

САНКТ-ПЕТЕРБУРГСКИЙ ГОСУДАРСТВЕННЫЙ
УНИВЕРСИТЕТ

Физический факультет

Кафедра
Вычислительной физики

Е. А. Яревский

Теоретические основы методов
компьютерного моделирования
Часть 1

Учебно-методическое пособие

Санкт Петербург
2010 г.

Теоретические основы методов компьютерного моделирования. Часть 1. – СПб., 2010

В учебно-методическом пособии рассмотрены методы решения дифференциальных уравнений в частных производных второго порядка, применяемые в системах компьютерного моделирования. Основное внимание уделяется методам вариационного типа, в частности методам конечных и граничных элементов. Подробно разобраны все аспекты этих подходов, в том числе постановка вариационных задач, построение базисных функций, триангуляция области, оценки погрешности. Пособие предназначено для студентов 3-4-го курсов.

1 Абстрактная задача минимизации.

В физике и технике состояние системы часто описывают в терминах *энергии* системы J . Энергия может зависеть от положения системы, ее ориентации, внутреннего состояния, взаимного положения еџ подсистем. В математическом описании переменная u , описывающая состояние системы, принадлежит некоторому пространству $u \in V$. Очень часто переменная описания не может принимать произвольные значения, а должна принадлежать некоторому множеству допустимых состояний U , являющемуся подмножеством пространства V . Состояние системы определяется минимумом ее энергии, что приводит к необходимости поиска $u \in U$, удовлетворяющей условию

$$J(u) = \inf_{v \in U} J(v).$$

Для различных систем, энергия системы может задаваться по-разному. Мы рассмотрим здесь наиболее простой, однако важный и распространенный случай квадратичного функционала энергии:

$$J(v) = \frac{1}{2}a(v, v) - f(v),$$

где $a(., .)$ – симметричная билинейная форма, а f – линейная форма.

Для математически корректной постановки задачи, мы должны потребовать выполнения следующих условий: пусть

- V – нормированное векторное пространство с нормой $\|.\|$,
- $a(., .)$ – непрерывная билинейная форма $a(., .) : V \times V \rightarrow \mathbf{R}$,
- f – непрерывная линейная форма $f : V \rightarrow \mathbf{R}$,
- U – непустое подмножество V .

Тогда **абстрактная задача минимизации** (АЗМ) формулируется так: найти u такое, что

$$u \in U \quad \text{и} \quad J(u) = \inf_{v \in U} J(v), \quad (1)$$

где функционал $J : V \rightarrow \mathbf{R}$ определен как

$$J(v) = \frac{1}{2}a(v, v) - f(v). \quad (2)$$

Теорема 1. *Предположим дополнительно:*

- (i) Пространство V полно
- (ii) U – замкнутое выпуклое подмножество из V
- (iii) билинейная форма $a(., .)$ симметрична и V -эллипична, т.е. существует $\alpha > 0$ такое, что для всех $v \in V$

$$a(v, v) \geq \alpha \|v\|^2. \quad (3)$$

Тогда задача (1) имеет единственное решение.

Доказательство. Билинейная форма $a(., .)$ является скалярным произведением на V . В силу ограниченности и V -эллиптичности, порожденная им норма $a(., .)$ эквивалентна норме $\|.\|$. Если мы наделим V этим скалярным произведением, то оно станет гильбертовым. По теореме Рисса о представлении линейного функционала, существует такой элемент из V , что

$$\forall v \in V \quad f(v) = a(\sigma f, v).$$

Теперь, в силу симметричности билинейной формы:

$$J(v) = \frac{1}{2}a(v, v) - a(\sigma f, v) = \frac{1}{2}a(v - \sigma f, v - \sigma f) - \frac{1}{2}a(\sigma f, \sigma f).$$

Таким образом, решение абстрактной задачи минимизации равносильно минимизации расстояния между элементом вне U и множеством U относительно нормы $\sqrt{a(., .)}$. Следовательно решение – это проекция σf элемента f на множество U относительно скалярного произведения $a(., .)$. По теореме о проекции такой элемент существует и единственен, так как U – замкнутое выпуклое подмножество пространства V . ■

1.1 Эквивалентные вариационные формулировки

Иногда бывает удобно переформулировать рассмотренную нами абстрактную задачу минимизации в вариационных терминах. Такая формулировка дается следующей теоремой.

Теорема 2. *Элемент u – решение абстрактной задачи минимизации (1) тогда и только тогда, если*

- $$u \in U \text{ и } \forall v \in U \quad a(u, v - u) \geq f(v - u) \quad (4)$$

в общем случае;

- $$u \in U \text{ и } \forall v \in U \quad a(u, v) \geq f(v); \quad a(u, u) = f(u) \quad (5)$$

если U – замкнутый выпуклый конус с вершиной в 0;

- $$u \in U \text{ и } \forall v \in U \quad a(u, v) = f(v) \quad (6)$$

если U – замкнутое подпространство.

Уравнение (6) обычно называют вариационным уравнением, а уравнения (4),(5) – вариационными неравенствами.

Доказательство. Уравнение (4). Решение u – проекция σf на U . В силу выпуклости U , проекция полностью характеризуется свойством

$$u \in U \text{ и } \forall v \in U \quad a(\sigma f - u, v - u) \leq 0.$$

Фактически, это означает что угол между векторами тупой. Переписывая неравенство в виде

$$a(u, v - u) \geq a(\sigma f, v - u) = f(v - u),$$

получаем уравнение (4).

Уравнение (5). Пусть U – замкнутый выпуклый конус с вершиной 0. Тогда $u + v$ принадлежит множеству U всякий раз, когда v принадлежит U . При замене v на $u + v$ в (4) получим

$$\forall v \in U \quad a(u, v) \geq f(v),$$

в том числе $a(u, u) \geq f(u)$. Выбирая $v = 0$ в уравнении (4), найдем

$$a(u, u) \leq f(u)$$

и, следовательно, $a(u, u) = f(u)$.

Уравнение (6). Если U – подпространство, то условие (5), записанное для u и $-v$, даст

$$\left. \begin{array}{l} a(u, v) \geq f(v) \\ a(u, v) \leq f(v) \end{array} \right\} \Rightarrow a(u, v) = f(v)$$

для всех $v \in U$. ■

Замечание 1. Проектирующее отображение линейно тогда и только тогда, когда U – подпространство. Таким образом задачи, соответствующие вариационным неравенствам, вообще говоря, нелинейны.

Замечание 2. Линейность на подпространстве обусловлена тем, что функционал (2) – квадратичный.

2 Абстрактная Вариационная Задача

В приложениях часто бывает полезным сформулировать вариационную задачу без отсылки к задаче минимизации. Например, это необходимо сделать, если условия Теоремы 1 не соблюдаются. Определим *Абстрактную Вариационную Задачу* (АВЗ), не обращаясь к функционалу энергии J .

АВЗ: найти такой элемент u , что

$$u \in U \text{ и } \forall v \in U \quad a(u, v - u) \geq f(v - u), \quad (7)$$

или найти такой элемент u , что

$$u \in U \text{ и } \left\{ \begin{array}{l} \forall v \in U \quad a(u, v) \geq f(v), \\ a(u, u) = f(u), \end{array} \right. \quad (8)$$

если U – конус с вершиной 0 ;
или найти такой элемент u , что;

$$u \in U \text{ и } \forall v \in U \quad a(u, v) = f(v), \quad (9)$$

если U – подпространство.

Теорема 2 показывает, что у таких задач существует единственное решение, если выполнены условия Теоремы 1.

Мы можем задаться вопросом, можно ли ослабить условия Теоремы 1 так, чтобы у АЗМ и АВЗ по прежнему существовало единственное решение. В общем случае, это невозможно. Однако ситуация меняется, если мы дополнительно предположим, что V – *гильбертово пространство*. Как мы увидим в дальнейшем, такое предположение выполняется для многих важных уравнений и приложений. В этом случае мы можем опустить требование о симметричности билинейной формы, но АВЗ будет по прежнему иметь единственное решение. Важно отметить, что при тех же условиях для АЗМ это не верно!

Основной результат о АВЗ с несимметричными билинейными формами дает лемма Лакса–Мильграма. Мы сформулируем ее для более простого случая $U = V$.

Теорема 3. (Лемма Лакса–Мильграма).

Пусть V – гильбертово пространство, билинейная форма

$$a(., .) : V \times V \rightarrow \mathbf{R}$$

непрерывна и V -эллипична, линейная форма $f : V \rightarrow \mathbf{R}$ непрерывна. Тогда АВЗ: Найти такой элемент u , что

$$u \in V \text{ и } \forall v \in V \quad a(u, v) = f(v) \quad (10)$$

имеет единственное решение.

Доказательство этой теоремы мы здесь приводить не будем. Оно основано на теореме о сжимающем отображении и может быть найдено, например, в книге [2]. ■

Нужно еще раз подчеркнуть, что АВЗ – это основная задача, которую мы будем изучать на протяжении нашего курса. Она даст представление для уравнений в частных производных, которое оказывается удобным для многих численных методов, в частности, для метода конечных элементов.

3 Формулы Грина

Для того, чтобы применить абстрактную теорию, развитую в предыдущих параграфах, к конкретным уравнениям, нам необходимо определить подходящие пространства функций и средства для работы с ними. В этом параграфе мы коротко сформулируем необходимые определения и результаты. Более подробное обсуждение может быть найдено в книгах [1, 2].

Пусть Ω – ограниченная замкнутая область в \mathbb{R}^n с гладкой границей Γ . Мы будем использовать следующие функциональные пространства в области Ω :

- пространство непрерывных функций $C(\Omega)$ с нормой

$$\|v\|_{C(\Omega)} = \max_{x \in \Omega} |v(x)|,$$

- пространство квадратично интегрируемых функций $L^2(\Omega)$ с нормой

$$\|v\|_{0,\Omega} = \left(\int_{\Omega} |v(x)|^2 dx \right)^{1/2},$$

- пространство Харди $H^1(\Omega)$ непрерывных дифференцируемых функций с нормой

$$\|v\|_{1,\Omega} = \left(\sum_{|\alpha| \leq 1} \int_{\Omega} |\partial^\alpha v(x)|^2 dx \right)^{1/2},$$

- и его подпространство $H_0^1(\Omega)$ – пространство функций, принадлежащих $H^1(\Omega)$ и обращающихся в нуль на границе области:

$$H_0^1(\Omega) = \{v \in H^1(\Omega) : v(x) = 0 \text{ при } x \in \Gamma\}.$$

Важным средством для работы с дифференциальными уравнениями являются формулы Грина. Они дают естественное многомерное обобщение для одномерной формулы интегрирования по частям.

Фундаментальная формула Грина:

$$\int_{\Omega} u \partial_i v dx = - \int_{\Omega} \partial_i u v dx + \int_{\Gamma} u v \nu_i d\gamma \quad (11)$$

для $u, v \in H^1(\Omega)$ и для всякого $i \in [1, n]$. В этом уравнении Γ – граница множества, ν – единичная внешняя нормаль, $d\gamma$ – поверхностная мера.

Заменяя в уравнении (11) u на $\partial_i u$ и беря сумму от 1 до n , получим

$$\int_{\Omega} \sum_{i=1}^n \partial_i u \partial_i v dx = - \int_{\Omega} \Delta u v dx + \int_{\Gamma} \partial_\nu u v d\gamma \quad (12)$$

для всех $u \in H^2(\Omega)$, $v \in H^1(\Omega)$, где производная по внешней нормали к поверхности

$$\partial_\nu = \sum_{i=1}^n \nu_i \partial_i.$$

Аналогично, заменяя v на $\partial_i v$, получим

$$\int_{\Omega} u \Delta v dx = - \int_{\Omega} \sum_{i=1}^n \partial_i u \partial_i v dx + \int_{\Gamma} u \partial_{\nu} v d\gamma$$

для всех $u \in H^1(\Omega)$, $v \in H^2(\Omega)$. Вычисляя разность этих двух равенств, найдем еще одну формулу Грина:

$$\int_{\Omega} (u \Delta v - \Delta uv) dx = \int_{\Gamma} (u \partial_{\nu} v - \partial_{\nu} uv) d\gamma. \quad (13)$$

4 Примеры краевых задач второго порядка. Граничные условия.

Для того, чтобы проиллюстрировать применение полученных ранее общих результатов, мы рассмотрим два примера краевых задач для дифференциальных уравнений второго порядка в частных производных.

4.1 Задача 1: Краевая задача второго порядка – задача Дирихле.

В этом примере пространства U и V в общей теории выбраны одинаковыми и равными $H_0^1(\Omega)$:

$$V = U = H_0^1(\Omega).$$

Билинейный функционал выберем в виде

$$A(u, v) = \int_{\Omega} \left(\sum_{i=1}^n \partial_i u(x) \partial_i v(x) + a(x) u(x) v(x) \right) dx,$$

а линейный функционал определим как

$$F(v) = \int_{\Omega} f(x) v(x) dx.$$

От функций $a(x)$ и $f(x)$ потребуем выполнения следующих условий:

$$a \in C(\Omega), \quad a(x) \text{ ограничена и } a(x) \geq 0 \text{ на } \Omega,$$

$$f \in L^2(\Omega).$$

Анализ задачи.

Дважды используя неравенство Коши-Буняковского $|(u, v)| \leq \|u\| \|v\|$, для любых $u, v \in H^1(\Omega)$ мы получаем следующую оценку:

$$\begin{aligned} |A(u, v)| &\leq \sum_{i=1}^n \|\partial_i u\|_{0,\Omega} \|\partial_i v\|_{0,\Omega} + \|a\|_{C(\Omega)} \|u\|_{0,\Omega} \|v\|_{0,\Omega} \leq \\ &\leq \max\{1, \|a\|_{C(\Omega)}\} \|u\|_{1,\Omega} \|v\|_{1,\Omega}. \end{aligned}$$

Таким образом мы видим, что билинейный функционал $A(u, v)$ является непрерывным в пространстве $H^1(\Omega)$. Необходимо отметить, что правильный выбор пространства очень важен: например, этот же функционал $A(u, v)$ не является непрерывным в норме пространства $L^2(\Omega)$, даже на гладких функциях.

В силу положительности функции $a(x)$, для всех $v \in H^1(\Omega)$ мы имеем

$$A(v, v) \geq \int_{\Omega} \sum_{i=1}^n (\partial_i v)^2 dx.$$

В силу *неравенства Пуанкаре-Фридрихса* для ограниченной области Ω

$$\|v\|_{0,\Omega} \leq C(\Omega) \left(\int_{\Omega} \sum_{i=1}^n (\partial_i v)^2 dx \right)^{1/2} \quad \forall v \in H_0^1(\Omega),$$

мы получаем оценку

$$A(v, v) \geq \frac{1}{2} \int_{\Omega} \sum_{i=1}^n (\partial_i v)^2 dx + \frac{1}{2C^2(\Omega)} \|v\|_{0,\Omega}^2 \geq \min \left\{ \frac{1}{2}, \frac{1}{2C^2(\Omega)} \right\} \|v\|_{1,\Omega}^2.$$

Таким образом, функционал $A(v, v)$ является также $H_0^1(\Omega)$ -эллиптическим.

Оценим теперь линейный функционал $F(v)$. Для любой функции $v \in H^1(\Omega)$ мы имеем:

$$|F(v)| \leq \|f\|_{0,\Omega} \|v\|_{0,\Omega} \leq \|f\|_{0,\Omega} \|v\|_{1,\Omega},$$

то есть наш линейный функционал непрерывен.

Таким образом, мы доказали все условия теорем Л1.1 и Л1.2. Следовательно, существует единственная функция $u \in H_0^1(\Omega)$, минимизирующая квадратичную форму

$$J(u) = \frac{1}{2} \int_{\Omega} \left(\sum_{i=1}^n (\partial_i u(x))^2 + a(x)u^2(x) \right) dx - \int_{\Omega} f(x)u(x) dx \quad (14)$$

на $H_0^1(\Omega)$.

По теореме Л1.2, функция $u(x)$ удовлетворяет также вариационному уравнению

$$\forall v \in H_0^1(\Omega) \quad \int_{\Omega} \left(\sum_{i=1}^n \partial_i u(x) \partial_i v(x) + a(x)u(x)v(x) \right) dx = \int_{\Omega} f(x)v(x) dx. \quad (15)$$

Связь с уравнением в частных производных:

Функция $u(x)$, решающая задачи (14) и (15), дает решение задачи минимизации и вариационной задачи, соответственно. Однако, поскольку эти задачи сформулированы в терминах функций в области, естественно попытаться найти дифференциальное уравнение, соответствующее этим задачам. Для этого мы

введем пространство гладких финитных функций $\mathcal{D}(\Omega)$. Применим к нашему билинейному функционалу формулу Грина:

$$\int_{\Omega} \sum_{i=1}^n \partial_i u(x) \partial_i v(x) dx = - \int_{\Omega} \Delta u(x) v(x) dx + \int_{\Gamma} \partial_\nu u v d\gamma. \quad (16)$$

Тогда для любой гладкой функции $\forall \phi \in \mathcal{D}(\Omega)$ мы получим:

$$A(u, \phi) = \int_{\Omega} (-\Delta u(x) + a(x)u(x))\phi dx = \langle -\Delta u + au, \phi \rangle,$$

и

$$f(\phi) = \langle f, \phi \rangle.$$

Таким образом, из вариационного уравнения (15) следует, что u – решение в $\mathcal{D}'(\Omega)$ уравнения

$$\begin{aligned} -\Delta u + au &= f \text{ в } \Omega \\ u &= 0 \text{ на } \Gamma. \end{aligned} \quad (17)$$

Следовательно, искомая дифференциальная задача является однородной задачей Дирихле. Граничные условия задачи (17) называются *главными* граничными условиями (*essential boundary conditions*).

Нужно отметить, что более общая неоднородная задача Дирихле

$$\begin{aligned} -\Delta u + au &= f \text{ в } \Omega \\ u &= g(x) \text{ на } \Gamma \end{aligned}$$

сводится к задаче (17) заменой функции u на $u - \tilde{u}$, где \tilde{u} выбирается так, чтобы

$$\tilde{u}|_{\Gamma} = g(x).$$

4.2 Задача 2: Краевая задача второго порядка – задача Неймана.

В этом примере пространства U и V вновь выбраны одинаковыми,

$$V = U = H^1(\Omega),$$

билинейный функционал выбран таким же как в задаче 1,

$$A(u, v) = \int_{\Omega} \left(\sum_{i=1}^n \partial_i u(x) \partial_i v(x) + a(x)u(x)v(x) \right) dx,$$

а линейный функционал включает интеграл по границе области:

$$F(v) = \int_{\Omega} f(x)v(x) dx + \int_{\Gamma} gv d\gamma.$$

Функции a , f и g удовлетворяют условиям

$$a \in C(\Omega), \text{ ограничена и } a \geq a_0 > 0 \text{ на } \Omega,$$

$$f \in L^2(\Omega), \quad g \in L^2(\Gamma).$$

Анализ задачи:

Проводя те же выкладки, что и в задаче 1, мы видим что функционал (u, v) непрерывен. С другой стороны, для всех $v \in H^1(\Omega)$ мы можем оценить:

$$A(v, v) = \int_{\Omega} \sum_{i=1}^n (\partial_i v(x))^2 dx + \int_{\Omega} a(x)v^2(x) dx \geq \min \{1, a_0\} \|v\|_{1,\Omega}^2.$$

Таким образом, функционал (v, v) является также и $H^1(\Omega)$ -эллиптическим.

Линейная форма $v \in H^1(\Omega) \rightarrow \int_{\Gamma} gv d\gamma$ непрерывна, т.к. в силу существования следа функции мы можем оценить ее как

$$\left| \int_{\Gamma} gv d\gamma \right| \leq \|g\|_{L^2(\Gamma)} \|v\|_{L^2(\Gamma)} \leq C(\Omega) \|g\|_{L^2(\Gamma)} \|v\|_{1,\Omega}.$$

Мы вновь можем использовать теорему Л1.1, в соответствии с ней существует единственная $u \in H^1(\Omega)$, минимизирующая функционал

$$J(v) = \frac{1}{2} \int_{\Omega} \left(\sum_{i=1}^n (\partial_i v(x))^2 + a(x)v^2(x) \right) dx - \int_{\Omega} f(x)v(x) dx - \int_{\Gamma} gv d\gamma$$

на пространстве $H^1(\Omega)$.

По теореме Л1.2, u удовлетворяет $\forall v \in H^1(\Omega)$

$$\int_{\Omega} \left(\sum_{i=1}^n \partial_i u \partial_i v + auv \right) dx = \int_{\Omega} fv dx + \int_{\Gamma} gv d\gamma.$$

Связь с уравнением с частными производными:

Для достаточно гладких решений (например, для решений с квадратично интегрируемыми вторыми производными $u \in H^2(\Omega)$) по формуле Грина:

$$a(u, v) = \int_{\Omega} (-\Delta u + au)v dx + \int_{\Gamma} \partial_{\nu} uv d\gamma = \int_{\Omega} fv dx - \int_{\Gamma} gv d\gamma.$$

Это соотношение должно выполняться для всех функций v , в том числе и функций, равных нулю на границе. Тогда мы получаем, что должно выполняться уравнение

$$-\Delta u + au = f \text{ в } \Omega. \tag{18}$$

Так как это уравнение должно быть верно для всех функций, мы получаем:

$$\forall v \in H^1(\Omega) \quad \int_{\Gamma} \partial_{\nu} uv d\gamma = \int_{\Gamma} gv d\gamma,$$

и, следовательно,

$$\partial_\nu u = g \text{ на } \Gamma. \quad (19)$$

Уравнения (18,19) определяют *неоднородную задачу Неймана*. Эта задача называется *однородной* при выполнении условия $g = 0$. Граничные условия в *однородной задаче Неймана* называются *естественными* граничными условиями (*natural boundary conditions*). Слагаемое

$$\int_{\Gamma} gv \, d\gamma$$

является способом введения в задачу Неймана дополнительных (ненулевых) краевых условий.

5 Методы аппроксимации

Рассмотрим стандартную вариационную задачу, т.е. будем искать функцию $u \in V$ такую, что

$$\forall v \in V \quad A(u, v) = F(v).$$

В дальнейшем мы предполагаем, что выполнены все условия леммы Лакса-Мильграма. Как правило, пространство V является бесконечномерным, $\dim V = \infty$, так что мы можем найти точное решение вариационной задаче только в самых простых случаях. Поскольку при компьютерном моделировании, мы можем работать только с конечномерными пространствами, представляется естественным искать решение в конечномерном подпространстве V_N пространства V , $\dim V_N = N$. Конечно, такое решение будет лишь приближенным, а не точным.

Пример 1. В качестве примера конечномерного пространства, мы отметим здесь широко используемое в дальнейшем пространство $P(k)$ всех полиномов степени меньше или равной k над пространством \mathbb{R}^n . Для наиболее важных для нас размерностей пространства n , размерность $P(k)$ такова:

$$\begin{aligned} \dim P(k) &= k + 1 && \text{в } R^1, \\ \dim P(k) &= (k + 1)(k + 2)/2 && \text{в } R^2, \\ \dim P(k) &= (k + 1)(k + 2)(k + 3)/6 && \text{в } R^3. \quad \blacksquare \end{aligned}$$

Когда мы используем конечномерную аппроксимацию исходной вариационной задачи, мы должны проверить выполнение некоторых условий для того, чтобы получить достоверные численные результаты. К этим условиям относятся:

- Существование и единственность конечномерного решения.

- Сходимость: по подходящей норме должно выполняться условие

$$\|u - u_N\| \rightarrow 0 \text{ когда } N \rightarrow \infty.$$

- Оценка скорости сходимости: как именно (с каким p) убывает норма погрешности:

$$\|u - u_N\| \sim C/N^p \text{ при } N \rightarrow \infty.$$

Для того чтобы математически корректно разрешить эти вопросы, мы должны в первую очередь аппроксимировать пространство V . Пусть $\{V_n\}_{n=1}^{\infty} \subset V$ – последовательность подпространств V таких, что $\overline{\bigcup_{i=1}^{\infty} V_n} = V$, где $V_n \subset V_{n+1} \subset V$, и $\dim V_n = N_n < \infty$. Определим теперь *дискретную задачу*: найти *дискретное решение* $u_n \in V_n$ такое, что

$$A(u_n, v) = F(v) \quad \text{для всех } v \in V_n. \quad (20)$$

Необходимо отметить, что если исходная вариационная задача удовлетворяет условиям леммы Лакса-Мильграма в пространстве V , то, как правило, соответствующая дискретная задача также удовлетворяет условиям леммы. Единственное, что необходимо проверить – выполнение оценок по норме уже в нормированном конечномерном пространстве. Если условие леммы выполнено, то дискретная задача также имеет единственное решение. Такой подход к конечномерной аппроксимации, и в частности задача (20), называется *методом Галеркина*.

Если билинейная форма $A(u, v)$ симметрична, мы можем рассмотреть дискретную задачу минимизации вместо дискретной вариационной задачи:

$$J(u_N) = \inf_{v_N \in V_N} \left(\frac{1}{2} A(v_N, v_N) - F(v_N) \right). \quad (21)$$

Такой подход, и задача (21), называются *методом Рунца*. Сравнивая теоремы 1.1 и 1.2, можно убедиться, что Метод Галеркина приводит к тем же самым уравнениям, что и метод Рунца, если билинейная форма симметрична и для u_N и v_N в уравнении (20) выбран одинаковый базис.

