Reconsidered error analysis in the finite element methods Interpolation error analysis using a new geometric parameter

Hiroki ISHIZAKA *
https://teamfem.github.io/hiroki_ishizaka/

April 20, 2025

Abstract

This article presents novel proof methods for estimating interpolation errors, predicated on the understanding that one has already studied foundational error analysis using the finite element method. This article summarizes References [29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39]. We are also correcting any typos found in each paper as we find them. The purpose is to make an easy-to-understand note of 'Special Topics in Finite Element Methods.'

Contents

*h.ishizaka005@gmail.com

1 Preliminalies

1.1 General Convention

Throughout this article, we denote by c a constant independent of h (defined later) and the angles and aspect ratios of simplices, unless specified otherwise all constants c are bounded if the maximum angle is bounded. These values vary across different contexts.

1.2 Basic Notation

d	The space dimension, $d \in \{2, 3\}$
$\mathbb{R}^d$	d-dimensional real Euclidean space
$\mathbb{N}_0$	$\mathbb{N}_0 := \mathbb{N} \cup \{0\}$
$\mathbb{R}_+$	The set of positive real numbers
$\lvert \cdot \rvert_d$	d-dimensional Hausdorff measure
$v\lvert_D$	Restriction of the function v to the set D
$\dim(V)$	Dimension of the vector space V
δ_{ij}	Kronecker delta: $\delta_{ij} = 1$ if $i = j$ and 0 otherwise
$(x_1, \dots, x_d)^T$	Cartesian coordinates in $\mathbb{R}^d$

1.3 Vectors and Matrices

$(v_1, \dots, v_d)^T$	Cartesian components of the vector v in $\mathbb{R}^d$
$x \cdot y$	Euclidean scalar product in $\mathbb{R}^d$: $x \cdot y := \sum_{i=1}^{d} x_i y_i$
$\lvert x \rvert_E$	Euclidean norm in $\mathbb{R}^d$: $\lvert x \rvert_E := (x \cdot x)^{1/2}$
$\mathbb{R}^{m \times n}$	Vector space $m \times n$ matrices with real-valued entries
A, B	Matrices
A_{ij} or $[A]_{ij}$	Entry of A in the ith and the jth column
$A^\top$	Transpose of the matrix A
$\mathrm{Tr}(A)$	Trace of A: For $A \in \mathbb{R}^{m \times n}$, $\mathrm{Tr}(A) := \sum_{i=1}^{d} A_{ii}$
$\det(A)$	Determinant of A
$\mathrm{diag}(A)$	Diagonal of A: For $A \in \mathbb{R}^{m \times n}$, $\mathrm{diag}(A)_{ij} := \delta_{ij} A_{ij}$, $1 \leq i, j \leq d$
Ax	Matrix-vector product: For $A \in \mathbb{R}^{m \times n}$ and $x \in \mathbb{R}^n$, $(Ax)_i := \sum_{j=1}^{d} A_{ij} x_j$ for $1 \leq i \leq d$
$A : B$	Double contraction: For $A \in \mathbb{R}^{m \times n}$ and $B \in \mathbb{R}^{m \times n}$, $A : B := \sum_{i=1}^{m} \sum_{j=1}^{n} A_{ij} B_{ij}$
$\lVert A \rVert_2$	Operator norm of A: For $A \in \mathbb{R}^{d \times d}$, $\lVert A \rVert_2 := \sup_{0 \neq x \in \mathbb{R}^d} \frac{\lvert Ax \rvert_E}{\lvert x \rvert_E}$
$\lVert A \rVert_{\max}$	Max norm of A: For $A \in \mathbb{R}^{d \times d}$, $\lVert A \rVert_{\max} := \max_{1 \leq i,j \leq d} \lvert A_{ij} \rvert$
$O(d)$	$O(d)$ consists of all orthogonal matrices of determinant ± 1

In this article, we use the following facts.

For $A \in \mathbb{R}^{m\times n}$, it holds that

$$\|A\|_{\max} \le \|A\|_2 \le \sqrt{mn}\|A\|_{\max}, \tag{1.1}$$

e.g., see [25, p. 56]. For $A, B \in \mathbb{R}^{m\times m}$, it holds that

$$\|AB\|_2 \le \|A\|_2\|B\|_2. \tag{1.2}$$

If $A^\top A$ is a positive definite matrix in $\mathbb{R}^{d\times d}$, the spectral norm of the matrix $A^\top A$ is the largest eigenvalue of $A^\top A$; i.e.,

$$\|A\|_2 = \big(\lambda_{\max}(A^\top A)\big)^{1/2} = \sigma_{\max}(A), \tag{1.3}$$

where $\lambda_{\max}(A)$ and $\sigma_{\max}(A)$ are respectively the largest eigenvalues and singular values of A.

If $A \in O(d)$, because $A^\top = A^{-1}$ and

$$|Ax|_E^2 = (Ax)^\top(Ax) = x^\top A^\top Ax = x^\top A^{-1}Ax = |x|_E^2,$$

it holds that

$$\|A\|_2 = \sup_{0\ne x\in\mathbb{R}^d} \frac{|Ax|_E}{|x|_E} = \sup_{0\ne x\in\mathbb{R}^d} \frac{|x|_E}{|x|_E} = 1.$$

1.4 Function Spaces

This article uses standard Sobolev spaces with associated norms (e.g., see [10, 20, 21]).

1.5 Function Space $H(\mathrm{div}; D)$

Let D be a Lipschitz domain of $\mathbb{R}^d$. We denote the function space by

$$H(\mathrm{div}; D) := \left\{v \in L^2(D)^d;\ \mathrm{div}\, v \in L^2(D)\right\},$$

which is a Hilbert space with the inner product and norm:

$$\begin{aligned}(u, v)_{H(\mathrm{div};D)} &:= (u, v) + (\mathrm{div}\, u, \mathrm{div}\, v),\\ \|v\|_{H(\mathrm{div};D)} &:= (v, v)^{1/2}_{H(\mathrm{div};D)} = \Big(\|v\|^2_{L^2(D)^d} + \|\,\mathrm{div}\, v\|^2_{L^2(D)}\Big)^{1/2}.\end{aligned}$$

Theorem 1.1. The space $\mathcal{C}^\infty(\overline{D})^d$ is dense in $H(\mathrm{div}; D)$.

Proof. A proof can be found in [27, Theorem 2.4]. The condition of "boundedness" is entered into the assumptions because we use the space $\mathcal{C}^\infty(\overline{D})^d$. □

Theorem 1.2. The trace operator $\gamma^d : \mathcal{C}^\infty(\overline{D})^d \to \mathcal{C}^\infty(\overline{\partial D})$ which maps $\varphi \mapsto \varphi \cdot n|_{\partial D}$ can be extended to a continuous, linear mapping

$$\gamma^d : H(\mathrm{div}; D) \to H^{-\frac{1}{2}}(\partial D),$$

where $H^{-\frac{1}{2}}(\partial D)$ is the dual space $H^{\frac{1}{2}}(\partial D)$.

Proof. A proof can be found in [27, Theorem 2.5]. □

Theorem 1.3. The trace theorem is optimal in the sense that

$$\gamma^d : H(\mathrm{div}; D) \to H^{-\frac{1}{2}}(\partial D),$$

is surjective.

Proof. Let $\mu \in H^{-\frac{1}{2}}(\partial D)$. To show is that there exists $v \in H(\operatorname{div}; D)$ such that

$$v \cdot n = \mu \quad \text{on } \partial D,$$
$$\|v\|_{H(\operatorname{div};D)} \leq \|v \cdot n\|_{H^{-\frac{1}{2}}(\partial D)}.$$

We know that the problem

$$-\Delta\varphi + \varphi = 0 \quad \text{in } D, \quad \frac{\partial\varphi}{\partial n} = \mu \quad \text{on } \partial D$$

has a unique solution $\varphi \in H^1(D)$ satisfying

$$\|\varphi\|^2_{H^1(D)} = \langle \mu, \varphi \rangle_{\partial D} \leq \|\mu\|_{H^{-\frac{1}{2}}(\partial D)} \|\varphi\|_{H^1(D)}, \tag{1.4}$$

see [27, Section 1.4 and (1.16)]. Setting $v = \nabla\varphi$, we have $v \in H(\operatorname{div}; D)$, $v \cdot n = \mu$, and

$$\begin{aligned} \|v\|_{H(\operatorname{div};D)} &= \left(\|v\|^2_{L^2(D)^d} + \|\operatorname{div} v\|^2_{L^2(D)}\right)^{1/2} = \|\varphi\|_{H^1(D)} \\ &\leq \|\mu\|_{H^{-\frac{1}{2}}(\partial D)} = \|\mu\|_{H^{-\frac{1}{2}}(\partial D)} = \|v \cdot n\|_{H^{-\frac{1}{2}}(\partial D)}. \end{aligned}$$

□

Theorem 1.4. It holds that

$$H_0(\operatorname{div}; D) := \ker(\gamma^d) = \{v \in H(\operatorname{div}; D) : \ v \cdot n|_{\partial D} = 0\}.$$

Proof. A proof can be found in [27, Theorem 2.6]. □

Theorem 1.5. Let

$$H_\sigma := \{v \in H_0(\operatorname{div}; D) : \operatorname{div} v = 0\}.$$

It then holds that

$$L^2(D)^d = H_\sigma \oplus H^\perp,$$

where $H^\perp$ denotes the orthogonal of H_σ in $L^2(D)^d$ for the scalar product, that is,

$$H^\perp := \{v = \nabla q : \ q \in H^1(D)\}.$$

Proof. A proof can be found in [27, Theorem 2.7]. Remark that D is open, bounded, connected, and a Lipschitz set, because D is a Lipschitz domain of $\mathbb{R}^d$. □

1.6 Finite-Element-Methods-Related Symbols

1.6.1 Symbols

$\mathbb{P}^k$	Vector space of polynomials in the variables $x_1, \ldots, x_d$ of global degree at most $k \in \mathbb{N}_0$
$N^{(d,k)}$	$N^{(d,k)} := \dim(\mathbb{P}^k) = \begin{pmatrix} d+k \\ k \end{pmatrix}$
$\mathbb{RT}^k$	The Raviart–Thomas polynomial space of order $k \in \mathbb{N}_0$ as $\mathbb{RT}^k := (\mathbb{P}^k)^d + x\mathbb{P}^k$ for any $x \in \mathbb{R}^d$

$N^{(RT)}$	$N^{(RT)} := \dim RT^k$
T, $\widetilde{T}$, $\widehat{T}$, K	Closed simplices in $\mathbb{R}^d$
$\mathbb{P}^k(T)$, $\mathbb{RT}^k(T)$	$\mathbb{P}^k(T)$ (or $\mathbb{RT}^k(T)$) is spanned by the restriction to T of polynomials in $\mathbb{P}^k$ (or $\mathbb{RT}^k$)

1.6.2 Meshes

Let $\Omega \subset \mathbb{R}^d$, $d \in \{2, 3\}$, be a bounded polyhedral domain. Furthermore, we assume that Ω is convex if necessary. Let $\mathbb{T}_h = \{T\}$ be a simplicial mesh of $\overline{\Omega}$ made up of closed d-simplices, such as

$$\overline{\Omega} = \bigcup_{T \in \mathbb{T}_h} T,$$

with $h := \max_{T \in \mathbb{T}_h} h_T$, where $h_T := \operatorname{diam}(T)$. We also use a symbol ρ_T which means the radius of the largest ball inscribed in T. We assume that each face of any d-simplex T_1 in $\mathbb{T}_h$ is either a subset of the boundary $\partial\Omega$ or a face of another d-simplex T_2 in $\mathbb{T}_h$. That is, $\mathbb{T}_h$ is a simplicial mesh of $\overline{\Omega}$ without hanging nodes. Such mesh $\mathbb{T}_h$ is said to be conformal. Let $\{\mathbb{T}_h\}$ be a family of conformal meshes.

Let T be a simplex of $\mathbb{T}_h$ which is a convex full of $d+1$ vertices, $p_1, \dots, p_{d+1}$, that do not belong to the same hyperplane. Let S_i be the face of a simplex T opposite to the vertex p_i. For $d = 3$, angles between faces of a tetrahedron are called *dihedral*, whereas angles between its edges are called *solid*.

1.6.3 Broken Sobolev Spaces, Mesh faces, Averages and Jumps

Let $\mathcal{F}_h^i$ be the set of interior faces, and $\mathcal{F}_h^\partial$ be the set of faces on boundary $\partial\Omega$. We set $\mathcal{F}_h := \mathcal{F}_h^i \cup \mathcal{F}_h^\partial$. For any $F \in \mathcal{F}_h$, we define the unit normal n_F to F as follows: (i) If $F \in \mathcal{F}_h^i$ with $F = T_\natural \cap T_\sharp$, $T_\natural, T_\sharp \in \mathbb{T}_h$, $\natural > \sharp$, let n_F be the unit normal vector from $T_\natural$ to $T_\sharp$. (ii) If $F \in \mathcal{F}_h^\partial$, n_F is the unit outward normal n to $\partial\Omega$. We also use the following set. For any $F \in \mathcal{F}_h$,

$$\mathbb{T}_F := \{T \in \mathbb{T}_h : \ F \subset T\}.$$

Furthermore, for a simplex $T \subset \mathbb{R}^d$, let $\mathcal{F}_T$ be the collection of the faces of T.

We consider $\mathbb{R}^q$-valued functions for some $q \in \mathbb{N}$. Let $p \in [1, \infty]$ and $s > 0$ be a positive real number. We define a broken (piecewise) Sobolev space as

$$W^{s,p}(\mathbb{T}_h; \mathbb{R}^q) := \{v \in L^p(\Omega; \mathbb{R}^q) : \ v|_T \in W^{s,p}(T; \mathbb{R}^q) \ \forall T \in \mathbb{T}_h\}$$

with the norms

$$\|v\|_{W^{s,p}(\mathbb{T}_h;\mathbb{R}^q)} := \left(\sum_{T \in \mathbb{T}_h} \|v\|^p_{W^{s,p}(T;\mathbb{R}^q)} \right)^{1/p} \quad \text{if } p \in [1, \infty),$$
$$\|v\|_{W^{s,\infty}(\mathbb{T}_h;\mathbb{R}^q)} := \max_{T \in \mathbb{T}_h} \|v\|_{W^{s,\infty}(T;\mathbb{R}^q)}.$$

When $q = 1$, we denote $W^{s,p}(\mathbb{T}_h) := W^{s,p}(\mathbb{T}_h; \mathbb{R})$. When $p = 2$, we write $H^s(\mathbb{T}_h)^q := H^s(\mathbb{T}_h; \mathbb{R}^q) := W^{s,2}(\mathbb{T}_h; \mathbb{R}^q)$ and $H^s(\mathbb{T}_h) := W^{s,2}(\mathbb{T}_h; \mathbb{R})$. We use the norm

$$|\varphi|_{H^1(\mathbb{T}_h)} := \left(\sum_{T \in \mathbb{T}_h} \|\nabla \varphi\|^2_{L^2(T)^d} \right)^{1/2} \quad \varphi \in H^1(\mathbb{T}_h).$$

Let $\varphi \in H^1(\mathbb{T}_h)$. Suppose that $F \in \mathcal{F}_h^i$ with $F = T_\natural \cap T_\sharp$, $T_\natural, T_\sharp \in \mathbb{T}_h$, $\natural > \sharp$. We set $\varphi_\natural := \varphi|_{T_\natural}$ and $\varphi_\sharp := \varphi|_{T_\sharp}$. The jump in φ across F is defined as

$$[\![\varphi]\!] := [\![\varphi]\!]_F := \varphi_\natural - \varphi_\sharp, \quad \natural > \sharp.$$

For a boundary face $F \in \mathcal{F}_h^\partial$ with $F = \partial T \cap \partial\Omega$, $[\![\varphi]\!]_F := \varphi|_T$. For any $v \in H^1(\mathbb{T}_h)^d$, the notations

$$\begin{aligned} [\![v \cdot n]\!] &:= [\![v \cdot n]\!]_F := v_\natural \cdot n_F - v_\sharp \cdot n_F, \quad \natural > \sharp, \\ [\![v]\!] &:= [\![v]\!]_F := v_\natural - v_\sharp, \quad \natural > \sharp, \end{aligned}$$

denote the jump in the normal component of v and the jump of v. Set two nonnegative real numbers $\omega_{T_\natural,F}$ and $\omega_{T_\sharp,F}$ such that

$$\omega_{T_\natural,F} + \omega_{T_\sharp,F} = 1.$$

The skew-weighted average of φ across F is then defined as

$$\{\{\varphi\}\}_{\overline{\omega}} := \{\{\varphi\}\}_{\overline{\omega},F} := \omega_{T_\sharp,F}\varphi_\natural + \omega_{T_\natural,F}\varphi_\sharp.$$

For a boundary face $F \in \mathcal{F}_h^\partial$ with $F = \partial T \cap \partial\Omega$, $\{\{\varphi\}\}_{\overline{\omega}} := \varphi|_T$. Furthermore,

$$\{\{v\}\}_\omega := \{\{v\}\}_{\omega,F} := \omega_{T_\natural,F} v_\natural + \omega_{T_\sharp,F} v_\sharp,$$

for the weighted average of v. For any $v \in H^1(\mathbb{T}_h)^d$ and $\varphi \in H^1(\mathbb{T}_h)$,

$$[\![(v\varphi) \cdot n]\!]_F = \{\{v\}\}_{\omega,F} \cdot n_F [\![\varphi]\!]_F + [\![v \cdot n]\!]_F \{\{\varphi\}\}_{\overline{\omega},F}.$$

We define a broken gradient operator as follows. Let $p \in [1, \infty]$. For $\varphi \in W^{1,p}(\mathbb{T}_h)$, the broken gradient $\nabla_h : W^{1,p}(\mathbb{T}_h) \to L^p(\Omega)^d$ is defined by

$$(\nabla_h \varphi)|_T := \nabla(\varphi|_T) \quad \forall T \in \mathbb{T}_h,$$

and we define the broken $H(\mathrm{div}; T)$ space by

$$H(\mathrm{div}; \mathbb{T}_h) := \left\{ v \in L^2(\Omega)^d;\ v|_T \in H(\mathrm{div}; T)\ \forall T \in \mathbb{T}_h \right\},$$

and the broken divergence operator $\mathrm{div}_h : H(\mathrm{div}; \mathbb{T}_h) \to L^2(\Omega)$ such that, for all $v \in H(\mathrm{div}; \mathbb{T}_h)$,

$$(\mathrm{div}_h\, v)|_T := \mathrm{div}(v|_T) \quad \forall T \in \mathbb{T}_h.$$

1.6.4 Barycentric Coodinates

For a simplex $T \subset \mathbb{R}^d$, let $\{p_i\}_{i=1}^{d+1}$ be vertices of T and $(x_1^{(i)}, \ldots, x_d^{(i)})^T$ coordinates of p_i. We set

$$\Delta := \det \begin{pmatrix} 1 & \cdots & 1 \\ x_1^{(1)} & \cdots & x_1^{(d+1)} \\ \vdots & \vdots & \vdots \\ x_d^{(1)} & \cdots & x_d^{(d+1)} \end{pmatrix} > 0.$$

The barycentric coordinates $\{\lambda_i\}_{i=1}^{d+1}:\mathbb{R}^d\to\mathbb{R}$ of the point $p(x_1,\dots,x_d)$ with respect to $\{p_i\}_{i=1}^{d+1}$ are then defined as

$$\lambda_i(x) := \frac{1}{\Delta}\det\begin{array}{c} \begin{array}{ccccc} & & \overbrace{}^{i} & & \end{array} \\ \begin{pmatrix} 1 & \cdots & 1 & \cdots & 1 \\ x_1^{(1)} & \cdots & x_1 & \cdots & x_1^{(d+1)} \\ \vdots & & \vdots & & \vdots \\ x_d^{(1)} & \cdots & x_d & \cdots & x_d^{(d+1)} \end{pmatrix} \end{array}.$$

The barycentric coordinates have the following properties:

$$\lambda_i(p_j)=\delta_{ij}, \quad \sum_{i=1}^{d+1}\lambda_i(x)=1.$$

1.7 Useful Tools for Analysis

1.7.1 Jensen-type Inequality

Let r,s be two nonnegative real numbers and $\{x_i\}_{i\in I}$ be a finite sequence of nonnegative numbers. It then holds that

$$\begin{cases} \left(\sum_{i\in I} x_i^s\right)^{\frac{1}{s}} \le \left(\sum_{i\in I} x_i^r\right)^{\frac{1}{r}} & \text{if } r\le s, \\ \left(\sum_{i\in I} x_i^s\right)^{\frac{1}{s}} \le \operatorname{card}(I)^{\frac{r-s}{rs}}\left(\sum_{i\in I} x_i^r\right)^{\frac{1}{r}} & \text{if } r>s, \end{cases} \tag{1.5}$$

see [21, Exercise 12.1].

1.7.2 Embedding Theorems

The following is well known as the Sobolev embedding theorem.

Theorem 1.6. Let $d\ge 2$, $s>0$, and $p\in[1,\infty]$. Let $D\subset\mathbb{R}^d$ be a bounded open subset of $\mathbb{R}^d$. If D is a Lipschitz set, we then have

$$W^{s,p}(D)\hookrightarrow \begin{cases} L^q(D) & \forall q\in[p,\frac{pd}{d-sp}] \text{ if } sp<d, \\ L^q(D) & \forall q\in[p,\infty), \text{ if } sp=d, \\ L^\infty(D)\cap\mathcal{C}^{0,\xi}(\overline{D}) & \xi=1-\frac{d}{sp} \text{ if } sp>d. \end{cases} \tag{1.6}$$

Furthermore,

$$W^{s,p}(D)\hookrightarrow L^\infty(D)\cap\mathcal{C}^0(\overline{D}) \quad (\text{case } s=d \text{ and } p=1). \tag{1.7}$$

Proof. See, for example, [20, Corollary B.43, Theorem B.40] and [21, Theorem 2.31] and the references therein. □

The following is the embedding theorem related to operator from $W^{s,p}(D)$ into $L^q(S_r)$, where S_r is some plane r-dimensional piece belonging to D with dimensions $r<d$.

Theorem 1.7. Let $p,q\in[1,+\infty]$ and $s\ge 1$ be an integer. Let $D\subset\mathbb{R}^d$ be a bounded open set having piecewise smooth boundaries. The following embeddings are then continuous:

$$W^{s,p}(D)\hookrightarrow \begin{cases} L^q(S_r) & \text{if } 1\le p<\frac{d}{s},\ r>d-sp \text{ and } q\le\frac{pr}{d-sp}, \\ L^q(S_r) & \text{if } p=\frac{d}{s} \text{ for } q<+\infty. \end{cases} \tag{1.8}$$

Proof. See, for example, [43, Theorem 2.1 (p. 61)] and the references therein. □

1.7.3 Trace Theorem

The following theorem is well-known, e.g., see [21].

Theorem 1.8 (Trace)**.** Let $p \in [1, \infty)$. Let $s > \frac{1}{p}$ if $p > 1$ or $s \geq 1$ if $p = 1$. Let D be a Lipschitz domain (e.g., see [21, Definition 3.2]) in $\mathbb{R}^d$. There exists a bounded linear operator $\gamma^g : W^{s,p}(D) \to L^p(\partial D)$ such that

1. $\gamma^g(\varphi) = \varphi|_{\partial D}$, whenever φ is smooth, e.g., $\varphi \in \mathcal{C}(\overline{D})$.
2. The kernel of γ^g is $W_0^{s,p}(D)$.
3. If $s = 1$ and $p = 1$, or if $s \in (\frac{1}{2}, \frac{3}{2})$ and $p = 2$, or if $s \in (\frac{1}{p}, 1]$ and $p \notin \{1, 2\}$, then $\gamma^g : W^{s,p}(D) \to W^{s-\frac{1}{p},p}(\partial D)$ is bounded and surjective, that is, there exists C^{γ^g} such that, for every functions $g \in W^{s-\frac{1}{p},p}(\partial D)$, one can find a function $\varphi_g \in W^{s,p}(D)$, called a lifting of g, such that

$$\gamma^g(\varphi_g) = g, \quad \|\varphi_g\|_{W^{s,p}(D)} \leq C^{\gamma^g} \ell_D^{\frac{1}{p}} \|g\|_{W^{s-\frac{1}{p},p}(\partial D)}, \tag{1.9}$$

where ℓ_D is a characteristic length of D, e.g., $\ell_D := \operatorname{diam}(D)$.

Proof. See [21, Theorem 3.10], and the references therein. □

Theorem 1.9 (Trace on low-dimensional manifolds)**.** Let $p \in [1, \infty)$ and let D be a Lipschitz domain in $\mathbb{R}^d$. Let M be a smooth, or polyhedral, manifold of dimension r in $\overline{D}$, $r \in \{0, \ldots, d\}$. Then, there exists a bounded trace operator from $W^{s,p}(D)$ to $L^p(M)$, provided $sp > d - r$, or $s \geq d - r$ if $p = 1$.

Proof. See [21, Theorem 3.15]. □

1.7.4 Bramble–Hilbert–type Lemma

The Bramble–Hilbert–type lemma (e.g., see [19, 14]) plays a major role in interpolation error analysis. We use the following estimates on anisotropic meshes proposed in [3, Lemma 2.1].

Lemma 1.10. Let $D \subset \mathbb{R}^d$ be a connected open set that is star-shaped concerning balls B. Let γ be a multi-index with $m := |\gamma|$ and $\varphi \in L^1(D)$ be a function with $\partial^\gamma \varphi \in W^{\ell-m,p}(D)$, where $\ell \in \mathbb{N}$, $m \in \mathbb{N}_0$, $0 \leq m \leq \ell$, $p \in [1, \infty]$. It then holds that

$$\|\partial^\gamma(\varphi - Q^{(\ell)}\varphi)\|_{W^{\ell-m,p}(D)} \leq C^{BH} |\partial^\gamma \varphi|_{W^{\ell-m,p}(D)}, \tag{1.10}$$

where C^{BH} depends only on d, ℓ, $\operatorname{diam} D$, and $\operatorname{diam} B$, and $Q^{(\ell)}\varphi$ is defined as

$$(Q^{(\ell)}\varphi)(x) := \sum_{|\delta| \leq \ell-1} \int_B \eta(y)(\partial^\delta \varphi)(y) \frac{(x-y)^\delta}{\delta!} dy \in \mathbb{P}^{\ell-1}, \tag{1.11}$$

where $\eta \in \mathcal{C}_0^\infty(B)$ is a given function with $\int_B \eta dx = 1$.

To give local interpolation error estimates on isotropic meshes, we use the inequalities given in [18, Theorem 1.1] and [19, 14, 50] which are variants of the Bramble–Hilbert lemma.

Lemma 1.11. Let $D \subset \mathbb{R}^d$ be a bounded convex domain. Let $\varphi \in W^{m,p}(D)$ with $m \in \mathbb{N}$ and $1 \leq p \leq \infty$. There exists a polynomial $\eta \in \mathbb{P}^{m-1}$ such that

$$|\varphi - \eta|_{W^{k,p}(D)} \leq C^{BH}(d, m) \operatorname{diam}(D)^{m-k} |\varphi|_{W^{m,p}(D)}, \quad k = 0, 1, \ldots, m. \tag{1.12}$$

Proof. The proof is found in [18, Theorem 1.1]. $\square$

Remark 1.12. In [14, Lemma 4.3.8], the Bramble–Hilbert lemma is given as follows. Let B be a ball in $D \subset \mathbb{R}^d$ such that D is star-shaped with respect to B and its radius $r > \frac{1}{2} r_{\max}$, where $r_{\max} := \sup\{r : D \text{ is star-shaped with respect to a ball of radius } r\}$. Let $\varphi \in W^{m,p}(D)$ with $m \in \mathbb{N}$ and $1 \le p \le \infty$. There exists a polynomial $\eta \in \mathbb{P}^{m-1}$ such that

$$|\varphi - \eta|_{W^{k,p}(D)} \le C^{BH}(d, m, \gamma)\, \mathrm{diam}(D)^{m-k} |\varphi|_{W^{m,p}(D)}, \quad k = 0, 1, \ldots, m. \tag{1.13}$$

Here, γ is called the chunkiness parameter of D, which is defined by

$$\gamma := \frac{\mathrm{diam}(D)}{r_{\max}}.$$

The main drawback is that the constant $C^{BH}(d, m, \gamma)$ depends on the chunkiness parameter. Meanwhile, the constant $C^{BH}(d, m)$ of the estimate (1.12) does not depend on the geometric parameter γ.

Remark 1.13. For general Sobolev spaces $W^{m,p}(\Omega)$, the upper bounds on the constant $C^{BH}(d, m)$ are not given, as far as we know. However, when $p = 2$, the following result has been obtained by Verf*ü*rth [50].

Let $D \subset \mathbb{R}^d$ be a bounded convex domain. Let $\varphi \in H^m(D)$ with $m \in \mathbb{N}$. There exists a polynomial $\eta \in \mathbb{P}^{m-1}$ such that

$$|\varphi - \eta|_{H^k(D)} \le C^{BH}(d, k, m)\, \mathrm{diam}(D)^{m-k} |\varphi|_{H^m(D)}, \quad k = 0, 1, \ldots, m-1. \tag{1.14}$$

Verf*ü*rth has given upper bounds on the constants in the estimates such that

$$C^{BH}(d, k, m) \le \pi^{k-m} \binom{d+k-1}{k}^{1/2} \frac{\{(m-k)!\}^{1/2}}{\{\left[\frac{m-k}{d}\right]!\}^{d/2}},$$

where $[x]$ denotes the largest integer less than or equal to x.

As an example, let us consider the case $d = 3$, $k = 1$, and $m = 2$. We then have

$$C^{BH}(3, 1, 2) \le \frac{\sqrt{3}}{\pi},$$

thus on the standard reference element $\widehat{T}$, we obtain

$$|\hat{\varphi} - \hat{\eta}|_{H^1(\widehat{T})} \le \frac{\sqrt{6}}{\pi} |\hat{\varphi}|_{H^2(\widehat{T})} \quad \forall \hat{\varphi} \in H^2(\widehat{T}),$$

because $\mathrm{diam}(\widehat{T}) = \sqrt{2}$.

