On the Reference Mapping for Quadrilateral and Hexahedral Finite Elements on Multilevel Adaptive Grids

Gunar Matthies* Friedhelm Schieweck†

15.11.2006

Abstract

We study the properties of the reference mapping for quadrilateral and hexahedral finite elements. We consider multilevel adaptive grids with possibly hanging nodes which are typically generated by adaptive refinement starting from a regular coarse grid. It turns out that for such grids the reference mapping behaves – up to a perturbation depending on the mesh size – like an affine mapping. As an application, we prove optimal estimates of the interpolation error for discontinuous mapped $\mathbb{P}_r$-elements on quadrilateral and hexahedral grids.

Keywords: quadrilateral and hexahedral finite elements, multilevel adaptive grids, hanging nodes, finite element analysis

Mathematics Subject Classification (2000): 65N15, 65N30, 65N50

1 Introduction

In the finite element analysis, often the assumption is used that the finite elements are affine equivalent [2]. For triangular or tetrahedral elements, this is not a restriction for

*Fakultät für Mathematik, Ruhr-Universität Bochum, Universitätsstraße 150, D-44780 Bochum, Germany, URL: http://www.ruhr-uni-bochum.de/jpnum

†Institut für Analysis und Numerik, Otto-von-Guericke-Universität Magdeburg, Postfach 4120, 39016 Magdeburg, Germany, URL: http://www-ian.math.uni-magdeburg.de/~schiewec/

the shape of the elements. However, for quadrilateral or hexahedral elements, an affine reference mapping means that only parallelograms or parallelepipeds are admitted for the shape of the elements. This is too restrictive for practical applications. In order to handle general quadrilateral or hexahedral elements, one needs a non-affine multilinear reference mapping.

For applications like discontinuous Galerkin methods or discontinuous pressure approximations for the Navier–Stokes problem, the approximation properties of discontinuous mapped $\mathbb{P}_r$-elements on quadrilateral or hexahedral grids are of interest. With mapped $\mathbb{P}_r$-elements we denote finite elements where the associated basis functions on the reference element are polynomials of a total degree not greater than r. Here it is known from the literature [1, 6] that there will be a dramatic decrease of the order of the interpolation error on general quadrilateral or hexahedral grids. In contrast, one gets the optimal order on grids consisting of affine equivalent elements, see [2]. In this paper, we investigate quadrilateral and hexahedral finite elements on *multilevel adaptive grids* which are generated from an arbitrary shape regular coarse grid by adaptive refinement allowing hanging nodes. In applications, these grids are well suited to resolve geometric structures of different scales. We prove that the reference mapping on multilevel adaptive grids behaves – up to a perturbation depending on the mesh size – like an affine mapping. As a consequence, we show the optimal order of the interpolation error for the discontinuous mapped $\mathbb{P}_r$-element. A rigorous proof of this fact was missing until now.

In this paper, we consider two types of adaptive refinement. The first one is suited for polygonal or polyhedral domains where the coarse grid describes the boundary exactly such that no boundary adaption is needed during refinement. The second type is suited for domains with curved boundaries where, for elements at the boundary, a modification of the refinement is necessary. For this refinement with boundary adaption, we prove – up to a logarithmic factor – the same optimal local estimates of the interpolation error as for the refinement without boundary adaption. Furthermore, we can show optimal global interpolation error estimates under additional assumptions.

The outline of this paper is as follows. In Section 2, we introduce some notations and prove that, for a fixed element K, certain quantities of the reference mapping F_K are invariant with respect to the different possible choices of F_K. Furthermore, we describe the concept of *multilevel adaptive grids* which are considered in this paper. We prove estimates for the derivatives of the reference mapping for refinement without boundary adaption in

Section 3 and for refinement with boundary adaption in Section 4. These estimates are used in Section 5 to prove optimal and quasi optimal estimates of the interpolation error under the assumption that the finite dimensional function space on the reference element contains the space $\mathbb{P}_r$ of the polynomials with total degree of at most r.

2 Preliminaries and Notations

2.1 General Notations

We do not distinguish between the notation of scalar and vector-valued functions $v : \Omega \to \mathbb{R}^n$ where $\Omega \subset \mathbb{R}^d$, $d \in \{2, 3\}$, is a bounded domain, $n = 1$ for scalar functions and $n = d$ for vector-valued functions. For a set $G \subset \Omega$, let $|\cdot|_{m,p,G}$ and $\|\cdot\|_{m,p,G}$ denote the usual semi-norm and norm in the Sobolev spaces $W^{m,p}(G)$ and $(W^{m,p}(G))^d$, respectively. For the Hilbert spaces $H^m(G)$ and $(H^m(G))^d$, we omit the index p and denote the semi-norm and norm by $|\cdot|_{m,G}$ and $\|\cdot\|_{m,G}$. By $\mathbb{P}_m(G)$ we denote the space of all polynomials on the domain $G \subset \mathbb{R}^d$ with total degree less than or equal to m and by $\mathbb{Q}_m(G)$ the space of those polynomials where the maximum power in each coordinate is less than or equal to m. For $p \in [1, \infty]$, both the usual p-norm of a vector in $\mathbb{R}^d$ and the induced p-norm of a $d \times d$-matrix are denoted by $\|\cdot\|_p$. For a set $G \subset \mathbb{R}^d$, we denote by $\mathrm{int}(G)$ and $\overline{G}$ the interior and closure of G, respectively. Throughout this paper, C will denote a generic constant which may have different values at different places. All these constants occurring inside of any estimate will be independent of the local and global mesh parameters h_K and h, respectively, which will be defined below.

2.2 Mesh Notation and Shape Regularity

Let the bounded domain $\Omega \subset \mathbb{R}^d$ be decomposed into a set of elements $K \in \mathcal{T}_h$ which are assumed to be open quadrilaterals in the 2D-case and open hexahedrons in the 3D-case such that $\overline{\Omega} = \cup_{K \in \mathcal{T}_h} \overline{K}$. For an element $K \in \mathcal{T}_h$, we denote by h_K the diameter of K. The *mesh size* h of $\mathcal{T}_h$ is defined by $h := \max_{K \in \mathcal{T}_h} h_K$. For each $K \in \mathcal{T}_h$, we denote by F_K the reference mapping between the reference element $\widehat{K} := (-1, 1)^d$ and the original element $K = F_K(\widehat{K})$. In this paper, we assume that F_K is multilinear, i.e. $F_K \in (\mathbb{Q}_1(\widehat{K}))^d$. Note that this implies in the 3D-case that the faces of a hexahedral element $K = F_K(\widehat{K})$ are curved in general. The reason is the following. If one prescribes the 8 vertices of a

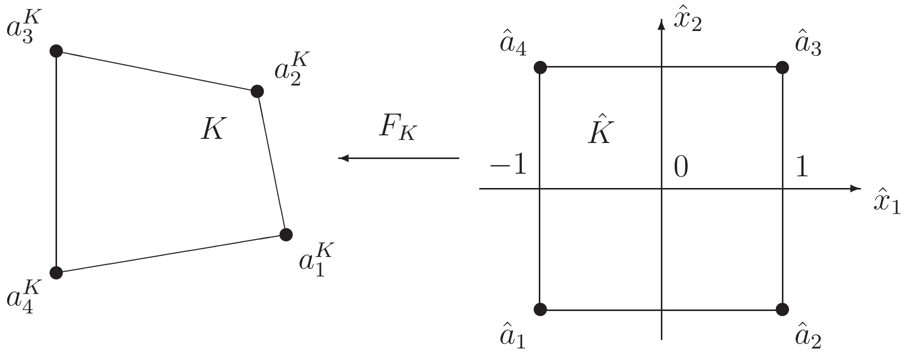

Figure 1: *Numbering of the vertices of element K and reference element $\widehat{K}$ in 2D.*

hexahedral element K then the set of the 4 vertices of a face E of K in general need not to be contained in a plane. Thus, such a face E must be curved. Therefore, we define in this paper an element K of the mesh $\mathcal{T}_h$ in the following way. At first, we prescribe (like most mesh generators do) the set of the vertices of K. Then, to each vertex we assign a local number such that the vertices of K can be denoted by a_j^K, $j \in \{1, \ldots, 2^d\}$. Finally, we determine in a unique way the reference mapping $F_K \in (\mathbb{Q}_1(\widehat{K}))^d$ by the conditions

$$F_K(\hat{a}_j) = a_j^K \qquad \forall\, j \in \{1, \ldots, 2^d\}, \tag{1}$$

where $\hat{a}_j$ denote the vertices of $\widehat{K}$, and define the element K by means of the reference mapping as $K := F_K(\widehat{K})$, see Figures 1 and 2. Thus, the set of the vertices of an element

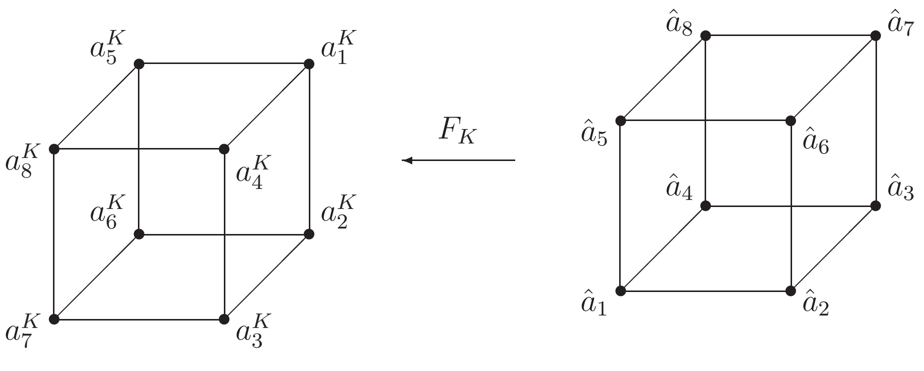

Figure 2: *Vertex numbering of original element K and reference element $\widehat{K}$ in 3D.*

K and their local numbering determine the reference mapping F_K. It is clear that not each numbering yields a reasonable bijective mapping $F_K : \widehat{K} \to K$. Therefore, we make the following assumption on the local vertex numbering of the elements which is not restrictive from the practical point of view. In the 2D-case, we only assume that the vertices are numbered counterclockwise, see Figure 1. In the 3D-case, we choose first one of the 6 faces of K, called E_1, as the face with the vertices $a_1^K, \ldots, a_4^K$. Then we take one of the 4 vertices of E_1 as the vertex a_1^K and number the remaining vertices of E_1 in a counterclockwise sense (looking from inside of K at E_1). The remaining 4 vertices on the face opposite to E_1 are numbered in relation to $a_1^K, \ldots, a_4^K$ as shown in Figure 2.

