ASC Report No. 39/2015

Interpolation and quasi-interpolation in h- and hp-version finite element spaces (extended version)

T. Apel, J.M. Melenk

Institute for Analysis and Scientific Computing — Vienna University of Technology — TU Wien
www.asc.tuwien.ac.at ISBN 978-3-902627-05-6

Most recent ASC Reports

38/2015 *H. Woraceki*
xxxx

37/2015 *A. Jüngel and N. Zamponi*
Quality behavior of solutions to cross-diffusion systems from population dynamics

36/2015 *A. Jüngel and W. Yue*
Discrete Beckner inequalities via the Bochner-Bakry-Emery approach for Markov chains

35/2015 *J. Burkotová, I. Rachunková, M. Hubner, E. Weinmüller*
Numerical evidence of Kneser solutions to a class of singular BVPs in ODEs.

34/2015 *O. Koch, S. Schirnhofer, E. Weinmüller*
Numerical simulation of the Korteweg-de Vries Equation for shallow water waves.

33/2015 *S. Börm and J.M. Melenk*
Approximation of the high-frequency Helmholtz kernel by nested directional interpolation

32/2015 *F. Achleitner, A. Arnold, E.A. Carlen*
On hypocoercive BGK models

31/2015 *M. Feischl, T. Führer, M. Niederer, S. Strommer, A. Steinboeck, and D. Praetorius*
Efficient numerical computation of direct exchange areas in thermal radiation analysis

30/2015 *B. Düring, P. Fuchs, and A. Jüngel*
A higher-order gradient flow scheme for a singular one-dimensional diffusion equation

29/2015 *C. Erath and D. Praetorius*
Adaptive finite volume methods with convergence rates

Institute for Analysis and Scientific Computing
Vienna University of Technology
Wiedner Hauptstraße 8–10
1040 Wien, Austria

E-Mail: admin@asc.tuwien.ac.at
WWW: http://www.asc.tuwien.ac.at
FAX: +43-1-58801-10196

ISBN 978-3-902627-05-6

Interpolation and quasi-interpolation in h- and hp-version finite element spaces (extended version†)

Thomas Apel[1] and Jens M. Melenk[2]

[1] *Universität der Bundeswehr München, Neubiberg, Germany*
[2] *Technische Universität Wien, Wien, Austria*

ABSTRACT

Interpolation operators map a function u to an element $\mathrm{I}u$ of a finite element space. Unlike more general approximation operators, interpolants are defined locally. Estimates of the *interpolation error*, i.e., the difference $u - \mathrm{I}u$, are of utmost importance in numerical analysis. These estimates depend on the size of the finite elements, the polynomial degree employed, and the regularity of u. In contrast to interpolation the term quasi-interpolation is used when the regularity is so low that interpolation has to be combined with regularization. This paper gives an overview of different interpolation operators and their error estimates. The discussion includes the h-version and the hp-version of the finite element method, interpolation on the basis of triangular/tetrahedral and quadrilateral/hexahedral meshes, affine and nonaffine elements, isotropic and anisotropic elements, and Lagrangian and other elements.

approximation, polynomial interpolation, nodal interpolation, quasi-interpolation, Clément interpolation, Scott-Zhang interpolation, isotropic finite element, shape-regular element, anisotropic element, h-version, p-version, hp-version

1. Introduction

1.1. Overview of this Chapter

The question of how well a given function u can be approximated from a finite-dimensional space is of fundamental importance for the understanding of most numerical methods for partial differential equations. Mostly, this basic question takes the form of approximating u from a space of piecewise polynomials, i.e., the approximating functions are defined on some *mesh* and they are polynomial on each element K of the mesh. The key parameters that control the approximation properties of such spaces are the mesh size h, i.e., a quantity that controls the size of the elements of the mesh, and the polynomial degree k. Increasing the accuracy can be achieved in several ways: first, by decreasing the mesh size, which we call the h-version; second, by increasing the approximation order, which we refer to as the p-version; third, by simultaneously varying h and k, which goes by the name of hp-version.

---

†This text is an extended version of the contribution *Interpolation and quasi-interpolation in h- and hp-version finite element spaces* to the *Encyclopedia of Computational Mechanics (second ed.)*, edited by E. Stein, R. de Borst, T. Hughes.

In this chapter, we present several locally defined approximation operators I. To be defined locally means that, on an element K, the approximation $(\mathrm{I}u)|_K$ is determined by $u|_K$ or $u|_{\omega_K}$, where ω_K is a union of elements adjacent to K. We call I an *interpolation* operator if $(\mathrm{I}u)|_K$ depends solely on $u|_K$. If $(\mathrm{I}u)|_K$ depends on the neighborhood ω_K, we refer to I as a *quasi-interpolation* operator. Possibly the simplest representative of an interpolation operator is the classical piecewise linear interpolation, which requires continuity of u to be well-defined. An example of a quasi-interpolation is one where the node values are obtained by averaging on elements close to that node; then, u merely needs to be in L^1. Both interpolation and quasi-interpolation operators have their established place in numerical analysis. While interpolation operators have tighter locality properties and are often simpler to use, the regularity requirements can be fairly stringent. In contrast, quasi-interpolation operators require minimal regularity. As an example for the need of quasi-interpolation operators we mention the analysis of residual error estimation in FEM. The approximation properties of such approximation operators depend crucially on the (local) regularity of u. In the present chapter, we capture this regularity in terms of membership in Sobolev spaces.

Let us highlight some key features of locally defined approximation operators. In the h-version, i.e., for fixed k, the error $(u - \mathrm{I}u)|_K$ is of the form $O(h^\alpha)$, where α is controlled in terms of *both* k and the Sobolev regularity of u on or near K. Even if u is smooth (C^∞), only an algebraic convergence (as measured in error versus number of degrees of freedom) can be achieved. In this respect, the p-version is different since for smooth (more precisely: analytic) functions u, exponential convergence is possible when $k \to \infty$. The locality of the operators I, i.e., the fact that the size of the error $(u - \mathrm{I}u)|_K$ can be assessed in terms of the regularity of u near K, is a strong asset. For example, it allows one to assess the approximation properties of spaces that are based on locally refined (*graded*, *adaptive*) meshes; we illustrate this feature for the h-version in Section 2.5 and for the hp-version in Section 3.1.3.

The topic of piecewise polynomial approximation is vast, and a selection has to be made. In the context of the h-version, we focus on local error estimates for both isotropic and anisotropic elements. For the p-version, we concentrate on the single-element case and emphasize the ability to achieve exponential convergence. We refer the reader to the chapter **Finite Element Methods** of **ECM2**[‡], and the chapter **The *p*-Version of the Finite Element Method** of **ECM2**, where these features are discussed in the context of the FEM. Several important aspects of polynomial approximation are not covered in this chapter. We mention in particular certain structure-preserving approximation operators with the *commuting diagram property*, which are discussed in the chapter **Mixed Finite Element Methods** of **ECM2**, and the chapter **Finite Element Methods for Maxwell Equations** of **ECM2**. Approximation theoretic questions that are closely related to those of the present chapter have to be addressed in the Virtual Element Method (VEM, see Beirão da Veiga, Brezzi, Cangiani, Manzini, Marini, and Russo (2013); Beirão da Veiga, Brezzi, Marini, and Russo (2014)) and the Isogeometric Analysis (IGA, see Cottrell, Hughes, and Bazilevs (2009), and the chapter **Mathematics of Isogeometric Analysis** of **ECM2**). Also not covered in this chapter are the approximation theoretic questions that arise in meshless methods and generalized FEM; see the chapter **Meshfree Methods** of **ECM2** and the chapter **Extended finite element methods** of **ECM2** and (Melenk, 2005b) for details.

The plan of this Chapter is to continue this introductory section with a description of the notion of finite elements in Subsection 1.2, with a sketch of key arguments of this chapter in Subsection 1.3, and with a discussion of various conditions on element shapes in Subsection 1.4. In two separate sections we then treat the h-version interpolation, see Section 2, and the hp-version interpolation, see Section 3.

Let $d = 2, 3$ be the space dimension and $x = (x_1, \ldots, x_d)$ a Cartesian coordinate system. We use standard multi-index notation with $\alpha := (\alpha_1, \ldots, \alpha_d)$, where the entries α_i are from the set $\mathbb{N}_0$ of

[‡] **ECM2**=Encyclopedia of Computational Mechanics (second ed.). Edited by E. Stein, R. de Borst, T. Hughes

nonnegative integers, and

$$x^\alpha := \prod_{i=1}^d x_i^{\alpha_i}, \quad \alpha! := \prod_{i=1}^d \alpha_i!, \quad |\alpha| := \sum_{i=1}^d \alpha_i,$$
$$D^\alpha := \frac{\partial^{\alpha_1}}{\partial x_1^{\alpha_1}} \cdots \frac{\partial^{\alpha_d}}{\partial x_d^{\alpha_d}}.$$

We denote by $\mathbb{P}_k$ the space of polynomials of degree $k \in \mathbb{N}_0$, i.e., we set $\mathbb{P}_k := \operatorname{span}\{x^\alpha \colon |\alpha| \le k\}$. The "tensor product space" $\mathbb{Q}_k$ is given by $\mathbb{Q}_k := \operatorname{span}\{x^\alpha \colon \alpha_i \le k, i = 1, \ldots, d\}$.

The notation $W^{\ell,p}(G)$, $\ell \in \mathbb{N}_0$, $p \in [1, \infty]$, is used for the classical Sobolev spaces with the norm and seminorm

$$\|v\|^p_{W^{\ell,p}(G)} := \sum_{|\alpha| \le \ell} \int_G |D^\alpha v|^p,$$
$$|v|^p_{W^{\ell,p}(G)} := \sum_{|\alpha| = \ell} \int_G |D^\alpha v|^p$$

for $p < \infty$ and the usual modification for $p = \infty$. In general, we will write $L^p(G)$ for $W^{0,p}(G)$ and $H^s(G)$ for $W^{s,2}(G)$. The symbol C is used for a generic positive constant, which may be different at each occurrence.

1.2. Finite Elements

In this subsection, we briefly introduce the notion of finite elements. While the chapter **Finite Element Methods** of **ECM2** presents this topic comprehensively, we focus on those pieces that are necessary for the the current chapter.

Finite element spaces over a domain $\Omega \subset \mathbb{R}^d$ are defined on the basis of a *finite element mesh* or *triangulation* $\mathcal{T}$ of Ω. This is a subdivision of $\overline{\Omega}$ into a finite number of closed bounded subsets K with nonempty interior and piecewise smooth boundary in such a way that the following properties are satisfied:

- The closure of the domain Ω is equal to the union of all the finite elements, $\overline{\Omega} = \bigcup_{K \in \mathcal{T}} K$.
- Distinct $K_1, K_2 \in \mathcal{T}$ have no common interior points.
- Any face of an element $K_1 \in \mathcal{T}$ is either a subset of the boundary $\partial\Omega$ or a face of another element $K_2 \in \mathcal{T}$. In particular, hanging nodes are not admitted.

The elements are usually triangles or quadrilaterals in 2D or tetrahedra or hexahedra in 3D but further shapes are possible. In fact, elements of different shape may be used side by side in a mesh. For reasons of implementation and analysis, elements are typically the image of a *reference* configuration under an *element map*. These observations lead us to the following assumption:

Assumption 1.1. *There is a fixed finite set* $\mathcal{R}$ *of* reference elements. *For each* $K \in \mathcal{T}$*, there is a reference element* $\widehat{K} \in \mathcal{R}$ *and a bijection, the* element map,

$$F_K \colon \widehat{x} \in \mathbb{R}^d \mapsto x = F_K(\widehat{x}) \in \mathbb{R}^d$$

with $K = F_K(\widehat{K})$.

The standard example is the affine element mapping F_K

$$F_K(\widehat{x}) = A\widehat{x} + a \quad \text{with } A \in \mathbb{R}^{d\times d},\ a \in \mathbb{R}^d; \tag{1.1}$$

of course, this puts restrictions on the number of possible shapes of an element. More general mappings are discussed in Subsection 1.4. The standard examples of $\mathcal{R}$ in the case $d = 2$ is the set $\mathcal{R} = \{T, S\}$ with the reference triangle T and the reference square S. For $d = 3$, $\mathcal{R}$, may consist of a tetrahedron, a hexahedron, a wedge, and a pyramid.

Let $\mathcal{P}_{\widehat{K}}$ be a finite-dimensional linear space of functions defined on the reference element $\widehat{K}$. On each element $K \in \mathcal{T}$ we define the space $\mathcal{P}_K$ via

$$u \in \mathcal{P}_K \quad \text{iff} \quad \widehat{u} := u \circ F_K \in \mathcal{P}_{\widehat{K}}. \tag{1.2}$$

In this chapter we will concentrate on spaces $\mathcal{P}_K$ of scalar functions; vector valued functions as they occur in Nédélec or Raviart–Thomas elements are not considered. The simplest example for triangular or tetrahedral elements K is that $\mathcal{P}_{\widehat{K}}$ is the space of polynomials of degree one. By an affine mapping F_K one obtains also for $\mathcal{P}_K$ polynomials of degree one. Further examples of finite elements are discussed in the chapter **Finite Element Methods** of **ECM2** (Examples 4-9 and 16-17).

The weak solution of elliptic boundary value problems of order $2m$ is generally sought in a subspace V of the Sobolev space $H^m(\Omega)$. The corresponding H^m-conforming finite element space is defined by

$$\mathrm{FE}_{\mathcal{T}} := \{v \in H^m(\Omega) : v_K := v|_K \in \mathcal{P}_K\ \forall K \in \mathcal{T}\}. \tag{1.3}$$

The dimension of $\mathrm{FE}_{\mathcal{T}}$ is denoted by N_+. Note that we usually consider a family of finite element spaces with $N_+ \to \infty$ but the finite set $\mathcal{R}$ of reference elements remains the same for all members of this family.

In the finite element method (FEM), the H^m-conforming space $\mathrm{FE}_{\mathcal{T}}$ is typically modified in several ways. First, essential boundary conditions have to be accounted for. Imposing for simplicity homogeneous essential boundary conditions we obtain the space $V \subset H^m(\Omega)$ and the N-dimensional finite element space $V_{\mathcal{T}}$. The method is called *conforming* if $V_{\mathcal{T}} \subset V$, i.e., when $V_{\mathcal{T}} = V \cap \mathrm{FE}_{\mathcal{T}}$. Second, *non-conforming* finite element spaces $V_{\mathcal{T}} \not\subset V$ are often employed.

A unified framework that includes many conforming and non-conforming finite elements was introduced by Ciarlet (1978): A finite element is taken to be the triple $(K, \mathcal{P}_K, \mathcal{N}_K)$ with $K \in \mathcal{T}$ and $\mathcal{P}_K$ as above and with $\mathcal{N}_K = \{\mathrm{N}_{i,K}\}_{i=1}^n$ being a *basis* of the dual space $\mathcal{P}_K'$. The linear functionals in $\mathcal{N}_K$ are sometimes called *nodal variables*, although $\mathrm{N} \in \mathcal{N}_K$ is not necessarily a function evaluation in nodes. The approach of Ciarlet (1978) leads in a natural way to an interpolation operator, the *nodal interpolation operator* studied in Subsection 2.1.1.

1.3. Arguments in a Nutshell

The aim of this subsection is to demonstrate which ingredients are needed for a simple proof of a local interpolation error estimate. The details are then worked out in Sections 2 and 3.

An interpolation operator $\mathrm{I}_{\mathcal{T}} : W^{\ell,p}(\Omega) \to \mathrm{FE}_{\mathcal{T}}$ is defined elementwise

$$(\mathrm{I}_{\mathcal{T}} u)|_K = \mathrm{I}_K u \quad \forall K \in \mathcal{T}, \tag{1.4}$$

where the element interpolation operator $\mathrm{I}_K : W^{\ell,p}(K) \to \mathcal{P}_K$ is defined via

$$\mathrm{I}_K u = \sum_{i=1}^n \mathrm{N}_{i,K}(u)\phi_{i,K}$$

by using a basis $\{\phi_{i,K}\}_{i=1}^n$ of $\mathcal{P}_K$ and a set of unisolvent linear functionals $\{\mathrm{N}_{i,K}\}_{i=1}^n$. Unisolvent means the following implication: if the values $\mathrm{N}_{i,K}(w)$, $i = 1\ldots,n$, vanish for function $w \in \mathcal{P}_K$ then $w \equiv 0$.

In order to obtain constants independent of the shape and size of the element K, the core estimate is proved on the reference element $\widehat{K}$ from Assumption 1.1 and transformed from $\widehat{K}$ to K by the mapping F_K. To do this one needs the relationship

$$(\mathrm{I}_K u) \circ F_K = \mathrm{I}_{\widehat{K}}(u \circ F_K) \quad \forall u \in W^{\ell,p}(K)$$

with an interpolation operator $\mathrm{I}_{\widehat{K}} \colon W^{\ell,p}(\widehat{K}) \to \mathcal{P}_{\widehat{K}}$. This relationship follows from $\mathrm{N}_{i,K}(u) = \mathrm{N}_{i,\widehat{K}}(u \circ F_K)$, which is satisfied for a large class of interpolants. The interpolant enjoys the interpolation property

$$\mathrm{I}_{\widehat{K}}\widehat{w} = \widehat{w} \quad \forall \widehat{w} \in \mathbb{P}_{\ell-1} \subset \mathcal{P}_{\widehat{K}},$$

where in this form a restriction on the parameter ℓ is hidden. Moreover, the interpolation operator is assumed to be continuous in the sense

$$|\widehat{v} - \mathrm{I}_{\widehat{K}}\widehat{v}|_{W^{m,q}(\widehat{K})} \le C(\mathcal{P}_{\widehat{K}})\|\widehat{v}\|_{W^{\ell,p}(\widehat{K})} \quad \forall v \in W^{\ell,p}(\widehat{K}),$$

which implies assumptions on the parameters m, q, ℓ and p.

The mapping F_K is also used to work out the dependence of the error on the size of the element $K = F_K(\widehat{K})$. In Lemma 2.8 we show for the case of an *affine* mapping the estimates

$$|u|_{W^{m,q}(K)} \le C|K|^{1/q}\varrho_K^{-m}|\widehat{u}|_{W^{m,q}(\widehat{K})},$$
$$|\widehat{u}|_{W^{\ell,p}(\widehat{K})} \le C|K|^{-1/p}h_K^{\ell}|u|_{W^{\ell,p}(K)},$$

where h_K is the diameter of K and ϱ_K is the diameter of the largest ball inscribed in K.

By using the polynomial w from the Deny-Lions lemma

$$\exists \widehat{w} \in \mathbb{P}_{\ell-1} : \quad \|\widehat{u} - \widehat{w}\|_{W^{\ell,p}(\widehat{K})} \le C|\widehat{u}|_{W^{\ell,p}(\widehat{K})},$$

see Subsection 2.2, we obtain via

$$\begin{aligned}
|u - \mathrm{I}_K u|_{W^{m,q}(K)} &\le C|K|^{1/q}\varrho_K^{-m}|\widehat{u} - \mathrm{I}_{\widehat{K}}\widehat{u}|_{W^{m,q}(\widehat{K})} \\
&= C|K|^{1/q}\varrho_K^{-m}|(\widehat{u} - \widehat{w}) - \mathrm{I}_{\widehat{K}}(\widehat{u} - \widehat{w})|_{W^{m,q}(\widehat{K})} \\
&\le CC(\mathcal{P}_{\widehat{K}})|K|^{1/q}\varrho_K^{-m}\|\widehat{u} - \widehat{w}\|_{W^{\ell,p}(\widehat{K})} \\
&\le CC(\mathcal{P}_{\widehat{K}})|K|^{1/q}\varrho_K^{-m}|\widehat{u}|_{W^{\ell,p}(\widehat{K})} \\
&\le CC(\mathcal{P}_{\widehat{K}})|K|^{1/q-1/p}\varrho_K^{-m}h_K^{\ell}|u|_{W^{\ell,p}(K)}
\end{aligned}$$

the desired error estimate.

In the h-version finite element method (h-FEM) the polynomial space $\mathcal{P}_{\widehat{K}}$ is fixed so that $C(\mathcal{P}_{\widehat{K}})$ is just a constant. Hence Section 2 focusses on investigating various interpolation operators and element shapes. In contrast to that, it is important in the hp-version finite element method (hp-FEM) to elaborate the dependence of $C(\mathcal{P}_{\widehat{K}})$ on the polynomial degree k. Section 3 focusses on this issue in the one-dimensional case and briefly in the multi-dimensional case.

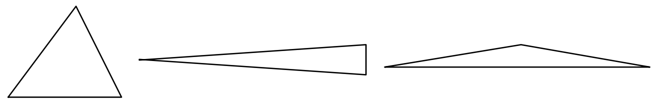

Figure 1. Isotropic and anisotropic triangles.

1.4. Shape Regularity and Conditions on the Element Maps

A typical parameter for the description of the elements K is the *aspect ratio* γ_K, the ratio of the diameter h_K of K and the diameter ϱ_K of the largest ball inscribed in K. We will call elements with a moderate aspect ratio *isotropic* and elements with a large aspect ratio *anisotropic*. For isotropic elements, we allow the quantity γ_K to be absorbed in constants in error estimates, whereas for anisotropic elements, the impact of the aspect ratio must be made explicit. Ideally, estimates that are uniform in the aspect ratio are sought.

Example 1.2. *(Isotropic and anisotropic simplices) tetrahedra and triangles with planar faces (edges) are sometimes called* shape-regular *if they are isotropic. Shape-regularity is generally used as a property that is easy to achieve in mesh generation and that allows for a numerical analysis at moderate technical expense, e.g., interpolation error estimation (uniform in* γ_K*), or the proof of a discrete inf-sup condition in mixed methods or efficient multilevel solvers.*

Zlámal (1968) has shown for triangles with straight edges that a lower bound on the interior angles is equivalent to an upper bound on the aspect ratio. Therefore, shape-regularity can be defined equivalently via a minimal angle condition*: There exists a constant* $\gamma_{\min} > 0$ *such that the angles of all triangles of a family of triangulations are bounded from below by* $\gamma_{\min}$*.*

Elements with large aspect ratio can be used advantageously for the approximation of anisotropic features in functions, for example, boundary layers or singularities in the neighborhood of concave edges of the domain. For the numerical analysis, it is often necessary to impose a maximal angle condition*: There exists a constant* $\gamma_{\max} > 0$ *such that the angles of all triangles of a family of triangulations are bounded from above by* $\gamma_{\max}$*. An analogous definition can be given for tetrahedra (see Apel (1999a), pages 54, 90f). Figure 1.2 shows an isotropic triangle and two anisotropic triangles, one that satisfies the maximal angle condition and one that does not. Note that if the angles are bounded from below away from zero, they are also bounded from above away from* π*, whereas the converse is not true. Therefore, estimates are usually easier to obtain for shape-regular elements (where* $\cot\gamma_{\min}$ *enters the constant) than for anisotropic elements (where, if necessary at all,* $\cot\gamma_{\max}$ *enters the constants).*

Most monographs consider only shape-regular elements, for example, Braess (1997), Brenner and Scott (1994), Ciarlet (1978, 1991), Hughes (1987), and Oswald (1994). Anisotropic elements are investigated mainly in research papers and in the book by Apel (1999a). The maximal angle condition was introduced first by Synge (1957), and later rediscovered by Gregory (1975), Babuška and Aziz (1976), and Jamet (1976). ∎

Example 1.3. *(Shape-regular quadrilaterals) From the theoretical point of view, it is important to distinguish between parallelograms and more general quadrilaterals. By an affine mapping* F_K *from a reference square* $\widehat{K}$ *one obtains parallelograms only, otherwise a more general sub- or isoparametric mapping has to be used. The affine case is of course simpler, and results for triangles can usually be extended to parallelograms. Shape-regularity is defined by a bounded aspect ratio* γ_K*. Parallelograms are sometimes even easier to handle than triangles since the edges point into two directions only. Similar statements can be made in the three-dimensional case for parallelepipeds.*

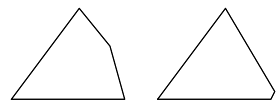

Figure 2. Degenerated isotropic quadrilaterals.

The situation changes for more general elements. A bounded aspect ratio is necessary but not sufficient for shape-regular quadrilaterals. For several estimates, it is advantageous to exclude quadrilaterals that degenerate to triangles (see Figure 1.3). The literature is not unanimous about an appropriate description. Ciarlet and Raviart (1972,b) demand a uniformly bounded ratio of the lengths of the longest and the shortest edge of the quadrilateral and that the interior angles are away from zero and π*. Girault and Raviart (1986) assume equivalently that the four triangles that can be formed from the vertices of the quadrilateral are shape-regular in the sense of Example 1.2.*

Weaker mesh conditions were derived by Jamet (1977) and Acosta and Durán (2000). Jamet proves that the elements shown in Figure 1.3 can be admitted, but he still relies on a bounded aspect ratio. Acosta and Durán formulate the regular decomposition property*, which is the weakest known condition that allows to prove the standard interpolation error estimate for* $\mathbb{Q}_1$ *elements. For a detailed review, we refer to Ming and Shi (2002,b).*

Further classes of meshes can be described as being asymptotically parallelograms (see also the papers by Ming and Shi). Some results that are valid for parallelograms can be extended to such meshes but not to general quadrilateral meshes, for example, superconvergence results and interpolation error estimates for certain serendipity elements (compare Remark 2.14). Meshes of one of these classes arise typically from a hierarchical refinement of a coarse initial mesh.

A more detailed discussion of all these conditions is beyond the frame of this chapter. We will restrict further discussion to affine elements and to elements that are shape-regular in the sense of Ciarlet/Raviart or Girault/Raviart. ▪

Remark 1.4. ***(curved elements)*** *We presented the arguments in Subsection 1.3 for* affine *element maps* F_K*. Non-affine element maps* F_K*, for example, the isoparametric mappings mentioned above or transfinite blending elements, (Gordon and Hall, 1973), may be required to ensure sufficiently accurate representations of the geometry. This is particularly pronounced in the context of high order methods and therefore addressed in the chapter* ***The*** p***-Version of the Finite Element Method*** *of* ***ECM2*** *(Sec. 2.4). Various issues arise from the use of non-affine* F_K *including the following two difficulties not encountered in the affine case: First, the parametrizations of common faces of neighboring elements have to be compatible so that the approximation space* $\mathrm{FE}_{\mathcal{T}}$ *is sufficiently large. Second, some conditions on the derivatives of* F_K *have to be imposed in order to make sure that the expected powers of* $|K|$ *arise by the transformation arguments of Subsection 1.3. In the present article, we do not elaborate these issues, only the latter is discussed briefly in Remarks 2.13 and 2.14, page 14, and Example 2.20, page 18, for the case of isoparametric mappings.* ▪

2. h-Version Interpolation

Section 2.1 is devoted to the definition of the interpolation operators. After a short discussion of the classical Deny-Lions lemma in Section 2.2, we derive error estimates for the nodal interpolation operator in Section 2.3, both for isotropic and anisotropic elements. We develop the theory in detail for affine elements and discuss briefly the nonaffine case. Quasi-interpolants are investigated in Section 2.4 for isotropic Lagrangian elements, whereas anisotropic elements are mentioned only in short. An example for a global interpolation error estimate is presented in Section 2.5. A typical solution of ellptic partial differential equations in corner domains with corner singularities is interpolated on a family of graded meshes, which are chosen such that the optimal order of convergence is obtained despite the irregular terms. We underline that the chapter is written in the spirit of the h-version of the finite element method. We do not investigate the dependence of constants on the polynomial degree of the functions and refer to Section 3 for this.

2.1. Definition of Interpolation Operators

2.1.1. Nodal Interpolation. As already mentioned at the end of Subsection 1.2, Ciarlet (1978) introduces finite elements in a more abstract way as the triple $(K, \mathcal{P}_K, \mathcal{N}_K)$, where $\mathcal{N}_K = \{\mathrm{N}_{i,K}\}_{i=1}^n$ is a *basis* of the dual space $\mathcal{P}_K'$. The linear functionals in $\mathcal{N}_K$ are sometimes called *nodal variables* and define a basis $\{\phi_{j,K}\}_{j=1}^n$ of $\mathcal{P}_K$ via

$$\mathrm{N}_{i,K}(\phi_{j,K}) = \delta_{i,j}, \quad i, j = 1, \dots, n, \tag{2.1}$$

which is called the *nodal basis*. It is now straightforward to introduce the nodal interpolation operator

$$\mathrm{I}_K u := \sum_{i=1}^n \mathrm{N}_{i,K}(u)\phi_{i,K}.$$

This interpolation operator is well defined for functions u for which the functionals $\mathrm{N}_{i,K}$ can be evaluated. For example, if these functionals include the pointwise evaluation of derivatives up to order s, then I_K is defined for functions from $C^s(K) \supset W^{s',p}(K)$ with $s' > s + d/p$. If the functionals involve the evaluation of integrals only, the required smoothness is correspondingly lower. The duality condition (2.1) yields $\mathrm{I}_K \phi_{j,K} = \phi_{j,K}$, $j = 1, \dots, n$, and thus

$$\mathrm{I}_K \phi = \phi \quad \forall \phi \in \mathcal{P}_K. \tag{2.2}$$

We now impose assumptions in addition to those already made in the introduction.

Assumption 2.1. *The finite element $(K, \mathcal{P}_K, \mathcal{N}_K)$ is constructed on the basis of a reference element $(\widehat{K}, \mathcal{P}_{\widehat{K}}, \mathcal{N}_{\widehat{K}})$ satisfying Assumption 1.1 and the relation* (1.2)*, see page 4. Furthermore, we assume that the nodal variables satisfy*

$$\mathrm{N}_{i,K}(u) = \mathrm{N}_{i,\widehat{K}}(u \circ F_K), \quad i = 1, \dots, n, \tag{2.3}$$

for all u for which the functionals are well defined.

Under this assumption, we obtain the property

$$\begin{aligned} (\mathrm{I}_K u) \circ F_K &= \left(\sum_{i=1}^n \mathrm{N}_{i,K}(u)\phi_{i,K} \right) \circ F_K \\ &= \sum_{i=1}^n \mathrm{N}_{i,\widehat{K}}(\widehat{u})\phi_{i,\widehat{K}} = \mathrm{I}_{\widehat{K}} \widehat{u}, \end{aligned}$$

which allows one to estimate the error on the reference element $\widehat{K}$ and then transform the estimate to K.

The definition of $\mathrm{FE}_{\mathcal{T}}$ in (1.3) allows us to introduce the global interpolation operator $\mathrm{I}_{\mathcal{T}}$ by (1.4), i.e., $(\mathrm{I}_{\mathcal{T}}\, u)|_K = \mathrm{I}_K(u|_K)$ for all $K \in \mathcal{T}$. With the basis $\{\phi_i\}_{i=1}^{N_+}$ of $\mathrm{FE}_{\mathcal{T}}$ and the globalized set of nodal variables $\mathcal{N}_{\mathcal{T},+} = \{\mathrm{N}_i\}_{i=1}^{N_+}$, we can write

$$\mathrm{I}_{\mathcal{T}}\, u = \sum_{i=1}^{N_+} \mathrm{N}_i(u)\phi_i. \tag{2.4}$$

Note again that $\mathrm{I}_{\mathcal{T}}$ acts only on sufficiently regular functions so that all functionals N_i are well defined.

