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 (4.20).

This section studies the approximation quality of finite element spaces. Estimates are proved for interpolants of functions. Interpolation estimates are of course upper bounds of 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 prescribed. Then, there is a uniquely determined polynomial $p \in P_m(\Omega)$ such that*

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

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

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

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

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

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

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

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

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

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

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 $v(\boldsymbol{x})$.

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 $\partial_{\boldsymbol{\alpha}} v(\boldsymbol{x})$.

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

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

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

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

It follows from the assumption that the second integral on the left-hand side vanishes that

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

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 (3.3), and the estimate $\|\boldsymbol{x} - \boldsymbol{y}\|_2 \le 2R$ yields

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

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

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

Applying the theorem of Fubini allows the commutation of the integration

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

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

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

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

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

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

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

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

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

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

Remark 6.5. On Lemma 6.4. 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 $v(\boldsymbol{x})$ 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[1]–Hilbert[2] lemma. The Bramble–Hilbert lemma considers a continuous linear functional that is defined on a Sobolev space and that vanishes for all polynomials of degree less than or equal to 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 $m \in \mathbb{N} \cup \{0\}$, $p \in [1, \infty]$, and $F : W^{m+1,p}(\Omega) \to \mathbb{R}$ be a continuous linear functional, and let the conditions of Lemma 6.2 and Lemma 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 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

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

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, more precisely Lemma 6.4, will be used for estimating the interpolation error for 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. □

---

[1] James H. Bramble, born 1930

[2] Stephen R. Hilbert

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

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

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

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

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

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

  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

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

  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
- It is possible to define $\hat{\Phi}(\hat{v}) = \hat{v}(\hat{\boldsymbol{x}}_0)$ for an arbitrary point $\hat{\boldsymbol{x}}_0 \in \hat{K}$. 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 $\hat{\boldsymbol{x}}_0$, see also Figure 6.1.

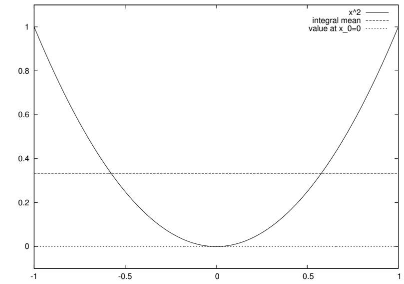

Fig. 6.1 Interpolation of x^2 in $[-1, 1]$ by a P_0 function with the integral mean value and with the value of the function at $x_0 = 0$.

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 $P_m(\hat{K}) \subset \hat{P}(\hat{K})$, let $p \in [1, \infty)$, and let $\hat{s} \in \mathbb{N} \cup \{0\}$ such that $(m + 1 - \hat{s})p > d \ge (m - \hat{s})p$ and $\hat{s} \ge s$, where s appears in the definition of the interpolation operator. Then there is a constant C that is independent of $\hat{v}(\hat{\boldsymbol{x}})$ 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. Since $\hat{K}$ is bounded, one has the Sobolev imbedding, Theorem 3.52,

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

Because $\hat{K}$ is convex, the imbedding $C^{\hat{s}}(\hat{K}) \to C^{s}(\hat{K})$ is compact, see (Adams, 1975, Theorem 1.31), such that 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 $\hat{q} \in P_m(\hat{K})$

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

Now, $\hat{q}(\hat{\boldsymbol{x}})$ is chosen such that

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

holds. Hence, the assumptions of Lemma 6.4 are satisfied. It follows that

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

■

Definition 6.11. Quasi-uniform and regular family of triangulations, (Brenner & Scott, 2008, Def. 4.4.13). Let $\{\mathcal{T}^h\}$ with $0 < h \le 1$, be a family of triangulations such that

$$\max_{K \in \mathcal{T}^h} h_K \le h \operatorname{diam}(\Omega),$$

where h_K is the diameter of $K = F_K(\hat{K})$, i.e., the largest distance of two points that are contained in K. The family is called to be quasi-uniform, if there exists a $C > 0$ such that

$$\min_{K \in \mathcal{T}^h} \rho_K \ge Ch \operatorname{diam}(\Omega) \tag{6.5}$$

for all $h \in (0, 1]$, where ρ_K is the diameter of the largest ball contained in K.