In the following, we will describe the correlation between two different reference mappings $F_K, \widetilde{F}_K : \widehat{K} \to K$ that satisfy the above assumptions on the vertex numbering of an element $K \in \mathcal{T}_h$. Let $\pi : \{1, \ldots, 2^d\} \to \{1, \ldots, 2^d\}$ denote a bijective mapping that connects the two different vertex numberings corresponding to F_K and $\widetilde{F}_K$ in the sense that

$$\tilde{a}_j^K = \widetilde{F}_K(\hat{a}_j) = a_{\pi(j)}^K = F_K(\hat{a}_{\pi(j)}) \qquad \forall\, j \in \{1, \ldots, 2^d\}. \tag{2}$$

Then, from the assumption of the vertex numbering, it follows that there exists an orthogonal matrix $U \in \mathbb{R}^{d\times d}$ such that $\hat{a}_{\pi(j)} = U\hat{a}_j$ for all $j = 1, \ldots, 2^d$. In the 2D-case, this matrix corresponds to a rotation of the coordinate system with an angle of $\frac{k\pi}{2}$, $k \in \{0, 1, 2, 3\}$, which moves the vertex $\hat{a}_1$ to the vertex $\hat{a}_{\pi(1)}$. Therefore, the matrix entries in each row and column of U are zero except for exactly one entry which is equal to $+1$ or -1. This implies

$$\|U\|_p \le 1 \qquad \forall\, p \in [1, \infty]. \tag{3}$$

In the 3D-case, U can be written as a product of two matrices $U = U_2U_1$ where U_1 corresponds to a rotation of the coordinate system that moves the "ground face" $\widehat{E}_1$ of $\widehat{K}$ with the vertices $\hat{a}_1, \ldots, \hat{a}_4$ to that face $\widetilde{E}_1$ of $\widehat{K}$ that contains the vertices $\hat{a}_{\pi(1)}, \ldots, \hat{a}_{\pi(4)}$. The matrix U_2 corresponds to a rotation around the coordinate axis orthogonal to $\widetilde{E}_1$ such that the vertex $U_1\hat{a}_1$ of $\widetilde{E}_1$ is moved to the right position, i.e. $\hat{a}_{\pi(1)} = U_2U_1\hat{a}_1$. Again the matrix entries in each row and column of U_1 and U_2 are zero except for exactly one entry which is equal to $+1$ or -1. Thus, we get $\|U_i\|_p \le 1$ for $i \in \{1, 2\}$ which implies (3). Now, the properties (2) and $\hat{a}_{\pi(j)} = U\hat{a}_j$, $j = 1, \ldots, 2^d$, as well as the interpolation conditions (1) applied to $F_K, \widetilde{F}_K \in (\mathbb{Q}_1(\widehat{K}))^d$ yield the following relation between the reference mappings F_K and $\widetilde{F}_K$:

$$\widetilde{F}_K(\hat{x}) = F_K(U\hat{x}) \qquad \forall\, x \in \widehat{K}. \tag{4}$$

This relation is important for the proof that certain quantities of the reference mapping are invariant with respect to the special choice of an admissible mapping F_K (see Lemma 2).

In order to guarantee that the reference mapping $F_K : \widehat{K} \to K$ is bijective for all $K \in \mathcal{T}_h$ and that the finite elements will have good approximation properties, we assume that the mesh is *shape regular* in the sense of [8]. In the following, we describe this shape regularity assumption in more detail. By a Taylor expansion of $F_K(\hat{x})$ at the point $\hat{x} = 0$ we get

$$F_K(\hat{x}) = b_K + B_K\,\hat{x} + G_K(\hat{x}), \tag{5}$$

with $b_K := F_K(0)$, $B_K := DF_K(0)$ and $G_K(\hat{x}) := F_K(\hat{x}) - F_K(0) - B_K\hat{x}$. We denote by $\widehat{S} \subset \widehat{K}$ the d-simplex with the vertices 0, $e_1, \ldots e_d$, where $e_i \in \mathbb{R}^d$ denotes the i-th unit vector. By S_K we denote the d-simplex which is the image of $\widehat{S}$ under the affine mapping $\hat{x} \mapsto B_K\,\hat{x} + b_K$. In [8] it is assumed that the simplices S_K satisfy the usual shape regularity condition

$$\frac{h_{S_K}}{\rho_{S_K}} \le C \qquad \forall K \in \mathcal{T}_h, \tag{6}$$

where h_{S_K} denotes the diameter of S_K and ρ_{S_K} the diameter of the largest ball that can be inscribed into S_K. However, we will use here another assumption on the matrices $B_K = DF_K(0)$ which can be proved to be equivalent to (6). We assume that, for any $p \in [1, \infty]$, there exists an h-independent constant C_s such that

$$\|B_K\|_p \le C_s h_{S_K}, \qquad \|B_K^{-1}\|_p \le C_s h_{S_K}^{-1} \qquad \forall K \in \mathcal{T}_h. \tag{7}$$

For an element $K \in \mathcal{T}_h$ and $p \in [1, \infty]$, we define the quantity $\gamma_p(K)$ by

$$\gamma_p(K) := \sup_{\hat{x} \in \widehat{K}} \|B_K^{-1}\, DF_K(\hat{x}) - I\|_p\,, \tag{8}$$

which is a measure of the deviation of K from a parallelogram or a parallelepiped, respectively. Note that $\gamma_p(K) = 0$ is equivalent to the fact that the mapping F_K is affine.

Definition 1 *A family $\{\mathcal{T}_h\}$ of meshes consisting of quadrilateral or hexahedral elements is called* **shape regular** *if there exist a fixed $p \in [1, \infty]$ and h-independent constants C_s and γ_0 such that* (7) *and the estimates*

$$\gamma_p(K) \le \gamma_0 < 1 \qquad \forall K \in \mathcal{T}_h \tag{9}$$

are satisfied.

In this paper, we assume that the mesh $\mathcal{T}_h$ is shape regular in the sense of Definition 1. It has been proven in [8] that for such a shape regular mesh the reference mappings $F_K : \widehat{K} \to K$ are bijective for all elements $K \in \mathcal{T}_h$. Furthermore, one can find in [8, Sect. 2] sufficient conditions that guarantee the fulfillment of the assumption (9). These conditions can be easily checked in practical computations. In the 2D-case, the condition $\gamma_1(K) < 1$ is equivalent to the fact that the quadrilateral element K is convex.

In the next lemma, the point is investigated that, for a given element K, one can use different reference mappings $F_K : \widehat{K} \to K$. We prove that certain quantities are invariant with respect to the special choice of F_K.

Lemma 2 *Let $F_K, \widetilde{F}_K \in (\mathbb{Q}_1(\widehat{K}))^d$ be two different reference mappings from $\widehat{K}$ to K which are defined by* (1) *with two admissible numberings of the vertices of $K \in \mathcal{T}_h$. Then, for all integers $m \geq 0$ it holds*

$$|\widetilde{F}_K|_{m,\infty,\widehat{K}} = |F_K|_{m,\infty,\widehat{K}}. \tag{10}$$

Furthermore, the value of $\gamma_p(K)$ defined in (8) *with the reference mapping F_K does not change if F_K and $B_K = DF_K(0)$ are replaced by $\widetilde{F}_K$ and $\widetilde{B}_K = D\widetilde{F}_K(0)$, respectively.*

Proof. We know from the derivation of (4) that there exists an orthogonal matrix U such that $\widetilde{F}_K(\hat{x}) = F_K(U\hat{x})$ for all $\hat{x} \in \widehat{K}$. At first we prove the lemma for the case where the matrix U has the special structure

$$U = (U_{k,j}) \in \mathbb{R}^{d\times d} \qquad \text{with} \qquad U_{k,j} = \begin{cases} s_j & \text{if} \quad k = \mu(j), \\ 0 & \text{if} \quad k \neq \mu(j), \end{cases} \qquad s_j \in \{-1, +1\}, \tag{11}$$

where $\mu(\cdot)$ denotes a bijective mapping $\mu : \{1, \dots, d\} \to \{1, \dots, d\}$. We are in this situation if $d = 2$ or if, in the case $d = 3$, where $U = U_2U_1$ (see the explanations preceding (4)) with $U_2 = I$.

Let us start to prove (10) in the case $m = 0$. Since the mapping $\hat{x} \mapsto U\hat{x}$ is bijective from $\widehat{K}$ onto $\widehat{K}$, we have for the components $\widetilde{F}_{K,i}$ and $F_{K,i}$, $i = 1, \dots, d$, of the reference mappings $\widetilde{F}_K$ and F_K that

$$\sup_{\hat{x}\in\widehat{K}} |\widetilde{F}_{K,i}(\hat{x})| = \sup_{U\hat{x}\in\widehat{K}} |F_{K,i}(U\hat{x})| = \sup_{\hat{x}\in\widehat{K}} |F_{K,i}(\hat{x})| \qquad \forall i \in \{1, \dots, d\}$$

which proves (10) for $m = 0$.

Now, in order to prove (10) for $m = 1$, we apply the chain rule to $\widetilde{F}_K(\hat{x}) = F_K(U\hat{x})$, utilize (11) and get for all $i, j \in \{1, \dots, d\}$

$$\frac{\partial \widetilde{F}_{K,i}}{\partial \hat{x}_j}(\hat{x}) = \frac{\partial F_{K,i}}{\partial \hat{x}_{\mu(j)}}(U\hat{x})s_j \qquad \forall\, \hat{x} \in \widehat{K}. \tag{12}$$

Hence,

$$\left|\frac{\partial \widetilde{F}_{K,i}}{\partial \hat{x}_j}\right|_{0,\infty,\widehat{K}} = \left|\frac{\partial F_{K,i}}{\partial \hat{x}_{\mu(j)}}\right|_{0,\infty,\widehat{K}} \qquad \forall i,j \in \{1,\dots,d\},$$

which implies (10) for $m = 1$. If we apply to (12) again the chain rule and (11) then we can conclude (10) in an analogous way for $m = 2$. Thus, by continuation of this procedure we obtain (10) for an arbitrary positive integer m.