1.7.5 Poincaré inequality

Theorem 1.14 (Poincaré inequality)**.** Let $D \subset \mathbb{R}^d$ be a convex domain with diameter $\mathrm{diam}(D)$. It then holds that, for $\varphi \in H^1(D)$ with $\int_D \varphi dx = 0$,

$$\|\varphi\|_{L^2(D)} \le \frac{\mathrm{diam}(D)}{\pi} |\varphi|_{H^1(D)}. \tag{1.15}$$

Proof. The proof is found in [45, Theorem 3.2], also see [47]. $\square$

Remark 1.15. The coefficient $\frac{1}{\pi}$ of (1.15) may be improved.

1.8 Abbreviated expression

FE	Finite Element
FEMs	Finite Element Methods
CR	Crouzeix–Raviart
RT	Raviart–Thomas

2 Isotropic and Anisotropic Mesh Elements

In the context of FEMs, mesh elements can be classified based on their geometric properties. *An isotropic mesh element* has equal or nearly equal edge lengths and angles, resulting in a balanced shape. In contrast, *an anisotropic mesh element* features significant variation in edge lengths and angles.

Consider the following examples: Let $s, \delta \in \mathbb{R}_+$, and $\varepsilon \geq 1$, $\varepsilon \in \mathbb{R}$.

Example 2.1. In the case of the simplex $T \subset \mathbb{R}^2$ defined by the vertices $p_1 := (0,0)^\top$, $p_2 := (2s,0)^\top$, and $p_3 := (s, \delta s)^\top$, the triangle is classified as follows:

- If $\delta \approx 1$, the triangle T is considered an isotropic mesh element.
- Conversely, if δ is much less than 1, i.e., $\delta \ll 1$, the triangle T becomes an anisotropic mesh element.

Example 2.2. In this case, consider the simplex $T \subset \mathbb{R}^2$ defined by the vertices $p_1 := (0,0)^\top$, $p_2 := (2s,0)^\top$, and $p_3 := (s, s^\varepsilon)^\top$. Here, the vertex p_3 introduces a parameter ε that can influence the shape of the simplex. The classification of this simplex as isotropic or anisotropic depends on the value of ε:

- If $\varepsilon = 1$, the triangle maintains a balanced shape, making it isotropic.
- If $\varepsilon > 1$, the triangle becomes flat when $s \ll 1$, resulting in an anisotropic mesh element.

Example 2.3. Consider the simplex $T \subset \mathbb{R}^2$ defined by the vertices $p_1 := (0,0)^\top$, $p_2 := (s,0)^\top$, and $p_3 := (0, \delta s)^\top$. In this configuration, the classification of the simplex as isotropic or anisotropic depends on the value of δ:

- If $\delta \approx 1$, the triangle is an isotropic mesh element.
- If $\delta \ll 1$, the triangle becomes an anisotropic mesh element.

Example 2.4. Let $T \subset \mathbb{R}^2$ be the simplex with vertices $p_1 := (0,0)^\top$, $p_2 := (s,0)^\top$, and $p_3 := (0, s^\varepsilon)^\top$. In this case, the classification of the simplex as isotropic or anisotropic depends on the value of ε:

- If $\varepsilon = 1$, the triangle is isotropic because the height from p_3 is equal to the base length.
- If $\varepsilon > 1$, the triangle will be classified as anisotropic, as the edge lengths will differ significantly when $s \ll 1$.

Example 2.5. Let $T \subset \mathbb{R}^2$ be the simplex with vertices $p_1 := (0,0)^\top$, $p_2 := (s,0)^\top$ and $p_3 := (s^\delta, s^\varepsilon)^\top$. The classification of the simplex into two types of anisotropic structures is determined by the values of δ and ε:

- If $1 < \varepsilon < \delta$, the triangle is flattened so that the point p_3 approaches the point p_1, i.e. the origin as $s \to 0$.
- If $1 < \delta < \varepsilon$, the triangle is flattened so that point p_3 approaches a point on the straight line $\overline{p_1 p_2}$ that does not include points p_1 and p_2 as $s \to 0$.

3 Classical Geometric Conditions

3.1 Classical Interpolation Error Estimate

Let $\widehat{T} \subset \mathbb{R}^d$ and $T \subset \mathbb{R}^d$ be a reference element and a simplex, respectively. Let these two elements be affine equivalent. Let us consider two finite elements $\{\widehat{T}, \widehat{P}, \widehat{\Sigma}\}$ and $\{T, P, \Sigma\}$ with associated normed vector spaces $V(\widehat{T})$ and $V(T)$. The transformation Φ_T takes the form

$$\Phi_T : \widehat{T} \ni \hat{x} \mapsto \Phi_T(\hat{x}) := B_T \hat{x} + b_T \in T,$$

where $B_T \in \mathbb{R}^{d\times d}$ is an invertible matrix and $b_T \in \mathbb{R}^d$. Let $I_T : V(T) := W^{2,p}(T) \to P := \mathbb{P}^1(T)$ with $p \in [1,\infty]$ be an interpolation on T with $I_T p = p$ for any $p \in \mathcal{P}^1(T)$. According to the classical theory (e.g., see [17, 20]), there exists a positive constant c, independent of h_T, such that

$$|\varphi - I_T\varphi|_{W^{1,p}(T)} \leq c \left(\|B_T\|_2 \|B_T^{-1}\|_2 \right) \|B_T\|_2 |\varphi|_{W^{2,p}(T)}.$$

Here, the quantity $\|B_T\|_2 \|B_T^{-1}\|_2$ is called the *Euclidean condition number* of B_T. By standard estimates (e.g., see [20, Lemma 1.100]), we have

$$\|B_T\|_2 \|B_T^{-1}\|_2 \leq c \frac{h_T}{\rho_T}, \quad \|B_T\|_2 \leq c h_T.$$

It thus holds that

$$|\varphi - I_T\varphi|_{W^{1,p}(T)} \leq c \frac{h_T}{\rho_T} h_T |\varphi|_{W^{2,p}(T)}. \tag{3.1}$$

As a geometric condition, the *shape-regularity condition* is well known to obtain global interpolation error estimates. This condition is stated as follows.

Condition 3.1 (Shape-regularity condition)**.** There exists a constant $\gamma_1 > 0$ such that

$$\rho_T \geq \gamma_1 h_T \quad \forall \mathbb{T}_h \in \{\mathbb{T}_h\}, \quad \forall T \in \mathbb{T}_h. \tag{3.2}$$

Under Condition 3.1, that is, when the quantity $\frac{h_T}{\rho_T}$ is bounded on each T, it holds that

$$|\varphi - I_h\varphi|_{W^{1,p}(\Omega)} \leq c h |\varphi|_{W^{2,p}(\Omega)},$$

where $I_h\varphi$ is the standard global linear interpolation of φ on $\mathbb{T}_h$.

3.2 Regular Mesh Conditions

Geometric conditions equivalent to the shape-regularity condition are known; that is, the following three conditions are equivalent to the shape-regularity condition (3.2). A proof can be found in [13, Theorem 1].

Condition 3.2 (Zlámal's condition)**.** There exists a constant $\gamma_2 > 0$ such that for any $\mathbb{T}_h \in \{\mathbb{T}_h\}$, any simplex $T \in \mathbb{T}_h$ and any dihedral angle ψ and for $d = 3$, also any solid angle θ of T, we have

$$\psi \geq \gamma_2, \quad \theta \geq \gamma_2. \tag{3.3}$$

Condition 3.3. There exists a constant $\gamma_3 > 0$ such that for any $\mathbb{T}_h \in \{\mathbb{T}_h\}$ and any simplex $T \in \mathbb{T}_h$, we have

$$|T|_d \geq \gamma_3 h_T^d. \tag{3.4}$$

Condition 3.4. There exists a constant $\gamma_4 > 0$ such that for any $\mathbb{T}_h \in \{\mathbb{T}_h\}$ and any simplex $T \in \mathbb{T}_h$, we have

$$|T|_d \geq \gamma_4 |B_d^T|, \tag{3.5}$$

where $B^T \supset T$ is the circumscribed ball of T.

Note 3.5. If Condition 3.1 or 3.2 or 3.3 or 3.4 holds, a family of simplicial partitions is called *regular*.

Note 3.6. Condition 3.2 was presented by Zlámal [52] in 1968. The condition is called the *minimum-angle condition* and guarantees the convergence of finite element methods for linear elliptic problems on $\mathbb{R}^2$. Zlámal's condition can be generalised into $\mathbb{R}^n$ for any $n \in \{2, 3, \ldots\}$. Later, the shape-regularity condition (the inscribed ball condition) was introduced; see [17]. Triangles or tetrahedra cannot be too flat in a shape-regular family of triangulations.

Note 3.7. Condition 3.3 seems to be simpler than Condition 3.1, Condition 3.2 and Condition 3.4. Therefore, it may be useful to analyse theoretical finite element methods and implement finite element codes to keep nondegenerate mesh partitions.

3.3 What happens when anisotropic meshes are used?

Using the equivalence conditions in Section 3.2, the error estimate (3.1) is rewritten as

$$|\varphi - I_T \varphi|_{W^{1,p}(T)} \leq c \frac{h_T^2}{|T|_2} h_T |\varphi|_{W^{2,p}(T)}. \tag{3.6}$$

We considered the following five anisotropic elements as in Section 2: Let $0 < s, \delta \ll 1$, $s, \delta \in \mathbb{R}$, and $\varepsilon > 1$, $\varepsilon \in \mathbb{R}$.

Example 3.8. Let $T \subset \mathbb{R}^2$ be the simplex with vertices $p_1 := (0,0)^\top$, $p_2 := (2s,0)^\top$, and $p_3 := (s, \delta s)^\top$. Then, we have that $h_T = 2s$, $|T|_2 = \delta s^2$, and

$$\frac{h_T^2}{|T|_2} = \frac{4}{\delta} < +\infty.$$

Therefore, the shape regularity is satisfied. The estimate (3.6) is as follows:

$$|\varphi - I_T \varphi|_{W^{1,p}(T)} \leq \frac{c}{\delta} h_T |\varphi|_{W^{2,p}(T)}.$$

When $\delta \ll 1$, the interpolation error (3.6) may be large.

Example 3.9. Let $T \subset \mathbb{R}^2$ be the simplex with vertices $p_1 := (0,0)^\top$, $p_2 := (2s,0)^\top$, and $p_3 := (s, s^\varepsilon)^\top$. Then, we have that $h_T = 2s$, $|T|_2 = s^{1+\varepsilon}$ and

$$\frac{h_T^2}{|T|_2} = 4 s^{1-\varepsilon} \to \infty \text{ as } s \to 0.$$

Therefore, the shape-regularity is not satisfied. In this case, when $\varepsilon > 2$, the estimate (3.6) diverges as $s \to 0$.

Example 3.10. Let $T \subset \mathbb{R}^2$ be the simplex with vertices $p_1 := (0,0)^\top$, $p_2 := (s,0)^\top$, and $p_3 := (0, \delta s)^\top$. Then, we have that $h_T = s\sqrt{1+\delta^2} \approx s$, $|T|_2 = \frac{1}{2}\delta s^2$ and

$$\frac{h_T^2}{|T|_2} = \frac{2(1+\delta^2)}{\delta} < +\infty.$$

Therefore, the shape regularity is satisfied. The estimate (3.6) is as follows:

$$|\varphi - I_T\varphi|_{W^{1,p}(T)} \leq \frac{c}{\delta} h_T |\varphi|_{W^{2,p}(T)}. \tag{3.7}$$

It is implied that the interpolation error (3.7) may be large when $\delta \ll 1$.

Example 3.11. Let $T \subset \mathbb{R}^2$ be the simplex with vertices $p_1 := (0,0)^\top$, $p_2 := (s,0)^\top$, and $p_3 := (0, s^\varepsilon)^\top$. Subsequently, we obtain $h_T = \sqrt{s^2 + s^{2\varepsilon}} \approx s$, $|T|_2 = \frac{1}{2}s^{1+\varepsilon}$ and

$$\frac{h_T^2}{|T|_2} = \frac{2(s^2 + s^{2\varepsilon})}{s^{1+\varepsilon}} \to \infty \text{ as } s \to 0.$$

Therefore, the shape-regularity condition is not satisfied. In this case, it is implied that the estimate (3.6) diverges as $s \to 0$.

Example 3.12. Let $T \subset \mathbb{R}^2$ be the simplex with vertices $p_1 := (0,0)^\top$, $p_2 := (s,0)^\top$ and $p_3 := (s^\delta, s^\varepsilon)^\top$. If $1 < \varepsilon < \delta$, we have $h_T = \sqrt{(s - s^\delta)^2 + s^{2\varepsilon}}$, $|T|_2 = \frac{1}{2}s^{1+\varepsilon}$ and

$$\frac{h_T^2}{|T|_2} = \frac{2(s - s^\delta)^2 + s^{2\varepsilon}}{s^{1+\varepsilon}} \leq cs^{1-\varepsilon} \to \infty \text{ as } s \to 0.$$

Therefore, the shape-regularity condition is not satisfied. In this case, it is implied that the estimate (3.6) diverges as $s \to 0$. If $1 < \delta < \varepsilon$, we have $h_T = \sqrt{(s - s^\delta)^2 + s^{2\varepsilon}}$, $|T|_2 = \frac{1}{2}s^{1+\varepsilon}$ and

$$\frac{h_T^2}{|T|_2} = \frac{2(s - s^\delta)^2 + s^{2\varepsilon}}{s^{1+\varepsilon}} \leq cs^{1-\varepsilon} \to \infty \text{ as } s \to 0.$$

Therefore, the shape-regularity condition is not satisfied. In this case, it is implied that the estimate (3.6) diverges as $s \to 0$.

Remark 3.13. As will be explained later, the factor $\frac{1}{\delta}$ in Example 3.10 is violated. The interpolation error estimate converges in the cases of Example 3.11 and Example 3.12 with $1 < \varepsilon < \delta$ using new precise interpolation error estimates under more relaxed geometric conditions.

4 Classical Relaxed Geometric Conditions

4.1 Semi-regular Mesh Conditions for $d = 2$

In 1957, Synge [48, Section 3.8] proposed the following condition.

Condition 4.1 (Synge's condition)**.** There exists $\frac{\pi}{3} \leq \gamma_5 < \pi$ such that, for any $\mathbb{T}_h \in \{\mathbb{T}_h\}$ and any simplex $T \in \mathbb{T}_h$,

$$\theta_{T,\max} \leq \gamma_5, \tag{4.1}$$

where $\theta_{T,\max}$ is the maximal angle of T.

Under Condition 4.1, Synge proved an optimal interpolation error estimate as follows.

$$\|\varphi - I_h\varphi\|_{W^{1,p}(\Omega)} \leq ch|\varphi|_{W^{2,p}(\Omega)} \quad \text{for } p = \infty.$$

The inequality (4.1) is called *Synge's condition* or the *maximum-angle condition*. In 1976, several author's [7, 9, 26, 40] independently proved the convergence of finite element for $p < \infty$. It ensures that finite elements converge effectively when the minimum angle approaches zero as the mesh size decreases. If this condition is not met, the accuracy of interpolation for linear triangular elements can suffer, similar to the absence of Zlámal's condition, see e.g. [7, p. 223]. This underscores the importance of keeping proper geometric constraints to ensure reliable outcomes in numerical methods. Synge's condition is essential in finite element analysis.

In [41], Křížek proposed the following circumscribed ball condition for $d = 2$ which is equivalent to Synge's condition.

Condition 4.2. There exists $\gamma_6 > 0$ such that, for any $\mathbb{T}_h \in \{\mathbb{T}_h\}$ and any simplex $T \in \mathbb{T}_h$,

$$\frac{R_2}{h_T} \leq \gamma_6, \tag{4.2}$$

where R_2 is the radius of the circumscribed ball of $T \subset \mathbb{R}^2$.

Note 4.3. If Condition 4.1 or 4.2 holds, the associated families of partitions are called *semi-regular*.

Remark 4.4. Assume that Condition 3.3 holds, that is, there exists a constant $\gamma_3 > 0$ such that for any $\mathbb{T}_h \in \{\mathbb{T}_h\}$ and any simplex $T \in \mathbb{T}_h$, we have

$$|T| \geq \gamma_3 h_T^2.$$

Let $T \subset \mathbb{R}^2$ be the triangle with vertices P_1, P_2 and P_3 such that the maximum angle $\theta_{T,\max}$ of T is $\angle P_2 P_1 P_3$. We then have $h_T = |P_2 P_3|$ and

$$\frac{R_2}{h_T} = \frac{|P_2 P_3|}{2 h_T \sin \theta_{T,\max}} = \frac{|P_1 P_2||P_1 P_3|}{2|P_1 P_2||P_1 P_3| \sin \theta_{T,\max}} \leq c \frac{h_T^2}{|T|} \leq \frac{c}{\gamma_3} =: \gamma_6.$$

This implies that each regular family is semi-regular. However, the converse implication does not hold.

4.2 Semi-regular Mesh Conditions for $d = 3$

Synge's condition (4.1) is extended to the case of tetrahedra in [42].

Condition 4.5. There exists a constant $0 < \gamma_7 < \pi$ such that for any $\mathbb{T}_h \in \{\mathbb{T}_h\}$ and any simplex $T \in \mathbb{T}_h$,

$$\theta_{T,\max} \leq \gamma_7, \tag{4.3a}$$

$$\psi_{T,\max} \leq \gamma_7, \tag{4.3b}$$

where $\theta_{T,\max}$ is the maximum angle of all triangular faces of the tetrahedron T and $\psi_{T,\max}$ is the maximum dihedral angle of T.

Remark 4.6. The theory of anisotropic interpolation has been advanced through extensive research ([4, 3, 15]).

Question 4.7. Is there a semi-regularity condition which equivalent to Synge's condition (4.3) for $d = 3$?

Remark 4.8. This article introduces a novel geometric condition intended to serve as an alternative to Synge's condition specifically for three-dimensional cases.

5 Settings for New Interpolation Theory

5.1 Reference Elements

We first define the reference elements $\widehat{T} \subset \mathbb{R}^d$.

Two-dimensional case

Let $\widehat{T} \subset \mathbb{R}^2$ be a reference triangle with vertices $\hat{p}_1 := (0,0)^\top$, $\hat{p}_2 := (1,0)^\top$, and $\hat{p}_3 := (0,1)^\top$.

Three-dimensional case

In the three-dimensional case, we consider the following two cases: (i) and (ii); see Condition 5.2.

Let $\widehat{T}_1$ and $\widehat{T}_2$ be reference tetrahedra with the following vertices:

(i) $\widehat{T}_1$ has vertices $\hat{p}_1 := (0,0,0)^\top$, $\hat{p}_2 := (1,0,0)^\top$, $\hat{p}_3 := (0,1,0)^\top$, and $\hat{p}_4 := (0,0,1)^T$;

(ii) $\widehat{T}_2$ has vertices $\hat{p}_1 := (0,0,0)^\top$, $\hat{p}_2 := (1,0,0)^\top$, $\hat{p}_3 := (1,1,0)^\top$, and $\hat{p}_4 := (0,0,1)^T$.

Therefore, we set $\widehat{T} \in \{\widehat{T}_1, \widehat{T}_2\}$. Note that the case (i) is called *the regular vertex property*, see [2].

5.2 Two-step Affine Mapping

To an affine simplex $T \subset \mathbb{R}^d$, we construct two affine mappings $\Phi_{\widetilde{T}} : \widehat{T} \to \widetilde{T}$ and $\Phi_T : \widetilde{T} \to T$. First, we define the affine mapping $\Phi_{\widetilde{T}} : \widehat{T} \to \widetilde{T}$ as

$$\Phi_{\widetilde{T}} : \widehat{T} \ni \hat{x} \mapsto \tilde{x} := \Phi_{\widetilde{T}}(\hat{x}) := A_{\widetilde{T}} \hat{x} \in \widetilde{T}, \tag{5.1}$$

where $A_{\widetilde{T}} \in \mathbb{R}^{d \times d}$ is an invertible matrix. We then define the affine mapping $\Phi_T : \widetilde{T} \to T$ as follows:

$$\Phi_T : \widetilde{T} \ni \tilde{x} \mapsto x := \Phi_T(\tilde{x}) := A_T \tilde{x} + b_T \in T, \tag{5.2}$$

where $b_T \in \mathbb{R}^d$ is a vector and $A_T \in O(d)$ denotes the rotation and mirror-imaging matrix. We define the affine mapping $\Phi : \widehat{T} \to T$ as

$$\Phi := \Phi_T \circ \Phi_{\widetilde{T}} : \widehat{T} \ni \hat{x} \mapsto x := \Phi(\hat{x}) = (\Phi_T \circ \Phi_{\widetilde{T}})(\hat{x}) = A\hat{x} + b_T \in T,$$

where $A := A_T A_{\widetilde{T}} \in \mathbb{R}^{d \times d}$.

Construct mapping $\Phi_{\widetilde{T}} : \widehat{T} \to \widetilde{T}$

We consider the affine mapping (5.1). We define the matrix $A_{\widetilde{T}} \in \mathbb{R}^{d \times d}$ as follows. We first define the diagonal matrix as

$$\widehat{A} := \operatorname{diag}(h_1, \ldots, h_d), \quad h_i \in \mathbb{R}_+ \quad \forall i, \tag{5.3}$$

where $\mathbb{R}_+$ denotes the set of positive real numbers.

For $d = 2$, we define the regular matrix $\widetilde{A} \in \mathbb{R}^{2 \times 2}$ as

$$\widetilde{A} := \begin{pmatrix} 1 & s \\ 0 & t \end{pmatrix}, \tag{5.4}$$

with the parameters

$$s^2 + t^2 = 1, \quad t > 0.$$

For the reference element $\widehat{T}$, let $\mathfrak{T}^{(2)}$ be a family of triangles.

$$\widetilde{T} = \Phi_{\widetilde{T}}(\widehat{T}) = A_{\widetilde{T}}(\widehat{T}), \quad A_{\widetilde{T}} := \widetilde{A}\widehat{A}$$

with the vertices $\tilde{p}_1 := (0,0)^\top$, $\tilde{p}_2 := (h_1, 0)^\top$ and $\tilde{p}_3 := (h_2 s, h_2 t)^\top$. Then, $h_1 = |\tilde{p}_1 - \tilde{p}_2| > 0$ and $h_2 = |\tilde{p}_1 - \tilde{p}_3| > 0$.

For $d = 3$, we define the regular matrices $\widetilde{A}_1, \widetilde{A}_2 \in \mathbb{R}^{3\times 3}$ as follows:

$$\widetilde{A}_1 := \begin{pmatrix} 1 & s_1 & s_{21} \\ 0 & t_1 & s_{22} \\ 0 & 0 & t_2 \end{pmatrix}, \ \widetilde{A}_2 := \begin{pmatrix} 1 & -s_1 & s_{21} \\ 0 & t_1 & s_{22} \\ 0 & 0 & t_2 \end{pmatrix} \tag{5.5}$$

with the parameters

$$\begin{cases} s_1^2 + t_1^2 = 1, \ s_1 > 0, \ t_1 > 0, \ h_2 s_1 \leq h_1/2, \\ s_{21}^2 + s_{22}^2 + t_2^2 = 1, \ t_2 > 0, \ h_3 s_{21} \leq h_1/2. \end{cases}$$

Therefore, we set $\widetilde{A} \in \{\widetilde{A}_1, \widetilde{A}_2\}$. For the reference elements $\widehat{T}_i$, $i = 1, 2$, let $\mathfrak{T}_i^{(3)}$, $i = 1, 2$, be a family of tetrahedra.

$$\widetilde{T}_i = \Phi_{\widetilde{T}_i}(\widehat{T}_i) = A_{\widetilde{T}_i}(\widehat{T}_i), \quad A_{\widetilde{T}_i} := \widetilde{A}_i \widehat{A}, \quad i = 1, 2,$$

with the vertices

$$\begin{aligned} &\tilde{p}_1 := (0,0,0)^\top, \ \tilde{p}_2 := (h_1, 0, 0)^\top, \ \tilde{p}_4 := (h_3 s_{21}, h_3 s_{22}, h_3 t_2)^\top, \\ &\begin{cases} \tilde{p}_3 := (h_2 s_1, h_2 t_1, 0)^\top & \text{for case (i)}, \\ \tilde{p}_3 := (h_1 - h_2 s_1, h_2 t_1, 0)^\top & \text{for case (ii)}. \end{cases} \end{aligned}$$

Subsequently, $h_1 = |\tilde{p}_1 - \tilde{p}_2| > 0$, $h_3 = |\tilde{p}_1 - \tilde{p}_4| > 0$, and

$$h_2 = \begin{cases} |\tilde{p}_1 - \tilde{p}_3| > 0 & \text{for case (i)}, \\ |\tilde{p}_2 - \tilde{p}_3| > 0 & \text{for case (ii)}. \end{cases}$$

Construct mapping $\Phi_T : \widetilde{T} \to T$

We determine the affine mapping (5.2) as follows. Let $T \in \mathbb{T}_h$ have vertices p_i ($i = 1, \ldots, d+1$). Let $b_T \in \mathbb{R}^d$ be the vector and $A_T \in O(d)$ be the rotation and mirror imaging matrix such that

$$p_i = \Phi_T(\tilde{p}_i) = A_T \tilde{p}_i + b_T, \quad i \in \{1, \ldots, d+1\},$$

where vertices p_i ($i = 1, \ldots, d+1$) satisfy the following conditions:

Condition 5.1 (Case in which $d = 2$)**.** Let $T \in \mathbb{T}_h$ have vertices p_i ($i = 1, \ldots, 3$). We assume that $\overline{p_2 p_3}$ is the longest edge of T, that is, $h_T := |p_2 - p_3|$. We set $h_1 = |p_1 - p_2|$ and $h_2 = |p_1 - p_3|$. We then assume that $h_2 \leq h_1$. Because $\frac{1}{2} h_T < h_1 \leq h_T$, $h_1 \approx h_T$.

Condition 5.2 (Case in which $d = 3$)**.** Let $T \in \mathbb{T}_h$ have vertices p_i ($i = 1, \ldots, 4$). Let L_i ($1 \leq i \leq 6$) be the edges of T. We denote by $L_{\min}$ the edge of T with the minimum length; that is, $|L_{\min}| = \min_{1 \leq i \leq 6} |L_i|$. We set $h_2 := |L_{\min}|$ and assume that

$$\text{the endpoints of } L_{\min} \text{ are either } \{p_1, p_3\} \text{ or } \{p_2, p_3\}.$$

Among the four edges sharing an endpoint with $L_{\min}$, we consider the longest edge $L_{\max}^{(\min)}$. Let p_1 and p_2 be the endpoints of edge $L_{\max}^{(\min)}$. Thus, we have

$$h_1 = |L_{\max}^{(\min)}| = |p_1 - p_2|.$$

We consider cutting $\mathbb{R}^3$ with a plane that contains the midpoint of the edge $L_{\max}^{(\min)}$ and is perpendicular to the vector $p_1 - p_2$. Thus, there are two cases.

(Type i) p_3 and p_4 belong to the same half-space;

(Type ii) p_3 and p_4 belong to different half-spaces.

In each case, we set

(Type i) p_1 and p_3 as the endpoints of $L_{\min}$, that is, $h_2 = |p_1 - p_3|$;

(Type ii) p_2 and p_3 as the endpoints of $L_{\min}$, that is, $h_2 = |p_2 - p_3|$.

Finally, we set $h_3 = |p_1 - p_4|$. We implicitly assume that p_1 and p_4 belong to the same half-space. Additionally, note that $h_1 \approx h_T$.

Note 5.3. As an example, we define the matrices A_T as

$$A_T := \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}, \quad A_T := \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

where θ denotes the angle.

Note 5.4. None of the lengths of the edges of a simplex or the measures of the simplex are changed by the transformation, i.e.,

$$h_i \leq h_T, \quad i = 1, \ldots, d. \tag{5.6}$$

5.3 Additional Notations and Assumptions

For convenience, we introduce the following additional notation. We define a parameter $\widetilde{\mathscr{H}_i}$, $i = 1, \ldots, d$, as

$$\begin{cases} \widetilde{\mathscr{H}_1} := h_1, \quad \widetilde{\mathscr{H}_2} := h_2 t \quad \text{if } d = 2, \\ \widetilde{\mathscr{H}_1} := h_1, \quad \widetilde{\mathscr{H}_2} := h_2 t_1, \quad \widetilde{\mathscr{H}_3} := h_3 t_2 \quad \text{if } d = 3, \end{cases}$$

see Fig. 1.

Assumption 5.5. In an anisotropic interpolation error analysis, we impose a geometric condition for the simplex $\widetilde{T}$:

1. If $d = 2$, there are no additional conditions;

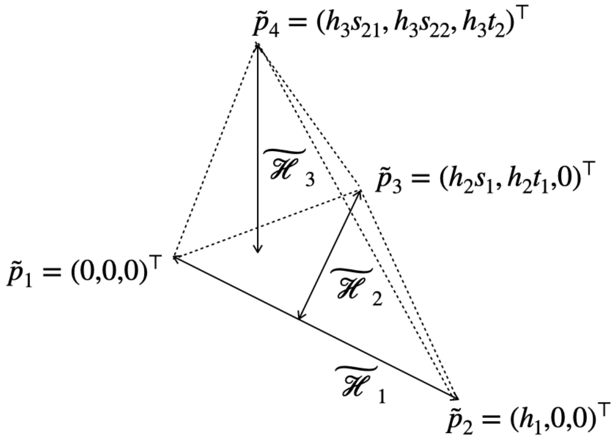

Fig. 1: New parameters $\widetilde{\mathscr{H}_i}$, $i = 1, 2, 3$

2. If $d = 3$, there exists a positive constant M independent of $h_{\widetilde{T}}$ such that $|s_{22}| \le M \frac{h_2 t_1}{h_3}$. Note that if $s_{22} \neq 0$, this condition means that the order concerning h_T of h_3 coincides with the order of h_2, and if $s_{22} = 0$, the order of h_3 may be different from that of h_2.

