Numerical Analysis of Elliptic and Parabolic Partial Differential Equations

Prof. Christoph Schwab
Seminar for Applied Mathematics
ETH Zürich

Lecture Notes for
Fall Term 2007

Contents

1 Finite Elements for the Poisson Equation **1**

1.1 One Dimensional Model Problem 1

1.1.1 Variational Formulation of the Problem 1

1.1.2 Sobolev Spaces in One Dimension 3

1.1.3 Finite Element Spaces 9

1.1.4 Finite Element Algorithm 11

1.1.5 Extensions 16

1.1.6 p-Element Shape Functions in Dimension 1 17

1.1.7 Global basis functions 21

1.1.8 Static Condensation 21

1.1.9 Convergence Analysis for the FEM in $\boldsymbol{1-d}$ 23

1.2 Linear Finite Elements in Two Dimensions 28

1.2.1 Sobolev Spaces in $\mathbb{R}^n$, $n > 1$ 28

1.2.2 Variational Formulation for the Poisson Problem 33

1.2.3 Linear Finite Elements for the Poisson Problem 35

1.3 Implementation 38

1.3.1 Geometry Data 39

1.3.2 Element Stiffness Matrix 40

1.3.3 Assembly of the global stiffness matrix 41

1.3.4 Assembly of the Load Vector 42

1.3.5 Homogeneous Dirichlet Data 43

1.4 Further Finite Elements 43

1.4.1 Simplicial Affine Elements 43

1.4.2 Bilinear Quadrilateral Elements 44

1.4.3 Bilinear Quadrilateral Elements of Order $p \geq 1$ 47

1.4.4 Isoparametric Quadrilateral Elements 48

1.5 Element Families in Three Dimensions 49

1.5.1 Reference Elements $\widehat{K}$ 50

1.5.2 Reference Element Shape Functions 50

1.6 Irregular Meshes and Hanging Nodes 53

1.6.1 T Matrices 53

1.6.2 S Matrices 55

1.6.3 Assembly Process 57

1.7 Examination Checklist 58

Bibliography 59

1 Finite Elements for the Poisson Equation

This section will introduce all basic concepts of the FEM for a linear diffusion problem in a domain $D \subset \mathbb{R}^n$ of dimension $n = 1, 2, 3$. We introduce the Dirichlet, Neumann and mixed boundary value problems, their variational formulation, the $L^2(D)$ and $H^k(D)$ Sobolev spaces of functions of finite energy, and all key algorithmic notions of the FEM: element stiffness matrices, load vectors, assembly, element shape functions, static condensation, isoparametric elements in dimensions $n = 1, 2, 3$.

1.1 One Dimensional Model Problem

1.1.1 Variational Formulation of the Problem

We consider the following problem:
find a function $u|D = (0,1) \longmapsto \mathbb{R}$ such that

$$-u''(x) = f(x) \;\text{ for }\; x \in D \tag{1.1.1}$$

together with the boundary conditions

$$u(0) = 0 \text{ and } u'(1) = 0. \tag{1.1.2}$$

Integrating the differential equation we obtain

$$\begin{aligned} u'(\xi) &= -\int_0^\xi f(\eta)\, d\eta + u'(0), \\ u(x) &= -\int_0^x \int_0^\xi f(\eta)\, d\eta\, d\xi + xu'(0) + u(0). \end{aligned}$$

Due to (1.1.2) we have $u(0) = 0$ and

$$0 = u'(1) = -\int_0^1 f(\eta)\, d\eta + u'(0),$$

which leads to the following expression for the solution

$$u(x) = x \int_0^1 f(\eta)\, d\eta - \int_0^x \int_0^\xi f(\eta)\, d\eta\, d\xi. \tag{1.1.3}$$

If the right hand side f is continuous, it follows from (1.1.3) that u is twice continuously differentiable, i. e.

$$f \in C^0(\overline{D}) \Longrightarrow u \in C^2(\overline{D}).$$

In general, u has two continuous derivatives more than f; formally,

$$f \in C^k(\overline{D}) \Longrightarrow u \in C^{k+2}(\overline{D}). \tag{1.1.4}$$

The number $k+2$ of continuous derivatives of u is called **regularity**, and the statement (1.1.4) is a **shift-theorem**.

Let us suppose now that f has finitely many discontinuity points in $(0, 1)$ which we denote by s_i with $i = 1, \ldots, M$ in $(0, 1)$, and that f is continuous otherwise. In this case we say that f is **piecewise continuous** (see Figure 1.1).

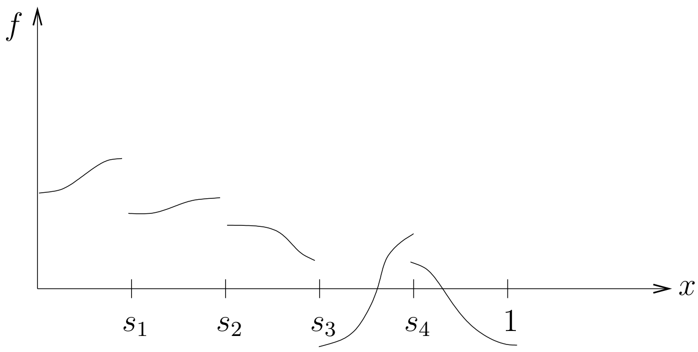


Figure 1.1: Piecewise continuous function.

Definition 1.1. *A function f is* ***piecewise*** *C^k in D, if there exists a partition of D into finitely many open subdomains $D_i \subset D$ with $f|_{D_i} \in C^k(\overline{D_i})$. We write: $f \in C^k_{pw}(D)$.*

From (1.1.3) we immediately get

Proposition 1.2. *Assume that $f \in C^0_{pw}(D)$. Then u as in (1.1.3) is in $C^2_{pw}(D) \cap C^1(\overline{D})$.*

The solution formula (1.1.3) makes sense even if the differential equation (1.1.1) is not defined at s_j.

The starting point of the FEM is the **variational formulation** of (1.1.1) and (1.1.2): Equation (1.1.1) is not satisfied pointwise, it holds only "in an averaged sense". We therefore **multiply** (1.1.1) with a **test function** v with $v(0) = 0$ and **integrate by parts**:

$$\begin{aligned}\int_0^1 f(x)v(x)\,dx &= -\int_0^1 v(x)u''(x)\,dx = \int_0^1 u'(x)v'(x)\,dx - u'(x)\,v(x)\Big|_{x=0}^{x=1} \\ &= \int_0^1 u'(x)v'(x)\,dx.\end{aligned}$$

The **variational formulation** of (1.1.1) reads:
Find $u \in V$ such that

$$a(u,v) := \int_0^1 u'v'\,dx = \int_0^1 fv\,dx =: l(v) \;\text{ for all }\; v \in V. \tag{1.1.5}$$

For $u, v, w \in C^2(\overline{D})$ and for $\lambda \in \mathbb{R}$ it holds:

$$\begin{aligned} a(u+v,w) &= a(u,w) + a(v,w), \\ a(u,v+w) &= a(u,v) + a(u,w), \\ a(\lambda u,v) &= \lambda a(u,v) = a(u,\lambda v), \\ l(v+\lambda w) &= l(v) + \lambda l(w). \end{aligned}$$

We call $l(.)$ a **linear form** and $a(.,.)$ a **bilinear form**. In applications $a(.,.)$ stands for the energy of the system, whereas $l(.)$ stands for external sources, respectively.

We now give a characterization of the space V in (1.1.5).

1.1.2 Sobolev Spaces in One Dimension

A linear function space V is a set of functions such that

$$\begin{aligned} u,v \in V &\Rightarrow u+v \in V \\ u \in V &\Rightarrow \lambda u \in V \quad \forall \lambda \in \mathbb{R} \text{ or } \lambda \in \mathbb{C}; \end{aligned}$$

we also say that V is a function space over $\mathbb{R}$ or $\mathbb{C}$. We will need to measure the size of a function $v \in V$. To this end, we introduce **norms** on V.

Definition 1.3 (Norm on a function space)**.** *The expression $\|.\| : V \to \mathbb{R}_+$ is a norm on V, if it satisfies*

$$\|u\| \geq 0 \text{ and } \|u\| = 0 \Longleftrightarrow u = 0 \tag{N1}$$

$$\|\lambda u\| = |\lambda|\,\|u\| \quad \forall \lambda \in \mathbb{R} \textit{ or } \lambda \in \mathbb{C} \tag{N2}$$

$$\|u+v\| \leq \|u\| + \|v\| \quad \forall u,v \in V. \tag{N3}$$

For example,

$$\|u\|_{C^0(\overline{D})} := \max\{|u(x)| : x \in \overline{D}\}$$

is a norm on $C^0(\overline{D})$ (verify the properties (N1)-(N3)).

The appropriate sets V in (1.1.5) are linear function spaces with functions of finite energy. These are the so-called **Sobolev spaces** $H^k(D)$. The spaces $H^k(D)$ are strictly larger than $C^k(\overline{D})$,

$$H^k(D) \supset C^k(\overline{D}),$$

Definition 1.4. *Let $D = (0,1)$ and $u \in C^0(\overline{D})$. Then we call v the* ***weak*** *or* ***distributional derivative*** *of u, if there holds*

$$\int_D v\varphi \, dx = -\int_D u\varphi' \, dx \quad \forall \varphi \in C_0^1(\overline{D}).$$

Here, $C_0^1(\overline{D}) := \{\varphi \in C^1(\overline{D}) : \varphi(0) = \varphi(1) = 0\}$.

Analogously, we call v the ***k-th weak derivative*** *of u if*

$$\int_D v\varphi \, dx = (-1)^k \int_D u\varphi^{(k)} \, dx \quad \forall \varphi \in C_0^k(\overline{D}), \tag{1.1.6}$$

where $C_0^k(\overline{D}) := \{\varphi \in C^k(\overline{D}) : \varphi^{(j)}(0) = \varphi^{(j)}(1) = 0 \text{ for } j = 0, \ldots, k-1\}$.

Remark 1.5. The integrals in (1.1.6) are Lebesgue-integrals. Therefore, weak derivatives are not defined pointwise - weak derivatives $v(x)$ may be changed on subsets of D of Lebesgue measure zero without effect on (1.1.6). Weak derivatives are therefore equivalence classes of functions v satisfying (1.1.6).

Example 1.6. Let $D = (0,1)$ and assume that $u \in C^0(\overline{D}) \cap C^1_{pw}(\overline{D})$. Then u has a weak derivative u' given by

$$u'(x) = \begin{cases} \frac{du}{dx}(x) & \text{if } s_{i-1} < x < s_i \\ \text{not defined} & \text{for } x = s_i. \end{cases}$$

Remark 1.7. If u is differentiable in the classical sense, i. e. if the limit

$$\frac{du}{dx}(x_0) = \lim_{|h| \to 0} \frac{u(x_0 + h) - u(x_0)}{h}$$

exists for all $x_0 \in \overline{D}$, i. e. $0 \leq x_0 \leq 1$, then $\frac{du}{dx}$ coincides with the weak derivative of u. If however u is continuous but only piecewise differentiable, the weak derivative of u is the discontinuous function that equals the classical derivative of u in each subdomain and is not defined at the discontinuities.

Henceforth we shall make the convention that all derivatives are to be understood in the weak sense.

Definition 1.8 (Sobolev spaces)**.** *Let $D = (0,1)$. Then*

$$H^1(D) := \left\{ u \in L^2(D) : \exists g \in L^2(D) \text{ s.t. } \int_D u\varphi' = -\int_D g\varphi \quad \forall \varphi \in C^1_{\text{comp}}(D) \right\},$$
$$H^k(D) := \left\{ u \in L^2(D) : u' \in H^{k-1}(D) \right\}.$$

We define $H^0(D) := L^2(D)$. Then obviously

$$L^2(D) \supset H^1(D) \supset \cdots \supset H^k(D).$$

Proposition 1.9. *$H^k(D)$ is a normed linear space with norm $\|.\|_{H^k(D)}$ given by*

$$\|u\|_{H^k(D)} := \Big(\sum_{j=0}^{k} \| \frac{d^j u}{dx^j} \|^2_{L^2(D)} \Big)^{\frac{1}{2}} \text{ for } k \geq 0. \tag{1.1.7}$$

We define by

$$|u|_{H^j(D)} := \left\| \frac{d^j u}{dx^j} \right\|_{L^2(D)} = \Big(\int_D \Big(\frac{d^j u}{dx^j}\,(x) \Big)^2 dx \Big)^{\frac{1}{2}} \tag{1.1.8}$$

*the so-called $H^j(D)$-**semi-norm**.*

Exercise 1.10. Verify the norm axioms for $\|.\|_{H^k(D)}$. Show, that $|.|_{H^k(D)}$ verifies (N2)–(N3), but does not verify (N1). Find all functions u, which do not verify (N1).

Example 1.11. Let $D = (0, 1)$.

1. $u(x) = x^\alpha \in L^2(D)$ for $\alpha > -1/2$, since

$$\|u\|^2_{L^2(D)} = \int_0^1 x^{2\alpha}\, dx = \frac{1}{2\alpha + 1} x^{2\alpha+1} \Big|_0^1 < \infty \iff \alpha > -1/2.$$

2. $u(x) = x^\alpha \in H^1(D)$ for $\alpha > \frac{1}{2}$, since u is strongly differentiable at x for all $0 < x \leq 1$:

$$\frac{du}{dx}(x) = \alpha x^{\alpha-1} \text{ for } 0 < x \leq 1,$$

   therefore $u(x) = x^\alpha \in H^1(D)$ follows from 1.

3. $u(x) = x^\alpha \in H^k(D)$ for $\alpha > -\frac{1}{2} + k$.

4.
$$u(x) = \begin{cases} 0 & \text{for } 0 \leq x \leq \frac{1}{2} \\ 1 & \text{for } x > \frac{1}{2} \end{cases}$$

   is in $L^2(D)$, but is not in $H^1(D)$. The proof follows from Theorem 1.13 below.

Without proof, we state the following important facts about Sobolev spaces.

Remark 1.12.

1. $H^k(D)$ is complete, and therefore a Banach space.
2. Smooth functions $C^\infty(\overline{D})$ are dense in $H^k(D)$ for all k.

For intervals D, functions in $H^1(D)$ are continuous:

Theorem 1.13. *Let $D = (0,1)$ and $u \in H^1(D)$. Then:*

1. *u can be modified in a set of Lebesgue measure zero to $\overline{u}$ s.t. $\overline{u} \in C^0(\overline{D})$; and*
2. *for each $x_0 \in \overline{D}$ the following* ***trace estimate***

$$|u(x_0)| \leq C(D) \, \|u\|_{H^1(D)} \tag{1.1.9}$$

holds, where C is independent of x_0.

For nonlinear problems, spaces of functions with $L^p(D)$-integrability for $1 \leq p \leq \infty$ are of interest.

Definition 1.14. *Let $1 \leq p \leq \infty$ and $D = (0,1)$. Then*

$$W^{1,p}(D) := \Big\{ u \in L^p(D) : \ \exists g \in L^p(D) \text{ such that } \int_D u\varphi' = -\int_D g\varphi \\ \text{for all testfunctions } \varphi \in C_0^1(\overline{D}) \Big\} .$$

Obviously,

$$H^1(D) = W^{1,2}(D) \, ,$$

and for $u \in W^{1,p}(D)$ we denote $u' = g$ (with equality understood in $L^p(D)$).

Moreover, if $u \in C^1(D) \cap L^p(D)$ and $u' \in L^p(D)$, then $u \in W^{1,p}(D)$, i.e. the classical derivative coincides with generalized one in $L^p(D)$.

The space $W^{1,p}(D)$ is equipped with the norms

$$\|u\|_{W^{1,p}(D)} := \big(\|u\|^p_{L^p(D)} + \|u'\|^p_{L^p(D)} \big)^{\frac{1}{p}}, \ 1 \leq p < \infty,$$
$$\|u\|_{W^{1,\infty}(D)} := \max \big\{ \|u\|_{L^\infty(D)}, \ \|u'\|_{L^\infty(D)} \big\} .$$

It holds (see, e.g., [Brezis: Analyse Fonctionnelle, Prop. VIII.1]

- $W^{1,p}(D)$ is a Banach space for $1 \leq p \leq \infty$,
- $W^{1,p}(D)$ is reflexive for $1 < p < \infty$,
- $W^{1,p}(D)$ is separable for $1 \leq p < \infty$.

Theorem 1.15. *Let $D = (a, b) \subset \mathbb{R}$ be bounded and let $u \in W^{1,p}(D)$. Then there is $\widetilde{u} \in C^0(\overline{D})$ such that*

$$1) \qquad u(x) = \widetilde{u}(x) \text{ a.e. } x \in D\,,$$

$$2) \qquad \widetilde{u}(x) - \widetilde{u}(y) = \int_y^x u'(\xi)\, d\xi \quad \forall x, y \in \overline{D}\,,$$

$$3) \qquad \sup_{x,y\in\overline{D}, x\neq y} \frac{|\widetilde{u}(x) - \widetilde{u}(y)|}{|x-y|^\alpha} \leq \|u'\|_{L^p(D)} \text{ for } \alpha = 1 - \frac{1}{p},\ 1 \leq p \leq \infty\,.$$

To prove Theorem 1.15, we need

Lemma 1.16. *Let $D = (0, 1)$, $f \in L^1_{\text{loc}}(D)$ such that*

$$\int_D f\varphi' = 0 \quad \forall \varphi \in C_0^1(\overline{D})\,. \tag{*}$$

Then there exists a constant C s.t. $f(x) \equiv C$ a.e. $x \in D$.

Proof. Fix $\psi(x) \in C_0^0(\overline{D})$, $\displaystyle\int_D \psi = 1$. Then $\forall w \in C_0^0(\overline{D})$ ex. $\varphi \in C_0^1(\overline{D})$ such that

$$\varphi' = w - \Big(\int_D w\Big)\psi\,.$$

Indeed, $h := w - (\int_D w)\psi \in C_0^0(D)$, $\int_D h = 0$.

$\Longrightarrow$ there exists an antiderivative $H(x) = \int^x h \in C_0^1(\overline{D})$. Now (*) implies:

$$\int_D f\Big\{w - \Big(\int_D w\Big)\psi\Big\} = 0 \quad \forall w \in C_0^1(\overline{D})\,,$$

i.e.

$$\int_D \Big\{f - \Big(\int_D f\psi\Big)\Big\} w = 0 \quad \forall w \in C_0^1(\overline{D})\,.$$

$\Longrightarrow$

$$f(x) = \int_D f\psi \text{ a.e. } x \in D$$

$\Longrightarrow$

$$f \equiv C \text{ a.e. } x \in D\,.$$

□

Lemma 1.17. *Let $g(x) \in L^1_{\text{loc}}(D)$, $y_0 \in D$ fixed. Put*

$$\nu(x) = \int_{y_0}^{x} g(t)\,dt\,, \quad x \in D\,.$$

Then $\nu \in C^0(\overline{D})$ and

$$\int_D \nu\varphi' = -\int_D g\varphi \quad \forall \varphi \in C^1_0(\overline{D})\,.$$

Proof. The Lemma is a consequence of Fubini's Theorem and left as exercise. □

Now we give the *Proof of Theorem 1.15*:

1) $w.l.o.g. D = (0.1)$. Fix $y_0 \in D$ and put

$$\overline{u}(x) := \int_{y_0}^{x} u'(t)\,dt\,.$$

Lemma 1.17 $\Longrightarrow$

$$\begin{aligned} &\int_D \overline{u}\,\varphi' &&= -\int_D u'\varphi &&\forall \varphi \in C^1_0(D)\,,\\ \Longrightarrow\ &\int_D (u-\overline{u})\,\varphi' &&= 0 &&\forall \varphi \in C^1_0(D)\,. \end{aligned}$$

Lemma 1.16 $\Longrightarrow$

$$(u-\overline{u})(x) = \text{const.} \quad \text{a.e. } x \in D\,,$$

$\Longrightarrow \qquad \widetilde{u}(x) := \overline{u}(x) +$ Const. satisfies 1) and 2).

To prove 3), we use that for $x, y \in D$, $x \neq y$, we have

$$\begin{aligned} |\widetilde{u}(x) - \widetilde{u}(y)| &= \Big|\int_y^x u'(t)\,dt\Big| \\ &\leq \Big(\int_y^x \wedge\,dt\Big)^{\frac{1}{q}} \Big(\int_D |u'(t)|^p\,dt\Big)^{\frac{1}{p}} \\ &= |x-y|^{\frac{1}{q}}\,\|u'\|_{L^p(D)}\,. \end{aligned}$$

□

Exercise 1.18. The constant C in (1.1.9) depends on the domain D. Find the constant C corresponding to an interval (a, b). Hint: use a scaling argument.

In particular, $C^1(D) \nsubseteq H^1(D)$. Moreover, Theorem 1.13 allows us to prescribe boundary values for functions in $H^1(D)$.

Definition 1.19. *Let $D = (0, 1)$. Then we define*

$$H_0^1(D) := \left\{u \in H^1(D) : u(0) = u(1) = 0\right\}. \tag{1.1.10}$$

Remark 1.20. Due to (1.1.9) this definition makes sense, while e. g. $u'(0)$ and $u'(1)$ are not defined. For $u \in H^1(D)$, $u' \in L^2(D)$ (and $u'(0)$ does not need to be finite):

$$u(x) = x^{\frac{3}{4}} \in H^1(D) \text{ but } u'(0) = \frac{3}{4}x^{-\frac{1}{4}}\Big|_{x=0} = \infty.$$

Exercise 1.21. $H_0^1(D) \subset H^1(D)$ is a closed linear subspace.

We can give now a precise description of the space V in (1.1.5):
Find $u \in V$ with $V = \{u \in H^1(D) : u(0) = 0\}$, such that

$$a(u, v) = l(v) \text{ for all } v \in V. \tag{1.1.11}$$

1.1.3 Finite Element Spaces

A FEM corresponds to a discretization of the variational formulation (1.1.11), in which the Ansatz- and test functions u and v are now elements of a finite dimensional subspace $V_N \subset V$ of V of finite dimension $N = \dim V_N < \infty$.

