу 8, с. 100). Далее при определении пространств конечных элементов, аппроксимирующих пространства Соболева, нам будут необходимы и другие, в некотором смысле эквивалентные, наборы узловых параметров ассоциированных элементов.

В этой связи дадим

**Определение 1.** Два эрмитовых конечных элемента  $(e, P_e, \Sigma_e)$  и  $(\tau, P_\tau, \Sigma_\tau)$  равны, если  $e = \tau$ ,  $P_e = P_\tau$ ,  $\pi_e = \pi_\tau$ .

Для каждого эрмитова элемента можно указать множество равных ему элементов. Приведем пример равных элементов, из которого будет ясно, как можно вообще определить эрмитов элемент, равный имеющемуся. Советуем читателю внимательно проанализировать этот пример, поскольку использованный в нем прием замены узловых параметров является характерным и необходимым при использовании эрмитовых элементов.

**Элемент, равный ассоциированному элементу Белла.**

Пусть элемент  $(e, P_e, \Sigma_e)$  является аффинно-эквивалентным базисному элементу Белла (см. пример на с. 97). Согласно следствию 2, с. 103, множество его узловых параметров имеет вид

$$\begin{aligned} \Sigma_e = \{ \phi_{i0}^e(p) = p(a_i^e), \quad \phi_{i1}^e(p) = Dp(a_i^e)(a_{i-1}^e - a_i^e), \\ \phi_{i2}^e(p) = Dp(a_i^e)(a_{i+1}^e - a_i^e), \quad \phi_{i,j+2}^e(p) = D^2p(a_i^e)(a_{j+1}^e - a_j^e)^2, \\ i, j = 1, 2, 3 \}, \end{aligned}$$

где, как и ранее, индексы справа от знака равенства берутся по модулю 3,  $a_i^e$  — координата  $i$ -той вершины треугольника  $e$ .

Сквозную нумерацию узловых параметров и базисных функций введем как на базисном элементе. Обозначим через  $\phi_i^e(\cdot) = (\phi_{i0}^e(\cdot), \dots, \phi_{i5}^e(\cdot))$  вектор узловых параметров, соответствующих вершине  $a_i^e$ , и пусть  $\phi^e(\cdot) = (\phi_1^e(\cdot), \phi_2^e(\cdot), \phi_3^e(\cdot))^T$  — вектор узловых параметров элемента  $e$ . Аналогично пронумеруем базисные функции, полагая

$$\varphi_i^e(x) = (\varphi_{i0}^e(x), \dots, \varphi_{i5}^e(x)), \quad \varphi^e(x) = (\varphi_1^e(x), \dots, \varphi_3^e(x)).$$

Оператор интерполяции определяется при этом равенством  $\pi_e u = \varphi^e(x)\phi^e(u)$ . Зададим новое множество узловых параметров

$$\begin{aligned} \Sigma_e^g &= \{\phi_{ik}^g(p), \quad i = 1, 2, 3, \quad k = 0, \dots, 5\} = \\ &= \{p(a_i^e), \partial_1 p(a_i^e), \partial_2 p(a_i^e), \partial_{11} p(a_i^e), \partial_{12} p(a_i^e), \partial_{22} p(a_i^e), \quad i = 1, 2, 3\}, \end{aligned}$$

и образуем конечный элемент  $(e, P_e, \Sigma_e^g)$ . Покажем, что он равен исходному. Для этого аналогично предыдущему введем вектор новых узловых параметров:

$$\begin{aligned} \phi^{eg}(p) &= (\phi_1^g(p), \phi_2^g(p), \phi_3^g(p))^T, \\ \phi_i^g(p) &= (p(a_i^e), \partial_1 p(a_i^e), \partial_2 p(a_i^e), \partial_{11} p(a_i^e), \partial_{12} p(a_i^e), \partial_{22} p(a_i^e)). \end{aligned}$$

Векторы  $\phi^e(p)$  и  $\phi^{eg}(p)$  связаны линейным преобразованием, вид которого мы установим, используя равенства:

$$\begin{aligned} Dp(a_i^e)(a_k^e - a_i^e) &= \sum_{\alpha=1}^2 (a_k^e - a_i^e)_\alpha \partial_\alpha p(a_i^e), \\ D^2 p(a_i^e)(a_{j+1}^e - a_j^e)^2 &= \sum_{\alpha, \beta=1}^2 (a_{j+1}^e - a_j^e)_\alpha (a_{j+1}^e - a_j^e)_\beta \partial_{\alpha\beta} p(a_i^e), \end{aligned}$$

где  $(a)_k$  —  $k$ -тая компонента вектора  $a$ . Введем матрицы

$$\begin{aligned} D_{1i}^e &= \begin{pmatrix} (a_{i-1}^e - a_i^e)_1 & (a_{i-1}^e - a_i^e)_2 \\ (a_{i+1}^e - a_i^e)_1 & (a_{i+1}^e - a_i^e)_2 \end{pmatrix}, \\ D_{2i}^e &= \begin{pmatrix} (a_2^e - a_1^e)_1^2 & 2(a_2^e - a_1^e)_1(a_2^e - a_1^e)_2 & (a_2^e - a_1^e)_2^2 \\ (a_3^e - a_2^e)_1^2 & 2(a_3^e - a_2^e)_1(a_3^e - a_2^e)_2 & (a_3^e - a_2^e)_2^2 \\ (a_1^e - a_3^e)_1^2 & 2(a_1^e - a_3^e)_1(a_1^e - a_3^e)_2 & (a_1^e - a_3^e)_2^2 \end{pmatrix}, \end{aligned}$$

и образуем блочно-диагональные матрицы

$$D_i^e = \text{diag}(1, D_{1i}^e, D_{2i}^e), \quad D^e = \text{diag}(D_1^e, D_2^e, D_3^e).$$

Нетрудно убедиться, что  $\phi^e(p) = D^e \phi^{eg}(p)$ . Следовательно,

$$\pi_e u = \varphi^e(x) \phi^e(u) = \varphi^e(x) D^e \phi^{eg}(u) = \varphi^{eg} \phi^{eg}(u), \quad \varphi^{eg} = \varphi^e D^e. \quad (8)$$

Отсюда вытекает, что компоненты  $\varphi_i^{eg}(x)$  вектор-функции  $\varphi^{eg}(x)$  образуют базис Лагранжа элемента  $(e, P_e, \Sigma_e^g)$ . Пусть  $\pi_e^g$  соответствующий ему оператор интерполирования. Тогда из (8) следует искомое равенство

$$\pi_e u = \varphi^e(x) \phi^e(u) = \varphi^{eg}(x) \phi_i^{eg}(u) = \pi_e^g u.$$

Отличительной особенностью нового набора узловых параметров, который назовем глобальным, является то, что он образован с использованием производных только по фиксированным, не зависящим от элемента, направлениям. Как мы увидим далее, введение множества  $\Sigma_e^g$  и соответствующего ему базиса Лагранжа необходимо, как при определении пространства эрмитовых конечных элементов, так и для программной реализации метода конечных элементов.

#### Задачи.

1. Пусть элемент  $(e, P_e, \Sigma_e)$  аффинно-эквивалентен базисному эрмитову треугольному кубическому элементу класса  $C^0$  (см. пример на с. 96). По аналогии с элементом Белла показать, что ему равен элемент  $(e, P_e, \Sigma_e^g)$  с множеством узловых параметров

$$\Sigma_e^g = \{p(a_i^e), \partial_1 p(a_i^e), \partial_2 p(a_i^e), p(a_4^e), \quad i = 1, 2, 3\},$$

где  $a_4^e = (a_1^e + a_2^e + a_3^e)/3$ . Проверить, что функции  $\{\varphi_i^{eg}(x)\}_{i=1}^{10}$ , определяемые равенствами

$$\varphi^{eg}(x) = \varphi^e(x) D^e, \quad D^e = \text{diag}(D_1^e, D_2^e, D_3^e, 1), \quad D_i^e = \text{diag}(1, D_{1i}^e), \quad i = 1, 2, 3,$$

образуют базис Лагранжа этого элемента.

2. Пусть элемент  $(e, P_e, \Sigma_e)$  является аффинно-эквивалентным базисному эрмитову бикубическому элементу класса  $C^1$  (см. пример на с. 95). По аналогии с элементом Белла показать, что ему равен элемент  $(e, P_e, \Sigma_e^g)$  с множеством узловых параметров

$$\Sigma_e^g = \{p(a_i^e), \partial_1 p(a_i^e), \partial_2 p(a_i^e), \partial_{12} p(a_i^e), \quad i = 1, \dots, 4\}.$$

Указать способ вычисления базиса Лагранжа  $\{\varphi_i^{eg}(x)\}_{i=1}^{16}$  этого элемента.

## § 5. Криволинейные конечные элементы

Криволинейные элементы естественным образом возникают при построении точной триангуляции области с кусочно гладкой криволинейной границей. На практике обычно используются элементы с прямолинейными сторонами внутри области, тогда как приграничные элементы имеют одну криволинейную сторону, совпадающую с участком границы рассматриваемой области. Рассмотрим один из

способов определения двумерных криволинейных лагранжевых треугольных элементов  $(e, P_e, \Sigma_e)$ , ассоциированных с базисным элементом  $(\hat{e}, \hat{P}, \hat{\Sigma})$  в  $R^2$ . Будем считать, что две стороны криволинейного треугольника  $e$  с вершинами  $a_1, a_2, a_3$  — отрезки прямых, а третья — дуга кривой  $\Gamma$ . Треугольник с вершинами  $a_1, a_2, a_3$  (с прямолинейными сторонами) обозначим через  $\tilde{e}$ . Будем предполагать, что элементы настолько малы, что криволинейная сторона  $\widehat{a_2 a_3}$  не имеет общих точек с прямолинейными сторонами этого треугольника за исключением точек  $a_2, a_3$  (см. рис. 12).

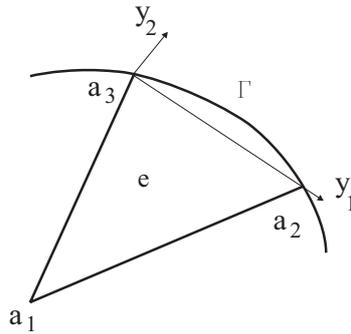


Рис. 12. Криволинейный конечный элемент.

Совокупность всех таким образом определенных криволинейных треугольных элементов обозначим через  $\mathcal{T}_h$ . Здесь  $h = \max_{e \in \mathcal{T}_h} \text{diam}(e)$ .

Через  $\tilde{\mathcal{T}}_h$  обозначим соответствующее множество треугольников  $\tilde{e}$ .

Для определения ассоциированного элемента мы должны задать отображение базисного треугольника  $\hat{e}$  на  $e \in \mathcal{T}_h$ . Если  $e$  — прямолинейный треугольник, то это — аффинное отображение  $x = B_e \hat{x} + b_e$ . При построении и исследовании отображения на криволинейный элемент оказывается полезной следующая теорема о свойствах непрерывных взаимно однозначных отображений в конечномерных пространствах.

**Теорема 1.** Пусть  $\Omega$  — ограниченное открытое множество в  $R^n$  такое, что  $\text{int } \bar{\Omega} = \Omega$ ,  $f$  — непрерывное, взаимно однозначное отображение  $\bar{\Omega}$  в  $R^n$ . Тогда

$$f(\bar{\Omega}) = \overline{f(\Omega)}, \quad f(\Omega) = \text{int } f(\bar{\Omega}), \quad f(\partial\Omega) = \partial f(\Omega) = \partial f(\bar{\Omega}).$$

Будем считать, что кривая  $\Gamma$  задана в параметрической форме

$$x = \chi(s), \quad 0 \leq s \leq L,$$

причем

$$\chi \in C^1[0, L], \quad |\chi'(s)| \geq c_0 > 0, \quad s \in [0, L]. \quad (1)$$

Последнее неравенство означает, что у  $\Gamma$  нет особых точек. Для определенности предположим, что при возрастании параметра  $s$  совершается обход кривой против часовой стрелки.

Пусть  $a_2 = \chi(s_2)$ ,  $a_3 = \chi(s_3)$ ,  $l_e = s_3 - s_2$ . Положим

$$\varphi(t) = \chi(s_2 + tl_e), \quad L_1(t) = a_2 + t(a_3 - a_2), \quad 0 \leq t \leq 1.$$

Ясно, что  $L_1$  — линейный интерполянт функции  $\varphi$  — дает параметрическое представление хорды  $a_2a_3$ . Определим отображение базисного элемента:

$$x = x_e(\hat{x}) = a_1 + B_e \hat{x} + \tilde{x}_e(\hat{x}), \quad \tilde{x}_e(\hat{x}) = \hat{x}_1 \Phi(\hat{x}_2), \quad (2)$$

где

$$\Phi(t) = \frac{\varphi(t) - L_1(t)}{1 - t}, \quad 0 \leq t \leq 1.$$

Отметим, что слагаемое  $\tilde{x}_e(\hat{x}) = a_1 + B_e \hat{x}$  в правой части (2) задает отображение  $\hat{e}$  на  $\tilde{e}$ , а слагаемое  $\tilde{x}_e(\hat{x})$  не имеет особенностей, более того, используя правило Лопиталья, нетрудно показать, что функция  $\Phi(t)$  непрерывна.

**Теорема 2.** Пусть выполнены условия (1), более того,  $\chi \in C^{k+1}$ ,  $k \geq 1$ ,  $h$  достаточно мало, семейство элементов  $\tilde{T}_h$  регулярно, т. е.

$$c^{-1}h^2 \leq \det(B_e) \leq ch^2, \quad \|B_e\| \leq ch, \quad \|B_e^{-1}\| \leq ch^{-1} \quad (3)$$

с постоянной  $c$ , не зависящей от  $h$ . Тогда отображение (2) — непрерывное взаимно однозначное отображение  $\hat{e}$  на  $e$ , причем

$$c_1 h^2 \leq \det(Dx_e(\hat{x})) \leq c_2 h^2 \quad \forall \hat{x} \in \hat{e}, \quad (4)$$

$$|x_e|_{j,\infty,\hat{e}} \leq c h^j, \quad j = 1, 2, \dots, k, \quad (5)$$

т. е.  $\tilde{T}_h$  —  $k$ -регулярное семейство элементов.

**ДОКАЗАТЕЛЬСТВО.** Отметим прежде всего, что отображение (2) переводит  $\partial \hat{e}$  в  $\partial e$ . Точнее,  $x_e(\hat{a}_1 \hat{a}_2) = a_1 a_2$ ,  $x_e(\hat{a}_1 \hat{a}_3) = a_1 a_3$ ,  $x_e(\hat{a}_2 \hat{a}_3) = a_2 a_3$ . Действительно, первые два равенства очевидны, поскольку  $x_e(\hat{x}) = a_1 + B_e \hat{x}$ , если  $\hat{x}_1 = 0$  или  $\hat{x}_2 = 0$ . Для обоснования последнего равенства заметим, что  $a_1 + B_e \hat{x} = a_1 + \hat{x}_1(a_2 - a_1) + \hat{x}_2(a_3 - a_1)$ . Если  $\hat{x} \in \hat{a}_2 \hat{a}_3$ , т. е.  $\hat{x}_1 = 1 - \hat{x}_2$ , то  $a_1 + B_e \hat{x} = a_2 + \hat{x}_2(a_3 - a_2) = L_1(\hat{x}_2)$ , и, следовательно, при  $\hat{x} \in \hat{a}_2 \hat{a}_3$  имеем  $x = a_1 + B_e \hat{x} + \varphi(\hat{x}_2) - L_1(\hat{x}_2) = \varphi(\hat{x}_2)$ .

Докажем, что отображение  $\tilde{x}_e(\hat{x})$  является малым возмущением аффинного преобразования, точнее, что выполняются оценки

$$|\tilde{x}_e|_{j,\infty,\hat{e}} = O(h^{\max(j,2)}), \quad j = 0, 1, \dots, k. \quad (6)$$

Для этого достаточно убедиться в том, что

$$\Phi \in C^k, \quad |\Phi(t)| \leq ch^2, \quad |\Phi^{(j)}(t)| \leq ch^{j+1}, \quad j = 1, 2, \dots, k. \quad (7)$$

Докажем оценки (7). Ясно, что

$$|\varphi^{(j)}(t)| = \left| \frac{d^j \chi(s_2 + tl_e)}{ds^j} l_e^j \right| \leq cl_e^j.$$

Вследствие условия (1)

$$|a_2 \widehat{a}_3| = \int_{s_2}^{s_3} |\chi'(s)| ds \geq c_0 l_e.$$

С другой стороны, записывая выражение для длины дуги  $a_2 \widehat{a}_3$  в местной декартовой системе координат  $y_1 y_2$  (см. рис. 12), получим  $|a_2 \widehat{a}_3| \leq c |a_2 a_3| \leq ch$ , следовательно,  $l_e \leq ch$  и  $|\varphi^{(j)}(t)| \leq ch^j$ .

Заметим теперь, что<sup>1)</sup>

$$\Phi(t) = \varphi(1) - \varphi(0) - \frac{\varphi(1) - \varphi(t)}{1-t} = \int_0^1 (\varphi'(\tau) - \varphi'(t + \tau(1-t))) d\tau.$$

Таким образом,  $\Phi \in C^k$  и

$$\begin{aligned} \Phi(t) &= \int_0^1 \int_{t+\tau(1-t)}^{\tau} \varphi''(\theta) d\theta d\tau, \\ \Phi^{(j)}(t) &= - \int_0^1 \varphi^{j+1}(t + \tau(1-t)) (1-\tau)^j d\tau. \end{aligned} \quad (8)$$

Отсюда вытекают неравенства (7), а с ними и оценки (6). Из (6), очевидно, следуют оценки (5), поскольку элементы матрицы  $B_e$  имеют порядок  $h$ .

<sup>1)</sup>Для  $x(t) \in R^n$  как обычно  $x'(t) = (x'_1(t), \dots, x'_n(t))^T$ ,  $\int x(t) dt = (\int x_1(t) dt, \dots, \int x_n(t) dt)^T$ .

Покажем теперь, что при достаточно малом  $h$  отображение  $x_e$  взаимно однозначно. Пусть  $x_e(\hat{x}_1) = x_e(\hat{x}_2)$ ,  $\hat{x}_1, \hat{x}_2 \in \hat{e}$ . Тогда

$$B_e \hat{z} + \tilde{x}_e(\hat{x}_1) - \tilde{x}_e(\hat{x}_2) = 0,$$

где  $\hat{z} = \hat{x}_1 - \hat{x}_2$ . Отсюда имеем

$$\hat{z} + B_e^{-1} \int_0^1 D\tilde{x}_e(\hat{x}_2 + t\hat{z}) \hat{z} dt = 0,$$

и, следовательно<sup>1)</sup>,

$$|\hat{z}| \left(1 - \|B_e^{-1}\| \max_{0 \leq t \leq 1} \|D\tilde{x}_e(\hat{x}_2 + t\hat{z})\|\right) \leq 0. \quad (9)$$

Поскольку из (6) вытекает оценка  $\max_{0 \leq t \leq 1} \|D\tilde{x}_e(\hat{x}_2 + t\hat{z})\| \leq ch^2$ , то множитель при  $|\hat{z}|$  в (9) положителен при достаточно малом  $h$ , следовательно,  $\hat{x}_1 = \hat{x}_2$ , т. е. отображение  $x_e$  взаимно однозначно. Кроме того, как показано выше, отображение  $x_e$  переводит  $\partial\hat{e}$  в  $\partial e$ , поэтому на основании теоремы 1 образом  $\hat{e}$  является  $e$ .

Учитывая обратимость матриц  $B_e$ , получим

$$Dx_e(\hat{x}) = B_e + D\tilde{x}_e(\hat{x}) = B_e(I + B_e^{-1}D\tilde{x}_e(\hat{x})).$$

Как отмечалось выше,  $\|B_e^{-1}D\tilde{x}_e(\hat{x})\| \leq ch \leq \gamma < 1$  при достаточно малом  $h$  равномерно по  $\hat{x} \in \hat{e}$ . Отсюда следует обратимость матриц  $Dx_e(\hat{x})$  для всех  $\hat{x} \in \hat{e}$ , а также соотношения:

$$Dx_e^{-1}(x) = (Dx_e(\hat{x}))^{-1} = (I + B_e^{-1}D\tilde{x}_e(\hat{x}))^{-1} B_e^{-1},$$

$$\max_{\hat{x} \in \hat{e}} \|(I + B_e^{-1}D\tilde{x}_e(\hat{x}))^{-1}\| \leq \frac{1}{1 - \gamma},$$

$$\max_{x \in e} \|Dx_e^{-1}(x)\| \leq \frac{\|B_e^{-1}\|}{1 - \gamma} \leq \frac{c}{h}.$$

Докажем, наконец, оценки (4). Элементарные вычисления приводят к равенству  $\det(Dx_e(\hat{x})) = \det(b_{ij} + \tilde{b}_{ij})$ , где  $b_{ij}$  — элементы матрицы  $B_e$ , а  $\tilde{b}_{ij} = O(h^2)$ , следовательно,  $\det(Dx_e(\hat{x})) = \det(B_e) + O(h^3)$ , откуда при достаточно малом  $h$  вытекает нужная оценка для  $\det(Dx_e(\hat{x}))$ <sup>1)</sup>.  $\square$

<sup>1)</sup>Напомним, что при  $x \in R^n$  справедливы оценки  $|\int_0^1 B(t)x dt| \leq \int_0^1 |B(t)x| dt \leq \max_{0 \leq t \leq 1} \|B(t)\| |x|$ .

<sup>1)</sup>Напомним, что из невырожденности матрицы Якоби следует лишь локальная обратимость отображения.

ЗАМЕЧАНИЕ 1. Следует подчеркнуть, что регулярность семейства  $\mathcal{T}_h$  криволинейных элементов лимитируется только гладкостью границы  $\Gamma$ .

Итак, требуемое отображение элемента  $\hat{e}$  на  $e$  построено. Оно определяется по формуле (2). Соответствующий криволинейный лагранжев конечный элемент определяется затем так, как это было описано на с. 100. Достаточные условия для выполнения оценок (3) были указаны ранее на с. 111.

Полезно отметить, что сужение любой функции из  $P_e$  на криволинейную сторону элемента  $e$  есть полином от параметра  $s \in [s_1, s_2]$ . Точнее, справедливо

**Следствие 1.** Пусть выполнены условия теоремы 2,  $\hat{P} = P_l$ . Тогда для любой функции  $p_e \in P_e$  выполнено равенство

$$p_e(x) = \sum_{k=0}^l \alpha_k s^k, \quad x \in \Gamma, \quad s \in [s_1, s_2], \quad (10)$$

где  $\alpha_0, \alpha_1, \dots, \alpha_l$  — постоянные.

ДОКАЗАТЕЛЬСТВО. Ясно, что если  $x \in \Gamma$ , то по определению

$$p_e(x) = \sum_{0 \leq k_1 + k_2 \leq l} a_{k_1 k_2} \hat{x}_1^{k_1} \hat{x}_2^{k_2},$$

где  $\hat{x}$  однозначно определяется по  $x$ , причем  $\hat{x}_1 + \hat{x}_2 = 1$ , т. е.

$$p_e(x) = \sum_{k=0}^l \beta_k \hat{x}_2^k,$$

но  $\hat{x}_2 = (s - s_2)/l_e$  и, следовательно, (10) выполнено.  $\square$

Простой анализ доказательства теоремы 2 показывает, что справедлива следующая теорема, из которой вытекает, в частности, обратимость отображений, являющихся достаточно малыми возмущениями обратимых.

**Теорема 3.** Пусть отображение  $x = x_e(\hat{x})$  элемента  $\hat{e}$  на  $e \in \mathcal{T}_h$  имеет вид

$$x = x_e(\hat{x}) = \check{x}_e(\hat{x}) + \tilde{x}_e(\hat{x}),$$

причем

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

$$|\check{x}_e|_{j,\infty,\hat{e}} \leq ch^j, \quad |\tilde{x}_e|_{j,\infty,\hat{e}} \leq ch^{\max(j,2)}, \quad j = 1, 2, \dots, k,$$

с постоянными, не зависящими от  $h$ . Тогда при достаточно малом  $h$  отображение  $x = x_e(\hat{x})$  — непрерывное взаимно однозначное отображение  $\hat{e}$  на  $e$ , а семейство  $\mathcal{T}_h$  есть  $k$ -регулярное семейство.

УПРАЖНЕНИЕ 4. Доказать теорему 3.

## § 6. Изопараметрические конечные элементы

**1. Лагранжевы элементы.** Опишем сначала общий подход к построению изопараметрических лагранжевых элементов. Пусть  $(\hat{e}, \hat{P}, \hat{\Sigma})$  — лагранжев базисный конечный элемент в  $R^n$  и  $\hat{\omega} = \{\hat{a}_i\}_{i=1}^N$  есть множество его узлов интерполяции. Напомним, что в этом случае  $(\hat{P}, \hat{\Sigma})$  — унисольвентная пара.

**Лемма 1.** Для любого множества  $\omega = \{a_i\}_{i=1}^N \subset R^n$  существует единственное отображение  $x = x_e(\hat{x}) = (x_1(\hat{x}), \dots, x_n(\hat{x}))^T \in [\hat{P}]^n$  такое, что

$$x_e(\hat{a}_i) = a_i, \quad i = 1, 2, \dots, N. \quad (1)$$

**ДОКАЗАТЕЛЬСТВО.** Пусть  $\hat{\varphi}_i(\hat{x})$  базис Лагранжа в  $\hat{P}$ , т. е.  $\hat{\varphi}_i(\hat{a}_j) = \delta_{ij}$ . Тогда отображение<sup>1)</sup>

$$x_e(\hat{x}) = \sum_{i=1}^N \hat{\varphi}_i(\hat{x}) a_i = \left( \sum_{i=1}^N a_{i1} \hat{\varphi}_i(\hat{x}), \dots, \sum_{i=1}^N a_{in} \hat{\varphi}_i(\hat{x}) \right)^T \quad (2)$$

удовлетворяет условиям (1).

С другой стороны, если отображение  $x_e$  удовлетворяет условиям (1), то

$$x_j(\hat{a}_i) = a_{ij} \quad (3)$$

для всех  $i = 1, 2, \dots, N$ , причем  $x_j \in \hat{P}$ . Но  $(\hat{P}, \hat{\Sigma})$  — унисольвентная пара, поэтому для заданного множества  $\{a_{ij}\}_{i=1}^N$  существует единственная функция  $x_j \in \hat{P}$ , удовлетворяющая условиям (3) и она, фактически, определена выше формулой (2):

$$x_j(\hat{x}) = \sum_{i=1}^N a_{ij} \hat{\varphi}_i(\hat{x}). \quad \square$$

Лемма 1 позволяет описать следующий подход к построению изопараметрических конечных элементов, ассоциированных с заданным базисным элементом.

<sup>1)</sup>Через  $a_{ik}$ ,  $k = 1, 2, \dots, n$ , обозначены координаты точки  $a_i$ .

1) По заданному множеству точек  $\omega = \{a_i\}_{i=1}^N \subset R^n$  и базисному элементу  $\hat{e}$  определяем при помощи формулы (2) отображение  $x = x_e(\hat{x})$ , а затем множество  $e = x_e(\hat{e})$ . Множество  $\omega$  следует выбирать так, чтобы  $e$  имело непустую внутренность, а отображение  $x_e : \hat{e} \rightarrow e$  было взаимно однозначным.

2) Используя отображение  $x = x_e(\hat{x})$ , определяем конечный элемент  $(e, P_e, \Sigma_e)$ , ассоциированный с базисным, согласно теореме 1, с. 100.

Отметим следующий тривиальный случай. Пусть задан элемент  $(\tilde{e}, P_{\tilde{e}}, \Sigma_{\tilde{e}})$  аффинно-эквивалентный элементу  $(\hat{e}, \hat{P}, \hat{\Sigma})$ , причем  $\hat{P}$  содержит  $P_1$ . Выберем в качестве узлов  $\tilde{\omega} = \{\tilde{a}_i\}$  узлы интерполяции на элементе  $\tilde{e}$  и построим отображение  $x = x_{\tilde{e}}$  по формуле (2):

$$x_{\tilde{e}}(\hat{x}) = \sum_{i=1}^N \hat{\varphi}_i(\hat{x}) \tilde{a}_i. \quad (4)$$

Это отображение оказывается аффинным, поскольку аффинное отображение  $\hat{e}$  на  $\tilde{e}$  удовлетворяет условиям (1), а в силу леммы 1 других отображений  $x_{\tilde{e}}(\hat{x}) \in [\hat{P}]^n$ , удовлетворяющих условиям (1), не существует. Таким образом,

$$x_{\tilde{e}}(\hat{x}) = b_{\tilde{e}} + B_{\tilde{e}} \hat{x}. \quad (5)$$

Это простое замечание позволяет надеяться, что можно удовлетворить условию 1), выбирая множества  $\omega$  так, что  $a_i = \tilde{a}_i + \varepsilon_i, i = 1, 2, \dots, N$ , где  $\tilde{a}_i$  — узлы интерполяции элемента, аффинно-эквивалентного  $\hat{e}$ , а векторы  $\varepsilon_i$  достаточно малы.

**ЗАМЕЧАНИЕ 1.** Польза от введения в рассмотрение изопараметрических элементов, в основном, заключается в том, что мы можем распорядиться выбором точек  $\omega$  так, чтобы получить более общие геометрические формы элементов по сравнению с аффинно-эквивалентными элементами, и более простые элементы (в смысле конструктивного определения отображения  $x = x_e$ ) по сравнению с криволинейными элементами. В случае изопараметрических элементов нет необходимости также в параметризации криволинейной границы, что существенно облегчает программную реализацию метода конечных элементов.

Рассмотрим примеры изопараметрических элементов.

### 1. Квадратичный изопараметрический элемент.

Пусть  $(\hat{e}, \hat{P}, \hat{\Sigma})$  — квадратичный треугольный базисный конечный элемент в  $R^2$ . Выберем на плоскости три точки  $a_1, a_3, a_5$ , считая, что треугольник  $\tilde{e}$  с вершинами в этих точках имеет диаметр, меньший,

чем  $h$ . Множество таких треугольников обозначим через  $\tilde{\mathcal{T}}_h$ . Пусть далее точки  $\tilde{a}_2, \tilde{a}_4, \tilde{a}_6$  расположены на серединах сторон треугольника  $\tilde{e}$ , а узлы  $a_2, a_4, a_6$  выбраны соответственно в  $O(h^2)$  окрестности этих точек (рис. 13), т. е.  $|\tilde{a}_i - a_i| = O(h^2), i = 2, 4, 6$ . Рассмотрим образ

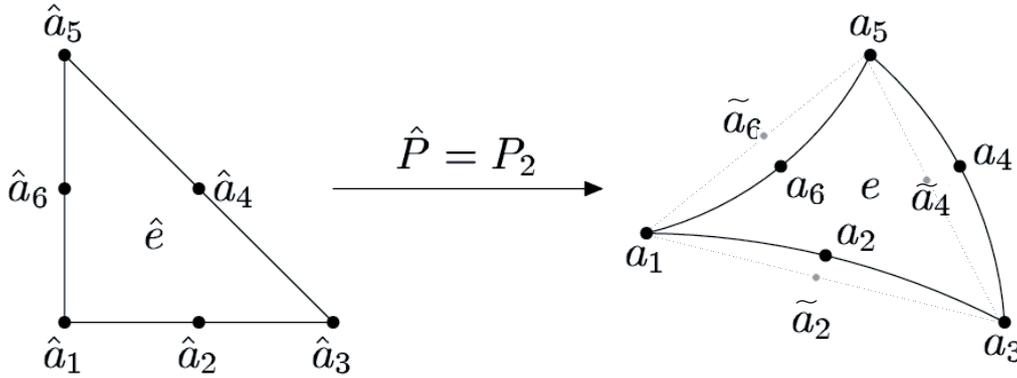


Рис. 13. Квадратичный изопараметрический элемент.

элемента  $\hat{e}$  при отображении, определяемом формулой (2) и введенными узлами  $\{a_i\}_{i=1}^6$ . По определению узлы  $\hat{a}_i$  переходят в узлы  $a_i$ . Образами сторон элемента  $\hat{e}$  являются дуги соответствующих парабол. В самом деле, например, образ стороны  $\hat{a}_1\hat{a}_3$  полностью определяется узлами  $a_1, a_2, a_3$  (поскольку базисные функции  $\varphi_i(\hat{x})$  с номерами, отличными от 1, 2, 3, обращаются в нуль на стороне  $\hat{a}_1\hat{a}_3$ ), и при этом каждая компонента отображения (2) является квадратичной функцией  $\hat{x}_1$ . Таким образом, возникает криволинейный треугольник  $e$ , ограниченный дугами парабол, проходящими через узлы  $a_i, i = 1, \dots, 6$ . Множество таких элементов обозначим через  $\mathcal{T}_h$  и назовем семейством изопараметрических квадратичных элементов.

Пусть  $\hat{\varphi}_i$  — базисные функции на базисном элементе,  $i = 1, \dots, 6$ . С учетом формул (4), (5) имеем

$$x_e(\hat{x}) = a_1 + B_e \hat{x} + \tilde{x}_e(\hat{x}), \quad \tilde{x}_e(\hat{x}) = \sum_{k=1}^3 (\tilde{a}_{2k} - a_{2k}) \hat{\varphi}_{2k}(\hat{x}), \quad (6)$$

где матрица  $B_e$  определяется узлами  $a_1, a_3, a_5$ . Следующая теорема является аналогом теоремы 2, с. 115 (см. также теорему 3, с. 118).

**Теорема 1.** Пусть  $h$  достаточно мало и семейство треугольников  $\tilde{\mathcal{T}}_h$  регулярно. Тогда отображение (6) — непрерывное взаимно

однозначное отображение  $\hat{e}$  на  $e$ , причем

$$c_1 h^2 \leq \det(Dx_e(\hat{x})) \leq c_2 h^2 \quad \forall \hat{x} \in \hat{e},$$

$$|x_e|_{j,\infty,\hat{e}} \leq c h^j, \quad j = 1, 2, \quad |x_e|_{k,\infty,\hat{e}} = 0, \quad k \geq 3,$$

т. е. семейство изопараметрических квадратичных элементов  $\mathcal{T}_h$  регулярно.

Доказательство теоремы опустим, поскольку оно проводится так же, как и доказательство теоремы 2, с. 115, для криволинейных элементов, которое было основано на малости нелинейного возмущения  $\tilde{x}_e(\hat{x})$ . В данном случае оно также имеет нужный порядок малости, поскольку для любых  $\alpha$ ,  $|\alpha| \leq 2$ , и  $\hat{x} \in \hat{e}$

$$|D^\alpha \tilde{x}_e(\hat{x})| = \left| \sum_{k=1}^3 (\tilde{a}_{2k} - a_{2k}) D^\alpha \hat{\varphi}_{2k}(\hat{x}) \right| \leq c \sum_{k=1}^3 |\tilde{a}_{2k} - a_{2k}| \leq c h^2,$$

причем  $D^\alpha \tilde{x}_e(\hat{x}) = 0$  для  $|\alpha| \geq 3$ .

ЗАМЕЧАНИЕ 2. Обратимость отображения  $x = x_e(\hat{x})$  при малых  $h$  можно гарантировать (проверьте!), если потребовать лишь, что  $|\tilde{a}_i - a_i| = o(h)$ ,  $i = 2, 4, 6$ . Мы накладываем более жесткие условия на расположение узлов, чтобы обеспечить регулярность семейства  $\mathcal{T}_h$ .

## 2. Треугольный изопараметрический элемент с одной криволинейной границей.

Если при определении элемента, рассмотренного выше, положить  $a_2 = \tilde{a}_2$ ,  $a_6 = \tilde{a}_6$ , то мы получим изопараметрический элемент с одной криволинейной стороной, изображенный на рис. 14. Такие элементы обычно используются как приграничные — при триангуляции области с криволинейной границей. В этом случае преобразование  $x_e(\hat{x})$  имеет

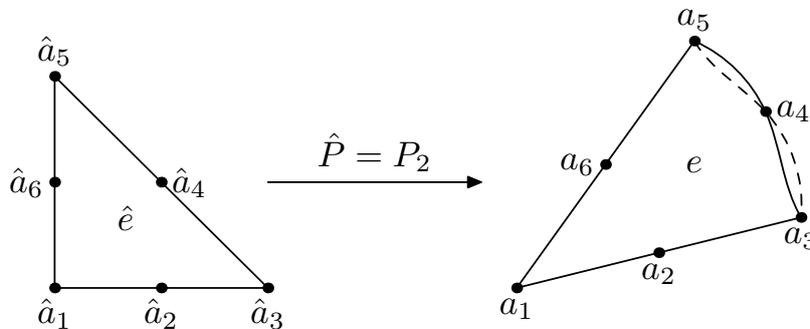


Рис. 14. Дуга параболы аппроксимирует границу области.

особенно простой вид:

$$x_e(\hat{x}) = a_1 + B_e \hat{x} + (\tilde{a}_4 - a_4) \hat{\varphi}_4(\hat{x}), \quad \tilde{a}_4 = (a_3 + a_5)/2.$$

Напомним, что  $\hat{\varphi}_4(\hat{x}) = 4\hat{x}_1\hat{x}_2$  (см. с. 90).

Рассмотрим аналогичную ситуацию в общем случае. Пусть  $(\hat{e}, \hat{P}, \hat{\Sigma})$  есть треугольный базисный конечный элемент степени  $m$  в  $R^2$ , и пусть  $\tilde{e}$  — криволинейный треугольник с двумя прямолинейными сторонами и криволинейной стороной  $\Gamma$ , параметризуемой так, как указано на с. 115, с диаметром меньшим  $h$ . Множество таких элементов  $\tilde{e}$  обозначим через  $\tilde{\mathcal{T}}_h$ . Выберем на  $\Gamma$  точки  $\chi(t_j)$ ,  $t_j = j/m$ ,  $j = 0, 1, \dots, m$ , и заменим функции  $\varphi_1(t)$ ,  $\varphi_2(t)$ , определяющие параметризацию  $\varphi(t) = (\varphi_1(t), \varphi_2(t))^T$  кривой  $\Gamma$ , интерполяционными полиномами  $L_m^1(t)$ ,  $L_m^2(t)$  степени  $m$ , полагая, что

$$\varphi_k(t_j) = L_m^k(t_j), \quad t_j = j/m, \quad j = 0, 1, \dots, m.$$

При этом нелинейное отображение  $x = x_e(\hat{x})$ , определяемое по формуле (2), с. 115, заменяется на полиномиальное отображение

$$x = \bar{x}_e(\hat{x}) = a_1 + B_e \hat{x} + \hat{x}_1 \Phi_{m-1}(\hat{x}_2), \quad (7)$$

где

$$\Phi_{m-1}(t) = \frac{\varphi_m(t) - L_1(t)}{1-t}, \quad \varphi_m(t) = (L_m^1(t), L_m^2(t))^T.$$

Каждая из компонент функции  $\Phi_{m-1}$  — полином степени  $m-1$ , поскольку точка  $t=1$  — корень каждого из полиномов  $L_m^k - L_1^k$ ,  $k=1, 2$ . Таким образом, компоненты вектор-функции  $\bar{x}_e$  — полиномы степени  $m$  (по совокупности переменных).

Заметим, что отображение (7) можно представить в виде

$$x = \bar{x}_e(\hat{x}) = x_e(\hat{x}) + \check{x}_e(\hat{x}), \quad \check{x}_e(\hat{x}) = \hat{x}_1(\Phi_{m-1}(\hat{x}_2) - \Phi(\hat{x}_2)),$$

причем добавок  $\check{x}_e(\hat{x})$  можно рассматривать как возмущение отображения  $x_e(\hat{x})$ , исследованного нами ранее (см. с. 115). Отметим также, что образом элемента  $\hat{e}$  при отображении  $x = \bar{x}_e(\hat{x})$  является элемент  $e$ , получающийся из элемента  $\tilde{e}$  заменой дуги  $\Gamma$  на кривую  $x = \varphi_m(t)$ ,  $0 \leq t \leq 1$ . Множество таких элементов обозначим через  $\mathcal{T}_h$  и назовем семейством треугольных изопараметрических элементов степени  $m$ .

**Теорема 2.** Пусть  $\Gamma \in C^{m+1}$ ,  $h$  достаточно мало, семейство криволинейных элементов  $\tilde{\mathcal{T}}_h$  есть  $m$ -регулярное семейство. Тогда

отображение (7) — непрерывное взаимно однозначное отображение  $\hat{e}$  на  $e$ , причем

$$\begin{aligned}\bar{x}_e(\hat{x}) &= x_e(\hat{x}) + O(h^{m+1}), \quad \bar{x}_e(\hat{x}) \in [\hat{P}]^2, \\ c_1 h^2 &\leq \det(D\bar{x}_e(\hat{x})) \leq c_2 h^2 \quad \forall \hat{x} \in \hat{e}, \\ |\bar{x}_e|_{j,\infty,\hat{e}} &\leq c h^j, \quad j = 1, 2, \dots, m, \quad |\bar{x}_e|_{j,\infty,\hat{e}} = 0, \quad j \geq m + 1,\end{aligned}$$

т. е. семейство треугольных изопараметрических элементов регулярно.

ДОКАЗАТЕЛЬСТВО. Если мы покажем, что

$$\left| \Phi^{(j)}(t) - \Phi_{m-1}^{(j)}(t) \right| \leq c h^{m+1}, \quad 0 \leq t \leq 1, \quad j = 0, 1, \dots, m,$$

то все утверждения теоремы, очевидно, будут непосредственно вытекать из теоремы 3, с. 118. Применяя формулу Лагранжа для остаточного члена интерполяционного полинома, получим

$$\Phi(t) - \Phi_{m-1}(t) = \frac{\varphi(t) - \varphi_m(t)}{1-t} = -\frac{t(t-t_1)(t-t_2)\dots(t-t_{m-1})}{(m+1)!} \varepsilon_{m+1},$$

где  $\varepsilon_{m+1} = (\varphi_1^{(m+1)}(\xi_1), \varphi_2^{(m+1)}(\xi_2))$ ,  $\xi_1, \xi_2 \in (0, 1)$ , т. е.

$$|\Phi(t) - \Phi_{m-1}(t)| \leq c \max_{0 \leq t \leq 1} |\varphi^{(m+1)}(t)| \leq c h^{m+1}.$$

Для оценки производных запишем  $\Phi(t) - \Phi_{m-1}(t)$  в виде

$$\Phi(t) - \Phi_{m-1}(t) = \frac{(\varphi(t) - L_1(t)) - (\varphi_m(t) - L_1(t))}{1-t}.$$

Ясно, что  $\varphi_m(0) = L_1(0)$ ,  $\varphi_m(1) = L_1(1)$ , поэтому аналогично (8), с. 116, при  $j \geq 1$  получим

$$|\Phi^{(j)}(t) - \Phi_{m-1}^{(j)}(t)| \leq c \max_{0 \leq t \leq 1} |\varphi^{(j+1)}(t) - \varphi_m^{(j+1)}(t)|,$$

откуда после применения хорошо известной оценки погрешности интерполирования (см., например, [1]) будем иметь

$$|\Phi^{(j)}(t) - \Phi_{m-1}^{(j)}(t)| \leq c \max_{0 \leq t \leq 1} |\varphi^{(m+1)}(t)| \leq c h^{m+1}. \quad \square$$

Используя теперь отображение  $x = \bar{x}_e(\hat{x})$ , можем определить изопараметрический конечный элемент  $(e, P_e, \Sigma_e)$ ,  $e \in \mathcal{T}_h$ , ассоциированный с базисным согласно теореме 1, с. 100 (см. рис. 15, соответствующий случаю  $m = 3$ ).

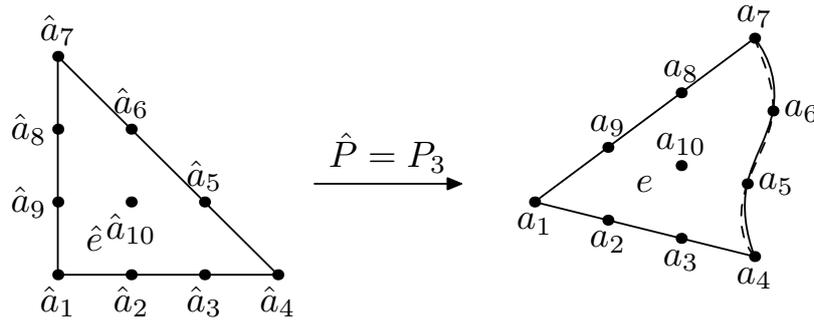


Рис. 15. Изопараметрический треугольный элемент третьей степени.

ЗАМЕЧАНИЕ 3. Условия  $m$ -регулярности семейства криволинейных элементов  $\mathcal{T}_h$  указаны в теореме 2 на с. 115.

В дальнейшем будет полезна следующая

**Теорема 3.** Пусть выполнены условия теоремы 2. Тогда для достаточно малых  $h$

$$\text{dist}(\Gamma, \Gamma_m) = \sup_{x \in \Gamma} \inf_{y \in \Gamma_m} |x - y| \leq ch^{m+1}, \quad (8)$$

$$\text{dist}(x(\hat{a}), \Gamma) = \inf_{y \in \Gamma} |x(\hat{a}) - y| \geq c \text{dist}(\hat{a}, \partial \hat{e})h, \quad (9)$$

где  $\hat{a}$  — произвольная внутренняя точка  $\hat{e}$ .

ДОКАЗАТЕЛЬСТВО. Пусть  $x \in \Gamma$ . Тогда существует точка  $\hat{x} \in \hat{a}_2\hat{a}_3$  такая, что  $x = x_e(\hat{x})$ . При этом  $y = \bar{x}_e(\hat{x}) \in \Gamma_m$ . Вследствие теоремы 2 имеем:  $|y - x| \leq ch^{m+1}$  равномерно по  $\hat{x} \in \hat{a}_2\hat{a}_3$ , следовательно, оценка (8) доказана. Далее, если  $\hat{a} \in \text{int } \hat{e}$ ,  $\hat{b} \in \hat{a}_2\hat{a}_3$ , то по теореме 1, с. 115, при достаточно малых  $h$  справедливы включения  $a = x_e(\hat{a}) \in \text{int } \tilde{e}$ ,  $b = x_e(\hat{b}) \in \Gamma$ , причем (см. с. 117)

$$|\hat{a} - \hat{b}| = |x_e^{-1}(a) - x_e^{-1}(b)| \leq \max_{x \in e} \|Dx_e^{-1}(x)\| |a - b| \leq \frac{c}{h} |a - b|,$$

т. е.  $|a - b| \geq ch|\hat{a} - \hat{b}| \geq c \text{dist}(\hat{a}, \partial \hat{e})h$ , откуда вытекает (9).  $\square$

**Следствие 1.** Если  $\hat{a}$  внутренняя точка  $\hat{e}$ , то при достаточно малых  $h$  точка  $\bar{x}_e(\hat{a})$  — внутренняя точка  $\tilde{e}$  и, следовательно, области  $\Omega$ .

### Задачи.

1. Показать, что при  $m = 2$  (см. рис. 14), функция  $\Phi_{m-1}(\hat{x}_2)$  имеет вид

$$\Phi_1(\hat{x}_2) = 4\left(a_4 - \frac{a_3 + a_5}{2}\right)\hat{x}_2.$$

2. Показать, что при  $m = 3$  (см. рис. 15)

$$\Phi_2(\hat{x}_2) = 9/2((a_4 - a_7 + 3(a_6 - a_5))\hat{x}_2 + 2a_5 - a_6 - a_4)\hat{x}_2.$$

### 3. Изопараметрический четырехугольный элемент первой степени.

Пусть  $(\hat{e}, \hat{P}, \hat{\Sigma})$  — билинейный базисный конечный элемент. Выберем на плоскости четыре точки  $a_1, \dots, a_4$ , считая, что четырехугольник  $e$  с вершинами в этих точках обладает следующими свойствами для заданного  $h > 0$ :

- 1)  $e$  — выпуклый;
- 2)  $\text{diam}(e) \leq h$ ;
- 3) углы  $e$  лежат в пределах  $[\theta_0, \pi - \theta_0]$ ,  $\theta_0 > 0$  не зависит от  $h$ ;
- 4) длины сторон лежат в пределах от  $c_1 h$  до  $c_2 h$ , где постоянные  $0 < c_1 \leq c_2$  не зависят от  $h$ .