We define the vectors $r_n \in \mathbb{R}^d$ and $n = 1, \dots, d$ as follows: If $d = 2$,

$$r_1 := \frac{p_2 - p_1}{|p_2 - p_1|}, \quad r_2 := \frac{p_3 - p_1}{|p_3 - p_1|},$$

see Fig. 2, and if $d = 3$,

$$r_1 := \frac{p_2 - p_1}{|p_2 - p_1|}, \quad r_3 := \frac{p_4 - p_1}{|p_4 - p_1|}, \quad \begin{cases} r_2 := \dfrac{p_3 - p_1}{|p_3 - p_1|}, & \text{for case (i)}, \\ r_2 := \dfrac{p_3 - p_2}{|p_3 - p_2|} & \text{for case (ii)}, \end{cases}$$

see Fig 3 for (Type i) and Fig 4 for (Type ii). Furthermore, we define the vectors $\tilde{r}_n \in \mathbb{R}^d$ and $n = 1, \dots, d$ as follows. If $d = 2$,

$$\tilde{r}_1 := (1, 0)^\top, \quad \tilde{r}_2 := (s, t)^\top,$$

and if $d = 3$,

$$\tilde{r}_1 := (1, 0, 0)^\top, \quad \tilde{r}_3 := (s_{21}, s_{22}, t_2)^\top, \quad \begin{cases} \tilde{r}_2 := (s_1, t_1, 0)^\top & \text{for case (i)}, \\ \tilde{r}_2 := (-s_1, t_1, 0)^\top & \text{for case (ii)}. \end{cases}$$

Remark 5.6. The vectors $\tilde{r}_i$, $i \in \{1, \dots, d\}$ are unit vectors. Indeed, if $d = 2$,

$$|\tilde{r}_1|_E = 1, \quad |\tilde{r}_2|_E = \sqrt{s^2 + t^2} = 1,$$

if $d = 3$,

$$|\tilde{r}_1|_E = 1, \quad |\tilde{r}_2|_E = \sqrt{s_1^2 + t_1^2} = 1, \quad |\tilde{r}_3|_E = \sqrt{s_{21}^2 + s_{22}^2 + t_2^2} = 1.$$

Remark 5.7. Let $A_T \in O(d)$ be the orthogonal matrix defined in (5.2). Then,

$$r_i = A_T \tilde{r}_i, \quad i = 1, \dots, d. \tag{5.7}$$

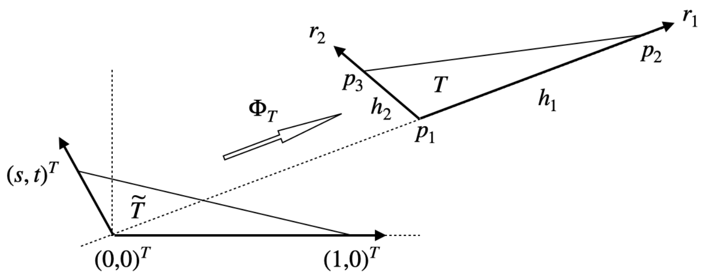

Fig. 2: Affine mapping Φ_T and vectors r_i, $i = 1, 2$

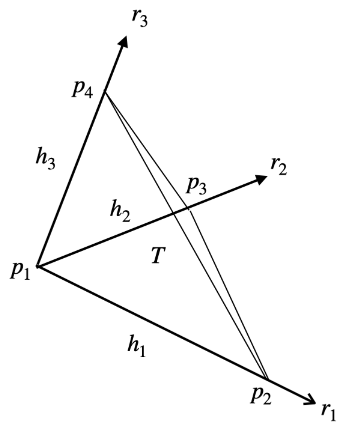

Fig. 3: (Type i) Vectors r_i, $i = 1, 2, 3$

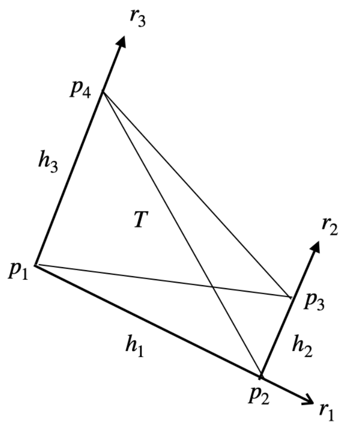

Fig. 4: (Type ii) Vectors r_i, $i = 1, 2, 3$

For a sufficiently smooth function φ and vector function $v := (v_1, \ldots, v_d)^\top$, we define the directional derivative of $i \in \{1, \ldots, d\}$ as:

$$\frac{\partial \varphi}{\partial r_i} := (r_i \cdot \nabla_x)\varphi = \sum_{i_0=1}^{d} (r_i)_{i_0} \frac{\partial \varphi}{\partial x_{i_0}},$$
$$\frac{\partial v}{\partial r_i} := \left(\frac{\partial v_1}{\partial r_i}, \ldots, \frac{\partial v_d}{\partial r_i} \right)^\top = ((r_i \cdot \nabla_x) v_1, \ldots, (r_i \cdot \nabla_x) v_d)^\top,$$

and for a sufficiently smooth function $\tilde{\varphi}$ and vector function $\tilde{v} := (\tilde{v}_1, \ldots, \tilde{v}_d)^\top$,

$$\frac{\partial \tilde{\varphi}}{\partial \tilde{r}_i} := (\tilde{r}_i \cdot \nabla_{\tilde{x}})\tilde{\varphi} = \sum_{i_0=1}^{d} (\tilde{r}_i)_{i_0} \frac{\partial \tilde{\varphi}}{\partial \tilde{x}_{i_0}},$$
$$\frac{\partial \tilde{v}}{\partial \tilde{r}_i} := \left(\frac{\partial \tilde{v}_1}{\partial \tilde{r}_i}, \ldots, \frac{\partial \tilde{v}_d}{\partial \tilde{r}_i} \right)^\top = ((\tilde{r}_i \cdot \nabla_{\tilde{x}}) \tilde{v}_1, \ldots, (\tilde{r}_i \cdot \nabla_{\tilde{x}}) \tilde{v}_d)^\top.$$

For a multi-index $\beta = (\beta_1, \ldots, \beta_d) \in \mathbb{N}_0^d$, we use the following notation.

$$\partial^\beta \varphi := \frac{\partial^{|\beta|} \varphi}{\partial x_1^{\beta_1} \ldots \partial x_d^{\beta_d}}, \quad \partial_r^\beta \varphi := \frac{\partial^{|\beta|} \varphi}{\partial r_1^{\beta_1} \ldots \partial r_d^{\beta_d}}.$$

Note that $\partial^\beta \varphi \neq \partial_r^\beta \varphi$.

6 New Semi-regularity Condition

6.1 New Geometric Parameter and Condition

We proposed a new geometric parameter H_T in [37].

Definition 6.1. Parameter H_T is defined as follows:

$$H_T := \frac{\prod_{i=1}^d h_i}{|T|_d} h_T.$$

We introduce geometric conditions to obtain the optimal convergence rate of the anisotropic error estimates.

Condition 6.2. A family of meshes $\{\mathbb{T}_h\}$ is semi-regular if there exists $\gamma_0 > 0$ such that

$$\frac{H_T}{h_T} \leq \gamma_0 \quad \forall \mathbb{T}_h \in \{\mathbb{T}_h\}, \quad \forall T \in \mathbb{T}_h. \tag{6.1}$$

Remark 6.3. The geometric condition in (6.1) is equivalent to the maximum angle condition (Section 7).

Remark 6.4. We consider the good elements on the meshes in Section 8. On anisotropic meshes, good elements may satisfy the following conditions:

$(d = 2)$ $h_2 \approx h_2 t$;

$(d = 3)$ $h_2 \approx h_2 t_1$ and $h_3 \approx h_3 t_2$.

6.2 Properties of the New Geometric Parameter

We first show the relation between h_T and H_T.

Lemma 6.5. It holds that

$$h_T \leq \frac{1}{2} H_T \quad \text{if } d = 2, \tag{6.2}$$

$$h_T < \frac{1}{6} H_T \quad \text{if } d = 3. \tag{6.3}$$

Proof. We consider for each dimension, $d = 2, 3$.

Two-dimensional case. By constructing the standard element in the two-dimensional case, the angle $\theta_{\max} := \angle p_2 p_1 p_3$ is the maximum angle of T. We then have $\frac{\pi}{3} < \theta_{\max} < \pi$, that is, $0 < \sin\theta_{\max} \leq 1$. Therefore, it holds that

$$H_T = \frac{h_1 h_2}{|T|_2} h_T = \frac{2}{\sin\theta_{\max}} h_T \geq 2 h_T.$$

We here used the fact that $|T|_2 = \frac{1}{2} h_1 h_2 \sin\theta_{\max}$.

Three-dimensional case. We denote by ϕ_T the angle between the base $\triangle p_1 p_2 p_3$ of T and the segment $\overline{p_1 p_4}$. Recall that there are two types of standard elements, (Type i) or (Type ii). We denote by θ_T

(Type i) the angle between the segments $\overline{p_1 p_2}$ and $\overline{p_1 p_3}$, that is, $\theta_T := \angle p_2 p_1 p_3$, or

(Type ii) the angle between the segments $\overline{p_2 p_1}$ and $\overline{p_2 p_3}$, that is, $\theta_T := \angle p_1 p_2 p_3$.

We set $t_1 := \sin\theta_T$ and $t_2 := \sin\phi_T$. By constructing the standard element in the three-dimensional case, the angle $\angle p_1p_3p_2$ is the maximum angle of the base $\triangle p_1p_2p_3$ of T. Therefore, we have $0 < \theta_T < \frac{\pi}{2}$. Because $0 < \phi_T < \pi$, it holds that

$$H_T = \frac{h_1h_2h_3}{|T|_3}h_T = \frac{6}{\sin\theta_T\sin\phi_T}h_T > 6h_T.$$

We here used the fact that $|T|_3 = \frac{1}{6}h_1h_2h_3\sin\theta_T\sin\phi_T$. $\square$

We introduce another geometric parameter regarding Definition 6.1.

Definition 6.6 (Another parameter H_T^*)**.** For $T \in \mathbb{T}_h$, we denote by L_i edges of the simplex T. We define the new parameter H_T^* as

$$H_T^* := \frac{h_T^2}{|T|_2}\min_{1\le i\le 3}|L_i| \quad \text{if } d=2, \quad H_T^* := \frac{h_T^2}{|T|_3}\min_{1\le i,j\le 6, i\neq j}|L_i||L_j| \quad \text{if } d=3. \tag{6.4}$$

The parameters H_T^* and H_T are equivalent.

Lemma 6.7. It holds that

$$\frac{1}{2}H_T^* < H_T < 2H_T^*. \tag{6.5}$$

Furthermore, H_T^* is equivalent to the circumradius R_2 of T in the two-dimensional case.

Proof. We consider for each dimension, $d = 2, 3$.

Two-dimensional case. Let L_i $(i = 1, 2, 3)$ denote edges of the triangle T with $|L_1| \le |L_2| \le |L_3|$. It obviously holds that $h_2 = |L_1|$ and $h_T = |L_3| = h_T$. Because $h_2 \le h_1 < 2h_T$ and $h_T < h_1 + h_2 \le 2h_1$ for the triangle $\triangle p_1p_2p_3$, it holds that

$$\frac{1}{2}h_T < h_1 = |L_2| < 2h_T = 2h_T.$$

We thus have

$$\frac{1}{2}H_T^* = \frac{1}{2}\frac{|L_1|}{|T|_2}h_T^2 < H_T = \frac{h_1h_2}{|T|_2}h_T < 2\frac{|L_1|}{|T|_2}h_T^2 = 2H_T^*.$$

Furthermore, it holds that

$$2R_2 = 2\frac{|L_1||L_2||L_3|}{4|T|_2} < H_T^* = \frac{|L_1|}{|T|_2}h_T^2 < 8\frac{|L_1||L_2||L_3|}{4|T|_2} = 8R_2.$$

Three-dimensional case. Let L_i $(i = 1, \dots, 6)$ denote edges of the triangle T with $|L_1| \le |L_2| \le \cdots \le |L_6|$. It obviously holds that $h_2 = |L_1|$ and $h_T = |L_6|$. Recall that there are two types of standard elements, (Type i) or (Type ii).

(Type i) We set $h_4 := |p_3 - p_4|$, $h_5 := |p_2 - p_4|$ and $h_6 := |p_2 - p_3|$. Because $h_1 = |E_{\max}^{(\min)}| = |p_1 - p_2|$ is the longest edge among the four edges that share an endpoint with L_1, it holds that

$$h_2 \le \min\{h_3, h_4, h_6\} \le \max\{h_3, h_4, h_6\} \le h_1. \tag{6.6}$$

Because p_1 and p_4 belong to the same half-space for the triangle $\triangle p_1p_2p_4$, it holds that

$$\begin{cases} h_3 \le h_5 \le h_1 = h_T & \text{or} \\ h_3 \le h_1 \le h_5 = h_T. \end{cases}$$

We thus have

$$\begin{cases} h_3 \le h_5 \le h_1 = h_T \quad \text{or} \\ h_3 \le h_1 \le h_T < 2h_1, \quad \frac{1}{2}h_T < h_1 \le h_T. \end{cases}$$

Because $h_3 \le h_5$, the length of the edge L_2 is equal to the one of h_3, h_4 or h_6.

Assume that $|L_2| = h_3$. We then have

$$\frac{1}{2}H_T^* = \frac{1}{2}\frac{|L_1||L_2|}{|T|_3}h_T^2 < H_T = \frac{h_1 h_2 h_3}{|T|_3}h_T \le \frac{|L_1||L_2|}{|T|_3}h_T^2 = H_T^*(< 2H_T^*).$$

Assume that $|L_2| = h_4$. We consider the triangle $\triangle p_1 p_3 p_4$. From the assumption, we have $h_2 \le h_4 \le h_3$ and $\frac{1}{2}h_3 < h_4 \le h_3$. We then obtain

$$\frac{1}{2}H_T^* = \frac{1}{2}\frac{|L_1||L_2|}{|T|_3}h_T^2 < H_T = \frac{h_1 h_2 h_3}{|T|_3}h_{T^s} < 2\frac{|L_1||L_2|}{|T|_3}h_T^2 = 2H_T^*.$$

Assume that $|L_2| = h_6$. We consider the triangle $\triangle p_1 p_2 p_3$. Because p_1 and p_3 belong to the same half-space for the triangle $\triangle p_1 p_2 p_3$, it holds that $h_2 \le h_6 \le h_1$ and $\frac{1}{2}h_1 < h_6 \le h_1$. From (6.6), we have

$$\frac{1}{2}h_3 \le \frac{1}{2}h_1 < h_6 \le h_1.$$

Because $h_6 \le h_3$, we then obtain

$$\frac{1}{2}H_T^* = \frac{1}{2}\frac{|L_1||L_2|}{|T|_3}h_T^2 < H_T = \frac{h_1 h_2 h_3}{|T|_3}h_T < 2\frac{|L_1||L_2|}{|T|_3}h_T^2 = 2H_T^*.$$

(Type ii) We set $h_4 := |p_3 - p_4|$, $h_5 := |p_2 - p_4|$, and $h_6 := |p_1 - p_3|$. Because $h_1 = |E_{\max}^{(\min)}| = |p_1 - p_2|$ is the longest edge among the four edges that share an endpoint with L_1, it holds that

$$h_2 \le \min\{h_4, h_5, h_6\} \le \max\{h_4, h_5, h_6\} \le h_1. \tag{6.7}$$

Because p_1 and p_4 belong to the same half-space for the triangle $\triangle p_1 p_2 p_4$ and (6.7), it holds that

$$h_3 \le h_5 \le h_1.$$

This implies that $h_1 = h_T$. Therefore, the length of the edge L_2 is equal to the one of h_3, h_4, or h_6.

Assume that $|L_2| = h_3$. We then have

$$\begin{aligned}\left(\frac{1}{2}H_T^* <\right) H_T^* = \frac{|L_1||L_2|}{|T|_3}h_T^2 = H_T &= \frac{h_1 h_2 h_3}{|T|_3}h_T \\ &= \frac{|L_1||L_2|}{|T|_3}h_T^2 = H_T^*(< 2H_T).\end{aligned}$$

Assume that $|L_2| = h_4$. For the triangle $\triangle p_2 p_3 p_4$, we have

$$h_2 \le h_4 \le h_5 < 2h_4.$$

Because $h_3 \leq h_5$ and $h_4 \leq h_3$, it holds that

$$\left(\frac{1}{2}H_T^* <\right) H_T^* = \frac{|L_1||L_2|}{|T|_3} h_T^2 \leq H_T = \frac{h_1 h_2 h_3}{|T|_3} h_T < 2\frac{|L_1||L_2|}{|T|_3} h_T^2 = 2H_T^*.$$

Assume that $|L_2| = h_6$. We have $h_1 < h_2 + h_6 < 2h_6$ for the triangle $\triangle p_1 p_2 p_3$. Therefore, since $h_6 \leq h_3 \leq h_1$, we obtain

$$\left(\frac{1}{2}H_T^* <\right) H_T^* = \frac{|L_1||L_2|}{|T|_3} h_T^2 \leq H_T = \frac{h_1 h_2 h_3}{|T|_3} h_T < 2\frac{|L_1||L_2|}{|T|_3} h_T^2 = 2H_T^*.$$

□

6.3 Euclidean Condition Number

Examining the Euclidean condition number is useful for deriving appropriate interpolation error estimates.

Lemma 6.8. It holds that

$$\|\widehat{A}\|_2 \leq h_T, \quad \|\widehat{A}\|_2 \|\widehat{A}^{-1}\|_2 = \frac{\max\{h_1, \cdots, h_d\}}{\min\{h_1, \cdots, h_d\}}, \tag{6.8a}$$

$$\|\widetilde{A}\|_2 \leq \begin{cases} \sqrt{2} & \text{if } d = 2, \\ 2 & \text{if } d = 3, \end{cases} \qquad \|\widetilde{A}\|_2 \|\widetilde{A}^{-1}\|_2 \leq \begin{cases} \frac{h_1 h_2}{|T|_2} = \frac{H_T}{h_T} & \text{if } d = 2, \\ \frac{2}{3}\frac{h_1 h_2 h_3}{|T|_3} = \frac{2}{3}\frac{H_T}{h_T} & \text{if } d = 3, \end{cases} \tag{6.8b}$$

$$\|A_T\|_2 = 1, \quad \|A_T^{-1}\|_2 = 1. \tag{6.8c}$$

where a parameter H_T is defined in Definition 6.1. Furthermore, we have

$$|\det(A_{\widetilde{T}})| = |\det(\widetilde{A})||\det(\widehat{A})| = \frac{|T|_d}{|\widetilde{T}|_d}\frac{|\widetilde{T}|_d}{|\widehat{T}|_d} = d!|T|_d, \quad |\det(A_T)| = 1. \tag{6.9}$$

Proof. We first show the equality (6.9). Because

$$\int_T dx = \int_{\widetilde{T}} |\det(A_T)| d\tilde{x}, \quad \int_{\widetilde{T}} d\tilde{x} = \int_{\widehat{T}} |\det(A_{\widetilde{T}})| d\hat{x},$$

and $|T|_d = |\widetilde{T}|_d$, we conclude (6.9).

We show the equality (6.8a). From

$$(\widehat{A})^\top \widehat{A} = \operatorname{diag}(h_1^2, \ldots, h_d^2), \quad \widehat{A}^{-1}\widehat{\mathcal{A}}^{-\top} = \operatorname{diag}(h_1^{-2}, \ldots, h_d^{-2}),$$

we have

$$\|\widehat{A}\|_2 = \lambda_{\max}(\widehat{A}^\top \widehat{A})^{\frac{1}{2}} = \max\{h_1, \cdots, h_d\} \leq h_T,$$

and

$$\|\widehat{A}\|_2 \|\widehat{A}^{-1}\|_2 = \lambda_{\max}(\widehat{A}^\top \widehat{A})^{\frac{1}{2}} \lambda_{\max}(\widehat{A}^{-1}\widehat{A}^{-\top})^{\frac{1}{2}} = \frac{\max\{h_1, \cdots, h_d\}}{\min\{h_1, \cdots, h_d\}},$$

which leads to (6.8a).

We next show the equality (6.8b). We consider for each dimension, $d = 2, 3$.

Two-dimensional case. Because

$$\widetilde{A}^\top \widetilde{A} = \begin{pmatrix} 1 & s \\ s & 1 \end{pmatrix}, \quad \widetilde{A}^{-1}\widetilde{A}^{-\top} = \frac{1}{t^2}\begin{pmatrix} 1 & -s \\ -s & 1 \end{pmatrix}, \quad |s| \leq 1,$$

we have

$$\|\widetilde{A}\|_2 = \lambda_{\max}(\widetilde{A}^\top \widetilde{A})^{\frac{1}{2}} \leq (1 + |s|)^{\frac{1}{2}} \leq \sqrt{2},$$

and

$$\|\widetilde{A}\|_2 \|\widetilde{A}^{-1}\|_2 = \lambda_{\max}(\widetilde{A}^\top \widetilde{A})^{\frac{1}{2}} \lambda_{\max}(\widetilde{A}^{-1}\widetilde{A}^{-\top})^{\frac{1}{2}} \leq \frac{2}{t} = \frac{h_1 h_2}{|T|_d},$$

which leads to (6.8b) for $d = 2$. Here, we used the fact that $|\widetilde{T}|_d = \frac{1}{2} h_1 h_2 t$ and $|T|_d = |\widetilde{T}|_d$.

Three-dimensional case. The matrices $\widetilde{A}_1$ and $\widetilde{A}_2$ introduced in (5.5) can be decomposed as $\widetilde{A}_1 = \widetilde{M}_0 \widetilde{M}_1$ and $\widetilde{A}_2 = \widetilde{M}_0 \widetilde{M}_2$ with

$$\widetilde{M}_0 := \begin{pmatrix} 1 & 0 & s_{21} \\ 0 & 1 & s_{22} \\ 0 & 0 & t_2 \end{pmatrix}, \ \widetilde{M}_1 := \begin{pmatrix} 1 & s_1 & 0 \\ 0 & t_1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \ \widetilde{M}_2 := \begin{pmatrix} 1 & -s_1 & 0 \\ 0 & t_1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

The eigenvalues of $\widetilde{M}_2^\top \widetilde{M}_2$ coincide with those of $\widetilde{M}_1^\top \widetilde{M}_1$, and only Case (i) is shown.

We have the inequalities

$$\begin{aligned} \|\widetilde{A}_1\|_2 &= \lambda_{\max}(\widetilde{A}_1^\top \widetilde{A}_1)^{\frac{1}{2}} \leq \lambda_{\max}(\widetilde{M}_0^\top \widetilde{M}_0)^{\frac{1}{2}} \lambda_{\max}(\widetilde{M}_1^\top \widetilde{M}_1)^{\frac{1}{2}} \\ &\leq \left(1 + \sqrt{s_{21}^2 + s_{22}^2}\right)^{\frac{1}{2}} (1 + |s_1|)^{\frac{1}{2}} \leq 2, \end{aligned}$$

and

$$\begin{aligned} \|\widetilde{A}_1\|_2 \|\widetilde{A}_1^{-1}\|_2 &= \lambda_{\max}(\widetilde{A}_1^\top \widetilde{A}_1)^{\frac{1}{2}} \lambda_{\max}(\widetilde{A}_1^{-1}\widetilde{A}_1^{-\top})^{\frac{1}{2}} \\ &\leq \frac{\left(1 + \sqrt{s_{21}^2 + s_{22}^2}\right)(1 + |s_1|)}{t_1 t_2} \leq \frac{4}{t_1 t_2} = \frac{2}{3} \frac{h_1 h_2 h_3}{|T|_d}, \end{aligned}$$

where we used the fact that $|\widetilde{T}|_d = \frac{1}{6} h_1 h_2 h_3 t_1 t_2$ and $|T|_d = |\widetilde{T}|_d$.

Because the length of all edges of a simplex and measure of the simplex is not changed by a rotation and mirror imaging matrix and $A_T, A_T^{-1} \in O(d)$,

$$\|A_T\|_2 = 1, \quad \|A_T^{-1}\|_2 = 1,$$

which is (6.8c). □

7 New Geometric Mesh Condition and the Maximum-angle Condition

7.1 Statements

We state the following theorems concerning the new condition.

Theorem 7.1. Condition 6.2 holds if and only if Condition 4.1 holds when $d = 2$.

Proof. In the case of $d = 2$, we use the previous result presented in [41]; i.e., there exists a constant $\gamma_6 > 0$ such that

$$\frac{R_2}{h_T} \leq \gamma_6 \quad \forall \mathbb{T}_h \in \{\mathbb{T}_h\}, \quad \forall T \in \mathbb{T}_h,$$

if and only if Condition 4.1 is satisfied. Combining this result with H_T being equivalent to the circumradius R_2 of T (Lemma 6.7), we have the desired conclusion. □

Theorem 7.2. Condition 6.2 holds if and only if Condition 4.5 holds when $d = 3$.

The proof can be found in [36]. Preparation is needed to prove the three-dimensional case. The following subsection shows the symbols used only in this section.

7.2 Notation

Let $T \in \mathbb{T}_h$ be the standard element in $\mathbb{R}^3$ with vertices, P_1, P_2, P_3 and P_4. Let F_i be the face of a simplex T opposite to the vertex P_i. We denote by $\psi^{i,j}$ (Table 5) the angle between the face F_i and the face F_j, see Figure 5. Note that $\psi^{i,j} = \psi^{j,i}$. Furthermore, we denote by θ^i_j (Table 6) the internal angle at the vertex P_j on the face F_i and by ϕ^i_j (Table 7) the angle between the face F_i and the segment $\overline{P_j P_i}$.

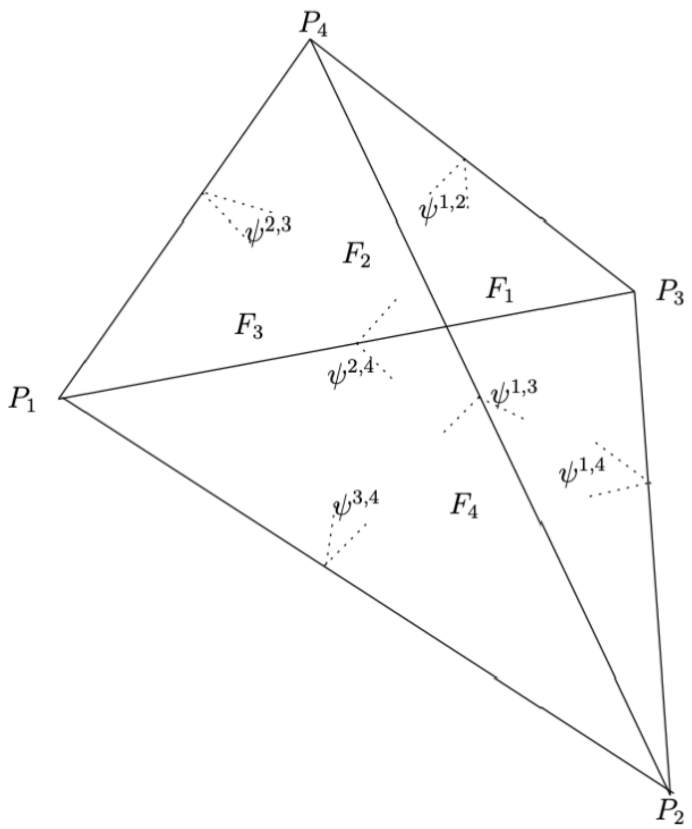

Fig. 5: Tetrahedra

7.3 Preliminaries: Part 1

We introduce three lemmata.