A linear space V_N is **N-dimensional**, if there exists a *basis* $\{b_i\}_{i=1}^N$, such that

$$V_N = \{u : u(x) = \alpha_1 b_1(x) + \cdots + \alpha_N b_N(x) \text{ for } \alpha_i \in \mathbb{R}\} = \operatorname{span}\{b_i\}_{i=1}^N. \tag{1.1.12}$$

Then

$$\sum_{i=1}^N \alpha_i b_i(x) = 0 \iff \alpha_i = 0 \text{ for all } i = 1, \ldots, N \text{ (linear independence)}. \tag{1.1.13}$$

Example 1.22.

1. $\mathbb{R}^N = \operatorname{span}\{e_i\}_{i=1}^N$ with basis vectors $e_1 = (1, 0, \ldots, 0)^\top$, $e_2 = (0, 1, 0, \ldots, 0)^\top$, etc.
2. Let $I \subset \mathbb{R}$ be an interval. Then

$$\begin{aligned}\mathcal{P}_d(I) &= \{u : u \text{ is polynomial of degree at most } d\}\\ &= \left\{u(x) = \alpha_0 + \alpha_1 x + \cdots \alpha_d x^d \,|\, \alpha_i \in \mathbb{R}\right\}\\ &= \operatorname{span}\{b_i\}_{i=0}^d \text{ with } b_i = x^i,\end{aligned}$$

   and $\dim \mathcal{P}_d(I) = d + 1$.

3. Let $D = (0, 1)$ and assume that

$$\mathcal{T} = \{0 = x_0 < x_1 < \cdots < x_M = 1\}$$

is an arbitrary mesh, i. e. a partition of the domain in M non-overlapping subdomains. We also use the notation

$$\mathcal{T} = \{K_j\}_{j=1}^{M} \text{ with elements } K_j = (x_{j-1}, x_j)$$

Then we have the following sequence of FE-spaces

$$S^{d,l}(D, \mathcal{T}) = \left\{u \in H^l(D) : u|_K \in \mathcal{P}_d \text{ for } K \in \mathcal{T}\right\} \text{ for } l = 0, 1, \ldots. \tag{1.1.14}$$

For $l = 1$ we drop the l and use the notation $S^d(D, \mathcal{T})$. For $d = 1$ and $l = 1$ the space S^1 is the space of **piecewise linear continuous functions**, since $S^1 = S^{1,1} \subset H^1 \subset C^0$ (see also Theorem 1.13).

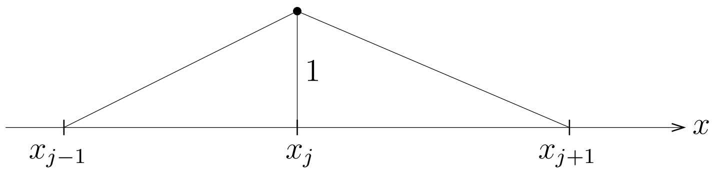


Figure 1.2: Hat function $b_j(x)$.

A basis for $S^1(D, \mathcal{T})$ is given by the so-called "hat functions" $\{b_i\}_{i=0}^{M}$ with

$$b_i(x) \in S^{1,1}(D, \mathcal{T}), \quad b_i(x_j) = \delta_{ij} \text{ for } 0 \le i, j \le M, \tag{1.1.15}$$

as illustrated in Figure 1.2. Therefore $N = \dim\left(S^{1,1}(D, \mathcal{T})\right) = M + 1$.

4. Further, we define

$$S_0^{1,1}(D, \mathcal{T}) = S^{1,1}(D, \mathcal{T}) \cap H_0^1(D). \tag{1.1.16}$$

Then,

$$S_0^{1,1}(D, \mathcal{T}) = \operatorname{span}\{b_i\}_{i=1}^{M-1} \text{ with } N = \dim S_0^{1,1} = M - 1.$$

We can give now the **Finite Element discretization** of (1.1.11):
Find $u_N \in V_N := \{u \in S^{1,1}(D, \mathcal{T}) : u(0) = 0\}$ such that

$$a(u_N, v) = l(v) \text{ for all } v \in V_N. \tag{1.1.17}$$

This problem corresponds to a linear system of equations with $N = M$ unknowns, since in terms of the basis functions $b_i(x)$ from (1.1.15)

$$u_N(x) = \sum_{j=1}^{N} \alpha_j b_j(x) = \underline{\alpha}^\top \underline{b} \text{ with } u_N(x_i) = \alpha_i. \tag{1.1.18}$$

Algorithm 1.1 Finite Element algorithm.

- Computing of element stiffness matrices and element load vectors,
- assembling of the (global) linear system,
- solving the resulting linear system of equations,
- post-processing.

Analogously,

$$v = \sum_{i=1}^{N} v_i b_i = \underline{v}^\top \underline{b},$$

where we use the notation $\underline{v}$ to indicate the vector of the nodal values of the FE test function v and we write $\underline{b}$ for the vector of the FE basis functions in V_N. Substituting this into (1.1.17) we obtain

$$\begin{aligned} a\Big(\sum_j \alpha_j b_j, \sum_i v_i b_i\Big) &= l\Big(\sum_i v_i b_i\Big) \\ \iff \sum_i v_i \Big(\sum_j a(b_j, b_i)\alpha_j - l(b_i)\Big) &= 0 \quad \forall \{v_i\} \in \mathbb{R}^N \\ \iff \underline{v}^\top (\boldsymbol{A}\underline{\alpha} - \underline{l}) &= 0 \quad \forall \underline{v} \in \mathbb{R}^N \\ \iff \boldsymbol{A}\underline{\alpha} &= \underline{l}. \end{aligned} \tag{1.1.19}$$

Here we denote by $\boldsymbol{A}$, with $A_{ij}\, a(b_j, b_i)$, the **(global) stiffness matrix**, and by $\underline{l}$ the **load vector**. Since $a(b_i, b_j) = 0$ for $|i - j| \geq 2$, the matrix $\boldsymbol{A}$ is tridiagonal. Furthermore, since $a(u, v) = a(v, u)$ the matrix $\boldsymbol{A}$ is also symmetric.

1.1.4 Finite Element Algorithm

We illustrate the FE Algorithm 1.1 for a more general problem

$$-(a(x)u')' + c(x)u = f \text{ in } D = (0, 1), \tag{1.1.20a}$$

$$u(0) = 0, \quad a(1)u'(1) = g. \tag{1.1.20b}$$

Again, we take $V = \{u \in H^1(D)\colon u(0) = 0\}$. The variational formulation then reads: Find $u \in V$ such that

$$a(u, v) = l(v) \quad \forall v \in V, \tag{1.1.21}$$

where

$$a(u, v) = \int_0^1 \big(a(x)u'v' + c(x)uv\big)\, dx \text{ and } l(v) = \int_0^1 fv\, dx + gv(1). \tag{1.1.22}$$

Exercise 1.23. Derive the variational formulation (1.1.22)! What happens to the boundary condition at $x = 1$?

The (abstract) FE discretization is again (1.1.17) and the goal of Algorithm 1.1 is to compute $\boldsymbol{A}$ and $\underline{l}$ in (1.1.19). We assume that

$$a(x) = a = \text{const} > 0, c(x) = c = \text{const} > 0 \text{ and } g = 0,$$

and consider piecewise linear, continuous elements, i. e. V_N is as in (1.1.17). Furthermore we use a uniform mesh, i. e.

$$\mathcal{T} = \{K_j\}_{j=1}^N, K_j = (x_{j-1}, x_j), x_j = jh, h = \frac{1}{N} \text{ and } j = 1, \ldots, N.$$

Then the basis function $b_j(x)$ in Figure 1.2 is given by

$$b_j(x) = \begin{cases} 1 - |x - x_j|/h & x_{j-1} \le x \le x_{j+1} \\ 0 & \text{otherwise,} \end{cases}$$

and $\operatorname{supp}(b_j) \subset [x_{j-1}, x_{j+1}]$.

For $1 \le i, j \le N$, we compute

$$\begin{aligned} A_{ij} = a(b_j, b_i) &= a \int_0^1 b_j'(x) b_i'(x)\,dx + c \int_0^1 b_j(x) b_i(x)\,dx \\ &= \frac{a}{h} \left\{ \begin{array}{rl} 2 & \text{if } i = j, \\ -1 & \text{if } |i-j| = 1, \\ 0 & \text{else} \end{array} \right\} + ch \left\{ \begin{array}{rl} 2/3 & \text{if } i = j, \\ 1/6 & \text{if } |i-j| = 1, \\ 0 & \text{else} \end{array} \right\}, \end{aligned}$$

and for each $i = 1, \ldots, N$ in (1.1.19) we obtain:

$$-a \frac{\alpha_{i-1} - 2\alpha_i + \alpha_{i+1}}{h^2} + c \frac{\alpha_{i-1} + 4\alpha_i + \alpha_{i+1}}{6} = \frac{1}{h} \int_0^1 f b_i\,dx.$$

The above procedure based on global shape-functions is not appropriate for computer implementation. It is more efficient to base the computation of the stiffness matrix and load vector on elements $K \in \mathcal{T}$.

Let now $\boldsymbol{\mathcal{T}}$ **be an arbitrary mesh**, and define

$$\mathcal{T} = \{K_j\}_{j=1}^N, K_j = (x_{j-1}, x_j), h_j := |K_j| = x_j - x_{j-1} \text{ and } m_j := \frac{x_{j-1} + x_j}{2}. \quad (1.1.23)$$

We split $a(.,.)$ and $l(.)$ into their element contributions:

$$\begin{aligned} a(u, v) &= \sum_K a_K(u, v), & a_K(u, v) &= \int_K a(x) u' v' + c(x) u v\,dx, \\ l(v) &= \sum_K l_K(v), & l_K(v) &= \int_K f v\,dx. \end{aligned}$$

The **element stiffness matrix** $\boldsymbol{A}_K$ is the matrix corresponding to the element bilinear form $a_K(u,v)$: Let u and $v \in V_N$. Then the restrictions $u|_{K_j}$ and $v|_{K_j}$ are linear,

$$u\big|_{K_j}(x) = u(x_{j-1})b_{j-1}(x)\big|_{K_j} + u(x_j)b_j(x)\big|_{K_j} = \underline{u}_{K_j}^\top \underline{N}_{K_j}(x),$$

where $\underline{u}_{K_j} = (u(x_{j-1}), u(x_j))^\top$ is the vector of **elemental degrees of freedom** and $\underline{N}_{K_j} = (b_{j-1}(x)|_{K_j}, b_j(x)|_{K_j})$ is the vector of **element shape-functions**. Similarly, for a generic element $K \in \mathcal{T}$ and $v \in V_N$,

$$v(x)\big|_K = \underline{v}_K^\top \underline{N}_K(x)$$

and therefore

$$a_K(u\big|_K, v\big|_K) = \underline{v}_K^\top \boldsymbol{A}_K \underline{u}_K.$$

The **element stiffness matrix** $\boldsymbol{A}_K$ can be computed independently for each element, therefore the computation of the element stiffness matrices can be done in parallel. To this end, we transform each element $K_j \in \mathcal{T}$ to the so-called **reference element** $\widehat{K} = (-1,1)$ via an **element mapping**

$$K_j \ni x = F_{K_j}(\xi) := m_j + h_j\,\xi/2,\ \xi \in \widehat{K}. \tag{1.1.24}$$

Then, for all $K \in \mathcal{T}$ it holds:

$$\underline{N}_K(F_K(\xi)) = \underline{\widehat{N}}(\xi) = (\widehat{N}_0(\xi),\, \widehat{N}_1(\xi))^\top,$$

where the **reference element shape-functions**

$$\widehat{N}_0(\xi) = \frac{1-\xi}{2},\ \ \widehat{N}_1(\xi) = \frac{1+\xi}{2} \tag{1.1.25}$$

are now independent of the element K!

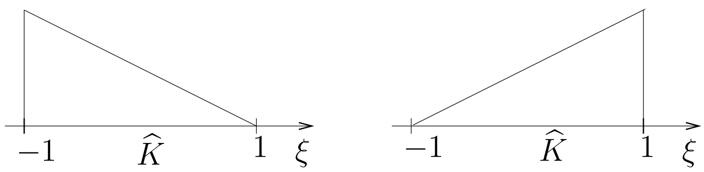


Figure 1.3: Reference element shape-functions $\widehat{N}_0$ and $\widehat{N}_1$ in the reference element $\widehat{K}$.

Furthermore, for $i,j = 0,1$ it holds:

$$\begin{aligned}(\boldsymbol{A}_K)_{ij} &:= a_K((\underline{N}_K)_i, (\underline{N}_K)_j)\\ &= \int_K (a(x)(\underline{N}_K)_i'(\underline{N}_K)_j' + c(x)(\underline{N}_K)_i(\underline{N}_K)_j)\,dx\\ &= \int_{\widehat{K}} \Big(\widehat{a}_K(\xi)\Big(\frac{d\xi}{dx}\Big)^2 \widehat{N}_i'\widehat{N}_j' + \widehat{c}_K(\xi)\widehat{N}_i\widehat{N}_j\Big)\frac{dx}{d\xi}\,d\xi,\end{aligned}$$

where

$$\widehat{a}_K(\xi) := a(F_K(\xi)), \;\; \widehat{c}_K(\xi) := c(F_K(\xi)),$$

Since $d\xi/dx = 2/h_K$, we obtain

$$(\boldsymbol{A}_K)_{ij} = \int_{\widehat{K}} \left(\widehat{a}_K(\xi)\, \frac{2}{h_K}\, \widehat{N}_i' \, \widehat{N}_j' + \widehat{c}_K(\xi)\, \frac{h_K}{2}\, \widehat{N}_i \, \widehat{N}_j \right) d\xi. \tag{1.1.26}$$

For general coefficients $a(x)$ and $c(x)$ these integrals cannot be computed exactly. Hence, they have to be approximated by a numerical quadrature. However, for piecewise constant $a(x)$, $c(x)$:

$$a(x)|_K = a_K = \text{const.}, \quad c(x)|_K = c_K = \text{const.}, \tag{1.1.27}$$

we obtain that

$$\boldsymbol{A}_K = a_K \frac{2}{h_K} \widehat{\boldsymbol{S}} + c_K \frac{h_K}{2} \widehat{\boldsymbol{M}}, K \in \mathcal{T}, \tag{1.1.28}$$

where $\widehat{\boldsymbol{S}}$, $\widehat{\boldsymbol{M}}$ given by

$$\widehat{\boldsymbol{S}} = \begin{pmatrix} 1/2 & -1/2 \\ -1/2 & 1/2 \end{pmatrix}, \widehat{\boldsymbol{M}} = \begin{pmatrix} 2/3 & 1/3 \\ 1/3 & 2/3 \end{pmatrix},$$

are independent of K. Analogously, we obtain the **element load vector**

$$\underline{l}_K = \int_{\widehat{K}} \widehat{f}_K(\xi)\, \underline{\widehat{N}}(\xi)\, \frac{h_K}{2}\, d\xi, \quad K \in \mathcal{T}.$$

Note: For a large number of elements $\boldsymbol{A}_K$ and $\underline{l}_K$ can be computed independently and therefore in parallel.

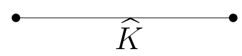


Figure 1.4: Reference element $\widehat{K}$ with 2 degrees of freedom.

The piecewise linear element has two degrees of freedom, namely the function values at the end points. This is illustrated in Figure 1.4. It remains now to construct the global stiffness matrix $\boldsymbol{A}$ and the global load vector $\underline{l}$ in (1.1.19) from the element stiffness matrices $\boldsymbol{A}_K$ and element load vectors $\underline{l}_K$, respectively. This procedure is called **assembly** of $\boldsymbol{A}$ and $\underline{l}$.

Assembly of the global linear system

At this stage we introduce the concept of global assembly. The idea is to express the global basis functions $b_i(x)$ in terms of the local expansion shape-functions. Let us consider the mesh

$$0 = x_0 < x_1 < \cdots < x_N = 1, \quad K_i = (x_{i-1}, x_i), \quad \mathcal{T} = \{K_i\},\, i \in \{1, \ldots, N\}$$

where the indices i correspond to global nodes $i \in \{0, 1, 2, \ldots, N\}$. The indices $j \in \{0, 1\}$ are used for the element shape functions. This correspondence is realized with the so-called $\boldsymbol{T}$ **matrices** defined as follows: for each element $K \in \mathcal{T}$ the matrix $\boldsymbol{T}_K \in \mathbb{R}^{2\times(N+1)}$ is given by:

$$[\boldsymbol{T}_K]_{ji} := \begin{cases} 1 & \text{if } i \text{ is the global index of the element end point } j \in \{0,1\}, \\ 0 & \text{otherwise.} \end{cases}$$

With this, we can assemble the vector of global basis functions in terms of the element shape-functions as follows

$$\underline{b}(x) = \sum_{K \in \mathcal{T}} \boldsymbol{T}_K^\top \underline{N}_K(x),$$

and from (1.1.19) we get

$$\begin{aligned} \boldsymbol{A} = a(\underline{b}, \underline{b}^\top) &= a\Big(\sum_K \boldsymbol{T}_K^\top \underline{N}_K, \Big(\sum_{K'} \boldsymbol{T}_{K'}^\top \underline{N}_{K'} \Big)^\top \Big) \\ &= \sum_{K,K'} \boldsymbol{T}_K^\top \underbrace{a_K(\underline{N}_K, \underline{N}_{K'}^\top)}_{=0 \text{ for } K' \neq K} \boldsymbol{T}_{K'} \\ &= \sum_K \boldsymbol{T}_K^\top \boldsymbol{A}_K \boldsymbol{T}_K. \end{aligned}$$

The global stiffness matrix $\boldsymbol{A}$ is now obtained via summation from $\boldsymbol{A}_K$:

$$\boldsymbol{A} = \begin{pmatrix} [\boldsymbol{A}_{K_1}] & & & \\ & 0 & & \\ & & \ddots & \\ & & & 0 \end{pmatrix} + \begin{pmatrix} 0 & & & & \\ & [\boldsymbol{A}_{K_2}] & & & \\ & & 0 & & \\ & & & \ddots & \\ & & & & 0 \end{pmatrix} + \cdots + \begin{pmatrix} 0 & & & \\ & \ddots & & \\ & & 0 & \\ & & & [\boldsymbol{A}_{K_M}] \end{pmatrix}. \tag{1.1.29}$$

Analogously, we obtain the global load vector $\underline{l}$. Symbolically, the assembly can be written as follows

$$\boldsymbol{A} = \mathcal{A}_{K\in\mathcal{T}}\boldsymbol{A}_K, \quad \underline{l} = \mathcal{A}_{K\in\mathcal{T}}\underline{l}_K, \quad \underline{b} = \mathcal{A}_{K\in\mathcal{T}}\underline{N}_K. \tag{1.1.30}$$

A more detailed explanation of the assembly process is given in Section 1.6 below.

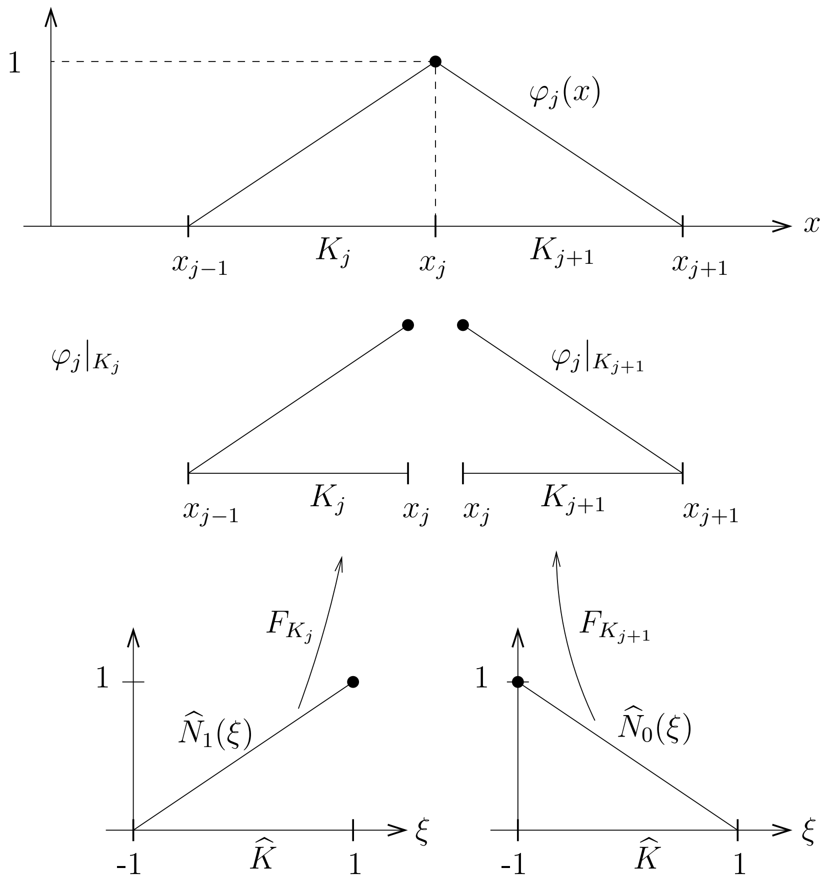


Figure 1.5: Assembly of the global basis function $b_j(x)$ from the element shape-functions $\underline{N}_{K_j}$, which are mapped from the reference shape-functions $\underline{\widehat{N}}(\xi)$.