In the following, we show that the value of $\gamma_p(K)$ defined in (8) is invariant of the choice of F_K. Let us denote by $\widetilde{\gamma}_p(K)$ the value of $\gamma_p(K)$ from definition (8) where the mapping F_K and the matrix B_K are replaced by $\widetilde{F}_K$ and $\widetilde{B}_K = D\widetilde{F}_K(0)$, respectively. Then, the relation $\widetilde{F}_K(\hat{x}) = F_K(U\hat{x})$ gives

$$D\widetilde{F}_K(\hat{x}) = DF_K(U\hat{x})\,U \qquad \forall\, \hat{x} \in \widehat{K}$$

which implies $\widetilde{B}_K = B_K U$ and

$$\widetilde{\gamma}_p(K) = \sup_{\hat{x}\in\widehat{K}} \|\widetilde{B}_K^{-1}\, D\widetilde{F}_K(\hat{x}) - I\|_p = \sup_{\hat{x}\in\widehat{K}} \|U^T\big(B_K^{-1}\, DF_K(\hat{x}) - I\big)U\|_p = \gamma_p(K).$$

Here we have used the fact that for our special orthogonal matrix U given in (11) it holds that $\|M\|_p = \|MU\|_p$ and $\|M\|_p = \|U^T M\|_p$ for an arbitrary matrix $M \in \mathbb{R}^{d\times d}$. This is a simple consequence of the identities

$$\|Ux\|_p = \|x\|_p = \|U^T x\|_p \qquad \forall\, x \in \mathbb{R}^d.$$

At the end, we prove the lemma for the remaining case when the matrix U has a representation $U = U_2 U_1$ where both matrices U_1 and U_2 have the special structure of the matrix U in (11). We write the mapping $\widetilde{F}_K(\hat{x}) = F_K(U\hat{x})$ in the form

$$\widetilde{F}_K(\hat{x}) = \overline{F}_K(U_1\hat{x}) \qquad \text{with} \qquad \overline{F}_K(\hat{z}) := F_K(U_2\hat{z}),$$

where also $\overline{F}_K : \widehat{K} \to K$ is a reference mapping that satisfies the assumptions of the lemma. Then, the pairs of reference mappings $\widetilde{F}_K$ and $\overline{F}_K$ as well as $\overline{F}_K$ and F_K belong to the above treated case. Hence, we have

$$|\widetilde{F}_K|_{m,\infty,\widehat{K}} = |\overline{F}_K|_{m,\infty,\widehat{K}} = |F_K|_{m,\infty,\widehat{K}}$$

and $\widetilde{\gamma}_p(K) = \overline{\gamma}_p(K) = \gamma_p(K)$ where $\overline{\gamma}_p(K)$ denotes the corresponding value associated with the reference mapping $\overline{F}_K$. $\square$

2.3 Multilevel Adaptive Grids

In the following, we will describe the type of grids that is treated in this paper. We assume that $\mathcal{T}_h$ is a *multilevel adaptive grid* generated by a refinement process in the following way. We start with a partition $\mathcal{T}^0$ of the domain Ω into elements $K \in \mathcal{T}^0$ of grid-level 0, i.e. $\Omega = \mathrm{int}\big(\bigcup_{K\in\mathcal{T}^0} \overline{K}\big)$. The grid $\mathcal{T}^0$ is assumed to be *regular* in the usual sense, i.e., for any two different elements $K_1, K_2 \in \mathcal{T}^0$ the intersection $\overline{K_1} \cap \overline{K_2}$ is either empty or a common $(d-m)$-dimensional face of K_1 and K_2 where $m \in \{1, \dots, d\}$. Now, starting with the elements $K \in \mathcal{T}^0$, an existing element K can be refined, i.e., it can be splitted into 2^d many new elements called *son-elements* which are denoted by $\sigma_i(K)$, $i = 1, \dots, 2^d$, see Figure 3. Let $\sigma(K)$ denote the set of all son-elements $\sigma_i(K)$, $i = 1, \dots, 2^d$.

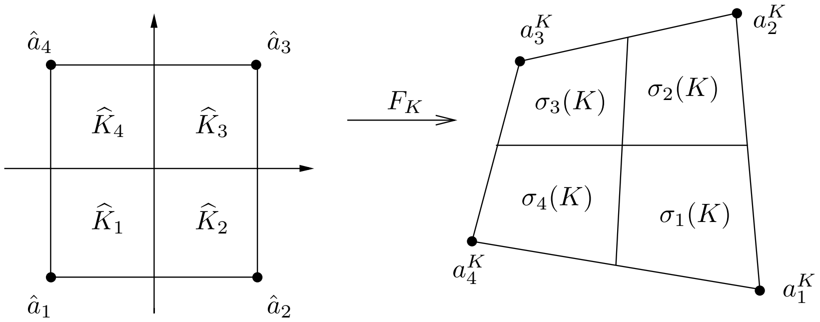

Figure 3: *Refinement of element K into son-elements $\sigma_i(K)$, $i = 1, \dots, 4$, without boundary adaption.*

These son-elements are defined as follows. Let $\widehat{K}_i$, $i \in \{1, \dots, 2^d\}$, be the son-elements of the reference element $\widehat{K} = (-1, 1)^d$ which are congruent d-cubes with $|\widehat{K}_i| = 1$. Then, the i-th son-element $\sigma_i(K)$ of the original element K is defined as

$$\sigma_i(K) := F_K(\widehat{K}_i) \qquad \forall\, i = 1, \dots, 2^d, \tag{13}$$

if K has no face E located at such a boundary part of the domain Ω which cannot be exactly represented by a $\mathbb{Q}_1$-element face. For elements K having such a boundary face, we will describe and analyze the decomposition of K into son-elements in Section 4. Note that in the 3D-case, even if an element K has only planar faces, it can happen that the son-elements $\sigma_i(K)$ have curved faces in the interior of K (see [7, Remark 3.3] for an example).

For a new element $K' = \sigma_i(K)$, we will say that K is the *father-element* of K' and we will write $K = \mathcal{F}(K')$. If an element K is refined then it is replaced in the partition of the domain Ω by the set $\sigma(K)$ of its son-elements $\sigma_i(K)$, $i = 1, \ldots, 2^d$. The new elements can be refined again and again until the final partition $\mathcal{T}_h$ of Ω is created. Examples of such grids used in practical computations can be found in [4, 5].

Definition 3 *For an element $K \in \mathcal{T}_h$, generated from the initial grid $\mathcal{T}^0$ by the refinement process described above, we define the* **refinement level** *$\ell(K)$ as $\ell(K) := 0$ if $K \in \mathcal{T}^0$ and $\ell(K) := m \geq 1$ if there exists a chain of m father-elements K_i, $i = m-1, \ldots, 0$, starting from $K_m := K$ and defined by $K_{i-1} := \mathcal{F}(K_i)$ for $i = m, \ldots, 1$, such that $K_0 \in \mathcal{T}^0$. The grid $\mathcal{T}_h$, generated by the above defined refinement process, is called* **m-irregular** *if the condition*

$$|\ell(K) - \ell(K')| \leq m$$

holds for any pair of different neighbored elements $K, K' \in \mathcal{T}_h$ where $\partial K \cap \partial K'$ is a one- or two-dimensional manifold.

The above defined *refinement level* $\ell(K)$ is equal to the number of refinement steps that are needed to generate the element K from an element of the coarsest grid $\mathcal{T}^0$. In most practical computations 1-irregular grids are used. They are sufficient to describe complicated structures in a reasonable way, see e.g. [4, 5]. An example that shows the capabilities of 1-irregular quadrilateral grids to resolve an inner fine scale structure is presented in Figure 4.

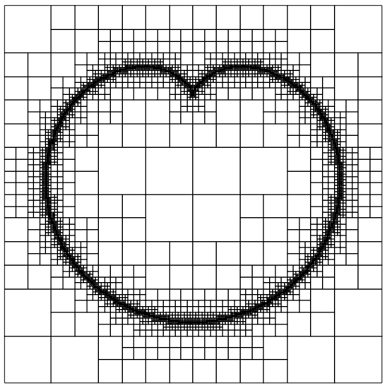

Figure 4: *Resolution of an inner fine scale structure by a multilevel adaptive 1-irregular grid.*

3 Refinement Without Boundary Adaption

In this section we assume that the coarse grid $\mathcal{T}^0$ describes the boundary of the domain exactly such that a boundary adaption during the refinement process is not necessary.

Let us introduce the set $\mathcal{T}_h^l$ of all grid cells K of the final mesh $\mathcal{T}_h$ that have been created by means of l refinement steps from a coarse grid cell of $\mathcal{T}^0$, i.e.

$$\mathcal{T}_h^l := \{K \in \mathcal{T}_h : \ \ell(K) = l\}. \tag{14}$$

In the following, we will represent the reference mapping F_K of a son-element $K \in \sigma(K_0)$ of an element K_0 by means of the reference mapping F_{K_0} and an additional affine mapping. This representation will be the key idea of the subsequent analysis. Let $\widehat{K}_s := F_{K_0}^{-1}(K) \subset \widehat{K}$ be that son-element of $\widehat{K}$ that corresponds to the original son-element K of K_0 by means of the mapping F_{K_0}, see Figure 5. Furthermore, let a_j^K, $\hat{v}_j := F_{K_0}^{-1}(a_j^K)$ and $\hat{a}_j$, $j = 1, \ldots, 2^d$,

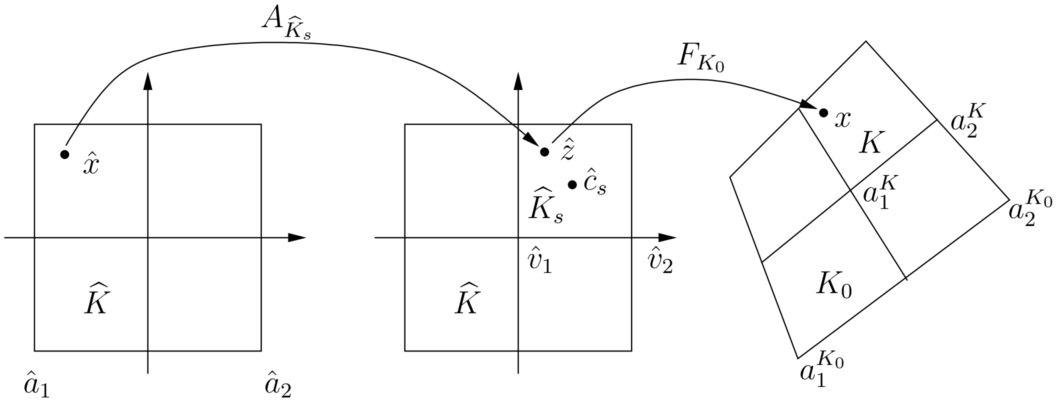

Figure 5: *Reference mapping for the son-element K of K_0.*

be the vertices of K, $\widehat{K}_s$ and $\widehat{K}$, respectively. We suppose that the vertex numbering of the son-element K is derived from the vertex numbering of K_0 in such a way that the vertices $\hat{v}_j$ of $\widehat{K}_s$ and $\hat{a}_j$ of $\widehat{K}$ are connected by the following affine mapping $A_{\widehat{K}_s} : \widehat{K} \to \widehat{K}_s$

$$\hat{v}_j = A_{\widehat{K}_s}(\hat{a}_j) \quad \forall\, j \in \{1, \ldots, 2^d\} \qquad \text{with} \qquad A_{\widehat{K}_s}(\hat{x}) := \hat{c}_s + \tfrac{1}{2}\,\hat{x} \quad \forall\, \hat{x} \in \widehat{K}, \tag{15}$$

where $\hat{c}_s$ denotes the center of $\widehat{K}_s$, see Figure 5. The reference mapping based on this vertex numbering will be denoted by F_K^*. The composed mapping $\hat{x} \mapsto \hat{z} = A_{\widehat{K}_s}(\hat{x}) \mapsto x = F_{K_0}(\hat{z})$ is multilinear in $\hat{x}$. Therefore, the mapping $F_{K_0} \circ A_{\widehat{K}_s}$, which maps the vertices $\hat{a}_j$ of $\widehat{K}$ onto the vertices a_j^K of K, is a mapping from the space $(\mathbb{Q}_1(\widehat{K}))^d$. Since the mapping

$F_K^* \in (\mathbb{Q}_1(\widehat{K}))^d$ is uniquely determined by $a_j^K = F_K^*(\hat{a}_j)$ for $j = 1, \ldots, 2^d$, we get that $F_K^* = F_{K_0} \circ A_{\widehat{K}_s}$, i.e.

$$F_K^*(\hat{x}) = F_{K_0}\big(\hat{c}_s + \tfrac{1}{2}\hat{x}\big) \qquad \forall\, \hat{x} \in \widehat{K}\,. \tag{16}$$

The following result plays a fundamental role for the subsequent analysis.