Table 5: $\psi^{i,j}$

	F_1	F_2	F_3	F_4
F_1	-	$\psi^{1,2}$	$\psi^{1,3}$	$\psi^{1,4}$
F_2	$\psi^{2,1}$	-	$\psi^{2,3}$	$\psi^{2,4}$
F_3	$\psi^{3,1}$	$\psi^{3,2}$	-	$\psi^{3,4}$
F_4	$\psi^{4,1}$	$\psi^{4,2}$	$\psi^{4,3}$	-

Table 6: θ^i_j

	F_1	F_2	F_3	F_4
P_1	-	θ^2_1	θ^3_1	θ^4_1
P_2	θ^1_2	-	θ^3_2	θ^4_2
P_3	θ^1_3	θ^2_3	-	θ^4_3
P_4	θ^1_4	θ^2_4	θ^3_4	-

Table 7: ϕ^i_j

	F_1	F_2	F_3	F_4
P_1	-	ϕ^2_1	ϕ^3_1	ϕ^4_1
P_2	ϕ^1_2	-	ϕ^3_2	ϕ^4_2
P_3	ϕ^1_3	ϕ^2_3	-	ϕ^4_3
P_4	ϕ^1_4	ϕ^2_4	ϕ^3_4	-

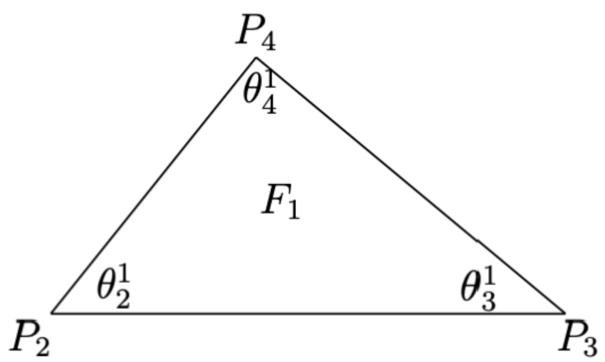

Fig. 6: Face 1

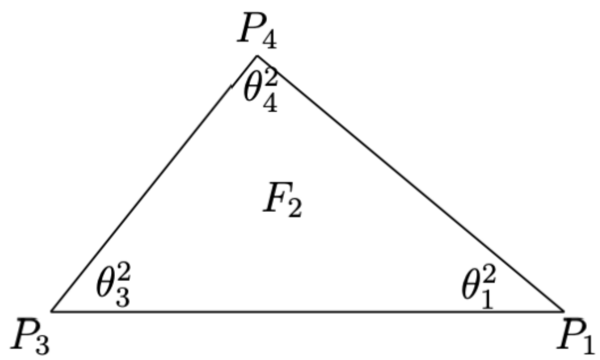

Fig. 7: Face 2

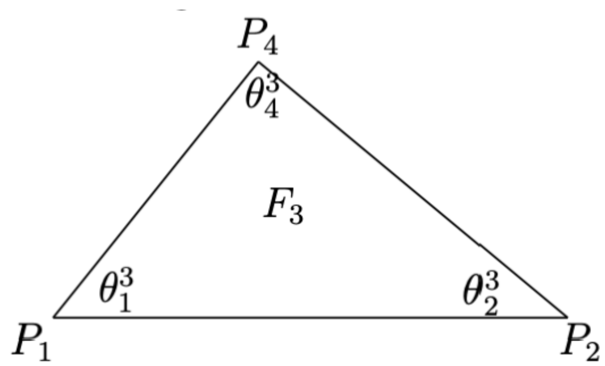

Fig. 8: Face 3

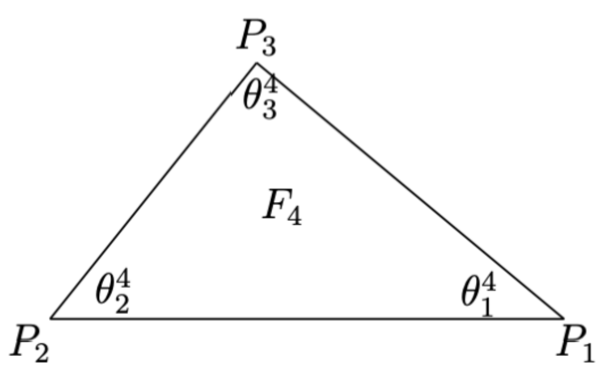

Fig. 9: Face 4

Lemma 7.3. Let $K \subset \mathbb{R}^2$ be a simplex and let θ_1, θ_2 and θ_3 be internal angles of K with $\theta_1 \le \theta_2 \le \theta_3$. If there exists $0 < \theta_0 < \pi$, $\theta_0 \in \mathbb{R}$, such that $\theta_3 \le \theta_0$, we then have

$$\sin\theta_2,\ \sin\theta_3 \ge \min\left\{\sin\frac{\pi-\theta_0}{2}, \sin\theta_0\right\}.$$

Proof. Because $\theta_1 + \theta_2 + \theta_3 = \pi$ and $\theta_1 \le \theta_2 \le \theta_3$, we have

$$\theta_0 \ge \theta_3 \ge \theta_2 \ge \frac{\theta_1+\theta_2}{2} \ge \frac{\pi-\theta_3}{2} \ge \frac{\pi-\theta_0}{2},$$

which leads to the target inequality. □

Lemma 7.4. Let $K \subset \mathbb{R}^2$ be a simplex with internal angles θ_1, θ_2 and θ_3. For any fixed $\gamma \in \mathbb{R}$ with $0 < \gamma < \pi$, we assume that $\pi - \gamma \le \theta_i$, $i \in \{1, 2, 3\}$. We then have $\theta_{i+1},\ \theta_{i+2} \le \gamma$, where the indices i, $i+1$ and $i+2$ have to be understood "mod 3".

Proof. Because $\theta_1 + \theta_2 + \theta_3 = \pi$, we have

$$\theta_{i+1} = \pi - \theta_i - \theta_{i+2} < \pi - \theta_i \le \pi - (\pi - \gamma) = \gamma.$$

□

Lemma 7.5. Let $\gamma \in \mathbb{R}$ with $\frac{\pi}{3} \le \gamma < \pi$. It then holds that

$$0 < \frac{\cos\gamma + 1}{\sin\frac{\gamma}{2} + 1} \le 1.$$

Proof. Because $\cos\gamma = 1 - 2\sin^2\frac{\gamma}{2}$, we have

$$\frac{\cos\gamma + 1}{\sin\frac{\gamma}{2} + 1} = \frac{2 - 2\sin^2\frac{\gamma}{2}}{\sin\frac{\gamma}{2} + 1} = 2\left(1 - \sin\frac{\gamma}{2}\right).$$

Therefore, for $\frac{\pi}{3} \le \gamma < \pi$, the target inequality holds. □

7.4 Preliminaries: Part 2

Lemma 7.6 (Cosine rules for the sides and for the angles)**.** It holds that

$$\begin{aligned}
\cos\theta_j^{j+3} &= \cos\theta_j^{j+1}\cos\theta_j^{j+2} + \sin\theta_j^{j+1}\sin\theta_j^{j+2}\cos\psi^{j+1,j+2}, && (7.1a)\\
\cos\theta_j^{j+1} &= \cos\theta_j^{j+2}\cos\theta_j^{j+3} + \sin\theta_j^{j+2}\sin\theta_j^{j+3}\cos\psi^{j+2,j+3}, && (7.1b)\\
\cos\theta_j^{j+2} &= \cos\theta_j^{j+3}\cos\theta_j^{j+1} + \sin\theta_j^{j+3}\sin\theta_j^{j+1}\cos\psi^{j+3,j+1}, && (7.1c)\\
\cos\psi^{j+1,j+2} &= \sin\psi^{j+2,j+3}\sin\psi^{j+3,j+1}\cos\theta_j^{j+3} - \cos\psi^{j+2,j+3}\cos\psi^{j+3,j+1}, && (7.1d)\\
\cos\psi^{j+2,j+3} &= \sin\psi^{j+3,j+1}\sin\psi^{j+1,j+2}\cos\theta_j^{j+1} - \cos\psi^{j+3,j+1}\cos\psi^{j+1,j+2}, && (7.1e)\\
\cos\psi^{j+3,j+1} &= \sin\psi^{j+1,j+2}\sin\psi^{j+2,j+3}\cos\theta_j^{j+2} - \cos\psi^{j+1,j+2}\cos\psi^{j+2,j+3}, && (7.1f)
\end{aligned}$$

where the indices j, $j+1$, $j+2$ and $j+3$ have to be understood "mod 4".

Proof. A proof can be found in [24, 49]. □

Lemma 7.7. Let $\gamma_{\max} \in \mathbb{R}$ with $\frac{\pi}{3} \le \gamma_{\max} < \pi$ satisfy Condition 4.5 for the maximum solid $\theta_{T,\max}$ and the maximum dihedral $\psi_{T,\max}$ of T. Assume that for each $j = 1, 2$, θ_j^4 is not the minimum angle of $\triangle P_1 P_2 P_3$ and $\theta_j^4 < \frac{\pi}{2}$. Then, setting $\delta := \delta(\gamma_{\max})$, $0 < \delta \le \frac{\pi}{2}$ such that

$$\sin\delta = \left(\frac{\cos\gamma_{\max} + 1}{\sin\frac{\gamma_{\max}}{2} + 1} \right)^{1/2},$$

it holds that

$$\psi^{j+1,4} \ge \delta, \quad \text{or} \quad \psi^{3,4} \ge \delta, \tag{7.2}$$

where the indices j and $j+1$ have to be understood "mod 2".

Proof. From Lemma 7.5, we have

$$0 < \frac{\cos\gamma_{\max} + 1}{\sin\frac{\gamma_{\max}}{2} + 1} \le 1,$$

because $\frac{\pi}{3} \le \gamma_{\max} < \pi$. Therefore, δ is well-defined.

We use proof by contradiction. Suppose that

$$0 < \psi^{j+1,4} < \delta, \quad 0 < \psi^{3,4} < \delta,$$

that is,

$$0 < \sin\psi^{j+1,4} \sin\psi^{3,4} < \sin^2\delta, \text{ and } 1 > \cos\psi^{j+1,4}\cos\psi^{3,4} > \cos^2\delta \ge 0.$$

From Lemma 7.3 and assumption, we have

$$\frac{\pi - \gamma_{\max}}{2} \le \theta_j^4 < \frac{\pi}{2},$$

which implies

$$0 < \cos\theta_j^4 \le \cos\left(\frac{\pi - \gamma_{\max}}{2} \right) = \sin\frac{\gamma_{\max}}{2}.$$

We thus obtain

$$\sin\psi^{j+1,4}\sin\psi^{3,4}\cos\theta_j^4 < \sin^2\delta \sin\frac{\gamma_{\max}}{2}.$$

Using the cosine rule (7.1d) with $j = 1$ and the above inequalities yield

$$\begin{aligned} \cos\psi_{2,3} &= \sin\psi^{3,4}\sin\psi^{4,2}\cos\theta_1^4 - \cos\psi^{3,4}\cos\psi^{4,2} \\ &< \sin^2\delta\sin\frac{\gamma_{\max}}{2} - (1 - \sin^2\delta) \\ &= \frac{\cos\gamma_{\max} + 1}{\sin\frac{\gamma_{\max}}{2} + 1}\left(\sin\frac{\gamma_{\max}}{2} + 1\right) - 1 = \cos\gamma_{\max}. \end{aligned}$$

This is contradiction for the maximum-angle condition $0 < \psi^{2,3} \le \gamma_{\max} < \pi$, that is, $\cos\psi^{2,3} \ge \cos\gamma_{\max}$.

Analogously, using the cosine rule (7.1f) with $j = 2$ and the above inequalities yield

$$\begin{aligned} \cos\psi^{1,3} &= \sin\psi^{3,4}\sin\psi^{4,1}\cos\theta_2^4 - \cos\psi^{3,4}\cos\psi^{4,1} \\ &< \cos\gamma_{\max}. \end{aligned}$$

This is contradiction for the maximum-angle condition $0 < \psi^{1,3} \le \gamma_{\max} < \pi$, that is, $\cos\psi^{1,3} \ge \cos\gamma_{\max}$. □

Corollary 7.8. For each $j = 1, 2$, under assumptions in Lemma 7.7, it holds that setting $C_0 := \min\{\delta, \gamma_{\max}\}$,

$$\sin \psi^{j+1,4} \geq C_0, \quad \text{or} \quad \sin \psi^{3,4} \geq C_0$$

where the indices j and $j+1$ have to be understood "mod 2".

Lemma 7.9. For any $i, j \in \{1, 2, 3, 4\}$, $i \neq j$ and $k \in \{1, 2, 3, 4\}$, $k \neq i, j$, it holds that

$$\sin \phi_j^i = \sin \theta_j^k \sin \psi^{k,i}.$$

Proof. We only show the case $i = 4$, $j = 1$ and $k = 2$. We then have

$$\sin \phi_1^4 = |\overline{P_1 P_4}| \sin \theta_1^2 \times \frac{1}{|\overline{P_1 P_4}|} \sin \psi^{2,4} = \sin \theta_1^2 \sin \psi^{2,4}.$$

□

Lemma 7.10. Assume that there exists a positive constant M_j $(j = 1, 2)$ with $0 < M_j < 1$ such that

$$\sin \theta_j^4 \sin \phi_1^4 > M_j, \quad j = 1, 2.$$

Setting $\gamma(M_j) := \pi - \sin^{-1} M_j$ $(j = 1, 2)$, we have $\frac{\pi}{2} < \gamma(M_j) < \pi$ and it holds that for each $j = 1, 2$,

$$\begin{aligned}
&\theta_1^4, \ \theta_2^4, \ \theta_3^4 < \gamma(M_j), \\
&\theta_2^3, \ \theta_4^3, \ \theta_3^2, \ \theta_4^2, \ \theta_1^2, \ \theta_1^3, \ \psi^{2,4}, \ \psi^{3,4} < \gamma(M_j).
\end{aligned}$$

Proof. From assumption, we have, for each $j = 1, 2$,

$$\begin{aligned}
&\sin \theta_j^4 \geq \sin \theta_j^4 \sin \phi_1^4 > M_j, \\
&\sin \phi_1^4 > M_j.
\end{aligned}$$

The definition of $\gamma(M_j)$ and Lemma 7.4 yield, for each $j = 1, 2$,

$$\begin{aligned}
&\pi - \gamma < \theta_j^4 < \gamma(M_j), \quad \theta_{j+1}^4 < \gamma(M_j), \quad \theta_{j+2}^4 < \gamma(M_j), \\
&\pi - \gamma < \phi_1^4 < \gamma(M_j),
\end{aligned}$$

where the indices j, $j+1$ and $j+2$ have to be understood "mod 3".

We obtain, from Lemma 7.9,

$$\sin \phi_1^4 = \sin \theta_1^2 \sin \psi^{2,4} = \sin \theta_1^3 \sin \psi^{3,4} > M_j, \quad j = 1, 2.$$

We then have, for each $j = 1, 2$,

$$\sin \theta_1^2, \ \sin \psi^{2,4}, \ \sin \theta_1^3, \ \sin \psi^{3,4} > M_j,$$

that is,

$$\pi - \gamma(M_j) < \theta_1^2, \ \theta_1^3, \ \psi^{2,4}, \ \psi^{3,4} < \gamma(M_j).$$

On $\triangle P_1 P_2 P_4$ and $\triangle P_1 P_3 P_4$, using Lemma 7.4 yields

$$\theta_2^3, \ \theta_4^3, \ \theta_3^2, \ \theta_4^2 < \gamma(M_j), \quad j = 1, 2.$$

□

By analogous argument with Lemma 7.10, we get the following two lemmata.

Lemma 7.11. Assume that there exists M_3 with $0 < M_3 < 1$ such that

$$\sin\theta_3^1 \sin\phi_3^1 > M_3.$$

Setting $\gamma(M_3) := \pi - \sin^{-1} M_3$, we have $\frac{\pi}{2} < \gamma(M_3) < \pi$ and it holds that

$$\theta_3^2,\ \theta_3^4,\ \theta_2^1,\ \theta_4^1,\ \theta_3^1,\ \psi^{2,1},\ \psi^{4,1} < \gamma(M_3).$$

Proof. From assumption, we have

$$\sin\theta_3^1 \geq \sin\theta_3^1 \sin\phi_3^1 > M_3, \quad \sin\phi_3^1 > M_3.$$

Using the definition of $\gamma(M_3)$ yields

$$\pi - \gamma < \theta_1^3 < \gamma(M_3), \quad \pi - \gamma < \phi_1^3 < \gamma(M_3).$$

We obtain, from Lemma 7.9,

$$\sin\phi_3^1 = \sin\theta_3^2 \sin\psi^{2,1} = \sin\theta_3^4 \sin\psi^{4,1} > M_3.$$

We then have

$$\sin\theta_3^2,\ \sin\psi^{2,1},\ \sin\theta_3^4,\ \sin\psi^{4,1} > M_3,$$

that is,

$$\pi - \gamma(M_3) < \theta_3^2,\ \theta_3^4,\ \psi^{2,1},\ \psi^{4,1} < \gamma(M_3).$$

Meanwhile, on $\triangle P_2P_3P_4$, using Lemma 7.4, we have

$$\theta_2^1,\ \theta_4^1 < \gamma(M_3).$$

□

Lemma 7.12. Assume that there exists M_4 with $0 < M_4 < 1$ such that

$$\sin\theta_2^1 \sin\phi_4^1 > M_4.$$

Setting $\gamma(M_4) := \pi - \sin^{-1} M_4$, we have $\frac{\pi}{2} < \gamma(M_4) < \pi$ and it holds that

$$\theta_4^2,\ \theta_4^3,\ \theta_2^1,\ \theta_3^1,\ \theta_4^1,\ \psi^{1,2},\ \psi^{1,3} < \gamma(M_4).$$

Proof. The proof is obtained by using an analogous argument with Lemma 7.11. □

7.5 Proof of Theorem 7.2 in (Type i)

7.5.1 Condition 4.5 ⇒ Condition 6.2

We set $t_1 := \sin\theta_1^4$ and $t_2 := \sin\phi_1^4$. We then have

$$\frac{H_T}{h_T} = \frac{h_1h_2h_3}{|T|_3} = \frac{6}{\sin\theta_1^4 \sin\phi_1^4}.$$

We here used the fact that $|T|_3 = \frac{1}{6} h_1 h_2 h_3 \sin \theta_1^4 \sin \phi_1^4$. By construct of the standard element (Type i), the angle θ_3^4 and θ_2^4 are respectively the maximum angle and the minimum angle of the base $\triangle P_1 P_2 P_3$ of T. We hence have $\theta_1^4 < \frac{\pi}{2}$. From Lemma 7.3, we have

$$\frac{\pi - \gamma_{11}}{2} \leq \theta_1^4 \leq \gamma_{11}, \quad \sin \theta_1^4 \geq \min \left\{ \sin \frac{\pi - \gamma_{11}}{2}, \sin \gamma_{11} \right\} =: C_1.$$

Due to Lemma 7.7, setting $\delta := \delta(\gamma_{11})$, $0 < \delta \leq \frac{\pi}{2}$ such that

$$\sin \delta = \left(\frac{\cos \gamma_{11} + 1}{\sin \frac{\gamma_{11}}{2} + 1} \right)^{1/2},$$

it holds that

$$\psi^{2,4} \geq \delta, \quad \text{or} \quad \psi^{3,4} \geq \delta.$$

Suppose that $\psi^{2,4} \geq \delta$. By Corollary 7.8 and Lemma 7.9, we have

$$\sin \phi_1^4 = \sin \theta_1^2 \sin \psi^{2,4} \geq C_0 \sin \theta_1^2.$$

By construct of the standard element (Type i), the angle θ_1^2 is not the minimum angle of $\triangle P_1 P_3 P_4$. From Lemma 7.3, we have

$$\frac{\pi - \gamma_{11}}{2} \leq \theta_1^2 \leq \gamma_{11}, \quad \sin \theta_1^2 \geq C_1.$$

We thus obtain

$$\sin \phi_1^4 \geq C_0 C_1.$$

Suppose that $\psi^{3,4} \geq \delta$. By Corollary 7.8 and Lemma 7.9, we have

$$\sin \phi_1^4 = \sin \theta_1^3 \sin \psi^{3,4} \geq C_0 \sin \theta_1^3.$$

By construct of the standard element (Type i), the angle θ_1^3 is not the minimum angle of $\triangle P_1 P_2 P_4$. From Lemma 7.3, we have

$$\frac{\pi - \gamma_{11}}{2} \leq \theta_1^3 \leq \gamma_{11}, \quad \sin \theta_1^3 \geq C_1.$$

We thus obtain

$$\sin \phi_1^4 \geq C_0 C_1.$$

In both cases

$$\psi^{2,4} \geq \delta, \quad \text{or} \quad \psi^{3,4} \geq \delta,$$

gathering the above results yield

$$\frac{H_T}{h_T} = \frac{6}{\sin \theta_1^4 \sin \phi_1^4} \leq \frac{6}{C_0 C_1^2} =: D_1 > 0,$$

that is, Condition 6.2 holds. □

7.5.2 Condition 6.2 $\Rightarrow$ Condition 4.5

From assumption, it holds that

$$\frac{H_T}{h_T} = \frac{h_1 h_2 h_3}{|T|_3} = \frac{6}{\sin\theta_1^4 \sin\phi_1^4} \le \gamma_9.$$

Remark that $\frac{6}{\gamma_9} < 1$ because $\theta_1^4 < \frac{\pi}{2}$ and $\sin\theta_1^4 \sin\phi_1^4 < 1$. Therefore, we have

$$\sin\theta_1^4 \sin\phi_1^4 \ge \frac{6}{\gamma_9} =: C_2.$$

From Lemma 7.10 with $j = 1$, setting $\gamma(C_2) := \pi - \sin^{-1} C_2$, we have $\frac{\pi}{2} < \gamma(C_2) < \pi$ and it holds that

$$\begin{aligned}&\theta_1^4,\ \theta_2^4,\ \theta_3^4 < \gamma(C_2),\\ &\theta_2^3,\ \theta_4^3,\ \theta_1^3,\ \theta_3^2,\ \theta_4^2,\ \theta_1^2,\ \psi^{2,4},\ \psi^{3,4} < \gamma(C_2).\end{aligned}$$

Furthermore, we write

$$\begin{aligned}|T|_3 &= \frac{1}{3} \times \frac{1}{2}|\overline{P_2P_3}||\overline{P_3P_4}| \sin\theta_3^1 \times h_2 \sin\phi_3^1 = \frac{1}{6} h_2 |\overline{P_2P_3}||\overline{P_3P_4}| \sin\theta_3^1 \sin\phi_3^1\\ &< \frac{1}{3} h_1 h_2 h_3 \sin\theta_3^1 \sin\phi_3^1,\end{aligned}$$

where we used the fact that $|\overline{P_3P_4}| < |\overline{P_1P_4}| + |\overline{P_1P_3}| \le 2h_3$ on $\triangle P_1P_3P_4$ and $|\overline{P_2P_3}| \le h_1$. We thus have

$$\gamma_9 \ge \frac{H_T}{h_T} > \frac{3}{\sin\theta_3^1 \sin\phi_3^1},$$

that is,

$$\sin\theta_3^1 \sin\phi_3^1 > \frac{3}{\gamma_9} =: C_3.$$

From Lemma 7.11, setting $\gamma(C_3) := \pi - \sin^{-1} C_3$, we have $\frac{\pi}{2} < \gamma(C_3) < \pi$ and it holds that

$$\theta_2^1,\ \theta_4^1,\ \theta_3^1,\ \psi^{2,1},\ \psi^{4,1} < \gamma(C_3).$$

Due to the cosine rule (7.1f) with $j = 2$, we get

$$\cos\psi^{1,3} = \sin\psi^{3,4} \sin\psi^{4,1} \cos\theta_2^4 - \cos\psi^{3,4} \cos\psi^{4,1}.$$

By constructing the standard element (Type i), the angle θ_2^4 is the minimum angle of $\triangle P_1P_2P_3$. Therefore, we have

$$\begin{aligned}&\cos\theta_2^4 \ge \frac{1}{2} \quad \text{because } \theta_2^4 \le \frac{\pi}{3},\\ &\sin\psi^{3,4} \sin\psi^{4,1} \cos\theta_2^4 > 0, \quad \text{because } \sin\psi^{3,4} \sin\psi^{4,1} > 0,\end{aligned}$$

and thus

$$\cos\psi^{1,3} > -\cos\psi^{3,4} \cos\psi^{4,1}.$$

Using $\sin \psi^{3,4} > C_2$ and $\sin \psi^{4,1} > C_3$ yields

$$\begin{aligned}\cos \psi^{1,3} &> - \cos \psi^{3,4} \cos \psi^{4,1} \\ &\geq -|\cos \psi^{3,4}||\cos \psi^{4,1}| = -\sqrt{1 - \sin^2 \psi^{3,4}}\sqrt{1 - \sin^2 \psi^{4,1}} \\ &> -\sqrt{1 - C_2^2}\sqrt{1 - C_3^2} =: C_4 > -1.\end{aligned}$$

Setting $\gamma(C_4) := \cos^{-1} C_4$, it holds that

$$\psi^{1,3} < \gamma(C_4) < \pi.$$

Due to the cosine rule (7.1d) with $j = 1$, we get

$$\cos \psi^{2,3} = \sin \psi^{3,4} \sin \psi^{4,2} \cos \theta_1^4 - \cos \psi^{3,4} \cos \psi^{4,2}.$$

By constructing the standard element (Type i), the angle θ_3^4 and θ_2^4 are respectively the maximum angle and the minimum angle of the base $\triangle P_1 P_2 P_3$ of T^s. We hence have $\theta_1^4 < \frac{\pi}{2}$. Therefore, we have

$$\begin{aligned}&\cos \theta_1^4 > 0 \quad \text{because } \theta_1^4 \leq \frac{\pi}{2}, \\ &\sin \psi^{3,4} \sin \psi^{4,2} \cos \theta_1^4 > 0, \quad \text{because } \sin \psi^{3,4} \sin \psi^{4,2} > 0,\end{aligned}$$

and thus

$$\cos \psi^{2,3} > - \cos \psi^{3,4} \cos \psi^{4,2}.$$

Using $\sin \psi^{3,4} > C_2$ and $\sin \psi^{4,2} > C_2$ yield

$$\begin{aligned}\cos \psi^{2,3} &> - \cos \psi^{3,4} \cos \psi^{4,2} \\ &\geq -|\cos \psi^{3,4}||\cos \psi^{4,2}| = -\sqrt{1 - \sin^2 \psi^{3,4}}\sqrt{1 - \sin^2 \psi^{4,2}} \\ &> -(1 - C_2^2) =: C_5 > -1.\end{aligned}$$

Setting $\gamma(C_5) := \cos^{-1} C_5$, it holds that

$$\psi^{2,3} < \gamma(C_5) < \pi.$$

We set $\gamma_{\max} := \max\{\gamma(C_3), \gamma(C_4), \gamma(C_5)\}$. We then have $0 < \gamma_{\max} < \pi$, that is, Condition 4.5 holds. □

7.6 Proof of Theorem 7.2 in (Type ii)

7.6.1 Condition 4.5 ⇒ Condition 6.2

We set $t_1 := \sin \theta_2^4$ and $t_2 := \sin \phi_1^4$. We then have

$$\frac{H_T}{h_T} = \frac{h_1 h_2 h_3}{|T|_3} = \frac{6}{\sin \theta_2^4 \sin \phi_1^4}.$$

We here used the fact that $|T|_3 = \frac{1}{6} h_1 h_2 h_3 \sin \theta_2^4 \sin \phi_1^4$. By construct of the standard element (Type ii), the angle θ_3^4 and θ_1^4 are respectively the maximum angle and the minimum angle of the base $\triangle P_1 P_2 P_3$ of T^s. We hence have $\theta_2^4 < \frac{\pi}{2}$. From Lemma 7.3, we have

$$\frac{\pi - \gamma_{11}}{2} \leq \theta_2^4 \leq \gamma_{11}, \quad \sin \theta_2^4 \geq C_1.$$

Due to Lemma 7.7, it holds that

$$\psi^{1,4} \geq \delta, \quad \text{or} \quad \psi^{3,4} \geq \delta.$$

Suppose that $\psi^{1,4} \geq \delta$. By Corollary 7.8 and Lemma 7.9, we have

$$\sin \phi_2^4 = \sin \theta_2^1 \sin \psi^{1,4} \geq C_0 \sin \theta_2^1.$$

Furthermore, it holds that

$$\sin \phi_1^4 = \frac{|\overline{P_2 P_4}| \sin \phi_2^4}{h_3}.$$

By construct of the standard element (Type ii), the angle θ_2^1 is not the minimum angle of $\triangle P_2 P_3 P_4$. From Lemma 7.3, we have

$$\frac{\pi - \gamma_{11}}{2} \leq \theta_2^1 \leq \gamma_{11}, \quad \sin \theta_2^1 \geq C_1.$$

Because $h_3 = |\overline{P_1 P_4}| < |\overline{P_2 P_4}|$ on $\triangle P_1 P_2 P_4$, we thus obtain

$$\sin \phi_1^4 = \frac{|\overline{P_2 P_4}|}{h_3} \sin \phi_2^4 > C_0 C_1.$$

Suppose that $\psi^{3,4} \geq \delta$. By Corollary 7.8 and Lemma 7.9, we have

$$\sin \phi_1^4 = \sin \theta_1^3 \sin \psi^{3,4} \geq C_0 \sin \theta_1^3.$$

By constructing the standard element (Type ii), the angle θ_1^3 is not the minimum angle of $\triangle P_1 P_2 P_4$. From Lemma 7.3, we have

$$\frac{\pi - \gamma_{11}}{2} \leq \theta_1^3 \leq \gamma_{11}, \quad \sin \theta_1^3 \geq C_1.$$

We thus obtain

$$\sin \phi_1^4 > C_0 C_1.$$

In both cases

$$\psi^{1,4} \geq \delta, \quad \text{or} \quad \psi^{3,4} \geq \delta,$$

gathering the above results yields

$$\frac{H_T}{h_T} = \frac{6}{\sin \theta_2^4 \sin \phi_1^4} \leq \frac{6}{C_0 C_1^2} = D_1 > 0,$$

that is, Condition 6.2 holds. □

7.6.2 Condition 6.2 ⇒ Condition 4.5

From assumption, it holds that

$$\frac{H_T}{h_T} = \frac{h_1 h_2 h_3}{|T|_3} = \frac{6}{\sin \theta_2^4 \sin \phi_1^4} \leq \gamma_9.$$

Remark that $\frac{6}{\gamma_9} < 1$ because $\theta_2^4 < \frac{\pi}{2}$ and $\sin\theta_2^4 \sin\phi_1^4 < 1$. Therefore, we have

$$\sin\theta_2^4 \sin\phi_1^4 \geq \frac{6}{\gamma_9} = C_2.$$

From Lemma 7.10 with $j = 2$, it holds that

$$\begin{aligned} &\theta_1^4,\ \theta_2^4,\ \theta_3^4 < \gamma(C_2), \\ &\theta_2^3,\ \theta_4^3,\ \theta_1^3,\ \theta_3^2,\ \theta_4^2,\ \theta_1^2,\ \psi^{2,4},\ \psi^{3,4} < \gamma(C_2). \end{aligned}$$

Furthermore, we write

$$\begin{aligned} |T|_3 &= \frac{1}{3} \times \frac{1}{2} |\overline{P_2P_4}||\overline{P_2P_3}| \sin\theta_2^1 \times h_3 \sin\phi_4^1 \\ &< \frac{1}{6} h_1 h_2 h_3 \sin\theta_2^1 \sin\phi_4^1, \end{aligned}$$

where we used the fact that $|\overline{P_3P_2}| = h_2$ and $|\overline{P_2P_4}| \leq h_1$. We thus have

$$\gamma_9 \geq \frac{H_{T^s}}{h_{T^s}} > \frac{6}{\sin\theta_2^1 \sin\phi_4^1},$$

that is,

$$\sin\theta_2^1 \sin\phi_4^1 > \frac{6}{\gamma_9} = C_2.$$

From Lemma 7.12, it holds that

$$\theta_2^1,\ \theta_4^1,\ \theta_3^1,\ \psi^{1,2},\ \psi^{1,3} < \gamma(C_2).$$

