Chapter 3

Introduction to Sobolev Spaces

Remark 3.1 *Contents.* Sobolev spaces are the basis of the theory of weak or variational forms of partial differential equations. A very popular approach for discretizing partial differential equations, the finite element method, is based on variational forms. In this chapter, a short introduction into Sobolev spaces will be given. Recommended literature are the books Adams (1975); Adams and Fournier (2003) and Evans (2010). □

3.1 Elementary Inequalities

Lemma 3.2 Inequality for strictly monotonically increasing function. *Let $f : \mathbb{R}_+ \cup \{0\} \to \mathbb{R}$ be a continuous and strictly monotonically increasing function with $f(0) = 0$ and $f(x) \to \infty$ for $x \to \infty$. Then, for all $a, b \in \mathbb{R}_+ \cup \{0\}$ it is*

$$ab \le \int_0^a f(x)\ dx + \int_0^b f^{-1}(y)\ dy,$$

where $f^{-1}(y)$ is the inverse of $f(x)$.

Proof: Since $f(x)$ is strictly monotonically increasing, the inverse function exists. The proof is based on a geometric argument, see Figure 3.1.

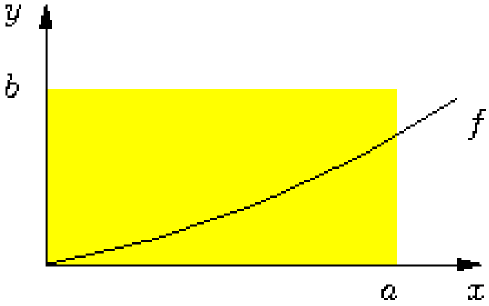

Figure 3.1: Sketch to the proof of Lemma 3.2.

Consider the interval $(0, a)$ on the x-axis and the interval $(0, b)$ on the y-axis. Then, the area of the corresponding rectangle is given by ab, $\int_0^a f(x)\ dx$ is the area below the curve, and $\int_0^b f^{-1}(y)\ dy$ is the area between the positive y-axis and the curve. From Figure 3.1, the inequality follows immediately. The equal sign holds only iff $f(a) = b$. ■

Remark 3.3 *Young's inequality.* Young's inequality

$$ab \le \frac{\varepsilon}{2}a^2 + \frac{1}{2\varepsilon}b^2 \quad \forall\ a, b, \varepsilon \in \mathbb{R}_+ \tag{3.1}$$

follows from Lemma 3.2 with $f(x) = \varepsilon x$, $f^{-1}(y) = \varepsilon^{-1}y$. It is also possible to derive this inequality from the binomial theorem. For proving the generalized Young inequality

$$ab \le \frac{\varepsilon^p}{p}a^p + \frac{1}{q\varepsilon^q}b^q, \quad \forall\ a, b, \varepsilon \in \mathbb{R}_+ \tag{3.2}$$

with $p^{-1} + q^{-1} = 1, p, q \in (1, \infty)$, one chooses $f(x) = x^{p-1}$, $f^{-1}(y) = y^{1/(p-1)}$ and applies Lemma 3.2 with intervals where the upper bounds are given by εa and $\varepsilon^{-1}b$. □

Remark 3.4 *Cauchy–Schwarz inequality.* The Cauchy[1]–Schwarz[2] inequality (for vectors, for sums)

$$|(\mathbf{x}, \mathbf{y})| \le \|\mathbf{x}\|_2 \|\mathbf{y}\|_2 \ \forall\ \mathbf{x}, \mathbf{y} \in \mathbb{R}^n, \tag{3.3}$$

where $(\cdot, \cdot)$ is the Euclidean product and $\|\cdot\|_2$ the Euclidean norm, is well known. One can prove this inequality with the help of Young's inequality.

First, it is clear that the Cauchy–Schwarz inequality is correct if one of the vectors is the zero vector. Now, let $\mathbf{x}, \mathbf{y}$ with $\|\mathbf{x}\|_2 = \|\mathbf{y}\|_2 = 1$. One obtains with the triangle inequality and Young's inequality (3.1)

$$|(\mathbf{x}, \mathbf{y})| = \left|\sum_{i=1}^n x_i y_i\right| \le \sum_{i=1}^n |x_i|\,|y_i| \le \frac{1}{2}\sum_{i=1}^n |x_i|^2 + \frac{1}{2}\sum_{i=1}^n |y_i|^2 = 1.$$

Hence, the Cauchy–Schwarz inequality is correct for $\mathbf{x}, \mathbf{y}$. Last, one considers arbitrary vectors $\tilde{\mathbf{x}} \ne \mathbf{0}, \tilde{\mathbf{y}} \ne \mathbf{0}$. Now, one can utilize the homogeneity of the Cauchy–Schwarz inequality. From the validity of the Cauchy–Schwarz inequality for $\mathbf{x}$ and $\mathbf{y}$, one obtains by a scaling argument

$$\big|(\underbrace{\|\tilde{\mathbf{x}}\|_2^{-1}\tilde{\mathbf{x}}}_{\mathbf{x}}, \underbrace{\|\tilde{\mathbf{y}}\|_2^{-1}\tilde{\mathbf{y}}}_{\mathbf{y}})\big| \le 1$$

Both vectors $\mathbf{x}, \mathbf{y}$ have the Euclidean norm 1, hence

$$\frac{1}{\|\tilde{\mathbf{x}}\|_2 \|\tilde{\mathbf{y}}\|_2} |(\tilde{\mathbf{x}}, \tilde{\mathbf{y}})| \le 1 \quad \Longleftrightarrow \quad |(\tilde{\mathbf{x}}, \tilde{\mathbf{y}})| \le \|\tilde{\mathbf{x}}\|_2 \|\tilde{\mathbf{y}}\|_2 .$$

The generalized Cauchy–Schwarz inequality or Hölder inequality

$$|(\mathbf{x}, \mathbf{y})| \le \left(\sum_{i=1}^n |x_i|^p\right)^{1/p} \left(\sum_{i=1}^n |y_i|^q\right)^{1/q}$$

with $p^{-1} + q^{-1} = 1, p, q \in (1, \infty)$, can be proved in the same way with the help of the generalized Young inequality. □

Definition 3.5 *Lebesgue spaces.* The space of functions which are Lebesgue integrable on Ω to the power of $p \in [1, \infty)$ is denoted by

$$L^p(\Omega) = \left\{ f \ :\ \int_\Omega |f|^p(\mathbf{x})\ d\mathbf{x} < \infty \right\},$$

[1] Augustin Louis Cauchy (1789 – 1857)

[2] Hermann Amandus Schwarz (1843 – 1921)

which is equipped with the norm

$$\|f\|_{L^p(\mathbf{x})} = \left(\int_\Omega |f|^p(\mathbf{x})\ d\mathbf{x}\right)^{1/p}.$$

For $p = \infty$, this space is

$$L^\infty(\Omega) = \{f\ :\ |f(\mathbf{x})| < \infty \text{ almost everywhere in } \Omega\}$$

with the norm

$$\|f\|_{L^\infty(\Omega)} = \text{ess sup}_{\mathbf{x}\in\Omega}|f(\mathbf{x})|.$$

□

Lemma 3.6 Hölder's inequality. *Let $p^{-1} + q^{-1} = 1, p, q \in [1,\infty]$. If $u \in L^p(\Omega)$ and $v \in L^q(\Omega)$, then it is $uv \in L^1(\Omega)$ and it holds that*

$$\|uv\|_{L^1(\Omega)} \le \|u\|_{L^p(\Omega)} \|v\|_{L^q(\Omega)}. \tag{3.4}$$

If $p = q = 2$, then this inequality is also known as Cauchy–Schwarz inequality

$$\|uv\|_{L^1(\Omega)} \le \|u\|_{L^2(\Omega)} \|v\|_{L^2(\Omega)}. \tag{3.5}$$

Proof: $p, q \in (1,\infty)$. First, one has to show that $|uv(\mathbf{x})|$ can be estimated from above by an integrable function. Setting in the generalized Young inequality (3.2) $\varepsilon = 1$, $a = |u(\mathbf{x})|$, and $b = |v(\mathbf{x})|$ gives

$$|u(\mathbf{x})v(\mathbf{x})| \le \frac{1}{p}|u(\mathbf{x})|^p + \frac{1}{q}|v(\mathbf{x})|^q.$$

Since the right hand side of this inequality is integrable, by assumption, it follows that $uv \in L^1(\Omega)$. In addition, Hölder's inequality is proved for the case $\|u\|_{L^p(\Omega)} = \|v\|_{L^q(\Omega)} = 1$ using this inequality

$$\int_\Omega |u(\mathbf{x})v(\mathbf{x})|\ d\mathbf{x} \le \frac{1}{p}\int_\Omega |u(\mathbf{x})|^p\ d\mathbf{x} + \frac{1}{q}\int_\Omega |v(\mathbf{x})|^q\ d\mathbf{x} = 1.$$

The general inequality follows, for the case that both functions do not vanish almost everywhere, with the same homogeneity argument as used for proving the Cauchy–Schwarz inequality of sums. In the case that one of the functions vanishes almost everywhere, (3.4) is trivially satisfied.

$p = 1, q = \infty$. It is

$$\int_\Omega |u(\mathbf{x})v(\mathbf{x})|\ d\mathbf{x} \le \int_\Omega |u(\mathbf{x})| \text{ess sup}_{\mathbf{x}\in\Omega}|v(\mathbf{x})|\ d\mathbf{x} = \|u\|_{L^1(\Omega)} \|v\|_{L^\infty(\Omega)}.$$

■

3.2 Weak Derivative and Distributions

Remark 3.7 *Contents.* This section introduces a generalization of the derivative which is needed for the definition of weak or variational problems. For an introduction to the topic of this section, see, e.g., Haroske and Triebel (2008)

Let $\Omega \subset \mathbb{R}^d$ be a domain with boundary $\Gamma = \partial\Omega$, $d \in \mathbb{N}$, $\Omega \neq \emptyset$. A domain is always an open set. □

Definition 3.8 The space $C_0^\infty(\Omega)$. The space of infinitely often differentiable real functions with compact (closed and bounded) support in Ω is denoted by $C_0^\infty(\Omega)$

$$C_0^\infty(\Omega) = \{v\ :\ v \in C^\infty(\Omega),\ \text{supp}(v) \subset \Omega\},$$

where

$$\text{supp}(v) = \overline{\{\mathbf{x} \in \Omega\ :\ v(\mathbf{x}) \neq 0\}}.$$

□

Definition 3.9 Convergence in $C_0^\infty(\Omega)$. The sequence of functions $\{\phi_n(\mathbf{x})\}_{n=1}^\infty$, $\phi_n \in C_0^\infty(\Omega)$ for all n, is said to convergence to the zero functions if and only if

a) $\exists K \subset \Omega, K$ compact (closed and bounded) with $\mathrm{supp}(\phi_n) \subset K$ for all n,

b) $D^{\boldsymbol{\alpha}}\phi_n(\mathbf{x}) \to 0$ for $n \to \infty$ on K for all multi-indices $\alpha = (\alpha_1, \ldots, \alpha_d)$, $|\alpha| = \alpha_1 + \ldots + \alpha_d$.

It is

$$\lim_{n\to\infty} \phi_n(\mathbf{x}) = \phi(\mathbf{x}) \quad \Longleftrightarrow \quad \lim_{n\to\infty} (\phi_n(\mathbf{x}) - \phi(\mathbf{x})) = 0.$$

□

Definition 3.10 Weak derivative. Let $f, F \in L^1_{\mathrm{loc}}(\Omega)$. ($L^1_{\mathrm{loc}}(\Omega)$: for each compact subset $\Omega' \subset \Omega$ it holds

$$\int_{\Omega'} |u(\mathbf{x})| \ d\mathbf{x} < \infty \ \forall \ u \in L^1_{\mathrm{loc}}(\Omega).)$$

If for all functions $g \in C_0^\infty(\Omega)$ it holds that

$$\int_\Omega F(\mathbf{x}) g(\mathbf{x}) \ d\mathbf{x} = (-1)^{|\boldsymbol{\alpha}|} \int_\Omega f(\mathbf{x}) D^{\boldsymbol{\alpha}} g(\mathbf{x}) \ d\mathbf{x},$$

then $F(\mathbf{x})$ is called weak derivative of $f(\mathbf{x})$ with respect to the multi-index $\boldsymbol{\alpha}$. □

Remark 3.11 *On the weak derivative.*

- One uses the same notations for the derivative as in the classical case : $F(\mathbf{x}) = D^{\boldsymbol{\alpha}} f(\mathbf{x})$.
- If $f(\mathbf{x})$ is classically differentiable on Ω, then the classical derivative is also the weak derivative.
- The assumptions on $f(\mathbf{x})$ and $F(\mathbf{x})$ are such that the integrals in the definition of the weak derivative are well defined. In particular, since the test functions vanish in a neighborhood of the boundary, the behavior of $f(\mathbf{x})$ and $F(\mathbf{x})$ if $\mathbf{x}$ approaches the boundary is not of importance.
- The main aspect of the weak derivative is due to the fact that the (Lebesgue) integral is not influenced from the values of the functions on a set of (Lebesgue) measure zero. Hence, the weak derivative is uniquely defined only up to a set of measure zero. It follows that $f(\mathbf{x})$ might be not classically differentiable on a set of measure zero, e.g., in a point, but it can still be weakly differentiable.
- The weak derivative is uniquely determined, in the sense described above.

□

Example 3.12 *Weak derivative.* The weak derivative of the function $f(x) = |x|$ is

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

In $x = 0$, one can use also any other real number. The proof of this statement follows directly from the definition and it is left as an exercise. □

Definition 3.13 Distribution. A continuous linear functional defined on $C_0^\infty(\Omega)$ is called distribution. The set of all distributions is denoted by $(C_0^\infty(\Omega))'$.

Let $u \in C_0^\infty(\Omega)$ and $\psi \in (C_0^\infty(\Omega))'$, then the following notation is used for the application of the distribution to the function

$$\psi(u(\mathbf{x})) = \langle \psi, u \rangle \in \mathbb{R}.$$

□

Remark 3.14 *On distributions.* Distributions are a generalization of functions. They assign each function from $C_0^\infty(\Omega)$ a real number. □

Example 3.15 *Regular distribution.* Let $u(\mathbf{x}) \in L^1_{\text{loc}}(\Omega)$. Then, a distribution is defined by

$$\int_\Omega u(\mathbf{x})\phi(\mathbf{x})\ d\mathbf{x} = \langle \psi, \phi \rangle,\ \forall \phi \in C_0^\infty(\Omega).$$

This distribution will be identified with $u(\mathbf{x}) \in L^1_{\text{loc}}(\Omega)$.

Distributions with such an integral representation are called regular, otherwise they are called singular. □

Example 3.16 *Dirac distribution.* Let $\boldsymbol{\xi} \in \Omega$ fixed, then

$$\langle \delta_{\boldsymbol{\xi}}, \phi \rangle = \phi(\boldsymbol{\xi})\ \forall\ \phi \in C_0^\infty(\Omega)$$

defines a singular distribution, the so-called Dirac distribution or δ-distribution. It is denoted by $\delta_{\boldsymbol{\xi}} = \delta(\mathbf{x} - \boldsymbol{\xi})$. □

Definition 3.17 *Derivatives of distributions.* Let $\phi \in (C_0^\infty(\Omega))'$ be a distribution. The distribution $\psi \in (C_0^\infty(\Omega))'$ is called derivative in the sense of distributions or distributional derivative of ϕ if

$$\langle \psi, u \rangle = (-1)^{|\boldsymbol{\alpha}|} \langle \phi, D^{\boldsymbol{\alpha}} u \rangle\ \forall u \in C_0^\infty(\Omega),$$

$\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_d)$, $\alpha_j \geq 0, j = 1, \ldots, d$, $|\boldsymbol{\alpha}| = \alpha_1 + \ldots + \alpha_d$. □

Remark 3.18 *On derivatives of distributions.* Each distribution has derivatives in the sense of distributions of arbitrary order.

If the derivative in the sense of distributions $D^\alpha u(\mathbf{x})$ of $u(\mathbf{x}) \in L^1_{\text{loc}}(\Omega)$ is a regular distribution, then also the weak derivative of $u(\mathbf{x})$ exists and both derivatives are identified. □

3.3 Lebesgue Spaces and Sobolev Spaces

Remark 3.19 *On the spaces $L^p(\Omega)$.* These spaces were introduced in Definition 3.5.

- The elements of $L^p(\Omega)$ are, strictly speaking, equivalence classes of functions which are different only on a set of Lebesgue measure zero.
- The spaces $L^p(\Omega)$ are Banach spaces (complete normed spaces). A space X is complete, if each so-called Cauchy sequence $\{u_n\}_{n=0}^\infty \in X$, i.e., for all $\varepsilon > 0$ there is an index $n_0(\varepsilon)$ such that for all $i, j > n_0(\varepsilon)$

  $$\|u_i - u_j\|_X < \varepsilon.$$

  converges and the limit is an element of X.
- The space $L^2(\Omega)$ becomes a Hilbert spaces with the inner product

  $$(f, g) = \int_\Omega f(\mathbf{x})g(\mathbf{x})\ d\mathbf{x}, \quad \|f\|_{L^2} = (f, f)^{1/2}, \quad f, g \in L^2(\Omega).$$

- The dual space of a space X is the space of all bounded linear functionals defined on X. Let Ω be a domain with sufficiently smooth boundary Γ. of the Lebesgue spaces $L^p(\Omega)$, $p \in [1, \infty]$, then

  $$\begin{aligned} (L^p(\Omega))' &= L^q(\Omega) \text{ with } p, q \in (1, \infty),\ \frac{1}{p} + \frac{1}{q} = 1, \\ \left(L^1(\Omega)\right)' &= L^\infty(\Omega), \\ (L^\infty(\Omega))' &\neq L^1(\Omega). \end{aligned}$$

The spaces $L^1(\Omega)$, $L^\infty(\Omega)$ are not reflexive, i.e., the dual space of the dual space is not the original space again.

□

Definition 3.20 Sobolev[3] spaces. Let $k \in \mathbb{N} \cup \{0\}$ and $p \in [1, \infty]$, then the Sobolev space $W^{k,p}(\Omega)$ is defined by

$$W^{k,p}(\Omega) := \{u \in L^p(\Omega) \ : \ D^{\boldsymbol{\alpha}} u \in L^p(\Omega) \ \forall \ \boldsymbol{\alpha} \text{ with } |\boldsymbol{\alpha}| \le k\}.$$

This space is equipped with the norm

$$\|u\|_{W^{k,p}(\Omega)} := \sum_{|\boldsymbol{\alpha}| \le k} \|D^{\boldsymbol{\alpha}} u\|_{L^p(\Omega)}. \tag{3.6}$$

□

Remark 3.21 *On the spaces $W^{k,p}(\Omega)$.*

- Definition 3.20 has the following meaning. From $u \in L^p(\Omega)$, $p \in [1, \infty)$, it follows in particular that $u \in L^1_{\text{loc}}(\Omega)$, such that $u(\mathbf{x})$ defines (represents) a distribution. Then, all derivatives $D^{\boldsymbol{\alpha}} u$ exist in the sense of distributions. The statement $D^{\boldsymbol{\alpha}} u \in L^p(\Omega)$ means that the distribution $D^{\boldsymbol{\alpha}} u \in (C_0^\infty(\Omega))'$ can be represented by a function from $L^p(\Omega)$.
- One can add elements from $W^{k,p}(\Omega)$ and one can multiply them with real numbers. The result is again a function from $W^{k,p}(\Omega)$. With this property, the space $W^{k,p}(\Omega)$ becomes a vector space (linear space). It is straightforward to check that (3.6) is a norm. (*exercise*)
- It is $D^{\boldsymbol{\alpha}} u(\mathbf{x}) = u(\mathbf{x})$ for $\boldsymbol{\alpha} = (0, \ldots, 0)$ and $W^{0,p}(\Omega) = L^p(\Omega)$.
- The spaces $W^{k,p}(\Omega)$ are Banach spaces.
- Sobolev spaces have for $p \in [1, \infty)$ a countable basis $\{\varphi_n(\mathbf{x})\}_{n=1}^\infty$ (Schauder basis), i.e., each element $u(\mathbf{x})$ can be written in the form

$$u(\mathbf{x}) = \sum_{n=1}^{\infty} u_n \varphi_n(\mathbf{x}), \quad u_n \in \mathbb{R} \ n = 1, \ldots, \infty.$$

- Sobolev spaces are uniformly convex for $p \in (1, \infty)$, i.e., for each $\varepsilon \in (0, 2]$ (note that the largest distance in the ball is equal to 2) there is a $\delta(\varepsilon) > 0$ such that for all $u, v \in W^{k,p}(\Omega)$ with $\|u\|_{W^{k,p}(\Omega)} = \|v\|_{W^{k,p}(\Omega)} = 1$, and $\|u - v\|_{W^{k,p}(\Omega)} > \varepsilon$ it holds that $\left\|\frac{u+v}{2}\right\|_{W^{k,p}(\Omega)} \le 1 - \delta(\varepsilon)$, see Figure 3.2 for an illustration.
- Sobolev spaces are reflexive for $p \in (1, \infty)$.
- On can show that $C^\infty(\Omega)$ is dense in $W^{k,p}(\Omega)$, e.g., see (Alt, 1999, Satz 1.21, Satz 2.10) or (Adams, 1975, Lemma 3.15). With this property, one can characterize the Sobolev spaces $W^{k,p}(\Omega)$ as completion of the functions from $C^\infty(\Omega)$ with respect to the norm (3.6). For domains with smooth boundary, one can even show that $C^\infty(\overline{\Omega})$ is dense in $W^{k,p}(\Omega)$.
- The Sobolev space $H^k(\Omega) = W^{k,2}(\Omega)$ is a Hilbert space with the inner product

$$(u, v)_{H^k(\Omega)} = \sum_{|\boldsymbol{\alpha}| \le k} \int_\Omega D^{\boldsymbol{\alpha}} u(\mathbf{x}) D^{\boldsymbol{\alpha}} v(\mathbf{x}) \ d\mathbf{x}$$

and the norm $\|u\|_{H^k(\Omega)} = (u, u)^{1/2}$.

□

[3] Sergei Lvovich Sobolev (1908 – 1989)

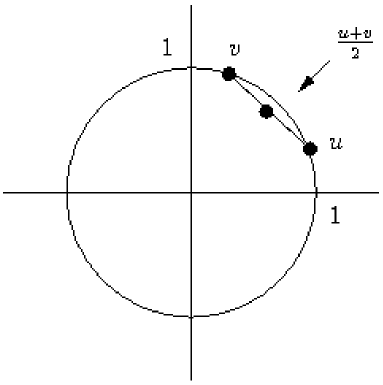

Figure 3.2: Illustration of the uniform convexity of Sobolev spaces.