Lemma 4 *Let $K \in \sigma(K_0)$ be a son-element of an element K_0 and let $F_K : \widehat{K} \to K$ and $F_{K_0} : \widehat{K} \to K_0$ denote the associated reference mappings from $(\mathbb{Q}_1(\widehat{K}))^d$ satisfying the assumptions from Section 2. Then, for each nonnegative integer m, the estimate*

$$|F_K|_{m,\infty,\widehat{K}} \le \big(\tfrac{1}{2}\big)^m |F_{K_0}|_{m,\infty,\widehat{K}} \tag{17}$$

holds. For each element $K \in \mathcal{T}_h^l$ generated from the coarse grid element $K_0 \in \mathcal{T}^0$ with $K \subset K_0$, the estimate

$$|F_K|_{m,\infty,\widehat{K}} \le \big(2^{-l}\big)^m\, |F_{K_0}|_{m,\infty,\widehat{K}} \tag{18}$$

holds.

Proof. At first we suppose that $F_K = F_K^*$. Using the notation $F_\kappa(\hat{x}) = (F_{\kappa,1}(\hat{x}), \ldots, F_{\kappa,d}(\hat{x}))^T$ for $\kappa \in \{K, K_0\}$, we obtain from (16) by means of the chain rule

$$\frac{\partial F_{K,i}}{\partial \hat{x}_j}(\hat{x}) = \frac{1}{2}\frac{\partial F_{K_0,i}}{\partial \hat{x}_j}\big(\hat{c}_s + \tfrac{1}{2}\hat{x}\big)$$

which gives by induction

$$D^\alpha F_{K,i}(\hat{x}) = \big(\tfrac{1}{2}\big)^{|\alpha|} D^\alpha F_{K_0,i}\big(\hat{c}_s + \tfrac{1}{2}\hat{x}\big)$$

for all $i \in \{1, \ldots, d\}$ and each multiindex $\alpha = (\alpha_1, \ldots, \alpha_d)$ with $|\alpha| := \alpha_1 + \cdots + \alpha_d$. Then, the norm definition

$$|F_K|_{m,\infty,\widehat{K}} := \max_{i=1,\ldots,d}\ \max_{|\alpha|=m}\ \sup_{\hat{x}\in\widehat{K}} |D^\alpha F_{K,i}(\hat{x})|$$

yields the estimate (17). Note that in the d-dimensional case, it holds $|F_K|_{m,\infty,\widehat{K}} = 0$ and $|F_{K_0}|_{m,\infty,\widehat{K}} = 0$ for $m \ge d+1$.

Next we consider the case that $F_K \ne F_K^*$ is based on an arbitrary vertex numbering of K satisfying the assumptions of Section 2.2. Using the already proved estimate (17) for $F_K = F_K^*$ and the property that $|F_K|_{m,\infty,\widehat{K}} = |F_K^*|_{m,\infty,\widehat{K}}$, see Lemma 2, we obtain the estimate (17) also for $F_K \ne F_K^*$.

Now, let $K \in \mathcal{T}_h^l$ be an element generated from the coarse grid element $K_0 \in \mathcal{T}^0$ with $K \subset K_0$. Then there exists a chain of elements $K_1, \ldots, K_l$ such that $K_i \in \sigma(K_{i-1})$ for all $i = 1, \dots, l$ and $K_l = K$. If we apply successively the estimate (17) to each pair of elements K_i and K_{i-1} for $i = l, l-1, \dots, 1$, we finally obtain the estimate (18). $\square$

As a consequence of Lemma 4 we obtain the following results.

Lemma 5 *Let $\mathcal{T}_h$ be a multilevel adaptive grid generated from a shape regular coarse grid $\mathcal{T}^0$ without boundary adaption as described in Section 2.3. Then, there are constants $C_1(\mathcal{T}^0)$ and $C_2(\mathcal{T}^0)$ such that the estimates*

$$C_1(\mathcal{T}^0)\, 2^{-l} \leq h_K \leq C_2(\mathcal{T}^0)\, 2^{-l} \qquad \forall\, K \in \mathcal{T}_h^l \tag{19}$$

hold. Furthermore, there exist constants $C_3(m, \mathcal{T}^0)$, $C_4(\mathcal{T}^0)$ and $C_5(\mathcal{T}^0)$ such that the following estimates are satisfied for all $K \in \mathcal{T}_h$

$$\begin{aligned} |F_K|_{m,\infty,\widehat{K}} &\leq C_3(m, \mathcal{T}^0)\, h_K^m, && (20)\\ \|DF_K(\widehat{x}) - B_K\|_p &\leq C_4(\mathcal{T}^0)\, h_K^2 \qquad \forall\, p \in [1, \infty], \quad \forall\, \widehat{x} \in \widehat{K}, && (21)\\ \gamma_p(K) &\leq C_5(\mathcal{T}^0)\, h_K \qquad \forall\, p \in [1, \infty], && (22) \end{aligned}$$

where $B_K := DF_K(0)$.

Proof. Let $K \in \mathcal{T}_h^l$ be an element with vertices a_i^K, $i \in \{1, \dots, 2^d\}$, which is contained in a coarse grid element $K_0 \in \mathcal{T}^0$. Then, due to the refinement process, the set $\widehat{W}_K := F_{K_0}^{-1}(K)$ is a d-cube in $\widehat{K}$ where the points $\hat{v}_i := F_{K_0}^{-1}(a_i^K)$, $i \in \{1, 2\}$, are the vertices of an edge of this cube with a length of 2^{-l+1}. Thus, we obtain using the mean value theorem

$$2^{-l+1} = \|\hat{v}_1 - \hat{v}_2\|_2 \leq d\, |F_{K_0}^{-1}|_{1,\infty,K_0}\, \|a_1^K - a_2^K\|_2 \leq d\, |F_{K_0}^{-1}|_{1,\infty,K_0}\, h_K$$

which implies the left inequality of (19) with

$$C_1(\mathcal{T}^0) := \min_{K_0 \in \mathcal{T}^0} \frac{2}{d} \left(|F_{K_0}^{-1}|_{1,\infty,K_0} \right)^{-1}.$$

Let $x, y \in \overline{K}$ be two points with $h_K = \|x - y\|_2$. Then, the points $\hat{x} := F_{K_0}^{-1}(x)$ and $\hat{y} := F_{K_0}^{-1}(y)$ are contained in the d-cube $\overline{\widehat{W}_K}$ with a diameter of $2^{-l+1}\sqrt{d}$. This gives

$$h_K = \|F_{K_0}(\hat{x}) - F_{K_0}(\hat{y})\|_2 \leq d\, |F_{K_0}|_{1,\infty,\widehat{K}}\, \|\hat{x} - \hat{y}\|_2 \leq d\, |F_{K_0}|_{1,\infty,\widehat{K}}\, 2^{-l+1}\sqrt{d}\,.$$

which proves the right inequality of (19) with

$$C_2(\mathcal{T}^0) := \max_{K_0 \in \mathcal{T}^0} 2d^{3/2} |F_{K_0}|_{1,\infty,\widehat{K}}.$$

The estimates (18) of Lemma 4 and (19) yield

$$|F_K|_{m,\infty,\widehat{K}} \le \big(h_K/C_1(\mathcal{T}^0)\big)^m |F_{K_0}|_{m,\infty,\widehat{K}}$$

which shows (20) with

$$C_3(m,\mathcal{T}^0) := C_1(\mathcal{T}^0)^{-m} \max_{K_0\in\mathcal{T}^0} |F_{K_0}|_{m,\infty,\widehat{K}}.$$

For the smooth function $F_K(\hat{x}) = (F_{K,1}(\hat{x}),\dots,F_{K,d}(\hat{x}))^T$, we have for arbitrary $i,j \in \{1,\dots,d\}$ and $\hat{x}\in\widehat{K}$ the estimate

$$\left|\frac{\partial F_{K,i}}{\partial \hat{x}_j}(\hat{x}) - \frac{\partial F_{K,i}}{\partial \hat{x}_j}(0)\right| \le |F_K|_{2,\infty,\widehat{K}} \sum_{\nu=1}^{d} |\hat{x}_\nu| \le C_3(2,\mathcal{T}^0)\, d\, h_K^2\,,$$

which implies for the matrix norm $\|\cdot\|_p$ with arbitrary $p\in[1,\infty]$ that

$$\|DF_K(\hat{x}) - DF_K(0)\|_p \le C_3(2,\mathcal{T}^0)\, d^{2+1/p}\, h_K^2 \le C_3(2,\mathcal{T}^0)\, d^3\, h_K^2.$$

This proves (21) with $C_4(\mathcal{T}^0) := C_3(2,\mathcal{T}^0)d^3$.