Рассмотрим подробнее, как именно применяется метод Галеркина. Так как пространство U_N конечномерно, в нем существует базис из N элементов $\phi_k|_{k=1}^N$, и любая функция u_N может быть разложена по этому базису: $u_N(x) = \sum_{k=1}^N \alpha_k \phi_k$. Вариационное равенство должно быть выполнено для всех функций v_N , в том числе и для произвольного базиса $\psi_k|_{k=1}^N$. Записывая эти N вариационных равенств, получим:

$$\sum_{k=1}^N A(\phi_k, \psi_i) \alpha_k = f(\psi_i), \quad i = 1 \dots N. \quad (22)$$

Это равенство удобно записать в матричной форме:

$$\hat{A} \hat{\alpha} = \hat{f}, \quad \text{где } \hat{A}_{ik} = A(\phi_k, \psi_i), \quad \hat{\alpha}_i = \alpha_i, \quad \hat{f}_i = f(\psi_i). \quad (23)$$

Таким образом, решение вариационного равенства и определение решения u_N сводится к решению линейной системы (23). Из леммы Лакса-Мильграма мы знаем, что решение этой системы всегда существует и единственно.

Следует особо подчеркнуть, что существование и единственность решения дискретной задачи не гарантируют сходимость этого решения к точному при $N \rightarrow \infty$. Действительно, рассмотрим следующий

Пример 2. Будем решать следующее дифференциальное уравнение на отрезке $[0, 1]$:

$$-\frac{d^2}{dx^2}u(x) = 2 \quad (24)$$

в классе функций $H_0^1([0, 1])$. Поскольку линейный оператор симметричен, мы можем использовать метод Ритца. В качестве конечномерных пространств V_N выберем линейные оболочки следующих функций: $\phi_k(x) = \sin(2\pi kx)$, очевидно принадлежащие $H_0^1([0, 1])$. Любая функция из $v_N \in V_N$ может быть представлена как линейная комбинация $v_N(x) = \sum_{k=1}^N \alpha_k \phi_k(x)$. Чтобы найти коэффициенты α_k , мы должны минимизировать следующий функционал:

$$J(v_N) = \frac{1}{2} \int_0^1 (v'_N(x))^2 dx - \int_0^1 2 \sum_{k=1}^N \alpha_k \phi_k dx. \quad (25)$$

Второй интеграл в этом выражении равен нулю, а первый всегда неотрицателен. Тем самым, его минимум равен нулю – если он достигается. В нашем случае это действительно происходит при всех $\alpha_k = 0$. Таким образом, решение $v_N(x) \equiv 0$ для любого N . В то же время легко видеть, что точное решение задачи (24) равно $v(x) = x(1 - x)$, и погрешность аппроксимации вовсе не зависит от N и не стремится к нулю. ■

5.1 Метод наименьших квадратов

Метод Галеркина является мощным и достаточно общим подходом к вариационным задачам. Для иллюстрации этого мы рассмотрим еще один пример. Именно, мы покажем как из метода Галеркина можно получить *метод наименьших квадратов* (МНК). Для этого, как обычно, рассмотрим вариационную задачу

$$(Au_N, v_N) = (f, v_N). \quad (26)$$

Заменим теперь функции v_N на Av_N . Это всегда можно сделать, если оператор A обратим. Тогда задача (26) может быть переписана в виде:

$$\begin{aligned} (Au_N, Av_N) &= (f, Av_N), \\ (A^*Au_N, v_N) &= (A^*f, v_N). \end{aligned} \quad (27)$$

Оператор A^*A симметричен, поэтому мы можем применить к задаче (27) метод Рунца: искомое решение определяется как минимум функционала

$$\begin{aligned} \frac{1}{2}(A^*Av_N, v_N) - (A^*f, v_N) &= \frac{1}{2}(Av_N, Av_N) - (f, Av_N) = \\ &= \frac{1}{2}(Av_N - f, Av_N - f) - \frac{1}{2}(f, f) = \frac{1}{2}\|Av_N - f\|^2 - \|f\|^2. \end{aligned}$$

Таким образом, минимум этого функционала совпадает с минимумом функционала МНК, т.е. $\|Av_N - f\|^2$.

6 Ортогональность ошибок и лемма Шеа (Sea)

Для эллиптических уравнений, рассмотренных нами в предыдущих лекциях, ошибка $e_n = u - u_n$ решения дискретной задачи обладает свойством *ортогональности*.

Лемма. Пусть $u \in V$ – точное решение, а u_n – решение дискретной задачи. Тогда

$$A(u - u_n, v) = A(e_n, v) = 0 \quad \text{для всех } v \in V_n. \quad (28)$$

Доказательство леммы элементарно. Так как u_n – решение дискретной задачи, оно удовлетворяет следующему равенству:

$$A(u_n, v) = F(v) \quad \text{для всех } v \in V_n.$$

Для точной задачи, вариационное уравнение должно быть выполнено для всех $v \in V$, и, в частности, для v из подпространства V_n :

$$A(u, v) = F(v) \quad \text{для всех } v \in V_n \subset V.$$

Вычитая эти два уравнения друг из друга и пользуясь линейностью формы, мы получаем требуемый результат. ■

Полезно обсудить геометрический смысл свойства ортогональности. Если форма $A(., .)$ симметрична, мы можем ввести энергетическое скалярное произведение

$$(u, v)_e = A(u, v),$$

и $(e_n, v)_e = 0$ для всех $v \in V_n$. Таким образом, ошибка e_n ортогональна Галеркинскому подпространству V_n в энергетическом скалярном произведении. Таким образом, u_n является ортогональной проекцией решения $u \in V$ на V_n и, следовательно, является ближайшим элементом к u :

$$\|u - u_n\|_e = \inf_{v \in V_n} \|u - v\|_e.$$

Теорема (лемма Шеа). Пусть V – гильбертово пространство, $A(., .) : V \times V \rightarrow \mathbb{R}$ – билинейная ограниченная V -эллиптическая форма. $u \in V$ – решение

точной задачи, и подпространство $V_n \subset V$, $u_n \in V_n$. Обозначим C_b и C_{el} константы непрерывности и V -эллиптичности формы A . Тогда

$$\|u - u_n\|_V \leq \frac{C_b}{C_{el}} \inf_{v \in V_n} \|u - v\|_V.$$

Важность леммы Шеа заключается в том, что она показывает независимость ошибки от конкретного выбора базиса в подпространстве V_n . Хотя доказательство леммы не очень сложное, мы не приводим его здесь, см.[?].

Лемма также может быть использована для доказательства сходимости решения дискретной задачи к решению точной.

Теорема. Пусть выполнены условия леммы Шеа. Пусть дополнительно последовательность подпространств V_n будет выбрана так, что $V_1 \subset V_2 \dots \subset V_n \dots \subset V$, и

$$\overline{\bigcup_{i=1}^{\infty} V_n} = V. \quad (29)$$

Тогда

$$\lim_{n \rightarrow \infty} \|u - u_n\|_V = 0.$$

Доказательство. Пусть $u \in V$ – точное решение вариационной задачи. В соответствии с уравнением (29), можно найти такую последовательность $\{v_n\}$, $v_n \in V_n$, что $\lim_{n \rightarrow \infty} \|u - v_n\|_V = 0$. Пусть $u_n \in V_n$ – решение дискретной задачи. Тогда мы можем оценить соответствующую ошибку с помощью леммы Шеа следующим образом:

$$\|u - u_n\|_V \leq \frac{C_b}{C_{el}} \inf_{v \in V_n} \|u - v\|_V \leq \frac{C_b}{C_{el}} \|u - v_n\|_V$$

для любого n . Так как правая часть неравенства стремится к нулю когда n стремится к бесконечности, левая часть так же стремится к нулю. ■

7 Основные аспекты метода конечных элементов (МКЭ)

Начиная с этого момента, мы будем обсуждать различные вычислительные варианты метода Галеркина. Основное внимание будет уделено методу конечных элементов (МКЭ) для дифференциальных уравнений второго порядка эллиптического типа. Такие уравнения являются сравнительно простыми, однако на их примере мы сможем рассмотреть все основные особенности и свойства МКЭ. В конце курса мы рассмотрим применение МКЭ для уравнений второго порядка других типов, а также другой метод, тесно связанный с МКЭ, – метод граничных элементов.

В методе Галеркина для дифференциальных уравнений, подобных рассмотренным в предыдущей лекции, мы решаем вариационную задачу: найти $u \in V$ такую, что для всех $v \in V$ выполнено вариационное уравнение

$$\int_{\Omega} \left(\sum_{i=1}^n \partial_i u(x) \partial_i v(x) + a(x) u(x) v(x) \right) dx = \int_{\Omega} f(x) v(x) dx.$$

Затем мы переходим к дискретной задаче, рассматривая подходящее подпространство V_N функционального пространства V . Различные способы выбора подпространства приводят нас к различным численным методам. Одним из таких методов является МКЭ, который обычно связывается со следующими основными подходами:

МКЭ 1. Осуществление **триангуляции** \mathcal{F}_h на области Ω . Это означает, что Ω разбивается на конечное число подмножеств K_i таких, что

1. $\Omega = \bigcup_{K \in \mathcal{F}_h} K$.
2. Граница каждого ∂K_i – кусочно-гладкая.
3. $(K_i \setminus \partial K_i) \cap (K_j \setminus \partial K_j) = \emptyset$, то есть подмножества могут пересекаться только по границам.

Выберем конечномерное пространство функций, называемое также *пространством конечных элементов* V_h . Определим пространства проекций V_h на элементы K_i :

$$P_{K_i} = \{v_h|_K : v_h \in V_h\}.$$

Тогда следующее свойство МКЭ можно сформулировать так:

МКЭ 2. Пространства P_{K_i} содержат многочлены, или функции "близкие" к многочленам.

Чтобы пояснить важность этого свойства, вспомним как решается дискретная задача. Пусть $\{w_k(x)\}_{k=1}^N$ – базис в V_h . Тогда решение

$$u_h(x) = \sum_{k=1}^N u_k w_k(x)$$

находится из дискретного вариационного уравнения

$$\sum_{k=1}^N u_k \int_{\Omega} \left(\sum_{i=1}^n \partial_i w_k(x) \partial_i w_l(x) + a(x) w_k(x) w_l(x) \right) dx = \int_{\Omega} f(x) w_l(x) dx$$

для всех l таких, что $1 \leq l \leq N$. Таким образом, сложность вычисления коэффициентов линейной системы определяется сложностью функций $w_k(x)$. Наиболее просто такое вычисление проводится для полиномов, как в аналитическом, так и в численном вариантах. Кроме того, как мы увидим далее, для полиномиальных функций достаточно просто оценить скорость сходимости дискретного решения к точному.

Третье свойство требует выполнения некоторого условия на базис в пространстве V_h , именно:

МКЭ 3. В V_h существует по крайней мере один базис из функций с "минимальными" носителями, то есть из функций, не равных нулю только на нескольких конечных элементах.

Это требование приводит к разреженности матрицы линейной системы, что существенно уменьшает трудоемкость ее решения. Такой базис существует всегда, однако в МКЭ 3 подразумевается, что такой базис можно легко и явно предъявить.

Среди разнообразных вариантов МКЭ выделяется специальный класс методов, анализ которых достаточно прост. Это так называемые *конформные МКЭ*. Такие методы выделяются тем, что в них

- V_h – подпространство V ,
- билинейная и линейная форма в дискретной вариационной задаче совпадают с исходными.

Следует отметить, что мы столкнемся и с ситуациями, когда такие методы неприменимы. Такое может случиться, например, когда

- 1 Граница области Ω криволинейна. Тогда ее точная триангуляция затруднена, используется аппроксимацию Ω_h . В этом случае V_h не является подпространством V .
- 2 Функции из V_h имеют разрывы, то есть вновь V_h не является подпространством V .
- 3 Для вычисления форм используется аппроксимация (например, численное интегрирование), следовательно они не совпадают с исходными.

Анализ неконформных МКЭ более сложен, и мы лишь обсудим в дальнейшем некоторые результаты для этих МКЭ.

8 Одномерный МКЭ.

Рассмотрим задачу Дирихле для простого одномерного уравнения второго порядка:

$$\begin{aligned} -(p(x)u'(x))' + q(x)u(x) &= f(x), \quad 0 \leq x \leq 1, \\ u(0) &= u(1) = 0, \end{aligned}$$

где все функции $p(x)$, $q(x)$ и $f(x)$ – гладкие, а две первые – положительные: $p(x) > 0$, $q(x) > 0$. Вариационная форма для этой задачи формулируется так:

найти $u(x)$ такую, что для любой $v(x)$ выполнено

$$\int_0^1 (u'(x)p(x)v'(x) + q(x)u(x)v(x)) dx = \int_0^1 f(x)v(x)dx. \quad (30)$$

Соответствующая задача минимизации имеет вид: найти $u(x)$, минимизирующую функционал

$$\int_0^1 (p(x)u'^2(x) + q(x)u^2(x) - 2f(x)u(x)) dx. \quad (31)$$

Из построенной нами ранее теории следует, что решения задач (30) и (31) существуют, являются единственными и совпадают друг с другом.

Давайте построим простейший МКЭ для этой задачи.

1. Триангуляция в одномерном случае задается естественным делением отрезка на части:

$$x_0 < x_1 < \dots < x_N.$$

2. Выбор элементных функций основан, как правило, на интерполяции Лагранжа.

В простейшем случае кусочно-линейных полиномов мы можем записать:

$$\phi_j(x) = \begin{cases} \frac{x-x_{j-1}}{x_j-x_{j-1}}, & \text{при } x_{j-1} \leq x < x_j, \\ \frac{x_{j+1}-x}{x_{j+1}-x_j}, & \text{при } x_j \leq x < x_{j+1}, \\ 0, & \text{для всех других } x. \end{cases}$$

По построению эти функции непрерывны и отличны от 0 только на двух интервалах каждая.

8.1 Лагранжевы элементы.

Для анализа лагранжевых конечных элементов (КЭ), рассмотрим вначале квадратичный случай. На каждом интервале $K_j = [x_{j-1}, x_j]$ длины $h_j = x_j - x_{j-1}$ добавим среднюю точку с полуцелым индексом $x_{j-1/2}$ и построим базис, состоящий из следующих функций:

$$\phi_j(x) = \begin{cases} 1 + 3\left(\frac{x-x_j}{h_j}\right) + 2\left(\frac{x-x_j}{h_j}\right)^2, & \text{при } x_{j-1} \leq x < x_j, \\ 1 - 3\left(\frac{x-x_j}{h_{j+1}}\right) + 2\left(\frac{x-x_j}{h_{j+1}}\right)^2, & \text{при } x_j \leq x < x_{j+1}, \\ 0, & \text{для всех других } x, \end{cases}$$

$$\phi_{j-1/2}(x) = \begin{cases} 1 - 4\left(\frac{x-x_{j-1/2}}{h_j}\right)^2, & \text{при } x_{j-1} \leq x < x_j, \\ 0, & \text{для всех других } x. \end{cases}$$

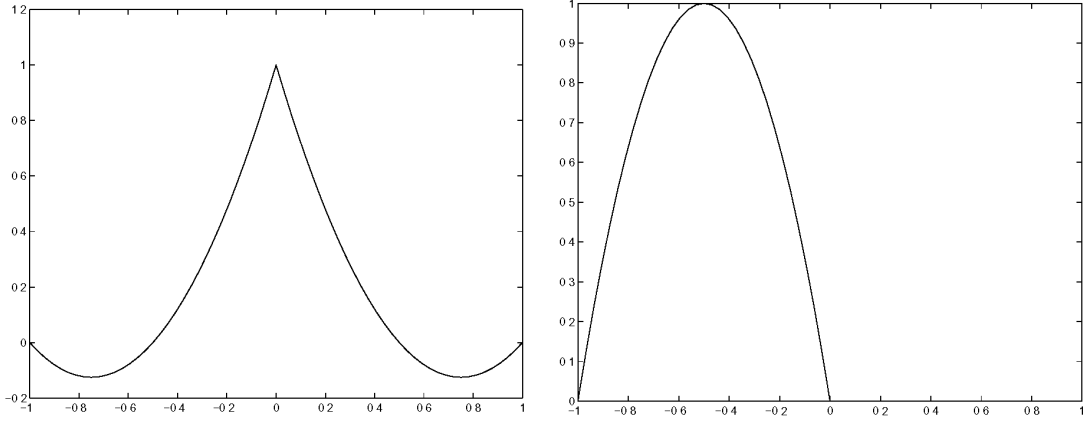


Рис. 1: Лагранжевы квадратичные функции (слева – для границы $x = 0$, справа – для средней точки $x = -0.5$.) Узлы $x_{j-1} = -1$, $x_{j-1/2} = -0.5$, $x_j = 0$, $x_{j+1} = 1$.

Пример таких функций приведен на рис. 1. Этот базис удовлетворяет свойствам МКЭ 2 и МКЭ 3, так как проекции на интервалы – полиномы, а перекрытие различных функций мало. Кроме того, функции, образующие этот базис, обладают свойством Лагранжа: они равны 1 в "своих" узлах, и 0 во всех остальных:

$$\phi_j(x_k) = \begin{cases} 1, & \text{при } j = k \\ 0, & \text{при } j \neq k \end{cases}, \quad j, k = 0, 1/2, 1, \dots, N.$$

Элементные базисные функции

Элементными базисными функциями называются проекции глобальных базисных функций ϕ_j на элементы K_i . В случае квадратичных базисных функций, введенных выше, на каждом элементе присутствуют три ненулевых функции:

$$\begin{aligned} N_{j-1,j}(x) &= 1 - 3 \left(\frac{x - x_{j-1}}{h_j} \right) + 2 \left(\frac{x - x_{j-1}}{h_j} \right)^2, \\ N_{j-1/2,j}(x) &= 1 - 4 \left(\frac{x - x_{j-1/2}}{h_j} \right)^2, \\ N_{j,j}(x) &= 1 + 3 \left(\frac{x - x_j}{h_j} \right) + 2 \left(\frac{x - x_j}{h_j} \right)^2. \end{aligned}$$

В качестве примера, эти функции для интервала $[0,1]$ изображены на рис. 2.

В общем случае, элементные базисные функции $N_{k,j}$ не равны нулю только на своем собственном интервале, или, другими словами, только если узел k принадлежит элементу K_j . Это обеспечивает малое ненулевое перекрытие между такими функциями.

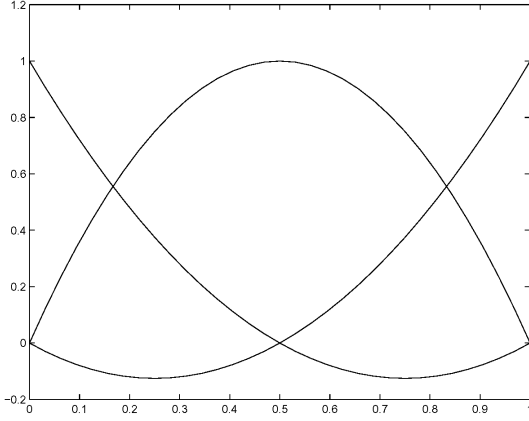


Рис. 2: Лагранжевы квадратичные элементные базисные функции на $[0,1]$.

8.2 Канонический элемент.

Мы можем определить на произвольном элементе элементные базисные функции произвольного порядка методом, подобным примененному для квадратичных функций. Технически, однако, оказывается более удобным определять такие функции на одном и том же, так называемом *каноническом* элементе, и затем пересчитывать эти функции на произвольный элемент с помощью линейного преобразования.

Для определения базисных функций порядка p введем канонический элемент $-1 \leq \xi \leq 1$. Координата произвольного *физического* элемента x связана с координатой на каноническом элементе линейным соотношением

$$x(\xi) = \frac{1 - \xi}{2}x_{j-1} + \frac{1 + \xi}{2}x_j \in K_j.$$

Введем на каноническом элементе $p + 1$ узел и пронумеруем их так, чтобы

$$-1 = \xi_0 < \xi_1 < \dots < \xi_{p-1} < \xi_p = 1.$$

Соответствующие им узлы на физическом элементе K_j задаются выражением

$$x_{j-1+i/p} = \frac{1 - \xi_i}{2}x_{j-1} + \frac{1 + \xi_i}{2}x_j, \quad i = 0, 1, \dots, p.$$

Определим Лагранжевы интерполяционные полиномы $N_{k,c}$, имеющие степень p и отвечающие k -му узлу:

$$N_{k,c}(\xi) = \prod_{i=0, i \neq k}^p \frac{\xi - \xi_i}{\xi_k - \xi_i}.$$

По определению, для них выполнено

$$N_{k,c}(\xi_i) = \delta_{ki}, \quad \deg N_{k,c}(\xi) = p.$$

Мы можем легко построить примеры таких полиномов:

1. Линейных функций, $p = 1$, существует две:

$$N_{0,c}(\xi) = \frac{1-\xi}{2}, \quad N_{1,c}(\xi) = \frac{1+\xi}{2}.$$

2. В квадратичном случае, $p = 2$, присутствуют три функции. Сами эти функции зависят от выбора ξ_1 . При $\xi_1 = 0$ они равны:

$$N_{0,c}(\xi) = \frac{\xi(\xi-1)}{2}, \quad N_{1,c}(\xi) = 1 - \xi^2, \quad N_{2,c}(\xi) = \frac{\xi(\xi+1)}{2}.$$

8.3 Иерархический базис.

В случае построения Лагранжевых конечных элементов, все функции являются полиномами одной и той же степени. Если по каким-то причинам мы захотим увеличить степень используемых полиномов, нам придется пересчитывать все функции и, соответственно, матричные элементы.

Идея иерархического базиса состоит в том, что базис строится рекуррентным образом: базис порядка $p+1$ получается из базиса порядка p добавлением одной новой функции. Таким образом, вычисленные ранее функции и их матричные элементы не пропадают и позволяют уменьшить необходимую вычислительную работу.

Рассмотрим пример квадратичного иерархического базиса на элементе K_j . Он строится добавлением квадратичной функции к линейному базису:

$$U^2(x) = U^1(x) + c_{j-1/2} N_{j-1/2,j}^2(x), \quad (32)$$

где $U^1(x)$ – линейный базис:

$$U^1(x) = c_{j-1} N_{j-1,j}^1(x) + c_j N_{j,j}^1(x),$$

$$N_{j-1,j}^1(x) = \begin{cases} \frac{x_j-x}{h_j}, & \text{для } x \in K_j \\ 0, & \text{для других } x \end{cases},$$

$$N_{j,j}^1(x) = \begin{cases} \frac{x-x_{j-1/2}}{h_j}, & \text{для } x \in K_j \\ 0, & \text{для других } x \end{cases}.$$

Условия для квадратичной добавки $N_{j-1/2,j}^2(x)$ не определяют её однозначно. Мы должны лишь потребовать, чтобы это была квадратичная функция на элементе, непрерывная везде в области и нулевая вне элемента K_j . В качестве такой функции мы можем, например, взять

$$N_{j-1/2,j}^2(x) = \begin{cases} 1 - 4 \left(\frac{x-x_{j-1/2}}{h_j} \right)^2, & \text{для } x \in K_j \\ 0, & \text{для других } x \end{cases}. \quad (33)$$

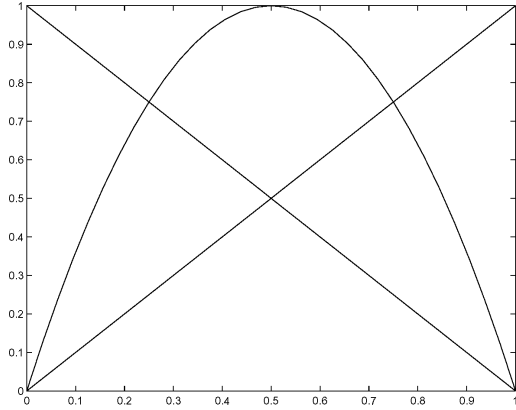


Рис. 3: Иерархические квадратичные базисные функции на $[0,1]$.

Квадратичный иерархический базис изображен на рис. 3.

Используя представление (32), мы можем связать значения аппроксимирующей функции с коэффициентами разложения по базису:

$$c_{j-1} = U^2(x_{j-1}), \quad c_j = U^2(x_j),$$

$$U^2(x_{j-1/2}) = \frac{c_{j-1} + c_j}{2} + c_{j-1/2}.$$

Используя выражение для квадратичной функции (33), можно вычислить коэффициент

$$c_{j-1/2} = -\frac{h_j^2}{8}(U^2)''(x_{j-1/2}).$$

Таким образом, в случае иерархического базиса значения некоторых коэффициентов разложения определяются значением производных функции (а не самой функции) в некоторой точке. Такое наблюдение приводит нас к следующим определениям:

Определение. Лагранжев конечный элемент – конечный элемент, все степени свободы (т.е. коэффициенты разложения решения по базису) которого определяются *значениями* решения в узлах.

Определение. Эрмитов конечный элемент – конечный элемент, у которого хотя бы одна степень свободы определяется *значением производных* решения в узлах.

Необходимо отметить, что существуют и другие элементы, не связывающие степени свободы с узлами. Как правило, такие элементы являются элементами высоких порядков, и рассмотрение решения в неузловых точках позволяет исключить из рассмотрения высокие производные.

Иерархический базис порядка p также удобнее строить на каноническом элементе. Разложение решения дискретной задачи по такому базису выглядит

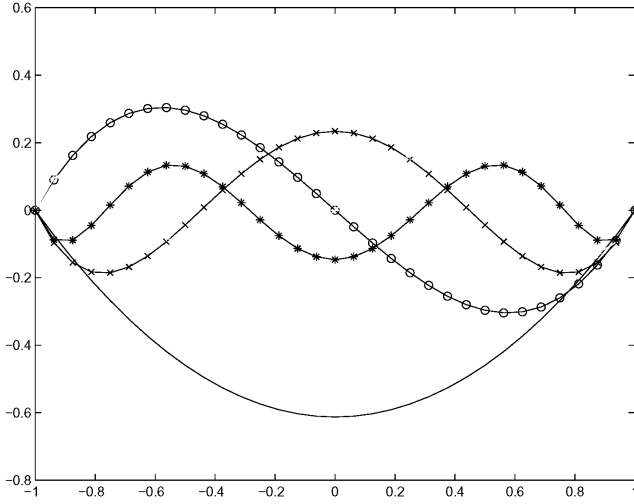


Рис. 4: Иерархические базисные функции порядка 2, 3, 4, 6 на каноническом элементе.