Definition 3.22 The space $W_0^{k,p}(\Omega)$. The Sobolev space $W_0^{k,p}(\Omega)$ is defined as the completion of $C_0^\infty(\Omega)$ in the norm of $W^{k,p}(\Omega)$

$$W_0^{k,p}(\Omega) = \overline{C_0^\infty(\Omega)}^{\|\cdot\|_{W^{k,p}(\Omega)}}.$$

□

3.4 The Trace of a Function from a Sobolev Space

Remark 3.23 *Motivation.* This class considers boundary value problems for partial differential equations. In the theory of weak or variational solutions, the solution of the partial differential equation is searched in an appropriate Sobolev space. Then, for the boundary value problem, this solution has to satisfy the boundary condition. However, since the boundary of a domain is a manifold of dimension $(d-1)$, and consequently it has Lebesgue measure zero, one has to clarify how a function from a Sobolev space is defined on this manifold. This definition will be presented in this section. □

Definition 3.24 Boundary of class $C^{k,\alpha}$. A bounded domain $\Omega \subset \mathbb{R}^d$ and its boundary Γ are of class $C^{k,\alpha}$, $0 \le \alpha \le 1$ if for all $\mathbf{x}_0 \in \Gamma$ there is a ball $B(\mathbf{x}_0, r)$ with $r > 0$ and a bijective map $\psi\ :\ B(\mathbf{x}_0, r) \to D \subset \mathbb{R}^d$ such that
1) $\psi\left(B(\mathbf{x}_0, r) \cap \Omega\right) \subset \mathbb{R}^d_+$,
2) $\psi\left(B(\mathbf{x}_0, r) \cap \Gamma\right) \subset \partial\mathbb{R}^d_+$,
3) $\psi \in C^{k,\alpha}(B(\mathbf{x}_0, r)), \psi^{-1} \in C^{k,\alpha}(D)$, are Hölder continuous.
That means, Γ is locally the graph of a function with $d-1$ arguments. (A function $u(\mathbf{x})$ is Hölder continuous if

$$\|u\|_{C^{k,\alpha}(\Omega)} = \sum_{|\boldsymbol{\alpha}| \le k} \|D^{\boldsymbol{\alpha}} u\|_{C(\overline{\Omega})} + \sum_{|\boldsymbol{\alpha}| = k} [D^{\boldsymbol{\alpha}} u]_{C^{0,\alpha}(\overline{\Omega})}$$

with

$$[D^{\boldsymbol{\alpha}} u]_{C^{0,\alpha}(\overline{\Omega})} = \sup_{\mathbf{x},\mathbf{y} \in \Omega} \left\{ \frac{|u(\mathbf{x}) - u(\mathbf{y})|}{|\mathbf{x} - \mathbf{y}|^\alpha} \right\}$$

is finite.) □

Remark 3.25 *Lipschitz boundary.* It will be generally assumed that the boundary of Ω is of class $C^{0,1}$. That means, the map is Lipschitz[4] continuous. Such a boundary

[4] Rudolf Otto Sigismund Lipschitz (1832 – 1903)

is simply called Lipschitz boundary and the domain is called Lipschitz domain. An important feature of a Lipschitz boundary is that the outer normal vector is defined almost everywhere at the boundary and it is almost everywhere continuous. □

Example 3.26 *On Lipschitz domains.*

- Domains with Lipschitz boundary are, for example, balls or polygonal domains in two dimensions where the domain is always on one side of the boundary.
- A domain which is not a Lipschitz domain is a circle with a slit

  $$\Omega = \{(x,y)\ :\ x^2+y^2<1\} \setminus \{(x,y)\ :\ x \ge 0, y=0\}.$$

  At the slit, the domain is on both sides of the boundary.
- In three dimension, a polyhedral domain is not not necessarily a Lipschitz domain. For instance, if the domain is build of two bricks which are laying on each other like in Figure 3.3, then the boundary is not Lipschitz continuous where the edge of one brick meets the edge of the other brick.

□

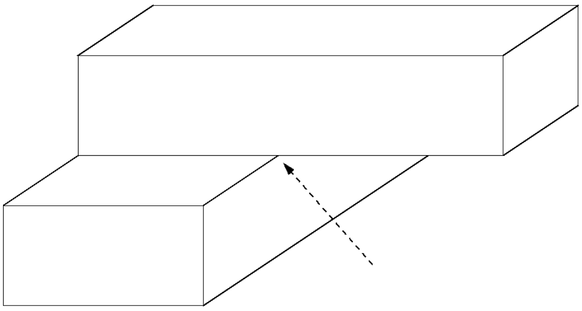

Figure 3.3: Polyhedral domain in three dimensions which is not Lipschitz continuous (at the corner where the arrow points to).

Theorem 3.27 Trace theorem. *Let $\Omega \subset \mathbb{R}^d$, $d \ge 2$, with a Lipschitz boundary. Then, there is exactly one linear and continuous operator $\gamma : W^{1,p}(\Omega) \to L^p(\Gamma)$, $p \in [1,\infty)$, which gives for functions $u \in C(\overline{\Omega}) \cap W^{1,p}(\Omega)$ the classical boundary values*

$$\gamma u(\mathbf{x}) = u(\mathbf{x}),\ \mathbf{x} \in \Gamma,\ \forall\ u \in C(\overline{\Omega}) \cap W^{1,p}(\Omega),$$

i.e., $\gamma u(\mathbf{x}) = u(\mathbf{x})|_{\mathbf{x}\in\Gamma}$.

Proof: The proof can be found in the literature, e.g., in Adams (1975); Adams and Fournier (2003). ■

Remark 3.28 *On the trace.* The operator γ is called trace or trace operator.

- Since a linear and continuous operator is bounded, there is a constant $C > 0$ with

  $$\|\gamma u\|_{L^p(\Gamma)} \le C \|u\|_{W^{1,p}(\Omega)}\ \forall\ u \in W^{1,p}(\Omega)$$

  or

  $$\|\gamma\|_{\mathcal{L}(W^{1,p}(\Omega),L^p(\Gamma))} \le C.$$

- By definition of the trace, one gets for $u \in C(\overline{\Omega})$ the classical boundary values. By the density of $C^\infty(\overline{\Omega})$ in $W^{1,p}(\Omega)$ for domains with smooth boundary, it follows that $C(\overline{\Omega})$ is also dense in $W^{1,p}(\Omega)$ such that for all $u \in W^{1,p}(\Omega)$ there is a sequence $\{u_n\}_{n=1}^\infty \in C(\overline{\Omega})$ with $u_n \to u$ in $W^{1,p}(\Omega)$. Then, the trace of u is defined to be $\gamma u = \lim_{k\to\infty}(\gamma u_k)$.

- It is

$$\begin{aligned} \gamma u(\mathbf{x}) &= 0 \quad \forall\, u \in W_0^{1,p}(\Omega), \\ \gamma D^{\boldsymbol{\alpha}} u(\mathbf{x}) &= 0 \quad \forall\, u \in W_0^{k,p}(\Omega), |\boldsymbol{\alpha}| \le k-1. \end{aligned} \tag{3.7}$$

□

3.5 Sobolev Spaces with Non-Integer and Negative Exponents

Remark 3.29 *Motivation.* Sobolev spaces with non-integer and negative exponents are important in the theory of variational formulations of partial differential equations.

Let $\Omega \subset \mathbb{R}^d$ be a domain and $p \in (1,\infty)$ mit $p^{-1} + q^{-1} = 1$. □

Definition 3.30 The space $W^{-k,q}(\Omega)$. The space $W^{-k,q}(\Omega), k \in \mathbb{N} \cup \{0\}$, contains distributions which are defined on $W^{k,p}(\Omega)$

$$W^{-k,q}(\Omega) = \left\{\varphi \in (C_0^\infty(\Omega))' \ : \ \|\varphi\|_{W^{-k,q}} < \infty\right\}$$

with

$$\|\varphi\|_{W^{-k,q}} = \sup_{u \in C_0^\infty(\Omega), u \neq 0} \frac{\langle \varphi, u\rangle}{\|u\|_{W^{k,p}(\Omega)}}.$$

□

Remark 3.31 *On the spaces $W^{-k,p}(\Omega)$.*

- It is $W^{-k,q}(\Omega) = \left[W_0^{k,p}(\Omega)\right]'$, i.e., $W^{-k,q}(\Omega)$ can be identified with the dual space of $W_0^{k,p}(\Omega)$. In particular it is $H^{-1}(\Omega) = \left(H_0^1(\Omega)\right)'$.
- It is

$$\ldots \subset W^{2,p}(\Omega) \subset W^{1,p}(\Omega) \subset L^p(\Omega) \subset W^{-1,q}(\Omega) \subset W^{-2,q}(\Omega) \ldots$$

□

Definition 3.32 Sobolev–Slobodeckij space. Let $s \in \mathbb{R}$, then the Sobolev–Slobodeckij or Sobolev space $H^s(\Omega)$ is defined as follows:

- $s \in \mathbb{Z}$. $H^s(\Omega) = W^{s,2}(\Omega)$.
- $s > 0$ with $s = k + \sigma$, $k \in \mathbb{N} \cup \{0\}$, $\sigma \in (0,1)$. The space $H^s(\Omega)$ contains all functions u for which the following norm is finite:

  $$\|u\|_{H^s(\Omega)}^2 = \|u\|_{H^k(\Omega)}^2 + |u|_{k+\sigma}^2,$$

  with

  $$\begin{aligned} (u,v)_{H^s(\Omega)} &= (u,v)_{H^k} + (u,v)_{k+\sigma}, \quad |u|_{k+\sigma}^2 = (u,u)_{k+\sigma}, \\ (u,v)_{k+\sigma} &= \sum_{|\boldsymbol{\alpha}|=k} \int_\Omega \int_\Omega \frac{\left(D^{\boldsymbol{\alpha}} u(\mathbf{x}) - D^{\boldsymbol{\alpha}} u(\mathbf{y})\right)\left(D^{\boldsymbol{\alpha}} v(\mathbf{x}) - D^{\boldsymbol{\alpha}} v(\mathbf{y})\right)}{\|\mathbf{x}-\mathbf{y}\|_2^{d+2\sigma}}\, d\mathbf{x} d\mathbf{y}, \end{aligned}$$

- $s < 0$. $H^s(\Omega) = \left[H_0^{-s}(\Omega)\right]'$ with $H_0^{-s}(\Omega) = \overline{C_0^\infty(\Omega)}^{\|\cdot\|_{H^{-s}(\Omega)}}$.

□

3.6 Theorem on Equivalent Norms

Definition 3.33 Equivalent norms. Two norms $\|\cdot\|_1$ and $\|\cdot\|_2$ on the linear space X are said to be equivalent if there are constants C_1 and C_2 such that

$$C_1 \|u\|_1 \le \|u\|_2 \le C_2 \|u\|_1 \;\; \forall\, u \in X.$$

$\square$

Remark 3.34 *On equivalent norms.* Many important properties, like continuity or convergence, do not change if an equivalent norm is considered. $\square$

Theorem 3.35 Equivalent norms in $W^{k,p}(\Omega)$. *Let $\Omega \subset \mathbb{R}^d$ be a domain with Lipschitz boundary Γ, $p \in [1,\infty]$, and $k \in \mathbb{N}$. Let $\{f_i\}_{i=1}^l$ be a system with the following properties:*
1) $f_i \,:\, W^{k,p}(\Omega) \to \mathbb{R}_+ \cup \{0\}$ is a semi norm,
2) $\exists C_i > 0$ with $0 \le f_i(v) \le C_i \|v\|_{W^{k,p}(\Omega)}$, $\forall\, v \in W^{k,p}(\Omega)$,
3) f_i is a norm on the polynomials of degree $k-1$, i.e., if for $v \in P_{k-1} = \left\{\sum_{|\boldsymbol{\alpha}|\le k-1} C_{\boldsymbol{\alpha}} x^{\boldsymbol{\alpha}}\right\}$ it holds that $f_i(v) = 0$, $i = 1,\ldots,l$, then it is $v \equiv 0$.
Then, the norm $\|\cdot\|_{W^{k,p}(\Omega)}$ defined in (3.6) *and the norm*

$$\begin{aligned} \|u\|'_{W^{k,p}(\Omega)} &:= \left(\sum_{i=1}^{l} f_i^p(u) + |u|^p_{W^{k,p}(\Omega)}\right)^{1/p} \quad \textit{with} \\ |u|_{W^{k,p}(\Omega)} &= \left(\sum_{|\boldsymbol{\alpha}|=k} \int_\Omega |D^{\boldsymbol{\alpha}} u(\mathbf{x})|^p \; d\mathbf{x}\right)^{1/p} \end{aligned}$$

are equivalent.

Remark 3.36 *On semi norms.* For a semi norm $f_i(\cdot)$, one cannot conclude from $f_i(v) = 0$ that $v = 0$. The third assumptions however states, that this conclusion can be drawn for all polynomials up to a certain degree. $\square$

Example 3.37 *Equivalent norms in Sobolev spaces.*

- The following norms are equivalent to the standard norm in $W^{1,p}(\Omega)$:

$$\begin{aligned} \text{a)} \quad \|u\|'_{W^{1,p}(\Omega)} &= \left(\left|\int_\Omega u \; d\mathbf{x}\right|^p + |u|^p_{W^{1,p}(\Omega)}\right)^{1/p}, \\ \text{b)} \quad \|u\|'_{W^{1,p}(\Omega)} &= \left(\left|\int_\Gamma u \; d\mathbf{s}\right|^p + |u|^p_{W^{1,p}(\Omega)}\right)^{1/p}, \\ \text{c)} \quad \|u\|'_{W^{1,p}(\Omega)} &= \left(\int_\Gamma |u|^p \; d\mathbf{s} + |u|^p_{W^{1,p}(\Omega)}\right)^{1/p}. \end{aligned}$$

- In $W^{k,p}(\Omega)$ it is

$$\|u\|'_{W^{k,p}(\Omega)} = \left(\sum_{i=0}^{k-1} \int_\Gamma \left|\frac{\partial^i u}{\partial \mathbf{n}^i}\right|^p \; d\mathbf{s} + |u|^p_{W^{k,p}(\Omega)}\right)^{1/p}$$

equivalent to the standard norm. Here, $\mathbf{n}$ denotes the outer normal on Γ with $\|\mathbf{n}\|_2 = 1$.

- In the case $W_0^{k,p}(\Omega)$, one does not need the regularity of the boundary. It is

$$\|u\|'_{W_0^{k,p}(\Omega)} = |u|_{W^{k,p}(\Omega)},$$

i.e., in the spaces $W_0^{k,p}(\Omega)$ the standard semi norm is equivalent to the standard norm.
In particular, it is for $u \in H_0^1(\Omega)$ $(k = 1, p = 2)$

$$C_1 \|u\|_{H^1(\Omega)} \le \|\nabla u\|_{L^2(\Omega)} \le C_2 \|u\|_{H^1(\Omega)}.$$

It follows that there is a constant $C > 0$ such that

$$\|u\|_{L^2(\Omega)} \le C \|\nabla u\|_{L^2(\Omega)} \quad \forall\, u \in H_0^1(\Omega). \tag{3.8}$$

□

3.7 Some Inequalities in Sobolev Spaces

Remark 3.38 *Motivation.* This section presents a generalization of the last part of Example 3.37. It will be shown that for inequalities of type (3.8) it is not necessary that the trace vanishes on the complete boundary.

Let $\Omega \subset \mathbb{R}^d$ be a bounded domain with boundary Γ and let $\Gamma_1 \subset \Gamma$ with $\text{meas}_{\mathbb{R}^{d-1}}(\Gamma_1) = \int_{\Gamma_1} d\mathbf{s} > 0$.

One considers the space

$$\begin{aligned} V_0 &= \left\{ v \in W^{1,p}(\Omega) \ : \ v|_{\Gamma_1} = 0 \right\} \subset W^{1,p}(\Omega) \text{ if } \Gamma_1 \subset \Gamma, \\ V_0 &= W_0^{1,p}(\Omega) \text{ if } \Gamma_1 = \Gamma \end{aligned}$$

with $p \in [1, \infty)$. □

Lemma 3.39 Friedrichs[5] inequality, Poincaré[6] inequality, Poincaré–Friedrichs inequality. *Let $p \in [1, \infty)$ and $\text{meas}_{\mathbb{R}^{d-1}}(\Gamma_1) > 0$. Then it is for all $u \in V_0$*

$$\int_\Omega |u(\mathbf{x})|^p \ d\mathbf{x} \le C_P \int_\Omega \|\nabla u(\mathbf{x})\|_2^p \ d\mathbf{x}, \tag{3.9}$$

where $\|\cdot\|_2$ is the Euclidean vector norm.

Proof: The inequality will be proved with the theorem on equivalent norms, Theorem 3.35. Let $f_1(u) \ : \ W^{1,p}(\Omega) \to \mathbb{R}_+ \cup \{0\}$ with

$$f_1(u) = \left(\int_{\Gamma_1} |u(\mathbf{s})|^p \ d\mathbf{s} \right)^{1/p}.$$

This functions has the following properties:
1) $f_1(u)$ is a semi norm.
2) It is

$$\begin{aligned} 0 &\le f_1(u) = \left(\int_{\Gamma_1} |u(\mathbf{s})|^p \ d\mathbf{s} \right)^{1/p} \le \left(\int_{\Gamma} |u(\mathbf{s})|^p \ d\mathbf{s} \right)^{1/p} \\ &= \|u\|_{L^p(\Gamma)} = \|\gamma u\|_{L^p(\Gamma)} \le C \|u\|_{W^{1,p}(\Omega)}. \end{aligned}$$

The last estimate follows from the continuity of the trace operator.

[5] Friedrichs
[6] Poincaré

3) Let $v \in P_0$, i.e., v is a constant. Then, one obtains from

$$0 = f_1(v) = \left(\int_{\Gamma_1} |v(\mathbf{s})|^p \; d\mathbf{s}\right)^{1/p} = |v| \left(\operatorname{meas}_{\mathbb{R}^{d-1}}(\Gamma_1)\right)^{1/p},$$

that $|v| = 0$.

Hence, all assumptions of Theorem 3.35 are satisfied. That means, there are two constants C_1 and $C_2$with

$$\underbrace{C_1 \left(\int_{\Gamma_1} |u(\mathbf{s})|^p \; d\mathbf{s} + \int_\Omega \|\nabla u(\mathbf{x})\|_2^p \; d\mathbf{x}\right)^{1/p}}_{\|u\|'_{W^{1,p}(\Omega)}} \le \|u\|_{W^{1,p}(\Omega)} \le C_2 \|u\|'_{W^{1,p}(\Omega)} \;\; \forall\, u \in W^{1,p}(\Omega).$$

In particular, it follows that

$$\int_\Omega |u(\mathbf{x})|^p \; d\mathbf{x} + \int_\Omega \|\nabla u(\mathbf{x})\|_2^p \; d\mathbf{x} \le C_2^p \left(\int_{\Gamma_1} |u(\mathbf{s})|^p \; d\mathbf{s} + \int_\Omega \|\nabla u(\mathbf{x})\|_2^p \; d\mathbf{x}\right)$$

or

$$\int_\Omega |u(\mathbf{x})|^p \; d\mathbf{x} \le C_P \left(\int_{\Gamma_1} |u(\mathbf{s})|^p \; d\mathbf{s} + \int_\Omega \|\nabla u(\mathbf{x})\|_2^p \; d\mathbf{x}\right)$$

with $C_P = C_2^p$. Since $u \in V_0$ vanishes on Γ_1, the statement of the lemma is proved. ■

Remark 3.40 *On the Poincaré–Friedrichs inequality.* In the space V_0 becomes $|\cdot|_{W^{1,p}}$ a norm which is equivalent to $\|\cdot\|_{W^{1,p}(\Omega)}$. The classical Poincaré–Friedrichs inequality is given for $\Gamma_1 = \Gamma$ and $p = 2$

$$\|u\|_{L^2} \le C_P \|\nabla u\|_{L^2} \;\; \forall\, u \in H_0^1(\Omega),$$

where the constant depends only on the diameter of the domain Ω. □

Lemma 3.41 Another inequality of Poincaré–Friedrichs type. *Let $\Omega' \subset \Omega$ with meas$_{\mathbb{R}^d}(\Omega') = \int_{\Omega'} d\mathbf{x} > 0$, then for all $u \in W^{1,p}(\Omega)$ it is*

$$\int_\Omega |u(\mathbf{x})|^p \; d\mathbf{x} \le C \left(\left|\int_{\Omega'} u(\mathbf{x}) \; d\mathbf{x}\right|^p + \int_\Omega \|\nabla u(\mathbf{x})\|_2^p \; d\mathbf{x}\right).$$

Proof: *Exercise.* ■

3.8 The Gaussian Theorem

Remark 3.42 *Motivation.* The Gaussian theorem is the generalization of the integration by parts from calculus. This operation is very important for the theory of weak or variational solutions of partial differential equations. One has to study, under which conditions on the regularity of the domain and of the functions it is well defined. □

Theorem 3.43 Gaussian theorem. *Let $\Omega \subset \mathbb{R}^d, d \ge 2$, be a bounded domain with Lipschitz boundary Γ. Then, the following identity holds for all $u \in W^{1,1}(\Omega)$*

$$\int_\Omega \partial_i u(\mathbf{x}) \; d\mathbf{x} = \int_\Gamma u(\mathbf{s}) \mathbf{n}_i(\mathbf{s}) \; d\mathbf{s}, \tag{3.10}$$

where $\mathbf{n}$ is the unit outer normal vector on Γ.

Proof: sketch. First of all, one proves the statement for functions from $C^1(\overline{\Omega})$. This proof is somewhat longer and it is referred to the literature, e.g., Evans (2010).

The space $C^1(\overline{\Omega})$ is dense in $W^{1,1}(\Omega)$, see Remark 3.21. Hence, for all $u \in W^{1,1}(\Omega)$ there is a sequence $\{u_n\}_{n=1}^{\infty} \in C^1(\overline{\Omega})$ with

$$\lim_{n\to\infty} \|u - u_n\|_{W^{1,1}(\Omega)} = 0$$

and (3.10) holds for all functions $u_n(\mathbf{x})$. It will be shown that the limit of the left hand side converges to the left hand side of (3.10) and the limit of the right hand side converges to the right hand side of (3.10).

From the convergence in $\|\cdot\|_{W^{1,1}(\Omega)}$, one has in particular

$$\lim_{n\to\infty} \int_\Omega \partial_i u_n(\mathbf{x})\ d\mathbf{x} = \int_\Omega \partial_i u(\mathbf{x})\ d\mathbf{x}.$$

On the other hand, the continuity of the trace operator gives

$$\lim_{n\to\infty} \|u - u_n\|_{L^1(\Gamma)} \le C \lim_{n\to\infty} \|u - u_n\|_{W^{1,1}} = 0,$$

from what follows that

$$\lim_{n\to\infty} \int_\Gamma u_n(\mathbf{s})\ d\mathbf{s} = \int_\Gamma u(\mathbf{s})\ d\mathbf{s}.$$

Since for a Lipschitz boundary, the normal $\mathbf{n}$ is almost everywhere continuous, it is

$$\lim_{n\to\infty} \int_\Gamma u_n(\mathbf{s})\mathbf{n}_i(\mathbf{s})\ d\mathbf{s} = \int_\Gamma u(\mathbf{s})\mathbf{n}_i(\mathbf{s})\ d\mathbf{s}.$$

Thus, the limits lead to (3.10). ■

Corollary 3.44 Vector field. *Let the conditions of Theorem 3.43 on the domain Ω be satisfied and let $\mathbf{u} \in \left(W^{1,1}(\Omega)\right)^d$ be a vector field. Then it is*

$$\int_\Omega \nabla \cdot \mathbf{u}(\mathbf{x})\ d\mathbf{x} = \int_\Gamma \mathbf{u}(\mathbf{s}) \cdot \mathbf{n}(\mathbf{s})\ d\mathbf{s}.$$

Proof: The statement follows by adding (3.10) from $i = 1$ to $i = d$. ■

Corollary 3.45 Integration by parts. *Let the conditions of Theorem 3.43 on the domain Ω be satisfied. Consider $u \in W^{1,p}(\Omega)$ and $v \in W^{1,q}(\Omega)$ with $p \in (1,\infty)$ and $\frac{1}{p} + \frac{1}{q} = 1$. Then it is*

$$\int_\Omega \partial_i u(\mathbf{x}) v(\mathbf{x})\ d\mathbf{x} = \int_\Gamma u(\mathbf{s})v(\mathbf{s})\mathbf{n}_i(\mathbf{s})\ d\mathbf{s} - \int_\Omega u(\mathbf{x})\partial_i v(\mathbf{x})\ d\mathbf{x}.$$

Proof: *exercise.* ■

Corollary 3.46 First Green[7]'s formula. *Let the conditions of Theorem 3.43 on the domain Ω be satisfied, then it is*

$$\int_\Omega \nabla u(\mathbf{x}) \cdot \nabla v(\mathbf{x})\ d\mathbf{x} = \int_\Gamma \frac{\partial u}{\partial \mathbf{n}}(\mathbf{s})v(\mathbf{s})\ d\mathbf{s} - \int_\Omega \Delta u(\mathbf{x}) v(\mathbf{x})\ d\mathbf{x}$$

for all $u \in H^2(\Omega)$ and $v \in H^1(\Omega)$.