Remark 2.2. *We distinguish between the finite element space* $\mathrm{FE}_{\mathcal{T}}$ *of dimension* N_+ *and its* N*-dimensional subspace* $V_{\mathcal{T}}$ *where boundary conditions are imposed,* $\mathrm{N}_i(u) = 0$ *for* $u \in V$ *and* $i = N+1, \ldots, N_+$*. Therefore, equation (2.4) is equivalent to*

$$\mathrm{I}_{\mathcal{T}}\, u = \sum_{i=1}^{N} \mathrm{N}_i(u)\phi_i, \tag{2.5}$$

i.e., boundary conditions pose no particular difficulty for the analysis of the nodal interpolation operator. ▪

2.1.2. Quasi-interpolation. A drawback of nodal interpolation is the required regularity of the functions the operator acts on. For example, for Lagrangian elements, we need $u \in W^{s',p}(K)$ with $s' > d/p$ to obtain well-defined point values via the Sobolev embedding theorem. This assumption may fail even for simple problems like the Poisson problem with mixed boundary conditions in concave three-dimensional domains, where a r^λ-singularity with λ close to 0.25 may occur; here, $r = r(x)$ is the distance of a point $x \in \Omega$ to the set of edges. Moreover, an interpolation operator for $W^{1,2}(\Omega)$-functions is needed for the analysis of a posteriori error estimators and multilevel methods.

The remedy is the introduction of a quasi-interpolation operator

$$\mathrm{Q}_{\mathcal{T}}\, u = \sum_{i=1}^{N} \mathrm{N}_i(\Pi_i u)\phi_i, \tag{2.6}$$

i.e., we replace the function u in (2.5) by the regularized functions $\Pi_i u$. The index i indicates that we may use for each functional N_i a different, locally defined averaging operator Π_i.

For simplicity of exposition, we restrict ourselves to Lagrangian finite elements, that is, the nodal variables have the form $\mathrm{N}_i(u) = u(a^i)$, where the points $a^i \in \overline{\Omega}$ are the nodes. (Nodes of Lagrangian elements are points where function values define the interpolant, see also the chapter **Finite Element Methods** of **ECM2**. Nodes should be distinguished from vertices, which are the corners of the elements of the mesh.) For quasi-interpolation of C^1-elements, we refer to Girault and Scott (2002), and for the definition of quasi-interpolants for lowest-order Nédélec elements of first type and lowest-order Raviart-Thomas elements that fulfill the *commuting diagram property* (de Rham diagram), we refer to Schöberl (2001) and Ern and Guermond (2015b). A new $L^1(\Omega)$-stable quasi-interpolation operator which is a projection and preserves homogeneous boundary conditions is introduced by Ern and Guermond (2015a) and illustrated in H^1-, $H(\mathrm{curl})$- and $H(\mathrm{div})$-conforming finite element spaces.

Each node a^i, $i = 1, \ldots, N$, is now related to a subdomain $\omega_i \subset \overline{\Omega}$ and a finite-dimensional space $\mathcal{P}_i$. Different authors prefer different choices. We present two main examples.

Example 2.3. *(Clément operator) Clément (1975) considers* $\mathcal{P}_K = \mathbb{P}_k$ *in simplicial elements* K *with plane faces. Each node* a^i*,* $i = 1, \ldots, N$*, is related to the subdomain* $\omega_i := \operatorname{int}\operatorname{supp}\phi_i$*, where* ϕ_i *is the corresponding nodal basis function and* int *stands for interior. The averaging operator*

$$\Pi_i \colon L^1(\omega_i) \to \mathbb{P}_{\ell-1} \tag{2.7}$$

is then defined by

$$\int_{\omega_i} (v - \Pi_i v)\phi = 0 \quad \forall \phi \in \mathbb{P}_{\ell-1}, \tag{2.8}$$

which is for $v \in L^2(\omega_i)$ *the* $L^2(\omega_i)$*-projection into* $\mathbb{P}_{\ell-1}$*. This operator has the important property*

$$\Pi_i \phi = \phi \quad \forall \phi \in \mathbb{P}_{\ell-1}. \tag{2.9}$$

One can choose the parameter ℓ *depending on* k*, for example,* $\ell = k+1$*, or depending on the regularity of* u*. For* $u \in W^{s,p}(\Omega)$*, the choice* $\ell = \min\{s, k+1\}$ *is appropriate. Note that the resulting operator* $\mathrm{Q}_{\mathcal{T}}$ *defined by (2.6) has the property that* $(\mathrm{Q}_{\mathcal{T}}\, v)|_K$ *is determined by* $v|_{\omega_K}$ *and not only by* $v|_K$*, where*

$$\omega_K := \bigcup_{i:a^i \in K} \omega_i. \tag{2.10}$$

We estimate the interpolation error for this operator in Subsection 2.4. ▪

Remark 2.4. *There are several modifications of the Clément operator from Example 2.3. Bernardi (1989) computes the average in some reference domain* $\widehat{\omega}_i$*, which is chosen from a fixed set of reference domains. This idea is used to treat meshes with curved (isoparametric) simplicial and quadrilateral elements. The particular difficulty is that the transformation that maps* $\widehat{\omega}_i$ *to* ω_i *is only piecewise smooth. Bernardi and Girault (1998) and Carstensen (1999) modify further and project into spaces of piecewise polynomial functions. Oswald (1994) uses the particularly simple choice* $\omega_i = \operatorname{int} K_i$ *with* $a^i \in K_i$ *over which to average. Verfürth (1999b) develops a new projection operator* $P_G^{\ell-1}$ *(see the last paragraph in Subsection 2.2) and uses it in Verfürth (1999a) as the averaging operator* Π_i *in the definition of a quasi-interpolation operator. This modification allows for making explicit the constants in the error estimates.* ▪

Remark 2.5. *The quasi-interpolant* $\mathrm{Q}_{\mathcal{T}}\, v$ *satisfies Dirichlet boundary conditions by construction since the sum in (2.6) extends only over the* N *degrees of freedom of* $V_{\mathcal{T}}$*. The nice property of* $\mathrm{I}_{\mathcal{T}}$ *mentioned in Remark 2.2 is not satisfied for the Clément operator, since* $\mathrm{N}_i(\Pi_i u)$*,* $i = N+1, \ldots, N_+$*, is not necessarily zero for* $u \in V$*. Consequently, the elements adjacent to the Dirichlet boundary must be treated separately in the analysis of the interpolation error. An alternative is developed by Scott and Zhang (1990).* ▪

Example 2.6. *(Scott-Zhang operator) The operator is introduced by Scott and Zhang (1990) similarly to the Clément operator (see Example 2.3). In particular, the projector* $\Pi_i\colon L^1(\omega_i) \to \mathbb{P}_{\ell-1}$ *is also defined by (2.8). The essential difference is that* ω_i *(still satisfying* $a^i \in \overline{\omega}_i$*) is allowed to be a* $(d-1)$*-dimensional face of an element* K_i*, and for Dirichlet boundary nodes, one chooses* ω_i *to be part of the Dirichlet boundary* Γ_D*. In this way, we obtain* $\mathrm{N}_i(\Pi_i u) = 0$ *if* $a^i \in \Gamma_D$*. The operator even preserves nonhomogeneous Dirichlet boundary conditions* $v|_{\Gamma_D} = g$ *if* $g \in \mathrm{FE}_{\mathcal{T}}|_{\Gamma_D}$ *and* $\ell = k+1$ *in (2.7). The Scott-Zhang operator is also projector onto the finite element space, i.e., it preserves finite element functions.*

To be specific about the choice of ω_i *for* $a^i \notin \Gamma_D$*, we recall from Scott and Zhang (1990) that* $\overline{\omega}_i = K \in \mathcal{T}$ *if* $a^i \in \operatorname{int} K$*, and* $\overline{\omega}_i \ni a^i$ *is a face of some element otherwise. Note that the face is not uniquely determined if* a^i *does not lie in the interior of a face or an element. For an illustration and an application of this operator in a context where the nodal interpolant is not applicable, we refer to Apel, Sändig, and Whiteman (1996).*

The operator can be applied to functions whose traces on $(d-1)$*-dimensional manifolds* ω_i *are in* $L^1(\omega_i)$*, i.e., for* $u \in W^{\ell,p}(\Omega)$ *with* $\ell \geq 1$ *and* $p = 1$*, or with* $\ell > 1/p$ *and* $p > 1$*. Consequently, it requires more regularity than the Clément operator, but, in general, less than the nodal interpolant.*

Finally, Verfürth (1999a) remarks that in certain interpolation error estimates that are valid for both the Scott-Zhang and the Clément operators, the constant is smaller for the Clément operator. ▪

2.2. The Deny-Lions Lemma

In this section, we discuss a result from functional analysis that turns out to be a classical approximation result, the Deny-Lions lemma (Deny and Lions, 1953/54), which is an essential ingredient of many error estimates in the finite element theory. It essentially states that the $W^{\ell,p}(G)$-seminorm is a norm in the quotient space $W^{\ell,p}(G)/\mathbb{P}_{\ell-1}$. We formulate it for domains G of unit size $\operatorname{diam} G = 1$.

Lemma 2.7. *(Deny and Lions) Let the domain $G \subset \mathbb{R}^d$ with $\operatorname{diam} G = 1$ be star-shaped with respect to a ball $B \subset G$, and let $\ell \geq 1$ be an integer and $p \in [1, \infty]$. For each $u \in W^{\ell,p}(G)$, there is a $w \in \mathbb{P}_{\ell-1}$ such that*

$$\|u - w\|_{W^{\ell,p}(G)} \leq C|u|_{W^{\ell,p}(G)}, \tag{2.11}$$

where the constant C depends only on d, ℓ, and $\gamma := \operatorname{diam} G/\operatorname{diam} B = 1/\operatorname{diam} B$.

One can find different versions of the lemma and its proof in the literature. Instead of giving one of them in full detail, we sketch some of them, thereby elucidating some important points.

A classical proof is to choose a basis $\{\sigma_\alpha\}_{|\alpha|\leq\ell-1}$ of $\mathbb{P}'_{\ell-1}$ and to prove that $|u|_{W^{\ell,p}(G)} + \sum_{|\alpha|\leq\ell-1} |\sigma_\alpha(u)|$ defines a norm in $W^{\ell,p}(G)$ that is equivalent to $\|u\|_{W^{\ell,p}(G)}$. Determining $w \in \mathbb{P}_{\ell-1}$ by $\sigma_\alpha(u - w) = 0$ for all α: $|\alpha| \leq \ell - 1$ leads to (2.11). For $\ell = 1$, there is only one functional σ to be used, typically $\sigma(u) := |G|^{-1} \int_G u$. For $\ell \geq 2$, one can take the nodal variables $\mathcal{N}_S$ of a simplicial Lagrange element $(S, \mathbb{P}_{\ell-1}, \mathcal{N}_S)$ with $S \subset G$ (Braess, 1997) or $\sigma_\alpha(u) := |G|^{-1} \int_G D^\alpha u$ (Bramble and Hilbert, 1970). The proof is based on the compact embedding of $W^{1,p}(G)$ in $L^p(G)$ and has the disadvantage that it only ensures that the constant is independent of u, but it can depend on all parameters, in particular, on the shape of G. The result is useful when applied only on a reference element $G = \widehat{K}$. Dobrowolski (1998) uses $\sigma_\alpha(u) := |G|^{-1} \int_G D^\alpha u$ as well and obtains with a different proof that the constant is independent of the shape of G (in particular also independent of γ) but he needs convexity of G. Related estimates can be found in (Karkulik and Melenk, 2015), Thm. 4.2.

Dupont and Scott (1980) choose w to be the *averaged Taylor polynomial* (also called *Sobolev polynomial*)

$$T_B^\ell u(x) := \int_B T_y^\ell u(x) \phi(y) \,\mathrm{d}y \in \mathbb{P}_{\ell-1},$$

where the *Taylor polynomial of order $\ell - 1$ evaluated at y* is given by

$$T_y^\ell u(x) := \sum_{|\alpha|\leq\ell-1} \frac{1}{\alpha!} D^\alpha u(y)(x - y)^\alpha \in \mathbb{P}_{\ell-1}$$

and where B is the ball from the lemma, and $\phi(\cdot)$ is a smooth cut-off function with support $\overline{B}$ and $\int_{\mathbb{R}^d} \phi = 1$; see also Brenner and Scott (1994, Section 4.1). This polynomial has several advantages, among them the property

$$D^\alpha T_B^\ell u = T_B^{\ell-|\alpha|} D^\alpha u \quad \forall u \in W^{|\alpha|,1}(B), \tag{2.12}$$

which will lead to simplifications later. This choice of w also allows one to characterize the constant in (2.11) as stated in Lemma 2.7. Moreover, several extensions can be made: the domain may be generalized to a union of overlapping domains that are each star-shaped with respect to a ball, and the Sobolev index ℓ may be noninteger, see the original paper by Dupont and Scott (1980).

Verfürth (1999b) defines in a recursive manner the projector $P_G^\ell\colon H^\ell(\Omega) \to \mathbb{P}_\ell$,

$$\begin{aligned} p_G^\ell(u) &:= \sum_{|\alpha|=\ell} \frac{x^\alpha}{\alpha!} \frac{1}{|G|} \int_G D^\alpha u, \\ p_G^{k-1}(u) &:= p_G^k(u) + \sum_{|\alpha|=k-1} \frac{x^\alpha}{\alpha!} \frac{1}{|G|} \int_G [D^\alpha(u - p_G^k(u))], \qquad k = \ell, \ell-1, \dots, 1, \\ P_G^\ell &:= p_G^0(u), \end{aligned} \tag{2.13}$$

which also commutes with differentiation in the sense of (2.12) and allows one to prove (2.11) for $w = P_G^{\ell-1}u$ with a constant C, depending only on ℓ and $p \in [2,\infty]$. The restriction $p \geq 2$ is outweighed by the fact that, for convex Ω, the constant C does *not* depend on the parameter $\gamma := \operatorname{diam} G/\operatorname{diam} B$.

2.3. Local Error Estimates for the Nodal Interpolant

2.3.1. Isotropic Elements. The proof of local interpolation error estimates was already sketched in Subsection 1.3 and shall be detailed now. Recall that we assume that for each element $K \in \mathcal{T}$ there is a bijective mapping $F_K\colon \hat{x} \in \mathbb{R}^d \mapsto x = F_K(\hat{x}) \in \mathbb{R}^d$, which maps $\widehat{K}$ to K. The following lemma provides transformation formulae for seminorms of functions if F_K is affine.

Lemma 2.8. *Let $F_K(\hat{x}) = A\hat{x} + a$ be an affine mapping with $K = F_K(\widehat{K})$. If $\hat{u} \in W^{m,q}(\widehat{K})$, then $u = \hat{u} \circ F_K^{-1} \in W^{m,q}(K)$ and*

$$|u|_{W^{m,q}(K)} \leq C|K|^{1/q} \varrho_K^{-m} |\hat{u}|_{W^{m,q}(\widehat{K})}. \tag{2.14}$$

If $u \in W^{\ell,p}(K)$, then $\hat{u} = u \circ F_K \in W^{\ell,p}(\widehat{K})$ and

$$|\hat{u}|_{W^{\ell,p}(\widehat{K})} \leq C|K|^{-1/p} h_K^\ell |u|_{W^{\ell,p}(K)}. \tag{2.15}$$

The constants depend on the shape and size of $\widehat{K}$.

Proof. We follow Ciarlet (1978). By examining the affine mapping, we get $\widehat{\nabla}\hat{v} = A^{\mathrm{T}}\nabla v$ and thus

$$\begin{aligned} |u|_{W^{m,q}(K)} &\leq C|K|^{1/q} \|A^{-1}\|_2^m \, |\hat{u}|_{W^{m,q}(\widehat{K})}, \\ |\hat{u}|_{W^{\ell,p}(\widehat{K})} &\leq C|K|^{-1/p} \|A\|_2^\ell \, |u|_{W^{\ell,p}(K)}. \end{aligned}$$

The factor with the power of $|K|$ arises from the Jacobi determinant of the transformation. This determinant is equal to the ratio of the areas of K and $\widehat{K}$. The norm of A can be estimated by considering the transformation of the largest sphere $\widehat{S}$ contained in $\widehat{K}$. For all $\hat{x} \in \mathbb{R}^d$ with $|\hat{x}| = \varrho_{\widehat{K}} = \operatorname{diam} \widehat{S}$, there are two points $\hat{y}, \hat{z} \in \widehat{S}$ such that $\hat{x} = \hat{y} - \hat{z}$. By observing that $|A\hat{x}| = |(A\hat{y} + a) - (A\hat{z} + a)| = |y - z| \leq h_K$, we get

$$\|A\|_2 := \sup_{|\hat{x}|=\varrho_{\widehat{K}}} \frac{|A\hat{x}|}{\varrho_{\widehat{K}}} \leq \frac{h_K}{\varrho_{\widehat{K}}}$$

and analogously $\|A^{-1}\|_2 \leq h_{\widehat{K}}/\varrho_K$. This finishes the proof. □

We are now ready to state the local error estimate.

Theorem 2.9. *Let* $(\widehat{K}, \mathcal{P}_{\widehat{K}}, \mathcal{N}_{\widehat{K}})$ *be a reference element with*

$$\mathbb{P}_{\ell-1} \subset \mathcal{P}_{\widehat{K}}, \tag{2.16}$$

$$\mathcal{N}_{\widehat{K}} \subset (C^s(\widehat{K}))'. \tag{2.17}$$

Assume that $(K, \mathcal{P}_K, \mathcal{N}_K)$ *is affine equivalent to* $(\widehat{K}, \mathcal{P}_{\widehat{K}}, \mathcal{N}_{\widehat{K}})$*, i.e., Assumption 2.1 holds with an affine mapping* F_K*. Let* $u \in W^{\ell,p}(K)$ *with* $\ell \in \mathbb{N}$*,* $p \in [1, \infty]$*, such that*

$$W^{\ell,p}(\widehat{K}) \hookrightarrow C^s(\widehat{K}), \quad i.e., \quad \ell > s + \frac{d}{p}, \tag{2.18}$$

and let $m \in \{0, \dots, \ell-1\}$ *and* $q \in [1, \infty]$ *be such that*

$$W^{\ell,p}(\widehat{K}) \hookrightarrow W^{m,q}(\widehat{K}). \tag{2.19}$$

Then the estimate

$$|u - \mathrm{I}_K u|_{W^{m,q}(K)} \le C|K|^{1/q-1/p} h_K^\ell \varrho_K^{-m} |u|_{W^{\ell,p}(K)} \tag{2.20}$$

holds.

Proof. From (2.17) and (2.18), we obtain

$$|\mathrm{N}_{i,\widehat{K}}(\widehat{v})| \le C\|\widehat{v}\|_{C^s(\widehat{K})} \le C\|\widehat{v}\|_{W^{\ell,p}(\widehat{K})} \tag{2.21}$$

and, thus, with $\|\phi_{i,\widehat{K}}\|_{W^{m,q}(\widehat{K})} \le C$, the boundedness of the interpolation operator:

$$\begin{aligned} |\mathrm{I}_{\widehat{K}}\widehat{v}|_{W^{m,q}(\widehat{K})} &= \left| \sum_{i=1}^n \mathrm{N}_{i,\widehat{K}}(\widehat{v}) \phi_{i,\widehat{K}} \right|_{W^{m,q}(\widehat{K})} \\ &\le \sum_{i=1}^n |\mathrm{N}_{i,\widehat{K}}(\widehat{v})| \, |\phi_{i,\widehat{K}}|_{W^{m,q}(\widehat{K})} \le C\|\widehat{v}\|_{W^{\ell,p}(\widehat{K})}, \end{aligned}$$

where the constant depends not only on $\widehat{K}$, s, m, q, ℓ, and p but also on $\mathcal{N}_{\widehat{K}}$. The embedding (2.19) yields

$$|\widehat{v}|_{W^{m,q}(\widehat{K})} \le C\|\widehat{v}\|_{W^{\ell,p}(\widehat{K})}.$$

Combining these estimates, choosing $\widehat{w} \in \mathbb{P}_{\ell-1}$ according to Lemma 2.7, and using $\widehat{w} = \mathrm{I}_{\widehat{K}}\widehat{w}$ due to (2.16), we get

$$\begin{aligned} |\widehat{u} - \mathrm{I}_{\widehat{K}}\widehat{u}|_{W^{m,q}(\widehat{K})} &= |(\widehat{u} - \widehat{w}) - \mathrm{I}_{\widehat{K}}(\widehat{u} - \widehat{w})|_{W^{m,q}(\widehat{K})} \\ &\le C\|\widehat{u} - \widehat{w}\|_{W^{\ell,p}(\widehat{K})} \\ &\le C|\widehat{u}|_{W^{\ell,p}(\widehat{K})}. \end{aligned} \tag{2.22}$$

By transforming this estimate to K (using Lemma 2.8), we obtain the desired result. ∎

Corollary 2.10. *Under the assumptions of Theorem 2.9 we obtain for isotropic elements in particular*

$$|u - \mathrm{I}_K u|_{W^{m,q}(K)} \le C|K|^{1/q-1/p} h_K^{\ell-m} |u|_{W^{\ell,p}(K)}. \tag{2.23}$$

Note that this theorem restricts for simplicity to affine elements, but is valid not only for Lagrangian elements but also for other types, including Hermite elements. Note further that we have used the assumption (2.18) for deriving the bound (2.21). It is indeed not sufficient to assume $\widehat{u} \in W^{\ell,p}(\widehat{K}) \cap C^s(\widehat{K})$ as we see in the following example.

Example 2.11. *Let $\widehat{K}$ be the triangle with vertices $(0,0)$, $(1,0)$, and $(0,1)$, let $\widehat{r} = \sqrt{\widehat{x}_1^2 + \widehat{x}_2^2}$, and consider the linear interpolation polynomial. For the family of functions $\widehat{u}_\varepsilon(x_1, x_2) = \min\{1, \varepsilon \ln|\ln(\widehat{r}/\mathrm{e})|\}$ one can compute that $I_{\widehat{K}}\widehat{u}_\varepsilon(\widehat{x}_1, \widehat{x}_2) = 1 - \widehat{x}_1 - \widehat{x}_2$ (independent of ε), $|\widehat{u}_\varepsilon|_{H^1(\widehat{K})} \to 0$ for $\varepsilon \to 0$, and $\|\widehat{u}_\varepsilon - I_{\widehat{K}}\widehat{u}_\varepsilon\|_{L^2(\widehat{K})} \not\to 0$ for $\varepsilon \to 0$, hence the estimate $\|\widehat{u} - I_{\widehat{K}}\widehat{u}\|_{L^2(\widehat{K})} \le C|\widehat{u}|_{H^1(\widehat{K})}$ does not hold.* ∎

Remark 2.12. *Interpolation error estimates can also be proved for functions from weighted Sobolev spaces, for example,*

$$H^{2,\alpha}(G) := \{u \in W^{1,2}(G) : r^\alpha D^\beta u \in L^2(G)\ \forall \beta : |\beta| = 2\},$$

where r is the distance to some point $x \in \overline{G} \subset \mathbb{R}^2$, and

$$\begin{aligned} |u|^2_{H^{2,\alpha}(G)} &:= \sum_{|\beta|=2} \|r^\alpha D^\beta u\|^2_{L^2(G)}, \\ \|u\|^2_{H^{2,\alpha}(G)} &:= \|u\|^2_{W^{1,2}(G)} + |u|^2_{H^{2,\alpha}(G)}. \end{aligned}$$

Grisvard (1985) shows in Lemma 8.4.1.3 the analog to the Deny-Lions lemma: For each $u \in H^{2,\alpha}(G)$ with $\alpha < 1$, there is a $w \in \mathbb{P}_1$ such that

$$\|u - w\|_{H^{2,\alpha}(G)} \le C(G)|u|_{H^{2,\alpha}(G)}. \tag{2.24}$$

The interpolation error estimate

$$|u - \mathrm{I}_K u|_{W^{1,2}(K)} \le C h_K^2 \varrho_K^{-1-\alpha} |u|_{H^{2,\alpha}(K)} \tag{2.25}$$

is then proved in Lemma 8.4.1.4 for triangles K. This result can be applied in the proof of mesh-grading techniques for singular solutions, where the singularity comes from corners in the domain $\Omega \subset \mathbb{R}^2$. ∎

Remark 2.13. *Second derivatives of an affine transformation F_K vanish. This leads to the special structure of the relations (2.14) and (2.15), where no low-order derivatives of $\widehat{u}$ and u, respectively, appear on the right-hand sides. This is no longer valid for nonaffine transformations. In the case that*

$$|\widehat{D}^\alpha F_K| \le C h_K^{|\alpha|} \quad \forall \alpha\colon |\alpha| \le \ell \tag{2.26}$$

we obtain

$$|\widehat{u}|_{W^{\ell,p}(\widehat{K})} \le C|K|^{-1/p} h_K^\ell \|u\|_{W^{\ell,p}(K)}, \tag{2.27}$$

which is weaker than (2.15), but is still sufficient for our purposes. The assumption (2.26) is satisfied when F_K differs only slightly from an affine mapping.

However, Estimate (2.26) is not valid for general quadrilateral meshes. Therefore, the theory has to be refined. For $\mathcal{P}_K = \mathbb{Q}_k$, this case can be treated with a sharper version of the Deny-Lions lemma: for each $u \in W^{\ell,p}(G)$ there is a $w \in \mathbb{Q}_{\ell-1}$ such that

$$\|u - w\|_{W^{\ell,p}(G)} \le C \left(\sum_{i=1}^d \left\| \frac{\partial^\ell u}{\partial x_i^\ell} \right\|_{L^p(G)} \right)^{1/p}$$

(see Bramble and Hilbert (1971)). For shape-regular elements (in the sense of Ciarlet/Raviart or Girault/Raviart, see Example 1.3) one can then prove (2.23). ∎

Remark 2.14. *Some results are weaker for general shape-regular quadrilateral elements than for (at least asymptotically) affine elements. For example, Arnold, Boffi, and Falk (2002) have shown for quadrilateral serendipity elements (here $\mathcal{P}_{\widehat{K}} = \mathbb{Q}'_k := \mathbb{P}_k \oplus \operatorname{span}\{\widehat{x}_1^k \widehat{x}_2, \widehat{x}_1 \widehat{x}_2^k\}$) that*

$$|u - \mathrm{I}_K u|_{W^{m,2}(K)} \le C h_K^{\lfloor k/2 \rfloor + 1 - m} |u|_{W^{k+1,2}(K)}, \quad m = 0, 1,$$

is sharp for general quadrilateral meshes, whereas for asymptotically parallelogram meshes, we get

$$|u - \mathrm{I}_K u|_{W^{m,2}(K)} \le C h_K^{k+1-m} |u|_{W^{k+1,2}(K)}, \quad m = 0, 1.$$

∎

2.3.2. Anisotropic Elements. Anisotropic elements are characterized by a large aspect ratio $\gamma_K := h_K/\varrho_K$. Estimate (2.20) can also be reformulated as

$$|u - \mathrm{I}_K u|_{W^{m,q}(K)} \leq C|K|^{1/q-1/p} h_K^{\ell-m} \gamma_K^m |u|_{W^{\ell,p}(K)},$$

which means that the quality of this estimate deteriorates if $m \geq 1$ and $\gamma_K \gg 1$. Let us examine whether the reason for this deterioration is formed by the anisotropic element (indicating that anisotropic elements should be avoided) or by the sharpness of the estimates (indicating that the estimates should be improved).

Example 2.15. *Consider the triangle K with the nodes $(-h, 0)$, $(h, 0)$, and $(0, \varepsilon h)$ and interpolate the function $u(x_1, x_2) = x_1^2$ in the vertices with polynomials of degree one. Then $\mathrm{I}_K u = h^2 - \varepsilon^{-1} h x_2$ and*

$$\begin{aligned} \frac{|u - \mathrm{I}_K u|_{W^{1,2}(K)}}{|u|_{W^{2,2}(K)}} &= \left(\frac{2h^4 \left[\frac{1}{3}\varepsilon + \frac{1}{2}\varepsilon^{-1}\right]}{4h^2 \varepsilon} \right)^{1/2} \\ &= h \left(\tfrac{1}{6} + \tfrac{1}{4}\varepsilon^{-2}\right)^{1/2} = c_\varepsilon h \end{aligned}$$

with $c_\varepsilon \to \infty$ for $\varepsilon \to 0$ and $c_\varepsilon \geq C\gamma_K$. We find that Estimate (2.20) cannot, in general, be essentially improved and that (2.23) is not valid. Estimate (2.20) can be improved only slightly by investigating in more detail the transformation from Lemma 2.8 (see e.g. Formaggia and Perotto (2001)). ■

Example 2.16. *Consider now the triangle with the nodes $(0, 0)$, $(h, 0)$, and $(0, \varepsilon h)$ and interpolate again the function $u(x_1, x_2) = x_1^2$ in $\mathbb{P}_1$. We get $I_K u = h x_1$ and*

$$\frac{|u - \mathrm{I}_K u|_{W^{1,2}(K)}}{|u|_{W^{2,2}(K)}} = \left(\frac{\frac{1}{6}\varepsilon h^4}{2\varepsilon h^2} \right)^{1/2} = \frac{1}{\sqrt{12}} h$$

where the constant is independent of ε. Estimate (2.23) is valid, although the element is anisotropic for small ε. ■

From the two examples, we can learn that the aspect ratio is not the right quantity to characterize elements that yield an interpolation error estimate of quality (2.23). Synge (1957) proved for triangles K and $\mathcal{P}_K = \mathbb{P}_1$ that

$$|u - \mathrm{I}_K u|_{W^{1,\infty}(K)} \leq C h_K |u|_{W^{2,\infty}(K)}$$

with a constant that depends linearly on $(\cos(\alpha/2))^{-1}$, where α is the maximal angle in K. The maximal angle condition was found (see also the comment in Example 1.2).

We will now elaborate that Estimate (2.23) *cannot* be obtained from (2.22) by just treating the transformation F_K more carefully (compare also Apel (1999a), Example 2.1). To do so, we consider again the triangle K from Example 2.16 with $\mathcal{P}_K = \mathbb{P}_1$. The transformation to the reference element $\widehat{K}$ with nodes $(0, 0)$, $(1, 0)$, and $(0, 1)$ is done via $x_1 = h\widehat{x}_1$, $x_2 = \varepsilon h \widehat{x}_2$. Transforming (2.22) in the special case $p = q = 2$, $m = 1$, $\ell = 2$ leads to

$$\begin{aligned} & h \left(\left\| \frac{\partial(u - \mathrm{I}_K u)}{\partial x_1} \right\|^2_{L^2(K)} + \varepsilon^2 \left\| \frac{\partial(u - \mathrm{I}_K u)}{\partial x_2} \right\|^2_{L^2(K)} \right)^{1/2} \\ & \quad \leq C h^2 \left(\left\| \frac{\partial^2 u}{\partial x_1^2} \right\|^2_{L^2(K)} + \varepsilon^2 \left\| \frac{\partial^2 u}{\partial x_1 \partial x_2} \right\|^2_{L^2(K)} + \varepsilon^4 \left\| \frac{\partial^2 u}{\partial x_2^2} \right\|^2_{L^2(K)} \right)^{1/2}, \end{aligned}$$

which can be simplified to

$$\begin{aligned} \left\| \frac{\partial (u - \mathrm{I}_K u)}{\partial x_1} \right\|_{L^2(K)} &\leq Ch|u|_{W^{2,2}(K)}, \\ \left\| \frac{\partial (u - \mathrm{I}_K u)}{\partial x_2} \right\|_{L^2(K)} &\leq C\varepsilon^{-1}h|u|_{W^{2,2}(K)}, \end{aligned} \tag{2.28}$$

but the independence of ε^{-1} is not obtainable in (2.28). This factor could only be avoided if we proved on the reference element the sharper estimate

$$\left\| \frac{\partial (\widehat{u} - \mathrm{I}_{\widehat{K}} \widehat{u})}{\partial \widehat{x}_2} \right\|_{L^2(\widehat{K})} \leq C \left(\left\| \frac{\partial^2 \widehat{u}}{\partial \widehat{x}_1 \partial \widehat{x}_2} \right\|^2_{L^2(\widehat{K})} + \left\| \frac{\partial^2 \widehat{u}}{\partial \widehat{x}_2^2} \right\|^2_{L^2(\widehat{K})} \right)^{1/2}.$$

The following lemma from Apel and Dobrowolski (1992) (see also Apel (1999a), Lemma 2.2) reduces the proof to finding functionals σ_i with certain properties.