как

$$U(\xi) = c_{-1}N_{-1}^1(\xi) + c_1N_1^1(\xi) + \sum_{i=2}^p c_i N^i(\xi).$$

Линейные функции $N_{-1}^1(\xi)$ и $N_1^1(\xi)$ определены стандартным образом:

$$N_{-1}^1(\xi) = \frac{1-\xi}{2}, \quad N_1^1(\xi) = \frac{1+\xi}{2}.$$

Полиномы $N^i(\xi)$ порядка i определяются требованиями правильного порядка ($\deg N^i(\xi) = i$) и равенству нулю на концах канонического элемента. Пример таких элементных базисных функции приведен на рис. 4.

9 Пример одномерного МКЭ для краевой задачи.

Рассмотрим иерархический базис порядка p на каноническом элементе. Элементная функция на интервале $[-1, 1]$ представляется суммой

$$U(\xi) = c_{-1}N_{-1}^1(\xi) + c_1N_1^1(\xi) + \sum_{i=2}^p c_i N^i(\xi). \quad (34)$$

Значения базисных функций на концах интервала ± 1 равны:

$$\begin{aligned} N_{-1}^1(-1) &= 1, & N_{-1}^1(1) &= 0, \\ N_1^1(-1) &= 0, & N_1^1(1) &= 1, \\ N^i(-1) &= 0, & N^i(1) &= 0, \quad i = 2 \dots p. \end{aligned}$$

Базисные функции могут быть пересчитаны с канонического элемента на произвольный элемент $[x_{j-1}, x_j]$ с помощью линейного преобразования

$$x(\xi) = \frac{1-\xi}{2}x_{j-1} + \frac{1+\xi}{2}x_j. \quad (35)$$

Выберем квадратичный иерархический базис:

$$N_{-1}^1(\xi) = \frac{1-\xi}{2}, \quad N^2(\xi) = \frac{3}{2\sqrt{6}}(\xi^2 - 1), \quad N_1^1(\xi) = \frac{1+\xi}{2}. \quad (36)$$

Очевидно, что квадратичная функция $N^2(\xi)$ удовлетворяет граничным условиям.

Пример 1. Будем решать следующую краевую задачу:

$$\begin{aligned} -pu''(x) + qu(x) &= f(x), \quad p > 0, \quad q > 0, \\ 0 \leq x \leq 1, \quad u(0) &= u(1) = 0. \end{aligned} \quad (37)$$

Функции $p(x)$ и $q(x)$ здесь выбраны константами, с целью упрощения вычислений и необходимости сконцентрироваться на идеях метода, а не на технических деталях. Как мы уже неоднократно поступали, перейдем от краевой задачи (37) к вариационной задаче: найти функцию $u(x)$ такую, что для произвольной функции $v(x)$ выполнено вариационное уравнение $(A(u,v)=(f,v))$:

$$\int_0^1 (u'(x)pv'(x) + u(x)qv(x)) dx = \int_0^1 f(x)v(x) dx. \quad (38)$$

Заменяя интеграл по интервалу $[0, 1]$ суммой интегралов по интервалам $[x_{j-1}, x_j]$ и подставляя вместо функций $u(x)$, $v(x)$ элементные функции (34), пересчитанные из канонического элемента с помощью преобразования (35), получим следующее выражение:

$$\sum_{j=1}^N [A_j^S(U, V) + A_j^M(U, V) - (f, V)_j] = 0 \quad \forall V. \quad (39)$$

В последнем выражении использованы несколько новых обозначений: внутренняя энергия A_j^S

$$A_j^S(U, V) = \int_{x_{j-1}}^{x_j} pU'(x)V'(x) dx = \frac{2p}{h_j} \int_{-1}^1 \frac{dU(\xi)}{d\xi} \frac{dV(\xi)}{d\xi} d\xi;$$

инерциальная и "внешняя" энергии A_j^M :

$$A_j^M(U, V) = \int_{x_{j-1}}^{x_j} qU(x)V(x) dx = \frac{qh_j}{2} \int_{-1}^1 U(\xi)V(\xi) d\xi;$$

а также вектор нагрузки:

$$(f, V)_j = \int_{x_{j-1}}^{x_j} f(x)V(x) dx = \frac{h_j}{2} \int_{-1}^1 f(x(\xi))V(\xi) d\xi.$$

Функции $U(\xi)$ и $V(\xi)$ могут быть двумя разными способами представлены как скалярные произведения:

$$U(\xi) = [c_{j-1}, c_{j-1/2}, c_j] \begin{bmatrix} N_{-1}^1 \\ N^2 \\ N_1^1 \end{bmatrix} = [N_{-1}^1, N^2, N_1^1] \begin{bmatrix} c_{j-1} \\ c_{j-1/2} \\ c_j \end{bmatrix},$$

$$V(\xi) = [d_{j-1}, d_{j-1/2}, d_j] \begin{bmatrix} N_{-1}^1 \\ N^2 \\ N_1^1 \end{bmatrix} = [N_{-1}^1, N^2, N_1^1] \begin{bmatrix} d_{j-1} \\ d_{j-1/2} \\ d_j \end{bmatrix},$$

Используя эти представления, мы можем записать представление для внутренней энергии в более удобном виде:

$$A_j^S(U, V) = [c_{j-1}, c_{j-1/2}, c_j] K_j \begin{bmatrix} d_{j-1} \\ d_{j-1/2} \\ d_j \end{bmatrix},$$

где матрица K_j определяется базисными функциями и может быть легко вычислена:

$$\begin{aligned} K_j &= \frac{2p}{h_j} \int_{-1}^1 \frac{d}{d\xi} \begin{bmatrix} N_{-1}^1 \\ N^2 \\ N_1^1 \end{bmatrix} \frac{d}{d\xi} [N_{-1}^1, N^2, N_1^1] d\xi = \\ &= \frac{2p}{h_j} \int_{-1}^1 \begin{bmatrix} -1/2 \\ \xi\sqrt{3/2} \\ 1/2 \end{bmatrix} [-1/2, \xi\sqrt{3/2}, 1/2] d\xi = \\ &= \frac{2p}{h_j} \int_{-1}^1 \begin{bmatrix} 1/4 & -\xi\sqrt{3/8} & -1/4 \\ -\xi\sqrt{3/8} & 3\xi^2/2 & \xi\sqrt{3/8} \\ -1/4 & \xi\sqrt{3/8} & 1/4 \end{bmatrix} d\xi = \\ &= \frac{p}{h_j} \begin{bmatrix} 1 & 0 & -1 \\ 0 & 2 & 0 \\ -1 & 0 & 1 \end{bmatrix}. \end{aligned}$$

Совершенно аналогично может быть получено и представление для энергии A_j^M :

$$A_j^M(U, V) = [c_{j-1}, c_{j-1/2}, c_j] M_j \begin{bmatrix} d_{j-1} \\ d_{j-1/2} \\ d_j \end{bmatrix},$$

где матрица M_j вычисляется как

$$M_j = \frac{qh_j}{2} \int_{-1}^1 \begin{bmatrix} N_{-1}^1 \\ N^2 \\ N_1^1 \end{bmatrix} [N_{-1}^1, N^2, N_1^1] d\xi = \frac{qh_j}{6} \begin{bmatrix} 2 & -\sqrt{3/2} & 1 \\ -\sqrt{3/2} & 6/5 & -\sqrt{3/2} \\ 1 & -\sqrt{3/2} & 2 \end{bmatrix}.$$

Вектор нагрузки также может быть записан в виде скалярного произведения:

$$(f, V)_j = \mathbf{l}_j [d_{j-1}, d_{j-1/2}, d_j]^\top,$$

где

$$\mathbf{l}_j = \frac{h_j}{2} \int_{-1}^1 f(x(\xi)) [N_{-1}^1, N^2, N_1^1] d\xi.$$

Для вычисления интеграла в последнем выражении мы используем линейную аппроксимацию функции $f(x)$:

$$f(x) \approx N_{-1}^1(\xi)f_{j-1} + N_1^1(\xi)f_j = [f_{j-1}, f_j] \begin{bmatrix} N_{-1}^1 \\ N_1^1 \end{bmatrix}, \quad \text{где } f_j = f(x_j).$$

Следует заметить, что необходимый порядок аппроксимации функции зависит от порядка используемых конечных элементов и должен быть выше для более высоких значений степени p . Таким образом, для вектора \mathbf{l}_j получаем:

$$\mathbf{l}_j = \frac{h_j}{2} \int_{-1}^1 [f_{j-1}, f_j] \begin{bmatrix} N_{-1}^1 \\ N_1^1 \end{bmatrix} [N_{-1}^1, N^2, N_1^1] d\xi = \frac{h_j}{6} \begin{bmatrix} 2f_{j-1} + f_j \\ -\sqrt{3/2}(f_{j-1} + f_j) \\ f_{j-1} + 2f_j \end{bmatrix}^\top.$$

Мы получили выражения для матриц энергии K_j , M_j и вектора нагрузки \mathbf{l}_j на элементе j . Для того, чтобы получить **глобальные** матрицы, мы должны учесть граничные условия в задаче (37). В данном случае мы просто требуем обращения в нуль коэффициентов при функциях, не равных нулю на концах интервала:

$$c_0 = c_N = d_0 = d_N = 0.$$

Поскольку все остальные функции равны нулю при $x = 0$ и $x = 1$ по построению, граничные условия (37) удовлетворены.

Для упрощения записи матриц, в дальнейшем мы будем считать, что все интервалы равны: $h_j \equiv h$. Необходимо отметить, что структура глобальной матрицы зависит от упорядочения. Если мы меняем порядок нумерации элементных функций, матрица тоже будет меняться. У нас выбрана следующая нумерация:

$$\mathbf{c} = [c_{1/2}, c_1, c_{3/2}, \dots, c_{N-1}, c_{N-1/2}]. \quad (40)$$

Коэффициенты c_0 и c_N отсутствуют в силу граничных условий.

Глобальные матрицы образуются комбинированием элементных матриц для троек коэффициентов $[c_j, c_{j+1/2}, c_{j+1}]$. Они перекрываются по элементам с целыми индексами:

$$\mathbf{K} = \frac{p}{h} \begin{bmatrix} 2 & 0 & 0 & & & & \\ 0 & 1+1 & 0 & -1 & & & \\ 0 & 0 & 2 & 0 & & & \\ & -1 & 0 & 1+1 & 0 & -1 & \\ & & 0 & 0 & 2 & 0 & \\ & & & -1 & 0 & 1+1 & \end{bmatrix},$$

$$\mathbf{M} = \frac{qh}{6} \begin{bmatrix} 6/5 & -\sqrt{3/2} & 0 & & & & \\ -\sqrt{3/2} & 2+2 & -\sqrt{3/2} & 1 & & & \\ 0 & -\sqrt{3/2} & 6/5 & -\sqrt{3/2} & & & \\ & 1 & -\sqrt{3/2} & 4 & -\sqrt{3/2} & 1 & \\ & & & -\sqrt{3/2} & 6/5 & -\sqrt{3/2} & \end{bmatrix},$$

$$\mathbf{l} = \frac{h}{6} \begin{bmatrix} -\sqrt{3/2}(f_0 + f_1) \\ f_0 + 4f_1 + f_2 \\ -\sqrt{3/2}(f_1 + f_2) \\ f_1 + 4f_2 + f_3 \end{bmatrix}^\top.$$

Так как для построения нами был выбран иерархический базис, мы можем легко получить и линейную аппроксимацию ($p = 1$) из полученных выражений. Для этого необходимо лишь исключить из матриц строки и столбцы, отвечающие коэффициентам с полуцелыми индексами:

$$\mathbf{K} = \frac{p}{h} \begin{bmatrix} 2 & -1 \\ -1 & 2 & -1 \\ & -1 & 2 & -1 \end{bmatrix}, \quad \mathbf{M} = \frac{qh}{6} \begin{bmatrix} 4 & 1 \\ 1 & 4 & 1 \\ & 1 & 4 & 1 \end{bmatrix},$$

$$\mathbf{l} = \frac{h}{6} \begin{bmatrix} f_0 + 4f_1 + f_2 \\ f_1 + 4f_2 + f_3 \end{bmatrix}^\top.$$

Окончательно, мы можем записать вариационную задачу (38) в виде:

$$\mathbf{c}(\mathbf{K} + \mathbf{M})\mathbf{d}^\top = \mathbf{l}\mathbf{d}^\top.$$

Поскольку это уравнение должно выполняться для любого \mathbf{d} , достаточно решить линейную задачу:

$$(\mathbf{K} + \mathbf{M})\mathbf{c}^\top = \mathbf{l}^\top.$$

Матрица $\mathbf{K} + \mathbf{M}$ обладает следующими свойствами, которые существенно упрощают решение соответствующей линейной системы:

симметричность,

N	$ e _\infty$	$ e _\infty/h^2$	$ e' _\infty$	$ e' _\infty/h$
4	0.269(-3)	0.430(-2)	0.111(0)	0.444
8	0.688(-4)	0.441(-2)	0.589(-1)	0.471
16	0.172(-4)	0.441(-2)	0.303(-1)	0.485
32	0.432(-5)	0.442(-2)	0.154(-1)	0.492
64	0.108(-5)	0.442(-2)	0.775(-2)	0.496
128	0.270(-6)	0.442(-2)	0.389(-2)	0.498

Рис. 5: Линейная аппроксимация, максимальные погрешности.

положительная определенность,
ленточность, ширина ленты $2p + 1$.

Численный пример краевой задачи. Рассмотрим численный пример реализации МКЭ. Выберем определенные значения для коэффициентов p и q в (37):

$$-u''(x) + u(x) = x, \quad 0 \leq x \leq 1, \quad u(0) = u(1) = 0.$$

Его точное решение легко сосчитать:

$$u(x) = x - \frac{\sinh x}{\sinh 1}.$$

Погрешность решения

$$e(x) = u(x) - U(x)$$

удобно характеризовать набором ее различных норм. Мы будем рассматривать следующие нормы:

$$|e|_\infty = \max_{0 \leq j \leq N} |e(x_j)|, \quad |e'|_\infty = \max_{0 \leq j \leq N} |e'(x_j)|,$$

$$\|e\|_0 = \left(\int_0^1 e^2(x) dx \right)^{1/2}, \quad \|e\|_A = \sqrt{A(e, e)} = \left(\int_0^1 [p(e')^2 + qe^2] dx \right)^{1/2},$$

10 Ошибки аппроксимации.

Как мы видели в численном примере предыдущего параграфа, погрешность МКЭ решения ведет себя регулярным образом. Для линейной аппроксимации,

N	$\ e\ _0$	$\ e\ _0/h^2$	$\ e\ _\Delta$	$\ e\ _\Delta/h$
4	0.265(-2)	0.425(-1)	0.390(-1)	0.156
8	0.656(-3)	0.426(-1)	0.195(-1)	0.157
16	0.167(-3)	0.427(-1)	0.979(-2)	0.157
32	0.417(-4)	0.427(-1)	0.490(-2)	0.157
64	0.104(-4)	0.427(-1)	0.245(-2)	0.157
128	0.260(-5)	0.427(-1)	0.122(-2)	0.157

Рис. 6: Линейная аппроксимация, средние погрешности.

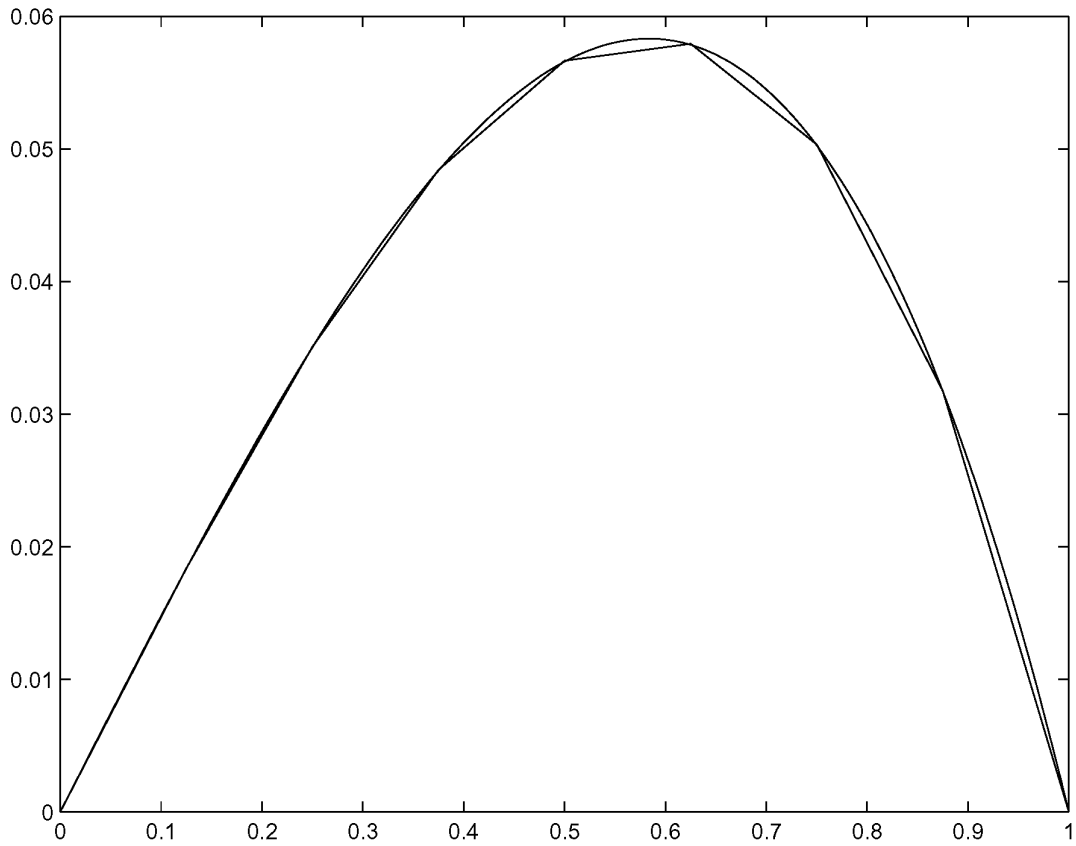


Рис. 7: Точное решение и линейная аппроксимация, 8 элементов.

N	Linear			Quadratic		
	DOF	$\ e\ _0$	$\ e\ _0/h^2$	DOF	$\ e\ _0$	$\ e\ _0/h^3$
4	3	0.265(-2)	0.425(-1)	7	0.126(-3)	0.807(-2)
8	7	0.656(-3)	0.426(-1)	15	0.158(-4)	0.809(-2)
16	15	0.167(-3)	0.427(-1)	31	0.198(-5)	0.809(-2)
32	31	0.417(-4)	0.427(-1)			

Рис. 8: Линейная и квадратичная аппроксимации, средние погрешности.

погрешность решения пропорциональна h^2 , а погрешность производной пропорциональна h . Для квадратичной аппроксимации, погрешности и для решения, и для производной на один порядок по h лучше. В этом параграфе мы докажем эти эмпирические наблюдения, а также сформулируем результаты для произвольного порядка аппроксимации.

Мы начнем с анализа погрешности на каноническом элементе:

$$U(\xi) = \sum_{k=0}^p c_k N_k(\xi).$$

Поскольку решение вариационной задачи единственно, мы можем использовать любой базис для его представления. В данном случае оказывается удобно использовать Лагранжев базис. Мы воспользуемся следующим результатом для оценки погрешности с помощью интерполяционных полиномов Лагранжа:

Теорема 1. Пусть $u(\xi) \in C^{p+1}[-1, 1]$, тогда для любой $\xi \in [-1, 1]$ существует такая $\zeta(\xi) \in (-1, 1)$ что

$$e(\xi) = \frac{u^{(p+1)}(\zeta)}{(p+1)!} \prod_{i=0}^p (\xi - \xi_i).$$

Доказательство этой теоремы может быть найдено, например, в [5]. Дополнительно мы будем использовать тот факт, что $\zeta(\xi)$ гладко зависит от ξ .

10.1 Линейная интерполяция.

Рассмотрим вначале простейший пример линейной интерполяции. В этом случае $\xi_0 = -1$, $\xi_1 = 1$, и ошибка численной аппроксимации может быть записана в соответствии с теоремой как

$$e(\xi) = \frac{u''(\zeta(\xi))}{2} (\xi + 1)(\xi - 1).$$

Мы можем оценить ее сверху как

$$|e(\xi)| \leq \frac{1}{2} \max_{-1 \leq \xi \leq 1} |u''(\xi)| \max_{-1 \leq \xi \leq 1} |\xi^2 - 1| \leq \frac{1}{2} \max_{-1 \leq \xi \leq 1} |u''(\xi)|.$$

Вводя обозначения

$$\|f\|_{\infty, j} := \max_{x \in [x_{j-1}, x_j]} |f(x)|$$

и используя соотношение

$$\frac{d^2 u(\xi)}{d\xi^2} = \frac{h_j^2}{4} \frac{d^2 u(x)}{dx^2},$$

находим оценку для ошибки на интервале

$$\|e\|_{\infty, j} \leq \frac{h_j^2}{8} \|u''\|_{\infty, j}.$$

Переходя к погрешностям на всей области решения нашей задачи,

$$\|e\|_{\infty} = \max_{1 \leq j \leq N} \|e\|_{\infty, j} = \max_{[x_0, x_N]} |e(x)|,$$

$$h = \max_{1 \leq j \leq N} h_j,$$

получаем окончательный результат

$$\|e\|_{\infty} \leq \frac{h^2}{8} \|u''\|_{\infty}. \quad (41)$$

Оценки в среднем.

Найдем теперь оценки для погрешности в среднем, то есть оценка в пространстве L^2 . Вычислим вначале интеграл от квадрата погрешности для каждого физического элемента:

$$\int_{x_{j-1}}^{x_j} e^2(x) dx = \frac{h_j}{2} \int_{-1}^1 \left[\frac{u''(\zeta(\xi))}{2} (\xi^2 - 1) \right]^2 d\xi.$$

Его можно оценить сверху как

$$\|e\|_{0, j}^2 \equiv \int_{x_{j-1}}^{x_j} e^2(x) dx \leq \frac{h_j}{8} \int_{-1}^1 [u''(\zeta(\xi))]^2 d\xi.$$

Функция $\zeta(\xi) \in (-1, 1)$ и монотонна, поэтому

$$\begin{aligned} \|e\|_{0, j}^2 &\leq \frac{h_j}{8} \int_{\zeta(-1)}^{\zeta(1)} [u''(\zeta(\xi))]^2 \frac{d\zeta(\xi)}{|\zeta'(\xi)|} \leq \\ &\leq \frac{h_j}{8} \int_{-1}^1 [u''(\zeta(\xi))]^2 \frac{d\zeta(\xi)}{|\zeta'(\xi)|} \leq C^2 \frac{h_j}{8} \int_{-1}^1 [u''(\zeta)]^2 d\zeta. \end{aligned}$$

Переходя от канонического элемента обратно к физическому, получим

$$\|e\|_{0,j}^2 \leq C^2 \frac{h_j^4}{64} \int_{x_{j-1}}^{x_j} [u''(x)]^2 dx = C^2 \frac{h_j^4}{64} \|u''\|_{0,j}^2.$$

Суммируя по всем интервалам и извлекая корень, получаем окончательный результат:

$$\|e\| \leq \tilde{C} h^2 \|u''\|. \quad (42)$$

Подобным же образом можно получить оценку и в пространстве H^1 , т.е. оценку, включающую производные:

$$\|e\|_1 \leq Ch \|u''\|_0.$$

10.2 Интерполяция полиномами степени p .

Вопрос об ошибке аппроксимации полиномами степени p технически более сложен. Мы приведем здесь основную теорему о такой аппроксимации:

Теорема 2. Фиксируем триангуляцию интервала $[a = x_0 < x_1 < x_2 \dots < x_N = b]$. Пусть $U(x)$ – полином степени не выше p на каждом элементе $[x_{j-1}, x_j]$ и $U(x) \in H^1[a, b]$. Пусть $U(x)$ так интерполирует $u(x) \in H^{p+1}[a, b]$, что интерполяция точна если $u(x)$ – любой полином степени не выше p . Тогда существует такая C_p , что

$$\|u - U\|_0 \leq C_p h^{p+1} \|u^{(p+1)}\|_0$$

и

$$\|u - U\|_1 \leq C_p h^p \|u^{(p+1)}\|_0,$$

где $h = \max_j(h_j)$.

Доказательство этой теоремы в деталях мы обсуждать здесь не будем. Идея состоит в применении оценки Теоремы 1:

$$|e(\xi)| \leq \max_{\zeta \in [a,b]} \left[\frac{|u^{(p+1)}(\zeta)|}{(p+1)!} \left| \prod_{i=0}^p (\xi - \xi_i) \right| \right] \leq \frac{\|u^{(p+1)}\|_\infty}{(p+1)!} \left| \prod_{i=0}^p (\xi - \xi_i) \right|.$$

Техническая сложность заключается в аккуратной оценке произведения

$$\prod_{i=0}^p (\xi - \xi_i).$$

Для одинаковых элементов на $[\xi_{k-1}, \xi_k]$ эта оценка может быть получена достаточно просто [5]:

$$\left| \prod_{i=0}^p (\xi - \xi_i) \right| \leq (p - k + 1)! k! h^{p+1}.$$

Дальнейшее доказательство аналогично линейному случаю. ■

11 Основные этапы использования МКЭ в многомерном случае.