1.1.5 Extensions

So far, we considered finite elements on triangulations $\mathcal{T}_h$ of meshwidth h, which approximate the exact solution u by piecewise polynomials on the elements $K \in \mathcal{T}_h$. To reduce the error $u - u_N$ in this Finite Element solution u_N, one can reduce the meshwidth h, thereby introducing more degrees of freedom (the error reduction will be precisely analyzed below), or, equivalently, increasing the dimension N of the FE-space from N_1 to $N_2 > N_1$, say.

The process of enlarging the dimension of the FE space is called an **extension**. Since mesh refinement reduces the meshwidth h of the triangulation, extension by mesh refinement is called ***h*-extension**. An alternative to h-extension which is often used in fluid-dynamics is ***p*-extension** where the polynomial degree p of the elements is increased on a fixed mesh $\mathcal{T}$. Mathematically, a sequence of subspaces obtained by p-extensions is

$$S^1(D,\mathcal{T}) \subset S^2(D,\mathcal{T}) \subset \cdots \subset S^p(D,\mathcal{T})\,. \tag{1.1.31}$$

1.1.6 p-Element Shape Functions in Dimension 1

To realize the p-extension (1.1.31), we must provide a family of elements, one for each polynomial degree $p \geq 1$. We consider again the domain $D = (0,1)$ and the problem

$$\begin{aligned} -(a(x)u')' + c(x)u &= f \quad \text{in } D \\ u(0) = u(1) &= 0\,. \end{aligned} \tag{1.1.32}$$

The variational formulation of (1.1.32) is:

$$u \in H_0^1(D) : \underbrace{\int_D (a(x)u'v' + c(x)uv)dx}_{=:a(u,v)} = \underbrace{\int_D fvdx}_{=:\ell(v)} \;\forall v \in H_0^1(D)\,. \tag{1.1.33}$$

As before, we let

$$0 = x_0 < x_1 < \cdots < x_M = 1$$

be a sequence of nodes in D and denote by $\mathcal{T} = \{K_j : j = 1,\ldots,M\}$ the FE-mesh with elements $K_j = (x_{j-1}, x_j)$ of size $h_j = x_j - x_{j-1}$.

We *assume* again that $a(x)$, $c(x)$ are piecewise constant:

$$a(x)\big|_{K_j} = a_j, \quad c(x)\big|_{K_j} = c_j\,. \tag{1.1.34}$$

Let $p \geq 1$ be a polynomial degree. Then

$$S^p(D,\mathcal{T}) = \{u \in C^0(D) : \, u\big|_K \in \mathcal{P}_p(K),\, K \in \mathcal{T}\} \tag{1.1.35}$$

is the FE-space of continuous functions which are polynomials of degree p in each element. Since $\dim(\mathcal{P}_p(K)) = p+1$, we have

$$N = \dim(S^p(D,\mathcal{T})) = M(p+1) - (M-1) = Mp + 1\,, \tag{1.1.36}$$

since continuity at x_j, $j = 1,\ldots,M-1$, imposes $M-1$ constraints.

The FE-space is then

$$V_N = S_0^p(D,\mathcal{T}) := S^p(D,\mathcal{T}) \cap H_0^1(D)\,, \tag{1.1.37}$$

and it has dimension $N = Mp - 1$.

p-Element Stiffness Matrix

As before, due to Assumption (1.1.34), the element stiffness matrix $\boldsymbol{A}_K$ for $K \in \mathcal{T}$ is given by

$$\boldsymbol{A}_K = \frac{2}{h_K}\, a_K\, \widehat{\boldsymbol{S}} + \frac{h_K}{2}\, c_K\, \widehat{\boldsymbol{M}} \tag{1.1.38}$$

where $\widehat{\boldsymbol{S}}, \widehat{\boldsymbol{M}}$ are the reference element stiffness and mass-matrices, respectively.

If $\underline{\widehat{N}} = (\widehat{N}_0, \ldots, \widehat{N}_p)^\top$ is a basis of $\mathcal{P}_p(\widehat{K})$, the polynomials of degree p on $\widehat{K} = (-1, 1)$, then

$$\widehat{\boldsymbol{S}} = \int_{-1}^{1} \underline{\widehat{N}}'(\xi)\, \underline{\widehat{N}}'(\xi)^\top\, d\xi,\ \widehat{\boldsymbol{M}} = \int_{-1}^{1} \underline{\widehat{N}}(\xi)\, \underline{\widehat{N}}(\xi)^\top d\xi\,. \tag{1.1.39}$$

It remains to specify the reference element shape-functions $\widehat{N}_j(\xi)$. Two choices are widely used:

1. Nodal Basis (“spectral” elements)

For $p \geq 1$, let

$$-1 =: \widehat{\xi}_0^p < \widehat{\xi}_1^p < \widehat{\xi}_2^p < \cdots < \widehat{\xi}_p^p := 1 \tag{1.1.40}$$

a set of $p+1$ points in $\widehat{K}$. Then the nodal (or Lagrange) shape-functions are given by

$$\widehat{N}_j^p(\xi) := \frac{\prod\limits_{i=0}^{p} (\xi - \widehat{\xi}_i^p)}{\prod\limits_{\substack{i=0\\ i\neq j}}^{p} (\widehat{\xi}_i^p - \widehat{\xi}_j^p)}\,. \tag{1.1.41}$$

Evidently, $\widehat{N}_j^p(\xi) \in \mathcal{P}_p(\widehat{K})$ and

$$\widehat{N}_j^p(\widehat{\xi}_i^p) = \delta_{ij}, \quad i, j - 0, \ldots, p \tag{1.1.42}$$

i.e. $\widehat{N}_j^p(\xi)$ vanishes in all nodes except node number j.

Remark 1.24.

i) many choices (1.1.40) of nodes are possible. The most common ones are:

equidistant nodes:

$$\widehat{\xi}_j^p := 1 + \frac{2j}{p},\ \ j = 0, \ldots, p\,. \tag{1.1.43}$$

These are not recommended for high p, since then $\widehat{\boldsymbol{S}}, \widehat{\boldsymbol{M}}$ become extremely ill-conditioned for large p.

Chebysev nodes:

$$\widehat{\xi}_j^p := -\cos\Big(j\,\frac{\pi}{p}\Big) \quad j = 0, 1, \ldots, p\,. \tag{1.1.44}$$

Here the conditioning is better.

Lobatto nodes: $\widehat{\xi}_j^p$ are the $p+1$ zeros of

$$(1-\xi^2)\,L_p'(\xi)\,, \tag{1.1.45}$$

where $L_p(\xi)$ is the p-th Legendre polynomial $\widehat{K}$. These points (which are used in spectral methods) are as well-conditioned as (1.1.44), but are in addition suitable for numerical integration.

ii) The $\widehat{N}_j^p(\xi)$ in (1.1.41) are not hierarchical - if p is increased, all reference element shape-functions change, and consequently $\widehat{\boldsymbol{S}}$, $\widehat{\boldsymbol{M}}$ must be reevaluated at each increase of the degree p.

2. Hierarchical Basis

A family $\big\{\{\widehat{N}_j^p\}_{j=0}^p\big\}_{p=1}^\infty$ is **hierarchical**, if

$$\widehat{N}_j^{p+1} = \widehat{N}_j^p\,, \quad j = 0, \ldots, p\,, \tag{1.1.46}$$

i.e. if the shape-functions of degree p form a subset of those of degree $p+1$.

Hierarchical reference element shape-functions can be built out of **Legendre polynomials, $\boldsymbol{\{L_p(\xi)\}_{p=0}^\infty}$**. They are obtained by applying the Gram-Schmidt process in $L^2(-1,1)$ to the monomials $\{\xi^p\}_{p=0}^\infty$. For classical reasons, the $\{L_p(\xi)\}$ are normalized by the condition

$$L_p(1) = 1\,. \tag{1.1.47}$$

Legendre polynomials satisfy the **Legendre differential equation**

$$\big((1-\xi^2)L_p'(\xi)\big)' + p(p+1)L_p(\xi) = 0 \text{ in } \widehat{K} = (-1,1), \tag{1.1.48}$$

and the **orthogonality**

$$\int_{-1}^1 L_n(\xi)L_m(\xi)d\xi = \left\{\begin{array}{ll} \dfrac{2}{2n+1} & \text{if } n = m \\ 0 & \text{else} \end{array}\right\}. \tag{1.1.49}$$

Also

$$L_p(\xi) = \frac{L_{p+1}'(\xi) - L_{p-1}'(\xi)}{2p+1}\,, \quad p \geq 1\,. \tag{1.1.50}$$

A hierarchic sequence of reference element shape-functions $\widehat{N}_j(\xi)$ is then given by

$$\widehat{N}_0(\xi) = \frac{1+\xi}{2}, \quad \widehat{N}_1(\xi) = \frac{1-\xi}{2} \tag{1.1.51}$$

$$\widehat{N}_j(\xi) = \sqrt{\frac{2j-1}{2}} \int_{-1}^{\xi} L_{j-1}(\eta) d\eta, j \geq 2\,. \tag{1.1.52}$$

In the hierarchical basis (1.1.51), (1.1.52), we have

$$\widehat{\boldsymbol{S}} = \begin{pmatrix} \frac{1}{2} & -\frac{1}{2} & & & \\ -\frac{1}{2} & \frac{1}{2} & & & \\ & & 1 & & \\ & & & \ddots & \\ & & & & \ddots \\ & & & & & 1 \end{pmatrix}$$

and

$$\widehat{\boldsymbol{M}} = \begin{pmatrix} \frac{2}{3} & \frac{1}{3} & -\frac{1}{\sqrt{6}} & -\frac{1}{3\sqrt{10}} & 0 \\ & \frac{2}{3} & -\frac{1}{\sqrt{6}} & -\frac{1}{3\sqrt{10}} & 0 \\ & & \frac{2}{5} & 0 & -\frac{1}{5\sqrt{21}} \\ & & & \frac{2}{21} & 0 \\ & & & & \ddots \\ \text{sym} & & & & \frac{2}{45} \end{pmatrix}.$$

Note that for $p \geq 2$ by (1.1.50)

$$\widehat{N}_p(\xi) = \frac{1}{\sqrt{2(2p-1)}} (L_p - L_{p-2})(\xi)\,, \tag{1.1.53}$$

hence

$$\widehat{N}_p(\pm 1) = 0 \quad p \geq 2\,. \tag{1.1.54}$$

We observe that the shape-functions $\widehat{N}_0, \widehat{N}_1$ in (1.1.51) are identical to the ones constituting the hat functions, whereas the $\widehat{N}_j, j \geq 2$ in (1.1.52) vanish at the endpoints of $\widehat{K}$.

1.1.7 Global basis functions

A basis $\{b(x)\}$ of $S^p(D,\mathcal{T})$ is obtained again by assembling the element shape-functions $\underline{N}_K(x)$, given by

$$\underline{N}_K(F_K(\xi)) = \underline{\widehat{N}}(\xi)\,; \tag{1.1.55}$$

with $\widehat{N}_j$ as in (1.1.51), (1.1.52).

For the construction of global basis functions $\underline{b}(x)$ for $S^p(D,\mathcal{T})$, we distinguish two types of element shape-functions:

i) external shape-functions $\underline{N}_K^0(x)$:

$$\underline{N}_K^0\Big|_{\partial K} \neq 0\,.$$

ii) Internal shape-functions $\underline{N}_K^1(x)$:

$$\underline{N}_K^1\Big|_{\partial K} = 0\,.$$

Clearly, (1.1.51) are external, (1.1.52) are internal. Likewise, external and internal reference element shape-functions are defined, $\underline{\widehat{N}}^0$, $\underline{\widehat{N}}^1$.

External element shape-functions are assembled as before, into global basisfunctions, whereas internal element shape-functions $\underline{N}_K^1(x)$ are simply extended by zero to D to give a global basis function. We partition the vector of all global basisfunctions $\underline{b}(x)$ into external ones $\underline{b}^0(x)$ and internal ones $\underline{b}^1(x)$; likewise the elemental degrees of freedom $\underline{u}_K$ are partitioned: if $u_{\mathrm{FE}}(x) \in S^p(D,\mathcal{T})$ is a FE-function, for any $K \in \mathcal{T}$

$$u_{\mathrm{FE}}\,\Big|_K = \underline{u}_K^\top\,\underline{N}_K(x) = (\underline{u}_K^0)^\top\,\underline{N}_K^0(x) + (\underline{u}_K^1)^\top\,\underline{N}_K^1(x) \tag{1.1.56}$$

where $\underline{u}_K^0$ are external and $\underline{u}_K^1$ are internal element degrees of freedom.

1.1.8 Static Condensation

As the elemental unknowns $\underline{u}_K$ were partitioned into internal and external ones, $\underline{u}_K^1, \underline{u}_K^0$, we can also partition the element matrices $\boldsymbol{A}_K$:

$$\boldsymbol{A}_K = \begin{pmatrix} \boldsymbol{A}_K^{00} & \boldsymbol{A}_K^{10} \\ \boldsymbol{A}_K^{01} & \boldsymbol{A}_K^{11} \end{pmatrix} \tag{1.1.57}$$

and the element load vector

$$\underline{l}_K = \begin{pmatrix} \underline{l}_K^0 \\ \underline{l}_K^1 \end{pmatrix}\,. \tag{1.1.58}$$

The definition

$$u_{\mathrm{FE}} \in S^p(D,\mathcal{T}) : a(u_{\mathrm{FE}}, v) = l(v) \quad \forall v \in S^p(D,\mathcal{T}) \tag{1.1.59}$$

of the FE solution gives in particular for all $v \in \underline{N}_K^1$:

$$l(\underline{N}_K^1) = a(u_{\mathrm{FE}}, \underline{N}_K^1) = a_K(u_{\mathrm{FE}}, \underline{N}_K^1)\,, \tag{1.1.60}$$

or, in matrix form

$$\underline{l}_K^1 = \boldsymbol{A}_K^{01}\, \underline{u}_K^0 + \boldsymbol{A}_K^{11}\, \underline{u}_K^1\,. \tag{1.1.61}$$

Hence, if the 2 external unknowns $\underline{u}_K^0$ on $K \in \mathcal{T}$ are known, $\underline{u}_K^1$ can be determined from (1.1.61):

$$\underline{u}_K^1 := (\boldsymbol{A}_K^{11})^{-1}(\underline{l}_K^1 - \boldsymbol{A}_K^{01}\, \underline{u}_K^0)\,. \tag{1.1.62}$$

It remains to find the $\underline{u}_K^0$. Rewrite (1.1.59) as

$$\begin{aligned}
0 &= a(u_{\mathrm{FE}}, v) - l(v) \\
&= \sum_K a_K(u|_K, v|_K) - l_K(v|_k) \\
&= \sum_K \begin{pmatrix} \underline{v}_K^0 \\ \underline{v}_K^1 \end{pmatrix}^{\top} \left\{ \begin{pmatrix} \boldsymbol{A}_K^{00} & \boldsymbol{A}_K^{10} \\ \boldsymbol{A}_K^{01} & \boldsymbol{A}_K^{11} \end{pmatrix} \begin{pmatrix} \underline{u}_K^0 \\ \underline{u}_K^1 \end{pmatrix} - \begin{pmatrix} \underline{l}_K^0 \\ \underline{l}_K^1 \end{pmatrix} \right\} \\
&= \sum_K (\underline{v}_K^0)^{\top} \left\{ \boldsymbol{A}_K^{00}\, \underline{u}_K^0 + \boldsymbol{A}_K^{10}\, \underline{u}_K^1 - \underline{l}_K^0 \right\} \\
&\quad + (\underline{v}_K^1)^{\top} \left\{ \boldsymbol{A}_K^{11}\, \underline{u}_K^1 + \boldsymbol{A}_K^{01}\, \underline{u}_K^0 - \underline{l}_K^1 \right\}
\end{aligned}$$

for every $\underline{v}_K^0 \in \mathbb{R}^2$, $\underline{v}_K^1 \in \mathbb{R}^{p-1}$.

Inserting here (1.1.62), we find

$$\begin{aligned}
0 &= \sum_K (\underline{v}_K^0)^{\top} \left\{ \boldsymbol{A}_K^{00}\, \underline{u}_K^0 - \boldsymbol{A}_K^{10} (\boldsymbol{A}_K^{11})^{-1}\, \boldsymbol{A}_K^{01}\, \underline{u}_K^0 - (\underline{l}_K^0 - A_K^{10} (\boldsymbol{A}_K^{11})^{-1}\, \underline{l}_K^1) \right\} \\
&= \sum_K (\underline{v}_K^0)^{\top} \left\{ \widetilde{\boldsymbol{A}}_K^{00}\, \underline{u}_K^0 - \widetilde{\underline{l}}_K^0 \right\}
\end{aligned}$$

with the **condensed element matrices**

$$\widetilde{\boldsymbol{A}}_K^{00} := \boldsymbol{A}_K^{00} - \boldsymbol{A}_K^{10} \left(\boldsymbol{A}_K^{11}\right)^{-1} \boldsymbol{A}_K^{01} \tag{1.1.63}$$

and the **condensed element load vectors**

$$\widetilde{\underline{l}}_K^0 := \underline{l}_K^0 - \boldsymbol{A}_K^{10} (\boldsymbol{A}_K^{11})^{-1}\, \underline{l}_K^1\,. \tag{1.1.64}$$

The **global condensed stiffness matrix** $\widetilde{\boldsymbol{A}}$ is then obtained by assembling the $\widetilde{\boldsymbol{A}}_K^{00}$, $\widetilde{\underline{l}}_K^0$ as in the case $p = 1$.

1.1.9 Convergence Analysis for the FEM in $1-d$

Consider (1.1.33): find

$$u \in H_0^1(D): \quad a(u,v) = \ell(v) \quad \forall v \in H_0^1(D). \tag{1.1.65}$$

For $V_N = S_0^{1,1}(D, \mathcal{T})$, the FEM reads: find

$$u_N \in V_N: \quad a(u_N, v) = \ell(v) \quad \forall v \in V_N. \tag{1.1.66}$$

Since $V_N \subset V := H_0^1(D)$, we have the **Galerkin orthogonality** for the error $e_N := u - u_N$: for $v \in V_N$,

$$\begin{aligned} a(e_N, v) &= a(u-u_N, v) = a(u,v) - a(u_N, v) \\ &= \ell(v) - \ell(v) = 0 \quad \forall v \in V_N. \end{aligned} \tag{1.1.67}$$

Assume now in (1.1.33) that

- $a(x), c(x) \in L^\infty(D)$,
- $\operatorname{ess}\inf\limits_{x\in D} a(x) \geq \underline{a} > 0$, $\operatorname{ess}\inf\limits_{x\in D} c(x) \geq 0$.

Then from the Cauchy Schwarz Inequality, we get $\forall v, w \in V$:

$$\begin{aligned} |a(v,w)| &= \left| \int_D (a(x)\, v'\, w' + c(x)\, vw)\, dx \right| \\ &\leq \|a\|_{L^\infty(D)} \|v'\|_{L^2(D)} \|w'\|_{L^2(D)} \\ &+ \ \|c\|_{L^\infty(D)} \|v\|_{L^2(D)} \|w\|_{L^2(D)} \\ &\leq \underbrace{\max\left(\|a\|_{L^\infty}, \|c\|_{L^\infty}\right)}_{=C_1} \ \|v\|_{H^1(D)} \|w\|_{H^1(D)}, \end{aligned} \tag{1.1.68}$$

i.e. the **continuity** of $a(\cdot,\cdot)$: $V \times V \to \mathbb{R}$.

Since $V_N \subset V$ implies $e_N = u - u_N \in V$ that for $0 < \gamma := \frac{1}{2}\, \underline{a} \min(1, C(D))$ holds that

$$\begin{aligned} \|e_N\|^2_{H^1(D)} &\leq \frac{1}{\gamma}\, a(e_N, e_N) \\ &= \gamma^{-1}\, a(e_N, u - u_N) \\ &= \gamma^{-1} \Big[a(e_N, u - u_N) + \underbrace{a(e_N, v_N)}_{=0\ \forall v_N \in V_N} \Big] \\ &= \gamma^{-1} a(e_N, u - u_N + u_N) \\ &= \gamma^{-1}\, a(e_N, u - w_N) \quad \forall w_N \in V_N \\ &\leq \gamma^{-1}\, C_1\, \|e_N\|_{H^1(D)} \|u - w_N\|_{H^1(D)}, \end{aligned}$$

for $C_1 = \max\left(\|a\|_{L^\infty(D)}, \|c\|_{L^\infty(D)}\right)$.

Hence we have proved

$$\forall w_N \in V_N: \quad \|e_N\|_{H^1(D)} \leq \gamma^{-1}\, C_1\, \|u - w_N\|_{H^1(D)}\,.$$

Taking the infimum over both sides, we get the **quasioptimality** of e_N:

$$\|e_N\|_{H^1(D)} \leq \gamma^{-1}\, C_1 \inf_{w_N \in V_N} \|u - w_N\|_{H^1(D)}\,. \tag{1.1.69}$$

We next prove **a-priori error estimates**, i.e. estimates on $\|e_N\|_{H^1(D)}$ in terms of $N = \dim(V_N)$. They will be **asymptotic**, i.e. as $N = \dim(V_N) \to \infty$. To this end, we use (1.1.69) and choose $w_N = Iu$ where $I : V \to V_N$ is a projector: $Iu_N = u_N$, i.e. $IV_N = V_N$ and we bound the **interpolation error** $\|u - Iu\|_{H^1(D)}$.

To this end, the following **basic pattern** is used:

1) Definition of $\widehat{I}$ on the reference element $\widehat{K}$,
2) scaling of $\widehat{K}$ to $K \in \mathcal{T}$ to obtain local estimates and
3) adding the local estimates to obtain the global estimate on D.