Lemma 2.17. *Let* $\mathrm{I}_{\widehat{K}} \colon C^s(\widehat{K}) \to \mathcal{P}_{\widehat{K}}$ *be a linear operator and assume that* $\mathbb{P}_k \subset \mathcal{P}_{\widehat{K}}$. *Fix* $m, \ell \in \mathbb{N}_0$ *and* $p, q \in [1, \infty]$ *such that* $0 \leq m \leq \ell \leq k+1$ *and*

$$W^{\ell-m,p}(\widehat{K}) \hookrightarrow L^q(\widehat{K}). \tag{2.29}$$

Consider a multi-index γ *with* $|\gamma| = m$ *and define* $J := \dim \widehat{D}^\gamma \mathcal{P}_{\widehat{K}}$. *Assume that there are linear functionals* σ_i, $i = 1, \ldots, J$, *such that*

$$\sigma_i \in (W^{\ell-m,p}(\widehat{K}))' \quad \forall i = 1, \ldots, J, \tag{2.30}$$

$$\begin{aligned} &\sigma_i(\widehat{D}^\gamma(\widehat{u} - \mathrm{I}_{\widehat{K}}\widehat{u})) = 0 \quad \forall i = 1, \ldots, J, \\ &\qquad \forall \widehat{u} \in C^s(\widehat{K}) \colon \widehat{D}^\gamma \widehat{u} \in W^{\ell-m,p}(\widehat{K}), \end{aligned} \tag{2.31}$$

$$\begin{aligned} &\widehat{w} \in \mathcal{P}_{\widehat{K}} \text{ and } \sigma_i(\widehat{D}^\gamma \widehat{w}) = 0 \quad \forall i = 1, \ldots, J, \\ &\qquad \Rightarrow \quad \widehat{D}^\gamma \widehat{w} = 0. \end{aligned} \tag{2.32}$$

Then the error can be estimated for all $\widehat{u} \in C^s(\widehat{K})$ *with* $\widehat{D}^\gamma \widehat{u} \in W^{\ell-m,p}(\widehat{K})$ *by*

$$\|\widehat{D}^\gamma(\widehat{u} - \mathrm{I}_{\widehat{K}}\widehat{u})\|_{L^q(\widehat{K})} \leq C|\widehat{D}^\gamma \widehat{u}|_{W^{\ell-m,p}(\widehat{K})}. \tag{2.33}$$

Proof. The proof is based on two ingredients. First, we conclude from (2.12) and the Deny-Lions lemma 2.7 that $\widehat{w} = T_B^\ell \widehat{u} \in \mathbb{P}_{\ell-1} \subset \mathbb{P}_k$ satisfies

$$\widehat{w}_\gamma := \widehat{D}^\gamma \widehat{w} = T_B^{\ell-|\gamma|} \widehat{D}^\gamma \widehat{u} \in \mathbb{P}_{k-m}$$

and

$$\|D^\gamma \widehat{u} - \widehat{w}_\gamma\|_{W^{\ell-m,p}(\widehat{K})} \leq C|\widehat{D}^\gamma \widehat{u}|_{W^{\ell-m,p}(\widehat{K})}. \tag{2.34}$$

Second, we see that $\widehat{D}^\gamma(\widehat{w} - \mathrm{I}_{\widehat{K}}\widehat{u}) \in \widehat{D}^\gamma \mathcal{P}_{\widehat{K}}$. Moreover, $\sum_{i=1}^J |\sigma_i(\cdot)|$ and $\|.\|_{L^q(\widehat{K})}$ are equivalent norms in $\widehat{D}^\gamma \mathcal{P}_{\widehat{K}}$. Therefore, we get with (2.31) and (2.30)

$$\begin{aligned} \|\widehat{D}^\gamma(\widehat{w} - \mathrm{I}_{\widehat{K}}\widehat{u})\|_{L^q(\widehat{K})} &\leq C \sum_{i=1}^J |\sigma_i(\widehat{D}^\gamma(\widehat{w} - \mathrm{I}_{\widehat{K}}\widehat{u}))| \\ &= C \sum_{i=1}^J |\sigma_i(\widehat{D}^\gamma(\widehat{w} - \widehat{u}))| \\ &\leq C\|\widehat{D}^\gamma(\widehat{w} - \widehat{u})\|_{W^{\ell-m,p}(\widehat{K})}. \end{aligned}$$

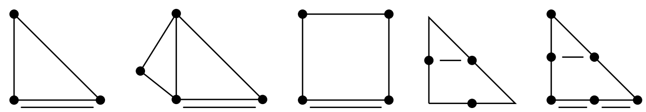

Figure 3. Functionals for Lagrangian elements.

Consequently, we obtain with (2.29) and (2.34)

$$\begin{aligned}\|\widehat{D}^\gamma(\widehat{u} - \mathrm{I}_{\widehat{K}}\widehat{u})\|_{L^q(\widehat{K})} &\leq \|\widehat{D}^\gamma(\widehat{u} - \widehat{w})\|_{L^q(\widehat{K})} \\ &\quad + \|\widehat{D}^\gamma(\widehat{w} - \mathrm{I}_{\widehat{K}}\widehat{u})\|_{L^q(\widehat{K})} \\ &\leq C\|\widehat{D}^\gamma(\widehat{w} - \widehat{u})\|_{W^{\ell-m,p}(\widehat{K})} \\ &\leq |\widehat{D}^\gamma\widehat{u}|_{W^{\ell-m,p}(\widehat{K})}\end{aligned}$$

which is the assertion. □

The creative part is now to find the functionals $\{\sigma_i\}_{i=1}^J$ with the properties (2.30)-(2.32).

Example 2.18. *For Lagrangian elements and $m = 1$, the functionals are integrals over some lines (see Figure 3). One can easily check (2.31) and (2.32). The critical condition is (2.30), which is satisfied for $\ell = 2$ only if $d = 2$ or $p > 2$. One can indeed give an example that shows that*

$$\|\widehat{D}^\gamma(\widehat{u} - \mathrm{I}_{\widehat{K}}\widehat{u})\|_{L^p(\widehat{K})} \leq C|\widehat{D}^\gamma\widehat{u}|_{W^{1,p}(\widehat{K})}$$

does not hold for $d = 3$, $p \leq 2$ (see Apel and Dobrowolski (1992)). ■

Let us transform Estimate (2.33) to the element $K = F_K(\widehat{K})$. We easily see that if $F_K(\widehat{x}) = A\widehat{x} + a$ with $A = \operatorname{diag}(h_{1,K}, \ldots, h_{d,K})$, then we get

$$\begin{aligned}&|K|^{-1/q}h_{1,K}^{\gamma_1}\cdots h_{d,K}^{\gamma_d}\|D^\gamma(u - \mathrm{I}_K u)\|_{L^q(K)} \\ &\quad\leq C|K|^{-1/p}h_{1,K}^{\gamma_1}\cdots h_{d,K}^{\gamma_d}\sum_{|\alpha|=\ell-m} h_{1,K}^{\alpha_1}\cdots h_{d,K}^{\alpha_d}\|D^{\alpha+\gamma}u\|_{L^p(K)}.\end{aligned}$$

Dividing by $|K|^{-1/q}h_{1,K}^{\gamma_1}\cdots h_{d,K}^{\gamma_d}$ and summing up over all γ with $|\gamma| = m$, we obtain the anisotropic interpolation error estimate

$$|u - \mathrm{I}_K u|_{W^{m,q}(K)} \leq C|K|^{1/q-1/p}\sum_{|\alpha|=\ell-m} h_{1,K}^{\alpha_1}\cdots h_{d,K}^{\alpha_d}|D^\alpha u|_{W^{m,p}(K)}. \tag{2.35}$$

Remark 2.19. *The question for which more general elements the estimate* (2.35) *holds, is tightly connected to the choice of a coordinate system in which the element K is described. A more detailed calculation shows that estimate* (2.35) *can also be obtained when the off-diagonal entries of $A = [a_{i,j}]_{i,j=1}^d$ in the affine mapping* (1.1) *are not zero but small,*

$$a_{i,j} \leq C\min\{h_{i,K}, h_{j,K}\}, \quad i,j = 1,\ldots,d, \quad i \neq j, \tag{2.36}$$

see Apel and Lube (1998). A geometrical description of two- and three-dimensional affine elements that satisfy (2.36) is given in terms of a maximal angle condition *and a* coordinate system condition *in Apel (1999a).*

Cao (2005) analyzes the matrix A in the affine mapping (1.1) *by using the singular value decomposition* $A = U\Sigma V^T$ *which essentially describes a rotation of the reference element, followed by a contraction and another rotation. He follows Formaggia and Perotto (2001) and uses the columns of U as the coordinate system for the partial derivatives on the right hand side of* (2.35)*. Later, Hetmaniuk and Knupp (2008) advocate to use the columns of A; these are directions of two sides of K but, in general, do not form an orthogonal coordinate system. If the maximal angle condition is not satisfied, this coordinate system does not form a stable basis.* ∎

The situation is more difficult for nonaffine elements.

Example 2.20. *Consider the quadrilateral K with nodes $(0,0)$, $(h_1,0)$, (h_1,h_2), (ε,h_2), $0<\varepsilon\le h_2\le h_1$, which is an ε-perturbation of the (affine) rectangle $(0,h_1)\times(0,h_2)$. We have*

$$F_K(\widehat{x}) = \begin{pmatrix} h_1\widehat{x}_1 + \varepsilon(1-\widehat{x}_1)\widehat{x}_2 \\ h_2\widehat{x}_2 \end{pmatrix}$$

and as in the affine theory

$$\begin{aligned}
|u - \mathrm{I}_K u|_{W^{1,2}(K)} &\le |K|^{1/2}\sum_{i=1}^{2} h_i^{-1}\left\|\frac{\partial(\widehat{u}-\mathrm{I}_{\widehat{K}}\widehat{u})}{\partial\widehat{x}_i}\right\|_{L^2(\widehat{K})} \\
&\le |K|^{1/2}\sum_{i=1}^{2} h_i^{-1}\left|\frac{\partial\widehat{u}}{\partial\widehat{x}_i}\right|_{W^{1,2}(\widehat{K})}, \\
\left\|\frac{\partial^2\widehat{u}}{\partial\widehat{x}_1^2}\right\|_{L^2(\widehat{K})} &\le C|K|^{-1/2}h_1^2\left\|\frac{\partial^2 u}{\partial x_1^2}\right\|_{L^2(K)}, \\
\left\|\frac{\partial^2\widehat{u}}{\partial\widehat{x}_2^2}\right\|_{L^2(\widehat{K})} &\le C|K|^{-1/2}h_2^2|u|_{W^{2,2}(K)},
\end{aligned}$$

but owing to $\partial^2 x_1/\partial\widehat{x}_1\partial\widehat{x}_2 = -\varepsilon \ne 0$, we get only

$$\left\|\frac{\partial^2\widehat{u}}{\partial\widehat{x}_1\partial\widehat{x}_2}\right\|_{L^2(\widehat{K})} \le C|K|^{-1/2}\left(h_1h_2\left|\frac{\partial u}{\partial x_1}\right|_{W^{1,2}(K)} + \varepsilon\left\|\frac{\partial u}{\partial x_1}\right\|_{L^2(K)}\right)$$

and, thus, by using again $\varepsilon\le h_2\le h_1$

$$|u - \mathrm{I}_K u|_{W^{1,2}(K)} \le C\left(\sum_{i=1}^{2} h_i\left|\frac{\partial u}{\partial x_i}\right|_{W^{1,2}(K)} + \frac{\varepsilon}{h_2}\left\|\frac{\partial u}{\partial x_1}\right\|_{L^2(K)}\right).$$

In Apel (1998), we concluded that we should allow only perturbations with $\varepsilon\le Ch_1h_2$, but later we found in Apel (1999a) a sharper estimate without the latter term: observing that $\mathbb{P}_1\in\mathcal{P}_K$, we get for $w\in\mathbb{P}_1$

$$\begin{aligned}
|u - \mathrm{I}_K u|_{W^{1,2}(K)} &= |(u-w) - \mathrm{I}_K(u-w)|_{W^{1,2}(K)} \\
&\le C\left(\sum_{i=1}^{2} h_i\left|\frac{\partial u}{\partial x_i}\right|_{W^{1,2}(K)} + \frac{\varepsilon}{h_2}\left\|\frac{\partial(u-w)}{\partial x_1}\right\|_{L^2(K)}\right).
\end{aligned}$$

By another Deny-Lions argument, compare (2.34), we get for appropriate w

$$\left\|\frac{\partial(u-w)}{\partial x_1}\right\|_{L^2(K)} \le C\sum_{i=1}^{2} h_i\left|\frac{\partial^2 u}{\partial x_1\partial x_i}\right|_{L^2(K)}$$

so that

$$|u - \mathrm{I}_K u|_{W^{1,2}(K)} \le C \sum_{i=1}^{2} h_i \left| \frac{\partial u}{\partial x_i} \right|_{W^{1,2}(K)}$$

can be proved for $\varepsilon \le Ch_2$. ∎

The approach from the example, where a second Deny-Lions argument is used, holds also for more general quadrilateral elements K with straight edges (subparametric elements) and $\mathcal{P}_{\widehat{K}} = \mathbb{Q}_k$ (see Apel (1999a)).

Summarizing this subsection, we can say that the anisotropic interpolation error estimate (2.35) can be proved for a large class of affine and subparametric elements (for details we refer to Apel (1999a)). Estimates for functions from weighted Sobolev spaces have been proved as well; see Apel and Nicaise (1996, 1998) and Apel, Nicaise, and Schöberl (2001).

Chen, Zhao, and Shi (2003, 2004) developed a method that supports finding the functionals $\{\sigma_i\}_{i=1}^J$ and used it to prove anisotropic interpolation error estimates for nonconforming elements, namely, the Adini, the Wilson element, and a variant of it. Chen, Yin, and Mao (2008) showed the same for a nonconforming plate element. The nonconforming lowest order Crouzeix-Raviart element (triangle and tetrahedron) and prismatic variants of it were investigated by Apel, Nicaise, and Schöberl (2001). These results were extended to higher polynomial degree by Apel and Matthies (2008). Anisotropic interpolation error estimates are proved by Acosta, Apel, Durán, and Lombardi (2008) for an interpolant defined in terms of moments for anisotropic triangles, quadrilaterals, tetrahedra, and hexahedra and arbitrarily high polynomial degree.

The anisotropic interpolation error estimates are suited to compensate large partial derivatives $D^\alpha u$ by an appropriate choice of the element sizes $h_{1,K}, \ldots, h_{d,K}$ in order to equilibrate the terms in the sum on the right-hand side of (2.35). The results can be applied to problems with anisotropic solutions; in particular, flow problems where first results on the resolution of all kinds of layers or shock fronts can be found, for example, in Peraire, Vahdati, Morgan, and Zienkiewicz (1987), Kornhuber and Roitzsch (1990), Zhou and Rannacher (1993), Zienkiewicz and Wu (1994), and Roos, Stynes, and Tobiska (1996).

2.4. Local Error Estimates for Quasi-Interpolants

Recall from Subsection 2.1.2 that we restrict ourselves here to Lagrangian elements, that is, $\mathrm{N}_{i,K}(u) = u(a^i)$ where a^i, $i = 1, \ldots, n$, are the nodes of K. The quasi-interpolants can be defined locally by

$$\mathrm{Q}_K u := \sum_{i=1}^{\widetilde{n}} \mathrm{N}_{i,K}(\Pi_{i,K} u)\phi_{i,K}$$

with projectors $\Pi_{i,K} u\colon L^1(\omega_{i,K}) \to \mathbb{P}_{\ell-1}$ and sets $\omega_{i,K}$ that are defined differently by Clément and Scott/Zhang (see Examples 2.3 and 2.6, respectively). The local number of degrees of freedom that defines the operator correctly is denoted by $\widetilde{n}$. For the Scott-Zhang operator, we can use $\widetilde{n} = n$. For the Clément operator, we also have $\widetilde{n} = n$ if $K \cap \Gamma_D = \emptyset$, but $\widetilde{n} < n$ if K touches the Dirichlet boundary. Let $\omega_K \subset \Omega$ be the interior of a union of finite elements with $K \subset \overline{\omega}_K$ and $\omega_{i,K} \subset \omega_K$, $i = 1, \ldots, \widetilde{n}$; typically, one defines

$$\overline{\omega}_K := \bigcup_{K_i \in \mathcal{T}: K_i \cap K \neq \emptyset} K_i.$$

We will prove error estimates in a uniform way for both operators and for triangles, tetrahedra, quadrilaterals, and hexahedra, but we restrict ourselves to affine isotropic elements.

To bound the interpolation error $u - \mathrm{Q}_K u$, we need several ingredients. The first one is the inclusion $\mathbb{P}_{\ell-1} \subset \mathcal{P}_K$ which is satisfied for the affine elements mentioned above if $\mathcal{P}_K = \mathbb{P}_k$ or $\mathcal{P}_K = \mathbb{Q}_k$ and $\ell \le k+1$. Then we obtain

$$\mathrm{Q}_K w = w \quad \forall w \in \mathbb{P}_{\ell-1} \tag{2.37}$$

because of $\mathrm{N}_{i,K}(\Pi_{i,K} w) = \mathrm{N}_{i,K}(w)$ for $w \in \mathbb{P}_{\ell-1}$.

Since $\Pi_{i,K} v$ is from a finite-dimensional space, we have

$$\|\Pi_{i,K} v\|_{L^\infty(\omega_{i,K})} \le C|\omega_{i,K}|^{-1/2} \|\Pi_{i,K} v\|_{L^2(\omega_{i,K})}.$$

Moreover, we get from the definition (2.8) with $\phi = \Pi_{i,K} u$ that

$$\begin{aligned} \|\Pi_{i,K} v\|^2_{L^2(\omega_{i,K})} &= \int_{\omega_{i,K}} v\, \Pi_{i,K} v \\ &\le \|v\|_{L^1(\omega_{i,K})} \|\Pi_{i,K} v\|_{L^\infty(\omega_{i,K})}, \end{aligned}$$

i.e.,

$$|\mathrm{N}_{i,K}(\Pi_{i,K} v)| \le \|\Pi_{i,K} v\|_{L^\infty(\omega_{i,K})} \le C|\omega_{i,K}|^{-1} \|v\|_{L^1(\omega_{i,K})}. \tag{2.38}$$

If $\omega_{i,K}$ is d-dimensional, we conclude

$$|\mathrm{N}_{i,K}(\Pi_{i,K} v)| \le C|K|^{-1} \|v\|_{L^1(\omega_K)}.$$

If $\omega_{i,K}$ is $(d-1)$-dimensional, namely, a face of an element $K_i \subset \omega_K$, we need to apply the trace theorem on the reference element. By transforming to K_i, we obtain

$$\begin{aligned} \|v\|_{L^1(\omega_{i,K})} &\le C|\omega_{i,K}|\, |K_i|^{-1} \left(\|v\|_{L^1(K_i)} + h_{K_i} |v|_{W^{1,1}(K_i)} \right), \\ |\mathrm{N}_{i,K}(\Pi_{i,K} v)| &\le C|K|^{-1} \left(\|v\|_{L^1(\omega_K)} + h_K |v|_{W^{1,1}(\omega_K)} \right), \end{aligned}$$

where we used the fact that adjacent isotropic elements are of comparable size: if $K_i \cap K_j \ne \emptyset$ then $h_{K_i} \le C_{\mathcal{T}} h_{K_j}$ with a constant $C_{\mathcal{T}}$ of moderate size, for example, $C_{\mathcal{T}} = 2$. We are now able to bound the norm of Q_K by

$$\begin{aligned} |\mathrm{Q}_K v|_{W^{m,q}(K)} &= \left| \sum_{i=1}^n \mathrm{N}_{i,K}(\Pi_{i,K} v) \phi_{i,K} \right|_{W^{m,q}(K)} \\ &\le \sum_{i=1}^n |\mathrm{N}_{i,K}(\Pi_{i,K} v)|\, |\phi_{i,K}|_{W^{m,q}(K)} \\ &\le C|K|^{-1} \sum_{j=0}^{\ell} h_K^j |v|_{W^{j,1}(\omega_K)}\, |K|^{1/q} h_K^{-m} \\ &\le C|K|^{1/q-1/p} \sum_{j=0}^{\ell} h_K^{j-m} |v|_{W^{j,p}(\omega_K)} \end{aligned} \tag{2.39}$$

with $\ell \ge 0$ for the Clément operator and $\ell \ge 1$ for the Scott-Zhang operator. We have also used the Hölder inequality $\|v\|_{L^1(K)} \le \|1\|_{L^{p'}(K)} \|v\|_{L^p(K)} = |K|^{1-1/p} \|v\|_{L^p(K)}$.

By transforming the embedding theorem $W^{\ell,p}(\widehat{K}) \hookrightarrow W^{m,q}(\widehat{K})$, which we assume to hold, we get

$$|v|_{W^{m,q}(K)} \le C|K|^{1/q-1/p} \sum_{j=0}^{\ell} h_K^{j-m} |v|_{W^{j,p}(K)}, \tag{2.40}$$

where we used the formulae from Lemma 2.8 and $h_K \le C\varrho_K$.

The next ingredient we need is a version of the Deny-Lions lemma 2.7. By the scaling $x = h_K \widetilde{x} \in \mathbb{R}^d$, we transform ω_K to $\widetilde{\omega}_K$, which is an isotropic domain with diameter of order 1. Thus, for any $\widetilde{u} \in W^{\ell,p}(\widetilde{\omega}_K)$, there is a polynomial $\widetilde{w} \in \mathbb{P}_{\ell-1}$ such that

$$\|\widetilde{u} - \widetilde{w}\|_{W^{\ell,p}(\widetilde{\omega}_K)} \le C|\widetilde{u}|_{W^{\ell,p}(\widetilde{\omega}_K)}.$$

Scaling back, we obtain

$$\sum_{j=0}^{\ell} h_K^j |u - w|_{W^{j,p}(\omega_K)} \le C h_K^{\ell} |u|_{W^{\ell,p}(\omega_K)}. \tag{2.41}$$

Finally, for Dirichlet nodes a^i, we need a sharper estimate for $|\mathrm{N}_{i,K}(\Pi_{i,K}u)|$ than (2.38). Consider an element K with a Dirichlet node a^i. Then there is a face $F_i \subset \Gamma_D$ at the Dirichlet part of the boundary with $a^i \in F_i$, and an element $K_i \subset \omega_T$ with $F_i \subset \partial K_i$. By using the inverse inequality, the identity $u|_{F_i} = 0$, and the trace theorem, we get for $\ell \ge 1$

$$\begin{aligned}
|\mathrm{N}_{i,K}(\Pi_{i,K}u)| &= |\Pi_{i,K}u(a^i)| \ \le\ \|\Pi_{i,K}u\|_{L^\infty(F_i)} \\
&\le C|F_i|^{-1}\|\Pi_{i,K}u\|_{L^1(F_i)} \\
&= C|F_i|^{-1}\|u - \Pi_{i,K}u\|_{L^1(F_i)} \\
&\le C|K_i|^{-1/p} \sum_{j=0}^{\ell} h_{K_i}^j |u - \Pi_{i,K}u|_{W^{j,p}(K_i)}.
\end{aligned}$$

Since $K_i \subset \omega_{i,K}$ and $\omega_{i,K}$ is isotropic, since $\Pi_{i,K}$ is bounded in $W^{j,p}(\omega_{i,K})$, and since $\Pi_{i,K}$ preserves polynomials $w \in \mathbb{P}_{\ell-1}$, we obtain in analogy to (2.41)

$$\begin{aligned}
\sum_{j=0}^{\ell} h_{K_i}^j |u - \Pi_{i,K}u|_{W^{j,p}(K_i)} &= \sum_{j=0}^{\ell} h_{K_i}^j |(u-w) - \Pi_{i,K}(u-w)|_{W^{j,p}(K_i)} \\
&\le C\sum_{j=0}^{\ell} h_K^j |u - w|_{W^{j,p}(\omega_K)} \\
&\le C h_K^{\ell} |u|_{W^{\ell,p}(\omega_K)}.
\end{aligned}$$

Thus, we have

$$|\mathrm{N}_{i,K}(\Pi_{i,K}u)| \le C|K|^{-1/p} h_K^{\ell} |u|_{W^{\ell,p}(\omega_K)}. \tag{2.42}$$

Note that this estimate holds also for $\ell = 0$ because of (2.38).

With these prerequisites, we obtain the final result.

Theorem 2.21. *Let $\mathcal{T}$ be an isotropic triangulation. Assume that each element $(K, \mathcal{P}_K, \mathcal{N}_K)$ is affine equivalent to $(\widehat{K}, \mathcal{P}_{\widehat{K}}, \mathcal{N}_{\widehat{K}})$ with $\mathbb{P}_{\ell-1} \subset \mathcal{P}_{\widehat{K}}$ and $\mathcal{N}_{\mathcal{T}} = \{\mathrm{N}_{i,K}\}_{i=1}^{N_+}$, where $\mathrm{N}_{i,K}(u) = u(a^i)$ and a^i are nodes. Let $u \in W^{\ell,p}(\omega_K)$, ω_K from (2.10), $\ell \ge 0$ for the Clément interpolant and $\ell \ge 1$ for the*

Scott-Zhang interpolant, $p \in [1,\infty]$. The numbers $m \in \{0,\dots,\ell-1\}$ and $q \in [1,\infty]$ are chosen such that $W^{\ell,p}(K) \hookrightarrow W^{m,q}(K)$. Then the estimate

$$|u - \mathrm{Q}_K u|_{W^{m,q}(K)} \le C|K|^{1/q-1/p} h_K^{\ell-m} |u|_{W^{\ell,p}(\omega_K)}$$

holds.

Proof. Consider first the case that $\mathrm{Q}_\mathcal{T}$ is the Scott-Zhang operator or, if $\mathrm{Q}_\mathcal{T}$ is the Clément operator, that K does not touch the Dirichlet part of the boundary. With (2.37), (2.39), (2.40), we obtain for $w \in \mathbb{P}_{\ell-1}$ from (2.41)

$$\begin{aligned} |u - \mathrm{Q}_K u|_{W^{m,q}(K)} &= |(u-w) - \mathrm{Q}_K(u-w)|_{W^{m,q}(K)} \\ &\le |u-w|_{W^{m,q}(K)} + |\mathrm{Q}_K(u-w)|_{W^{m,q}(K)} \\ &\le C|K|^{1/q-1/p} h_K^{-m} \sum_{j=0}^{\ell} h_K^j |u-w|_{W^{j,p}(\omega_K)}. \end{aligned}$$

By using (2.41), we obtain the desired result.

In the remaining case, we consider the Clément operator and an element K with nodes a^i, $i = 1,\dots,n$, where the nodes with $i = \widetilde{n}+1,\dots,n$ are at the boundary. Then we write

$$|u - \mathrm{Q}_K u|_{W^{m,q}(K)} \le \left| u - \sum_{i=1}^{n} \mathrm{N}_{i,K}(\Pi_{i,K} u)\phi_{i,K} \right|_{W^{m,q}(K)} + \sum_{i=\widetilde{n}+1}^{n} |\mathrm{N}_{i,K}(\Pi_{i,K} u)|\, |\phi_{i,K}|_{W^{m,q}(K)}.$$

The first term at the right-hand side has just been estimated; the remaining terms are bounded by (2.42) and $|\phi_{i,K}|_{W^{m,q}(K)} \le C|K|^{1/q} h_K^{-m}$. Consequently, the assertion is also proved in this case. □

Remark 2.22. *The proof given above extends to shape-regular quadrilateral elements (see Example 1.3), where $F_K(\cdot) \in (\mathbb{Q}_1)^2$ when $m \le 1$. In this case, the relations (2.37), (2.39), (2.40), and (2.41) hold as well. In the same way, we can treat the Clément operator for hexahedral elements K with $F_K(\cdot) \in (\mathbb{Q}_1)^3$. The Scott-Zhang operator can also be treated if all faces are planar. For elements with curved faces, one can use a projection operator Π_i on a reference configuration $\widehat{\omega}$ (see e.g. Bernardi (1989)).* ■

Remark 2.23. *For error estimates of the quasi-interpolants on anisotropic meshes, we refer to Apel (1999b) and Apel, Lombardi, and Winkler (2014). The main results are the following:*

- *In a suitable coordinate system (compare Remark 2.19), an anisotropic version of Theorem 2.21 holds for $m = 0$. This is obtained by a proper scaling.*
- *An example shows that both quasi-interpolation operators are not suited for deriving anisotropic error estimates in the sense of (2.35) if $m \ge 1$.*
- *Modifications of the Scott-Zhang operators have been suggested such that error estimates of type (2.35) can be obtained under certain assumptions on the mesh.* ■

2.5. Example for a Global Interpolation Error Estimate

The effectiveness of numerical methods for differential and integral equations depends on the choice of the mesh. Since singularities due to the geometry of the domain are known a priori, it is advantageous to adapt the finite element mesh $\mathcal{T}$ to these singularities. In this section, we define such meshes for a class of singularities and estimate the global interpolation error in the (broken) norm of the Sobolev space $W^{m,p}(\Omega)$. Such an error estimate is used in the estimation of the discretization error of various finite element methods. Note that this generality of the norm includes estimates in $L^2(\Omega)$, $L^\infty(\Omega)$, and $W^{1,2}(\Omega)$.

Assumption 2.24. *Let $\Omega \subset \mathbb{R}^2$ be a two-dimensional polygonal domain with corners c_j, $j = 1, \dots, J$. The solution u of an elliptic boundary value problem has, in general, singularities near the corners c_j, that is, the solution can be represented by*

$$u = u_0 + \sum_{j=1}^{J} u_j$$

with a regular part

$$u_0 \in W^{\ell,p}(\Omega) \tag{2.43}$$

and singular parts (corner singularities) u_j satisfying

$$|D^\alpha u_j| \le C r_j^{\lambda_j - |\alpha|} \quad \forall \alpha \colon |\alpha| \le \ell, \tag{2.44}$$

where $r_j = r_j(x) := \operatorname{dist}(x, c_j)$. The integer ℓ and the real numbers p and λ_j, $j = 1, \dots, J$, are defined by the data.

This assumption is realistic for a large class of elliptic problems, including those for the Poisson equation, the Lamé system, and the biharmonic equation. The numbers λ_j depend on the geometry of Ω (in particular on the internal angles at c_j), the differential operator, and the boundary conditions. For problems with mixed boundary conditions, there are, in general, further singular terms, but since they can also be characterized by (2.44), this poses no extra difficulty for the forthcoming analysis.