Так же как в одномерном случае, мы можем выделить несколько основных этапов построения МКЭ в многомерном случае, хотя их относительная важность и сложность не обязательно будет совпадать с одномерным МКЭ. Перечислим эти этапы:

1. Вариационная формулировка задачи в области Ω .

Исходное уравнение должно быть сформулировано как вариационное уравнение:

$$A(v, u) = (v, f) + \langle v, \beta \rangle \quad \forall v \in \mathcal{H}. \quad (43)$$

В отличие от одномерного случая, в нем явно присутствуют граничные условия в виде функции β . Учет граничных условий в многомерном МКЭ заметно сложнее, чем в одномерном, в основном из-за этапа триангуляции области.

2. Триангуляция области.

Триангуляция (разделение области Ω на элементы) в одномерном МКЭ тривиально: любой интервал можно точно разделить на любое число элементов. В многомерном случае соответствующая проблема представляет собой сложную геометрическую задачу. Более того, часто эта задача не имеет *точного* решения – например, если мы пытаемся разделить область с криволинейной границей $\partial\Omega$ на треугольники или прямоугольники. Таким образом, здесь мы сталкиваемся с новым типом ошибок – ошибками триангуляции.

3. Генерация элементных матриц жесткости и векторов нагрузки.

Так же как и в одномерном случае, мы заменяем интеграл по области Ω в (43) на сумму интегралов по элементам:

$$\sum_{i=1}^N [A_i(V, U) - (V, f)_i - \langle V, \beta \rangle_i] = 0 \quad \forall V. \quad (44)$$

Необходимо отметить, что в интегралах по границе области интегрирование ведется не по точной границе области $\partial\Omega$, а по ее аппроксимации $\tilde{\partial\Omega}$. Как мы уже отмечали, они не обязательно совпадают.

4. Сборка глобальной матрицы жесткости и вектора нагрузки.

Так же как в одномерном случае, мы должны пронумеровать все используемые базисные функции и составить глобальную матрицу жесткости и вектор нагрузки из интегралов по элементам в (44). В одномерном случае мы имели естественное и эффективное упорядочение – по возрастанию координаты. В многомерном МКЭ такое упорядочение отсутствует, и существуют различные стратегии нумераций глобальных функций.

5. Решение алгебраической системы.

Последний этап – этап решения алгебраической системы – мало отличается от одномерного МКЭ. Поскольку вариационное уравнение

$$\mathbf{d}^T [(\mathbf{K} + \mathbf{M})\mathbf{c} - \mathbf{l}] = 0$$

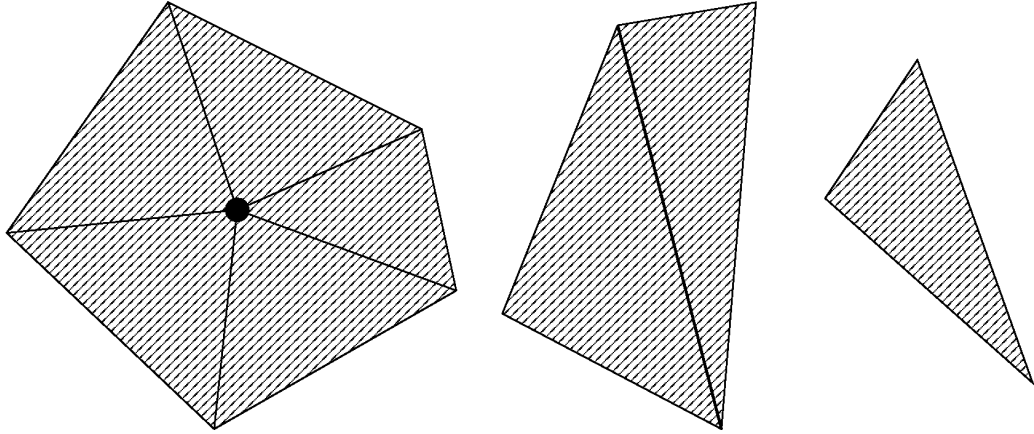


Рис. 9: Области, где функции, связанные с вершиной, ребром и элементом, не равны 0.

должно быть выполнено для любых \mathbf{d} , необходимо решить систему линейных уравнений:

$$(\mathbf{K} + \mathbf{M})\mathbf{c} = \mathbf{l},$$

где \mathbf{K} , \mathbf{M} – матрицы энергии, а \mathbf{l} – вектор нагрузки.

Так же как в одномерном случае, мы начнем анализ многомерного МКЭ с построения глобальных базисных функций. Необходимо отметить, что в отличие от одномерного МКЭ, сейчас базисные функции могут быть связаны с разными типами объектов. Например, в трехмерном пространстве \mathbb{R}^3 базисные функции могут быть связаны с вершинами, ребрами и гранями. Соответственно, элементные функции ϕ_j , связанные с объектом j , не равны нулю только на элементах, содержащих объект j .

12 Лагранжевы элементы в треугольнике.

Мы начнем рассмотрение с двумерного случая. Это наиболее простой случай, который, однако, содержит все характерные трудности многомерного МКЭ.

Построим вначале линейные Лагранжевы элементы. Для этого рассмотрим треугольник с вершинами 1, 2, 3, имеющими координаты (x_j, y_j) , $j = 1, 2, 3$. Искомые функции должны быть линейными и удовлетворять условию:

$$N_j(x_k, y_k) = \delta_{jk}, \quad j, k = 1, 2, 3. \quad (45)$$

Так как любая линейная функция может быть задана в виде

$$N_j(x, y) = a + bx + cy, \quad x, y \in \Omega_e,$$

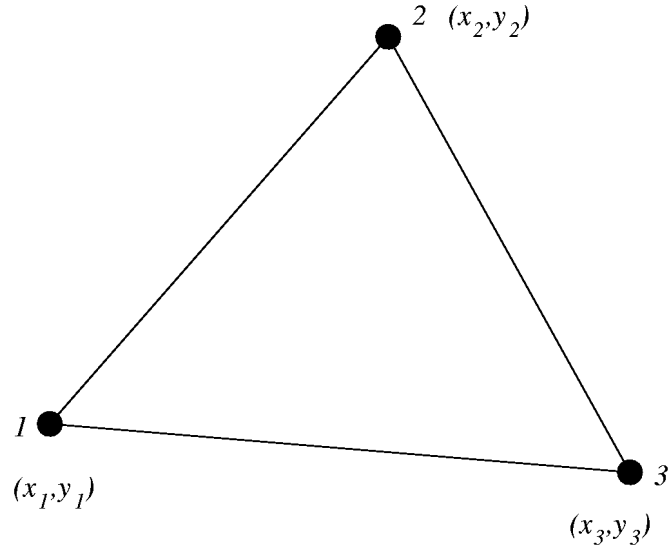


Рис. 10: Треугольный элемент с вершинами 1, 2, 3.

мы должны лишь найти коэффициенты a , b и c . Условие (45) можно записать в виде:

$$\begin{bmatrix} 1 & x_j & y_j \\ 1 & x_k & y_k \\ 1 & x_l & y_l \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad k \neq l \neq j.$$

Решение можно найти, воспользовавшись теоремой Крамера:

$$N_j(x, y) = \frac{D_{kl}(x, y)}{C_{jkl}}, \quad k \neq l \neq j,$$

где

$$D_{kl} = \det \begin{bmatrix} 1 & x & y \\ 1 & x_k & y_k \\ 1 & x_l & y_l \end{bmatrix}, \quad C_{jkl} = \det \begin{bmatrix} 1 & x_j & y_j \\ 1 & x_k & y_k \\ 1 & x_l & y_l \end{bmatrix}.$$

Сужение $U(x, y)$ на элемент e ($x, y \in \Omega_e$) равно

$$U(x, y) = c_1 N_1(x, y) + c_2 N_2(x, y) + c_3 N_3(x, y).$$

Условие (45) определяет коэффициенты c_j :

$$c_j = U(x_j, y_j), \quad j = 1, 2, 3.$$

12.1 Лагранжевы элементы порядка p .

Определим n_p функций $N_j(x, y)$:

$$j = 1, 2, \dots, n_p = \frac{(p+1)(p+2)}{2},$$

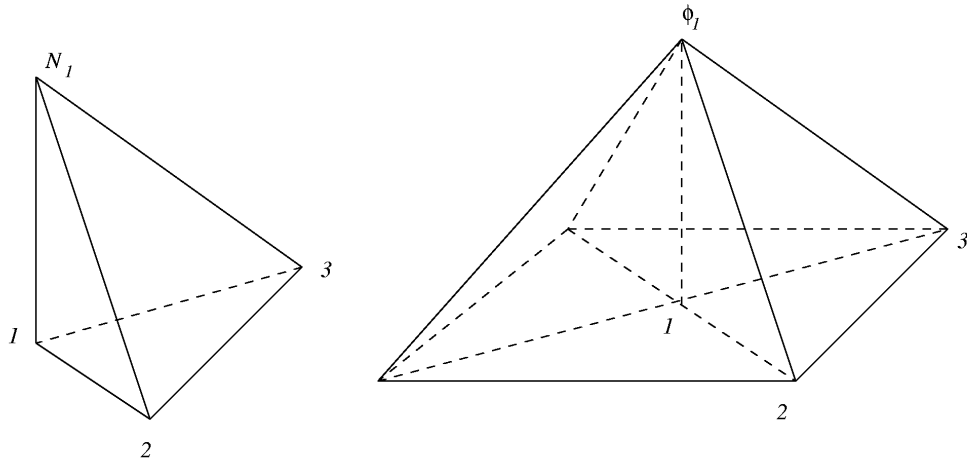


Рис. 11: Элементарная функция N_1 и базисная функция ϕ_1 .

следующим образом

$$N_j(x, y) = \sum_{i=1}^{n_p} a_i q_i(x, y) = \mathbf{a}^\top \mathbf{q}(x, y).$$

Вектор $\mathbf{q}(x, y)$ составлен из последовательно увеличивающихся степеней переменных x, y :

$$\mathbf{q}^\top(x, y) = [1, x, y, x^2, xy, y^2, \dots, y^p].$$

Например, для $p = 2$, $n_2 = 6$ и

$$\mathbf{q}^\top(x, y) = [1, x, y, x^2, xy, y^2].$$

При конструировании конечного элемента необходимо включать в сумму все n_p членов, так как только их полная линейная комбинация гарантирует изотропию функции, т.е. сохранение степени полинома при сдвигах и вращениях.

Для удобства нумерации и конструирования элементарных функций, введем нумерацию узлов, как показано на рис. 12.

Квадратичная функция задается как

$$N_j = a_1 + a_2x + a_3y + a_4x^2 + a_5xy + a_6y^2.$$

Она должна обладать свойством Лагранжевых интерполяционных полиномов:

$$N_j(x_k, y_k) = \delta_{jk}, \quad j, k = 1, 2, \dots, 6.$$

Значение функция на ребре определяется только значениями в точках, лежащих на этом же ребре, что гарантирует непрерывность базисных функций, определенных на соседних элементах.

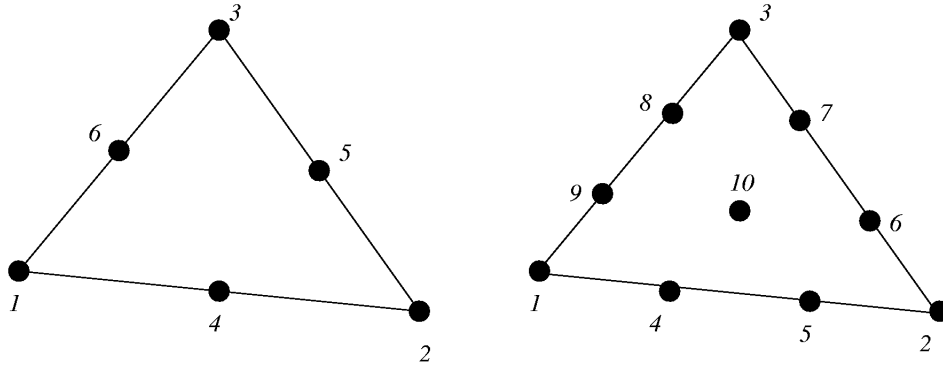


Рис. 12: Расположение узлов для квадратичной и кубической аппроксимаций.

Теперь мы можем определить значения коэффициентов a_i с помощью того же подхода, что и для линейных элементов. Хотя идейно этот подход прост, его техническая реализация весьма затруднительна. Вместо такого подхода, обычно используются другие способы конструирования элементных функций, к обсуждению которых мы сейчас и перейдем.

12.1.1 Канонический элемент

Используя ту же идею, что и в одномерном случае, мы введем в дополнение к физической плоскости с физическими координатами (x, y) – вычислительную плоскость с координатами (ξ, η) . Мы построим отображение – координатное преобразование – связывающее канонический элемент на плоскости (ξ, η) и произвольный физический элемент. В результате, нам будет достаточно конструировать элементные функции только на каноническом элементе, а затем пересчитывать их на физические элементы с помощью координатных преобразований.

Построим уравнения, описывающие отображение произвольного элемента на канонический, см. рис. 13. Уравнение линии, соединяющей вершины (1) и (3), записывается как $N_2(x, y) = 0$. Линия, параллельная $N_2(x, y) = 0$ и проходящая через вершину (2), задается уравнением $N_2(x, y) = 1$. Таким образом, отображение физической линии $N_2(x, y) = 0$ в линию $\xi = 0$ задается соотношением

$$\xi = N_2(x, y).$$

Аналогично, отображение линии, соединяющей вершины (1), (2) записывается как

$$\eta = N_3(x, y).$$

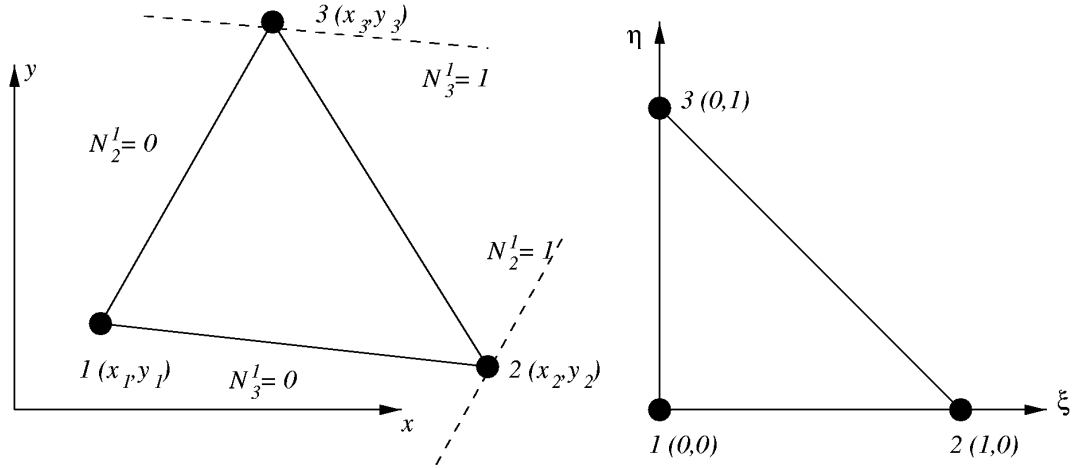


Рис. 13: Отображение произвольного треугольного элемента на канонический элемент.

Используя определение функций $N_i(x, y)$, можем записать

$$\xi = \frac{\det \begin{bmatrix} 1 & x & y \\ 1 & x_1 & y_1 \\ 1 & x_3 & y_3 \end{bmatrix}}{\det \begin{bmatrix} 1 & x_2 & y_2 \\ 1 & x_1 & y_1 \\ 1 & x_3 & y_3 \end{bmatrix}}, \quad \eta = \frac{\det \begin{bmatrix} 1 & x & y \\ 1 & x_1 & y_1 \\ 1 & x_2 & y_2 \end{bmatrix}}{\det \begin{bmatrix} 1 & x_3 & y_3 \\ 1 & x_1 & y_1 \\ 1 & x_2 & y_2 \end{bmatrix}}. \quad (46)$$

Как было отмечено ранее, с помощью преобразования (46) мы можем пересчитывать элементные функции из вычислительных в физические координаты.

12.1.2 Баричесентрические координаты

Идея этого подхода состоит во введении новой системы координат. В качестве этих новых координат мы возьмем значения функций $\{N_1, N_2, N_3\}$. Такие координаты называются *баричесентрическими* (треугольными), см. рис. 14. Необходимо особо отметить, что эта координатная система избыточна – она содержит 3 координаты вместо двух, необходимых на плоскости.

Мы будем использовать следующее обозначение для баричесентрических координат:

$$\zeta_1 = N_1(x, y), \quad \zeta_2 = N_2(x, y), \quad \zeta_3 = N_3(x, y). \quad (47)$$

Обратное преобразование координат, из баричесентрических в физические, задается уравнением

$$\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} \zeta_1 \\ \zeta_2 \\ \zeta_3 \end{bmatrix}. \quad (48)$$

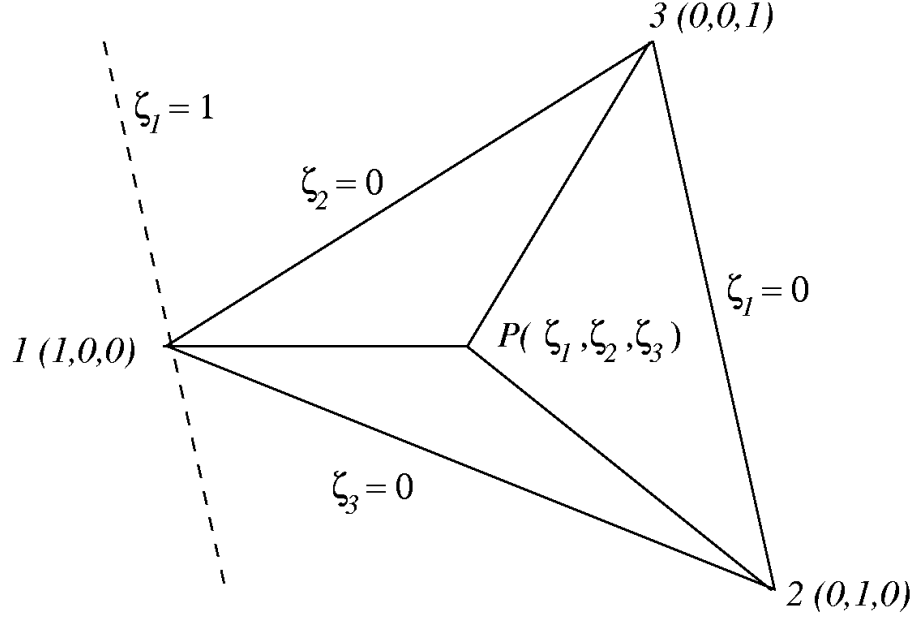


Рис. 14: Барицентрическая координатная система.

Последняя строка в этом матричном соотношении отражает избыточность координат. Соотношение (48) легко проверяется сравнением значений его правой и левой частей в вершинах треугольника

$$\begin{aligned} \text{вершина (1): } & (\zeta_1, \zeta_2, \zeta_3) = (1, 0, 0), \\ \text{вершина (2): } & (\zeta_1, \zeta_2, \zeta_3) = (0, 1, 0), \\ \text{вершина (3): } & (\zeta_1, \zeta_2, \zeta_3) = (0, 0, 1), \end{aligned}$$

и использованием линейности отображения.

Другое распространенное определение барицентрических координат позволяет придать им простой геометрический смысл:

$$\zeta_1 = \frac{A_{P23}}{A_{123}}, \quad \zeta_2 = \frac{A_{P31}}{A_{123}}, \quad \zeta_3 = \frac{A_{P12}}{A_{123}},$$

где A_{ijk} – площадь треугольника с соответствующими вершинами.

Пример Лагранжевых элементов для $p = 3$. Лагранжевы элементы в барицентрических координатах строятся очень просто: мы строим произведения линий, проходящих через те вершины, в которых значение элементной функции должно равняться нулю. Количество таких линий равно степени элементной функции. Нужно также не забыть нормировать значение функции в единственной точке с ненулевым значением на единицу. Приведем несколько примеров функций третьего порядка:

$$N_1^3(x, y) = \frac{9}{2} \zeta_1 (\zeta_1 - 1/3) (\zeta_1 - 2/3) = \frac{9}{2} N_1^1 (N_1^1 - 1/3) (N_1^1 - 2/3),$$

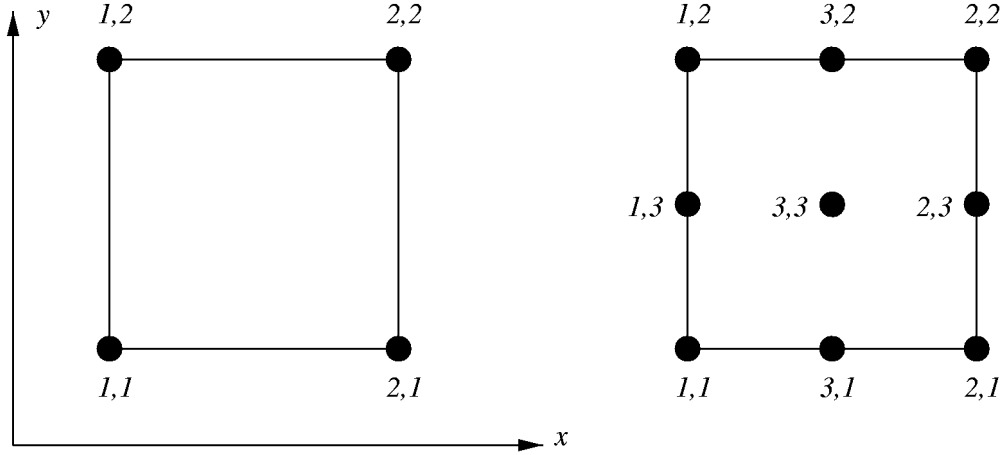


Рис. 15: Канонический квадрат для линейных и квадратичных полиномов.

$$N_4^3(x, y) = \frac{27}{2} \zeta_1 \zeta_2 (\zeta_1 - 1/3) = \frac{27}{2} N_1^1 N_2^1 (N_1^1 - 1/3),$$

$$N_{10}^3(x, y) = 27 \zeta_1 \zeta_2 \zeta_3 = 27 N_1^1 N_2^1 N_3^1.$$

13 Лагранжевы элементы в прямоугольнике.

Треугольник имеет минимум ребер (граней в трехмерном пространстве), так что он оптимален для определения непрерывных базисных функций. Однако, в качестве элементов разбиения могут использоваться и другие геометрические фигуры, обеспечивающие покрытие произвольной области. В качестве такой фигуры часто используется прямоугольник.

Для анализа Лагранжевых элементов в прямоугольнике, мы введем *канонический квадрат* $\{(\xi, \eta) : -1 \leq \xi, \eta \leq 1\}$ в вычислительной плоскости, изображенный на рис. 15. Многомерные Лагранжевы полиномы в прямоугольнике строятся как произведение одномерных.

Билинейный Лагранжев полином представляется в виде:

$$U(\xi, \eta) = c_{1,1} N_{1,1}(\xi, \eta) + c_{2,1} N_{2,1}(\xi, \eta) + c_{2,2} N_{2,2}(\xi, \eta) + c_{1,2} N_{1,2}(\xi, \eta).$$

Для него по-прежнему выполняются соотношения

$$N_{i,j}(\xi_k, \eta_l) = \delta_{ik} \delta_{jl}, \quad U(\xi_k, \eta_l) = c_{k,l}.$$

Однако, теперь $N_{i,j}(\xi, \eta)$ представляется как произведение одномерных функций (как говорят, факторизуется):

$$N_{i,j}(\xi, \eta) = N_i^1(\xi) N_j^1(\eta),$$

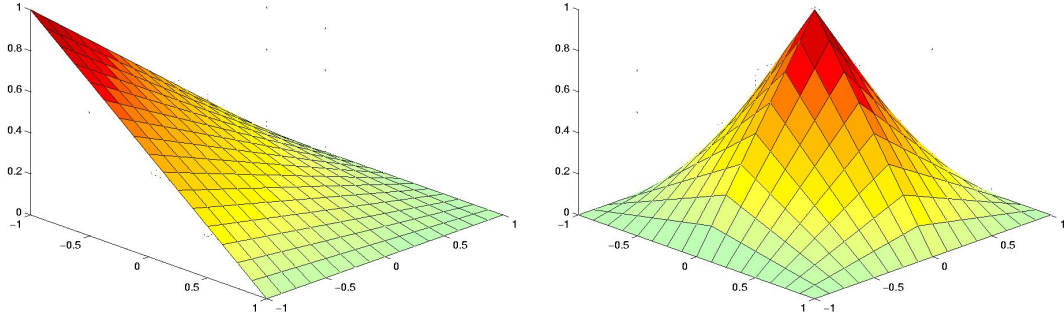


Рис. 16: Билинейная элементная функция $N_{1,1}$ и соответствующая базисная функция.

где

$$N_1^1(\xi) = \frac{1 - \xi}{2}, \quad N_2^1(\xi) = \frac{1 + \xi}{2}, \quad -1 \leq \xi \leq 1.$$

Таким образом, $N_{i,j}(\xi, \eta)$ может быть записана как:

$$N_{i,j}(\xi, \eta) = a_1 + a_2\xi + a_3\eta + a_4\xi\eta.$$

Это представление показывает, что в билинейных функциях на квадрате имеются квадратичные слагаемые. Они, однако, не увеличивают точность аппроксимации, поскольку квадратичная часть является неполной.