Proof: From the definition of the Sobolev spaces it follows that the integrals are well defined. Now, the proof follows the proof of Corollary 3.45, where one has now to sum over the components. ■

[7] Georg Green (1793 – 1841)

Remark 3.47 *On the first Green's formula.* The first Green's formula is the formula of integrating by parts once. The boundary integral can be equivalently written in the form

$$\int_\Gamma \nabla u(\mathbf{s}) \cdot \mathbf{n}(\mathbf{s}) v(\mathbf{s})\ d\mathbf{s}.$$

The formula of integrating by parts twice is called second Green's formula. □

Corollary 3.48 Second Green's formula. *Let the conditions of Theorem 3.43 on the domain Ω be satisfied, then one has*

$$\int_\Omega \big(\Delta u(\mathbf{x})v(\mathbf{x}) - \Delta v(\mathbf{x})u(\mathbf{x})\big)\ d\mathbf{x} = \int_\Gamma \left(\frac{\partial u}{\partial \mathbf{n}}(\mathbf{s})v(\mathbf{s}) - \frac{\partial v}{\partial \mathbf{n}}(\mathbf{s})u(\mathbf{s})\right)\ d\mathbf{s}$$

for all $u, v \in H^2(\Omega)$.

3.9 Sobolev Imbedding Theorems

Remark 3.49 *Motivation.* This section studies the question which Sobolev spaces are subspaces of other Sobolev spaces. With this property, called imbedding, it is possible to estimate the norm of a function in the subspace by the norm in the larger space. □

Lemma 3.50 Imbedding of Sobolev spaces with same integration power p and different orders of the derivative. *Let $\Omega \subset \mathbb{R}^d$ be a domain with $p \in [1, \infty)$ and $k \le m$, then it is $W^{m,p}(\Omega) \subset W^{k,p}(\Omega)$.*

Proof: The statement of this lemma follows directly from the definition of Sobolev spaces, see Definition 3.20. ■

Lemma 3.51 Imbedding of Sobolev spaces with the same order of the derivative k and different integration powers. *Let $\Omega \subset \mathbb{R}^d$ be a bounded domain, $k \ge 0$, and $p, q \in [1, \infty]$ with $q > p$. Then it is $W^{k,q}(\Omega) \subset W^{k,p}(\Omega)$.*

Proof: *exercise.* ■

Remark 3.52 *Imbedding of Sobolev spaces with the same order of the derivative k and the same integration power p in imbedded domains.* Let $\Omega \subset \mathbb{R}^d$ be a domain with sufficiently smooth boundary Γ, $k \ge 0$, and $p \in [1, \infty]$. Then there is a map $E : W^{k,p}(\Omega) \to W^{k,p}(\mathbb{R}^d)$, the so-called (simple) extension, with

- $Ev|_\Omega = v$,
- $\|Ev\|_{W^{k,p}(\mathbb{R}^d)} \le C \|v\|_{W^{k,p}(\Omega)}$, with $C > 0$,

e.g., see (Adams, 1975, Chapter IV) for details. Likewise, the natural restriction $e : W^{k,p}(\mathbb{R}^d) \to W^{k,p}(\Omega)$ can be defined and it is $\|ev\|_{W^{k,p}(\Omega)} \le \|v\|_{W^{k,p}(\mathbb{R}^d)}$. □

Theorem 3.53 A Sobolev inequality. *Let $\Omega \subset \mathbb{R}^d$ be a bounded domain with Lipschitz boundary Γ, $k \ge 0$, and $p \in [1, \infty)$ with*

$$\begin{array}{ll} k \ge d & \textit{for } p = 1, \\ k > d/p & \textit{for } p > 1. \end{array}$$

Then there is a constant C such that for all $u \in W^{k,p}(\Omega)$ it follows that $u \in C_B(\Omega)$, where

$$C_B(\Omega) = \{v \in C(\Omega)\ :\ v \textit{ is bounded}\},$$

and it is

$$\|u\|_{C_B(\Omega)} = \|u\|_{L^\infty(\Omega)} \le C \|u\|_{W^{k,p}(\Omega)}. \tag{3.11}$$

Proof: See literature, e.g., Adams (1975); Adams and Fournier (2003). ■

Remark 3.54 *On the Sobolev inequality.* The Sobolev inequality states that each function with sufficiently many weak derivatives (the number depends on the dimension of Ω and the integration power) can be considered as a continuous and bounded function in Ω. One says that $W^{k,p}(\Omega)$ is imbedded in $C_B(\Omega)$. It is

$$C\left(\overline{\Omega}\right) \subset C_B(\Omega) \subset C(\Omega).$$

Consider $\Omega = (0,1)$ and $f_1(x) = 1/x$ and $f_2(x) = \sin(1/x)$. Then, $f_1 \in C(\Omega)$, $f_1 \notin C_B(\Omega)$ and $f_2 \in C_B(\Omega)$, $f_2 \notin C(\overline{\Omega})$.

Of course, it is possible to apply this theorem to weak derivatives of functions. Then, one obtains imbeddings like $W^{k,p}(\Omega) \to C_B^s(\Omega)$ for $(k-s)p > d, p > 1$. A comprehensive overview on imbeddings can be found in Adams (1975); Adams and Fournier (2003). □

Example 3.55 $H^1(\Omega)$ *in one dimension.* Let $d = 1$ and Ω be a bounded interval. Then, each function from $H^1(\Omega)$ $(k = 1, p = 2)$ is continuous and bounded in Ω. □

Example 3.56 $H^1(\Omega)$ *in higher dimensions.* The functions from $H^1(\Omega)$ are in general not continuous for $d \geq 2$. This property will be shown with the following example.

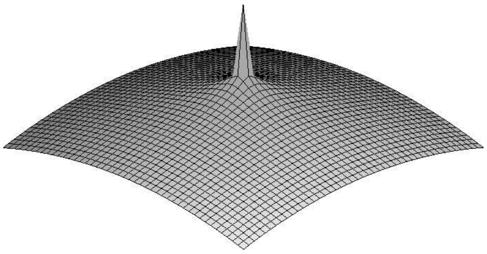

Figure 3.4: The function $f(\mathbf{x})$ of Example 3.56 for $d = 2$.

Let $\Omega = \{\mathbf{x} \in \mathbb{R}^d \ : \ \|\mathbf{x}\|_2 < 1/2\}$ and $f(\mathbf{x}) = \ln|\ln\|\mathbf{x}\|_2|$, see Figure 3.4. For $\|\mathbf{x}\|_2 < 1/2$ it is $|\ln\|\mathbf{x}\|_2| = -\ln\|\mathbf{x}\|_2$ and one gets for $\mathbf{x} \neq \mathbf{0}$

$$\partial_i f(\mathbf{x}) = -\frac{1}{\ln\|\mathbf{x}\|_2}\frac{1}{\|\mathbf{x}\|_2}\frac{x_i}{\|\mathbf{x}\|_2} = -\frac{x_i}{\|\mathbf{x}\|_2^2\ln\|\mathbf{x}\|_2}.$$

For $p \leq d$, one obtains

$$\left|\frac{\partial f}{\partial x_i}(\mathbf{x})\right|^p = \underbrace{\left|\frac{x_i}{\|\mathbf{x}\|_2}\right|^p}_{\leq 1}\underbrace{\left|\frac{1}{\|\mathbf{x}\|_2\ln\|\mathbf{x}\|_2}\right|^p}_{\geq e} \leq \left|\frac{1}{\|\mathbf{x}\|_2\ln\|\mathbf{x}\|_2}\right|^d.$$

The estimate of the second factor can be obtained, e.g., with a discussion of the curve. Using now spherical coordinates, $\rho = e^{-t}$ and S^{d-1} is the unit sphere, yields

$$\begin{aligned}\int_\Omega |\partial_i f(\mathbf{x})|^p \, d\mathbf{x} &\leq \int_\Omega \frac{d\mathbf{x}}{\|\mathbf{x}\|_2^d\,|\ln\|\mathbf{x}\|_2|^d} = \int_{S^{d-1}}\int_0^{1/2}\frac{\rho^{d-1}}{\rho^d\,|\ln\rho|^d}\,d\rho d\omega \\ &= \operatorname{meas}\left(S^{d-1}\right)\int_0^{1/2}\frac{d\rho}{\rho\,|\ln\rho|^d} = -\operatorname{meas}\left(S^{d-1}\right)\int_\infty^{\ln 2}\frac{dt}{t^d} < \infty,\end{aligned}$$

because of $d \geq 2$.

It follows that $\partial_i f \in L^p(\Omega)$ with $p \le d$. Analogously, one proves that $f \in L^p(\Omega)$ with $p \le d$. Altogether, one has $f \in W^{1,p}(\Omega)$ with $p \le d$. However, it is $f \notin L^\infty(\Omega)$. This example shows that the condition $k > d/p$ for $p > 1$ is sharp.

In particular, it was proved for $p = 2$ that from $f \in H^1(\Omega)$ in general it does not follow that $f \in C(\Omega)$. □

Example 3.57 *The assumption of a Lipschitz boundary.* Also the assumption that Ω is a Lipschitz domain is of importance.

Consider $\Omega = \{(x,y) \in \mathbb{R}^2 \ : \ 0 < x < 1, \ |y| < x^r, r > 1\}$, see Figure 3.5 for $r = 2$.

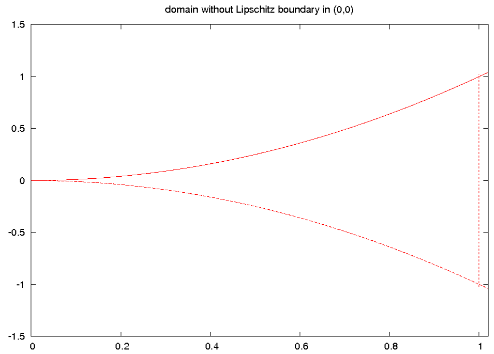

Figure 3.5: Domain of Example 3.57.

For $u(x,y) = x^{-\varepsilon/p}$ with $0 < \varepsilon < r$ it is

$$\partial_x u = x^{-\varepsilon/p-1}\left(-\frac{\varepsilon}{p}\right) = C(\varepsilon,p)x^{-\varepsilon/p-1}, \ \partial_y u = 0.$$

It follows that

$$\begin{aligned}\sum_{|\boldsymbol{\alpha}|=1}\int_\Omega |D^{\boldsymbol{\alpha}} u(x,y)|^p \ dxdy &= C(\varepsilon,p)\int_\Omega x^{-\varepsilon-p}\ dxdy \\ &= C(\varepsilon,p)\int_0^1 x^{-\varepsilon-p}\left(\int_{-x^r}^{x^r} dy\right) dx \\ &= \tilde{C}(\varepsilon,p)\int_0^1 x^{-\varepsilon-p+r}\ dx.\end{aligned}$$

This value is finite for $-\varepsilon - p + r > -1$ or for $p < 1 + r - \varepsilon$, respectively. If one chooses $r \ge \varepsilon > 0$, then it is $u \in W^{1,p}(\Omega)$. But for $\varepsilon > 0$ the function $u(\mathbf{x})$ is not bounded in Ω, i.e., $u \notin L^\infty(\Omega)$.

The unbounded values of the function are compensated in the integration by the fact that the neighborhood of the singular point $(0,0)$ possesses a small measure. □

Chapter 4

The Ritz Method and the Galerkin Method

Remark 4.1 *Contents.* This chapter studies variational or weak formulations of boundary value problems of partial differential equations in Hilbert spaces. The existence and uniqueness of an appropriately defined weak solution will be discussed. The approximation of this solution with the help of finite-dimensional spaces is called Ritz method or Galerkin method. Some basic properties of this method will be proved.

In this chapter, a Hilbert space V will be considered with inner product $a(\cdot,\cdot)$: $V \times V \to \mathbb{R}$ and norm $\|v\|_V = a(v,v)^{1/2}$. □

4.1 The Theorems of Riesz and Lax–Milgram

Theorem 4.2 Representation theorem of Riesz. *Let $f \in V'$ be a continuous and linear functional, then there is a uniquely determined $u \in V$ with*

$$a(u,v) = f(v) \quad \forall\, v \in V. \tag{4.1}$$

In addition, u is the unique solution of the variational problem

$$F(v) = \frac{1}{2}a(v,v) - f(v) \to \min \ \forall\, v \in V. \tag{4.2}$$

Proof: First, the existence of a solution u of the variational problem will be proved. Since f is continuous, it holds

$$|f(v)| \le c\,\|v\|_V \quad \forall\, v \in V,$$

from what follows that

$$F(v) \ge \frac{1}{2}\,\|v\|_V^2 - c\,\|v\|_V \ge -\frac{1}{2}c^2,$$

where in the last estimate the necessary criterion for a local minimum of the expression of the first estimate is used. Hence, the function $F(\cdot)$ is bounded from below and

$$d = \inf_{v \in V} F(v)$$

exists.

Let $\{v_k\}_{k\in\mathbb{N}}$ be a sequence with $F(v_k) \to d$ for $k \to \infty$. A straightforward calculation (parallelogram identity in Hilbert spaces) gives

$$\|v_k - v_l\|_V^2 + \|v_k + v_l\|_V^2 = 2\,\|v_k\|_V^2 + 2\,\|v_l\|_V^2\,.$$

Using the linearity of $f(\cdot)$ and $d \le F(v)$ for all $v \in V$, one obtains

$$\begin{aligned}
&\|v_k - v_l\|_V^2 \\
&= 2\|v_k\|_V^2 + 2\|v_l\|_V^2 - 4\left\|\frac{v_k+v_l}{2}\right\|_V^2 - 4f(v_k) - 4f(v_l) + 8f\left(\frac{v_k+v_l}{2}\right) \\
&= 4F(v_k) + 4F(v_l) - 8F\left(\frac{v_k+v_l}{2}\right) \\
&\le 4F(v_k) + 4F(v_l) - 8d \to 0
\end{aligned}$$

for $k, l \to \infty$. Hence $\{v_k\}_{k\in\mathbb{N}}$ is a Cauchy sequence. Because V is a complete space, there exists a limit u of this sequence with $u \in V$. Because $F(\cdot)$ is continuous, it is $F(u) = d$ and u is a solution of the variational problem.

In the next step, it will be shown that each solution of the variational problem (4.2) is also a solution of (4.1). It is

$$\begin{aligned}
\Phi(\varepsilon) &= F(u+\varepsilon v) = \frac{1}{2}a(u+\varepsilon v, u+\varepsilon v) - f(u+\varepsilon v) \\
&= \frac{1}{2}a(u,u) + \varepsilon a(u,v) + \frac{\varepsilon^2}{2}a(v,v) - f(u) - \varepsilon f(v).
\end{aligned}$$

If u is a minimum of the variational problem, then the function $\Phi(\varepsilon)$ has a local minimum at $\varepsilon = 0$. The necessary condition for a local minimum leads to

$$0 = \Phi'(0) = a(u,v) - f(v) \quad \text{for all } v \in V.$$

Finally, the uniqueness of the solution will be proved. It is sufficient to prove the uniqueness of the solution of the equation (4.1). If the solution of (4.1) is unique, then the existence of two solutions of the variational problem (4.2) would be a contradiction to the fact proved in the previous step. Let u_1 and u_2 be two solutions of the equation (4.1). Computing the difference of both equations gives

$$a(u_1 - u_2, v) = 0 \quad \text{for all } v \in V.$$

This equation holds, in particular, for $v = u_1 - u_2$. Hence, $\|u_1 - u_2\|_V = 0$, such that $u_1 = u_2$. ■

Definition 4.3 Bounded bilinear form, coercive bilinear form, V-elliptic bilinear form. Let $b(\cdot,\cdot) : V \times V \to \mathbb{R}$ be a bilinear form on the Banach space V. Then it is bounded if

$$|b(u,v)| \le M \|u\|_V \|v\|_V \quad \forall\, u, v \in V, M > 0, \tag{4.3}$$

where the constant M is independent of u and v. The bilinear form is coercive or V-elliptic if

$$b(u,u) \ge m \|u\|_V^2 \quad \forall\, u \in V, m > 0, \tag{4.4}$$

where the constant m is independent of u. □

Remark 4.4 *Application to an inner product.* Let V be a Hilbert space. Then the inner product $a(\cdot,\cdot)$ is a bounded and coercive bilinear form, since by the Cauchy–Schwarz inequality

$$|a(u,v)| \le \|u\|_V \|v\|_V \quad \forall\, u, v \in V,$$

and obviously $a(u,u) = \|u\|_V^2$. Hence, the constants can be chosen to be $M = 1$ and $m = 1$.

Next, the representation theorem of Riesz will be generalized to the case of coercive and bounded bilinear forms. □

Theorem 4.5 Theorem of Lax–Milgram. *Let $b(\cdot,\cdot) : V \times V \to \mathbb{R}$ be a bounded and coercive bilinear form on the Hilbert space V. Then, for each bounded linear functional $f \in V'$ there is exactly one $u \in V$ with*

$$b(u,v) = f(v) \quad \forall\, v \in V. \tag{4.5}$$

Proof: One defines linear operators $T, T' \,:\, V \to V$ by

$$a(Tu,v) = b(u,v)\ \forall\, v \in V, \quad a(T'u,v) = b(v,u)\ \forall\, v \in V. \tag{4.6}$$

Since $b(u,\cdot)$ and $b(\cdot,u)$ are continuous linear functionals on V, it follows from Theorem 4.2 that the elements Tu and $T'u$ exist and they are defined uniquely. Because the operators satisfy the relation

$$a(Tu,v) = b(u,v) = a(T'v,u) = a(u,T'v), \tag{4.7}$$

T' is called adjoint operator of T. Setting $v = Tu$ in (4.6) and using the boundedness of $b(\cdot,\cdot)$ yields

$$\|Tu\|_V^2 = a(Tu,Tu) = b(u,Tu) \le M\,\|u\|_V\,\|Tu\|_V \implies \|Tu\|_V \le M\,\|u\|_V$$

for all $u \in V$. Hence, T is bounded. Since T is linear, it follows that T is continuous. Using the same argument, one shows that T' is also bounded and continuous.

Define the bilinear form

$$d(u,v) := a(TT'u,v) = a(T'u,T'v) \quad \forall\, u,v \in V,$$

where (4.7) was used. Hence, this bilinear form is symmetric. Using the coercivity of $b(\cdot,\cdot)$ and the Cauchy–Schwarz inequality gives

$$m^2\,\|v\|_V^4 \le b(v,v)^2 = a(T'v,v)^2 \le \|v\|_V^2\,\|T'v\|_V^2 = \|v\|_V^2\,a(T'v,T'v) = \|v\|_V^2\,d(v,v).$$

Applying now the boundedness of $a(\cdot,\cdot)$ and of T' yields

$$m^2\,\|v\|_V^2 \le d(v,v) = a(T'v,T'v) = \|T'v\|_V^2 \le M\,\|v\|_V^2\,. \tag{4.8}$$

Hence, $d(\cdot,\cdot)$ is also coercive and, since it is symmetric, it defines an inner product on V. From (4.8) one has that the norm induced by $d(v,v)^{1/2}$ is equivalent to the norm $\|v\|_V$. From Theorem 4.2 it follows that there is a exactly one $w \in V$ with

$$d(w,v) = f(v) \quad \forall\, v \in V.$$

Inserting $u = T'w$ into (4.5) gives with (4.6)

$$b(T'w,v) = a(TT'w,v) = d(w,v) = f(v) \quad \forall\, v \in V,$$

hence $u = T'w$ is a solution of (4.5).

The uniqueness of the solution is proved analogously as in the symmetric case. ■

4.2 Weak Formulation of Boundary Value Problems

Remark 4.6 *Model problem.* Consider the Poisson equation with homogeneous Dirichlet boundary conditions

$$\begin{array}{rcll} -\Delta u & = & f & \text{in } \Omega \subset \mathbb{R}^d, \\ u & = & 0 & \text{on } \partial\Omega. \end{array} \tag{4.9}$$

□

Definition 4.7 Weak formulation of (4.9)**.** Let $f \in L^2(\Omega)$. A weak formulation of (4.9) consists in finding $u \in V = H_0^1(\Omega)$ such that

$$a(u,v) = (f,v) \quad \forall\, v \in V \tag{4.10}$$

with

$$a(u,v) = (\nabla u, \nabla v) = \int_\Omega \nabla u(\mathbf{x}) \cdot \nabla v(\mathbf{x})\ d\mathbf{x}$$

and $(\cdot,\cdot)$ is the inner product in $L^2(\Omega)$. □

Remark 4.8 *On the weak formulation.*

- The weak formulation is also called variational formulation.
- As usual in mathematics, 'weak' means that something holds for all appropriately chosen test functions.
- Formally, one obtains the weak formulation by multiplying the strong form of the equation (4.9) with the test function, by integrating the equation on Ω, and applying integration by parts. Because of the Dirichlet boundary condition, on can use as test space $H_0^1(\Omega)$ and therefore the integral on the boundary vanishes.
- The ansatz space for the solution and the test space are defined such that the arising integrals are well defined.
- The weak formulation reduces the necessary regularity assumptions for the solution by the integration and the transfer of derivatives to the test function. Whereas the solution of (4.9) has to be in $C^2(\overline{\Omega})$, the solution of (4.10) has to be only in $H_0^1(\Omega)$. The latter assumption is much more realistic for problems coming from applications.
- The regularity assumption on the right hand side can be relaxed to $f \in H^{-1}(\Omega)$.

□

Theorem 4.9 Existence and uniqueness of the weak solution. *Let $f \in L^2(\Omega)$. There is exactly one solution of* (4.10)*.*

Proof: Because of the Poincaré inequality (3.9), there is a constant c with

$$\|v\|_{L^2(\Omega)} \le c \|\nabla v\|_{L^2(\Omega)} \quad \forall\, v \in H_0^1(\Omega).$$

It follows for $v \in H_0^1(\Omega) \subset H^1(\Omega)$ that

$$\begin{aligned} \|v\|_{H^1(\Omega)} &= \left(\|v\|_{L^2(\Omega)}^2 + \|\nabla v\|_{L^2(\Omega)}^2\right)^{1/2} \le \left(c\|\nabla v\|_{L^2(\Omega)}^2 + \|\nabla v\|_{L^2(\Omega)}^2\right)^{1/2} \\ &\le C \|\nabla v\|_{L^2(\Omega)} \le C \|v\|_{H^1(\Omega)}. \end{aligned}$$

Hence, $a(\cdot,\cdot)$ is an inner product on $H_0^1(\Omega)$ with the induced norm

$$\|v\|_{H_0^1(\Omega)} = a(v,v)^{1/2},$$

which is equivalent to the norm $\|\cdot\|_{H^1(\Omega)}$.