Множество таких элементов обозначим через  $\mathcal{T}_h$  (см. рис. 16). Не-

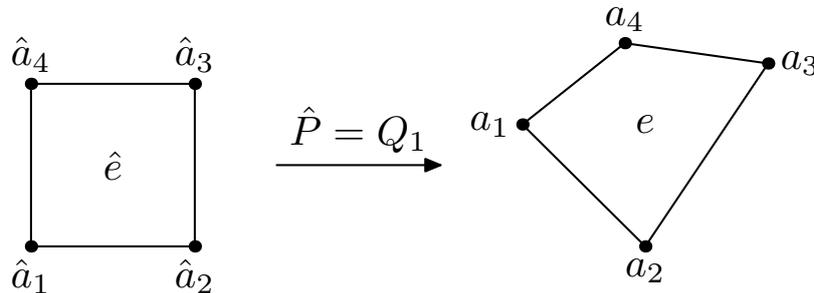


Рис. 16. Изопараметрический четырехугольный элемент первой степени.

посредственно проверяется в силу аффинности функций из  $Q_1$  по направлению каждой координатной оси, что отображение

$$x = x_e(\hat{x}) = \sum_{k=1}^4 a_k \hat{\varphi}_k(\hat{x}), \quad (10)$$

определяемое согласно формуле (2), отображает стороны  $\hat{e}$  на соответствующие стороны  $e$  и внутренность  $\hat{e}$  — на внутренность  $e$ . Напомним, что функции  $\varphi_k(\hat{x})$  в этом случае определяются формулами (??), с. ??, следовательно, преобразование (10) можно записать в виде

$$x_e(\hat{x}) = (1 - \hat{x}_1)(1 - \hat{x}_2)a_1 + \hat{x}_1(1 - \hat{x}_2)a_2 + \hat{x}_1\hat{x}_2a_3 + (1 - \hat{x}_1)\hat{x}_2a_4. \quad (11)$$

**Теорема 4.** *Отображение (11) — непрерывное взаимно однозначное отображение  $\hat{e}$  на  $e \in \mathcal{T}_h$ , причем*

$$c_1 h^2 \leq \det(Dx_e(\hat{x})) \leq c_2 h^2 \quad \forall \hat{x} \in \hat{e}, \quad [x_e]_{1,\infty,\hat{e}} \leq c h, \quad (12)$$

$$|x_e|_{k,\infty,\hat{e}} \leq c h, \quad k = 1, 2, \quad |x_e|_{k,\infty,\hat{e}} = 0, \quad k \geq 3, \quad [x_e]_{j,\infty,\hat{e}} = 0, \quad j \geq 2, \quad (13)$$

т. е. семейство четырехугольных элементов  $\mathcal{T}_h$  регулярно.

**ДОКАЗАТЕЛЬСТВО.** Пусть  $d = a_1 + a_3 - a_2 - a_4$ . Из формулы (11) следует, что

$$x_e(\hat{x}) = a_1 + \hat{x}_1(a_2 - a_1) + \hat{x}_2(a_4 - a_1) + \hat{x}_1\hat{x}_2d,$$

т. е.  $x_e(\hat{x}) = a_1 + B_e\hat{x} + \hat{x}_1\hat{x}_2d$ . Матрица Якоби этого отображения имеет вид

$$Dx_e(\hat{x}) = \begin{pmatrix} (a_2 - a_1)_1 + d_1\hat{x}_2 & (a_4 - a_1)_1 + d_1\hat{x}_1 \\ (a_2 - a_1)_2 + d_2\hat{x}_2 & (a_4 - a_1)_2 + d_2\hat{x}_1 \end{pmatrix}.$$

Определитель этой матрицы является аффинной функцией от  $\hat{x}_1, \hat{x}_2$  и поэтому достигает своих экстремальных значений в вершинах элемента  $\hat{e}$ , т. е. в узлах  $\hat{a}_i$ . Нетрудно проверить, что  $\det(Dx_e(\hat{a}_i))$  равен удвоенной площади треугольника, построенного на сторонах  $e$ , имеющих общей вершиной  $a_i$ . Отсюда следует первая оценка (12). Поскольку  $\text{diam}(e) \leq h$ , то  $|a_2 - a_1| \leq h$ ,  $|a_4 - a_1| \leq h$ ,  $|d| \leq h$ . Отсюда следует вторая оценка в (12) и первая оценка в (13).  $\square$

Для того, чтобы семейство  $\mathcal{T}_h$  было сильно регулярным, дополнительно необходимо потребовать выполнения неравенства  $\|D^2x_e(\hat{x})\| \leq c h^2$  для любых  $\hat{x} \in \hat{e}$ . Поскольку,

$$\frac{\partial^2 x_e}{\partial \hat{x}_i^2} \equiv 0, \quad i = 1, 2, \quad \frac{\partial^2 x_e}{\partial \hat{x}_1 \partial \hat{x}_2} \equiv d,$$

то оценки  $\|D^2x_e(\hat{x})\| \leq c h^2$  и  $|d| = |a_1 + a_3 - a_2 - a_4| \leq c h^2$  эквивалентны.

Нетрудно видеть, что условие  $|d| \leq c h^2$  приводит к следующему дополнительному ограничению на выбор вершин элементов из  $\mathcal{T}_h$ : точка  $a_i$  должна лежать в  $O(h^2)$ -окрестности точки  $\tilde{a}_i$  такой, что  $\tilde{a}_i$ ,  $i = 1, \dots, 4$ , являются вершинами некоторого параллелограмма. Таким образом, требование сильной регулярности триангуляции приводит в рассматриваемом случае к элементам, «близким» к параллелограммам.

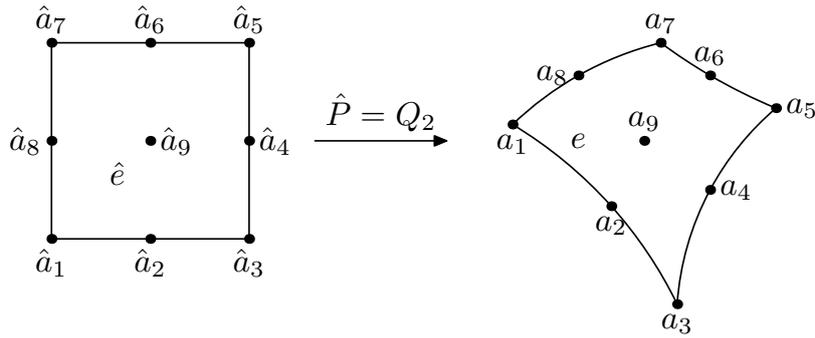


Рис. 17. Прямоугольный изопараметрический элемент второй степени.

**Задачи.**

1. Пусть  $\Omega$  — произвольный выпуклый четырехугольник. Определим разбиение  $\Omega$  на семейство  $\mathcal{T}_h$  изопараметрических четырехугольников первой степени следующим образом. Каждую сторону  $\Omega$  разобьем на  $n$  равных частей и соединим соответствующие точки разбиения противоположных сторон четырехугольника отрезками прямых. В результате образуется  $n^2$  «маленьких» четырехугольников, которые и отнесем к  $\mathcal{T}_h$  ( $h = h(n)$ ). Доказать, что семейство  $\mathcal{T}_h$  является сильно  $k$ -регулярным,  $k \geq 2$ .

2. Исследовать изопараметрический элемент, изображенный на рис. 17. Показать, что если узлы  $A = \{a_2, a_4, a_6, a_8, a_9\}$  выбрать в  $O(h^2)$  окрестности соответствующих середин сторон и центра тяжести четырехугольника  $a_1 a_3 a_5 a_7$  (с прямолинейными сторонами), удовлетворяющего условиям 1)–4), с. 126, то преобразование, определяемое согласно формуле (2), с. 119, является при достаточно малом  $h$  взаимно однозначным отображением  $\hat{e}$  на  $e$ , причем

$$\begin{aligned} c_1 h^2 &\leq \det(Dx_e(\hat{x})) \leq c_2 h^2 \quad \forall \hat{x} \in \hat{e}, \\ |x_e|_{j,\infty,\hat{e}} &\leq c h, \quad 1 \leq j \leq 4, \quad |x_e|_{j,\infty,\hat{e}} = 0, \quad j \geq 5, \\ [x_e]_{j,\infty,\hat{e}} &\leq c h^j, \quad j = 1, 2, \quad [x_e]_{k,\infty,\hat{e}} = 0, \quad k \geq 3, \end{aligned}$$

т. е. семейство элементов указанного вида регулярно.

Указание. Обратите внимание, что если радиус, указанной выше окрестности равен нулю, то преобразование  $x = x_e(\hat{x})$  имеет вид (11). Представьте искомое преобразование как  $O(h^2)$ -возмущение (определяемое множеством узлов  $A$ ) отображения (11) и воспользуйтесь теоремой 3, с. 118.

3. Исследовать изопараметрический элемент, являющийся следующим частным случаем элемента из предыдущего упражнения: узлы  $A = \{a_2, a_4, a_8, a_9\}$  совпадают с соответствующими серединами сторон и центром тяжести четырехугольника  $a_1 a_3 a_5 a_7$ . Такие элементы естественно использовать как приграничные. Проверить, что

$$x_e(\hat{x}) = \tilde{x}_e(\hat{x}) + (a_6 - \tilde{a}_6) \hat{x}_1 \hat{x}_2 (1 - \hat{x}_1), \quad \tilde{a}_6 = (a_5 + a_7)/2,$$

где  $\tilde{x}_e(\hat{x}) \in [Q_1]^2$  определяется вершинами  $a_1, a_3, a_5, a_7$  и имеет вид (11). Показать, что условие сильной  $k$ -регулярности,  $k \geq 3$ , равносильно требованию  $O(h^2)$ -близости к параллелограмму, а также условию  $|a_6 - \tilde{a}_6| \leq c h^3$ . Последнее ограничение не является излишне жестким и выполняется, например, если точки  $a_5, a_7$  лежат на некоторой кривой  $\Gamma \subset C^3$ , а точка  $a_6$  выбрана подходящим образом.

**2. Эрмитовы изопараметрические элементы.** Рассмотрим кратко эрмитовы изопараметрические элементы, узловые параметры

которых включают производные по направлениям порядка не выше первого, т. е. элементы класса  $C^0$ . Пусть  $(\hat{e}, \hat{P}, \hat{\Sigma})$  — базисный элемент такого типа в  $R^n$ . Из множества всех его узлов интерполяции выделим подмножество  $\hat{\omega}^0$ , которому соответствуют функционалы «значение функции в точке», и подмножество  $\hat{\omega}^1$ , которому соответствуют производные по направлениям. Пусть  $\hat{\omega}^0 = \{\hat{a}_i^0\}_{i=1}^{N_0}$ ,  $\hat{\omega}^1 = \{\hat{a}_i^1\}_{i=1}^{N_1} \subset \hat{\omega}^0$ . С каждой точкой  $\hat{a}_i^1 \in \hat{\omega}^1$  свяжем подпространство  $\hat{\Phi}_i \subset R^n$  размерности  $d_i$  с базисом  $\hat{\nu}_{i,1}, \dots, \hat{\nu}_{i,d_i}$ . Пусть далее множество  $\hat{\Sigma}$  узловых параметров  $\hat{\phi}_i$  имеет вид

$$\begin{aligned}\hat{\phi}_i^0(p) &= p(\hat{a}_i^0), \quad i = 1, 2, \dots, N_0, \\ \hat{\phi}_{i,k}^1(p) &= Dp(\hat{a}_i^1)\hat{\nu}_{i,k}, \quad i = 1, 2, \dots, N_1, \quad k = 1, 2, \dots, d_i,\end{aligned}$$

и  $\dim \hat{P} = N_0 + d_1 + \dots + d_{N_1-1}$ .

**Лемма 2.** Для любого множества  $\omega^0 = \{a_i^0\}_{i=1}^{N_0}$ , и любых  $N_1$  подпространств  $\Phi_i \subset R^n$  с базисами  $\nu_{i,1}, \dots, \nu_{i,d_i}$  существует, и притом единственное, отображение  $x_e(\hat{x}) = (x_1(\hat{x}), \dots, x_n(\hat{x}))^T$  такое, что

- a)  $x_i \in \hat{P}$  для всех  $i$ ;
- b)  $x_e(\hat{a}_i^0) = a_i^0$ ,  $i = 1, 2, \dots, N_0$ ;
- c)  $Dx_e(\hat{a}_i^1)\hat{\nu}_{i,k} = \nu_{i,k}$ ,  $i = 1, 2, \dots, N_1$ ,  $k = 1, 2, \dots, d_i$ .

ДОКАЗАТЕЛЬСТВО. Определим отображение

$$x_e = \sum_{j=1}^{N_0} a_j^0 \hat{\varphi}_j^0 + \sum_{j=1}^{N_1} \sum_{k=1}^{d_j} \nu_{j,k} \hat{\varphi}_{j,k}^1, \quad (14)$$

где  $\hat{\varphi}_j^0$ ,  $\hat{\varphi}_{j,k}^1$  базис Лагранжа в  $\hat{P}$ , т. е.

$$\hat{\varphi}_j^0(\hat{a}_i^0) = \delta_{ji}, \quad \hat{\varphi}_{j,k}^1(\hat{a}_i^0) = 0,$$

$$D\hat{\varphi}_j^0(\hat{a}_i^1)\hat{\nu}_{i,l} = 0, \quad D\hat{\varphi}_{j,k}^1(\hat{a}_i^1) = \delta_{ji}\delta_{kl}$$

для всех допустимых значений индексов  $i, j, k, l$ .

Нетрудно проверить, что это отображение удовлетворяет условиям a)–c). Вследствие унисольвентности  $\hat{\Sigma}$  и  $\hat{P}$  оно единственно.  $\square$

В качестве примера рассмотрим эрмитов кубический треугольный элемент. Имеем  $\hat{P} = P_3$ , множество  $\omega^0 = \omega$  состоит из вершин и центра тяжести треугольника  $\hat{e}$ ,  $\omega^1$  — множество его вершин. Далее,  $\hat{\xi}_{i,k} = \hat{a}_k - \hat{a}_i$ ,  $i \neq k$ ,  $i, k = 1, 2, 3$  (см. левый рис. 18).

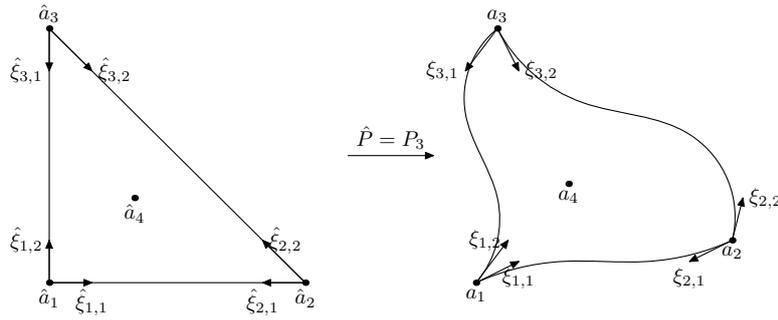


Рис. 18. Эрмитов изопараметрический кубический элемент.

Зададим на плоскости три точки  $a_1, a_2, a_3$  и пусть  $a_4 = (a_1 + a_2 + a_3)/3$ . Положим  $\omega_0 = \{a_1, a_2, a_3, a_4\}$ . В первых трех точках определим по два вектора  $\xi_{i,k}$  и зададим отображение  $x = x_e(\hat{x})$  по этим данным согласно формуле (14). Оно переводит узлы  $\hat{a}_i$  в  $a_i$ , стороны элемента  $\hat{e}$  — в дуги кубических кривых. Отметим, что согласно условию *c*) векторы  $\xi_{i,k}$ , образы базисных векторов  $\hat{\xi}_{i,k}$ , являются касательными к этим кривым. Таким образом, приходим к криволинейному треугольнику  $e$ , изображенному справа на рис. 18. Как и ранее, можно заметить, что если векторы  $\xi_{i,k}$  выбрать в виде  $\xi_{i,k} = B_e \hat{\xi}_{i,k}$  (т. е.  $\hat{\xi}_{i,k} = a_k - a_i$ ), где  $B_e$  — матрица соответствующего аффинного преобразования, то отображение  $x = x_e(\hat{x})$  будет аффинным и, соответственно, построенный изопараметрический элемент будет аффинно-эквивалентным базисному. Теперь ясно, что представляя преобразование  $x = x_e(\hat{x})$  как малое возмущение аффинного, можно провести анализ его свойств. Отметим важный частный случай, представлен-

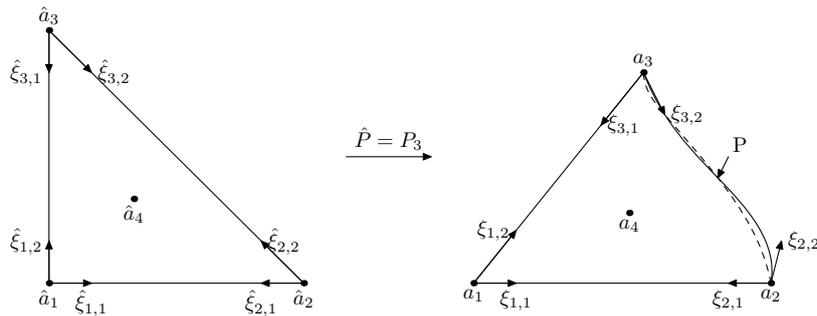


Рис. 19. Приграничный изопараметрический эрмитов элемент.

ный на рис. 19. Такие элементы используются как приграничные — при триангуляции областей с криволинейными границами.

## § 7. Пространства конечных элементов

Целью настоящего параграфа является построение для различных типов областей  $\Omega \subset R^n$ ,  $n = 1, 2$ , конечномерных пространств  $V_h$ , аппроксимирующих пространства Соболева  $V = H^k(\Omega)$ ,  $k = 1, 2$ .

Для всех рассматриваемых нами областей (отрезок, многоугольник, область, составленная из прямоугольников, область с кусочно гладкой границей) построения будут вестись по единой схеме, включающей следующие шаги:

1) множество  $\bar{\Omega}$  представляется в виде такого объединения конечных элементов  $e$  семейства  $\mathcal{T}_h$  выбранной формы, что различные элементы из  $\mathcal{T}_h$  либо не пересекаются, либо имеют общую вершину или общую грань; указанное разбиение будем называть правильной триангуляцией области  $\Omega$ , или просто триангуляцией области  $\Omega$  (см. рис. 20);

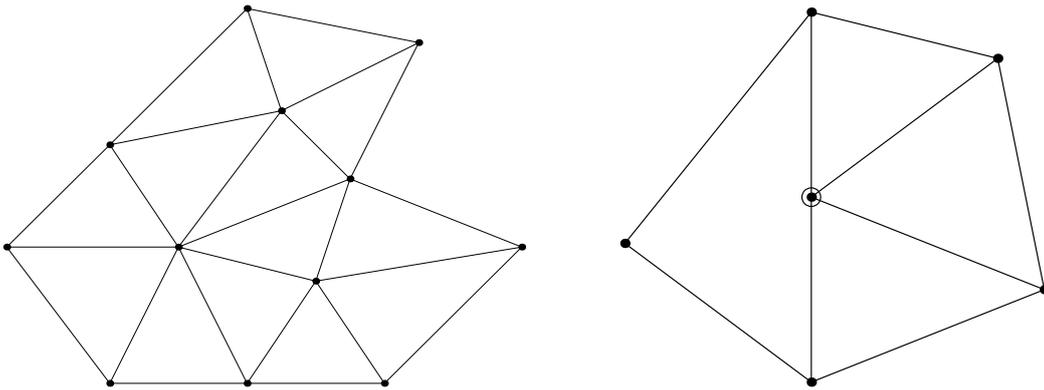


Рис. 20. Пример правильной (слева) и неправильной (справа) триангуляции области.

2) с каждым элементом  $e$  триангуляции связывается конечный элемент  $(e, P_e, \Sigma_e)$ , и таким образом получается семейство конечных элементов  $\mathcal{E}_h = \{(e, P_e, \Sigma_e), e \in \mathcal{T}_h\}$ ; предполагается, что  $P_e \subset C^k(e)$ ,  $k = 1, 2$ , для любого элемента  $e$ ;

3) определяется пространство конечных элементов:

$$V_h = \{v_h \in C^{k-1}(\Omega_h) : v_h|_e \in P_e \quad \forall e \in \mathcal{T}_h\}, \quad k = 1, 2,$$

где  $\Omega_h = \bigcup_e e$ ; конструируется базис этого пространства<sup>1)</sup>.

В силу теоремы 8, с. 44, имеем вложение  $V_h \subset H^k(\Omega_h)$ ,  $k = 1, 2$ .

Прокомментируем подробнее первый шаг. Выбор конфигурации элементов зависит, обычно, от формы области  $\Omega$ . Так, многоугольник

<sup>1)</sup>Обозначения типа  $\bigcup_e$ ,  $\sum_e$ ,  $\max_e$ , ..., понимаются как  $\bigcup_{e \in \mathcal{T}_h}$ ,  $\sum_{e \in \mathcal{T}_h}$ ,  $\max_{e \in \mathcal{T}_h}$  и т.д..

можно покрыть без налеганий множеством треугольников с прямолинейными сторонами, область, составленная из прямоугольников, легко разбивается на элементы прямоугольной формы (а также на треугольники). Произвольная область с кусочно гладкой границей уже не может быть покрыта без налеганий отмеченными выше элементами. В этом случае используются как треугольники с прямолинейными сторонами, так и треугольники с криволинейными сторонами (обычно это приграничные элементы). Во всех указанных случаях имеем

$$\bar{\Omega} = \bigcup_e e.$$

В случае областей с кусочно гладкой границей будем рассматривать также триангуляции, образованные «изопараметрическими» элементами. С одной стороны, это упрощает конструкцию пространства конечных элементов  $V_h$  (в сравнении с криволинейными элементами), с другой — позволяет использовать произвольные четырехугольники с прямолинейными сторонами при разбиении области. Поскольку в этом случае область не точно разбивается на элементы из  $\mathcal{T}_h$  (граница приближается, как правило, кусочно полиномиальной кривой), то

$$\Omega_h = \bigcup_e e \neq \bar{\Omega}.$$

Обычно, область разбивается на элементы одного типа (например, треугольные или четырехугольные), хотя это и не обязательно. Независимо от формы конечных элементов  $e$  семейство  $\mathcal{T}_h$  будем называть триангуляцией области  $\Omega$ .

Значок  $h$  (в обозначениях  $\mathcal{T}_h$ ,  $V_h$ , и т.д.) характеризует максимальный размер элементов в триангуляции и определяется равенством

$$h = \max_e \text{diam}(e).$$

**ЗАМЕЧАНИЕ 1.** Построение триангуляций произвольных двумерных областей является самостоятельной и нетривиальной задачей. Известные алгоритмы и программы триангуляции плоской области различаются как по способу задания области  $\Omega$ , так и по типу элементов, на которые производится разбиение. Например, произвольная область  $\Omega$  может быть задана своей характеристической функцией или как композиция простых (канонических) областей; односвязная многоугольная область  $\Omega$  — перечислением координат вершин; произвольная односвязная область может быть задана параметрическим описанием ее границы и т. д. Существующие алгоритмы позволяют строить разбиение области на треугольные элементы (с прямолинейными сторонами или «изопараметрические»). Известны также алгоритмы разбиения области на аналогичные четырехугольные элементы. Поскольку произвольную область достаточно трудно покрыть четырехугольниками, обеспечив при этом нужное качество триангуляции, то обычно в такую триангуляцию включается и небольшое число треугольных элементов.

Перейдем к описанию конструкции пространств  $V_h$ .

Пусть  $\mathcal{T}_h$  — триангуляция области  $\Omega$  и фиксировано целое число  $m \geq 1$ , определяющее порядок полиномов, используемых при построении конечных элементов. В одномерном случае  $\mathcal{T}_h$  — совокупность отрезков, в двумерном — совокупность треугольников и (или) четырехугольников как с прямолинейными, так и с криволинейными сторонами. С каждым элементом  $e \in \mathcal{T}_h$  свяжем конечный элемент  $(e, P_e, \Sigma_e)$ , которые в совокупности образуют семейство конечных элементов

$$\mathcal{E}_h = \{(e, P_e, \Sigma_e), e \in \mathcal{T}_h\}.$$

ЗАМЕЧАНИЕ 2. Процедура триангуляции области тесно связана с образованием семейства  $\mathcal{E}_h$  конечных элементов. Можно считать, что это происходит, практически, одновременно. Хорошим подтверждением сказанному служит семейство изопараметрических лагранжевых конечных элементов: для определения формы множества  $e$  нужны узлы интерполяции  $a_i^e$ , которые, в свою очередь, однозначно определяют конечный элемент  $(e, P_e, \Sigma_e)$ . Поэтому часто употребляются высказывания типа: триангуляция области на линейные треугольные конечные элементы (четырёхугольные изопараметрические конечные элементы степени  $m$ , прямоугольные бикубические эрмитовы конечные элементы и т.д.). Мы также будем употреблять такие высказывания, понимая под этим то, что сначала область триангулируется, а потом строится семейство конечных элементов  $\mathcal{E}_h$ .

В дальнейшем предполагаются выполненными следующие условия.

1) Все элементы триангуляции являются либо лагранжевыми класса  $C^0$ , либо эрмитовыми (класса  $C^0$  или  $C^1$ ).

2) Лагранжевы конечные элементы могут быть ассоциированы с различными базисными элементами. Допускается использование как аффинно-эквивалентных, так и криволинейных или изопараметрических конечных элементов (в зависимости от формы границы множества  $e$ ). Для всех базисных элементов  $\hat{P} \supseteq P_m$ , на каждой стороне базисного элемента задан  $m + 1$  узел интерполяции, включая вершины элемента. Для каждой стороны способ распределения узлов один и тот же. Предполагается, что узлы расположены симметрично относительно середины стороны.

3) Все элементы эрмитова типа ассоциированы с одним базисным конечным элементом и являются аффинно-эквивалентными,  $\hat{P} \supseteq P_m$ . Узловые параметры эрмитова элемента  $(e, P_e, \Sigma_e)$  являются глобальными, т. е. используются элементы  $(e, P_e, \Sigma_e^g)$ , равные  $(e, P_e, \Sigma_e)$  (см. с. 111). Для единообразия обозначений верхний индекс  $g$  в обозначениях базисных функций и узловых параметров элемента  $(e, P_e, \Sigma_e^g)$ , как правило, будем опускать.

Примеры точного разбиения области  $\Omega$  ( $\Omega_h = \bar{\Omega}$ ), соответственно, на аффинно-эквивалентные треугольные лагранжевы конечные элементы ( $m = 2$ ), кубические эрмитовы конечные элементы ( $m = 3$ ), конечные элементы Белла ( $m = 4$ ) и прямоугольные бикубические эрмитовы конечные элементы ( $m = 3$ ) даны далее в §8 (см. рис. 22–25). На рисунке 26 представлен пример неточного разбиения области на лагранжевы конечные элементы различных типов ( $\Omega_h \neq \bar{\Omega}$ ,  $m = 2$ ).

На каждом элементе  $e \in \mathcal{T}_h$  введем нумерацию узлов интерполяции и узловых параметров. Пусть  $\omega^e = \{a_i^e, i = 1, 2, \dots, n_e\}$  есть множество узлов интерполяции на элементе  $e$ , а  $\phi_{ik}^e, k = 0, 1, \dots, m_i^e$ , — множество всех степеней свободы, соответствующих узлу  $a_i^e$ . Пару  $(e, i)$  будем называть локальным номером узла  $a_i^e$ . Например, для произвольных лагранжевых элементов  $m_i^e = 0$  для всех  $e$  и  $i$ ; для элементов Белла имеем  $n_e = 3$ ,  $m_i^e = 5$  также для всех  $e$  и  $i$ . Таким образом, произвольная функция  $p$  из  $P_e$  представляется в виде

$$p(x) = \sum_{i=1}^{n_e} \sum_{k=0}^{m_i^e} p_{ik}^e \varphi_{ik}^e(x), \quad p_{ik}^e = \phi_{ik}^e(p), \quad \dim P_e = n_e \sum_{i=1}^{n_e} (m_i^e + 1).$$

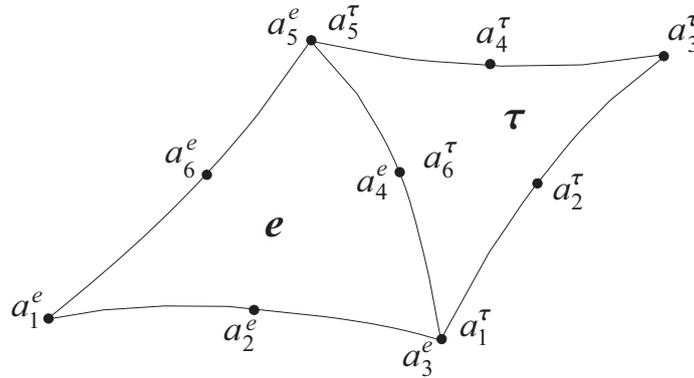


Рис. 21. Соседние квадратичные изопараметрические элементы.

ЗАМЕЧАНИЕ 3. Если  $e$  и  $\tau$  — два соседних элемента из  $\mathcal{T}_h$  с общей стороной  $\gamma = e \cap \tau$  (см. рис. 21, соответствующий случаю изопараметрических квадратичных элементов), то узлы  $a_i^e$  и  $a_k^\tau$  одинакового геометрического положения на  $\gamma$  (такие, как  $a_3^e$  и  $a_1^\tau$ ,  $a_4^e$  и  $a_6^\tau$  и т.д.) в силу определения конечных элементов  $e$  и  $\tau$  будут совпадать.

При определении пространства  $V_h$  будет использована также сквозная нумерация узлов и узловых параметров. Пусть  $\omega_h$  — подмножество всех различных узлов из  $\bigcup_e \omega^e$ ,  $N_h = \text{card } \omega_h$ . Множество  $\omega_h$  будем называть множеством узлов интерполяции в области  $\Omega_h$ . Прономеруем узлы из  $\omega_h$  произвольным образом и пусть  $a_l$  есть  $l$ -тый

узел из  $\omega_h$  в этой нумерации. Число  $l$  будем называть глобальным номером узла  $a_l$ .

Любому локальному номеру  $(e, i)$  узла  $a_i^e$  соответствует точно один глобальный номер  $l$ . Это соответствие мы будем указывать в виде  $l = s(e, i)$ . Обратное, вообще говоря, неверно: одному глобальному номеру  $l$  могут соответствовать несколько локальных номеров  $(e, i)$ . Равенство  $a_l = a_{s(e,i)}$  означает, что узел интерполяции  $a_l$  является  $i$ -тым узлом на элементе  $e$ .

Пусть в узле  $a_l$  задано  $m_l + 1$  узловых параметров. Будем считать, что  $m_l = m_i^e$ , где  $l = s(e, i)$ , и  $m_l$  не зависит от выбора локального номера  $(e, i)$ , соответствующего  $l$  (легко проверить, что это так для всех рассмотренных выше примеров конечных элементов).

В каждом узле  $a_l \in \omega_h$  зададим произвольно  $m_l + 1$  чисел  $v_{lk}$ ,  $k = 0, 1, \dots, m_l$ . По этим числам построим функцию  $v_h$ , определенную на  $\Omega_h$  так, что для любого  $e \in \mathcal{T}_h$  имеет место равенство:

$$v_h(x)|_e = \sum_{i=1}^{n_e} \sum_{k=0}^{m_i^e} v_{s(e,i),k} \varphi_{ik}^e(x). \quad (1)$$

По построению  $v_h|_e \in P_e$  для любого  $e \in \mathcal{T}_h$ . Покажем, что  $v_h \in C(\Omega_h)$ .

Действительно, пусть  $e$  и  $\tau$  — два соседних элемента из  $\mathcal{T}_h$  с общей гранью  $\gamma$ . Вследствие (1) функции  $v_h|_e$  и  $v_h|_\tau$  непрерывны и на  $\gamma$  имеют равные значения узловых параметров. Отсюда следует, что равны и следы этих функций на общей грани (они однозначно определяются узловыми параметрами этой грани, поскольку рассматриваемые элементы принадлежат классу  $C^0$ ), поэтому  $v_h \in C(\Omega_h)$ .

Если все элементы, использованные при определении функции  $v_h$ , относятся к семейству лагранжевых элементов, то, вообще говоря,  $v_h \notin C^1(\Omega_h)^1$ . Использование же эрмитовых элементов класса  $C^1$ , имеющих вторые производные в качестве узловых параметров, гарантирует, что  $v_h \in C^1(\Omega_h)$  (совпадение на  $\gamma$  первых производных функций  $v_h|_e$  и  $v_h|_\tau$  следует из определения элементов класса  $C^1$ ). Таким образом,  $v_h \in C^{k-1}(\Omega_h)$  для соответствующего  $k$ ,  $k = 1, 2$ .

Итак, мы определили способ представления произвольной функции из пространства

$$V_h = \{v_h \in C^{k-1}(\Omega_h) : v_h|_e \in P_e \quad \forall e \in \mathcal{T}_h\}, \quad k = 1, 2,$$

которое будем называть пространством конечных элементов. Если все элементы, использованные при определении  $V_h$ , имеют один и тот же

<sup>1)</sup>В этом легко убедиться, например, в случае аффинно-эквивалентных линейных ( $m = 1$ ) треугольных конечных элементов

тип, то в название пространства включается название использованного элемента (или соответствующего базисного элемента). Например, говорят: пространство лагранжевых прямоугольных изопараметрических конечных элементов степени  $m$ , пространство конечных элементов Белла и т. д.

В силу теоремы 8, с. 44, имеем  $V_h \subset H^k(\Omega_h)$  с соответствующим  $k = 1, 2$ .

Укажем теперь конструкцию базиса в  $V_h$ , состоящего из функций с малыми носителями. Для этого фиксируем узел с глобальным номером  $i$ ,  $i = 1, 2, \dots, N_h$ , и узловой параметр с номером  $j$ , соответствующий этому узлу,  $j = 0, 1, \dots, m_i$ . По этой паре  $(i, j)$  зададим набор чисел  $v_{lk}$ , полагая

$$v_{lk} = \delta_{li}\delta_{kj}, \quad l = 1, 2, \dots, N_h, \quad k = 0, \dots, m_l.$$

Пусть, далее, функция  $\varphi_{ij}$  определяется на  $\Omega_h$  так, что

$$\varphi_{ij}(x)|_e = \sum_{i=1}^{n_e} \sum_{k=0}^{m_i^e} v_{s(e,i)k} \varphi_{ik}^e(x) \quad \forall e \in \mathcal{T}_h. \quad (2)$$

Из соотношения (2) сразу вытекает, что  $\varphi_{ij}(x) \in V_h$  и, кроме того,

$$\varphi_{ij}(x)|_e = \begin{cases} \varphi_{kj}^e(x), & \text{если найдется такое } k, \text{ что } i = s(e, k), \\ 0, & \text{в противном случае,} \end{cases} \quad (3)$$

т. е. функция  $\varphi_{ij}(x)$  является непрерывной «склейкой» соответствующих базисных функций элементов и функции, тождественно равной нулю.

ЗАМЕЧАНИЕ 4. Для иллюстрации процесса построения базисных функций пространства  $V_h$  полезно обратиться к простейшему примеру, разобранному на с. 16.

Убедимся в том, что система  $\{\varphi_{ij}, i = 1, \dots, N_h, j = 0, \dots, m_i\}$  образует базис в  $V_h$ . Проверим сначала, что она линейно независима. Пусть

$$v_h(x) = \sum_{i=1}^{N_h} \sum_{j=0}^{m_i} c_{ij} \varphi_{ij}(x) = 0 \quad \text{в } \Omega_h.$$

Фиксируем глобальный номер  $l = 1, 2, \dots, N_h$  и пусть  $a_l \in e$  для некоторого  $e \in \mathcal{T}_h$  и  $l = s(e, k)$ . Возьмем сужение функции  $v_h$  на  $e$ . Тогда

$$v_h(x)|_e = \sum_{i=1}^{n_e} \sum_{j=0}^{m_i^e} c_{s(e,i)j} \varphi_{ij}^e(x) = 0.$$

Так как функции  $\varphi_{ij}^e(x)$  линейно независимы, то отсюда следует, что все числа  $c_{s(e,i),j} = 0$ , в частности,  $c_{lj} = 0$  для  $j = 0, \dots, m_i$ . В силу произвольности номера  $l$  отсюда вытекает, что все  $c_{ij} = 0$ .

Покажем теперь, что любая функция  $v_h \in V_h$  раскладывается по системе  $\{\varphi_{lk}\}$ . Для этого заметим, что для любого  $e \in \mathcal{T}_h$  из формул (1), (3) следует:

$$v_h(x)|_e = \sum_{i=1}^{n_e} \sum_{k=0}^{m_i} v_{s(e,i),k} \varphi_{ik}^e(x) = \sum_{l=1}^{N_h} \sum_{k=0}^{m_l} v_{lk} \varphi_{lk}(x)|_e.$$

Отсюда получаем, что

$$v_h(x) = \sum_{l=1}^{N_h} \sum_{k=0}^{m_l} v_{lk} \varphi_{lk}(x).$$

Таким образом, система  $\{\varphi_{lk}, l = 1, 2, \dots, N_h, k = 0, 1, \dots, m_l\}$  действительно образует базис в  $V_h$ . Будем называть его базисом Лагранжа пространства  $V_h$ .

По построению базисные функции равны нулю на большей части области  $\Omega_h$ . В самом деле, функция  $\varphi_{lk}$  отлична от нуля только на тех элементах, которые содержат узел  $a_l$ . Объединение всех таких элементов будем называть звездой элементов и обозначать через  $Z_l$ . Таким образом,

$$Z_l = \text{supp } \varphi_{lk} = \bigcup \{e \in \mathcal{T}_h : a_l \in e\}.$$

Например, если узел  $a_l$  является внутренним узлом некоторого элемента  $e$ , то  $Z_l = e$ . Если узел  $a_l$  принадлежит общей стороне двух элементов, но не является вершиной какого-либо элемента, то в  $Z_l$  входят все точки этих двух элементов. В любом случае функции  $\varphi_{lk}$  имеют «малый» носитель. Действительно,  $|Z_l| = \text{mes } Z_l = O(h^n)$ , и мера носителя  $\varphi_{lk}$  при малых  $h$  значительно меньше меры области  $\Omega_h$ .

**Определение 1.** Пусть  $u(x) \in C^{k-1}(\Omega_h)$ . Функция  $u_I \in V_h$  называется  $V_h$ -интерполянт функции  $u$ , если для любого  $e \in \mathcal{T}_h$

$$u_I(x)|_e = \sum_{i=1}^{n_e} \sum_{k=0}^{m_i} \phi_{ik}^e(u) \varphi_{ik}^e(x).$$

Оператор, ставящий в соответствие функции  $u \in C^{k-1}(\Omega_h)$  ее  $V_h$ -интерполянт, будем называть оператором  $V_h$ -интерполяции и обозначать через  $\pi_h$ .

Справедливость следующей формулы вытекает из определения базисных функций Лагранжа:  $(\pi_h u)|_e = \pi_e(u|_e)$ .

## § 8. Примеры пространств конечных элементов

В предыдущем параграфе была описана общая схема построения пространств конечных элементов. Рассмотрим примеры пространств конечных элементов, конкретизируя в общей схеме форму области  $\Omega$  и семейство конечных элементов.

### 1. Пространство лагранжевых треугольных элементов степени $m$ .

Пусть  $\Omega$  — многоугольник в  $R^2$ ,  $\mathcal{T}_h$  — правильное разбиение  $\Omega$  на треугольные элементы с прямолинейными сторонами,  $\{(e, P_e, \Sigma_e), e \in \mathcal{T}_h\}$  — семейство аффинно-эквивалентных лагранжевых треугольных конечных элементов степени  $m$ , ассоциированных с

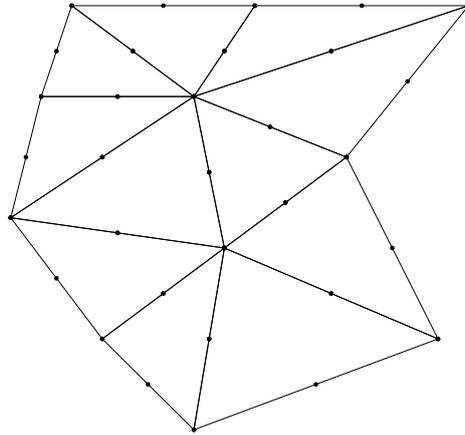


Рис. 22. Триангуляция области на аффинно-эквивалентные квадратичные лагранжевы элементы.

одним и тем же базисным элементом  $(\hat{e}, \hat{P}, \hat{\Sigma})$  (см. пример 3, с. 90). На рис. 22 представлен пример триангуляции при  $m = 2$ . Для каждого элемента  $e \in \mathcal{T}_h$  имеем в этом случае

$$P_e = P_m, \quad \Sigma_e = \{p(a_i^e), i = 1, 2, \dots, n_e\}, \quad n_e = (m + 1)(m + 2)/2,$$

и, кроме того,

$$\Omega_h = \bigcup_e e = \bar{\Omega}, \quad m_i^e = 0, \quad m_l = 0 \quad \forall i, l, e,$$

$$V_h = \{v_h \in C(\bar{\Omega}) : v_h|_e \in P_m \quad \forall e \in \mathcal{T}_h\} \subset H^1(\Omega).$$

Базисные функции Лагранжа  $\{\varphi_i(x)\}$  в пространстве  $V_h$  определяются из условий

$$\varphi_i \in V_h, \quad \varphi_i(a_l) = \delta_{il}, \quad i, l = 1, 2, \dots, N_h,$$

а произвольная функция  $v_h \in V_h$  может быть представлена в виде  $v_h(x) = \sum_{i=1}^{N_h} v_h(a_i) \varphi_i(x)$ . При  $m = 1$  пространство  $V_h$  будем называть пространством линейных конечных элементов.

УПРАЖНЕНИЕ 5. Описать пространство конечных элементов, основанное на использовании прямоугольного лагранжева элемента (см. с. 89) для двумерной области, составленной из прямоугольников.

## 2. Пространство эрмитовых кубических элементов.

Пусть  $\Omega$  — многоугольник в  $R^2$ ,  $\mathcal{T}_h$  — правильное разбиение  $\Omega$  на треугольные элементы с прямолинейными сторонами,  $\{(e, P_e, \Sigma_e), e \in \mathcal{T}_h\}$  — семейство аффинно-эквивалентных эрмитовых кубических треугольных конечных элементов. На рис. 23 представлен пример такой триангуляции. Напомним, что на каждом эле-

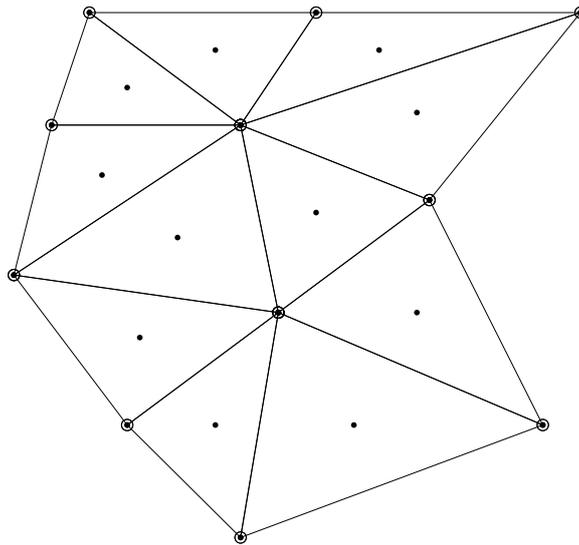


Рис. 23. Триангуляция области на аффинно-эквивалентные эрмитовы кубические элементы. Окружность указывает на задание в узле всех производных до первого порядка включительно.

менте  $e \in \mathcal{T}_h$  имеется четыре узла интерполяции (вершины элемента и его центр масс),  $\omega_e = \{a_i^e, i = 1, \dots, 4\}$ , произвольная функция из  $P_e = P_3$  может быть представлена в виде (см. определение элемента  $(e, P_e, \Sigma_e^g)$  в задаче 1 с. 113):

$$p = \sum_{i=1}^4 p_{i0} \varphi_{i0}^{eg}(x) + \sum_{i=1}^3 \sum_{k=1}^2 p_{ik} \varphi_{ik}^{eg}(x),$$

где  $p_{i0} = p(a_i^e)$ ,  $p_{ik} = \partial p(a_i^e) / \partial x_k$ .

Для каждого элемента  $e \in \mathcal{T}_h$  имеем, таким образом,  $P_e = P_3$ ,  $n_e = 4$ ,

$$\Sigma_e = \{p(a_i^e), i = 1, \dots, n_e, \frac{\partial p}{\partial x_k}(a_i^e), i = 1, 2, 3, k = 1, 2\},$$

кроме того,

$$\Omega_h = \bigcup_e e = \bar{\Omega}, \quad m_i^e = 2, \quad i = 1, 2, 3, \quad m_4^e = 0 \quad \forall e,$$

$$V_h = \{v_h \in C(\bar{\Omega}) : v_h|_e \in P_3 \quad \forall e \in \mathcal{T}_h\} \subset H^1(\Omega).$$

Пусть  $\{a_l\}_{l=1}^{N_1}$  — множество вершин,  $\{a_l\}_{l=N_1+1}^{N_h}$  — множество центров масс элементов из  $\mathcal{T}_h$ ,  $\omega_h = \{a_l\}_{l=1}^{N_h}$  — множество всех узлов интерполяции в области  $\Omega$ . Произвольная функция  $v_h \in V_h$  может быть представлена в виде

$$v_h(x) = \sum_{i=1}^{N_h} v_i \varphi_{i0}^g(x) + \sum_{i=1}^{N_1} \sum_{k=1}^2 v_{ik} \varphi_{ik}^g(x), \quad (1)$$

где  $v_i = v_h(a_i)$ ,  $v_{ik} = \partial v_h(a_i) / \partial x_k$ . Отметим, что первые производные функций из  $V_h$  непрерывны в узлах интерполяции  $\omega_h$ .