For the unit vector $e_1 = (1,0,\dots,0)^T \in \overline{\widehat{K}}$, we define the points $m_1 := F_K(e_1)$ and $m_2 := F_K(-e_1)$ as the midpoints of opposite faces of K. Then, the points $\hat{v}_i := F_{K_0}^{-1}(m_i)$, $i=1,2$, are the midpoints of opposite faces of the d-cube $\widehat{W}_K = F_{K_0}^{-1}(K)$ with an edge length of 2^{-l+1}. Therefore, we have

$$2^{-l+1} = \|\hat{v}_1 - \hat{v}_2\|_2 \le d|F_{K_0}^{-1}|_{1,\infty,K_0}\|m_1 - m_2\|_2. \tag{23}$$

Let F_K^{aff} denote the affine mapping associated with F_K, i.e.,

$$F_K^{\mathrm{aff}}(\widehat{x}) := F_K(0) + B_K\widehat{x} \qquad \forall\,\widehat{x}\in\widehat{K}.$$

Then we have $m_1 = F_K^{\mathrm{aff}}(e_1)$, $m_2 = F_K^{\mathrm{aff}}(-e_1)$ and $S_K = F_K^{\mathrm{aff}}(\widehat{S})$ (see the explanations before (6)) which implies

$$h_{S_K} \ge \|F_K^{\mathrm{aff}}(e_1) - F_K^{\mathrm{aff}}(0)\|_2 = \frac{1}{2}\|F_K^{\mathrm{aff}}(e_1) - F_K^{\mathrm{aff}}(-e_1)\|_2 = \frac{1}{2}\|m_1 - m_2\|_2.$$

Together with (23) and (19) this leads to

$$h_{S_K} \ge d^{-1}|F_{K_0}^{-1}|_{1,\infty,K_0}^{-1} 2^{-l} \ge C_6(\mathcal{T}^0)h_K,$$

where $C_6(\mathcal{T}^0) := d^{-1}\min_{K_0\in\mathcal{T}^0}\big\{|F_{K_0}^{-1}|_{1,\infty,K_0}^{-1}\big\}(C_2(\mathcal{T}^0))^{-1}$.

Using the condition (7) for the shape regularity of the mesh we get

$$\|B_K^{-1}\|_p \leq C_s h_{S_K}^{-1} \leq C_s (C_6(\mathcal{T}^0))^{-1} h_K^{-1}.$$

Thus, we obtain

$$\gamma_p(K) \leq \sup_{\hat{x}\in\widehat{K}} \|B_K^{-1}\|_p \|DF_K(\hat{x}) - DF_K(0)\|_p \leq C_s (C_6(\mathcal{T}^0))^{-1} C_4(\mathcal{T}^0)\, h_K$$

which proves (22). $\square$

In the following we will present for the 2D-case both a sharp lower and a sharp upper bound for the behavior of $\gamma_p(K)$ during a refinement process without boundary adaption as described in Section 2.3.

Lemma 6 *Let $K \in \sigma(K_0)$ be a son-element of a two-dimensional convex element K_0. Furthermore, let $F_K : \widehat{K} \to K$ and $F_{K_0} : \widehat{K} \to K_0$ denote the associated reference mappings from $(\mathbb{Q}_1(\widehat{K}))^d$ satisfying the assumptions of Section 2. Then, the estimate*

$$\frac{1}{2} \cdot \frac{\gamma_p(K_0)}{1 + \gamma_p(K_0)/2} \leq \gamma_p(K) \leq \frac{1}{2} \cdot \frac{\gamma_p(K_0)}{1 - \gamma_p(K_0)/2} \tag{24}$$

holds true. For each element $K \in \mathcal{T}_h^l$ generated from the coarse grid element $K_0 \in \mathcal{T}^0$ with $K \subset K_0$, the estimate

$$\frac{\gamma_p(K_0)}{1 + \gamma_p(K_0)} \cdot 2^{-l} \leq \gamma_p(K) \leq \frac{\gamma_p(K_0)}{1 - \gamma_p(K_0)} \cdot 2^{-l} \tag{25}$$

is satisfied.

Proof. Since $\gamma_p(K)$ is invariant with respect to the special choice of F_K depending on a certain vertex numbering of K, we can restrict ourselves to the case $F_K = F_K^*$ where formula (16) holds, i.e.,

$$F_K(\hat{x}) = F_{K_0}\left(\hat{c}_s + \frac{\hat{x}}{2}\right), \qquad \hat{x} \in \hat{K}.$$

The midpoint $\hat{c}_s$ of $F_{K_0}^{-1}(K)$ can be described as

$$\hat{c}_s = \frac{1}{2}(\varepsilon_1, \varepsilon_2)^T \quad \text{where} \quad \varepsilon_1, \varepsilon_2 \in \{-1, +1\}.$$

We want to represent $\gamma_p(K)$ by expressions involving $\gamma_p(K_0)$. Using the chain rule we obtain

$$DF_K(\hat{x}) = \frac{1}{2} DF_{K_0}\left(\hat{c}_s + \frac{\hat{x}}{2}\right).$$

Applying the definition of $B_\kappa = DF_\kappa(0)$ and $DG_\kappa(\hat{x}) = DF_\kappa(\hat{x}) - B_\kappa$ for $\kappa \in \{K, K_0\}$, we get

$$B_K = \frac{1}{2} DF_{K_0}(\hat{c}_s) = \frac{1}{2}\left[B_{K_0} + DG_{K_0}(\hat{c}_s)\right]$$

and

$$DG_K(\hat{x}) = \frac{1}{2}\left[DG_{K_0}\left(\hat{c}_s + \frac{\hat{x}}{2}\right) - DG_{K_0}(\hat{c}_s)\right] = \frac{1}{4} DG_{K_0}(\hat{x})$$

since DG_{K_0} is linear in the considered 2D-case. Hence, $B_K^{-1} DG_K(\hat{x})$ can be written as

$$B_K^{-1} DG_K(\hat{x}) = \frac{1}{2}\left[\left(I + B_{K_0}^{-1} DG_{K_0}(\hat{c}_s)\right)^{-1}\left(B_{K_0}^{-1} DG_{K_0}(\hat{x})\right)\right].$$

Since $F_{K_0} \in \left(\mathbb{Q}_1(\hat{K})\right)^2$, one easily checks that there are real numbers δ_1 and δ_2 such that

$$B_{K_0}^{-1} DG_{K_0}(\hat{x}) = \begin{pmatrix} \delta_1 \hat{x}_2 & \delta_1 \hat{x}_1 \\ \delta_2 \hat{x}_2 & \delta_2 \hat{x}_1 \end{pmatrix} = \begin{pmatrix} \delta_1 \\ \delta_2 \end{pmatrix} \begin{pmatrix} \hat{x}_2 & \hat{x}_1 \end{pmatrix} \qquad \forall (\hat{x}_1, \hat{x}_2) \in \hat{K} \tag{26}$$

Hence, for $\delta := (\delta_1, \delta_2)^T$ and $1/p + 1/q = 1$ we get

$$\gamma_p(K_0) = \sup_{\hat{x} \in \hat{K}} \left\| \begin{pmatrix} \delta_1 \\ \delta_2 \end{pmatrix} \begin{pmatrix} \hat{x}_2 & \hat{x}_1 \end{pmatrix} \right\|_p = 2^{1/q}\, \|\delta\|_p. \tag{27}$$

Since the convexity of K_0 is equivalent to the fact that $\gamma_1(K_0) < 1$ (see [8, p. 125]), we get from (27) that $\|\delta\|_1 < 1$. For arbitrary vectors $\delta, y, z \in \mathbb{R}^2$ with $|y^T \delta| < 1$ we have

$$(I + \delta y^T)^{-1} = I - \frac{1}{1 + y^T \delta}\, \delta y^T, \qquad (I + \delta y^T)^{-1} \delta z^T = \frac{1}{1 + y^T \delta}\, \delta z^T. \tag{28}$$

Applying (28) with $y = (\varepsilon_2/2, \varepsilon_1/2)^T$ and $z = (\hat{x}_2, \hat{x}_1)^T$, we get

$$\left(I + B_{K_0}^{-1} DG_{K_0}(\hat{c}_s)\right)^{-1}\left(B_{K_0}^{-1} DG_{K_0}(\hat{x})\right) = \frac{1}{1 + (\delta_1 \varepsilon_2 + \delta_2 \varepsilon_1)/2} B_{K_0}^{-1} DG_{K_0}(\hat{x})$$

and

$$B_K^{-1} DG_K(\hat{x}) = \frac{1}{2 + \delta_1 \varepsilon_2 + \delta_2 \varepsilon_1} B_{K_0}^{-1} DG_{K_0}(\hat{x}). \tag{29}$$

Note that $|(\delta_1, \delta_2)\,(\varepsilon_2/2, \varepsilon_1/2)^T| < 1$ due to $|\varepsilon_1| = |\varepsilon_2| = 1$ and $\|\delta\|_1 < 1$. Analogously to (26) we conclude that there exist real numbers δ_1^K, δ_2^K such that

$$B_K^{-1} DG_K(\hat{x}) = \begin{pmatrix} \delta_1^K \\ \delta_2^K \end{pmatrix} \begin{pmatrix} \hat{x}_2 & \hat{x}_1 \end{pmatrix}$$

which implies together with (29) and (27) the relation

$$\delta_i^K = \frac{1}{2+\delta_1\varepsilon_2+\delta_2\varepsilon_1}\delta_i, \qquad i=1,2.$$

Now, the definition of $\gamma_p(K)$ results in

$$\gamma_p(K) = \frac{1}{2+\delta_1\varepsilon_2+\delta_2\varepsilon_1}\,\gamma_p(K_0). \tag{30}$$

Using $\varepsilon_1, \varepsilon_2 \in \{-1,+1\}$ and (27), we obtain

$$|\delta_1\varepsilon_2+\delta_2\varepsilon_1| \le \|\delta\|_p\|(\varepsilon_1,\varepsilon_2)\|_q = \gamma_p(K_0)$$

which yields together with (30) the estimate (24).

Now, let $K \in \mathcal{T}_h^l$ be an element generated from the coarse grid element $K_0 \in \mathcal{T}^0$ with $K \subset K_0$. Then, there exists a chain of elements K_i, $i=1,\ldots,l$, such that $K_l = K$ and $K_{i-1} = \mathcal{F}(K_i)$ for all $i=1,\ldots,l$. The estimate (24) can be written as

$$\varphi_+\big(\gamma_p(K_i)\big) \le \gamma_p(K_{i+1}) \le \varphi_-\big(\gamma_p(K_i)\big) \quad \text{with} \quad \varphi_\pm(t) := \frac{1}{2}\cdot\frac{t}{1\pm t/2}.$$

Furthermore, we have

$$\varphi_\pm\big(\varphi_\pm(t)\big) = \frac{1}{4}\cdot\frac{t}{1\pm t(1-1/4)} \quad \text{and} \quad \underbrace{\varphi_\pm(\varphi_\pm(\cdots\varphi_\pm(t)))}_{l\text{-times}} = 2^{-l}\cdot\frac{t}{1\pm t(1-2^{-l})}$$

which can be proven easily by induction. Hence, we obtain

$$2^{-l}\cdot\frac{\gamma_p(K_0)}{1+\gamma_p(K_0)(1-2^{-l})} \le \gamma_p(K) \le 2^{-l}\cdot\frac{\gamma_p(K_0)}{1-\gamma_p(K_0)(1-2^{-l})} \qquad \forall K \in \mathcal{T}_h^l.$$

Exploiting the fact that $0 \le \gamma_p(K_0)$, we finally get (25). $\square$

Together with $h_K \sim 2^{-l}$ for $K \in \mathcal{T}_h^l$, see Lemma 5, we obtain from (25) that $\gamma_p(K)$ behaves like the local mesh size h_K, i.e., there exist constants $C_7(\mathcal{T}^0)$, $C_8(\mathcal{T}^0)$ such that for all $p \in [1,\infty]$ the estimates

$$C_7(\mathcal{T}^0)h_K \le \gamma_p(K) \le C_8(\mathcal{T}^0)h_K \qquad \forall K \in \mathcal{T}_h$$

hold.