Define for $f \in L^2(\Omega)$ the linear functional

$$\tilde{f}(v) := \int_\Omega f(\mathbf{x})v(\mathbf{x})\ d\mathbf{x} \quad \forall\, v \in H_0^1(\Omega).$$

Applying the Cauchy–Schwarz inequality (3.5) and the Poincaré inequality (3.9)

$$\left|\tilde{f}(v)\right| = |(f,v)| \le \|f\|_{L^2(\Omega)} \|v\|_{L^2(\Omega)} \le c\|f\|_{L^2(\Omega)} \|\nabla v\|_{L^2(\Omega)} = c\|f\|_{L^2(\Omega)} \|v\|_{H_0^1(\Omega)}$$

shows that this functional is continuous on $H_0^1(\Omega)$. Applying the representation theorem of Riesz, Theorem 4.2, gives the existence and uniqueness of the weak solution of (4.10). In addition, $u(\mathbf{x})$ solves the variational problem

$$F(v) = \frac{1}{2}\|\nabla v\|_2^2 - \int_\Omega f(\mathbf{x})v(\mathbf{x})\ d\mathbf{x} \to \min \quad \text{for all } v \in H_0^1(\Omega).$$

■

Example 4.10 *A more general elliptic problem.* Consider the problem

$$\begin{aligned} -\nabla \cdot (A(\mathbf{x})\nabla u) + c(\mathbf{x})u &= f \quad \text{in } \Omega \subset \mathbb{R}^d, \\ u &= 0 \quad \text{on } \partial\Omega, \end{aligned} \tag{4.11}$$

with $A(\mathbf{x}) \in \mathbb{R}^{d\times d}$ for each point $\mathbf{x} \in \Omega$. It will be assumed that the coefficients $a_{i,j}(\mathbf{x})$ and $c(\mathbf{x}) \geq 0$ are bounded, $f \in L^2(\Omega)$, and that the matrix (tensor) $A(\mathbf{x})$ is for all $\mathbf{x} \in \Omega$ uniformly elliptic, i.e., there are positive constants m and M such that

$$m \left\|\mathbf{y}\right\|_2^2 \leq \mathbf{y}^T A(\mathbf{x})\mathbf{y} \leq M \left\|\mathbf{y}\right\|_2^2 \quad \forall\, \mathbf{y} \in \mathbb{R}^d,\ \forall\, \mathbf{x} \in \Omega.$$

The weak form of (4.11) is obtained in the usual way by multiplying (4.11) with test functions $v \in H_0^1(\Omega)$, integrating on Ω, and applying integration by parts: Find $u \in H_0^1(\Omega)$, such that

$$a(u,v) = f(v) \quad \forall\, v \in H_0^1(\Omega)$$

with

$$a(u,v) = \int_\Omega \left(\nabla u(\mathbf{x})^T A(\mathbf{x})\nabla v(\mathbf{x}) + c(\mathbf{x})u(\mathbf{x})v(\mathbf{x})\right)\, d\mathbf{x}.$$

This bilinear form is bounded (*exercise*). The coercivity of the bilinear form is proved by using the uniform ellipticity of $A(\mathbf{x})$ and the non-negativity of $c(\mathbf{x})$:

$$\begin{aligned} a(u,u) &= \int_\Omega \nabla u(\mathbf{x})^T A(\mathbf{x})\nabla u(\mathbf{x}) + c(\mathbf{x})u(\mathbf{x})u(\mathbf{x})\, d\mathbf{x} \\ &\geq \int_\Omega m\nabla u(\mathbf{x})^T\nabla u(\mathbf{x})\, d\mathbf{x} = m\left\|u\right\|_{H_0^1(\Omega)}^2. \end{aligned}$$

Applying the Theorem of Lax–Milgram, Theorem 4.5, gives the existence and uniqueness of a weak solution of (4.11).

If the tensor is not symmetric, $a_{ij}(\mathbf{x}) \neq a_{ji}(\mathbf{x})$ for one pair i, j, then the solution cannot be characterized as the solution of a variational problem. □

4.3 The Ritz Method and the Galerkin Method

Remark 4.11 *Idea of the Ritz method.* Let V be a Hilbert space with the inner product $a(\cdot,\cdot)$. Consider the problem

$$F(v) = \frac{1}{2}a(v,v) - f(v) \to \min, \tag{4.12}$$

where $f\ :\ V \to \mathbb{R}$ is a bounded linear functional. As already proved in Theorem 4.2, there is a unique solution $u \in V$ of this variational problem which is also the unique solution of the equation

$$a(u,v) = f(v) \quad \forall\, v \in V. \tag{4.13}$$

For approximating the solution of (4.12) or (4.13) with a numerical method, it will be assumed that V has a countable orthonormal basis (Schauder basis). Then, there are finite-dimensional subspaces $V_1, V_2, \ldots \subset V$ with $\mathrm{dim}V_k = k$, which has the following property: for each $u \in V$ and each $\varepsilon > 0$ there is a $K \in \mathbb{N}$ and a $u_k \in V_k$ with

$$\left\|u - u_k\right\|_V \leq \varepsilon \quad \forall\, k \geq K. \tag{4.14}$$

Note that it is not required that there holds an inclusion of the form $V_k \subset V_{k+1}$.

The Ritz approximation of (4.12) and (4.13) is defined by: Find $u_k \in V_k$ with

$$a(u_k, v_k) = f(v_k) \quad \forall\, v_k \in V_k. \tag{4.15}$$

□

Lemma 4.12 Existence and uniqueness of a solution of (4.15)**.** *There exists exactly one solution of* (4.15)*.*

Proof: Finite-dimensional subspaces of Hilbert spaces are Hilbert spaces as well. For this reason, one can apply the representation theorem of Riesz, Theorem 4.2, to (4.15) which gives the statement of the lemma. In addition, the solution of (4.15) solves a minimization problem on V_k. ■

Lemma 4.13 Best approximation property. *The solution of* (4.15) *is the best approximation of* u *in* V_k*, i.e., it is*

$$\|u - u_k\|_V = \inf_{v_k \in V_k} \|u - v_k\|_V . \tag{4.16}$$

Proof: Since $V_k \subset V$, one can use the test functions from V_k in the weak equation (4.13). Then, the difference of (4.13) and (4.15) gives the orthogonality, the so-called Galerkin orthogonality,

$$a(u - u_k, v_k) = 0 \quad \forall\, v_k \in V_k. \tag{4.17}$$

Hence, the error $u - u_k$ is orthogonal to the space V_k: $u - u_k \perp V_k$. That means, u_k is the orthogonal projection of u onto V_k with respect of the inner product of V.

Let now $w_k \in V_k$ be an arbitrary element, then it follows with the Galerkin orthogonality (4.17) and the Cauchy–Schwarz inequality that

$$\begin{aligned} \|u - u_k\|_V^2 &= a(u - u_k, u - u_k) = a(u - u_k, u - \underbrace{(u_k - w_k)}_{v_k}) = a(u - u_k, u - v_k) \\ &\le \|u - u_k\|_V \|u - v_k\|_V . \end{aligned}$$

Since $w_k \in V_k$ was arbitrary, also $v_k \in V_k$ is arbitrary. If $\|u - u_k\|_V > 0$, division by $\|u - u_k\|_V$ gives the statement of the lemma. If $\|u - u_k\|_V = 0$, the statement of the lemma is trivially true. ■

Theorem 4.14 Convergence of the Ritz approximation. *The Ritz approximation converges*

$$\lim_{k\to\infty} \|u - u_k\|_V = 0.$$

Proof: The best approximation property (4.16) and property (4.14) give

$$\|u - u_k\|_V = \inf_{v_k \in V_k} \|u - v_k\|_V \le \varepsilon$$

for each $\varepsilon > 0$ and $k \ge K(\varepsilon)$. Hence, the convergence is proved. ■

Remark 4.15 *Formulation of the Ritz method as linear system of equations.* One can use an arbitrary basis $\{\phi_i\}_{i=1}^k$ of V_k for the computation of u_k. First of all, the equation for the Ritz approximation (4.15) is satisfied for all $v_k \in V_k$ if and only if it is satisfied for each basis function ϕ_i. This statement follows from the linearity of both sides of the equation with respect to the test function and from the fact that each function $v_k \in V_k$ can be represented as linear combination of the basis functions. Let $v_k = \sum_{i=i}^k \alpha_i \phi_i$, then from (4.15) it follows that

$$a(u_k, v_k) = \sum_{k=1}^{k} \alpha_i a(u_k, \phi_i) = \sum_{k=1}^{k} \alpha_i f(\phi_i) = f(v_k).$$

This equation is satisfied if $a(u_k, \phi_i) = f(\phi_i)$, $i = 1, \ldots, k$. On the other hand, if (4.15) holds then it holds in particular for each basis function ϕ_i.

Then, one uses as ansatz for the solution also a linear combination of the basis functions

$$u_k = \sum_{j=1}^{k} u^j \phi_j$$

with unknown coefficients $u^j \in \mathbb{R}$. Using as test functions now the basis functions yields

$$\sum_{j=1}^{k} a(u^j\phi_j, \phi_i) = \sum_{j=1}^{k} a(\phi_j, \phi_i)u^j = f(\phi_i), \quad i = 1, \ldots, k.$$

This equation is equivalent to the linear system of equations $A\mathbf{u} = \mathbf{f}$, where

$$A = (a_{ij})_{i,j=1}^{k} = a(\phi_j, \phi_i)_{i,j=1}^{k}$$

is called stiffness matrix. Note that the order of the indices is different for the entries of the matrix and the arguments of the inner product. The right hand side is a vector of length k with the entries $f_i = f(\phi_i)$, $i = 1, \ldots, k$.

Using the one-to-one mapping between the coefficient vector $(v^1, \ldots, v^k)^T$ and the element $v_k = \sum_{i=1}^{k} v^i\phi_i$, one can show that the matrix A is symmetric and positive definite (*exercise*)

$$\begin{aligned} A = A^T \quad &\Longleftrightarrow \quad a(v, w) = a(w, v) \quad \forall\, v, w \in V_k, \\ \mathbf{x}^T A\mathbf{x} > 0 \text{ for } \mathbf{x} \neq \mathbf{0} \quad &\Longleftrightarrow \quad a(v, v) > 0 \quad \forall\, v \in V_k, v \neq 0. \end{aligned}$$

□

Remark 4.16 *The case of a bounded and coercive bilinear form.* If $b(\cdot, \cdot)$ is bounded and coercive, but not symmetric, it is possible to approximate the solution of (4.5) with the same idea as for the Ritz method. In this case, it is called Galerkin method. The discrete problem consists in finding $u_k \in V_k$ such that

$$b(u_k, v_k) = f(v_k) \quad \forall\, v_k \in V_k. \tag{4.18}$$

□

Lemma 4.17 Existence and uniqueness of a solution of (4.18)**.** *There is exactly one solution of* (4.18)*.*

Proof: The statement of the lemma follows directly from the Theorem of Lax–Milgram, Theorem 4.5. ■

Remark 4.18 *On the discrete solution.* The discrete solution is not the orthogonal projection into V_k in the case of a bounded and coercive bilinear form, which is not the inner product of V. □

Lemma 4.19 Lemma of Cea, error estimate. *Let $b :\ V \times V \to \mathbb{R}$ be a bounded and coercive bilinear form on the Hilbert space V and let $f \in V'$ be a bounded linear functional. Let u be the solution of* (4.5) *and u_k be the solution of* (4.18)*, then the following error estimate holds*

$$\|u - u_k\|_V \leq \frac{M}{m} \inf_{v_k \in V_k} \|u - v_k\|_V\,, \tag{4.19}$$

where the constants M and m are given in (4.3) *and* (4.4)*.*

Proof: Considering the difference of the continuous equation (4.5) and the discrete equation (4.18), one obtains the error equation

$$b(u - u_k, v_k) = 0 \quad \forall\, v_k \in V_k,$$

which is also called Galerkin orthogonality. With (4.4), the Galerkin orthogonality, and (4.3) it follows that

$$\begin{aligned} \|u - u_k\|_V^2 \quad &\leq \quad \frac{1}{m} b(u - u_k, u - u_k) = \frac{1}{m} b(u - u_k, u - v_k) \\ &\leq \quad \frac{M}{m} \|u - u_k\|_V \|u - v_k\|_V\,, \quad \forall\, v_k \in V_k, \end{aligned}$$

from what the statement of the lemma follows immediately. ■

Remark 4.20 *On the best approximation error.* It follows from estimate (4.19) that the error is bounded by a multiple of the best approximation error, where the factor depends on properties of the bilinear form $b(\cdot,\cdot)$. Thus, concerning error estimates for concrete finite-dimensional spaces, the study of the best approximation error will be of importance. □

Remark 4.21 *The corresponding linear system of equations.* The corresponding linear system of equations is derived analogously to the symmetric case. The system matrix is still positive definite but not symmetric. □

Remark 4.22 *Choice of the basis.* The most important issue of the Ritz and Galerkin method is the choice of the spaces V_k, or more concretely, the choice of an appropriate basis $\{\phi_i\}_{i=1}^k$ that spans the space V_k. From the point of view of numerics, there are the requirements that it should be possible to compute the entries a_{ij} of the stiffness matrix efficiently and that the matrix A should be sparse. □

Chapter 5

Finite Element Methods

5.1 Finite Element Spaces

Remark 5.1 *Mesh cells, faces, edges, vertices.* A mesh cell K is a compact polyhedron in $\mathbb{R}^d$, $d \in \{2,3\}$, whose interior is not empty. The boundary ∂K of K consists of m-dimensional linear manifolds (points, pieces of straight lines, pieces of planes), $0 \le m \le d-1$, which are called m-faces. The 0-faces are the vertices of the mesh cell, the 1-faces are the edges, and the $(d-1)$-faces are just called faces. □

Remark 5.2 *Finite dimensional spaces defined on K.* Let $s \in \mathbb{N}$. Finite element methods use finite dimensional spaces $P(K) \subset C^s(K)$ which are defined on K. In general, $P(K)$ consists of polynomials. The dimension of $P(K)$ will be denoted by $\dim P(K) = N_K$. □

Example 5.3 *The space $P(K) = P_1(K)$.*The space consisting of linear polynomials on a mesh cell K is denoted by $P_1(K)$:

$$P_1(K) = \left\{ a_0 + \sum_{i=1}^d a_i x_i \ : \ \mathbf{x} = (x_1, \ldots, x_d)^T \in K \right\}.$$

There are $d+1$ unknown coefficients a_i, $i = 0, \ldots, d$, such that $\dim P_1(K) = N_K = d+1$. □

Remark 5.4 *Linear functionals defined on $P(K)$.* For the definition of finite elements, linear functional which are defined on $P(K)$ are of importance.

Consider linear and continuous functionals $\Phi_{K,1}, \ldots, \Phi_{K,N_K} \ : C^s(K) \to \mathbb{R}$ which are linearly independent. There are different types of functionals which can be utilized in finite element methods:

- point values: $\Phi(v) = v(\mathbf{x})$, $\mathbf{x} \in K$,
- point values of a first partial derivative: $\Phi(v) = \partial_i v(\mathbf{x})$, $\mathbf{x} \in K$,
- point values of the normal derivative on a face E of K: $\Phi(v) = \nabla v(\mathbf{x}) \cdot \mathbf{n}_E$, $\mathbf{n}_E$ is the outward pointing unit normal vector on E,
- integral mean values on K: $\Phi(v) = \frac{1}{|K|} \int_K v(\mathbf{x}) \, d\mathbf{x}$,
- integral mean values on faces E: $\Phi(v) = \frac{1}{|E|} \int_E v(\mathbf{s}) \, d\mathbf{s}$.

The smoothness parameter s has to be chosen in such a way that the functionals $\Phi_{K,1}, \ldots, \Phi_{K,N_K}$ are continuous. If, e.g., a functional requires the evaluation of a partial derivative or a normal derivative, then one has to choose at least $s = 1$. For the other functionals given above, $s = 0$ is sufficient. □

Definition 5.5 Unisolvence of $P(K)$ with respect to the functionals $\Phi_{K,1}, \ldots, \Phi_{K,N_K}$. The space $P(K)$ is called unisolvent with respect to the functionals $\Phi_{K,1}, \ldots, \Phi_{K,N_K}$ if there is for each $\mathbf{a} \in \mathbb{R}^{N_K}$, $\mathbf{a} = (a_1, \ldots, a_{N_K})^T$, exactly one $p \in P(K)$ with

$$\Phi_{K,i}(p) = a_i, \quad 1 \le i \le N_K.$$

□

Remark 5.6 *Local basis.* Unisolvence means that for each vector $\mathbf{a} \in \mathbb{R}^{N_K}$, $\mathbf{a} = (a_1, \ldots, a_{N_K})^T$, there is exactly one element in $P(K)$ such that a_i is the image of the i-th functional, $i = 1, \ldots, N_K$.

Choosing in particular the Cartesian unit vectors for $\mathbf{a}$, then it follows from the unisolvence that a set $\{\phi_{K,i}\}_{i=1}^{N_K}$ exists with $\phi_{K,i} \in P(K)$ and

$$\Phi_{K,i}(\phi_{K,j}) = \delta_{ij}, \quad i, j = 1, \ldots, N_K.$$

Consequently, the set $\{\phi_{K,i}\}_{i=1}^{N_K}$ forms a basis of $P(K)$. This basis is called local basis. □

Remark 5.7 *Transform of an arbitrary basis to the local basis.* If an arbitrary basis $\{p_i\}_{i=1}^{N_K}$ of $P(K)$ is known, then the local basis can be computed by solving a linear system of equations. To this end, represent the local basis in terms of the known basis

$$\phi_{K,j} = \sum_{k=1}^{N_K} c_{jk} p_k, \quad c_{jk} \in \mathbb{R},\ j = 1, \ldots, N_K,$$

with unknown coefficients c_{jk}. Applying the definition of the local basis leads to the linear system of equations

$$\Phi_{K,i}(\phi_{K,j}) = \sum_{k=1}^{N_K} c_{jk} a_{ik} = \delta_{ij}, \quad i, j = 1, \ldots, N_K, \quad a_{ik} = \Phi_{K,i}(p_k).$$

Because of the unisolvence, the matrix $A = (a_{ij})$ is non-singular and the coefficients c_{jk} are determined uniquely. □

Example 5.8 *Local basis for the space of linear functions on the reference triangle.* Consider the reference triangle $\hat{K}$ with the vertices $(0,0)$, $(1,0)$, and $(0,1)$. A linear space on $\hat{K}$ is spanned by the functions $1, \hat{x}, \hat{y}$. Let the functionals be defined by the values of the functions in the vertices of the reference triangle. Then, the given basis is not a local basis because the function 1 does not vanish at the vertices.

Consider first the vertex $(0,0)$. A linear basis function $a\hat{x} + b\hat{y} + c$ which has the value 1 in $(0,0)$ and which vanishes in the other vertices has to satisfy the following set of equations

$$\begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}.$$

The solution is $a = -1, b = -1, c = 1$. The two other basis functions of the local basis are $\hat{x}$ and $\hat{y}$, such that the local basis has the form $\{1 - \hat{x} - \hat{y}, \hat{x}, \hat{y}\}$. □

Remark 5.9 *Triangulation, grid, mesh, grid cell.* For the definition of global finite element spaces, a decomposition of the domain Ω into polyhedrons K is needed. This decomposition is called triangulation $\mathcal{T}^h$ and the polyhedrons K are called mesh cells. The union of the polyhedrons is called grid or mesh.

A triangulation is called regular, see the definition in Ciarlet Ciarlet (1978), if:

- It holds $\overline{\Omega} = \cup_{K \in \mathcal{T}^h} K$.

- Each mesh cell $K \in \mathcal{T}^h$ is closed and the interior $\overset{\circ}{K}$ is non-empty.
- For distinct mesh cells K_1 and K_2 there holds $\overset{\circ}{K}_1 \cap \overset{\circ}{K}_2 = \emptyset$.
- For each $K \in \mathcal{T}^h$, the boundary ∂K is Lipschitz-continuous.
- The intersection of two mesh cells is either empty or a common m-face, $m \in \{0, \ldots, d-1\}$.

□

Remark 5.10 *Global and local functionals.* Let $\Phi_1, \ldots, \Phi_N : C^s(\overline{\Omega}) \to \mathbb{R}$ continuous linear functionals of the same types as given in Remark 5.4. The restriction of the functionals to $C^s(K)$ defines local functionals $\Phi_{K,1}, \ldots, \Phi_{K,N_K}$, where it is assumed that the local functionals are unisolvent on $P(K)$. The union of all mesh cells K_j, for which there is a $p \in P(K_j)$ with $\Phi_i(p) \neq 0$, will be denoted by ω_i. □

Example 5.11 *On subdomains ω_i.* Consider the two-dimensional case and let Φ_i be defined as nodal value of a function in $\mathbf{x} \in K$. If $\mathbf{x} \in \overset{\circ}{K}$, then $\omega_i = K$. In the case that $\mathbf{x}$ is on a face of K but not in a vertex, then ω_i is the union of K and the other mesh cell whose boundary contains this face. Last, if $\mathbf{x}$ is a vertex of K, then ω_i is the union of all mesh cells which possess this vertex, see Figure 5.1. □

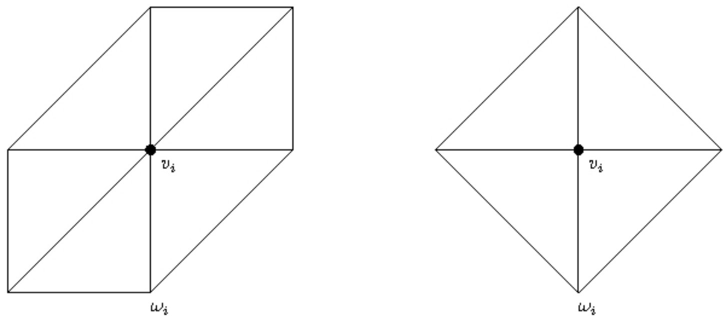

Figure 5.1: Subdomains ω_i.

Definition 5.12 Finite element space, global basis. A function $v(\mathbf{x})$ defined on Ω with $v|_K \in P(K)$ for all $K \in \mathcal{T}^h$ is called continuous with respect to the functional $\Phi_i : \Omega \to \mathbb{R}$ if

$$\Phi_i(v|_{K_1}) = \Phi_i(v|_{K_2}), \quad \forall\, K_1, K_2 \in \omega_i.$$

The space

$$\begin{aligned} S &= \Big\{ v \in L^\infty(\Omega) \ : \ v|_K \in P(K) \text{ and } v \text{ is continuous with respect to} \\ &\qquad \Phi_i, i = 1, \ldots, N \Big\} \end{aligned}$$

is called finite element space.