Глобальная базисная функция является объединением четырех функций с одной общей вершиной для четырех квадратов. Поскольку каждая функция линейна на сторонах квадрата, глобальная базисная функция непрерывна, смотри рис. 16.

Биквадратичные элементные функции могут быть записаны в виде

$$U(\xi, \eta) = \sum_{i=1}^3 \sum_{j=1}^3 c_{i,j} N_{i,j}(\xi, \eta),$$

где

$$N_{i,j}(\xi, \eta) = N_i^2(\xi) N_j^2(\eta), \quad i, j = 1, 2, 3,$$

$$N_1^2(\xi) = -\xi(1 - \xi)/2, \quad N_2^2(\xi) = \xi(1 + \xi)/2, \quad N_3^2(\xi) = (1 - \xi^2),$$

на интервале $-1 \leq \xi \leq 1$.

Функция $N_{i,j}(\xi, \eta)$ может быть записана как

$$N_{i,j}(\xi, \eta) = a_1 + a_2\xi + a_3\eta + a_4\xi^2 + a_5\xi\eta + a_6\eta^2 +$$

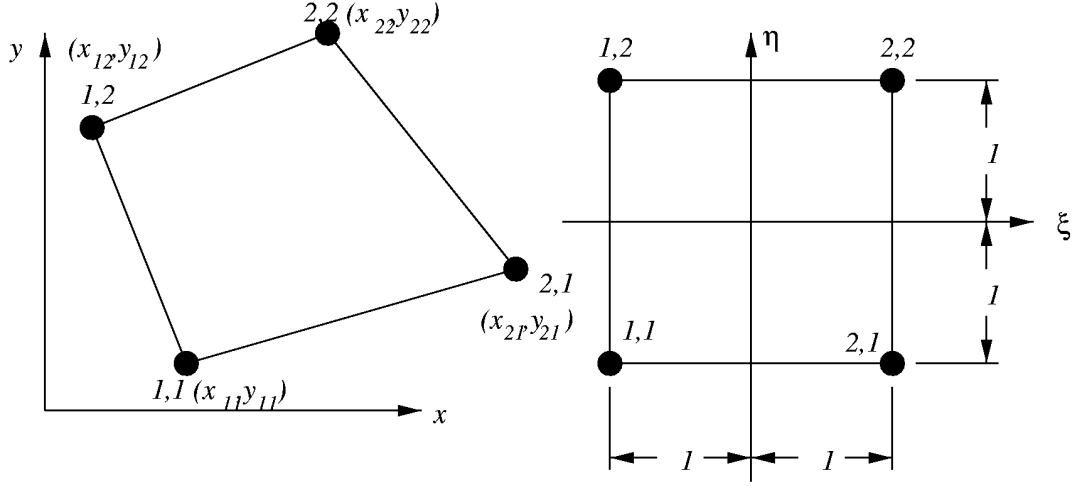


Рис. 17: Отображение канонического квадрата на 4x-угольник.

$$+a_7\xi^2\eta + a_8\xi\eta^2 + a_9\xi^2\eta^2,$$

что явно показывает наличие кубических членов и даже членов четвертого порядка в представлении. Однако, полный полином (т.е. полином, включающий все мономы фиксированной степени) имеет степень 2. Таким образом, аппроксимация является квадратичной, остальные слагаемые практически бесполезны, напротив лишь увеличивают размеры глобальной матрицы жесткости.

Мы построили элементные базисные функции на каноническом квадрате. Для того, чтобы найти эти функции на физическом элементе, мы должны найти отображение канонического квадрата на четырехугольник (x_{ij}, y_{ij}) . Такое отображение может быть построено с помощью функций $N_i^1(\xi)$:

$$\begin{bmatrix} x \\ y \end{bmatrix} = \sum_{i=1}^2 \sum_{j=1}^2 \begin{bmatrix} x_{ij} \\ y_{ij} \end{bmatrix} N_i^1(\xi) N_j^1(\eta).$$

Можно, например, проверить что ребро $\eta = -1$ отображается в ребро $(x_{11}, y_{11}) - (x_{21}, y_{21})$:

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x_{11} \\ y_{11} \end{bmatrix} \frac{1-\xi}{2} + \begin{bmatrix} x_{21} \\ y_{21} \end{bmatrix} \frac{1+\xi}{2}. \quad (49)$$

Важно, что вершины (x_{12}, y_{12}) , (x_{22}, y_{22}) не влияют на отображение (49) не содержащей их стороны, так как это гарантирует непрерывность построенных глобальных базисных функций при переходе через ребро четырехугольника.

14 Двумерные иерархические элементы.

Для изучения методов построения многомерных иерархических элементов мы ограничимся более простым случаем прямоугольника. Мы опишем базис на ка-

ноническом квадрате, а затем эти функции могут быть пересчитаны на произвольный прямоугольник с помощью преобразования, рассмотренного в предыдущем параграфе.

Базис порядка p состоит из нескольких различных типов функций, связанных с вершинами, ребрами и центром квадрата. Опишем их последовательно: *Элементные функции для вершин (билинейные, 4 функции):*

$$N_{ij}^1(\xi, \eta) = N_i^1(\xi)N_j^1(\eta), \quad i, j = 1, 2.$$

Элементные функции для ребер (связаны с серединами ребер, $4(p-1)$ функций для $p \geq 2$):

$$\begin{aligned} N_{31}^k(\xi, \eta) &= N_1^1(\eta)N^k(\xi), & N_{13}^k(\xi, \eta) &= N_1^1(\xi)N^k(\eta), \\ N_{32}^k(\xi, \eta) &= N_2^1(\eta)N^k(\xi), & N_{23}^k(\xi, \eta) &= N_2^1(\xi)N^k(\eta). \end{aligned}$$

Здесь индекс k пробегает значения $k = 2, 3, \dots, p$, а функции $N^k(y)$ определены как

$$N^k(y) = \sqrt{\frac{2k-1}{2}} \int_{-1}^y P_{k-1}(t) dt,$$

где $P_{k-1}(t)$ – полином Лежандра степени $k-1$. Легко убедиться, что $N^k(-1) = N^k(1) = 0$, так что каждая элементная функция для ребер тождественно равна 0 на трех ребрах из четырех.

Элементные внутренние функции (связаны с центром (3,3), $(p-2)(p-3)/2$ функций для $p \geq 4$) могут быть описаны в терминах функции

$$N_{33}^{400} = (1 - \xi^2)(1 - \eta^2).$$

Только эта функция присутствует для случая $p = 4$. Для более высоких степеней p , иерархические функции определяются как

$$\begin{aligned} N_{33}^{510} &= N_{33}^{400} P_1(\xi), & N_{33}^{501} &= N_{33}^{400} P_1(\eta), \\ N_{33}^{620} &= N_{33}^{400} P_2(\xi), & N_{33}^{611} &= N_{33}^{400} P_1(\xi)P_1(\eta), & N_{33}^{602} &= N_{33}^{400} P_2(\eta), \end{aligned}$$

и аналогично для более высоких степеней. Здесь в верхнем индексе $k\lambda\mu$ у функций последовательно стоят общая степень полинома и степени дополнительных полиномов по ξ и η , так что $k = \lambda + \mu + 4$.

Решение в терминах иерархического базиса записывается следующим образом:

$$U(\xi, \eta) = \sum_{i=1}^2 \sum_{j=1}^2 c_{ij}^1 N_{ij}^1 + \sum_{k=2}^p \left[\sum_{j=1}^2 c_{3j}^k N_{3j}^k + \sum_{i=1}^2 c_{i3}^k N_{i3}^k \right] + \sum_{k=4}^p \sum_{\lambda+\mu=k-4} c_{33}^{k\lambda\mu} N_{33}^{k\lambda\mu}.$$

Полное число функций в этом представлении равно:

$$4 + 4(p-1)_+ + \frac{(p-2)_+(p-3)_+}{2}, \quad \text{где } q_+ = \max(q, 0).$$

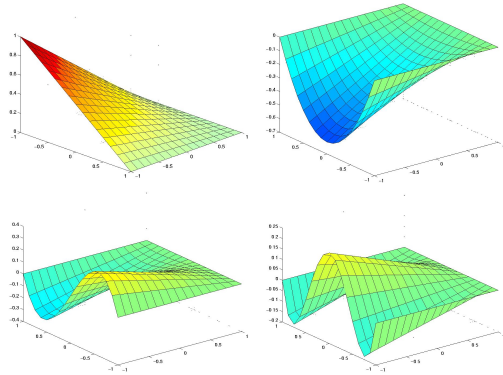


Рис. 18: Иерархические функции для вершины и для ребер ($k = 1, 2, 3, 4$).

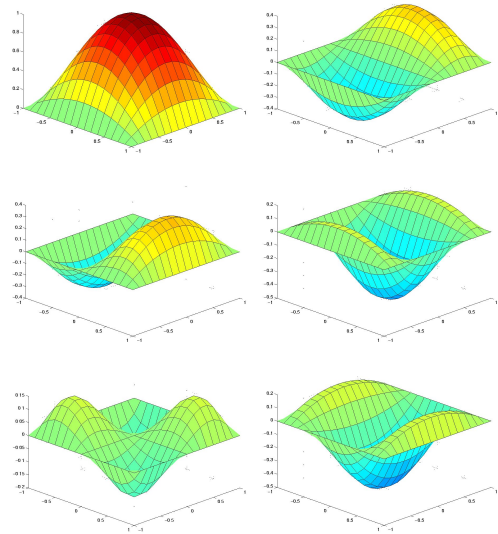


Рис. 19: Иерархические внутренние функции $N_{33}^{k\lambda\mu}$, $\lambda + \mu = k - 4$ ($k = 4, 5, 6$).

Таблица 1: Количество функций в различных КЭ

p	1	2	3	4
треугольник	3	6	10	15
Лагранжев базис	4	9	16	25
иерархический базис	4	8	12	17

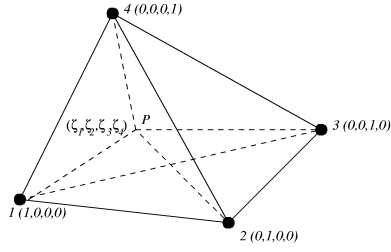


Рис. 20: Узлы для линейной элементной функции на тетраэдре и барицентрические координаты.

Интересно сравнить число функций порядка p для Лагранжева базиса в каноническом квадрате, иерархического базиса и минимальное количество функций, которое совпадает с числом функций в треугольнике.

Мы видим, что оба базиса на квадрате неоптимальны по числу функций. Для малых значений p , число функций в них примерно одинаково и заметно больше оптимального значения. С ростом p , однако, неоптимальность иерархического базиса уменьшается, и асимптотически число функций в нем стремится к оптимальному. Количество функций в Лагранжевом базисе остается слишком большим.

15 Трехмерные элементы.

В трехмерном случае, число полиномов степени p (число степеней свободы) равно:

$$n_p = \frac{(p+1)(p+2)(p+3)}{6}.$$

Построение элементарных функций происходит совершенно аналогично двумерным элементам, поэтому мы приведем лишь формулы для линейных Лагранжевых функций.

В случае тетраэдра, из Лагранжевых условий:

$$N_j(x_k, y_k, z_k) = \delta_{jk}, \quad j, k = 1, 2, 3, 4,$$

мы получаем для четырех линейных функций следующее представление:

$$N_j(x, y, z) = \frac{D_{klm}(x, y, z)}{C_{jklm}},$$

где $(jklm)$ – перестановка чисел $(1,2,3,4)$, а определители D_{klm} и C_{jklm} даются формулами

$$D_{klm} = \det \begin{bmatrix} 1 & x & y & z \\ 1 & x_k & y_k & z_k \\ 1 & x_l & y_l & z_l \\ 1 & x_m & y_m & z_m \end{bmatrix}, \quad C_{jklm} = \det \begin{bmatrix} 1 & x_j & y_j & z_j \\ 1 & x_k & y_k & z_k \\ 1 & x_l & y_l & z_l \\ 1 & x_m & y_m & z_m \end{bmatrix}.$$

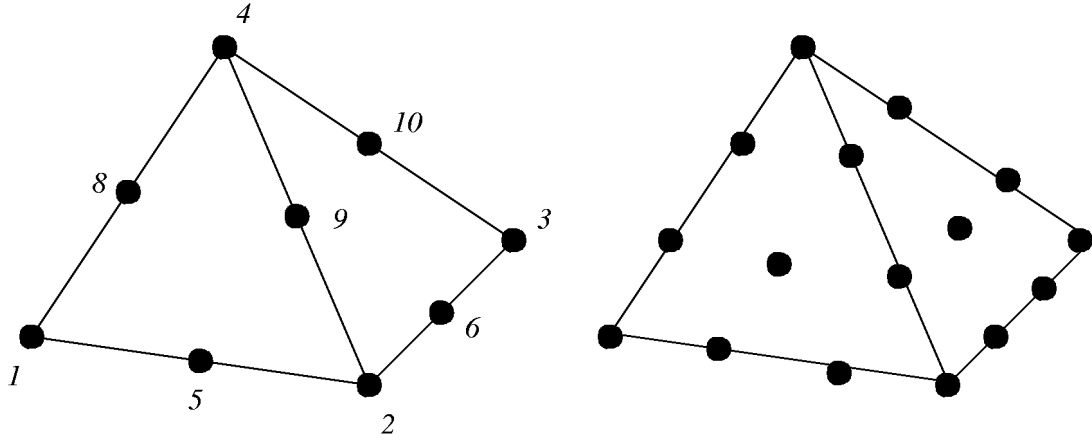


Рис. 21: Узлы для квадратичной и кубической элементной функции на тетраэдре.

Проекция решения U на элемент может быть разложена по этим функциям как:

$$U(x, y, z) = \sum_{j=1}^4 c_j N_j(x, y, z).$$

Барицентрические координаты (координаты объема), которые вновь являются избыточными, определяются формулами:

$$\zeta_j = N_j(x, y, z), \quad j = 1, 2, 3, 4.$$

и также дают объем соответствующих тетраэдров с вершиной в точке P :

$$\zeta_1 = \frac{V_{P234}}{V_{1234}}, \quad \zeta_2 = \frac{V_{1P34}}{V_{1234}}, \quad \zeta_3 = \frac{V_{12P4}}{V_{1234}}, \quad \zeta_4 = \frac{V_{123P}}{V_{1234}}.$$

Обратное преобразование координат задается соотношением:

$$\begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} = \begin{bmatrix} x_1 & x_2 & x_3 & x_4 \\ y_1 & y_2 & y_3 & y_4 \\ z_1 & z_2 & z_3 & z_4 \\ 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} \zeta_1 \\ \zeta_2 \\ \zeta_3 \\ \zeta_4 \end{bmatrix}.$$

Для прямоугольных элементов, вводится канонический куб

$$\{(\xi, \eta, \zeta) : -1 \leq \xi, \eta, \zeta \leq 1\}.$$

Элементные функции для узла (ijk) выбираются в виде произведения функций по каждой из координат:

$$N_{ijk} = N_i(\xi)N_j(\eta)N_k(\zeta).$$

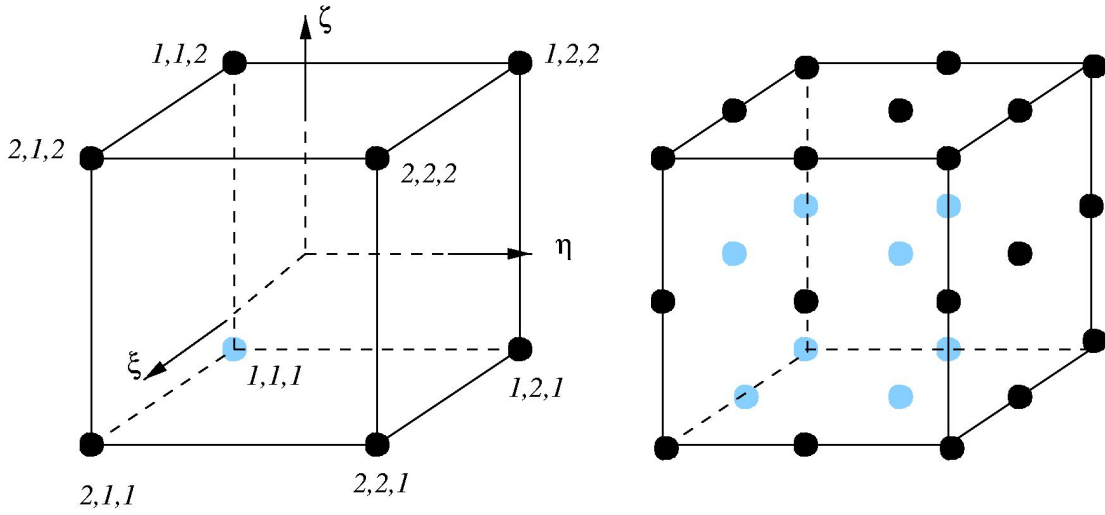


Рис. 22: Узлы для трилинейной и три-квадратичной элементной функции на кубе.

16 Погрешность аппроксимации для многомерных КЭ.

Мы оценим погрешность аппроксимации в два этапа. Вначале будет проведена оценка погрешности аппроксимации полиномами на каноническом элементе, а затем мы произведем пересчет ошибок из канонического на физический элемент.

С целью простоты изложения, рассмотрим двумерный случай. Пусть дискретное решение разлагается по элементным функциям в виде:

$$U(\xi, \eta) = \sum_{j=1}^n c_j N_j(\xi, \eta). \quad (50)$$

Тогда для оценки погрешности аппроксимации верна следующая

Теорема 1. Пусть p – наибольшее целое, для которого представление (50) точно для любого полинома степени p . Тогда для канонического элемента Ω_0 существует $C > 0$ такая, что:

$$|u - U|_{s, \Omega_0} \leq C |u|_{p+1, \Omega_0},$$

для любого $u \in H^{p+1}(\Omega_0)$ и $s = 0, 1, \dots, p+1$. Здесь $|u|_{s, \Omega_0}$ обозначена Соболевская полунорма.

Рассмотрим теперь преобразование канонического элемента в физический на примере треугольных элементов. Преобразование канонического треуголь-

ника задается уравнением:

$$\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} \zeta_1 \\ \zeta_2 \\ \zeta_3 \end{bmatrix} = \begin{bmatrix} x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 - \xi - \eta \\ \xi \\ \eta \end{bmatrix}.$$

Якобиан этого преобразования определен как:

$$J_e = \begin{bmatrix} x_\xi & x_\eta \\ y_\xi & y_\eta \end{bmatrix},$$

и его определитель равен

$$\det J_e = (x_2 - x_1)(y_3 - y_1) - (x_3 - x_1)(y_2 - y_1).$$

Оценку этого определителя можно провести с помощью следующей леммы:

Лемма. Пусть h_e – самая длинная сторона элемента e , а α_e – самый маленький угол. Тогда выполнена оценка:

$$\frac{h_e^2}{2} \sin(\alpha_e) < \det J_e < h_e^2 \sin(\alpha_e).$$

Также можно доказать следующую теорему [4]:

Теорема 2. Пусть $\theta(x, y) \in H^s(\Omega_e)$, $\tilde{\theta}(x, y) \in H^s(\Omega_0)$, где Ω_0 – канонический элемент. Тогда существуют такие константы c_s, C_s , что

$$c_s \sin^{s-1/2} \alpha_e h_e^{s-1} |\theta|_{s, \Omega_e} \leq |\tilde{\theta}|_{s, \Omega_0} \leq C_s \sin^{-1/2} \alpha_e h_e^{s-1} |\theta|_{s, \Omega_e}.$$

Из теорем 1 и 2 следуют основные результаты для оценок погрешности аппроксимации.

Теорема для треугольников. Пусть Ω разбита на треугольные элементы Ω_e . Пусть h – самая длинная сторона элемента в разбиении, а α – самый маленький угол. Для порядка интерполяции p существует такая константа $C > 0$, независимая от $u \in H^{p+1}$ и разбиения, что

$$\|u - U\|_s \leq \frac{Ch^{p+1-s}}{\sin^s \alpha} \|u\|_{p+1}, \quad \forall u \in H^{p+1}(\Omega), \quad s = 0, 1.$$

Теорема для прямоугольников. Пусть Ω разбита на прямоугольные элементы Ω_e . Пусть h – самая длинная сторона элемента в разбиении, а β – самое малое отношение сторон. Для порядка интерполяции p существует такая константа $C > 0$, независимая от $u \in H^{p+1}$ и разбиения, что

$$\|u - U\|_s \leq \frac{Ch^{p+1-s}}{\beta^s} \|u\|_{p+1}, \quad \forall u \in H^{p+1}(\Omega), \quad s = 0, 1.$$

В оценки двух последних сформулированных теорем входят дополнительные геометрические характеристики конечных элементов, и оценки погрешности зависят от этих характеристик. Во многих случаях это нежелательно. Чтобы исключить такую зависимость, мы введем сейчас важный в приложениях класс конечных элементов.

Определение. Семейство конечных элементов называется *регулярным*, если все углы всех элементов отделены от 0 и π (β отделено от 0), когда размер элементов $h \rightarrow 0$.

Для регулярных семейств две вышеприведенные теоремы упрощаются и могут быть сведены в одну.

Теорема. Пусть Ω разбита на элементы Ω_e из регулярного семейства. Пусть h – самая длинная сторона элемента в разбиении. Для порядка интерполяции r существует такая константа $C > 0$, независимая от $u \in H^{p+1}(\Omega)$ и разбиения, что

$$\|u - U\|_s \leq Ch^{p+1-s} \|u\|_{p+1}, \quad s = 0, 1. \quad (51)$$

Хотелось бы отметить, что скорость сходимости, гарантированная уравнением (51), достигается только если решение является достаточно гладким. Если гладкость решения меньше необходимой (имеются скачки производной, разрывна производная определенного порядка и т.д.), оценка погрешности ухудшается. Именно, если решение $u \in H^{q+1}(\Omega)$, $q < p$, то оценка сходимости (51) должна быть заменена на

$$\|u - U\|_s \leq Ch^{q+1-s} \|u\|_{q+1}, \quad s = 0, 1.$$

17 Триангуляция в многомерном МКЭ.

В отличие от одномерного случая, в многомерном МКЭ вопрос о разбиении области на элементы является непростым и требует специального рассмотрения. Эта задача осложняется еще тем, что для получения приемлемой погрешности решения мы должны пользоваться *регулярными семействами* МКЭ, что исключает из разбиения слишком плоские треугольники и прямоугольники. Вопрос о триангуляции области относится к вычислительной геометрии. Время разбиения может составлять заметную часть времени, необходимого для решения задачи с использованием МКЭ. В данном разделе, мы коротко опишем идеи основных методов, применяемых для триангуляции. Более подробная информация может быть найдена в [4, 6].

17.1 Разбиение на блоки, метод Кука

В этом подходе, триангуляция осуществляется в три этапа:

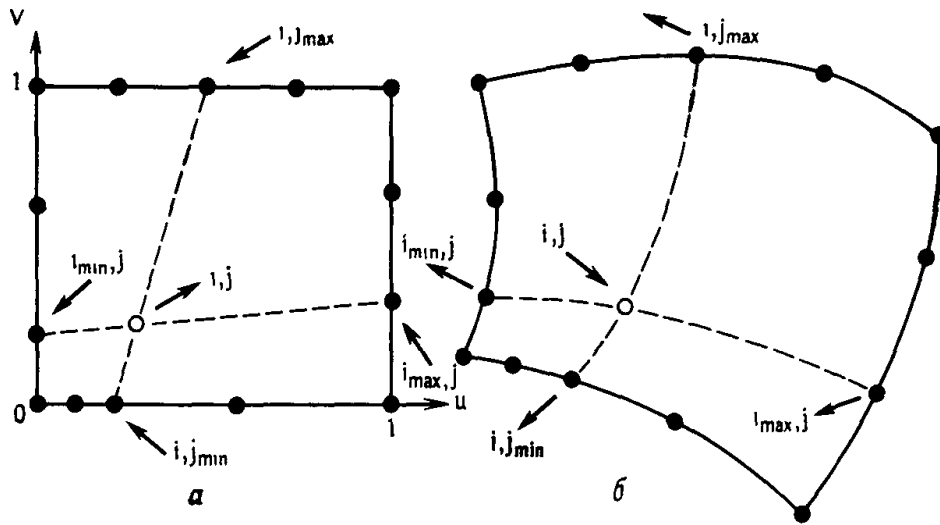


Рис. 23: Разбиение четырехугольника по методу Кука.

- Описание геометрии объекта
- Разбиение объекта на блоки
- Разбиение блоков на конечные элементы.

Основным является второй этап. Здесь мы помещаем некоторое количество точек на границы области, и затем соединяем их прямыми линиями. Затем эти точки могут быть пересчитаны в искривленную область с использованием координатных преобразований, см. рис. 23. Точки пересечения линий вычисляются как

$$u_{ij} = \frac{(u_{i,j_{max}} - u_{i,j_{min}})(v_{i_{min},j} + u_{i,j_{min}})}{1 - (u_{i,j_{max}} - u_{i,j_{min}})(v_{i_{max},j} + v_{i_{min},j})},$$

$$v_{ij} = \frac{(v_{i_{max},j} - v_{i_{min},j})(u_{i,j_{min}} + v_{i_{min},j})}{1 - (u_{i,j_{max}} - u_{i,j_{min}})(v_{i_{max},j} + v_{i_{min},j})}.$$

Получающиеся блоки могут быть выбраны существенно меньше, чем полная исходная область, что облегчает их разбиение на конечные элементы.