Example 1.25. a) Let $\widehat{K} = (0,1)$, $\widehat{v} \in L^2(\widehat{K})$. Then the average operator

$$\widehat{I}\,\widehat{v} := \frac{1}{|\widehat{K}|} \int_{\widehat{K}} \widehat{v}(\xi)\, d\xi \in \mathbb{P}_0(\widehat{K})$$

maps $L^2(\widehat{K})$ into constants, i.e. polynomials of degree zero. It holds:

$$\forall \widehat{v} \in L^2(\widehat{K}): \ \|\widehat{I}\widehat{v}\|_{L^2(\widehat{K})} \leq \|\widehat{v}\|_{L^2(\widehat{K})}$$

i.e. $\widehat{I} : L^2(\widehat{K}) \to \mathbb{P}_0(\widehat{K}) \subset L^2(\widehat{K})$ is bounded and

$$\|\widehat{v} - \widehat{I}\,\widehat{v}\|^2_{L^2(\widehat{K})} = \|\widehat{v}\|^2_{L^2(\widehat{K})} - \|\widehat{I}\,\widehat{v}\|^2_{L^2(\widehat{K})} \leq \|\widehat{v}\|^2_{L^2(\widehat{K})}\,. \tag{1.1.70}$$

b) Let now $\widehat{w} \in H^1(\widehat{K})$. Then $\widehat{w} - \widehat{I}\,\widehat{w} = \widehat{w} - ⨍_{\widehat{K}} \widehat{w}$ has zero mean. Since the smallest positive eigenvalue λ_1 of the problem: find $u \in H^1(0,1)$, $u \neq 0$, and $\lambda \in \mathbb{C}$:

$$\begin{aligned} -u'' &= \lambda u \ \text{ in } \ (0,1) = \widehat{K}\,, \\ u'(0) &= u'(1) = 0 \end{aligned}$$

satisfies

$$\lambda_1 = \min_{\psi \in H^1(\widehat{K})/\mathbb{R}} \frac{\displaystyle\int_{\widehat{K}} (\psi')^2}{\displaystyle\int_{\widehat{K}} \psi^2} = \pi^2\,, \tag{1.1.71}$$

we have more generally:

Lemma 1.26. *Let* $\widehat{K} = (0,1)$. *Then*

i)
$$\inf_{\widehat{w}\in H_0^1(\widehat{K})} \frac{\|\widehat{w}'\|^2_{L^2(\widehat{K})}}{\|\widehat{w}\|^1_{L^2(\widehat{K})}} \geq \pi^2\,, \tag{1.1.72}$$

ii) *in* $H^1(\widehat{K})/\mathbb{R} := \left\{ w \in H^1(\widehat{K}) : \int_{\widehat{K}} w = 0 \right\}$ *holds*
$$\inf_{\widehat{w}\in H^1(\widehat{K})/\mathbb{R}} \frac{\|\widehat{w}'\|^2_{L^2(\widehat{K})}}{\|\widehat{w}\|^2_{L^2(\widehat{K})}} \geq \pi^2\,. \tag{1.1.73}$$

We will exploit (1.1.72), (1.1.73) to estimate the interpolation errors in $H^1(\widehat{K})$:

Lemma 1.27. *Define for* $\widehat{w} \in H^1(\widehat{K})$, *its projection* $\widehat{\Pi}^0$ *onto* $\mathbb{P}_0(\widehat{K})$, *i.e.*
$$\widehat{\Pi}^0(\widehat{w}) := \fint_{\widehat{K}} \widehat{w} = \int_0^1 \widehat{w}\, dx : \ L^2(\widehat{K}) \to \mathbb{P}_0(\widehat{K})\,.$$

Then there holds

i) $(\widehat{\Pi}^0)^2 = \widehat{\Pi}^0$ *on* $L^2(\widehat{K})$, *i.e.* $\widehat{\Pi}^0$ *is a projection,*

ii) $\|\widehat{\Pi}^0\, \widehat{v}\|_{L^2(\widehat{K})} \leq \|\widehat{v}\|_{L^2(\widehat{K})} \quad \forall \widehat{v} \in L^2(\widehat{K})$,

iii) $\|\widehat{v} - \widehat{\Pi}^0\, \widehat{v}\|_{L^2(\widehat{K})} \leq \dfrac{1}{\pi}\, \|\widehat{v}'\|_{L^2(\widehat{K})}$.

Proof. i) and ii) are immediate, and iii) follows from (1.1.73) if we put $\widehat{w} = \widehat{v} - \widehat{\Pi}^0\, \widehat{v} \in H^1(\widehat{K})/\mathbb{R}$. Next, we address the approximation in $H^1(\widehat{K})$. □

Lemma 1.28. *For* $v \in H^1(\widehat{K})$, *define*
$$(\widehat{\Pi}^1\, v)(x) = \widetilde{v}(0) + x \int_0^1 v' = \widetilde{v}(0) + x\, \widehat{\Pi}^0(v') \in \mathbb{P}_1(\widehat{K})\,. \tag{1.1.74}$$

Then

i) *on* $H^1(\widehat{K})$: $(\widehat{\Pi}^1)^2 = \widehat{\Pi}^1$,

ii) $\|(\Pi^1\, v)'\|_{L^2(\widehat{K})} = |\widehat{\Pi}^1\, v|_{H^1(\widehat{K})} \leq \|v'\|_{L^2(\widehat{K})} \quad \forall v \in H^1(\widehat{K})$,

iii) $\|(v - \widehat{\Pi}^1\, v)'\|_{L^2(\widehat{K})} \leq \dfrac{1}{\pi}\, \|v''\|_{L^2(\widehat{K})}$

iv) $\|v - \widehat{\Pi}^1\, v\|_{L^2(\widehat{K})} \leq \dfrac{1}{\pi}\, \|v'\|_{L^2(\widehat{K})}$

v) $\|v - \widehat{\Pi}^1\, v\|_{L^2(\widehat{K})} \leq \dfrac{1}{\pi^2}\, \|v''\|_{L^2(\widehat{K})}$.

Proof. i) is evident, ii) follows from (1.1.74); since

$$(\widehat{\Pi}^1 v)' = \int_0^1 v' = \widehat{\Pi}^0(v') \in \mathbb{P}_0(\widehat{K}),$$

and Lemma 1.27, ii) implies ii).

iii) follows from Lemma 1.27, iii):

$$\|(v - \widehat{\Pi}^1 v)'\|_{L^2(\widehat{K})} = \|v' - \widehat{\Pi}^0(v')\|_{L^2(\widehat{K})} \leq \frac{1}{\pi} \|v''\|_{L^2(\widehat{K})} .$$

iv) follows from $(\widehat{\Pi}^1 v)(1) = \widetilde{v}(0) + \int_0^1 v' = \widetilde{v}(1)$, implying $v - \widehat{\pi}^1 v \in H_0^1(\widehat{K})$, and from (1.1.72):

$$\begin{aligned} \|v - \widehat{\Pi}^1 v\|_{L^2(\widehat{K})} &\overset{(1.1.72)}{\leq} \frac{1}{\pi} \|(v - \widehat{\Pi}^1 v)'\|_{L^2(\widehat{K})} \\ &= \frac{1}{\pi} \|v' - (\widehat{\Pi}^1 v)'\|_{L^2(\widehat{K})} \\ &= \frac{1}{\pi} \|v' - \widehat{\Pi}^0 (v')\|_{L^2(\widehat{K})} . \end{aligned} \tag{1.1.75}$$

Now Lemma 1.27, ii) implies

$$\|v - \widehat{\Pi}^1 v\|_{L^2(\widehat{K})} \leq \frac{1}{\pi} \|v'\|_{L^2(\widehat{K})} ,$$

and (1.1.75) with Lemma 1.27, iii) implies for $v \in H^2(\widehat{K})$

$$\|v - \widehat{\Pi}^1 v\|_{L^2(\widehat{K})} \leq \frac{1}{\pi^2} \|v''\|_{L^2(\widehat{K})} ,$$

i.e. v). □

Remark 1.29. We can combine Lemma 1.28, iii) - v) into:

$$|v - \widehat{\Pi}^1 v|_{H^s(\widehat{K})} \leq \frac{1}{\pi^{t-s}} |v|_{H^t(\widehat{K})}, \; s = 0, 1, \; s \leq t \leq 2, \tag{1.1.76}$$

where the $H^s(\widehat{K})$-seminorm is understood as $L^2(\widehat{K})$ norm if $s = 0$.

Next, we scale (1.1.76) to a physical element $K_j = (x_{j-1}, x_j)$ of width $h_j = x_j - x_{j-1} > 0$.

Lemma 1.30. *Let $K_j = (x_{j-1}, x_j)$, $h_j := |K_j| = x_j - x_{j-1}$. Then for $v \in H^1(K_j)$, define*

$$\Pi^1(v) := \widetilde{v}(x_{j-1}) + (x - x_{j-1})\, h_j^{-1} \int_{x_{j-1}}^{x_j} v' \in \mathbb{P}_1(K_j) . \tag{1.1.77}$$

Then holds

$$
\begin{aligned}
&\text{i)} \quad \Pi^1(v)(x_{j-1}) = \widetilde{v}(x_{j-1}),\ \Pi^1(v)(x_j) = \widetilde{v}(x_j), \\
&\text{ii)} \quad v - \Pi^1(v) \in H_0^1(K_j), \\
&\text{iii)} \quad |v - \Pi^1 v|_{H^s(K_j)} \le \left(\frac{h_j}{\pi}\right)^{t-s} |v|_{H^t(K_j)},\, s = 0, 1 \le t \le 2\,.
\end{aligned}
\tag{1.1.78}
$$

Proof. Without loss of generality assume $x_{j-1} = 0$, $x_j = h_j$. Then change variables $x \longmapsto x\, h_j$ in (1.1.76) to obtain (1.1.78). □

Theorem 1.31. *Let $D = (0,1)$, $\mathcal{T} = \{K_j : j = 1, \dots, N\}$ a mesh in D. Then, for all $v \in H^1(D)$, define $\Pi_{\mathcal{T}}^1 v : H^1(D) \to S^{1,1}(D, \mathcal{T})$ by*

$$
(\Pi_{\mathcal{T}}^1 v)_{K_j} := \Pi^1(v|_{K_j}), \quad K_j \in \mathcal{T}\,. \tag{1.1.79}
$$

Then

$$
\begin{aligned}
&\text{i)} \quad (\Pi_{\mathcal{T}}^1 v)(x_j) = \widetilde{v}(x_j), \quad j = 0, 1, \dots, N\,, \\
&\text{ii)} \quad |v - \Pi_{\mathcal{T}}^1 v|_{H^s(D)}^2 \le \sum_{j=1}^{N} \left(\frac{h_j}{\pi}\right)^{2(t-s)} |v|_{H^t(K_j)}^2\,,
\end{aligned}
$$

for $s = 0, 1 \le t \le 2$.

Proof. Square (1.1.78) and sum over $j = 1, \dots, N$ to get ii), using (1.1.79). □

Corollary 1.32. *If $\mathcal{T}$ is uniform, i.e. $h_j = h, j = 1, \dots, N$, we get*

$$
\begin{aligned}
\|v - \Pi_{\mathcal{T}}^1 v\|_{L^2(D)} &\le \frac{h^2}{\pi^2}\, \|v''\|_{L^2(D)}\,, \\
|v - \Pi_{\mathcal{T}}^1 v|_{H^1(D)} &\le \frac{h}{\pi}\, \|v''\|_{L^2(D)}\,.
\end{aligned}
\tag{1.1.80}
$$

Corollary 1.33. *Let $h(x) \in S^{1,1}(D, \mathcal{T})$ be such that*

$$
h(x_j) = h_j + h_{j+1}, \quad j = 1, \dots, N,\ h_{n+1} := 0\,,
$$

where $h_j := x_j - x_{j-1}$, $j = 1, \dots, N$.

Then

$$
h(x) \ge h_j \ \ \textit{for}\ \ x \in [x_{j-1}, x_j] \tag{1.1.81}
$$

and

$$
\|v - \Pi_{\mathcal{T}}^1 v\|_{H^1(D)}^2 \le \frac{1}{\pi^2}\, \|h v''\|_{L^2(D)}^2\,. \tag{1.1.82}
$$

1.2 Linear Finite Elements in Two Dimensions

We consider now two and three dimensional domains $D \subset \mathbb{R}^n$ with $n = 2, 3$. We proceed exactly as in the one dimensional case – in fact, all the concepts presented there have direct analogs in $\mathbb{R}^n$. We start with the definition of the Sobolev spaces needed for the variational formulation and we discuss then the simplest and most used finite elements.

1.2.1 Sobolev Spaces in $\mathbb{R}^n$, $n > 1$

Surface Integrals Let $n > 1$ and assume that $D \subset \mathbb{R}^n$ is an open bounded polyhedral domain in $\mathbb{R}^n$, i.e. its boundary $\Gamma = \partial D$ consists of finitely many open pieces of $n-1$ dimensional hyperplanes Γ_j and a set of vertices S (in the two dimensional case), or of vertices and edges (in the three dimensional case), respectively. Each surface piece Γ_j is parameterized by an affine mapping $\underline{\chi}_j : \hat{\Gamma}_j \to \Gamma_j$, where the domain $\hat{\Gamma}_j$ is an interval for $n = 2$ or a polygon for $n = 3$. In the latter case, the surface integral of f over Γ_j is then

$$\int_{\Gamma_j} f(\underline{x})\, d\underline{s}_{\underline{x}} := \int_{\hat{\Gamma}_j} f\big(\underline{\chi}_j(\underline{\xi})\big)\sqrt{g_j}\, d\underline{\xi},$$

where $\sqrt{g_j}$ denotes the surface element: let $\boldsymbol{G}_j(\underline{\xi})$ denote the Gram-matrix of $\underline{\chi}_j$:

$$\boldsymbol{G}_j(\underline{\xi}) = \left(\frac{\partial(\underline{\chi}_j)_k}{\partial \xi_\ell}\right)_{\substack{1 \le k \le 3 \\ 1 \le \ell \le 2}} \in \mathbb{R}^{3\times 2}$$

then

$$g_j(\underline{\xi}) = \det\left(\boldsymbol{G}_j^\top \boldsymbol{G}_j\right)(\underline{\xi}).$$

Since the surface measure of S is zero, it holds (all surface integrals are understood as Lebesgue integrals):

$$\int_{\Gamma} f(\underline{x})\, d\underline{s}_{\underline{x}} := \sum_j \int_{\Gamma_j} f(\underline{x})\, d\underline{s}_{\underline{x}}. \tag{1.2.1}$$

We say also that S is a **zero surface measure set**.

Derivatives and Spaces of continuous functions:
Let $\underline{\alpha} \in \mathbb{N}_0^n$ be a **multi-index**, i.e., a vector of n non-negative integers:

$$\underline{\alpha} = (\alpha_1, \ldots, \alpha_n) \in \mathbb{N}_0^n.$$

Set $|\underline{\alpha}| := \alpha_1 + \cdots + \alpha_n$ and denote by

$$\partial^{\underline{\alpha}} := \frac{\partial^{\alpha_1}}{\partial x_1^{\alpha_1}} \cdots \frac{\partial^{\alpha_n}}{\partial x_n^{\alpha_n}}$$

the partial derivative of order $|\underline{\alpha}|$. For any integer $k \geq 0$ we define

$$C^k(\overline{D}) := \left\{u : \partial^{\underline{\alpha}} u \in C^0(\overline{D}),\ \forall\, |\underline{\alpha}| \leq k\right\}$$

and

$$C_0^k(\overline{D}) := \left\{u \in C^k(\overline{D}) : \partial^{\underline{\alpha}} u|_{\Gamma} = 0,\ \forall\, |\underline{\alpha}| \leq k\right\}.$$

$C^k(\overline{D})$ is a normed linear space with norm

$$\|u\|_{C^k(\overline{D})} := \sum_{|\underline{\alpha}| \leq k} \|\partial^{\underline{\alpha}} u\|_{C^0(\overline{D})},$$

where

$$\|u\|_{C^0(\overline{D})} := \max_{x \in \overline{D}} \{|u(x)|\},$$

and $C_0^k(\overline{D}) \subset C^k(\overline{D})$ is a closed linear subspace.

Weak derivatives and Sobolev Spaces $H^k(D)$:

$$L^2(D) := \left\{u : D \to \mathbb{R} \text{ measurable}, \int_D |u(\underline{x})|^2 \, d\underline{x} < \infty\right\}$$

is a normed linear space with norm

$$\|u\|_0 := \Big(\int_D |u(\underline{x})|^2 \, d\underline{x}\Big)^{\frac{1}{2}}.$$

We use the notations

$$(u, v)_0 := \int_D u(\underline{x}) v(\underline{x}) \, d\underline{x} \text{ for } u, v \in L^2(D),$$

$$(\underline{u}, \underline{v})_0 := \sum_{i=1}^{m} (u_i, v_i)_0 \text{ for } \underline{u}, \underline{v} \in L^2(D)^m.$$

Definition 1.34. *Let $u \in L^2(D)$ and $\underline{\alpha} \in \mathbb{N}_0^n$. A function $v \in L^2(D)$ is called the **weak derivative** of order $|\underline{\alpha}| = k$ of u, if*

$$(v, \varphi)_0 = (-1)^k (u, \partial^{\underline{\alpha}} \varphi)_0 \quad \forall \varphi \in C_0^k(\overline{D}).$$

Definition 1.35 (Sobolev spaces)**.** *Let $k \geq 1$. Then*

$$H^k(D) := \left\{u \in L^2(D) : \partial^{\underline{\alpha}} u \in L^2(D), \forall\, |\underline{\alpha}| \leq k\right\}$$

is a normed, linear space with norm

$$\|u\|_k := \Big(\sum_{|\underline{\alpha}| \leq k} \|\partial^{\underline{\alpha}} u\|_0^2\Big)^{\frac{1}{2}},$$

and we set

$$(u, v)_k := \sum_{|\underline{\alpha}| \leq k} (\partial^{\underline{\alpha}} u, \partial^{\underline{\alpha}} v)_0 \text{ for } u, v \in H^k(D).$$

A vector function $\underline{u} : D \to \mathbb{R}^m$ is in $H^k(D)^m \iff u_i \in H^k(D)$ for all $i = 1, \dots, m$.

Remark 1.36. Boundary Values of Sobolev functions. Traces.: For $u \in C^0(\overline{D})$, we denote by $u|_\Gamma$ its *restriction to* Γ, the *trace of* u *on* Γ, i.e. the function $u(\underline{x})$ for $\underline{x} \in \Gamma$. The trace of $u \in C^0(\overline{D})$ is again a continuous function on Γ and the following trace inequality holds true:

$$\|u|_\Gamma\|_{C^0(\Gamma)} \le \|u\|_{C^0(\overline{D})}.$$

If $u \in H^0(D) = L^2(D)$, this is no longer true (see also Remark 1.12).

Remark 1.37. For $u \in L^2(D)$ the trace $u|_\Gamma$ (resp. the trace operator) is in general not defined. In particular, a trace inequality of the form

$$\|u|_\Gamma\|_{L^2(\Gamma)} \le C\,\|u\|_{L^2(D)} \tag{1.2.2}$$

is false. Indeed, let $D = (0,1)^2$ and, for $0 < h < 1$, let

$$\varphi(x_1, x_2) := \begin{cases} 0 & h \le x_1 \le 1, 0 \le x_2 \le 1, \\ 1 - {}^{x_1}\!/_h & 0 \le x_1 \le h, 0 \le x_2 \le 1. \end{cases}$$

Then it holds that

$$1 = \int_0^1 |\varphi(0, x_2)|^2\, dx_2 \le \|\varphi|_\Gamma\|^2_{L^2(\Gamma)}$$

and

$$\|\varphi\|^2_{L^2(D)} = \int_0^h \Big(1 - \frac{x_1}{h}\Big)^2 dx_1 \overset{y = x_1/h}{=} h \int_0^1 (1-y)^2\, dy = {}^h\!/_3.$$

If (1.2.2) were true, then there would exist a constant $C > 0$ such that $1 \le \frac{Ch}{3}$. For $h \to 0$ we obtain a contradiction.

A trace inequality can only be obtained if the order of the norm in the right hand side of (1.2.2) is increased. It holds

Theorem 1.38 (Multiplicative trace inequality)**.** *Let $D \subset \mathbb{R}^n$, $n \ge 2$, be a convex polyhedral domain. The trace mapping $\gamma_0 : u \mapsto u|_\Gamma$ is continuous from $H^1(D)$ into $L^2(\Gamma)$:*

$$\|\gamma_0 u\|_{L^2(\Gamma)} \le C(D)\, \|u\|_{H^1(D)} \qquad \forall u \in H^1(D).$$

More precisely, the following multiplicative trace inequality holds:

$$\|\gamma_0 u\|^2_{L^2(\Gamma)} \le C(D) \big\{ \|u\|^2_{L^2(D)} + \|u\|_{L^2(D)} \|\nabla u\|_{L^2(D)} \big\}. \tag{1.2.3}$$

Proof. Denote by $\widetilde{x}$ the center of the largest $n-$dimensional ball inscribed into D and by ρ_D its radius. Without loss of generality, we suppose that $\widetilde{x}$ is the origin of the coordinate system. We start from the following relation

$$\int_{\partial D} v^2 \underline{x} \cdot \underline{n}\, dS = \int_D \nabla \cdot (v^2 \underline{x})\, dx, \qquad v \in H^1(D). \tag{1.2.4}$$

Let n_j be the outer unit normal to D on the edge Γ_j, $j \in S$. Then

$$\underline{x} \cdot n_j = |\underline{x}||n_j| \cos\alpha = |\underline{x}| \cos\alpha = t_j, \quad \underline{x} \in \Gamma_j,\ j \in S, \tag{1.2.5}$$

where t_j is the distance from $\widetilde{x}$ to Γ_j, see Figure 1.6. Obviously,

$$t_j \geq \rho_D, \quad \forall j \in S. \tag{1.2.6}$$

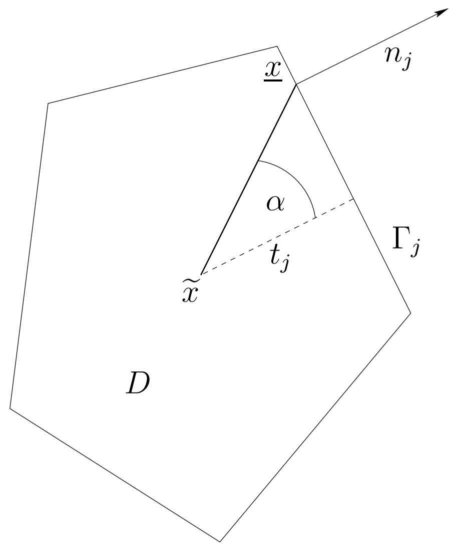


Figure 1.6: Distances t_j.