The global basis $\{\phi_j\}_{j=1}^N$ of S is defined by the condition

$$\phi_j \in S, \quad \Phi_i(\phi_j) = \delta_{ij}, \quad i, j = 1, \ldots, N.$$

□

Example 5.13 *Piecewise linear global basis function.* Figure 5.2 shows a piecewise linear global basis function in two dimensions. Because of its form, such a function is called hat function. □

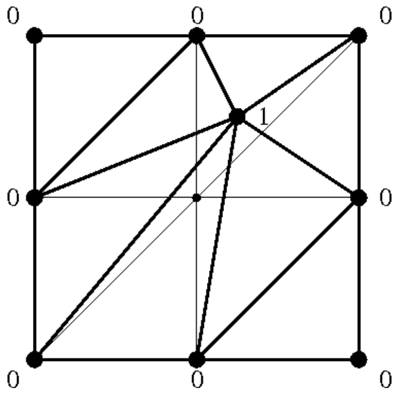

Figure 5.2: Piecewise linear global basis function (boldface lines), hat function.

Remark 5.14 *On global basis functions.* A global basis function coincides on each mesh cell with a local basis function. This property implies the uniqueness of the global basis functions.

For many finite element spaces it follows from the continuity with respect to $\{\Phi_i\}_{i=1}^N$, the continuity of the finite element functions themselves. Only in this case, one can speak of values of finite element functions on m-faces with $m < d$. □

Definition 5.15 Parametric finite elements. Let $\hat K$ be a reference mesh cell with the local space $P(\hat K)$, the local functionals $\hat\Phi_1, \ldots, \hat\Phi_{\hat N}$, and a class of bijective mappings $\{F_K \ : \hat K \to K\}$. A finite element space is called a parametric finite element space if:

- The images $\{K\}$ of $\{F_K\}$ form the set of mesh cells.
- The local spaces are given by

$$P(K) = \left\{p \ : \ p = \hat p \circ F_K^{-1}, \hat p \in \hat P(\hat K)\right\}. \tag{5.1}$$

- The local functionals are defined by

$$\Phi_{K,i}(v(\mathbf{x})) = \hat\Phi_i\left(v(F_K(\hat{\mathbf{x}}))\right), \tag{5.2}$$

  where $\hat{\mathbf{x}} = (\hat x_1, \ldots, \hat x_d)^T$ are the coordinates of the reference mesh cell and it holds $\mathbf{x} = F_K(\hat{\mathbf{x}})$.

□

Remark 5.16 *Motivations for using parametric finite elements.* Definition 5.12 of finite elements spaces is very general. For instance, different types of mesh cells are allowed. However, as well the finite element theory as the implementation of finite element methods become much simpler if only parametric finite elements are considered. □

5.2 Finite Elements on Simplices

Definition 5.17 *d-simplex.* **A d-simplex $K \subset \mathbb{R}^d$ is the convex hull of $(d+1)$ points $\mathbf{a}_1, \ldots, \mathbf{a}_{d+1} \in \mathbb{R}^d$ which form the vertices of K.** □

Remark 5.18 *On d-simplices.* It will be always assumed that the simplex is not degenerated, i.e., its d-dimensional measure is positive. This property is equivalent to the non-singularity of the matrix

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1,d+1} \\ a_{21} & a_{22} & \dots & a_{2,d+1} \\ \vdots & \vdots & \ddots & \vdots \\ a_{d1} & a_{d2} & \dots & a_{d,d+1} \\ 1 & 1 & \dots & 1 \end{pmatrix},$$

where $\mathbf{a}_i = (a_{1i}, a_{2i}, \dots, a_{di})^T$, $i = 1, \dots, d+1$.

For $d = 2$, the simplices are the triangles and for $d = 3$ they are the tetrahedrons. □

Definition 5.19 Barycentric coordinates. Since K is the convex hull of the points $\{\mathbf{a}_i\}_{i=1}^{d+1}$, the parametrization of K with a convex combination of the vertices reads as follows

$$K = \left\{ \mathbf{x} \in \mathbb{R}^d \ : \ \mathbf{x} = \sum_{i=1}^{d+1} \lambda_i \mathbf{a}_i, \ 0 \le \lambda_i \le 1, \ \sum_{i=1}^{d+1} \lambda_i = 1 \right\}.$$

The coefficients $\lambda_1, \dots, \lambda_{d+1}$ are called barycentric coordinates of $\mathbf{x} \in K$. □

Remark 5.20 *On barycentric coordinates.* From the definition it follows that the barycentric coordinates are the solution of the linear system of equations

$$\sum_{i=1}^{d+1} a_{ji} \lambda_i = x_j, \quad 1 \le j \le d, \quad \sum_{i=1}^{d+1} \lambda_i = 1.$$

Since the system matrix is non-singular, see Remark 5.18, the barycentric coordinates are determined uniquely.

The barycentric coordinates of the vertex $\mathbf{a}_i$, $i = 1, \dots, d+1$, of the simplex is $\lambda_i = 1$ and $\lambda_j = 0$ if $i \ne j$. Since $\lambda_i(\mathbf{a}_j) = \delta_{ij}$, the barycentric coordinate λ_i can be identified with the linear function which has the value 1 in the vertex $\mathbf{a}_i$ and which vanishes in all other vertices $\mathbf{a}_j$ with $j \ne i$.

The barycenter of the simplex is given by

$$S_K = \frac{1}{d+1} \sum_{i=1}^{d+1} \mathbf{a}_i = \sum_{i=1}^{d+1} \frac{1}{d+1} \mathbf{a}_i.$$

Hence, its barycentric coordinates are $\lambda_i = 1/(d+1)$, $i = 1, \dots, d+1$. □

Remark 5.21 *Simplicial reference mesh cells.* A commonly used reference mesh cell for triangles and tetrahedrons is the unit simplex

$$\hat{K} = \left\{ \hat{\mathbf{x}} \in \mathbb{R}^d \ : \ \sum_{i=1}^{d} \hat{x}_i \le 1, \ \hat{x}_i \ge 0, \ i = 1, \dots, d \right\},$$

see Figure 5.3. The class $\{F_K\}$ of admissible mappings are the bijective affine mappings

$$F_K \hat{\mathbf{x}} = B\hat{\mathbf{x}} + \mathbf{b}, \qquad B \in \mathbb{R}^{d \times d}, \ \det(B) \ne 0, \ \mathbf{b} \in \mathbb{R}^d.$$

The images of these mappings generate the set of the non-degenerated simplices $\{K\} \subset \mathbb{R}^d$. □

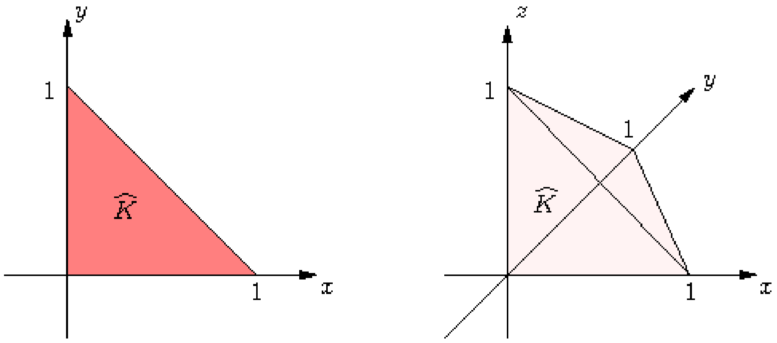

Figure 5.3: The unit simplices in two and three dimensions.

Definition 5.22 Affine family of simplicial finite elements. Given a simplicial reference mesh cell $\hat{K}$, affine mappings $\{F_K\}$, and an unisolvent set of functionals on $\hat{K}$. Using (5.1) and (5.2), one obtains a local finite element space on each non-degenerated simplex. The set of these local spaces is called affine family of simplicial finite elements. □

Definition 5.23 Polynomial space P_k. Let $\mathbf{x} = (x_1, \ldots, x_d)^T$, $k \in \mathbb{N} \cup \{0\}$, and $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_d)^T$. Then, the polynomial space P_k is given by

$$P_k = \text{span}\left\{ \prod_{i=1}^{d} x_i^{\alpha_i} = \mathbf{x}^{\boldsymbol{\alpha}} \ : \ \alpha_i \in \mathbb{N} \cup \{0\} \ \text{ for } \ i = 1, \ldots, d, \ \sum_{i=1}^{d} \alpha_i \le k \right\}.$$

□

Remark 5.24 *Lagrangian finite elements.* In all examples given below, the linear functionals on the reference mesh cell $\hat{K}$ are the values of the polynomials with the same barycentric coordinates as on the general mesh cell K. Finite elements whose linear functionals are values of the polynomials on certain points in K are called Lagrangian finite elements. □

Example 5.25 P_0 *: piecewise constant finite element.* The piecewise constant finite element space consists of discontinuous functions. The linear functional is the value of the polynomial in the barycenter of the mesh cell, see Figure 5.4. It is $\dim P_0(K) = 1$. □

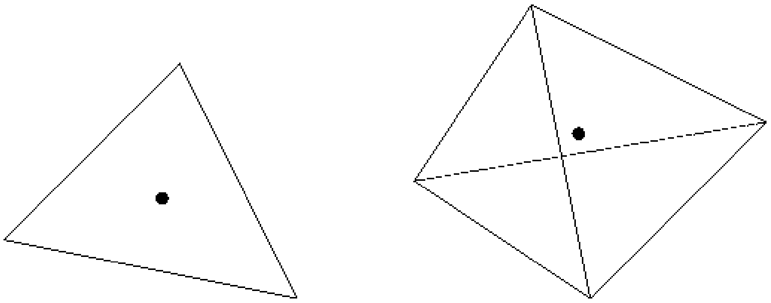

Figure 5.4: The finite element $P_0(K)$.

Example 5.26 P_1 : *conforming piecewise linear finite element.* This finite element space is a subspace of $C(\overline{\Omega})$. The linear functionals are the values of the function in the vertices of the mesh cells, see Figure 5.5. It follows that $\dim P_1(K) = d + 1$.

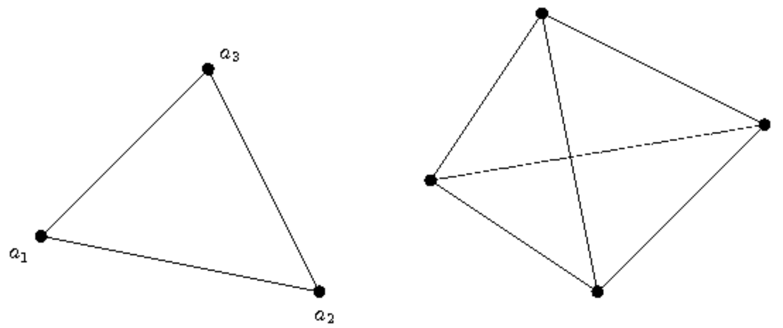

Figure 5.5: The finite element $P_1(K)$.

The local basis for the functionals $\{\Phi_i(v) = v(\mathbf{a}_i),\ i = 1, \ldots, d+1\}$, is $\{\lambda_i\}_{i=1}^{d+1}$ since $\Phi_i(\lambda_j) = \delta_{ij}$, see Remark 5.20. Since a local basis exists, the functionals are unisolvent with respect to the polynomial space $P_1(K)$.

Now, it will be shown that the corresponding finite element space consists of continuous functions. Let K_1, K_2 be two mesh cells with the common face E and let $v \in P_1(= S)$. The restriction of v_{K_1} on E is a linear function on E as well as the restriction of v_{K_2} on E. It has to be shown that both linear functions are identical. A linear function on the $(d-1)$-dimensional face E is uniquely determined with d linearly independent functionals which are defined on E. These functionals can be chosen to be the values of the function in the d vertices of E. The functionals in S are continuous, by the definition of S. Thus, it must hold that both restrictions on E have the same values in the vertices of E. Hence, it is $v_{K_1}|_E = v_{K_2}|_E$ and the functions from P_1 are continuous. □

Example 5.27 P_2 *: conforming piecewise quadratic finite element.* This finite element space is also a subspace of $C(\overline{\Omega})$. It consists of piecewise quadratic functions. The functionals are the values of the functions in the $d+1$ vertices of the mesh cell and the values of the functions in the centers of the edges, see Figure 5.6. Since each vertex is connected to each other vertex, there are $\sum_{i=1}^{d} i = d(d+1)/2$ edges. Hence, it follows that $\dim P_2(K) = (d+1)(d+2)/2$.

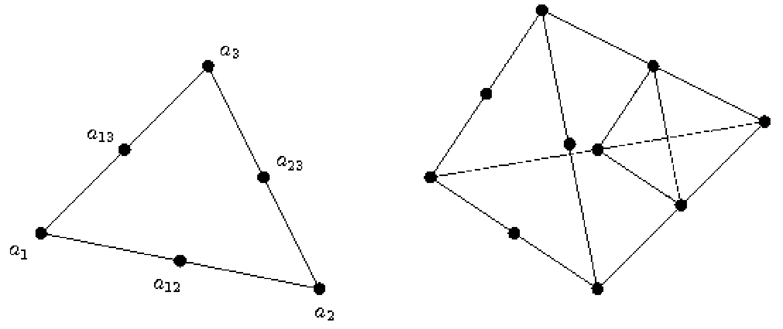

Figure 5.6: The finite element $P_2(K)$.

The part of the local basis which belongs to the functionals $\{\Phi_i(v) = v(\mathbf{a}_i),\ i = 1, \ldots, d+1\}$, is given by

$$\{\phi_i(\lambda) = \lambda_i(2\lambda_i - 1), \quad i = 1, \ldots, d+1\}.$$

Denote the center of the edges between the vertices $\mathbf{a}_i$ and $\mathbf{a}_j$ by $\mathbf{a}_{ij}$. The corresponding part of the local basis is given by

$$\{\phi_{ij} = 4\lambda_i\lambda_j, \quad i,j = 1,\ldots,d+1,\ i<j\}.$$

The unisolvence follows from the fact that there exists a local basis. The continuity of the corresponding finite element space is shown in the same way as for the P_1 finite element. The restriction of a quadratic function in a mesh cell to a face E is a quadratic function on that face. Hence, the function on E is determined uniquely with $d(d+1)/2$ linearly independent functionals on E.

The functions ϕ_{ij} are called in two dimensions edge bubble functions. □

Example 5.28 P_3 *: conforming piecewise cubic finite element.* This finite element space consists of continuous piecewise cubic functions. It is a subspace of $C(\overline{\Omega})$. The functionals in a mesh cell K are defined to be the values in the vertices ($(d+1)$ values), two values on each edge (dividing the edge in three parts of equal length) ($2\sum_{i=1}^{d} i = d(d+1)$ values), and the values in the barycenter of the 2-faces of K, see Figure 5.7. Each 2-face of K is defined by three vertices. If one considers for each vertex all possible pairs with other vertices, then each 2-face is counted three times. Hence, there are $(d+1)(d-1)d/6$ 2-faces. The dimension of $P_3(K)$ is given by

$$\dim P_3(K) = (d+1) + d(d+1) + \frac{(d-1)d(d+1)}{6} = \frac{(d+1)(d+2)(d+3)}{6}.$$

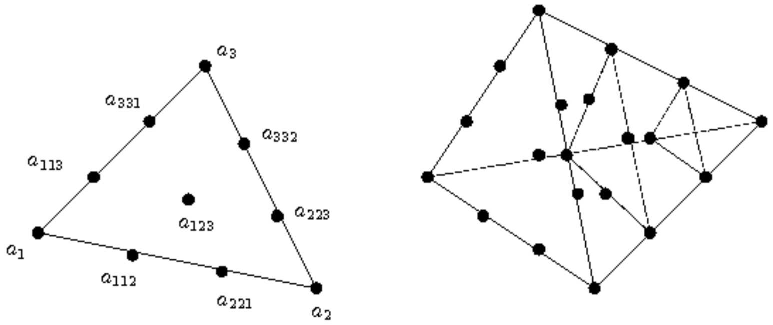

Figure 5.7: The finite element $P_3(K)$.

For the functionals

$$\left\{\begin{array}{rcll}\Phi_i(v) & = & v(\mathbf{a}_i),\ i=1,\ldots,d+1, & \text{(vertex)},\\ \Phi_{iij}(v) & = & v(\mathbf{a}_{iij}),\ i,j=1,\ldots,d+1, i\neq j, & \text{(point on edge)},\\ \Phi_{ijk}(v) & = & v(\mathbf{a}_{ijk}),\ i=1,\ldots,d+1, i<j<k & \text{(point on 2-face)}\end{array}\right\},$$

the local basis is given by

$$\left\{\begin{array}{rcl}\phi_i(\lambda) & = & \frac{1}{2}\lambda_i(3\lambda_i-1)(3\lambda_i-2),\\ \phi_{iij}(\lambda) & = & \frac{9}{2}\lambda_i\lambda_j(3\lambda_i-1),\\ \phi_{ijk}(\lambda) & = & 27\lambda_i\lambda_j\lambda_k\end{array}\right\}.$$

In two dimensions, the function $\phi_{ijk}(\lambda)$ is called cell bubble function. □

Example 5.29 *Cubic Hermite element.* The finite element space is a subspace of $C(\overline{\Omega})$, its dimension is $(d+1)(d+2)(d+3)/6$ and the functionals are the values of the function in the vertices of the mesh cell ($(d+1)$ values), the value of the barycenter at the 2-faces of K ($(d+1)(d-1)d/6$ values), and the partial derivatives at the vertices ($d(d+1)$ values), see Figure 5.8. The dimension is the same as for the P_3 element. Hence, the local polynomials can be defined to be cubic.

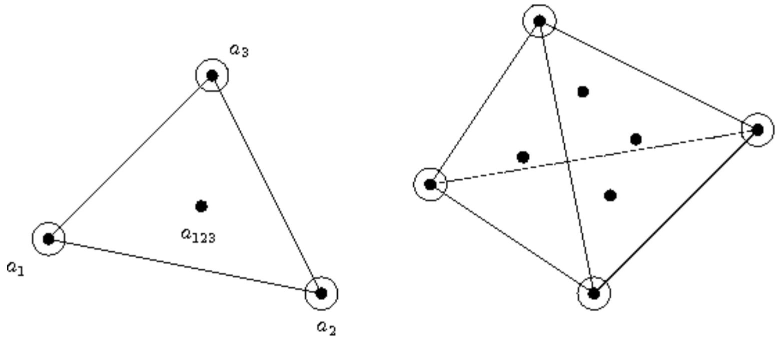

Figure 5.8: The cubic Hermite element.

This finite element does not define an affine family in the strict sense, because the functionals for the partial derivatives $\hat{\Phi}_i(\hat{v}) = \partial_i \hat{v}(\mathbf{0})$ on the reference cell are mapped to the functionals $\Phi_i(v) = \partial_{\mathbf{t}_i} v(\mathbf{a})$, where $\mathbf{a} = F_K(\mathbf{0})$ and $\mathbf{t}_i$ are the directions of edges which are adjacent to $\mathbf{a}$, i.e., $\mathbf{a}$ is an end point of this edge. This property suffices to control all first derivatives. On has to take care of this property in the implementation of this finite element.

Because of this property, one can use the derivatives in the direction of the edges as functionals

$$\begin{array}{rcll} \Phi_i(v) & = & v(\mathbf{a}_i), & \text{(vertices)} \\ \Phi_{ij}(v) & = & \nabla v(\mathbf{a}_i)\cdot(\mathbf{a}_j - \mathbf{a}_i),\ i,j = 1,\ldots,d-1, i \neq j, & \text{(directional derivative)} \\ \Phi_{ijk}(v) & = & v(\mathbf{a}_{ijk}),\ i<j<k, & \text{(2-faces)} \end{array}$$

with the corresponding local basis

$$\begin{array}{rcl} \phi_i(\lambda) & = & -2\lambda_i^3 + 3\lambda_i^2 - 7\lambda_i \sum_{j<k, j\neq i, k\neq i} \lambda_j \lambda_k, \\ \phi_{ij}(\lambda) & = & \lambda_i \lambda_j (2\lambda_i - \lambda_j - 1), \\ \phi_{ijk}(\lambda) & = & 27\lambda_i\lambda_j\lambda_k. \end{array}$$

The proof of the unisolvence can be found in the literature.

Here, the continuity of the functions will be shown only for $d = 2$. Let K_1, K_2 be two mesh cells with the common edge E and the unit tangential vector $\mathbf{t}$. Let V_1, V_2 be the end points of E. The restriction $v|_{K_1}, v|_{K_2}$ to E satisfy four conditions

$$v|_{K_1}(V_i) = v|_{K_2}(V_i), \quad \partial_{\mathbf{t}} v|_{K_1}(V_i) = \partial_{\mathbf{t}} v|_{K_2}(V_i),\ i = 1,2.$$

Since both restrictions are cubic polynomials and four conditions have to be satisfied, their values coincide on E.

The cubic Hermite finite element possesses an advantage in comparison with the P_3 finite element. For $d = 2$, it holds for a regular triangulation $\mathcal{T}^h$ that

$$\#(K) \approx 2\#(V), \quad \#(E) \approx 2\#(V),$$

where $\#(\cdot)$ denotes the number of triangles, nodes, and edges, respectively. Hence, the dimension of P_3 is approximately $7\#(V)$, whereas the dimension of the cubic

Hermite element is approximately $5\#(V)$. This difference comes from the fact that both spaces are different. The elements of both spaces are continuous functions, but for the functions of the cubic Hermite finite element, in addition, the first derivatives are continuous at the nodes. That means, these two spaces are different finite element spaces whose degree of the local polynomial space is the same (cubic). One can see at this example the importance of the functionals for the definition of the global finite element space. □

Example 5.30 *$P_1^{\rm nc}$: nonconforming linear finite element, Crouzeix–Raviart finite element Crouzeix and Raviart (1973).* This finite element consists of piecewise linear but discontinuous functions. The functionals are given by the values of the functions in the barycenters of the faces such that $\dim P_1^{\rm nc}(K) = (d+1)$. It follows from the definition of the finite element space, Definition 5.12, that the functions from $P_1^{\rm nc}$ are continuous in the barycenter of the faces

$$P_1^{\rm nc} = \big\{v \in L^2(\Omega) \ : \ v|_K \in P_1(K), \ v(\mathbf{x}) \text{ is continuous at the barycenter of all faces}\big\}. \tag{5.3}$$

Equivalently, the functionals can be defined to be the integral mean values on the faces and then the global space is defined to be

$$P_1^{\rm nc} = \left\{ v \in L^2(\Omega) \ : \ v|_K \in P_1(K), \int_E v|_K \ d\mathbf{s} = \int_E v|_{K'} \ d\mathbf{s} \ \forall \ E \in \mathcal{E}(K) \cap \mathcal{E}(K') \right\}, \tag{5.4}$$

where $\mathcal{E}(K)$ is the set of all $(d-1)$ dimensional faces of K.

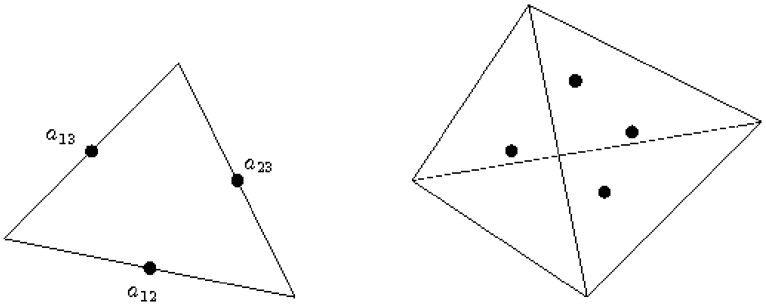

Figure 5.9: The finite element $P_1^{\rm nc}$.