### 3. Пространство конечных элементов Белла.

Пусть  $\Omega$  — многоугольник в  $R^2$ ,  $\mathcal{T}_h$  — правильное разбиение  $\Omega$  на треугольные элементы с прямолинейными сторонами,  $\{(e, P_e, \Sigma_e), e \in \mathcal{T}_h\}$  — семейство аффинно-эквивалентных конечных элементов Белла. На рис. 24 представлен пример триангуляции. Для каждого элемента  $e \in \mathcal{T}_h$  имеем  $P_e \supset P_4$ ,  $n_e = 18$ ,

$$\Sigma_e = \Sigma_e^g = \{p(a_i^e), \frac{\partial p}{\partial x_k}(a_i^e), \frac{\partial^2 p}{\partial x_k \partial x_l}(a_i^e), \quad i = 1, 2, 3, \quad l, k = 1, 2\}.$$

Кроме того,  $\Omega_h = \bigcup_e e = \bar{\Omega}$ ,  $m_i^e = 5$ ,  $m_l = 5$ ,  $i = 1, 2, 3$ ,  $\forall l, e$ . Пусть

$\omega_h = \{a_l\}_{l=1}^{N_h}$  — множество всех узлов интерполяции в области  $\Omega$ . Произвольная функция  $v_h \in V_h \subset H^2(\Omega)$  может быть представлена в виде

$$v_h(x) = \sum_{i=1}^{N_h} \sum_{k=0}^5 v_{ik} \varphi_{ik}(x),$$

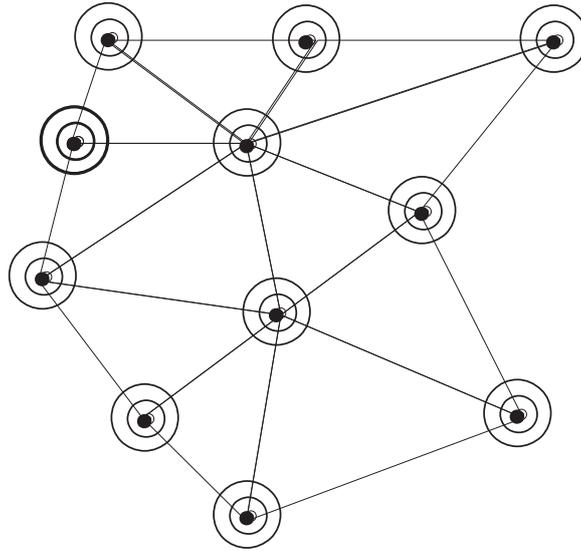


Рис. 24. Двойная окружность указывает на задание в узле всех производных до второго порядка включительно.

где  $v_{i0} = v_h(a_i)$ ,  $v_{ik} = \partial v_h(a_i)/\partial x_k$  при  $k = 1, 2$ ,

$$v_{i3} = \frac{\partial^2 v_h}{\partial x_1^2}(a_i), \quad v_{i4} = \frac{\partial^2 v_h}{\partial x_1 \partial x_2}(a_i), \quad v_{i5} = \frac{\partial^2 v_h}{\partial x_2^2}(a_i).$$

Отметим, что  $V_h \in C^1(\bar{\Omega})$ , и все производные второго порядка функций из  $V_h$  непрерывны в узлах интерполяции  $\omega_h$ .

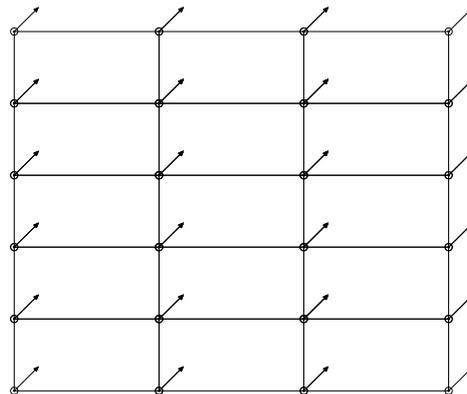


Рис. 25. Триангуляция области на аффинно-эквивалентные бикубические эрмитовы элементы. Окружность указывает на задание первых производных в узле, диагональная стрелка — смешанных вторых производных.

#### 4. Пространство прямоугольных бикубических конечных элементов.

Пусть  $\Omega$  — область в  $R^2$ , составленная из прямоугольников со сторонами, параллельными осям координат,  $\mathcal{T}_h$  — правильное разбиение  $\Omega$  на прямоугольные элементы,  $\mathcal{E}_h = \{(e, P_e, \Sigma_e), e \in \mathcal{T}_h\}$  есть семейство аффинно-эквивалентных бикубических конечных элементов. На рис. 25 представлен пример триангуляции области. Для всех элементов  $e \in \mathcal{T}_h$  имеем  $P_e = Q_3$ ,  $n_e = 16$ ,  $\Sigma_e = \Sigma_e^g = \{p(a_i^e), \frac{\partial p}{\partial x_k}(a_i^e), \frac{\partial^2 p}{\partial x_1 \partial x_2}(a_i^e), i = 1, \dots, 4, k = 1, 2\}$ ,  $\Omega_h = \bigcup_e \bar{\Omega}_e$ ,  $m_i^e = 3$ ,  $m_l = 3, i = 1, \dots, 4, \forall l, e$ . Произвольная функция  $v_h \in V_h \subset H^2(\Omega)$  может быть представлена в виде

$$v_h(x) = \sum_{i=1}^N \sum_{k=0}^3 v_{ik} \varphi_{ik}(x),$$

где  $v_{i0} = v_h(a_i)$ ,  $v_{ik} = \partial v_h(a_i) / \partial x_k$  при  $k = 1, 2$ ,  $v_{i3} = \frac{\partial^2 v_h}{\partial x_1 \partial x_2}(a_i)$ .

Отметим, что  $V_h \in C^1(\bar{\Omega})$ , а смешанные производные второго порядка функций из  $V_h$  непрерывны в узлах интерполяции  $\omega_h$ .

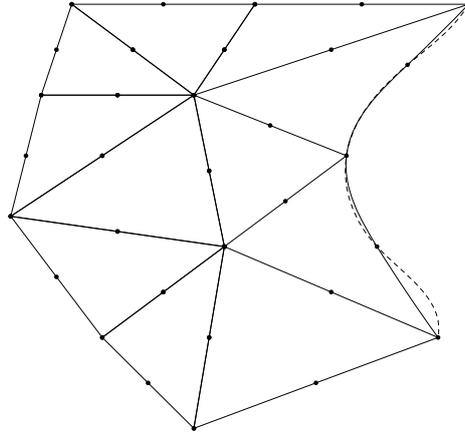


Рис. 26. Триангуляция области на элементы различных типов. Штриховая линия соответствует криволинейной части границы области.

**УПРАЖНЕНИЕ 6.** Следуя приведенным выше примерам, дать определение пространства конечных элементов в случае, когда область с частично криволинейной границей разбивается на треугольные элементы различных типов: треугольные изопараметрические квадратичные элементы и аффинно-эквивалентные элементы второй степени (см. рис. 26).

## § 9. Аппроксимация подпространств $H^k(\Omega)$ , $k = 1, 2$

Рассмотрим важный для МКЭ вопрос о том, как учесть главные граничные условия при построении пространства конечных элемен-

ТОВ.

### 1. Аппроксимация подпространств $H^1(\Omega)$ .

Пусть

$$V^0 = \{v \in H^1(\Omega) : v(x) = 0, x \in \Gamma_0\},$$

где  $\Gamma_0$  — часть границы  $\Gamma$  области  $\Omega$  ненулевой меры и конечной связности. Если  $\Gamma_0 = \Gamma$ , то по определению  $V^0 = H_0^1(\Omega)$ .

Опишем способы построения конечно-элементных аппроксимаций пространства  $V^0$  по некоторым конкретным аппроксимациям  $V_h$  пространства  $H^1(\Omega)$ .

**1. Лагранжевы элементы.** Рассмотрим сначала наиболее простую ситуацию. Пусть  $\Omega$  — двумерная многоугольная область,  $V_h$  — пространство лагранжевых конечных элементов (построенное на основе аффинно-эквивалентных или (и) изопараметрических элементов) такое, что  $V_h \subset H^1(\Omega)$ . Будем предполагать, что граничные точки  $\Gamma_0$  являются вершинами каких-либо конечных элементов, т. е.  $\bar{\Gamma}_0$  является объединением некоторого количества сторон конечных элементов.

Определим подпространство  $V_h^0$ :

$$V_h^0 = \{v_h \in V_h : v_h(x) = 0, x \in \Gamma_0\}. \quad (1)$$

Очевидно, что  $V_h^0 \subset V^0$ . Пусть  $\omega_h$  — множество узлов интерполяции в области  $\bar{\Omega}$ ,  $\gamma_h = \omega_h \cap \bar{\Gamma}_0$  — узлы интерполяции, лежащие на  $\bar{\Gamma}_0$ . Нетрудно видеть, что пространство  $V_h^0$  может быть также определено следующим образом:

$$V_h^0 = \{v_h \in V_h : v_h(a_i) = 0, a_i \in \gamma_h\}, \quad (2)$$

поскольку на каждой стороне конечного элемента  $\partial e \cap \Gamma \subset \bar{\Gamma}_0$  условия  $v_h(x) = 0$  на  $\partial e \cap \Gamma$  и  $v_h(a_i) = 0$ ,  $a_i \in \gamma_h \cap \omega_h^e$ , эквивалентны (см. лемму 1, с. 85).

Пусть теперь  $\{\varphi_i(x)\}_{i=1}^N$  — базис Лагранжа в  $V_h$  и  $I_h = \{i : a_i \in \omega_h \setminus \gamma_h\}$ . Из определения (2) следует, что произвольная функция  $v_h$  из  $V_h^0$  допускает представление

$$v_h(x) = \sum_{i \in I_h} v_i \varphi_i(x), \quad v_i = v_h(a_i), \quad (3)$$

т. е. система функций  $\{\varphi_i(x), i \in I_h\}$  образует базис Лагранжа в  $V_h^0$ .

Все сказанное выше остается справедливым также в случае, если  $\Omega$  — область с кусочно гладкой границей, и для точной триангуляции области используются лагранжевы криволинейные приграничные конечные элементы. В этом случае, как и выше, имеем  $V_h^0 \subset V^0$ .

Если используются изопараметрические приграничные конечные элементы, то область, вообще говоря, триангулируется не точно, и  $\Omega_h \neq \Omega$ ,  $V_h \subset H^1(\Omega_h)$ . Предполагая опять, что граничные точки  $\Gamma_0$  и угловые точки  $\Gamma$  (если они имеются) являются вершинами каких-либо конечных элементов, обозначим через  $\Gamma_0^h$  объединение тех граней элементов из  $\mathcal{T}_h$ , которые имеют с  $\Gamma_0$  не менее двух общих точек. Непрерывная кривая  $\Gamma_0^h$  является в данном случае естественной аппроксимацией  $\Gamma_0$  (см. рис. 27). Пространство

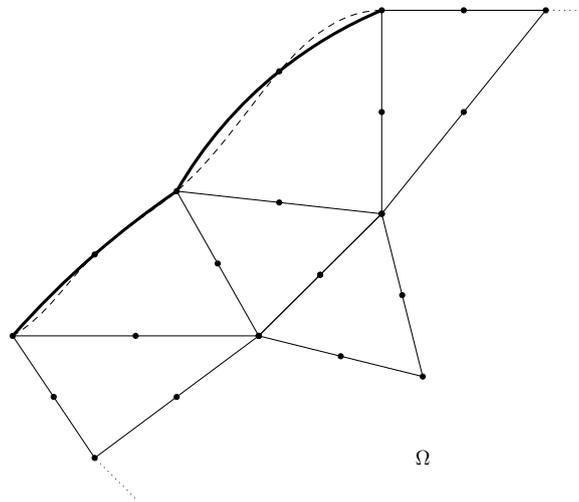


Рис. 27. Использование лагранжевых треугольных изопараметрических элементов ( $m \geq 2$ ) вблизи криволинейной части  $\Gamma_0$  границы (штриховая линия). Жирной линией выделен участок  $\Gamma_0^h$ .

$$V_h^0 = \{v_h \in V_h : v_h(x) = 0, x \in \Gamma_0^h\}$$

назовем пространством конечных элементов, аппроксимирующим пространство  $V^0$ . Имеем

$$V_h^0 \subset \{v \in H^1(\Omega_h) : v(x) = 0, x \in \Gamma_0^h\}.$$

Как и ранее,  $V_h^0$  допускает эквивалентное определение (2), справедлива также формула (3).

**2. Эрмитовы элементы.** Определение аппроксимации пространства  $V^0$  при помощи соотношения (1) сохраняется и при использовании эрмитовых элементов. Рассмотрим для примера пространство эрмитовых кубических элементов. По-прежнему, в узлах интерполяции  $a_i$ , принадлежащих  $\gamma_h$ , мы должны потребовать выполнения

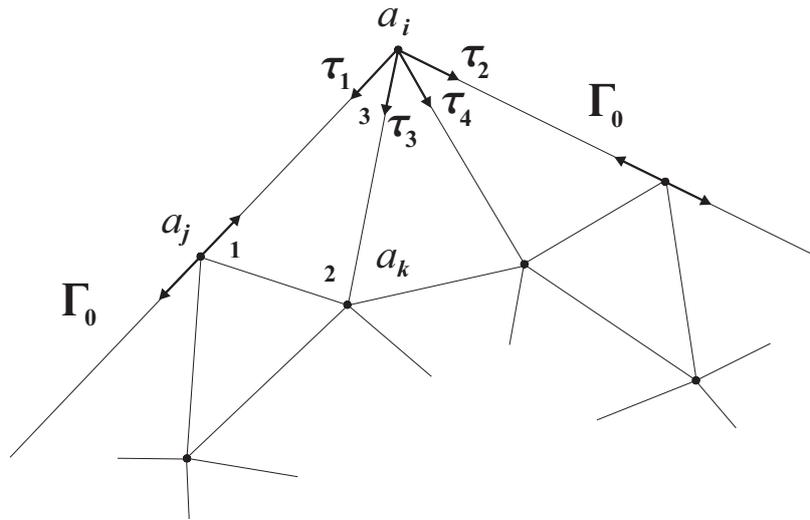


Рис. 28. Выделены направления, производные по которым полагаются равными нулю с целью учесть краевое условие на  $\Gamma_0$ .

условия  $v_h(a_i) = 0$ . Далее, очевидно также, что в узлах  $\gamma_h$  должны быть равны нулю и производные вдоль направлений сторон, лежащих на  $\Gamma_0$ . Множество таких направлений в узле  $a$  обозначим через  $\tau(a)$  (см. рис. 28).

Таким образом, пространство  $V_h^0$  может быть определено также соотношениями:

$$V_h^0 = \{v_h \in V_h : v_h(a_i) = 0, \partial_\tau v_h(a_i) = 0, \tau \in \tau(a_i), a_i \in \gamma_h\}. \quad (4)$$

Опишем теперь пространство  $V_h^0$  в терминах условий на узловые параметры приграничных элементов. Из условия  $\partial_\tau v_h(a) = 0$ ,  $a \in \gamma_h$  вытекает, что <sup>1)</sup>

$$\frac{\partial v_h(a)}{\partial x_2} = -\operatorname{tg} \alpha(a) \frac{\partial v_h(a)}{\partial x_1}, \quad (5)$$

где  $\alpha(a)$  — угол наклона соответствующей стороны области  $\Omega$  к оси  $x_2$ . Если  $a$  — угловая точка области  $\Omega$ , то условие (5) выполняется для двух непараллельных направлений, следовательно,

$$\frac{\partial v_h(a)}{\partial x_2} = \frac{\partial v_h(a)}{\partial x_1} = 0.$$

Таким образом, произвольная функция  $v_h \in V_h^0$  допускает представ-

<sup>1)</sup>Для упрощения записей считаем, что  $\alpha(a) \neq \pi/2$  для любой точки  $a \in \gamma_h$

ление (см. также формулу (1), с. 140):

$$v_h(x) = \sum_{a_i \in \gamma_h / \gamma_h^*} v_{i1}(\varphi_{i1}^g(x) - \operatorname{tg} \alpha(a_i) \varphi_{i2}^g(x)) + \\ + \sum_{a_i \in \omega_h / \gamma_h} v_i \varphi_{i0}^g(x) + \sum_{a_i \in \omega_h^{(1)} / \gamma_h} \sum_{k=1}^2 v_{ik} \varphi_{ik}^g(x).$$

Здесь  $\omega_h^{(1)}$  — множество вершин конечных элементов,  $\gamma_h^*$  — множество угловых точек области  $\Omega$ .

УПРАЖНЕНИЕ 7. Рассмотреть аналогичные вопросы в случаях, когда  $V_h$  — пространство конечных элементов Белла или пространство прямоугольных бикубических конечных элементов.

## 2. Аппроксимация подпространств $H^2(\Omega)$ .

Пусть

$$V^0 = \{v \in H^2(\Omega) : v(x) = 0, \partial_\nu v(x) = 0, x \in \Gamma_0\},$$

где  $\Gamma_0$  — часть границы  $\Gamma$  области  $\Omega$ , имеющая ненулевую меру и конечную связность,  $\nu$  — единичный вектор внешней нормали к  $\Gamma_0$ . Если  $\Gamma_0 = \Gamma$ , то по определению  $V^0 = H_0^2(\Omega)$ .

Рассмотрим сначала случай, когда область  $\Omega$  составлена из прямоугольников со сторонами, параллельными координатным осям, а пространство  $V_h$  — пространство прямоугольных бикубических конечных элементов. Как и ранее, будем считать, что  $\Gamma_0$  в точности совпадает с объединением некоторого множества сторон конечных элементов, входящих в триангуляцию. Положим

$$V_h^0 = \{v_h \in V_h : v_h(x) = 0, \partial_\nu v_h(x) = 0, x \in \Gamma_0\}. \quad (6)$$

Очевидно,  $V_h^0$  — подпространство  $V_h$ . Пусть  $e'$  — сторона некоторого конечного элемента, лежащая на  $\Gamma_0$ ,  $\tau$  — направление вдоль  $e'$  и  $v_h \in V_h^0$ . Тогда из определения  $V_h^0$  следует, что на  $e'$  и, в частности, в узлах интерполяции на  $e'$ , имеют место равенства

$$v_h(x) = \partial_\tau v_h(x) = \partial_\nu v_h(x) = \partial_{\tau\tau} v_h(x) = \partial_{\nu\tau} v_h(x) = 0. \quad (7)$$

Пусть  $\gamma_h$  — множество узлов интерполяции на  $\bar{\Gamma}_0$ . Из равенств (7) следует, что все узловые параметры, связанные с узлами  $\gamma_h$ , должны быть равны нулю. Таким образом, приходим к следующему эквивалентному определению:

$$V_h^0 = \{v_h \in V_h : v_h(x) = \frac{\partial v_h(x)}{\partial x_1} = \frac{\partial v_h(x)}{\partial x_2} = \frac{\partial^2 v_h(x)}{\partial x_1 \partial x_2} = 0, x \in \gamma_h\}.$$

В случае, когда  $\Omega$  — многоугольник в  $R^2$ , а пространство  $V_h$  — пространство конечных элементов Белла, определение (6) и свойства (7) также сохраняют силу. Как и в случае подпространств  $H^1$ , в вершинах многоугольника  $\Omega$  (принадлежащих  $\gamma_h$ ) необходимо потребовать выполнения дополнительного условия, а именно, в указанных узлах все узловые параметры должны быть равны нулю. В остальных узлах  $\gamma_h$  единственной неопределенной степенью свободы является  $\partial_{\nu\nu}v_h(x)$ .

## § 10. Оценки погрешности интерполяции

**1. Оценки на базисном конечном элементе.** Пусть  $(\hat{e}, \hat{P}, \hat{\Sigma})$  — базисный конечный элемент в  $R^n$  и

$$\hat{\omega} = \{\hat{a}_i\}, \quad \hat{\phi}_i, \quad \hat{\varphi}_i(\hat{x}), \quad \hat{x} \in \hat{e}, \quad \hat{\pi}$$

есть его узлы интерполяции, степени свободы, базисные функции и оператор интерполяции,  $\dim \hat{P} = N$ .

Пусть максимальная степень производных, использованных в определении  $\hat{\Sigma}$ , равна  $r \geq 0$ , функционалы  $\hat{\phi}_i$  непрерывны в  $C^r(\hat{e})$ , т. е.

$$|\hat{\phi}_i(u)| \leq c \|u\|_{\infty, r, \hat{e}}, \quad \|u\|_{\infty, r, \hat{e}} = \max_{\hat{x} \in \hat{e}} \sum_{|\alpha| \leq r} |D^\alpha u(\hat{x})|.$$

Очевидно, что это предположение справедливо для любых лагранжевых или эрмитовых конечных элементов. Будем также считать, что базисные функции Лагранжа  $\hat{\varphi}_i(\hat{x})$  являются достаточно гладкими функциями (во всех примерах, которые мы рассматривали, они были полиномами).

Пусть  $\hat{u} \in H_p^{s+1}(\hat{e})$  при целом  $s$  таком, что  $(s+1-r)p > n$ . По теореме вложения Соболева (см. теорему 3, с. 40) имеем включение  $\hat{u} \in C^r(\hat{e})$ . Отсюда следует, что по функции  $\hat{u}$  можно построить ее  $\hat{P}$ -интерполянт  $\hat{u}_I = \hat{\pi}\hat{u}$ . Напомним, что функция  $\hat{u}_I \in \hat{P}$  называется  $\hat{P}$ -интерполянтом функции  $\hat{u}$ , если  $\hat{\phi}_i(\hat{u}_I) = \hat{\phi}_i(\hat{u})$ ,  $i = 1, 2, \dots, N$ .

**Теорема 1.** Пусть  $\hat{u}, \hat{u}_I \in H_p^{s+1}(\hat{e})$ ,  $(s+1-r)p > n$ . Тогда для любого неотрицательного целого  $m$  такого, что  $m \leq s+1$ , справедлива оценка

$$|\hat{u} - \hat{u}_I|_{m, p, \hat{e}} \leq c_1 |\hat{u}|_{s+1, p, \hat{e}}, \quad (1)$$

если пространство  $\hat{P}$  содержит  $P_s$ , и

$$[\hat{u} - \hat{u}_I]_{m, p, \hat{e}} \leq c_2 [\hat{u}]_{s+1, p, \hat{e}}, \quad (2)$$

если  $\hat{P}$  содержит  $Q_s$ . Здесь постоянные  $c_1, c_2$  не зависят от функции  $\hat{u}$ .

ДОКАЗАТЕЛЬСТВО. Обоснуем оценку (1). Пусть  $q$  — показатель, сопряженный к  $p$ , т. е.  $1/p + 1/q = 1$ ,  $\hat{w} \in L_q(\hat{e})$  — произвольная фиксированная функция,  $|\alpha| = m$ . Докажем, что выражение

$$f(\hat{u}) = \int_{\hat{e}} D^\alpha(\hat{u} - \hat{u}_I)\hat{w} \, d\hat{x}$$

определяет линейный непрерывный функционал на пространстве  $H_p^{s+1}(\hat{e})$ . Линейность функционала непосредственно следует из его определения и линейности оператора интерполяции. Докажем ограниченность  $f$ . Используя неравенство Гельдера, получим

$$|f(\hat{u})| \leq |D^\alpha(\hat{u} - \hat{u}_I)|_{0,p,\hat{e}} |\hat{w}|_{0,q,\hat{e}} \leq (|D^\alpha \hat{u}|_{0,p,\hat{e}} + |D^\alpha \hat{u}_I|_{0,p,\hat{e}}) |\hat{w}|_{0,q,\hat{e}}.$$

Для оценки полунормы  $|D^\alpha \hat{u}_I|_{0,p,\hat{e}}$  используем представление  $\hat{u}_I$  в базисе Лагранжа  $\{\hat{\varphi}_i(\hat{x})\}_{i=1}^N$  пространства  $\hat{P}$ . Имеем:

$$\begin{aligned} |D^\alpha \hat{u}_I|_{0,p,\hat{e}} &= \left| \sum_{i=1}^N \hat{\varphi}_i(\hat{u}) \hat{D}^\alpha \varphi_i \right|_{0,p,\hat{e}} \leq \\ &\leq \max_{i=1,\dots,N} |\hat{\varphi}_i(\hat{u})| \sum_{i=1}^N |D^\alpha \hat{\varphi}_i|_{0,p,\hat{e}} \leq c \max_{i=1,\dots,N} |\hat{\varphi}_i(\hat{u})|. \end{aligned}$$

Из непрерывности вложения  $H_p^{s+1}(\hat{e})$  в  $C^r(\hat{e})$  и предположения о непрерывности функционалов, принадлежащих  $\hat{\Sigma}$ , следует, что

$$\max_{i=1,\dots,N} |\hat{\varphi}_i(\hat{u})| \leq c \|\hat{u}\|_{\infty,r,\hat{e}} \leq c \|\hat{u}\|_{s+1,p,\hat{e}},$$

откуда и вытекает ограниченность функционала  $f$  в пространстве  $H_p^{s+1}$ :

$$|f(\hat{u})| \leq c \|\hat{u}\|_{s+1,p,\hat{e}} |\hat{w}|_{0,q,\hat{e}}. \quad (3)$$

Заметим теперь, что  $f(\hat{p}) = 0$  при  $\hat{p} \in P_s \subset \hat{P}$ , поскольку в этом случае  $\hat{p}_I = \hat{p}$ . Учитывая это, а также оценку (3), мы можем использовать лемму Брамбла — Гильберта (с. 50). В результате получим, что

$$|f(\hat{u})| \leq c |\hat{u}|_{s+1,p,\hat{e}} |\hat{w}|_{0,q,\hat{e}},$$

или в силу (8), с. 35,

$$|D^\alpha(\hat{u} - \hat{u}_I)|_{0,p,\hat{e}} = \sup_{w \in L_q(\hat{e})} |f(\hat{u})| / |\hat{w}|_{0,q,\hat{e}} \leq c |\hat{u}|_{s+1,p,\hat{e}}.$$

Отсюда следует оценка (1), поскольку

$$|\hat{u} - \hat{u}_I|_{m,p,\hat{e}} = \left( \sum_{|\alpha|=m} |D^\alpha(\hat{u} - \hat{u}_I)|_{0,p,\hat{e}}^p \right)^{1/p}.$$

Оценка (2) доказывается аналогично. Функционал  $f$  в этом случае определяется равенством

$$f(\hat{u}) = \int_{\hat{e}} \sum_{i=1}^m \frac{\partial^m(\hat{u} - \hat{u}_I)}{\partial x_i^m} \frac{\partial^m \hat{w}}{\partial x_i^m} d\hat{x},$$

а в дальнейших выкладках полунорма  $|\cdot|$  заменяется на полунорму  $[\cdot]$ , затем используется соответствующая лемма Брамбла — Гильберта (с. 52).  $\square$

**Теорема 2.** Пусть  $\hat{u}, \hat{u}_I \in H_p^{s+1}(\hat{e})$ ,  $(s+1-r)p > n$ , пространство  $\hat{P}$  содержит  $P_s$ . Тогда для любого неотрицательного целого  $m$  такого, что  $(s+1-m)p > n$ , справедлива оценка

$$|\hat{u} - \hat{u}_I|_{\infty,m,\hat{e}} \leq c |\hat{u}|_{s+1,p,\hat{e}}. \quad (4)$$

**ДОКАЗАТЕЛЬСТВО.** Пусть  $x^* \in \hat{e}$ . Для любого фиксированного  $\alpha$ ,  $|\alpha| \leq m$ , функционал  $f(\hat{u}) = D^\alpha(\hat{u} - \hat{u}_I)(x^*)$  является линейным и непрерывным в  $H_p^{s+1}(\hat{e})$ . Действительно, рассуждая аналогично доказательству предыдущей теоремы, можем написать

$$\begin{aligned} |f(\hat{u})| &\leq \max_{\hat{x} \in \hat{e}} |D^\alpha \hat{u}(\hat{x})| + \max_{\hat{x} \in \hat{e}} |D^\alpha \hat{u}_I(\hat{x})| \leq \\ &\leq c \|\hat{u}\|_{s+1,p,\hat{e}} + \max_{i=1,\dots,N} |\hat{\phi}_i(\hat{u})| \max_{\hat{x} \in \hat{e}} \sum_{i=1}^N |D^\alpha \hat{\phi}_i(\hat{x})| \leq c \|\hat{u}\|_{s+1,p,\hat{e}}. \end{aligned}$$

С другой стороны, рассматриваемый функционал обращается в нуль на полиномах из  $P_s \subset \hat{P}$ , так что в силу леммы Брамбла — Гильберта

$$|D^\alpha(\hat{u} - \hat{u}_I)(x^*)| \leq c |\hat{u}|_{s+1,p,\hat{e}}.$$

Из этого неравенства в силу произвольности  $x^* \in \hat{e}$  и  $\alpha$ ,  $|\alpha| \leq m$  следует (4).  $\square$

**2. Оценка погрешности интерполяции для аффинно-эквивалентных элементов.** Пусть  $(\hat{e}, \hat{P}, \hat{\Sigma})$  — базисный элемент в  $R^n$  (либо лагранжев, либо эрмитов) и  $\mathcal{E}_h = \{(e, P_e, \Sigma_e), e \in \mathcal{T}_h\}$

есть семейство аффинно-эквивалентных ему элементов, диаметры которых не превосходят  $h$ . Пусть далее для фиксированного элемента  $e \in \mathcal{T}_h$  отображение  $\hat{e}$  на  $e$  задано формулой  $x_e(\hat{x}) = B_e \hat{x} + b_e$ ,  $u \in H_p^s(e)$ , и функция  $\hat{u}(\hat{x})$  связана с функцией  $u(x)$  соотношением  $\hat{u}(\hat{x}) = u(x_e(\hat{x}))$ . Согласно следствию 3, с. 56,

$$|\hat{u}|_{s,p,\hat{e}} \leq c |\det B_e|^{-1/p} \|B_e\|^s |u|_{s,p,e}, \quad (5)$$

$$|u|_{s,p,e} \leq c |\det B_e|^{1/p} \|B_e^{-1}\|^s |\hat{u}|_{s,p,\hat{e}}, \quad (6)$$

где  $\|B_e\|$  — спектральная норма матрицы  $B_e$ .

**Теорема 3.** Пусть  $u, u_I \in H_p^{s+1}(e)$ ,  $(s+1-r)p > n$  и пространство  $\hat{P}$  содержит  $P_s$ . Тогда для любого  $e \in \mathcal{T}_h$  и неотрицательного  $m$  такого, что  $m \leq s+1$ , справедлива оценка

$$|u - u_I|_{m,p,e} \leq c \|B_e^{-1}\|^m \|B_e\|^{s+1} |u|_{s+1,p,e}, \quad (7)$$

где  $u_I$  есть  $P_e$ -интерполяция функции  $u$ .

**ДОКАЗАТЕЛЬСТВО.** Используем теорему 1 и оценки (5), (6). Тогда

$$\begin{aligned} |u - u_I|_{m,p,e} &\leq c |\det B_e|^{1/p} \|B_e^{-1}\|^m |\hat{u} - \hat{u}_I|_{m,p,\hat{e}} \leq \\ &\leq c |\det B_e|^{1/p} \|B_e^{-1}\|^m |\hat{u}|_{s+1,p,\hat{e}} \leq c \|B_e^{-1}\|^m \|B_e\|^{s+1} |u|_{s+1,p,e}, \end{aligned}$$

откуда, очевидно, следует (7). Здесь мы учли, что

$$\widehat{(u_I)} = \widehat{(\pi_e u)} = \hat{\pi} \hat{u} = \hat{u}_I. \quad \square$$

Пусть теперь  $\Omega$  — многоугольная область в  $R^n$  и  $\mathcal{T}_h$  — точная триангуляция области  $\Omega$  семейством аффинно-эквивалентных конечных элементов. Напомним, что при этом:

1) различные элементы  $e \in \mathcal{T}_h$  либо не пересекаются, либо имеют общую вершину или общую грань;

2)  $\text{diam } e \leq h \quad \forall e \in \mathcal{T}_h$ ;

3)  $\bar{\Omega} = \bigcup_{e \in \mathcal{T}_h} e$ .

Предположим дополнительно, что триангуляция  $\mathcal{T}_h$  является регулярной (как семейство элементов), т. е.

4) для радиуса  $\rho_e$  вписанного в  $e$  шара справедлива оценка

$$\rho_e \geq c h \quad \forall e \in \mathcal{T}_h.$$

**Теорема 4.** Пусть  $\mathcal{T}_h$  точная регулярная триангуляция области  $\Omega$ , и  $\mathcal{E}_h = \{(e, P_e, \Sigma_e), e \in \mathcal{T}_h\}$  — семейство аффинно-эквивалентных элементов, ассоциированных с базисным элементом  $(\hat{e}, \hat{P}, \hat{\Sigma})$ . Пусть, далее, пространство конечных элементов  $V_h$ , построенное на основе этого семейства, является подпространством  $H_p^m(\Omega) \subset H_p^{s+1}(\Omega)$  с  $m \geq 0$ ,  $s \geq 1$ , максимальный порядок производных в определении  $\hat{\Sigma}$  равен  $r \geq 0$  и  $P_s \subset \hat{P}$ . Тогда для любой функции  $u \in H_p^{s+1}(\Omega)$ ,  $(s+1-r)p > n$ ,  $m \leq s+1$ , справедлива оценка погрешности интерполяции

$$|u - u_I|_{m,p,\Omega} \leq c h^{s+1-m} |u|_{s+1,p,\Omega}, \quad (8)$$

где  $u_I$  есть  $V_h$ -интерполянт функции  $u$ .

**ДОКАЗАТЕЛЬСТВО.** В силу регулярности семейства элементов  $\mathcal{T}_h$  имеем  $\|B_e\| \leq ch$ ,  $\|B_e^{-1}\| \leq ch^{-1}$  для любого  $e \in \mathcal{T}_h$ . Комбинируя эти неравенства с оценкой (7), будем иметь

$$|u - u_I|_{m,p,e}^p \leq c h^{p(s+1-m)} |u|_{s+1,p,e}^p.$$

Суммируя эти оценки по всем  $e \in \mathcal{T}_h$ , получим неравенство (8).  $\square$

Отметим некоторые частные случаи:

Если область  $\Omega \subset R^2$  разбита на треугольники с прямолинейными сторонами или на параллелограммы, и на каждом элементе использована интерполяция Лагранжа ( $r = 0$ ), или Эрмита ( $r > 0$ ) порядка  $s > r$  (полиномы  $P_s$  или  $Q_s$ ) при построении пространства  $V_h$ , то справедливы оценки ( $p = 2$ ):

$$|u - u_I|_{0,\Omega} \leq c h^{s+1} |u|_{s+1,\Omega}, \quad |u - u_I|_{1,\Omega} \leq c h^s |u|_{s+1,\Omega}.$$

Мы не можем оценить  $|u - u_I|_{m,\Omega}$  при  $m > 1$  для лагранжевых элементов, поскольку разность  $u - u_I$  в этом случае, вообще говоря, принадлежит только пространству  $H^1$ . При использовании эрмитовых элементов класса  $C^1$  справедливо вложение  $V_h \subset H^2(\Omega)$ , поэтому в оценке (8) можно выбрать  $m = 0, 1, 2$ . Например, для пространства бикубических эрмитовых элементов  $V_h$  ( $\hat{P} = Q_3 \supset P_3$ ) мы можем оценить  $u - u_I$  в пространстве  $H^2$ . Из теоремы 4 следует, что

$$|u - u_I|_{m,\Omega} \leq c h^{4-m} |u|_{4,\Omega}, \quad m = 0, 1, 2.$$

**3. Оценка погрешности интерполяции для криволинейных и изопараметрических элементов.** Исследование погрешности в данном случае проводится по той же схеме, что и в случае

аффинно-эквивалентных элементов. А именно, осуществляется переход с исходного элемента на базисный в определении функционала погрешности, затем используется лемма Брамбла — Гильберта и в полученной оценке делается обратный переход на исходный конечный элемент. Различие между криволинейными и изопараметрическими элементами при таком подходе отсутствует, поскольку эти элементы различаются только видом преобразования, но не его свойствами. Имеется, однако, некоторое различие в исследовании треугольных и четырехугольных элементов в силу разных определений регулярности.

Рассмотрим вначале треугольные элементы.

**Теорема 5.** Пусть  $(\hat{e}, \hat{P}, \hat{\Sigma})$  — базисный треугольный лагранжес конечный элемент в  $R^2$  и  $(e, P_e, \Sigma_e)$  — ассоциированный с ним посредством преобразования  $x = x_e(\hat{x})$  криволинейный (изопараметрический) элемент. Предположим, что  $\hat{P} \supseteq P_s$ ,  $s \geq 1$ , элемент  $e$  принадлежит семейству  $(s+1)$ -регулярных элементов, т. е. выполнены условия теоремы 2, с.115, при  $k = s+1$  (теоремы 2, с.123, при  $m = s+1$ ). Тогда

$$|\hat{u}|_{s,p,\hat{e}} \leq c h^{s-2/p} \|u\|_{s,p,e}, \quad s \geq 0, \quad (9)$$

$$|u|_{s,p,e} \leq c h^{2/p-s} |\hat{u}|_{s,p,\hat{e}}, \quad s \geq 0, \quad (10)$$

$$|u - u_I|_{m,p,e} \leq c h^{s+1-m} \|u\|_{s+1,p,e}, \quad s \geq 1. \quad (11)$$

Здесь  $\hat{u}(\hat{x}) = u(x_e(\hat{x}))$ ,  $u_I$  есть  $P_e$ -интерполюнт функции  $u$ ,  $0 \leq m \leq s+1$ ,  $p \in [1, \infty]$ .

**ДОКАЗАТЕЛЬСТВО.** Установим оценку (9) при  $p < \infty$ . Для этого воспользуемся оценкой (32), с. 55,

$$|\hat{u}|_{s,p,\hat{e}}^p \leq c \max_{x \in e} |\det(Dx_e^{-1}(x))| \sum_{l=0}^s |u|_{l,p,e}^p \sum_{i \in I(l,s)} \max_{\hat{x} \in \hat{e}} (\|Dx_e(\hat{x})\|^{p i_1} \cdots \|D^s x_e(\hat{x})\|^{p i_s}), \quad (12)$$

где множество индексов  $I(l, s)$  определяется соотношением:

$$I(l, s) = \{i = (i_1, \dots, i_s) : i_1 + \dots + i_s = l, i_1 + 2i_2 + \dots + si_s = s\}.$$

В силу предположения о регулярности семейства конечных элементов

величины  $\max_{\hat{x} \in \hat{e}} \|D^k x_e(\hat{x})\|$  имеют порядок  $h^k$ . Поэтому

$$\begin{aligned} \sum_{i \in I(l,s)} \max_{\hat{x} \in \hat{e}} (\|Dx_e(\hat{x})\|^{p_{i_1}} \|D^2x_e(\hat{x})\|^{p_{i_2}} \cdots \|D^s x_e(\hat{x})\|^{p_{i_s}}) \\ \leq c \sum_{i \in I(l,s)} h^{p_{i_1}} h^{p_{i_2}} \cdots h^{p_{i_s}} \leq c \sum_{i \in I(l,s)} h^{ps} \leq c h^{ps}. \end{aligned}$$

Поскольку  $|\det(Dx_e^{-1}(x))| = O(h^{-2})$ , то из (12) получаем искомую оценку

$$|\hat{u}|_{s,p,\hat{e}}^p \leq c h^{ps-2} \sum_{l=0}^s |u|_{l,p,e}^p = c h^{ps-2} \|u\|_{s,p,e}^p.$$

Аналогично рассматривается случай  $p = \infty$ . Для доказательства (10) воспользуемся следствием 2, с. 56. Имеем при конечных  $p$ :

$$\begin{aligned} |u|_{s,p,e}^p &\leq c \max_{\hat{x} \in \hat{e}} |\det(Dx_e(\hat{x}))| \sum_{l=0}^s |\hat{u}|_{l,p,\hat{e}}^p \times \\ &\times \sum_{i \in I(l,s)} \max_{x \in e} (\|Dx_e^{-1}(x)\|^{p_{i_1}} \cdots \|D^s x_e^{-1}(x)\|^{p_{i_s}}) \leq c h^2 \sum_{l=0}^s |\hat{u}|_{l,p,\hat{e}}^p \times \\ &\times \sum_{i \in I(l,s)} h^{-p(i_1+i_2+\dots+i_s)} \leq c h^2 \sum_{l=0}^s |\hat{u}|_{l,p,\hat{e}}^p h^{-pl} \leq c h^{2-ps} \|\hat{u}\|_{s,p,\hat{e}}^p. \end{aligned}$$

Здесь мы воспользовались оценкой (8) теоремы 2, с. 105. Доказательство оценки (10) при  $p = \infty$  проводится аналогично.

Заключительная оценка (11) получается комбинированием уже найденных оценок и оценки из теоремы 1, с. 147, поскольку по условию  $\hat{u}_I \in \hat{P} \supseteq P_s$ :

$$\begin{aligned} |u - u_I|_{m,p,e} &\leq c h^{2/p-m} |\hat{u} - \hat{u}_I|_{m,p,\hat{e}} \leq \\ &\leq c h^{2/p-m} |\hat{u}|_{s+1,p,\hat{e}} \leq c h^{s+1-m} \|u\|_{s+1,p,e}. \quad \square \end{aligned}$$

Пусть теперь  $V_h$  — пространство криволинейных конечных элементов, аппроксимирующее пространство Соболева  $H^1(\Omega)$  в случае ограниченной области  $\Omega \subset R^2$  с кусочно гладкой границей  $\Gamma$  такое, что  $V_h \subset H^1(\Omega)$ .

Напомним кратко его определение, основанное на точной триангуляции  $\mathcal{T}_h$  области  $\Omega$ . Каждый элемент  $e$  является либо треугольником, либо криволинейным треугольником, причем две стороны криволинейного треугольника — отрезки прямых, третья — дуга кривой  $\Gamma$ .

Триангуляция считается  $(s+1)$ -регулярной для некоторого  $s \geq 1$ . Последнее предполагает регулярность триангуляции, которая получается из описанной заменой криволинейных сторон треугольников соответствующими хордами. Это предположение обеспечивает справедливость условия (3) теоремы 2, с. 115.

С каждым элементом триангуляции свяжем конечный элемент (аффинно-эквивалентный или криволинейный) ассоциированный с базисным лагранжевым конечным элементом  $(\hat{e}, \hat{P}, \hat{\Sigma})$  класса  $C^0$ , причем  $\hat{P} = P_s$ . На основе этого семейства конечных элементов стандартным образом определяется далее пространство  $V_h$  (пространство криволинейных конечных элементов).

Дадим оценку погрешности интерполирования функциями из  $V_h$ .

**Теорема 6.** Пусть  $V_h$  — пространство треугольных криволинейных конечных элементов, ассоциированных с лагранжевым базисным элементом  $(\hat{e}, \hat{P}, \hat{\Sigma})$  класса  $C^0$ , причем  $\hat{P} = P_s$  для некоторого  $s \geq 1$ . Пусть далее  $u \in H_p^{s+1}(\Omega)$ ,  $u_I$  есть  $V_h$ -интерполянты функции  $u$ , триангуляция  $(s+1)$ -регулярна. Тогда

$$|u - u_I|_{m,p,\Omega} \leq c h^{s+1-m} \|u\|_{s+1,p,\Omega}, \quad m = 0, 1.$$

ДОКАЗАТЕЛЬСТВО. Понятно, что в представлении погрешности

$$|u - u_I|_{m,p,\Omega}^p = \sum_{e \in \mathcal{T}_h} |u - u_I|_{m,p,e}^p \quad (13)$$

достаточно оценить лишь вклады от каждого из элементов. Если элемент  $e$  является аффинно-эквивалентным, то соответствующая оценка получена в теореме 3, с. 150. В силу регулярности триангуляции она имеет вид

$$|u - u_I|_{m,p,e} \leq c h^{s+1-m} |u|_{s+1,p,e}.$$

Если элемент  $e$  является криволинейным, то требуемая оценка получена выше (см. оценку (11)). Из этих оценок и представления (13) следует утверждение теоремы.  $\square$

ЗАМЕЧАНИЕ 1. Если в определении пространства  $V_h$ , сформулированном выше, заменить криволинейные элементы на изопараметрические, то мы получим аппроксимацию пространства  $H^1(\Omega_h)$ . Очевидно, что в этом случае справедлива теорема, получающаяся из теоремы 6 заменой области  $\Omega$  на  $\Omega_h$ .

Рассмотрим теперь прямоугольные изопараметрические элементы.

**Теорема 7.** Пусть  $(\hat{e}, \hat{P}, \hat{\Sigma})$  — базисный прямоугольный лагранжев конечный элемент в  $R^2$ , а  $(e, P_e, \Sigma_e)$  — ассоциированный с ним

посредством преобразования  $x = x_e(\hat{x})$  изопараметрический элемент. Если элемент  $e$  принадлежит семейству  $(s+1)$ -регулярных элементов, пространство  $\hat{P}$  содержит  $Q_s$ , то

$$[\hat{u}]_{s,p,\hat{e}} \leq c h^{s-2/p} \|u\|_{s,p,e}, \quad s \geq 0, \quad (14)$$

$$|u|_{s,p,e} \leq c h^{2/p-s} |\hat{u}|_{s,p,\hat{e}}, \quad s \geq 0, \quad (15)$$

$$|u - u_I|_{m,p,e} \leq c h^{s+1-m} \|u\|_{s+1,p,e}, \quad s \geq 1. \quad (16)$$

Здесь  $u_I$  есть  $P_e$ -интерполянт функции  $u$ ,  $0 \leq m \leq s+1$ ,  $p \in [1, \infty]$ .

ДОКАЗАТЕЛЬСТВО. Докажем оценку (14). Для этого воспользуемся леммой 5, с. 57. При конечных  $p$  имеем

$$[\hat{u}]_{s,p,\hat{e}}^p \leq c \max_{x \in e} |\det(Dx_e^{-1}(x))| \sum_{l=0}^s |u|_{l,p,e}^p \sum_{i \in I(l,s)} \max_{\hat{x} \in \hat{e}} ([Dx_e(\hat{x})]^{p i_1} \cdots [D^s x_e(\hat{x})]^{p i_s}). \quad (17)$$

В силу предположения о регулярности семейства конечных элементов величины  $\max_{\hat{x} \in \hat{e}} [D^k x_e(\hat{x})]$  имеют порядок  $h^k$ . Поэтому

$$\begin{aligned} \sum_{i \in I(l,s)} \max_{\hat{x} \in \hat{e}} ([Dx_e(\hat{x})]^{p i_1} [D^2 x_e(\hat{x})]^{p i_2} \cdots [D^s x_e(\hat{x})]^{p i_s}) \\ \leq c \sum_{i \in I(l,s)} h^{p i_1} h^{p 2 i_2} h^{p s i_s} \leq c \sum_{i \in I(l,s)} h^{p s} \leq c h^{p s}. \end{aligned}$$

Поскольку  $|\det(Dx_e^{-1}(x))| = O(h^{-2})$ , то из (17) получаем искомую оценку

$$[\hat{u}]_{s,p,\hat{e}}^p \leq c h^{p s - 2} \sum_{l=0}^s |u|_{l,p,e}^p = c h^{p s - 2} \|u\|_{s,p,e}^p.$$

Аналогично рассматривается случай  $p = \infty$ . Для доказательства (15) воспользуемся следствием 2, с. 56. Имеем при конечных  $p$ :

$$\begin{aligned} |u|_{s,p,e}^p &\leq c \max_{\hat{x} \in \hat{e}} |\det(Dx_e(\hat{x}))| \sum_{l=0}^s |\hat{u}|_{l,p,\hat{e}}^p \times \\ &\times \sum_{i \in I(l,s)} \max_{x \in e} (\|Dx_e^{-1}(x)\|^{p i_1} \cdots \|D^s x_e^{-1}(x)\|^{p i_s}) \leq c h^2 \sum_{l=0}^s |\hat{u}|_{l,p,\hat{e}}^p \times \\ &\times \sum_{i \in I(l,s)} h^{-p(i_1 + 2i_2 + \cdots + s i_s)} \leq c h^2 \sum_{l=0}^s |\hat{u}|_{l,p,\hat{e}}^p h^{-ps} \leq c h^{2-ps} \|\hat{u}\|_{s,p,\hat{e}}^p. \end{aligned}$$

Здесь мы воспользовались оценкой (9) теоремы 2, с. 105. Заключительная оценка (11) получается комбинированием уже найденных оценок и оценки из теоремы 1, с. 147, поскольку по условию  $\hat{u}_I \in \hat{P} \supseteq Q_s$ :

$$\begin{aligned} |u - u_I|_{m,p,e} &\leq c h^{2/p-m} |\hat{u} - \hat{u}_I|_{m,p,\hat{e}} \leq \\ &\leq c h^{2/p-m} [\hat{u}]_{s+1,p,\hat{e}} \leq c h^{s+1-m} \|u\|_{s+1,p,e}. \quad \square \end{aligned}$$

Непосредственным следствием доказанной теоремы является

**Теорема 8.** Пусть  $V_h$  — пространство прямоугольных изопараметрических конечных элементов, ассоциированных с лагранжесвым базисным элементом  $(\hat{e}, \hat{P}, \hat{\Sigma})$  класса  $C^0$ , причем  $\hat{P} = Q_s$  для некоторого  $s \geq 1$ . Пусть далее  $u \in H_p^{s+1}(\Omega_h)$ ,  $\Omega_h = \bigcup\{e : e \in \mathcal{T}_h\}$ ,  $u_I$  есть  $V_h$ -интерполянты функции  $u$ , триангуляция  $(s+1)$ -регулярна. Тогда

$$|u - u_I|_{m,p,\Omega_h} \leq c h^{s+1-m} \|u\|_{s+1,p,\Omega_h}, \quad m = 0, 1.$$

---

---

ГЛАВА 4

**Метод конечных элементов для эллиптических уравнений**

В настоящей главе построены и исследованы методы конечных элементов для эллиптических уравнений второго и четвертого порядков в плоской двумерной области с кусочно гладкой границей.