17.2 Триангуляция наложением сетки

В этом подходе, мы вначале подготавливаем регулярную сетку, состоящую из треугольников (или из квадратов, тетраэдров и т.д.). Затем мы покрываем изучаемую область такой сеткой. В результате получается хорошая равномерная триангуляция области – везде, кроме окрестности границы. Узлы и границы нашей сетки могут не лежать на границе. Чтобы добиться привязки разбиения к границе, мы можем:

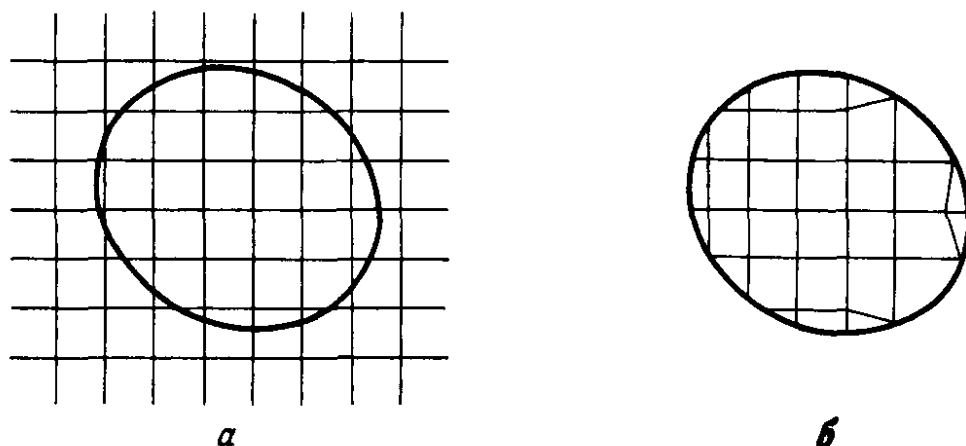


Рис. 24: Триангуляция наложением сетки.

1. переместить наиболее близкие узлы к границе.
2. рассчитать точки пересечения сетки и границы, и добавить эти точки к нашему разбиению.

В обоих случаях, мы получаем более сложные элементы разбиения, которые мы должны затем разделить на простые.

Однако даже с такими улучшениями сетки в окрестности границы, средний диаметр элементов остается одинаковым во всей области, что может приводить к неточному описанию границы области. Для того, чтобы получить более точную аппроксимацию в окрестности границы, используется так называемое квадратичное (или восьмиричное в трехмерном случае) дерево. Идея состоит в делении прямоугольников, имеющих пересечение с границей, на 4(8) меньших прямоугольника, дающих лучшее описание области в окрестности границы, рис. 25. Эта процедура может быть при необходимости повторена несколько раз.

17.3 Фронтальное распространение

Этот подход используется в основном для решения двумерных задач. Элементы строятся послойно, начиная с одной из границ области. Элементы нового слоя сдвигают фронт триангулированной части области, до заполнения всей области, рис. 26. Как правило, в таком подходе используются элементы, близкие к равносторонним треугольникам.

Для того, чтобы применить этот подход к неодносвязным областям, мы должны модифицировать его. Простейший способ модификации состоит во введении искусственных разрезов, рис. 27. В зависимости от связности области, нам может понадобиться один или несколько разрезов.

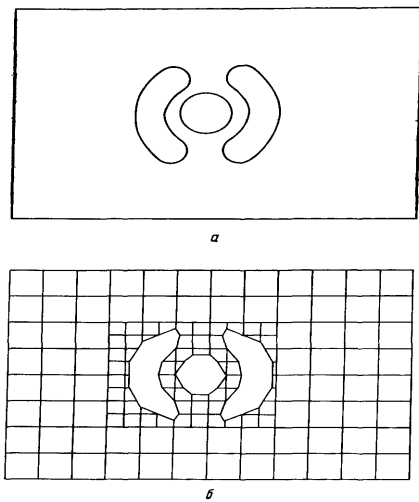


Рис. 25: Использование квадратичного дерева.

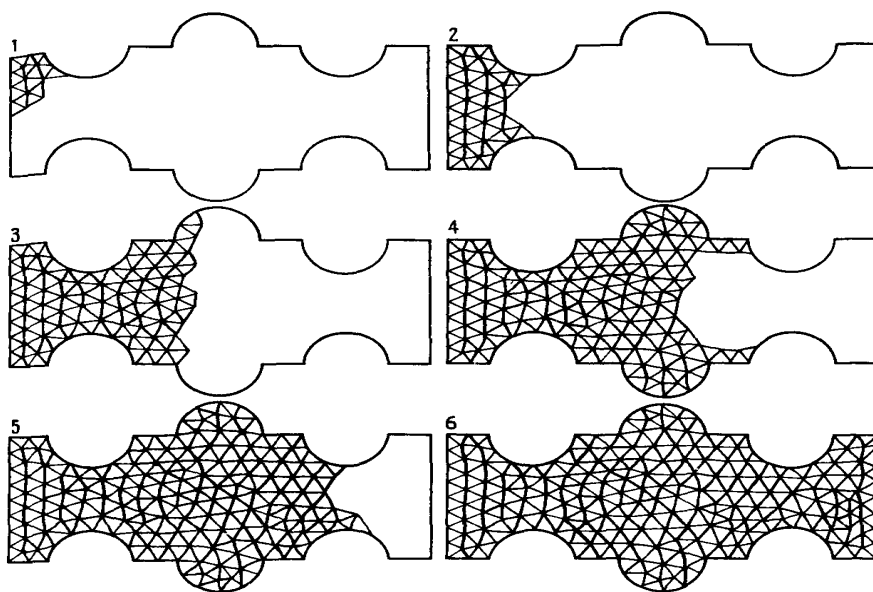


Рис. 26: Пример работы фронтального алгоритма.

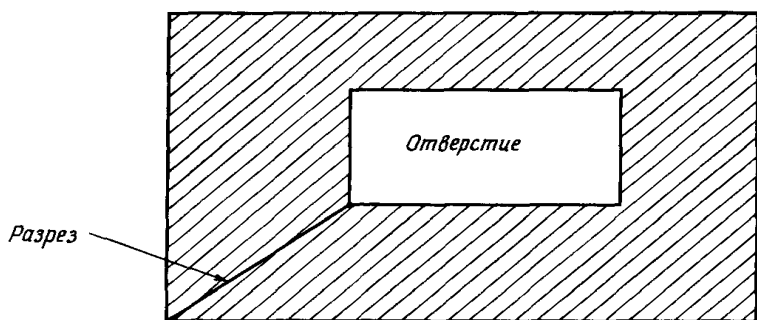


Рис. 27: Неодносвязная область с разрезом.

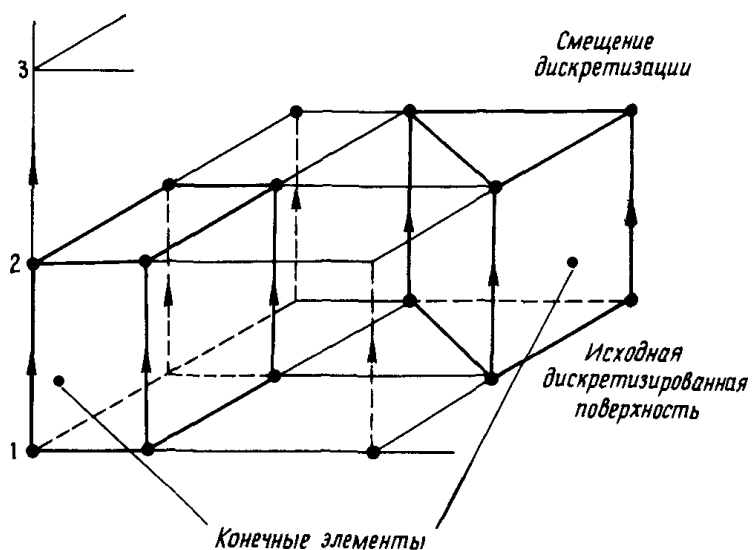


Рис. 28: Перемещение триангулированной поверхности.

17.4 Триангуляция наложением слоев

Этот метод является подходящим для трехмерных областей. Идея заключается в перемещении триангулированной двумерной поверхности вдоль линии в трехмерном пространстве. В связи с этим, такой подход называют еще 2.5-мерным методом.

Мы предполагаем, что триангулированная поверхность "плавно" меняется при перемещении вдоль линии. К новому слою добавляют узлы равномерно перемещенного предыдущего слоя, см. рис. 28. В результате оказывается, что две поверхности разбиты на треугольники, и мы можем простым образом разбить пространство между ними на тетраэдры.

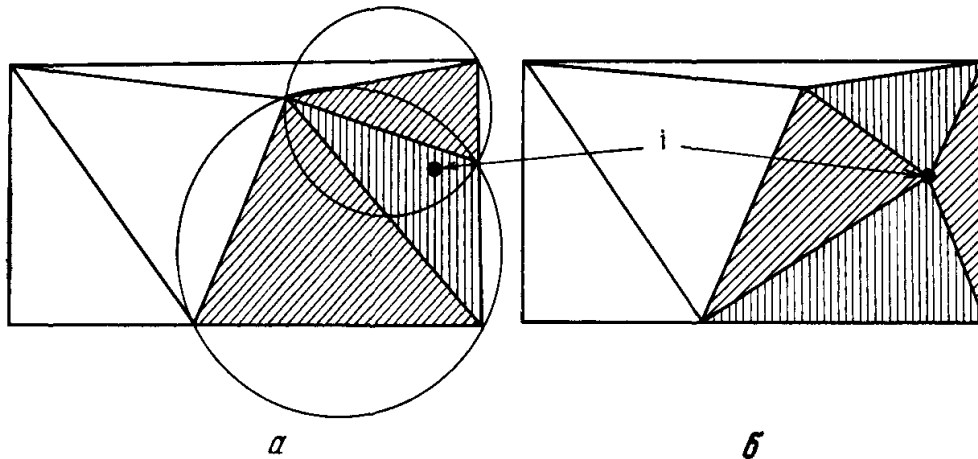


Рис. 29: Итерационная триангуляция по Делоне.

17.5 Глобальное разбиение на треугольники и тетраэдры (метод Делоне)

Этот метод является методом вычислительной геометрии, предназначенным для триангуляции области и/или для улучшения уже имеющейся триангуляции. Метод можно описать как итеративную процедуру, состоящую из следующих этапов:

- Мы стартуем с начальной триангуляцией D_0 всей системы. Она состоит из 2 треугольника в R^2 и 8 тетраэдров в R^3 .
- Каждая последующая триангуляция D_i строится добавлением нового i -го узла и объединением с предыдущей триангуляцией D_{i-1} так что:
 1. элементы, описанная окружность (сфера) которых содержит i , исключаются,
 2. грани (плоскости) соединяются с узлом i ,
 3. удаляются элементы, не пересекающиеся с нашей областью.

Метод Делоне имеет следующие преимущества:

- полученные треугольники и тетраэдры "в среднем" имеют приемлемую форму (т.е. углы не слишком близки к нулю).
- новые узлы могут добавляться процесс триангуляции.

Триангуляция Делоне обычно строится следующим образом:

- Задается первый набор узлов, "разумно" распределенный на границах.
- Строится первая триангуляция на основе этих узлов.

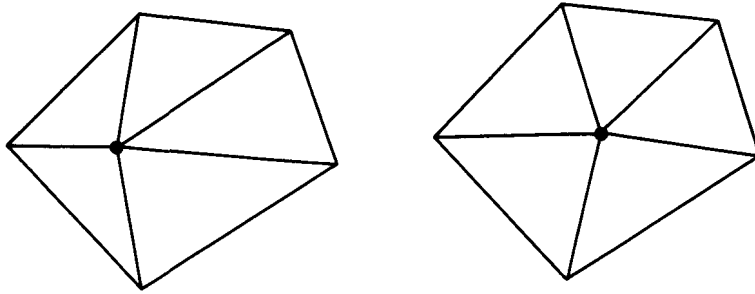


Рис. 30: Исключение плоского элемента.

- Удаляются слишком большие элементы. Для этого добавляется узел в барицентр таких элементов и применяется алгоритм Делоне.
- Предыдущий шаг повторяется до тех пор, пока не будет достигнут необходимый максимальный диаметр элемента.

Хотя метод Делоне дает хорошее разбиение в среднем, отдельные элементы все же могут иметь плохую форму. Для того, чтобы улучшить качество триангуляции, мы должны улучшить форму элементов и оптимизировать их размеры.

Оптимизация размера состоит из исключения больших элементов (с помощью добавления узлов, как описано выше) и исключения малых элементов с помощью их объединения с соседними. Улучшение формы происходит с помощью исключения слишком плоских элементов перемещением внутреннего узла к барицентру некоторого объемлющего элемента, рис. 30. Эта процедура, конечно, может быть реализована по-разному, и ее эффективность и скорость зависит от деталей реализации.

18 Координатные преобразования.

Мы ранее научились строить элементные базисные функции на каноническом элементе. Для решения вариационной задачи, однако, нам необходим базис на физических элементах, которые имеют различные положения, формы, размеры. Следовательно, нам необходимо координатное отображение, которое позволяет пересчитать элементные функции с канонического на физический элемент. Мы можем строить различные преобразования, однако естественно потребовать, чтобы они

- просто вычислялись,
- сохраняли непрерывность,

- были обратимыми, т.е. чтобы Якобиан преобразования не обращался в ноль. Например, для двумерного случая мы требуем

$$\det \begin{bmatrix} x_\xi & x_\eta \\ y_\xi & y_\eta \end{bmatrix} \neq 0.$$

Поскольку элементарные функции являются полиномами, естественно описывать координатные преобразования также кусочно-полиномиальными функциями. В зависимости от соотношения между степенями элементарных функций n_{func} и преобразований n_{trans} , выделяют три класса координатных преобразований:

субпараметрические: $n_{trans} < n_{func}$,

изопараметрические: $n_{trans} = n_{func}$,

суперпараметрические: $n_{trans} > n_{func}$.

В предыдущем изложении мы уже рассматривали линейные и билинейные преобразования. Они всегда удовлетворяют первым двум требованиям, но их обратимость должна быть проверена дополнительно.

Рассмотрим два примера координатных преобразований.

18.1 Билинейное отображение квадрата с квадратичными функциями (Субпараметрическое преобразование)

Решение на элементе записывается в терминах биквадратичных функций:

$$U(\xi, \eta) = \sum_{i=1}^3 \sum_{j=1}^3 c_{ij} N_{ij}^2(\xi, \eta),$$

где

$$N_{ij}^2(\xi, \eta) = N_i^2(\xi) N_j^2(\eta), \quad i, j = 1, 2, 3,$$

$$N_i^2(\xi) = \begin{cases} -\xi(1-\xi)/2, & i = 1, \\ \xi(1+\xi)/2, & i = 2, \\ 1 - \xi^2, & i = 3. \end{cases}$$

Линейное отображение записывается в терминах билинейных функций:

$$\begin{bmatrix} x(\xi, \eta) \\ y(\xi, \eta) \end{bmatrix} = \sum_{i=1}^2 \sum_{j=1}^2 \begin{bmatrix} x_{ij} \\ y_{ij} \end{bmatrix} N_{ij}^1(\xi, \eta),$$

где

$$N_{ij}^1(\xi, \eta) = N_i^1(\xi) N_j^1(\eta), \quad i, j = 1, 2,$$

$$N_i^1(\xi) = \begin{cases} (1-\xi)/2, & i = 1, \\ (1+\xi)/2, & i = 2. \end{cases}$$

Можно проверить, что это преобразование неособое, если все углы физического четырехугольника меньше π , то есть физический четырехугольник – выпуклый.

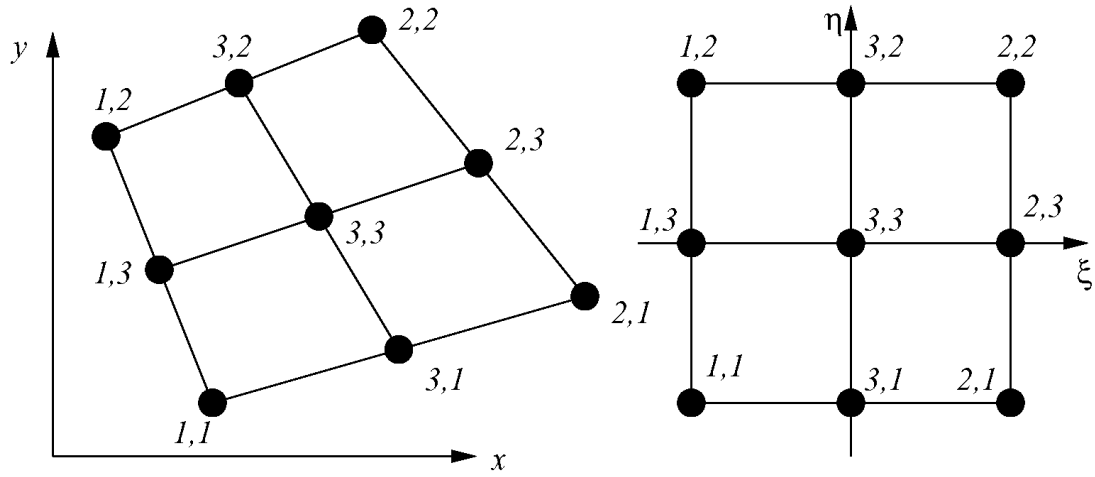


Рис. 31: Билинейное отображение квадрата с квадратичными функциями.

18.2 Квадратичное отображение треугольника с квадратичными функциями

Вобщем случае, квадратичное отображение записывается как

$$\begin{bmatrix} x(\xi, \eta) \\ y(\xi, \eta) \end{bmatrix} = \sum_{i=1}^6 \begin{bmatrix} x_i \\ y_i \end{bmatrix} N_i^2(\xi, \eta).$$

Здесь функции $N_i^2(\xi, \eta)$ задаются следующими выражениями:

$$\begin{aligned} N_i^2 &= 2\zeta_i(\zeta_i - 1/2), \quad i = 1, 2, 3, \\ N_4^2 &= 4\zeta_1\zeta_2, \quad N_5^2 = 4\zeta_2\zeta_3, \quad N_6^2 = 4\zeta_3\zeta_1, \\ \zeta_1 &= 1 - \xi - \eta, \quad \zeta_2 = \xi, \quad \zeta_3 = \eta. \end{aligned}$$

Рассмотрим случай, когда только одна сторона треугольника может быть кривой после преобразования, а две других остаются прямыми линиями. Это означает, что точки 4 и 6 сохраняют свои позиции в центрах соответствующих сторон:

$$\begin{aligned} x_4 &= (x_1 + x_2)/2, & y_4 &= (y_1 + y_2)/2, \\ x_6 &= (x_1 + x_3)/2, & y_6 &= (y_1 + y_3)/2. \end{aligned}$$

Такое отображение может быть записано в виде:

$$\begin{bmatrix} x(\xi, \eta) \\ y(\xi, \eta) \end{bmatrix} = \begin{bmatrix} x_1 \\ y_1 \end{bmatrix} (1 - \xi - \eta) + \begin{bmatrix} x_2 \\ y_2 \end{bmatrix} \xi(1 - 2\eta) + \begin{bmatrix} x_3 \\ y_3 \end{bmatrix} \eta(1 - 2\xi) + \begin{bmatrix} x_5 \\ y_5 \end{bmatrix} 4\xi\eta.$$

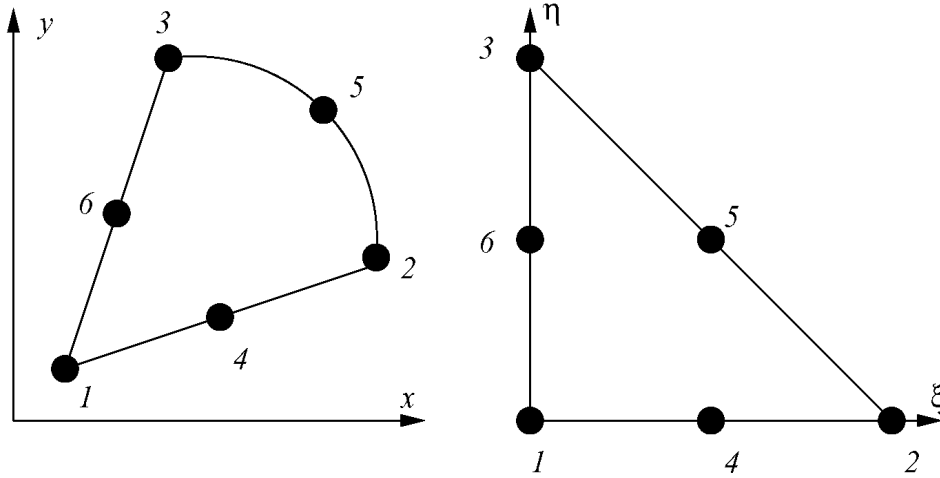


Рис. 32: Квадратичное отображение треугольника с двумя прямыми сторонами.

Легко видеть, что отображения сторон (1-2) и (1-3) – линейные:

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x_1 \\ y_1 \end{bmatrix} (1 - \xi) + \begin{bmatrix} x_2 \\ y_2 \end{bmatrix} \xi,$$

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x_1 \\ y_1 \end{bmatrix} (1 - \eta) + \begin{bmatrix} x_3 \\ y_3 \end{bmatrix} \eta.$$

Якобианы этих преобразований – константы, так что они не могут иметь особенностей. Однако, Якобиан может обнулиться при преобразовании третьей, кривой стороны, и это зависит от значения точки (x_5, y_5) . Разобьем искомое преобразование на 2 шага:

1. Линейное преобразование к каноническому элементу с кривой стороной. Определитель Якобиана постоянный, т.е. преобразование всегда обратимое.
2. "Выпрямление" кривой стороны треугольника.

Подставляя значения $(x_1, y_1) = (0, 0)$, $(x_2, y_2) = (1, 0)$, $(x_3, y_3) = (0, 1)$, в общую формулу, получим преобразование:

$$\begin{bmatrix} x(\xi, \eta) \\ y(\xi, \eta) \end{bmatrix} = \begin{bmatrix} \xi(1 - 2\eta) \\ \eta(1 - 2\xi) \end{bmatrix} + 4\xi\eta \begin{bmatrix} x_5 \\ y_5 \end{bmatrix}.$$

Его Якобиан $\mathbf{J}(\xi, \eta)$ равен:

$$\begin{bmatrix} x_\xi & x_\eta \\ y_\xi & y_\eta \end{bmatrix} = \begin{bmatrix} 1 - 2\eta + 4x_5\eta & -2\xi + 4x_5\xi \\ -2\eta + 4y_5\eta & 1 - 2\xi + 4y_5\xi \end{bmatrix}.$$

Определитель Якобиана также легко вычисляется:

$$\det \mathbf{J}(\xi, \eta) = 1 + (4x_5 - 2)\eta + (4y_5 - 2)\xi.$$

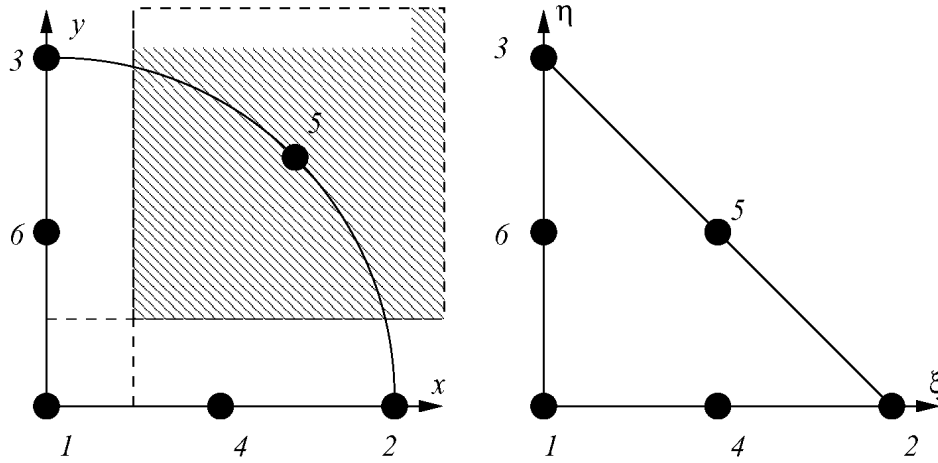


Рис. 33: Линеаризация канонического треугольника.

Так как определитель Якобиана – линейная функция ξ и η , достаточно убедиться, что его значения во всех вершинах канонического элемента имеют один и тот же знак::

$$\det \mathbf{J}(0, 0) = 1, \quad \det \mathbf{J}(1, 0) = 4y_5 - 1, \quad \det \mathbf{J}(0, 1) = 4x_5 - 1.$$

Таким образом, определитель не равен 0 (отображение обратимо) при выполнении условий

$$x_5 > 1/4, \quad y_5 > 1/4. \quad (52)$$

Следовательно, мы можем преобразовать канонический элемент в треугольник с произвольной выпуклой стороной. Вогнутые стороны тоже допустимы, однако максимальная степень вогнутости ограничена и задается условием (52).

19 Численное интегрирование в МКЭ.

Как мы уже видели, для формирования матричных элементов нужно вычислять квадратичные формы на элементных базисных функциях:

$$I = \int_{\Omega_e} A(U_i(\mathbf{x}), U_j(\mathbf{x})) d\mathbf{x}.$$

Рассмотрим вычисление таких интегралов на примере плоскости R^2 . Переходя к каноническому элементу, мы получим для типичного интеграла следующее выражение:

$$I = \int_{\Omega_0} \int \alpha(\xi, \eta) N_s(\xi) N_t^\top(\eta) \det(\mathbf{J}_e) d\xi d\eta,$$

где производная берется по одной из переменных $s, t \in \{\theta, \xi, \eta\}$, а

$$N^T = [N_1, N_2, \dots, N_{n_p}].$$

Таким образом, интеграл представляет собой матрицу, размерность которой равна числу элементарных функций n_p :

$$I = \int \int_{\Omega_0} \begin{bmatrix} (N_1)_s(N_1)_t & (N_1)_s(N_2)_t & & \\ (N_2)_s(N_1)_t & (N_2)_s(N_2)_t & & \\ & & \ddots & \\ & & & \ddots \end{bmatrix} \alpha(\xi, \eta) \det(\mathbf{J}_e) d\xi d\eta. \quad (53)$$

Иногда интеграл (53) может быть вычислен аналитически, однако в общем случае, как правило, приходится прибегать к численному интегрированию.