Due to the cosine rule (7.1e) with $j = 2$, we get

$$\cos\psi^{4,1} = \sin\psi^{1,3} \sin\psi^{3,4} \cos\theta_2^3 - \cos\psi^{1,3} \cos\psi^{3,4}.$$

By constructing the standard element (Type ii), the angle θ_2^3 is the minimum angle of $\triangle P_1P_2P_4$. Therefore, we have

$$\begin{aligned} &\cos\theta_2^3 \geq \frac{1}{2} \quad \text{because } \theta_2^3 \leq \frac{\pi}{3}, \\ &\sin\psi^{1,3} \sin\psi^{3,4} \cos\theta_2^3 > 0, \quad \text{because } \sin\psi^{1,3} \sin\psi^{3,4} > 0, \end{aligned}$$

and thus

$$\cos\psi^{4,1} > -\cos\psi^{1,3} \cos\psi^{3,4}.$$

Using $\sin\psi^{1,3} > C_2$ and $\sin\psi^{3,4} > C_2$ yield

$$\begin{aligned} \cos\psi^{4,1} &> -\cos\psi^{1,3} \cos\psi^{3,4} \\ &\geq -\sqrt{1 - \sin^2\psi^{1,3}} \sqrt{1 - \sin^2\psi^{3,4}} \\ &> -(1 - C_2^2) = C_5 > -1. \end{aligned}$$

It then holds that

$$\psi^{4,1} < \gamma(C_5) < \pi.$$

Due to the cosine rule (7.1d) with $j = 1$, we get

$$\cos\psi^{2,3} = \sin\psi^{3,4}\sin\psi^{4,2}\cos\theta_1^4 - \cos\psi^{3,4}\cos\psi^{4,2}.$$

By constructing the standard element (Type ii), the angle θ_1^4 is the minimum angle of $\triangle P_1P_2P_3$. We hence have $\theta_1^4 < \frac{\pi}{3}$. Therefore, we have

$$\begin{aligned}
&\cos\theta_1^4 \geq \frac{1}{2} \quad \text{because } \theta_1^4 \leq \frac{\pi}{3}, \\
&\sin\psi^{3,4}\sin\psi^{4,2}\cos\theta_1^4 > 0, \quad \text{because } \sin\psi^{3,4}\sin\psi^{4,2} > 0,
\end{aligned}$$

and thus

$$\cos\psi^{2,3} > -\cos\psi^{3,4}\cos\psi^{4,2}.$$

Using $\sin\psi^{3,4} > C_2$ and $\sin\psi^{4,2} > C_2$ yield

$$\begin{aligned}
\cos\psi^{2,3} &> -\cos\psi^{3,4}\cos\psi^{4,2} \\
&\geq -\sqrt{1-\sin^2\psi^{3,4}}\sqrt{1-\sin^2\psi^{4,2}} \\
&> -(1-C_2^2) = C_5 > -1.
\end{aligned}$$

It then holds that

$$\psi^{2,3} < \gamma(C_5) < \pi.$$

We set $\gamma_{\max} := \max\{\gamma(C_2), \gamma(C_5)\}$. We then have $0 < \gamma_{\max} < \pi$, that is, Condition 4.5 holds. $\square$

8 Good Elements or not for $d = 2, 3$?

In this subsection, we consider good elements on meshes. In this paper, we define 'good elements' on meshes as the existence of a positive constant $\gamma_0 > 0$ satisfying (6.1). We treat a "Right-angled triangle", "Blade" and "Dagger" for $d = 2$, and "Spire", "Spear", "Spindle", "Spike", "Splinter" and "Sliver" for $d = 3$, which are introduced in [16]. We give the quantities $h_{\max}/h_{\min}$ and H_T/h_T for those elements. The parameters $h_{\max}$ and $h_{\min}$ are defined as

$$h_{\max} := \max\{h_1, \dots, h_d\}, \quad h_{\min} := \min\{h_1, \dots, h_d\}. \tag{8.1}$$

8.1 Isotropic Mesh Elements

Recall that an isotropic mesh element has equal or nearly equal edge lengths and angles, resulting in a balanced shape. Then, the geometric condition (3.4) is satisfied. Therefore, it holds that

$$\frac{H_T}{h_T} \leq \frac{h_T^d}{|T|_d} \leq \frac{1}{\gamma_3}, \quad \frac{h_{\max}}{h_{\min}} \leq c\frac{h_T^d}{|T|_d} \leq \frac{c}{\gamma_3}.$$

In this case, elements satisfying the geometric condition (3.4) are "good."

8.2 Anisotropic mesh: two-dimensional case

Let $S \subset \mathbb{R}^2$ be a triangle. Let $0 < s \ll 1$, $s \in \mathbb{R}$ and $\varepsilon, \delta, \gamma \in \mathbb{R}$.

Example 8.1 (Right-angled triangle). Let $S \subset \mathbb{R}^2$ be the simplex with vertices $p_1 := (0,0)^\top$, $p_2 := (s,0)^\top$ and $p_3 := (0, s^\varepsilon)^\top$ with $1 < \varepsilon$. We then have $h_1 = s$ and $h_2 = s^\varepsilon$; i.e.,

$$\frac{h_{\max}}{h_{\min}} \leq s^{1-\varepsilon} \to \infty \quad \text{as } s \to 0, \quad \frac{H_S}{h_S} = 2.$$

In this case, the element S is "good."

Example 8.2 (Dagger). Let $S \subset \mathbb{R}^2$ be the simplex with vertices $p_1 := (0,0)^\top$, $p_2 := (s,0)^\top$ and $p_3 := (s^\delta, s^\varepsilon)^\top$ with $1 < \varepsilon < \delta$. We then have $h_1 = \sqrt{(s - s^\delta)^2 + s^{2\varepsilon}}$ and $h_2 = \sqrt{s^{2\delta} + s^{2\varepsilon}}$; i.e.,

$$\begin{aligned} \frac{h_{\max}}{h_{\min}} &= \frac{\sqrt{(s - s^\delta)^2 + s^{2\varepsilon}}}{\sqrt{s^{2\delta} + s^{2\varepsilon}}} \leq c s^{1-\varepsilon} \to \infty \quad \text{as } s \to 0, \\ \frac{H_S}{h_S} &= \frac{\sqrt{(s - s^\delta)^2 + s^{2\varepsilon}} \sqrt{s^{2\delta} + s^{2\varepsilon}}}{\frac{1}{2} s^{1+\varepsilon}} \leq c. \end{aligned}$$

In this case, the element S is "good."

Remark 8.3. In the above examples, $h_2 \approx \widetilde{\mathscr{H}_2}$ holds. That is, the good element $S \subset \mathbb{R}^2$ may satisfy conditions such as $h_2 \approx \widetilde{\mathscr{H}_2}$.

Example 8.4 (Blade). Let $S \subset \mathbb{R}^2$ be the simplex with vertices $p_1 := (0,0)^\top$, $p_2 := (2s,0)^\top$ and $p_3 := (s, s^\varepsilon)^\top$ with $1 < \varepsilon$. We then have $h_1 = h_2 = \sqrt{s^2 + s^{2\varepsilon}}$; i.e.,

$$\frac{h_{\max}}{h_{\min}} = 1, \quad \frac{H_S}{h_S} = \frac{s^2 + s^{2\varepsilon}}{s^{1+\varepsilon}} \to \infty \quad \text{as } s \to 0.$$

In this case, the element S is "not good."

Example 8.5 (Dagger). Let $S \subset \mathbb{R}^2$ be the simplex with vertices $p_1 := (0,0)^\top$, $p_2 := (s,0)^\top$ and $p_3 := (s^\delta, s^\varepsilon)^\top$ with $1 < \delta < \varepsilon$. We then have $h_1 = \sqrt{(s - s^\delta)^2 + s^{2\varepsilon}}$ and $h_2 = \sqrt{s^{2\delta} + s^{2\varepsilon}}$; i.e.,

$$\begin{aligned} \frac{h_{\max}}{h_{\min}} &= \frac{\sqrt{(s - s^\delta)^2 + s^{2\varepsilon}}}{\sqrt{s^{2\delta} + s^{2\varepsilon}}} \leq c s^{1-\delta} \to \infty \quad \text{as } s \to 0, \\ \frac{H_S}{h_S} &= \frac{\sqrt{(s - s^\delta)^2 + s^{2\varepsilon}} \sqrt{s^{2\delta} + s^{2\varepsilon}}}{\frac{1}{2} s^{1+\varepsilon}} \leq c s^{\delta-\varepsilon} \to \infty \quad \text{as } s \to 0. \end{aligned}$$

In this case, the element S is "not good."

Anisotropic elements in the next two examples are also "good." However, these examples differ slightly from Examples 8.1 and 8.4.

Example 8.6 (Right-angled triangle). Let $S \subset \mathbb{R}^2$ be the simplex with vertices $p_1 := (0,0)^\top$, $p_2 := (s,0)^\top$ and $p_3 := (0, \delta s)^\top$ with $\delta \ll 1$. We then have $h_1 = s$ and $h_2 = \delta s$; i.e.,

$$\frac{h_{\max}}{h_{\min}} = \frac{1}{\delta}, \quad \frac{H_S}{h_S} = 2.$$

In this case, the element S is "good." However, the factor $\frac{1}{\delta}$ is very large.

Example 8.7 (Blade). Let $S \subset \mathbb{R}^2$ be the simplex with vertices $p_1 := (0,0)^\top$, $p_2 := (2s,0)^\top$ and $p_3 := (s, \delta s)^\top$ with $\delta \ll 1$. We then have $h_1 = h_2 = s\sqrt{1 + \delta^2}$; i.e.,

$$\frac{h_{\max}}{h_{\min}} = 1, \quad \frac{H_S}{h_S} = \frac{s^2(1 + \delta^2)}{\delta s^2} \leq \frac{c}{\delta},$$

In this case, the element S is "good." However, the factor $\frac{1}{\delta}$ is very large.

8.3 Anisotropic mesh: three-dimensional case

Example 8.8. Let $T \subset \mathbb{R}^3$ be a tetrahedron. Let S be the base of T; i.e., $S = \triangle p_1 p_2 p_3$. Recall that

$$\frac{H_T}{h_T} = \frac{h_1 h_2 h_3}{|T|_3} = \frac{h_1 h_2}{\frac{1}{2} h_1 h_2 t_1} \frac{h_3}{\frac{1}{3} h_3 t_2} \leq \frac{H_S}{h_S} \frac{h_3}{\frac{1}{3}\widetilde{\mathscr{H}_3}}. \tag{8.2}$$

If the triangle S is "not good" such as in Examples 8.4 and 8.5, the quantity (8.2) may diverge. In the following, we consider the case that the triangle S is "good".

Assume that there exists a positive constant M such that $\frac{H_S}{h_S} \leq M$. For simplicity, we set $p_1 := (0,0,0)^\top$, $p_2 := (2s, 0, 0)^\top$, and $p_3 := (2s - \sqrt{4s^2 - s^{2\gamma}}, s^\gamma, 0)^\top$ with $1 < \gamma$. Then,

$$h_1 = 2s, \quad h_2 = \sqrt{\frac{4s^{2\gamma}}{2 + \sqrt{4 - s^{2\gamma - 2}}}},$$

and because $h_{\max} \approx cs$,

$$\frac{h_{\max}}{h_{\min}} \leq \frac{cs}{h_2} \leq cs^{1-\gamma} \to \infty \quad \text{as } s \to 0.$$

If we set $p_4 := (s, 0, s^\varepsilon)^\top$ with $1 < \varepsilon$, the triangle $\triangle p_1 p_2 p_4$ is the blade (Example 8.4). Then,

$$h_3 = \sqrt{s^2 + s^{2\varepsilon}}.$$

We thus have

$$\frac{H_T}{h_T} \leq c \frac{s^{2+\gamma}}{s^{1+\gamma+\varepsilon}} \leq cs^{1-\varepsilon} \to \infty \quad \text{as } s \to 0.$$

In this case, the element T is "not good."

If we set $p_4 := (s^\delta, 0, s^\varepsilon)^\top$ with $1 < \delta < \varepsilon < \gamma$, the triangle $\triangle p_1 p_2 p_4$ is the dagger (Example 8.5, Fig. 10). Then,

$$h_3 = \sqrt{s^{2\delta} + s^{2\varepsilon}}.$$

We thus have

$$\frac{H_T}{h_T} \leq c \frac{s^{1+\gamma+\delta}}{s^{1+\gamma+\varepsilon}} \leq cs^{\delta-\varepsilon} \to \infty \quad \text{as } s \to 0.$$

In this case, the element T is "not good."

If we set $p_4 := (s^\delta, 0, s^\varepsilon)^\top$ with $1 < \varepsilon < \delta < \gamma$, the triangle $\triangle p_1 p_2 p_4$ is the dagger (Example 8.2). Then,

$$h_3 = \sqrt{s^{2\delta} + s^{2\varepsilon}}.$$

We thus have

$$\frac{H_T}{h_T} \leq c \frac{s^{1+\gamma+\varepsilon}}{s^{1+\gamma+\varepsilon}} \leq c.$$

In this case, the element T is "good" and $h_3 \approx h_3 t_2 = \widetilde{\mathscr{H}_3}$ holds.

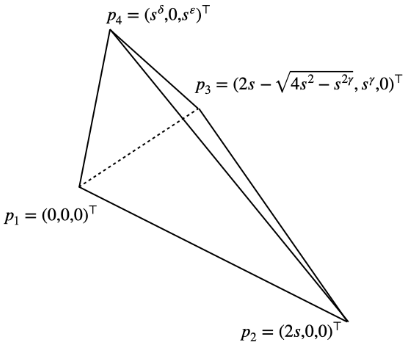

Fig. 10: Example 8.8

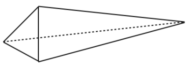

Fig. 11: Spire

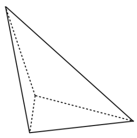

Fig. 12: Spear

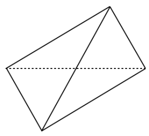

Fig. 13: Spindle

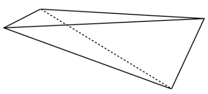

Fig. 14: Spike

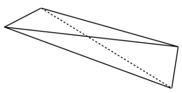

Fig. 15: Splinter

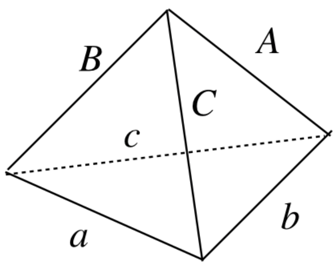

Fig. 16: R_3

Example 8.9. In [16], the spire has a cycle of three daggers among its four triangles; see Figure 11. The splinter has four daggers; see Figure 15. The spear and spike have two daggers and two blades as triangles; see Figures 12, 14. The spindle has four blades as triangles; see Figure 13.

Remark 8.10. The above examples reveal that the good element $T \subset \mathbb{R}^3$ may satisfy conditions such as $h_2 \approx \tilde{\mathscr{H}_2}$ and $h_3 \approx \tilde{\mathscr{H}_3}$.

Example 8.11. Using an element T called *Sliver*, we compare the three quantities $\frac{h_T^3}{|T|_3}$, $\frac{H_T}{h_T}$, and $\frac{R_3}{h_T}$, where the formulation of the circumradius R_3 of a tetrahedron T is as follows, e.g., see [28]. Let a, b and c be the lengths of the three edges of T and A, B, C the length of the opposite edges of a, b, c, respectively. Then,

$$R_3 = \frac{\sqrt{(aA+bB+cC)(aA+bB-cC)(aA-bB+cC)(-aA+bB+cC)}}{24|T|_3},$$

see Fig. 16.

Let $T \subset \mathbb{R}^3$ be the simplex with vertices $p_1 := (s^{\varepsilon_2}, 0, 0)^\top$, $p_2 := (-s^{\varepsilon_2}, 0, 0)^\top$, $p_3 := (0, -s, s^{\varepsilon_1})^\top$, and $p_4 := (0, s, s^{\varepsilon_1})^\top$ $(\varepsilon_1, \varepsilon_2 > 1)$, where $s := \frac{1}{N}$, $N \in \mathbb{N}$, see Fig. 17. Let L_i $(1 \le i \le 6)$ be the edges of T with $h_{\min} = L_1 \le L_2 \le \cdots \le L_6 = h_T$. Recall that $h_{\max} \approx h_T$ and

$$\frac{h_{\max}}{h_{\min}} \le c \frac{L_6}{L_1}, \quad \frac{H_T}{h_T} = \frac{L_1 L_2}{|T|_3} h_T.$$

Table 8: $h_T^3/|T|_3$, H_T/h_T and R_3/h_T $(\varepsilon_1 = 1.5,\ \varepsilon_2 = 1.0)$

N	s	L_6/L_1	$h_T^3/\|T\|_3$	H_T/h_T	R_3/h_T
32	3.1250e-02	1.4033	6.7882e+01	3.4471e+01	5.0195e-01
64	1.5625e-02	1.4087	9.6000e+01	4.8375e+01	5.0098e-01
128	7.8125e-03	1.4115	1.3576e+02	6.8147e+01	5.0049e-01

Table 9: $h_T^3/|T|_3$, H_T/h_T and R_3/h_T $(\varepsilon_1 = 1.0,\ \varepsilon_2 = 1.5)$

N	s	L_6/L_1	$h_T^3/\|T\|_3$	H_T/h_T	R_3/h_T
32	3.1250e-02	5.6569	6.7882e+01	8.5513	5.0006e-01
64	1.5625e-02	8.0000	9.6000e+01	8.5184	5.0002e-01
128	7.8125e-03	1.1314e+01	1.3576e+02	8.5018	5.0000e-01

Table 10: $h_T^3/|T|_3$, H_T/h_T and R_3/h_T $(\varepsilon_1 = 1.5,\ \varepsilon_2 = 1.5)$

N	s	L_6/L_1	$h_T^3/\|T\|_3$	H_T/h_T	R_3/h_T
32	3.1250e-02	5.6569	3.8400e+02	3.4986e+01	1.4170
64	1.5625e-02	8.0000	7.6800e+02	4.8744e+01	2.0010
128	7.8125e-03	1.1314e+01	1.5360e+03	6.8411e+01	2.8288

In Table 8, the angle between $\triangle p_1p_2p_3$ and $\triangle p_1p_2p_4$ tends to π as $s \to 0$, and the simplex T is "not good." In Table 9, the angle between $\triangle p_1p_3p_4$ and $\triangle p_2p_3p_4$ tends to 0 as $s \to 0$, the simplex T is "good." In Table 10, from the numerical result, the simplex T is "not good."

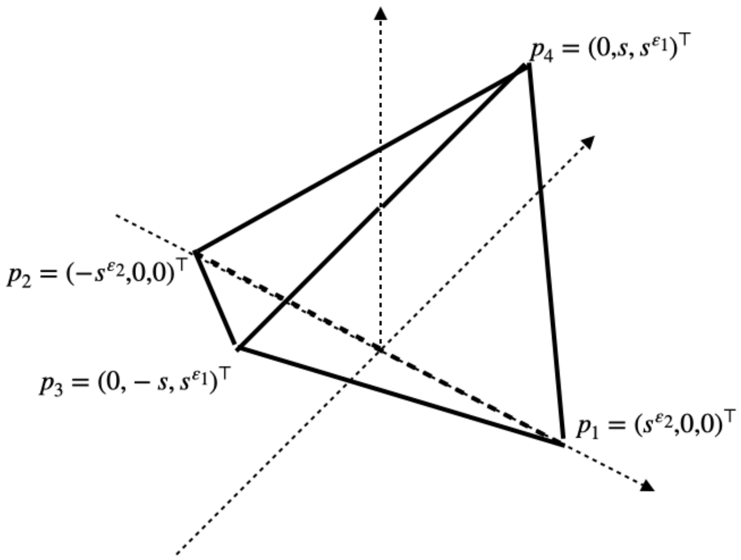

Fig. 17: Sliver

9 FE Generation

We follow the procedure described in [21, Chapter 9] and [20, Section 1.4.1 and 1.2.1]; also see [37, Section 3.5]. The definition of a *finite element* can be found in [17, p. 78] and [21, Definition 5.2].

For the reference element $\widehat{T}$ defined in Sections 5.1, let $\{\widehat{T}, \widehat{P}, \widehat{\Sigma}\}$ be a fixed reference finite element, where $\widehat{P}$ is a vector space of functions $\hat{q} : \widehat{T} \to \mathbb{R}^n$ for some positive integer n (typically $n = 1$ or $n = d$) and $\widehat{\Sigma}$ is a set of n_0 linear forms $\{\hat{\chi}_1, \ldots, \hat{\chi}_{n_0}\}$ such that

$$\widehat{P} \ni \hat{q} \mapsto (\hat{\chi}_1(\hat{q}), \ldots, \hat{\chi}_{n_0}(\hat{q}))^\top \in \mathbb{R}^{n_0}$$

is bijective; i.e., $\widehat{\Sigma}$ is a basis for $\mathcal{L}(\widehat{P}; \mathbb{R})$. Further, we denote by $\{\hat{\theta}_1, \ldots, \hat{\theta}_{n_0}\}$ in $\widehat{P}$ the local ($\mathbb{R}^n$-valued) shape functions such that

$$\hat{\chi}_i(\hat{\theta}_j) = \delta_{ij}, \quad 1 \le i, j \le n_0.$$

Let $V(\widehat{T})$ be a normed vector space of functions $\hat{\varphi} : \widehat{T} \to \mathbb{R}^n$ such that $\widehat{P} \subset V(\widehat{T})$ and the linear forms $\{\hat{\chi}_1, \ldots, \hat{\chi}_{n_0}\}$ can be extended to $V(\widehat{T})'$, i.e., there exist $\{\bar{\chi}_1, \ldots, \bar{\chi}_{n_0}\}$ and c_χ such that $\hat{\chi}_i(\hat{q}) = \bar{\chi}_i(\hat{q})$ for any $\hat{q} \in \widehat{P}$, and $|\bar{\chi}_i(\hat{v})| \le c_\chi \|\hat{v}\|_{V(\widehat{T})}$ and for $i \in \{1, \ldots, n_0\}$. We use the same symbol $\hat{\chi}_i$ instead of $\bar{\chi}_i$. The local interpolation operator $I_{\widehat{T}}$ is then defined by

$$I_{\widehat{T}} : V(\widehat{T}) \ni \hat{\varphi} \mapsto \sum_{i=1}^{n_0} \hat{\chi}_i(\hat{\varphi}) \hat{\theta}_i \in \widehat{P}. \tag{9.1}$$

It obviously holds that, for any $\hat{\varphi} \in V(\widehat{T})$,

$$\hat{\chi}_i(I_{\widehat{T}} \hat{\varphi}) = \hat{\chi}_i(\hat{\varphi}) \quad i = 1, \ldots, n_0. \tag{9.2}$$

Proposition 9.1. $\widehat{P}$ is invariant under $I_{\widehat{T}}$, that is,

$$I_{\widehat{T}} \hat{q} = \hat{q} \quad \forall \hat{q} \in \widehat{P}. \tag{9.3}$$

Proof. Let $\hat{q} = \sum_{j=1}^{n_0} \alpha_j \hat{\theta}_j$ for $\alpha_j \in \mathbb{R}$, $1 \leq j \leq n_0$. Then,

$$I_{\widehat{T}} \hat{q} = \sum_{i,j=1}^{n_0} \alpha_j \hat{\chi}_i(\hat{\theta}_j) \hat{\theta}_i = \hat{q}.$$

$\square$

Let $\Phi_{\widetilde{T}} : \widehat{T} \to \widetilde{T}$ and $\Phi_T : \widetilde{T} \to T$ be the two affine mappings defined in Section 5.2. For any $T \in \mathbb{T}_h$ with $T = \Phi(\widehat{T}) = (\Phi_T \circ \Phi_{\widetilde{T}})(\widehat{T})$, we define a Banach space $V(T)$ of $\mathbb{R}^n$-valued functions that is the counterpart of $V(\widehat{T})$ and define a linear bijection mapping by

$$\psi := \psi_{\widehat{T}} \circ \psi_{\widetilde{T}} : V(T) \ni \varphi \mapsto \hat{\varphi} := \psi(\varphi) := \varphi \circ \Phi \in V(\widehat{T}),$$

with two linear bijection mappings:

$$\begin{aligned} \psi_{\widetilde{T}} &: V(T) \ni \varphi \mapsto \tilde{\varphi} := \psi_{\widetilde{T}}(\varphi) := \varphi \circ \Phi_T \in V(\widetilde{T}), \\ \psi_{\widehat{T}} &: V(\widetilde{T}) \ni \tilde{\varphi} \mapsto \hat{\varphi} := \psi_{\widehat{T}}(\tilde{\varphi}) := \tilde{\varphi} \circ \Phi_{\widetilde{T}} \in V(\widehat{T}). \end{aligned}$$

Triples $\{\widetilde{T}, \widetilde{P}, \widetilde{\Sigma}\}$ and $\{T, P, \Sigma\}$ are defined as follows:

$$\begin{cases} \widetilde{T} = \Phi_{\widetilde{T}}(\widehat{T}); \\ \widetilde{P} = \{\psi_{\widehat{T}}^{-1}(\hat{q});\ \hat{q} \in \widehat{P}\}; \\ \widetilde{\Sigma} = \{\{\tilde{\chi}_i\}_{1 \leq i \leq n_0};\ \tilde{\chi}_i = \hat{\chi}_i(\psi_{\widehat{T}}(\tilde{q})), \forall \tilde{q} \in \widetilde{P}, \hat{\chi}_i \in \widehat{\Sigma}\}, \end{cases}$$

and

$$\begin{cases} T = \Phi_T(\widetilde{T}); \\ P = \{\psi_{\widetilde{T}}^{-1}(\tilde{q});\ \tilde{q} \in \widetilde{P}\}; \\ \Sigma = \{\{\chi_i\}_{1 \leq i \leq n_0};\ \chi_i = \tilde{\chi}_i(\psi_{\widetilde{T}}(q)), \forall q \in P, \tilde{\chi}_i \in \widetilde{\Sigma}\}. \end{cases}$$

Proposition 9.2. The triples $\{\widetilde{T}, \widetilde{P}, \widetilde{\Sigma}\}$ and $\{T, P, \Sigma\}$ are finite elements.

Proof. A proof can be obtained similarly for [21, Proposition 9.2]. $\square$

The local shape functions are $\tilde{\theta}_i = \psi_{\widehat{T}}^{-1}(\hat{\theta}_i)$ and $\theta_i = \psi_{\widetilde{T}}^{-1}(\tilde{\theta}_i)$, $1 \leq i \leq n_0$, and the associated local interpolation operators are respectively defined by

$$I_{\widetilde{T}} : V(\widetilde{T}) \ni \tilde{\varphi} \mapsto I_{\widetilde{T}} \tilde{\varphi} := \sum_{i=1}^{n_0} \tilde{\chi}_i(\tilde{\varphi}) \tilde{\theta}_i \in \widetilde{P}, \tag{9.4}$$

$$I_T : V(T) \ni \varphi \mapsto I_T \varphi := \sum_{i=1}^{n_0} \chi_i(\varphi) \theta_i \in P. \tag{9.5}$$

The following diagrams play an important role in analysing the interpolation error.

Proposition 9.3 (Commuting diagrams)**.** The diagrams

$$\begin{tikzcd} V(T) \arrow[r, "\psi_{\widetilde{T}}"] \arrow[d, "I_T"'] & V(\widetilde{T}) \arrow[r, "\psi_{\widehat{T}}"] \arrow[d, "I_{\widetilde{T}}"'] & V(\widehat{T}) \arrow[d, "I_{\widehat{T}}"] \\ P \arrow[r, "\psi_{\widetilde{T}}"'] & \widetilde{P} \arrow[r, "\psi_{\widehat{T}}"'] & \widehat{P} \end{tikzcd}$$

commute. Furthermore, $\widetilde{P}$ and P are respectively invariant under $I_{\widetilde{T}}$ and I_T.

Proof. A proof can be obtained similarly for [21, Proposition 9.3].

Let $\tilde{\varphi} \in V(\widetilde{T})$. The definition of $\{\widetilde{T}, \widetilde{P}, \widetilde{\Sigma}\}$ implies that

$$I_{\widehat{T}}(\psi_{\widehat{T}}(\tilde{\varphi})) = \sum_{i=1}^{n_0} \hat{\chi}_i(\psi_{\widehat{T}}(\tilde{\varphi}))\hat{\theta}_i = \sum_{i=1}^{n_0} \tilde{\chi}_i(\tilde{\varphi})\psi_{\widehat{T}}(\tilde{\theta}_i) = \psi_{\widehat{T}}(I_{\widetilde{T}}\tilde{\varphi}).$$

Here, we used the linearity of $\psi_{\widehat{T}}$. Therefore, the right diagram commutes.

Let $\tilde{q} \in \widetilde{P}$. Because $\psi_{\widehat{T}}(\tilde{q}) \in \widehat{P}$ and $\widehat{P}$ is invariant under $I_{\widehat{T}}$,

$$I_{\widetilde{T}}(\tilde{q}) = \psi_{\widehat{T}}^{-1}(I_{\widehat{T}}(\psi_{\widehat{T}}(\tilde{q}))) = \psi_{\widehat{T}}^{-1}(\psi_{\widehat{T}}(\tilde{q})) = \tilde{q}.$$

Another diagram can be proved in the same way. $\square$

Example 9.4. Let $\{\widehat{T}, \widehat{P}, \widehat{\Sigma}\}$ be a finite element.

1. For the Lagrange finite element of degree k, we set $V(\widehat{T}) := \mathcal{C}^0(\widehat{T})$.
2. For the Hermite finite element, we set $V(\widehat{T}) := \mathcal{C}^1(\widehat{T})$.
3. For the Crouzeix–Raviart finite element with $k = 1$, we set $V(\widehat{T}) := W^{1,1}(\widehat{T})$.