The family is called to be regular, if there is exists a $C > 0$ such that for all $K \in \mathcal{T}^h$ and for all $h \in (0, 1]$

$$\rho_K \ge Ch_K.$$

□

Remark 6.12. Assumptions on the reference mapping and 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 mapping from K to $\hat{K}$ and of the inverse mapping. Here, only the case of an affine family of finite elements whose mesh cells are generated by affine mappings

$$F_K \hat{\boldsymbol{x}} = B_K \hat{\boldsymbol{x}} + \boldsymbol{b},$$

will be considered, see (5.3), where B_K is a non-singular $d \times d$ matrix and $\boldsymbol{b}$ is a d vector. For the global estimate, a quasi-uniform family of triangulations will be considered. □

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

$$\|B_K\| \le Ch_K, \quad \left\| B_K^{-1} \right\| \le Ch_K^{-1}, \tag{6.6}$$

where the constants depend on the matrix norm.

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

$$\boldsymbol{x}_0 = B_K \hat{\boldsymbol{x}}_0 + \boldsymbol{b}, \quad \boldsymbol{x} = B_K(\hat{\boldsymbol{x}}_0 + \hat{\boldsymbol{y}}) + \boldsymbol{b} = \boldsymbol{x}_0 + B_K \hat{\boldsymbol{y}}$$

are contained in K. Hence, one obtains for all $\hat{\boldsymbol{y}}$

$$\|B_K\hat{\boldsymbol{y}}\|_2 = \|\boldsymbol{x} - \boldsymbol{x}_0\|_2 \le h_K.$$

Now, it holds for the spectral norm that

$$\|B_K\|_2 = \sup_{\|\hat{\boldsymbol{z}}\|_2=1} \|B_K\hat{\boldsymbol{z}}\|_2 = \frac{1}{r}\sup_{\|\hat{\boldsymbol{z}}\|_2=r} \|B_K\hat{\boldsymbol{z}}\|_2 \le \frac{h_K}{r}.$$

A bound of this form, with a possible different constant, holds also for all other matrix norms since all matrix norms are equivalent, see Remark 3.34.

The estimate for $\left\|B_K^{-1}\right\|$ 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 $v \in W^{m+1,p}(K)$, $p \in [1,\infty)$, there is a constant C, which is independent of v, such that*

$$\left\|D^k(v - I_K v)\right\|_{L^p(K)} \le C h_K^{m+1-k} \left\|D^{m+1}v\right\|_{L^p(K)}, \quad 0 \le k \le m+1. \tag{6.7}$$

Proof. The idea of the proof consists in transforming the left-hand side of (6.7) to the reference cell, using the interpolation estimate on the reference cell, and transforming back.

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

$$|b_{ij}| \le Ch_K, \quad \left|b_{ij}^{(-1)}\right| \le Ch_K^{-1}. \tag{6.8}$$

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

$$|\det B_K| \le Ch_K^d, \quad \left|\det B_K^{-1}\right| \le Ch_K^{-d}. \tag{6.9}$$

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

$$I_K v = \sum_{i=1}^N \Phi_{K,i}(v)\phi_{K,i}, \tag{6.10}$$

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

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

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

$$\Phi_{K,i}(v) = \hat{\Phi}_i(v \circ F_K) = \hat{\Phi}_i\left(\hat{v}\right). \tag{6.11}$$

Hence, it follows for $v = \hat{\phi}_j \circ F_K^{-1}$ 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 (6.11) and (6.10), one gets

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

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

iii). One obtains with the chain rule

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

It follows with (6.8) that (with each derivative one obtains an additional factor of B_K or B_K^{-1}, respectively)

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

One gets with (6.9)

$$
\int_K \left\|D_{\boldsymbol{x}}^k v(\boldsymbol{x})\right\|_2^p \, d\boldsymbol{x} \le C h_K^{-kp} \left|\det B_K\right| \int_{\hat{K}} \left\|D_{\hat{\boldsymbol{x}}}^k \hat{v}(\hat{\boldsymbol{x}})\right\|_2^p \, d\hat{\boldsymbol{x}} \le C h_K^{-kp+d} \int_{\hat{K}} \left\|D_{\hat{\boldsymbol{x}}}^k \hat{v}(\hat{\boldsymbol{x}})\right\|_2^p \, d\hat{\boldsymbol{x}} \tag{6.12}
$$

and

$$
\int_{\hat{K}} \left\|D_{\hat{\boldsymbol{x}}}^k \hat{v}(\hat{\boldsymbol{x}})\right\|_2^p \, d\hat{\boldsymbol{x}} \le C h_K^{kp} \left|\det B_K^{-1}\right| \int_K \left\|D_{\boldsymbol{x}}^k v(\boldsymbol{x})\right\|_2^p \, d\boldsymbol{x} \le C h_K^{kp-d} \int_K \left\|D_{\boldsymbol{x}}^k v(\boldsymbol{x})\right\|_2^p \, d\boldsymbol{x}. \tag{6.13}
$$