4 Refinement With Boundary Adaption

In order to simplify the notation, we restrict ourselves to the two-dimensional case. Generalizations to the three-dimensional case are possible but require a high technical effort.

We will now describe the modifications of the refinement procedure for elements located at a curved boundary part of the domain Ω. For simplicity we restrict ourselves to the case where only one edge of the element belongs to a curved boundary part. Generalizations to the case of more than one curved boundary parts are straightforward. We assume that the parameterization of all boundary points $x \in \Gamma := \partial\Omega$ is given as

$$x = g(t), \quad t \in I := \bigcup_{k=1}^{m} [a_k, b_k]\,, \tag{31}$$

where the parameter set I is built by some nonoverlapping closed intervals and the function $g : I \to \mathbb{R}^2$ is smooth in the sense that the second derivatives of g are continuous at each subinterval $[a_k, b_k]$ of I.

In order to describe the refinement process, we consider an element K on grid level $\ell \geq 0$ where the local numbering of its vertices a_i, $i = 1, \ldots, 4$, is assumed to be such that the edge $\overline{a_3 a_4}$ corresponds to a curved part of the boundary Γ, see Figure 6. Let the

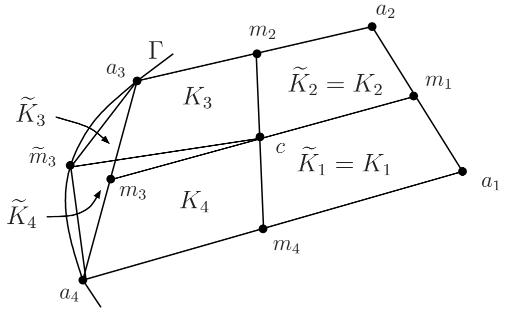

Figure 6: *Refinement of an element K with the vertices a_i, $i = 1, \ldots, 4$, into son-elements $\widetilde{K}_1$, $\ldots$, $\widetilde{K}_4$, with boundary adaption.*

parameterization of points x of this boundary part be given as

$$\begin{aligned} &x = g(t), \quad t \in [t_1, t_2] \subset I, \\ &a_3 = g(t_1), \;\; a_4 = g(t_2). \end{aligned} \tag{32}$$

The refinement of K into son-elements with boundary adaption is defined in two steps. In a first step, K is refined into son-elements $K_i = \sigma_i(K)$, $i = 1, \ldots, 4$, without boundary adaption as described in Section 2.3. For $i \in \{1, \ldots, 4\}$, let $m_i := \frac{1}{2}(a_i + a_{i+1})$ with $a_5 := a_1$ denote the midpoint of the K-edge $\overline{a_i a_{i+1}}$ and $c := \frac{1}{4}\sum_{i=1}^{4} a_i$ be the barycenter of K. Then, the son-element K_i is the quadrilateral with the vertices a_i, m_i, c and m_{i-1} where $m_0 := m_4$, see Figure 6. Let $\sigma(K)$ denote the set of all son-elements K_i, $i = 1, \ldots, 4$, of element K. In a second step, to each son-element K_i, $i \in \{1, \ldots, 4\}$, there will be assigned a possibly modified son-element $\widetilde{K}_i$ where the vertex m_3, if it belongs to the vertices of K_i, is replaced by a corresponding point $\widetilde{m}_3$ at the curved boundary defined as

$$\widetilde{m}_3 := g\big(\tfrac{1}{2}(t_1 + t_2)\big)\,.$$

Note that only the son-elements K_3 and K_4 at the curved boundary are affected by this modification, i.e., we have $\widetilde{K}_1 = K_1$ and $\widetilde{K}_2 = K_2$. So, in the refinement procedure with boundary adaption, an element K on grid level ℓ located at a curved boundary part will be replaced by the son-elements $\widetilde{K}_1, \ldots, \widetilde{K}_4$ on grid level $\ell + 1$. In the following, let $\tilde{\sigma}(K)$ denote the set of all possibly modified son-elements $\widetilde{K}_i$, $i = 1, \ldots, 4$, of element K.

Now, we can apply the same procedure to the case where the element K is the new element $\widetilde{K}_3$ or $\widetilde{K}_4$ from the previous refinement step. If we apply this procedure recursively starting from the initial coarse mesh $\mathcal{T}^0$, we obtain a multilevel adaptive mesh $\mathcal{T}_h$ with boundary adaption. An example of such meshes is shown in Figure 7.

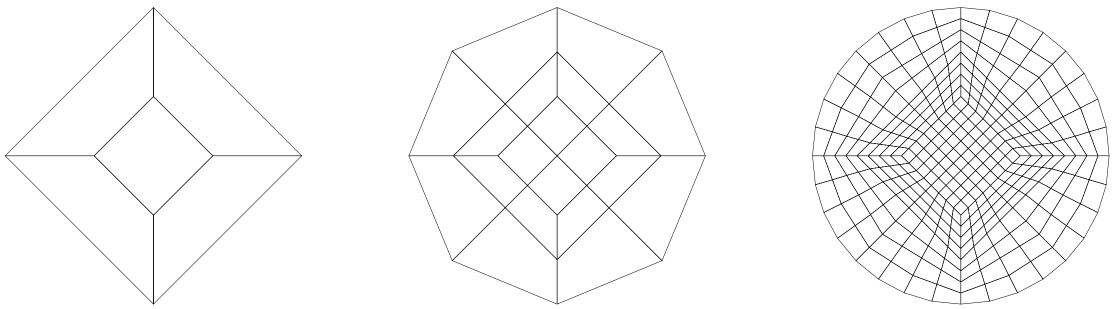

Figure 7: *Meshes during refinement with boundary adaption on grid level 0, 1, and 3.*

Lemma 7 *Let $K_0 \in \mathcal{T}^0$ be a coarse grid element having one edge located at a curved boundary part of the domain Ω and let $K \in \mathcal{T}_h$ be an element that has been generated from K_0 by means of $\ell = \ell(K)$ refinement steps with boundary adaption. Then, for the corresponding reference mappings it holds*

$$|F_K|_{2,\infty,\widehat{K}} \leq |F_{K_0}|_{2,\infty,\widehat{K}} \big(\tfrac{1}{4}\big)^{\ell} + C_g \ell \big(\tfrac{1}{4}\big)^{\ell}\,, \tag{33}$$

where C_g is a constant independent of K_0, K and ℓ.

Proof. Since K has been generated from K_0 by means of ℓ refinement steps with boundary adaption there is a unique chain of elements $K_0 = \widetilde{K}_0 \to K_1 \to \widetilde{K}_1 \to \cdots \to K_\ell \to \widetilde{K}_\ell = K$ such that $K_{i+1} \in \sigma(\widetilde{K}_i)$ and $\widetilde{K}_{i+1} \in \tilde{\sigma}(\widetilde{K}_i)$ for all $i \in \{0, 1, \ldots, \ell - 1\}$.

In the following, we will introduce some more notation for the description of the elements $\widetilde{K}_i$, K_{i+1} and $\widetilde{K}_{i+1}$. If $\widetilde{K}_i$ has an edge located at a curved boundary part of Ω we assume that the vertices $a_j^{(i)}$, $j \in \{1, \ldots, 4\}$, of $\widetilde{K}_i$ are denoted such that the edge from $a_3^{(i)}$ to $a_4^{(i)}$ corresponds to the curved boundary part, see Figure 8. According to the

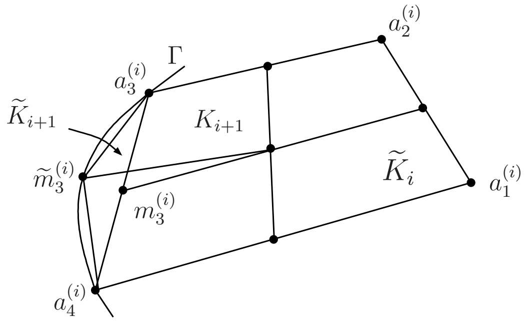

Figure 8: *Refinement with boundary adaption of element $\widetilde{K}_i$ with the vertices $a_j^{(i)}$, $j = 1, \ldots, 4$, and son-elements K_{i+1} and $\widetilde{K}_{i+1}$.*

parameterization of the boundary, we assume that

$$a_3^{(i)} = g(t_1^{(i)})\,, \quad a_4^{(i)} = g(t_2^{(i)}) \quad \text{and} \quad t_1^{(i)}, t_2^{(i)} \in [a_k, b_k]$$

for some $k \in \{1, \ldots, m\}$, see (31). We denote by $m_3^{(i)}$ the midpoint of the edge from $a_3^{(i)}$ to $a_4^{(i)}$ and by $\widetilde{m}_3^{(i)}$ the corresponding point at the curved boundary defined as

$$\widetilde{m}_3^{(i)} := g(t_m^{(i)}) \qquad \text{with} \qquad t_m^{(i)} := \tfrac{1}{2}\big(t_1^{(i)} + t_2^{(i)}\big)\,. \tag{34}$$

Now, $K_{i+1} \in \sigma(\widetilde{K}_i)$ is that son-element of $\widetilde{K}_i$ which occurs in the above mentioned unique chain of elements from K_0 to K. There are two cases. In the first case, where $m_3^{(i)}$ is a vertex of K_{i+1}, the element $\widetilde{K}_{i+1}$ is obtained from K_{i+1} by exchanging the vertex $m_3^{(i)}$ with the boundary point $\widetilde{m}_3^{(i)}$. In the second case, where $m_3^{(i)}$ is not a vertex of K_{i+1}, we have $\widetilde{K}_{i+1} = K_{i+1}$. If $\widetilde{K}_i$ has no edge located at a curved boundary part, we define $\widetilde{K}_{i+1} := K_{i+1}$.