10 New Scaling Argument: Part 1

This section gives estimates related to a scaling argument corresponding to [20, Lemma 1.101].

10.1 Preliminalies

10.1.1 Additional New Condition

The following condition is used for obtaining optimal interpolation error estimates.

Condition 10.1. In anisotropic interpolation error analysis, we impose the following geometric condition for the simplex T:

1. If $d = 2$, there are no additional conditions;
2. If $d = 3$, there must exist a positive constant M independent of h_T such that $|s_{22}| \le M\frac{h_2 t_1}{h_3}$. Note that if $s_{22} \neq 0$, this condition means that the order of h_3 with respect to h_T coincides with the order of h_2, and if $s_{22} = 0$, the order of h_3 may be different from that of h_2.

Recall that

$$\begin{aligned} &|s| \le 1, \quad h_2 \le h_1 \quad \text{if } d = 2, \\ &|s_1| \le 1,\ |s_{21}| \le 1, \quad h_2 \le h_3 \le h_1 \quad \text{if } d = 3. \end{aligned}$$

When $d = 3$, if Condition 10.1 is imposed, there exists a positive constant M independent of h_T such that $|s_{22}| \le M\frac{h_2 t_1}{h_3}$. We thus have, if $d = 2$,

$$h_1|[\widetilde{A}]_{j1}| \le \widetilde{\mathscr{H}_j}, \quad h_2|[\widetilde{A}]_{j2}| \le \widetilde{\mathscr{H}_j}, \quad j = 1, 2,$$

and, if $d = 3$, for $\widetilde{A} \in \{\widetilde{A}_1, \widetilde{A}_2\}$ and $j = 1, 2, 3$,

$$h_1|[\widetilde{A}]_{j1}| \le \widetilde{\mathscr{H}_j}, \quad h_2|[\widetilde{A}]_{j2}| \le \widetilde{\mathscr{H}_j}, \quad h_3|[\widetilde{A}]_{j3}| \le \max\{1, M\}\widetilde{\mathscr{H}_j}, \quad j = 1, 2, 3.$$

10.1.2 Calculations 1

We use the following calculations in (10.2). Recall that $\tilde{x} = A_{\widetilde{T}}\hat{x}$ with $A_{\widetilde{T}} = \widetilde{A}\widehat{A}$ and $x = A_T\tilde{x} + b_T$. For any multi-indices β and γ, we have

$$
\begin{aligned}
\partial_{\hat{x}}^{\beta+\gamma} &= \frac{\partial^{|\beta|+|\gamma|}}{\partial \hat{x}_1^{\beta_1}\cdots\partial \hat{x}_d^{\beta_d}\partial \hat{x}_1^{\gamma_1}\cdots\partial \hat{x}_d^{\gamma_d}} \\
&= \underbrace{\sum_{i_1^{(1)},i_1^{(0,1)}=1}^{d} h_1[\widetilde{A}]_{i_1^{(1)}1}[A_T]_{i_1^{(0,1)}i_1^{(1)}}\cdots \sum_{i_{\beta_1}^{(1)},i_{\beta_1}^{(0,1)}=1}^{d} h_1[\widetilde{A}]_{i_{\beta_1}^{(1)}1}[A_T]_{i_{\beta_1}^{(0,1)}i_{\beta_1}^{(1)}}\cdots}_{\beta_1\text{times}} \\
&\quad \underbrace{\sum_{i_1^{(d)},i_1^{(0,d)}=1}^{d} h_d[\widetilde{A}]_{i_1^{(d)}d}[A_T]_{i_1^{(0,d)}i_1^{(d)}}\cdots \sum_{i_{\beta_d}^{(d)},i_{\beta_d}^{(0,d)}=1}^{d} h_d[\widetilde{A}]_{i_{\beta_d}^{(d)}d}[A_T]_{i_{\beta_d}^{(0,d)}i_{\beta_d}^{(d)}}}_{\beta_d\text{times}} \\
&\quad \underbrace{\sum_{j_1^{(1)},j_1^{(0,1)}=1}^{d} h_1[\widetilde{A}]_{j_1^{(1)}1}[A_T]_{j_1^{(0,1)}j_1^{(1)}}\cdots \sum_{j_{\gamma_1}^{(1)},j_{\gamma_1}^{(0,1)}=1}^{d} h_1[\widetilde{A}]_{j_{\gamma_1}^{(1)}1}[A_T]_{j_{\gamma_1}^{(0,1)}j_{\gamma_1}^{(1)}}\cdots}_{\gamma_1\text{times}} \\
&\quad \underbrace{\sum_{j_1^{(d)},j_1^{(0,d)}=1}^{d} h_d[\widetilde{A}]_{j_1^{(d)}d}[A_T]_{j_1^{(0,d)}j_1^{(d)}}\cdots \sum_{j_{\gamma_d}^{(d)},j_{\gamma_d}^{(0,d)}=1}^{d} h_d[\widetilde{A}]_{j_{\gamma_d}^{(d)}d}[A_T]_{j_{\gamma_d}^{(0,d)}j_{\gamma_d}^{(d)}}}_{\gamma_d\text{times}} \\
&\quad \underbrace{\frac{\partial^{\beta_1}}{\partial x_{i_1^{(0,1)}}\cdots\partial x_{i_{\beta_1}^{(0,1)}}}}_{\beta_1\text{times}}\cdots\underbrace{\frac{\partial^{\beta_d}}{\partial x_{i_1^{(0,d)}}\cdots\partial x_{i_{\beta_d}^{(0,d)}}}}_{\beta_d\text{times}}\underbrace{\frac{\partial^{\gamma_1}}{\partial x_{j_1^{(0,1)}}\cdots\partial x_{j_{\gamma_1}^{(0,1)}}}}_{\gamma_1\text{times}}\cdots\underbrace{\frac{\partial^{\gamma_d}}{\partial x_{j_1^{(0,d)}}\cdots\partial x_{j_{\gamma_d}^{(0,d)}}}}_{\gamma_d\text{times}}.
\end{aligned}
$$

Let $\hat{\varphi} \in \mathcal{C}^2(\widehat{T})$ with $\tilde{\varphi} = \hat{\varphi}\circ\Phi_{\widetilde{T}}^{-1}$ and $\varphi = \tilde{\varphi}\circ\Phi_T^{-1}$. Then, for $1 \le i \le d$,

$$
\begin{aligned}
\left|\frac{\partial\hat{\varphi}}{\partial\hat{x}_i}\right| &= \left|\sum_{i_1^{(1)}=1}^{d}\sum_{i_1^{(0,1)}=1}^{d} h_i[\widetilde{A}]_{i_1^{(1)}i}[A_T]_{i_1^{(0,1)}i_1^{(1)}}\frac{\partial\varphi}{\partial x_{i_1^{(0,1)}}}\right| \\
&= h_i\left|\sum_{i_1^{(1)}=1}^{d}\sum_{i_1^{(0,1)}=1}^{d}[A_T]_{i_1^{(0,1)}i_1^{(1)}}(\tilde{r}_i)_{i_1^{(1)}}\frac{\partial\varphi}{\partial x_{i_1^{(0,1)}}}\right| = h_i\left|\frac{\partial\varphi}{\partial r_i}\right| \\
&\le h_i\|\widetilde{A}\|_{\max}\|A_T\|_{\max}\sum_{i_1^{(1)}=1}^{d}\sum_{i_1^{(0,1)}=1}^{d}\left|\frac{\partial\varphi}{\partial x_{i_1^{(0,1)}}}\right|,
\end{aligned}
$$

and for $1 \le i, j \le d$,

$$
\begin{aligned}
\left|\frac{\partial^2\hat{\varphi}}{\partial\hat{x}_i\partial\hat{x}_j}\right| = \Bigg|&\sum_{i_1^{(1)},j_1^{(1)}=1}^{d}\sum_{i_1^{(0,1)},j_1^{(0,1)}=1}^{d} h_ih_j[\widetilde{A}]_{i_1^{(1)}i}[\widetilde{A}]_{j_1^{(1)}j} \\
&[A_T]_{i_1^{(0,1)}i_1^{(1)}}[A_T]_{j_1^{(0,1)}j_1^{(1)}}\frac{\partial^2\varphi}{\partial x_{i_1^{(0,1)}}\partial x_{j_1^{(0,1)}}}\Bigg| = h_ih_j\left|\frac{\partial^2\varphi}{\partial r_i\partial r_j}\right|
\end{aligned}
$$

$$\le h_i h_j \sum_{j_1^{(1)}=1}^{d} |[\widetilde{A}]_{j_1^{(1)} j}| \left| \sum_{j_1^{(0,1)}=1}^{d} [A_T]_{j_1^{(0,1)} j_1^{(1)}} \frac{\partial^2 \varphi}{\partial r_i \partial x_{j_1^{(0,1)}}} \right|$$

$$\le h_i h_j \|\widetilde{A}\|_{\max} \|A_T\|_{\max} \sum_{j_1^{(0,1)}=1}^{d} \left| \frac{\partial^2 \varphi}{\partial r_i \partial x_{j_1^{(0,1)}}} \right|$$

$$\le h_i h_j \|\widetilde{A}\|_{\max}^2 \|A_T\|_{\max}^2 \sum_{i_1^{(0,1)}, j_1^{(0,1)}=1}^{d} \left| \frac{\partial^2 \varphi}{\partial x_{i_1^{(0,1)}} \partial x_{j_1^{(0,1)}}} \right|.$$

10.1.3 Calculations 2

We use the following calculations in (10.3). Recall that $\tilde{x} = A_{\widetilde{T}} \hat{x}$ with $A_{\widetilde{T}} = \widetilde{A}\widehat{A}$. For any multi-indices β and γ, we have

$$\begin{aligned}
\partial_{\hat{x}}^{\beta+\gamma} &= \frac{\partial^{|\beta|+|\gamma|}}{\partial \hat{x}_1^{\beta_1} \cdots \partial \hat{x}_d^{\beta_d} \partial \hat{x}_1^{\gamma_1} \cdots \partial \hat{x}_d^{\gamma_d}} \\
= &\underbrace{\sum_{i_1^{(1)}=1}^{d} h_1 [\widetilde{A}]_{i_1^{(1)} 1} \cdots \sum_{i_{\beta_1}^{(1)}=1}^{d} h_1 [\widetilde{A}]_{i_{\beta_1}^{(1)} 1}}_{\beta_1 \text{times}} \cdots \underbrace{\sum_{i_1^{(d)}=1}^{d} h_d [\widetilde{A}]_{i_1^{(d)} d} \cdots \sum_{i_{\beta_d}^{(d)}=1}^{d} h_d [\widetilde{A}]_{i_{\beta_d}^{(d)} d}}_{\beta_d \text{times}} \\
&\underbrace{\sum_{j_1^{(1)}=1}^{d} h_1 [\widetilde{A}]_{j_1^{(1)} 1} \cdots \sum_{j_{\gamma_1}^{(1)}=1}^{d} h_1 [\widetilde{A}]_{j_{\gamma_1}^{(1)} 1}}_{\gamma_1 \text{times}} \cdots \underbrace{\sum_{j_1^{(d)}=1}^{d} h_d [\widetilde{A}]_{j_1^{(d)} d} \cdots \sum_{j_{\gamma_d}^{(d)}=1}^{d} h_d [\widetilde{A}]_{j_{\gamma_d}^{(d)} d}}_{\gamma_d \text{times}} \\
&\underbrace{\frac{\partial^{\beta_1}}{\partial \tilde{x}_{i_1^{(1)}} \cdots \partial \tilde{x}_{i_{\beta_1}^{(1)}}}}_{\beta_1 \text{times}} \cdots \underbrace{\frac{\partial^{\beta_d}}{\partial \tilde{x}_{i_1^{(d)}} \cdots \partial \tilde{x}_{i_{\beta_d}^{(d)}}}}_{\beta_d \text{times}} \underbrace{\frac{\partial^{\gamma_1}}{\partial \tilde{x}_{j_1^{(1)}} \cdots \partial \tilde{x}_{j_{\gamma_1}^{(1)}}}}_{\gamma_1 \text{times}} \cdots \underbrace{\frac{\partial^{\gamma_d}}{\partial \tilde{x}_{j_1^{(d)}} \cdots \partial \tilde{x}_{j_{\gamma_d}^{(d)}}}}_{\gamma_d \text{times}}.
\end{aligned}$$

Let $\hat{\varphi} \in \mathcal{C}^2(\widehat{T})$ with $\tilde{\varphi} = \hat{\varphi} \circ \Phi_{\widetilde{T}}^{-1}$. Then, for $1 \le i \le d$,

$$\left| \frac{\partial \hat{\varphi}}{\partial \hat{x}_i} \right| \le \sum_{i_1^{(1)}=1}^{d} h_i \left| [\widetilde{A}]_{i_1^{(1)} i} \right| \left| \frac{\partial \tilde{\varphi}}{\partial \tilde{x}_{i_1^{(1)}}} \right| \le \begin{cases} h_i \|\widetilde{A}\|_{\max} \sum_{i_1^{(1)}=1}^{d} \left| \frac{\partial \tilde{\varphi}}{\partial \tilde{x}_{i_1^{(1)}}} \right| & \text{or,} \\ c \sum_{i_1^{(1)}=1}^{d} \widetilde{\mathscr{H}}_{i_1^{(1)}} \left| \frac{\partial \tilde{\varphi}}{\partial \tilde{x}_{i_1^{(1)}}} \right|, \end{cases}$$

and for $1 \le i, j \le d$,

$$\left| \frac{\partial^2 \hat{\varphi}}{\partial \hat{x}_i \partial \hat{x}_j} \right| = \left| \sum_{i_1^{(1)}, j_1^{(1)}=1}^{d} h_i h_j [\widetilde{A}]_{i_1^{(1)} i} [\widetilde{A}]_{j_1^{(1)} j} \frac{\partial^2 \tilde{\varphi}}{\partial \tilde{x}_{i_1^{(1)}} \partial \tilde{x}_{j_1^{(1)}}} \right|$$

$$\le \begin{cases} h_i h_j \|\widetilde{A}\|_{\max}^2 \sum_{i_1^{(1)}, j_1^{(1)}=1}^{d} \left| \frac{\partial^2 \tilde{\varphi}}{\partial \tilde{x}_{i_1^{(1)}} \partial \tilde{x}_{j_1^{(1)}}} \right| & \text{or,} \\ h_j \sum_{j_1^{(1)}=1}^{d} |[\widetilde{A}]_{j_1^{(1)} j}| \left| \sum_{i_1^{(1)}=1}^{d} h_i [\widetilde{A}]_{i_1^{(1)} i} \frac{\partial^2 \tilde{\varphi}}{\partial \tilde{x}_{i_1^{(1)}} \partial \tilde{x}_{j_1^{(1)}}} \right| & \\ \quad \le c h_j \|\widetilde{A}\|_{\max} \sum_{j_1^{(1)}=1}^{d} \sum_{i_1^{(1)}=1}^{d} \widetilde{\mathscr{H}}_{i_1^{(1)}} \left| \frac{\partial^2 \tilde{\varphi}}{\partial \tilde{x}_{i_1^{(1)}} \partial \tilde{x}_{j_1^{(1)}}} \right| & \text{or,} \\ c \sum_{i_1^{(1)}=1}^{d} \sum_{j_1^{(1)}=1}^{d} \widetilde{\mathscr{H}}_{i_1^{(1)}} \widetilde{\mathscr{H}}_{j_1^{(1)}} \left| \frac{\partial^2 \tilde{\varphi}}{\partial \tilde{x}_{i_1^{(1)}} \partial \tilde{x}_{j_1^{(1)}}} \right|. & \end{cases}$$

10.1.4 Calculations 3

We use the following calculations in (10.1). Recall that $\hat{x} = A_{\widetilde{T}}^{-1}\tilde{x}$ with $A_{\widetilde{T}} = \widetilde{A}\widehat{A}$. For any multi-indices β, we have

$$\partial_{\tilde{x}}^{\beta} = \frac{\partial^{|\beta|}}{\partial \tilde{x}_1^{\beta_1} \cdots \partial \tilde{x}_d^{\beta_d}}$$

$$= \underbrace{\sum_{i_1^{(1)}=1}^{d} h_{i_1^{(1)}}^{-1} [\widetilde{A}^{-1}]_{i_1^{(1)}1} \cdots \sum_{i_{\beta_1}^{(1)}=1}^{d} h_{i_{\beta_1}^{(1)}}^{-1} [\widetilde{A}^{-1}]_{i_{\beta_1}^{(1)}1}}_{\beta_1 \text{times}} \cdots \underbrace{\sum_{i_1^{(d)}=1}^{d} h_{i_1^{(d)}}^{-1} [\widetilde{A}^{-1}]_{i_1^{(d)}d} \cdots \sum_{i_{\beta_d}^{(d)}=1}^{d} h_{i_{\beta_d}^{(d)}}^{-1} [\widetilde{A}^{-1}]_{i_{\beta_d}^{(d)}d}}_{\beta_d \text{times}}$$

$$\underbrace{\frac{\partial^{\beta_1}}{\partial \hat{x}_{i_1^{1)}} \cdots \partial \hat{x}_{i_{\beta_1}^{(1)}}}}_{\beta_1 \text{times}} \cdots \underbrace{\frac{\partial^{\beta_d}}{\partial \hat{x}_{i_1^{(d)}} \cdots \partial \hat{x}_{i_{\beta_d}^{(d)}}}}_{\beta_d \text{times}}.$$

Let $\tilde{\varphi} \in \mathcal{C}^2(\widetilde{T})$ with $\hat{\varphi} = \tilde{\varphi} \circ \Phi_{\widetilde{T}}$. Then, for $1 \leq i \leq d$,

$$\left| \frac{\partial \tilde{\varphi}}{\partial \tilde{x}_i} \right| \leq \sum_{i_1^{(1)}=1}^{d} h_{i_1^{(1)}}^{-1} \left| [\widetilde{A}^{-1}]_{i_1^{(1)}i} \right| \left| \frac{\partial \hat{\varphi}}{\partial \hat{x}_{i_1^{(1)}}} \right| \leq \|\widetilde{A}^{-1}\|_{\max} \sum_{i_1^{(1)}=1}^{d} h_{i_1^{(1)}}^{-1} \left| \frac{\partial \hat{\varphi}}{\partial \hat{x}_{i_1^{(1)}}} \right|,$$

and for $1 \leq i, j \leq d$,

$$\left| \frac{\partial^2 \tilde{\varphi}}{\partial \tilde{x}_i \partial \tilde{x}_j} \right| = \left| \sum_{i_1^{(1)}, j_1^{(1)}=1}^{d} h_{i_1^{(1)}}^{-1} h_{j_1^{(1)}}^{-1} [\widetilde{A}^{-1}]_{i_1^{(1)}i} [\widetilde{A}^{-1}]_{j_1^{(1)}j} \frac{\partial^2 \hat{\varphi}}{\partial \hat{x}_{i_1^{(1)}} \partial \hat{x}_{j_1^{(1)}}} \right|$$
$$\leq \|\widetilde{A}^{-1}\|_{\max}^2 \sum_{i_1^{(1)}, j_1^{(1)}=1}^{d} h_{i_1^{(1)}}^{-1} h_{j_1^{(1)}}^{-1} \left| \frac{\partial^2 \hat{\varphi}}{\partial \hat{x}_{i_1^{(1)}} \partial \hat{x}_{j_1^{(1)}}} \right|.$$

10.2 Main Results

Lemma 10.2. Let $m, \ell \in \mathbb{N}_0$ with $\ell \geq m$. Let $\beta := (\beta_1, \ldots, \beta_d) \in \mathbb{N}_0^d$ and $\gamma := (\gamma_1, \ldots, \gamma_d) \in \mathbb{N}_0^d$ be multi-indices with $|\beta| = m$ and $|\gamma| = \ell - m$. Then, for any $\hat{\varphi} \in W^{m,p}(\widehat{T})$ with $\tilde{\varphi} = \hat{\varphi} \circ \Phi_{\widetilde{T}}^{-1}$, it holds that

$$|\tilde{\varphi}|_{W^{m,p}(\widetilde{T})} \leq c |\det(A_{\widetilde{T}})|^{\frac{1}{p}} \|\widetilde{A}^{-1}\|_2^m \left(\sum_{|\beta|=m} (h^{-\beta})^p \|\partial_{\hat{x}}^{\beta} \hat{\varphi}\|_{L^p(\widehat{T})}^p \right)^{\frac{1}{p}} \quad \text{if } p \in [1, \infty), \tag{10.1a}$$

$$|\tilde{\varphi}|_{W^{m,\infty}(\widetilde{T})} \leq c \|\widetilde{A}^{-1}\|_2^m \max_{|\beta|=m} \left(h^{-\beta} \|\partial_{\hat{x}}^{\beta} \hat{\varphi}\|_{L^{\infty}(\widehat{T})} \right) \quad \text{if } p = \infty. \tag{10.1b}$$

Let $p \in [0, \infty]$. Furthermore, for any $\hat{\varphi} \in W^{\ell,p}(\widehat{T})$ with $\tilde{\varphi} = \hat{\varphi} \circ \Phi_{\widetilde{T}}^{-1}$ and $\varphi = \tilde{\varphi} \circ \Phi_{T}^{-1}$, it holds that

$$\|\partial_{\hat{x}}^{\beta} \partial_{\hat{x}}^{\gamma} \hat{\varphi}\|_{L^p(\widehat{T})} \leq c |\det(A_{\widetilde{T}})|^{-\frac{1}{p}} \|\widetilde{A}\|_2^m h^{\beta} \sum_{|\epsilon|=|\gamma|} h^{\varepsilon} |\partial_r^{\epsilon} \varphi|_{W^{m,p}(T)}. \tag{10.2}$$

In particular, if Condition 10.1 is imposed, then for any $\hat{\varphi} \in W^{\ell,p}(\widehat{T})$ with $\tilde{\varphi} = \hat{\varphi} \circ \Phi_{\widetilde{T}}^{-1}$, it holds that

$$\|\partial_{\hat{x}}^{\beta} \partial_{\hat{x}}^{\gamma} \hat{\varphi}\|_{L^p(\widehat{T})} \leq c |\det(A_{\widetilde{T}})|^{-\frac{1}{p}} \|\widetilde{A}\|_2^m h^{\beta} \sum_{|\epsilon|=|\gamma|} \widetilde{\mathscr{H}}^{\varepsilon} |\partial_{\tilde{x}}^{\varepsilon} \tilde{\varphi}|_{W^{m,p}(\widetilde{T})}. \tag{10.3}$$

Here, for $p = \infty$ and any positive real x, $x^{-\frac{1}{p}} = 1$.

Proof. We divide the proof into three parts.

Proof of (10.1)**.** Let $p \in [1, \infty)$. Because the space $\mathcal{C}^m(\widehat{T})$ is dense in the space $W^{m,p}(\widehat{T})$, we show (10.1) for $\hat{\varphi} \in \mathcal{C}^m(\widehat{T})$ with $\tilde{\varphi} = \hat{\varphi} \circ \Phi_{\widetilde{T}}^{-1}$. Through the calculation (Section 10.1.4) and (1.1), we have for any multi-index γ with $|\gamma| = m$,

$$|\partial_{\tilde{x}}^{\gamma} \tilde{\varphi}| \leq c \|\widetilde{A}^{-1}\|_2^m \sum_{|\beta|=m} h^{-\beta} |\partial_{\hat{x}}^{\beta} \hat{\varphi}|.$$

Through a change in a variable, we obtain

$$|\tilde{\varphi}|^p_{W^{m,p}(\widetilde{T})} = \sum_{|\gamma|=m} \|\partial_{\tilde{x}}^{\gamma} \tilde{\varphi}\|^p_{L^p(\widetilde{T})} \leq c |\det(A_{\widetilde{T}})| \|\widetilde{A}^{-1}\|_2^{mp} \sum_{|\beta|=m} (h^{-\beta})^p \|\partial_{\hat{x}}^{\beta} \hat{\varphi}\|^p_{L^p(\widehat{T})},$$

which leads to the inequality (10.1a). We consider the case that $p = \infty$. A function $\hat{\varphi} \in W^{m,\infty}(\widehat{T})$ belongs to the space $W^{m,p}(\widehat{T})$ for any $p \in [1, \infty)$. It therefore holds that $\tilde{\varphi} \in W^{m,p}(\widetilde{T})$ for any $p \in [1, \infty)$ and, from (1.5),

$$\begin{aligned}
\|\partial_{\tilde{x}}^{\gamma} \tilde{\varphi}\|_{L^p(\widetilde{T})} &\leq |\tilde{\varphi}|_{W^{|\gamma|,p}(\widetilde{T})} \\
&\leq c |\det(A_{\widetilde{T}})|^{\frac{1}{p}} \|\widetilde{A}^{-1}\|_2^m \left(\sum_{|\beta|=|\gamma|} (h^{-\beta})^p \|\partial_{\hat{x}}^{\beta} \hat{\varphi}\|^p_{L^p(\widehat{T})} \right)^{\frac{1}{p}} \\
&\leq c \left(\sup_{1 \leq p} |\det(A_{\widetilde{T}})|^{\frac{1}{p}} \right) \|\widetilde{A}^{-1}\|_2^m \sum_{|\beta|=|\gamma|} h^{-\beta} \|\partial_{\hat{x}}^{\beta} \hat{\varphi}\|_{L^p(\widehat{T})} \\
&\leq c \left(\sup_{1 \leq p} |\det(A_{\widetilde{T}})|^{\frac{1}{p}} \right) \|\widetilde{A}^{-1}\|_2^m \sum_{|\beta|=|\gamma|} h^{-\beta} \|\partial_{\hat{x}}^{\beta} \hat{\varphi}\|_{L^\infty(\widehat{T})} < +\infty, \qquad (10.4)
\end{aligned}$$

for multi-index $\gamma \in \mathbb{N}_0^d$ with $|\gamma| \leq m$. This implies that the function $\partial_{\tilde{x}}^{\gamma} \tilde{\varphi}$ is in the space $L^\infty(\widetilde{T})$ for each $|\gamma| \leq m$. We therefore have $\tilde{\varphi} \in W^{m,\infty}(\widetilde{T})$. By passing to the limit $p \to \infty$ in (10.4) and because $\lim_{p \to \infty} \|\cdot\|_{L^p(\widetilde{T})} = \|\cdot\|_{L^\infty(\widetilde{T})}$, we have

$$|\tilde{\varphi}|_{W^{m,\infty}(\widetilde{T})} \leq c \|\widetilde{A}^{-1}\|_2^m \max_{|\beta|=m} \left(h^{-\beta} \|\partial_{\hat{x}}^{\beta} \hat{\varphi}\|_{L^\infty(\widehat{T})} \right),$$

which is (10.1b).