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

$$
\left\|D_{\hat{\boldsymbol{x}}}^k (\hat{v} - I_{\hat{K}}\hat{v})\right\|_{L^p(\hat{K})}^p \le C \left\|D_{\hat{\boldsymbol{x}}}^{m+1} \hat{v}\right\|_{L^p(\hat{K})}^p, \quad 0 \le k \le m+1. \tag{6.14}
$$

It follows with (6.12), (6.14), and (6.13) that

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

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

Remark 6.15. On estimate (6.7).

- 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

$$\left\| D^k(v - I^h v) \right\|_{L^p(\Omega)} = \left(\sum_{K \in \mathcal{T}^h} \left\| D^k(v - I_K v) \right\|_{L^p(K)}^p \right)^{1/p}$$
$$\leq \left(\sum_{K \in \mathcal{T}^h} C h_K^{(m+1-k)p} \left\| D^{m+1} v \right\|_{L^p(K)}^p \right)^{1/p}$$
$$\leq C h^{(m+1-k)} \left\| D^{m+1} v \right\|_{L^p(\Omega)} . \qquad (6.15)$$

□

Corollary 6.16. Finite element error estimate. *Let $u(\boldsymbol{x})$ be the solution of the model problem* (4.10) *with $u \in H^{m+1}(\Omega)$ and let $u^h(\boldsymbol{x})$ 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*

$$\left\| \nabla(u - u^h) \right\|_{L^2(\Omega)} \leq C h^m \left\| D^{m+1} u \right\|_{L^2(\Omega)} = C h^m \left| u \right|_{H^{m+1}(\Omega)} . \qquad (6.16)$$

Proof. The statement follows by combining Lemma 4.13 (for $V = H_0^1(\Omega)$) and (6.15)

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

■

Remark 6.17. To (6.16). Note that Lemma 4.13 provides only information about the error in the norm on the left-hand side of (6.16), but not in other norms. □

6.2 Inverse Estimate

Remark 6.18. On inverse estimates. The approach for proving interpolation error estimates can be used also to prove so-called inverse estimates. With inverse estimates, a norm of a higher order derivative of a finite element function is estimated by a norm of a lower order derivative of this function. Likewise, norms in different Lebesgue spaces are estimated. One obtains as penalty a factor with negative powers of the diameter of the mesh cell. □

Theorem 6.19. Inverse estimate. *Let $0 \leq k \leq l$ 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*

$$\left\| D^l v^h \right\|_{L^q(K)} \leq C_{\text{inv}} h_K^{(k-l)-d(p^{-1}-q^{-1})} \left\| D^k v^h \right\|_{L^p(K)} \quad \forall\, v^h \in P(K). \quad (6.17)$$

Proof. In the first step, (6.17) 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

$$\left\|D^l \hat{v}^h\right\|_{L^q(\hat{K})} \le \left\|\hat{v}^h\right\|_{W^{l,q}(\hat{K})} \le C \left\|\hat{v}^h\right\|_{L^p(\hat{K})} \quad \forall\, \hat{v}^h \in \hat{P}(\hat{K}). \tag{6.18}$$

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 (6.18) 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, compare the proof of Theorem 6.14. From the estimates (6.12) and (6.13) 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 is the equivalence of all norms in finite-dimensional spaces. Such a property does not hold in infinite-dimensional spaces. □

Corollary 6.21. Global inverse estimate. *Let $p = q$ and let $\left\{\mathcal{T}^h\right\}$ be a quasi-uniform family of triangulations of Ω, then*

$$\left\|D^l v^h\right\|_{L^{p,h}(\Omega)} \le C_{\text{inv}} h^{k-l} \left\|D^k v^h\right\|_{L^{p,h}(\Omega)}, \tag{6.19}$$

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 $k \ge 2$ or $l \ge 2$ since in these cases 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. □