From (1.2.5) and (1.2.6) we have

$$\begin{aligned}&\int_{\partial D} v^2 \underline{x} \cdot \underline{n}\, dS = \sum_{j \in S} \int_{\Gamma_j} v^2 \underline{x} \cdot n_j \, dS = \sum_{j\in S} t_j \int_{\Gamma_j} v^2\, dS \\ &\geq \rho_D \sum_{j \in S} \int_{\Gamma_j} v^2 \, dS = \rho_D \|v\|^2_{L^2(\partial D)}.\end{aligned} \tag{1.2.7}$$

Moreover,

$$\begin{aligned}&\int_D \nabla \cdot (v^2 \underline{x})\, dx = \int_D \left(v^2 \nabla \cdot \underline{x} + \underline{x} \cdot \nabla v^2\right) dx \\ &= n \int_D v^2\, dx + 2\int_D v\underline{x} \cdot \nabla v \, dx \leq n\|v\|^2_{L^2(D)} + 2\int_D |v\underline{x} \cdot \nabla v|\ dx.\end{aligned} \tag{1.2.8}$$

With the Cauchy inequality the second term in the right hand side of (1.2.8) is estimated as

$$2\int_D |v\underline{x} \cdot \nabla v|\, dx \leq 2 \sup_{\underline{x}\in D} |\underline{x}| \int_D |v||\nabla v|\, dx \leq 2h_D \|v\|_{L^2(D)} |v|_{H^1(D)}. \tag{1.2.9}$$

Then $h_D/\rho_D \le C_1$, (1.2.4), (1.2.7), (1.2.8) and (1.2.9) give

$$\begin{aligned} \|v\|^2_{L^2(\partial D)} &\le \frac{1}{\rho_D}\left[2h_D\|v\|_{L^2(D)}|v|_{H^1(D)} + n\|v\|^2_{L^2(D)}\right] \\ &\le C_1\left[2\|v\|_{L^2(D)}|v|_{H^1(D)} + \frac{n}{h_D}\|v\|^2_{L^2(D)}\right], \end{aligned} \tag{1.2.10}$$

which yields (1.2.3) with $C = C_1 \max\{2, n\}$. $\square$

The trace result Theorem 1.38 allows us now to perform a partial integration on a polygonal / polyhedral D: The following **Green formula** holds

$$\int_D u \operatorname{div} \underline{v}\, d\underline{x} = -\int_D \underline{v} \cdot \underline{\nabla} u\, d\underline{x} + \int_\Gamma u\underline{v} \cdot \underline{n}\, d\underline{s}, \tag{1.2.11}$$

that is

$$(u, \operatorname{div} \underline{v})_{0,D} = -(\underline{v}, \underline{\nabla} u)_{0,D} + \langle u, \underline{v} \cdot \underline{n}\rangle_{0,\Gamma} \tag{1.2.12}$$

for $u \in H^1(D)$ and $\underline{v} \in H^1(D)^n$. The boundary integral above is well-defined, since by the trace Theorem 1.38 and since $|\underline{n}| = 1$ on $\Gamma \setminus S$ where S denotes the set of edges and vertices of Γ and $\int_S d\underline{s} = 0$ we have that:

$$\left|\int_\Gamma u\underline{v} \cdot \underline{n}\, d\underline{s}\right| \le \Big(\int_\Gamma |u|^2 \, d\underline{s}\Big)^{\frac{1}{2}} \Big(\int_\Gamma |\underline{v}|^2 \, d\underline{s}\Big)^{\frac{1}{2}} \le C(D)\, \|u\|_{1,D}\, \|\underline{v}\|_{1,D} < \infty.$$

In Theorem 1.38 we considered only boundary values on the whole boundary Γ. However, the statement in Theorem 1.38 remains true if, instead of the whole boundary, we take only a part $\Gamma_i \subset \Gamma$.

Remark 1.39.

1. For each side $\Gamma_i \subset \Gamma$ it holds: the trace operator γ_0 is continuous from $H^1(D) \to L^2(\Gamma_i)$ and
$$\|\gamma_0 u\|_{L^2(\Gamma_i)} \le C(D)\, \|u\|_{1,D} \quad \forall u \in H^1(D).$$

2. The trace operator $\gamma_0 : H^1(D) \to L^2(\Gamma)$ is not surjective, i. e. $\gamma_0(H^1(D)) \subsetneqq L^2(\Gamma)$. $\gamma_0(H^1(D))$ is a linear subspace of $L^2(\Gamma)$ and is denoted by $H^{\frac{1}{2}}(\Gamma)$. A norm on $H^{\frac{1}{2}}(\Gamma)$ is given by
$$\|u\|_{H^{\frac{1}{2}}(D)} := \inf_{\substack{v \in H^1(D) \\ \gamma_0 v = u}} \|v\|_{H^1(D)}\,.$$
$\gamma_0(H^1_0(D))$ is closed with respect to $\|u\|_{H^{\frac{1}{2}}(\partial D)}$.

Exercise 1.40. Verify the norm axioms (N1)–(N3) in Definition 1.3 for $H^{\frac{1}{2}}(\Gamma)$.

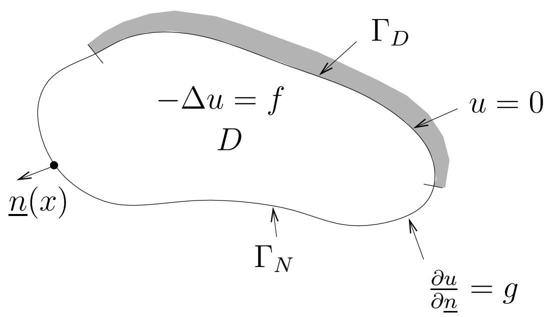


Figure 1.7: Boundary conditions on Γ_D and Γ_N.

1.2.2 Variational Formulation for the Poisson Problem

We consider the Poisson equation

$$-\Delta u = -\operatorname{div}(\underline{\nabla} u) = f \text{ in } D \tag{1.2.13}$$

with the following boundary conditions:

$$u = 0 \text{ on } \Gamma_D \text{ (Dirichlet boundary conditions)} \tag{1.2.14}$$

and

$$\frac{\partial u}{\partial \underline{n}} = \underline{\sigma} \cdot \underline{n} = g \text{ on } \Gamma_N \text{ (Neumann boundary conditions),} \tag{1.2.15}$$

where $\Gamma = \partial D$ is partitioned into Γ_D and Γ_N (see also Figure 1.7). We assume here that

$$|\Gamma_D| = \int_{\Gamma_D} d\underline{s} > 0,$$

i. e. the length of the Dirichlet boundary Γ_D in (1.2.14) is strictly positive.

In order to obtain the variational formulation we multiply (1.2.13) with a test function $v \in H^1(D)$ and integrate by parts. Then, from (1.2.12) it follows that:

$$\int_D f v \, d\underline{x} = -\int_D v \Delta u \, d\underline{x} = \int_D \underline{\nabla} u \cdot \underline{\nabla} v \, d\underline{x} - \int_\Gamma v \frac{\partial u}{\partial \underline{n}} \, d\underline{s}.$$

We plug in the boundary conditions (1.2.14) and (1.2.15) and get

$$\int_\Gamma v \frac{\partial u}{\partial \underline{n}} \, d\underline{s} = \int_{\Gamma_D} v \frac{\partial u}{\partial \underline{n}} \, d\underline{s} + \int_{\Gamma_N} v g \, d\underline{s}.$$

For $u \in H^1(D)$, $\underline{\nabla} u \in L^2(D)^n$, and therefore $\underline{n} \cdot \underline{\nabla} u|_\Gamma = \left.\frac{\partial u}{\partial \underline{n}}\right|_\Gamma$ is not defined (see also Remark 1.37). We prescribe therefore the **essential boundary conditions** (1.2.14) also for v: $v|_{\Gamma_D} = 0$.

This leads to the following **variational formulation** of (1.2.13)–(1.2.15):
Find $u \in V = \{u \in H^1(D) : \gamma_0(u)|_{\Gamma_D} = 0\}$ such that

$$a(u,v) = l(v) \quad \forall v \in V$$

where

$$\begin{aligned} a(u,v) &= \int_D \underline{\nabla} u \cdot \underline{\nabla} v \, d\underline{x} = (\underline{\nabla} u, \underline{\nabla} v)_{0,D} \\ l(v) &= \int_D f v \, d\underline{x} + \int_{\Gamma_N} g v \, d\underline{s} = (f,v)_{0,D} + \langle v, g \rangle_{0,\Gamma_N} . \end{aligned} \tag{1.2.16}$$

From Remark 1.39 we get that $V \subset H^1(D)$ is a closed, linear subspace.

The bilinear form $a(.,.) : V \times V \to \mathbb{R}$ and the linear form $l(.) : V \to \mathbb{R}$ are well-defined: For all $u, v \in H^1(D)$ and for $f \in L^2(D)$ and $g \in L^2(\Gamma_D)$ there holds

$$|a(u,v)| = |(\underline{\nabla} u, \underline{\nabla} v)_0| \leq \|\underline{\nabla} u\|_D \, \|\underline{\nabla} v\|_D \leq \|u\|_1 \, \|v\|_1 < \infty$$

and

$$\begin{aligned} |l(v)| &\leq |(f,v)_0| + \left| \langle g, v \rangle_{0,\Gamma_D} \right| \leq \|f\|_{0,D} \, \|v\|_{0,D} + \|g\|_{0,\Gamma_N} \, \|\gamma_0 v\|_{0,\Gamma_N} \\ &\leq \|f\|_{0,D} \, \|v\|_{1,D} + \|g\|_{0,\Gamma_N} \, \|\gamma_0 v\|_{0,\Gamma_N} \\ &\overset{(1.39)}{\leq} \left(\|f\|_{0,D} + C(D) \, \|g\|_{0,\Gamma_N} \right) \|v\|_{1,D} < \infty. \end{aligned}$$

The **FE discretization** of (1.2.16) follows now similarly to the one-dimensional case: Let $V_N \subset V$ be a linear subspace and $N = \dim(V_N) < \infty$. A finite element-approximation u_N of $u \in V$ is defined by:
Find $u_N \in V_N$ such that

$$a(u_N, v) = l(v) \quad \forall v \in V_N. \tag{1.2.17}$$

We describe several FE-spaces V_N that are most common in practice. As in the one dimensional case the starting point of a **Finite Element** algorithm is based on

- a reference element $\widehat{K}$,
- an element mapping $F_K : \widehat{K} \to K \in \mathcal{T}$ and
- reference element shape-functions $\widehat{\underline{N}}$.

1.2.3 Linear Finite Elements for the Poisson Problem

We consider first two-dimensional problems, i. e. $n = 2$. Let us further assume that $D \subset \mathbb{R}^2$ is a polygonal domain. Similarly to the one-dimensional case we triangulate D in finitely many elements.

Definition 1.41 (Regular mesh)**.** *A mesh $\mathcal{T} = \{K\}$ on the polygon $D \subset \mathbb{R}^2$ is a partition of D into open triangles K. $\mathcal{T}$ is called **regular**, if the intersection of two different triangles $\overline{K}$, $\overline{K'} \in \mathcal{T}$ is the empty set, a vertex or a whole edge.*

We define the FE-space of piecewise linear continuous functions with respect to a regular mesh $\mathcal{T}$ consisting of triangles:

$$S^1(D, \mathcal{T}) = \{u \in C^0(D) \,|\, u(x)|_K = a_K + b_K x_1 + c_K x_2 \quad \forall K \in \mathcal{T}\}. \tag{1.2.18}$$

Proposition 1.42 (Properties of $S^1(D, \mathcal{T})$)**.**

- *$S^1(D, \mathcal{T}) \subset H^1(D)$,*
- *$u \in S^1(D, \mathcal{T})$ is uniquely determined by the nodal values $u(P)$ in the vertices $P \in \mathcal{N}(\mathcal{T})$, the set of all vertices of $\mathcal{T}$,*
- *$N = \dim S^1(D, \mathcal{T}) = |\mathcal{N}(\mathcal{T})| < \infty$,*
- *$S^1(D, \mathcal{T}) = \operatorname{span}\{b_P(x) : P \in \mathcal{N}(\mathcal{T})\}$, where the **hat basis functions** are defined by*

$$b_P \in S^1(D, \mathcal{T}), b_P(\underline{P'}) = \begin{cases} 1 & \text{for } P' = P, \\ 0 & \text{for } P' \neq P \in \mathcal{N}(\mathcal{T}) \end{cases} \tag{1.2.19}$$

(see Figure 1.8).

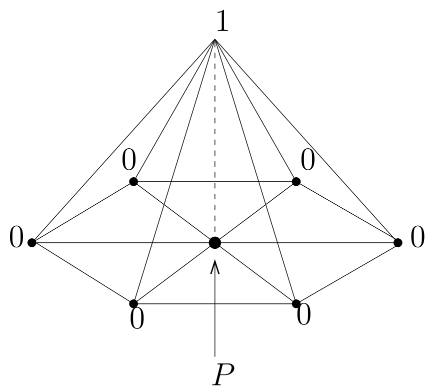


Figure 1.8: FE basis function $b_P(x)$.

Let $\underline{b}$ be the vector of all basis functions b_P (with respect to an arbitrary, but fixed numbering of the vertices $P \in \mathcal{N}(\mathcal{T})$):

$$\underline{b} = \{b_P : P \in \mathcal{N}(\mathcal{T})\}.$$

Again, we can write an arbitrary FE function $v \in S^1(D, \mathcal{T})$ uniquely as

$$v(x) = \sum_{P \in \mathcal{N}(\mathcal{T})} v(\underline{P})\, b_P(x) = \underline{v}^\top\, \underline{b}(x). \tag{1.2.20}$$

In particular, for the FE solution u_N it holds:

$$u_N(x) = \sum_{P \in \mathcal{N}(\mathcal{T})} u_N(\underline{P})\, b_P(x) = \underline{u}_N^\top\, \underline{b}(x). \tag{1.2.21}$$

The vector $\underline{u}_N \in \mathbb{R}^N$ is again solution of a linear (algebraic) system of equations (1.1.19). The global stiffness matrix $\boldsymbol{A}$ is obtained as before via an assembly procedure from the element stiffness matrices. It holds

$$\boldsymbol{A} = \{a(b_P, b_{P'}) : P, P' \in \mathcal{N}(\mathcal{T})\}. \tag{1.2.22}$$

For $K \in \mathcal{T}$, the **element shape-functions** are restrictions to K of those global basis functions $b_P(x)$ that are not identically zero in K:

$$\underline{N}_K(x) = \{b_P(x)|_K : b_P(x) \neq 0 \text{ for } x \in K\}. \tag{1.2.23}$$

For each triangle $K \in \mathcal{T}$ there are exactly three element shape-functions, and the element stiffness matrix is defined by

$$\boldsymbol{A}_K = \{a_K(b_P, b_{P'}) = \int_K \nabla b_P \cdot \nabla b_{P'} dx :\ P, P' \in \mathcal{N}(K)\}.$$

In order to compute the matrix $\boldsymbol{A}_K$ we first transform K to the reference element $\widehat{K}$ (see Figure 1.9). Each triangle K of a regular triangular mesh $\mathcal{T}$ of D is the image of the **reference element**

$$\widehat{K} = \{\underline{\xi} = (\xi_1, \xi_2) : 0 < \xi_1 < 1, 0 < \xi_2 < 1 - \xi_1\} \tag{1.2.24}$$

by an affine mapping $\underline{F}_K$: Let $\underline{P}_0$, $\underline{P}_1$, $\underline{P}_2$ be the vectors associated with the vertices P_0, P_1, P_2 of K (see
Figure 1.9) and let

$$\underline{x} = \underline{F}_K(\underline{\xi}) = \underline{P}_0 + \xi_1(\underline{P}_1 - \underline{P}_0) + \xi_2(\underline{P}_2 - \underline{P}_0) = \underline{b}_K + \boldsymbol{B}_K\, \underline{\xi} \tag{1.2.25}$$

be the affine element mapping with

$$\underline{b}_K = \underline{P}_0, \boldsymbol{B}_K = [\underline{P}_1 - \underline{P}_0, \underline{P}_2 - \underline{P}_0], K \in \mathcal{T}.$$

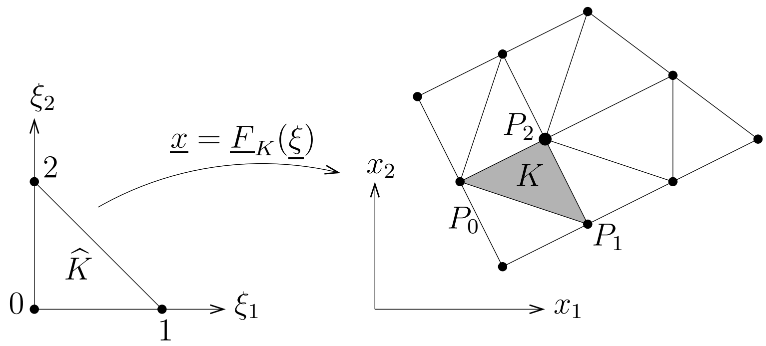


Figure 1.9: Numbering of regular mesh $\mathcal{T}$ of D and reference element $\widehat{K}$.

All derivatives are transformed to the reference coordinates in $\widehat{K}$. With the notation

$$\nabla = \left(\frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}\right)^\top, \quad \widehat{\nabla} = \left(\frac{\partial}{\partial \xi_1}, \frac{\partial}{\partial \xi_2}\right)^\top$$

and by using the chain rule we have (check!)

$$\nabla = \boldsymbol{B}_K^{-\top}\,\widehat{\nabla}, \tag{1.2.26}$$

and from (1.2.25) it follows that

$$dx = \det(\boldsymbol{B}_K)\,d\xi,\ x \in K,\ K \in \mathcal{T}.$$

As in the one-dimensional case $\underline{N}_K(\underline{F}_K(\underline{\xi})) = \widehat{\underline{N}}(\xi)$ with the (independent of K) **reference element shape-functions**

$$\widehat{N}_0(\xi) = 1 - \xi_1 - \xi_2, \quad \widehat{N}_i(\underline{\xi}) = \xi_i, i = 1, 2. \tag{1.2.27}$$

With the notation in Figure 1.9 for $u, v \in S^1(D, \mathcal{T})$ we obtain

$$a_K(b_{P_i}, b_{P_j}) = \int_K \nabla b_{P_i} \cdot \nabla b_{P_j}\,dx = \int_{\widehat{K}} (\widehat{\nabla}\widehat{N}_i)^\top\,\boldsymbol{B}_K^{-1}\,\boldsymbol{B}_K^{-\top}\,\widehat{\nabla}\widehat{N}_j|\boldsymbol{B}_K|d\xi, \tag{1.2.28}$$

$i, j = 0, 1, 2$. This integral is taken over the reference element $\widehat{K}$ and in general it has to be evaluated numerically. Since here however $\boldsymbol{B}_K$ is independent of $\underline{\xi}$ and since

$$\widehat{\nabla}\widehat{N}_0 = \begin{pmatrix}-1\\-1\end{pmatrix}, \widehat{\nabla}\widehat{N}_1 = \begin{pmatrix}1\\0\end{pmatrix}, \widehat{\nabla}\widehat{N}_2 = \begin{pmatrix}0\\1\end{pmatrix}$$

are constant, we can compute $\boldsymbol{A}_K$ exactly (details are given below):

$$\boldsymbol{A}_K = \frac{1}{4|K|}\begin{pmatrix}|\underline{P}_1 - \underline{P}_2|^2 & (\underline{P}_1 - \underline{P}_2)\cdot(\underline{P}_2 - \underline{P}_0) & (\underline{P}_1 - \underline{P}_2)\cdot(\underline{P}_0 - \underline{P}_1)\\ & |\underline{P}_2 - \underline{P}_0|^2 & (\underline{P}_2 - \underline{P}_0)\cdot(\underline{P}_0 - \underline{P}_1)\\ \text{sym.} & & |\underline{P}_1 - \underline{P}_0|^2\end{pmatrix}. \tag{1.2.29}$$

It remains to assemble the element stiffness matrices $\boldsymbol{A}_K$:

$$\boldsymbol{A} = \mathcal{A}_{K\in\mathcal{T}}\,\boldsymbol{A}_K.$$

The definition of the **assembly operator** $\mathcal{A}_{K\in\mathcal{T}}$ is the following: The nodes $P \in \mathcal{N}(\mathcal{T})$ are numbered globally from $1,\ldots,N$. For a triangle element $K \in \mathcal{T}$ we denote by $\boldsymbol{T}_K$ the following $N \times 3$ matrix: $(\boldsymbol{T}_K)_{ij} = 1$, if the node P_j in element K (see Figure 1.9) corresponds to a global index i, otherwise $(\boldsymbol{T}_K)_{ij} = 0$. With this we obtain an explicit representation of the **assembly operator**

$$\boldsymbol{A} = \sum_K \boldsymbol{T}_K^\top \boldsymbol{A}_K \boldsymbol{T}_K = \mathcal{A}_{K\in\mathcal{T}}\,\boldsymbol{A}_K. \tag{1.2.30}$$

A more detailed explanation of the assembly process is given in Section 1.6 below.