В рамках МКЭ, можно указать следующие особенности использования численного интегрирования:

- Численное интегрирование дает *точные* результаты в простых случаях (т.е. простые координатные преобразования и простая функция $\alpha(\xi, \eta)$)
- Для достижения оптимального порядка сходимости, точное вычисление интегралов *не обязательно*.

Прежде чем перейти к особенностям использования численного интегрирования в МКЭ, напомним базовые факты о квадратурных формулах. Квадратурной формулой называется приближение интеграла конечной суммой

$$I = \int \int_{\Omega_0} f(\xi, \eta) d\xi d\eta \approx \sum_{i=1}^n W_i f(\xi_i, \eta_i).$$

Говорят, что квадратурная формула имеет порядок точности q , если для любой $q + 1$ раз дифференцируемой функции, $f(\xi, \eta) \in H^{q+1}(\Omega_0)$, выполняется оценка:

$$|I - \sum_{i=1}^n W_i f(\xi_i, \eta_i)| \leq C \|f\|_{q+1}.$$

В одномерном случае,

$$I = \int_{-1}^1 f(\xi) d\xi \approx \sum_{i=1}^n W_i f(\xi_i),$$

наиболее распространенными квадратурными формулами являются формулы Ньютона-Котеса (имеющие порядок точности n) и формулы Гаусса-Лежандра (порядок точности $2n - 1$). Узлы и веса для этих формул низких порядков приведены в Таблице 2. Оценки погрешности и формулы более высоких порядков можно найти в стандартных руководствах по численным методам.

Таблица 2: Узлы и веса для формул Гаусса-Лежандра.

n	$\pm\xi_i$	W_i
1	0	2
2	$1/\sqrt{3}$	1
3	0	8/9
	$\sqrt{3/5}$	5/9

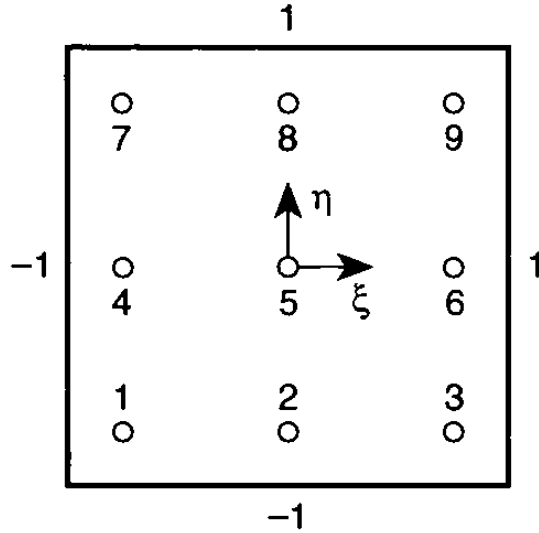


Рис. 34: Узлы интегрирования для $n = 3$ в квадрате (точные для полиномов степени 5).

Рассмотрение многомерного случая мы начнем с более простого случая канонического квадрата. В этом случае, простейшая квадратурная формула представляет собой тензорное произведение одномерных квадратурных формул:

$$\begin{aligned}
 I &= \int_{-1}^1 \int_{-1}^1 f(\xi, \eta) d\xi d\eta \approx \int_{-1}^1 \sum_{i=1}^n W_i f(\xi_i, \eta) d\eta = \\
 &= \sum_{i=1}^n W_i \int_{-1}^1 f(\xi_i, \eta) d\eta \approx \sum_{i=1}^n \sum_{j=1}^m W_i \tilde{W}_j f(\xi_i, \eta_j).
 \end{aligned}$$

Необходимо отметить, что построенные таким образом квадратурные формулы *не оптимальны*. Они содержат больше узлов, чем минимально необходимое количество. Эта разница становится особенно заметной для высоких порядков элементов в R^3 .

Таблица 3: Число узлов одномерных формул n , порядок точности, и количество узлов для тензорных N_{tens} и оптимальных N_{opt} формул.

n	степ.	R^2		R^3	
		N_{tens}	N_{opt}	N_{tens}	N_{opt}
1	1	1	1	1	1
2	3	4	4	8	5
3	5	9	7	27	14
4	7	16	12	64	30

Квадратурные формулы для треугольников:

Для вывода квадратурных формул низких порядков возможно применить метод неопределенных коэффициентов. Например, мы можем аппроксимировать интеграл следующей формулой:

$$\int \int_{\Omega_0} f(\xi, \eta) d\xi d\eta = W_1 f(\xi_1, \eta_1) + E.$$

Здесь имеется три свободных коэффициента, так что формула может быть точна для линейных функций. Проверяем ее для трех линейно независимых линейных функций $[1, \xi, \eta]^T$:

$$\int_0^1 \int_0^{1-\xi} \begin{bmatrix} 1 \\ \xi \\ \eta \end{bmatrix} d\eta d\xi = \begin{bmatrix} 1/2 \\ 1/6 \\ 1/6 \end{bmatrix} = \begin{bmatrix} W_1 \\ W_1 \xi_1 \\ W_1 \eta_1 \end{bmatrix}.$$

Решение этой системы находится, $W_1 = 1/2$, $\xi_1 = \eta_1 = 1/3$, что дает нам следующую квадратурную формулу:

$$\int \int_{\Omega_0} f(\xi, \eta) d\xi d\eta = \frac{1}{2} f(1/3, 1/3) + O(f'').$$

Квадратурные формулы высоких порядков более сложны для вывода. Некоторые примеры таких формул для треугольника и тетраэдра можно найти, например, в [7].

20 Ошибки дискретизации. Ошибки возмущений.

Мы уже обсуждали ошибки аппроксимации при использовании конечных элементов, и установили следующую

Теорему: Пусть Ω разбита на элементы Ω_e из регулярного семейства. Пусть h – самая длинная сторона элемента в разбиении. Для порядка интерполяции p существует $C > 0$ (независимая от $u \in H^{p+1}(\Omega)$ и разбиения):

$$\|u - U\|_s \leq Ch^{p+1-s} \|u\|_{p+1}, \quad s = 0, 1. \quad (54)$$

Однако, оценка ошибки (54) не учитывает все источники погрешностей при реальном применении МКЭ. В частности, оказываются неучтенными:

1. Ошибки численного интегрирования.
2. Ошибки аппроксимации граничных условий.
3. Ошибки триангуляции границы.

Сейчас мы обсудим, как можно оценить соответствующие погрешности.

20.1 Ошибки интегрирования

Исходная вариационная задача:

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

Метод Галеркина с точным интегрированием:

$$A(U, V) = (f, V) \quad \forall V \in H^N$$

Для численного интегрирования:

$$A_*(U^*, V) = (f, V)_* \quad \forall V \in H^N$$

Эффект численного интегрирования в A_* , $(\cdot, \cdot)_*$:

$$(f, V)_* = \sum_{e=1}^{N_\Delta} (f, V)_{e,*} = \sum_{e=1}^{N_\Delta} \sum_{k=1}^n W_k V(x_k, y_k) f(x_k, y_k).$$

Теорема 1: Пусть $A(u, v)$ и $A_*(U, V)$ – билинейные формы, A – непрерывная, A_* – положительно определенная, т.е. $\exists \alpha$ и β :

$$|A(u, v)| \leq \alpha \|u\|_1 \|v\|_1, \quad \forall u, v \in H,$$

$$A_*(U, U) \geq \beta \|U\|_1^2, \quad \forall U \in H^N.$$

Тогда

$$\|u - U^*\|_1 \leq C \left\{ \|u - V\|_1 + \sup_{W \in H^N} \frac{|A(W, V) - A_*(W, V)|}{\|W\|_1} + \sup_{W \in H^N} \frac{|(f, W) - (f, W)_*|}{\|W\|_1} \right\}, \quad \forall V \in H^N.$$

Теорема 2: Пусть $\mathbf{J}(\xi, \eta)$ – Якобиан преобразования из плоскости (ξ, η) в (x, y) . Пусть на регулярном разбиении Δ_h , $\det(\mathbf{J}(\xi, \eta))W_x(\xi, \eta)$ и $\det(\mathbf{J}(\xi, \eta))W_y(\xi, \eta)$ – кусочно-полиномиальные функции степени не более r_1 , и $\det(\mathbf{J}(\xi, \eta))W(\xi, \eta)$ – кусочно-полиномиальная функция степени не более r_0 . Тогда:

1. Если квадратурная формула точна в (ξ, η) для полиномов степени $r_1 + r_0$,

$$\frac{|A(W, V) - A_*(W, V)|}{\|W\|_1} \leq Ch^{r+1} \|V\|_{r+2}, \quad \forall V, W \in H^N.$$

2. Если квадратурная формула точна в (ξ, η) для полиномов степени $r_0 + r - 1$,

$$\frac{|(f, W) - (f, W)_*|}{\|W\|_1} \leq Ch^{r+1} \|f\|_{r+1}, \quad \forall W \in H^N.$$

Пример 1:

Координатное преобразование – линейно,

$\det(\mathbf{J}(\xi, \eta))$ – постоянный.

H^N – кусочно-полиномиальные функции степени p .

Тогда $r_1 = p - 1$, $r_0 = p$.

$$\|u - V\|_1 = O(h^p)$$

Квадратурная формула точна для степени ρ :

$$\rho = r_1 + r \Rightarrow r = \rho - p + 1.$$

$$\frac{|A(W, V) - A_*(W, V)|}{\|W\|_1} \leq Ch^{\rho-p+2} \|V\|_{\rho-p+3}.$$

$$\rho = r_0 + r - 1 \Rightarrow r = \rho - p + 1.$$

$$\frac{|(f, W) - (f, W)_*|}{\|W\|_1} \leq Ch^{\rho-p+2} \|f\|_{\rho-p+2}.$$

Варианты соотношения p и ρ :

1. Если $\rho = 2(p - 1) \Rightarrow r = p - 1$.

Все погрешности одинакового порядка:

$$\|u - U^*\|_1 = O(h^p).$$

Оптимальная ситуация.

2. Если $\rho > 2(p - 1) \Rightarrow r > p - 1$.

Интерполяционная ошибка – основная:

$$\|u - U^*\|_1 = O(h^p).$$

Доп. затраты на интегрирование не оправданы.

3. Если $\rho < 2(p - 1) \Rightarrow r < p - 1$.

Ошибки интегрирования – основные:

$$\|u - U^*\|_1 = O(h^{\rho-p+2}).$$

При $\rho \leq p - 2$ – нет сходимости!

Пример 2:

Изопараметрические элементы:

(степени отображения и элементных функций совпадают.) Оптимальный порядок интегрирования:

$$\rho = 4(p - 1).$$

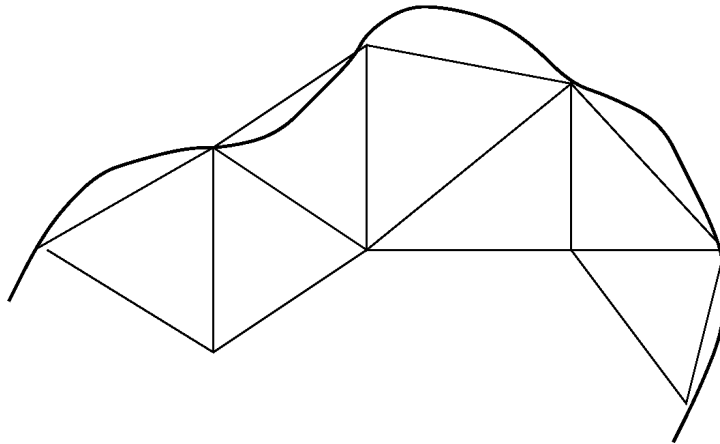


Рис. 35: Аппроксимация границы области полиномами.

20.2 Ошибки аппроксимации граничных условий

Если интегрирование выполняется точно и граница триангулирована точно, то для полиномов степени p

$$\|u - U\|_1 \leq \{h^p \|u\|_{p+1} + h^{p+1/2} \|u\|_{p+1}\}, \quad (55)$$

если решение $u \in H^{p+1}(\Omega)$.

Если Ω не гладкая (например, содержит углы), то решение не принадлежит $H^{p+1}(\Omega)$, и (3) не выполнено.

20.3 Ошибки аппроксимации границы

Пространство аппроксимирующих функций не содержится в исходном пространстве.

Для линейной аппроксимации:

$$\|u - U\|_1 = O(h),$$

Для квадратичной аппроксимации:

$$\|u - U\|_1 = O(h^{3/2}).$$

Ошибки сконцентрированы в окрестности $\partial\Omega$, вне ее – обычные.

Другой результат:

при аппроксимации полиномами степени p :

$$\|u - U\|_1 = O(h^p),$$

если расстояние между границами $\partial\Omega$ и $\partial\tilde{\Omega} \sim h^{p+1}$. Достаточно аппроксимировать границу полиномами степени p .

Т.о., нужно использовать изопараметрические элементы.

21 Априорные и апостериорные оценки погрешностей

Для оценки погрешности различных численных методов (не только МКЭ) широко применяют два типа методов: априорные и апостериорные. В соответствии с названиями, априорный метод позволяет получить информацию о погрешности еще до того, как решение получено, в то время как апостериорный метод требует получения решения (или нескольких решений) до того, как появится возможно оценить погрешность.

В рамках МКЭ, априорные оценки ошибок хорошо известны, и во многих случаях могут быть описаны как

$$|u - U|_s \leq Ch^{p+1-s}|u|_{p+1}, \quad (56)$$

$$s = 0, 1, \quad u \in H^{p+1}(\Omega),$$

где h – максимальный диаметр конечного элемента. В уравнении (56) приближение U к точному решению u ищется в области Ω , и оценки погрешностей приведены для решения и его производных. К достоинству этих оценок относится то, что мы знаем скорость сходимости приближенного решения к точному еще до получения решения U . Недостатком же является то, что мы, как правило, не знаем константу в правой части уравнения и норму точного решения u и, следовательно, не можем оценить погрешность количественно.

Как уже отмечалось, апостериорные оценки ошибок используют уже полученные решения для количественной оценки погрешности. Возможно предложить различные апостериорные оценки, поэтому разумно сформулировать те требования, которым должны удовлетворять эти оценки. Как правило, от них требуют, чтобы они:

- давали аккуратную оценку погрешности для различных триангуляций и степеней полиномов,
- были эффективны в вычислительном смысле,
- могли быть применены для вычислений оценок в различных нормах.

Основные типы апостериорных оценок основаны на:

- экстраполяции решений,
- вычислении невязки решений (solution residual),
- восстановлении решений (solution recovery).

Рассмотрим различные типы апостериорных оценок.

21.1 Оценки, основанные на экстраполяции решений

1. Экстраполяция по пространству

Построим два приближенных решения $U_h^p(x)$ и $U_{h/2}^p(x)$, соответствующих диаметрам КЭ h и $h/2$. Далее мы предполагаем, что оценка сверху погрешности в

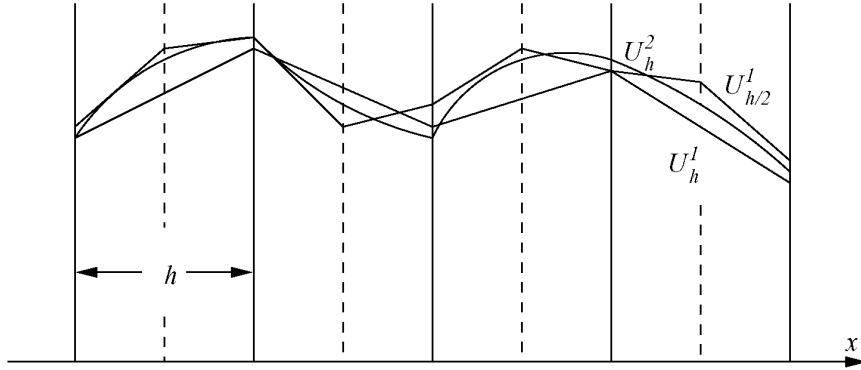


Рис. 36: Решения для линейных функций с элементами U_h^1 и $U_{h/2}^1$, и для квадратичных функций U_h^2 .

правой части уравнения (56) на самом деле является точной:

$$u(x) - U_h^p(x) = C_{p+1}h^{p+1} + O(h^{p+2}), \quad (57)$$

$$u(x) - U_{h/2}^p(x) = C_{p+1} \left(\frac{h}{2}\right)^{p+1} + O(h^{p+2}).$$

Вычитая уравнения друг из друга, найдем:

$$U_{h/2}^p(x) - U_h^p(x) = C_{p+1}h^{p+1} \left(1 - \frac{1}{2^{p+1}}\right) + O(h^{p+2}).$$

Пренебрегая высокими порядками по h , получим оценку погрешности:

$$u(x) - U_h^p(x) \approx \frac{U_{h/2}^p(x) - U_h^p(x)}{1 - 1/2^{p+1}}.$$

Таким образом, мы получили приближенное выражение для погрешности решения $U_h^p(x)$ в терминах его самого и решения с половинным диаметром объемам КЭ. Такой подход к построению оценки называется экстраполяцией по Ричардсону, или h -экстраполяцией.

2. Экстраполяция по степеням полиномов

Добавим к уравнению (57) уравнение для погрешности решения U_h^{p+1} с тем же объемом, но с увеличенной степенью полиномов:

$$u(x) - U_h^{p+1}(x) = C_{p+2}h^{p+2} + O(h^{p+3}).$$

Тогда, прибавляя и вычитая U_h^{p+1} из уравнения (57), найдем

$$u(x) - U_h^p(x) = [u(x) - U_h^{p+1}(x)] + [U_h^{p+1}(x) - U_h^p(x)] = [U_h^{p+1}(x) - U_h^p(x)] + O(h^{p+2}).$$

Таким образом, погрешность может быть вычислена как

$$u(x) - U_h^p(x) \approx [U_h^{p+1}(x) - U_h^p(x)].$$

Этот подход к оценке погрешностей называется p -экстраполяцией.

21.2 Оценки, основанные на невязке решений

В этом подходе, мы строим приближенное решение U , и затем используем его для вычисления локальной погрешности. Общая идея состоит в том, что мы используем граничные условия, полученные из приближенного решения, и решаем наше уравнение с большей точностью на каждом КЭ независимо. В силу этой независимости, вычислительные затраты для получения такого решения невелики. Разность между решением U и полученными локальными решениями (*невязка*) и будет оценкой погрешности. Конечно, более точное решение на элементе можно получить как разбивая элемент на подэлементы, так и увеличивая степень полиномов на элементе. Мы разберем подробнее второй случай.

Мы ищем решение U вариационного уравнения на всей области в пространстве H^p полиномов степени p :

$$A(U, V) = (f, V), \quad \forall V.$$

На каждом элементе для точного решения u выполняется:

$$A_e(u, v) = (f, v)_e, \quad \forall v \in H(\Omega_e),$$

и невязка e определяется как

$$u = U + e.$$

Из этих соотношений мы находим уравнение для невязки:

$$A_e(e, v) = (f, v)_e - A_e(U, v).$$

Для вычисления невязки, мы заменим e и v на их конечномерные аппроксимации E, V в подходящем пространстве:

$$A_e(E, V) = (f, V)_e - A_e(U, V) \quad \forall V \in H^L(\Omega_e). \quad (58)$$

Если пространство H^L содержит полиномы тех же степеней, что и U , то $H^L = H^p$ и $E = 0$. Обычно выбирают $H^L = H^{p+1}$, и тогда можно приближенно положить $E \approx e$. Для того, чтобы получить погрешность во всей области, необходимо решить уравнение (58) для всех элементов. Поскольку эти уравнения для каждого элемента независимы (в отличие от исходного уравнения!), такое решение обычно не составляет проблемы.

Методы, основанные на восстановлении решений, используют свойство сверхсходимости МКЭ, которое мы обсудим подробнее.

22 Сверхсходимость решений в МКЭ.

Для иллюстрации идеи сверхсходимости, рассмотрим уравнение 2го порядка на оси:

$$-u''(x) + \beta u(x) + q = 0.$$

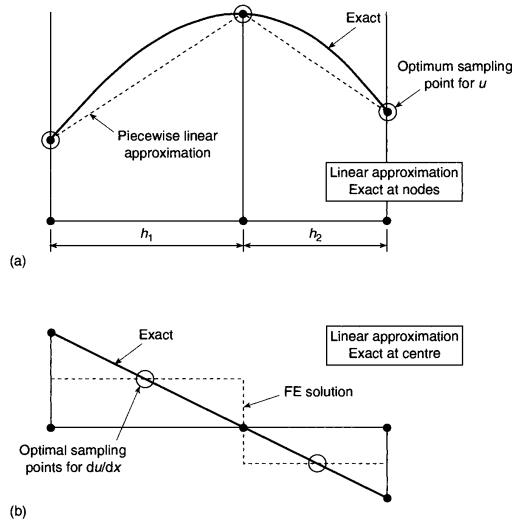


Рис. 37: Оптимальные точки для функции и производной (линейные элементы).

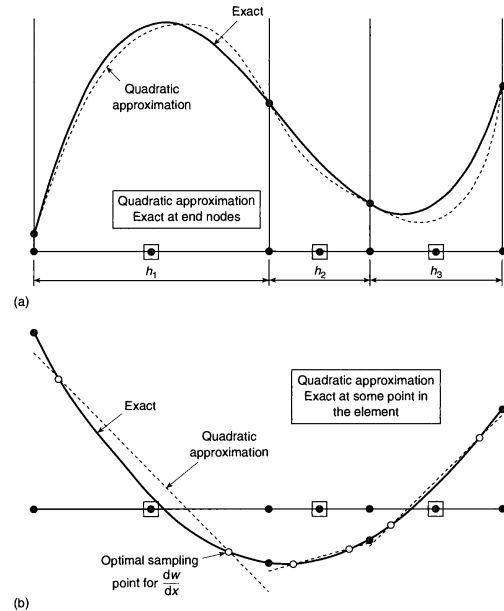


Рис. 38: Оптимальные точки для функции и производной (квадратичные элементы).

Построим его решение с помощью линейного и квадратичного МКЭ и построим графики точных и приближенных решений (приведены на Рис. 37 и 38.)

Можно легко заметить, что существуют точки, в которых погрешность заметно меньше, чем в произвольно выбранной точке интервала. Такие точки существуют как для решения, так и для его производной, причем эти точки разные: решение u точнее всего в узлах, а производные u' точнее всего во внутренних точках. Подобные наблюдения привели к введению понятия *сверхсходимости*, т.е. более быстрой сходимости по сравнению с гарантированной общими теоремами. Здесь нет никакого противоречия, так как общие теоремы гарантируют сходимость для **всех** точек, в то время как сверхсходимость наблюдается только для **некоторых** точек.

Анализируя аппроксимацию полиномами внимательнее, можно доказать следующие факты:

- дополнительный порядок сходимости равен 1.
- В R^1 :
оптимальные точки для функций – узлы.
оптимальные точки для производных – точки Гауссова интегрирования.
- В R^N :
оптимальные точки в кубах – прямое произведение одномерных точек

(есть сверхсходимоссть).

оптимальные точки в тетраэдрах – (сверхсходимости нет!).

С использованием свойства сверхсходимости могут быть построены оценки погрешности, основанные на восстановлении решений.

Рассмотрим двумерную задачу, и будем восстанавливать решение U^* методом наименьших квадратов (МНК):

$$U^* = \mathbf{p}\mathbf{a} = [1, x, y, \dots, y^p]\mathbf{a}, \quad \mathbf{a} = [a_1, a_2, \dots, a_m]^\top.$$

Необходимо определить набор коэффициентов \mathbf{a} , что достигается минимизацией суммы квадратов разностей в точках сверхсходимости (x_k, y_k) :

$$\Pi = \sum_{k=1}^n (U(x_k, y_k) - \mathbf{p}_k\mathbf{a})^2, \quad \mathbf{p}_k = \mathbf{p}(x_k, y_k).$$

Решение \mathbf{a} определяется формулами МНК:

$$\mathbf{a} = A^{-1}\mathbf{b},$$

где

$$A = \mathbf{p}_k^\top \mathbf{p}_k, \quad \mathbf{b} = \sum_{k=1}^n \mathbf{p}_k^\top U(x_k, y_k).$$

Определив коэффициенты \mathbf{a} , мы можем вычислить решение U^* в произвольной точке. Нужно отметить, что для применения МНК число точек n не может быть меньше, чем количество полиномов в \mathbf{p} , и следовательно, всегда будут использоваться значения решения на соседних элементах. Таким образом, данный метод является нелокальным, в отличие от метода невязки.

Пренебрегая, как обычно, разностью между точным решением и лучшим приближением, можем оценить погрешность как:

$$|u - U| \approx |U^* - U|.$$

Нужно также отметить, что свойство сверхсходимости может использоваться и для других целей, например для уточнения решений. Особенно привлекательно уточнение производной решения, так как в этом случае точность восстановленной производной будет совпадать с точностью самого решения.

23 Адаптивное уточнение решений: h -, p -, и hp -уточнение.

Построения решения уравнения с помощью МКЭ на каком-либо предопределенном разбиении, как правило, недостаточно для полного решения задачи. Дело в том, что должна быть проконтролирована точность полученного решения, и нет гарантии, что она сразу же окажется достаточно высокой. Таким образом, обычно решение задачи с помощью МКЭ состоит из следующих этапов:

- выбор триангуляции и элементных функций (т.е. выбор h и p)
- построение решения на выбранном разбиении
- анализ решения и вычисление погрешностей
- изменение триангуляции и/или элементных функций.