We remark that there can be terms that are best described by

$$|D^\alpha u_j| \le C r_j^{\lambda_j - |\alpha|} |\ln r_j|^{\beta_j}.$$

These terms can be treated either by a slight modification of the forthcoming analysis or by decreasing the exponent in (2.44) slightly. Note that $|\ln r_j|^{\beta_j} \le C_\varepsilon r^{-\varepsilon}$ for all $\varepsilon > 0$.

Assumption 2.25. *Let $\mathcal{T}$ be a finite element mesh, which is described by parameters h and $\mu_j \in (0, 1]$, $j = 1, \dots, J$. We assume that the diameter h_K of the element $K \in \mathcal{T}$ relates to the distances $r_{K,j} := \operatorname{dist}(K, c_j)$, $j = 1, \dots, J$, according to*

$$\begin{aligned} h_K &\le C h^{1/\mu_j} && \text{if } r_{K,j} = 0, \\ h_K &\le C h r_{K,j}^{1-\mu_j} && \text{if } r_{K,j} > 0 \quad \forall j = 1, \dots, J. \end{aligned} \tag{2.45}$$

For the purpose of interpolation, isotropic and anisotropic elements are admitted.

This assumption can be satisfied when isotropic elements are used, and when the elements in the neighborhoods U_j of the corners c_j satisfy

$$\begin{aligned} C_1 h^{1/\mu_j} \le h_K \le C_2 h^{1/\mu_j} \quad & \text{if } r_{K,j} = 0, \\ C_1 h r_{K,j}^{1-\mu_j} \le h_K \le C_2 h r_{K,j}^{1-\mu_j} \quad & \text{if } r_{K,j} > 0, \end{aligned} \tag{2.46}$$

$j = 1, \dots, J$, and when $C_1 h \le h_K \le C_2 h$ for all other elements. The size of the neighborhoods U_j of c_j should be independent of h but small enough such that $c_i \notin U_j$ for $i \ne j$. In this case one can prove that the number of elements of such meshes is of the order h^{-2}. To do this it is sufficient to show that the number of elements $K \subset U_j$ with $c_j \notin K$ is bounded by Ch^{-2}. By using $\int_K 1 = C_K h_K^2$ and the relations for h_K and r_K, we get

$$\begin{aligned}\sum_{K\subset U_j, c_j\notin K} 1 &= \sum_{K\subset U_j, c_j\notin K} C_K^{-1} h_K^{-2} \int_K 1 \\ &\le Ch^{-2} \sum_{K\subset U_j, c_j\notin K} r_{K,j}^{-2(1-\mu_j)} \int_K 1 \\ &\le Ch^{-2} \sum_{K\subset U_j, c_j\notin K} \int_K r_j^{-2(1-\mu_j)} \\ &\le Ch^{-2} \int_{U_j} r_j^{-2(1-\mu_j)} \le Ch^{-2},\end{aligned}$$

since $\mu_j > 0$.

Finally, we remark that meshes with property (2.46) can be created in different ways. If the neighborhood U_j is a circular sector of radius R_j, one can just move the nodes of a uniform mesh according to the coordinate transformation

$$\frac{r_j}{R_j} \mapsto \left(\frac{r_j}{R_j}\right)^{1/\mu_j}$$

(see e.g. Raugel (1978); Oganesyan and Rukhovets (1979); or Apel and Milde (1996)). A second possibility is to start with a uniform mesh of mesh size h and to split all elements recursively until (2.46) is satisfied; see Fritzsch (1990) or Apel and Milde (1996).

Assumption 2.26. *Let $V_\mathcal{T}$ be a finite element space corresponding to the triangulation $\mathcal{T}$. Let $\mathrm{I}_\mathcal{T} : W^{\ell,p} \to V_\mathcal{T}$ be the corresponding nodal interpolation operator satisfying $(\mathrm{I}_\mathcal{T} u)|_K = \mathrm{I}_K(u|_K)$ for all $K \in \mathcal{T}$. We assume that it permits the local interpolation error estimate*

$$|u - \mathrm{I}_K u|_{W^{m,p}(K)} \le C h_K^{\ell-m} |u|_{W^{\ell,p}(K)} \tag{2.47}$$

with ℓ, p from (2.43), and some $m \in \{0, \ldots, \ell-1\}$.

Note that this assumption relates the regularity of u_0 to the polynomial degree. The estimate (2.47) is proved in Theorem 2.9 only if $\mathbb{P}_{\ell-1} \subset \mathcal{P}_{\widehat{K}}$. So, if the regularity is low, then a large polynomial degree does not pay; if the polynomial degree is too low, then the regularity (2.43)-(2.44) is not fully exploited.

Theorem 2.27. *Let the function u, the mesh $\mathcal{T}$, and the interpolation operator $\mathrm{I}_\mathcal{T}$ satisfy Assumptions 2.24, 2.25, and 2.26 respectively. Then the error estimate*

$$\left(\sum_{K\in\mathcal{T}} |u - \mathrm{I}_K u|^p_{W^{m,p}(K)}\right)^{1/p} \le Ch^{\ell-m} \tag{2.48}$$

holds if $m \le \ell$ and, for all $j = 1, \ldots, J$,

$$m - \frac{2}{p} \le \lambda_j, \tag{2.49}$$

$$\mu_j < \frac{\lambda_j - m + \frac{2}{p}}{\ell - m} \quad \text{if} \quad \lambda_j < \ell - \frac{2}{p}, \tag{2.50}$$

$$\mu_j \le 1 \quad \text{if} \quad \lambda_j \ge \ell - \frac{2}{p}. \tag{2.51}$$

Note that condition (2.49) restricts m and p in such a way that $u_j \in W^{m,p}(\Omega)$ is ensured, since

$$|u_j|^p_{W^{m,p}(\Omega)} = \sum_{|\alpha|=m} \int_\Omega |D^\alpha u_j| \le \int_\Omega r^{(\lambda_j - m)p} < \infty$$

if $(\lambda_j - m)p > -2$. With this argument, we see also that $u_j \in W^{\ell,p}(\Omega)$ if $\lambda_j \ge \ell - 2/p$, that is, the function u_j is as regular as u_0 in this case, and no refinement ($\mu_j = 1$) is necessary in U_j.

The left-hand side of (2.48) is formulated with this broken Sobolev norm in order to cover the case that $\mathrm{I}_\mathcal{T}\, u \notin W^{m,p}(\Omega)$. Important applications are discretization error estimates for nonconforming finite element methods. If $\mathrm{I}_\mathcal{T}\, u \in W^{m,p}(\Omega)$, then estimate (2.48) can be written in the form

$$|u - \mathrm{I}_K u|_{W^{m,p}(\Omega)} \le Ch^{\ell-m}.$$

Proof. Consider the neighborhood U_j of c_j with $j \in \{1,\dots,J\}$ arbitrary but fixed. By Assumption 2.24, we have

$$u_i \in W^{\ell,p}(U_j), \quad i = 0,\dots,J, \quad i \ne j$$

and, therefore,

$$\begin{aligned}\left(\sum_{K\in U_j} |u_i - \mathrm{I}_K u_i|^p_{W^{m,p}(K)}\right)^{1/p} &\le \left(\sum_{K\in U_j} Ch_K^{(\ell-m)p} |u_i|^p_{W^{\ell,p}(K)}\right)^{1/p} \\ &\le Ch^{\ell-m}, \qquad i = 0,\dots,J, \quad i \ne j. \end{aligned} \tag{2.52}$$

Since the same argument can be applied for $\Omega \setminus \bigcup_{j=1}^J U_j$, it remains to be shown that (2.52) holds also for $i = j$. Note that we can assume that $\lambda_j < \ell - (2/p)$, since, otherwise, $u_j \in W^{\ell,p}(\Omega)$ and no refinement is necessary.

If $c_j \in K$, we estimate the interpolant simply by $\|\mathrm{I}_K u_j\|_{L^\infty(K)} \le C\|u_j\|_{L^\infty(K)} \le Ch_K^{\lambda_j}$. Using the triangle inequality, a direct computation of $|u_j|_{W^{m,p}(K)}$, and the inverse inequality for $|\mathrm{I}_K u_j|_{W^{m,p}(K)}$ (if $m > 0$, the case $m = 0$ is even direct), we get

$$\begin{aligned}|u_j - \mathrm{I}_K u_j|_{W^{m,p}(K)} &\le |u_j|_{W^{m,p}(K)} + |\mathrm{I}_K u_j|_{W^{m,p}(K)} \\ &\le C\left(\int_K r_j^{(\lambda_j - m)p}\right)^{1/p} + Ch_K^{-m} |K|^{1/p} \|\mathrm{I}_K u_j\|_{L^\infty(K)} \\ &\le Ch_K^{\lambda_j - m} |K|^{1/p} \le Ch^{(\lambda_j - m + 2/p)/\mu_j} \le Ch^{\ell - m},\end{aligned}$$

where we have used the inequalities $|K| \le Ch_K^2$, (2.44), (2.45), and (2.50). Note that the number of elements K with $c_j \in K$ is bounded by a constant that does not depend on h. So we get

$$\left(\sum_{K\in U_j, c_j \in K} |u_j - \mathrm{I}_K u_j|^p_{W^{m,p}(K)}\right)^{1/p} \le Ch^{\ell-m}. \tag{2.53}$$

Consider now an element $K \in U_j$ with $c_j \notin K$, that is, with $r_{K,j} > 0$. Then, $u_j \in W^{\ell,p}(K)$ and we can use the interpolation error estimate (2.47). So we get

$$\begin{aligned}|u_j - \mathrm{I}_K u_j|_{W^{m,p}(K)} &\leq Ch_K^{\ell-m}|u_j|_{W^{\ell,p}(K)} \\ &\leq Ch^{\ell-m} r_{j,K}^{(\ell-m)(1-\mu_j)} \left(\int_K r_j^{(\lambda_j-\ell)p}\right)^{1/p} \\ &\leq Ch^{\ell-m}\left(\int_K r_j^{[\lambda_j-\ell+(\ell-m)(1-\mu_j)]p}\right)^{1/p} \\ &= Ch^{\ell-m}\left(\int_K r_j^{[\lambda_j-m-\mu_j(\ell-m)]p}\right)^{1/p},\end{aligned}$$

since $r_{j,K} \leq r_j$ in K. Hence,

$$\left(\sum_{K\in U_j, c_j\notin K} |u_j - \mathrm{I}_K u_j|^p_{W^{m,p}(K)}\right)^{1/p} \leq Ch^{\ell-m}\left(\int_{U_j} r_j^{[\lambda_j-m-\mu_j(\ell-m)]p}\right)^{1/p}.$$

The integral on the right-hand side is finite if $[\lambda_j - m - \mu_j(\ell - m)]p > -2$, which is equivalent to (2.50). With (2.52) and (2.53), we conclude (2.48). □

Remark 2.28. *The given proof is an improved version of a proof for a more specific function u in a paper by Fritzsch and Oswald (1988). In this paper, the authors also address the question of the optimal choice of* μ_j*. They obtain for* $\mu_j = [\lambda_j - m + (2/p)]/[\ell - m + (2/p)]$ *the equidistribution of the element-wise interpolation error in the sense* $|r_j^{\lambda_j} - \mathrm{I}_K r_j^{\lambda_j}|_{W^{m,p}(K)} \approx const.$ ▪

Remark 2.29. *The given proof has the advantage that it needs minimal knowledge from functional analysis. A more powerful approach is to use weighted Sobolev spaces (see Remark 2.12 on page 14 for an example as well as the procedure in Section 3.1.3). The solutions of elliptic boundary value problems are often described by analysts in terms of different versions of such spaces; see, for example, the monographs by Grisvard (1985), Kufner and Sändig (1987), Dauge (1988), Nazarov and Plamenevsky (1994), or Kozlov, Maz'ya, and Roßmann (2001). For local and global interpolation error estimates for functions of such spaces see, for example, Grisvard (1985), Apel, Sändig, and Whiteman (1996), Apel and Nicaise (1998), Apel, Nicaise, and Schöberl (2001), Băcuţă, Nistor, and Zikatanov (2007) or Apel, Lombardi, and Winkler (2014). The advantage is that this approach covers the three-dimensional case, whereas Assumption 2.24 is too simple to cover edge singularities.* ▪

3. p-Version Approximation

In the h-version discussed so far, the underlying idea is to approximate (locally) by fixed degree polynomials. In the p or hp-version, the polynomial degree is not fixed but allowed to tend to infinity. Possibly the most striking feature of the p-version is the ability to achieve very rapid and even exponential convergence for analytic and, on suitably selected meshes, piecewise analytic functions. This setting is very often encountered in computational mechanics.

A hallmark of the h-version techniques is the Deny-Lions lemma discussed in Section 2.2. That is, norm equivalences of finite-dimensional spaces on the reference element are exploited and powers of h are obtained by scaling arguments. In the p-version, this is fundamentally different as the dimension of the approximation spaces on the reference element is not fixed. Thus, different tools have to be used.

Good references concerning the approximation properties of p and hp-FEM are the book by Schwab (1998) and the survey article by Babuška and Suri (1994). A wealth of techniques and results is available in the very closely related spectral element method discussed in the chapter **Spectral Methods** of **ECM2** and in the survey by Bernardi and Maday (1997). Excellent expositions emphasizing the approximation theoretic viewpoint and aspects are the books by DeVore and Lorentz (1993) and Davis (1974). The highly recommendable book by Trefethen (2013) emphasizes the power of Chebyshev polynomials and how to use them algorithmically.

In this chapter, we discuss mostly the single-element case, i.e., the polynomial approximation on the reference element; only the final Section 3.4 touches on H^1-conforming approximations on meshes with several elements. We do not discuss the choice of the polynomial basis, for which we refer to the chapter **The p-Version of the Finite Element Method** of **ECM2** and the books by Schwab (1998), Karniadakis and Sherwin (1999), and Demkowicz, Kurtz, Pardo, Paszyński, Rachowicz, and Zdunek (2008).

3.1. Approximation in 1D

Approximation by high order polynomials has a long history. For example, Taylor expansions of a function f converge rapidly if the domain of holomorphy of f is sufficiently large. But also functions that not as "nice" can be approximated by polynomials. Indeed, Weierstrass proved in 1885 that the space of polynomials is dense in $C(I)$ for $I = [-1, 1]$.

3.1.1. Stability of Lagrange Interpolation Processes. A natural idea to approximate a function u is by polynomial interpolation. Care has to be taken regarding the interpolation points. The classical Runge example shows that interpolation in equidistant nodes may diverge even for analytic functions whose domain of analyticity is not sufficiently large (cf., e.g., (Trefethen, 2013), Sec. 13 for a detailed discussion of this phenomenon). In order to ensure convergence of the interpolation process for a large, practically relevant class of functions, the interpolation points need to cluster suitably near the endpoints of the interval $I = [-1, 1]$ under consideration here. Classical examples include the Chebyshev points (i.e., the zeros of the Chebyshev polynomials T_n), the Gauss points (the zeros of the Legendre polynomials L_n) and the Gauss-Lobatto points (the extrema of the Legendre polynomials including the endpoints ± 1, i.e., the zeros of $x \mapsto (1 - x^2)L_n'(x)$). The quality of the interpolation process is usually measured in terms of the L^∞-stability constant of the interpolation operator, the L^∞-*Lebesgue constant* Λ_k^∞. That is, given $k + 1$ distinct nodes x_i, $i = 0, \dots, k$, they determine the interpolation operator $\mathrm{I}_k : C(I) \to \mathbb{P}_k$, and the corresponding Lebesgue constant Λ_k^∞ is

$$\Lambda_k^\infty := \sup_{0 \neq f \in C(I)} \frac{\|\mathrm{I}_k f\|_{L^\infty(I)}}{\|f\|_{L^\infty(I)}}.$$

Since $\mathrm{I}_k v = v$ for all polynomials $v \in \mathbb{P}_k$, it is an easy application of the triangle inequality to relate the interpolation error to the best approximation error:

$$\|f - \mathrm{I}_k f\|_{L^\infty(I)} \leq (1 + \Lambda_k^\infty) \inf_{v \in \mathbb{P}_k} \|f - v\|_{L^\infty(I)}. \tag{3.1}$$

Theorem 3.1 (L^∞-stability of Lagrange interpolation) *Let $I = [-1, 1]$. There exists a constant $C > 0$ independent of $k \geq 1$ such that the following holds:*

(i) For any node set $\{x_i\}_{i=0}^k \subset I$ one has $\Lambda_k^\infty \geq \frac{2}{\pi} \ln(k + 1)$. Furthermore, for any sequence $(\{x_i^k\}_{i=0}^k)_{k=1}^\infty$ of interpolation points, there exists a function u such that the sequence $(\mathrm{I}_k u)_{k=1}^\infty$ of polynomial interpolants diverges almost everywhere.

(ii) The uniform node distribution $\{x_i\}_{i=0}^k$ *with* $x_i = -1 + i/(2k)$ *satisfies* $\Lambda_k^\infty \sim 2^k/(\mathrm{e}k\ln k)$ *as* $k \to \infty$.

(iii) For the Chebyshev points $\{x_i^C\}_{i=0}^k$*, there holds* $\Lambda_k^\infty \le 1 + \frac{2}{\pi}\ln(k+1)$.

(iv) For the Gauss-Lobatto points $\{x_i^{GL}\}_{i=0}^k$*, there holds* $\Lambda_k^\infty \le C\ln(k+1)$.

(v) For the Gauss points $\{x_i^G\}_{i=0}^k$*, there holds* $\Lambda_k^\infty \le C\sqrt{k+1}$.

Proof. See (Trefethen, 2013), Thm. 15.2 for (i), (iii). The sharp asymptotics of (ii) are due to Turetskii (1940). Item (iv) is shown in (Sündermann, 1980), and (v) is from (Szegő, 1975), p. 338. We finally refer to reader to the survey by Brutman (1997) for an extensive discussion of Lebesgue constants. □

Theorem 3.1, (i) shows that the logarithmic factor in the stability bound is unavoidable in an L^∞-setting. In practice, its presence is acceptable for interpolation in the Chebyshev or Gauss-Lobatto points. Stability that is uniform in k may be available in other norms:

Theorem 3.2 (L^2- and H^1-stability) *Denote by* I_k^{GL} *and* I_k^G *the interpolation in Gauss-Lobatto and Gauss points, respectively. Then there exists a* $C > 0$ *independent of* k *such that*

$$\|\mathrm{I}_k^{GL}u\|_{H^1(I)} \le C\|u\|_{H^1(I)} \qquad \forall u \in H^1(I), \tag{3.2}$$

$$\|\mathrm{I}_k^{GL}u\|_{L^2(I)} + \|\mathrm{I}_k^{G}u\|_{L^2(I)} \le C\left[\|u\|_{L^2(I)} + k^{-1/2}\|u\|_{L^2(I)}^{1/2}\|u'\|_{L^2(I)}^{1/2}\right] \qquad \forall u \in H^1(I), \tag{3.3}$$

$$\|\mathrm{I}_k^{GL}u\|_{L^2(I)} + \|\mathrm{I}_k^{G}u\|_{L^2(I)} \le C(1 + k'/k)\|u\|_{L^2(I)} \qquad \forall u \in \mathbb{P}_{k'}. \tag{3.4}$$

Proof. Estimate (3.2) is shown in (13.27) of (Bernardi and Maday, 1997). The estimate (3.3) follows from an appropriate modification of the procedure in Thms. 13.1, 13.4 of (Bernardi and Maday, 1997), the details of which are worked out below in Theorem A.2.2 (cf. eqn. (A.2.14)) and Lemma A.2.3 (cf. eqn. (A.2.25)). Finally, (3.4) is taken from Rems. 13.1, 13.5 of (Bernardi and Maday, 1997). □

Remark 3.3. *The bound (3.3) implies the following estimate with the Besov space* $B_{2,1}^{1/2}(I) = (L^2(I), H^1(I))_{1/2,1}$:

$$\|\mathrm{I}_k^{G}\|_{L^2(I)} + \|\mathrm{I}_k^{GL}u\|_{L^2(I)} \le C\left[\|u\|_{L^2(I)} + k^{-1/2}\|u\|_{B_{2,1}^{1/2}(I)}\right]; \tag{3.5}$$

the details of the argument are worked out in Theorem A.2.2 (cf. (A.2.15)) and Lemma A.2.3 (cf. (A.2.26)).

While estimate (3.2) states H^1*-stability for* I_k^{GL}*, it is shown in eqn. (13.14) of (Bernardi and Maday, 1997) that* I_k^G *is not* H^1*-stable.* ■

3.1.2. Exponential Convergence. A key feature of polynomial approximation with $k \to \infty$ is the possibility of exponential convergence. This is the case when real analytic functions are approximated. A function $f \in C^\infty(a,b)$ is said to be real analytic on an (open) interval $(a,b) \subset \mathbb{R}$ if for each $x_0 \in (a,b)$, the Taylor series of f around x_0 converges to f in some neighborhood of x_0. Equivalently, f is real analytic if $f = F|_{(a,b)}$ for a function F that is holomorphic in neighborhood $\mathcal{G} \subset \mathbb{C}$ of (a,b). Corresponding to these two characterizations are two types of techniques to show exponential convergence in polynomial approximation, which we call *complex variables* and *real variables techniques.*

Complex Variables Techniques. We denote by $\mathcal{E}_\rho$, $\rho > 1$, the Bernstein ellipse

$$\mathcal{E}_\rho := \{z \in \mathbb{C}\colon |z-1| + |z+1| < \rho + \rho^{-1}\}. \tag{3.6}$$

Theorem 3.4. *Let $\rho > 1$ and let f be holomorphic on $\mathcal{E}_\rho$. Then for every $1 < \rho_1 < \rho$ we have*

$$\inf_{v \in \mathbb{P}_k} \|f - v\|_{L^\infty(-1,1)} \leq \frac{2}{\rho_1 - 1} \rho_1^{-k} \|f\|_{L^\infty(\mathcal{E}_{\rho_1})}.$$

If additionally $f|_{[-1,1]}$ is real-valued, then the polynomials in the infimum may be assumed to be real-valued on $\mathbb{R}$. Furthermore, for every $r \in \mathbb{N}$ and every $1 < \rho_1 < \rho$ there exists $C > 0$ such that

$$\inf_{v \in \mathbb{P}_k} \|f - v\|_{W^{r,\infty}(I)} \leq C \rho_1^{-k} \|f\|_{L^\infty(\mathcal{E}_{\rho_1})}.$$

Proof. Let $(T_n)_{n=0}^\infty$ be the Chebyshev polynomials. We approximate f by its truncated Chebyshev expansion $\Pi_k^C f$ and therefore have to show for the coefficients $(a_n)_{n=0}^\infty$ of the Chebyshev expansion $f(x) = \sum_{n=0}^\infty a_n T_n(x)$ an exponential decay; more precisely, we have to establish

$$|a_n| \leq \rho_1^{-n} \|f\|_{L^\infty(\mathcal{E}_{\rho_1})} \qquad \forall 1 < \rho_1 < \rho. \tag{3.7}$$

Step 1: The map $w \mapsto \Omega(w) := \frac{1}{2}(w + w^{-1})$ maps the annulus $A_\rho := \{w \in \mathbb{C}\colon \rho^{-1} < |w| < \rho\}$ conformally onto $\mathcal{E}_\rho$. In fact, Ω is a bijection between $\mathbb{C} \setminus B_1(0)$ and $\mathbb{C} \setminus [-1,1]$. For $\rho > 1$, both circles $\partial B_\rho(0)$ and $\partial B_{1/\rho}(0)$ are mapped to the ellipse $\partial \mathcal{E}_\rho$ (see, e.g., (Davis, 1974), Sec. 1.13).
Step 2: We consider the Chebyshev-transformed function $\tilde{f}(\theta) := f(\cos\theta)$. It is continuously differentiable, 2π periodic, and even. Hence, it can be represented as a Fourier series

$$\tilde{f}(\theta) = \frac{1}{2} a_0 + \sum_{n=1}^\infty a_n \cos n\theta, \qquad a_n = \frac{1}{\pi} \int_{-\pi}^{\pi} \tilde{f}(\theta) \cos n\theta \, d\theta. \tag{3.8}$$

We check that the integral for a_n is the parametrization of the following complex integral:

$$a_n = \frac{1}{\pi \mathbf{i}} \oint_{|w|=1} f\Big(\frac{w + w^{-1}}{2}\Big) \frac{w^n + w^{-n}}{2} \frac{1}{w} \, dw. \tag{3.9}$$

Step 3: From the first step, we have $\Omega(A_{\rho_1}) = \mathcal{E}_{\rho_1}$. The function $w \mapsto f(\frac{1}{2}(w + w^{-1})) = f(\Omega(w))$ is holomorphic on the annulus A_{ρ_1}, and we may use Cauchy's integral theorem to write

$$a_n = \frac{1}{\pi \mathbf{i}} \oint_{|w|=\rho_1} f\Big(\frac{w + w^{-1}}{2}\Big) \frac{w^{-n}}{2} \frac{1}{w} \, dw + \frac{1}{\pi \mathbf{i}} \oint_{|w|=1/\rho_1} f\Big(\frac{w + w^{-1}}{2}\Big) \frac{w^n}{2} \frac{1}{w} \, dw \tag{3.10}$$

Hence, we can bound

$$|a_n| \leq \rho_1^{-n} \|f\|_{L^\infty(\mathcal{E}_{\rho_1})} + \rho_1^{-n} \|f\|_{L^\infty(\mathcal{E}_{\rho_1})} \leq 2 \rho_1^{-n} \|f\|_{L^\infty(\mathcal{E}_{\rho_1})}.$$

Finally, using the fact that for the Chebyshev polynomials $|T_n(x)| \leq 1$ for all $x \in [-1,1]$, we get for $x \in [-1,1]$

$$\left| f(x) - \sum_{n=0}^k a_n T_n(x) \right| \leq \sum_{n=k+1}^\infty |a_n| \leq 2 \|f\|_{L^\infty(\mathcal{E}_{\rho_1})} \sum_{n=k+1}^\infty \rho_1^{-n} = \frac{2}{\rho_1 - 1} \|f\|_{L^\infty(\mathcal{E}_{\rho_1})} \rho_1^{-k}.$$

The construction shows that if f is real valued on $[-1,1]$, then the coefficients a_n are real. Finally, the estimate in the $W^{r,\infty}$-norm follows by approximating $f^{(r)}$ by a polynomial of degree $k - r$ and integrating r times. ∎

Remark 3.5. *The polynomial approximation constructed in Theorem 3.4 is the truncated Chebyshev expansion* Π_k^C. *Indeed, when expressing the coefficients* a_n *in (3.8) in terms of* f *instead of the transplanted function* $\tilde{f}$, *one has* $a_n = \pi^{-1}\int_{-1}^{1} f(x)T_n(x)(1-x^2)^{-1/2}\,dx$. *Exponential convergence results similar to Theorem 3.4 can be obtained by truncating other expansions, e.g., the Legendre expansion, see (Davis, 1974), Sec. 12.4.* ▪

The approximation of Theorem 3.4 is the truncated Chebyshev expansion. Complex variables techniques can also be used to estimate the error incurred by interpolating in the Chebyshev points:

Remark 3.6. ***(Hermite interpolation formula)*** *For holomorphic functions the Hermite interpolation formula (3.11) provides a very powerful tool to obtain error bounds when interpolating holomorphic functions. We refer to Thm. 3.6.1 of (Davis, 1974) or to Chap. 11 of (Trefethen, 2013) for details. For* $k+1$ *distinct interpolation points* $x_i \in \mathbb{C}$, $i = 0,\ldots,k$, *and any simple closed contour* $\mathcal{C}$ *that encircles* z *and the interpolation points* x_i, *the Hermite formula asserts the error representation*

$$f(z) - \mathrm{I}_k f(z) = \frac{1}{2\pi\mathbf{i}} \oint_{\mathcal{C}} \frac{f(\zeta)}{\zeta - z} \frac{\omega_k(z)}{\omega_k(\zeta)}\, d\zeta, \qquad \omega_k(z) = \prod_{i=0}^{k} (z - x_i). \tag{3.11}$$

As an application of (3.11) and an illustration of how knowledge of the asymptotic distribution of the interpolation points can be used, we show the exponential convergence of the interpolation in the Chebyshev points. We follow (Davis, 1974), Sec. 4.4. For the Chebyshev points x_i, $i = 0,\ldots,k$, *we have* $\omega_k(z) = 2^k T_{k+1}(z)$. *Let* $\Omega : \mathbb{C}\setminus B_1(0) \to \mathbb{C}\setminus[-1,1]$ *be the conformal map of step 1 of the proof of Theorem 3.4. A calculation reveals* $T_k(\Omega(w)) = 1/2(w^k + w^{-k})$. *For* $z \in \mathcal{E}_\rho$ *we have* $w = \Omega^{-1}(z) \in \partial B_\rho(0)$ *so that* $\lim_{k\to\infty} |T_k(z)|^{1/k} = \rho$ *and that, in fact, this convergence is uniform on* $\partial\mathcal{E}_\rho$. *Fix* $1 < \rho_1 < \rho$, *take the contour* $\mathcal{C}$ *in (3.11) as* $\partial\mathcal{E}_\rho$ *and consider* $z \in \partial\mathcal{E}_{\rho_1}$. *The above observations show* $\lim_{k\to\infty}(|\omega_k(z)/\omega_k(\zeta)|)^{1/k} = \rho_1/\rho < 1$ *uniformly for* $\zeta \in \partial\mathcal{E}_\rho$ *and* $z \in \partial\mathcal{E}_{\rho_1}$. *Inserting this in (3.11), one can infer an exponential convergence result for functions* f *that are holomorphic in* $\mathcal{G} \subset \mathbb{C}$ *with* $\mathrm{closure}(\mathcal{E}_\rho) \subset \mathcal{G}$. ▪

The convergence rate in the exponential convergence result in Theorem 3.4 depends on the size of the domain of holomorphy of the function f to be approximated. Geometric considerations show that for every domain $\mathcal{G} \subset \mathbb{C}$ with $[-1,1] \subset \mathcal{G}$, one can find $\rho > 1$ such that $[-1,1] \subset \mathcal{E}_\rho \subset \mathcal{G}$ so that exponential convergence in polynomial approximation is given if f is holomorphic on $\mathcal{G}$.

It may be of interest to check whether a function $f \in C^\infty(-1,1)$ has a holomorphic extension to some domain $\mathcal{G} \subset \mathbb{C}$. The following lemma answers that question:

Lemma 3.7. *Let* $f \in C^\infty(-1,1)$. *Then statements (i) and (ii) are equivalent:*

(i) There exist C_f, $\gamma_f > 0$ *such that* $\|f^{(n)}\|_{L^2(-1,1)} \le C_f \gamma_f^n n! \qquad \forall n \in \mathbb{N}_0$.

(ii) There exists $\rho > 1$ *such that* f *has a holomorphic extension to* $\mathcal{E}_\rho$ *with* $\|f\|_{L^\infty(\mathcal{E}_\rho)} < \infty$.