Для уравнений второго порядка рассматриваются методы, основанные на использовании пространства лагранжевых криволинейных элементов с точной аппроксимацией границы области, а также схемы для плоской многоугольной области, основанные на использовании эрмитова треугольного элемента класса  $C^0$ . Для уравнений четвертого порядка — схемы, основанные на использовании прямоугольного бикубического элемента и треугольника Белла.

Подробно проанализированы схемы, возникающие при применении квадратурных формул для вычисления элементов матриц и правых частей системы линейных алгебраических уравнений, определяющей приближенное решение в методе конечных элементов.

**§ 1. Схемы с точной аппроксимацией границы для уравнений второго порядка**

**1. Задача Дирихле. Лагранжевы элементы.** Рассмотрим неоднородную задачу Дирихле.

$$-\operatorname{div}(A\nabla u) + a\nabla u + a_0u = f, \quad x \in \Omega, \quad (1)$$

$$u(x) = g(x), \quad x \in \Gamma, \quad (2)$$

предполагая выполненными условия теоремы 1, с. 64. Кроме того, функцию  $g$  будем считать непрерывной на  $\Gamma$  и допускающей продолжение в пространство  $H^1(\Omega)$ . При этих условиях задача (1), (2) имеет единственное обобщенное решение из  $H^1(\Omega)$ .

Введем в рассмотрение пространство  $V_h$  лагранжевых конечных элементов с точной аппроксимацией границы, построенное на основе триангуляции  $\mathcal{T}_h$  области  $\Omega$  треугольными элементами. Напомним, что приграничные элементы — треугольники, одна сторона каждого

которых, вообще говоря, криволинейна и совпадает с соответствующим участком кривой  $\Gamma$ . Каждый конечный элемент ассоциирован с треугольным базисным элементом  $(\hat{e}, \hat{P}, \hat{\Sigma})$  класса  $C^0$ , причем  $\hat{P} \supseteq P_m$  для некоторого  $m \geq 1$ . Отметим, что в простейшем случае, когда область  $\Omega$  — многоугольник, пространство  $V_h$  строится на основе семейства аффинно-эквивалентных треугольных элементов степени  $m$ .

Наряду с пространством  $V_h$  будем рассматривать его подпространство

$$V_h^0 = \{v_h \in V_h : v_h(x) = 0, x \in \Gamma\}.$$

Его эквивалентное определение (см. с. 143):

$$V_h^0 = \{v_h \in V_h : v_h(x) = 0, x \in \gamma_h\},$$

где  $\gamma_h$  — множество узлов интерполяции, лежащих на  $\Gamma$ .

**Определение 1.** Функцию  $u_h \in V_h$  будем называть приближенным решением задачи (1), (2), если она удовлетворяет интегральному тождеству

$$a(u_h, v_h) = f(v_h) \quad \forall v_h \in V_h^0 \quad (3)$$

и граничному условию

$$u_h(x) = g(x), \quad x \in \gamma_h. \quad (4)$$

Напомним, что

$$a(u, v) = \int_{\Omega} (A \nabla u \cdot \nabla v + a \nabla u v + a_0 u v) dx \quad \forall u, v \in H^1, \quad (5)$$

$$f(v) = \int_{\Omega} f v dx. \quad (6)$$

Обозначим, как обычно, через  $\varphi_i(x)$ ,  $i = 1, 2, \dots, N$ , базис Лагранжа в  $V_h$ , причем  $\varphi_i(x)$  с номерами  $i = 1, 2, \dots, N_0$  соответствуют узлам, не лежащим на  $\Gamma$ . Тогда

$$u_h(x) = \sum_{j=1}^{N_0} y_j \varphi_j(x) + \sum_{j=N_0+1}^N g(x_j) \varphi_j(x), \quad (7)$$

$y_j = u_h(x_j)$ , а произвольная функция  $v_h \in V_h^0$  записывается в виде

$$v_h(x) = \sum_{i=1}^{N_0} v_i \varphi_i(x), \quad v_i = v_h(x_i). \quad (8)$$

Подставляя разложения (7) и (8) в (3), приходим к системе линейных алгебраических уравнений для отыскания  $y = (y_1, y_2, \dots, y_{N_0})$ :

$$Ay = F,$$

где  $A$  — матрица с элементами

$$A_{ij} = a(\varphi_i, \varphi_j), \quad i, j = 1, 2, \dots, N_0,$$

$F$  — вектор с элементами

$$F_i = \int_{\Omega} f(x)\varphi_i(x)dx - \sum_{j=N_0+1}^N A_{ij}g(x_j), \quad i = 1, 2, \dots, N_0.$$

Способы вычисления элементов матрицы и правой части этой системы будут подробно описаны далее в §2 приложения. Здесь мы займемся исследованием ее разрешимости и оценкой точности приближенного решения  $u_h$ .

Отметим, что вследствие неоднородности граничного условия (4) мы не можем непосредственно сослаться на результаты §1 гл. 2, относящиеся к методу Галеркина, но будем использовать развитые там способы исследования.

**Теорема 1.** Пусть выполнены условия теоремы 1, с. 64. Тогда задача (3), (4) имеет единственное решение.

**ДОКАЗАТЕЛЬСТВО.** Убедимся, что соответствующая однородная система уравнений

$$a(u_h, v_h) = 0 \quad \forall v_h \in V_h^0, \tag{9}$$

$$u_h(x) = 0, \quad x \in \gamma_h, \tag{10}$$

имеет только тривиальное решение. Действительно, любое решение  $u_h$  системы (9), (10) принадлежит пространству  $V_h^0 \subset H_0^1$ . Поэтому в тождестве (9) можно положить  $v_h = u_h$ , следовательно,  $a(u_h, u_h) = 0$ . На основании леммы 2, с. 64, форма  $a(\cdot, \cdot)$  положительно определена на  $H_0^1$ , таким образом,  $u_h = 0$ .  $\square$

Оценка точности приближенного решения наиболее просто получается в случае однородных граничных условий.

**Теорема 2.** Пусть  $u$  — решение задачи (1), (2),  $u_h$  — решение задачи (3), (4),  $g(x) = 0$ ,  $x \in \Gamma$ , выполнены условия теоремы 1, с. 64. Тогда существует постоянная  $c > 0$  такая, что

$$\|u - u_h\|_{1,\Omega} \leq c \|u - v_h\|_{1,\Omega} \tag{11}$$

для любой функции  $v_h \in V_h^0$ .

ДОКАЗАТЕЛЬСТВО. В рассматриваемом случае задачу Дирихле (1), (2) можно трактовать как задачу об отыскании функции  $u \in H_0^1$ , такой что

$$a(u, v) = f(v) \quad \forall v \in H_0^1, \quad (12)$$

а метод конечных элементов — как метод Галеркина приближенного решения задачи (12). Именно, мы разыскиваем функцию  $u_h$  из  $V_h^0 \subset H_0^1$ , удовлетворяющую интегральному тождеству

$$a(u_h, v_h) = f(v_h) \quad \forall v_h \in V_h^0.$$

Таким образом, утверждение теоремы — непосредственное следствие леммы Сеа (см. с. 25).  $\square$

**Теорема 3.** Пусть выполнены условия теоремы 2. Область  $\Omega$  — многоугольник, триангуляция  $\mathcal{T}_h$  регулярна,  $\hat{P} \supseteq P_m$ ,  $m \geq 1$ , решение задачи (1), (2) принадлежит пространству  $H^{l+1}(\Omega)$ ,  $1 \leq l \leq m$ . Тогда

$$\|u - u_h\|_{1,\Omega} \leq ch^l \|u\|_{l+1,\Omega}.$$

ДОКАЗАТЕЛЬСТВО. В неравенстве (11) положим  $v_h = u_I$ , где  $u_I$  есть  $V_h^0$ -интерполянт функции  $u$  и воспользуемся оценкой теоремы 4, с. 151, при  $m = 1$ ,  $s = l$ ,  $p = 2$ .  $\square$

ЗАМЕЧАНИЕ 1. Условие принадлежности решения пространству  $H^l(\Omega)$  налагает определенные ограничения на исходные данные задачи (1), (2). В общем случае из-за наличия углов области они имеют сложный интегральный характер. В простейшем случае, когда область  $\Omega$  — выпуклый многоугольник,  $l = 2$ , достаточно потребовать выполнения следующих условий:  $a_{ij} \in W_\infty^1(\Omega)$ ,  $a_i \in L_\infty(\Omega)$ ,  $f \in L_2(\Omega)$ .

Для криволинейных областей с гладкой границей, опираясь на теорему 4, с. 72, можно сформулировать простые достаточные условия в терминах исходных данных задачи (1), (2), обеспечивающие заданный порядок точности метода конечных элементов.

**Теорема 4.** Пусть выполнены условия теоремы 2,  $\hat{P} \supseteq P_m$ ,  $m \geq 1$ , и справедливы следующие условия гладкости исходных данных задачи (1), (2):

$$a_{ij} \in W_\infty^l(\Omega), \quad a_i \in W_\infty^{l-1}(\Omega), \quad f \in W_2^{l-1}(\Omega), \quad \Gamma \in C^{l+1} \quad (13)$$

при некотором  $l$ ,  $1 \leq l \leq m$ . Тогда  $\|u - u_h\|_{1,\Omega} \leq ch^l \|f\|_{l-1,\Omega}$ .

ДОКАЗАТЕЛЬСТВО. В неравенстве (11) положим  $v_h = u_I$ , где  $u_I$  есть  $V_h^0$ -интерполянт функции  $u$ . Поскольку  $\Gamma \in C^{l+1}$ , то точная

триангуляция области  $\Omega$  является  $(l+1)$ -регулярной, поэтому можно воспользоваться оценкой теоремы 6, с. 154, при  $m = 1$ ,  $s = l$ ,  $p = 2$ . Получим  $\|u - u_h\|_{1,\Omega} \leq ch^l \|u\|_{l+1,\Omega}$ . Теперь остается учесть, что согласно теореме 4, с. 72, справедливо неравенство  $\|u\|_{l+1,\Omega} \leq c \|f\|_{l-1,\Omega}$ .  $\square$

Получим оценки точности для неоднородной задачи Дирихле.

**Теорема 5.** Пусть выполнены условия теоремы 1, с. 64,  $u$  — решение задачи (1), (2),  $u_h$  — решение задачи (3), (4). Тогда

$$\|u_h - u\|_{1,\Omega} \leq c \|u_I - u\|_{1,\Omega}.$$

ДОКАЗАТЕЛЬСТВО. Поскольку  $u_I - u_h \in H_0^1$ , справедлива оценка

$$c |u_I - u_h|_{1,\Omega}^2 \leq a(u_I - u_h, u_I - u_h).$$

Далее,

$$\begin{aligned} a(u_I - u_h, u_I - u_h) &= a(u_I - u, u_I - u_h) + a(u, u_I - u_h) - a(u_h, u_I - u_h) = \\ &= a(u_I - u, u_I - u_h) \leq c \|u_I - u\|_{1,\Omega} |u_I - u_h|_{1,\Omega}, \end{aligned}$$

следовательно,

$$|u_I - u_h|_{1,\Omega} \leq c \|u_I - u\|_{1,\Omega},$$

и, наконец,  $\|u_h - u\|_{1,\Omega} \leq \|u_h - u_I\|_{1,\Omega} + \|u_I - u\|_{1,\Omega} \leq c \|u_I - u\|_{1,\Omega}$ .  $\square$

Используя теперь оценки погрешности интерполяции, полученные в теоремах 4, 6 §10 гл. 3, можно сформулировать оценки погрешности, полностью аналогичные случаю однородной задачи. Отметим только, что дополнительно к условиям (13) нужно, естественно, предположить, что функция  $g$  есть след на  $\Gamma$  некоторой функции из  $H^{l+1}(\Omega)$ .

**2. Задача Дирихле. Эрмитов кубический элемент класса  $C^0$ .** При построении и исследовании метода конечных элементов для задачи (1), (2) с использованием эрмитова кубического элемента класса  $C^0$  будем предполагать, что область  $\Omega$  — многоугольник и рассматриваемая задача имеет решение, непрерывно дифференцируемое в замыкании области  $\Omega$ . Отсюда вытекает, что функция  $g$  непрерывно дифференцируема вдоль каждой стороны многоугольника  $\Omega$ . Причем  $u_\tau(x) = g_\tau(x)$ , где  $\tau$  — единичный вектор касательной к  $\Gamma$  в точке  $x$ . Ясно, что если  $x$  — угловая точка области  $\Omega$ , то  $\partial u(x)/\partial x_1$ ,  $\partial u(x)/\partial x_2$  выражаются через производные функции  $g$  по направлениям соответствующих сторон.

Построим на области  $\Omega$  триангуляцию и введем в рассмотрение пространство  $V_h$  эрмитовых кубических элементов (см. с. 139), а также его подпространство  $V_h^0$  (см. с. 144).

**Определение 2.** Функцию  $u_h \in V_h$  назовем приближенным решением задачи (1), (2), если она удовлетворяет интегральному тождеству

$$a(u_h, v_h) = f(v_h) \quad \forall v_h \in V_h^0 \quad (14)$$

и граничным условиям

$$u_h(x) = g(x), \quad x \in \gamma_h, \quad u_{h\tau}(x) = g_\tau(x), \quad \tau \in \tau(x), \quad x \in \gamma_h/\gamma_h^*, \quad (15)$$

$$\frac{\partial u_h(x)}{\partial x_1} = \frac{\partial u(x)}{\partial x_1}, \quad \frac{\partial u_h(x)}{\partial x_2} = \frac{\partial u(x)}{\partial x_2}, \quad x \in \gamma_h^*. \quad (16)$$

Напомним, что, как обычно, через  $\gamma_h$  обозначено множество всех узлов элементов, принадлежащих  $\Gamma$ ,  $\gamma_h^*$  — угловые точки области  $\Omega$ .

Рассуждая так же, как при описании пространства  $V_h^0$  на с. 145, получим, что функция  $u_h \in V_h$ , удовлетворяющая граничным условиям (15), (16), представима в виде

$$\begin{aligned} u_h(x) = & \sum_{a_i \in \gamma_h/\gamma_h^*} u_{i1}(\varphi_{i1}^g(x) - \operatorname{tg} \alpha(a_i)\varphi_{i2}^g(x)) + \sum_{a_i \in \omega_h/\gamma_h} u_i \varphi_{i0}^g(x) + \\ & + \sum_{a_i \in \omega_h^{(1)}/\gamma_h} \sum_{k=1}^2 u_{ik} \varphi_{ik}^g(x) + \sum_{a_i \in \gamma_h/\gamma_h^*} g_\tau(a_i)(\cos \alpha(a_i))^{-1} \varphi_{i2}^g(x) + \\ & + \sum_{a_i \in \gamma_h^*} \sum_{k=1}^2 \frac{\partial u(a_i)}{\partial x_k} \varphi_{ik}^g(x) + \sum_{a_i \in \gamma_h} u(a_i) \varphi_{i0}^g(x). \quad (17) \end{aligned}$$

Параметры  $u_i$ ,  $u_{i1}$ ,  $u_{i2}$  подлежат определению из системы линейных алгебраических уравнений, порождаемой тождеством (14).

**Теорема 6.** Пусть выполнены условия теоремы 1, с. 64. Тогда задача (14), (15) имеет единственное решение.

**ДОКАЗАТЕЛЬСТВО.** Как показано на с. 144, функция  $u_h \in V_h$ , удовлетворяющая однородным граничным условиям (15), принадлежит  $V_h^0 \subset H_0^1$ . Поэтому в однородном тождестве, соответствующем (14), можно положить  $v_h = u_h$ , откуда вследствие положительной определенности квадратичной формы  $a(\cdot, \cdot)$  на пространстве  $H_0^1$  получаем, что  $u_h = 0$ , т. е. однородная задача (14), (15) может иметь только тривиальное решение.  $\square$

**Теорема 7.** Пусть выполнены условия теоремы 1, с. 64, триангуляция, использованная при построении пространства,  $V_h$  регулярна, решение задачи (1), (2) принадлежит пространству  $H^l(\Omega)$ ,  $l = 3, 4$ . Тогда

$$\|u - u_h\|_{1,\Omega} \leq ch^{l-1} \|u\|_{l,\Omega}.$$

Доказательство проводится вполне аналогично доказательству теоремы 5 с последующим использованием оценки погрешности интерполирования, полученной в теореме 4, с. 151. Отметим, что условие  $l \geq 3$  вызвано тем, что в число узловых параметров в рассматриваемом случае входят значения производных в угловых точках элементов, а это значит, что оператор интерполирования определен лишь на функциях из  $C^1$ .

**3. Третья краевая задача.** Рассмотрим третью краевую задачу для эллиптического уравнения второго порядка как задачу об отыскании функции  $u \in H^1(\Omega)$ , удовлетворяющей интегральному тождеству

$$\begin{aligned} a(u, v) &= \int_{\Omega} (A \nabla u \cdot \nabla \eta + a \nabla u \eta + a_0 u \eta) dx + \int_{\Gamma} \sigma u \eta dx = \\ &= \int_{\Omega} f \eta dx + \int_{\Gamma} g \eta dx \quad \forall \eta \in H^1(\Omega). \end{aligned} \quad (18)$$

Предположим выполненными условия теоремы 2, с. 66, обеспечивающие ограниченность и положительную определенность квадратичной формы  $a(\cdot, \cdot)$ .

Построение и исследование метода конечных элементов в рассматриваемом случае выполняется проще, чем для задачи Дирихле. Оно целиком укладывается в схему метода Галеркина. Для определенности остановимся на использовании конечноэлементного пространства  $V_h$  лагранжевых треугольных, вообще говоря, криволинейных элементов.

Функцию  $u_h \in V_h$  назовем приближенным решением третьей краевой задачи, если она удовлетворяет интегральному тождеству

$$a(u_h, v_h) = \int_{\Omega} f v_h dx + \int_{\Gamma} g v_h dx \quad \forall v_h \in V_h.$$

Представляя функции  $u_h, v_h$  в виде разложений по базисным функциям Лагранжа пространства  $V_h$ :

$$u_h(x) = \sum_{j=1}^N y_j \varphi_j(x), \quad v_h(x) = \sum_{j=1}^N v_j \varphi_j(x), \quad y_j = u_h(x_j), \quad v_j = v_h(x_j),$$

получим систему линейных алгебраических уравнений для отыскания вектора  $y = (y_1, y_2, \dots, y_N)$ :

$$Ay = F. \quad (19)$$

Элементы матрицы  $A$  и вектора  $F$  определяются соотношениями:

$$A_{ij} = a(\varphi_i, \varphi_j), \quad i, j = 1, \dots, N,$$

$$F_i = \int_{\Omega} f \varphi_i(x) dx + \int_{\Gamma} g \varphi_i(x) dx, \quad i = 1, \dots, N.$$

Способы фактического построения системы уравнений (19) обсуждаются в §2 приложения.

Однозначная разрешимость системы (19) обеспечивается положительной определенностью формы  $a(\cdot, \cdot)$ . Из леммы Сеа (см. с. 25) вытекает оценка погрешности приближенного решения:

$$\|u - u_h\|_{1,\Omega} \leq c \|u - v_h\|_{1,\Omega} \quad \forall v_h \in V_h,$$

откуда согласно теореме 6, с. 154, следует, что если триангуляция области  $\Omega$  является  $(l+1)$ -регулярной,  $\hat{P} \supseteq P_m$ ,  $m \geq 1$ , точное решение задачи (18) принадлежит пространству  $H^{l+1}(\Omega)$ ,  $1 \leq l \leq m$ , то

$$\|u - u_h\|_{1,\Omega} \leq ch^l \|u\|_{l+1,\Omega}.$$

## § 2. Схемы с точной аппроксимацией границы для уравнений четвертого порядка

В этом параграфе будет построен и исследован метод конечных элементов решения первой краевой задачи для эллиптических уравнений четвертого порядка:

$$\sum_{i,j,k,l=1}^2 \frac{\partial^2}{\partial x_i \partial x_j} \left( a_{ijkl} \frac{\partial^2 u}{\partial x_k \partial x_l} \right) - \sum_{i,j=1}^2 \frac{\partial}{\partial x_i} a_{ij} \frac{\partial u}{\partial x_j} +$$

$$+ \sum_{i=1}^2 a_i \frac{\partial u}{\partial x_i} + a_0 u = f(x), \quad x \in \Omega, \quad (1)$$

$$u(x) = g_1(x), \quad \frac{\partial u}{\partial \nu} = g_2(x), \quad x \in \Gamma. \quad (2)$$

Будем предполагать, что  $f \in L_2(\Omega)$ , функции  $g_1(x), g_2(x)$  допускают продолжение в  $H^2(\Omega)$  и  $H^1(\Omega)$  соответственно, коэффициенты уравнения удовлетворяют условиям, обеспечивающим ограничен-

ность и положительную определенность квадратичной формы

$$a(u, u) = \int_{\Omega} \left( \sum_{i,j,k,l=1}^2 a_{ijkl} \frac{\partial^2 u}{\partial x_i \partial x_j} \frac{\partial^2 u}{\partial x_k \partial x_l} + \sum_{i,j=1}^2 a_{ij} \frac{\partial u}{\partial x_i} \frac{\partial u}{\partial x_j} + \sum_{i=1}^2 a_i \frac{\partial u}{\partial x_i} u + a_0 u^2 \right) dx$$

на пространстве  $H_0^2(\Omega) \times H_0^2(\Omega)$  (см. с. 74). Можно показать, что при этих условиях задача (1), (2) имеет единственное обобщенное решение из пространства  $H^2(\Omega)$ .

При построении метода конечных элементов, как и в случае уравнений второго порядка при использовании эрмитовых элементов, на граничные функции  $g_1, g_2$  будут налагаться некоторые дополнительные требования гладкости.

Будем рассматривать два случая — схемы, построенные на основе прямоугольных бикубических элементов, и схемы с использованием треугольных элементов Белла.

**1. Прямоугольные бикубические элементы.** Пусть область  $\Omega$  составлена из прямоугольников, решение задачи (1), (2) дважды непрерывно дифференцируемо в замыкании области  $\Omega$ . Тогда функции  $g_1, g_2$  непрерывны на  $\Gamma$  и непрерывно дифференцируемы на каждой прямой, образующей границу области  $\Omega$ , причем производная функции  $g_2$  вдоль  $\Gamma$  непрерывна в угловых точках области  $\Omega$ , а функция  $g_2$  дважды непрерывно дифференцируема вдоль каждой прямой, образующей  $\Gamma$ .

Введем в рассмотрение пространство  $V_h$ , образованное семейством аффинно-эквивалентных прямоугольных бикубических элементов (см. с. 141), а также его подпространство  $V_h^0 \subset H_0^2$  (см. с. 146):

$$V_h^0 = \{v_h \in V_h : v_h(x) = \partial_{\tau} v_h(x) = \partial_{\nu} v_h(x) = \partial_{\nu\tau} v_h(x) = 0, x \in \gamma_h\}.$$

Под приближенным решением задачи (1), (2) будем понимать функцию  $u_h \in V_h$ , удовлетворяющую интегральному тождеству

$$a(u_h, v_h) = \int_{\Omega} f v_h dx \quad \forall v_h \in V_h^0 \quad (3)$$

и граничным условиям

$$u_h(x) = g_1(x), \quad \partial_{\tau} u_h(x) = g_{1\tau}(x), \quad \partial_{\nu} u_h(x) = g_2(x), \\ \partial_{\nu\tau} u_h(x) = g_{2\tau}(x), \quad x \in \gamma_h. \quad (4)$$

Граничные условия (4) определяют все узловые параметры функции  $u_h$  в узлах, принадлежащих  $\gamma_h$ .

**Теорема 1.** Пусть выполнены условия положительной определенности билинейной формы  $a(\cdot, \cdot)$  на пространстве  $H_0^2(\Omega) \times H_0^2(\Omega)$ . Тогда задача (3), (4) имеет единственное решение при любых  $f, g_1, g_2$ .

Доказательство вполне аналогично доказательству теоремы 1 и использует тот факт, что функция  $u_h \in V_h$ , удовлетворяющая однородным граничным условиям (4), принадлежит пространству  $V_h^0$ .

Аналогично теореме 5, с. 161, доказывается

**Теорема 2.** Пусть выполнены условия положительной определенности билинейной формы  $a(\cdot, \cdot)$ ,  $u$  — решение задачи (1), (2),  $u_h$  — решение задачи (3), (4),  $u_I \in V_h$  — интерполянт функции  $u$ . Тогда

$$\|u - u_h\|_{2,\Omega} \leq c \|u - u_I\|_{2,\Omega}.$$

Используя теперь оценку погрешности интерполяции из теоремы 4, с. 151, получим

**Следствие 1.** Пусть  $u \in H^4(\Omega)$ , триангуляция области  $\Omega$  регулярна. Тогда

$$\|u - u_h\|_{2,\Omega} \leq ch^2 \|u\|_{4,\Omega}. \quad (5)$$

**ЗАМЕЧАНИЕ 1.** Условие принадлежности точного решения задачи пространству  $H^4(\Omega)$  при получении оценки (5) необходимо, поскольку в число узловых параметров, определяющих интерполянт функции  $u$ , в данном случае входят ее вторые производные.

**2. Треугольные элементы Белла.** Предположим, что область  $\Omega$  — многоугольник. Под пространством  $V_h$  будем теперь понимать пространство, образованное семейством афинно-эквивалентных элементов Белла, соответствующим некоторой регулярной триангуляции области  $\Omega$  (см. с. 140). Через  $V_h^0 \subset V_h$  обозначим подпространство (см. с. 146), состоящее из функций, удовлетворяющих граничным условиям:

$$\begin{aligned} v_h(x) = \partial_\tau v_h(x) = \partial_\nu v_h(x) = \partial_{\tau\tau} v_h(x) = \partial_{\nu\tau} v_h(x) = 0, \quad x \in \gamma_h / \gamma_h^*, \\ v_h(x) = 0, Dv_h(x)h_1 = 0, Dv_h(x)h_2 = 0, \\ D^2v_h(x)h_1^2 = 0, D^2v_h(x)h_2^2 = 0, D^2v_h(x)h_1h_2 = 0, \quad x \in \gamma_h^*, \end{aligned} \quad (6)$$

где  $h_1, h_2$  — векторы, совпадающие по длине и направлению со сторонами элемента, пересекающимися в точке  $x \in \gamma_h^*$ .

Ясно, что если  $u$  — решение задачи (1), (2), то значения

$$u(x), Du(x)h_1, Du(x)h_1, D^2u(x)h_1^2, D^2u(x)h_2^2, D^2u(x)h_1h_2$$

в любой точке  $x \in \gamma_h^*$  и для любой пары векторов  $h_1, h_2$  выражаются через значения функций  $g_1, g_2$  и их производных не выше второго порядка вдоль сторон многоугольника  $\Omega$ , пересекающихся в точке  $x$ .

Под приближенным решением задачи (1), (2) будем понимать функцию  $u_h \in V_h$ , удовлетворяющую интегральному тождеству

$$a(u_h, v_h) = \int_{\Omega} f v_h dx \quad \forall v_h \in V_h^0$$

и граничным условиям

$$\begin{aligned} u_h(x) &= g_1(x), \quad \partial_{\tau} u_h(x) = g_{1\tau}(x), \quad \partial_{\nu} u_h(x) = g_2(x), \quad \partial_{\tau\tau} u_h(x) = g_{1\tau\tau}(x), \\ &\quad \partial_{\nu\tau} u_h(x) = g_{2\tau}(x), \quad x \in \gamma_h / \gamma_h^*, \\ u_h(x) &= u(x), \quad Du_h(x)h_1 = Du(x)h_1, \quad Du_h(x)h_2 = Du(x)h_2, \\ D^2u_h(x)h_1^2 &= Du(x)h_1^2, \quad D^2u_h(x)h_2^2 = Du(x)h_2^2, \\ D^2u_h(x)h_1h_2 &= Du(x)h_1h_2, \quad x \in \gamma_h^*. \end{aligned}$$

Здесь векторы  $h_1, h_2$  имеют тот же смысл, что и в равенствах (6).

#### Задача.

Построить представление функции  $u_h$ , аналогичное представлению (17) на с. 162.

Исследование разрешимости и получение оценки точности протекает здесь по той же схеме, что и при использовании прямоугольных бикубических элементов. Если  $u \in H^s$ ,  $s = 4, 5$ , то вследствие включения  $\hat{P} \supset P_4$  справедлива оценка:  $\|u - u_h\|_{2,\Omega} \leq ch^{s-2} \|u\|_{s,\Omega}$ .

### § 3. Дополнительные замечания

**1. Сходимость к обобщенному решению.** Приближенные решения, построенные при помощи методов, рассмотренных в предыдущих параграфах настоящей главы, обладают свойством сходимости к соответствующим обобщенным решениям без каких-либо дополнительных предположений об их гладкости. Докажем это утверждение для однородной задачи Дирихле (1), (2), т. е. при  $g = 0$ , применительно к методу с лагранжевыми треугольными криволинейными, вообще говоря, элементами (см. с. 157). Мы предполагаем, что используется 2-регулярная триангуляция,  $\hat{P} = P_l$ ,  $l \geq 1$ . Поскольку в

данном случае метод конечных элементов совпадает с методом Галлеркина, достаточно установить, что семейство подпространств  $V_h^0$  предельно полно в пространстве  $H_0^1$ . Иными словами, нужно показать, что для любой функции  $u \in H_0^1$  существует последовательность функций  $v_h \in V_h^0$ ,  $h \rightarrow 0$ , такая, что  $\lim_{h \rightarrow 0} \|u - v_h\|_{1,\Omega} \rightarrow 0$ . По определению пространства  $H_0^1(\Omega)$  для любого  $\varepsilon > 0$  можно указать функцию  $v \in C_0^\infty(\Omega)$  такую, что  $\|u - v\|_{1,\Omega} \leq \varepsilon/2$ . Положим  $v_h = v_I$ , где  $v_I \in V_h$  есть  $V_h$ -интерполюант функции  $v$ . Применяя полученные ранее оценки погрешности интерполяции (см. теорему 4, с. 151), будем иметь, что  $\|v - v_h\|_{1,\Omega} \leq h\|v\|_{2,\Omega}$ , причем  $\|v - v_h\|_{1,\Omega} \leq \varepsilon/2$ , если  $h$  достаточно мало, т. е.  $\|u - v_h\|_{1,\Omega} \leq \varepsilon$  и сходимость метода конечных элементов доказана.

**2. Оценки погрешности метода конечных элементов в слабых нормах.** В предыдущих параграфах настоящей главы были получены оценки погрешности метода конечных элементов в энергетических нормах, а именно, в норме пространства  $H^1(\Omega)$  для уравнений второго порядка и в норме пространства  $H^2(\Omega)$  для уравнений четвертого порядка. Из них тривиальным образом вытекают оценки скорости сходимости с тем же порядком относительно  $h$  в более слабых нормах, например, в норме  $L_2(\Omega)$  для уравнений второго порядка.

Эти оценки, однако, не являются оптимальными. Как установлено в теоремах аппроксимации, в пространстве  $V_h$  можно указать элемент, приближающий решение в соответствующей слабой норме лучше, чем это гарантируется доказанными нами теоремами о сходимости метода конечных элементов.

Оказывается, что при некоторых естественных дополнительных предположениях о гладкости исходных данных задачи можно получить оценки погрешности решения в слабых нормах того же порядка, что и оценки погрешности интерполяции.

Мы продемонстрируем это на примере решения однородной задачи Дирихле для уравнения второго порядка в области с достаточно гладкой границей.

Итак, пусть  $u \in H_0^1(\Omega)$  и удовлетворяет интегральному тождеству

$$a(u, v) = f(v) \quad \forall v \in H_0^1(\Omega), \quad (1)$$

где формы  $a(\cdot, \cdot)$  и  $f(\cdot)$  определены соотношениями (5), (6), с. 158.

Наряду с задачей (1) будем рассматривать сопряженную задачу, состоящую в отыскании функции  $u^* \in H_0^1(\Omega)$ , удовлетворяющей ин-

тегральному тождеству

$$a(v, u^*) = f(v) \quad \forall v \in H_0^1(\Omega). \quad (2)$$

Будем говорить, что задача (2) регулярна, если она однозначно разрешима при любой правой части  $f \in L_2(\Omega)$ , причем  $u^* \in H^2(\Omega)$  и справедлива оценка  $\|u^*\|_{2,\Omega} \leq c\|f\|_{0,\Omega}$ .

**ЗАМЕЧАНИЕ 1.** Тождество (2) соответствует обобщенной постановке однородной задачи Дирихле для эллиптического уравнения с оператором, сопряженным оператору, порождаемому билинейной формой  $a(\cdot, \cdot)$ :

$$-\sum_{i,j=1}^2 \frac{\partial}{\partial x_i} \left( a_{ji} \frac{\partial u}{\partial x_j} \right) - \sum_{i=1}^n \frac{\partial (a_i u)}{\partial x_i} + a_0 u = f, \quad x \in \Omega.$$

Условия регулярности налагают определенные ограничения на гладкость коэффициентов уравнения и границы области  $\Omega$  (см. теорему 4, с. 72).

Пусть  $V_h^0$  — конечноэлементное пространство, построенное при помощи лагранжевых треугольных элементов (приграничные элементы, вообще говоря, криволинейны). Пусть далее  $u_h$  — приближенное решение задачи (1), т. е.  $u_h \in V_h^0$  и

$$a(u_h, v_h) = f(v_h) \quad \forall v_h \in V_h^0. \quad (3)$$

**Теорема 1.** Пусть выполнены условия теоремы 1, с. 64. Задача (2) регулярна, триангуляция, использованная при построении пространства  $V_h$ , является 1-регулярной,  $\hat{P} \supset P_1$ . Тогда для разности решений задач (1) и (3) справедлива оценка:

$$\|u - u_h\|_{0,\Omega} \leq ch \|u - u_h\|_{1,\Omega}. \quad (4)$$

**ДОКАЗАТЕЛЬСТВО.** Поскольку  $u - u_h \in L_2(\Omega)$ , можно определить функцию  $u^* \in H_0^1$  как решение задачи

$$a(v, u^*) = (u - u_h, v)_{0,\Omega} \equiv \int_{\Omega} (u - u_h)v \, dx \quad \forall v \in H_0^1. \quad (5)$$

Вследствие регулярности этой задачи имеем

$$\|u^*\|_{2,\Omega} \leq c \|u - u_h\|_{0,\Omega}. \quad (6)$$

Полагая  $v = u - u_h$  в равенстве (5), получим

$$a(u - u_h, u^*) = \|u - u_h\|_{0,\Omega}^2.$$

Заметим теперь, что

$$a(u - u_h, u^*) = a(u - u_h, u^* - v_h) \quad \forall v_h \in V_h^0,$$

следовательно,

$$\|u - u_h\|_{0,\Omega}^2 \leq c \|u - u_h\|_{1,\Omega} \|u^* - v_h\|_{1,\Omega}.$$

Выбирая в качестве  $v_h$  интерполянт функции  $u^*$  и используя оценки погрешности из теоремы 6, с. 154, получим:

$$\|u - u_h\|_{0,\Omega}^2 \leq ch \|u - u_h\|_{1,\Omega} \|u^*\|_{2,\Omega},$$

откуда вследствие неравенства (6) вытекает (4).  $\square$

**Следствие 1.** *Предположим дополнительно к условиям теоремы 1, что триангуляция  $(m+1)$ -регулярна,  $\hat{P} \supset P_m$ ,  $u \in H^{m+1}(\Omega)$ ,  $m \geq 1$ . Тогда*

$$\|u - u_h\|_{0,\Omega} \leq ch^{m+1} \|u\|_{m+1,\Omega}. \quad (7)$$

Оценка (7) согласуется с полученной ранее оценкой погрешности интерполяции в норме пространства  $L_2(\Omega)$ .

**3. Об асимптотической неулучшаемости оценок погрешности метода конечных элементов.** Оценки точности метода конечных элементов, полученные в настоящем параграфе, асимптотически неулучшаемы на классе задач, в том смысле, что можно указать такие коэффициенты, правые части рассматриваемого уравнения и граничных условий, что для погрешности приближенного решения справедлива оценка снизу того же порядка.

Проведем доказательство этого утверждения применительно к задаче Дирихле для эллиптического уравнения второго порядка. При этом будет использована следующая элементарная

**Лемма 1.** *Пусть  $\Omega$  — ограниченная область в пространстве  $R^n$ ,  $p_{k+1} \in P_{k+1}$  и  $p_{k+1} \notin P_k$ . Тогда*

$$\inf_{q_k \in P_k} \|p_{k+1} - q_k\|_{L_2(\Omega)} = c_0 > 0.$$

**ДОКАЗАТЕЛЬСТВО.** Множество всех полиномов  $P_k$  степени не выше  $k$  по совокупности переменных — линейное замкнутое подпространство пространства  $L_2(\Omega)$ , поэтому существует полином  $q_k^0 \in P_k$  такой, что

$$\inf_{q_k \in P_k} \|p_{k+1} - q_k\|_{L_2(\Omega)} = \|p_{k+1} - q_k^0\|_{L_2(\Omega)}.$$

Ясно, что  $\|p_{k+1} - q_k^0\|_{L_2(\Omega)} \neq 0$ , иначе  $p_{k+1} = q_k^0 \in P_k$ .  $\square$

**Теорема 2.** *Пусть  $\Omega$  — многоугольник в  $R^2$ ,  $V_h$  — пространство аффинно-эквивалентных треугольных лагранжевых конечных*

элементов, построенное с использованием регулярной триангуляции области  $\Omega$ ,  $\hat{P} = P_m$ ,  $m \geq 1$ . Тогда существует  $f \in C_0^\infty(\Omega)$  такая, что для обобщенного решения задачи Дирихле

$$-\Delta u = f, \quad x \in \Omega, \quad (8)$$

$$u(x) = 0, \quad x \in \Gamma, \quad (9)$$

для всех достаточно малых  $h$  справедливы оценки

$$c_1 h^m \leq |u - u_h|_{1,\Omega} \leq c_2 h^m.$$

Здесь  $u_h$  — приближенное решение по методу конечных элементов, т. е.  $u_h \in V_h^0$  и

$$\int_{\Omega} \nabla u_h \cdot \nabla v_h dx = \int_{\Omega} f v_h \quad \forall v_h \in V_h^0.$$

**ДОКАЗАТЕЛЬСТВО.** Пусть  $B_R$  — круг радиуса  $R > 0$  такой, что  $\bar{B}_R \subset \Omega$ . Фиксируем функцию  $\varphi \in C_0^\infty(\Omega)$  такую, что  $\varphi(x) = 1$ ,  $x \in B_R$ . Положим  $u(x) = x_1^{m+1} \varphi(x) \in C_0^\infty(\Omega)$ . Ясно, что  $u$  — обобщенное решение задачи (8), (9) при  $f(x) = -\Delta(x_1^{m+1} \varphi(x)) \in C_0^\infty(\Omega)$ . В рассматриваемом случае метод конечных элементов можно трактовать как метод Рунта, поэтому

$$|u - u_h|_{1,\Omega} = \inf_{v_h \in V_h^0} |u - v_h|_{1,\Omega},$$

и оценка сверху для  $|u - u_h|_{1,\Omega}$  следует из теоремы 3, с. 160, в которой можно выбрать максимально возможное  $l = m$ .

Получим оценку снизу. Рассмотрим произвольный треугольник  $e \in \mathcal{T}_h$ , принадлежащий  $B_R$ . При достаточно малом  $h$  вследствие регулярности триангуляции множество таких треугольников не пусто. Более того, можно считать, что  $\text{mes} \left( \bigcup_{e \in B_R} e \right)$  ограничена снизу постоянной, не зависящей от  $h$ . Для любой функции  $v_h \in V_h^0$

$$|u - v_h|_{1,e}^2 = \int_e \left( (m+1)x_1^m - \frac{\partial v_h}{\partial x_1} \right)^2 dx \equiv I_e.$$

Выполним в интеграле  $I_e$  замену переменных, полагая  $x_1 = a_1^e + t_1 h$ ,  $x_2 = a_2^e + t_2 h$ , где  $a_1^e$ ,  $a_2^e$  — координаты центра окружности, вписанной в треугольник  $e$ . При такой замене переменных треугольник  $e$

переходит в подобный треугольник  $\tilde{e}$ , центр вписанной окружности которого лежит в начале координат плоскости  $t_1 t_2$ , причем

$$I_e = h^{2m+2} \int_{\tilde{e}} ((m+1)t_1^m - q_{m-1}(t))^2 dt,$$

где  $q_{m-1}$  — полином степени не выше  $m-1$  по совокупности переменных. Вследствие регулярности триангуляции, очевидно, существует  $r_0 > 0$ , не зависящее от  $h$  и выбора треугольника  $e \in \mathcal{T}_h$  такое, что круг  $B_{r_0}(0)$  радиуса  $r_0$  с центром в начале координат принадлежит  $\tilde{e}$ . Отсюда и из леммы 1 вытекает, что

$$I_e \geq h^{2m+2} \int_{B_{r_0}(0)} ((m+1)t_1^m - q_{m-1}(t))^2 dt \geq ch^{2m+2}.$$

Вновь используя регулярность триангуляции, можем написать, что  $\text{mes}(e) \leq ch^2 \forall e \in \mathcal{T}_h$ , поэтому

$$|u - v_h|_{1,\Omega}^2 \geq ch^{2m} \sum_{e \in B_R} h^2 \geq ch^{2m} \sum_{e \in B_R} \text{mes}(e) \geq ch^{2m}. \quad \square$$

#### 4. О методе конечных элементов для систем уравнений.

Остановимся вкратце на построении и исследовании схем метода конечных элементов для двумерных задачи теории упругости (см. с. 71). Будем рассматривать однородную первую краевую задачу, состоящую в отыскании вектор-функции  $u = (u_1, u_2) \in V = H_0^1(\Omega) \times H_0^1(\Omega)$ , удовлетворяющей интегральному тождеству

$$a(u, v) = f(v) \quad \forall v \in V, \quad (10)$$

где

$$a(u, v) = \int_{\Omega} \sum_{i,j=1}^2 \sigma_{ij}(u) \varepsilon_{ij}(\eta) dx, \quad \sigma_{ij} = \sigma_{ij}(u) = \sum_{k,l=1}^2 a_{ijkl} \varepsilon_{kl},$$

$$\varepsilon_{ij} = \varepsilon_{ij}(u) = \frac{1}{2} \left( \frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} \right), \quad i, j = 1, 2, \quad f = (f_1, f_2) \in (L_2(\Omega))^2.$$

Относительно элементов  $a_{ijkl}$  матрицы Гука предполагаются выполненными условия симметрии и положительной определенности, так что и билинейная форма  $a(\cdot, \cdot)$  оказывается симметричной и положительно определенной.

Введем в рассмотрение пространство  $\mathbf{V}_h^0 = V_h^0 \times V_h^0$ , где  $V_h^0$  — какая-либо из рассмотренных ранее конечноэлементных аппроксимаций пространства  $H_0^1$ , и определим приближенное решение задачи (10) как вектор-функцию  $u_h \in \mathbf{V}_h^0$ , удовлетворяющую интегральному тождеству

$$a(u_h, v_h) = f(v_h) \quad \forall v_h \in \mathbf{V}_h^0. \quad (11)$$

Вследствие симметрии и положительной определенности билинейной формы  $a(\cdot, \cdot)$  задача (11) имеет единственное решение при любой правой части  $f \in (L_2(\Omega))^2$ . При этом

$$\|u - u_h\|_V \leq c \|u - v_h\|_V \quad \forall v_h \in \mathbf{V}_h,$$

откуда вытекают оценки точности, такие же, как для эллиптических уравнений второго порядка (см. §1).

**ЗАМЕЧАНИЕ 2.** При построении метода конечных элементов для системы уравнений теории упругости можно было бы использовать пространство  $\mathbf{V}_h^0 = V_{h,1}^0 \times V_{h,2}^0$ , где  $V_{h,1}^0, V_{h,2}^0$  — различные конечноэлементные аппроксимации пространства  $H_1^0$ . Такой подход редко используется при решении задач теории упругости, однако является типичным при решении многих систем уравнений механики сплошной среды, например, уравнений теории вязкой жидкости, теории пластин и оболочек, где различные компоненты искомого решения по самой постановке задачи принадлежат различным функциональным пространствам.

**5. О точности аппроксимации области.** Рассмотрим однородную задачу Дирихле для уравнения Пуассона в произвольной области  $\Omega$  с границей  $\Gamma \in C^\infty$ . Обобщенная формулировка этой задачи имеет вид

$$u \in V^0 \equiv H_0^1(\Omega) : \int_{\Omega} \nabla u \cdot \nabla v \, dx = \int_{\Omega} f v \, dx \quad \forall v \in V^0. \quad (12)$$

Схемы МКЭ на основе криволинейных элементов для решения этой задачи были изучены нами в предыдущем параграфе. Учитывая большую сложность криволинейных элементов по сравнению с элементами с прямолинейными сторонами, естественно поставить вопрос: нельзя ли использовать аффинно-эквивалентные элементы для решения этой задачи? Ответим на этот вопрос, предполагая, для упрощения рассуждений, что область  $\Omega$  является выпуклой,  $f \in C^\infty(\bar{\Omega})$  и, следовательно,  $u \in C^\infty(\bar{\Omega})$ .

Пусть  $\mathcal{T}_h$  — регулярная триангуляция  $\Omega$ . Будем предполагать, что все элементы  $\mathcal{T}_h$  — треугольники с прямолинейными сторонами,

$$\Omega_h = \bigcup_{e \in \mathcal{T}_h} e \subset \bar{\Omega},$$

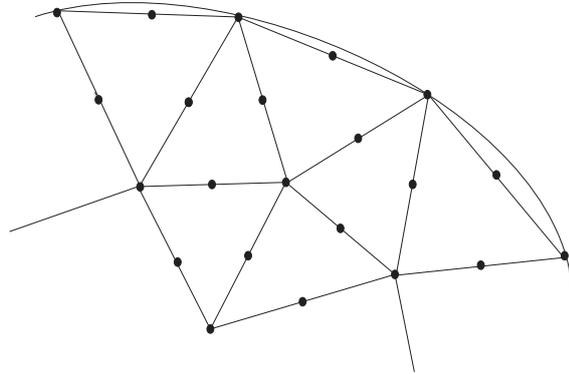


Рис. 1. Аппроксимация  $\Omega$  многоугольной областью  $\Omega_h$ :  $\text{mes}(\Omega \setminus \Omega_h) = O(h^2)$ .