For the description of this finite element, one defines the functionals by

$$\Phi_i(v) = v(\mathbf{a}_{i-1,i+1}) \text{ for } d = 2, \quad \Phi_i(v) = v(\mathbf{a}_{i-2,i-1,i+1}) \text{ for } d = 3,$$

where the points are the barycenters of the faces with the vertices that correspond to the indices. This system is unisolvent with the local basis

$$\phi_i(\lambda) = 1 - d\lambda_i, \quad i = 1, \ldots, d+1.$$

□

5.3 Finite Elements on Parallelepipeds

Remark 5.31 *Reference mesh cells, reference map.* On can find in the literature two reference cells: the unit cube $[0,1]^d$ and the large unit cube $[-1,1]^d$. It does not matter which reference cell is chosen. Here, the large unit cube will be used: $\hat{K} = [-1,1]^d$. The class of admissible reference maps $\{F_K\}$ consists of bijective affine mappings of the form

$$F_K\hat{\mathbf{x}} = B\hat{\mathbf{x}} + \mathbf{b}, \quad B \in \mathbb{R}^{d\times d},\ \mathbf{b} \in \mathbb{R}^d.$$

If B is a diagonal matrix, then $\hat{K}$ is mapped to d-rectangles.

The class of mesh cells which are obtained in this way is not sufficient to triangulate general domains. If one wants to use more general mesh cells than parallelepipeds, then the class of admissible reference maps has to be enlarged, see Section 5.4. □

Definition 5.32 Polynomial space Q_k. Let $\mathbf{x} = (x_1, \ldots, x_d)^T$ and denote by $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_d)^T$ a multi-index. Then, the polynomial space Q_k is given by

$$Q_k = \operatorname{span}\left\{\prod_{i=1}^{d} x_i^{\alpha_i} = \mathbf{x}^{\boldsymbol{\alpha}} \ :\ 0 \le \alpha_i \le k \ \text{ for } \ i = 1, \ldots, d\right\}.$$

□

Example 5.33 Q_1 *vs.* P_1. The space Q_1 consists of all polynomials which are d-linear. Let $d = 2$, then it is

$$Q_1 = \operatorname{span}\{1, x, y, xy\},$$

whereas

$$P_1 = \operatorname{span}\{1, x, y\}.$$

□

Remark 5.34 *Finite elements on d-rectangles.* For simplicity of presentation, the examples below consider d-rectangles. In this case, the finite elements are just tensor products of one-dimensional finite elements. In particular, the basis functions can be written as products of one-dimensional basis functions. □

Example 5.35 Q_0 *: piecewise constant finite element.* Similarly to the P_0 space, the space Q_0 consists of piecewise constant, discontinuous functions. The functional is the value of the function in the barycenter of the mesh cell K and it holds $\dim Q_0(K) = 1$. □

Example 5.36 Q_1 *: conforming piecewise d-linear finite element.* This finite element space is a subspace of $C(\overline{\Omega})$. The functionals are the values of the function in the vertices of the mesh cell, see Figure 5.10. Hence, it is $\dim Q_1(K) = 2^d$.

The one-dimensional local basis functions, which will be used for the tensor product, are given by

$$\hat{\phi}_1(\hat{x}) = \frac{1}{2}(1 - \hat{x}), \quad \hat{\phi}_2(\hat{x}) = \frac{1}{2}(1 + \hat{x}).$$

With these functions, e.g., the basis functions in two dimensions are computed by

$$\hat{\phi}_1(\hat{x})\hat{\phi}_1(\hat{y}),\ \hat{\phi}_1(\hat{x})\hat{\phi}_2(\hat{y}),\ \hat{\phi}_2(\hat{x})\hat{\phi}_1(\hat{y}),\ \hat{\phi}_2(\hat{x})\hat{\phi}_2(\hat{y}).$$

The continuity of the functions of the finite element space Q_1 is proved in the same way as for simplicial finite elements. It is used that the restriction of a function from $Q_k(K)$ to a face E is a function from the space $Q_k(E)$, $k \ge 1$. □

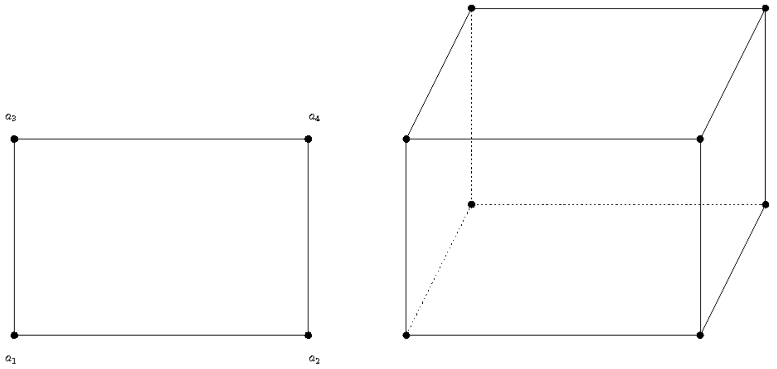

Figure 5.10: The finite element Q_1.

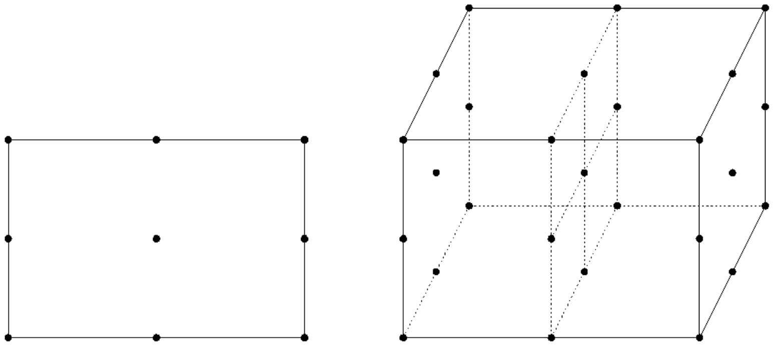

Figure 5.11: The finite element Q_2.

Example 5.37 Q_2 *: conforming piecewise d-quadratic finite element.* It holds that $Q_2 \subset C(\overline{\Omega})$. The functionals in one dimension are the values of the function at both ends of the interval and in the center of the interval, see Figure 5.11. In d dimensions, they are the corresponding values of the tensor product of the intervals. It follows that $\dim Q_2(K) = 3^d$.

The one-dimensional basis function on the reference interval are defined by

$$\hat{\phi}_1(\hat{x}) = -\frac{1}{2}\hat{x}(1-\hat{x}), \quad \hat{\phi}_2(\hat{x}) = (1-\hat{x})(1+\hat{x}), \quad \hat{\phi}_3(\hat{x}) = \frac{1}{2}(1+\hat{x})\hat{x}.$$

The basis function $\prod_{i=1}^{d} \hat{\phi}_2(\hat{x}_i)$ is called cell bubble function. □

Example 5.38 Q_3 *: conforming piecewise d-quadratic finite element.* This finite element space is a subspace of $C(\overline{\Omega})$. The functionals on the reference interval are given by the values at the end of the interval and the values at the points $\hat{x} = -1/3$, $\hat{x} = 1/3$. In multiple dimensions, it is the corresponding tensor product, see Figure 5.12. The dimension of the local space is $\dim Q_3(K) = 4^d$.

The one-dimensional basis functions in the reference interval are given by

$$\begin{aligned}
\hat{\phi}_1(\hat{x}) &= -\frac{1}{16}(3\hat{x}+1)(3\hat{x}-1)(\hat{x}-1), \\
\hat{\phi}_2(\hat{x}) &= \frac{9}{16}(\hat{x}+1)(3\hat{x}-1)(\hat{x}-1), \\
\hat{\phi}_3(\hat{x}) &= -\frac{9}{16}(\hat{x}+1)(3\hat{x}+1)(\hat{x}-1), \\
\hat{\phi}_4(\hat{x}) &= \frac{1}{16}(3\hat{x}+1)(3\hat{x}-1)(\hat{x}+1).
\end{aligned}$$

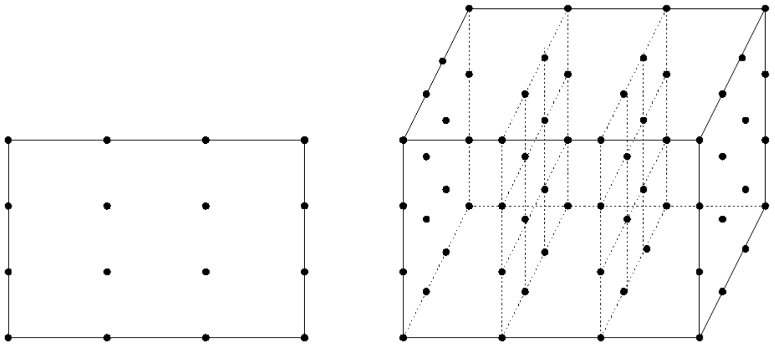

Figure 5.12: The finite element Q_3.

□

Example 5.39 Q_1^{rot} *: rotated nonconforming element of lowest order, Rannacher–Turek element Rannacher and Turek (1992):* This finite element space is a generalization of the P_1^{nc} finite element to quadrilateral and hexahedral mesh cells. It consists of discontinuous functions which are continuous at the barycenter of the faces. The dimension of the local finite element space is $\dim Q_1^{\text{rot}}(K) = 2d$. The space on the reference mesh cell is defined by

$$\begin{aligned} Q_1^{\text{rot}}\left(\hat{K}\right) &= \left\{\hat{p} \;:\; \hat{p} \in \operatorname{span}\{1, \hat{x}, \hat{y}, \hat{x}^2 - \hat{y}^2\}\right\} && \text{for } d = 2, \\ Q_1^{\text{rot}}\left(\hat{K}\right) &= \left\{\hat{p} \;:\; \hat{p} \in \operatorname{span}\{1, \hat{x}, \hat{y}, \hat{z}, \hat{x}^2 - \hat{y}^2, \hat{y}^2 - \hat{z}^2\}\right\} && \text{for } d = 3. \end{aligned}$$

Note that the transformed space

$$Q_1^{\text{rot}}(K) = \{p = \hat{p} \circ F_K^{-1}, \hat{p} \in Q_1^{\text{rot}}(\hat{K})\}$$

contains polynomials of the form $ax^2 - by^2$, where a, b depend on F_K.

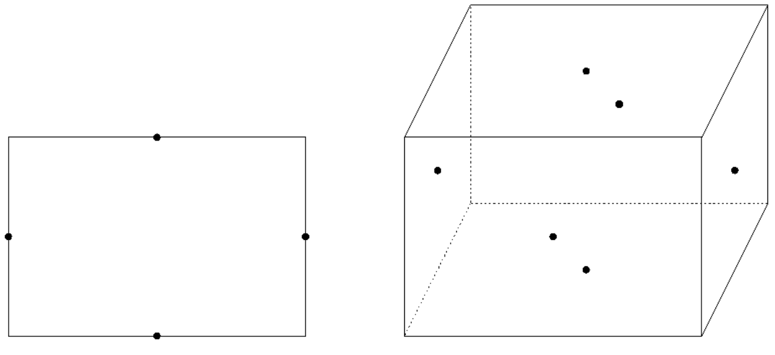

Figure 5.13: The finite element Q_1^{rot}.

For $d = 2$, the local basis on the reference cell is given by

$$\begin{aligned} \phi_1(\hat{x}, \hat{y}) &= -\frac{3}{8}(\hat{x}^2 - \hat{y}^2) - \frac{1}{2}\hat{y} + \frac{1}{4}, \\ \phi_2(\hat{x}, \hat{y}) &= \frac{3}{8}(\hat{x}^2 - \hat{y}^2) + \frac{1}{2}\hat{x} + \frac{1}{4}, \\ \phi_3(\hat{x}, \hat{y}) &= -\frac{3}{8}(\hat{x}^2 - \hat{y}^2) + \frac{1}{2}\hat{y} + \frac{1}{4}, \\ \phi_4(\hat{x}, \hat{y}) &= \frac{3}{8}(\hat{x}^2 - \hat{y}^2) - \frac{1}{2}\hat{x} + \frac{1}{4}. \end{aligned}$$

Analogously to the Crouzeix–Raviart finite element, the functionals can be defined as point values of the functions in the barycenters of the faces, see Figure 5.13, or as integral mean values of the functions at the faces. Consequently, the finite element spaces are defined in the same way as (5.3) or (5.4), with $P_1^{\mathrm{nc}}(K)$ replaced by $Q_1^{\mathrm{rot}}(K)$.

In the code MooNMD John and Matthies (2004), the mean value oriented Q_1^{rot} finite element space is implemented fro two dimensions and the point value oriented Q_1^{rot} finite element space for three dimensions. For $d = 3$, the integrals on the faces of mesh cells, whose equality is required in the mean value oriented Q_1^{rot} finite element space, involve a weighting function which depends on the particular mesh cell K. The computation of these weighting functions for all mesh cells is an additional computational overhead. For this reason, Schieweck (Schieweck, 1997, p. 21) suggested to use for $d = 3$ the simpler point value oriented form of the Q_1^{rot} finite element. □

5.4 Parametric Finite Elements on General d-Dimensional Quadrilaterals

Remark 5.40 *Parametric mappings.* The image of an affine mapping of the reference mesh cell $\hat{K} = [-1, 1]^d$, $d \in \{2, 3\}$, is a parallelepiped. If one wants to consider finite elements on general q-quadrilaterals, then the class of admissible reference maps has to be enlarged.

The simplest parametric finite element on quadrilaterals in two dimensions uses bilinear mappings. Let $\hat{K} = [-1, 1]^2$ and let

$$F_K(\hat{\mathbf{x}}) = \begin{pmatrix} F_K^1(\hat{\mathbf{x}}) \\ F_K^2(\hat{\mathbf{x}}) \end{pmatrix} = \begin{pmatrix} a_{11} + a_{12}\hat{x} + a_{13}\hat{y} + a_{14}\hat{x}\hat{y} \\ a_{21} + a_{22}\hat{x} + a_{23}\hat{y} + a_{24}\hat{x}\hat{y} \end{pmatrix}, F_K^i \in Q_1,\ i = 1, 2,$$

be a bilinear mapping from $\hat{K}$ on the class of admissible quadrilaterals. A quadrilateral K is called admissible if

- the length of all edges of K is larger than zero,
- the interior angles of K are smaller than π, i.e. K is convex.

This class contains, e.g., trapezoids and rhombi. □

Remark 5.41 *Parametric finite element functions.* The functions of the local space $P(K)$ on the mesh cell K are defined by $p = \hat{p} \circ F_K^{-1}$. These functions are in general rational functions. However, using d-linear mappings, then the restriction of F_K on an edge of $\hat{K}$ is an affine map. For instance, in the case of the Q_1 finite element, the functions on K are linear functions on each edge of K for this reason. It follows that the functions of the corresponding finite element space are continuous, see Example 5.26. □

5.5 Transform of Integrals

Remark 5.42 *Motivation.* The transform of integrals from the reference mesh cell to mesh cells of the grid and vice versa is used as well for analysis as for the implementation of finite element methods. This section provides an overview of the most important formulae for transforms.

Let $\hat{K} \subset \mathbb{R}^d$ be the reference mesh cell, K be an arbitrary mesh cell, and $F_K : \hat{K} \to K$ with $\mathbf{x} = F_K(\hat{\mathbf{x}})$ be the reference map. It is assumed that the reference map is a continuous differentiable one-to-one map. The inverse map is denoted by $F_K^{-1} : K \to \hat{K}$. For the integral transforms, the derivatives (Jacobians) of F_K and F_K^{-1} are needed

$$DF_K(\hat{\mathbf{x}})_{ij} = \frac{\partial x_i}{\partial \hat{x}_j}, \quad DF_K^{-1}(\mathbf{x})_{ij} = \frac{\partial \hat{x}_i}{\partial x_j}, \quad i,j = 1,\ldots,d.$$

□

Remark 5.43 *Integral with a function without derivatives.* This integral transforms with the standard rule of integral transforms

$$\int_K v(\mathbf{x})\ d\mathbf{x} = \int_{\hat{K}} \hat{v}(\hat{\mathbf{x}})\left|\det DF_K(\hat{\mathbf{x}})\right|\ d\hat{\mathbf{x}}, \tag{5.5}$$

where $\hat{v}(\hat{\mathbf{x}}) = v(F_K(\hat{\mathbf{x}}))$. □

Remark 5.44 *Transform of derivatives.* Using the chain rule, one obtains

$$\begin{aligned}\frac{\partial v}{\partial x_i}(\mathbf{x}) &= \sum_{j=1}^d \frac{\partial \hat{v}}{\partial \hat{x}_j}(\hat{\mathbf{x}})\frac{\partial \hat{x}_j}{\partial x_i} = \nabla_{\hat{\mathbf{x}}}\hat{v}(\hat{\mathbf{x}}) \cdot \left(\left(DF_K^{-1}(\mathbf{x})\right)^T\right)_i \\ &= \nabla_{\hat{\mathbf{x}}}\hat{v}(\hat{\mathbf{x}}) \cdot \left(\left(DF_K^{-1}(F_K(\hat{\mathbf{x}}))\right)^T\right)_i, \qquad (5.6)\\ \frac{\partial \hat{v}}{\partial \hat{x}}(\hat{\mathbf{x}}) &= \sum_{j=1}^d \frac{\partial v}{\partial x_j}(\mathbf{x})\frac{\partial x_j}{\partial \hat{x}_i} = \nabla v(\mathbf{x}) \cdot \left(\left(DF_K(\hat{\mathbf{x}})\right)^T\right)_i \\ &= \nabla v(\mathbf{x}) \cdot \left(\left(DF_K(F_K^{-1}(\mathbf{x}))\right)^T\right)_i. \qquad (5.7)\end{aligned}$$

The index i denotes the i-th row of a matrix. Derivatives on the reference mesh cell are marked with a symbol on the operator. □

Remark 5.45 *Integrals with a gradients.* Using the rule for transforming integrals and (5.6) gives

$$\begin{aligned}&\int_K \mathbf{b}(\mathbf{x}) \cdot \nabla v(\mathbf{x})\ d\mathbf{x} \\ &= \int_{\hat{K}} \mathbf{b}\left(F_K(\hat{\mathbf{x}})\right) \cdot \left[\left(DF_K^{-1}\right)^T\left(F_K(\hat{\mathbf{x}})\right)\right]\nabla_{\hat{\mathbf{x}}}\hat{v}(\hat{\mathbf{x}})\left|\det DF_K(\hat{\mathbf{x}})\right|\ d\hat{\mathbf{x}}. \qquad (5.8)\end{aligned}$$

Similarly, one obtains

$$\begin{aligned}&\int_K \nabla v(\mathbf{x}) \cdot \nabla w(\mathbf{x})\ d\mathbf{x} \\ &= \int_{\hat{K}} \left[\left(DF_K^{-1}\right)^T\left(F_K(\hat{\mathbf{x}})\right)\right]\nabla_{\hat{\mathbf{x}}}\hat{v}(\hat{\mathbf{x}}) \cdot \left[\left(DF_K^{-1}\right)^T\left(F_K(\hat{\mathbf{x}})\right)\right]\nabla_{\hat{\mathbf{x}}}\hat{w}(\hat{\mathbf{x}}) \\ &\quad \times \left|\det DF_K(\hat{\mathbf{x}})\right|\ d\hat{\mathbf{x}}. \qquad (5.9)\end{aligned}$$

□

Remark 5.46 *Integral with the divergence.* Integrals of the following type are important for the Navier–Stokes equations

$$\begin{aligned}&\int_K \nabla \cdot v(\mathbf{x})q(\mathbf{x})\ d\mathbf{x} = \int_K \sum_{i=1}^d \frac{\partial v_i}{\partial x_i}(\mathbf{x})q(\mathbf{x})\ d\mathbf{x} \\ &= \int_{\hat{K}} \sum_{i=1}^d \left[\left(\left(DF_K^{-1}(F_K(\hat{\mathbf{x}}))\right)^T\right)_i \cdot \nabla_{\hat{\mathbf{x}}}\hat{v}_i(\hat{\mathbf{x}})\right]\hat{q}(\hat{\mathbf{x}})\left|\det DF_K(\hat{\mathbf{x}})\right|\ d\hat{\mathbf{x}} \\ &= \int_{\hat{K}} \left[\left(DF_K^{-1}(F_K(\hat{\mathbf{x}}))\right)^T : D_{\hat{\mathbf{x}}}\mathbf{v}(\hat{\mathbf{x}})\right]\hat{q}(\hat{\mathbf{x}})\left|\det DF_K(\hat{\mathbf{x}})\right|\ d\hat{\mathbf{x}}. \qquad (5.10)\end{aligned}$$

In the derivation, (5.6) was used. □

Example 5.47 *Affine transform.* The most important class of reference maps are affine transforms

$$\mathbf{x} = B\hat{\mathbf{x}} + \mathbf{b}, \quad B \in \mathbb{R}^{d\times d}, \mathbf{b} \in \mathbb{R}^{d},$$

where the invertible matrix B and the vector $\mathbf{b}$ are constants. It follows that

$$\hat{\mathbf{x}} = B^{-1}\left(\mathbf{x} - \mathbf{b}\right) = B^{-1}\mathbf{x} - B^{-1}\mathbf{b}.$$

In this case, there are

$$DF_K = B, \quad DF_K^{-1} = B^{-1}, \quad \det DF_K = \det(B).$$

One obtains for the integral transforms from (5.5), (5.8), (5.9), and (5.10)

$$\begin{aligned}
\int_K v(\mathbf{x})\ d\mathbf{x} &= |\det(B)| \int_{\hat{K}} \hat{v}(\hat{\mathbf{x}})\ d\hat{\mathbf{x}}, && (5.11)\\
\int_K \mathbf{b}(\mathbf{x}) \cdot \nabla v(\mathbf{x})\ d\mathbf{x} &= |\det(B)| \int_{\hat{K}} \mathbf{b}\left(F_K(\hat{\mathbf{x}})\right) \cdot B^{-T}\nabla_{\hat{\mathbf{x}}}\hat{v}(\hat{\mathbf{x}})\ d\hat{\mathbf{x}}, && (5.12)\\
\int_K \nabla v(\mathbf{x}) \cdot \nabla w(\mathbf{x})\ d\mathbf{x} &= |\det(B)| \int_{\hat{K}} B^{-T}\nabla_{\hat{\mathbf{x}}}\hat{v}(\hat{\mathbf{x}}) \cdot B^{-T}\nabla_{\hat{\mathbf{x}}}\hat{w}(\hat{\mathbf{x}})\ d\hat{\mathbf{x}}, && (5.13)\\
\int_K \nabla \cdot v(\mathbf{x}) q(\mathbf{x})\ d\mathbf{x} &= |\det(B)| \int_{\hat{K}} \left[B^{-T} : D_{\hat{\mathbf{x}}}\mathbf{v}(\hat{\mathbf{x}})\right] \hat{q}(\hat{\mathbf{x}})\ d\hat{\mathbf{x}}. && (5.14)
\end{aligned}$$

□

Chapter 6

Interpolation

Remark 6.1 *Motivation.* Variational forms of partial differential equations use functions in Sobolev spaces. The solution of these equations shall be approximated with the Ritz method in finite dimensional spaces, the finite element spaces. The best possible approximation of an arbitrary function from the Sobolev space by a finite element function is a factor in the upper bound for the finite element error, e.g., see the Lemma of Cea, estimate (4.19).

