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 $c_{\boldsymbol{\beta}} \ne 0$ with maximal value $|\boldsymbol{\beta}|$. Then, it is $\partial_{\boldsymbol{\beta}} q(\mathbf{x}) = C c_{\boldsymbol{\beta}} = \mathit{const} \ne 0$, 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

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

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

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

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

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

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 $\hat{\Phi}(1) = 1 \ne 0$. 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

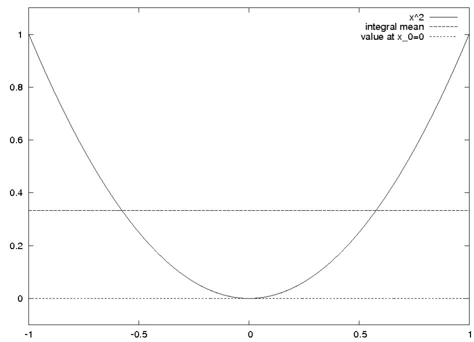

Figure 6.1: Interpolation of x^2 in $[-1,1]$ by a P_0 function with the integral mean value and with the value of the function at $x_0 = 0$.

- 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 $(m+1-s)p \le d$, 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

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

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*

$$\|B\| \le c h_K, \quad \left\|B^{-1}\right\| \le c h_K^{-1},$$

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 $\hat{\mathbf{y}}$

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

Now, it holds for the spectral norm that

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

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*

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

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

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

■

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*

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

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

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

■

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*

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

where

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

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 $\mathcal{T}^h$. $\square$

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

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

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

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*

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

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 $D^2 P^h_{\mathrm{Cle}} v(\mathbf{x})|_K = 0$.

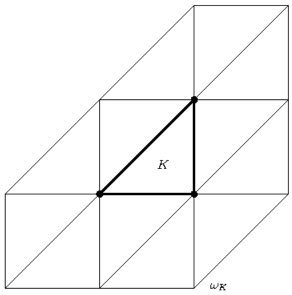

Figure 6.2: A subdomain ω_K.

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

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

■

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

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

  follows directly from (6.14).

□