Два основных типа уточнения:

1. h -уточнение:

Тип элементов неизменен, меняется размер: уменьшается в одних областях, увеличивается в других.

2. p -уточнение:

Размер элемента неизменен, увеличивается степень полиномов (как правило, иерархические элементы).

Подтипы h -уточнения:

1. Разбиение элементов.

Элементы с большой локальной ошибкой делятся на подэлементы, с малой – объединяются.

Проблемы:

- висящие узлы
- сложная структура объединения.

2. Полная ре-генерация триангуляции.

На основе вычисленной ошибки предсказывается новый размер элемента, происходит новая триангуляция.

Проблемы:

- может быть дорогим в применении.
- сложно переносить результаты предыдущего этапа.

3. r -уточнение.

Число узлов неизменно, меняется их положение.

Может использоваться для описания движущихся особенностей.

Подтипы p -уточнения:

1. Равномерное увеличение порядка во всех элементах.
2. Локальное увеличение порядка с помощью иерархических элементов.

Критерий уточнения.

Критерии будут разными в разных нормах!

Потребуем чтобы относительная ошибка $\eta = \|e\|/\|u\|$:

$$\eta \leq \bar{\eta}.$$

"Оптимальная" триангуляция: ошибки на всех элементах одинаковы. Допустимая ошибка:

$$\bar{\eta}\|u\| \approx \bar{\eta} (\|U\|^2 + \|e\|^2)^{1/2}.$$

Требуем:

$$\|e\|_k < \bar{\eta} \left(\frac{\|U\|^2 + \|e\|^2}{m} \right)^{1/2} = \bar{e}_m,$$

m – число элементов.

$$\xi_k = \|e\|_k / \bar{e}_m.$$

Должны уточнять решение, если

$$\xi_k > 1.$$

(Первый тип)

Генерация новой триангуляции с $\xi_k \leq 1$.

Предположим, что

$$\|e\|_k \sim h_k^p.$$

Тогда

$$h_{new} = \min_k \left(\xi_k^{-1/p} h_k \right).$$

Стратегия hp -уточнения.

1. h -уточнение с низким порядком ($p = 1$ или 2) до разумной точности. (5%).
Ошибки распределены равномерно.
2. Равномерное p -уточнение для всех элементов.
Экстраполяция по p :

$$\|e\| \leq CN^{-\beta}.$$

$$\|u\|^2 - \|U_q\|^2 = C^2 N_q^{-2\beta}, \quad q = p-2, p-1, p.$$

$$\frac{\|u\|^2 - \|U_p\|^2}{\|u\|^2 - \|U_{p-1}\|^2} = \left(\frac{\|u\|^2 - \|U_{p-1}\|^2}{\|u\|^2 - \|U_{p-2}\|^2} \right)^{\frac{\log(N_{p-1}/N_p)}{\log(N_{p-2}/N_{p-1})}}.$$

Теперь знаем ошибку для любого q :

$$\|e\|^2 = \|u\|^2 - \|U_q\|^2, \quad q = 1, 2 \dots p.$$

24 Метод граничных элементов (МГЭ)

Метод конечных элементов успешно зарекомендовал себя для решения задач в произвольных областях с разнообразными граничными условиями и дифференциальными операторами. Однако, матрицы линейных систем, получающиеся в МКЭ, весьма велики и решение таких систем может представлять определенные трудности. В связи с этим, для определенного класса задач был предложен другой метод – метод граничных элементов (МГЭ), который позволяет преодолеть эту трудность.

Идея МГЭ весьма проста и естественна. Исходное дифференциальное уравнение в частных производных в некоторой области переписывается как интегральное уравнение на границе этой области. Решая это интегральное уравнение, мы находим решение на границе, а затем, если необходимо, восстанавливаем решение во всей исходной области. Таким образом, мы получаем сокращение

размерности на единицу, что существенно уменьшает размерность линейной системы.

Сформулировать уравнения МГЭ можно по-разному. Мы рассмотрим две формулировки МГЭ: прямую и непрямую на простейшем примере уравнения Лапласа в области.

24.1 Непрямая формулировка МГЭ

Пусть у нас задано уравнение Лапласа в некоторой области Ω . Граничные условия на границе области Γ могут быть как условиями Дирихле (на части Γ_1), так и условиями Неймана (на части Γ_2), причем $\Gamma = \Gamma_1 + \Gamma_2$. Вводя обозначения $u(x)$ для решения и $q(x) = \partial u(x)/\partial n$ для производной, запишем нашу задачу:

$$\begin{cases} \Delta u(x) = 0, & x \in \Omega \\ u(x) = \bar{u}(x), & x \in \Gamma_1 \\ q(x) = \bar{q}(x), & x \in \Gamma_2. \end{cases} \quad (59)$$

Здесь вектор \vec{n} – единичный вектор внешней нормали к границе области. Известно, что поскольку функция $u(x)$ – гармоническая функция, ей соответствует некоторое распределение потенциала на границе, которым она, в свою очередь, определяется. Рассмотрим отдельно задачи Неймана и Дирихле.

24.1.1 Задача Неймана

Известно, что решение задачи Неймана для уравнения Лапласа может быть записано в виде потенциала простого слоя с неизвестной плотностью $\sigma(x)$:

$$u(x) = \int_{\Gamma} \sigma(\xi) G(\xi, x) d\Gamma(\xi), \quad x \in \Omega. \quad (60)$$

В этом уравнении $G(\xi, x)$ – фундаментальное решение уравнения Лапласа:

$$G(\xi, x) = \begin{cases} 1/|\xi - x| & \text{в } R^3, \\ \log(1/|\xi - x|) & \text{в } R^2. \end{cases} \quad (61)$$

Вычисляя производную представления (60) по внешней нормали, находим:

$$q(x) = -\alpha\pi\sigma(x) + \int_{\Gamma} \sigma(\xi) \frac{\partial G(\xi, x)}{\partial n(x)} d\Gamma(\xi), \quad x \in \Gamma, \quad (62)$$

где $\alpha = 1$ в R^2 и $\alpha = 2$ в R^3 . Левая часть этого уравнения известна, поэтому мы получили уравнение Фредгольма второго рода относительно неизвестной функции $\sigma(x)$. Решив это уравнение, мы затем сможем восстановить решение $u(x)$ во всей области Ω с помощью представления (60).

Необходимо отметить, что решение уравнения (62) существует лишь при выполнении условия Гаусса:

$$\int_{\Gamma} q(x) d\Gamma(x) = 0.$$

24.1.2 Задача Дирихле

Решение задачи Дирихле в области может быть записано в виде потенциала двойного слоя с неизвестной плотностью $\mu(x)$:

$$u(x) = \int_{\Gamma} \mu(\xi) \frac{\partial G(\xi, x)}{\partial n(x)} d\Gamma(\xi), \quad x \in \Omega. \quad (63)$$

Рассматривая это уравнение на границе и используя скачок потенциала двойного слоя, можем записать (63) в виде:

$$u(x) = -\alpha\pi\mu(x) + \int_{\Gamma} \mu(\xi) \frac{\partial G(\xi, x)}{\partial n(x)} d\Gamma(\xi), \quad x \in \Gamma,$$

Это уравнение вновь является уравнением Фредгольма второго рода, может быть решено относительно $\mu(x)$, а решение во всей области затем может быть восстановлено с помощью (63).

Нужно отметить, что решение задачи Дирихле можно также искать и в виде потенциала простого слоя:

$$u(x) = \int_{\Gamma} \sigma(\xi) G(\xi, x) d\Gamma(\xi), \quad x \in \Omega.$$

В силу непрерывности потенциала простого слоя на границе, мы немедленно получаем соотношение

$$u(x) = \int_{\Gamma} \sigma(\xi) G(\xi, x) d\Gamma(\xi), \quad x \in \Gamma.$$

Это соотношение является уравнением Фредгольма первого рода относительно $\sigma(x)$. Как правило, такие уравнения плохо обусловлены, и их решение получить не просто. В данном случае, однако, ядро интегрального оператора сингулярно, и вследствие этого уравнение становится достаточно хорошо обусловленным.

24.2 Прямая формулировка МГЭ

Неудобством не прямой формулировки МГЭ является необходимость введения формальных плотностей поверхностных потенциалов. В некоторых задачах (например, в задачах электростатики) такие потенциалы имеют физический смысл, однако часто эти вспомогательные функции нефизичны. Для таких задач естественнее использовать прямую формулировку МГЭ, в которую входит только само решение и его производные. Для получения соответствующих уравнений можно воспользоваться третьей формулой Грина или методом невязок.

Мы по-прежнему решаем уравнение (59). Применяя третью формулу Грина для произвольной функции $u^*(\xi, x)$, мы немедленно получаем из уравнения Лапласа:

$$\int_{\Omega} \Delta u^*(\xi, x) u(x) d\Omega = - \int_{\Gamma} q(x) u^*(\xi, x) d\Gamma(x) + \int_{\Gamma} u(x) q^*(\xi, x) d\Gamma(x), \quad (64)$$

где

$$q^*(\xi, x) = \frac{\partial u^*(\xi, x)}{\partial n(x)}.$$

Выберем теперь в качестве функции $u^*(\xi, x)$ фундаментальное решение G уравнения Лапласа:

$$\Delta G(\xi, x) = -2\alpha\pi\delta(\xi - x).$$

Подставляя его в уравнение (64), получим:

$$2\alpha\pi u(\xi) + \int_{\Gamma} u(x) \frac{\partial G(\xi, x)}{\partial n(x)} d\Gamma(x) = \int_{\Gamma} q(x) G(\xi, x) d\Gamma(x). \quad (65)$$

Следовательно, теперь у нас в уравнении имеется сумма двух потенциальных слоев, однако их плотности выражены в терминах решения уравнения. В зависимости от граничных условий, это уравнение может быть уравнением Фредгольма первого или второго рода. Мы можем также применять его к задачам с граничными условиями смешанного типа.

Более внимательное изучение уравнения (65) показывает, что требования к гладкости границы в нем могут быть ослаблены. В частности, допускается наличие углов и ребер.

25 Решение нестационарных уравнений параболического типа.

25.1 Слабая формулировка нестационарных уравнений

Рассмотрим модельную начально-граничную задачу (initial-boundary value problem):

$$\begin{aligned} \frac{\partial}{\partial t} u(x, t) - \Delta u(x, t) &= f(x, t) \quad \text{в } \Omega \times (0, T) \\ u(\bullet, 0) &= u_0 \quad \text{на } \bar{\Omega} \\ u(x, t) &= 0 \quad \text{на } \partial\Omega \times (0, T). \end{aligned} \quad (66)$$

В этом уравнении Ω – ограниченная область в R^n , $n \geq 1$, с кусочно-гладкой границей $\partial\Omega$. Задача решается на временном интервале $[0, T]$, $T > 0$.

Умножая уравнение (66) на гладкую финитную функцию $v \in C_0^\infty(\Omega)$ и интегрируя по частям, получаем:

$$\frac{d}{dt} \int_{\Omega} u(x, t) v(x) dx + \int_{\Omega} \nabla u(x, t) \nabla v(x) dx = \int_{\Omega} f(x, t) v(x) dx. \quad (67)$$

В дальнейшем мы будем использовать обозначение $u(x, t) \equiv u(t) \in V = H_0^1(\Omega)$. С его использованием, задача (66) может быть сформулирована в следующем

виде:

Найти функцию $u(t) \in V$ такую, что $u(0) = u_0$, и

$$\frac{d}{dt}(u(t), v) + a(u(t), v) = (f(t), v) \quad \text{для любой } v \in V. \quad (68)$$

В этом уравнении (\bullet, \bullet) – скалярное произведение в $L_2(\Omega)$, а $a(u(t), v) = (\nabla u(t), \nabla v)$. Точная формулировка задачи выходит за рамки наших лекций [1]. Именно, пространство V должно быть некоторым пространством Соболева: $H_0^1(\Omega) \subset V \subset H^1(\Omega)$, форма $a(\bullet, \bullet) : V \times V \rightarrow R$ – V -эллиптической билинейной формой, а $f \in L_2$ в подходящем пространстве. Тогда можно показать [1], что задача (68) имеет единственное решение.

Выбирая $v = u(t)$, мы можем записать уравнение (68) в виде:

$$\frac{1}{2} \frac{d}{dt} \|u(t)\|^2 + a(u(t), u(t)) = (f(t), u(t)), \quad (69)$$

где $\|\bullet\| = \|\bullet\|_{L_2(\Omega)}$. Используя V -эллиптичность (с константой α) формы $a(\bullet, \bullet)$ и неравенство Коши-Буняковского, получим:

$$\frac{1}{2} \frac{d}{dt} \|u(t)\|^2 + \alpha \|u(t)\|^2 \leq \|f(t)\| \|u(t)\|,$$

что для функций с ненулевой нормой можно переписать как

$$\frac{d}{dt} \|u(t)\| + \alpha \|u(t)\| \leq \|f(t)\|.$$

Интегрируя последнее неравенство по времени, получим:

$$\|u(t)\| \leq \|u_0\| e^{-\alpha t} + \int_0^t e^{-\alpha(t-s)} \|f(s)\| ds \quad \text{для } 0 \leq t \leq T. \quad (70)$$

Данная оценка показывает, что задача такого типа имеет *сглаживающее* свойство, т.е. решение становится более регулярным (гладким) с увеличением времени t . Однако, в окрестности $t = 0$ могут возникать проблемы со сходимостью или гладкостью. Примером результатов такого типа может быть следующее утверждение:

Лемма. Рассмотрим задачу

$$\begin{aligned} \frac{\partial}{\partial t} u(x, t) - \Delta u(x, t) &= f(x, t) \quad \text{в } \Omega \times (0, T) \\ u(\bullet, 0) &= u_0 \quad \text{на } \bar{\Omega} \\ u(x, t) &= 0 \quad \text{на } \partial\Omega \times (0, T). \end{aligned}$$

Если $u_0 \in L_2(\Omega)$, тогда для $t \geq \delta > 0$ и любого натурального числа k , $H^k(\Omega)$ норма функции $u(t)$ ограничена. Более точно,

$$\|u(t)\|_{H^k(\Omega)} \leq Ct^{-\frac{1}{2}k} \|u_0\| \quad \text{для } t > 0.$$

25.2 Полудискретизация конечными элементами

При численном исследовании нестационарных уравнений, возможна конечно-элементная дискретизация одновременно временной и пространственных координат. Однако, есть определенные преимущества у схем, где только пространство (или только время) дискретизовано.

Мы обсудим здесь метод, в котором дискретизируются только пространственные координаты. Такой метод называется (вертикальным) методом прямых (method of lines, MOL). Мы рассматриваем начально-граничную задачу:

$$\begin{aligned} \frac{\partial}{\partial t}u + Lu &= f \quad \text{в} \quad \Omega \times (0, T) \\ u(\bullet, 0) &= u_0 \quad \text{при} \quad t = 0 \text{ и } x \in \Omega \\ u(x, t) &= 0 \quad \text{на} \quad \partial\Omega \times (0, T). \end{aligned} \quad (71)$$

Здесь L – равномерно-эллиптический оператор в Ω . Область $\Omega \subset R^n$ и имеет гладкую границу. Когда дифференциальное уравнение в задаче (71) дискретизируется любым методом (например, конечно-разностным, или методом конечных элементов), мы получаем систему обыкновенных дифференциальных уравнений.

Пример. Рассмотрим следующее уравнение в области $\Omega = [0, 1]$.

$$\begin{aligned} u_t - u_{xx} &= f(x, t) \\ u|_{x=0} &= u|_{x=1} = 0 \\ u|_{t=0} &= u_0. \end{aligned} \quad (72)$$

Разобьем интервал $[0, 1]$ на одинаковые подинтервалы точками x_i , $i = 0 \dots N$, так, что $x_0 = 0$, $x_N = 1$, $h = x_i - x_{i-1}$. Введем следующие обозначения: $u_i(t) = u(x_i, t)$, $f_i(t) = f(x_i, t)$.

Для стандартной конечно-разностной аппроксимации, получается следующая система уравнений:

$$\begin{aligned} \frac{du_i}{dt} &= \frac{u_{i-1} - 2u_i + u_{i+1}}{h^2} + f_i(t), \quad i = 1 \dots N-1, \\ u_0 &= u_N = 0, \quad u_i(0) = u_0(x_i) \quad i = 0 \dots N. \end{aligned} \quad (73)$$

В случае конечно-элементной аппроксимации, мы должны использовать слабую формулировку задачи:

$$\frac{d}{dt}(u(t), v) + a(u(t), v) = (f(t), v) \quad \text{для любой } v \in V.$$

Выбор пространства V , а также начальных и граничных условий обсуждались выше. Введем пространство V_h конформных конечных элементов $V_h \subset V$. Тогда

дискретный аналог слабой формулировки задачи можно записать так: найти функцию $u_h(t) \in V_h$ такую, что

$$\frac{d}{dt}(u_h(t), v_h) + a(u_h(t), v_h) = (f(t), v_h) \quad \text{для любой } v_h \in V_h, \quad (74)$$

и $u_h(0) = u_h^0 \in V_h$, где u_h^0 – аппроксимация начальных условий u_0 .

Воспользуемся методом Галеркина. Пусть $\{\varphi_1, \dots, \varphi_M\}$ – базис в пространстве V_h . Тогда численное решение раскладывается по этому базису:

$$u_h(x, t) = \sum_{i=1}^M u_i(t) \varphi_i(x). \quad (75)$$

Подставляя разложение (75) в уравнение (74), мы получим следующую систему обыкновенных дифференциальных уравнений:

$$\sum_{i=1}^M u_i'(t) (\varphi_i, \varphi_j) + \sum_{i=1}^M u_i(t) a(\varphi_i, \varphi_j) = (f(t), \varphi_j), \quad j = 1 \dots M. \quad (76)$$

Введем следующие матрицы и вектора:

$$\begin{aligned} D &= \{d_{ij}\}, & d_{ij} &= (\varphi_i, \varphi_j), \\ A &= \{a_{ij}\}, & a_{ij} &= a(\varphi_i, \varphi_j), \\ \hat{f}(t) &= \{f_j\}, & f_j &= (f, \varphi_j), \\ \hat{u}(t) &= \{u_j\}. \end{aligned}$$

С использованием этих обозначений, система уравнений (76) может быть записана в виде:

$$D(\hat{u}(t))' + A\hat{u}(t) = \hat{f}(t). \quad (77)$$

Мы должны также учесть начальные условия. Для этого потребуем, чтобы были равны проекции начального условия и решения (метод L_2 -проекций):

$$(u_h^0, v_h) = (u_0, v_h) \quad \text{для любой } v_h \in V_h.$$

Тогда

$$D\hat{u}(0) = u_0^*, \quad \text{где } u_0^* = \{(u_0, \varphi_j)\}.$$

Для примера (72), используем линейные Лагранжевы конечные элементы на разбиении h . Матрицы D и A просто вычисляются:

$$D = \frac{h}{6} \begin{bmatrix} 4 & 1 & 0 & \dots & 0 \\ 1 & 4 & 1 & & \vdots \\ 0 & & \ddots & & 0 \\ \vdots & & & 1 & 4 & 1 \\ 0 & \dots & 0 & 1 & 4 \end{bmatrix}, \quad A = \frac{1}{h} \begin{bmatrix} 2 & -1 & 0 & \dots & 0 \\ -1 & 2 & -1 & & \vdots \\ 0 & & \ddots & & 0 \\ \vdots & & & -1 & 2 & -1 \\ 0 & \dots & 0 & -1 & 2 \end{bmatrix}.$$

Если поделить полученные выражения на h , то матрица A совпадет с конечно-разностной аппроксимацией (73). Однако матрица D не является больше единичной матрицей! Таким образом, в общем случае использование метода конечных элементов приводит к системе уравнений (77), где матрица D не является диагональной.

25.3 Оценка погрешностей

Пусть мы выбрали пространство конечных элементов V_h так, что оно обладает следующим аппроксимационным свойством: для некоторого $r > 1$ выполнено

$$\|v - \Pi_h v\| + h \|\nabla(v - \Pi_h v)\| \leq Ch^s \|v\|_s, \quad 1 \leq s \leq r, \quad v \in H^s(\Omega), \quad (78)$$

где $\Pi_h : V \rightarrow V_h$ – интерполирующий оператор, h – размер разбиения и $\|\bullet\| = \|\bullet\|_{L_2(\Omega)}$. Например, для конформных линейных конечных элементов $r = 2$.

Теорема. Если u достаточно гладкая, тогда $L_2(\Omega)$ ошибка для полудискретизации с конечными элементами V_h , удовлетворяющими свойству (78), может быть записана для $0 \leq t \leq T$ как

$$\|u(t) - u_h(t)\| \leq Ce^{-\alpha t} h^r \|u_0\|_r + Ch^r \left\{ \|u(t)\|_r + \int_0^t e^{-\alpha(t-s)} \|u_t(s)\|_r ds \right\}. \quad (79)$$

Для линейных конечных элементов, эта теорема дает ожидаемую оценку погрешности $O(h^2)$. Существуют аналогичные результаты и для других норм, отличных от L_2 нормы.

25.4 Mass lumping

Процедура mass lumping (рассмотрение в целом) применяется для упрощения решения матричной системы уравнений (77). Для конформных линейных конечных элементов это стандартно применяемый путь решения. В соответствии с этой процедурой, матрица D заменяется на диагональную матрицу, построенную по следующему правилу:

$$D \rightarrow \bar{D} = \{\bar{d}_{ii}\}, \quad \bar{d}_{ii} = \sum_k d_{ik}.$$

Можно показать, что такая процедура эквивалентна приближенному вычислению интегралов (φ_i, φ_j) в слабой форме записи нашей задачи. Результаты о скорости сходимости аналогичны результатам, приведенным в предыдущем параграфе.

Список литературы

- [1] Solin P., *Partial Differential Equations and the Finite Element Method*, Wiley, 2005.
- [2] Сьярле Ф., Метод конечных элементов для эллиптических задач (Мир, 1980).
- [3] Митчелл А.Р., Уэйт Р., Метод конечных элементов для уравнений с частными производными (Мир, 1981).
- [4] Flaherty J.E., *Finite element analysis*, Rensselaer lecture notes, 2000.
- [5] Буслов В.А., Яковлев С.Л., Численные методы I. Исследование функций (СПбГУ, 2001). Глава 2.
- [6] Сабоннадьер Ж.-К., Кулон Ж.-Л., *Метод конечных элементов и САПР* (Мир, 1989). Глава 7.
- [7] Zienkiewicz O.C., Taylor R.L., *The finite element method. Vol. 1. The basis*, 2000.

Содержание

1	Абстрактная задача минимизации.	1
1.1	Эквивалентные вариационные формулировки	2
2	Абстрактная Вариационная Задача	3
3	Формулы Грина	4
4	Примеры краевых задач второго порядка. Граничные условия.	6
4.1	Задача 1: Краевая задача второго порядка – задача Дирихле.	6
4.2	Задача 2: Краевая задача второго порядка – задача Неймана.	8
5	Методы аппроксимации	10
5.1	Метод наименьших квадратов	12
6	Ортогональность ошибок и лемма Шеа (Sea)	13
7	Основные аспекты метода конечных элементов (МКЭ)	14
8	Одномерный МКЭ.	16
8.1	Лагранжевы элементы.	17
8.2	Канонический элемент.	19
8.3	Иерархический базис.	20
9	Пример одномерного МКЭ для краевой задачи.	22
10	Ошибки аппроксимации.	27
10.1	Линейная интерполяция.	29
10.2	Интерполяция полиномами степени p	31
11	Основные этапы использования МКЭ в многомерном случае.	32
12	Лагранжевы элементы в треугольнике.	33
12.1	Лагранжевы элементы порядка p	34
12.1.1	Канонический элемент	36
12.1.2	Барицентрические координаты	37
13	Лагранжевы элементы в прямоугольнике.	39
14	Двумерные иерархические элементы.	41
15	Трехмерные элементы.	44
16	Погрешность аппроксимации для многомерных КЭ.	46

17	Триангуляция в многомерном МКЭ.	48
17.1	Разбиение на блоки, метод Кука	48
17.2	Триангуляция наложением сетки	49
17.3	Фронтальное распространение	50
17.4	Триангуляция наложением слоев	52
17.5	Глобальное разбиение на треугольники и тетраэдры (метод Делоне)	53
18	Координатные преобразования.	54
18.1	Билинейное отображение квадрата с квадратичными функциями (Суб- параметрическое преобразование)	55
18.2	Квадратичное отображение треугольника с квадратичными функциями	56
19	Численное интегрирование в МКЭ.	58
20	Ошибки дискретизации. Ошибки возмущений.	61
20.1	Ошибки интегрирования	62
20.2	Ошибки аппроксимации граничных условий	64
20.3	Ошибки аппроксимации границы	64
21	Априорные и апостериорные оценки погрешностей	65
21.1	Оценки, основанные на экстраполяции решений	65
21.2	Оценки, основанные на невязке решений	67
22	Сверхсходимость решений в МКЭ.	67
23	Адаптивное уточнение решений: h-, p-, и hp-уточнение.	69
24	Метод граничных элементов (МГЭ)	71
24.1	Непрямая формулировка МГЭ	72
24.1.1	Задача Неймана	72
24.1.2	Задача Дирихле	73
24.2	Прямая формулировка МГЭ	73
25	Решение нестационарных уравнений параболического типа.	74
25.1	Слабая формулировка нестационарных уравнений	74
25.2	Полудискретизация конечными элементами	76
25.3	Оценка погрешностей	78
25.4	Mass lumping	78