$\Gamma_h = \partial\Omega_h$ , все вершины треугольников, принадлежащие  $\Gamma_h$ , лежат на  $\Gamma$  (см. рис. 1, соответствующий случаю  $m = 2$ ). Пусть далее  $V_h \subset H^1(\Omega_h)$  — пространство аффинно-эквивалентных лагранжевых треугольных конечных элементов степени  $m \geq 1$ ,  $V_h^0 \subset H_0^1(\Omega_h)$  — подпространство  $V_h$ . Схему МКЭ определим тождеством

$$u_h \in V_h^0 : \int_{\Omega_h} \nabla u_h \cdot \nabla v_h \, dx = \int_{\Omega_h} f v_h \, dx \quad \forall v_h \in V_h^0. \quad (13)$$

**Лемма 2.** Если  $u$  — решение задачи (12),  $u_h$  — решение задачи (13), то

$$|u - u_h|_{1, \Omega_h} = \inf_{v_h \in V_h^0} |u - v_h|_{1, \Omega_h}. \quad (14)$$

**ДОКАЗАТЕЛЬСТВО.** Продолжим функции из  $V_h^0$  нулем в приграничную полосу  $\Omega \setminus \Omega_h$ , т. е. введем в рассмотрение пространство

$$\tilde{V}_h^0 = \{\tilde{v}_h : \tilde{v}_h = v_h \in V_h^0 \text{ на } \Omega_h, \tilde{v}_h \equiv 0 \text{ на } \Omega \setminus \Omega_h\}.$$

Очевидно, что  $\tilde{V}_h^0 \subset V^0$ , а схема МКЭ (13) эквивалентна задаче

$$\tilde{u}_h \in \tilde{V}_h^0 : \int_{\Omega} \nabla \tilde{u}_h \cdot \nabla v_h \, dx = \int_{\Omega} f v_h \, dx \quad \forall v_h \in \tilde{V}_h^0.$$

Отсюда следует, что  $\tilde{u}_h$  является приближением к  $u$  по методу Галеркина, или в силу симметрии оператора исходной задачи — по методу Ритца (см. с. 22). Поэтому

$$|u - \tilde{u}_h|_{1, \Omega} = \inf_{v_h \in \tilde{V}_h^0} |u - v_h|_{1, \Omega},$$

откуда непосредственно следует утверждение леммы.  $\square$

**Лемма 3.** *Справедлива оценка*

$$|u - u_h|_{1, \Omega_h} \leq c h^{3/2}. \quad (15)$$

**ДОКАЗАТЕЛЬСТВО.** Пусть  $u_I$  есть  $V_h$ -интерполянт решения  $u$ . Поскольку  $u = 0$  на  $\Gamma_h$  только в вершинах элементов, то  $u_I$  не принадлежит, вообще говоря,  $V_h^0$  и мы не можем из (14) получить неравенство  $|u - u_h|_{1, \Omega_h} \leq |u - u_I|_{1, \Omega_h}$  и непосредственно воспользоваться оценкой погрешности интерполяции. Поэтому поступим следующим образом.

Обозначим через  $\gamma_h$  множество всех узлов интерполяции, лежащих на  $\Gamma_h$ , а через  $\omega_h^0$  — множество всех остальных узлов. Имеем  $u_I(x) = \bar{u}_I + \tilde{u}_I$ , где

$$\bar{u}_I = \sum_{a_i \in \omega_h^0} u(a_i) \varphi_i(x), \quad \tilde{u}_I = \sum_{a_i \in \gamma_h} u(a_i) \varphi_i(x).$$

По определению  $\bar{u}_I \in V_h^0$ . Носитель  $\tilde{u}_I$  (обозначим его через  $\Omega_\Gamma$ ) состоит из объединения  $O(1/h)$  тех приграничных элементов, которые имеют сторону, принадлежащую  $\Gamma_h$ . Из (14) следует, что

$$|u - u_h|_{1, \Omega_h} \leq |u - \bar{u}_I|_{1, \Omega_h} \leq |u - u_I|_{1, \Omega_h} + |u_I - \bar{u}_I|_{1, \Omega_h}. \quad (16)$$

Поскольку  $|u - u_I|_{1, \Omega_h} \leq c h^m |u|_{m+1, \Omega_h}$ , то остается оценить второе слагаемое в правой части (16). Имеем

$$\begin{aligned} |u_I - \bar{u}_I|_{1, \Omega_h}^2 &= |\tilde{u}_I|_{1, \Omega_h}^2 = \sum_{e \in \Omega_\Gamma} \int_e \left| \sum_{a_i \in \gamma_h \cap e} u(a_i) \nabla \varphi_i \right|^2 dx \leq \\ &\leq (m-1) \max_{a_i \in \gamma_h} |u(a_i)|^2 \sum_{e \in \Omega_\Gamma} \sum_{a_i \in \gamma_h \cap e} \int_e |\nabla \varphi_i|^2 dx \leq \frac{c}{h} \max_{a_i \in \gamma_h} |u(a_i)|^2. \end{aligned} \quad (17)$$

Здесь мы учли оценку  $\int_e |\nabla \varphi_i|^2 dx \leq c$ , которая легко устанавливается переходом на базисный элемент. Обозначим через  $x_i \in \Gamma$  точку, ближайшую к  $a_i \in \gamma_h$ . Учитывая, что  $|a_i - x_i| = O(h^2)$ , получим

$$|u(a_i)| = \left| \int_0^1 Du(x_i + t(a_i - x_i))(a_i - x_i) dt \right| \leq c h^2 \max_{x \in \Omega} |\nabla u(x)| \leq c h^2.$$

Подставляя эту оценку в (17) и используя затем (16), получим оценку (15).  $\square$

Оценка (15) неулучшаема по порядку  $h$ . Не останавливаясь на доказательстве этого факта, отметим, что отсюда следует нецелесообразность использования в случае произвольной области аффинно-эквивалентных элементов степени  $m \geq 3$ . В следующих параграфах мы покажем, что схемы, построенные на основе изопараметрических элементов, являются наиболее подходящими в этой ситуации: они существенно проще схем, использующих криволинейные элементы, при той же оценке точности. Одновременно мы проанализируем влияние численного интегрирования на точность схем метода конечных элементов.

#### § 4. Численное интегрирование в методе конечных элементов. Общие построения

В §1 мы рассмотрели различные схемы МКЭ для эллиптического уравнения второго порядка. Каждая из схем в конечном итоге сводила исходную задачу к решению системы линейных алгебраических уравнений. При этом мы предполагали, что элементы матрицы системы, определяемые формулами

$$A_{kl} = a(\varphi_l, \varphi_k) \equiv \int_{\Omega_h} \left( \sum_{i,j=1}^2 a_{ij} \frac{\partial \varphi_l}{\partial x_j} \frac{\partial \varphi_k}{\partial x_i} + \sum_{i=1}^2 a_i \frac{\partial \varphi_l}{\partial x_i} \varphi_k + a_0 \varphi_l \varphi_k \right) dx,$$

а также компоненты вектора правой части

$$F_l = f(\varphi_l) \equiv \int_{\Omega_h} f \varphi_l dx$$

вычисляются точно. Это требование трудно выполнить, даже если функции  $a_{ij}$ ,  $a_i$  и  $f$  постоянны.

Поясним сказанное подробнее. Положим, как и ранее,  $A = \{a_{ij}\}_{i,j=1}^2$  и пусть  $a_i \equiv 0$ ,  $i = 0, 1, 2$ . Введем для сокращения записей обозначения:  $p = \varphi_l$ ,  $q = \varphi_k$ , и будем считать, что область  $\Omega$  не является многоугольником. Тогда  $A_{kl}$  является суммой интегралов

$$I_e = \int_e A \nabla p \cdot \nabla q dx$$

по всем  $e \in \mathcal{T}_h$ . Среди элементов  $e \in \mathcal{T}_h$  обязательно найдется криволинейный (или изопараметрический) элемент. Пусть для определенности  $e$  — треугольный изопараметрический элемент степени  $m$ .

О функциях  $p$  и  $q$  нам известно лишь следующее:  $p(x) = \hat{p}(x_e^{-1}(x))$ ,  $q(x) = \hat{q}(x_e^{-1}(x))$ ,  $\hat{p}, \hat{q} \in P_m$ , где  $\hat{p}, \hat{q}$  — базисные функции элемента  $\hat{e}$ . Обратим внимание на то, что хотя мы имеем явные выражения для  $\hat{p}, \hat{q}$ , нам не известно в явном виде обратное отображение  $\hat{x} = x_e^{-1}(x)$ , следовательно, мы не имеем явных формул для вычисления функций  $p$  и  $q$ .

Выходом из создавшегося положения является замена переменных  $x = x_e(\hat{x})$  в интеграле  $I_e$ . Пусть  $x = x_e(\hat{x}) = (x_1(\hat{x}), x_2(\hat{x}))^T$  — преобразование  $\hat{e}$  на  $e$ ,  $J_e(\hat{x}) = Dx_e(\hat{x})$  — матрица Якоби,  $J_e^{-T}(\hat{x}) = (J_e^{-1}(\hat{x}))^T$ . Отметим, что

$$J_e^{-1}(\hat{x}) = \frac{1}{\det(J_e(\hat{x}))} \begin{pmatrix} \frac{\partial x_2(\hat{x})}{\partial \hat{x}_2} & -\frac{\partial x_1(\hat{x})}{\partial \hat{x}_2} \\ -\frac{\partial x_2(\hat{x})}{\partial \hat{x}_1} & \frac{\partial x_1(\hat{x})}{\partial \hat{x}_1} \end{pmatrix}.$$

Будем предполагать, что  $\det(J_e(\hat{x})) > 0$  для всех  $\hat{x} \in \hat{e}$ . При замене переменных выполняются следующие преобразования

$$e \rightarrow \hat{e}, \quad dx = \det(J_e(\hat{x})) d\hat{x}, \quad A(x) = \hat{A}(\hat{x}),$$

$$\nabla p = J_e^{-T}(\hat{x}) \hat{\nabla} \hat{p}, \quad \hat{\nabla} \hat{p} = \left( \frac{\partial \hat{p}}{\partial \hat{x}_1}, \frac{\partial \hat{p}}{\partial \hat{x}_2} \right)^T.$$

Таким образом,

$$I_e = \int_{\hat{e}} \det(J_e(\hat{x})) \hat{A} J_e^{-T}(\hat{x}) \hat{\nabla} \hat{p} \cdot J_e^{-T}(\hat{x}) \hat{\nabla} \hat{q} d\hat{x}. \quad (1)$$

Нетрудно заметить, что под интегралом в равенстве (1) стоит сумма слагаемых вида

$$\hat{b}(\hat{x}) \frac{\partial \hat{p}}{\partial \hat{x}_i} \frac{\partial \hat{q}}{\partial \hat{x}_j}, \quad \hat{b}(\hat{x}) = \frac{1}{\det(J_e(\hat{x}))} \frac{\partial x_\alpha(\hat{x})}{\partial \hat{x}_k} \frac{\partial x_\beta(\hat{x})}{\partial \hat{x}_l} \hat{a}_{st}.$$

Например, при постоянных коэффициентах  $a_{ij}$  имеем

$$\hat{b}(\hat{x}) = \frac{\text{полином степени } 2m - 2}{\text{полином степени } 2m - 2}, \quad \frac{\partial \hat{p}}{\partial \hat{x}_i} \frac{\partial \hat{q}}{\partial \hat{x}_j} \in P_{2m-2}.$$

Таким образом, задача вычисления элементов матрицы системы МКЭ в рассматриваемой ситуации сводится к интегрированию дробно-рациональных функций по треугольной области.

Ясно, что в случае переменных коэффициентов  $a_{ij}$  и использования криволинейных элементов эта задача может быть решена только

приближенно, с использованием формул численного интегрирования, называемых квадратурными (кубатурными) формулами.

Опишем общую схему применения квадратурных формул в МКЭ. Для приближенного вычисления интеграла  $I_e$  определим сначала некоторую квадратурную формулу на базисном элементе  $\hat{e}$ :

$$\int_{\hat{e}} \hat{\phi}(\hat{x}) d\hat{x} \approx \sum_{l=1}^L \hat{c}_l \hat{\phi}(\hat{b}_l) \equiv \hat{S}(\hat{\phi}). \quad (2)$$

Числа  $\hat{c}_l$  и точки  $\hat{b}_l \in \hat{e}$  называются соответственно коэффициентами и узлами квадратурной формулы  $\hat{S}$ . Преобразование переменных  $x = x_e(\hat{x})$  по этой формуле нетрудно определить квадратуру  $S_e$  для произвольного элемента  $e \in \mathcal{T}_h$ . Действительно,

$$\begin{aligned} \int_e \phi(x) dx &= \int_{\hat{e}} \det(J_e(\hat{x})) \phi(x_e(\hat{x})) d\hat{x} \approx \\ &\approx \sum_{l=1}^L \hat{c}_l \det(J_e(\hat{b}_l)) \phi(x_e(\hat{b}_l)) = \sum_{l=1}^L c_l^e \phi(b_l^e) \equiv S_e(\phi), \end{aligned}$$

где  $c_l^e = \det(J_e(\hat{b}_l)) \hat{c}_l$ ,  $b_l^e = x_e(\hat{b}_l)$ . Отметим, что если  $\hat{\phi}(\hat{x}) = \phi(x_e(\hat{x}))$ , то

$$S_e(\phi) = \hat{S}(\det(J_e)\hat{\phi}). \quad (3)$$

Таким образом,

$$I_e \approx S_e(A\nabla p \cdot \nabla q),$$

или, после перехода на базисный элемент,

$$\begin{aligned} I_e &\approx \hat{S}(\det(J_e) \hat{A} J_e^{-T} \hat{\nabla} \hat{p} \cdot J_e^{-T} \hat{\nabla} \hat{q}) = \\ &= \sum_{l=1}^L \hat{c}_l \det(J_e(\hat{b}_l)) (\hat{A} J_e^{-T} \hat{\nabla} \hat{p} \cdot J_e^{-T} \hat{\nabla} \hat{q})(\hat{b}_l). \end{aligned}$$

Полученная приближенная формула легко реализуется: все, что необходимо для этого при заданной квадратуре  $\hat{S}$ , это уметь вычислять  $x_e(\hat{x})$  и  $J_e(\hat{x})$  в узлах квадратуры  $\hat{x} = \hat{b}_l$ . Последнее в методе конечных элементов всегда гарантируется.

В результате вычислений элементов матрицы и правой части системы алгебраических уравнений МКЭ по приближенным формулам мы приходим к системе уравнений, отличной от системы уравнений

МКЭ. Таким образом, использование квадратурных формул приводит к новым конечномерным задачам. Эти задачи мы будем называть схемами МКЭ с численным интегрированием.

При использовании численного интегрирования в МКЭ квадратурные формулы естественно выбирать так, чтобы:

1) сохранялись условия разрешимости исходной задачи; например, если исходная задача была однозначно разрешима, то схема с численным интегрированием также должна быть однозначно разрешима; если исходная задача имела множество решений вида  $\text{const} + u(x)$ , с единственным  $u(x)$  (как в случае задачи Неймана), то аналогичное множество решений должна иметь и схема с численным интегрированием и т.д.;

2) сохранялся порядок оценки точности, которым обладала схема МКЭ.

Мы будем рассматривать только случай однозначно разрешимых схем и исследуем вначале обозначенные проблемы для схем МКЭ с аффинно-эквивалентными элементами, а затем приведем обобщение на криволинейные и изопараметрические элементы. В последнем случае дополнительную трудность составляет аппроксимация границы области. Во всех случаях будем стремиться трактовать изучаемые схемы как возмущенный метод Галеркина и для оценки их точности использовать лемму Стрэнга (см. с. 27). Будем разделять также исследование схем с треугольными и четырехугольными элементами: изложение материала при исследовании схем с треугольными элементами будет проведено достаточно подробно, в случае четырехугольных элементов рассуждения могут быть проведены по аналогии и потому более кратко.

При обсуждении основных вопросов, возникающих при использовании квадратурных формул в методе конечных элементов, ограничимся рассмотрением следующей двумерной задачи Дирихле:

$$\begin{aligned} - \sum_{i,j=1}^2 \frac{\partial}{\partial x_i} (a_{ij} \frac{\partial u}{\partial x_j}) &= f(x), \quad x \in \Omega, \\ u &= 0, \quad x \in \Gamma. \end{aligned}$$

Будем предполагать, что функции  $a_{ij} \in L_\infty(\Omega)$ ,  $f \in L_2(\Omega)$ , всюду определены на  $\bar{\Omega}$ , и

$$\sum_{i,j=1}^2 a_{ij}(x) \xi_j \xi_i \geq c |\xi|^2 \quad \forall \xi \in R^2, \quad \forall x \in \bar{\Omega}. \quad (4)$$

При указанных условиях существует единственное обобщенное решение  $u \in H_0^1(\Omega)$  этой краевой задачи. Функция  $u$  удовлетворяет интегральному тождеству

$$a(u, v) \equiv \int_{\Omega} \sum_{i,j=1}^2 a_{ij} \frac{\partial u}{\partial x_j} \frac{\partial v}{\partial x_i} dx = \int_{\Omega} f v dx \equiv f(v) \quad \forall v \in H_0^1(\Omega). \quad (5)$$

При определении пространств конечных элементов и квадратурных формул будем предполагать для заданного  $m \geq 1$ , что:

- 1) все конечные элементы  $(e, P_e, \Sigma_e)$  ассоциированы с одним базисным конечным элементом  $(\hat{e}, \hat{P}, \hat{\Sigma})$ ,  $\hat{P} \supseteq P_m$ ;
- 2) семейство триангуляций  $\mathcal{T}_h$  является  $(m+1)$ -регулярным;
- 3) все квадратурные формулы  $S_e$  индуцированы одной квадратурной формулой  $\hat{S}$  на базисном элементе.

## § 5. Схемы с численным интегрированием для многоугольных областей

**1. Аффинно-эквивалентные треугольные элементы.** Рассмотрим случай многоугольной области  $\Omega$ . Построим конечноэлементную схему, используя лагранжевы или эрмитовы аффинно-эквивалентные треугольные элементы.

Пусть  $\mathcal{T}_h$  — точная регулярная триангуляция области  $\bar{\Omega}$ , т. е.  $\bigcup\{e, e \in \mathcal{T}_h\} = \bar{\Omega}$ ,  $h$  — максимальный из диаметров элементов  $e \in \mathcal{T}_h$ . Пусть, далее,  $(\hat{e}, \hat{P}, \hat{\Sigma})$ ,  $\hat{P} = P_m$ , — некоторый базисный треугольный элемент класса  $C^0$  или  $C^1$  и  $\mathcal{E}_h = \{(e, P_e, \Sigma_e), e \in \mathcal{T}_h\}$  — семейство ассоциированных с ним треугольных аффинно-эквивалентных элементов,  $x = x_e(\hat{x}) = B_e \hat{x} + b_e$  — аффинное отображение  $\hat{e}$  на  $e$ , причем  $\det(B_e) > 0$ . Предположение  $\det(B_e) > 0$  означает, что выбрана подходящая нумерация вершин  $e$  при определении отображения  $x_e$ . Регулярность триангуляции обеспечивает выполнение оценок  $\|B_e\| \leq ch$ ,  $\|B_e^{-1}\| \leq ch^{-1}$  для всех  $e \in \mathcal{T}_h$ .

Пусть  $V_h \subset V = H_0^1(\Omega)$  — соответствующее  $\mathcal{E}_h$  пространство конечных элементов:

$$V_h = \{v_h \in C(\bar{\Omega}) : v_h|_e \in P_m \quad \forall e \in \mathcal{T}_h, v_h(x) = 0, x \in \Gamma\}.$$

Решение схемы МКЭ есть функция  $u_h \in V_h$  такая, что

$$a(u_h, v_h) = f(v_h) \quad \forall v_h \in V_h. \quad (1)$$

Условия на коэффициенты дифференциального уравнения обеспечивают однозначную разрешимость задачи (1).

Выберем некоторую формулу численного интегрирования  $\hat{S}$  на базисном конечном элементе  $\hat{e}$  (см. (2), с. 178). Поскольку  $\det(Dx_e) = \det(B_e)$ , то

$$S_e(\phi) = \sum_{l=1}^L c_l^e \phi(b_l^e) = \det(B_e) \sum_{l=1}^L \hat{c}_l \phi(x_e(\hat{b}_l)),$$

где  $c_l^e = \det(B_e)\hat{c}_l$ ,  $b_l^e = B_e\hat{b}_l + b_e \in e$ , и формула (3), с. 178, принимает вид

$$S_e(\phi) = \det(B_e)\hat{S}(\hat{\phi}),$$

где  $\det(B_e) = 2 \text{mes}(e)$ .

Используя квадратуры  $S_e$ , построим составную квадратурную формулу <sup>1)</sup>:

$$\int_{\Omega} \phi(x) dx = \sum_e \int_e \phi(x) dx \approx \sum_e S_e(\phi) \equiv S_{\Omega}(\phi).$$

Рассмотрим теперь конечноэлементную схему (1) и заменим все входящие в нее интегралы квадратурной формулой  $S_{\Omega}$ . В итоге получим схему МКЭ с численным интегрированием:

$$\text{найти } u_h \in V_h : \quad a_h(u_h, v_h) = f_h(v_h) \quad \forall v_h \in V_h. \quad (2)$$

Здесь

$$\begin{aligned} a_h(u, v) &= S_{\Omega} \left( \sum_{i,j=1}^2 a_{ij} \frac{\partial u}{\partial x_j} \frac{\partial v}{\partial x_i} \right) = \\ &= \sum_e \det(B_e) \sum_{l=1}^L \hat{c}_l \left( \sum_{i,j=1}^2 a_{ij} \frac{\partial u}{\partial x_j} \frac{\partial v}{\partial x_i} \right) (b_l^e), \quad (3) \end{aligned}$$

$$f_h(v) = S_{\Omega}(fv) = \sum_e \det(B_e) \sum_{l=1}^L \hat{c}_l (fv)(b_l^e). \quad (4)$$

Исследуем разрешимость и точность полученной схемы. Для этого понадобятся дополнительные понятия и утверждения.

---

<sup>1)</sup>Напомним, что  $\sum_e$  означает суммирование по всем  $e \in \mathcal{T}_h$ .

**Лемма 1.** Пусть  $A, B$  — симметричные неотрицательные матрицы порядка  $m$ ,  $a(x, x) = Ax \cdot x$ ,  $b(x, x) = Bx \cdot x$  — соответствующие им квадратичные формы;

$$X = \ker(b) = \{x \in R^m : b(x, x) = 0\} \subseteq \ker(a).$$

Тогда существует такая положительная постоянная  $c_{ab}$ , что

$$a(x, x) \leq c_{ab} b(x, x) \quad \forall x \in R^m. \quad (5)$$

Если  $\ker(b) = \ker(a)$ , то формы  $a$  и  $b$  эквивалентны:

$$c_{ba} b(x, x) \leq a(x, x) \leq c_{ab} b(x, x) \quad \forall x \in R^m. \quad (6)$$

**ДОКАЗАТЕЛЬСТВО.** Отметим, прежде всего, что  $X$  — подпространство  $R^m$ , поэтому для любого  $x \in R^m$  справедливо представление  $x = x_0 + x_1$ , где  $x_0 \in X$ ,  $x_1 \in X^\perp = \{x \in R^m : x \cdot y = 0 \quad \forall y \in X\}$ . Если  $x \in X$ , то неравенство  $a(x, x) \leq c_{ab} b(x, x)$  выполняется тривиальным образом. Если  $x \notin X$ , то  $b(x, x) = Bx_0 \cdot x_0 + 2Bx_0 \cdot x_1 + Bx_1 \cdot x_1 = Bx_1 \cdot x_1$ , точно так же  $a(x, x) = Ax_0 \cdot x_0 + 2Ax_0 \cdot x_1 + Ax_1 \cdot x_1 = Ax_1 \cdot x_1$ . Таким образом,  $a(x, x) \leq \|A\| |x_1|^2$ ,  $b(x, x) \geq \lambda_B |x_1|^2$ , где  $\lambda_B$  — минимальное ненулевое собственное число матрицы  $B$ , и неравенство (5) выполнено, причем  $c_{ab} = \|A\|/\lambda_B$ . Неравенства (6) непосредственно следуют из (5).  $\square$

**УПРАЖНЕНИЕ 1.** Доказать, что множество  $X$ , введенное в лемме 1, — подпространство пространства  $R^m$ .

Определим функционалы погрешности квадратурных формул:

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

$$E_\Omega(\phi) = \int_\Omega \phi(x) dx - S_\Omega(\phi) = \sum_e E_e(\phi).$$

Нетрудно видеть, что  $E_e(\phi) = \det(B_e) \hat{E}(\hat{\phi})$ .

Будем говорить, что квадратура  $\hat{S}$  точна на полиномах из  $P_s$ ,  $s \geq 0$ , если  $\hat{E}(\hat{\phi}) = 0$  для любого  $\hat{\phi} \in P_s$ . Легко проверить, что в этом случае

$$\sum_{l=1}^L \hat{c}_l = \text{mes}(\hat{e}). \quad (7)$$

Введем обозначения для дискретных аналогов норм  $|\cdot|_{0,\Omega}$ ,  $|\cdot|_{1,\Omega}$  в пространстве  $V_h$ :

$$|u_h|_{0,h}^2 = S_\Omega(u_h^2), \quad |u_h|_{1,h}^2 = S_\Omega(|\nabla u_h|^2). \quad (8)$$

**Лемма 2.** Пусть квадратурная формула  $\hat{S}$  такова, что

1)  $\hat{c}_l > 0$ ,  $l = 1, \dots, L$ ;

и выполнено одно из следующих двух условий:

2) множество узлов  $\{\hat{b}_l\}_{l=1}^L$  содержит подмножество, унисоль-  
вентное с  $P_{m-1}$ <sup>1)</sup>,

2') квадратурная формула  $\hat{S}$  точна на полиномах из  $P_{2m-2}$ .

Тогда существуют не зависящие от  $h$  положительные постоянные  $c_1$  и  $c_2$  такие, что

$$c_1 |u_h|_{1,\Omega} \leq |u_h|_{1,h} \leq c_2 |u_h|_{1,\Omega}, \quad (9)$$

$$|u_h|_{0,h} \leq c_2 |u_h|_{0,\Omega} \quad (10)$$

для любого  $u_h \in V_h$ .

**ДОКАЗАТЕЛЬСТВО.** Установим сначала справедливость оценок (9). Для этого достаточно проверить, что для любого конечного элемента  $e \in \mathcal{T}_h$  и любой функции  $u_h \in V_h$  справедливы неравенства:

$$c_1^2 \int_e |\nabla u_h(x)|^2 dx \leq S_e(|\nabla u_h|^2) \leq c_2^2 \int_e |\nabla u_h(x)|^2 dx$$

с положительными, не зависящими от  $h$  постоянными  $c_1, c_2$ .

Выполняя замену переменных  $x = B_e \hat{x} + b_e$  и используя обозначение  $\hat{u}(\hat{x}) = u_h(B_e \hat{x} + b_e)$ , получаем (см. (1), с. 177):

$$|\nabla u_h(x)|^2 = B_e^{-T} \hat{\nabla} \hat{u}(\hat{x}) \cdot B_e^{-T} \hat{\nabla} \hat{u}(\hat{x}) = |B_e^{-T} \hat{\nabla} \hat{u}(\hat{x})|^2.$$

Таким образом,

$$\int_e |\nabla u_h(x)|^2 dx = \det(B_e) \int_{\hat{e}} |B_e^{-T} \hat{\nabla} \hat{u}(\hat{x})|^2 d\hat{x},$$

$$S_e(|\nabla u_h|^2) = \det(B_e) \hat{S}(|B_e^{-T} \hat{\nabla} \hat{u}|^2),$$

и, следовательно, достаточно доказать неравенства

$$c_1^2 \int_{\hat{e}} |B_e^{-T} \hat{\nabla} \hat{u}(\hat{x})|^2 d\hat{x} \leq \hat{S}(|B_e^{-T} \hat{\nabla} \hat{u}|^2) \leq c_2^2 \int_{\hat{e}} |B_e^{-T} \hat{\nabla} \hat{u}(\hat{x})|^2 d\hat{x}. \quad (11)$$

Предположим сначала, что существуют не зависящие от  $h$  постоянные  $k_1, k_2$  такие, что

$$k_1 \int_{\hat{e}} |\hat{\nabla} \hat{u}(\hat{x})|^2 d\hat{x} \leq \hat{S}(|\hat{\nabla} \hat{u}|^2) \leq k_2 \int_{\hat{e}} |\hat{\nabla} \hat{u}(\hat{x})|^2 d\hat{x}. \quad (12)$$

<sup>1)</sup>Предполагается, что множество линейных функционалов вида  $p(\hat{b}_l)$  отождествляется с множеством узлов  $\hat{b}_l$ .

Тогда, учитывая равенство  $\|B_e\| = \|B_e^T\|$ <sup>2)</sup> и оценки

$$\|B_e\| \|B_e^{-1}\| \leq c, \quad |y| = |B_e^T B_e^{-T} y| \leq \|B_e^T\| |B_e^{-T} y|,$$

получим неравенства:

$$\begin{aligned} \int_{\hat{e}} |B_e^{-T} \hat{\nabla} \hat{u}(\hat{x})|^2 d\hat{x} &\leq \|B_e^{-T}\|^2 \int_{\hat{e}} |\hat{\nabla} \hat{u}(\hat{x})|^2 d\hat{x} \leq \\ &\leq \|B_e^{-T}\|^2 / k_1 \hat{S}(|\hat{\nabla} \hat{u}|^2) \leq c^2 / k_1 \hat{S}(|B_e^{-T} \hat{\nabla} \hat{u}|^2), \end{aligned}$$

т. е. оценку снизу в (11) с постоянной  $c_1^2 = k_1 / c^2$ . Аналогично доказывается оценка сверху с постоянной  $c_2^2 = k_2 c^2$ . Поскольку  $c$  не зависит от  $h$ , отсюда получаем неравенства (9).

Обратимся теперь к доказательству оценок (12). Имеем  $\hat{u} \in \hat{P} = P_m$ , значит  $\hat{u} = \sum_{i=1}^M \hat{u}_i \hat{\varphi}_i(\hat{x})$ , где  $M = \dim P_m$ ,  $\hat{u}_i = \hat{u}(\hat{a}_i)$ ,  $\hat{\varphi}_i$  — базисные функции Лагранжа. Следовательно,

$$\begin{aligned} \hat{I}_1 &= \int_{\hat{e}} |\hat{\nabla} \hat{u}(\hat{x})|^2 d\hat{x} = \sum_{i,j=1}^M \hat{a}_{ij} \hat{u}_i \hat{u}_j, \quad \text{где } \hat{a}_{ij} = \int_{\hat{e}} \hat{\nabla} \hat{\varphi}_j \cdot \hat{\nabla} \hat{\varphi}_i d\hat{x}, \\ \hat{I}_2 &= \hat{S}(|\hat{\nabla} \hat{u}|^2) = \sum_{i,j=1}^M \hat{b}_{ij} \hat{u}_i \hat{u}_j, \quad \text{где } \hat{b}_{ij} = \hat{S}(\hat{\nabla} \hat{\varphi}_j \cdot \hat{\nabla} \hat{\varphi}_i) \end{aligned}$$

есть неотрицательные квадратичные формы, причем неотрицательность  $\hat{I}_2$  обеспечивается положительностью коэффициентов квадратуры  $\hat{S}$ .

Покажем, что ядра этих квадратичных форм совпадают. Если  $\hat{I}_1 = 0$ , то  $\hat{u} = \text{const}$ , т. е.  $\hat{u}_1 = \dots = \hat{u}_M$ . Вычисляя ядро квадратичной формы  $\hat{I}_2$ , мы должны различать два случая. Если выполнено условие 2'), то  $\hat{I}_2 = \hat{I}_1$ , поскольку  $|\hat{\nabla} \hat{u}|^2 \in P_{2m-2}$ . Если выполнено условие 2) то, используя тот факт, что из равенства  $\hat{I}_2$  нулю вытекает, что  $\hat{\nabla} \hat{u}(\hat{b}_l) = 0$ ,  $l = 1, 2, \dots, L$ , а  $\partial \hat{u} / \partial x_k \in P_{m-1}$ ,  $k = 1, 2$ , получим, что  $\hat{\nabla} \hat{u} \equiv 0$ , следовательно, как и для квадратичной формы  $\hat{I}_1$ , ядро определяется уравнениями  $\hat{u}_1 = \dots = \hat{u}_M$ .

Таким образом, вследствие леммы 1 формы  $\hat{I}_1$  и  $\hat{I}_2$  эквивалентны. Постоянные эквивалентности  $k_1, k_2$  зависят от выбранной квадратуры  $\hat{S}$  и базисного элемента  $(\hat{e}, \hat{P}, \hat{\Sigma})$ , но не зависят от  $h$ , что означает справедливость неравенств (12).

<sup>2)</sup> Напомним, что  $\|B\|^2$  совпадает с максимальным собственным числом матрицы  $B^T B$ .

При доказательстве оценки (10) будем рассуждать аналогично. Из леммы 1 следует, что

$$\hat{S}(\hat{u}^2) \leq c \int_{\hat{e}} \hat{u}^2(\hat{x}) d\hat{x},$$

с постоянной, не зависящей от  $h$ , поскольку квадратичная форма, определяемая интегралом в правой части последнего неравенства, имеет тривиальное ядро. Таким образом,

$$S_e(u_h^2) = \det(B_e) \hat{S}(\hat{u}^2) \leq c \det(B_e) \int_{\hat{e}} \hat{u}^2(\hat{x}) d\hat{x} = c \int_e u_h^2(x) dx \quad \forall e \in \mathcal{T}_h.$$

Суммируя эти неравенства по всем  $e \in \mathcal{T}_h$ , получим (10).  $\square$

Достаточные условия однозначной разрешимости схемы МКЭ с численным интегрированием очевидным образом вытекают из следующей теоремы.

**Теорема 1.** Пусть квадратурная формула  $\hat{S}$  удовлетворяет условиям леммы 2. Тогда билинейная форма  $a_h$  равномерно по  $h$  положительно определена и ограничена на  $V_h \times V_h$ , линейная форма  $f_h$  равномерно непрерывна на  $V_h$ , т. е. существуют положительные, не зависящие от  $h$ , постоянные  $m$ ,  $M$  и  $c$  такие, что

$$a_h(u_h, u_h) \geq m |u_h|_{1,\Omega}^2, \quad |a_h(u_h, v_h)| \leq M |u_h|_{1,\Omega} |v_h|_{1,\Omega},$$

$$|f_h(v_h)| \leq c |v_h|_{1,\Omega}$$

для любых  $u_h, v_h \in V_h$ .

**ДОКАЗАТЕЛЬСТВО.** Из условия (4), с. 179, и леммы 2 следует, что

$$a_h(u_h, u_h) = S_\Omega \left( \sum_{i,j=1}^2 a_{ij} \frac{\partial u}{\partial x_j} \frac{\partial v}{\partial x_i} \right) \geq c S_\Omega (|\nabla u_h|^2) = c |u_h|_{1,h}^2 \geq m |u_h|_{1,\Omega}^2.$$

Далее, пусть  $A(x) = \{a_{ij}(x)\}_{i,j=1}^2$ ,  $\beta = \max_{x \in \bar{\Omega}} \|A(x)\|$ . Используя положительность коэффициентов квадратуры и неравенство Коши — Буняковского, получим (см. далее задачу 3, с. 193)

$$|a_h(u_h, v_h)| = |S_\Omega(A \nabla u_h \cdot \nabla v_h)| \leq \beta S_\Omega(|\nabla u_h| |\nabla v_h|) \leq$$

$$\leq \beta |u_h|_{1,h} |v_h|_{1,h} \leq M |u_h|_{1,\Omega} |v_h|_{1,\Omega}.$$

Наконец, аналогично предыдущему,

$$|f_h(v_h)| = |S_\Omega(fv_h)| \leq |f|_{0,h}|v_h|_{0,h} \leq c|v_h|_{0,\Omega} \leq c|v_h|_{1,\Omega}.$$

В последней оценке использовано неравенство Фридрихса.  $\square$

**ЗАМЕЧАНИЕ 1.** Условие положительности коэффициентов квадратурной формулы не является необходимым для разрешимости сеточной схемы, и в некоторых случаях разрешимость может быть доказана и при использовании квадратурных формул, не удовлетворяющих этому условию (см. далее задачи 2 а) и 4). Однако подавляющее большинство используемых на практике квадратурных формул имеет положительные коэффициенты. Примеры квадратурных формул будут приведены ниже.

Итак, исходная задача (5) аппроксимирована дискретной задачей (2), найдены условия, при которых последняя однозначно разрешима.

Перейдем к оценке точности схемы (2). Благодаря теореме 1, мы можем использовать теорию возмущенного метода Галеркина для оценки погрешности этой схемы. Именно, мы используем оценку из леммы Стренга (см. неравенство (17), с. 27), которая в рассматриваемой ситуации принимает вид :

$$|u - u_h|_{1,\Omega} \leq C \left\{ \inf_{v_h \in V_h} \left( |u - v_h|_{1,\Omega} + \sup_{w_h \in V_h} \frac{|a(v_h, w_h) - a_h(v_h, w_h)|}{|w_h|_{1,\Omega}} \right) + \sup_{w_h \in V_h} \frac{|f(w_h) - f_h(w_h)|}{|w_h|_{1,\Omega}} \right\}. \quad (13)$$

Здесь  $u$  — решение исходной задачи,  $u_h$  — решение задачи (2). Предполагая функцию  $u$  достаточно гладкой, мы можем оценить сверху правую часть неравенства (13), заменив  $v_h$  на  $u_I \in V_h$ , где  $u_I$  есть  $V_h$ -интерполянт точного решения  $u(x)$ . Получим:

$$|u - u_h|_{1,\Omega} \leq C \{ |u - u_I|_{1,\Omega} + E_a(u_I) + E_f \}, \quad (14)$$

где

$$E_a(u_I) = \sup_{w_h \in V_h} \frac{|a(u_I, w_h) - a_h(u_I, w_h)|}{|w_h|_{1,\Omega}}, \quad E_f = \sup_{w_h \in V_h} \frac{|f(w_h) - f_h(w_h)|}{|w_h|_{1,\Omega}}.$$

Таким образом, требуется оценить три составляющие погрешности: погрешность интерполяции  $|u - u_I|_{1,\Omega}$ ; погрешность квадратуры  $E_a(u_I)$  при вычислении формы  $a$ , и, наконец, погрешность квадратуры  $E_f$  при вычислении формы  $f$ .

Для оценки погрешности интерполяции можно использовать доказанную ранее теорему 4, с. 151, в которой надо положить  $s = m$ ,

$p = 2, n = 2$ . Ясно, что условие  $m > r$  выполняется как для лагранжевых, так и для эрмитовых элементов. Согласно этой теореме, если  $u \in V \cap W_2^{m+1}(\Omega)$ , то

$$|u - u_I|_{1,\Omega} \leq c h^m |u|_{m+1,\Omega}. \quad (15)$$

В дальнейшем будем часто использовать следующую лемму, доказательство которой оставляем читателю (см. задачу 5, с. 193).

**Лемма 3.** Пусть  $u \in W_p^k(\Omega)$ ,  $\varphi \in W_\infty^k(\Omega)$ ,  $p \in [1, \infty]$ . Тогда  $\varphi u \in W_p^k(\Omega)$ , существует постоянная  $c$  такая, что

$$|\varphi u|_{k,p,\Omega} \leq c \sum_{l=0}^k |\varphi|_{l,\infty,\Omega} |u|_{k-l,p,\Omega}, \quad [\varphi u]_{k,p,\Omega} \leq c \sum_{l=0}^k [\varphi]_{l,\infty,\Omega} [u]_{k-l,p,\Omega}.$$

Оценки погрешности квадратур при вычислении форм  $a$  и  $f$  непосредственно вытекают из следующей леммы.

**Лемма 4.** Пусть квадратурная формула  $\hat{S}$  точна на полиномах из  $P_{2m-2}$ , т. е.  $\hat{E}(\hat{\phi}) = 0$  для всех  $\hat{\phi} \in P_{2m-2}$ ,  $a \in W_\infty^m(e)$ ,  $f \in W_r^m(e)$ ,  $rm > 2$ ,  $r \geq 1$ . Тогда для любых функций  $p, q \in P_m$  и любого элемента  $e \in \mathcal{T}_h$  справедливы оценки

$$|E_e(a \frac{\partial q}{\partial x_j} \frac{\partial p}{\partial x_i})| \leq c h^m \|a\|_{m,\infty,e} \|q\|_{m,e} \|p\|_{1,e}, \quad (16)$$

$$|E_e(fp)| \leq c h^m (\text{mes}(e))^{1/2-1/r} \|f\|_{m,r,e} \|p\|_{1,e} \quad (17)$$

с постоянной  $c$ , не зависящей от  $h$ .

**ДОКАЗАТЕЛЬСТВО.** Докажем неравенство (16). Введем обозначения:

$$u = \frac{\partial q}{\partial x_j} \in P_{m-1}, \quad v = \frac{\partial p}{\partial x_i} \in P_{m-1}, \quad \phi = au.$$

Тогда

$$E \equiv E_e\left(a \frac{\partial q}{\partial x_j} \frac{\partial p}{\partial x_i}\right) = \det(B_e) \hat{E}(\hat{\phi} \hat{v}), \quad (18)$$

где  $\hat{\phi} \in W_\infty^m(\hat{e})$ ,  $\hat{v} \in P_{m-1}$ . Учитывая равенство (7), получим:

$$|\hat{E}(\hat{\phi} \hat{v})| = \left| \int_{\hat{e}} \hat{\phi}(\hat{x}) \hat{v}(\hat{x}) d\hat{x} - \sum_{l=1}^L \hat{c}_l(\hat{\phi} \hat{v})(\hat{b}_l) \right| \leq |\hat{\phi} \hat{v}|_{0,\infty,\hat{e}} \leq |\hat{\phi}|_{0,\infty,\hat{e}} |\hat{v}|_{0,\infty,\hat{e}}.$$

Используя эквивалентность норм на  $P_{m-1}$ , нетрудно показать, что (см. задачу 6, с. 193)

$$|\hat{v}|_{j,\infty,\hat{e}} \leq c |\hat{v}|_{j,\hat{e}}, \quad j = 0, 1, \dots, m-1. \quad (19)$$

Таким образом,

$$|\hat{E}(\hat{\phi}\hat{v})| \leq c \|\hat{\phi}\|_{m,\infty,\hat{e}} |\hat{v}|_{0,\hat{e}},$$

откуда следует, что линейный функционал  $f(\hat{\phi}) = \hat{E}(\hat{\phi}\hat{v})$  является непрерывным в  $W_\infty^m(\hat{e})$  и в силу условия на точность квадратуры обращается в нуль на пространстве  $P_{m-1}$ . Используя лемму Брамбла — Гильберта, получим:

$$|\hat{E}(\hat{\phi}\hat{v})| \leq c |\hat{\phi}|_{m,\infty,\hat{e}} |\hat{v}|_{0,\hat{e}}.$$

Из леммы 3 и неравенства (19) следует, что

$$|\hat{\phi}|_{m,\infty,\hat{e}} = |\hat{a}\hat{u}|_{m,\infty,\hat{e}} \leq c \sum_{j=0}^{m-1} |\hat{a}|_{m-j,\infty,\hat{e}} |\hat{u}|_{j,\infty,\hat{e}} \leq c \sum_{j=0}^{m-1} |\hat{a}|_{m-j,\infty,\hat{e}} |\hat{u}|_{j,\hat{e}}.$$

Здесь мы учли, что  $\hat{u} \in P_{m-1}$ . Таким образом,

$$|\hat{E}(\hat{\phi}\hat{v})| \leq c \left( \sum_{j=0}^{m-1} |\hat{a}|_{m-j,\infty,\hat{e}} |\hat{u}|_{j,\hat{e}} \right) |\hat{v}|_{0,\hat{e}}.$$

Перейдем в полученной оценке с базисного элемента  $\hat{e}$  на элемент  $e$ . Согласно следствию 3, с. 56, имеем:  $|\hat{a}|_{m-j,\infty,\hat{e}} \leq c h^{m-j} |a|_{m-j,\infty,e}$ ,  $|\hat{u}|_{j,\hat{e}} \leq c h^j (\det(B_e))^{-1/2} |u|_{j,e}$ ,  $|\hat{v}|_{0,\hat{e}} \leq c (\det(B_e))^{-1/2} |v|_{0,e}$ , т. е.

$$\begin{aligned} |\hat{E}(\hat{\phi}\hat{v})| &\leq c h^m (\det(B_e))^{-1} \left( \sum_{j=0}^{m-1} |a|_{m-j,\infty,e} |u|_{j,e} \right) |v|_{0,e} \leq \\ &\leq c h^m (\det(B_e))^{-1} \|a\|_{m,\infty,e} \|u\|_{m-1,e} |v|_{0,e}. \end{aligned}$$

Эта оценка и равенство (18) приводят к неравенству (16).

Докажем теперь оценку (17). Пусть  $\hat{a}$  — некоторый узел интерполяции на  $\hat{e}$ . По функции  $\hat{p} \in P_m$  определим постоянную на  $\hat{e}$  функцию  $\hat{p}_0(\hat{x}) = \hat{p}(\hat{a})$ . Отметим, что  $\hat{p}_0$  является  $P_0$ -интерполянтom функции  $\hat{p}$ . Отсюда и из конечномерности  $P_m$  следует, что

$$|\hat{p}_0| \leq |\hat{p}|_{0,\infty,\hat{e}} \leq c |\hat{p}|_{0,\hat{e}}, \quad |\hat{p} - \hat{p}_0|_{0,\hat{e}} \leq c |\hat{p}|_{1,\hat{e}}.$$

Ясно, что

$$E_e(fp) = \det(B_e)\hat{E}(f\hat{p}), \quad \hat{p} \in P_m, \quad (20)$$

причем

$$\hat{E}(f\hat{p}) = \hat{E}(f\hat{p}_0) + \hat{E}(f(\hat{p} - \hat{p}_0)) \equiv E_0(f) + E_1(f). \quad (21)$$

Оценим величину  $E_0(f)$ . Учитывая вложение  $W_r^m(\hat{e}) \subset C(\hat{e})$  при  $mr > 2$ , получим

$$|E_0(\hat{f})| = |\hat{E}(\hat{f})| |\hat{p}_0| \leq c |\hat{f}|_{0,\infty,\hat{e}} |\hat{p}|_{0,\hat{e}} \leq c \|\hat{f}\|_{m,r,\hat{e}} |\hat{p}|_{0,\hat{e}}.$$

Поскольку  $E_0(P_{m-1}) = 0$ , то из леммы Брамбла — Гильберта следует оценка  $|E_0(\hat{f})| \leq c |\hat{f}|_{m,r,\hat{e}} |\hat{p}|_{0,\hat{e}}$ . Аналогично, учитывая, что  $E_1(P_{m-2}) = 0$ , имеем:

$$|E_1(\hat{f})| \leq c \|\hat{f}\|_{m,r,\hat{e}} |p - \hat{p}_0|_{0,\hat{e}} \leq c (|\hat{f}|_{m-1,r,\hat{e}} + |\hat{f}|_{m,r,\hat{e}}) |p|_{1,\hat{e}}.$$

Далее, поскольку  $|\hat{f}|_{j,r,\hat{e}} \leq ch^j (\det(B_e))^{-1/r} |f|_{j,r,e}$ ,  $j = m - 1, m$ ,  $|\hat{p}|_{k,\hat{e}} \leq ch^k (\det(B_e))^{-1/2} |p|_{k,e}$ ,  $k = 0, 1$ , а  $\det(B_e) = 2 \text{mes}(e)$ , то

$$|E_0(\hat{f})| + |E_1(\hat{f})| \leq ch^m (\text{mes}(e))^{-1/2-1/r} \|f\|_{m,r,e} \|p\|_{1,e}.$$

Последняя оценка вместе с равенствами (20), (21) приводит к (17).  $\square$

**ЗАМЕЧАНИЕ 2.** Необходимо сделать несколько пояснений к доказательству оценки (17). Во-первых, всегда можно указать подходящее  $r$ . При  $m = 1$  имеем  $r > 2$ , при  $m > 1$  можно считать  $r = 2$ . Отметим, что при  $r \geq 2$  имеем:  $1/s \equiv 1/2 - 1/r \geq 0$ ,  $1/s + 1/r + 1/2 = 1$ . С какой целью был введен полином  $\hat{p}_0$ ? Иначе бы мы имели оценку

$$|\hat{E}(\hat{f})| \leq c \|\hat{f}\|_{m,r,\hat{e}} |\hat{p}|_{0,\hat{e}} \leq c (|\hat{f}|_{m-1,r,\hat{e}} + |\hat{f}|_{m,r,\hat{e}}) |p|_{0,\hat{e}},$$

из которой в правой части неравенства (17) получили бы множитель  $h^{m-1}$  вместо  $h^m$ .