Proof of (10.3)**.** Let $\varepsilon = (\varepsilon_1, \ldots, \varepsilon_d) \in \mathbb{N}_0^d$ and $\delta = (\delta_1, \ldots, \delta_d) \in \mathbb{N}_0^d$ be multi-indies with $|\varepsilon| = |\gamma|$ and $|\delta| = |\beta|$. Let $p \in [1, \infty)$. Because the space $\mathcal{C}^\ell(\widehat{T})$ is dense in the space $W^{\ell,p}(\widehat{T})$, we show (10.3) for $\hat{\varphi} \in \mathcal{C}^\ell(\widehat{T})$ with $\tilde{\varphi} = \hat{\varphi} \circ \Phi_{\widetilde{T}}^{-1}$. Through a simple calculation, we have

$$\begin{aligned}
&|\partial_{\hat{x}}^{\beta+\gamma} \hat{\varphi}| = \left| \frac{\partial^\ell \hat{\varphi}}{\partial \hat{x}_1^{\beta_1} \cdots \partial \hat{x}_d^{\beta_d} \partial \hat{x}_1^{\gamma_1} \cdots \partial \hat{x}_d^{\gamma_d}} \right| \\
&\leq c h^{\beta} \|\widetilde{A}\|_{\max}^{|\beta|} \underbrace{\sum_{i_1^{(1)}=1}^{d} \cdots \sum_{i_{\beta_1}^{(1)}=1}^{d}}_{\beta_1 \text{times}} \cdots \underbrace{\sum_{i_1^{(d)}=1}^{d} \cdots \sum_{i_{\beta_d}^{(d)}=1}^{d}}_{\beta_d \text{times}} \underbrace{\sum_{j_1^{(1)}=1}^{d} \cdots \sum_{j_{\gamma_1}^{(1)}=1}^{d}}_{\gamma_1 \text{times}} \cdots \underbrace{\sum_{j_1^{(d)}=1}^{d} \cdots \sum_{j_{\gamma_d}^{(d)}=1}^{d}}_{\gamma_d \text{times}} \\
&\underbrace{\widetilde{\mathscr{H}}_{j_1^{(1)}} \cdots \widetilde{\mathscr{H}}_{j_{\varepsilon_1}^{(1)}}}_{\gamma_1 \text{times}} \cdots \underbrace{\widetilde{\mathscr{H}}_{j_1^{(d)}} \cdots \widetilde{\mathscr{H}}_{j_{\varepsilon_d}^{(d)}}}_{\gamma_d \text{times}} \\
&\left| \underbrace{\frac{\partial^{\beta_1}}{\partial \tilde{x}_{i_1^{(1)}} \cdots \partial \tilde{x}_{i_{\beta_1}^{(1)}}}}_{\beta_1 \text{times}} \cdots \underbrace{\frac{\partial^{\beta_d}}{\partial \tilde{x}_{i_1^{(d)}} \cdots \partial \tilde{x}_{i_{\beta_d}^{(d)}}}}_{\beta_d \text{times}} \underbrace{\frac{\partial^{\gamma_1}}{\partial \tilde{x}_{j_1^{(1)}} \cdots \partial \tilde{x}_{j_{\gamma_1}^{(1)}}}}_{\gamma_1 \text{times}} \cdots \underbrace{\frac{\partial^{\gamma_d}}{\partial \tilde{x}_{j_1^{(d)}} \cdots \partial \tilde{x}_{j_{\gamma_d}^{(d)}}}}_{\gamma_d \text{times}} \tilde{\varphi} \right|
\end{aligned}$$

$$\leq ch^{\beta} \|\widetilde{A}\|_{\max}^{|\beta|} \sum_{|\delta|=|\beta|} \sum_{|\varepsilon|=|\gamma|} \widetilde{\mathscr{H}^{\varepsilon}} |\partial_{\tilde{x}}^{\delta} \partial_{\tilde{x}}^{\varepsilon} \tilde{\varphi}|.$$

We then have, using (1.1),

$$\begin{aligned}\int_{\widehat{T}} |\partial_{\hat{x}}^{\beta} \partial_{\hat{x}}^{\gamma} \hat{\varphi}|^p d\hat{x} &\leq c \|\widetilde{A}\|_2^{mp} h^{\beta p} \sum_{|\delta|=|\beta|} \sum_{|\epsilon|=|\gamma|} \widetilde{\mathscr{H}^{\varepsilon p}} \int_{\widehat{T}} |\partial_{\tilde{x}}^{\delta} \partial_{\tilde{x}}^{\varepsilon} \tilde{\varphi}|^p d\hat{x} \\ &= c |\det(A_{\widetilde{T}})|^{-1} \|\widetilde{A}\|_2^{mp} h^{\beta p} \sum_{|\delta|=|\beta|} \sum_{|\epsilon|=|\gamma|} \widetilde{\mathscr{H}^{\varepsilon p}} \int_{\widetilde{T}} |\partial_{\tilde{x}}^{\delta} \partial_{\tilde{x}}^{\varepsilon} \tilde{\varphi}|^p d\tilde{x}.\end{aligned}$$

Therefore, using (1.5), we have

$$\|\partial_{\hat{x}}^{\beta} \partial_{\hat{x}}^{\gamma} \hat{\varphi}\|_{L^p(\widehat{T})} \leq c |\det(A_{\widetilde{T}}))|^{-\frac{1}{p}} \|\widetilde{A}\|_2^m h^{\beta} \sum_{|\epsilon|=|\gamma|} \widetilde{\mathscr{H}^{\varepsilon}} |\partial_{\tilde{x}}^{\epsilon} \tilde{\varphi}|_{W^{m,p}(\widetilde{T})},$$

which concludes (10.3). We consider the case that $p = \infty$. A function $\varphi \in W^{\ell,\infty}(T)$ belongs to the space $W^{\ell,p}(T)$ for any $p \in [1, \infty)$. It therefore holds that $\hat{\varphi} \in W^{\ell,p}(\widehat{T})$ for any $p \in [1, \infty)$ and thus

$$\begin{aligned}\|\partial_{\hat{x}}^{\beta} \partial_{\hat{x}}^{\gamma} \hat{\varphi}\|_{L^p(\widehat{T})} &\leq c |\det(A_{\widetilde{T}})|^{-\frac{1}{p}} \|\widetilde{A}\|_2^m h^{\beta} \sum_{|\epsilon|=|\gamma|} \widetilde{\mathscr{H}^{\varepsilon}} |\partial_{\tilde{x}}^{\epsilon} \tilde{\varphi}|_{W^{m,p}(\widetilde{T})} \\ &\leq c \|\widetilde{A}\|_2^m h^{\beta} \sum_{|\epsilon|=|\gamma|} \widetilde{\mathscr{H}^{\varepsilon}} |\partial_{\tilde{x}}^{\epsilon} \tilde{\varphi}|_{W^{m,\infty}(\widetilde{T})} < \infty. \end{aligned} \tag{10.5}$$

This implies that the function $\partial_{\hat{x}}^{\beta} \partial_{\hat{x}}^{\gamma} \hat{\varphi}$ is in the space $L^{\infty}(\widehat{T})$. Inequality (10.3) for $p = \infty$ is obtained by passing to the limit $p \to \infty$ in (10.5) on the basis that $\lim_{p\to\infty} \|\cdot\|_{L^p(\widehat{T})} = \|\cdot\|_{L^{\infty}(\widehat{T})}$.

Proof of (10.2)**.** We follow the proof of (10.3). Let $p \in [1, \infty)$. Because the space $\mathcal{C}^{\ell}(\widehat{T})$ is dense in the space $W^{\ell,p}(\widehat{T})$, we show (10.2) for $\hat{\varphi} \in \mathcal{C}^{\ell}(\widehat{T})$ with $\tilde{\varphi} = \hat{\varphi} \circ \Phi_{\widetilde{T}}^{-1}$ and $\varphi = \tilde{\varphi} \circ \Phi_T^{-1}$, it holds that, for $1 \leq i, k \leq d$,

$$\left| \partial_{\hat{x}}^{\beta+\gamma} \hat{\varphi} \right| \leq ch^{\beta} \|\widetilde{A}\|_{\max}^{|\beta|} \|A_T\|_{\max}^{|\beta|} \sum_{|\delta|=|\beta|} \sum_{|\varepsilon|=|\gamma|} h^{\varepsilon} \left| \partial_x^{\delta} \partial_r^{\varepsilon} \varphi \right|.$$

Using (6.8c) and (1.1), we obtain (10.2) for $p \in [1, \infty]$ by an argument analogous to the proof of (10.3). □

Remark 10.3. In inequality (10.3), it is possible to obtain the estimates in T by specifically determining the matrix $\mathcal{A}_T$.

Let $\ell = 2$, $m = 1$ and $p = q = 2$. Recall that

$$\Phi_T : \widetilde{T} \ni \tilde{x} \mapsto x = \Phi_T(\tilde{x}) = A_T \tilde{x} + b_T \in T.$$

For $\tilde{\varphi} \in \mathcal{C}^2(\widetilde{T})$ with $\varphi = \tilde{\varphi} \circ \Phi_T^{-1}$ and $1 \leq i, j \leq d$, we have

$$\left| \frac{\partial^2 \tilde{\varphi}}{\partial \tilde{x}_i \partial \tilde{x}_j}(x) \right| = \left| \sum_{i_1^{(1)}, j_1^{(1)}=1}^{2} [A_T]_{i_1^{(1)} i} [A_T]_{j_1^{(1)} j} \frac{\partial^2 \varphi}{\partial x_{i_1^{(1)}} \partial x_{j_1^{(1)}}}(x) \right|.$$

Let $d = 2$. We define the matrix A_T as

$$A_T := \begin{pmatrix} \cos\frac{\pi}{2} & -\sin\frac{\pi}{2} \\ \sin\frac{\pi}{2} & \cos\frac{\pi}{2} \end{pmatrix}.$$

Because $\|A_T\|_{\max} = 1$, we have

$$\left| \frac{\partial^2 \tilde{\varphi}}{\partial \tilde{x}_i \partial \tilde{x}_j}(x) \right| \leq \left| \frac{\partial^2 \varphi}{\partial x_{i+1} \partial x_{j+1}}(x) \right|,$$

where the indices i, $i+1$ and j, $j+1$ have to be understood mod 2. Because $|\det(A_T)| = 1$, it holds that

$$\left\| \frac{\partial^2 \tilde{\varphi}}{\partial \tilde{x}_i \partial \tilde{x}_j} \right\|_{L^2(\widetilde{T})} \leq \left\| \frac{\partial^2 \varphi}{\partial x_{i_{i+1}} \partial x_{j+1}} \right\|_{L^2(T)}.$$

We then have

$$\sum_{j=1}^{2} \widetilde{\mathscr{H}_j} \left| \frac{\partial \tilde{\varphi}}{\partial \tilde{x}_j} \right|_{H^1(\widetilde{T})} \leq \sum_{j=1}^{2} \widetilde{\mathscr{H}_j} \left| \frac{\partial \varphi}{\partial x_{j+1}} \right|_{H^1(T)},$$

where the indices j, $j+1$ have to be understood mod 2.

We define the matrix A_T as

$$A_T := \begin{pmatrix} \cos \frac{\pi}{4} & -\sin \frac{\pi}{4} \\ \sin \frac{\pi}{4} & \cos \frac{\pi}{4} \end{pmatrix}.$$

We then have

$$\left| \frac{\partial^2 \tilde{\varphi}}{\partial \tilde{x}_i \partial \tilde{x}_j}(x) \right| \leq \frac{1}{\sqrt{2}} \sum_{i_1^{(1)}, j_1^{(1)}=1}^{2} \left| \frac{\partial^2 \varphi}{\partial x_{i_1^{(1)}} \partial x_{j_1^{(1)}}}(x) \right|,$$

which leads to

$$\left\| \frac{\partial^2 \tilde{\varphi}}{\partial \tilde{x}_i \partial \tilde{x}_j} \right\|_{L^2(\widetilde{T})}^2 \leq c \sum_{i_1^{(1)}, j_1^{(1)}=1}^{2} \left\| \frac{\partial^2 \varphi}{\partial x_{i_1^{(1)}} \partial x_{j_1^{(1)}}} \right\|_{L^2(T)}^2 \leq c |\varphi|_{H^2(T)}^2.$$

We then have, using (1.5),

$$\sum_{j=1}^{2} \widetilde{\mathscr{H}_j} \left| \frac{\partial \varphi^s}{\partial x_j^s} \right|_{H^1(T^s)} \leq \sum_{j=1}^{2} \widetilde{\mathscr{H}_j} |\varphi|_{H^2(T)} \leq c h_T |\varphi|_{H^2(T)}.$$

In this case, anisotropic interpolation error estimates cannot be obtained.

Remark 10.4. We consider a general case. Let $p = q = 2$. The space $\mathcal{C}^1(\widetilde{T})$ is dense in the space $H^1(\widetilde{T})$. For $\tilde{\varphi} \in \mathcal{C}^1(\widetilde{T})$ with $\varphi = \tilde{\varphi} \circ \Phi_T^{-1}$ and $1 \leq i \leq d$, we have

$$\left| \frac{\partial \tilde{\varphi}}{\partial \tilde{x}_i}(\tilde{x}) \right| = \left| \sum_{i_1^{(1)}=1}^{d} [A_T]_{i_1^{(1)} i} \frac{\partial \varphi}{\partial x_{i_1^{(1)}}}(x) \right|.$$

Let $d = 2$. We define a rotation matrix A_T as

$$A_T := \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix},$$

where θ denotes the angle. We then have

$$\left|\frac{\partial \tilde{\varphi}}{\partial \tilde{x}_1}(\tilde{x})\right| = \left|\cos\theta \frac{\partial \varphi}{\partial x_1}(x) + \sin\theta \frac{\partial \varphi}{\partial x_2}(x)\right|,$$
$$\left|\frac{\partial \tilde{\varphi}}{\partial \tilde{x}_2}(\tilde{x})\right| = \left|-\sin\theta \frac{\partial \varphi}{\partial x_1}(x) + \cos\theta \frac{\partial \varphi}{\partial x_2}(x)\right|.$$

If $|\sin\theta| \leq c\frac{\widetilde{\mathscr{H}_2}}{\widetilde{\mathscr{H}_1}}$ and $\widetilde{\mathscr{H}_2} \leq c\widetilde{\mathscr{H}_1}$, we can deduce

$$\left|\frac{\partial \tilde{\varphi}}{\partial \tilde{x}_1}(\tilde{x})\right| \leq \left|\frac{\partial \varphi}{\partial x_1}(x)\right| + c\frac{\widetilde{\mathscr{H}_2}}{\widetilde{\mathscr{H}_1}}\left|\frac{\partial \varphi}{\partial x_2}(x)\right|,$$
$$\left|\frac{\partial \tilde{\varphi}}{\partial \tilde{x}_2}(\tilde{x})\right| \leq c\left|\frac{\partial \varphi}{\partial x_1}(x)\right| + \left|\frac{\partial \varphi}{\partial x_2}(x)\right|.$$

As $|\det(A_T)| = 1$, it holds that for $i = 1, 2$,

$$\widetilde{\mathscr{H}_i}\left\|\frac{\partial \tilde{\varphi}}{\partial \tilde{x}_i}\right\|_{L^2(\widetilde{T})} \leq c\sum_{j=1}^{2}\widetilde{\mathscr{H}_j}\left\|\frac{\partial \varphi}{\partial x_j}\right\|_{L^2(T)}.$$

Lemma 10.5. Let Φ_T be the affine mapping defined in (5.2). Let $s \geq 0$ and $1 \leq p \leq \infty$. There exists positive constants c_1 and c_2 such that, for all $T \in \mathbb{T}_h$ and $\varphi \in W^{s,p}(T)$,

$$c_1|\varphi|_{W^{s,p}(T)} \leq |\tilde{\varphi}|_{W^{s,p}(\widetilde{T})} \leq c_2|\varphi|_{W^{s,p}(T)}, \tag{10.6}$$

with $\tilde{\varphi} = \varphi \circ \Phi_T$.

Proof. The following inequalities are found in [20, Lemma 1.101]. There exists a positive constant c such that, for all $T \in \mathbb{T}_h$ and $\varphi \in W^{s,p}(T)$,

$$|\tilde{\varphi}|_{W^{s,p}(\widetilde{T})} \leq c\|A_T\|_2^s|\det(A_T)|^{-\frac{1}{p}}|\varphi|_{W^{s,p}(T)}, \tag{10.7}$$
$$|\varphi|_{W^{s,p}(T)} \leq c\|A_T^{-1}\|_2^s|\det(A_T)|^{\frac{1}{p}}|\tilde{\varphi}|_{W^{s,p}(\widetilde{T})}. \tag{10.8}$$

Because the length of all edges of a simplex and measure of the simplex is not changed by a rotation and mirror imaging matrix and $A_T, A_T^{-1} \in O(d)$,

$$|\det(A_T)| = \frac{|T|_d}{|\widetilde{T}|_d} = 1, \quad \|A_T\|_2 = 1, \quad \|A_T^{-1}\|_2 = 1. \tag{10.9}$$

From (10.7), (10.8), and (10.9), we obtain the desired inequality (10.6). □

11 Classical Interpolation Error Estimates

11.1 Local Interpolation Error Estimates

The following theorem is another representation of the standard interpolation error estimates, e.g., see [20, Theorem 1.103].

Theorem 11.1. Let $1 \leq p \leq \infty$ and assume that there exists a nonnegative integer k such that

$$\mathbb{P}^k \subset \widehat{P} \subset W^{k+1,p}(\widehat{T}) \subset V(\widehat{T}).$$

Let ℓ $(0 \leq \ell \leq k)$ be such that $W^{\ell+1,p}(\widehat{T}) \subset V(\widehat{T})$ with continuous embedding. Furthermore, assume that $\ell, m \in \mathbb{N} \cup \{0\}$ and $p, q \in [1, \infty]$ such that $0 \leq m \leq \ell + 1$ and

$$W^{\ell+1,p}(\widehat{T}) \hookrightarrow W^{m,q}(\widehat{T}). \tag{11.1}$$

It holds that, for any $m \in \{0, \dots, \ell+1\}$ and any $\varphi \in W^{\ell+1,p}(T)$,

$$|\varphi - I_T \varphi|_{W^{m,q}(T)} \leq C_*^I |T|_d^{\frac{1}{q}-\frac{1}{p}} \left(\frac{h_{\max}}{h_{\min}} \right)^m \left(\frac{H_T}{h_T} \right)^m h_T^{\ell+1-m} |\varphi|_{W^{\ell+1,p}(T)}, \tag{11.2}$$

where C_*^I is a positive constant independent of h_T and H_T, and the parameters $h_{\max}$ and $h_{\min}$ are defined by (8.1), that is,

$$h_{\max} = \max\{h_1, \dots, h_d\}, \quad h_{\min} = \min\{h_1, \dots, h_d\}.$$

Proof. Let $\hat{\varphi} \in W^{\ell+1,p}(\widehat{T})$. Because $0 \leq \ell \leq k$, $\mathbb{P}^\ell \subset \mathbb{P}^k \subset \widehat{P}$. Therefore, for any $\hat{\eta} \in \mathbb{P}^\ell$, we have $I_{\widehat{T}} \hat{\eta} = \hat{\eta}$. Using (9.3) and (11.1), we obtain

$$\begin{aligned}
|\tilde{\varphi} - I_{\widehat{T}} \hat{\varphi}|_{W^{m,q}(\widehat{T})} &\leq |\tilde{\varphi} - \hat{\eta}|_{W^{m,q}(\widehat{T})} + |I_{\widehat{T}}(\hat{\eta} - \hat{\varphi})|_{W^{m,q}(\widehat{T})} \\
&\leq c \|\tilde{\varphi} - \hat{\eta}\|_{W^{\ell+1,p}(\widehat{T})},
\end{aligned}$$

where we used the stability of the interpolation operator $I_{\widehat{T}}$, that is,

$$|I_{\widehat{T}}(\hat{\eta} - \hat{\varphi})|_{W^{m,q}(\widehat{T})} \leq \sum_{i=1}^{n_0} |\hat{\chi}_i(\hat{\eta} - \hat{\varphi})| |\hat{\theta}_i|_{W^{m,q}(\widehat{T})} \leq c \|\hat{\eta} - \tilde{\varphi}\|_{W^{\ell+1,p}(\widehat{T})}.$$

Using the classic Bramble–Hilbert–type lemma (e.g., [14, Lemma 4.3.8]), we obtain

$$|\tilde{\varphi} - I_{\widehat{T}} \hat{\varphi}|_{W^{m,q}(\widehat{T})} \leq c \inf_{\hat{\eta} \in \mathbb{P}^\ell} \|\hat{\eta} - \tilde{\varphi}\|_{W^{\ell+1,p}(\widehat{T})} \leq c |\hat{\varphi}|_{W^{\ell+1,p}(\widehat{T})}. \tag{11.3}$$

The inequalities (10.6), (10.1), (1.5), and (11.3) yield

$$\begin{aligned}
|\varphi - I_T \varphi|_{W^{m,q}(T)} &\leq c |\tilde{\varphi} - I_{\widetilde{T}} \tilde{\varphi}|_{W^{m,q}(\widetilde{T})} \\
&\leq c |\det(A_{\widetilde{T}})|^{\frac{1}{q}} \|\widetilde{A}^{-1}\|_2^m \left(\sum_{|\beta|=m} (h^{-\beta})^q \|\partial^\beta (\hat{\varphi} - I_{\widehat{T}} \hat{\varphi})\|_{L^q(\widehat{T})}^q \right)^{\frac{1}{q}} \\
&\leq c |\det(A_{\widetilde{T}})|^{\frac{1}{q}} \|\widetilde{A}^{-1}\|_2^m \max\{h_1^{-1}, \dots, h_d^{-1}\}^{|\beta|} |\tilde{\varphi} - I_{\widehat{T}} \hat{\varphi}|_{W^{m,q}(\widehat{T})} \\
&\leq c |\det(A_{\widetilde{T}})|^{\frac{1}{q}} \|\widetilde{A}^{-1}\|_2^m h_{\min}^{-|\beta|} |\hat{\varphi}|_{W^{\ell+1,p}(\widehat{T})}.
\end{aligned} \tag{11.4}$$

Using the inequalities (1.5), (10.6) and (10.2), we have

$$\begin{aligned}
|\hat{\varphi}|_{W^{\ell+1,p}(\widehat{T})} &\leq \sum_{|\gamma|=\ell+1-m} \sum_{|\beta|=m} \|\partial^\beta \partial^\gamma \hat{\varphi}\|_{L^p(\widehat{T})} \\
&\leq c |\det(A_{\widetilde{T}})|^{-\frac{1}{p}} \|\widetilde{A}\|_2^m \sum_{|\gamma|=\ell+1-m} \sum_{|\beta|=m} h^\beta \sum_{|\epsilon|=|\gamma|} h^\varepsilon |\partial_r^\epsilon \varphi|_{W^{m,p}(T)} \\
&\leq c |\det(A_{\widetilde{T}})|^{-\frac{1}{p}} \|\widetilde{A}\|_2^m \max\{h_1, \dots, h_d\}^{|\beta|} h_T^{\ell+1-m} |\varphi|_{W^{\ell+1,p}(T)} \\
&\leq c |\det(A_{\widetilde{T}})|^{-\frac{1}{p}} \|\widetilde{A}\|_2^m h_{\max}^{|\beta|} h_T^{\ell+1-m} |\varphi|_{W^{\ell+1,p}(T)}.
\end{aligned} \tag{11.5}$$

From (11.4) and (11.5) together with (6.8) and (6.9), we have the desired estimate (11.2). □

Remark 11.2. We introduced the estimate (1.12), a variant of the Bramble–Hilbert lemma. However, because we prove estimate (11.3) with $p = q$ using the reference element, it is sufficient to use the standard estimate (e.g., [19, 14]) to achieve our goal.

Example 11.3. As the examples in [20, Example 1.106], we get local interpolation error estimates for a Lagrange finite element of degree k, a more general finite element, and the Crouzeix–Raviart finite element with $k = 1$.

1. For a Lagrange finite element of degree k, we set $V(\widehat{T}) := \mathcal{C}^0(\widehat{T})$. The condition on ℓ in Theorem 11.1 is $\frac{d}{p} - 1 < \ell \leq k$ because $W^{\ell+1,p}(\widehat{T}) \subset \mathcal{C}^0(\widehat{T})$ if $\ell + 1 > \frac{d}{p}$ according to the Sobolev imbedding theorem.

2. For a general finite element with $V(\widehat{T}) := \mathcal{C}^t(\widehat{T})$ and $t \in \mathbb{N}$. The condition on ℓ in Theorem 11.1 is $\frac{d}{p} - 1 + t < \ell \leq k$. When $t = 1$, there is a Hermite finite element.

3. For the Crouzeix–Raviart finite element with $k = 1$, we set $V(\widehat{T}) := W^{1,1}(\widehat{T})$. The condition on ℓ in Theorem 11.1 is $0 \leq \ell \leq 1$.

11.2 Examples of Anisotropic Elements

When $m = \ell = 1$ and $q = p$ in (11.2) of Theorem 11.1, the estimate is written as

$$|\varphi - I_T\varphi|_{W^{1,p}(T)} \leq C_*^I \frac{h_{\max}}{h_{\min}} \frac{H_T}{h_T} h_T |\varphi|_{W^{2,p}(T)}. \tag{11.6}$$

Let $T \subset \mathbb{R}^2$ be a triangle. As described in Section 8.1, an isotropic mesh element has equal or nearly equal edge lengths and angles, resulting in a balanced shape. Then, the estimate (11.6) becomes

$$|\varphi - I_T\varphi|_{W^{1,p}(T)} \leq c h_T |\varphi|_{W^{2,p}(T)}. \tag{11.7}$$

We introduce typical examples of the quantities $\frac{h_{\max}}{h_{\min}}$ and $\frac{H_T}{h_T}$ in anisotropic elements. We considered the following five anisotropic elements as in Section 8.2: Let $0 < s \ll 1$, $s \in \mathbb{R}$ and $\varepsilon, \delta, \gamma \in \mathbb{R}$.

Example 11.4 (Right-angled triangle)**.** Let $T \subset \mathbb{R}^2$ be the simplex with vertices $p_1 := (0,0)^\top$, $p_2 := (s,0)^\top$ and $p_3 := (0,s^\varepsilon)^\top$ with $1 < \varepsilon$. We then have $h_1 = s$, $h_2 = s^\varepsilon$ and $h_T = \sqrt{s^2 + s^{2\varepsilon}}$; i.e.,

$$\frac{h_{\max}}{h_{\min}} \leq s^{1-\varepsilon} \to \infty \quad \text{as } s \to 0, \qquad \frac{H_T}{h_T} = 2.$$

In this case, the estimate (11.6) becomes

$$|\varphi - I_T\varphi|_{W^{1,p}(T)} \leq 2C_*^I s^{2-\varepsilon} |\varphi|_{W^{2,p}(T)}.$$

When $\varepsilon > 2$, this implies that the estimate diverges as $s \to 0$. However, new interpolation error estimates will be shown to converge, see Example 14.3.

Example 11.5 (Dagger)**.** Let $T \subset \mathbb{R}^2$ be the simplex with vertices $p_1 := (0,0)^\top$, $p_2 := (s,0)^\top$ and $p_3 := (s^\delta, s^\varepsilon)^\top$ with $1 < \varepsilon < \delta$. We then have $h_1 = \sqrt{(s - s^\delta)^2 + s^{2\varepsilon}}$, $h_2 = \sqrt{s^{2\delta} + s^{2\varepsilon}}$ and $h_T = s$; i.e.,

$$\begin{aligned} \frac{h_{\max}}{h_{\min}} &= \frac{\sqrt{(s - s^\delta)^2 + s^{2\varepsilon}}}{\sqrt{s^{2\delta} + s^{2\varepsilon}}} \leq c s^{1-\varepsilon} \to \infty \quad \text{as } s \to 0, \\ \frac{H_T}{h_T} &= \frac{\sqrt{(s - s^\delta)^2 + s^{2\varepsilon}}\sqrt{s^{2\delta} + s^{2\varepsilon}}}{\frac{1}{2} s^{1+\varepsilon}} \leq c. \end{aligned}$$

In this case, the estimate (11.6) becomes

$$|\varphi - I_T\varphi|_{W^{1,p}(T)} \le cs^{2-\varepsilon}|\varphi|_{W^{2,p}(T)}.$$

When $\varepsilon > 2$, this implies that the estimate diverges as $s \to 0$. However, new interpolation error estimates will be shown to converge, see Example 14.4.

Example 11.6 (Blade)**.** Let $T \subset \mathbb{R}^2$ be the simplex with vertices $p_1 := (0,0)^\top$, $p_2 := (2s,0)^\top$ and $p_3 := (s, s^\varepsilon)^\top$ with $1 < \varepsilon$. We then have $h_1 = h_2 = \sqrt{s^2 + s^{2\varepsilon}}$ and $h_T = 2s$; i.e.,

$$\frac{h_{\max}}{h_{\min}} = 1, \quad \frac{H_T}{h_T} = \frac{s^2 + s^{2\varepsilon}}{s^{1+\varepsilon}} \to \infty \quad \text{as } s \to 0.$$

In this case, the estimate (11.6) becomes

$$|\varphi - I_T\varphi|_{W^{1,p}(T)} \le cs^{2-\varepsilon}|\varphi|_{W^{2,p}(T)}.$$

When $\varepsilon > 2$, this implies that the estimate diverges as $s \to 0$. In this case, the interpolation error estimate can not be improved, , see Example 14.5.

Example 11.7 (Dagger)**.** Let $T \subset \mathbb{R}^2$ be the simplex with vertices $p_1 := (0,0)^\top$, $p_2 := (s,0)^\top$ and $p_3 := (s^\delta, s^\varepsilon)^\top$ with $1 < \delta < \varepsilon$. We then have $h_1 = \sqrt{(s - s^\delta)^2 + s^{2\varepsilon}}$, $h_2 = \sqrt{s^{2\delta} + s^{2\varepsilon}}$ and $h_T = s$; i.e.,

$$\begin{aligned} \frac{h_{\max}}{h_{\min}} &= \frac{\sqrt{(s - s^\delta)^2 + s^{2\varepsilon}}}{\sqrt{s^{2\delta} + s^{2\varepsilon}}} \le cs^{1-\delta} \to \infty \quad \text{as } s \to 0, \\ \frac{H_T}{h_T} &= \frac{\sqrt{(s - s^\delta)^2 + s^{2\varepsilon}}\sqrt{s^{2\delta} + s^{2\varepsilon}}}{\frac{1}{2}s^{1+\varepsilon}} \le cs^{\delta-\varepsilon} \to \infty \quad \text{as } s \to 0. \end{aligned}$$

In this case, the estimate (11.6) becomes

$$|\varphi - I_T\varphi|_{W^{1,p}(T)} \le cs^{2-\varepsilon}|\varphi|_{W^{2,p}(T)}.$$

When $\varepsilon > 2$, this implies that the estimate diverges as $s \to 0$. In this case, the interpolation error estimate can not be improved, see Example 14.6.

Example 11.8 (Right-angled triangle)**.** Let $T \subset \mathbb{R}^2$ be the simplex with vertices $p_1 := (0,0)^\top$, $p_2 := (s,0)^\top$ and $p_3 := (0, \delta s)^\top$ with $\delta \ll 1$. We then have $h_1 = s$, $h_2 = \delta s$ and $h_T = s\sqrt{1+\delta^2}$; i.e.,

$$\frac{h_{\max}}{h_{\min}} = \frac{1}{\delta}, \quad \frac{H_T}{h_T} = 2.$$

In this case, the estimate (11.6) becomes

$$|\varphi - I_T\varphi|_{W^{1,p}(T)} \le \frac{c}{\delta} s|\varphi|_{W^{2,p}(T)}.$$

This implies that the estimate converges as $s \to 0$ and the error may be large. However, new interpolation error estimates remove the factor $\frac{1}{\delta}$, see Example 14.7.

Example 11.9 (Blade)**.** Let $T \subset \mathbb{R}^2$ be the simplex with vertices $p_1 := (0,0)^\top$, $p_2 := (2s,0)^\top$ and $p_3 := (s, \delta s)^\top$ with $\delta \ll 1$. We then have $h_1 = h_2 = s\sqrt{1+\delta^2}$ and $h_T = 2s$; i.e.,

$$\frac{h_{\max}}{h_{\min}} = 1, \quad \frac{H_T}{h_T} = \frac{s^2(1+\delta^2)}{\delta s^2} \le \frac{c}{\delta},$$

In this case, the estimate (11.6) becomes

$$|\varphi - I_T\varphi|_{W^{1,p}(T)} \le \frac{c}{\delta} s|\varphi|_{W^{2,p}(T)}.$$

This implies that the estimate converges as $s \to 0$ and the error may be large. Unfortunately, new interpolation error estimates do not remove the factor $\frac{1}{\delta}$, see Example 14.8.

Example 11.10 ($\mathbb{P}^1$ + bubble finite element in $\mathbb{R}^2$)**.** We give a numerical example which is not optimal in the usual sense. Let $T \subset \mathbb{R}^2$ be the triangle with vertices $p_1 := (0,0)^\top$, $p_2 := (s,0)^\top$, $p_3 := (0,s^\varepsilon)^\top$ (Example 11.4), where $s := \frac{1}{N}$, $N \in \mathbb{N}$ and $\varepsilon \in \mathbb{R}$, $1 < \varepsilon \leq 2$. Let p_4 be the barycentre of T.