Proof. (ii) $\Longrightarrow$ (i) follows from the Cauchy integral theorem for derivatives. To see the implication (i) $\Longrightarrow$ (ii), we use the 1D Sobolev embedding theorem $C([-1,1]) \subset H^1(-1,1)$ to assert the existence of C, $\widetilde{\gamma}_f$ with $\|f^{(n)}\|_{L^\infty(-1,1)} \le C C_f \widetilde{\gamma}_f^n n!$ for all $n \in \mathbb{N}_0$. Hence, the Taylor series of f about each point $x \in [-1,1]$ converges on a ball of radius $1/\widetilde{\gamma}_f$. An appeal to (Börm, Löhndorf, and Melenk, 2005), Lemma 3.14 concludes the proof. □

As a corollary, we have from the combination of Lemma 3.7 and Theorem 3.4:

Corollary 3.8. *Let r, $k' \in \mathbb{N}_0$ and let the function $f \in C^\infty(-1,1)$ satisfy for some C_f, $\gamma_f > 0$ the bounds*

$$\|f^{(n)}\|_{L^2(-1,1)} \leq C_f \gamma_f^n n! \qquad \forall n \in \mathbb{N}_0, \quad n \geq k'+1.$$

Then there exist C, $b > 0$ depending solely on γ_f such that for $k \geq k'$ one has

$$\inf_{v \in \mathbb{P}_k} \|f - v\|_{W^{r,\infty}(-1,1)} \leq C C_f \mathrm{e}^{-bk}.$$

Real Variables Techniques. The decay rate of the coefficients $(a_n)_{n=0}^\infty$ of the Chebyshev expansion can also be estimated from (3.8) if f is not analytic. The basic mechanism to exploit the regularity of $\widetilde{f}$ (which is related to that of f) is integrating by parts. One can then show for both the truncated Chebyshev expansion and the Chebyshev interpolation error the following:

Theorem 3.9 ((Trefethen, 2013), Thm. 7.2) *Let $f \in W^{r,\infty}([-1,1])$ for some $r \geq 2$. Denote by $\Pi_k^C f$ the truncated Chebyshev expansion of f and by $\mathrm{I}_k^C f$ the Chebyshev interpolant. Then for every $k \geq r$*

$$\|f - \Pi_k^C f\|_{L^\infty(-1,1)} \leq \frac{2}{\pi r (k-r+1)^{r-1}} \|f^{(r)}\|_{L^\infty(-1,1)},$$
$$\|f - \mathrm{I}_k^C f\|_{L^\infty(-1,1)} \leq \frac{4}{\pi r (k-r+1)^{r-1}} \|f^{(r)}\|_{L^\infty(-1,1)}.$$

An interesting corollary to Theorem 3.9 is obtained for $f \in C^\infty(-1,1)$ that are not analytic but in Gevrey classes. Such functions can be approximated by polynomials at superalgebraic ("root exponential") rates:

Corollary 3.10 (approximation of functions of Gevrey classes) *Let C_f, $\gamma_f > 0$, $\alpha \geq 1$. Let $f \in C^\infty(-1,1)$ satisfy $\|f^{(n)}\|_{L^\infty(-1,1)} \leq C_f \gamma_f^n n^{n\alpha}$ for all $n \in \mathbb{N}_0$. Then there are constants C, $b > 0$ depending only on γ_f and α such that*

$$\inf_{v \in \mathbb{P}_k} \|f - v\|_{L^\infty(-1,1)} \leq C C_f \mathrm{e}^{-bk^{1/\alpha}}.$$

Proof. Theorem 3.4 covers the case $\alpha = 1$. Hence assume $\alpha > 1$. Note that Theorem 3.9 is applicable with any $r \in \mathbb{N}$. We fix $\beta \in (0,1)$ with $\gamma_f \beta^\alpha < 1$ and apply Theorem 3.9 with $r = \lfloor \beta k^{1/\alpha} \rfloor + 1$. ∎

3.1.3. The Geometric Mesh. In computational mechanics, the functions to be approximated (e.g., solutions of PDEs, geometries that need to be approximated as they cannot be realized exactly) are often not analytic but piecewise analytic. The regularity of such functions can suitably be described in terms of *countably normed spaces*, (Babuška and Guo, 1988); (Costabel, Dauge, and Nicaise, 2012). A fairly general strategy to approximate such piecewise analytic functions by high order polynomials is to employ a suitably graded mesh, namely, the *geometric mesh*, see Example 3.11. The key features can already be seen in 1D for the approximation of the function $u(x) = x^\alpha$ by piecewise polynomials in the H^1-norm say, which is worked out in detail by Babuška and Suri (1994). Here, we consider a slightly more general case by studying for $\Omega = (0,1)$ the approximation of a function $u \in C^\infty(\Omega) \cap H^1(\Omega)$ that satisfies for some $\beta \in (0,1)$ and constants C_u, $\gamma_u > 0$

$$\|u\|_{H^1(\Omega)} \leq C_u, \qquad \|x^{\beta+n} u^{(n+2)}\|_{L^2(\Omega)} \leq C_u \gamma_u^n n! \qquad \forall n \in \mathbb{N}_0. \tag{3.12}$$

It can be checked that for $\alpha \in (1/2, 1)$ and $\beta > 3/2 - \alpha$ the function $x \mapsto x^\alpha$ satisfies (3.12) for some constants C_u, γ_u. We emphasize that the geometric mesh idea is successful for the approximation of elliptic boundary value problems in 2D polygonal domains, (Babuška and Guo, 1986), and in 3D polyhedral domains, (Schötzau, Schwab, and Wihler, 2013), where, however, *anisotropic refinement* is a key new ingredient over the 2D case. We refer to the chapter **The p-Version of the Finite Element Method** of **ECM2** and (Babuška and Suri, 1994) for examples illustrating the success of the geometric mesh idea.

We use meshes $\mathcal{T}_n = \{K_i\}_{i=1}^n$ with n elements on the domain $\Omega = (0,1)$. With each element K_i, we associate a polynomial degree $k_i \in \mathbb{N}$. Since we consider below approximation in $H^1(\Omega)$, we define the approximation space $S^{\mathbf{k},1}(\mathcal{T}_n) := \{u \in H^1(\Omega)\colon u|_{K_i} \in \mathbb{P}_{k_i},\ i = 1, \ldots, n\}$. The characterizing features of the *geometric mesh* that is refined towards a singularity, which we take to be at the origin, are:

- The element at the singularity is small, and this is exploited ("h-FEM").
- All elements $K \in \mathcal{T}_n$ with $0 \notin K$ satisfy $\operatorname{diam} K \sim \operatorname{dist}(K, 0)$.

The proof of Theorem 3.13 will clarify how these two ingredients are used. The archetypal geometric mesh is given in the following example.

Example 3.11. *For a parameter $\sigma \in (0,1)$ let the grid points x_i of the mesh $\mathcal{T}_n$ be given by*

$$x_0 = 0, \quad x_i = \sigma^{n-i}, \quad i = 1, \ldots, n; \tag{3.13}$$

correspondingly, the elements are $K_i = (x_{i-1}, x_i)$, $i = 1, \ldots, n$. The element diameters are $h_i := |K_i| = x_i - x_{i-1}$. The parameter σ is called the grading factor*, and the parameter n is referred to as the number of* layers *of geometric refinement. We note that the condition $\operatorname{diam} K_i \sim \operatorname{dist}(K_i, 0)$ for all elements K_i with $0 \notin K_i$ is satisfied. In fact*

$$h_1 = \sigma^{n-1}, \qquad h_i = \sigma^{n-i}(1-\sigma) = \sigma^{n-(i-1)}(1/\sigma - 1), \tag{3.14}$$

$$\frac{\operatorname{diam} K_i}{\operatorname{dist}(K_i, 0)} = 1/\sigma - 1, \qquad i \geq 2. \tag{3.15}$$

■

We will analyze below the *linear degree distribution* in conjunction with geometric meshes:

Example 3.12. *Let $\mathcal{T}_n$ be a the geometric mesh of Example 3.11. For a parameter $s > 0$, we set*

$$k_i := 1 + \lfloor s(i-1) \rfloor, \qquad i = 1, \ldots, n.$$

The parameter s is called the slope *of the linear degree distribution.* ■

Theorem 3.13. *Let $u \in C^\infty(\Omega) \cap H^1(\Omega)$ satisfy (3.12) for some $\beta \in [0,1)$. Consider the geometric mesh of Example 3.11 with grading factor $\sigma \in (0,1)$ and the linear degree vector of Example 3.12 with slope $s > 0$. Then there exist constants C, $b > 0$ (depending only on γ_u, σ, β, s) such that the space $S^{\mathbf{k},1}(\mathcal{T}_n)$ has the following approximation properties: There exists $v \in S^{\mathbf{k},1}(\mathcal{T}_n)$ such that*

$$\|u - v\|^2_{H^1(\Omega)} \leq C C_u \left[\mathrm{e}^{-bn} + \sigma^{(1-\beta)n} \right], \tag{3.16}$$

$$u(x_i) = v(x_i) \qquad i = 0, \ldots, n. \tag{3.17}$$

Furthermore, $\dim S^{\mathbf{k},1}(\mathcal{T}_n) \sim n^2$.

Proof. The approximation $v \in S^{\mathbf{k},1}(\mathcal{T}_n)$ is defined using the Gauss-Lobatto interpolation operator I_k^{GL}. (Operators other than I_k^{GL} could equally well be used, e.g., the operator Π_k^{BS} from (3.25)). Denote by $I = [-1,1]$ the reference element and by $F_{K_i} : I \to K_i$ an affine bijection between I and K_i. Define $v \in S^{\mathbf{k},1}(\mathcal{T}_n)$ elementwise by

$$v|_{K_i} := (\mathrm{I}_{k_i}^{GL}(u \circ F_{K_i})) \circ F_{K_i}^{-1}$$

Note in particular that $v(x_i) = u(x_i)$ for $i = 0, \dots, n$ since the endpoints ± 1 of I are interpolation nodes in the Gauss-Lobatto interpolation scheme. Also note that v coincides with the linear interpolant of u in the element K_1 since $k_1 = 1$. The analysis distinguishes between the element K_1 and the remaining elements K_i, $i \geq 2$.

Step 1: We claim that on the first element K_1, the linear interpolant $\mathrm{I}_1^{GL} u$ satisfies

$$\|u - \mathrm{I}_1^{GL} u\|_{L^2(K_1)} \leq C h_1^{2-\beta}, \qquad \|(u - \mathrm{I}_1^{GL} u)'\|_{L^2(K_1)} \leq C h_1^{1-\beta}.$$

Such an estimate is obtained in the standard h-FEM fashion by a scaling argument, if the estimate

$$\|\widehat{u} - \mathrm{I}_1^{GL} \widehat{u}\|_{H^1(I)} \leq C \|(1+x)^\beta \widehat{u}''\|_{L^2(I)} \tag{3.18}$$

is proved. Estimate (3.18) is a consequence of a variant of the Deny-Lions lemma as in (2.24), (2.25). Alternatively, a direct proof starts from the observation that Rolle's theorem gives the existence of $\xi \in I$ with $(\widehat{u} - \mathrm{I}_1^{GL} \widehat{u})'(\xi) = 0$. Hence, for $x \in I$ we get

$$(\widehat{u} - \mathrm{I}_1^{GL} \widehat{u})'(x) = \int_\xi^x \widehat{u}''(t)\, dt = \int_\xi^x (1+t)^{-\beta} (1+t)^\beta \widehat{u}''(t)\, dt.$$

Applying the Cauchy-Schwarz inequality to the integral (note that $\beta \in (0,1)$) and then integrating in x gives $\|(\widehat{u} - \mathrm{I}_1^{GL} \widehat{u})'\|_{L^2(I)} \leq C \|(1+x)^\beta \widehat{u}''\|_{L^2(I)}$. Since $(\widehat{u} - \mathrm{I}_1^{GL} \widehat{u}) \in H_0^1(I)$, the Poincaré inequality in fact gives (3.18).

Step 2: Consider an element K_i, $i = 2, \dots, n$. We define the pull-back $\widehat{u} := u|_{K_i} \circ F_{K_i}$ and observe, since $\mathrm{I}_{k_i}^{GL}$ is an H^1-stable projector (Theorem 3.2),

$$\|\widehat{u} - \mathrm{I}_{k_i}^{GL} \widehat{u}\|_{H^1(I)} \leq C \inf_{v \in \mathbb{P}_{k_i}} \|\widehat{u} - v\|_{H^1(I)}. \tag{3.19}$$

In order to estimate this infimum, we calculate $\widehat{u}^{(n)}$ for $n \geq 2$, bearing in mind that the element map F_{K_i} is affine with $F_{K_i}' = \frac{h_i}{2}$ and $K_i = [x_{i-1}, x_i]$:

$$\begin{aligned}
\|\widehat{u}^{(n)}\|_{L^2(I)} &= (h_i/2)^{n-1/2} \|u^{(n)}\|_{L^2(K_i)} \leq (h_i/2)^{n-1/2} x_{i-1}^{-\beta-(n-2)} \|x^{\beta+(n-2)} u^{((n-2)+2)}\|_{L^2(K_i)} \\
&\leq \left(\frac{\operatorname{diam} K_i}{2 \operatorname{dist}(K_i, 0)}\right)^{n-1/2} (\operatorname{dist}(K_i, 0))^{-\beta+3/2} \|x^{\beta+(n-2)} u^{((n-2)+2)}\|_{L^2(K_i)} \\
&\leq C C_u h_i^{3/2-\beta} \left(\gamma_u (1/\sigma - 1)/2\right)^{n-2} (n-2)!,
\end{aligned}$$

where $C > 0$ depends only on σ, γ_u, and β; here, we employed that (3.12). By Corollary 3.8, we can therefore conclude together with (3.19)

$$\|\widehat{u} - \mathrm{I}_{k_i}^{GL} \widehat{u}\|_{H^1(I)} \quad \leq \quad C C_u h_i^{3/2-\beta} \mathrm{e}^{-b k_i},$$

where the constants C, $b > 0$ are *independent* of the element K_i (they depend only on γ_u, σ, β). Pushing forward to the element K_i, we get that the polynomial $v_{K_i} := (\mathrm{I}_{k_i}^{GL} \widehat{u}) \circ F_{K_i}^{-1} \in \mathbb{P}_{k_i}$ satisfies $u(x_{i-1}) = v(x_{i-1})$ and $u(x_i) = v(x_i)$ together with

$$\begin{aligned}
\|u - v_{K_i}\|_{L^2(K_i)} &= \left(\frac{h_i}{2}\right)^{1/2} \|\widehat{u} - \mathrm{I}_{k_i}^{GL} \widehat{u}\|_{L^2(I)}, \\
\|(u - v_{K_i})'\|_{L^2(K_i)} &= \left(\frac{h_{K_i}}{2}\right)^{-1/2} \|(\widehat{u} - \mathrm{I}_{k_i}^{GL} \widehat{u})'\|_{L^2(I)}.
\end{aligned}$$

We arrive at

$$\|u-v\|_{L^2(K_i)} \le Ch_i^{2-\beta}\mathrm{e}^{-bk_i}, \qquad \|(u-v)'\|_{L^2(K_i)} \le Ch_i^{1-\beta}\mathrm{e}^{-bk_i}.$$

The error $u-v$ is assessed by summing over all elements K_i, $i=1,\dots,n$:

$$\|u-v\|^2_{H^1(\Omega)} \quad \le \quad Ch_1^{2(1-\beta)} + C\sum_{i=2}^{n} h_i^{2-2\beta}\mathrm{e}^{-2bk_i}. \tag{3.20}$$

We recall that $1-\beta>0$. Using $h_i \le \sigma^{n-i}$ and $k_i = \lfloor 1+s(i-1)\rfloor \ge s(i-1)$ one can check by elementary techniques that $i \mapsto h_i^{2(1-\beta)}\mathrm{e}^{-2bk_i}$ is convex so that

$$\max_{i=1,\dots,n} h_i^{2-2\beta}\mathrm{e}^{-2bk_i} \le C\mathrm{e}^{-\lambda n}, \qquad \lambda := \min\{-2(1-\beta)\ln\sigma, 2bs\} > 0. \tag{3.21}$$

Inserting (3.21) into (3.20) gives $\|u-v\|^2_{H^1(\Omega)} \le C\sigma^{2(1-\beta)n} + Cn\mathrm{e}^{-\lambda n}$. Selecting $b\in(0,\lambda)$ and observing $\sup_{n>0} ne^{-n(b-\lambda)} < \infty$ allows us to conclude the proof. □

Remark 3.14. *Theorem 3.13 asserts exponential convergence for* all *grading factors $\sigma \in (0,1)$ and* all *slopes $s>0$, but the constant $b>0$ depends on σ and s. The optimal choice of the grading factor σ for the approximation of singularity functions of the form $x \mapsto x^\alpha$ is $\sigma = (\sqrt{2}-1)^2 \approx 0.17$; see (Babuška and Suri, 1994), Thm. 2.6 and (Scherer, 1981) for details.* ■

3.1.4. Approximation by Truncated Series. Polynomial approximations are often generated from truncated series expansions or, more generally, quasi-interpolation operators. The reasons not to interpolate are manifold: the function may not be continuous (so that interpolation is not possible), interpolation operators may not have good *simultaneous* approximation or stability properties in scales of Sobolev norms, or the selection of the interpolation points may not be clear (while good choices in 1D and thus also for tensor product elements such as squares and hexahedra are available, their choice for triangles and tetrahedra is less clear).

Consider the L^2-projection $\Pi_k^{L^2} : L^2(I) \to \mathbb{P}_k$. Since the Legendre polynomials $(L_n)_{n=0}^\infty$ are orthogonal polynomials with respect to the $L^2(I)$-inner product, $\Pi_k^{L^2}$ is given explicitly as a truncated Legendre expansion, i.e.,

$$\Pi_k^{L^2} u = \sum_{n=0}^{k} u_n L_n,$$

where $u \in L^2(I)$ is written as

$$u(x) = \sum_{n=0}^{\infty} u_n L_n(x), \qquad u_n = \frac{2n+1}{2}\int_I L_n(x)u(x)\,dx; \tag{3.22}$$

here, the Legendre polynomials are normalized by the condition $L_n(1)=1$. The error is

$$\inf_{v\in\mathbb{P}_k} \|u-v\|^2_{L^2(-1,1)} = \|u-\Pi_k^{L^2}u\|^2_{L^2(-1,1)} = \sum_{n=k+1}^{\infty} \frac{2}{2n+1}|u_n|^2.$$

Approximation properties of $\Pi_k^{L^2}$ can thus be inferred from decay properties of the expansion coefficients $(u_n)_{n=0}^\infty$. The decay of the coefficients is closely related to regularity properties of the function u. In fact, weighted Sobolev spaces appear naturally through properties of orthogonal polynomials as we now show. We start by noting that the derivatives of orthogonal polynomials are again orthogonal polynomials, but with respect to a different weight: As shown in eqn. (4.21.7) of (Szegő, 1975) we have for the Jacobi polynomials $P_n^{(\alpha,\beta)}$ the relationship

$$\frac{d}{dx}P_n^{(\alpha,\beta)} = \frac{1}{2}(n+\alpha+\beta+1)P_{n-1}^{(\alpha+1,\beta+1)}.$$

This can be used to show (recall that $P_n^{(0,0)} = L_n$ and see, e.g., (Schwab, 1998), Lemma 3.10 for details) that for the Legendre expansion $u = \sum_{n=0}^{\infty} u_n L_n$ one has

$$|u|^2_{V_r^r(-1,1)} := \int_{-1}^{1} (1-x^2)^r |u^{(r)}|^2\, dx = \sum_{n=r}^{\infty} \frac{2}{2n+1} \frac{(n+r)!}{(n-r)!} |u_n|^2. \tag{3.23}$$

Hence, the error incurred by truncating the Legendre series is for $k+1 \geq r$

$$\begin{aligned} \|u - \Pi_k^{L^2} u\|^2_{L^2(-1,1)} &= \sum_{n=k+1}^{\infty} \frac{2}{2n+1} |u_n|^2 = \sum_{n=k+1}^{\infty} \frac{(n-r)!}{(n+r)!} \frac{2}{2n+1} \frac{(n+r)!}{(n-r)!} |u_n|^2 \\ &\leq \frac{(k+1-r)!}{(k+1+r)!} \|(1-x^2)^{r/2} u^{(r)}\|^2_{L^2(-1,1)} \leq C k^{-2r} |u|^2_{V_r^r(-1,1)}, \end{aligned} \tag{3.24}$$

where we exploited asymptotic properties of the Γ-function (Stirling's formula) in the last step. We flag at this point already that the weight in the definition of $|\cdot|_{V_r^r}$ is precisely the weight that appears in the more general approximation result (3.35). Since obviously $|u|_{V_r^r(-1,1)} \leq |u|_{H^r(-1,1)}$, we see that Sobolev regularity of u implies convergence rates for the L^2-projection. The presence of the weight $(1-x^2)^{r/2}$ in $|\cdot|_{V_r^r(-1,1)}$, however, leads to even better convergence rates for certain types of functions, namely, those with singularities at the endpoints ± 1 of the interval I, as the following example shows:

Example 3.15. *Let $u(x) = (1+x)^\alpha$ for $\alpha \in (0,1)$. Then $u \in H^{\alpha+1/2-\varepsilon}(-1,1)$ for every $\varepsilon > 0$. It can be checked by direct calculations that the L^2-approximation error by piecewise constant functions on a quasi-uniform mesh with mesh size h is $O(h^{\alpha+1/2})$ and not better. We next claim that $\inf_{v \in \mathbb{P}_k} \|u - v\|_{L^2(-1,1)} \leq C k^{-2(\alpha+1/2)}$, which is twice the rate of convergence (error versus dimension of the approximation space) compared to the h-version case of piecewise constant approximation on quasi-uniform meshes. To see this, let $r := \lfloor 1 + 2\alpha \rfloor \in \mathbb{N}$ be the integer with $2\alpha < r \leq 2\alpha + 1$. Let $(\chi_\delta)_{\delta>0} \subset C^\infty(\mathbb{R})$ be a family of functions with a) $\operatorname{supp} \chi_\delta \subset [-1-2\delta, -1+2\delta]$ and b) $\chi_\delta \equiv 1$ on $[-1-\delta, -1+\delta]$, and c) $\|\chi_\delta^{(j)}\|_{L^\infty(\mathbb{R})} \leq C\delta^{-j}$ for $j = 0, \ldots, r$. We write $u = \chi_\delta u + (1-\chi_\delta) u$ and note that by the support properties of χ_δ we have $(1-\chi_\delta) u \in V_{r'}^{r'}$ for any r'. We estimate with (3.24)*

$$\begin{aligned} \inf_{v \in \mathbb{P}_k} \|u - v\|_{L^2(-1,1)} &= \inf_{v_1, v_2 \in \mathbb{P}_k} \|u - (v_1 + v_2)\|_{L^2(-1,1)} \\ &\leq \inf_{v \in \mathbb{P}_k} \|\chi_\delta u - v\|_{L^2(-1,1)} + \inf_{v \in \mathbb{P}_k} \|(1-\chi_\delta) u - v\|_{L^2(-1,1)} \\ &\leq C k^{-r} |\chi_\delta u|_{V_r^r} + C k^{-(r+1)} |(1-\chi_\delta) u|_{V_{r+1}^{r+1}} \\ &\leq C \left[k^{-r} \delta^{(\alpha - r/2)+1/2} + k^{-(r+1)} \delta^{\alpha - r/2} \right], \end{aligned}$$

where the last estimate exploited the properties of χ_δ as well as $2\alpha < r \leq 2\alpha + 1$. Selecting $\delta = k^{-2}$ gives the result. For similar approximation results in H^1 instead of L^2 and singular function of the form $x \mapsto (1+x)^\alpha (\log(1+x))^\beta$, see (Schwab, 1998), Thm. 3.26. For the analog in 2D, see (Babuška and Suri, 1987). ∎

Remark 3.16. *Polynomial approximations can also be constructed from truncating expansions in other systems of polynomials. Often, expansions in Jacobi polynomials $(P_n^{(\alpha,\beta)})_{n=0}^\infty$ are employed. The choice $\alpha = \beta = 0$ corresponds to the expansion in Legendre polynomials, the choice $\alpha = \beta = 1/2$ to the Chebyshev expansion, which we encountered already in Section 3.1.2.* ∎

Example 3.17. *The operator $\Pi_k^{L^2}$ can be used to design operators for the approximation in other norms, e.g., in H^1. The operator*

$$\Pi_k^{BS} u(x) := u(-1) + \int_{-1}^{x} (\Pi_{k-1}^{L^2} u')(t)\, dt, \tag{3.25}$$

which is often associated with the names of Babuška and Szabó, has optimal approximation properties in the H^1-seminorm. We note that by construction Π_k^{BS} has a commuting diagram property*: $(\Pi_k^{BS} u)' = \Pi_{k-1}^{L^2} u'$. Obviously, $u(-1) = (\Pi_N^{BS} u)(-1)$; furthermore, orthogonality properties of the Legendre polynomials give $u(1) = (\Pi_k^{BS} u)(1)$. We remark that both $\Pi_k^{L^2} : L^2(I) \to \mathbb{P}_k$ and $\Pi_k^{BS} : H^1(I) \to \mathbb{P}_k$ are projections. Since $(u - \Pi_k^{BS} u)(\pm 1) = 0$, the Poincaré inequality gives the stability estimate*

$$\|\Pi_k^{BS} u\|_{H^1(-1,1)} \leq C\|u\|_{H^1(-1,1)}.$$

In fact, a sharper stability estimate is available:

$$\|(1-x^2)^{-1/2}(u - \Pi_k^{BS} u)\|_{L^2(-1,1)} + \|u - \Pi_k^{BS} u\|_{H^1(-1,1)} \leq Ck^{-r}|u'|_{V_r^r}, \qquad k \geq r, \tag{3.26}$$

$$\|\Pi_k^{BS} u\|_{L^2(-1,1)} \leq C\left[\|u\|_{L^2(-1,1)} + k^{-1}\|u'\|_{L^2(-1,1)}\right]. \tag{3.27}$$

The starting point for the proof of (3.26) is that by the above developments we have the bound $|(u - \Pi_k^{BS} u)'\|_{L^2(-1,1)} \leq Ck^{-r}|u'|_{V_r^r}$. For the L^2-part of $\|u - \Pi_N^{BS} u\|_{H^1(-1,1)}$, properties of the Legendre polynomials (see the proof of (Schwab, 1998), Thm. 3.14 for details) yield the estimate $\|(1-x^2)^{-1/2}(u - \Pi_k^{BS} u)\|_{L^2(-1,1)} \leq Ck^{-1}\|(u - \Pi_k^{BS} u)'\|_{L^2(-1,1)}$, which concludes the proof of (3.26). The stability estimate (3.27) follows from (3.26) with $r = 0$. ∎

3.1.5. Quasi-Interpolation. A caveat about the L^2-projection $\Pi_k^{L^2}$ or, more generally, best approximation operator is that these operators may not be stable in other norms and thus do not produce the optimal rate in these norms (see, e.g., the discussion in Sec. 2.6 of (Babuška and Suri, 1994)). Then, they are not well-suited for the problem of simultaneous approximation. If approximation operators are constructed based on truncating polynomial expansions, a common device (which can be traced back at least to work by de la Vallée-Poussin) is not to truncate but to smoothly attenuate the coefficients. The following two examples present such operators:

Example 3.18. *(cf., e.g., (Bernardi and Maday, 1999)) Let $\chi \in W^{1,\infty}(\mathbb{R})$ with* $\operatorname{supp}\chi \subset [-2,2]$, *$0 \leq \chi \leq 1$, and $\chi \equiv 1$ on $[0,1]$. Define the attenuated truncated L^2-projection by*

$$\widetilde{\mathrm{Q}}_k u := \sum_{n=0}^{\infty} \chi(n/k) u_n L_n, \tag{3.28}$$

where u is written as $u = \sum_{n=0}^{\infty} u_n L_n$. The properties of χ readily imply $\widetilde{\mathrm{Q}}_k : L^2(I) \to \mathbb{P}_{2k}$ with $\|\widetilde{Q}_k u\|_{L^2(I)} \leq \|u\|_{L^2(I)}$ as well as $\widetilde{Q}_k u = u$ for all $u \in \mathbb{P}_k$. A non-trivial calculation (cf. Theorem A.3.1 for details) shows $\|\widetilde{Q}^k u\|_{H^1(I)} \leq C\|u\|_{H^1(I)}$ for a constant $C > 0$ depending solely on χ. ∎

The previous example produces an operator that maps into $\mathbb{P}_{2k}$. A modification from (Braess, Pillwein, and Schöberl, 2009) improves this in that the operators Q_k maps into $\mathbb{P}_k$ instead of $\mathbb{P}_{2k}$:

Example 3.19. *(cf. (Braess, Pillwein, and Schöberl, 2009)) Let χ be as in Example 3.18. Define the operator*

$$\mathrm{Q}_k u := \sum_{n=0}^{k} \chi(n/k) u_n L_n + \sum_{n=k+1}^{2k} \chi(n/k) u_n L_{2k-n}. \tag{3.29}$$

This operator satisfies $\mathrm{Q}_k : L^2(I) \to \mathbb{P}_k$, $\mathrm{Q}_k u = u$ for all $u \in \mathbb{P}_k$ and the stability bounds

$$\|\mathrm{Q}_k u\|_{L^2(I)} \leq C\|u\|_{L^2(I)}, \qquad \|\mathrm{Q}_k u\|_{H^1(I)} \leq C\|u\|_{H^1(I)}$$

for a $C > 0$ independent of k (cf. Theorem A.3.2 for details). ∎

We observe that the operators $\widetilde{\mathrm{Q}}_k$ and Q_k, being stable (both in L^2 and H^1) projections have simultaneous approximation properties in the norms L^2 and H^1. Such results are not restricted to L^2-based Sobolev spaces. The following variant of simultaneous approximation operators gives a flavor of what may be expected; it is formulated with seminorms on the right-hand side to emphasize that scaling arguments (which lead to powers of the local mesh size) are possible:

Lemma 3.20. *Let $R \in \mathbb{N}$ and $q \in [1,\infty]$. For every $k \in \mathbb{N}_0$ there exists a bounded linear operator $\mathrm{J}_{R,k} : L^1(I) \to \mathbb{P}_k$ and a constant $C > 0$, which depends only on R and q, such that for each r with $0 \le r \le R$*

$$\|u - \mathrm{J}_{R,k}u\|_{W^{j,q}(I)} \le C(k+1)^{-(r-j)}\|u\|_{W^{r,q}(I)}, \qquad j = 0,\ldots,r.$$

Furthermore, $\mathrm{J}_{R,k}$ can be constructed such that for $0 \le r \le R$ and $k \ge R-1$

$$\mathrm{J}_{R,k}u = u \qquad \forall u \in \mathbb{P}_{R-1},$$
$$\|u - \mathrm{J}_{R,k}u\|_{W^{j,q}(I)} \le C(k+1)^{-(r-j)}|u|_{W^{r,q}(I)}, \qquad j = 0,\ldots,r.$$

Proof. Such results can be found in (DeVore and Lorentz, 1993). The particular form is taken from (Melenk, 2005a), Prop. A.2. For a slightly different approach, see also (Karkulik and Melenk, 2015), Thm. 3.3. ∎

The simultaneous approximation result of Lemma 3.20 can be combined with the stability properties of various polynomial approximation operators:

Corollary 3.21. *Let $\widetilde{\mathrm{Q}}_k$, Q_k be the operators of Examples 3.18, 3.19, and let I_k^G, I_k^{GL} be the Gauss and Gauss-Lobatto interpolation operators, respectively. Then:*

$$\|u - \mathrm{Q}_k u\|_{H^r(I)} + \|u - \widetilde{\mathrm{Q}}_k u\|_{H^r(I)} \le C_{r,s}k^{-(s-r)}\|u\|_{H^s(I)}, \quad 0 \le r \le 1, \quad 0 \le r \le s, \tag{3.30}$$
$$\|u - \mathrm{I}_k^{GL} u\|_{H^r(I)} \le C_{r,s}k^{-(s-r)}\|u\|_{H^s(I)}, \quad 0 \le r \le 1, \quad s > (1+r)/2, \tag{3.31}$$
$$\|u - \mathrm{I}_k^{G} u\|_{L^2(I)} \le C_s k^{-s}\|u\|_{H^s(I)}, \qquad 1/2 < s. \tag{3.32}$$

The constants C_s, $C_{r,s}$ are independent of k.