This section studies the approximation quality of finite element spaces. Estimates are proved for interpolants of functions. Interpolation estimates are of course upper bounds for the best approximation error and they can serve as factors in finite element error estimates. □

6.1 Interpolation in Sobolev Spaces by Polynomials

Lemma 6.2 Unique determination of a polynomial with integral conditions. *Let Ω be a bounded domain in $\mathbb{R}^d$ with Lipschitz boundary. Let $m \in \mathbb{N} \cup \{0\}$ be given and let for all derivatives with multi-index $\boldsymbol{\alpha}$, $|\boldsymbol{\alpha}| \le m$, a value $a_{\boldsymbol{\alpha}} \in \mathbb{R}$ be given. Then, there is a uniquely determined polynomial $p \in P_m(\Omega)$ such that*

$$\int_\Omega \partial_{\boldsymbol{\alpha}} p(\mathbf{x})\ d\mathbf{x} = a_{\boldsymbol{\alpha}}, \quad |\boldsymbol{\alpha}| \le m. \tag{6.1}$$

Proof: Let $p \in P_m(\Omega)$ be an arbitrary polynomial. It has the form

$$p(\mathbf{x}) = \sum_{|\boldsymbol{\beta}| \le m} b_{\boldsymbol{\beta}} \mathbf{x}^{\boldsymbol{\beta}}.$$

Inserting this representation into (6.1) leads to a linear system of equations $M\mathbf{b} = \mathbf{a}$ with

$$M = (M_{\boldsymbol{\alpha}\boldsymbol{\beta}}),\ M_{\boldsymbol{\alpha}\boldsymbol{\beta}} = \int_\Omega \partial_{\boldsymbol{\alpha}} \mathbf{x}^{\boldsymbol{\beta}}\ d\mathbf{x},\ \mathbf{b} = (b_{\boldsymbol{\beta}}),\ \mathbf{a} = (a_{\boldsymbol{\alpha}}),$$

for $|\boldsymbol{\alpha}|, |\boldsymbol{\beta}| \le m$. Since M is a squared matrix, the linear system of equations possesses a unique solution if and only if M is non-singular.

The proof is performed by contradiction. Assume that M is singular. Then there exists a non-trivial solution of the homogeneous system. That means, there is a polynomial $q \in P_m(\Omega) \setminus \{0\}$ with

$$\int_\Omega \partial_{\boldsymbol{\alpha}} q(\mathbf{x})\ d\mathbf{x} = 0 \text{ for all } |\boldsymbol{\alpha}| \le m.$$

The polynomial $q(\mathbf{x})$ has the representation $q(\mathbf{x}) = \sum_{|\boldsymbol{\beta}| \le m} c_{\boldsymbol{\beta}} \mathbf{x}^{\boldsymbol{\beta}}$. Now, one can choose a $c_{\boldsymbol{\beta}} \neq 0$ with maximal value $|\boldsymbol{\beta}|$. Then, it is $\partial_{\boldsymbol{\beta}} q(\mathbf{x}) = C c_{\boldsymbol{\beta}} = \mathit{const} \neq 0$, where $C > 0$ comes from the differentiation rule for polynomials, which is a contradiction to the vanishing of the integral for $\partial_{\boldsymbol{\beta}} q(\mathbf{x})$. ■

Remark 6.3 *To Lemma 6.2.* Lemma 6.2 states that a polynomial is uniquely determined if a condition on the integral on Ω is prescribed for each derivative. □

Lemma 6.4 Poincaré-type inequality. *Denote by $D^k v(\mathbf{x})$, $k \in \mathbb{N} \cup \{0\}$, the total derivative of order k of a function $v(\mathbf{x})$, e.g., for $k = 1$ the gradient of $v(\mathbf{x})$. Let Ω be convex and be included into a ball of radius R. Let $k, l \in \mathbb{N} \cup \{0\}$ with $k \le l$ and let $p \in \mathbb{R}$ with $p \in [1, \infty]$. Assume that $v \in W^{l,p}(\Omega)$ satisfies*

$$\int_\Omega \partial_{\boldsymbol{\alpha}} v(\mathbf{x})\ d\mathbf{x} = 0 \text{ for all } |\boldsymbol{\alpha}| \le l - 1,$$

then it holds the estimate

$$\left\| D^k v \right\|_{L^p(\Omega)} \le C R^{l-k} \left\| D^l v \right\|_{L^p(\Omega)},$$

where the constant C does not depend on Ω and on $v(\mathbf{x})$.

Proof: There is nothing to prove if $k = l$. In addition, it suffices to prove the lemma for $k = 0$ and $l = 1$, since the general case follows by applying the result to $\partial_{\boldsymbol{\alpha}} v(\mathbf{x})$. Only the case $p < \infty$ will be discussed here in detail.

Since Ω is assumed to be convex, the integral mean value theorem can be written in the form

$$v(\mathbf{x}) - v(\mathbf{y}) = \int_0^1 \nabla v(t\mathbf{x} + (1-t)\mathbf{y}) \cdot (\mathbf{x} - \mathbf{y})\ dt, \quad \mathbf{x}, \mathbf{y} \in \Omega.$$

Integration with respect to $\mathbf{y}$ yields

$$v(\mathbf{x}) \int_\Omega d\mathbf{y} - \int_\Omega v(\mathbf{y})\ d\mathbf{y} = \int_\Omega \int_0^1 \nabla v(t\mathbf{x} + (1-t)\mathbf{y}) \cdot (\mathbf{x} - \mathbf{y})\ dt\ d\mathbf{y}.$$

It follows from the assumption that the second integral on the left hand side vanishes. Hence, one gets

$$v(\mathbf{x}) = \frac{1}{|\Omega|} \int_\Omega \int_0^1 \nabla v(t\mathbf{x} + (1-t)\mathbf{y}) \cdot (\mathbf{x} - \mathbf{y})\ dt\ d\mathbf{y}.$$

Now, taking the absolute value on both sides, using that the absolute value of an integral is estimated from above by the integral of the absolute value, applying the Cauchy–Schwarz inequality for vectors and the estimate $\|\mathbf{x} - \mathbf{y}\|_2 \le 2R$ yields

$$\begin{aligned} |v(\mathbf{x})| &= \frac{1}{|\Omega|} \left| \int_\Omega \int_0^1 \nabla v(t\mathbf{x} + (1-t)\mathbf{y}) \cdot (\mathbf{x} - \mathbf{y})\ dt\ d\mathbf{y} \right| \\ &\le \frac{1}{|\Omega|} \int_\Omega \int_0^1 \left| \nabla v(t\mathbf{x} + (1-t)\mathbf{y}) \cdot (\mathbf{x} - \mathbf{y}) \right|\ dt\ d\mathbf{y} \\ &\le \frac{2R}{|\Omega|} \int_\Omega \int_0^1 \left\| \nabla v(t\mathbf{x} + (1-t)\mathbf{y}) \right\|_2\ dt\ d\mathbf{y}. \end{aligned} \tag{6.2}$$

Then (6.2) is raised to the power p and then integrated with respect to $\mathbf{x}$. One obtains with Hölder's inequality (3.4), with $p^{-1} + q^{-1} = 1 \implies p/q - p = p(1/q - 1) = -1$, that

$$\begin{aligned} \int_\Omega |v(\mathbf{x})|^p\ d\mathbf{x} &\le \frac{CR^p}{|\Omega|^p} \int_\Omega \left(\int_\Omega \int_0^1 \left\| \nabla v(t\mathbf{x} + (1-t)\mathbf{y}) \right\|_2\ dt\ d\mathbf{y} \right)^p d\mathbf{x} \\ &\le \frac{CR^p}{|\Omega|^p} \int_\Omega \Bigg[\underbrace{\left(\int_\Omega \int_0^1 1^q\ dt\ d\mathbf{y} \right)^{p/q}}_{|\Omega|^{p/q}} \\ &\qquad \times \left(\int_\Omega \int_0^1 \left\| \nabla v(t\mathbf{x} + (1-t)\mathbf{y}) \right\|_2^p\ dt\ d\mathbf{y} \right) \Bigg] d\mathbf{x} \\ &= \frac{CR^p}{|\Omega|} \int_\Omega \left(\int_\Omega \int_0^1 \left\| \nabla v(t\mathbf{x} + (1-t)\mathbf{y}) \right\|_2^p\ dt\ d\mathbf{y} \right) d\mathbf{x}. \end{aligned}$$

Applying the theorem of Fubini allows the commutation of the integration

$$\int_\Omega |v(\mathbf{x})|^p \ d\mathbf{x} \le \frac{CR^p}{|\Omega|} \int_0^1 \int_\Omega \left(\int_\Omega \|\nabla v(t\mathbf{x} + (1-t)\mathbf{y})\|_2^p \ d\mathbf{y} \right) d\mathbf{x}\ dt.$$

Using the integral mean value theorem in one dimension gives that there is a $t_0 \in [0,1]$, such that

$$\int_\Omega |v(\mathbf{x})|^p \ d\mathbf{x} \le \frac{CR^p}{|\Omega|} \int_\Omega \left(\int_\Omega \|\nabla v(t_0\mathbf{x} + (1-t_0)\mathbf{y})\|_2^p \ d\mathbf{y} \right) d\mathbf{x}.$$

The function $\|\nabla v(\mathbf{x})\|_2^p$ will be extended to $\mathbb{R}^d$ by zero and the extension will be also denoted by $\|\nabla v(\mathbf{x})\|_2^p$. Then, it is

$$\int_\Omega |v(\mathbf{x})|^p \ d\mathbf{x} \le \frac{CR^p}{|\Omega|} \int_\Omega \left(\int_{\mathbb{R}^d} \|\nabla v(t_0\mathbf{x} + (1-t_0)\mathbf{y})\|_2^p \ d\mathbf{y} \right) d\mathbf{x}. \tag{6.3}$$

Let $t_0 \in [0,1/2]$. Since the domain of integration is $\mathbb{R}^d$, a substitution of variables $t_0\mathbf{x} + (1-t_0)\mathbf{y} = \mathbf{z}$ can be applied and leads to

$$\int_{\mathbb{R}^d} \|\nabla v(t_0\mathbf{x} + (1-t_0)\mathbf{y})\|_2^p \ d\mathbf{y} = \frac{1}{1-t_0} \int_{\mathbb{R}^d} \|\nabla v(\mathbf{z})\|_2^p \ d\mathbf{z} \le 2 \|\nabla v\|_{L^p(\Omega)}^p,$$

since $1/(1-t_0) \le 2$. Inserting this expression into (6.3) gives

$$\int_\Omega |v(\mathbf{x})|^p \ d\mathbf{x} \le 2CR^p \|\nabla v\|_{L^p(\Omega)}^p.$$

If $t_0 > 1/2$ then one changes the roles of $\mathbf{x}$ and $\mathbf{y}$, applies the theorem of Fubini to change the sequence of integration, and uses the same arguments.

The estimate for the case $p = \infty$ is also based on (6.2). ■

Remark 6.5 *On Lemma 6.4.* The Lemma 6.4 proves an inequality of Poincaré-type. It says that it is possible to estimate the $L^p(\Omega)$ norm of a lower derivative of a function $v(\mathbf{x})$ by the same norm of a higher derivative if the integral mean values of some lower derivatives vanish.

An important application of Lemma 6.4 is in the proof of the Bramble–Hilbert lemma. The Bramble–Hilbert lemma considers a continuous linear functional which is defined on a Sobolev space and which vanishes for all polynomials of degree less or equal than m. It states that the value of the functional can be estimated by the Lebesgue norm of the $(m+1)$th total derivative of the functions from this Sobolev space. □

Theorem 6.6 ***Bramble–Hilbert lemma.*** *Let $m \in \mathbb{N} \cup \{0\}$, $m \ge 0$, $p \in [1,\infty]$, and $F : W^{m+1,p}(\Omega) \to \mathbb{R}$ be a continuous linear functional, and let the conditions of Lemma 6.2 and 6.4 be satisfied. Let*

$$F(p) = 0 \quad \forall \ p \in P_m(\Omega),$$

then there is a constant $C(\Omega)$, which is independent of $v(\mathbf{x})$ and F, such that

$$|F(v)| \le C(\Omega) \left\| D^{m+1} v \right\|_{L^p(\Omega)} \quad \forall \ v \in W^{m+1,p}(\Omega).$$

Proof: Let $v \in W^{m+1,p}(\Omega)$. It follows from Lemma 6.2 that there is a polynomial from $P_m(\Omega)$ with

$$\int_\Omega \partial_{\boldsymbol{\alpha}}(v+p)(\mathbf{x}) \ d\mathbf{x} = 0 \text{ for } |\boldsymbol{\alpha}| \le m.$$

Lemma 6.4 gives, with $l = m+1$ and considering each term in $\|\cdot\|_{W^{m+1,p}(\Omega)}$ individually, the estimate

$$\|v+p\|_{W^{m+1,p}(\Omega)} \le C(\Omega) \left\| D^{m+1}(v+p) \right\|_{L^p(\Omega)} = C(\Omega) \left\| D^{m+1} v \right\|_{L^p(\Omega)}.$$

From the vanishing of F for $p \in P_m(\Omega)$ and the continuity of F it follows that

$$|F(v)| = |F(v+p)| \le c \, \|v+p\|_{W^{m+1,p}(\Omega)} \le C(\Omega) \left\| D^{m+1} v \right\|_{L^p(\Omega)} .$$

■

Remark 6.7 *Strategy for estimating the interpolation error.* The Bramble–Hilbert lemma will be used for estimating the interpolation error for an affine family of finite elements. The strategy is as follows:

- Show first the estimate on the reference mesh cell $\hat{K}$.
- Transform the estimate on an arbitrary mesh cell K to the reference mesh cell $\hat{K}$.
- Apply the estimate on $\hat{K}$.
- Transform back to K.

One has to study what happens if the transforms are applied to the estimate. □

Remark 6.8 *Assumptions, definition of the interpolant.* Let $\hat{K} \subset \mathbb{R}^d, d \in \{2,3\}$, be a reference mesh cell (compact polyhedron), $\hat{P}(\hat{K})$ a polynomial space of dimension N, and $\hat{\Phi}_1, \ldots, \hat{\Phi}_N \; : \; C^s(\hat{K}) \to \mathbb{R}$ continuous linear functionals. It will be assumed that the space $\hat{P}(\hat{K})$ is unisolvent with respect to these functionals. Then, there is a local basis $\hat{\phi}_1, \ldots, \hat{\phi}_N \in \hat{P}(\hat{K})$.

Consider $\hat{v} \in C^s(\hat{K})$, then the interpolant $I_{\hat{K}} \hat{v} \in \hat{P}(\hat{K})$ is defined by

$$I_{\hat{K}} \hat{v}(\hat{\mathbf{x}}) = \sum_{i=1}^{N} \hat{\Phi}_i(\hat{v}) \hat{\phi}_i(\hat{\mathbf{x}}).$$

The operator $I_{\hat{K}}$ is a continuous and linear operator from $C^s(\hat{K})$ to $\hat{P}(\hat{K})$. From the linearity it follows that $I_{\hat{K}}$ is the identity on $\hat{P}(\hat{K})$

$$I_{\hat{K}} \hat{p} = \hat{p} \quad \forall \;\; \hat{p} \in \hat{P}(\hat{K}).$$

□

Example 6.9 *Interpolation operators.*

- Let $\hat{K} \subset \mathbb{R}^d$ be an arbitrary reference cell, $\hat{P}(\hat{K}) = P_0(\hat{K})$, and

$$\hat{\Phi}(\hat{v}) = \frac{1}{\left|\hat{K}\right|} \int_{\hat{K}} \hat{v}(\hat{\mathbf{x}}) \; d\hat{\mathbf{x}}.$$

The functional $\hat{\Phi}$ is continuous on $C^0(\hat{K})$ since

$$\left|\hat{\Phi}(\hat{v})\right| \le \frac{1}{\left|\hat{K}\right|} \int_{\hat{K}} |\hat{v}(\hat{\mathbf{x}})| \;\; d\hat{\mathbf{x}} \le \frac{\left|\hat{K}\right|}{\left|\hat{K}\right|} \max_{\hat{\mathbf{x}} \in \hat{K}} |\hat{v}(\hat{\mathbf{x}})| = \|\hat{v}\|_{C^0(\hat{K})} .$$

For the constant function $1 \in P_0(\hat{K})$ it is $\hat{\Phi}(1) = 1 \neq 0$. Hence, $\{\hat{\phi}\} = \{1\}$ is the local basis and the space is unisolvent with respect to $\hat{\Phi}$. The operator

$$I_{\hat{K}} \hat{v}(\hat{\mathbf{x}}) = \hat{\Phi}(\hat{v}) \hat{\phi}(\hat{\mathbf{x}}) = \frac{1}{\left|\hat{K}\right|} \int_{\hat{K}} \hat{v}(\hat{\mathbf{x}}) \; d\hat{\mathbf{x}}$$

is an integral mean value operator, i.e., each continuous function on $\hat{K}$ will be approximated by a constant function whose value equals the integral mean value, see Figure 6.1

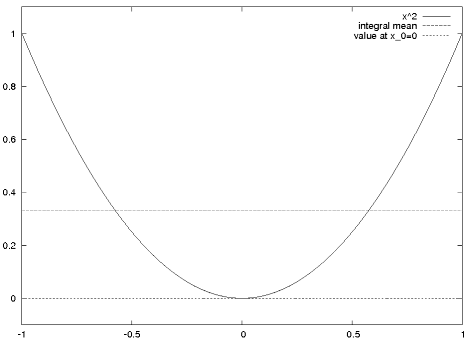

Figure 6.1: Interpolation of x^2 in $[-1,1]$ by a P_0 function with the integral mean value and with the value of the function at $x_0 = 0$.

- It is possible to define $\hat{\Phi}(\hat{v}) = \hat{v}(\hat{\mathbf{x}}_0)$ for an arbitrary point $\hat{\mathbf{x}}_0 \in \hat{K}$. This functional is also linear and continuous in $C^0(\hat{K})$. The interpolation operator $I_{\hat{K}}$ defined in this way interpolates each continuous function by a constant function whose value is equal to the value of the function at $\hat{\mathbf{x}}_0$, see also Figure 6.1.

  Interpolation operators which are defined by using values of functions, are called Lagrangian interpolation operators.

This example demonstrates that the interpolation operator $I_{\hat{K}}$ depends on $\hat{P}(\hat{K})$ and on the functionals $\hat{\Phi}_i$. □

Theorem 6.10 Interpolation error estimate on a reference mesh cell. *Let $P_m(\hat{K}) \subset \hat{P}(\hat{K})$ and $p \in [1,\infty]$ with $(m+1-s)p > d$. Then there is a constant C that is independent of $\hat{v}(\hat{\mathbf{x}})$ such that*

$$\left\|\hat{v} - I_{\hat{K}}\hat{v}\right\|_{W^{m+1,p}(\hat{K})} \le C \left\|D^{m+1}\hat{v}\right\|_{L^p(\hat{K})} \quad \forall\, \hat{v} \in W^{m+1,p}(\hat{K}). \tag{6.4}$$

Proof: Because of the Sobolev imbedding, Theorem 3.53, ($\lambda = 0, j = s, m$ (of Sobolev imbedding) $= m+1-s$) it holds that

$$W^{m+1,p}(\hat{K}) \to C^s(\hat{K})$$

if $(m+1-s)p > d$. That means, the interpolation operator is well defined in $W^{m+1,p}(\hat{K})$. From the identity of the interpolation operator in $P_m(\hat{K})$, the triangle inequality, the boundedness of the interpolation operator (it is a linear and continuous operator mapping $C^s(\hat{K}) \to \hat{P}(\hat{K}) \subset W^{m+1,p}(\hat{K})$), and the Sobolev imbedding, one obtains for $\hat{q} \in P_m(\hat{K})$

$$\begin{aligned}
\|\hat{v} - I_{\hat{K}}\hat{v}\|_{W^{m+1,p}(\hat{K})} &= \|\hat{v} + \hat{q} - I_{\hat{K}}(\hat{v} + \hat{q})\|_{W^{m+1,p}(\hat{K})} \\
&\le \|\hat{v} + \hat{q}\|_{W^{m+1,p}(\hat{K})} + \|I_{\hat{K}}(\hat{v} + \hat{q})\|_{W^{m+1,p}(\hat{K})} \\
&\le \|\hat{v} + \hat{q}\|_{W^{m+1,p}(\hat{K})} + c\,\|\hat{v} + \hat{q}\|_{C^s(\hat{K})} \\
&\le c\,\|\hat{v} + \hat{q}\|_{W^{m+1,p}(\hat{K})}\,.
\end{aligned}$$

Choosing $\hat{q}(\hat{\mathbf{x}})$ in Lemma 6.2 such that

$$\int_{\hat{K}} \partial_{\boldsymbol{\alpha}}(\hat{v} + \hat{q})\ d\hat{\mathbf{x}} = 0 \quad \forall\ |\boldsymbol{\alpha}| \le m,$$

the assumptions of Lemma 6.4 are satisfied. It follows that

$$\|\hat{v} + \hat{q}\|_{W^{m+1,p}(\hat{K})} \le c \left\|D^{m+1}(\hat{v} + \hat{q})\right\|_{L^p(\hat{K})} = c \left\|D^{m+1}\hat{v}\right\|_{L^p(\hat{K})}.$$

■

Remark 6.11 *On Theorem 6.10.*

- One can construct examples which show that the Sobolev imbedding is not valid if $(m+1-s)p > d$ is not satisfied. In the case $(m+1-s)p \leq d$, the statement of Theorem 6.10 is not true.
  Consider the interpolation of continuous functions ($s = 0$) with piecewise linear elements ($m = 1$) in Sobolev spaces that are also Hilbert spaces ($p = 2$). Then $(m+1-s)p = 4$ and it follows that the theorem can be applied only for $d \in \{2, 3\}$. For piecewise constant finite elements, the statement of the theorem is true only for $d = 1$.
- The theorem requires only that $P_m(\hat{K}) \subset \hat{P}(\hat{K})$. This requirement does not exclude that $\hat{P}(\hat{K})$ contains polynomials of higher degree, too. However, this property is not utilized and also not needed if the other assumptions of the theorem are satisfied.

□

Remark 6.12 *Assumptions on the triangulation.* For deriving the interpolation error estimate for arbitrary mesh cells K, and finally for the finite element space, one has to study the properties of the affine mapping from K to $\hat{K}$ and of the back mapping.

Consider an affine family of finite elements whose mesh cells are generated by affine mappings

$$F_K \hat{\mathbf{x}} = B\hat{\mathbf{x}} + \mathbf{b},$$

where B is a non-singular $d \times d$ matrix and $\mathbf{b}$ is a d vector.

Let h_K be the diameter of $K = F_K(\hat{K})$, i.e., the largest distance of two points that are contained in K. The images $\{K = F_K(\hat{K})\}$ are assumed to satisfy the following conditions:

- $K \subset \mathbb{R}^d$ is contained in a ball of radius $C_R h_K$,
- K contains a ball of radius $C_R^{-1} h_K$,

where the constant C_R is independent of K. Hence, it follows for all K that

$$\frac{\text{radius of circumcircle}}{\text{radius of inscribed circle}} \leq C_R^2.$$

A triangulation with this property is called a quasi-uniform triangulation. □

Lemma 6.13 Estimates of matrix norms. *For each matrix norm $\|\cdot\|$ one has the estimates*

$$\|B\| \leq ch_K, \quad \left\|B^{-1}\right\| \leq ch_K^{-1},$$

where the constants depend on the matrix norm and on C_R.