1.3 Implementation

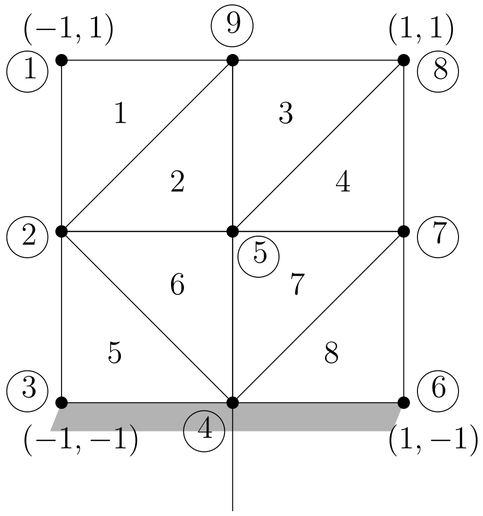


Figure 1.10: Example of a triangulation in $D = (-1,1)^2$.

We describe the implementation of Algorithm 1.1 in the two-dimensional case. Let D be a polygon and let $\mathcal{T}$ be a regular mesh consisting of triangles K. We illustrate the ideas with the following example: consider the Poisson problem (1.2.13) in $D = (-1,1)^2$ with Dirichlet-boundary $\Gamma_D = \{(x_1,x_2) : x_2 = -1\}$ (Γ_D is outlined in Figure 1.10 by the shaded region).

i	P_i	
1	-1	1
2	-1	0
3	-1	-1
4	0	-1
5	0	0
6	1	-1
7	1	0
8	1	1
9	0	1

K_j	P_{1j}	P_{2j}	P_{3j}
1	1	2	9
2	2	5	9
3	5	8	9
4	5	7	8
5	3	4	2
6	4	5	2
7	4	7	5
8	4	6	7

Table 1.1: Contents of the files `coord.dat` and `elem.dat` for $\mathcal{T}$ in Figure 1.10.

edge number	vertex indices	
1	3	4
2	4	6

edge number	vertex indices	
1	6	7
2	7	8
3	8	9
4	9	1
5	1	2
6	2	3

Table 1.2: `Diri.dat` and `Neum.dat` for Figure 1.10.

1.3.1 Geometry Data

A regular triangular mesh $\mathcal{T}$ in a polygon D will be described by two lists: the **coordinate list** `coord.dat` and the **element list** `elem.dat`. `coord.dat` contains the coordinates of the vertices $P \in \mathcal{N}(T)$. For the example in Figure 1.10 we have: $N = |\mathcal{N}(T)| = 9$ and the data in Table 1.1.

Each element $K \in \mathcal{T}$ is uniquely determined by its three vertices. This correspondence is realized in the element list. The ith line of `elem.dat` consists of the three vertices of element i, i. e. `elem.dat` corresponds to the matrix $\boldsymbol{T}_K$. The three vertices of the triangle $T \in \mathcal{T}$ are numbered counterclockwise, such that the normal vector on the element boundary ∂T points to the exterior of T. For the triangulation $\mathcal{T}$ in Figure 1.10 we obtain Table 1.1.

In MATLAB the data is read with

```
load coord.dat; coord(:,1) = [ ];
load elem.dat;  elem(:,1) = [ ];
```

For the boundary conditions (1.2.14) and (1.2.15) the boundary edges $e \subset \partial T$ for which $e \cap \Gamma \neq \emptyset$ will be distributed in 2 groups: we obtain two further lists `Diri.dat` and `Neum.dat`, for the storage of the start- and end points of the edges contained in Γ_D and Γ_N, respectively.

The ordering of the vertex indices i, j in a line is such that the direction $\underline{P}_i \to \underline{P}_j$ and the exterior normal to e form a 'right hand system'. As before, the data is read in MATLAB as follows:

```
load Neum.dat; Neum(:,1) = [ ];
load Diri.dat; Diri(:,1) = [ ];
```

1.3.2 Element Stiffness Matrix

Let $T \in \mathcal{T}$ be a triangle with vertices P_0, P_1, P_2 and coordinates (x_i, y_i), $i = 0, 1, 2$. Let further $b_{P_i}(x)$ be the restrictions of the basis functions associated with node P_i to element $T \in \mathcal{T}$. The element stiffness matrix $\boldsymbol{A}_T \in \mathbb{R}^{3\times 3}_{\mathrm{sym}}$ is defined by

$$A_{ij} := \int_T \nabla b_{P_i} \cdot \nabla b_{P_j}\, dx \qquad 0 \le i, j \le 2. \tag{1.3.1}$$

We prove (1.2.29) directly, i.e. without any transformation to the reference element coordinates. We use that

$$b_{P_i}\big|_T \text{ linear on } T, b_{P_i}(P_j) = \delta_{ij}, \ \ 0 \le i, j \le 2.$$

Then for $(x, y) \in T$ (indices to be taken modulo 3!)

$$b_{P_i}(x,y) = \frac{\begin{vmatrix} 1 & x & y \\ 1 & x_{i+1} & y_{i+1} \\ 1 & x_{i+2} & y_{i+2} \end{vmatrix}}{\begin{vmatrix} 1 & x_i & y_i \\ 1 & x_{i+1} & y_{i+1} \\ 1 & x_{i+2} & y_{i+2} \end{vmatrix}} = \frac{1}{|\boldsymbol{B}_T|} \begin{vmatrix} 1 & x & y \\ 1 & x_{i+1} & y_{i+1} \\ 1 & x_{i+2} & y_{i+2} \end{vmatrix}, \tag{1.3.2}$$

and we calculate

$$\frac{\partial b_{P_i}}{\partial x} = -\frac{1}{|\boldsymbol{B}_T|} \begin{vmatrix} 1 & y_{i+1} \\ 1 & y_{i+2} \end{vmatrix} = \frac{y_{i+1} - y_{i+2}}{|\boldsymbol{B}_T|} \quad \frac{\partial b_{P_i}}{\partial y} = \frac{1}{|\boldsymbol{B}_T|} \begin{vmatrix} 1 & x_{i+1} \\ 1 & x_{i+2} \end{vmatrix} = \frac{x_{i+2} - x_{i+1}}{|\boldsymbol{B}_T|}.$$

We obtain:

$$\nabla b_{P_i} = \frac{1}{2|T|} \begin{pmatrix} y_{i+1} - y_{i+2} \\ x_{i+2} - x_{i+1} \end{pmatrix}, \quad i = 0, 1, 2, \tag{1.3.3}$$

where the indices are taken modulo 3, i.e. $\underline{P}_3 = \underline{P}_0$, $\underline{P}_4 = \underline{P}_1$ etc., and

$$|T| = \int_T dx = {}^1\!/_2 \det(\boldsymbol{B}_T) = {}^1\!/_2 \det \begin{pmatrix} x_1 - x_0 & x_2 - x_0 \\ y_1 - y_0 & y_2 - y_0 \end{pmatrix} = {}^1\!/_2 \det \begin{pmatrix} 1 & 1 & 1 \\ x_0 & x_1 & x_2 \\ y_0 & y_1 & y_2 \end{pmatrix}$$

Now, we show (1.2.29): let

$$VERT = \begin{pmatrix} x_0 & y_0 \\ x_1 & y_1 \\ x_2 & y_2 \end{pmatrix}$$

be the coordinates matrix for the vertices P_i of T and define $\boldsymbol{G}$ by

$$\boldsymbol{G} = \begin{pmatrix} 1 & 1 & 1 \\ x_0 & x_1 & x_2 \\ y_0 & y_1 & y_2 \end{pmatrix}^{-1} \begin{pmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Then the element stiffness matrix is given by

$$\boldsymbol{A}_T = \frac{|T|}{2} \boldsymbol{G}\boldsymbol{G}^\top$$

This is realized by the following MATLAB function which is written more generally for linear, simplicial elements in dimension n:

```
function S = stiff(vertices)
%get dimension n of the problem
n = size(vertices, 2);
G = [ones(1, n+1); vertices'] \ [zeros(1, n); eye(n)];
A_T = det([ones(1, n+1); vertices']) * G * G' / prod(1:n);
```

The stiffness matrix for the element $T_j \in \mathcal{T}$ is obtained by

$$\texttt{stiff}(\underbrace{\texttt{coord}(\texttt{elem}(j,:),:)}_{\texttt{VERT}\text{ of element } T_j})$$

1.3.3 Assembly of the global stiffness matrix

We realize the assembly of the matrix $\boldsymbol{A} = \mathcal{A}_{T\in\mathcal{T}} \boldsymbol{A}_T$ in MATLAB.

First of all we allocate some space for the sparse matrix $\boldsymbol{A}$ with

```
N = size (coord,1);
A = sparse (N,N);
```

The assembly is then realized as follows

```
for j=1:size(elem,1)
   A(elem(j,:),elem(j,:)) = A(elem(j,:),elem(j,:)) + ...
   stiff(coord(elem(j,:),:));
end
```

1.3.4 Assembly of the Load Vector

We have (see (1.2.16))

$$l(v) = \int_D fv \, d\underline{x} + \int_{\Gamma_N} gv \, ds,$$

and therefore

$$l_j = l(b_{P_j}) = \int_D f b_{P_j} \, d\underline{x} + \int_{\Gamma_N} g b_{P_j} \, ds, j = 1, \ldots, N.$$

For a Dirichlet problem, $\Gamma_N = \emptyset$ and the boundary integral vanishes. We assume that $f(\underline{x})$ and $g(\underline{x})$ are defined by MATLAB functions. Similarly to $\boldsymbol{A}$, the load vector $\underline{l}$ will be assembled from elemental contributions: For general $f(x), g(x)$ a numerical quadrature is needed. The simplest quadrature rule is

$$\int_T f \, dx \approx |T| f(s_T), \int_e g \, ds \approx |e| g(s_e),$$

where s_T and s_e denote the centre of gravity of T and e, respectively:

$$\underline{s}_T = {}^1\!/_3(\underline{P}_0 + \underline{P}_1 + \underline{P}_2), \quad \underline{s}_e = {}^1\!/_2(\underline{P} + \underline{Q})$$

for an edge $e \in \overline{PQ}$.

The load vector is initialized by

```
L = sparse(size(coord,1),1);
```

We obtain the right hand side $\underline{l}$ (in MATLAB-Code `L`) by assembling $\int_T f b_{P_j} \, dx$, $\int_{\overline{T} \cap \Gamma_N} g b_{P_j} \, ds$; this is realized as follows:

```
for j = 1: size(elem,1)
  area = det ([1 1 1; coord(elem(j,:),:)'])/2;
  L(elem(j,:)) = L(elem(j,:)) + area * f( sum(coord(elem(j,:),:)/3)/3;
end

for j = 1: size(Neum,1)
  L(Neum(j,:)) = L(Neum(j,:)) + ...
                 norm(coord(Neum(j,1),:) - coord(Neum(j,2),:)) * ...
                 g(sum(coord(Neum(j,:),:))/2)/2;
end
```

Here, `sum(coord(elem(j,:),:)` $= 3\underline{s}_T$.

1.3.5 Homogeneous Dirichlet Data

Until now we did not enforce any Dirichlet boundary conditions (1.2.14); the matrix $\boldsymbol{A}$ has full dimension $N \times N$. (1.2.14) means that for $P_i \in \Gamma_D$ it holds:

$$u_N(P_i) = 0, \quad P_i \in \Gamma_D.$$

This is enforced by the substitution

$$A_{ii} = 1, A_{ij} = A_{ji} = 0 \text{ for } 1 \leq j \leq N, j \neq i, L_i = 0 \tag{1.3.4}$$

if $\underline{P}_i \in \Gamma_D$.

Exercise 1.43. Write a MATLAB routine for (1.3.4).

1.4 Further Finite Elements

We present here further elements for the Poisson problem. To this end, we first give the element mapping and the Ansatz space on the reference element and the element shape-functions on the reference element.

1.4.1 Simplicial Affine Elements

Definition 1.44 (Regular mesh)**.** *A partition $\mathcal{T}$ of a polyhedron $D \subset \mathbb{R}^3$ into simplices K is **regular**, if the intersection $K \cap K'$ of any two different simplices $K, K' \in \mathcal{T}$ is a vertex, an entire edge or an entire side.*

Let $D \subset \mathbb{R}^n$, $n = 3$, be a polyhedron, $\mathcal{T}$ a regular simplicial triangulation, $\widehat{K}$ the unit simplex

$$\widehat{K} := \Big\{ \underline{\xi} \in \mathbb{R}^n : \xi_i > 0, \sum_{i=1}^{n} \xi_i < 1 \Big\}.$$

The reference element shape-functions $\widehat{N}_i(\xi)$, $i = 0, 1, ..., n$ are given by

$$\widehat{N}_0(\underline{\xi}) = 1 - \sum_{i=1}^{n} \xi_i, \quad \widehat{N}_i(\underline{\xi}) = \xi_i, \quad i = 1, \ldots, n,$$

and $\underline{F}_K : \widehat{K} \to K \in \mathcal{T}$ denotes the affine element mapping

$$\mathcal{T} \ni K \ni \underline{x} = \underline{F}_K(\underline{\xi}) = \sum_{i=0}^{n} \widehat{N}_i(\underline{\xi})\, \underline{P}_i.$$

Let further

$$S^p(D, \mathcal{T}) = \left\{ u \in H^1(D) : u|_K \in \mathcal{P}_p(K) \text{ for } K \in \mathcal{T} \right\}, \tag{1.4.1}$$

where

$$\mathcal{P}_p(K) = \left\{ u : K \to \mathbb{R} | u(\underline{x}) = \sum_{|\underline{\alpha}| \le p} c_{\underline{\alpha}} x^{\underline{\alpha}} \right\}$$

is the set of polynomials of total degree p. The local degrees of freedom are again the **function values at the vertices**; for dimension $n = 2$ and $n = 3$ this is illustrated in Figure 1.11.

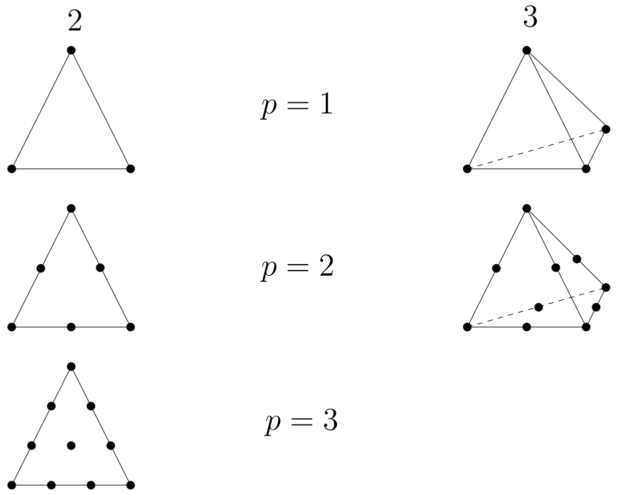


Figure 1.11: Higher order simplicial elements.

1.4.2 Bilinear Quadrilateral Elements

Let $D \subset \mathbb{R}^2$ be a polygon and let $\mathcal{T}$ be a regular mesh of quadrilaterals K as illustrated in Figure 1.12.

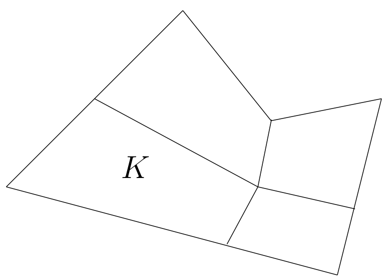


Figure 1.12: Regular quadrilateral mesh in D.

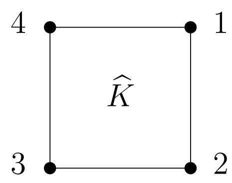


Figure 1.13: Local degrees of freedom for bilinear shape-functions (1.4.2) for quadrilateral elements.

Reference element shape-functions

The bilinear element shape-functions on the reference element $\widehat{K} = (-1,1)^2$ are:

$$\begin{aligned} \widehat{N}_1(\underline{\xi}) &= 1/4\,(1+\xi_1)(1+\xi_2), \quad \widehat{N}_2(\underline{\xi}) = 1/4\,(1+\xi_1)(1-\xi_2), \\ \widehat{N}_3(\underline{\xi}) &= 1/4\,(1-\xi_1)(1-\xi_2), \quad \widehat{N}_4(\underline{\xi}) = 1/4\,(1-\xi_1)(1+\xi_2). \end{aligned} \tag{1.4.2}$$

The degrees of freedom are indicated in Figure 1.13.

Element mapping

The bilinear element mapping is given by

$$\underline{F}_K(\underline{\xi}) = \sum_{i=1}^{4} N_i(\underline{\xi})\underline{P}_i : \widehat{K} \to K \in \mathcal{T}, \tag{1.4.3}$$

as in Figure 1.14. The element mappings $\underline{F}_K$ are collected in $\boldsymbol{F} := \{\underline{F}_K : K \in \mathcal{T}\}$.

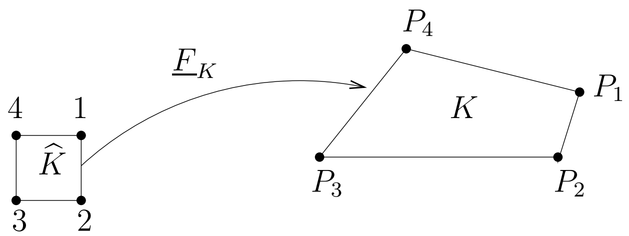


Figure 1.14: Bilinear element mapping for quadrilateral elements.

FE-Space $S^1(D, \mathcal{T}, \boldsymbol{F})$

The Finite Element space is defined by transporting the bilinear functions in $\widehat{K}$ onto $K \in \mathcal{T}$ by $\underline{F}_K(\xi)$. Let

$$\mathcal{Q}_1 = \operatorname{span}\left\{\widehat{N}_1, \widehat{N}_2, \widehat{N}_3, \widehat{N}_4\right\} = \operatorname{span}\left\{\underline{\xi}^{\underline{\alpha}} : |\underline{\alpha}| \le 1\right\} \tag{1.4.4}$$

be the polynomials of degree 1 in each variable. Then

$$S^1(D, \mathcal{T}, \boldsymbol{F}) = \left\{u \in H^1(D) : u|_K \circ \underline{F}_K(\underline{\xi}) \in \mathcal{Q}_1(\widehat{K}), K \in \mathcal{T}\right\}. \tag{1.4.5}$$

In general, $u_N(\underline{x})$ on $K \in \mathcal{T}$ is not a polynomial anymore, but the image of a polynomial by $\underline{F}_K$.

In case the element mapping is clear, as e. g. for affine elements, we drop the argument $\boldsymbol{F}$ in (1.4.5).

Element stiffness matrix

The mapping $\underline{F}_K(\underline{\xi})$ in (1.4.3) is affine if and only if K is a parallelogram. Therefore, the element stiffness matrix for bilinear quadrilaterals can be computed only numerically with the aid of numerical integration. This is realized as follows: first of all we split the bilinear form in a sum of element bilinear forms

$$a(u,v) = \sum_{K \in \mathcal{T}} a_K(u,v), \quad u,v \in S^1(D,\mathcal{T},\underline{F})$$
$$a_K(u,v) = \int_K \underline{\nabla} v \cdot \boldsymbol{a}(x)\, \underline{\nabla} u \, dx = (\underline{u}_K)^\top \boldsymbol{A}_K \, \underline{v}_K$$

with symmetric, positive definite diffusion matrix $\boldsymbol{a}(x)$ and

$$(\boldsymbol{A}_K)_{ij} = \int_K (\underline{\nabla}(\underline{N}_K))_i^\top \cdot \boldsymbol{a}(x)\, \underline{\nabla}(\underline{N}_K)_j \, dx \tag{1.4.6}$$

and element shape-functions

$$\underline{N}_K(\underline{x}) = \underline{\widehat{N}}(\underline{F}_K^{-1}(\underline{x})), \ \underline{\widehat{N}} = (\widehat{N}_1, \widehat{N}_2, \widehat{N}_3, \widehat{N}_4)^\top$$

with reference element shape-functions $\underline{\widehat{N}}_i$ as defined in (1.4.2). We have that

$$K \ni \underline{x} = \underline{F}_K(\underline{\xi}), \ \ \underline{\xi} \in \widehat{K} = (-1,1)^2,$$

and (see also (1.2.26)) (1.4.6) transforms to

$$\underline{\nabla} = \boldsymbol{B}_K^{-\top} \, \underline{\widehat{\nabla}},$$

where $\boldsymbol{B}_K(\xi)$ is a polynomial matrix function in $\underline{\xi}$ and

$$dx = \det\left(\frac{\partial \underline{F}_K}{d\underline{\xi}}\right) d\xi = \det(\boldsymbol{B}_K(\xi))\, d\xi.$$

We obtain that

$$(\boldsymbol{A}_K)_{ij} = \int_{\widehat{K}} (\widehat{\nabla}\widehat{N}_i)^\top (\boldsymbol{B}_K(\underline{\xi}))^{-1} \, \widehat{\boldsymbol{a}}_K(\underline{\xi})\, (\boldsymbol{B}_K(\xi))^{-\top} (\widehat{\nabla}\widehat{N}_j)(\underline{\xi})\, |\boldsymbol{B}_K| \, d\underline{\xi},$$

where we define

$$\widehat{\boldsymbol{a}}_K(\underline{\xi}) := \boldsymbol{a}(\underline{F}_K(\underline{\xi})).$$

The inverse $\boldsymbol{B}_K^{-1}(\underline{\xi})$ is a $n \times n$ matrix whose entries are rational functions of $\underline{\xi}$—therefore, $\int_{\widehat{K}} d\underline{\xi}$ has to be approximated by numerical integration, even if $\boldsymbol{a}$ is constant.