Proof. To show (3.30), one notes that the projection and stability properties of Q_k and $\widetilde{\mathrm{Q}}_k$ give approximation properties with optimal rates in $L^2(I)$ and $H^1(I)$ for integer s. The extension to non-integer r and s follows from interpolation arguments.

We show (3.31) in the restricted setting $0 \le r \le 1$ and $s \ge 1$; see (Bernardi and Maday, 1997), Thm. 13.4 for the full proof. For arbitrary $v \in \mathbb{P}_k$ one has $\mathrm{I}_k^{GL}v = v$. Hence,

$$u - \mathrm{I}_k^{GL}u = (u - v) - \mathrm{I}_k^{GL}(u - v).$$

For the cases $r = 0$ and $r = 1$ as well as integer $s \ge 1$, it is then easy to combine the stability assertions of Theorem 3.2 with Lemma 3.20 to get (3.31). The extension to non-integer r and s follows by interpolation arguments, which are worked out in Lemma A.2.4. The estimate (3.32) follows by similar arguments (again, details can be found in Lemma A.2.4). We refer the reader to (Bernardi and Maday, 1997), p. 299 for a discussion of bounds for $\|u - \mathrm{I}_k^G u\|_{H^r(I)}$, $r \ge 0$. ∎

3.1.6. Inverse Estimates in 1D and Multi-d. A fundamental difference between the h-version and the p-version is the mismatch between approximation results (“direct estimates”,“Jackson estimates”) and inverse estimates (“Bernstein estimates”) in scales of standard (unweighted) Sobolev spaces. The classical Markov type inverse estimates for $q \in [1,\infty]$ are (see, e.g., (Bernardi, Dauge, and Maday, 2007), Chap. III, Props. 3.2, 3.3)

$$\|v'\|_{L^q(I)} \le Ck^2\|v\|_{L^q(I)} \qquad \forall v \in \mathbb{P}_k, \tag{3.33}$$

which do not match the approximation property

$$\inf_{v\in\mathbb{P}_k} \|u-v\|_{L^q(I)} \le Ck^{-1}\|u'\|_{L^q(I)} \qquad \forall u \in W^{1,q}(I). \tag{3.34}$$

The general situation of matching direct and inverse estimates in 1D is as follows: With the weight function $\phi(x) := \sqrt{1-x^2}$ we have for $q \in [1,\infty)$ and fixed $r \in \mathbb{N}$ (cf. (DeVore and Lorentz, 1993), Chap. 8, eqns. (7.1), (7.2))

$$\inf_{v\in\mathbb{P}_k} \|f-v\|_{L^q(I)} \le C_{q,r}k^{-r}\|\phi^r f^{(r)}\|_{L^q(I)}, \quad k \ge r, \tag{3.35}$$

$$\|\phi^r v^{(r)}\|_{L^q(I)} \le C_{q,r}k^r\|v\|_{L^q(I)} \qquad \forall v \in \mathbb{P}_k. \tag{3.36}$$

We note that, due to the presence of the weight ϕ, the function f needs less regularity than $f \in H^r(I)$ to yield approximation order r. We observed this phenomenon in Subsection 3.1.4 already for the special case of $L^2(I)$ and illustrated in Example 3.15 how this is responsible for the doubling of the convergence rate when approximating certain singular functions. The presence of the weight ϕ is closely related to delicate endpoint behavior of polynomial approximation, for which we refer the reader to (DeVore and Lorentz, 1993), Chap. 8.

Remark 3.22. *The combination of the direct estimate (3.35) and the inverse estimate (3.36) is at the root of characterizing regularity in terms of approximability by polynomials. Essentially, if a function u can be approximated at a certain (algebraic) rate by polynomials, then it is in a certain weighted Sobolev or Besov space; see (DeVore and Lorentz, 1993), Chap. 8, Thm. 7.7 for the precise statement.* ▪

Remark 3.23. ***(inverse estimates in multi-d)*** *Inverse estimates in unweighted Sobolev spaces take a form similar to* (3.33), *e.g., for* $q \in [1,\infty]$

$$\|\nabla v_k\|_{L^q(K)} \le Ck^2\|v_k\|_{L^q(K)} \qquad \forall v_k \in \mathbb{P}_k,$$

where $K \subset \mathbb{R}^d$ *is simplex or a hypercube. This and further estimates can be found in (Schwab, 1998) and (Bernardi, Dauge, and Maday, 2007). The 1D weighted inverse estimate* (3.36) *can be generalized to hypercubes, e.g., for the square* $S = [-1,1]^2$, *one can derive from the 1D estimate the bound*

$$\int_S (1-x^2)|\partial_x v_k|^2 + (1-y^2)|\partial_y v_k|^2 \,dx\,dy \le Ck^2\|v_k\|^2_{L^2(S)} \qquad \forall v_k \in \mathbb{Q}_k.$$

For the case of triangles, see (Braess and Schwab, 2000). ▪

3.2. Tensor Product Approximation

The simplest way to extend 1D approximation results to higher dimensions is by tensor product constructions. The approximation properties of the tensor product operator are then controlled in terms of both the univariate approximation properties and certain stability properties. We first illustrate the basic structure in an L^∞-setting in 2D:

Theorem 3.24. *Let* $S = I \times I = [-1,1]^2$. *Let* $\mathrm{I}_k : C(I) \to \mathbb{P}_k$ *be a (univariate) interpolation operator with* L^∞*-Lebesgue constant* Λ_k^∞. *Then the bivariate interpolation operator* $\mathrm{I}_k^{2D} := \mathrm{I}_k^x \circ \mathrm{I}_k^y : C(S) \to \mathbb{Q}_k$ *satisfies*

$$\|f - \mathrm{I}_k^{2D} f\|_{L^\infty(S)} \le \sup_{y\in I} \|f(\cdot,y) - \mathrm{I}_k^x f(\cdot,y)\|_{L^\infty(I)} + \Lambda_k^\infty \sup_{x\in I} \|f(x,\cdot) - \mathrm{I}_k^y f(x,\cdot)\|_{L^\infty(I)}. \tag{3.37}$$

In view of $\mathrm{I}_k^{2D} = \mathrm{I}_k^x \circ \mathrm{I}_k^y = \mathrm{I}_k^y \circ \mathrm{I}_k^x$*, the same estimate holds with the roles of* x *and* y *interchanged. Here, the superscripts* x *and* y *in* I_k^x *and* I_k^y *indicate the variable with respect to which the univariate interpolation operator* I_k *acts.*

Proof. We start by showing that $\mathrm{I}_k^x \circ \mathrm{I}_k^y = \mathrm{I}_k^y \circ \mathrm{I}_k^x$ indeed maps into $\mathbb{Q}_k$. To see this, we write the univariate operator I_k as $(\mathrm{I}_k f)(x) = \sum_{i=0}^k f(x_i)\ell_i(x)$, where $\{x_i\}_{i=0}^k$ are the interpolation nodes and the polynomials $\ell_i \in \mathbb{P}_k$ are the Lagrange interpolation polynomials for these points. Then $(\mathrm{I}_k^x \circ \mathrm{I}_k^y f)(x,y) = (\mathrm{I}_k^x(\sum_{i=0}^k f(\cdot, x_i)\ell_i(y)))(x) = \sum_{j=0}^k (\sum_{i=0}^k f(x_j, x_i)\ell_i(y))\ell_j(x)$.

For the error estimate, we compute

$$\begin{aligned}\|u - \mathrm{I}_k^{2D} u\|_{L^\infty(S)} &\le \|u - \mathrm{I}_k^x u\|_{L^\infty(S)} + \|\mathrm{I}_k^x(u - \mathrm{I}_k^y u)\|_{L^\infty(S)} \\ &\le \sup_{y\in I} \|u(\cdot, y) - \mathrm{I}_k^x u(\cdot, y)\|_{L^\infty(I)} + \|\mathrm{I}_k^x\| \sup_{x \in I} \|u(x,\cdot) - \mathrm{I}_k^y u(x, \cdot)\|_{L^\infty(I)}.\end{aligned}$$

Noting that $\|\mathrm{I}_k^x\| = \Lambda_k^\infty$ concludes the proof. □

The roles of (simultaneous) approximation and stability properties of the univariate operator I_k are even more clearly seen in an L^2- and an H^1-setting. The Gauss and Gauss-Lobatto interpolation operators I_k^G and I_k^{GL} satisfy the condition (3.38) of the following theorem by Theorem 3.2. The H^1-stability of the univariate operator required in the second part of Theorem 3.25 is true for the Gauss-Lobatto interpolation operator.

Theorem 3.25. *Let* $u \in H^s(S)$ *with* $s > 1$*. Assume that the projector* $\mathrm{I}_k : H^1(I) \to \mathbb{P}_k$ *has the following stability property:*

$$\|\mathrm{I}_k v\|_{L^2(I)} \le C\left[\|v\|_{L^2(I)} + k^{-1/2}\|v\|_{L^2(I)}^{1/2}\|v'\|_{L^2(I)}^{1/2}\right] \qquad \forall v \in H^1(I). \tag{3.38}$$

Then the tensor product approximation $\mathrm{I}_k^{2D} := \mathrm{I}_k^x \circ \mathrm{I}_k^y$ *satisfies*

$$\|u - \mathrm{I}_k^{2D} u\|_{L^2(S)} \le C k^{-s} \|u\|_{H^s(S)}. \tag{3.39}$$

If I_k *is (uniformly in* k*)* H^1*-stable, i.e.,* $\|\mathrm{I}_k u\|_{H^1(I)} \le C\|u\|_{H^1(I)}$ *for all* $u \in H^1(I)$*, then*

$$\|u - \mathrm{I}_k^{2D} u\|_{H^1(S)} \le C k^{-(s-1)} \|u\|_{H^s(S)}. \tag{3.40}$$

Proof. We restrict our presentation to the case of integer $s \in \mathbb{N}$ with $s \ge 2$ and refer for the general case to the interpolation arguments worked out for the case of Gauss- and Gauss-Lobatto interpolation in (Bernardi and Maday, 1997), Thm. 14.1. Inspection of the proof of Corollary 3.21 shows that it is the stability property (3.38) that implies the approximation property $\|v - \mathrm{I}_k v\|_{L^2(I)} \le C_t k^{-t}\|v\|_{H^t(I)}$ for $t \ge 1$. We estimate

$$\begin{aligned}&\|u - \mathrm{I}_k^x \circ \mathrm{I}_k^y u\|_{L^2(S)} \le \|u - \mathrm{I}_k^x u\|_{L^2(S)} + \|\mathrm{I}_k^x(u - \mathrm{I}_k^y u)\|_{L^2(S)} \\ &\le C\left[\|u - \mathrm{I}_k^x u\|_{L^2(S)} + \|u - \mathrm{I}_k^y u\|_{L^2(S)} + k^{-1/2}\|u - \mathrm{I}_k^y u\|_{L^2(S)}^{1/2}\|\partial_x(u - \mathrm{I}_k^y u)\|_{L^2(S)}^{1/2}\right] \\ &\le C\left[\|u - \mathrm{I}_k^x u\|_{L^2(S)} + \|u - \mathrm{I}_k^y u\|_{L^2(S)} + k^{-1/2}\|u - \mathrm{I}_k^y u\|_{L^2(S)}^{1/2}\|\partial_x u - \mathrm{I}_k^y \partial_x u\|_{L^2(S)}^{1/2}\right] \\ &\le C\left[k^{-s}\|\partial_x^s u\|_{L^2(S)} + k^{-s}\|\partial_y^s u\|_{L^2(S)} + k^{-1/2}k^{-s/2}\|\partial_x^s u\|_{L^2(S)}^{1/2} k^{-(s-1)/2}\|\partial_x \partial_y^{s-1} u\|_{L^2(S)}^{1/2}\right] \\ &\le C k^{-s}\|u\|_{H^s(S)}.\end{aligned}$$

For the H^1-estimate, we merely consider $\partial_y(u - \mathrm{I}_k u)$. The key step is that the assumed H^1-stability of the projection I_k brings about the additional approximation property $\|v - \mathrm{I}_k v\|_{H^1(I)} \le$

$C_t k^{-(t-1)}\|v\|_{H^t(I)}$ for $t \geq 1$. Reasoning similarly as in the L^2-case, we get

$$\begin{aligned}
&\|\partial_y(u - \mathrm{I}_k^x \circ \mathrm{I}_k^y u)\|_{L^2(S)} \leq \|\partial_y u - \mathrm{I}_k^x \partial_y u\|_{L^2(S)} + \|\mathrm{I}_k^x \partial_y(u - \mathrm{I}_k^y u)\|_{L^2(S)} \\
&\leq C\left[\|\partial_y u - \mathrm{I}_k^x \partial_y u\|_{L^2(S)} + \|\partial_y(u - \mathrm{I}_k^y u)\|_{L^2(S)} + k^{-1/2}\|\partial_y(u - \mathrm{I}_k^y u)\|_{L^2(S)}^{1/2}\|\partial_y(\partial_x u - \mathrm{I}_k^y \partial_x u)\|_{L^2(S)}^{1/2}\right] \\
&\leq C\Big[k^{-(s-1)}\|\partial_x^{s-1}\partial_y u\|_{L^2(S)} + k^{-(s-1)}\|\partial_y^s u\|_{L^2(S)} \\
&\qquad + k^{-1/2}k^{-(s-1)/2}\|\partial_y^s u\|_{L^2(S)}^{1/2} k^{-(s-2)/2}\|\partial_x \partial_y^{s-1} u\|_{L^2(S)}^{1/2}\Big] \\
&\leq C k^{-(s-1)}\|u\|_{H^s(S)},
\end{aligned}$$

which completes the proof. □

Remark 3.26. *The proof shows that not all partial derivatives are required for the tensor product argument but only certain combinations of mixed derivatives.*

Theorem 3.25 is formulated for the 2D case and requires $s > 1$. For d-dimensional hypercubes, $s > d/2$ is the necessary regularity requirement to ensure that the interpolation operator is meaningful, and then the same approximation results hold (this is worked out for the Gauss and Gauss-Lobatto interpolants in (Bernardi and Maday, 1997), Thms. 14.1, 14.2).

We presented the proof for $s \geq 2$ and could have then worked with the weaker stability estimate $\|\mathrm{I}_k u\|_{L^2(I)} \leq C\left[\|u\|_{L^2(I)} + k^{-1}\|u'\|_{L^2(I)}\right]$. Such a stability result also holds for the operator Π_k^{BS} by (3.27). The same error estimates as in (3.39), (3.40) hold then on hypercubes for Π_k^{BS}, albeit under the more restrictive regularity assumption $s \geq d$ (instead of $s > d/2$ as for Gauss- and Gauss-Lobatto interpolation). ■

The operators $\widetilde{\mathrm{Q}}_k$, Q_k, and $\mathrm{J}_{R,k}$ of Examples 3.18, 3.19 and Lemma 3.20 have both simultaneous approximation properties and stability properties. A tensor product argument therefore leads to the following result:

Lemma 3.27. *Let $S^d = [-1,1]^d$ and $s \geq 0$. There are linear operators $\Pi_k : L^2(S^d) \to \mathbb{Q}_k$ such that*

$$\|u - \Pi_k u\|_{H^t(S^d)} \leq C k^{-(s-t)}\|u\|_{H^s(S^d)}, \qquad 0 \leq t \leq s.$$

The constant $C > 0$ depends solely on s, t, and d.

Proof. The operator Π_k can be taken to be any of the tensor products of the operators $\widetilde{\mathrm{Q}}_k$, Q_k, or $\mathrm{J}_{R,k}$. The convergence result for integers s, t follows by arguments similar to the tensor product arguments above for $d = 2$. The extension to fractional Sobolev spaces follows by interpolation arguments. □

The following example illustrates how simultaneous stability properties can be useful to construct approximation operators that have the commuting diagram property; we refer to the chapter **Finite Element Methods for Maxwell Equations** of **ECM2** and (Demkowicz and Buffa, 2005); (Demkowicz, 2008) for more details.

Example 3.28. *(cf. (Braess, Pillwein, and Schöberl, 2009)) Let $\mathrm{Q}_{k+1} : L^2(I) \to \mathbb{P}_{k+1}$ be the operator of Example 3.19. Define the anti-derivative operator $\mathrm{A} : v \mapsto \int_{-1}^x v(t)\,dt$. Define the operator $\widehat{\mathrm{Q}}_k : L^2(I) \to \mathbb{P}_k$ by $v \mapsto (\mathrm{Q}_{k+1}\mathrm{A}v)'$. One has*

$$(\mathrm{Q}_{k+1}u)' = \widehat{\mathrm{Q}}_k u',$$

and the H^1-stability properties of Q_{k+1} imply that $\widehat{\mathrm{Q}}_k$ is stable in $L^2(I)$. Set $S = I \times I$. Define on the space $(L^2(S))^2$ the operator $\mathrm{Q}_k^\Sigma \mathbf{u} := (\mathrm{Q}_{k+1}^x \circ \widehat{\mathrm{Q}}_k^y u_x, \widehat{\mathrm{Q}}_k^x \circ \mathrm{Q}_{k+1}^y u_y)$, where the superscripts indicate *again the variable the operator acts on and* $\mathbf{u} = (u_x, u_y)^\top$. *Then, for* $\mathbf{u} \in H(\mathrm{div}, S)$ *we have*

$$\mathrm{div}\, \mathrm{Q}_k^\Sigma \mathbf{u} = (\widehat{\mathrm{Q}}_k^x \circ \widehat{\mathrm{Q}}_k^y)\, \mathrm{div}\, \mathbf{u},$$

and the operator Q_k^Σ *is stable in* $L^2(S)$ *as well as* $H(\mathrm{div}, S)$. ■

3.3. Triangles and Tetrahedra

Interpolation operators on triangles with good stability properties are rather hard to construct, although some have been proposed in the literature (see (Rapetti, Sommariva, and Vianello, 2012) for a recent review). We restrict our attention to approximation operators. Two approaches to the construction of polynomial approximants are frequently encountered:

1. The function u defined on the reference triangle (or, more generally, the reference simplex) T is extended to the reference square (or hypercube) S. The Stein extension operator, (Stein, 1970), Chap.VI.3 for example, has the mapping property $E : W^{r,p}(T) \to W^{r,p}(S)$ (with norm depending only on r, p, and the precise choice of T and S). The extended function Eu can be approximated from $\mathbb{Q}_k$ using, for example, tensor product operators. Note that $\mathbb{Q}_k \subset \mathbb{P}_{dk}$ so that by ultimately restricting to T, one obtains an approximation operator defined on T that maps into $\mathbb{P}_{dk}$. This process is particularly suited to construct operators for the approximation of functions with finite Sobolev regularity, leading to *algebraic* convergence rates in k.
2. Certain polynomial bases on the triangle are available, in particular an L^2-orthogonal basis associated with the names of Koornwinder (1975) or Dubiner (1991). For functions that are real analytic in a neighborhood of T, truncating the expansion in this orthogonal basis leads to *exponential* convergence results, see (Melenk, 2002), Sec. 3.2.3 and (Eibner and Melenk, 2007). A noteworthy feature of this basis and related ones such as those proposed by Karniadakis and Sherwin (1999) is the product structure of the basis after transformation to the reference square via the Duffy transformation; this can be exploited algorithmically (see, e.g., the chapter **Spectral Methods** of **ECM2** and Karniadakis and Sherwin (1999)).

3.4. The Multi-Element Approximation and the Extension/Lifting Problem

So far, we have studied the approximation on a single reference element, e.g., the triangle or the square. Quite often, e.g., in H^1-conforming FEMs, piecewise polynomial approximations have to be constructed on a mesh consisting of several/many elements, and the approximation is required to be *continuous* across element interfaces. A key ingredient for the construction of suitable approximation operators is a discrete lifting. We mention that the existence of such discrete liftings is of interest outside the realm of approximation theory, e.g., in the analysis of iterative solvers such as *iterative substructuring* discussed, for example, in the chapter **Domain Decomposition Methods and Preconditioning** of **ECM2** and in (Toselli and Widlund, 2005).

One basic building of the lifting operators is described in the following example, which goes back at least to Gagliardo (1957):

Example 3.29. *Let* $e = (-1, 1)$, *which we identify with* $(-1, 1) \times \{0\} \subset \mathbb{R}^2$. *Let* ρ_ε *be any (symmetric) mollifier with* $\mathrm{supp}\, \rho_\varepsilon \subset [-\varepsilon, \varepsilon]$ *and choose a parameter* $\alpha \in (0, 1)$. *Define for* $(x, y) \in T_\alpha := \{(x, y) \in \mathbb{R}^2 \colon |x| < 1, 0 < y < \alpha(1 - |x|)\}$ *the lifting operator*

$$(Eu)(x, y) := \int_{-1}^{1} \rho_{\alpha y}(x - t) u(t)\, dt.$$

Thus, E maps a function defined on $(-1,1)$ to a function defined on the triangle T_α. We have:

1. *It is a lifting, i.e., $\lim_{y\to 0}(Eu)(x,y) = u(x)$ if u is sufficiently smooth.*
2. *It is polynomial preserving: If $u \in \mathbb{P}_k$ (in the variable x), then $Eu \in \mathbb{P}_k$ (in the variable (x,y)).*
3. *It is linear and continuous in suitable norms. For example, $E : H^{1/2}(e) \to H^1(T_\alpha)$ (see, e.g., (Babuška, Craig, Mandel, and Pitkäranta, 1991)).*

The first two properties follow by inspection, the stability bound is non-trivial. ■

The operator of Example 3.29 is suitable for lifting from a single edge of the reference triangle. For the full strength of the following Theorem 3.30, one has to construct liftings for two and three edges simultaneously, for which we refer to (Babuška, Craig, Mandel, and Pitkäranta, 1991).

Theorem 3.30 (Babuška, Craig, Mandel, and Pitkäranta (1991)) *Let $T \subset \mathbb{R}^2$ be the reference triangle. Then there exists a bounded linear operator $E : H^{1/2}(\partial T) \to H^1(T)$ that is a lifting, i.e., $(Eu)|_{\partial T} = u$ for all $u \in H^{1/2}(\partial T)$. Additionally, E is polynomial preserving: for any $u \in C(\partial T)$ that is edgewise a polynomial of degree k, one has $Eu \in \mathbb{P}_k$.*

Remark 3.31. *A corresponding result also holds for the reference square. Extensions to 3D are available in (Muñoz-Sola, 1997) and (Ben Belgacem, 1994). Key observations in this directions have been made earlier by Maday (1989).* ■

We next illustrate one way to construct polynomial approximations for the case $d = 2$ and triangular elements that naturally lead to H^1-conforming approximations in a multi-element setting. In fact, the construction is such that, for each vertex and edge of the triangulation, the polynomial approximation is completely determined by the restriction of u to that vertex or edge, respectively. In other words, the approximation is defined element by element.

Theorem 3.32. *Let $T \subset \mathbb{R}^2$ be the reference triangle, $s > 1$. Then there is a linear operator $\Pi_k : H^s(T) \to \mathbb{P}_k$ such that*

1. *$\|u - \Pi_k u\|_{H^1(T)} \le Ck^{-(s-1)}\|u\|_{H^s(T)}$.*
2. *For each vertex V of T there holds $u(V) = (\Pi_k u)(V)$.*
3. *For each edge e of T there holds: $(\Pi_k u)|_e$ depends only on $u|_e$.*

Proof. In order to simplify some notation, we assume $1 < s < 3/2$. For edges e and $t \in (0,1)$, the fractional Sobolev spaces $H^t(e)$, $H^t_0(e)$, and $H^{1/2}_{00}(e)$ and some of their properties have to be employed. We refer to (Grisvard, 1985) for their definition.

Step 1: We fix $(\Pi_k u)|_{\partial T}$ by prescribing $(\Pi_k u)|_e$ for each edge e of T. Additionally, we will ensure that $\|u - \Pi_k u\|_{H^{1/2}(\partial T)} \le Ck^{-(s-1)}\|u\|_{H^s(T)}$. We present the arguments merely for one edge e, which we assume to be $(-1,1)\times\{0\}$ and which we identify with $(-1,1)$ when convenient. Let Π^e_k be a 1D approximation operator of the form given in Lemma 3.20. It satisfies

$$\|u - \Pi^e_k u\|_{H^t(e)} \le Ck^{-(s-1/2-t)}\|u\|_{H^{s-1/2}(e)} \le Ck^{-(s-1/2-t)}\|u\|_{H^s(T)}, \qquad 0 \le t \le s - 1/2, \quad (3.41)$$

where the last estimate follows from the trace inequality. We want to modify $\Pi_k^e u$ in such a way that the modified function coincides with u in the two endpoints of e. To that end, we note $H^{s-1/2}(e) \subset C(\overline{e})$ by the Sobolev embedding theorem; in fact the sharper multiplicative interpolation inequality $\|v\|_{L^\infty(e)} \leq C\|v\|_{L^2(e)}^{1-1/(2s-1)}\|v\|_{H^{s-1/2}(e)}^{1/(2s-1)}$ is valid (see (Canuto and Quarteroni, 1982), p. 85) so that we obtain additionally

$$\|u - \Pi_k^e u\|_{L^\infty(e)} \leq Ck^{-(s-1)}\|u\|_{H^s(T)}. \tag{3.42}$$

Recalling that we identify e with $(-1,1)$, we introduce the functions $\ell_{-1}(x) := 2^{-k}(1-x)^k$ and $\ell_1(x) := 2^{-k}(1+x)^k$. These functions satisfy

$$\|\ell_{\pm 1}\|_{H^t(e)} \leq C_t k^{-1/2+t}, \qquad t \geq 0; \tag{3.43}$$

for integer t, this follows by a direct calculation and for fractional t by an interpolation argument. We fix $(\Pi_k u)|_e$ by setting

$$(\Pi_k u)|_e(x) := \Pi_k^e(u|_e)(x) - ((\Pi_k^e u|_e)(-1) - u(-1,0))\ell_{-1}(x) - ((\Pi_k^e u|_e)(1) - u(1,0))\ell_1(x).$$

Combining (3.41), (3.42), (3.43) produces

$$\|u - \Pi_k u\|_{L^2(e)} \leq Ck^{-(s-1/2)}\|u\|_{H^s(T)}, \tag{3.44}$$

$$\|u - \Pi_k u\|_{H^{s-1/2}(e)} \leq C\|u\|_{H^s(T)}, \tag{3.45}$$

$$(u - \Pi_k u)(\pm 1, 0) = 0. \tag{3.46}$$

The combination of (3.45) and (3.46) implies in view of our assumption $1 < s < 3/2$ the stronger estimate $\|u - \Pi_k u\|_{H_0^{s-1/2}(e)} \leq C\|u\|_{H^s(T)}$. The interpolation inequality $\|v\|_{H_{00}^{1/2}(e)} \leq \|v\|_{L^2(e)}^{1-1/(2s-1)}\|v\|_{H_0^{s-1/2}(e)}^{1/(2s-1)}$ gives

$$\|u - \Pi_k^e u\|_{H_{00}^{1/2}(e)} \leq k^{-(s-1)}\|u\|_{H^s(T)}. \tag{3.47}$$

In this way, we define $(\Pi_k u)|_e$ for each edge e. That is, we have fixed the function $\widetilde{u}_k := (\Pi_k u)|_{\partial T}$. The bound (3.47) implies the key estimate

$$\|u - \widetilde{u}_k\|_{H^{1/2}(\partial T)} = \|u - \Pi_k u\|_{H^{1/2}(\partial T)} \leq Ck^{-(s-1)}\|u\|_{H^s(T)}. \tag{3.48}$$

Step 2: The previous step fixed $\widetilde{u}_k = (\Pi_k u)|_{\partial T}$. For the final approximation $\Pi_k u$ on T, let $\widetilde{\Pi}_k : H^1(T) \to \mathbb{P}_k$ be an approximation operator with the property

$$\|v - \widetilde{\Pi}_k v\|_{H^1(T)} \leq Ck^{-(s-1)}\|v\|_{H^s(T)}. \tag{3.49}$$

For the present case of a triangle, such an operator can, for example, be constructed as described at the beginning of Section 3.3. We finally set $\Pi_k u := \widetilde{\Pi}_k u - E(\widetilde{\Pi}_k u - \widetilde{u}_k)$, where $E : H^{1/2}(\partial T) \to H^1(T)$ is the lifting of Theorem 3.30. We note that $(\Pi_k u)|_{\partial T} = \widetilde{u}_k$ and

$$\begin{aligned}
\|u - \Pi_k u\|_{H^1(T)} &\leq \|u - \widetilde{\Pi}_k u\|_{H^1(T)} + \|E(\widetilde{\Pi}_k u - \widetilde{u}_k)\|_{H^1(T)} \\
&\leq C\left[\|u - \widetilde{\Pi}_k u\|_{H^1(T)} + \|\widetilde{\Pi}_k u - u\|_{H^{1/2}(\partial T)} + \|u - \widetilde{u}_k\|_{H^{1/2}(\partial T)}\right] \\
&\leq C\left[\|\widetilde{\Pi}_k u - u\|_{H^1(T)} + \|u - \widetilde{u}_k\|_{H^{1/2}(\partial T)}\right].
\end{aligned}$$

With (3.48) and (3.49) we obtain the desired error estimate. ∎

Remark 3.33. *Several variants of the above procedure, including extensions to 3D, can be found in the literature, e.g., (Babuška and Suri, 1987), (Muñoz-Sola, 1997), the* projection-based interpolation*, (Demkowicz, 2008); (Demkowicz, Kurtz, Pardo, Paszyński, Rachowicz, and Zdunek, 2008) and (Melenk and Sauter, 2010), Appendix B and (Melenk, Parsania, and Sauter, 2013), Appendix B. It is worth stressing that the condition $s > 1$ is specific to the case $d = 2$. In general, the above arguments extend to $d > 2$ under the constraint that the underlying regularity is $u \in H^s$, $s > d/2$ (so that point evaluations in the vertices are admissible). The restriction $s > d/2$ can be removed by the following two-step procedure: First, the function u is suitably regularized on a length scale that is determined by the local mesh size and the local approximation order. The thus obtained* smooth *function can be approximated by the above schemes. This procedure is worked out in (Karkulik and Melenk, 2015), Thm. 3.3 and may be understood as one extension of the classical Clément interpolant to the hp-context. However, the polynomial approximation on an element K will not only depend on $u|_K$ but $u|_{\omega_K}$, where ω_K is the union of elements that touch K.* ▪

Remark 3.34. ***(interelement continuity for squares/hexahedra)*** *The approximation operator constructed in Theorem 3.32 for triangles has analogs in 2D for squares. It is worth pointing out a special case in tensor product approximation. For $d = 2$, for example, if the underlying 1D operator I_k satisfies $(\mathrm{I}_k u)(\pm 1) = u(\pm 1)$, then the restriction to the edges of S of the tensor product approximation $\mathrm{I}_k^x \circ \mathrm{I}_k^y u$ reduces to the 1D operator, e.g., $(\mathrm{I}_k^x \circ \mathrm{I}_k^y u)|_{x=-1} = \mathrm{I}_k u(-1, \cdot)$. This implies, for example, for meshes consisting of affine quadrilaterals only that the elementwise approximation of a function u by its tensor-product Gauss-Lobatto interpolant is an H^1-conforming approximation as it is globally continuous.*

Another approach that is often taken in connection with squares/hexahedra is the use of meshes with hanging nodes. This leads to the problem of constrained approximation; see (Demkowicz, Kurtz, Pardo, Paszyński, Rachowicz, and Zdunek, 2008), Chap. 3 for the algorithmic aspects. ▪

Remark 3.35. *The proof of Theorem 3.32 presented emphasizes the importance of simultaneous approximation. The underlying reason for its usefulness is that Sobolev and Besov spaces are interpolation spaces so that* multiplicative *interpolation inequalities are available. Such multiplicative inequalities are useful in approximation if simultaneous approximation in the two pertinent norms is available. This underlies our treatment of L^∞-estimates in (3.42) and was also exploited in the proof of Theorem 3.25.* ▪

Remark 3.36. *Using the techniques of the proof of Theorem 3.32, it is possible to construct piecewise polynomial approximants in an element-by-element fashion, where the polynomial degree is allowed to vary from element to element. The key condition is the ratio of the polynomial degrees of two neighboring elements is bounded by a fixed number that enters the constants in the final estimates.* ▪

3.5. Extensions and Remarks

In the case of multivariate approximation, we have mostly focussed on the case $d = 2$ as an example and, with the exception of Section 3.4, restricted our attention to the approximation on the reference element. For shape-regular meshes, the appropriate powers of the local mesh size can be obtained as in the h-version by scaling arguments. For high order anisotropic elements, care has to be taken in the scaling arguments; see (Melenk, 2002) for the 2D case in the context of boundary layer resolution and (Schötzau, Schwab, and Wihler, 2013) for 3D elliptic boundary value problems in polyhedra.