Proof: Since $\hat{K}$ is a Lipschitz domain with polyhedral boundary, it contains a ball $B(\hat{\mathbf{x}}_0, r)$ with $\hat{\mathbf{x}}_0 \in \hat{K}$ and some $r > 0$. Hence, $\hat{\mathbf{x}}_0 + \hat{\mathbf{y}} \in \hat{K}$ for all $\|\hat{\mathbf{y}}\|_2 = r$. It follows that the images

$$\mathbf{x}_0 = B\hat{\mathbf{x}}_0 + \mathbf{b}, \quad \mathbf{x} = B(\hat{\mathbf{x}}_0 + \hat{\mathbf{y}}) + \mathbf{b} = \mathbf{x}_0 + B\hat{\mathbf{y}}$$

are contained in K. Since the triangulation is assumed to be quasi-uniform, one obtains for all $\hat{\mathbf{y}}$

$$\|B\hat{\mathbf{y}}\|_2 = \|\mathbf{x} - \mathbf{x}_0\|_2 \leq C_R h_K.$$

Now, it holds for the spectral norm that

$$\|B\|_2 = \sup_{\hat{\mathbf{z}} \neq \mathbf{0}} \frac{\|B\hat{\mathbf{z}}\|_2}{\|\hat{\mathbf{z}}\|_2} = \frac{1}{r} \sup_{\|\hat{\mathbf{z}}\|_2 = r} \|B\hat{\mathbf{z}}\|_2 \leq \frac{C_R}{r} h_K.$$

An estimate of this form, with a possible different constant, holds also for all other matrix norms since all matrix norms are equivalent.

The estimate for $\left\|B^{-1}\right\|$ proceeds in the same way with interchanging the roles of K and $\hat{K}$. ■

Theorem 6.14 Local interpolation estimate. *Let an affine family of finite elements be given by its reference cell $\hat{K}$, the functionals $\{\hat{\Phi}_i\}$, and a space of polynomials $\hat{P}(\hat{K})$. Let all assumptions of Theorem 6.10 be satisfied. Then, for all $v \in W^{m+1,p}(K)$ there is a constant C, which is independent of $v(\mathbf{x})$ such that*

$$\left\|D^k(v - I_K v)\right\|_{L^p(K)} \le C h_K^{m+1-k} \left\|D^{m+1} v\right\|_{L^p(K)}, \quad k \le m+1. \tag{6.5}$$

Proof: The idea of the proof consists in transforming left hand side of (6.5) to the reference cell, using the interpolation estimate on the reference cell and transforming back.

i). Denote the elements of the matrices B and B^{-1} by b_{ij} and $b_{ij}^{(-1)}$, respectively. Since $\|B\|_\infty = \max_{i,j} |b_{ij}|$ is also a matrix norm, it holds that

$$|b_{ij}| \le C h_K, \quad \left|b_{ij}^{(-1)}\right| \le C h_K^{-1}. \tag{6.6}$$

Using element-wise estimates for the matrix B (Leibniz formula for determinants), one obtains

$$|\det B| \le C h_K^d, \quad \left|\det B^{-1}\right| \le C h_K^{-d}. \tag{6.7}$$

ii). The next step consists in proving that the transformed interpolation operator is equal to the natural interpolation operator on K. The latter one is given by

$$I_K v = \sum_{i=1}^N \Phi_{K,i}(v) \phi_{K,i}, \tag{6.8}$$

where $\{\phi_{K,i}\}$ is the basis of the space

$$P(K) = \{p \ : \ K \to \mathbb{R} \ : \ p = \hat{p} \circ F_K^{-1}, \hat{p} \in \hat{P}(\hat{K})\},$$

which satisfies $\Phi_{K,i}(\phi_{K,j}) = \delta_{ij}$. The functionals are defined by

$$\Phi_{K,i}(v) = \hat{\Phi}_i(v \circ F_K)$$

Hence, it follows with $v = \hat{\phi}_j \circ F_K^{-1}$ from the condition on the local basis on $\hat{K}$ that

$$\Phi_{K,i}(\hat{\phi}_j \circ F_K^{-1}) = \hat{\Phi}_i(\hat{\phi}_j) = \delta_{ij},$$

i.e., the local basis on K is given by $\phi_{K,j} = \hat{\phi}_j \circ F_K^{-1}$. Using (6.8), one gets

$$\begin{aligned} I_{\hat{K}} \hat{v} &= \sum_{i=1}^N \hat{\Phi}_i(\hat{v})\hat{\phi}_i = \sum_{i=1}^N \Phi_{K,i}(\underbrace{\hat{v} \circ F_K^{-1}}_{=v})\ \phi_{K,i} \circ F_K = \left(\sum_{i=1}^N \Phi_{K,i}(v)\phi_{K,i}\right) \circ F_K \\ &= I_K v \circ F_K. \end{aligned}$$

Hence, $I_{\hat{K}} \hat{v}$ is transformed correctly.

iii). One obtains with the chain rule

$$\frac{\partial v(\mathbf{x})}{\partial \mathbf{x}_i} = \sum_{j=1}^d \frac{\partial \hat{v}(\hat{\mathbf{x}})}{\partial \hat{\mathbf{x}}_j} b_{ji}^{(-1)}, \quad \frac{\partial \hat{v}(\hat{\mathbf{x}})}{\partial \hat{\mathbf{x}}_i} = \sum_{j=1}^d \frac{\partial v(\mathbf{x})}{\partial \mathbf{x}_j} b_{ji}.$$

It follows with (6.6) that (with each derivative one obtains an additional factor of B or B^{-1}, respectively)

$$\left\|D_{\mathbf{x}}^k v(\mathbf{x})\right\|_2 \le C h_K^{-k} \left\|D_{\hat{\mathbf{x}}}^k \hat{v}(\hat{\mathbf{x}})\right\|_2, \quad \left\|D_{\hat{\mathbf{x}}}^k \hat{v}(\hat{\mathbf{x}})\right\|_2 \le C h_K^{k} \left\|D_{\mathbf{x}}^k v(\mathbf{x})\right\|_2.$$

One gets with (6.7)

$$\int_K \left\|D_{\mathbf{x}}^k v(\mathbf{x})\right\|_2^p \, d\mathbf{x} \le C h_K^{-kp} |\det B| \int_{\hat{K}} \left\|D_{\hat{\mathbf{x}}}^k \hat{v}(\hat{\mathbf{x}})\right\|_2^p \, d\hat{\mathbf{x}} \le C h_K^{-kp+d} \int_{\hat{K}} \left\|D_{\hat{\mathbf{x}}}^k \hat{v}(\hat{\mathbf{x}})\right\|_2^p \, d\hat{\mathbf{x}}$$

and

$$\int_{\hat{K}} \left\|D_{\hat{\mathbf{x}}}^k \hat{v}(\hat{\mathbf{x}})\right\|_2^p \, d\hat{\mathbf{x}} \le C h_K^{kp} \left|\det B^{-1}\right| \int_K \left\|D_{\mathbf{x}}^k v(\mathbf{x})\right\|_2^p \, d\mathbf{x} \le C h_K^{kp-d} \int_K \left\|D_{\mathbf{x}}^k v(\mathbf{x})\right\|_2^p \, d\mathbf{x}.$$

Using now the interpolation estimate on the reference cell (6.4) yields

$$\left\| D_{\hat{\mathbf{x}}}^k (\hat{v} - I_{\hat{K}} \hat{v}) \right\|_{L^p(\hat{K})}^p \le C \left\| D_{\hat{\mathbf{x}}}^{m+1} \hat{v} \right\|_{L^p(\hat{K})}^p , \quad 0 \le k \le m+1.$$

It follows that

$$\begin{aligned} \left\| D_{\mathbf{x}}^k (v - I_K v) \right\|_{L^p(K)}^p &\le C h_K^{-kp+d} \left\| D_{\hat{\mathbf{x}}}^k (\hat{v} - I_{\hat{K}} \hat{v}) \right\|_{L^p(\hat{K})}^p \\ &\le C h_K^{-kp+d} \left\| D_{\hat{\mathbf{x}}}^{m+1} \hat{v} \right\|_{L^p(\hat{K})}^p \\ &\le C h_K^{(m+1-k)p} \left\| D_{\mathbf{x}}^{m+1} v \right\|_{L^p(K)}^p . \end{aligned}$$

Taking the p-th root proves the statement of the theorem. ■

Remark 6.15 *On estimate* (6.5)*.*

- Note that the power of h_K does not depend on p and d.
- Consider a quasi-uniform triangulation and define

$$h = \max_{K \in \mathcal{T}^h} \{h_K\}.$$

Then, one obtains by summing over all mesh cells an interpolation estimate for the global finite element space

$$\begin{aligned} \left\| D^k (v - I_h v) \right\|_{L^p(\Omega)} &= \left(\sum_{K \in \mathcal{T}^h} \left\| D^k (v - I_K v) \right\|_{L^p(K)}^p \right)^{1/p} \\ &\le \left(\sum_{K \in \mathcal{T}^h} c h_K^{(m+1-k)p} \left\| D^{m+1} v \right\|_{L^p(K)}^p \right)^{1/p} \\ &\le c h^{(m+1-k)} \left\| D^{m+1} v \right\|_{L^p(\Omega)} . \end{aligned} \qquad (6.9)$$

For linear finite elements P_1 $(m = 1)$ it is, in particular,

$$\left\| v - I_h v \right\|_{L^p(\Omega)} \le c h^2 \left\| D^2 v \right\|_{L^p(\Omega)} , \quad \left\| \nabla (v - I_h v) \right\|_{L^p(\Omega)} \le c h \left\| D^2 v \right\|_{L^p(\Omega)} ,$$

if $v \in W^{2,p}(\Omega)$.

□

Corollary 6.16 Finite element error estimate. *Let $u(\mathbf{x})$ be the solution of the model problem* (4.9) *with $u \in H^{m+1}(\Omega)$ and let $u^h(\mathbf{x})$ be the solution of the corresponding finite element problem. Consider a family of quasi-uniform triangulations and let the finite element spaces V^h contain polynomials of degree m. Then, the following finite element error estimate holds*

$$\left\| \nabla (u - u^h) \right\|_{L^2(\Omega)} \le c h^m \left\| D^{m+1} u \right\|_{L^2(\Omega)} = c h^m \left| u \right|_{H^{m+1}(\Omega)} . \qquad (6.10)$$

Proof: The statement follows by combining Lemma 4.13 (for $V = H_0^1(\Omega)$) and (6.9)

$$\left\| \nabla (u - u^h) \right\|_{L^2(\Omega)} \le \inf_{v^h \in V^h} \left\| \nabla (u - v^h) \right\|_{L^2(\Omega)} \le \left\| \nabla (u - I_h u) \right\|_{L^2(\Omega)} \le c h^m \left| u \right|_{H^{m+1}(\Omega)} .$$

■

Remark 6.17 *To* (6.10). Note that Lemma 4.13 provides only information about the error in the norm on the left-hand side of (6.10), but not in other norms. □

6.2 Inverse Estimate

Remark 6.18 *On inverse estimates.* The approach for proving interpolation error estimates can be uses also to prove so-called inverse estimates. In contrast to interpolation error estimates, a norm of a higher order derivative of a finite element function will be estimated by a norm of a lower order derivative of this function. One obtains as penalty a factor with negative powers of the diameter of the mesh cell. □

Theorem 6.19 Inverse estimate. *Let $0 \le k \le l$ be natural numbers and let $p, q \in [1, \infty]$. Then there is a constant C_{inv}, which depends only on $k, l, p, q, \hat{K}, \hat{P}(\hat{K})$ such that*

$$\left\| D^l v^h \right\|_{L^q(K)} \le C_{\text{inv}} h_K^{(k-l)-d(p^{-1}-q^{-1})} \left\| D^k v^h \right\|_{L^p(K)} \quad \forall\, v^h \in P(K). \tag{6.11}$$

Proof: In the first step, (6.11) is shown for $h_{\hat{K}} = 1$ and $k = 0$ on the reference mesh cell. Since all norms are equivalent in finite dimensional spaces, one obtains

$$\left\| D^l \hat{v}^h \right\|_{L^q(\hat{K})} \le \left\| \hat{v}^h \right\|_{W^{l,q}(\hat{K})} \le C \left\| \hat{v}^h \right\|_{L^p(\hat{K})} \quad \forall\, \hat{v}^h \in \hat{P}(\hat{K}).$$

If $k > 0$, then one sets

$$\tilde{P}(\hat{K}) = \left\{ \partial_{\boldsymbol{\alpha}} \hat{v}^h \ : \ \hat{v}^h \in \hat{P}(\hat{K}), |\boldsymbol{\alpha}| = k \right\},$$

which is also a space consisting of polynomials. The application of the first estimate of the proof to $\tilde{P}(\hat{K})$ gives

$$\begin{aligned} \left\| D^l \hat{v}^h \right\|_{L^q(\hat{K})} &= \sum_{|\boldsymbol{\alpha}|=k} \left\| D^{l-k} \left(\partial_{\boldsymbol{\alpha}} \hat{v}^h \right) \right\|_{L^q(\hat{K})} \le C \sum_{|\boldsymbol{\alpha}|=k} \left\| \partial_{\boldsymbol{\alpha}} \hat{v}^h \right\|_{L^p(\hat{K})} \\ &= C \left\| D^k \hat{v}^h \right\|_{L^p(\hat{K})}. \end{aligned}$$

This estimate is transformed to an arbitrary mesh cell K analogously as for the interpolation error estimates. From the estimates for the transformations, one obtains

$$\begin{aligned} \left\| D^l v^h \right\|_{L^q(K)} &\le C h_K^{-l+d/q} \left\| D^l \hat{v}^h \right\|_{L^q(\hat{K})} \le C h_K^{-l+d/q} \left\| D^k \hat{v}^h \right\|_{L^p(\hat{K})} \\ &\le C_{\text{inv}} h_K^{k-l+d/q-d/p} \left\| D^k v^h \right\|_{L^p(K)}. \end{aligned}$$

■

Remark 6.20 *On the proof.* The crucial point in the proof was the equivalence of all norms in finite dimensional spaces. Such a property does not exist in infinite dimensional spaces. □

Corollary 6.21 Global inverse estimate. *Let $p = q$ and let $\mathcal{T}^h$ be a regular triangulation of Ω, then*

$$\left\| D^l v^h \right\|_{L^{p,h}(\Omega)} \le C_{\text{inv}} h^{k-l} \left\| D^k v^h \right\|_{L^{p,h}(\Omega)},$$

where

$$\|\cdot\|_{L^{p,h}(\Omega)} = \left(\sum_{K \in \mathcal{T}^h} \|\cdot\|_{L^p(K)}^p \right)^{1/p}.$$

Remark 6.22 *On* $\|\cdot\|_{L^{p,h}(\Omega)}$. The cell wise definition of the norm is important for $l \ge 2$ since in this case finite element functions generally do not possess the regularity for the global norm to be well defined. It is also important for $l \ge 1$ and non-conforming finite element functions. □

6.3 Interpolation of Non-Smooth Functions

Remark 6.23 *Motivation.* The interpolation theory of Section 6.1 requires that the interpolation operator is continuous on the Sobolev space to which the function belongs that should be interpolated. But if one, e.g., wants to interpolate discontinuous functions with continuous, piecewise linear elements, then Section 6.1 does not provide estimates.

A simple remedy seems to be first to apply some smoothing operator to the function to be interpolated and then to interpolate the smoothed function. However, this approach leads to difficulties at the boundary of Ω and it will not be considered further.

There are two often used interpolation operators for non-smooth functions. The interpolation operator of Clément (1975) is defined for functions from $L^1(\Omega)$ and it can be generalized to more or less all finite elements. The interpolation operator of Scott and Zhang (1990) is more special. It has the advantage that it preserves homogeneous Dirichlet boundary conditions. Here, only the interpolation operator of Clément, for linear finite elements, will be considered.

Let $\mathcal{T}^h$ be a regular triangulation of the polyhedral domain $\Omega \subset \mathbb{R}^d, d \in \{2,3\}$, with simplicies K. Denote by P_1 the space of continuous, piecewise linear finite elements on $\mathcal{T}^h$. □

Remark 6.24 *Construction of the interpolation Operator of Clément.* For each vertex V_i of the triangulation, the union of all grid cells which possess V_i as vertex will be denoted by ω_i, see Figure 5.1.

The interpolation operator of Clément is defined with the help of local $L^2(\omega_i)$ projections. Let $v \in L^1(\Omega)$ and let $P_1(\omega_i)$ be the space of continuous piecewise linear finite elements on ω_i. Then, the local $L^2(\omega_i)$ projection of $v \in L^1(\omega_i)$ is the solution $p_i \in P_1(\omega_i)$ of

$$\int_{\omega_i} (v - p_i)(\mathbf{x})q(\mathbf{x})\ d\mathbf{x} = 0 \quad \forall\, q \in P_1(\omega_i) \tag{6.12}$$

or equivalently of

$$(v - p_i, q)_{L^2(\omega_i)} = 0 \quad \forall\, q \in P_1(\omega_i).$$

Then, the Clément interpolation operator is defined by

$$P^h_{\text{Cle}} v(\mathbf{x}) = \sum_{i=1}^{N} p_i(V_i)\phi_i^h(\mathbf{x}), \tag{6.13}$$

where $\{\phi_i^h\}_{i=1}^N$ is the standard basis of the global finite element space P_1. Since $P^h_{\text{Cle}}v(\mathbf{x})$ is a linear combination of basis functions of P_1, it defines a map $P^h_{\text{Cle}} : L^1(\Omega) \to P_1$. □

Theorem 6.25 Interpolation estimate. *Let $k,l \in \mathbb{N} \cup \{0\}$ and $q \in \mathbb{R}$ with $k \le l \le 2$, $1 \le q \le \infty$ and let ω_K be the union of all subdomains ω_i that contain the mesh cell K, see Figure 6.2. Then it holds for all $v \in W^{l,q}(\omega_K)$ the estimate*

$$\left\|D^k(v - P^h_{\text{Cle}}v)\right\|_{L^q(K)} \le Ch^{l-k}\left\|D^l v\right\|_{L^q(\omega_K)}, \tag{6.14}$$

with $h = diam(\omega_K)$, where the constant C is independent of $v(\mathbf{x})$ and h.

Proof: The statement of the lemma is obvious in the case $k = l = 2$ since it is $D^2 P^h_{\text{Cle}} v(\mathbf{x})|_K = 0$.

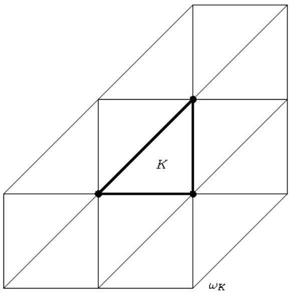

Figure 6.2: A subdomain ω_K.

Let $k \in \{0,1\}$. Because the $L^2(\Omega)$ projection gives an element with best approximation, one gets with (6.12)

$$P^h_{\mathrm{Cle}} p = p \quad \text{in } K \quad \forall\, p \in P_1(\omega_K). \tag{6.15}$$

One says that P^h_{Cle} is a consistent operator.

The next step consists in proving the stability of P^h_{Cle}. One obtains with the inverse inequality (6.11)

$$\|p\|_{L^\infty(\omega_i)} \le ch^{-d/2} \|p\|_{L^2(\omega_i)} \qquad \text{for all } p \in P_1(\omega_i).$$

The inverse inequality and definition (6.12) of the local L^2 projection with the test function $q = p_i$ gives

$$\|p_i\|^2_{L^\infty(\omega_i)} \le ch^{-d} \|p_i\|^2_{L^2(\omega_i)} \le ch^{-d} \|v\|_{L^1(\omega_i)} \|p_i\|_{L^\infty(\omega_i)}.$$

Dividing by $\|p_i\|_{L^\infty(\omega_i)}$ and applying Hölder's inequality, one obtains for $p^{-1} = 1 - q^{-1}$ (*exercise*)

$$|p_i(V_i)| \quad \le \quad ch^{-d/q} \|v\|_{L^q(\omega_i)} \tag{6.16}$$

for all $V_i \in K$. From the regularity of the triangulation, it follows for the basis functions that (inverse estimate)

$$\left\| D^k \phi_i \right\|_{L^\infty(K)} \le ch^{-k}, \quad k = 0, 1. \tag{6.17}$$

Using the triangle inequality, combining (6.16) and (6.17) yields the stability of P^h_{Cle}

$$\begin{aligned} \left\| D^k P^h_{\mathrm{Cle}} v \right\|_{L^q(K)} &\le \sum_{V_i \in K} |p_i(V_i)| \left\| D^k \phi_i \right\|_{L^q(K)} \\ &\le c \sum_{V_i \in K} h^{-d/q} \|v\|_{L^q(\omega_i)} \left\| D^k \phi_i \right\|_{L^\infty(K)} \|1\|_{L^q(K)} \\ &\le c \sum_{V_i \in K} h^{-d/q} \|v\|_{L^q(\omega_i)} h^{-k} h^{d/q} \\ &= ch^{-k} \|v\|_{L^q(\omega_K)}. \end{aligned} \tag{6.18}$$

The remainder of the proof follows the proof of the interpolation error estimate for the polynomial interpolation, Theorem 6.10, apart from the fact that a reference cell is not used for the Clément interpolation operator. Using Lemma 6.2 and 6.4, one can find a polynomial $p \in P_1(\omega_K)$ with

$$\left\| D^j (v - p) \right\|_{L^q(\omega_K)} \le ch^{l-j} \left\| D^l v \right\|_{L^q(\omega_K)}, \quad 0 \le j \le l \le 2. \tag{6.19}$$

With (6.15), the triangle inequality, $\|\cdot\|_{L^q(K)} \le \|\cdot\|_{L^q(\omega_K)}$, (6.18), and (6.19), one obtains

$$\begin{aligned}
\left\| D^k \left(v - P_{\mathrm{Cle}}^h v \right) \right\|_{L^q(K)} &= \left\| D^k \left(v - p + P_{\mathrm{Cle}}^h p - P_{\mathrm{Cle}}^h v \right) \right\|_{L^q(K)} \\
&\le \left\| D^k (v-p) \right\|_{L^q(K)} + \left\| D^k P_{\mathrm{Cle}}^h (v-p) \right\|_{L^q(K)} \\
&\le \left\| D^k (v-p) \right\|_{L^q(\omega_K)} + ch^{-k} \left\| v-p \right\|_{L^q(\omega_K)} \\
&\le ch^{l-k} \left\| D^l v \right\|_{L^q(\omega_K)} + ch^{-k} h^l \left\| D^l v \right\|_{L^q(\omega_K)} \\
&= ch^{l-k} \left\| D^l v \right\|_{L^q(\omega_K)} .
\end{aligned}$$

■

Remark 6.26 *Uniform meshes.*

- If all mesh cells in ω_K are of the same size, then one can replace h by h_K in the interpolation error estimate (6.14). This property is given in many cases.
- If one assumes that the number of mesh cells in ω_K is bounded uniformly for all considered triangulations, the global interpolation estimate

  $$\left\| D^k (v - P_{\mathrm{Cle}}^h v) \right\|_{L^q(\Omega)} \le Ch^{l-k} \left\| D^l v \right\|_{L^q(\Omega)}, \quad 0 \le k \le l \le 2,$$

  follows directly from (6.14).

□