Теперь все готово для доказательства основной теоремы настоящего параграфа.

**Теорема 2.** Пусть квадратурная формула  $\hat{S}$  имеет положительные коэффициенты и является точной на полиномах из  $P_{2m-2}$ . Пусть, кроме того,

$$u \in W_2^{m+1}(\Omega), \quad a_{ij} \in W_\infty^m(\Omega), \quad i, j = 1, 2, \quad f \in W_r^m(\Omega), \quad rm > 2, \quad r \geq 2.$$

Тогда справедлива следующая оценка точности схемы (2):

$$\|u - u_h\|_{1,\Omega} \leq ch^m (|u|_{m+1,\Omega} + \sum_{i,j=1}^2 \|a_{ij}\|_{m,\infty,\Omega} \|u\|_{m+1,\Omega} + \|f\|_{m,r,\Omega}), \quad (22)$$

где постоянная  $c$  не зависит от  $h$ .

ДОКАЗАТЕЛЬСТВО. Вследствие условия на точность квадратуры  $\hat{S}$  можно использовать теорему 1, а, значит, и оценку (14), для первого слагаемого в правой части которой справедливо неравенство (15). При этом надо заметить, что в силу граничного условия Дирихле нормы  $\|\cdot\|_{1,\Omega}$  и  $|\cdot|_{1,\Omega}$  на пространстве  $V$  эквивалентны.

Два других слагаемых в правой части неравенства (14) оценим, используя лемму 4. Применяя неравенство Коши — Буняковского, получим, что для любого  $w_h \in V_h$

$$\begin{aligned} |a(u_I, w_h) - a_h(u_I, w_h)| &\leq \sum_e \sum_{i,j=1}^2 \left| E_e \left( a_{ij} \frac{\partial u_I}{\partial x_j} \frac{\partial w_h}{\partial x_i} \right) \right| \leq \\ &\leq c h^m \sum_e \sum_{i,j=1}^2 \|a_{ij}\|_{m,\infty,e} \|u_I\|_{m,e} |w_h|_{1,e} \leq \\ &\leq c h^m \left( \sum_{i,j=1}^2 \|a_{ij}\|_{m,\infty,\Omega} \right) \sum_e \|u_I\|_{m,e} |w_h|_{1,e} \leq \\ &\leq c h^m \left( \sum_{i,j=1}^2 \|a_{ij}\|_{m,\infty,\Omega} \right) \left( \sum_e \|u_I\|_{m,e}^2 \right)^{1/2} |w_h|_{1,\Omega}. \quad (23) \end{aligned}$$

По теореме 3, с. 150, имеем:

$$|u - u_I|_{j,e} \leq c h^{m+1-j} |u|_{m+1,e} \leq c |u|_{m+1,e}, \quad j = 0, 1, \dots, m+1.$$

Следовательно,  $|u_I|_{j,e} \leq |u|_{j,e} + |u - u_I|_{j,e} \leq c \|u\|_{m+1,e}$ , и

$$\|u_I\|_{m,e} \leq c \|u\|_{m+1,e}, \quad \sum_e \|u_I\|_{m,e}^2 \leq c \sum_e \|u\|_{m+1,e}^2 = c \|u\|_{m+1,\Omega}^2.$$

Таким образом, после деления неравенства (23) на  $|w_h|_{1,\Omega}$  и перехода к точной верхней грани по всем  $w_h \in V_h$ , получим:

$$E_a(u_I) \leq c h^m \sum_{i,j=1}^2 \|a_{ij}\|_{m,\infty,\Omega} \|u\|_{m+1,\Omega}. \quad (24)$$

Аналогично, при  $s = 1/2 - 1/r$  имеем

$$\begin{aligned} |f(w_h) - f_h(w_h)| &\leq \sum_e |E_e(f w_h)| \leq c h^m \sum_e (\text{mes}(e))^s \|f\|_{m,r,e} \|w_h\|_{1,e} \leq \\ &\leq c h^m (\text{mes}(\Omega))^s \|f\|_{m,r,\Omega} \|w_h\|_{1,\Omega} \leq c h^m \|f\|_{m,r,\Omega} |w_h|_{1,\Omega}. \quad (25) \end{aligned}$$

Здесь использовано неравенство Гельдера (см. с. 35)

$$\sum_e |a_e b_e c_e| \leq \left( \sum_e |a_e|^\alpha \right)^{1/\alpha} \left( \sum_e |b_e|^\beta \right)^{1/\beta} \left( \sum_e |c_e|^\gamma \right)^{1/\gamma}, \quad (26)$$

справедливое для любых  $\alpha, \beta, \gamma \geq 1, 1/\alpha + 1/\beta + 1/\gamma = 1$  ( $1/\alpha = s, \beta = r, \gamma = 2$ ). Из (25), очевидно, следует, что

$$E_f \leq c h^m \|f\|_{m,r,\Omega}. \quad (27)$$

Подставляя оценки (15), (24) и (27) в (14), приходим к (22).  $\square$

ЗАМЕЧАНИЕ 3. Требование точности квадратурной формулы  $\hat{S}$  на полиномах из  $P_{2m-2}$  для аффинно-эквивалентных треугольных элементов равносильно требованию того, чтобы интегралы

$$\int_e a_{ij} \frac{\partial u_h}{\partial x_j} \frac{\partial v_h}{\partial x_i} dx$$

по этой квадратурной формуле вычислялись точно для любых  $u_h, v_h \in V_h$ , если коэффициенты  $a_{ij}$  — постоянные.

**Примеры квадратурных формул.** Приведем примеры квадратурных формул для треугольного базисного конечного элемента  $\hat{e}$ . При этом будем помнить, что  $\text{mes}(\hat{e}) = 1/2$ .

1)  $m = 1$ . Формула численного интегрирования с одним узлом

$$\int_{\hat{e}} \hat{\phi}(\hat{x}) d\hat{x} \approx \frac{1}{2} \hat{\phi}(\hat{b}_1) = \hat{S}(\hat{\phi}), \quad \hat{b}_1 \in \hat{e},$$

при любом выборе  $\hat{b}_1 \in \hat{e}$  точна на  $P_0$ . Поэтому она может быть использована в схемах МКЭ с линейными лагранжевыми элементами. Заметим, что если  $\hat{b}_1$  совпадает с центром тяжести треугольника  $\hat{e}$ , т. е.  $\hat{b}_1 = (1/3, 1/3)$ , то эта квадратура точна на  $P_1$ .

В схемах МКЭ с линейными лагранжевыми элементами можно применять также квадратуру

$$\int_{\hat{e}} \hat{\phi}(\hat{x}) d\hat{x} \approx \frac{1}{6} \sum_{l=1}^3 \hat{\phi}(\hat{b}_l) = \hat{S}(\hat{\phi}),$$

где  $\hat{b}_1 = (0, 0), \hat{b}_2 = (1, 0), \hat{b}_3 = (0, 1)$  (вершины треугольника  $\hat{e}$ ). Эта квадратура точна на  $P_1$ , и, тем более, на  $P_0$ .

2)  $m = 2$ . Формула

$$\int_{\hat{e}} \hat{\phi}(\hat{x}) d\hat{x} \approx \frac{1}{6} \sum_{l=1}^3 \hat{\phi}(\hat{b}_l) = \hat{S}(\hat{\phi}),$$

где  $\hat{b}_1 = (1/2, 0)$ ,  $\hat{b}_2 = (0, 1/2)$ ,  $\hat{b}_3 = (1/2, 1/2)$  (середины сторон треугольника  $\hat{e}$ ), точна на  $P_2$ . Таким образом, эта квадратурная формула может быть использована в конечно-элементных схемах с  $P_2$ -интерполяцией, т. е. с квадратичными элементами.

Примеры более сложных квадратур приведены в задаче 2.

### Задачи.

1. Проверить все утверждения о точности приведенных выше квадратур на треугольном элементе  $\hat{e}$ . Убедиться, что достаточно проверить равенства

$$\int_{\hat{e}} \hat{x}_1^{\alpha_1} \hat{x}_2^{\alpha_2} d\hat{x} = \hat{S}(\hat{x}_1^{\alpha_1} \hat{x}_2^{\alpha_2}), \quad 0 \leq \alpha_1 + \alpha_2 \leq s,$$

если утверждается, что  $\hat{S}$  точна на  $P_s$ .

2. Приведем примеры более сложных квадратурных формул на треугольном базисном элементе. Для сокращения записей будем придерживаться следующих соглашений. Узел квадратуры будем определять тройкой чисел  $\hat{b}$  и кратностью  $k$ . В тройке  $\hat{b} = (x_1, x_2, x_3)$  всегда  $x_3 = 1 - x_1 - x_2$ , и пара  $(x_1, x_2)$  определяет координаты узла квадратуры. Если кратность узла равна трем, то пары  $(x_1, x_3)$ ,  $(x_3, x_2)$  также определяют узлы квадратуры. Кратным узлам соответствуют равные коэффициенты квадратуры  $\hat{c}$ .

а) Точность  $P_3$ , 4 узла:

$$\begin{array}{lll} \hat{b}_{1-3} = (1/5, 1/5, 3/5), & k = 3, & \hat{c}_{1-3} = 25/96, \\ \hat{b}_4 = (1/3, 1/3, 1/3), & k = 1, & \hat{c}_4 = -9/32. \end{array}$$

б) Точность  $P_3$ , 6 узлов:

$$\begin{array}{lll} \hat{b}_{1-3} = (1/2, 1/2, 0), & k = 3, & \hat{c}_{1-3} = 1/60, \\ \hat{b}_{4-6} = (1/6, 1/6, 2/3), & k = 3, & \hat{c}_{4-6} = 9/60. \end{array}$$

в) Точность  $P_3$ , 7 узлов:

$$\begin{array}{lll} \hat{b}_{1-3} = (0, 0, 1), & k = 3, & \hat{c}_{1-3} = 1/40, \\ \hat{b}_{4-6} = (1/2, 1/2, 0), & k = 3, & \hat{c}_{4-6} = 1/15, \\ \hat{b}_7 = (1/3, 1/3, 1/3), & k = 1, & \hat{c}_7 = 27/120. \end{array}$$

д) Точность  $P_4$ , 6 узлов (узлы и коэффициенты даны с точностью 16 знаков):

$$\begin{array}{lll} \hat{b}_{1-3} = (0.09157621350977073, 0.09157621350977073, x_3), & k = 3, & \hat{c}_{1-3} = 0.1099517436553218/2, \\ \hat{b}_{4-6} = (0.4459484909159649, 0.4459484909159649, x_3), & k = 3, & \hat{c}_{4-6} = 0.2233815896780115/2. \end{array}$$

е) Точность  $P_5$ , 7 узлов:

$$\begin{array}{lll} \hat{b}_{1-3} = \left( \frac{6 - \sqrt{15}}{21}, \frac{6 - \sqrt{15}}{21}, \frac{9 + 2\sqrt{15}}{21} \right), & k = 3, & \hat{c}_{1-3} = \frac{155 - \sqrt{15}}{2400}, \\ \hat{b}_{4-6} = \left( \frac{6 + \sqrt{15}}{21}, \frac{6 + \sqrt{15}}{21}, \frac{9 - 2\sqrt{15}}{21} \right), & k = 3, & \hat{c}_{4-6} = \frac{155 + \sqrt{15}}{2400}, \\ \hat{b}_7 = (1/3, 1/3, 1/3), & k = 1, & \hat{c}_7 = 9/80. \end{array}$$

Проверить правильность указанных формул. Может быть проще составить программу для ЭВМ?

3. Пусть коэффициенты квадратуры  $\hat{S}$  неотрицательны. Показать, что

$$\hat{S}(\hat{f}) \leq \hat{S}(\hat{g}), \text{ если } \hat{f}(\hat{x}) \leq \hat{g}(\hat{x}) \quad \forall \hat{x} \in \hat{e},$$

$$|\hat{S}(\hat{f}\hat{g})|^2 \leq \hat{S}(\hat{f}^2) \hat{S}(\hat{g}^2).$$

4. Пусть квадратура  $\hat{S}$  точна на полиномах  $P_{2m-2}$ , а ее коэффициенты не обязательно все положительны. Доказать, что для любого  $u_h \in V_h$

$$|a(u_h, u_h) - a_h(u_h, u_h)| \leq ch \sum_{i,j=1}^2 |a_{ij}|_{1,\infty,\Omega} |u_h|_{1,\Omega}^2,$$

и, как следствие, форма  $a_h$  равномерно положительно определена на  $V_h$  при достаточно малых  $h$ , если положительно определена форма  $a$ .

5. Используя формулу

$$D^\alpha(\varphi u) = \sum_{j=0}^m \sum_{|\beta|=j, \beta+\gamma=\alpha} D^\beta \varphi D^\gamma u, \quad |\alpha| = m, \quad m \geq 1,$$

и неравенства вида

$$\left| \sum_{j=1}^k a_j f_j \right|_{0,p,\Omega} \leq \sum_{j=1}^k |a_j|_{0,\infty,\Omega} |f_j|_{0,p,\Omega},$$

доказать лемму 3.

6. Пусть  $\hat{v} \in P_m$ . Пользуясь оценкой  $|\hat{v}|_{0,\infty,\hat{e}} \leq c|\hat{v}|_{0,\hat{e}}$ , доказать, что  $|\hat{v}|_{j,\infty,\hat{e}} \leq c|\hat{v}|_{j,\hat{e}}$  для  $j = 0, 1, \dots, m$ .

7. Пусть  $\Omega$  — единичный квадрат и его триангуляция  $\mathcal{T}_h$  построена следующим образом. Область  $\Omega$  разбивается на квадраты размеров  $h \times h$  и затем каждый квадрат — на два треугольника биссектрисой левого нижнего угла.

а) Рассмотреть задачу Дирихле для уравнения Пуассона

$$-\Delta u = f(x) \text{ при } x \in \Omega, \quad u(x) = 0 \text{ при } x \in \Gamma$$

и построить схему МКЭ с численным интегрированием, используя линейные лагранжевы конечные элементы и квадратурные формулы с одним узлом, совпадающим с вершиной при прямом угле. Убедиться, что построенная схема совпадает со стандартной пятиточечной разностной схемой

$$h^{-2}(4u_{i,j} - u_{i-1,j} - u_{i+1,j} - u_{i,j-1} - u_{i,j+1}) = f_{i,j},$$

где  $u_{i,j}$  — значение сеточной функции в узле сетки  $(i, j) = (ih, jh)$ .

б) Сделать то же для задачи Неймана:

$$-\Delta u = f(x) \text{ при } x \in \Omega, \quad \frac{\partial u(x)}{\partial n} = 0 \text{ при } x \in \Gamma.$$

## 2. Аффинно-эквивалентные прямоугольные элементы.

Полагая, что читатель знаком с материалом предыдущего параграфа, в котором было подробно проанализировано влияние численного интегрирования на МКЭ в случае треугольных аффинно-эквивалентных элементов, рассмотрим те же вопросы в случае прямоугольных аффинно-эквивалентных элементов, придерживаясь аналогичной схемы рассуждений.

Пусть область  $\Omega$  составлена из прямоугольников и (или) параллелограммов. В этом случае она может быть точно триангулирована семейством  $\mathcal{T}_h$  четырехугольных элементов, каждый из которых является либо прямоугольником, либо параллелограммом.

Пусть  $(\hat{e}, \hat{P}, \hat{\Sigma})$  — базисный прямоугольный элемент класса  $C^0$  или  $C^1$ ,  $P_m \subset \hat{P} \subseteq Q_m$ , и  $\mathcal{E}_h = \{(e, P_e, \Sigma_e), e \in \mathcal{T}_h\}$  — семейство ассоциированных с ним четырехугольных аффинно-эквивалентных элементов;  $x = x_e(\hat{x}) = B_e \hat{x} + b_e$  — аффинное отображение  $\hat{e}$  на  $e$  такое, что  $\det(B_e) > 0$ <sup>1)</sup> и  $\|B_e\| \leq ch$ ,  $\|B_e^{-1}\| \leq ch^{-1}$  для всех  $e \in \mathcal{T}_h$ , что равносильно требованию регулярности  $\mathcal{T}_h$ .

Определим  $V_h \subset V = H_0^1(\Omega)$  как соответствующее  $\mathcal{E}_h$  пространство конечных элементов:

$$V_h = \{v_h \in C(\bar{\Omega}) : v_h|_e \in P_e \ \forall e \in \mathcal{T}_h, v_h(x) = 0, x \in \Gamma\}.$$

Пусть далее  $\hat{S}$  — некоторая квадратурная формула на базисном элементе (т. е. на квадрате  $[0, 1] \times [0, 1]$ ),  $S_e$  и  $S_\Omega$  — индуцированные  $\hat{S}$  квадратуры на  $e \in \mathcal{T}_h$  и  $\Omega$  соответственно. Напомним, что  $\det(B_e) = \text{mes}(e)$ .

Рассмотрим схему МКЭ с численным интегрированием:

$$\text{найти } u_h \in V_h : \quad a_h(u_h, v_h) = f_h(v_h) \quad \forall v_h \in V_h, \quad (28)$$

где формы  $a_h$  и  $f_h$  определяются формулами (3) и (4).

Оставим пока в стороне вопрос о разрешимости схемы (28) и установим аналог леммы 4, что позволит нам сразу описать класс допустимых квадратур  $\hat{S}$ , обеспечивающих нужную точность.

**Лемма 5.** Пусть квадратурная формула  $\hat{S}$  точна на полиномах из  $Q_{2m-1}$ , т. е.  $\hat{E}(\hat{\phi}) = 0 \ \forall \hat{\phi} \in Q_{2m-1}$ ,

$$a \in W_\infty^m(e), \quad f \in W_r^m(e), \quad rm > 2, \quad r \geq 1.$$

<sup>1)</sup>Предположение  $\det(B_e) > 0$  означает, что выбрана подходящая нумерация вершин  $e$  при определении отображения  $x_e$ .

Тогда для любых функций  $p, q \in P_e$  и любого элемента  $e \in \mathcal{T}_h$  справедливы оценки

$$\left| E_e \left( a \frac{\partial q}{\partial x_j} \frac{\partial p}{\partial x_i} \right) \right| \leq ch^m \|a\|_{m,\infty,e} \|q\|_{m,e} \|p\|_{1,e}, \quad (29)$$

$$|E_e(fp)| \leq ch^m (\text{mes}(e))^{1/2-1/r} \|f\|_{m,r,e} \|p\|_{1,e} \quad (30)$$

с постоянной  $c$ , не зависящей от  $h$ .

ДОКАЗАТЕЛЬСТВО. В основном, рассуждения проводятся аналогично доказательству леммы 4. Докажем неравенство (29). Введем обозначения:

$$u = \frac{\partial q}{\partial x_j}, \quad v = \frac{\partial p}{\partial x_i}, \quad \varphi = au.$$

Тогда

$$E \equiv E_e \left( a \frac{\partial q}{\partial x_j} \frac{\partial p}{\partial x_i} \right) = \det(B_e) \hat{E}(\hat{\varphi}\hat{v}),$$

причем  $\hat{\varphi} \in W_\infty^m(\hat{e})$ ,  $\hat{v} \in \hat{P}$ ,  $P_m \subset \hat{P} \subseteq Q_m$ . Имеем

$$|\hat{E}(\hat{\varphi}\hat{v})| \leq 2|\hat{\varphi}\hat{v}|_{0,\infty,\hat{e}} \leq 2|\hat{\varphi}|_{0,\infty,\hat{e}} |\hat{v}|_{0,\infty,\hat{e}} \leq c \|\hat{\varphi}\|_{m,\infty,\hat{e}} |\hat{v}|_{0,\hat{e}},$$

откуда следует, что линейный функционал  $f(\hat{\varphi}) = \hat{E}(\hat{\varphi}\hat{v})$  непрерывен на  $W_\infty^m(\hat{e})$  и в силу предположения о точности квадратуры обращается в нуль на пространстве  $Q_{m-1}$  и, тем более, на  $P_{m-1}$ . Используя лемму Брамбла — Гильберта, получим  $|\hat{E}(\hat{\varphi}\hat{v})| \leq c |\hat{\varphi}|_{m,\infty,\hat{e}} |\hat{v}|_{0,\hat{e}}$ . Далее рассуждения проводятся так же, как и при доказательстве леммы 4.

При доказательстве оценки (30) теперь нет необходимости во введении функции  $\hat{p}_0$ . В самом деле,

$$E_e(fp) = \det(B_e) \hat{E}(f\hat{p}) \equiv \det(B_e) F(\hat{f}), \quad \hat{p} \in \hat{P}, \quad P_m \subset \hat{P} \subseteq Q_m.$$

Учитывая вложение  $W_r^m(\hat{e}) \subset C(\hat{e})$  при  $mr > 2$ , получим

$$|F(\hat{f})| \leq c |\hat{f}|_{0,\infty,\hat{e}} |\hat{p}|_{0,\infty,\hat{e}} \leq c \|\hat{f}\|_{m,r,\hat{e}} |\hat{p}|_{0,\hat{e}}.$$

Поскольку  $F(Q_{m-1}) = 0$ , и, тем более,  $F(P_{m-1}) = 0$ , то из леммы Брамбла — Гильберта следует, что

$$|F(\hat{f})| \leq c |\hat{f}|_{m,r,\hat{e}} |\hat{p}|_{0,\hat{e}}.$$

Далее рассуждения проводятся по аналогии с доказательством леммы 4.  $\square$

Как и в случае треугольных элементов, из этой леммы следует оценка

$$E_a(u_I) + E_f \leq ch^m \left( \sum_{i,j=1}^2 \|a_{ij}\|_{m,\infty,\Omega} \|u\|_{m+1,\Omega} + \|f\|_{m,r,\Omega} \right).$$

**Примеры квадратурных формул.** Можно указать два типа квадратурных формул на базисном элементе, удовлетворяющих условию леммы 5.

Пусть задана одномерная квадратура

$$\int_0^1 \hat{\varphi}(\hat{x}) d\hat{x} \approx \sum_{i=1}^L \hat{c}_i \hat{\varphi}(\hat{b}_i), \quad \hat{b}_i \in [0, 1], \quad i = 1, 2, \dots, L, \quad (31)$$

точная на полиномах из  $P_s$ . Вычисляя интеграл по элементу  $\hat{e}$  как повторный, можем написать

$$\begin{aligned} \int_{\hat{e}} \hat{\phi}(\hat{x}) d\hat{x} &= \int_0^1 \left( \int_0^1 \hat{\phi}(\hat{x}_1, \hat{x}_2) d\hat{x}_1 \right) d\hat{x}_2 \equiv \int_0^1 \hat{\varphi}(\hat{x}_2) d\hat{x}_2 \approx \sum_{j=1}^L \hat{c}_j \hat{\varphi}(\hat{b}_j) = \\ &= \sum_{j=1}^L \hat{c}_j \int_0^1 \hat{\phi}(\hat{x}_1, \hat{b}_j) d\hat{x}_1 \approx \sum_{i,j=1}^L \hat{c}_i \hat{c}_j \hat{\phi}(\hat{b}_i, \hat{b}_j) = \hat{S}(\hat{\phi}). \end{aligned}$$

Говорят, что квадратура  $\hat{S}$  получена суперпозицией одномерных квадратур (31). Ясно, что она точна на полиномах из  $Q_s$ . Таким образом, достаточно построить квадратуру (31), точную на  $P_{2m-1}$ .

С вычислительной точки зрения естественно применять квадратуры с минимально возможным  $L$ . Приведем примеры таких квадратур.

Пусть

$$L_m(t) = \frac{1}{2^m m!} \frac{d^m}{dt^m} (t^2 - 1)^m \quad (32)$$

есть полином Лежандра степени  $m^1$ ,  $t_i \in (-1, 1)$ ,  $i = 1, \dots, m$ , — его корни,  $t^{(k)} \in (-1, 1)$ ,  $k = 2, \dots, m$ , — корни  $L'_m(t)$ .

а) Квадратура Гаусса. Узлы и коэффициенты определяются по формулам

$$\hat{b}_i = (1 + t_i)/2, \quad \hat{c}_i = \frac{1}{(1 - t_i^2)(L'_m(t_i))^2}, \quad i = 1, 2, \dots, m.$$

<sup>1)</sup> Система полиномов  $L_0, L_1, L_2, \dots$  ортогональна в  $L_2(-1, 1)$ , т. е.  $\int_{-1}^1 L_i L_j dt = 0$ , если  $i \neq j$ .

б) Квадратура Лобатто. Узлы и коэффициенты определяются следующим образом:  $\hat{b}_1 = 0$ ,  $\hat{b}_{m+1} = 1$ ,  $\hat{b}_k = (1 + t^{(k)})/2$ ,  $k = 2, \dots, m$ , — точки экстремума полинома  $L_m(2t - 1)$ ,

$$\hat{c}_i = \frac{1}{m(m+1)(L_m(t^{(i)}))^2}, \quad i = 1, 2, \dots, m+1.$$

ЗАМЕЧАНИЕ 4. Хорошо известно, что не существует квадратуры вида (31) с числом узлов, меньшим  $m$ , точной на любом полиноме степени  $2m - 1$ . Если два узла квадратуры совпадают с концами отрезка интегрирования, то единственная квадратура вида (31), имеющая  $m + 1$  узел и точная на любом полиноме степени  $2m - 1$ , есть квадратура Лобатто.

Квадратуру  $\hat{S}$ , полученную суперпозицией одномерных  $m$ -точечных квадратур Гаусса ( $(m + 1)$ -точечных квадратур Лобатто), будем называть  $m \times m$ -точечной квадратурой Гаусса ( $(m + 1) \times (m + 1)$ -точечной квадратурой Лобатто). Эти квадратуры точны на полиномах из  $Q_{2m-1}$ , а их коэффициенты положительны.

**Лемма 6.** Пусть  $\hat{u} \in Q_m$ ,  $m \geq 1$ ,  $\hat{S}$  является  $m \times m$ -точечной квадратурой Гаусса или Лобатто,  $\hat{S}(|\hat{\nabla}\hat{u}|^2) = 0$ . Тогда

$$\hat{u} = c_1 + c_2 \prod_{i,j=1}^m (\hat{x}_1 - \hat{b}_i)(\hat{x}_2 - \hat{b}_j), \quad c_1, c_2 \in R.$$

ДОКАЗАТЕЛЬСТВО. Из условий леммы сразу вытекает, что

$$\frac{\partial \hat{u}}{\partial \hat{x}_1}(\hat{b}_i, \hat{b}_j) = \frac{\partial \hat{u}}{\partial \hat{x}_2}(\hat{b}_i, \hat{b}_j) = 0, \quad i, j = 1, 2, \dots, m.$$

Поскольку  $\hat{u} \in Q_m$ , то для фиксированного  $\hat{x}_2 = \hat{b}_j$  производная  $\partial \hat{u} / \partial \hat{x}_1 \equiv p(\hat{x}_1) \in P_{m-1}$  и  $p(\hat{b}_i) = 0$ ,  $i = 1, \dots, m$ , следовательно,

$$\frac{\partial \hat{u}}{\partial \hat{x}_1}(\hat{x}_1, \hat{b}_j) = 0 \text{ для всех } \hat{x}_1 \in [0, 1], \quad j = 1, 2, \dots, m.$$

Это означает, что для любого фиксированного  $\hat{x}_1$  полином  $q(\hat{x}_2) \equiv \partial \hat{u} / \partial \hat{x}_1 \in P_m$  обращается в нуль в  $m$  точках  $\hat{x}_2 = \hat{b}_j$  и потому

$$\frac{\partial \hat{u}}{\partial \hat{x}_1}(\hat{x}) = p_1(\hat{x}_1) \prod_{j=1}^m (\hat{x}_2 - \hat{b}_j), \quad p_1 \in P_{m-1}.$$

Аналогично,

$$\frac{\partial \hat{u}}{\partial \hat{x}_2}(\hat{x}) = q_1(\hat{x}_2) \prod_{i=1}^m (\hat{x}_1 - \hat{b}_i), \quad q_1 \in P_{m-1}.$$

Интегрируя эти соотношения и сравнивая полученные результаты, получим утверждение леммы.  $\square$

**Лемма 7.** Пусть квадратурная формула  $\hat{S}$  совпадает с  $s \times s$ -точечной квадратурой Гаусса или Лобатто и

- i)  $s \geq m + 1$ , если  $\hat{P} = Q_m$ ;
- ii)  $s \geq m$ , если  $P_m \subset \hat{P} \subset Q_m$ .

Тогда существуют не зависящие от  $h$  постоянные  $c_1$  и  $c_2$  такие, что для любого  $u_h \in V_h$

$$c_1 |u_h|_{1,\Omega} \leq |u_h|_{1,h} \leq c_2 |u_h|_{1,\Omega}, \quad (33)$$

$$|u_h|_{0,h} \leq c_2 |u_h|_{0,\Omega}, \quad (34)$$

где  $|u_h|_{0,h}$  и  $|u_h|_{1,h}$  — величины, определенные равенствами (8).

**ДОКАЗАТЕЛЬСТВО.** Будем рассуждать, как при доказательстве леммы 2. Обоснование оценки (34) дословно повторяет обоснование неравенства (10). Чтобы убедиться в справедливости неравенств (33), достаточно доказать, что для любого конечного элемента  $e \in \mathcal{T}_h$  и любой функции  $u_h \in P_e$

$$c_1^2 \int_e |\nabla u_h(x)|^2 dx \leq S_e(|\nabla u_h|^2) \leq c_2^2 \int_e |\nabla u_h(x)|^2 dx$$

с постоянными  $c_1, c_2$ , не зависящими от  $h$ . При доказательстве леммы 2 эта задача была сведена к установлению справедливости неравенств

$$k_1 \int_{\hat{e}} |\hat{\nabla} \hat{u}(\hat{x})|^2 d\hat{x} \leq \hat{S}(|\hat{\nabla} \hat{u}|^2) \leq k_2 \int_{\hat{e}} |\hat{\nabla} \hat{u}(\hat{x})|^2 d\hat{x},$$

где  $\hat{u} = \sum_{i=1}^M \hat{u}_i \hat{\varphi}_i(\hat{x}) \in \hat{P}$ . Теперь достаточно доказать, что  $u(\hat{x}) \equiv \text{const}$ ,

если  $\hat{I}_2 \equiv \hat{S}(|\hat{\nabla} \hat{u}|^2) = 0$ . Ясно, что из равенства  $\hat{I}_2 = 0$  вытекает, что  $\partial \hat{u} / \partial \hat{x}_k = 0$ ,  $k = 1, 2$  в узлах квадратуры. В случае i) множество узлов квадратуры содержит подмножество, унисольвентное с  $Q_m$ , и  $\partial \hat{u} / \partial \hat{x}_k \in Q_m$ . Отсюда следует, что  $\partial \hat{u} / \partial \hat{x}_k \equiv 0$ , и  $\hat{u} \equiv \text{const}$ . В случае ii) мы находимся в условиях леммы 6, из которой в силу условия  $P_m \subset \hat{P} \subset Q_m$  вытекает, что  $\hat{u} \equiv c_1$ .  $\square$

Доказательство следующих двух теорем проводится так же, как и для треугольных элементов.

**Теорема 3.** Пусть для квадратурной формулы  $\hat{S}$  выполнены условия леммы 7. Тогда существуют такие положительные, не зависящие от  $h$  постоянные  $m$ ,  $M$  и  $c$ , что

$$a_h(u_h, u_h) \geq m |u_h|_{1,\Omega}^2, \quad |a_h(u_h, v_h)| \leq M |u_h|_{1,\Omega} |v_h|_{1,\Omega},$$

$$|f_h(v_h)| \leq c |v_h|_{1,\Omega}$$

для произвольных  $u_h, v_h \in V_h$ .

**Теорема 4.** Пусть квадратурная формула  $\hat{S}$  удовлетворяет условиям леммы 7,

$$u \in W_2^{m+1}(\Omega), \quad a_{ij} \in W_\infty^m(\Omega), \quad i, j = 1, 2, \quad f \in W_r^m(\Omega), \quad rm > 2, \quad r \geq 2.$$

Тогда справедлива оценка точности схемы (28):

$$\|u - u_h\|_{1,\Omega} \leq c h^m \left( |u|_{m+1,\Omega} + \sum_{i,j=1}^2 \|a_{ij}\|_{m,\infty,\Omega} \|u\|_{m+1,\Omega} + \|f\|_{m,r,\Omega} \right),$$

где постоянная  $c$  не зависит от  $h$ .

#### Задачи.

1. Целью этой задачи является проверка точности условий леммы 7 и, соответственно, теоремы 3. Пусть  $\Omega$  — единичный квадрат и его триангуляция  $\mathcal{T}_h$  состоит из прямоугольных элементов размера  $h \times h$ ;  $\hat{P} = Q_m$ ; узлы интерполяции образуют ортогональную равномерную сетку на  $\hat{e}$ ;  $\hat{S}$  есть  $m \times m$ -точечная квадратура Гаусса;  $V_h \subset H_0^1(\Omega)$  — пространство аффинно-эквивалентных прямоугольных элементов. Показать, что:

а)  $a_h(u_h, u_h) > 0$  для любого  $u_h \in V_h, u_h \neq 0$ ;

б) найдется элемент  $u_h \in V_h$  такой, что  $a_h(u_h, u_h)/|u_h|_{1,\Omega}^2 = O(h)$ , т. е. форма  $a_h$  положительно определена, но не равномерно по  $h$ .

2. Используя формулу (32), найти корни  $L_m$ , точки экстремума  $L'_m$ , определить узлы и коэффициенты квадратур Гаусса и Лобатто для  $m = 1, 2, 3$ . Убедиться, что при  $m = 1$  квадратура Лобатто совпадает с формулой трапеций:

$$\int_0^1 \hat{\phi}(\hat{x}) d\hat{x} \approx \frac{1}{2} (\hat{\phi}(0) + \hat{\phi}(1)),$$

а при  $m = 2$  — с формулой Симпсона:

$$\int_0^1 \hat{\phi}(\hat{x}) d\hat{x} \approx \frac{1}{6} (\hat{\phi}(0) + 4\hat{\phi}(1/2) + \hat{\phi}(1)).$$

3. Пусть  $\Omega$  — единичный квадрат, и его триангуляция  $\mathcal{T}_h$  состоит из квадратов со стороной  $h$ . Рассмотрим задачу Дирихле для уравнения Пуассона:

$$-\Delta u = f(x) \quad \text{при } x \in \Omega, \quad u(x) = 0 \quad \text{при } x \in \Gamma.$$

а) Построить сеточную схему, используя лагранжевы конечные элементы первого порядка ( $Q_1$ -интерполяция) и квадратурную формулу трапеций ( $m = 1$ ). Сравнить построенную сеточную схему со стандартной пятиточечной разностной схемой.

б) Прodelать то же самое, используя 1-точечную формулу Гаусса.

## § 6. Схемы с численным интегрированием для областей с произвольной границей

**1. Изопараметрические треугольные элементы.** Рассмотрим теперь случай произвольной области  $\Omega$  с кусочно гладкой липшицевой границей  $\Gamma$ . Начнем с построения области  $\bar{\Omega}_h = \bigcup e$  — аппроксимации области  $\Omega$  — как объединения элементов  $e \in \mathcal{T}_h$ . Будем предполагать, что  $\mathcal{E}_h = \{(e, P_e, \Sigma_e), e \in \mathcal{T}_h\}$  — семейство треугольных изопараметрических конечных элементов лагранжева типа, ассоциированных с одним базисным элементом  $(\hat{e}, \hat{P}, \hat{\Sigma})$ ,  $\hat{P} = P_m$ ,  $m \geq 1$ . Исходя из соображений простоты вычислений, естественно стремиться обойтись минимальным числом «истинно изопараметрических элементов», используя треугольники с прямолинейными сторонами в качестве «внутренних» элементов, и треугольники с одной криволинейной стороной — в качестве приграничных элементов (определение таких элементов дано на с. 122). Пример триангуляции области на элементы указанного типа приведен на рис. 2.

Для элементов с прямолинейными сторонами преобразование  $x = x_e(\hat{x}) = B_e \hat{x} + b_e \in [P_1]^2$  и изопараметрический элемент  $(e, P_e, \Sigma_e)$  является в то же время аффинно-эквивалентным, поскольку  $x = x_e(\hat{x}) \in [\hat{P}]^2$ . Будем предполагать далее, что триангуляция  $\mathcal{T}_h$  яв-

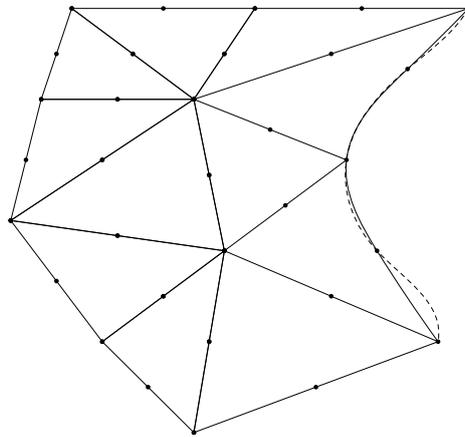


Рис. 2. Триангуляция области с частично криволинейной границей,  $m = 2$ . Используются лагранжевы аффинно-эквивалентные элементы второй степени внутри области и квадратичные изопараметрические — у криволинейной границы  $\Gamma$  (обозначенной пунктирной линией).

ляется  $(m + 1)$ -регулярной. Напомним, это означает, что для любого

$e \in \mathcal{T}_h$  (см. определение и теорему 2 на с. 105):

$$\begin{aligned} c_1 h^2 &\leq \det(Dx_e(\hat{x})) \leq c_2 h^2 \quad \forall \hat{x} \in \hat{e}, \\ c_1 h^{-2} &\leq \det(Dx_e^{-1}(x)) \leq c_2 h^{-2} \quad \forall x \in e, \\ |x_e|_{k,\infty,\hat{e}} &\leq c h^k, \quad |x_e^{-1}|_{k,\infty,e} \leq c h^{-k}, \quad k = 1, 2, \dots, m+1, \\ |\det(Dx_e)|_{k,\infty,\hat{e}} &\leq c h^{2+k}, \quad k = 1, 2, \dots, m+1. \end{aligned}$$

Пусть  $V_h \subset V = H_0^1(\Omega_h)$  — соответствующее  $\mathcal{E}_h$  пространство конечных элементов:

$$V_h = \{v_h \in C(\bar{\Omega}_h) : v_h|_e \in P_e \quad \forall e \in \mathcal{T}_h, v_h(x) = 0, x \in \Gamma_h\},$$

где  $\Gamma_h = \partial\Omega_h$  — граница области  $\Omega_h$ . Отметим, что в силу теоремы 3, с. 125, расстояние между  $\Gamma$  и  $\Gamma_h$  есть величина  $O(h^{m+1})$ , если  $\Gamma$  состоит из кривых класса  $C^{m+1}$ .

Определим схему МКЭ с численным интегрированием для задачи (5). Учитывая проведенный нами анализ точности аффинно-эквивалентных элементов, ограничимся далее квадратурами  $\hat{S}$ , удовлетворяющими следующим условиям:

$K_1$ ) коэффициенты  $\hat{S}$  положительны;

$K_2$ )  $\hat{E}(\hat{\phi}) = 0$  для любого  $\hat{\phi} \in P_{2m-2}$ .

Кроме того, будем предполагать, что:

$K_3$ ) если узел квадратуры  $\hat{S}$  лежит на границе  $\hat{e}$ , то он является узлом интерполяции элемента  $(\hat{e}, \hat{P}, \hat{\Sigma})$ .

УПРАЖНЕНИЕ 1. Считая, что узлы интерполяции на границе  $\hat{e}$  расположены равномерно, установить, какие из приведенных на с. 191 и в задаче 2, с. 192, квадратурных формул удовлетворяют ограничениям  $K_1 - K_3$ .

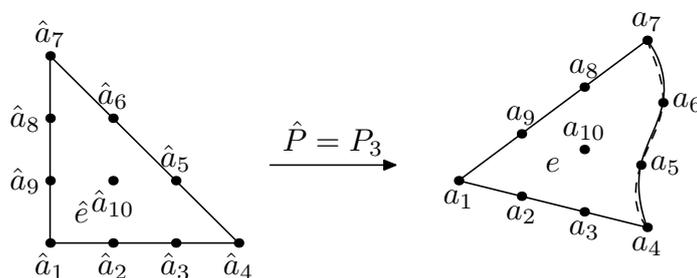


Рис. 3. Пример приграничного кубического изопараметрического элемента.

Условие  $K_3$  позволяет утверждать, что узлы  $b_i^e$  квадратуры  $S_e$ , индуцированной  $\hat{S}$ , принадлежат области  $\bar{\Omega}$  при достаточно малых  $h$  для любого элемента  $e \in \mathcal{T}_h$ . В самом деле, это так для элементов,

имеющих с  $\Gamma$  не более одной общей точки (вершины элемента). Если  $e$  — приграничный элемент (см. рис. 3) и  $b_i^e \in \partial e \cap \Gamma_h$ , то  $b_i^e$  является узлом интерполяции и также лежит на  $\Gamma$  вследствие условия  $K_3$ . Если  $b_i^e$  — внутренняя точка  $e$ , то  $b_i^e \in \Omega$  в силу следствия 1, с. 125.

Определим формы  $a_h$  и  $f_h$ , аппроксимирующие формы  $a$  и  $f$  соответственно:

$$\begin{aligned} a_h(u, v) &= S_{\Omega_h} \left( \sum_{i,j=1}^2 a_{ij} \frac{\partial u}{\partial x_j} \frac{\partial v}{\partial x_i} \right) = \sum_e S_e \left( \sum_{i,j=1}^2 a_{ij} \frac{\partial u}{\partial x_j} \frac{\partial v}{\partial x_i} \right) = \\ &= \sum_e \det(J_e) \sum_{l=1}^L \hat{c}_l \left( \sum_{i,j=1}^2 a_{ij} \frac{\partial u}{\partial x_j} \frac{\partial v}{\partial x_i} \right) (b_l^e), \quad (1) \end{aligned}$$

$$f_h(v) = S_{\Omega_h}(fv) = \sum_e \det(J_e) \sum_{l=1}^L \hat{c}_l(fv)(b_l^e), \quad (2)$$

где  $J_e(\hat{x}) = Dx_e(\hat{x})$  — матрица Якоби преобразования  $x_e^1$ . В итоге придем к схеме МКЭ с численным интегрированием:

$$\text{найти } u_h \in V_h : \quad a_h(u_h, v_h) = f_h(v_h) \quad \forall v_h \in V_h. \quad (3)$$

Положим по аналогии с принятыми ранее обозначениями

$$|u_h|_{0,h}^2 = S_{\Omega_h}(u_h^2), \quad |u_h|_{1,h}^2 = S_{\Omega_h}(|\nabla u_h|^2).$$

**Лемма 1.** *Существуют не зависящие от  $h$  постоянные  $c_1$  и  $c_2$  такие, что*

$$c_1 |u_h|_{1,\Omega_h} \leq |u_h|_{1,h} \leq c_2 |u_h|_{1,\Omega_h}, \quad (4)$$

$$|u_h|_{0,h} \leq c_2 |u_h|_{0,\Omega_h} \quad (5)$$

для любого  $u_h \in V_h$ .

**ДОКАЗАТЕЛЬСТВО.** Докажем (4). Для этого достаточно установить, что для любого конечного элемента  $e \in \mathcal{T}_h$  и любой функции  $u_h \in V_h$  справедливы неравенства

$$c_1^2 \int_e |\nabla u_h(x)|^2 dx \leq S_e(|\nabla u_h|^2) \leq c_2^2 \int_e |\nabla u_h(x)|^2 dx \quad (6)$$

с постоянными  $c_1, c_2$ , не зависящими от  $h$ .

<sup>1)</sup>Считаем, что  $\det(J_e(\hat{x})) > 0$  для любого  $\hat{x} \in \hat{e}$  и  $e \in \mathcal{T}_h$ .

Выполняя замену переменных  $x = x_e(\hat{x})$  и используя обозначение  $\hat{u}(\hat{x}) = u_h(x_e(\hat{x}))$ , получаем  $J_e^{-T}(\hat{x}) = (J_e^{-1}(\hat{x}))^T$ ,  $J_e^{-1}(\hat{x}) = Dx_e^{-1}(x)$ ,

$$|\nabla u_h(x)|^2 = |J_e^{-T} \hat{\nabla} \hat{u}(\hat{x})|^2.$$

Таким образом,

$$I_1 \equiv \int_e |\nabla u_h(x)|^2 dx = \int_{\hat{e}} \det(J_e) |J_e^{-T} \hat{\nabla} \hat{u}(\hat{x})|^2 d\hat{x},$$

$$S_1 \equiv S_e(|\nabla u_h|^2) = \hat{S}(\det(J_e) |J_e^{-T} \hat{\nabla} \hat{u}|^2).$$

Ранее было доказано (см. с. 183), что

$$k_1 \int_{\hat{e}} |\hat{\nabla} \hat{u}(\hat{x})|^2 d\hat{x} \leq \hat{S}(|\hat{\nabla} \hat{u}|^2) \leq k_2 \int_{\hat{e}} |\hat{\nabla} \hat{u}(\hat{x})|^2 d\hat{x}.$$

Учитывая оценку  $|y| \leq \|J_e^T\| |J_e^{-T} y|$  и равенство  $\|J_e\| = \|J_e^T\|$ , получим

$$I_1 \leq |\det(J_e)|_{0,\infty,\hat{e}} \|J_e^{-T}\|_{0,\infty,\hat{e}}^2 \int_{\hat{e}} |\hat{\nabla} \hat{u}(\hat{x})|^2 d\hat{x} \leq$$

$$\leq |\det(J_e)|_{0,\infty,\hat{e}} \|J_e^{-1}\|_{0,\infty,\hat{e}}^2 / k_1 \hat{S}(|\hat{\nabla} \hat{u}|^2) \leq$$

$$\leq d/k_1 \hat{S}(\det(J_e) |J_e^{-T} \hat{\nabla} \hat{u}|^2) = d/k_1 S_1,$$

т. е. оценку снизу в (6) с постоянной  $c_1^2 = k_1/d$ . Здесь

$$d = \frac{|\det(J_e)|_{0,\infty,\hat{e}}}{\min_{\hat{x} \in \hat{e}} \det(J_e(\hat{x}))} \|J_e^{-1}\|_{0,\infty,\hat{e}}^2 \|J_e\|_{0,\infty,\hat{e}}^2.$$

В силу регулярности  $\mathcal{T}_h$  имеем  $\|J_e\|_{0,\infty,\hat{e}} \leq ch$ ,  $c^{-1}h^2 \leq \det(J_e) \leq ch^2$ ,

$$\|J_e^{-1}\|_{0,\infty,\hat{e}} = \|Dx_e^{-1}\|_{0,\infty,e} \leq ch^{-1},$$

поэтому  $d \leq c$ , причем  $c$  не зависит от  $h$ . Аналогично доказывается оценка сверху в (6) с постоянной  $c_2 = k_2 d$ . Доказательство неравенства (5) оставляем читателю в качестве упражнения.  $\square$

Из леммы 1 вытекает

**Теорема 1.** Пусть квадратурная формула  $\hat{S}$  удовлетворяет условиям  $K_1$ – $K_3$ . Тогда существуют положительные не зависящие от  $h$  постоянные  $m$ ,  $M$  и  $c$  такие, что  $\forall u_h, v_h \in V_h$

$$a_h(u_h, u_h) \geq m |u_h|_{1,\Omega_h}^2, \quad |a_h(u_h, v_h)| \leq M |u_h|_{1,\Omega_h} |v_h|_{1,\Omega_h}, \quad (7)$$

$$|f_h(v_h)| \leq c |v_h|_{1, \Omega_h}.$$

**Следствие 1.** *Схема МКЭ с численным интегрированием (3) однозначно разрешима.*

Перейдем к оценке точности схемы (3). При этом мы сразу сталкиваемся с двумя затруднениями, не встречавшимися ранее.