We noted in Example 3.15 that the p-version can approximate certain singular functions at twice the rate as the h-version on quasi-uniform meshes. In this example, the singularity was at an endpoint of the domain of interest. The corresponding result extends to 2D, if the singularity is at a mesh point, (Babuška and Suri, 1987); see also the chapter **The p-Version of the Finite Element Method** of **ECM2**.

Remark 3.37. *All results cited in this note can be recovered from the books and general articles that are quoted in the References.* ∎

APPENDIX

A.1. Further inverse estimates

Many inverse estimates can be found in (Schwab, 1998), Sections 3.6, 4.6 or in (Bernardi and Maday, 1997), Chap. I, Sec. 5.

It is possible to bound $\|u\|_{L^p(-1,1)}$ by $\|u\|_{L^q(-1,1)}$ for $u \in \mathbb{P}_k$:

Lemma A.1.1. *For every $p \in [2, \infty]$ there holds*

$$\|u\|_{L^p(-1,1)} \leq (\sqrt{2}k)^{1-2/p} \|u\|_{L^2(-1,1)} \qquad \forall u \in \mathbb{P}_k. \tag{A.1.1}$$

Proof. We start with the case $p = \infty$. Write $u = \sum_{i=0}^k u_i L_i$ and use $\|L_i\|_{L^\infty(-1,1)} \leq 1$ together with $\|u\|^2_{L^2(-1,1)} = \sum_{i=0}^k \frac{2}{2i+1}|u_i|^2$. This yields (A.1.1) for $p = \infty$. The case $p = 2$ is trivial. The case $p \in (2, \infty)$ follows from the log-convexity of the L^p-norms, specifically, the bound[§]

$$\|f\|_{L^{p_\theta}} \leq \|f\|^{1-\theta}_{L^{p_0}} \|f\|^\theta_{L^{p_1}}, \qquad \frac{1}{p_\theta} = \frac{1-\theta}{p_0} + \frac{\theta}{p_1}.$$

The choice $p_0 = \infty$ and $p_1 = 2$ and $\theta = 2/p$ finishes the argument. □

An interesting variant of polynomial inverse estimates is given by the following result due to Bernstein:

Lemma A.1.2 ((DeVore and Lorentz, 1993), Chap. 4, Thm. 2.2) *Let $\mathcal{E}_\rho$, $\rho > 1$ denote the ellipse of (3.6). Then for every $\pi_k \in \mathbb{P}_k$ (with complex coefficients) and every $\rho > 1$*

$$\|\pi_k\|_{L^\infty(\mathrm{int}(\mathcal{E}_\rho))} \leq \rho^k \|\pi_k\|_{L^\infty(-1,1)}.$$

A.2. Properties of the Gauß-Lobatto interpolation operator

A.2.1. Stability of the Gauß-Lobatto interpolation operator (cf. Remark 3.3)

For $k \in \mathbb{N}$ we define the Gauß-Lobatto points x_i, $i = 0, \ldots, k$, as the zeros of the polynomial $x \mapsto (1-x^2)L_k'(x)$, where $L(x) = P_k^{(0,0)}(x)$ is the *Legendre polynomial* of degree k. The polynomials

[§]Write $|f| = |f|^\theta |f|^{(1-\theta)}$ and apply Hölder to $|f|^{p_\theta}$

L_n' are orthogonal polynomials (with respect to the weight $(1-x^2)$) so that its zeros are distinct and lie in the interval $[-1,1]$. The Gauß-Lobatto quadrature has the form

$$\sum_{i=0}^{k} \rho_i f(x_i), \tag{A.2.1}$$

where the positive quadrature weights are given by $\rho_i = 2/(k(k+1)L_k^2(\xi_i)$, (Bernardi and Maday, 1997), (4.24). We collect some properties of the Gauß-Lobatto quadrature:

Lemma A.2.1. *(i) The quadrature rule (A.2.1) is exact for polynomials of degree* $2k-1$*. Additionally, one has*

$$\|f\|^2_{L^2(-1,1)} \le \sum_{i=0}^{k} \rho_i |f(x_i)|^2 \le (2+1/k)\|f\|^2_{L^2(-1,1)} \qquad \forall f \in \mathbb{P}_k. \tag{A.2.2}$$

(ii) If the quadrature points x_i*,* $i=0,\ldots,k$ *are sorted in descending order* $-1 = x_k < x_{k-1} < \cdots < x_0 = 1$ *and written in the form* $x_i = \cos\theta_i$ *with* $\theta_i \in [0,\pi]$*, then the corresponding* θ_i *satisfy*

$$\frac{i}{k+1/2}\pi \le \theta_i \le \frac{i+1/2}{k+1/2}\pi, \qquad i=0,\ldots,k. \tag{A.2.3}$$

(iii) The nodes x_i *are distributed symmetrically around* 0 *and satisfy*

$$x_i - x_{i+1} = \int_{\theta_i}^{\theta_{i+1}} \sin\theta \, d\theta \sim \frac{1}{k} \begin{cases} \sin\widetilde{\theta}_i, & 1 \le i \le \lfloor k/2 \rfloor + 1, \\ \frac{1}{k}, & i=0. \end{cases} \quad \forall \widetilde{\theta}_i \in [\theta_i, \theta_{i+1}], \tag{A.2.4}$$

Here, the implied constant is independent of k*,* i*, and* $\widetilde{\theta}_i$*. In particular,*

$$x_i - x_{i+1} \sim k^{-1}\left[k^{-1} + \sqrt{1-\widetilde{x}_i^2}\right], \qquad i=0,\ldots,\lfloor k/2\rfloor+1, \qquad \forall \widetilde{x}_i \in [x_{i+1}, x_i] \tag{A.2.5}$$

with implied constants independent of i*,* k*, and* $\widetilde{x}_i$*.*

(iv) Define the function

$$\omega_k(x) := \frac{1}{k}\left[\frac{1}{k} + \sqrt{1-x^2}\right]. \tag{A.2.6}$$

Then there is a constant $C>0$ *independent of* k *such that*

$$C^{-1}\rho_i \le \omega_k(x_i) \le C\rho_i, \qquad i=0,\ldots,k. \tag{A.2.7}$$

(v) Define $I_i := (\min\{x_i, x_{i+1}\}, \max\{x_i, x_{i+1}\})$*,* $i=0,\ldots,k-1$*. Then*

$$C^{-1}\max\{\rho_i,\rho_{i+1}\} \le \omega_k(x) \le C\min\{\rho_i,\rho_{i+1}\} \qquad \forall x \in I_i, \quad i=0,\ldots,k-1. \tag{A.2.8}$$

Proof. These properties are known in the literature.
Proof of (i): cf., e.g., (Bernardi and Maday, 1992), Chap. III, Cor. 1.13.

Proof of (ii): This is the key result of the lemma and due to Sündermann, (Sündermann, 1980); (Sündermann, 1983). See also (Canuto, Hussaini, Quarteroni, and Zang, 2006), Sec. 2.3.1.

Proof of (iii): Follows from $x_i - x_{i+1} = \cos\theta_i - \cos\theta_{i+1} = \int_{\theta_i}^{\theta_{i+1}} \sin\theta\, d\theta$ and (A.2.3).

Proof of (iv): This is stated in (Bernardi and Maday, 1992), Chap. III, Lemma 1.14 but is extracted from (Szegő, 1975), (15.3.14) combined with the property (A.2.3).

Proof of (v): It suffices to consider the case $i \le \lfloor k/2 \rfloor$ and establish, in view of monotonicity properties of ω_k,

$$\omega_k(x_i) \le \omega_k(x_{i+1}) \lesssim \omega_k(x_i). \tag{A.2.9}$$

The bound $\omega_k(x_i) \le \omega_k(x_{i+1})$ follows from the monotonicity properties of ω_k. For the upper bound we introduce the variable θ with $x = \cos\theta$ and have to establish

$$C^{-1} \le \frac{k^{-1} + \sin\theta_i}{k^{-1} + \sin\theta_{i+1}} \le C.$$

For $i = 0$, this follows by inspection and (A.2.3). For $i \ge 1$, this follows again from (A.2.3) and the observation

$$\frac{\sin(\theta+\delta)}{\sin\theta} = \cos\delta - \frac{\cos\theta}{\sin\theta}\sin\delta \le C \quad \text{ if } \theta \in [1/k, \pi/2] \text{ and } \delta \le c/k.$$

∎

Define the Gauß-Lobatto interpolation operator $\mathrm{I}_k^{GL} : C([-1,1]) \to \mathbb{P}_k$ by

$$(\mathrm{I}_k^{GL} u)(x) := \sum_{i=0}^{k} u(x_i)\ell_i(x), \qquad \ell_i(x) = \prod_{\substack{j=0\\ j\ne i}}^{k} \frac{x - x_j}{x_i - x_j} \tag{A.2.10}$$

Some important stability properties of the Gauß-Lobatto interpolation operator I_k^{GL} are collected in the following proposition:

Theorem A.2.2. *Define $I = [-1,1]$ and the weight function*

$$\omega_k(x) := k^{-1}\left[k^{-1} + \sqrt{1-x^2}\right], \qquad x \in I.$$

Fix $\theta \in [0,1]$. Then there exists $C > 0$ independent of k, $k' \in \mathbb{N}$ such that:

$$\|\mathrm{I}_k^{GL} u\|_{L^\infty(I)} \le C(1+\log k)\|u\|_{L^\infty(-1,1)} \qquad \forall u \in C(\overline{I}) \tag{A.2.11}$$

$$\|\mathrm{I}_k^{GL} u\|_{H^1(I)} \le C\|u\|_{H^1(-1,1)} \qquad \forall u \in H^1(I), \tag{A.2.12}$$

$$\|\mathrm{I}_k^{GL} u\|_{L^2(I)} \le C\left[\|u\|_{L^2(-1,1)} + \frac{1}{k}\|u'\|_{L^2(I)}\right] \ \forall u \in H^1(I), \tag{A.2.13}$$

$$\|\mathrm{I}_k^{GL} u\|_{L^2(I)} \le C\left[\|u\|_{L^2(-1,1)} + \|\omega_k^\theta u\|_{L^2(I)}^{1/2}\|\omega_k^{1-\theta} u'\|_{L^2(I)}^{1/2}\right] \ \forall u \in H^1(I), \tag{A.2.14}$$

$$\|\mathrm{I}_k^{GL} u\|_{L^2(I)} \le C\left[\|u\|_{L^2(-1,1)} + \frac{1}{\sqrt{k}}\|u\|_{B^{1/2}_{2,1}(I)}\right] \ \forall u \in B^{1/2}_{2,1}(I), \tag{A.2.15}$$

$$\|\mathrm{I}_k^{GL} u\|_{L^2(I)} \le C(1+k'/k)\|u\|_{L^2(I)} \qquad \forall u \in \mathbb{P}_{k'} \tag{A.2.16}$$

$$\left\|\frac{u - \mathrm{I}_k^{GL} u}{\sqrt{1-x^2}}\right\|_{L^2(I)} \le C\frac{1}{k}\|u'\|_{L^2(I)} \qquad \forall u \in H^1(I). \tag{A.2.17}$$

Proof. The stability result (A.2.11) is due to Sündermann, (Sündermann, 1980); (Sündermann, 1983). The estimates (A.2.12), (A.2.13), (A.2.17), can be obtained from (Bernardi and Maday, 1997), Thm. 13.4. The estimate (A.2.16) is shown in (Bernardi and Maday, 1997), Rem. 13.5. The estimates (A.2.14), (A.2.15) will be shown below.

The key ingredient for the proof of the estimates cited from (Bernardi and Maday, 1997), Thm. 13.4 and the (A.2.14), (A.2.15) is (A.2.3), i.e., the fact that values θ_i of the representation $x_i = \cos\theta_i$ are nearly uniformly distributed in $[0,\pi]$.

Let us be more specific about (A.2.14) and (A.2.15). With the quadrature weights ρ_i for Gauss-Lobatto quadrature, we note

$$\|\mathrm{I}_k^{GL} u\|^2_{L^2(I)} \overset{(A.2.2)}{\leq} \sum_{i=0}^{p} \rho_i |u(x_i)|^2 \leq \sum_{i=0}^{\lfloor p/2\rfloor+1} \rho_i |u(x_i)|^2 + \sum_{\lfloor k/2\rfloor+1}^{k} \rho_i |u(x_i)|^2.$$

By symmetry of the distribution of the Gauss-Lobatto points around the midpoint $x=0$, it suffices to consider the sum $\sum_{i=0}^{\lfloor k/2\rfloor+1} \rho_i |u(x_i)|^2$—the other one is treated similarly. For each x_i, we consider $I_i := (x_{i+1}, x_i)$. The 1D Sobolev embedding then gives

$$\|u\|^2_{L^\infty(I_i)} \lesssim \frac{1}{|I_i|}\|u\|^2_{L^2(I_i)} + \|u\|_{L^2(I_i)}\|u'\|_{L^2(I_i)} \tag{A.2.18}$$

Next, we note that (cf. (iv), (v) and (Bernardi and Maday, 1997), Thm. 4.5)

$$\rho_i \lesssim \frac{1}{k}\left[k^{-1} + \sqrt{1-x_i^2}\right]$$

so that

$$\frac{\rho_i}{|I_i|} \overset{(A.2.5)}{\lesssim} 1.$$

Upon writing $\rho_i = \rho_i^\theta \rho_i^{1-\theta}$ and noting that ρ_i is equivalent to ω_k on I_i (cf. (A.2.8)), we get

$$\|\mathrm{I}_k^{GL} u\|^2_{L^2(I)} \leq \sum_{i=0}^{k} \rho_i |u(x_i)|^2 \lesssim \sum_{i=0}^{k} \|u\|^2_{L^2(I_i)} + \rho_i \|u\|_{L^2(I_i)}\|u'\|_{L^2(I_i)} \tag{A.2.19}$$

$$\lesssim \|u\|^2_{L^2(I)} + \|\omega_k^\theta u\|_{L^2(I)}\|\omega_k^{1-\theta} u'\|_{L^2(I)}, \tag{A.2.20}$$

which is (A.2.14). We note that the following simplified estimates can be derived:

$$\|\mathrm{I}_k^{GL} u\|^2_{L^2(I)} \lesssim \|u\|^2_{L^2(I)} + k^{-1}\|u\|_{L^2(I)}\|u'\|_{L^2(I)} \lesssim \|u\|^2_{L^2(I)} + k^{-2}\|u'\|^2_{L^2(I)}. \tag{A.2.21}$$

We now show (A.2.15) using (A.2.21). Let $\Pi_k : L^2(I) \to \mathbb{P}_k$ be a quasi-interpolation operator with the following properties:

$$\|\Pi_k u\|_{H^j(I)} \lesssim \|u\|_{H^j(I)}, \qquad j \in \{0,1\}, \tag{A.2.22}$$

$$\|u - \Pi_k u\|_{L^2(I)} \lesssim k^{-1}\|u\|_{H^1(I)}. \tag{A.2.23}$$

Such an operator is constructed, for example, in Examples 3.18, 3.19 or in (Karkulik and Melenk, 2014). The simultaneous stability in L^2 and H^1 implies stability in $B^{1/2}_{2,1}(I)$ (cf. (Tartar, 2007), Lemma 25.3):

$$\|\Pi_k u\|_{B^{1/2}_{2,1}(I)} \lesssim \|u\|_{B^{1/2}_{2,1}(I)}. \tag{A.2.24}$$

Next, we estimate

$$\begin{aligned}\|\mathrm{I}_k^{GL} u\|_{L^2(I)} &\lesssim \|\mathrm{I}_k^{GL}(u - \Pi_k u)\|_{L^2(I)} + \|\Pi_k u\|_{L^2(I)} \\ &\lesssim \|\mathrm{I}_k^{GL}(u - \Pi_k u)\|_{L^2(I)} + \|u\|_{L^2(I)}.\end{aligned}$$

We now study the mapping $u \mapsto \mathrm{I}_k^{GL}(u - \Pi_k u)$ and estimate

$$\begin{aligned}\|\mathrm{I}_k^{GL}(u-\Pi_k u)\|_{L^2(I)} &\lesssim \|u-\Pi_k u\|_{L^2(I)} + k^{-1/2}\|u-\Pi_k u\|^{1/2}_{L^2(I)}\|u-\Pi_k u\|^{1/2}_{H^1(I)}\\ &\lesssim k^{-1/2}\|u\|_{B^{1/2}_{2,1}(I)} + k^{-1/2}\|u\|^{1/2}_{L^2(I)}\|u\|^{1/2}_{H^1(I)},\\ &\lesssim k^{-1/2}\|u\|^{1/2}_{L^2(I)}\|u\|^{1/2}_{H^1(I)},\end{aligned}$$

where we used the classical interpolation inequality in the last step. An appeal to (Tartar, 2007), Lemma 25.3 concludes the proof of (A.2.15). □

A.2.2. Stability of the Gauß interpolation operator (cf. Remark 3.3)

We now show that the estimates (A.2.14), (A.2.15), also hold for the Gauss points:

Lemma A.2.3. *Define the function $\omega_G(x) := \sqrt{1-x^2}$. Fix $\theta \in [0,1]$. Let ζ_i, $i = 1, \ldots, k$, be the Gauss points, i.e., the zeros of L_k. Denote by $\mathrm{I}_{k-1}^G : C([-1,1]) \to \mathbb{P}_{p-1}$ the interpolation in these points. Then there is $C > 0$ independent of p such that*

$$\|\mathrm{I}_k^G u\|^2_{L^2(-1,1)} \le C\left[\|u\|^2_{L^2(-1,1)} + k^{-1}\|\omega_G^\theta u\|_{L^2(-1,1)}\|\omega_G^{1-\theta}u'\|_{L^2(-1,1)}\right], \tag{A.2.25}$$

$$\|\mathrm{I}_k^G u\|^2_{L^2(-1,1)} \le C\left[\|u\|^2_{L^2(-1,1)} + k^{-1/2}\|u\|_{B^{1/2}_{2,1}(-1,1)}\right]. \tag{A.2.26}$$

Proof. The proof parallels that of (A.2.14), (A.2.15). We define $\vartheta_i := \arccos\zeta_i$, $i = 1, \ldots, k$. In fact, the key steps can be found in the proof of (Bernardi and Maday, 1997), Thm. 13.1.

1. step: According to (Szegő, 1975), Thm. 6.21.3 we have for the values ϑ_i:

$$\frac{i-1/2}{k}\pi \le \vartheta_i \le \frac{i}{k+1}\pi \qquad i = 1, \ldots, \lfloor k/2 \rfloor. \tag{A.2.27}$$

2. step: The Gauß quadrature with weights ρ_i^G has the form $\sum_{i=1}^k \rho_i^G f(\zeta_i)$. It is exact for polynomials $f \in \mathbb{P}_{2k-1}$. Furthermore, according to (Bernardi and Maday, 1997), Thm. 4.4, Rem. 4.4 we have

$$C^{-1}k^{-1}\sqrt{1-\zeta_i^2} \le \rho_i^G \le Ck^{-1}\sqrt{1-\zeta_i^2}. \tag{A.2.28}$$

3. step: For each i select an interval I_i of length proportional to ρ_i^G, e.g.,

$$I_i := (\cos(\vartheta_i - \delta/k), \cos(\vartheta_i + \delta/k)), \qquad i = 1, \ldots, k,$$

where $\delta > 0$ is sufficiently small so that the intervals I_i do not overlap (this is possible in view of (A.2.27)). Next, one checks that

$$|I_i| \sim k^{-1}\sqrt{1-\zeta_i^2} \sim \rho_i^G \tag{A.2.29}$$

uniformly in i and k.

4. step:

$$\begin{aligned}\|\mathrm{I}_{k-1}^G u\|^2_{L^2(I)} = \sum_{i=1}^k \rho_i^G |u(\zeta_i)|^2 &\overset{(A.2.18)}{\lesssim} \sum_{i=1}^k \|u\|^2_{L^2(I_i)} + \rho_i^G\|u\|_{L^2(I_i)}\|u'\|_{L^2(I_i)}\\ &\overset{(A.2.28)}{\lesssim} \|u\|^2_{L^2(I)} + k^{-1}\|\omega_G^\theta u\|_{L^2(I)}\|\omega_G^{1-\theta}u'\|_{L^2(I)}.\end{aligned}$$

This is (A.2.25). The proof of (A.2.26) follows from this in exactly the same way as in (A.2.15) follows from (A.2.14). □

A.2.3. Simultaneous approximation in Gauß-Lobatto and Gauß interpolation (cf. Cor. 3.21)

As an alternative to the procedure in (Bernardi and Maday, 1997), Thm. 13.4 we present a proof of of Corollary 3.21 that is based on interpolation arguments.

Lemma A.2.4. *Let I_k^{GL} and I_k^{G} represent interpolation in the Gauß-Lobatto and Gauß points. Then*

$$\|u - \mathrm{I}_k^{GL} u\|_{H^r(I)} \le C_{r,s} k^{-(s-r)} \|u\|_{H^s(I)}, \qquad 0 \le r \le 1, \quad s > (1+r)/2, \tag{A.2.30}$$

$$\|u - \mathrm{I}_k^{G} u\|_{L^2(I)} \le C_s k^{-s} \|u\|_{H^s(I)}, \qquad 1/2 < s. \tag{A.2.31}$$

Proof. *Proof of (A.2.30):* Corollary 3.21 has shown the result in the range $0 \le r \le 1$ together with $1 \le s$. Thus, we have to show the remaining cases $(1+r)/2 < s \le 1$. The key to using interpolation arguments for this case is the refined stability estimate (A.2.15) for the limiting case. From (A.2.15), we obtain the approximation result

$$\|u - \mathrm{I}_k^{GL} u\|_{L^2(I)} \le C k^{-1/2} \|u\|_{B^{1/2}_{2,1}(I)}. \tag{A.2.32}$$

From the $H^1(I)$-stability of I_k^{GL}, we infer (again first for integer s and then, by interpolation, all $s' \ge 1$)

$$\|u - \mathrm{I}_k^{GL} u\|_{H^1(I)} \le C_{s'} k^{-(s'-1)} \|u\|_{H^{s'}(I)}, \qquad s' \ge 1. \tag{A.2.33}$$

For $r \in (0,1)$ and $1 > s > (1+r)/2$ select

$$s' := \frac{s - (1-r)/2}{r}.$$

Note that $s' > 1$. The choice of s' become clear once one observes that the reinterpolation theorem allows us to identify the Sobolev space $H^s(I)$, $s \in (1/2, 1)$ as $H^s(I) = (B^{1/2}_{2,1}(I), H^{s'}(I))_{r,2}$. Then, interpolating between (A.2.32), (A.2.33) yields

$$\|u - \mathrm{I}_k^{GL} u\|_{H^r(I)} \le C k^{-(1-r)/2} k^{-rs'} \|u\|_{(B^{1/2}_{2,1}(I), H^{s'}(I)_{r,2})} \le C k^{-s} \|u\|_{H^s(I)}.$$

Proof of (A.2.31): The estimate (A.2.26) provides the limiting case for interpolation arguments. We have

$$\|u - \mathrm{I}_k^{G} u\|_{L^2(I)} \le C k^{-1/2} \|u\|_{B^{1/2}_{2,1}(I)}$$
$$\|u - \mathrm{I}_k^{G} u\|_{L^2(I)} \le C k^{-s} \|u\|_{H^s(I)}, \qquad s \ge 1.$$

Interpolation between these two estimates and using again that the reinterpolation theorem yields $H^{\theta/2 + (1-\theta)s}(I) = (B^{1/2}_{2,2}(I), H^s(I))_{\theta,2}$ concludes the proof. □

A.3. Operators stable both in L^2 and H^1 (cf. Examples 3.18, 3.19)

Examples 3.18, 3.19 claimed the existence of some projection operators that are simultaneously stable in $L^2(-1,1)$ and $H^1(-1,1)$. Here, we provide self-contained proofs of these claims.

Theorem A.3.1 (cf. Example 3.18) *Let $I = [-1,1]$. For each $k \in \mathbb{N}_0$ there is a linear operator $\widetilde{\mathrm{Q}}_k : L^2(I) \to \mathbb{P}_{2k}$ with the following properties:*

(i) $\|\widetilde{Q}_k u\|_{L^2(I)} \leq \|u\|_{L^2(I)}$.

(ii) $\|\widetilde{Q}_k u\|_{H^1(I)} \leq C\|u\|_{H^1(I)}$ *for a* $C > 0$ *independent of* k.

(iii) $\widetilde{Q}_k u = u$ *for all* $u \in \mathbb{P}_k$.

Proof. Such operators of the "de la Vallée-Poussin" type have been constructed repeatedly in the literature. We mention here (Bernardi and Maday, 1999); (Braess, Pillwein, and Schöberl, 2009). Fix a smooth function $\chi \in C^\infty(\mathbb{R})$ with $\operatorname{supp}\chi \subset [-2, 2]$, $\chi \equiv 1$ on $[0, 1]$ and $0 \leq \chi \leq 1$. We expand the function u as well as u' in Legendre series:

$$u = \sum_{n=0}^{\infty} u_n L_n, \tag{A.3.1}$$

$$u' = \sum_{n=0}^{\infty} b_n L_n. \tag{A.3.2}$$

We will repeatedly use the following definitions and properties:

$$\gamma_n = \|L_n\|^2_{L^2(I)} = \frac{2}{2n+1}, \tag{A.3.3}$$

$$(2n+1)L_n = L'_{n+1} - L'_{n-1}, \tag{A.3.4}$$

$$\|L'_n\|^2_{L^2(I)} = n(n+1); \tag{A.3.5}$$

here, we employed (Bernardi and Maday, 1997), Thm. 3.2 in (A.3.3), (Bernardi and Maday, 1997), Thm. 3.3 in (A.3.4), and (Bernardi and Maday, 1997), (5.3) in (A.3.5). We recall the relation between the coefficients u_n and b_n (cf. (Houston, Schwab, and Süli, 2002), Lemma 3.5 or (Melenk and Wurzer, 2014), (1.6)):

$$u_n = \frac{b_{n-1}}{2n-1} - \frac{b_{n+1}}{2n+3}, \qquad n \geq 1. \tag{A.3.6}$$

Furthermore,

$$\|u\|^2_{L^2(I)} = \sum_{n=0}^{\infty} \gamma_n |u_n|^2, \qquad \|u'\|^2_{L^2(I)} = \sum_{n=0}^{\infty} \gamma_n |b_n|^2.$$

We define the operator $\widetilde{Q}$ by

$$\widetilde{Q}_k u := \sum_{n=0}^{\infty} \chi(n/k) u_n L_n. \tag{A.3.7}$$

From $\chi|_{[0,1]} \equiv 1$, it is clear that $\widetilde{Q}u = u$ for $u \in \mathbb{P}_k$ and $\|\chi\|_{L^\infty(\mathbb{R})} \leq 1$ implies $\|\widetilde{Q}_k u\|_{L^2(I)} \leq \|u\|_{L^2(I)}$. From $\operatorname{supp}\chi \subset [-2, 2]$ it is clear that $\widetilde{Q}_k u \in \mathbb{P}_k$. It thus remains to study $\|(\widetilde{Q}u)'\|_{L^2(I)}$. The key idea is to exploit smoothness properties of χ through a summation by parts argument:

$$
\begin{aligned}
(\widetilde{Q}_k u)' &= \sum_{n=1}^{\infty} \chi(n/k) u_n L_n' = \sum_{n=1}^{\infty} \chi(n/k) \left(\frac{b_{n-1}}{2(n-1)+1} - \frac{b_{n+1}}{2(n+1)+1} \right) L_n' \\
&= \sum_{n=3}^{\infty} \frac{b_{n-1}}{2(n-1)+1} \left[\chi(n/k) L_n' - \chi((n-2)/k) L_{n-2}' \right] + \sum_{n=1}^{2} \chi(n/k) \frac{b_{n-1}}{2(n-1)+1} L_n' \\
&= \sum_{n=3}^{\infty} \frac{b_{n-1}}{2(n-1)+1} \chi(n/k) \underbrace{\left[L_n' - L_{n-2}' \right]}_{\stackrel{(A.3.4)}{=} (2(n-1)+1) L_{n-1}} \\
&\quad - \sum_{n=3}^{\infty} \frac{b_{n-1}}{2(n-1)+1} \left[\chi((n-2)/k) - \chi(n/k) \right] L_{n-2}' + \sum_{n=1}^{2} \chi(n/k) \frac{b_{n-1}}{2(n-1)+1} L_n' \\
&=: S_1 + S_2 + S_3.
\end{aligned}
$$

We estimate these three terms in turn:

$$
\begin{aligned}
\|S_1\|_{L^2(I)}^2 &= \sum_{n=3}^{\infty} |\chi(n/k)|^2 \gamma_{n-1} |b_{n-1}|^2 \le \|u'\|_{L^2(I)}^2, \\
\|S_2\|_{L^2(I)} &\lesssim \sum_{n=3}^{2k} \frac{|b_{n-1}|}{2n-1} \frac{1}{k} \|L_{n-2}'\|_{L^2(I)} \lesssim \sum_{n=1}^{2k} \frac{\sqrt{\gamma_{n-1}} |b_{n-1}|}{(n+1)\sqrt{\gamma_{n-1}}} k^{-1} n \\
&\lesssim \sqrt{\sum_{n=1}^{2k} \gamma_{n-1} |b_{n-1}|^2} \sqrt{\sum_{n=1}^{2k} \frac{1}{\gamma_{n-1}} k^{-2}} \lesssim \|u'\|_{L^2(I)}, \\
\|S_3\|_{L^2(I)} &\lesssim |b_0| + |b_1| \lesssim \|u'\|_{L^2(I)}.
\end{aligned}
$$

□

We now turn to Example 3.19.

Theorem A.3.2 (cf. Example 3.19) *Let $I = [-1, 1]$. For each $k \in \mathbb{N}_0$ there is a linear operator $\mathrm{Q}_k : L^2(I) \to \mathbb{P}_k$ with the following properties:*

(i) $\|\mathrm{Q}_k u\|_{L^2(I)} \le C \|u\|_{L^2(I)}$.