Numerical integration

The integral $\int_{-1}^{1} f(\xi)d\xi$ will be approximated by a q-point Gauss-quadrature: it holds

$$\int_{-1}^{1} f(\xi)\,d\xi - \sum_{k=1}^{q} w_k^{(q)} f(\xi_k^{(q)}) = C_q\, f^{(2q)}(\zeta)$$

for a $\zeta \in (-1,1)$: the quadrature is exact for $f \in \mathcal{P}_{2q-1}$. $w_k^{(q)}$ are called **weights** and $\xi_k^{(q)}$ **nodes** of the q-point Gauss formula. Quadrature rules in $\widehat{K} = (-1,1)^2$ are obtained by a tensor product rule:

$$\int_{\widehat{K}} f(\underline{\xi})\,d\underline{\xi} \approx \sum_{k,l=1}^{q} w_k^{(q)} w_l^{(q)} f(\xi_k^{(q)}, \xi_l^{(q)}) =: G^{(q)}[f].$$

We obtain the approximation for $\boldsymbol{A}_K$:

$$(\boldsymbol{A}_K)_{ij} \approx G^{(q)}\left[(\widehat{\underline{\nabla}}\widehat{N}_i)^\top\, \boldsymbol{B}_K^{-1}\, \widehat{\boldsymbol{a}}_K\, \boldsymbol{B}_K^{-\top}\, \widehat{\underline{\nabla}}\, \underline{N}_j\, |\boldsymbol{B}_K|\right],$$

where $\boldsymbol{B}_K^{-1}(\xi_k^{(q)}, \xi_l^{(q)})$ has to be evaluated at each Gauss point (inverse of a 2×2 matrix). How to compute the weights and abscissas for a Gaussian quadrature rule is shown in Algorithm 1.2.

Project

For the quadrilateral element K in Figure 1.14 the element mapping $\underline{F}_K$ can degenerate, if e. g. three vertices are on the same line. Prove, that the bilinear mapping $\underline{F}_K$ in (1.4.3) is a bijection from $\widehat{K}$ to K, provided that K is convex and that K does not degenerate to a triangle.

Hint: Prove that $\det(\partial \underline{F}_K / \partial \underline{\xi})$ does not vanish in $\overline{\widehat{K}}$.

Project (RW)

Modify the MATLAB routines for the quadrilateral FE mesh. Use numerical integration in “stiff”.

1.4.3 Bilinear Quadrilateral Elements of Order $p \geq 1$

As in (1.4.3), the element mapping $\underline{F}_K(\underline{\xi})$ is again bilinear. The polynomial space on the reference element $\widehat{K} = (-1,1)^n$ is now the space of tensor product polynomials of degree p in each variable:

$$\mathcal{Q}_p := \operatorname{span}\left\{\xi_1^i \xi_2^j\right\}_{0\leq i,j\leq p}.$$

For $p > 1$, $S^p(D, \mathcal{T}, \boldsymbol{F})$

$$S^p(D,\mathcal{T},\boldsymbol{F}) = \{u \in H^1(D) : u|_K \circ F_K \in Q_p,\ K \in \mathcal{T}\}\,. \tag{1.4.7}$$

Algorithm 1.2 Weights and Abscissas for the Gaussian Quadrature.

```
function [x,w]=gauleg(n)
%  input: the order of the Gauss formula; n
%  output: the abscissas x and the weights w for the Gauss formula; x are
%          simply the zeros of the nth Legendre polynomial
%
x=zeros(n,1);
w=zeros(n,1);
  m=(n+1)/2;
  xm=0.0;
  xl=1.0;
  for i=1:m
    z=cos(pi*(i-0.25)/(n+0.5));
    while 1
      p1=1.0;
      p2=0.0;
      for j=1:n
        p3=p2;
        p2=p1;
        p1=((2.0*j-1.0)*z*p2-(j-1.0)*p3)/j;
      end
      pp=n*(z*p1-p2)/(z*z-1.0);
      z1=z;
      z=z1-p1/pp;
      if (abs(z-z1)<eps), break, end
    end
    x(i)=xm-xl*z;
    x(n+1-i)=xm+xl*z;
    w(i)=2.0*xl/((1.0-z*z)*pp*pp);
    w(n+1-i)=w(i);
  end
```

1.4.4 Isoparametric Quadrilateral Elements

Our problem: linear and bilinear element mappings $\underline{F}_K$ allow for an exact mapping only for polygons and polyhedra. However, many domains in practical applications have a curvilinear boundary. **Isoparametric elements** have element mappings $\underline{F}_K : \widehat{K} \to K \in \mathcal{T}$ in span $\{N_i(\underline{\xi})\}$. $S^1(D, \mathcal{T})$ is obviously an isoparametrical FE-space for both triangles/simplices and quadrilaterals (bilinear Ansatzes/element mappings). For $p > 1$, however, this is no longer true. There again can affine mappings reproduce only polyhedral domains, when the FE solution may however be a polynom of degree $p > 1$.

The simplest example of a curvilinear element is illustrated in Figure 1.15.

Example 1.45 (Isoparametrical $\mathcal{Q}_2$-element shape-functions)**.** Let $\widehat{K} = (-1,1)^2$ and $\psi(z) = 1 - z^2$, $\psi_-(z) = \frac{1}{2}(z-1)$ and $\psi_+(z) = \frac{1}{2}(z+1)$. Then we have (see also

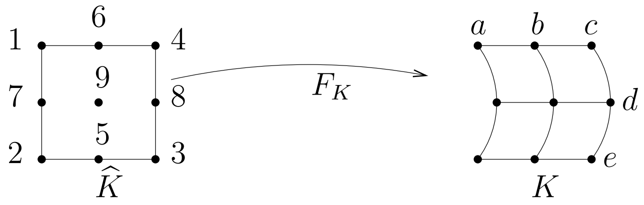


Figure 1.15: Local degrees of freedom and element mapping of the isoparametrical $\mathcal{Q}_2$-element.

Figure 1.15) the (bilinear) shape-functions for the vertices,

$$\widehat{N}_1(\underline{\xi}) = \psi_-(\xi_1)\psi_+(\xi_2), \qquad \widehat{N}_2(\underline{\xi}) = \psi_-(\xi_1)\psi_-(\xi_2),$$
$$\widehat{N}_3(\underline{\xi}) = \psi_+(\xi_1)\psi_-(\xi_2), \qquad \widehat{N}_4(\underline{\xi}) = \psi_+(\xi_1)\psi_+(\xi_2),$$

the edge shape-functions

$$\widehat{N}_5(\underline{\xi}) = \psi(\xi_1)\psi_-(\xi_2), \qquad \widehat{N}_6(\underline{\xi}) = \psi(\xi_1)\psi_+(\xi_2),$$
$$\widehat{N}_7(\underline{\xi}) = \psi_-(\xi_1)\psi(\xi_2), \qquad \widehat{N}_8(\underline{\xi}) = \psi_+(\xi_1)\psi(\xi_2)$$

and the bubble function

$$\widehat{N}_9(\underline{\xi}) = \psi(\xi_1)\psi(\xi_2).$$

The element mapping is given by

$$\underline{F}_K(\underline{\xi}) = \sum_{i=1}^{9} \widehat{N}_i(\underline{\xi})\underline{P}_i.$$

In particular, the element mapping is not affine anymore. Consequently, after the transformation into the reference coordinates, the entries of the element stiffness matrix have to be evaluated numerically, using a numerical quadrature rule.

1.5 Element Families in Three Dimensions

In two dimensions, polygonal domains are meshed by triangles and/or quadrilaterals. In three-dimensional applications, the analogs to quadrilateral elements are hexahedra which provide often better approximations than tetrahedral elements with the same number of nodes. With hexahedra, however, it is difficult to mesh arbitrary volumes without getting some irregular (i. e. hanging) nodes. What is therefore usually done is the mixture of different types of elements in the same mesh, also to avoid unduely

tetrahedron	$a_0 = (0,0,0),\ \ a_1 = (1,0,0),\ \ a_2 = (0,1,0),\ \ a_3 = (0,0,1)$
pyramid	$a_0 = (0,0,0),\ \ a_1 = (1,0,0),\ \ a_2 = (1,1,0),\ \ a_3 = (0,1,0),$ $a_4 = (0,0,1)$
prism	$a_0 = (0,0,0),\ \ a_1 = (1,0,0),\ \ a_2 = (0,1,0),\ \ a_3 = (0,0,0),$ $a_4 = (1,0,1),\ \ a_5 = (0,1,1)$
hexahedron	$a_0 = (0,0,0),\ \ a_1 = (1,0,0),\ \ a_2 = (1,1,0),\ \ a_3 = (0,1,0),$ $a_4 = (0,0,1)\ \ a_5 = (1,0,1),\ \ a_6 = (1,1,1),\ \ a_7 = (0,1,1)\,.$

Table 1.3: Nodes of the reference elements in three dimensions.

distorted elements; however, whereas in $2-d$ regular meshes in general polygons can be constructed only from triangles and quadrilaterals, in $3-d$, (a few) pyramids must be included.

On regular meshes, continuous FE-approximations are required to construct FE-spaces V_N which are contained in $H^1(D)$. In the present section, we present Reference-Element shape-functions for piecewise linear ($p=1$) and quadratic ($p=2$) approximations on general, regular meshes in polyhedra $D \subset \mathbb{R}^3$. We follow [83].

For hexahedra and tetrahedra, standard nodal shape-functions are used, whereas special care is needed for pyramids and prisms. Our goal is to obtain globally continuous approximations. For pyramids, there are in general no polynomial shape-functions satisfying this requirement. We give a construction which requires a split into two tetrahedra.

1.5.1 Reference Elements $\widehat{K}$

We define the shape-functions for the reference elements $\widehat{K} \subset \mathbb{R}^3$ with the nodes $a_0, \dots, a_n$: Table 1.3.

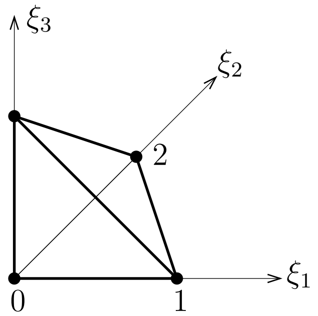
1.5.2 Reference Element Shape Functions

The shape-functions for tetrahedrons and hexahedrons are well known and can be found in many textbooks, e. g. [28]. The conforming $\widehat{N}_1$ elements are defined in Figures 1.16–1.19.

Pyramid elements

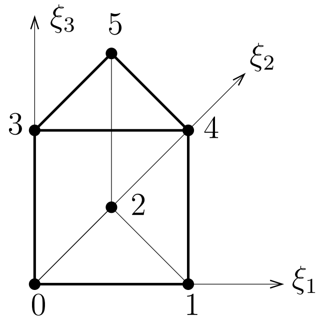
The construction of the shape-functions for pyramids is more involved and not unique. In particular, a conforming family of elements cannot be achieved with polynomial shape-functions.

Theorem 1.46. *There exist no continuously differentiable conforming shape functions for the pyramid which are linear resp. bilinear on the faces.*



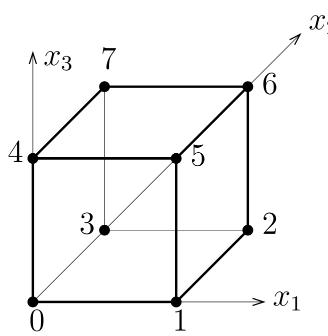
$$\widehat{N}_0(\xi_1,\xi_2,\xi_3) = 1-\xi_1-\xi_2-\xi_3$$
$$\widehat{N}_1(\xi_1,\xi_2,\xi_3) = \xi_1$$
$$\widehat{N}_2(\xi_1,\xi_2,\xi_3) = \xi_2$$
$$\widehat{N}_3(\xi_1,\xi_2,\xi_3) = \xi_3$$

Figure 1.16: Tetrahedral $\mathcal{P}_1$ reference element.



$$\widehat{N}_0(\xi_1,\xi_2,\xi_3) = (1-\xi_1-\xi_2)(1-\xi_3)$$
$$\widehat{N}_1(\xi_1,\xi_2,\xi_3) = \xi_1(1-\xi_3)$$
$$\widehat{N}_2(\xi_1,\xi_2,\xi_3) = \xi_2(1-\xi_3)$$
$$\widehat{N}_3(\xi_1,\xi_2,\xi_3) = (1-\xi_1-\xi_2)\xi_3$$
$$\widehat{N}_4(\xi_1,\xi_2,\xi_3) = \xi_1\xi_3$$
$$\widehat{N}_5(\xi_1,\xi_2,\xi_3) = \xi_2\xi_3$$

Figure 1.17: Prismatic reference element.



$$\widehat{N}_0(\xi_1,\xi_2,\xi_3) = (1-\xi_1)(1-\xi_2)(1-\xi_3)$$
$$\widehat{N}_1(\xi_1,\xi_2,\xi_3) = \xi_1(1-\xi_2)(1-\xi_3)$$
$$\widehat{N}_2(\xi_1,\xi_2,\xi_3) = \xi_1\xi_2(1-\xi_3)$$
$$\widehat{N}_3(\xi_1,\xi_2,\xi_3) = (1-\xi_1)\xi_2(1-\xi_3)$$
$$\widehat{N}_4(\xi_1,\xi_2,\xi_3) = (1-\xi_1)(1-\xi-2)\xi_3$$
$$\widehat{N}_5(\xi_1,\xi_2,\xi_3) = \xi_1(1-\xi_2)\xi_3$$
$$\widehat{N}_6(\xi_1,\xi_2,\xi_3) = \xi_1\xi_2\xi_3$$
$$\widehat{N}_7(\xi_1,\xi_2,\xi_3) = (1-\xi_1)\xi_2\xi_3$$

Figure 1.18: Hexahedral $\mathcal{Q}_1$ reference element.

Proof. Let p be a conforming shape-function in the reference pyramid for the corner $(0,0,0)$. Then, restricted to the faces we have

(a)	$p(x,y,z) = (1-x)(1-y)$	in $\operatorname{conv}\{(0,0,0),(1,0,0),(1,1,0),(0,1,0)\}$
(b)	$p(x,y,z) = (1-x-z)$	in $\operatorname{conv}\{(0,0,0),(1,0,0),(0,0,1)\}$
(c)	$p(x,y,z) = (1-y-z)$	in $\operatorname{conv}\{(0,0,0),(0,1,0),(0,0,1)\}$
(d)	$p(x,y,z) = 0$	in $\operatorname{conv}\{(1,0,0),(1,1,0),(0,0,1)\}$
(e)	$p(x,y,z) = 0$	in $\operatorname{conv}\{(1,1,0),(0,1,0),(0,0,1)\}$

Using (d) and (e), we get

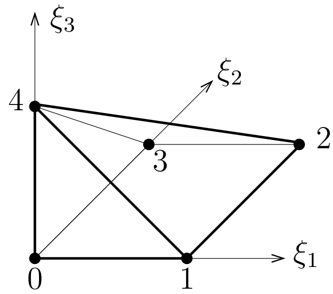
$$\nabla p(0,0,1)\begin{pmatrix}0\\-1\\1\end{pmatrix} = \nabla p(0,0,1)\begin{pmatrix}-1\\0\\1\end{pmatrix} = \nabla p(0,0,1)\begin{pmatrix}-1\\-1\\1\end{pmatrix} = 0.$$

If ∇p is continuous at $(0,0,1)$, this implies $\nabla p(0,0,1)=0$, but (b) and (c) give

$$\nabla p(0,0,1)\begin{pmatrix}0\\0\\1\end{pmatrix} = -1.$$

This is a contradiction. □

Of course, the same result holds for quadratic shape-functions on pyramids. Here, we describe one possibility for the construction of appropriate pyramid shape-functions based on a splitting of the pyramid into two tetrahedra.



$$\widehat{N}_0(\xi_1,\xi_2,\xi_3) = \begin{cases} (1-\xi_1)(1-\xi_2)+ \\ \quad \cdots + \xi_3(\xi_2-1) & \text{for } \xi_1 > \xi_2 \\ (1-\xi_1)(1-\xi_2)+ \\ \quad \cdots + \xi_3(\xi_1-1) & \text{for } \xi_1 \le \xi_2 \end{cases}$$

$$\widehat{N}_2(\xi_1,\xi_2,\xi_3) = \begin{cases} \xi_1(1-\xi_2) - \xi_3\xi_2 & \text{for } \xi_1 > \xi_2 \\ \xi_1(1-\xi_2) - \xi_3\xi_1 & \text{for } \xi_1 \le \xi_2 \end{cases}$$

$$\widehat{N}_3(\xi_1,\xi_2,\xi_3) = \begin{cases} (1-\xi_1)\xi_2 - \xi_3\xi_2 & \text{for } \xi_1 > \xi_2 \\ (1-\xi_1)\xi_2 - \xi_3\xi_1 & \text{for } \xi_1 \le \xi_2 \end{cases}$$

$$\widehat{N}_4(\xi_1,\xi_2,\xi_3) = \xi_3$$

Figure 1.19: Pyramidal reference element.

1.6 Irregular Meshes and Hanging Nodes

Until now, FE meshes had to be regular. Often, however, irregular meshes are of interest, in particular in the context of adaptive mesh refinements which generate hanging nodes (cf. Figure 1.6).

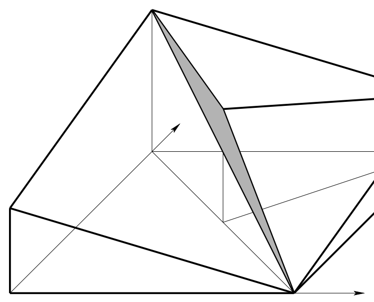


Figure 1.20: Hanging node with a real degree of freedom leading to a discontinuous FE function.

Treating a hanging node like a real degree of freedom (as in Figure 1.6) can result in a discontinuous FE function. To avoid this, the hanging node has to be treated as constrained by the neighbouring degrees of freedom in the assembly process. This can be done by manipulating the T matrices.

1.6.1 T Matrices

Definition 1.47 (T matrix). *Let m_K be the number of local shape functions $\left\{N_j^K\right\}_{j=1}^{m_K}$ in the element K and N the number of global basis functions $\{b_i\}_{i=1}^{N}$. The T matrix $\boldsymbol{T}_K \in \mathbb{R}^{m_K \times N}$ of the element K describes how the restriction of the global basis functions $\{b_i\}_{i=1}^{N}$ onto the element K are constructed from the local shape functions:*

$$b_i|_K = \sum_{j=1}^{m_K} [\boldsymbol{T}_K]_{ji}\, \varphi_j^K$$

and in vector notation: $\underline{b}|_K = \boldsymbol{T}_K^\top \underline{N}^K$.

As in one dimension (1.1.30), this leads to the following assembling procedure for the stiffness matrix:

$$\boldsymbol{A} = \sum_K \boldsymbol{T}_K^\top \boldsymbol{A}_K \boldsymbol{T}_K.$$

Regular Mesh

There is a "one-to-one" correspondence of local shape functions and global basis functions: Every local shape function contributes to at most one global basis function. On the other hand, the restriction of a global basis function onto an element gets contributions from at most one local shape function.

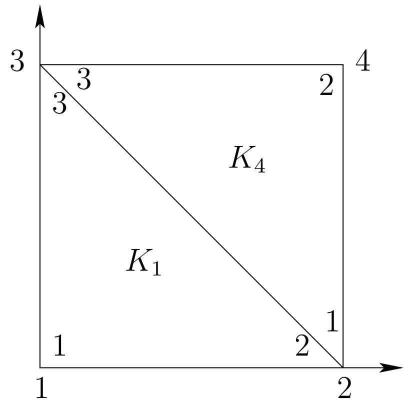


Figure 1.21: Numbering of regular mesh with two elements with three local shape functions each and four global degrees of freedom.

Example 1.48. Consider the regular mesh shown in Figure 1.21. The elements K_1 and K_4 have the T matrices

$$\boldsymbol{T}_{K_1} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \text{ and } \boldsymbol{T}_{K_4} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}. \tag{1.6.1}$$

As a result of the "one-to-one" correspondence of local shape functions and global basis functions, every row and every column of the two T matrices in (1.6.1) have at most one entry equal to 1.

Irregular Mesh

Consider the irregular mesh shown in Figure 1.22. Clearly, there is no such "one-to-one" correspondence of local shape functions and global basis function as is was the case for the regular mesh.

Example 1.49. Using Definition 1.47, the T matrices of the elements K_2 and K_3 are:

$$\boldsymbol{T}_{K_2} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1/2 & 1/2 & 0 \end{pmatrix} \text{ and } \boldsymbol{T}_{K_3} = \begin{pmatrix} 0 & 1/2 & 1/2 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}. \tag{1.6.2}$$

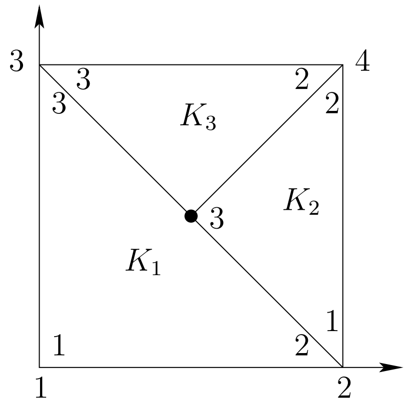


Figure 1.22: Irregular mesh with three elements with three local shape functions each and four global degrees of freedom. The hanging node is marked with a $\circ$.

1.6.2 S Matrices

In the regular case, the T matrices can be computed by counting and associating local and global degrees of freedom. In the irregular case, this is too complicated. If the irregularity in the mesh results form a refinement without topological closure, it is possible to compute the T matrices of the small elements with irregularities from their parent's T matrix using S matrices.