Первое затруднение, однако, не очень серьезное и часто встречающееся при обосновании приближенных методов, связано с понятием близости решений  $u_h \in V_h$  и  $u \in V$  и вызвано тем, что  $u_h$  и  $u$  определены на разных множествах:  $\Omega_h$  и  $\Omega$  соответственно.

Второе затруднение более серьезно. Поскольку  $\Omega_h \neq \Omega$ , то  $V_h$  не является подпространством  $V$  и, как следствие, мы не можем трактовать схему (3) как метод Галеркина с возмущениями для задачи (5) и использовать лемму Стренга для оценки ее точности.

Первая трудность преодолевается следующим образом. Так как по построению  $\Gamma_h$  и  $\Gamma$  близки, то ясно, что существует такое ограниченное открытое множество  $\tilde{\Omega}$ , что  $\Omega, \Omega_h \subset \tilde{\Omega}$  для всех рассматриваемых триангуляций  $\mathcal{T}_h$  (при достаточно малых  $h$ ). Следовательно, согласно теореме 4, с. 40, функцию  $u \in W_2^{m+1}(\Omega)$  можно продолжить на область  $\tilde{\Omega}$  с сохранением нормы. Точнее, существует функция  $\tilde{u} \in W_2^{m+1}(\tilde{\Omega})$  такая, что  $\tilde{u}(x) = u(x)$  при  $x \in \Omega$  и

$$\|\tilde{u}\|_{m+1, \tilde{\Omega}} \leq c \|u\|_{m+1, \Omega},$$

где  $c$  — постоянная, зависящая от областей  $\Omega$  и  $\tilde{\Omega}$ , но не зависящая от  $u$ . Положим  $\tilde{V} = W_2^1(\Omega_h)$ . По определению имеем  $V_h \subset \tilde{V}$ . Поскольку функции  $\tilde{u}$  и  $u_h$  принадлежат  $\tilde{V}$ , то под погрешностью схемы естественно понимать  $\|\tilde{u} - u_h\|_{1, \Omega_h}$ .

Для преодоления второй трудности получим оценку, аналогичную установленной в лемме Стренга, с. 27. С этой целью определим такие продолжения  $\tilde{a}_{ij}(x)$  коэффициентов исходного уравнения, что

$$\|\tilde{a}_{ij}\|_{m, \infty, \tilde{\Omega}} \leq c \|a_{ij}\|_{m, \infty, \Omega} \leq M = \text{const}, \quad i, j = 1, 2, \quad (8)$$

и положим

$$\tilde{f}(x) = - \sum_{i,j=1}^2 \frac{\partial}{\partial x_i} \left( \tilde{a}_{ij}(x) \frac{\partial \tilde{u}(x)}{\partial x_j} \right), \quad x \in \tilde{\Omega}. \quad (9)$$

Отметим, что  $\tilde{f}$  является продолжением  $f$  на  $\tilde{\Omega}$ . Умножим обе части равенства (9) на произвольную функцию  $v_h \in V_h$  и проинтегрируем почленно по области  $\Omega_h$ . Используя формулу интегрирования по

частям и учитывая, что  $v_h = 0$  на  $\Gamma_h$ , получим тождество:

$$\tilde{u} \in \tilde{V} : \quad \tilde{a}(\tilde{u}, v_h) = \tilde{f}(v_h) \quad \forall v_h \in V_h, \quad (10)$$

где

$$\tilde{a}(\tilde{u}, v_h) = \sum_{i,j=1}^2 \int_{\Omega_h} \tilde{a}_{ij}(x) \frac{\partial \tilde{u}(x)}{\partial x_j} \frac{\partial v_h}{\partial x_i} dx, \quad \tilde{f}(v_h) = \int_{\Omega_h} \tilde{f} v_h dx.$$

Теперь нетрудно заметить, что исследуемую схему МКЭ с численным интегрированием можно трактовать как аппроксимацию тождества (10). Действительно, заменяя в (10) пространство  $\tilde{V}$  его подпространством  $V_h$ , а интегралы по области  $\Omega_h$  — составной квадратурой формулой  $S_{\Omega_h}$ , придем к схеме (3). Справедливость равенств

$$S_{\Omega_h} \left( \tilde{a}_{ij} \frac{\partial u_h}{\partial x_j} \frac{\partial v_h}{\partial x_i} \right) = S_{\Omega_h} \left( a_{ij} \frac{\partial u_h}{\partial x_j} \frac{\partial v_h}{\partial x_i} \right), \quad S_{\Omega_h}(\tilde{f} v_h) = S_{\Omega_h}(f v_h)$$

для произвольных  $u_h, v_h \in V_h$  обеспечивается тем, что все узлы квадратуры  $S_{\Omega_h}$  лежат в  $\bar{\Omega}$ .

ЗАМЕЧАНИЕ 1. Подчеркнем, что тождество (10) лишь характеризует функцию  $\tilde{u}$  и не служит для ее фактического определения. В частности, функции из  $\tilde{V}$ , в отличие от элементов  $V_h$ , не удовлетворяют краевому условию  $\tilde{u} = 0$  на  $\Gamma_h$ .

Сказанное выше позволяет, фактически, воспроизвести рассуждения, применявшиеся при доказательстве леммы Стренга, с. 27, и на основании теоремы 1 получить оценку:

$$|\tilde{u} - u_h|_{1, \Omega_h} \leq c \left\{ \inf_{v_h \in V_h} \left( |\tilde{u} - v_h|_{1, \Omega_h} + \sup_{w_h \in V_h} \frac{|\tilde{a}(v_h, w_h) - a_h(v_h, w_h)|}{|w_h|_{1, \Omega_h}} \right) + \sup_{w_h \in V_h} \frac{|\tilde{f}(w_h) - f_h(w_h)|}{|w_h|_{1, \Omega_h}} \right\}. \quad (11)$$

Действительно, для любого элемента  $v_h \in V_h$  справедливо тождество

$$a_h(u_h - v_h, u_h - v_h) = \tilde{a}(\tilde{u} - v_h, u_h - v_h) + \{ \tilde{a}(v_h, u_h - v_h) - a_h(v_h, u_h - v_h) \} + \{ f_h(u_h - v_h) - \tilde{f}(u_h - v_h) \},$$

откуда, используя оценки (7), (8), получим

$$m|u_h - v_h|_{1, \Omega_h}^2 \leq \left\{ M|\tilde{u} - v_h|_{1, \Omega_h} + \frac{|\tilde{a}(v_h, u_h - v_h) - a_h(v_h, u_h - v_h)|}{|u_h - v_h|_{1, \Omega_h}} + \frac{|f_h(u_h - v_h) - \tilde{f}(u_h - v_h)|}{|u_h - v_h|_{1, \Omega_h}} \right\} |u_h - v_h|_{1, \Omega_h},$$

или

$$m|u_h - v_h|_{1,\Omega_h} \leq M|\tilde{u} - v_h|_{1,\Omega_h} + \sup_{w_h \in V_h} \frac{|\tilde{a}(v_h, w_h) - a_h(v_h, w_h)|}{|w_h|_{1,\Omega_h}} + \sup_{w_h \in V_h} \frac{|f_h(w_h) - \tilde{f}(w_h)|}{|w_h|_{1,\Omega_h}}.$$

Используя неравенство треугольника

$$|\tilde{u} - u_h|_{1,\Omega_h} \leq |\tilde{u} - v_h|_{1,\Omega_h} + |u_h - v_h|_{1,\Omega_h}$$

и вычисляя нижнюю грань по  $v_h \in V_h$ , получим (11).

Отсюда, полагая  $v_h = \tilde{u}_I$ , где  $\tilde{u}_I$  есть  $V_h$ -интерполюнт функции  $\tilde{u}$ , будем иметь<sup>1)</sup>

$$|\tilde{u} - u_h|_{1,\Omega_h} \leq c \{ |\tilde{u} - \tilde{u}_I|_{1,\Omega_h} + E_a(\tilde{u}_I) + E_f \}, \quad (12)$$

$$E_a(\tilde{u}_I) = \sup_{w_h \in V_h} \frac{|\tilde{a}(\tilde{u}_I, w_h) - a_h(\tilde{u}_I, w_h)|}{|w_h|_{1,\Omega_h}}, \quad E_f = \sup_{w_h \in V_h} \frac{|\tilde{f}(w_h) - f_h(w_h)|}{|w_h|_{1,\Omega_h}}.$$

Оценка погрешности интерполяции  $|\tilde{u} - \tilde{u}_I|_{1,\Omega_h}$  вытекает из теоремы 6, с. 154, (см. также замечание к ней) при  $s = m$ ,  $n = 2$ ,  $r = 0$ ,  $p = 2$ ,  $\Omega = \Omega_h$ . Согласно этой теореме

$$|\tilde{u} - \tilde{u}_I|_{1,\Omega_h} \leq c h^m \|\tilde{u}\|_{m+1,\Omega_h}.$$

Учитывая, что

$$\|\tilde{u}\|_{m+1,\Omega_h} \leq \|\tilde{u}\|_{m+1,\tilde{\Omega}} \leq c \|u\|_{m+1,\Omega},$$

для первой составляющей оценки погрешности (12) получим

$$|\tilde{u} - \tilde{u}_I|_{1,\Omega_h} \leq c h^m \|u\|_{m+1,\Omega}. \quad (13)$$

Оценки остальных слагаемых вытекают из следующей леммы, являющейся аналогом леммы 4, с. 187.

**Лемма 2.** Пусть квадратурная формула  $\hat{S}$  точна на полиномах из  $P_{2m-2}$ ,

$$a_{ij} \in W_\infty^m(e), \quad i, j = 1, 2, \quad f \in W_r^m(e), \quad rm > 2, \quad r \geq 1.$$

Тогда для любых функций  $p, q \in P_e$  и любого  $e \in \mathcal{T}_h$  справедливы оценки:

$$\left| E_e \left( \sum_{i,j=1}^2 a_{ij} \frac{\partial q}{\partial x_j} \frac{\partial p}{\partial x_i} \right) \right| \leq c h^m \sum_{i,j=1}^2 \|a_{ij}\|_{m,\infty,e} \|q\|_{m,e} |p|_{1,e}, \quad (14)$$

<sup>1)</sup>Мы учли здесь, что  $\tilde{u}_I \in V_h$ , поскольку  $\tilde{u}$  равна нулю в узлах, принадлежащих  $\Gamma$ .

$$|E_e(fp)| \leq ch^m (\text{mes}(e))^{1/2-1/r} \|f\|_{m,r,e} \|p\|_{1,e} \quad (15)$$

с постоянной  $c$ , не зависящей от  $h$ .

ДОКАЗАТЕЛЬСТВО. Докажем сначала неравенство (14). Перейдем на базисный элемент (см. также (1), с. 177). Получим:

$$I_e \equiv E_e \left( \sum_{i,j=1}^2 a_{ij} \frac{\partial q}{\partial x_j} \frac{\partial p}{\partial x_i} \right) = \hat{E} \left( \sum_{i,j=1}^2 \hat{b}_{ij} \frac{\partial \hat{q}}{\partial \hat{x}_j} \frac{\partial \hat{p}}{\partial \hat{x}_i} \right),$$

где

$$\hat{B}(\hat{x}) = \{\hat{b}_{ij}\}_{i,j=1}^2 = \det(J_e(\hat{x})) J_e^{-1}(\hat{x}) \hat{A} J_e^{-T}(\hat{x}).$$

Принимая во внимание выражения для элементов матрицы  $J_e^{-1}$ , будем иметь отсюда, что  $\hat{b}_{ij}$  является суммой слагаемых вида

$$\hat{a}(\hat{x}) = \frac{1}{\det(J_e(\hat{x}))} \frac{\partial x_i(\hat{x})}{\partial \hat{x}_k} \frac{\partial x_j(\hat{x})}{\partial \hat{x}_l} \hat{a}_{kl}.$$

Следовательно, оценка  $I_e$  сводится к оценке величины  $E \equiv \hat{E} \left( \hat{a} \frac{\partial \hat{q}}{\partial \hat{x}_j} \frac{\partial \hat{p}}{\partial \hat{x}_i} \right)$ ,

где  $\hat{p}, \hat{q} \in P_m$ . Введем обозначения:

$$\hat{u} = \frac{\partial \hat{q}}{\partial \hat{x}_j} \in P_{m-1}, \quad \hat{v} = \frac{\partial \hat{p}}{\partial \hat{x}_i} \in P_{m-1}, \quad \hat{\phi} = \hat{a} \hat{u}.$$

Тогда  $E = \hat{E}(\hat{\phi} \hat{v})$ , причем  $\hat{\phi} \in W_\infty^m(\hat{e})$ ,  $\hat{v} \in P_{m-1}$ . Оценка функционала  $E$  была получена ранее при рассмотрении аффинных элементов (см. доказательство леммы 4, с. 187):

$$\begin{aligned} |E| &\leq c \left( \sum_{j=0}^{m-1} |\hat{a}|_{m-j,\infty,\hat{e}} |\hat{u}|_{j,\hat{e}} \right) |\hat{v}|_{0,\hat{e}} \leq \\ &\leq c \left( \sum_{j=0}^{m-1} |\hat{a}|_{m-j,\infty,\hat{e}} |\hat{q}|_{j+1,\hat{e}} \right) |\hat{p}|_{1,\hat{e}}. \quad (16) \end{aligned}$$

Осталось перейти в полученной оценке на элемент  $e$ . Из теоремы 5, с. 152, следует, что

$$|\hat{q}|_{j+1,\hat{e}} \leq c h^j \|q\|_{j+1,e}, \quad |\hat{p}|_{1,\hat{e}} \leq c |p|_{1,e}. \quad (17)$$

Далее, используя лемму 3, с. 187, получим

$$|\hat{a}|_{s,\infty,\hat{e}} \leq c \sum_{t=0}^s \left| \frac{1}{\det(J_e(\hat{x}))} \frac{\partial x_i(\hat{x})}{\partial \hat{x}_k} \frac{\partial x_j(\hat{x})}{\partial \hat{x}_l} \right|_{t,\infty,\hat{e}} |\hat{a}_{kl}|_{s-t,\infty,\hat{e}}.$$

Аналогично неравенствам (17) доказывается оценка:  $|\hat{a}_{kl}|_{s-t, \infty, \hat{e}} \leq c h^{s-t} \|a_{kl}\|_{s-t, \infty, e}$ . Из условия  $(m+1)$ -регулярности триангуляции нетрудно вывести, что

$$\left| \frac{1}{\det(J_e(\hat{x}))} \frac{\partial x_i(\hat{x})}{\partial \hat{x}_k} \frac{\partial x_j(\hat{x})}{\partial \hat{x}_l} \right|_{t, \infty, \hat{e}} \leq c h^t, \quad 0 \leq t \leq m.$$

Таким образом,

$$|\hat{a}|_{s, \infty, \hat{e}} \leq c h^s \|a_{kl}\|_{s, \infty, e}, \quad 0 \leq s \leq m. \quad (18)$$

Наконец, используя неравенства (17) и (18) для оценки правой части неравенства (16), получим, что

$$|E| = |\hat{E}(\hat{\phi}\hat{v})| \leq c h^m \|a_{kl}\|_{m, \infty, e} \|q\|_{m, e} |p|_{1, e},$$

откуда вытекает неравенство (14).

Схема доказательства оценки (15) аналогична аффинному случаю. Имеем:

$$E_e(fp) = \hat{E}(\det(Dx_e)\hat{f}\hat{p}) = \hat{E}(\hat{\phi}\hat{p}_0) + \hat{E}(\hat{\phi}(\hat{p} - \hat{p}_0)) \equiv E_0(\hat{\phi}) + E_1(\hat{\phi}),$$

где  $\hat{p} \in P_m$ ,  $\hat{\phi} = \det(Dx_e)\hat{f}$ ,  $\hat{p}_0$  есть  $P_0$ -интерполюант  $\hat{p}$ . Оценка  $E_0$  также соответствует аффинному случаю:

$$|E_0(\hat{\phi})| \leq c |\hat{\phi}|_{m, r, \hat{e}} |\hat{p}|_{0, \hat{e}}.$$

Поскольку  $|\hat{p}|_{k, \hat{e}} \leq c h^{k-1} \|p\|_{k, e}$ ,  $k \geq 0$ ,

$$|\hat{\phi}|_{m, r, \hat{e}} = |\det(Dx_e)\hat{f}|_{m, r, \hat{e}} \leq c \sum_{j=0}^m |\det(Dx_e)|_{j, r, \hat{e}} |\hat{f}|_{m-j, r, \hat{e}},$$

и, кроме того,

$$\begin{aligned} |\det(Dx_e)|_{j, r, \hat{e}} &\leq c |\det(Dx_e)|_{j, \infty, \hat{e}} \leq c h^{j+2}, \\ |\hat{f}|_{m-j, r, \hat{e}} &\leq c h^{m-j-2/r} \|f\|_{m-j, r, e}, \end{aligned}$$

то

$$|E_0(\hat{\phi})| \leq c h^{m+1-2/r} \|f\|_{m, r, e} \|p\|_{1, e}.$$

Учитывая, что  $h^{1-2/r} \leq c (\text{mes}(e))^{1/2-1/r}$ , получаем требуемую оценку  $E_0$ . Оценку  $E_1$  оставляем читателю в качестве упражнения.  $\square$

УПРАЖНЕНИЕ 2. В условиях леммы 2 доказать справедливость оценки

$$\left| E_e \left( \sum_{i,j=1}^2 a_{ij} \frac{\partial q}{\partial x_j} \frac{\partial p}{\partial x_i} \right) \right| \leq ch \sum_{i,j=1}^2 \|a_{ij}\|_{1, \infty, e} |q|_{1, e} |p|_{1, e}.$$

Используя этот факт, дать другое доказательство равномерной положительной определенности формы  $a_h$  на  $V_h$  при достаточно малых  $h$ .

**Теорема 2.** Пусть квадратурная формула  $\hat{S}$  удовлетворяет ограничениям  $K_1-K_3$ ,

$$u \in W_2^{m+1}(\Omega), \quad a_{ij} \in W_\infty^m(\Omega), \quad i, j = 1, 2, \quad \tilde{f} \in W_r^m(\Omega), \quad rm > 2, \quad r \geq 2.$$

Тогда справедлива следующая оценка точности схемы (3)

$$\|\tilde{u} - u_h\|_{1,\Omega} \leq ch^m \left( \|u\|_{m+1,\Omega} + \sum_{i,j=1}^2 \|a_{ij}\|_{m,\infty,\Omega} \|u\|_{m+1,\Omega} + \|\tilde{f}\|_{m,r,\tilde{\Omega}} \right), \quad (19)$$

где постоянная  $c$  не зависит от  $h$ .

ДОКАЗАТЕЛЬСТВО. Достаточно оценить  $E_a(\tilde{u}_I)$  и  $E_f$  в правой части неравенства (12). Используя лемму 2, получим

$$\begin{aligned} |\tilde{a}(\tilde{u}_I, w_h) - a_h(\tilde{u}_I, w_h)| &= \left| \sum_e E_e \left( \sum_{i,j=1}^2 \tilde{a}_{ij} \frac{\partial \tilde{u}_I}{\partial x_j} \frac{\partial w_h}{\partial x_i} \right) \right| \leq \\ &\leq ch^m \sum_e \sum_{i,j=1}^2 \|\tilde{a}_{ij}\|_{m,\infty,e} \|\tilde{u}_I\|_{m,e} |w_h|_{1,e} \leq \\ &\leq ch^m \sum_{i,j=1}^2 \|\tilde{a}_{ij}\|_{m,\infty,\tilde{\Omega}} \sum_e \|\tilde{u}_I\|_{m,e} |w_h|_{1,e} \leq \\ &\leq ch^m \sum_{i,j=1}^2 \|\tilde{a}_{ij}\|_{m,\infty,\tilde{\Omega}} \|\tilde{u}_I\|_{m,\Omega_h} |w_h|_{1,\Omega_h} \leq \\ &\leq ch^m \sum_{i,j=1}^2 \|a_{ij}\|_{m,\infty,\Omega} \|u\|_{m+1,\Omega} |w_h|_{1,\Omega_h}. \quad (20) \end{aligned}$$

Здесь мы учли, что

$$\|\tilde{u}_I\|_{m,\Omega_h} \leq \|\tilde{u} - \tilde{u}_I\|_{m,\Omega_h} + \|\tilde{u}\|_{m,\Omega_h} \leq c \|\tilde{u}\|_{m+1,\Omega_h} \leq c \|u\|_{m+1,\Omega}.$$

Из неравенства (20) следует, что

$$E_a(\tilde{u}_I) \leq ch^m \sum_{i,j=1}^2 \|a_{ij}\|_{m,\infty,\Omega} \|u\|_{m+1,\Omega}. \quad (21)$$

Аналогично, используя неравенство Гельдера (26), с. 191, и эквивалентность норм  $\|w_h\|_{1,\Omega_h}$  и  $|w_h|_{1,\Omega_h}$  на пространстве  $V_h$ , получим:

$$\begin{aligned} |\tilde{f}(w_h) - f_h(w_h)| &= \left| \sum_e E_e(\tilde{f}w_h) \right| \leq c h^m \sum_e |e|^{1/2-1/r} \|\tilde{f}\|_{m,r,e} \|w_h\|_{1,e} \leq \\ &\leq c h^m |\Omega_h|^{1/2-1/r} \|\tilde{f}\|_{m,r,\Omega_h} \|w_h\|_{1,\Omega_h} \leq c h^m \|\tilde{f}\|_{m,r,\tilde{\Omega}} |w_h|_{1,\Omega_h}, \end{aligned}$$

где  $|e| = \text{mes}(e)$ . Отсюда вытекает требуемая оценка:

$$E_f \leq c h^m \|\tilde{f}\|_{m,r,\tilde{\Omega}}. \quad (22)$$

Подставляя теперь оценки (13), (21), (22) в неравенство (12), получим (19).  $\square$

**2. Произвольная область. Изопараметрические прямоугольные элементы.** Предполагается, что читатель знаком с материалом, изложенным в предыдущих двух параграфах. Будем использовать все принятые в предыдущем параграфе обозначения и определения, за исключением определения базисного конечного элемента и квадратурной формулы  $\hat{S}$ .

Как и ранее, начнем с построения области  $\bar{\Omega}_h = \bigcup e$ , аппроксимирующей область  $\Omega$  семейством изопараметрических прямоугольных элементов  $e \in \mathcal{T}_h$ , и определим  $\mathcal{E}_h = \{(e, P_e, \Sigma_e), e \in \mathcal{T}_h\}$  — семейство прямоугольных изопараметрических конечных элементов лагранжевого типа, ассоциированных с базисным элементом  $(\hat{e}, \hat{P}, \hat{\Sigma})$ ,  $\hat{P} = Q_m$ ,  $m \geq 1$ . Будем предполагать далее, что триангуляция  $\mathcal{T}_h$  является  $(m+1)$ -регулярной. Напомним, это означает, что для любого  $e \in \mathcal{T}_h$  (см. определение и теорему 2 на с. 105):

$$\begin{aligned} c_1 h^2 &\leq \det(Dx_e(\hat{x})) \leq c_2 h^2 \quad \forall \hat{x} \in \hat{e}, \\ c_1 h^{-2} &\leq \det(Dx_e^{-1}(x)) \leq c_2 h^{-2} \quad \forall x \in e, \\ [x_e]_{k,\infty,\hat{e}} &\leq c h^k, \quad |x_e|_{k,\infty,\hat{e}} \leq c h, \\ |x_e^{-1}|_{k,\infty,e} &\leq c h^{-k}, \quad k = 1, 2, \dots, m+1. \end{aligned}$$

Пусть  $V_h \subset H_0^1(\Omega_h)$  — соответствующее  $\mathcal{E}_h$  пространство конечных элементов:

$$V_h = \{v_h \in C(\bar{\Omega}_h) : v_h|_e \in P_e \quad \forall e \in \mathcal{T}_h, v_h(x) = 0, x \in \Gamma_h\}.$$

При определении схемы МКЭ с численным интегрированием для задачи (5) учтем проведенный нами анализ точности аффинно-эквивалентных прямоугольных элементов. Будем предполагать, что квадратура  $\hat{S}$  является  $s \times s$ -точечной квадратурой Гаусса,  $s \geq m+1$ .

В этом случае все узлы  $b_i^e$  квадратуры  $S_e$ , индуцированной  $\hat{S}$ , для любого элемента  $e \in \mathcal{T}_h$  принадлежат области  $\Omega$ , если  $h$  достаточно мало.

Определяя дискретные формы соотношениями (1), (2), с. 202, придем к схеме МКЭ с численным интегрированием:

$$\text{найти } u_h \in V_h : \quad a_h(u_h, v_h) = f_h(v_h) \quad \forall v_h \in V_h. \quad (23)$$

Справедливость следующей теоремы устанавливается аналогично случаю треугольных изопараметрических элементов.

**Теорема 3.** *Существуют положительные не зависящие от  $h$  постоянные  $m$ ,  $M$  и  $c$  такие, что*

$$a_h(u_h, u_h) \geq m |u_h|_{1, \Omega_h}^2, \quad |a_h(u_h, v_h)| \leq M |u_h|_{1, \Omega_h} |v_h|_{1, \Omega_h},$$

$$|f_h(v_h)| \leq c |v_h|_{1, \Omega_h}$$

для произвольных  $u_h, v_h \in V_h$ .

Доказательство теоремы основывается на лемме 1, с. 202. Эта лемма справедлива и в рассматриваемом случае, поскольку при ее доказательстве использовалась лишь 1-регулярность триангуляции, что для треугольных и прямоугольных изопараметрических элементов означает одно и то же. Из теоремы 3 вытекает

**Следствие 2.** *Схема МКЭ с численным интегрированием (23) однозначно разрешима.*

Оценка точности схемы (23), как и в случае треугольных изопараметрических элементов, имеет вид

$$|\tilde{u} - u_h|_{1, \Omega_h} \leq c \{ |\tilde{u} - \tilde{u}_I|_{1, \Omega_h} + E_a(\tilde{u}_I) + E_f \}, \quad (24)$$

где

$$E_a(\tilde{u}_I) = \sup_{w_h \in V_h} \frac{|\tilde{a}(\tilde{u}_I, w_h) - a_h(\tilde{u}_I, w_h)|}{|w_h|_{1, \Omega_h}}, \quad E_f = \sup_{w_h \in V_h} \frac{|\tilde{f}(w_h) - f_h(w_h)|}{|w_h|_{1, \Omega_h}}.$$

Оценку погрешности интерполяции  $|\tilde{u} - \tilde{u}_I|_{1, \Omega_h}$  получим на основании теоремы 8, с. 156, в которой надо положить  $m = 1$ ,  $s = m$ ,  $n = 2$ ,  $p = 2$ ,  $r = 0$ ,  $\Omega = \Omega_h$ . В результате будем иметь

$$|\tilde{u} - \tilde{u}_I|_{1, \Omega_h} \leq c h^m \|\tilde{u}\|_{m+1, \Omega_h} \leq c h^m \|u\|_{m+1, \Omega}. \quad (25)$$

Оценки погрешности квадратурных формул даны в следующей лемме.

**Лемма 3.** Пусть квадратурная формула  $\hat{S}$  точна на полиномах из  $Q_{3m-2}$ , триангуляция  $\mathcal{T}_h$  является  $(m+1)$ -регулярной. Тогда

$$\left| E_e \left( \sum_{i,j=1}^2 a_{ij} \frac{\partial q}{\partial x_j} \frac{\partial p}{\partial x_i} \right) \right| \leq ch^m |e|^{1/2-1/r} \sum_{i,j=1}^2 \|a_{ij}\|_{m,\infty,e} \|q\|_{m+1,r,e} |p|_{1,e}. \quad (26)$$

Если  $\hat{S}$  точна на полиномах из  $Q_{4m-3}$ , то

$$|E_e(fp)| \leq ch^m |e|^{1/2-1/r} \|f\|_{m,r,e} \|p\|_{1,e}. \quad (27)$$

Здесь  $|e| = \text{mes}(e)$ ,  $p, q \in P_e$ ,  $e \in \mathcal{T}_h$ ,

$$a_{ij} \in W_\infty^m(e), \quad i, j = 1, 2, \quad f \in W_r^m(e), \quad rm > 2, \quad r \geq 1.$$

**ДОКАЗАТЕЛЬСТВО.** Докажем (26). Достаточно оценить слагаемое  $E_e(v\partial p/\partial x_i)$ , где  $v = a_{ij}\partial q/\partial x_j$ . Перейдем на базисный элемент. Имеем:

$$\widehat{\nabla} p = J_e^{-T} \widehat{\nabla} \hat{p}, \quad E_e(v\partial p/\partial x_i) = \hat{E}(\det(J_e) \widehat{v} \widehat{\partial p/\partial x_i}),$$

$$\begin{aligned} \det(J_e) \frac{\widehat{\partial p}}{\partial x_1} &= \frac{\partial x_2}{\partial \hat{x}_2} \frac{\partial \hat{p}}{\partial \hat{x}_1} - \frac{\partial x_2}{\partial \hat{x}_1} \frac{\partial \hat{p}}{\partial \hat{x}_2}, \\ \det(J_e) \frac{\widehat{\partial p}}{\partial x_2} &= -\frac{\partial x_1}{\partial \hat{x}_2} \frac{\partial \hat{p}}{\partial \hat{x}_1} + \frac{\partial x_1}{\partial \hat{x}_1} \frac{\partial \hat{p}}{\partial \hat{x}_2}. \end{aligned}$$

Отсюда следует, что величина  $E_e(v\partial p/\partial x_i)$  представляет собой сумму двух слагаемых вида

$$\hat{J}(\hat{v}) = \hat{E}(\hat{v}\hat{g}), \quad \hat{g} = \frac{\partial x_k}{\partial \hat{x}_i} \frac{\partial \hat{p}}{\partial \hat{x}_j}, \quad i \neq j,$$

причем  $\hat{g} \in Q_{2m-1}$ . Оценим функционал  $\hat{J}(\hat{v})$ , учитывая, что  $\hat{J}(Q_{m-1}) = 0$ . Имеем

$$|\hat{J}(\hat{v})| \leq 2|\hat{v}|_{0,\infty,\hat{e}} |\hat{g}|_{0,\infty,\hat{e}} \leq c \|\hat{v}\|_{m,r,\hat{e}} |\hat{g}|_{0,\hat{e}} \leq c [\hat{v}]_{m,r,\hat{e}} |\hat{g}|_{0,\hat{e}}.$$

Осталось перейти в полученной оценке на элемент  $e$ , принимая во внимание, что  $|\hat{g}|_{0,\hat{e}} \leq ch|\hat{p}|_{1,\hat{e}}$ . Таким образом,

$$|\hat{J}(\hat{v})| \leq ch^m (\text{mes}(e))^{1/2-1/r} \|v\|_{m,r,e} |p|_{1,e}.$$

Используя здесь лемму 3, с. 187, получим оценку (26). Аналогично,

$$E_e(fp) = \hat{E}(\det(J_e)\hat{f}\hat{p}) = \hat{E}(\hat{f}\det(J_e)\hat{p}_0) + \hat{E}(\hat{f}\det(J_e)(\hat{p} - \hat{p}_0)) = J_0(\hat{f}) + J_1(\hat{f}). \quad (28)$$

Здесь  $\hat{p}_0$  есть  $P_0$ -интерполянт  $\hat{p}$ ,  $\det(J_e) \in Q_{2m-1}$ . Введем обозначения:

$$g_0 = \det(J_e)\hat{p}_0 \in Q_{2m-1}, \quad g_1 = \det(J_e)(\hat{p} - \hat{p}_0) \in Q_{3m-1}.$$

Поскольку  $\det(J_e) = O(h^2)$ , то

$$|g_0|_{0,\hat{e}} \leq ch^2|\hat{p}_0|_{0,\hat{e}} \leq ch^2|\hat{p}|_{0,\hat{e}}, \quad |g_1|_{0,\hat{e}} \leq ch^2|\hat{p} - \hat{p}_0|_{0,\hat{e}} \leq ch^2|\hat{p}|_{1,\hat{e}}.$$

Далее,  $J_0(Q_{m-1}) = 0$ ,  $J_1(Q_\alpha) = 0$ , где  $\alpha = \max\{0, m-2\}$ , поэтому

$$|\hat{J}_0(\hat{f})| \leq c\|\hat{f}\|_{m,r,\hat{e}}|\hat{g}_0|_{0,\hat{e}} \leq c[\hat{f}]_{m,r,\hat{e}}|\hat{g}_0|_{0,\hat{e}} \leq \leq ch^2[\hat{f}]_{m,r,\hat{e}}|\hat{p}|_{0,\hat{e}} \leq ch^m(\text{mes}(e))^{1/2-1/r}\|f\|_{m,r,e}\|p\|_{0,e},$$

$$|\hat{J}_1(\hat{f})| \leq c\|\hat{f}\|_{m,r,\hat{e}}|\hat{g}_1|_{0,\hat{e}} \leq c([\hat{f}]_{m-1,r,\hat{e}} + [\hat{f}]_{m,r,\hat{e}})|\hat{g}_1|_{0,\hat{e}} \leq \leq ch^2([\hat{f}]_{m-1,r,\hat{e}} + [\hat{f}]_{m,r,\hat{e}})|\hat{p}|_{1,\hat{e}} \leq ch^m(\text{mes}(e))^{1/2-1/r}\|f\|_{m,r,e}\|p\|_{1,e}.$$

Из этих оценок и равенства (28) следует (27).  $\square$

Поскольку  $4m-3 \geq 3m-2$  для любого  $m \geq 1$ , то из доказанной леммы следует, что квадратурная формула, используемая для вычисления формы  $f$ , должна быть точнее, чем квадратура, соответствующая форме  $a$ . Если используется одна и та же  $s \times s$ -точечная квадратура Гаусса и для  $f$ , и для  $a$ , то необходимо положить  $2s-1 \geq 4m-3$ ,  $s \geq m+1$  (для обеспечения эллиптичности формы  $a_h$ ). Отсюда получаем условие  $s \geq \max\{m+1, 2m-1\}$ .

Используя рассуждения, аналогичные, применявшимся при доказательстве теоремы 2, с. 209, нетрудно показать, что из (24), (25) и леммы 3 следует

**Теорема 4.** Пусть  $\hat{S}$  есть  $s \times s$ -точечная квадратура Гаусса,  $s \geq \max\{m+1, 2m-1\}$ , триангуляция  $\mathcal{T}_h$  является  $(m+1)$ -регулярной,

$$u \in W_r^{m+1}(\Omega), \quad a_{ij} \in W_\infty^m(\Omega), \quad i, j = 1, 2, \quad \tilde{f} \in W_r^m(\tilde{\Omega}), \quad rm > 2, \quad r \geq 2.$$

Тогда справедлива следующая оценка точности схемы (23):

$$\|\tilde{u} - u_h\|_{1,\Omega_h} \leq ch^m(\|u\|_{m+1,\Omega} + \sum_{i,j=1}^2 \|a_{ij}\|_{m,\infty,\Omega}\|u\|_{m+1,r,\Omega} + \|\tilde{f}\|_{m,r,\tilde{\Omega}}).$$

ЗАМЕЧАНИЕ 2. Интересно сравнить требования на точность квадратурной формулы  $\hat{S}$  в случае использования аффинно-эквивалентных и изопараметрических элементов. Если элементы треугольные, то как в том, так и в другом случае может использоваться одна и та же квадратура, точная на полиномах из  $P_{2m-2}$ . Если же элементы прямоугольные, то во втором случае квадратура должна быть точнее, когда  $m \geq 3$  (для  $s \times s$ -точечной квадратуры Гаусса  $s \geq m + 1$  и  $s \geq \max\{m + 1, 2m - 1\}$  соответственно). Причина такого различия между треугольными и прямоугольными изопараметрическими элементами кроется в различном определении их регулярности. Если предполагать выполненными условия сильной регулярности пространства прямоугольных изопараметрических элементов, то можно налагать такие же требования на точность квадратурных формул, как и для аффинно-эквивалентных элементов.

---

---

## Приложение. Алгоритмические аспекты метода конечных элементов

В предыдущих главах мы рассмотрели основные теоретические вопросы, касающиеся метода конечных элементов, а именно, вопросы, посвященные построению и исследованию точности дискретных аналогов краевых задач. Дискретная задача представляет собой не что иное, как систему алгебраических уравнений, линейных, если исходная задача была линейной. С практической точки зрения основным является:

- 1) формирование системы алгебраических уравнений, т. е. вычисление элементов матрицы и правой части системы;
- 2) решение полученной системы уравнений.

Сразу надо отметить, что эти два вопроса тесно взаимосвязаны, в особенности, если используются прямые методы решения системы линейных алгебраических уравнений. Дело в том, что возникающие в приложениях матрицы имеют очень высокую размерность (средней считается задача, включающая около десяти тысяч неизвестных) и нетривиальной является даже организация хранения таких матриц в памяти ЭВМ. Спасает положение то, что эти матрицы являются разреженными (подавляющее число их элементов — нули). Поэтому достаточно каким-либо образом хранить только индексы и значения ненулевых элементов. Ясно, что соответствующий способ хранения матриц должен позволять эффективно реализовать метод решения системы уравнений (например, метод Гаусса). Указанные вопросы (кодирование разреженных матриц, решение систем уравнений с разреженными матрицами) составляют предмет технологии разреженных матриц. Ему посвящены специальные монографии (см., напр. [4], [16]). Учитывая сказанное, основное внимание мы уделим вопросам формирования систем алгебраических уравнений. Для простоты изложения ограничимся рассмотрением лагранжевых конечных элементов класса  $C^0$  для двумерных эллиптических уравнений.

## § 1. Форма данных для представления триангуляции в ЭВМ

**1. Простейший случай. Линейные треугольные элементы.** Рассмотрим триангуляцию  $\mathcal{T}_h$ , изображенную на рис. 1. Предполагается, что с каждым элементом (треугольником) связан линейный конечный элемент. Какие данные необходимы для однозначного представления триангуляции в ЭВМ? Прежде всего, это два числа:  $N$  — общее число узлов сетки (вершин треугольников) и  $n_e$  — общее число конечных элементов.

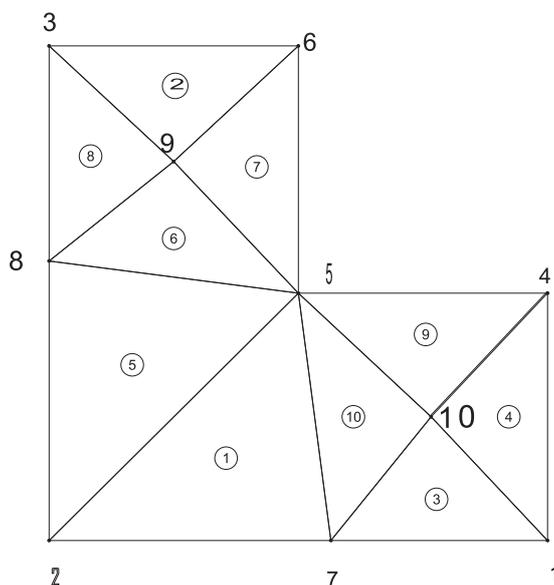


Рис. 1. Триангуляция  $\mathcal{T}_h$ :  $N = 10$ ,  $n_e = 10$ . Рядом с вершиной треугольника проставлен ее глобальный номер, номера треугольников указаны в кружочках.

Далее, очевидно, необходимо также указать

- координаты узлов;
- информацию о соединении узлов между собой.

Для этих целей достаточно ввести два массива: массив координат узлов  $x$ , размера  $2 \times N$  и так называемую матрицу связности элементов  $M$  размера  $3 \times n_e$ . Элемент  $x_{ij}$  матрицы  $x$  — это  $i$ -тая координата  $j$ -го узла сетки; элемент  $m_{ij}$  матрицы  $M$  — глобальный номер  $i$ -го узла (в локальной нумерации) на  $j$ -том конечном элементе. Напомним, что локальная нумерация узлов<sup>1)</sup> вводится парой чисел  $(j, i)$ , где  $j$  — номер элемента,  $i$  — номер узла на этом элементе. Таким

<sup>1)</sup>Точнее, узловых параметров. Для лагранжевых элементов между узлами и узловыми параметрами имеется взаимно однозначное соответствие.

образом, матрица  $M$  определяет соответствие между локальными и глобальными номерами узлов. Для рассматриваемой нами триангуляции матрица связности элементов имеет следующий вид:

$$M = \begin{pmatrix} 5 & 6 & 7 & 1 & 2 & 8 & 5 & 3 & 4 & 5 \\ 2 & 3 & 1 & 4 & 5 & 5 & 6 & 8 & 5 & 7 \\ 7 & 9 & 10 & 10 & 8 & 9 & 9 & 9 & 10 & 10 \end{pmatrix}.$$

Мы воспользовались следующим общепринятым способом локальной нумерации: узлы каждого элемента нумеруются против часовой стрелки.

**ЗАМЕЧАНИЕ 1.** На локальную нумерацию узлов на элементах не накладывается никаких ограничений, кроме того, что правило нумерации должно быть одинаковым для всех элементов (включая базисный элемент). Использованное нами правило "против часовой стрелки" удобно тем, что обеспечивает положительность якобиана преобразования произвольного элемента в базисный. Глобальная нумерация узлов определяет положение ненулевых элементов матрицы системы алгебраических уравнений, однозначно определяемой по матрице  $M$ .

Числа  $N$ ,  $n_e$  и матрицы  $x$  и  $M$  однозначно определяют триангуляцию. Используя их, легко

- выполнить чертеж триангуляции, подобный рис. 1, на экране графического дисплея;
- найти координаты вершин заданного элемента;
- вычислить площадь или периметр элемента и т. д.

#### Задачи.

1. Нарисовать триангуляцию области, соответствующую данным:

$$M = \begin{pmatrix} 2 & 2 & 6 & 3 & 4 & 4 & 1 & 5 \\ 7 & 5 & 5 & 6 & 6 & 9 & 7 & 7 \\ 5 & 8 & 8 & 8 & 5 & 5 & 9 & 9 \end{pmatrix},$$

$$x = \begin{pmatrix} 0.0 & 1.0 & 1.0 & 0.0 & 0.5 & 0.5 & 0.5 & 1.00.0 \\ 0.0 & 0.0 & 1.0 & 1.0 & 0.5 & 1.0 & 0.0 & 0.00.5 \end{pmatrix}.$$

2. Написать программу (на любом языке программирования) триангуляции прямоугольной области  $\Omega = (0, l_x) \times (0, l_y)$ , вычисляющую числа  $N$  и  $n_e$  и массивы  $x$ ,  $M$  по следующему правилу:  $\Omega$  разбивается на  $n_x \times n_y$  равных прямоугольников, которые в свою очередь разбиваются на два треугольника одной из диагоналей.

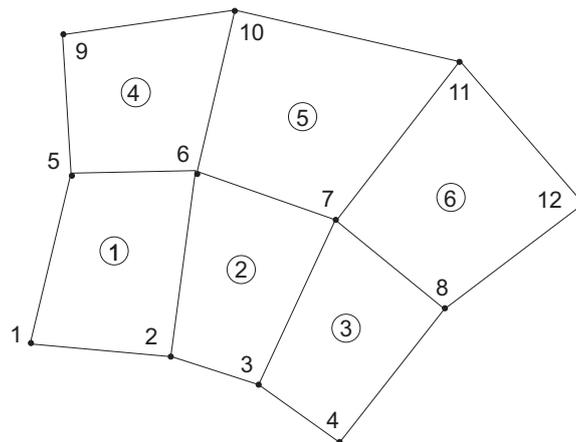


Рис. 2. Разбиение области на четырехугольные элементы:  $N = 12$ ,  $n_e = 6$ . Рядом с вершиной элемента проставлен его глобальный номер, номера элементов указаны в кружочках.

## 2. Другие типы триангуляций и конечных элементов.

Рассмотрим случай разбиения области на четырехугольные элементы первой степени ( $Q_1$ ), изображенный на рис. 2. В данном случае матрица связности элементов имеет вид:

$$M = \begin{pmatrix} 1 & 2 & 3 & 5 & 6 & 7 \\ 2 & 3 & 4 & 6 & 7 & 8 \\ 6 & 7 & 8 & 10 & 11 & 12 \\ 5 & 6 & 7 & 9 & 10 & 11 \end{pmatrix},$$

координаты узлов  $x$  определяются как и ранее. Для локальной нумерации узлов мы опять воспользовались правилом «против часовой стрелки» для всех элементов.

Для элементов более высокой степени в матрицу связности элементов удобно включать глобальные номера всех узловых параметров элемента, а не только узлов интерполяции. Например, для лагранжевых элементов второй степени (рис. 3) матрица  $M$  будет иметь размер  $6 \times n_e$ , поскольку на каждом элементе имеется шесть узловых параметров, соответствующих значению функции в точке. В данном конкретном случае матрица связности элементов имеет вид:

$$M = \begin{pmatrix} 1 & 2 & 3 & 1 & 2 & 3 \\ 2 & 3 & 4 & 6 & 7 & 8 \\ 6 & 7 & 8 & 5 & 6 & 7 \\ 9 & 10 & 11 & 13 & 15 & 17 \\ 14 & 16 & 18 & 19 & 20 & 21 \\ 13 & 15 & 17 & 12 & 14 & 16 \end{pmatrix}$$

(какое правило использовано при локальной нумерации узлов?).

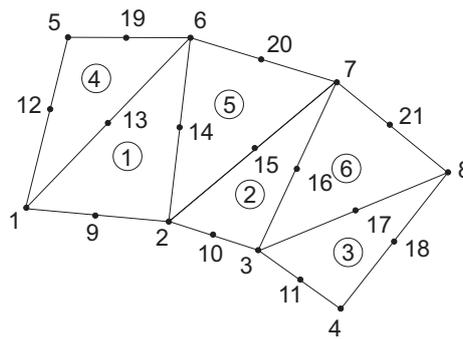


Рис. 3. Разбиение области на лагранжевые треугольные элементы второй степени:  $N = 21, n_e = 6$ .

## § 2. Алгоритм сборки системы уравнений

Отметим прежде всего, что алгоритм формирования системы алгебраических уравнений в методе конечных элементов является единым для всех типов конечных элементов и допускает простую и эффективную программную реализацию. Перейдем к его рассмотрению.

**1. Краевые условия 3-го рода.** Рассмотрим следующую задачу: найти функцию  $u(x) \in H^1(\Omega)$  такую, что

$$\int_{\Omega} (A \nabla u \cdot \nabla v + a uv) dx + \int_{\Gamma} \sigma uv dx = \int_{\Omega} f v dx + \int_{\Gamma} g v dx \quad \forall v \in H^1(\Omega), \quad (1)$$

где  $\Omega$  область в  $R^2$  с достаточно гладкой границей  $\Gamma$ . Исходные данные, а именно, функции  $A, a, f, \sigma, g$ , будем считать, для простоты изложения, определенными на всей плоскости.