(ii) $\|\mathrm{Q}_k u\|_{H^1(I)} \le C \|u\|_{H^1(I)}$ *for a $C > 0$ independent of k.*

(iii) $\mathrm{Q}_k u = u$ *for all $u \in \mathbb{P}_k$.*

Proof. As in the proof of Theorem A.3.1 we expand u and u' in the Legendre series (A.3.1), (A.3.2). We also use the cut-off function χ of the proof of Theorem A.3.1. The operator Q_k is defined by

$$\mathrm{Q}_k u := \sum_{n=0}^{k} \chi(n/k) u_n L_n + \sum_{n=k+1}^{2k} \chi(n/k) u_n L_{2k-n}. \tag{A.3.8}$$

By construction, it is a projection onto $\mathbb{P}_k$ so that only the stability in $L^2(I)$ and $H^1(I)$ have to be shown. In view of the stability properties of the operator $\widetilde{Q}_k$ constructed in the proof of Theorem A.3.1, it suffices to establish the stability of the operator

$$\widetilde{\mathrm{Q}}_k u - \mathrm{Q}_k u = \sum_{n=k+1}^{2k} \chi(n/k) u_n \left[L_n - L_{2k-n} \right]. \tag{A.3.9}$$

We start with the $L^2(I)$-stability. We write

$$\widetilde{\mathrm{Q}}_k u - \mathrm{Q}_k u = \sum_{n=k+1}^{2k} \chi(n/k) u_n L_n + \sum_{n=0}^{k-1} \chi((2k-n)/k) u_{2k-n} L_n$$

so that

$$\|\widetilde{\mathrm{Q}}_k u - \mathrm{Q}_k u\|^2_{L^2(I)} = \sum_{n=k}^{2k} |\chi(n/k)|^2 |u_n|^2 \gamma_n + \sum_{n=0}^{k-1} |\chi((2k-n)/k)|^2 \frac{\gamma_n}{\gamma_{2k-n}} |u_{2k-n}|^2 \gamma_{2k-n}.$$

By the smoothness and the support properties of χ, we have

$$|\chi((2k-n)/k)| \leq Cn/k \tag{A.3.10}$$

so that for $0 \leq n \leq k$ we have

$$|\chi((2k-n)/k)| \frac{\gamma_n}{\gamma_{2k-n}} \leq C \frac{n}{k} \frac{2(2k-n)+1}{2n+1} \leq C.$$

Hence, we obtain

$$\|\widetilde{\mathrm{Q}}_k u - \mathrm{Q}_k u\|^2_{L^2(I)} \leq \sum_{n=k}^{2k} |\chi(n/k)|^2 |u_n|^2 \gamma_n + C \sum_{n=0}^{k-1} |u_{2k-n}|^2 \gamma_{2k-n} \leq C \|u\|^2_{L^2(I)}.$$

For the $H^1(I)$-stability analysis of $\widetilde{\mathrm{Q}}_k - \mathrm{Q}_k$ we write with (A.3.9) and (A.3.6):

$$(\widetilde{Q}_k u - \mathrm{Q}_k u)' = \sum_{n=k+1}^{2k} \chi(n/k) u_n \left[L_n - L_{2k-n}\right]' = \sum_{n=k+1}^{2k} \chi(n/k) \left(\frac{b_{n-1}}{2n-1} - \frac{b_{n+1}}{2n+3} \right) \left[L_n - L_{2k-n}\right]'.$$

The summation by parts formula of Lemma A.3.3 gives

$$\begin{aligned}
(\widetilde{Q}_k u - \mathrm{Q}_k u)' = {} & \sum_{n=k+3}^{2k} \frac{b_{n-1}}{2n-1} \left[\chi(n/k)(L_n - L_{2k-n})' - \chi((n-2)/k)(L_{n-2} - L_{2k-n+2})' \right] \\
& + \frac{b_{k-1}}{2k-1} \chi(k/k)(L_k - L_{2k-k})' + \frac{b_k}{2k+3} \chi((k+1)/k)(L_{k+1} - L_{2k-k-1})' \\
& - \frac{b_{2k-1}}{2(2k)-1} \chi((2k-1)/k)(L_{2k-1} - L_{2k-(2k-1)})' - \frac{b_{2k}}{2(2k+1)-1} \chi(2k/k)(L_{2k} - L_{2k-(2k)})' \\
& =: S_1 + S_2 + S_3.
\end{aligned}$$

We deal with the terms S_1, S_2, S_3 in turn.

The term S_1: We write

$$\begin{aligned}
S_1 = {} & \sum_{n=k+3}^{2k} \frac{b_{n-1}}{2n-1} \chi(n/k)(L_n - L_{n-2})' - \sum_{n=k+3}^{2k} \frac{b_{n-1}}{2n-1} \chi(n/k)(L_{2k-n} - L_{2k-n+2})' \\
& - \sum_{n=k+3}^{2k} \frac{b_{n-1}}{2n-1} (\chi((n-2)/k) - \chi(n/k))(L_{n-2} - L_{2k-n+2})' \\
& =: S_{1,1} + S_{1,2} + S_{1,3}.
\end{aligned}$$

For the terms $S_{1,1}$, $S_{1,2}$ we employ (A.3.4) to get

$$S_{1,1} + S_{1,2} = \sum_{n=k+3}^{2k} \chi(n/k) b_{n-1} L_{n-1} - \sum_{n=k+3}^{2k} \frac{b_{n-1}}{2n-1} \chi(n/k)(2(2k-n+1)+1) L_{2k-n+1}.$$

This implies

$$\|S_{1,1}+S_{1,2}\|_{L^2(I)} \lesssim \|u'\|_{L^2(I)}.$$

For the term $S_{1,3}$ we use the Lipschitz continuity of χ to infer $|\chi((n-2)/k)-\chi(n/k)| \leq Ck^{-1}$ (uniformly in n) and we use (A.3.5) to get

$$\|S_{1,3}\|_{L^2(I)} \lesssim \frac{1}{k}\sum_{n=k+3}^{2k}\frac{|b_{n-1}|}{2n-1}\underbrace{[n+(2k-n+1)]}_{\lesssim k} \lesssim \frac{1}{k}\sqrt{\sum_n \gamma_{n-1}|b_{n-1}|^2}\sqrt{\sum_{n=k+3}^{2k} k} \lesssim \|u'\|_{L^2(I)}.$$

The term S_2:

$$S_2 = \frac{b_k}{2k+3}\chi((k+1)/k)\left[L_{k+1}-L_{k-1}\right]' \overset{(A.3.5)}{=} \frac{b_k}{2k+3}\chi((k+1)/k)(2k+1)L_k$$

so that

$$\|S_1\|_{L^2(I)} = \frac{|b_k|}{2k+3}|\chi((k+1)/k)(2k+1)|\sqrt{\gamma_k} \leq \|u'\|_{L^2(I)}.$$

The term S_3: Since $\operatorname{supp}\chi \subset [-2,2]$, we have $\chi(2k/k)=0$ so that

$$S_3 = \frac{b_{2k-1}}{4k-1}\chi((2k-1)/k)\left[L_{2k-1}-L_1\right]'.$$

From the support property of χ we also get $|\chi((2k-1)/k)| \leq ck^{-1}$ and in view of (A.3.5) we arrive at

$$\|S_3\|_{L^2(I)} \lesssim \frac{|b_{2k-1}|}{4k-1}k^{-1}\left[k+1\right] \lesssim |b_{2k-1}|k^{-1} \lesssim k^{-1/2}\|u'\|_{L^2(I)}.$$

∎

Lemma A.3.3 (summation by parts) *Let $(a_n)_{n=0}^\infty$, $(b_n)_{n=0}^\infty \subset \mathbb{R}$ and p, $q \in \mathbb{N}_0$ with $q \leq q$. Then:*

$$\sum_{n=p+1}^{q}(a_{n-1}-a_{n+1})b_n = \sum_{n=p+3}^{q} a_{n-1}(b_n-b_{n-2}) + a_pb_{p+1} + a_{p+1}b_{p+2} - a_qb_{q-1} - a_{q+1}b_q.$$

Proof.

$$\begin{aligned}\sum_{n=p+1}^{q}(a_{n-1}-a_{n+1})b_n &= \sum_{n=p+1}^{q} a_{n-1}b_n - \sum_{n=p+1}^{q} a_{n+1}b_n = \sum_{n=p+1}^{q} a_{n-1}b_n - \sum_{n=p+3}^{q+2} a_{n-1}b_{n-2}\\ &= \sum_{n=p+3}^{q} a_{n-1}(b_n-b_{n-2}) + \sum_{n=p+1}^{p+2} a_{n-1}b_n - \sum_{n=q+1}^{q+2} a_{n-1}b_{n-2}\\ &= \sum_{n=p+3}^{q} a_{n-1}(b_n-b_{n-2}) + a_pb_{p+1} + a_{p+1}b_{p+2} - a_qb_{q-1} - a_{q+1}b_q.\end{aligned}$$

This concludes the proof. ∎

REFERENCES

G. Acosta, Th. Apel, R. G. Durán, and A. L. Lombardi (2008). 'Anisotropic error estimates for an interpolant defined via moments'. *Computing* **82**:1–9.

G. Acosta and R. G. Durán (2000). 'Error estimates for $\mathcal{Q}_1$ isoparametric elements satisfying a weak angle condition'. *SIAM J. Numer. Anal.* **38**:1073–1088.

Th. Apel (1998). 'Anisotropic interpolation error estimates for isoparametric quadrilateral finite elements'. *Computing* **60**:157–174.

Th. Apel (1999a). *Anisotropic finite elements: Local estimates and applications.* Advances in Numerical Mathematics. Teubner, Stuttgart.

Th. Apel (1999b). 'Interpolation of non-smooth functions on anisotropic finite element meshes'. *Math. Modeling Numer. Anal.* **33**:1149–1185.

Th. Apel and M. Dobrowolski (1992). 'Anisotropic interpolation with applications to the finite element method'. *Computing* **47**:277–293.

Th. Apel, A. L. Lombardi, and M. Winkler (2014). 'Anisotropic mesh refinement in polyhedral domains: error estimates with data in $L^2(\Omega)$'. *Math. Modeling Numer. Anal. (M2AN)* **48**:1117–1145.

Th. Apel and G. Lube (1998). 'Anisotropic mesh refinement for a singularly perturbed reaction diffusion model problem'. *Appl. Numer. Math.* **26**:415–433.

Th. Apel and G. Matthies (2008). 'Non-conforming, anisotropic, rectangular finite elements of arbitrary order for the Stokes problem'. *SIAM J. Numer. Anal.* **46**:1867–1891.

Th. Apel and F. Milde (1996). 'Comparison of several mesh refinement strategies near edges'. *Comm. Numer. Methods Engrg.* **12**:373–381. Shortened version of Preprint SPC94_15, TU Chemnitz-Zwickau, 1994.

Th. Apel and S. Nicaise (1996). 'Elliptic problems in domains with edges: anisotropic regularity and anisotropic finite element meshes'. In J. Cea, D. Chenais, G. Geymonat, and J. L. Lions (eds.), *Partial Differential Equations and Functional Analysis (In Memory of Pierre Grisvard)*, pp. 18–34. Birkhäuser, Boston. Shortened version of Preprint SPC94_16, TU Chemnitz-Zwickau, 1994.

Th. Apel and S. Nicaise (1998). 'The finite element method with anisotropic mesh grading for elliptic problems in domains with corners and edges'. *Math. Methods Appl. Sci.* **21**:519–549.

Th. Apel, S. Nicaise, and J. Schöberl (2001). 'Crouzeix-Raviart type finite elements on anisotropic meshes'. *Numer. Math.* **89**:193–223.

Th. Apel, A.-M. Sändig, and J. R. Whiteman (1996). 'Graded mesh refinement and error estimates for finite element solutions of elliptic boundary value problems in non-smooth domains.'. *Math. Methods Appl. Sci.* **19**:63–85.

D. N. Arnold, D. Boffi, and R. S. Falk (2002). 'Approximation by quadrilateral finite elements'. *Math. Comput.* **239**:909–922.

I. Babuška, A. Craig, J. Mandel, and J. Pitkäranta (1991). 'Efficient preconditioning for the p version finite element method in two dimensions'. *SIAM J. Numer. Anal.* **28**(3):624–661.

I. Babuška and B.Q. Guo (1986). 'The $h-p$ version of the finite element method. Part 1: The basic approximation results'. *Computational Mechanics* **1**:21–41.

I. Babuška and B. Guo (1988). 'Regularity of the Solution of Elliptic Problems with Piecewise Analytic Data. Part I. Boundary Value Problems for Linear Elliptic Equations of Second Order'. *SIAM J. Math. Anal* **19**(1):172–203.

I. Babuška and Manil Suri (1987). 'The *h-p* version of the finite element method with quasi-uniform meshes'. *RAIRO Modél. Math. Anal. Numér.* **21**(2):199–238.

I. Babuška and M. Suri (1987). 'The optimal convergence rate of the p-version of the finite element method'. *SIAM J. Numer. Anal.* **24**:750–776.

I. Babuška and M. Suri (1994). 'The p and h-p versions of the finite element method, basic principles and properties'. *SIAM review* **36**(4):578–632.

I. Babuška and A. K. Aziz (1976). 'On the angle condition in the finite element method'. *SIAM J. Numer. Anal.* **13**:214–226.

C. Băcuţă, V. Nistor, and L. T. Zikatanov (2007). 'Improving the rate of convergence of high-order finite elements in polyhedra II: mesh refinements and interpolation'. *Numer. Funct. Anal. Optimization* **28**:775–824.

L. Veiga, F. Brezzi, A. Cangiani, G. Manzini, L. D. Marini, and A. Russo (2013). 'Basic principles of virtual element methods'. *Math. Models Methods Appl. Sci.* **23**(1):199–214.

L. Veiga, F. Brezzi, L. D. Marini, and A. Russo (2014). 'The hitchhiker's guide to the virtual element method'. *Math. Models Methods Appl. Sci.* **24**(8):1541–1573.

F. Ben Belgacem (1994). 'Polynomial extensions of compatible polynomial traces in three dimensions'. *Comput. Meth. Appl. Mech. Engrg.* **116**:235–241.

C. Bernardi (1989). 'Optimal finite-element interpolation on curved domains'. *SIAM J. Numer. Anal.* **26**:1212–1240.

C. Bernardi, M. Dauge, and Y. Maday (2007). 'Polynomials in the Sobolev World (Version 2)'. Tech. Rep. 14, IRMAR.

C. Bernardi and V. Girault (1998). 'A local regularization operator for triangular and quadrilateral finite elements'. *SIAM J. Numer. Anal.* **35**:1893–1916.

C. Bernardi and Y. Maday (1992). *Approximations spectrales de problèmes aux limites elliptiques.* Mathématiques & Applications. Springer Verlag.

C. Bernardi and Y. Maday (1997). 'Spectral Methods'. In P.G. Ciarlet and J.L. Lions (eds.), *Handbook of Numerical Analysis, Vol. 5.* North Holland, Amsterdam.

C. Bernardi and Y. Maday (1999). 'Uniform inf-sup conditions for the spectral discretization of the Stokes problem'. *Math. Models Methods Appl. Sci.* **9**(3):395–414.

S. Börm, M. Löhndorf, and J.M. Melenk (2005). 'Approximation of Integral Operators by Variable-Order Interpolation'. *Numer. Math.* **99**(4):605–643.

D. Braess (1997). *Finite Elemente.* Springer, Berlin.

Dietrich Braess, Veronika Pillwein, and Joachim Schöberl (2009). 'Equilibrated residual error estimates are p-robust'. *Comput. Methods Appl. Mech. Engrg.* **198**(13-14):1189–1197.

D. Braess and C. Schwab (2000). 'Approximation on simplices with respect to weighted Sobolev norms'. *J. Approx. Theory* **103**:329–337.

J. H. Bramble and S. R. Hilbert (1970). 'Estimation of linear functionals on Sobolev spaces with applications to Fourier transforms and spline interpolation'. *SIAM J. Numer. Anal.* **7**:112–124.

J. H. Bramble and S. R. Hilbert (1971). 'Bounds for a class of linear functionals with applications to Hermite interpolation'. *Numer. Math.* **16**:362–369.

S. C. Brenner and L. R. Scott (1994). *The mathematical theory of finite element methods.* Springer, New York.

L. Brutman (1997). 'Lebesgue functions for polynomial interpolation—a survey'. *Ann. Numer. Math.* **4**(1-4):111–127. The heritage of P. L. Chebyshev: a Festschrift in honor of the 70th birthday of T. J. Rivlin.

C. Canuto, M. Y. Hussaini, A. Quarteroni, and T. A. Zang (2006). *Spectral methods.* Scientific Computation. Springer-Verlag, Berlin. Fundamentals in single domains.

C. Canuto and A. Quarteroni (1982). 'Approximation results for orthogonal polynomials in Sobolev spaces'. *Math. Comp.* **38**(257):67–86.

W. Cao (2005). 'On the error of linear interpolation and the orientation, aspect ratio, and internal angles of a triangle'. *SIAM J. Numer. Anal.* **43**:19–40.

C. Carstensen (1999). 'Quasi-interpolation and a posteriori error analysis in finite element methods'. *Math. Model. Numer. Anal.* **33**:1187–1202.

S. Chen, D. Shi, and Y. Zhao (2004). 'Anisotropic interpolation and quasi-Wilson element for narrow quadrilateral meshes'. *IMA J. Numer. Anal.* **24**:77–95.

S. Chen, L. Yin, and S. Mao (2008). 'An anisotropic, superconvergent nonconforming plate finite element'. *J. Comput. Appl. Math.* **220**:96–110.

S. Chen, Y. Zhao, and D. Shi (2003). 'Anisotropic interpolations with application to nonconforming elements'. *Appl. Numer. Math.* **49**:135–152.

P. Ciarlet and P.-A. Raviart (1972). 'General Lagrange and Hermite interpolation in $\mathbb{R}^n$ with application to finite elements'. *Arch. Ration. Mech. Anal.* **46**:177–199.

P. G. Ciarlet (1978). *The finite element method for elliptic problems.* North-Holland, Amsterdam. Reprinted by SIAM, Philadelphia, 2002.

P. G. Ciarlet (1991). 'Basic error estimates for elliptic problems'. In P. G. Ciarlet and J. L. Lions (eds.), *Finite element methods (Part 1)*, vol. II of *Handbook of Numerical Analysis*, pp. 17–351. North-Holland, Amsterdam.

P. Clément (1975). 'Approximation by finite element functions using local regularization'. *RAIRO Anal. Numer.* **2**:77–84.

Martin Costabel, Monique Dauge, and Serge Nicaise (2012). 'Analytic regularity for linear elliptic systems in polygons and polyhedra'. *Math. Models Methods Appl. Sci.* **22**(8):1250015, 63.

A. Cottrell, T.J.R. Hughes, and Y. Bazilevs (2009). *Isogeometric Analysis: Toward Integration of CAD and FEA.* Wiley.

M. Dauge (1988). *Elliptic boundary value problems on corner domains – smoothness and asymptotics of solutions*, vol. 1341 of *Lecture Notes in Mathematics*. Springer, Berlin.

P.J. Davis (1974). *Interpolation and Approximation.* Dover.

L. Demkowicz (2008). 'Polynomial exact sequences and projection-based interpolation with applications to Maxwell's equations'. In D. Boffi, F. Brezzi, L. Demkowicz, L.F. Durán, R. Falk, and M. Fortin (eds.), *Mixed Finite Elements, Compatibility Conditions, and Applications*, vol. 1939 of *Lectures Notes in Mathematics*. Springer Verlag.

L. Demkowicz and A. Buffa (2005). 'H^1, H(curl) and H(div)-conforming projection-based interpolation in three dimensions. Quasi-optimal p-interpolation estimates'. *Comput. Methods Appl. Mech. Engrg.* **194**(2-5):267–296.

L. Demkowicz, J. Kurtz, D. Pardo, M. Paszyński, W. Rachowicz, and A Zdunek (2008). *Computing with hp-adaptive finite elements. Vol. 2.* Chapman & Hall/CRC Applied Mathematics and Nonlinear Science Series. Chapman & Hall/CRC, Boca Raton, FL. Frontiers: three dimensional elliptic and Maxwell problems with applications.

J. Deny and J.-L. Lions (1953/54). 'Les espaces du type de Beppo Levi'. *Ann. Inst. Fourier* **5**:305–370.

R.A. DeVore and G.G. Lorentz (1993). *Constructive Approximation.* Springer Verlag.

M. Dobrowolski (1998). 'Finite Elemente'. Lecture Notes, Universität Würzburg.

M. Dubiner (1991). 'Spectral methods on triangles and other domains'. *J. Sci. Comp.* **6**:345–390.

T. Dupont and R. Scott (1980). 'Polynomial approximation of functions in Sobolev spaces'. *Math. Comp.* **34**:441–463.

T. Eibner and J.M. Melenk (2007). 'An adaptive strategy for hp-FEM based on testing for analyticity'. *Computational Mechanics* **39**:575–595.

A. Ern and J.-L. Guermond (2015a). 'Finite element quasi-interpolation and best approximation'. Preprint arXiv:1505.0693v2 [math.NA], arXiv.

A. Ern and J.-L. Guermond (2015b). ‘Mollification in strongly Lipschitz domains with application to continuous and discrete de Rham complex’. Preprint arXiv:1509.01325 [math.NA], arXiv.

L. Formaggia and S. Perotto (2001). ‘New anisotropic a priori error estimates’. *Numer. Math.* **89**:641–667.

R. Fritzsch (1990). *Optimale Finite-Elemente-Approximationen für Funktionen mit Singularitäten.* Ph.D. thesis, TU Dresden.

R. Fritzsch and P. Oswald (1988). ‘Zur optimalen Gitterwahl bei Finite-Elemente-Approximationen’. *Wissenschaftliche Zeitschrift TU Dresden* **37**(3):155–158.

E. Gagliardo (1957). ‘Caratterizzazione delle tracce sulla frontiera relative ad alcune classi di funzioni in n variabili’. *Rend. Sem. Mat. Univ. Padova* **27**:284–305.

V. Girault and P.-A. Raviart (1986). *Finite element methods for Navier-Stokes equations. Theory and algorithms*, vol. 5 of *Springer Series in Computational Mathematics*. Springer, Berlin.

V. Girault and L. R. Scott (2002). ‘Hermite interpolation of nonsmooth functions preserving boundary conditions’. *Math. Comp.* **71**:1043–1074.

W.J. Gordon and Ch.A. Hall (1973). ‘Transfinite Element Methods: Blending Function Interpolation over Arbitrary Curved Element Domains’. *Numer. Math.* **21**:109–129.

J. A. Gregory (1975). ‘Error bounds for linear interpolation in triangles’. In J. R. Whiteman (ed.), *The Mathematics of Finite Elements and Applications II*, pp. 163–170. Academic Press, London.

P. Grisvard (1985). *Elliptic problems in nonsmooth domains*, vol. 24 of *Monographs and Studies in Mathematics*. Pitman, Boston.

U. Hetmaniuk and P. Knupp (2008). ‘Local anisotropic interpolation error estimates based on directional derivatives along edges’. *SIAM J. Numer. Anal.* **47**:575–595.

P. Houston, C. Schwab, and E. Süli (2002). ‘Discontinuous *hp*-finite element methods for advection-diffusion-reaction problems’. *SIAM J. Numer. Anal.* **39**(6):2133–2163.

T. J. R. Hughes (1987). *The finite element method. Linear static and dynamic finite element analysis.* Prentice Hall, Englewood Cliffs, NJ.

P. Jamet (1976). ‘Estimations d’erreur pour des éléments finis droits presque dégénérés’. *R.A.I.R.O. Anal. Numér.* **10**:43–61.

P. Jamet (1977). ‘Estimation of the interpolation error for quadrilateral finite elements which can degenerate into triangles’. *SIAM J. Numer. Anal.* **14**:925–930.

M. Karkulik and J.M. Melenk (2014). ‘local high order regularization and applications to *hp*-methods (extended version)’. Tech. rep. arXiv:1411.5209.

M. Karkulik and J. M. Melenk (2015). ‘Local high-order regularization and applications to *hp*-methods’. *Comput. Math. Appl.* **70**(7):1606–1639.

G.E. Karniadakis and S.J. Sherwin (1999). *Spectral/hp Element Methods for CFD.* Oxford University Press, New York.

T. Koornwinder (1975). ‘Two-variable analogues of the classical orthogonal polynomials’. In *Theory and application of special functions (Proc. Advanced Sem., Math. Res. Center, Univ. Wisconsin, Madison, Wis., 1975)*, pp. 435–495. Math. Res. Center, Univ. Wisconsin, Publ. No. 35. Academic Press, New York.

R. Kornhuber and R. Roitzsch (1990). ‘On adaptive grid refinement in the presence of internal and boundary layers’. *IMPACT of Computing in Sci. and Engrg.* **2**:40–72.

V. A. Kozlov, V. G. Maz’ya, and J. Roßmann (2001). *Spectral Problems Associated with Corner Singularities of Solutions to Elliptic Equations.* American Mathematical Society, Providence, RI.

A. Kufner and A.-M. Sändig (1987). *Some Applications of Weighted Sobolev Spaces.* Teubner, Leipzig.

Y. Maday (1989). ‘Relèvement de traces polyômiales et interpolations hilbertiennes entres espaces de polynômes’. *C.R. Acad. Sci. Paris, Série I* **309**:463–468.

J.M. Melenk (2002). *hp finite element methods for singular perturbations*, vol. 1796 of *Lecture Notes in Mathematics*. Springer Verlag.

J.M. Melenk (2005a). '*hp*-interpolation of nonsmooth functions and an application to *hp* a posteriori error estimation'. *SIAM J. Numer. Anal.* **43**:127–155.

J.M. Melenk (2005b). 'On approximation in meshless methods'. In J. Blowey and A. Craig (eds.), *Frontier in Numerical Analysis, Durham 2004*, pp. 65–141. Springer Verlag.

J.M. Melenk, A. Parsania, and S. Sauter (2013). 'General DG-methods for highly indefinite Helmholtz problems'. *J. Sci. Comp.* **57**:536–581.

J.M. Melenk and S. Sauter (2010). 'Convergence Analysis for Finite Element Discretizations of the Helmholtz equation with Dirichlet-to-Neumann boundary conditions'. *Math. Comp.* **79**:1871–1914.

J.M. Melenk and T. Wurzer (2014). 'On the stability of the boundary trace of the polynomial L^2-projection on triangles and tetrahedra'. *Comput. Math. Appl.* **67**:944–965.

P. Ming and Z. Shi (2002). 'Quadrilateral mesh'. *Chin. Ann. of Math., Ser. B* **23**:235–252.

R. Muñoz-Sola (1997). 'Polynomial liftings on a tetrahedron and applications to the *hp*-version of the finite element method in three dimensions'. *SIAM J. Numer. Anal.* **34**(1):282–314.

S. A. Nazarov and B. A. Plamenevsky (1994). *Elliptic problems in domains with piecewise smooth boundary*, vol. 13 of *de Gruyter Expositions in Mathematics*. de Gruyter, Berlin.

L. A. Oganesyan and L. A. Rukhovets (1979). *Variational-difference methods for the solution of elliptic equations*. Izd. Akad. Nauk Armyanskoi SSR, Jerevan. In Russian.

P. Oswald (1994). *Multilevel Finite Element Approximation: Theory and Applications*. Teubner, Stuttgart.

J. Peraire, M. Vahdati, K. Morgan, and O. C. Zienkiewicz (1987). 'Adaptive remeshing for compressible flow computation'. *J. Comp. Phys.* **72**:449–466.

F. Rapetti, A. Sommariva, and M. Vianello (2012). 'On the generation of symmetric Lebesgue-like points in the triangle'. *J. Comput. Appl. Math.* **236**(18):4925–4932.

G. Raugel (1978). 'Résolution numérique par une méthode d'éléments finis du problème de Dirichlet pour le Laplacien dans un polygone'. *C. R. Acad. Sci. Paris, Sér. A* **286**(18):A791–A794.

H.-G. Roos, M. Stynes, and L. Tobiska (1996). *Numerical methods for singularly perturbed differential equations. Convection-diffusion and flow problems*. Springer, Berlin.

K. Scherer (1981). 'On optimal global error bounds obtained by scaled local error estimates'. *Numer. Math.* **36**:151–176.

J. Schöberl (2001). 'Commuting quasi-interpolation operators for mixed finite elements'. Preprint ISC-01-10-MATH, Texas A&M University.

D. Schötzau, Ch. Schwab, and T. P. Wihler (2013). '*hp*-DGFEM for second order elliptic problems in polyhedra II: Exponential convergence'. *SIAM J. Numer. Anal.* **51**(4):2005–2035.

Ch. Schwab (1998). *p- and hp-finite element methods*. Numerical Mathematics and Scientific Computation. The Clarendon Press Oxford University Press, New York. Theory and applications in solid and fluid mechanics.

L. R. Scott and S. Zhang (1990). 'Finite element interpolation of non-smooth functions satisfying boundary conditions'. *Math. Comp.* **54**:483–493.

E.M. Stein (1970). *Singular integrals and differentiability properties of functions*. Princeton University Press.

B. Sündermann (1980). 'Lebesgue constants in Lagrangian interpolation at the Fekete points'. Ergebnisberichte der Lehrstühle Mathematik III und VIII (Angewandte Mathematik) 44, Universität Dortmund.

B. Sündermann (1983). 'Lebesgue constants in Lagrangian interpolation at the Fekete points'. *Mitt. Math. Ges. Hamb.* **11**:204–211.

J. L. Synge (1957). *The hypercircle in mathematical physics.* Cambridge University Press, Cambridge.

G. Szegő (1975). *Orthogonal Polynomials.* American Mathematical Society, fourth edn.

L. Tartar (2007). *An introduction to Sobolev spaces and interpolation spaces*, vol. 3 of *Lecture Notes of the Unione Matematica Italiana*. Springer, Berlin.

A. Toselli and O. Widlund (2005). *Domain Decomposition Methods — Algorithms and Theory.* Springer Verlag.

L. N. Trefethen (2013). *Approximation theory and approximation practice.* Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA.

A.H. Turetskii (1940). 'The bounding of polynomials prescribed at equally distributed points'. *Proc. Pedag. Inst. Vitebsk* **3**:117127. (Russian).

R. Verfürth (1999a). 'Error estimates for some quasi-interpolation operators'. *Math. Model. Numer. Anal.* **33**:695–713.

R. Verfürth (1999b). 'A note on polynomial approximation in Sobolev spaces'. *Math. Model. Numer. Anal.* **33**:715–719.

G. Zhou and R. Rannacher (1993). 'Mesh orientation and anisotropic refinement in the streamline diffusion method'. In M. Křížek, P. Neitaanmäki, and R. Stenberg (eds.), *Finite Element Methods: Fifty Years of the Courant Element*, vol. 164 of *Lecture Notes in Pure and Applied Mathematics*, pp. 491–500. Marcel Dekker, Inc., New York. Also published as Preprint 93-57, Universität Heidelberg, IWR, SFB 359, 1993.

O. C. Zienkiewicz and J. Wu (1994). 'Automatic directional refinement in adaptive analysis of compressible flows'. *Int. J. Numer. Methods Engrg.* **37**:2189–2210.

M. Zlámal (1968). 'On the finite element method'. *Numer. Math.* **12**:394–409.