We will now prove by induction the estimate

$$|F_{\widetilde{K}_i}|_{2,\infty,\widehat{K}} \le |F_{K_0}|_{2,\infty,\widehat{K}}\big(\tfrac{1}{4}\big)^i + C_g\, i\big(\tfrac{1}{4}\big)^i\,, \tag{35}$$

for a suitable constant C_g and $i = 0, \ldots, \ell$. In the case $i = 0$, estimate (35) is obviously satisfied for any constant C_g. Supposing that (35) is valid for some $i \in \{0, 1, \ldots, \ell - 1\}$, we will show that it holds also for $i+1$ instead of i. By the triangle inequality we have

$$|F_{\widetilde{K}_{i+1}}|_{2,\infty,\widehat{K}} \leq |F_{K_{i+1}}|_{2,\infty,\widehat{K}} + |F_{\widetilde{K}_{i+1}} - F_{K_{i+1}}|_{2,\infty,\widehat{K}} \,. \tag{36}$$

Since K_{i+1} is a son-element of $\widetilde{K}_i$ without boundary adaption, we know from (17) in Lemma 4 that

$$|F_{K_{i+1}}|_{2,\infty,\widehat{K}} \leq \tfrac{1}{4}\, |F_{\widetilde{K}_i}|_{2,\infty,\widehat{K}} \,. \tag{37}$$

In the case when $\widetilde{K}_{i+1}$ and K_{i+1} are different, we easily get the estimate

$$|F_{\widetilde{K}_{i+1}} - F_{K_{i+1}}|_{2,\infty,\widehat{K}} \leq \tfrac{1}{4}\, \|\widetilde{m}_3^{(i)} - m_3^{(i)}\|_\infty \tag{38}$$

by using explicit expressions for the reference mappings $F_{\widetilde{K}_{i+1}}(\hat{x})$ and $F_{K_{i+1}}(\hat{x})$. By means of (34) we obtain

$$\delta^{(i)} := \widetilde{m}_3^{(i)} - m_3^{(i)} = g(t_m^{(i)}) - \left\{ \tfrac{1}{2}\, g(t_1^{(i)}) + \tfrac{1}{2}\, g(t_2^{(i)}) \right\}. \tag{39}$$

The j-th component $\delta_j^{(i)}$ of the vector $\delta^{(i)}$ can be regarded as the error of a linear interpolation of the function $g_j(t)$ at the point $t = t_m^{(i)}$. Thus, from the well-known error representation we know that there exists a $\xi_j^{(i)} \in \big(t_1^{(i)}, t_2^{(i)}\big)$ such that

$$\delta_j^{(i)} = \tfrac{1}{2}\, g_j''(\xi_j^{(i)})(t_m^{(i)} - t_1^{(i)})\,(t_m^{(i)} - t_2^{(i)}) = -\tfrac{1}{8}\, g_j''(\xi_j^{(i)})\big(t_2^{(i)} - t_1^{(i)}\big)^2 \,. \tag{40}$$

For the parameter values $t_{1/2}^{(k)}$, we have by construction that

$$t_2^{(k+1)} - t_1^{(k+1)} = \tfrac{1}{2}\left(t_2^{(k)} - t_1^{(k)}\right) \qquad \text{for} \quad k = 0, \ldots, i-1,$$

which implies

$$t_2^{(i)} - t_1^{(i)} = \left(\tfrac{1}{2}\right)^i \left(t_2^{(0)} - t_1^{(0)}\right) \leq \left(\tfrac{1}{2}\right)^i \max_{k=1,\ldots,m} (b_k - a_k) \,. \tag{41}$$

From (40) and (41) and the fact that $\xi_j^{(i)} \in [a_k, b_k]$, we obtain

$$|\delta_j^{(i)}| \leq C_g \left(\tfrac{1}{4}\right)^i, \qquad j = 1, 2, \tag{42}$$

with

$$C_g := \max_{k=1,\ldots,m} \left\{ \max_{j=1,2}\ \max_{\xi \in [a_k, b_k]} \tfrac{1}{8}\, |g_j''(\xi)|\, (b_k - a_k)^2 \right\}.$$

Combining (38), (39) and (42), we can conclude the estimate

$$|F_{\widetilde{K}_{i+1}} - F_{K_{i+1}}|_{2,\infty,\widehat{K}} \leq C_g \big(\tfrac{1}{4}\big)^{i+1}. \tag{43}$$

This estimate obviously holds true in the case when $\widetilde{K}_{i+1} = K_{i+1}$. If we now combine (35)–(37) and (43), we get

$$\begin{aligned} |F_{\widetilde{K}_{i+1}}|_{2,\infty,\widehat{K}} &\leq \tfrac{1}{4}\Big\{|F_{K_0}|_{2,\infty,\widehat{K}} \big(\tfrac{1}{4}\big)^i + C_g\, i \big(\tfrac{1}{4}\big)^i\Big\} + C_g \big(\tfrac{1}{4}\big)^{i+1} \\ &= |F_{K_0}|_{2,\infty,\widehat{K}} \big(\tfrac{1}{4}\big)^{i+1} + C_g\,(i+1)\big(\tfrac{1}{4}\big)^{i+1}, \end{aligned}$$

which proves the estimate (35) for $i+1$ instead of i. Finally, the estimate (33) follows from the estimate (35) for $i = \ell$ and the fact that $\widetilde{K}_\ell = K$. $\square$

Remark 8 *One can construct an example of a domain with a curved boundary which shows that the factor $\ell = \ell(K)$ cannot be omitted in the estimate* (33).

5 Interpolation Properties of Mapped Polynomials

In this section, we will derive estimates of the interpolation error for mapped finite elements by using the estimates for the derivatives of the reference mappings from the previous sections.

Mapped finite elements are defined using the reference mapping F_K and a function space on a reference cell $\widehat{K}$. To be precise, let $\widehat{V}$ denote a finite dimensional function space on $\widehat{K}$. Then, the local function space $V(K)$ on K is defined by

$$V(K) := \big\{v = \hat{v} \circ F_K^{-1} \,:\, \hat{v} \in \widehat{V}\big\}. \tag{44}$$

If $\widehat{\Pi} : H^{r+1}(\widehat{K}) \to \widehat{V}$ denotes a continuous interpolation operator with

$$\widehat{\Pi}\hat{v} = \hat{v} \qquad \forall \hat{v} \in \widehat{V} \tag{45}$$

then the local interpolation operator $\Pi_K : H^{r+1}(K) \to V(K)$ is defined by

$$\Pi_K v := \big(\widehat{\Pi}(v \circ F_K)\big) \circ F_K^{-1}.$$

In the following, we will present estimates of the interpolation error.

5.1 Without boundary adaption

Let Ω be a polygonal or polyhedral domain such that a boundary adaption during the refinement is not necessary. In the case where all reference mappings are affine, an optimal order of the interpolation error is obtained if $\mathbb{P}_r(\widehat{K}) \subset \widehat{V}$, see [2]. In the non-affine case where $F_K \in \mathbb{Q}_1(\widehat{K})^d$, the situation is more complicated. On arbitrary but shape regular meshes, it turns out that the condition $\mathbb{Q}_r(\widehat{K}) \subset \widehat{V}$ is necessary and sufficient to have optimal order in L^2 and H^1, see [1] for the two-dimensional case and [6] for the three-dimensional case.

In applications like flow simulations with a discontinuous pressure approximation or a non-conforming velocity approximation, the case where $\mathbb{Q}_r(\widehat{K}) \not\subset \widehat{V}$ but $\mathbb{P}_r(\widehat{K}) \subset \widehat{V}$ is of interest. For this situation, there will be a dramatic decrease of the order of the interpolation error on general families of meshes, see [1, 6]. However, on multilevel adaptive grids, it was reported in literature that $\mathbb{P}_r(\widehat{K}) \subset \widehat{V}$ is sufficient for optimal order, although a rigorous proof was missing until now.

Lemma 9 *Let $\mathbb{P}_r(\widehat{K}) \subset \widehat{V}$ and $\mathcal{T}_h$ a shape regular multilevel adaptive grid generated from the coarse grid $\mathcal{T}^0$ without boundary adaption. Then, the local interpolation error estimate*

$$|v - \Pi_K v|_{m,K} \leq C(\mathcal{T}^0)\, h_K^{r+1-m} \|v\|_{r+1,K}, \qquad \forall v \in H^{r+1}(K),$$

holds true for all $K \in \mathcal{T}_h$ and $m \in \{0,1\}$.

Proof. The definition of Π_K and the transformation of norms between the Sobolev spaces $H^{r+1}(K)$ and $H^{r+1}(\widehat{K})$ yield

$$\|v - \Pi_K v\|_{0,K} \leq C \|\det DF_K\|_{0,\infty,\widehat{K}}^{1/2} \|\hat{v} - \widehat{\Pi}\hat{v}\|_{0,\widehat{K}}$$

where $\hat{v} = v \circ F_K$ and DF_K denotes the Jacobi matrix associated with the reference mapping F_K. Using the Bramble-Hilbert lemma, we get

$$\|\hat{v} - \widehat{\Pi}\hat{v}\|_{0,\widehat{K}} \leq C |\hat{v}|_{r+1,\widehat{K}}$$

since $\widehat{\Pi}$ is a continuous operator on $H^{r+1}(\widehat{K})$ and $\mathbb{P}_r(\widehat{K})$ remains unchanged under $\widehat{\Pi}$ due to $\mathbb{P}_r(\widehat{K}) \subset \widehat{V}$ and (45). The norm of the higher order derivatives of the composite function $\hat{v} = v \circ F_K$ can be estimated in the following way

$$|v \circ F_K|_{r+1,\widehat{K}} \leq C \|\det D(F_K^{-1})\|_{0,\infty,K}^{1/2} \|v\|_{r+1,K} \sum_{\substack{\alpha+2\beta+3\gamma \\ =r+1}} |F_K|_{1,\infty,\widehat{K}}^{\alpha} |F_K|_{2,\infty,\widehat{K}}^{\beta} |F_K|_{3,\infty,\widehat{K}}^{\gamma},$$

see [3, p. 222] where α, β, γ denote nonnegative integers. Putting together all estimates, we obtain

$$\|v - \Pi_K v\|_{0,K} \leq C\|v\|_{r+1,K} \sum_{\substack{\alpha+2\beta+3\gamma \\ =r+1}} |F_K|^{\alpha}_{1,\infty,\widehat{K}} |F_K|^{\beta}_{2,\infty,\widehat{K}} |F_K|^{\gamma}_{3,\infty,\widehat{K}}, \tag{46}$$

where we have used that

$$\|\det DF_K\|_{0,\infty,\widehat{K}} \, \|\det D(F_K^{-1})\|_{0,\infty,K} \leq C$$

due to the shape regularity of $\mathcal{T}_h$. For the two-dimensional case note that the third derivatives of F_K are identically zero. Hence, the sum reduces to a sum with $\alpha + 2\beta = r + 1$. Using (20) we obtain

$$\|v - \Pi_K v\|_{0,K} \leq C(\mathcal{T}^0) h_K^{r+1} \|v\|_{r+1,K}$$

which is the assertion for $m = 0$. The proof for $m = 1$ follows the same lines but we start with

$$|v - \Pi_K v|_{1,K} \leq C\|\det DF_K\|^{1/2}_{0,\infty,\widehat{K}} |F_K^{-1}|_{1,\infty,K} |\hat{v} - \widehat{\Pi}\hat{v}|_{1,\widehat{K}}.$$

Observing $|F_K^{-1}|_{1,\infty,K} \leq Ch_K^{-1}$, we get the desired estimate for $m = 1$. $\square$

5.2 With boundary adaption

Now we consider the refinement with boundary adaption for a domain with curved boundary. We will restrict ourselves to the case of first order elements since higher order elements would require also a higher order boundary approximation than the one that we have used in Section 4. Furthermore, all presented results will be two-dimensional.