Напомним, что задача (1) представляет собой обобщенную формулировку третьей краевой задачи:

$$\begin{aligned} -\operatorname{div}(A(x) \nabla u) + a(x)u &= f(x), & x \in \Omega, \\ A(x) \nabla u \cdot \nu(x) + \sigma(x)u &= g(x), & x \in \Gamma. \end{aligned}$$

Здесь  $\nu(x)$  — единичный вектор внешней нормали в точке  $x \in \Gamma$ .

Пусть  $\mathcal{T}_h$  — триангуляция области  $\Omega$  на конечные элементы  $e$  какого-либо типа (треугольные или четырехугольные),  $\Omega_h = \cup_e e$ ,  $\Gamma_h$  — граница области  $\Omega_h$ ,  $V_h$  — соответствующее пространство лагранжевых конечных элементов (аффинно-эквивалентных, изопараметрических или криволинейных элементов).

Построение приближенного решения задачи (1) по методу конечных элементов сводится к отысканию функции  $u_h \in V_h$  такой, что

$$\begin{aligned} \int_{\Omega_h} (A \nabla u_h \cdot \nabla v_h + a u_h v_h) dx + \int_{\Gamma_h} \sigma u_h v_h dx = \\ = \int_{\Omega_h} f v_h dx + \int_{\Gamma_h} g v_h dx \quad \forall v_h \in V_h. \end{aligned} \quad (2)$$

Пусть  $\varphi_i(x)$ ,  $i = 1, 2, \dots, N$  ( $N$  зависит от  $h$ ), есть базис Лагранжа в  $V_h$ ,

$$u_h(x) = \sum_{i=1}^N y_i \varphi_i(x).$$

Вектор  $y = \{y_1, \dots, y_N\}^T$ , вектор узловых параметров функции  $u_h$ , определяется как решение системы линейных алгебраических уравнений

$$Ay = F, \quad (3)$$

где «глобальная матрица жесткости»  $A = \{a_{ij}\}_{i,j=1}^N$  задается элементами

$$A_{ij} = \int_{\Omega_h} (A(x) \nabla \varphi_j \cdot \nabla \varphi_i + a(x) \varphi_j \varphi_i) dx + \int_{\Gamma_h} \sigma(x) \varphi_j \varphi_i dx, \quad (4)$$

а «глобальный вектор сил»  $F = \{F_i\}_{i=1}^N$  имеет компоненты

$$F_i = \int_{\Omega_h} f(x) \varphi_i dx + \int_{\Gamma_h} g(x) \varphi_i dx, \quad 1 \leq i \leq N. \quad (5)$$

Формулы (4), (5) дают некоторый способ вычисления элементов матрицы  $A$  и вектора  $F$ , но ими непосредственно не пользуются при практических вычислениях, поскольку существует более удобный и более экономичный метод, который мы и рассмотрим.

Из формулы (4) следует, что матрица  $A$  является суммой двух матриц:  $A = A_\Omega + A_\Gamma$ ,

$$A_\Omega = \left\{ \int_{\Omega_h} (A \nabla \varphi_j \cdot \nabla \varphi_i + a \varphi_j \varphi_i) dx \right\}_{i,j=1}^N, \quad A_\Gamma = \left\{ \int_{\Gamma_h} \sigma \varphi_j \varphi_i dx \right\}_{i,j=1}^N.$$

Аналогично,

$$F = F_\Omega + F_\Gamma, \quad F_\Omega = \left\{ \int_{\Omega_h} f \varphi_i dx \right\}_{i=1}^N, \quad F_\Gamma = \left\{ \int_{\Gamma_h} g \varphi_i dx \right\}_{i=1}^N.$$

**Вычисление  $A_\Omega$  и  $F_\Omega$ .** Согласно определению

$$A_\Omega y \cdot v = \int_{\Omega_h} (A \nabla u_h \cdot \nabla v_h + a u_h v_h) dx,$$

где  $v$  — вектор узловых параметров функции  $v_h$ , или

$$A_\Omega y \cdot v = \sum_e \int_e (A \nabla u_h \cdot \nabla v_h + a u_h v_h) dx. \quad (6)$$

Пусть  $m_e$  — количество узловых параметров функции из  $V_h$  на элементе  $e$ . Воспользуемся представлением функций из  $V_h$  на элементе  $e$ :

$$u_h(x) = \sum_{\beta=1}^{m_e} y_\beta^e \varphi_\beta^e(x), \quad v_h(x) = \sum_{\alpha=1}^{m_e} v_\alpha^e \varphi_\alpha^e(x).$$

Подставляя эти разложения в (6), получим

$$A_\Omega y \cdot v = \sum_e \sum_{\alpha, \beta=1}^{m_e} a_{\alpha\beta}^e y_\beta^e v_\alpha^e.$$

Образовавшаяся здесь матрица  $A^e = \{a_{\alpha\beta}^e\}_{\alpha, \beta=1}^{m_e}$ , где

$$a_{\alpha\beta}^e = \int_e (A(x) \nabla \varphi_\beta^e \cdot \nabla \varphi_\alpha^e + a(x) \varphi_\beta^e \varphi_\alpha^e) dx$$

называется матрицей жесткости элемента  $e$  (или просто локальной матрицей жесткости).

Используя связь между локальной и глобальной нумерациями (задаваемую матрицей связности элементов), можем написать  $y_\beta^e = y_{m_{\beta e}}$ ,  $v_\alpha^e = v_{m_{\alpha e}}$ , и, следовательно,

$$A_\Omega y \cdot v = \sum_e \sum_{\alpha, \beta=1}^{m_e} a_{\alpha\beta}^e y_{m_{\beta e}} v_{m_{\alpha e}}. \quad (7)$$

Преобразуем это равенство. Введем в рассмотрение матрицу  $\tilde{A}^e$  размера  $N \times N$ , состоящую из нулей, за исключением  $m_e^2$  элементов, которые определим соотношением

$$\tilde{A}_{m_{\alpha e}, m_{\beta e}}^e = a_{\alpha\beta}^e, \quad \alpha, \beta = 1, 2, \dots, m_e.$$

Тогда формула (7) примет вид

$$\begin{aligned} A_{\Omega} y \cdot v &= \sum_e \sum_{\alpha, \beta=1}^{m_e} \tilde{A}_{m_{\alpha e}, m_{\beta e}}^e y_{m_{\beta e}} v_{m_{\alpha e}} = \sum_e \sum_{i, j=1}^N \tilde{A}_{ij}^e y_j v_i = \\ &= \sum_e \tilde{A}^e y \cdot v = \left( \sum_e \tilde{A}^e \right) y \cdot v, \end{aligned}$$

откуда в силу произвольности векторов  $y$  и  $v$  вытекает, что

$$A_{\Omega} = \sum_e \tilde{A}^e.$$

Последнее равенство показывает, что глобальную матрицу жесткости можно получить суммированием по всем конечным элементам локальных матриц жесткости. Поскольку нули не имеет смысла суммировать, приходим к следующему алгоритму, известному как

#### Алгоритм сборки матрицы жесткости.

- Положить  $A = 0$  ( $A$  — матрица размера  $N \times N$ ).
- Для каждого  $e \in \mathcal{T}_h$ :
  - вычислить  $A^e$  (размера  $m_e \times m_e$ ).
  - Для  $\alpha, \beta = 1, 2, \dots, m_e$  суммировать:

$$A_{m_{\alpha e}, m_{\beta e}} = A_{m_{\alpha e}, m_{\beta e}} + a_{\alpha \beta}^e.$$

Аналогичные соображения приводят к алгоритму вычисления вектора  $F_{\Omega}$ .

#### Алгоритм сборки вектора сил.

- Положить  $F = 0$  ( $F$  — вектор длины  $N$ ).
- Для каждого  $e \in \mathcal{T}_h$ :
  - вычислить  $F^e$  (вектор длины  $m_e$ ).
  - Для  $\alpha = 1, 2, \dots, m_e$  суммировать:

$$F_{m_{\alpha e}} = F_{m_{\alpha e}} + F_{\alpha}^e.$$

Здесь  $F^e$  — вектор сил элемента  $e$ , имеет компоненты

$$F_\alpha^e = \int_e f(x) \varphi_\alpha^e dx.$$

**Вычисление  $A_\Gamma$  и  $F_\Gamma$ .** Граница  $\Gamma_h$  области  $\Omega_h$  является объединением граней элементов из  $\mathcal{T}_h$ . Пусть  $\Gamma_h = \cup \tau$ ,  $\tau \subset \Gamma_h$  — грань некоторого конечного элемента,  $m_\tau$  — количество узловых параметров, соответствующих грани  $\tau$ . В рассматриваемом нами случае — это число узлов интерполяции на  $\tau$ . По определению функции из  $V_h$ , сужение ее на грань  $\tau$  определяются узловыми параметрами только этой грани. Поэтому

$$u_h(x)|_\tau = \sum_{\alpha=1}^{m_\tau} y_\alpha^\tau \varphi_\alpha^\tau(x),$$

где  $(\alpha, \tau)$  — локальный номер узлового параметра на грани  $\tau$ . Точно так же, как при вычислении матрицы  $A_\Omega$ , можем написать, что

$$A_\Gamma y \cdot v = \sum_{\tau \subset \Gamma_h} \int_\tau \sigma(x) u_h v_h dx = \sum_{\tau \subset \Gamma_h} \sum_{\alpha, \beta=1}^{m_\tau} a_{\alpha\beta}^\tau y_\beta^\tau v_\alpha^\tau,$$

где  $A^\tau = \{a_{\alpha\beta}^\tau\}_{\alpha, \beta=1}^{m_\tau}$  — «матрица жесткости» элемента  $\tau$ :

$$a_{\alpha\beta}^\tau = \int_\tau \sigma(x) \varphi_\beta^\tau \varphi_\alpha^\tau dx.$$

Вводя расширенную матрицу  $\tilde{A}^\tau$ , как и ранее, получим:

$$A_\Gamma = \sum_{\tau \subset \Gamma_h} \tilde{A}^\tau.$$

Нет необходимости хранить матрицу  $A_\Gamma$  отдельно. Поэтому ее элементы по мере их вычисления сразу прибавляются к соответствующим элементам матрицы  $A_\Omega$ .

Аналогично строится и вектор  $F_\Gamma$ , при этом вводится «вектор сил»  $F^\tau$  грани  $\tau$ , определяемый компонентами

$$F_\alpha^\tau = \int_\tau g(x) \varphi_\alpha^\tau dx.$$

Итак, если матрица  $A = A_\Omega$  и вектор  $F = F_\Omega$  уже вычислены при помощи описанного выше алгоритма сборки, то затем применяется следующий

### Алгоритм учета краевых условий 3-го рода.

- Для каждого  $\tau \in \Gamma_h$ :
  - вычислить  $A^\tau$  и  $F^\tau$  (размера  $m_\tau \times m_\tau$  и  $m_\tau$  соответственно).
  - Для  $\alpha, \beta = 1, \dots, m_\tau$  суммировать:

$$A_{m_{\alpha\tau}, m_{\beta\tau}} = A_{m_{\alpha\tau}, m_{\beta\tau}} + a_{\alpha\beta}^\tau.$$

- Для  $\alpha = 1, \dots, m_\tau$  суммировать:

$$F_{m_{\alpha\tau}} = F_{m_{\alpha\tau}} + F_\alpha^\tau.$$

Здесь  $m_{\alpha\tau}$  — глобальный номер узла  $\alpha$  на грани  $\tau$ .

#### Задачи.

1. Получить из тождества (2) систему (3).
2. Выполнить подробно построения, приводящие к алгоритму сборки вектора сил.

**2. Главное краевое условие.** Предположим, что на части  $\Gamma^0$  границы  $\partial\Omega$  задано краевое условие Дирихле (главное краевое условие)

$$u(x) = g(x), \quad x \in \Gamma^0.$$

Введем обозначения:

$$V^g = \{v \in H^1(\Omega) : v(x) = g(x), \quad x \in \Gamma^0\},$$

$$V^0 = \{v \in H^1(\Omega) : v(x) = 0, \quad x \in \Gamma^0\}.$$

Рассмотрим задачу: найти функцию  $u(x) \in V^g$  такую, что<sup>1)</sup>

$$\int_{\Omega} (A(x) \nabla u \cdot \nabla v + a(x) uv) dx = \int_{\Omega} f v dx \quad \forall v \in V^0. \quad (8)$$

УПРАЖНЕНИЕ 3. Дать классическую формулировку задачи (8).

Пусть  $\mathcal{T}_h$  — триангуляция области  $\Omega$  на конечные элементы  $e$  какого-либо типа (треугольные или четырехугольные), согласованная с границей  $\Gamma^0$  (граничные точки  $\Gamma^0$  являются вершинами каких-либо конечных элементов). Пусть  $\Omega_h = \bigcup_e e$ ,  $\Gamma_h^0$  — соответствующая  $\Gamma^0$  часть границы области  $\Omega_h$ . Пусть, далее,  $V_h$  — соответствующее  $\mathcal{T}_h$

<sup>1)</sup>Для простоты изложения мы ограничились рассмотрением однородного граничного условия на  $\Gamma^1 = \Gamma \setminus \Gamma^0$  при  $\sigma \equiv 0$ .

пространство конечных элементов (аффинно-эквивалентных, изопараметрических или криволинейных), аппроксимирующее  $H^1(\Omega)$ . Обозначим через  $\gamma_0$  множество узлов интерполяции, принадлежащих  $\Gamma_h^0$ . Аппроксимации аффинного множества функций  $V^g$  и пространства  $V^0$  определяются следующим образом:

$$\begin{aligned} V_h^g &= \{v \in V_h : v(x) = g(x), \quad x \in \gamma_0\}, \\ V_h^0 &= \{v \in V_h : v(x) = 0, \quad x \in \gamma_0\}. \end{aligned}$$

Построение приближенного решения задачи (8) по методу конечных элементов сводится к отысканию функции  $u_h \in V_h^g$  такой, что

$$\int_{\Omega_h} (A(x)\nabla u_h \cdot \nabla v_h + a(x)u_h v_h) dx = \int_{\Omega_h} f v_h dx \quad \forall v_h \in V_h^0. \quad (9)$$

Пусть  $\varphi_i(x)$ ,  $i = 1, 2, \dots, N$ , есть базис Лагранжа в  $V_h$ , причем  $\varphi_i(x)$  с номерами  $i = 1, 2, \dots, N_0$  соответствуют узлам, не лежащим на  $\Gamma_h^0$ . Тогда

$$u_h(x) = \sum_{j=1}^{N_0} y_j \varphi_j(x) + \sum_{j=N_0+1}^N g(x_j) \varphi_j(x), \quad (10)$$

а произвольная функция  $v_h \in V_h^0$  имеет представление

$$v_h(x) = \sum_{i=1}^{N_0} v_i \varphi_i(x). \quad (11)$$

Подставляя разложения (10) и (11) в тождество (9), приходим к системе линейных алгебраических уравнений:

$$A^0 y = F^0. \quad (12)$$

Определим, как и ранее, глобальную матрицу жесткости  $A = (A_{ij})_{i,j=1}^N$  и вектор сил  $F = (F_i)_{i=1}^N$  (без учета краевого условия Дирихле),

$$A_{ij} = \int_{\Omega_h} (A(x)\nabla \varphi_j \cdot \nabla \varphi_i + a(x)\varphi_j \varphi_i) dx, \quad F_i = \int_{\Omega_h} f \varphi_i dx.$$

Тогда матрица  $A^0$  размера  $N_0 \times N_0$  и вектор  $F^0$  будут иметь компоненты

$$A_{ij}^0 = A_{ij}, \quad 1 \leq i, j \leq N_0,$$

$$F_i^0 = F_i - \sum_{j=N_0+1}^N A_{ij}g(x_j), \quad 1 \leq i \leq N_0.$$

Видно, что в вычислениях участвуют, вообще говоря, все столбцы матрицы  $A$  с номерами, большими  $N_0$ .

Укажем простой способ получения системы алгебраических уравнений, эквивалентный системе (12), но не требующий специальной нумерации неизвестных. При этом, однако, нужно предварительно построить вектор  $d$  длины  $N$  такой, что  $d_i = 0$ , если узел с глобальным номером  $i$  не принадлежит границе;  $d_i = 1$ , если в  $i$ -том узле задано условие Дирихле, и  $d_i = 2$  в противном случае. Далее, через  $a_{*j}$  будем обозначать  $j$ -тый столбец матрицы  $A$ , через  $g_i$  — значение функции  $g$  в  $i$ -той граничной точке.

### Алгоритм учета главных краевых условий.

- Вычислить матрицу  $A$  и вектор  $F$ , используя алгоритмы сборки матрицы жесткости и вектора сил.
- Для каждого  $i$  такого, что  $d_i = 1$ :
  - суммировать:  $F = F - a_{*i}g_i$  ( $a_{*i}$  —  $i$ -й столбец матрицы  $A$ ).
  - все элементы  $i$ -той строки и  $i$ -того столбца положить равными нулю.
  - положить:  $A_{ii} = 1, F_i = g_i$ .

ЗАМЕЧАНИЕ 1. Если все же необходимо, как, например, при решении задачи на собственные значения, получить непосредственно матрицу  $A^0$ , то можно поступить следующим образом.

(1) Определим число  $N_0$  как число не равных единице элементов вектора  $d$ .

(2) Перенумеруем узлы. Для этого определим вектор  $p$  длины  $N$  такой, что  $p_i = 0$ , если в  $i$ -том узле задано условие Дирихле, иначе  $p_i$  определим как новый номер  $i$ -го узла ( $p_i \in [1, N_0]$ ). Для этого используем алгоритм:

- $j=0$ ;
- Для каждого  $i = 1, \dots, N$ :
  - если  $d_i \neq 1$  то  $\{j = j + 1; p_i = j\}$  иначе  $\{p_i = 0\}$ .
- (3) Вычислим матрицу  $A^0$ , используя

**Модифицированный алгоритм сборки матрицы жесткости.**

- Положить  $A^0 = 0$  ( $A^0$  — матрица размера  $N_0 \times N_0$ ).
- Для каждого  $e \in \mathcal{T}_h$ :
  - вычислить  $A^e$  (размера  $m_e \times m_e$ ).
  - Для  $\alpha, \beta = 1, \dots, m_e$ :
    - $m_\alpha = p(m_{\alpha e}), \quad m_\beta = p(m_{\beta e});$
    - $A_{m_\alpha, m_\beta}^0 = A_{m_\alpha, m_\beta}^0 + a_{\alpha\beta}^e.$

**3. Вычисление матрицы жесткости элемента.** Рассмотрим, наконец, задачу вычисления матрицы жесткости элемента  $e$ :

$$a_{\alpha\beta}^e = \int_e (A(x) \nabla \varphi_\beta^e \cdot \nabla \varphi_\alpha^e + a(x) \varphi_\beta^e \varphi_\alpha^e) dx. \quad (13)$$

Будем предполагать, что элемент  $(e, P_e, \Sigma_e)$  является ассоциированным с базисным элементом  $(\hat{e}, \hat{P}, \hat{\Sigma})$  посредством преобразования:

$$x = x_e(\hat{x}) = (x_1^e(\hat{x}), x_2^e(\hat{x}))^T.$$

Определитель матрицы

$$J_e(\hat{x}) \equiv (Dx_e(\hat{x}))^T = \begin{pmatrix} \frac{\partial x_1^e}{\partial \hat{x}_1} & \frac{\partial x_2^e}{\partial \hat{x}_1} \\ \frac{\partial x_1^e}{\partial \hat{x}_2} & \frac{\partial x_2^e}{\partial \hat{x}_2} \end{pmatrix} (\hat{x})$$

этого преобразования будем считать положительным. Перейдем в интеграле (13) от элемента  $e$  к элементу  $\hat{e}$ , учитывая, что

$$A(x) = \hat{A}(\hat{x}) = A(x_e(\hat{x})), \quad \varphi_\alpha^e(x) = \hat{\varphi}_\alpha(\hat{x}),$$

$$\nabla = J_e^{-1}(\hat{x}) \hat{\nabla}, \quad \hat{\nabla} = (\partial/\partial \hat{x}_1, \partial/\partial \hat{x}_2)^T, \quad dx = |J_e(\hat{x})| d\hat{x},$$

где  $\hat{\varphi}_\alpha(\hat{x})$  — базисные функции Лагранжа в пространстве  $\hat{P}$ . Будем иметь

$$a_{\alpha\beta}^e = \int_{\hat{e}} (\hat{A}(\hat{x}) J_e^{-1}(\hat{x}) \hat{\nabla} \hat{\varphi}_\beta \cdot J_e^{-1}(\hat{x}) \hat{\nabla} \hat{\varphi}_\alpha + \hat{a}(\hat{x}) \hat{\varphi}_\beta \hat{\varphi}_\alpha) |J_e(\hat{x})| d\hat{x}. \quad (14)$$

Подынтегральная функция в (14) вычисляется по явным формулам в любой точке  $\hat{x}$ , если известны явные формулы для преобразования

координат и базисных функций  $\hat{\varphi}_\alpha$ . Конечно, только в частных случаях этот интеграл может быть вычислен точно (например, в случае аффинно-эквивалентных элементов и постоянных  $A(x)$  и  $a(x)$ ).

В общем случае необходимо использовать квадратурные формулы для приближенного вычисления<sup>1)</sup> интеграла (14). Если

$$\int_{\hat{e}} \hat{f}(\hat{x}) d\hat{x} \approx \sum_{i=1}^s c_i \hat{f}(b_i),$$

то окончательно получим допускающую простую программную реализацию формулу для элементов матрицы жесткости:

$$a_{\alpha\beta}^e = \sum_{i=1}^s c_i \left( (\hat{A} J_e^{-1} \hat{\nabla} \hat{\varphi}_\beta \cdot J_e^{-1} \hat{\nabla} \hat{\varphi}_\alpha + \hat{a} \hat{\varphi}_\beta \hat{\varphi}_\alpha) |J_e| \right) (b_i). \quad (15)$$

#### Задачи.

1. Указать способ вычисления компонент вектора сил элемента

$$F_\alpha^e = \int_e f(x) \varphi_\alpha^e dx.$$

2. Конкретизировать формулу (15) для линейных треугольных конечных элементов ( $\hat{P} = P_1$ ). Какую квадратурную формулу следует использовать в этом случае?

3. Решить предыдущую задачу в случае, когда  $\hat{P} = Q_1$  (изопараметрические четырехугольные элементы).

---

<sup>1)</sup>Напомним, что в этом случае мы приходим к схеме МКЭ с численным интегрированием

---

---

## Литература

1. **Бабенко К.И.** Основы численного анализа. — М.: Наука, 1986.
2. **Гильбарг Д., Трудингер Н.** Эллиптические дифференциальные уравнения с частными производными второго порядка. — М.: Наука, 1989.
3. **Деклу Ж.** Метод конечных элементов. — М.: Мир, 1976.
4. **Джордж А., Лю Дж.** Численное решение больших разреженных систем уравнений. — М.: Мир, 1984.
5. **Зенкевич О.К.** Метод конечных элементов в технике. — М.: Мир, 1989.
6. **Корнеев В.Г.** Схемы метода конечных элементов высоких порядков точности. — Л.: ЛГУ, 1977.
7. **Ладыженская О.А.** Краевые задачи математической физики. — М.: Наука, 1973.
8. **Марчук Г.И.** Методы вычислительной математики. — М.: Наука, 1980.
9. **Марчук Г.И., Агошков В.И.** Введение в проекционно-сеточные методы. — М.: Наука, 1981.
10. **Митчел Э., Уэйт Р.** Метод конечных элементов для уравнений с частными производными. — М.: Мир, 1981.
11. **Михлин С.Г.** Вариационные методы в математической физике. — М.: Гостехиздат, 1957.
12. **Михлин С.Г.** Численная реализация вариационных методов. — М.: Наука, 1966.
13. **Обен Ж.-П.** Приближенное решение эллиптических краевых задач. — М.: Мир, 1988.
14. **Оганесян Л.А., Руховец Л.А.** Вариационно-разностные методы решения эллиптических уравнений. — Ереван: Изд-во АН АрмССР, 1979.

15. **Павлова М.Ф., Тимербаев М.Р.** Пространства Соболева. — Казань, 2002.
16. **Писсанецки С.** Технология разреженных матриц. — М.: Мир, 1988.
17. **Самарский А.А., Николаев Е.С.** Методы решения сеточных уравнений. — М.: Наука, 1978.
18. **Соболев С.Л.** Некоторые применения функционального анализа в математической физике. — М.: Наука, 1988.
19. **Стейн И.** Сингулярные интегралы и дифференциальные свойства функций. — М.: Мир, 1973
20. **Стренг Г., Фикс Дж.** Теория метода конечных элементов. — М.: Мир, 1977.
21. **Сьярле Ф.** Метод конечных элементов для эллиптических задач. — М.: Мир, 1980.
22. **Сьярле Ф.** Математическая теория упругости. — М.: Мир, 1992.
23. **Темам Р.** Уравнения Навье — Стокса. Теория и численный анализ. — М.: Мир, 1981.
24. **Трибель Х.** Теория интерполяции, функциональные пространства, дифференциальные операторы. — М.: Мир, 1980.
25. **Шайдуров В.В.** Многосеточные методы конечных элементов. — М.: Наука, 1989.

---

---

## Основные обозначения

### Общие обозначения

$\delta_{ij}$  — символ Кронекера.

$R$  — множество всех вещественных чисел.

$\text{card } A$  — число элементов конечного множества  $A$ .

$\dim B$  — размерность конечномерного линейного пространства  $B$ .

$R^n$  —  $n$ -мерное евклидово пространство.

$A^T$  — матрица, полученная транспонированием матрицы  $A$ .

$x = (x_1, \dots, x_n)^T$ ,  $x \in R$  — элементы пространства  $R^n$ .

$x \cdot y = x^T y = x_1 y_1 + \dots + x_n y_n$  — скалярное произведение в  $R^n$ .

$|x| = (x \cdot x)^{1/2}$  — длина вектора  $x \in R^n$ .

$[R^n]^k = R^n \times R^n \times \dots \times R^n$  — прямое произведение  $k$  экземпляров евклидова пространства.

$A(\xi^1, \dots, \xi^k)$ ,  $\xi^i \in R^n$ ,  $i = 1, \dots, k$  — полилинейная форма (полилинейное отображение  $[R^n]^k$  в  $R$ ).

$A\xi^k = A(\xi, \dots, \xi)$  (для симметричной полилинейной формы).

$\Omega$  — ограниченная область евклидова пространства  $R^n$ .

$\Gamma$  или  $\partial\Omega$  — граница области  $\Omega$ .

$\nu$  — единичный вектор внешней нормали к  $\Gamma$ .

$\text{mes}(\Omega)$  или  $|\Omega|$  — мера области  $\Omega$ .

$\text{int } B$  — внутренность множества  $B \subset R^n$ .

### Производные и дифференциальные операторы

$\frac{\partial^k u}{\partial x_{i_1} \partial x_{i_2} \dots \partial x_{i_k}}$ ,  $i_1, i_2, \dots, i_k \in \{1, 2, \dots, n\}$ ,  $k \geq 1$  — частная производная порядка  $k$  функции  $u = u(x_1, \dots, x_n)$ .

$\alpha = (\alpha_1, \alpha_2, \dots, \alpha_n)$  — мультииндекс (вектор с целочисленными неотрицательными компонентами).

$|\alpha| = \alpha_1 + \alpha_2 + \dots + \alpha_n$  — длина мультииндекса  $\alpha$ .

$D^\alpha u = \frac{\partial^{|\alpha|} u}{\partial x_1^{\alpha_1} \partial x_2^{\alpha_2} \dots \partial x_n^{\alpha_n}}$  — производная порядка  $|\alpha|$  (классическая или обобщенная).

$D^k u(x)(h^1, h^2, \dots, h^k) = \sum_{i_1, i_2, \dots, i_k=1}^n \frac{\partial^k u(x)}{\partial x_{i_1} \partial x_{i_2} \dots \partial x_{i_k}} h_{i_1}^1 h_{i_2}^2 \dots h_{i_k}^k$  — симметричная полилинейная форма, порожденная частными производными порядка  $k$  функции  $u$ .

$D^k u(x)h^k = D^k u(x)(h, h, \dots, h)$  — дифференциал порядка  $k$  функции  $u$  в точке  $x$ .

$\nabla u(x) = \left( \frac{\partial u(x)}{\partial x_1}, \frac{\partial u(x)}{\partial x_2}, \dots, \frac{\partial u(x)}{\partial x_n} \right)^T$  — градиент функции  $u$  в точке  $x$ .

$Du(x)h \equiv D^1 u(x)h = (\nabla u(x))^T h \quad \forall h \in R^n$ .

$u^+(x) = (u(x) + |u(x)|)/2$  — положительная часть функции  $u$ .

$u^-(x) = (u(x) - |u(x)|)/2$  — отрицательная часть функции  $u$ .

$\Delta u$  — оператор Лапласа.

$\text{supp } u = \{x \in R^n : u(x) \neq 0\}$  — носитель функции  $u$ .

### Линейные нормированные пространства

$H$  — вещественное пространство Гильберта.

$(\cdot, \cdot)$  — скалярное произведение в гильбертовом пространстве.

$\|\cdot\| = (\cdot, \cdot)^{1/2}$  — норма в гильбертовом пространстве.

$f(\cdot): H \rightarrow R$  — непрерывная линейная форма (линейный ограниченный функционал).

$a(\cdot, \cdot): H \times H \rightarrow R$  — непрерывная билинейная форма.

$(u, v)_a = a(u, v)$  (для симметричной положительно определенной формы  $a$ ) — энергетическое скалярное произведение.

$H_n$  — замкнутое подпространство гильбертова пространства  $H$  (обычно конечномерное).

$L_p(\Omega)$ ,  $p \geq 1$ , — пространство Лебега функций, измеримых на  $\Omega$  и суммируемых со степенью  $p$ .

$\|u\|_{p,\Omega} = \left( \int_{\Omega} |u|^p dx \right)^{1/p}$  — норма на пространстве  $L_p(\Omega)$ .

$\text{ess sup}_{x \in \Omega} |u(x)|$  — существенная верхняя грань функции  $u$  на  $\Omega$ .

$L_{\infty}(\Omega)$  — линейное пространство функций с нормой  $\|u\|_{\infty,\Omega} = \text{ess sup}_{x \in \Omega} |u(x)|$ .

$C(\bar{\Omega})$  — линейное пространство непрерывных на  $\bar{\Omega}$  вещественных функций.

$\|u\|_{C(\bar{\Omega})} = \max_{x \in \bar{\Omega}} |u(x)|$  — норма на пространстве  $C(\bar{\Omega})$ .

$C^k(\Omega)$  — линейное пространство  $k$  раз дифференцируемых на области  $\Omega$  функций.

$C^{k,\lambda}(\Omega)$  — линейное пространство  $k$  раз дифференцируемых на области  $\Omega$  функций, все  $k$ -е производные которых удовлетворяют условию Липшица с показателем  $\lambda \in (0, 1]$ .

$C_0^{\infty}(\Omega)$  — множество бесконечно дифференцируемых финитных на  $\Omega$  функций.

$C^{\infty}(\bar{\Omega})$  — линейное пространство, получаемое сужением на  $\bar{\Omega} = \Omega \cup \Gamma$  множества бесконечно дифференцируемых на  $R^n$  функций.

$W_p^s(\Omega)$  — пространство Соболева.

$\|u\|_{s,p,\Omega} = \left( \int_{\Omega} \sum_{|\alpha| \leq s} |D^{\alpha} u|^p dx \right)^{1/p}$  — норма на пространстве Соболева при

$1 \leq p < \infty$ .

$\|u\|_{s,\infty,\Omega} = \max_{|\alpha| \leq s} \text{ess sup}_{x \in \Omega} |D^{\alpha} u(x)|$  — норма на пространстве Соболева при  $p = \infty$ .

$|u|_{s,p,\Omega} = \left( \int_{\Omega} \sum_{|\alpha|=s} |D^{\alpha} u|^p dx \right)^{1/p}$ ,  $[u]_{s,p,\Omega} = \left( \int_{\Omega} \sum_{i=1}^n \left| \frac{\partial^s u}{\partial x_i^s} \right|^p \right)^{1/p}$  — полунормы на про-

странстве Соболева при  $1 \leq p < \infty$ .

$|u|_{s,\infty,\Omega} = \max_{|\alpha|=s} \text{ess sup}_{x \in \Omega} |D^{\alpha} u(x)|$ ,  $[u]_{s,\infty,\Omega} = \max_{i=1,\dots,n} \text{ess sup}_{x \in \Omega} \left| \frac{\partial^s u}{\partial x_i^s}(x) \right|$  — полунормы на пространстве Соболева при  $p = \infty$ .

$(u, v)_{s,\Omega} = \int_{\Omega} \sum_{|\alpha| \leq s} D^{\alpha} u D^{\alpha} v dx$  — скалярное произведение на пространстве  $W_2^s(\Omega)$ .

$\|\cdot\|_{s,\Omega}$  — норма на пространстве  $W_2^s(\Omega)$ .

$H^s(\Omega) = W_2^s(\Omega)$ .

## Конечные элементы

$\mathcal{T}_h = \{e\}$  — триангуляция области  $\Omega$ .

$$\Omega_h = \bigcup_{e \in \mathcal{T}_h} e.$$

$(e, P_e, \Sigma_e)$  или  $e$  — конечный элемент.

$\rho_e$  — максимальный радиус шара, содержащегося в  $e$ .

$h_e$  — диаметр  $e$ .

$\phi_i^e$  — степени свободы конечного элемента.

$\Sigma_e$  — множество всех степеней свободы конечного элемента.

$\varphi_i^e(x)$  — базисные функции Лагранжа конечного элемента.

$\omega^e$  — множество узлов интерполяции конечного элемента

$(\hat{e}, P_{\hat{e}}, \Sigma_{\hat{e}}) = (\hat{e}, \hat{P}, \hat{\Sigma})$  или  $\hat{e}$  — базисный конечный элемент.

$\hat{\rho}$  — максимальный радиус шара, содержащегося в  $\hat{e}$ .

$\hat{h}$  — диаметр  $\hat{e}$ .

$\hat{\phi}_i$  — степени свободы базисного конечного элемента.

$\hat{\Sigma}$  — множество всех степеней свободы базисного конечного элемента.

$\hat{\varphi}_i(x)$  — базисные функции Лагранжа базисного конечного элемента.

$\hat{\omega}$  — множество узлов интерполяции базисного конечного элемента

$x = x_e(\hat{x})$  — отображение  $\hat{e} \rightarrow e \in \mathcal{T}_h$ .

$J_e = Dx_e(\hat{x})$  — матрица Якоби отображения  $x_e$ .

$\det(Dx_e)$  — якобиан отображения  $x_e$ .

$\mathcal{E}_h = \{(e, P_e, \Sigma_e), e \in \mathcal{T}_h\}$  — семейство конечных элементов на области  $\Omega$ .

$x^\alpha = x_1^{\alpha_1} \cdots x_n^{\alpha_n}$  — моном степени  $|\alpha| = \alpha_1 + \cdots + \alpha_n$ .

$P_m = P_m(x) = \left\{ p : p = \sum_{|\alpha| \leq m} c_\alpha x^\alpha \right\}$  — пространство полиномов степени не выше  $m$  по совокупности переменных  $x_1, \dots, x_n$ .

$Q_m = Q_m(x) = \left\{ p : p = \sum_{\alpha_1, \dots, \alpha_n=0}^m c_\alpha x^\alpha \right\}$  — пространство полиномов степени не выше  $m$  по каждой переменной.

$u_I$  — интерполянт функции  $u$ .

$\pi$  — оператор интерполирования.

$\hat{u}_I$  — интерполянт функции  $\hat{u}$  на базисном элементе.

$\hat{\pi}$  — оператор интерполирования на базисном элементе.

$V_h = \{v_h \in C^{k-1}(\Omega_h) : v_h|_e \in P_e \quad \forall e \in \mathcal{T}_h\}$ ,  $k = 1, 2$ , — пространство конечных элементов.

## Квадратурные формулы

$\hat{b}_l$  — узлы квадратурной формулы на базисном элементе  $\hat{e}$ .

$\hat{c}_l$  — коэффициенты квадратурной формулы на базисном элементе  $\hat{e}$ .

$\hat{S}(\hat{\phi}) = \sum_{l=1}^L \hat{c}_l \hat{\phi}(\hat{b}_l)$  — квадратурная формула на базисном элементе.

$b_l^e$  — узлы квадратурной формулы на элементе  $e$ .

$c_l^e$  — коэффициенты квадратурной формулы на элементе  $e$ .

$S_e(\phi) = \sum_{l=1}^L c_l^e \phi(b_l^e)$  — квадратурная формула на элементе  $e$ .

$S_{\Omega}(\phi) = \sum_e S_e(\phi)$  — квадратурная формула на области  $\Omega$ .

$a_h(u, v) = S_{\Omega} \left( \sum_{i,j=1}^2 a_{ij} \frac{\partial u}{\partial x_j} \frac{\partial v}{\partial x_i} \right)$  — билинейная форма, порожденная квадратурной формулой.

$f_h(v) = S_{\Omega}(fv)$  — линейная форма, порожденная квадратурной формулой.

$\hat{E}(\hat{\phi}) = \int_{\hat{e}} \hat{\phi}(\hat{x}) d\hat{x} - \hat{S}(\hat{\phi})$  — погрешность квадратурной формулы на базисном элементе.

$E_e(\phi) = \int_e \phi(x) dx - S_e(\phi)$  — погрешность квадратурной формулы на элементе.

$E_{\Omega}(\phi) = \int_{\Omega} \phi(x) dx - S_{\Omega}(\phi)$  — погрешность квадратурной формулы на области.

---

---

## Предметный указатель

- Алгоритм  
— сборки  
— вектора сил, 222  
— матрицы жесткости, 222  
— матрицы жесткости (модифицированный), 227  
— учета главных краевых условий, 226  
— учета краевых условий 3-города, 224
- Базис Лагранжа, 81
- Вектор  
— сил  
— глобальный, 220  
— узловых параметров, 220
- Вектора  
— длина, 28
- Граничные условия  
— Дирихле, 68  
— второго рода, 68  
— главные, 66  
— естественные, 66
- Гаука закон, 68
- Дискретные аналоги норм в  $V_h$ , 182
- Диффеоморфизм  
—  $C^l$ , 53
- Задача  
— Дирихле, 14, 59, 179  
— неоднородная, 65, 157  
— интерполяции в  $R^n$ , 77  
— краевая  
— первая, 62  
— третья, 65  
— минимизации, 64
- Интерполяция  
— полиномиальная  
— лагранжева, 78  
— эрмитова, 78
- Квадрат единичный, 14
- Квадратура  
— Гаусса, 197  
— Лобатто, 197
- Квадратурной формулы  
— коэффициенты, 178  
— узлы, 178
- Квадратурных формул  
— примеры, 191, 196
- Конечного элемента  
— базисные функции, 83  
— степени свободы, 83, 84  
— узлы интерполяции, 83, 84
- Конечные элементы  
— ассоциированные, 100  
— аффинно-эквивалентные, 107  
— эрмитовы  
— равные, 112
- Конечный элемент, 75  
—  $m$ -мерный, 83  
— ассоциированный с базисным, 101  
— аффинно-эквивалентный, 101  
— базисный, 77, 87  
— прямоугольный, 87  
—  $(m, s)$ -неполный, 93  
— симплицальный, 87  
— изопараметрический, 101, 120  
— квадратичный, 121  
— прямоугольный второй степени, 128  
— четырехугольный первой степени, 126  
— эрмитов, 129  
— эрмитов кубический, 130
- Белла (класса  $C^1$ ), 97  
— изопараметрический с одной криволинейной границей, 122  
— класса  $C^k$ , 84  
— криволинейный, 102, 114  
— лагранжев, 83  
— линейный, 91  
— одномерный  
— лагранжев степени  $m$ , 88  
— эрмитов кубический класса  $C^1$ , 94  
— прямоугольный  
— лагранжев степени  $m$ , 88  
— максимально неполный степени  $m$ , 91  
— равный ассоциированному элементу Белла, 112  
— треугольный  
— лагранжев степени  $m$ , 90  
— эрмитов, 84  
— бикубический класса  $C^1$ , 95

- треугольный кубический класса  $C^0$ , 96
- Конечных элементов
  - Белла пространство, 140
  - лагранжевых треугольных степени  $m$  пространство, 138
  - пространство, 131
  - прямоугольных бикубических пространство, 141
  - семейство, 75
  - эрмитовых кубических треугольных пространство, 139
- Ламе постоянные, 69
- Лемма
  - Брамбла — Гильберта, 51, 52
  - Лакса — Мильграма, 21, 25, 27, 64
  - Сеа, 26
  - Стренга, 27, 179
- Матрица
  - Грама, 24
  - Якоби, 32
  - жесткости
    - глобальная, 220
    - локальная, 221
  - положительно определенная
    - равномерно, 62
  - разреженная, 15, 215
  - связности элементов, 216, 218
- Матрицы
  - жесткости элемента
    - вычисление, 227
- Метод
  - Галеркина, 8, 14, 25
  - возмущенный, 27, 179
  - Ритца, 22, 25
  - конечных элементов, 9
    - для системы уравнений теории упругости, 172
    - для уравнений второго порядка, 158, 163
    - для уравнений четвертого порядка, 165, 167
    - с численным интегрированием, 178
- Мультииндекс, 29, 78
- Мультииндекса
  - длина, 29
- Неравенство
  - Гельдера, 34, 56
  - обобщенное, 35
  - Коши — Буняковского, 35
  - Корна, 70
- Пуанкаре, 49
- Фридрихса, 47, 50
- треугольника, 28
- Норма
  - $\|u\|_{s,\infty,\Omega}$ , 38
  - $\|u\|_{s,p,\Omega}$ , 38
  - на пространстве Гильберта, 20
  - полилинейной формы, 29
  - симметричной, 30
- Нормы
  - эквивалентные, 22
- Нумерация
  - узлов
    - глобальная, 135, 216
    - локальная, 134, 216
- Область
  - $\Omega_h$ , 75
  - класса  $C^{k,\lambda}$ , 38
  - липшицева, 39
- Оператор
  - Лапласа, 65
  - вложения, 40
    - вполне непрерывный, 40
  - интерполяции, 82
  - положительно определенный, 20
  - самосопряженный, 21
  - энергетически эквивалентный, 25
- Остаточный член
  - в интегральной форме, 32
  - в форме Лагранжа, 33
  - в форме Пеано, 33
- Отображение
  - полилинейное, 30
- Отображения
  - якобиан, 32
- Оценка погрешности метода конечных элементов
  - асимптотическая неумлучшаемость, 1170
  - для уравнений второго порядка, 160, 161, 163, 164
  - для уравнений четвертого порядка, 167
  - для уравнений второго порядка в норме  $L_2$ , 170
- Оценка погрешности схемы МКЭ с численным интегрированием, 190, 199, 209
- Погрешности интерполяции
  - оценка, 147, 149–152, 154, 156
- Полунорма

- $[u]_{s,\infty,\Omega}$ , 38
- $[u]_{s,p,\Omega}$ , 38
- $|u|_{s,\infty,\Omega}$ , 39
- $|u|_{s,p,\Omega}$ , 38
- Проектор, 82
- Произведение
  - скалярное, 20
  - $(u, v)_{s,\Omega}$ , 40
  - энергетическое, 22
- Производная
  - конормальная, 65
  - обобщенная, 36
- Пространств
  - семейство
  - предельно полное, 23
- Пространство
  - $C(\bar{\Omega})$ , 34
  - $C[0, 1]$ , 7
  - $C^\infty(\bar{\Omega})$ , 39
  - $C^k[0, 1]$ , 7
  - $C_0^k[0, 1]$ , 7
  - $C_0^\infty(\Omega)$ , 36
  - $H^s(\Omega)$ , 40
  - $H_0^s(\Omega)$ , 45
  - $L_\infty(\Omega)$ , 34
  - $L_p(\Omega)$ , 34
  - $R^n$ , 28
  - $V_h^0$ , 144
  - $V_h$ , 74
  - $W_p^s(\Omega)$ , 38, 40
  - $[R^n]^k$ , 28
  - $\overset{\circ}{W}_p^s(\Omega)$ , 45
  - Гильберта, 20
  - Соболева, 28, 38
  - полиномов
    - степени не выше  $m$  по каждой переменной, 78
    - степени не выше  $m$  по совокупности переменных, 78
- Регулярность
  - решений эллиптических уравнений, 72
- Решение
  - классическое, 61
  - обобщенное, 8, 61, 63, 65, 66
- Семейства аффинно-эквивалентных элементов регулярность
  - условия, 109
- Семейства конечных элементов
  - $k$ -регулярность, 104
  - $k$ -регулярность сильная, 104
- Семейства криволинейных элементов  $k$ -регулярность
  - условия, 116
- Семейства треугольных аффинно-эквивалентных элементов регулярность
  - условия, 111
- Семейство подпространств
  - предельно полное, 26
- Сетка, 9
- След функции, 43
- Существенная верхняя грань функции, 34
- Схема разностная, 11
- Тензор
  - деформаций, 68
  - напряжений, 68
- Теорема
  - Рисса, 20, 25
  - Соболева об эквивалентных нормировках, 46
- Теории упругости
  - задачи двумерные, 71
  - система уравнений, 67
- Теория Фредгольма, 64
- Тождество интегральное, 7, 14, 63
- Триангуляция, 75
- Узловой параметр, 84
- Унисольвентная пара, 80, 81
- Уравнение
  - Пуассона, 14
  - Эйлера, 59
  - бигармоническое, 74
  - дифференциальное
    - обыкновенное второго порядка, 7
    - эллиптическое
      - четвертого порядка, 73
- Условие
  - Липшица, 39
  - положительной определенности
  - равномерной, 27
- Условия
  - разрешимости схемы МКЭ с численным интегрированием, 185, 204
- Форма
  - билинейная, 27
  - симметричная, 21
  - непрерывная, 20
  - линейная, 27
  - непрерывная, 20

- полилинейная, 28
- кососимметричная, 30
- симметричная, 28
- положительно определенная, 20
- Формула
  - Тейлора, 32
  - интегрирования по частям, 44
  - квадратурная
    - на базисном элементе, 178
    - на произвольном элементе, 178
- Функции
  - $P$ -интерполянт, 82
  - $V_h$ -интерполянт, 137
  - базисные, 8, 14
    - Куранта, 16
    - с локальным носителем, 10
  - гессиан, 32
  - градиент, 32
  - дифференциал порядка  $k$ , 31
  - отрицательная часть, 64
  - положительная часть, 64
  - продолжение с сохранением нормы, 41
- Функционал
  - линейный
    - ограниченный, 20
    - квадратичный, 21
  - погрешности квадратурной формулы, 182
  - энергетический, 65
- Функция
  - класса
    - $C^{k,\lambda}$ , 39
    - финитная, 36
  - функция  $l$  раз непрерывно дифференцируемая, 31
- Эквивалентные нормировки соболевских пространств, 45
- Элемент конечный, 9, 75
- Элемент наилучшего приближения, 23
- Энергия
  - потенциальная, 71

Даутов Рафаил Замилович  
Карчевский Михаил Миронович

**ВВЕДЕНИЕ В ТЕОРИЮ  
МЕТОДА КОНЕЧНЫХ ЭЛЕМЕНТОВ**

Главный редактор Н.И. Колосова

Подписано в печать 15.08.2004  
Бумага офсетная. Печать ризографическая  
Формат 60x84 1/16. Усл. печ. л. 13,89  
Уч.-изд. л. 14,93. Тираж 250 экз. Заказ 7/68

Издательство  
"Казанский государственный университет  
им. В.И. Ульянова–Ленина"  
42008, г. Казань, ул. Кремлевская, 18

---

Отпечатано с готового оригинал-макета  
в типографии Издательского центра Казанского  
государственного университета им. В.И. Ульянова–Ленина  
42008, г. Казань, ул. Университетская, 17  
Тел. 38–05–96