Definition 1.50 (S matrix)**.** *Let $K' \subset K$ be the result of a refinement of element K. The S matrix $\boldsymbol{S}_{K'K} \in \mathbb{R}^{m_{K'} \times m_K}$ describes how the restriction of the shape functions $\left\{N_j^K\right\}_{j=1}^{m_K}$ onto K' are constructed from the shape functions $\left\{N_l^{K'}\right\}_{l=1}^{m_{K'}}$ of K':*

$$N_j^K\big|_{K'} = \sum_{l=1}^{m_{K'}} \left[\boldsymbol{S}_{K'K}\right]_{lj} N_l^{K'}$$

and in vector notation: $\underline{N}^K\big|_{K'} = \boldsymbol{S}_{K'K}^\top \underline{N}^{K'}$. In the trivial case $K = K'$ (i. e. no refinement), the S matrix $\boldsymbol{S}_{KK}$ is defined as the identity matrix.

Consider a mesh $\mathcal{M}$ for which all elements and associated T matrices have been generated. Suppose the mesh $\mathcal{M}'$ is the result of a subdivision of several elements of $\mathcal{M}$. The basis functions $B := \{b_1, \ldots, b_N\}$ defined for $\mathcal{M}$ may be partitioned into two sets—one denoted by B_{repl} containing all basis functions that can be described solely by elements of $\mathcal{M}'$ that are not part of $\mathcal{M}$ and another one denoted by B_{keep} representing the rest:

$$B = B_{\text{repl}} \cup B_{\text{keep}}.$$

Note that B_{repl} is easily determined; the support of basis functions in B_{repl} consists entirely of newly inserted elements.

The set of basis functions B' related to mesh $\mathcal{M}'$ contains all basis functions in B_{keep} plus an additional set B_{ins} of basis functions generated by consistent components of mesh $\mathcal{M}'$ formed by elements not part of $\mathcal{M}$:

$$B' = B_{\text{ins}} \cup B_{\text{keep}}.$$

The meshes of Examples 1.48 and 1.49 have $B = B'$ since $B_{\text{ins}} = \emptyset$.

Proposition 1.51. *Let $K' \subset K$ be the result of a refinement of an element K. Then, the T matrix of K' can be computed as*

$$\boldsymbol{T}_{K'} = \boldsymbol{S}_{K'K}\boldsymbol{T}_K^{keep} + \boldsymbol{T}_K^{ins}.$$

where $\boldsymbol{T}_K^{keep}$ denotes the T matrix of element K with columns not related to functions in B_{keep} set to zero and $\boldsymbol{T}_K^{ins}$ the T matrix for functions in B_{ins} with respect to K'.

Proof. If $K = K'$, nothing is to be proved. Let now $K' \subset K$ with strict inclusion. Consider the global basis function $N_i \in B_{\text{keep}}$ restricted to the element K':

$$\begin{aligned} b_i|_{K'} = \ b_i|_K|_{K'} &= \sum_{l=1}^{m_{K'}} [\boldsymbol{T}_{K'}]_{li}\, N_l^{K'} \\ &= \sum_{j=1}^{m_K} [\boldsymbol{T}_K]_{ji}\, N_j^K\big|_{K'} \\ &= \sum_{j=1}^{m_K} [\boldsymbol{T}_K]_{ji} \sum_{l=1}^{m_{K'}} [\boldsymbol{S}_{K'K}]_{lj}\, N_l^{K'}. \\ \Rightarrow [\boldsymbol{T}_{K'}]_{li} = \sum_{j=1}^{m_K} [\boldsymbol{T}_K]_{ji}\, [\boldsymbol{S}_{K'K}]_{lj} & \quad \forall l = 1, \ldots, m_{K'} \end{aligned}$$

This holds for all $b_i \in B_{\text{keep}}$. For $b_i \in B_{\text{ins}}$, the assertion holds by definition. □

Exercise 1.52. Show that the S matrices only depend on the mesh topology but not on the particular element shapes.

Example 1.53. Consider the mesh given in Figure 1.22 where the elements K_2 and K_3 result from a refinement of K_4 in Figure 1.21. Therefore, we should be able to get the T matrices $\boldsymbol{T}_{K_2}$ and $\boldsymbol{T}_{K_3}$ in (1.6.2) from the T matrix $\boldsymbol{T}_{K_2}$ in (1.6.1) and the S matrices $\boldsymbol{S}_{K_2K_4}$ and $\boldsymbol{S}_{K_3K_4}$.

Figure 1.23 denotes the numbering of the shape functions of the elements K_2 and K_3 and of their parent K_4. The corresponding S matrices are:

$$\boldsymbol{S}_{K_2K_4} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1/2 & 0 & 1/2 \end{pmatrix} \text{ and } \boldsymbol{S}_{K_2K_4} = \begin{pmatrix} 1/2 & 0 & 1/2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Verify $\boldsymbol{T}_{K_2} = \boldsymbol{S}_{K_2K_4}\boldsymbol{T}_{K_4}$ and $\boldsymbol{T}_{K_3} = \boldsymbol{S}_{K_3K_4}\boldsymbol{T}_{K_4}$!

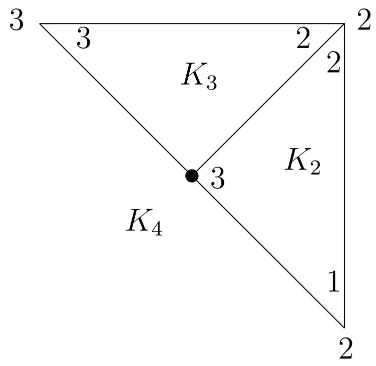


Figure 1.23: Element K_4 with children K_2 and K_3 and the numbering of the shape functions.

1.6.3 Assembly Process

The assembly process was already mentioned in the one and two dimensional FE algorithms in (1.1.30) and (1.2.30) respectively. Here, the algorithm is presented more generally.

Definition 1.54 (Assembly operator)**.** *The* ***assembly operator*** $\mathcal{A}$ *is defined as follows*

- *for assembling matrices*

$$\boldsymbol{A} = \mathcal{A}_{K,\tilde{K}\in\mathcal{T}} \boldsymbol{A}_{K\tilde{K}} := \sum_{K,\tilde{K}\in\mathcal{T}} \boldsymbol{T}_K^\top \boldsymbol{A}_{K\tilde{K}} \boldsymbol{T}_{\tilde{K}} \tag{1.6.3}$$

- *for assembling vectors*

$$\underline{l} = \mathcal{A}_{K\in\mathcal{T}} \underline{l}_K := \sum_{K\in\mathcal{T}} \boldsymbol{T}_K^\top \underline{l}_K \tag{1.6.4}$$

Remark 1.55.

- If $K \subseteq K'$, the T matrix of K is obtained by multiplying the T matrix of K' with the S matrix of K and K': $\boldsymbol{T}_K = \boldsymbol{S}_{KK'} \boldsymbol{T}_{K'}$ (Proposition 1.51).

- For the FEM considered in this chapter, $\boldsymbol{A}_{K\tilde{K}} = 0$ for $K \neq \tilde{K}$ and (1.6.3) becomes

$$\boldsymbol{A} = \mathcal{A}_{K\in\mathcal{T}} \boldsymbol{A}_K := \textstyle\sum_{K\in\mathcal{T}} \boldsymbol{T}_K^\top \boldsymbol{A}_K \boldsymbol{T}_K \text{ where } \boldsymbol{A}_{KK} = \boldsymbol{A}_K.$$

  There are other FEM (Discontinuous Galerkin FEM, short: DGFEM) where this is no longer the case.

1.7 Examination Checklist

You should be able to

- know the variational formulation as a boundary value problem for the Poisson equation with the simplest finite elements in one and two-dimensions.
- describe the implementation of the Finite Element Method for the Poisson equation in a polygon **(RW)**.
- prove the multiplicative trace inequality.
- explain the notions
  - global basis function
  - element shape-function
  - reference element shape-function
  - element mapping
- describe other Finite Elements (isoparametric, p-elements).
- explain the difference between regular and irregular meshes.
- explain the assembly process for regular and irregular meshes using S and T matrices.
- explain internal external shape-functions
- explain static condensation

Bibliography

[1] A. Adams A. *Sobolev Spaces.* Academic Press, 1978.

[2] I. Babuška and K. Aziz. On the angle condition in the finite element method. *SIAM Journal on Numerical Analysis*, 13(2):214-226, 1976.

[3] I. Babuška and M. Vogelius. Feeback and adaptive finite element solution of one-dimensional boundary value problems. *Numerische Mathematik*, 44:75-102, 1984.

[4] I. Babuška and K. Aziz. *Survey lectures on the mathematical foundation of the finite element method* in *"The mathematical foundations of the finite element method with applications to partial differential equations"*, A.K. Aziz, Hrsg., Academic Press, New York, 1972, 3-343 (R).

[5] I. Babuška and J. Osborn. *Eigenvalue Problems,* in: Handbook of Numerical Analysis II, pp. 645-785, P.G. Ciarlet and J.L. Lions (eds.), Elsevier Publ., Amsterdam, 1991 (R).

[6] C. Bacuta, L. Chen and J. Xu. An adaptive grid refinement procedure and its asymptotic convergence estimate. *Preprint*, 2005.

[7] C. Bacuta, J.H. Bramble and J. Xu. Regularity estimates for elliptic boundary value problems in Besov spaces. *Mathematics of Computation*, 72(244):1577-1595, 2002.

[8] C. Bacuta, J.H. Bramble and J. Xu. Regularity estimates for elliptic boundary value problems with smooth data on polygonal domains. *Numerische Mathematik*, 11(2):75-94, 2003.

[9] R.E. Bank, A.H. Sherman and A. Weiser. Refinement algorithms and data structures for regular local mesh refinement. In *Scientific Computing*, pages 3-17. IMACS/North-Holland Publishing Company, Amsterdam, 1983.

[10] R.E. Bank and R.K. Smith. A posteriori estimates based on hierarchical basis. *SIAM Journal on Numerical Analysis*, 30:921-935, 1993.

[11] E. Bänsch. Local mesh refinement in 2 and 3 dimensions. *Impact of Computing in Science and Engineering*, 3:181-191, 1991.

[12] E. Bänsch, P. Morin and R.H. Nochetto. An adaptive Uzawa FEM for the Stokes problem: Convergence without the inf-sup condition. *SIAM Journal on Numerical Analysis*, 40(4):1207-1229, 2002.

[13] R. Becker and R. Rannacher. Weighted a-posteriori error control in FEM.

[14] J. Bergh and J. Löfström. *Interpolation Spaces*, Springer, 1976.

[15] W.L. Biggs. *A multigrid tutorial*, SIAM Publications, Philadelphia, 1987 (I)

[16] P. Binev, W. Dahmen and R. DeVore. Adaptive finite element methods with convergence rates. *Numerische Mathematik*, 97(2):219-268, 2004.

[17] P. Binev and R. DeVore. Fast computation in adaptive tree approximation. *Numerische Mathematik*, 97:193-217, 2004.

[18] D. Boffi D, F. Brezzi and L. Gastaldi. *On the problem of spurious eigenvalues in the approximation of linear elliptic problems in mixed form*, Math. Comp. (1999) (to appear) (R).

[19] F. Bornemann, B. Erdmann and R. Kornhuber.A posteriori error estimates for elliptic problems in two and three space dimensions. *SIAM Journal on Numerical Analysis*, 33(3):1188-1204, 1996.

[20] D. Braess. *Finite Elements,* Cambridge University Press, 1997 (I).

[21] J.H. Bramble and X. Zhang. The analysis of multigrid methods. In *Handbook of Numerical Analysis*, Vol. VII, pages 173-415. North-Holland, Amsterdam, 2000.

[22] J.H. Bramble. A proof of the inf-sup condition for the Stokes equations on Lipschitz domains. 13(3), 361-371 (2003).

[23] S. Brenner and L.R. Scott. *A mathematical introduction into the theory of finite elements,* Springer-Verlag, Heidelberg, New-York (1994) (A).

[24] C. Carstensen and R.H. Hoppe. Error reduction and convergence for an adaptive mixed finite element method. *Mathematics of Computation*, 2005.

[25] C. Carstensen and R.H. Hoppe. Convergence analysis of an adaptive nonconforming finite element methods. *Numerische Mathematik*, 103(2):251-266, 2006.

[26] C. Carstensen and R. Verfürth. Edge residuals dominate a posteriori error estimates for low order finite element methods. *SIAM Journal on Numerical Analysis*, 36(5):1571-1587, 1999.

[27] J.M. Cascón and R.H. Nochetto. Convergence and optimal complexity of AFEM for general elliptic PDE and applications. *To appear*, 2006.

[28] P.G. Ciarlet. *The finite element methods for elliptic problems,* North Holland, Amsterdam, 1980.

[29] L. Chen, M. Holst and J. Xu. Convergence and optimality of adaptive mixed finite element methods. *Preprint*, 2006.

[30] B. Cockburn.Discontinuous Galerkin Methods for Convection Dominated Problems. In: High–Order Methods for Computational Physics. Lecture Notes in Computational Science and Engineering 9, Springer, Berlin, 1999.

[31] B. Cockburn, G.,E. Karniadakis and C.-W. Shu (eds.). Discontinuous Galerkin Methods. Lecture Notes in Computational Science and Engineering 11. Springer, Berlin, 2000.

[32] S. Dahlke. Besov regularity for elliptic boundary value problems on polygonal domains. *Appl. Math. Lett.*, 12:31-36, 1999.

[33] S. Dahlke and R.A. DeVore. Besov regularity for elliptic boundary value problems. *Comm. Partial Differential Equations*, 22(1&2):1-16, 1997.

[34] J.R. Dennis Jr. Jr. and R. Schnabel. *Numerical methods for unconstrained optimization,* Prentice-Hall, New York, 1983.

[35] R.A. DeVore and R.C. Sharpley. Besov spaces on domains in $\mathbb{R}^d$. *Transactions of the American Mathematical Society*, 335(2):843-864, 1993.

[36] V. Dolejší, M. Feistauer andC.Schwab. A Finite Volume Discontinous Galerkin Scheme for Nonlinear Convection-Diffusion Problems. Preprint, Forschungsinstitut für Mathematik ETH Zürich, 2001.

[37] V. Dolejší, M. Feistauer and C.Schwab. On some Aspects of the Discontinous Galerkin Finite Element Method for Conservation Laws. Preprint, Mathematics and Computers in Simulation, 2001.

[38] R.A. DeVore. Nonlinear approximation. *Acta Numerica*, 51-150, 1998.

[39] R.A. DeVore and G.G. Lorentz. *Constructive Approximation*, Springer-Verlag, New York, 1993.

[40] J.J. Dongarra, I.S. Duff, D.C. Sorensen and H.A. van der Vorst. *Numerical linear algebra for high-performance computers,* SIAM Publ., Philadelphia, 1998.

[41] W. Dörfler. A convergent adaptive algorithm for Poisson's equation. *SIAM Journal on Numerical Analysis*, 33:1106-1124, 1996.

[42] W. Dörfler and R.H. Nochetto. Small data oscillation implies the saturation assumption. *Numerische Mathematik*, 91(1):1-12, 2002..

[43] K. Eriksson, D. Estep, P. Hansbo and C. Johnson. *Computational Differential Equations,* Cambridge University Press, 1996.

[44] P. Frauenfelder, and C. Lage. Concepts – An Object-Oriented Software Package for Partial Differential Equations. Research Report No. 2002-09, Seminar für Angewandte Mathematik ETH Zürich, 2002.

[45] I. Fried. Condition of finite element matrices generated from nonuniform meshes. *AIAA J.*, 10:219-221,1972.

[46] A. George and J.W.H. Liu. *Computer solution of large, sparse, psotive definite systems,* Prentice Hall Inc. Englewood Cliffs, 1983.

[47] E. Godlewski and P.-A. Raviart. Numerical Approximation of Hyperbolic Systems of Conservation Laws. Applied Mathematical Sciences 118, Springer, New York, 1996.

[48] G.H. Golub and Ch.F. van Loan. *Matrix Computations,* (3rd ed.), John Hopkins University Press, Baltimore, 1996.

[49] G.H. Golub and Ch.F. van Loan. *Matrix Computations,* 4th Ed., Johns Hopkins Univ. Press, Baltimore, (1996) (I).

[50] P. Grisvard. *Elliptic problems in nonsmooth domains*. Pitman, Boston, 1995.

[51] B. Gustafsson, H.O. Kreiss and J. Oliger. *Time dependent problems and difference methods*, Wiley Publ. New York, (1995) (A).

[52] W. Hackbusch. *Multigrid methods and applications*, Springer Verlag, 1985 (A).

[53] W. Hackbusch. *Theorie und Numerik elliptischer Differentialgleichungen*, Teubner Verlag, Stuttgart, 1986 (A).

[54] E. Hairer, S.P. Norsett and G. Wanner. *Solving ordinary differential equations*, Vol. 1 (2nd Ed.) and Volume 2, Springer Series in Computational Mathematics, Springer Verlag, Heidelberg, (1991) (A).

[55] C. Hirsch. *Numerical Computation of Internal and External flows*, Vol. 1 and 2. Wiley Publ. New York, (1990), (A).

[56] D. Kröner. Numerical Schemes for Conservation Laws. Teubner, Stuttgart, 1997.

[57] D. Levi and E. Tadmor. *From Semidiscrete to fully discrete: Stability of Runge-Kutta schemes by the energy method*, SIAM Review **40** (1998), 40-73 (A).

[58] K. Mekchay and R. Nochetto. Convergence of adaptive finite element methods for general second order linear elliptic PDE. *SIAM Journal on Numerical Analysis*, 43(5):1803-1827, 2005.

[59] W.F. Mitchell. A comparison of adaptive refinement techniques for elliptic problems. *ACM Transactions on Mathematical Software (TOMS) archive*, 15(4):326-347, 1989.

[60] P. Morin, R.H. Nochetto and K.G. Siebert. Data oscillation and convergence of adaptive FEM. *SIAM Journal on Numerical Analysis*, 38(20:466-488, 2000.

[61] P. Morin, R.H. Nochetto and K.G. Siebert. Convergence of adaptive finite element methods. *SIAM Review*, 44(4):631-658, 2002.

[62] P. Morin, R.H. Nochetto and K.G. Siebert. Local problems on stars: A posteriori error estimators, convergence, and performance. *Mathematics of Computation*, 72:1067-1097, 2003.

[63] R. Nochetto. Removing the saturation assumption in a posteriori error analysis. *Istit. Lombardo Accad. Sci. Lett. Rend.*, A 127:67-82, 1993.

[64] J. Nečas. *Les Méthodes Directes en Théorie des Equations Elliptiques*, Masson, Paris, 1967 (A).

[65] O.A. Oleinik, A.S. Shamaev and G.A. Yosifian. *Mathematical Problems in Elasticity and Homogenization,* Studies in Mathematics and its Applications Vol. 26, North-Holland Publ. Amsterdam, New York, 1992 (A).

[66] Quarteroni A. and Valli, A., Numerical Approximation of Partial Differential Equations, Springer Verlag, 1991 (I).

[67] Raviart P.A. and J.M. Thomas. *Introduction à l'analyse numérique des équations aux dérivées partielles*. Dunod 1998.

[68] R.D. Richtmyer and K.W. Morton. *Difference methods for initial value problems*, 2nd Ed., Interscience Publ. New York, (1967) (A).

[69] M.C. Rivara. Design and data structure for fully adaptive, multigrid finite element software. *ACM Trans. Math. Soft.*, 10:242-264, 1984.

[70] M.C. Rivara. Mesh refinement processes based on the generalized bisection of simplices. *SIAM Journal on Numerical Analysis*, 21:604-613, 1984.

[71] Y. Saad. *Iterative methods for sparse linear systems.* PWS Publishing Comp., Boston, London, 1996.

[72] J. Schöberl. NETGEN, `http://www.sfb013.uni-linz.ac.at/~joachim/netgen`

[73] C. Schwab. *p- and hp-Finite Element Methods,* Oxford Univ. Press, 1998 (A).

[74] R. Scott and S. Zhang. Finite element interpolation of nonsmooth functions satisfying boundary conditions. *Mathematics of Computation*, 54:483-493, 1990.

[75] E.G. Sewell.Automatic generation of triangulations for piecewise polynomial approximation. In: Ph.D. Dissertation, Purdue University, West Lafayette, Ind., 1972.

[76] C.-W. Shu. Total-Variation-Diminishing Time Discretizations, SIAM J. Sci. Statist. Comput. 9 (1988), 1073-1084.

[77] R. Stevenson. *Optimality of a standard adaptive finite element method.* Department of Mathematics, 2005.

[78] J.C. Strikwerda. *Finite difference schemes and partial differential equations*. Wadsworth & Brooks/Cole Publ. Pacific Grove, California, (1989) (I).

[79] V. Thomée. *Galerkin Finite Element Methods for Parabolic Problems*, Springer Series Comp. Math. **25**, Springer Verlag, (1997) (A).

[80] L.N. Trefethen. *Numerical solution of ordinary and partial differential equations*, SIAM Publ. Philadelphia, (1997) (I).

[81] H. Triebel. *Interpolation theory, function spaces, and differential operators.* North-Holland, Amsterdam, 1978.

[82] R. Verfürth. *A review of a-posteriori error estimation and adaptive mesh refinement techniques,* Wiley Teubner, New York, Stuttgart, 1996 (A).

[83] Ch. Wieners. *Conforming discretizations on tetrahedrons, pyramids, prisms and hexahedrons,* Tech. Report, Institut für Computeranwendungen III, Universität Stuttgart, 1998.

[84] J. Xu. Iterative methods by space decomposition and subspace correction. *SIAM Review*, 34:581-613, 1992.

[85] J. Xu and L. Zikatanov. The method of alternating projections and the method of subspace corrections in Hilbert space. *Journal of the American Mathematical Society*, 15:573-597, 2002.

[86] D. Yu. Asymptotically exact a posteriori error estimator for elements of bi-even degree. *Chinese Journal of Numerical Mathematics and Applications*, 13:82-90, 1991.

[87] D. Yu. Asymptotically exact a posteriori error estimator for elements of bi-odd degree. *Chinese Journal of Numerical Mathematics and Applications*, 13:64-78, 1991.