Lemma 10 *Let $\mathbb{P}_1(\widehat{K}) \subset \widehat{V}$ and $\mathcal{T}_h$ be a multilevel adaptive grid with boundary adaption as explained in Section 4. Then, the local interpolation error estimate*

$$|v - \Pi_K v|_{m,K} \leq \left\{ Ch_K^2 + \big(C(\mathcal{T}^0) + C_g \ell(K)\big)\big(2^{-\ell(K)}\big)^2 \right\} h_K^{-m} \, \|v\|_{2,K}, \tag{47}$$

holds true for all $v \in H^2(K)$, $K \in \mathcal{T}_h$ and $m \in \{0, 1\}$.

Proof. Proceeding as in the proof of Lemma 9, we obtain like in (46) the estimate

$$\|v - \Pi_K v\|_{0,K} \leq C\|v\|_{2,K} \Big(|F_K|^2_{1,\infty,\widehat{K}} + |F_K|_{2,\infty,\widehat{K}} \Big) \qquad \forall v \in H^2(K).$$

Using the estimates $|F_K|_{1,\infty,\widehat{K}} \leq Ch_K$ and (33) from Lemma 7, we obtain the assertion (47) for $m = 0$. The case $m = 1$ is obtained as in the proof of Lemma 9. $\square$

Remark 11 *For an element K that has been generated from a coarse grid element $K_0 \in \mathcal{T}^0$ by means of $\ell(K)$ refinement steps, it is clear from the geometrical imagination that the mesh size h_K and the refinement level $\ell(K)$ are related in the following way*

$$\left(\tfrac{1}{2}\right)^{\ell(K)} \sim h_K \qquad \textit{and} \qquad \ell(K) \sim |\log h_K|\,, \tag{48}$$

although a rigorous proof of (48) *is still an open problem. Using* (48)*, the inequality* (47) *of Lemma 10 yields the estimate*

$$|v - \Pi_K v|_{m,K} \leq \Big\{C + C'|\log h_K|\Big\} h_K^{2-m} \|v\|_{2,K} \qquad \forall\, v \in H^2(K),$$

where $m \in \{0, 1\}$. This says that in the case of refinement with boundary adaption, we obtain a local interpolation error estimate which is optimal up to a logarithmic factor $|\log h_K| \sim \ell(K)$.

Finally, we will show that the logarithmic factor $\ell(K) \sim |\log h_K|$ which occurs in the local estimate (47) often will become invisible in a global error estimate. To this end, we assume for simplicity that the coarse grid $\mathcal{T}^0 = \{K_0\}$ consists of only one cell K_0 and that we have uniform refinement with boundary adaption. Refining K_0 ℓ-times gives 4^ℓ cells which form the triangulation $\mathcal{T}_h = \mathcal{T}_h^\ell$ where $\mathcal{T}_h^\ell$ is defined in (14). Let Ω_h^ℓ denote the domain which is covered by the cells from $\mathcal{T}_h^\ell$. Furthermore, we define

$$W_{h,\ell} := \left\{ v \in L^2\big(\Omega_h^\ell\big) \;:\; v|_K \in V(K) \;\; \forall K \in \mathcal{T}_h^\ell \right\}$$

as a finite element space where no continuity across cell boundaries is required. Let Π denote the global finite element interpolation operator which is obtained by putting together the local interpolation operators Π_K. Obviously, Π maps into $W_{h,\ell}$.

Furthermore, we make the assumption that the function v which has to be interpolated satisfies

$$\|v\|_{2,K}^2 \leq C \frac{1}{4^\ell} \|v\|_{2,\Omega_h^\ell}^2 \qquad \forall K \in \mathcal{T}_h^\ell \tag{49}$$

where 4^ℓ is the number of cells on grid level ℓ. Note that (49) excludes functions with sharp layers.

Lemma 12 *Let $\mathcal{T}_h^\ell$ be the grid with boundary adaption as described above and let $v \in H^2(\Omega_h^\ell)$ satisfy* (49). *Then, the global interpolation error estimate*

$$\|v - \Pi v\|_{0,\Omega_h^\ell} \le \left\{ Ch^2 + C'\left(2^{-\ell}\right)^2 \right\} \|v\|_{2,\Omega_h^\ell} \tag{50}$$

holds true.

Proof. In order to prove this lemma, we have a closer look to the refinement with boundary adaption. For a cell at the boundary, the refinement process with boundary adaption is carried out. As its result, one gets 4 cells where 2 cells are again located at the boundary and 2 cells are away from the boundary. The cells of the first group will be refined further with boundary adaption while for cells from the latter group the usual refinement is applied.

This means that after ℓ refinement steps there will be cells which are refined m-times with boundary adaption and $(\ell - m)$-times without boundary adaption, $m = 0, \dots, \ell$. A counting gives that one gets $(1/2)^{m+1}\, 4^\ell$ cells that have been generated by m refinements with and $(\ell - m)$ refinements without boundary adaption, where $m = 0, \dots, \ell - 1$. There are 2^ℓ cells generated by ℓ refinements with boundary adaption. Naturally, the refinements with boundary adaptions are performed before the refinements without boundary adaption. Hence, by combining the results of Lemmata 4 and 7, the estimate

$$|F_K|_{2,\infty,\widehat{K}} \le \left(\frac{1}{4}\right)^\ell \left[C_0 + mC_g\right], \tag{51}$$

holds true for all cells $K \in \mathcal{T}_h^\ell$ which were refined m-times with boundary adaption and $(\ell - m)$-times without boundary adaption where $m = 0, \dots, \ell$ and $C_0 := \max_{K_0 \in \mathcal{T}^0} |F_{K_0}|_{2,\infty,\widehat{K}}$.

For estimating the global interpolation error, we start with

$$\|v - \Pi v\|_{0,\Omega_h^\ell}^2 = \sum_{K \in \mathcal{T}_h^\ell} \|v - \Pi_K v\|_{0,K}^2 \le C \sum_{K \in \mathcal{T}_h^\ell} \|v\|_{2,K}^2 \left[|F_K|_{1,\infty,\hat{K}}^4 + |F_K|_{2,\infty,\hat{K}}^2 \right], \tag{52}$$

which is obtained in the same way as (46) in the proof of Lemma 9. The term with the first derivatives of F_K gives the optimal second order estimate

$$\sum_{K \in \mathcal{T}_h^\ell} \|v\|_{2,K}^2 |F_K|_{1,\infty,\hat{K}}^4 \le Ch^4 \|v\|_{2,\Omega_h^\ell}^2. \tag{53}$$

Using (49), the term which contains the second order derivatives of the reference mapping can be estimated as

$$\sum_{K \in \mathcal{T}_h^\ell} \|v\|_{2,K}^2 |F_K|_{2,\infty,\hat{K}}^2 \le C \frac{\|v\|_{2,\Omega_h^\ell}^2}{4^\ell} \sum_{K \in \mathcal{T}_h^\ell} |F_K|_{2,\infty,\hat{K}}^2$$

The sum will be split up according to the number of refinements with boundary adaption. Using (51) we get for $\ell \geq 1$

$$
\begin{aligned}
\sum_{K\in\mathcal{T}_h^\ell} |F_K|^2_{2,\infty,\hat{K}} &\leq \sum_{m=0}^{\ell-1}(1/2)^{m+1}4^\ell\big((1/4)^\ell(C_0+mC_g)\big)^2 + 2^\ell\big((1/4)^\ell(C_0+\ell C_g)\big)^2 \\
&\leq 2\sum_{m=0}^{\ell-1}(1/2)^{m+1}4^\ell(1/4)^{2\ell}(C_0^2+m^2C_g^2) + 2(1/8)^\ell(C_0^2+\ell^2C_g^2) \\
&= C_0^2\Big[(1/4)^\ell\sum_{m=0}^{\ell-1}(1/2)^m + 2(1/8)^\ell\Big] \\
&\quad + C_g^2\Big[(1/4)^\ell\sum_{m=0}^{\ell-1}m^2(1/2)^m + 2\ell^2(1/8)^\ell\Big] \\
&= C_0^2\Big[(1/4)^\ell 2\big(1-(1/2)^\ell\big) + 2(1/8)^\ell\Big] \\
&\quad + C_g^2\Big[(1/4)^\ell\big\{6-2\ell^2(1/2)^\ell - 6(1/2)^\ell - 4\ell(1/2)^\ell\big\} + 2\ell^2(1/8)^\ell\Big].
\end{aligned}
$$

This implies

$$
\sum_{K\in\mathcal{T}_h^\ell} |F_K|^2_{2,\infty,\hat{K}} \leq 2C_0^2(1/4)^\ell + 6C_g^2(1/4)^\ell. \tag{54}
$$

The estimate (54) holds true also in the trivial case $\ell = 0$. Combining (52)–(54), we get the assertion (50). $\square$

Lemma 12 and the relations (48) in Remark 11 say that even in the case of refinement with boundary adaption, the optimal global interpolation error estimate

$$
\|v-\Pi v\|_{0,\Omega_h^\ell} \leq Ch^2\,\|v\|_{2,\Omega_h^\ell}
$$

holds true under the assumption (49).

References

[1] D. N. Arnold, D. Boffi, and R. S. Falk. Approximation by quadrilateral finite elements. *Math. Comput.*, 71(239):909–922, 2002.

[2] P. G. Ciarlet. *The finite element method for elliptic problems.* Studies in Mathematics and its Applications. Vol. 4. Amsterdam - New York - Oxford: North-Holland Publishing Company, 1978.

[3] H. Federer. *Geometric measure theory.* Springer-Verlag New York Inc., New York, 1969.

[4] V. Heuveline and R. Rannacher. A posteriori error control for finite element approximations of elliptic eigenvalue problems. *Adv. Comput. Math.*, 15(1-4):107–138, 2001.

[5] V. Heuveline and R. Rannacher. Duality-based adaptivity in the *hp*-finite element method. *J. Numer. Math.*, 11(2):95–113, 2003.

[6] G. Matthies. Mapped finite elements on hexahedra. Necessary and sufficient conditions for optimal interpolation errors. *Numer. Algorithm*, 27(4):317–327, 2001.

[7] G. Matthies. *Finite element methods for free boundary value problems with capillary surfaces.* PhD thesis, Fakultät für Mathematik, Otto-von-Guericke-Universität Magdeburg, 2002. published at Shaker-Verlag Aachen.

[8] G. Matthies and L. Tobiska. The inf-sup condition for the mapped Q_k-P_{k-1}^{disc} element in arbitrary space dimensions. *Computing*, 69:119–139, 2002.