Using the barycentric coordinates $\lambda_i : \mathbb{R}^2 \to \mathbb{R}$, $i = 1, \ldots, 3$, we define the local basis functions as

$$\theta_4(x) := 27\lambda_1(x)\lambda_2(x)\lambda_3(x), \quad \theta_i(x) := \lambda_i(x) - \frac{1}{3}\theta_4(x), \quad i = 1,2,3.$$

The interpolation operator I_T^b defined by

$$I_T^b : H^2(T) \ni \varphi \mapsto I_T^b\varphi := \sum_{i=1}^{4} \varphi(x_i)\theta_i \in \operatorname{span}\{\theta_1, \theta_2, \theta_3, \theta_4\}.$$

From Theorem 11.1, we have

$$|\varphi - I_T^b\varphi|_{H^1(T)} \leq c h_T^{2-\varepsilon} |\varphi|_{H^2(T^s)} \quad \forall \varphi \in H^2(T).$$

Let φ be a function such that

$$\varphi(x,y) := 2x^2 - xy + 3y^2.$$

We compute the convergence order concerning the H^1 norm defined by

$$Err_s^b(H^1) := \frac{|\varphi - I_T^b\varphi|_{H^1(T)}}{|\varphi|_{H^2(T)}},$$

for the cases: $\varepsilon = 1.5$ (Table 13) and $\varepsilon = 2.0$ (Table 14). The convergence indicator r is defined by

$$r = \frac{1}{\log(2)} \log\left(\frac{Err_t^b(H^1)}{Err_{t/2}^b(H^1)}\right).$$

Table 11: Error of the local interpolation operator ($\varepsilon = 1.5$)

N	s	$Err_s^b(H^1)$	r
128	7.8125e-03	2.9951e-02	
256	3.9062e-03	2.1101e-02	5.0529e-01
512	1.9531e-03	1.4874e-02	5.0452e-01
1024	9.7656e-04	1.0491e-02	5.0364e-01

Table 12: Error of the local interpolation operator ($\varepsilon = 2.0$)

N	s	$Err_s^b(H^1)$	r
128	7.8125e-03	3.3397e-01	
256	3.9062e-03	3.3366e-01	1.3398e-03
512	1.9531e-03	3.3350e-01	6.9198e-04
1024	9.7656e-04	3.3341e-01	3.8939e-04

Remark 11.11. If we are concerned with anisotropic elements, it would be desirable to remove the quantity $h_{\max}/h_{\min}$ from estimate (11.2).

12 Anisotropic Interpolation on the Reference Element

We introduce estimates on the reference element due to [4, 3] to obtain anisotropic interpolation error estimates.

For the reference element $\widehat{T}$ defined in Sections 5.1 and 5.1, let the triple $\{\widehat{T}, \widehat{P}, \widehat{\Sigma}\}$ be the reference finite element with associated normed vector space $V(\widehat{T})$.

Theorem 12.1. Let $I_{\widehat{T}} : \mathcal{C}(\widehat{T}) \to \mathbb{P}^k(\widehat{T})$ be a linear operator. Fix $m, \ell \in \mathbb{N}$ and $p, q \in [1, \infty]$ such that $0 \leq m \leq \ell \leq k+1$ and

$$W^{\ell-m,p}(\widehat{T}) \hookrightarrow L^q(\widehat{T}). \tag{12.1}$$

Let β be a multi-index with $|\beta| = m$. We set $j := \dim(\partial_{\hat{x}}^{\beta} \mathcal{P}^k)$. Assume that there exist linear functionals $\mathscr{F}_i$, $i = 1, \ldots, j$, such that

$$\mathscr{F}_i \in W^{\ell-m,p}(\widehat{T})', \quad \forall i = 1, \ldots, j, \tag{12.2a}$$

$$\mathscr{F}_i(\partial_{\hat{x}}^{\beta}(\hat{\varphi} - I_{\widehat{T}}\hat{\varphi})) = 0 \quad \forall i = 1, \ldots, j, \quad \forall \hat{\varphi} \in \mathcal{C}(\widehat{T}) : \ \partial_{\hat{x}}^{\beta}\hat{\varphi} \in W^{\ell-m,p}(\widehat{T}), \tag{12.2b}$$

$$\hat{\eta} \in \mathbb{P}^k, \quad \mathscr{F}_i(\partial_{\hat{x}}^{\beta}\hat{\eta}) = 0 \quad \forall i = 1, \ldots, j \quad \Rightarrow \quad \partial^{\beta}\hat{\eta} = 0. \tag{12.2c}$$

It holds that for all $\hat{\varphi} \in \mathcal{C}(\widehat{T})$ with $\partial^{\beta}\hat{\varphi} \in W^{\ell-m,p}(\widehat{T})$,

$$\|\partial_{\hat{x}}^{\beta}(\hat{\varphi} - I_{\widehat{T}}\hat{\varphi})\|_{L^q(\widehat{T})} \leq C^F |\partial_{\hat{x}}^{\beta}\hat{\varphi}|_{W^{\ell-m,p}(\widehat{T})}. \tag{12.3}$$

Proof. We follow [3, Lemma 2.2].

For all $\hat{\eta} \in \mathbb{P}^{\ell-1}$, we have

$$\|\partial_{\hat{x}}^{\beta}(\hat{\varphi} - I_{\widehat{T}}\hat{\varphi})\|_{L^q(\widehat{T})} \leq \|\partial_{\hat{x}}^{\beta}(\hat{\varphi} - \hat{\eta})\|_{L^q(\widehat{T})} + \|\partial_{\hat{x}}^{\beta}(\hat{\eta} - I_{\widehat{T}}\hat{\varphi})\|_{L^q(\widehat{T})}. \tag{12.4}$$

Note that $\hat{\eta} - I_{\widehat{T}}\hat{\varphi} \in \mathbb{P}^k$, because $\ell \leq k+1$. That is, $\partial_{\hat{x}}^{\beta}(\hat{\eta} - I_{\widehat{T}}\hat{\varphi}) \in \partial_{\hat{x}}^{\beta}\mathbb{P}^k$. Because the polynomial spaces are finite-dimensional all norms are equivalent, that is, by the fact $\sum_{i=1}^{j} |\mathscr{F}_i(\hat{\eta})|$ is a norm on $\partial_{\hat{x}}^{\beta}\mathbb{P}^k$, together with (12.2a), (12.2b) and (12.2c), we have for any $\hat{\eta} \in \mathbb{P}^{\ell-1}$,

$$\begin{aligned}
\|\partial_{\hat{x}}^{\beta}(\hat{\eta} - I_{\widehat{T}}\hat{\varphi})\|_{L^q(\widehat{T})} &\leq c \sum_{i=1}^{j} |\mathscr{F}_i(\partial_{\hat{x}}^{\beta}(\hat{\eta} - I_{\widehat{T}}\hat{\varphi}))| = c \sum_{i=1}^{j} |\mathscr{F}_i(\partial_{\hat{x}}^{\beta}(\hat{\eta} - \hat{\varphi}))| \\
&\leq c \|\partial_{\hat{x}}^{\beta}(\hat{\eta} - \hat{\varphi})\|_{W^{\ell-m,p}(\widehat{T})}.
\end{aligned}$$

Using (12.4) and (16.5), it holds that for any $\hat{\eta} \in \mathbb{P}^{\ell-1}$,

$$\begin{aligned}
\|\partial_{\hat{x}}^{\beta}(\hat{\varphi} - I_{\widehat{T}}\hat{\varphi})\|_{L^q(\widehat{T})} &\leq \|\partial_{\hat{x}}^{\beta}(\hat{\varphi} - \hat{\eta})\|_{L^q(\widehat{T})} + \|\partial_{\hat{x}}^{\beta}(\hat{\eta} - I_{\widehat{T}}\hat{\varphi})\|_{L^q(\widehat{T})} \\
&\leq c \|\partial_{\hat{x}}^{\beta}(\hat{\eta} - \hat{\varphi})\|_{W^{\ell-m,p}(\widehat{T})}.
\end{aligned}$$

By Lemma 1.10, we have

$$\begin{aligned}
\|\partial_{\hat{x}}^{\beta}(\hat{\varphi} - I_{\widehat{T}}\hat{\varphi})\|_{L^q(\widehat{T})} &\leq c \inf_{\hat{\eta} \in \mathbb{P}^{\ell-1}} \|\partial_{\hat{x}}^{\beta}(\hat{\eta} - \hat{\varphi})\|_{W^{\ell-m,p}(\widehat{T})} \\
&\leq c |\partial_{\hat{x}}^{\beta}\hat{\varphi}|_{W^{\ell-m,p}(\widehat{T})}.
\end{aligned}$$

□

Remark 12.2. Note that it is not required $I_{\widehat{T}}\hat{\eta} = \hat{\eta}$ for any $\hat{\eta} \in \mathbb{P}^{\ell-1}$.

13 Remarks on Anisotropic Interpolation Analysis

Let $\widehat{T} \subset \mathbb{R}^2$ be the reference element defined in Section 5.1. We set $k = m = 1$, $\ell = 2$, and $p = 2$. For $\hat{\varphi} \in H^2(\widehat{T})$, we set $\tilde{\varphi} = \hat{\varphi} \circ \Phi_{\widetilde{T}}^{-1}$ and $\varphi = \tilde{\varphi} \circ \Phi_T^{-1}$. Inequalities (10.1) and (10.6) yield

$$|\varphi - I_T \varphi|_{H^1(T)} \leq c |\det(A_{\widetilde{T}})|^{\frac{1}{2}} \|\widetilde{A}^{-1}\|_2 \left(\sum_{i=1}^{2} h_i^{-2} \|\partial_{\hat{x}_i} (\hat{\varphi} - I_{\widehat{T}} \hat{\varphi})\|^2_{L^2(\widehat{T})} \right)^{\frac{1}{2}}. \tag{13.1}$$

The coefficient h_i^{-2} appears on the right-hand side of Eq. (13.1). A further assumption is required for this. Using Eq. (1.11) and the triangle inequality, we have

$$\|\partial_{\hat{x}_i} (\hat{\varphi} - I_{\widehat{T}} \hat{\varphi})\|^2_{L^2(\widehat{T})} \leq 2 \|\partial_{\hat{x}_i} (\hat{\varphi} - Q^{(2)} \hat{\varphi})\|^2_{L^2(\widehat{T})} + 2 \|\partial_{\hat{x}_i} (Q^{(2)} \hat{\varphi} - I_{\widehat{T}} \hat{\varphi})\|^2_{L^2(\widehat{T})}.$$

We use inequality (1.10) to remove the coefficient h_i^{-2}. To this end, we have to show that

$$\|\partial_{\hat{x}_i} (Q^{(2)} \hat{\varphi} - I_{\widehat{T}} \hat{\varphi})\|_{L^2(\widehat{T})} \leq c \|\partial_{\hat{x}_i} (\hat{\varphi} - Q^{(2)} \hat{\varphi})\|_{H^1(\widehat{T})}. \tag{13.2}$$

However, this is unlikely to hold because Eqs. (9.1) and (9.3) yield

$$\begin{aligned} \|\partial_{\hat{x}_i} (Q^{(2)} \hat{\varphi} - I_{\widehat{T}} \hat{\varphi})\|_{L^2(\widehat{T})} &= \|\partial_{\hat{x}_i} (I_{\widehat{T}} (Q^{(2)} \hat{\varphi}) - I_{\widehat{T}} \hat{\varphi})\|_{L^2(\widehat{T})} \\ &\leq c \|Q^{(2)} \hat{\varphi} - \hat{\varphi}\|_{H^2(\widehat{T})} \leq c |\hat{\varphi}|_{H^2(\widehat{T})}. \end{aligned}$$

Using the classical scaling argument (see [20, Lemma 1.101]), we have

$$|\hat{\varphi}|_{H^2(\widehat{T})} \leq c |\det(A)|^{-\frac{1}{2}} \|A\|_2 |\varphi|_{H^2(T)},$$

which does not include the quantity h_i. Therefore, the quantity h_i^{-1} in Eq. (13.1) remains.

To overcome this problem, we use Theorem 12.1. That is, we assume that there exists a linear functional $\mathscr{F}_1$ such that

$$\begin{aligned} &\mathscr{F}_1 \in H^1(\widehat{T})', \\ &\mathscr{F}_1(\partial_{\hat{x}_i}(\hat{\varphi} - I_{\widehat{T}} \hat{\varphi})) = 0 \quad i = 1, 2, \quad \forall \hat{\varphi} \in \mathcal{C}(\widehat{T}) : \ \partial_{\hat{x}_i} \hat{\varphi} \in H^1(\widehat{T}), \\ &\hat{\eta} \in \mathbb{P}^1, \quad \mathscr{F}_1(\partial_{\hat{x}_i} \hat{\eta}) = 0 \quad i = 1, 2, \quad \Rightarrow \quad \partial_{\hat{x}_i} \hat{\eta} = 0. \end{aligned}$$

Because the polynomial spaces are finite-dimensional, all norms are equivalent; i.e., because $|\mathscr{F}_1(\partial_{\hat{x}_i}(\hat{\eta} - I_{\widehat{T}} \hat{\varphi}))|$ $(i = 1, 2)$ is a norm on $\mathbb{P}^0$, we have that, for $i = 1, 2$,

$$\begin{aligned} \|\partial_{\hat{x}_i} (\hat{\eta} - I_{\widehat{T}} \hat{\varphi})\|_{L^2(\widehat{T})} &\leq c |\mathscr{F}_1(\partial_{\hat{x}_i}(\hat{\eta} - I_{\widehat{T}} \hat{\varphi}))| = c |\mathscr{F}_1(\partial_{\hat{x}_i}(\hat{\eta} - \hat{\varphi}))| \\ &\leq c \|\partial_{\hat{x}_i} (\hat{\eta} - \hat{\varphi})\|_{H^1(\widehat{T})}. \end{aligned}$$

Setting $\hat{\eta} := Q^{(2)} \hat{\varphi}$, we obtain Eq. (13.2). Using inequality (1.10) yields

$$\|\partial_{\hat{x}_i} (\hat{\varphi} - I_{\widehat{T}} \hat{\varphi})\|^2_{L^2(\widehat{T})} \leq c |\partial_{\hat{x}_i} \hat{\varphi}|^2_{H^1(\widehat{T})},$$

and so inequality (13.1) together with Eq. (1.5) can be written as

$$|\varphi - I_T \varphi|_{H^1(T)} \leq c |\det(A_{\widetilde{T}})|^{\frac{1}{2}} \|\widetilde{A}^{-1}\|_2 \sum_{i,j=1}^{2} h_i^{-1} \|\partial_{\hat{x}_i} \partial_{\hat{x}_j} \hat{\varphi}\|_{L^2(\widehat{T})}. \tag{13.3}$$

Inequality (10.2) yields

$$\|\partial_{\hat{x}_i} \partial_{\hat{x}_j} \hat{\varphi}\|_{L^2(\widehat{T})} \leq c |\det(A_{\widetilde{T}})|^{-\frac{1}{2}} \|\widetilde{A}\|_2 h_i \sum_{n=1}^{2} h_n \left| \frac{\partial \varphi}{\partial r_n} \right|_{H^1(T)}. \tag{13.4}$$

Therefore, the quantity h_i^{-1} in Eq. (13.3) and the quantity h_i in Eq. (13.4) cancel out.

14 New Interpolation Error Estimates

14.1 Local Interpolation Error Estimates

The new scaling arguments in Section 10 are the heart of the following local interpolation error estimates.

Theorem 14.1 (Local interpolation)**.** Let $\{\widehat{T}, \widehat{P}, \widehat{\Sigma}\}$ be a finite element with the normed vector space $V(\widehat{T}) := \mathcal{C}(\widehat{T})$ and $\widehat{P} := \mathcal{P}^k(\widehat{T})$ with $k \geq 1$. Let $I_{\widehat{T}} : V(\widehat{T}) \to \widehat{P}$ be a linear operator. Fix $\ell \in \mathbb{N}$, $m \in \mathbb{N}_0$, and $p, q \in [1, \infty]$ such that $0 \leq m \leq \ell \leq k+1$, $\ell - m \geq 1$, and the embeddings (1.6) and (1.7) with $s := \ell - m$ hold. Let β be a multi-index with $|\beta| = m$. We set $j := \dim(\partial^\beta \mathcal{P}^k)$. Assume that there exist linear functionals $\mathscr{F}_i$, $i = 1, \ldots, j$, satisfying the conditions (12.2). It then holds that, for all $\hat{\varphi} \in W^{\ell,p}(\widehat{T}) \cap \mathcal{C}(\widehat{T})$ with $\varphi := \hat{\varphi} \circ \Phi^{-1}$,

$$|\varphi - I_T \varphi|_{W^{m,q}(T)} \leq C_1^I |T|_d^{\frac{1}{q} - \frac{1}{p}} \left(\frac{H_T}{h_T} \right)^m \sum_{|\varepsilon| = \ell - m} h^\varepsilon |\partial_r^\varepsilon \varphi|_{W^{m,p}(T)}, \tag{14.1}$$

where C_1^I is a positive constant independent of h_T and H_T. In particular, if Condition 10.1 is imposed, it holds that, for all $\hat{\varphi} \in W^{\ell,p}(\widehat{T}) \cap \mathcal{C}(\widehat{T})$ with $\varphi = \hat{\varphi} \circ \Phi^{-1}$,

$$|\varphi - I_T \varphi|_{W^{m,q}(T)} \leq C_2^I |T|_d^{\frac{1}{q} - \frac{1}{p}} \left(\frac{H_T}{h_T} \right)^m \sum_{|\varepsilon| = \ell - m} \widetilde{\mathscr{H}^\varepsilon} |\partial_{\tilde{x}}^\varepsilon (\varphi \circ \Phi_T)|_{W^{m,p}(\Phi_T^{-1}(T))}, \tag{14.2}$$

where C_2^I is a positive constant independent of h_{T^s} and H_{T^s}.

Proof. The introduction of the functionals $\mathscr{F}_i$ follows from [4, 3], also see Theorem 12.1. Actually, under the same assumptions as in Theorem 14.1, we have

$$\|\partial_{\hat{x}}^\beta (\hat{\varphi} - I_{\widehat{T}} \hat{\varphi})\|_{L^q(\widehat{T})} \leq C^B |\partial_{\hat{x}}^\beta \hat{\varphi}|_{W^{\ell-m,p}(\widehat{T})}, \tag{14.3}$$

where $|\beta| = m$, $\hat{\varphi} \in \mathcal{C}(\widehat{T})$, and $\partial_{\hat{x}}^\beta \hat{\varphi} \in W^{\ell-m,p}(\widehat{T})$.

The inequalities in (10.6), (1.5), (10.1), and (14.3) yield

$$\begin{aligned}
|\varphi - I_T \varphi|_{W^{m,q}(T)} &\leq c |\varphi - I_T \varphi|_{W^{m,q}(T)} \\
&\leq c |\det(A_{\widetilde{T}})|^{\frac{1}{q}} \|\widetilde{A}^{-1}\|_2^m \left(\sum_{|\beta|=m} (h^{-\beta})^q \|\partial_{\hat{x}}^\beta (\hat{\varphi} - I_{\widehat{T}} \hat{\varphi})\|_{L^q(\widehat{T})}^q \right)^{1/q} \\
&\leq c |\det(A_{\widetilde{T}})|^{\frac{1}{q}} \|\widetilde{A}^{-1}\|_2^m \sum_{|\beta|=m} (h^{-\beta}) \|\partial_{\hat{x}}^\beta (\hat{\varphi} - I_{\widehat{T}} \hat{\varphi})\|_{L^q(\widehat{T})} \\
&\leq c |\det(A_{\widetilde{T}})|^{\frac{1}{q}} \|\widetilde{A}^{-1}\|_2^m \sum_{|\beta|=m} (h^{-\beta}) |\partial_{\hat{x}}^\beta \hat{\varphi}|_{W^{\ell-m,p}(\widehat{T})}.
\end{aligned} \tag{14.4}$$

Inequalities (1.5) and (10.2) yield

$$\begin{aligned}
&\sum_{|\beta|=m} (h^{-\beta}) |\partial_{\hat{x}}^\beta \hat{\varphi}|_{W^{\ell-m,p}(\widehat{T})} \\
&\quad \leq \sum_{|\gamma|=\ell-m} \sum_{|\beta|=m} (h^{-\beta}) \|\partial_{\hat{x}}^\beta \partial_{\hat{x}}^\gamma \hat{\varphi}\|_{L^p(\widehat{T})} \\
&\quad \leq c |\det(A_{\widetilde{T}})|^{-\frac{1}{p}} \|\widetilde{A}\|_2^m \sum_{|\gamma|=\ell-m} \sum_{|\beta|=m} (h^{-\beta}) h^\beta \sum_{|\epsilon|=|\gamma|} h^\varepsilon |\partial_r^\epsilon \varphi|_{W^{m,p}(T)} \\
&\quad \leq c |\det(A_{\widetilde{T}})|^{-\frac{1}{p}} \|\widetilde{A}\|_2^m \sum_{|\epsilon|=\ell-m} h^\varepsilon |\partial_r^\varepsilon \varphi|_{W^{m,p}(T)}.
\end{aligned} \tag{14.5}$$

From (6.8), (6.9), (14.4), and (14.5), we have

$$|\varphi - I_T\varphi|_{W^{m,q}(T)} \leq c|T|_d^{\frac{1}{q}-\frac{1}{p}} \left(\frac{H_T}{h_T}\right)^m \sum_{|\varepsilon|=\ell-m} h^{\varepsilon} |\partial_r^{\varepsilon}\varphi|_{W^{m,p}(T)},$$

which is the inequality (14.1).

Assume that Condition 10.1 is imposed. Inequality (10.3) yields

$$\begin{aligned}
&\sum_{|\beta|=m} (h^{-\beta}) |\partial_{\hat{x}}^{\beta}\hat{\varphi}|_{W^{\ell-m,p}(\widehat{T})} \\
&\quad \leq \sum_{|\gamma|=\ell-m} \sum_{|\beta|=m} (h^{-\beta}) \|\partial_{\hat{x}}^{\beta}\partial_{\hat{x}}^{\gamma}\hat{\varphi}\|_{L^p(\widehat{T})} \\
&\quad \leq c|\det(A_{\widetilde{T}})|^{-\frac{1}{p}} \|\widetilde{A}\|_2^m \sum_{|\gamma|=\ell-m} \sum_{|\beta|=m} (h^{-\beta}) h^{\beta} \sum_{|\epsilon|=|\gamma|} \widetilde{\mathscr{H}}^{\varepsilon} |\partial_{\tilde{x}}^{\epsilon}\tilde{\varphi}|_{W^{m,p}(\widetilde{T})} \\
&\quad \leq c|\det(A_{\widetilde{T}})|^{-\frac{1}{p}} \|\widetilde{A}\|_2^m \sum_{|\epsilon|=\ell-m} \widetilde{\mathscr{H}}^{\varepsilon} |\partial_{\tilde{x}}^{\epsilon}\tilde{\varphi}|_{W^{m,p}(\widetilde{T})}.
\end{aligned} \tag{14.6}$$

From (6.8), (6.9), (14.4), and (14.6), we have

$$|\varphi - I_T\varphi|_{W^{m,q}(T)} \leq c|T|_d^{\frac{1}{q}-\frac{1}{p}} \left(\frac{H_T}{h_T}\right)^m \sum_{|\varepsilon|=\ell-m} \widetilde{\mathscr{H}}^{\varepsilon} |\partial_{\tilde{x}}^{\epsilon}\tilde{\varphi}|_{W^{m,p}(\widetilde{T})},$$

which is the inequality (14.2) using $\widetilde{T} = \Phi_T^{-1}(T)$ and $\tilde{\varphi} = \varphi \circ \Phi_T$. □

14.2 Global Interpolation Error Estimates

A global interpolation operator I_h is constructed as follows (e.g., see [20, Section 1.4.2]). Its domain is defined by

$$D(I_h) := \{\varphi \in L^1(\Omega);\ \varphi|_T \in V(T),\ \forall T \in \mathbb{T}_h\}.$$

For $T \in \mathbb{T}_h$ and $\varphi \in D(I_h)$, the quantities $\chi_i(\varphi|_T)$ are meaningful on all the mesh elements and $1 \leq i \leq n_0$. The global interpolation $I_h\varphi$ can be specified elementwise using the local interpolation operators, that is,

$$(I_h\varphi)|_T := I_T(\varphi|_T) = \sum_{i=1}^{n_0} \chi_i(\varphi|_T)\theta_i \quad \forall T \in \mathbb{T}_h, \quad \forall \varphi \in D(I_h).$$

The global interpolation operator $I_h : D(I_h) \to V_h$ is defined as

$$I_h : D(I_h) \ni \varphi \mapsto I_h\varphi := \sum_{T \in \mathbb{T}_h} \sum_{i=1}^{n_0} \chi_i(\varphi|_T)\theta_i \in V_h^n,$$

where V_h^n is defined as

$$V_h^n := \{\varphi_h \in L^1(\Omega)^n;\ \varphi_h|_T \in P,\ \forall T \in \mathbb{T}_h\}. \tag{14.7}$$

Corollary 14.2. Suppose that the assumptions of Theorem 14.1 are satisfied. We impose Condition 6.2. Let I_h be the corresponding global interpolation operator. It then holds that, for any $\varphi \in W^{\ell,p}(\Omega)$;

(I) if Condition 10.1 is not imposed,

$$|\varphi - I_h \varphi|_{W^{m,q}(\Omega)} \leq c \sum_{T \in \mathbb{T}_h} |T|_d^{\frac{1}{q}-\frac{1}{p}} \sum_{|\varepsilon|=\ell-m} h^{\varepsilon} |\partial_r^{\varepsilon} \varphi|_{W^{m,p}(T)}. \tag{14.8}$$

(II) if Condition 10.1 is imposed,

$$|\varphi - I_h \varphi|_{W^{m,q}(\Omega)} \leq c \sum_{T \in \mathbb{T}_h} |T|_d^{\frac{1}{q}-\frac{1}{p}} \sum_{|\varepsilon|=\ell-m} \widetilde{\mathscr{H}}^{\varepsilon} |\partial_{\tilde{x}}^{\varepsilon} (\varphi \circ \Phi_T)|_{W^{m,p}(\Phi_T^{-1}(T))}. \tag{14.9}$$

Proof. If Condition 10.1 is not imposed, then using the local interpolation error (14.1),

$$\begin{aligned}
|\varphi - I_h \varphi|^q_{W^{m,q}(\Omega)} &= \sum_{T \in \mathbb{T}} |\varphi - I_T \varphi|^q_{W^{m,q}(T)} \\
&\leq c \sum_{T \in \mathbb{T}} |T|_d^{q\left(\frac{1}{q}-\frac{1}{p}\right)} \left(\frac{H_T}{h_T} \right)^{qm} \left(\sum_{|\varepsilon|=\ell-m} h^{\varepsilon} |\partial_r^{\varepsilon} \varphi|_{W^{m,p}(T)} \right)^q,
\end{aligned}$$

which leads to the desired result together with (1.5) and Condition 6.2.

If Condition 10.1 is not imposed, then using the local interpolation error (14.2),

$$\begin{aligned}
|\varphi - I_h \varphi|^q_{W^{m,q}(\Omega)} &= \sum_{T \in \mathbb{T}} |\varphi - I_T \varphi|^q_{W^{m,q}(T)} \\
&\leq c \sum_{T \in \mathbb{T}} |T|_d^{q\left(\frac{1}{q}-\frac{1}{p}\right)} \left(\frac{H_T}{h_T} \right)^{qm} \left(\sum_{|\varepsilon|=\ell-m} \widetilde{\mathscr{H}}^{\varepsilon} |\partial_{\tilde{x}}^{\varepsilon} (\varphi \circ \Phi_T)|_{W^{m,p}(\Phi_T^{-1}(T))} \right)^q,
\end{aligned}$$

which leads to the desired result together with (1.5) and Condition 6.2. □

14.3 Examples of Anisotropic Elements

When $k = 1$, $ell = 2$, $m = 1$ and $q = p$ in (14.1) of Theorem 14.1, the estimate is written as

$$|\varphi - I_T \varphi|_{W^{1,p}(T)} \leq C_1^I \frac{H_T}{h_T} \sum_{i=1}^{d} h_i \left| \frac{\partial \varphi}{\partial r_i} \right|_{W^{1,p}(T)}. \tag{14.10}$$

Let $T \subset \mathbb{R}^2$ be a triangle. As described in Section 8.1, an isotropic mesh element has equal or nearly equal edge lengths and angles, resulting in a balanced shape. Then, the estimate (14.10) becomes

$$|\varphi - I_T \varphi|_{W^{1,p}(T)} \leq c h_T |\varphi|_{W^{2,p}(T)}. \tag{14.11}$$

We considered the following five anisotropic elements as in Section 8.2: Let $0 < s \ll 1$, $s \in \mathbb{R}$ and $\varepsilon, \delta, \gamma \in \mathbb{R}$.

Example 14.3 (Right-angled triangle)**.** Let $T \subset \mathbb{R}^2$ be the simplex with vertices $p_1 := (0,0)^\top$, $p_2 := (s,0)^\top$ and $p_3 := (0, s^{\varepsilon})^\top$ with $1 < \varepsilon$. We then have $h_1 = s$, $h_2 = s^{\varepsilon}$ and $h_T = \sqrt{s^2 + s^{2\varepsilon}}$; i.e.,

$$\frac{H_T}{h_T} = 2.$$

In this case, the estimate (14.10) becomes

$$|\varphi - I_T \varphi|_{W^{1,p}(T)} \leq 2 C_1^I \sum_{i=1}^{2} h_i \left| \frac{\partial \varphi}{\partial r_i} \right|_{W^{1,p}(T)},$$

which is the anisotropic interpolation error estimate.

Example 14.4 (Dagger)**.** Let $T \subset \mathbb{R}^2$ be the simplex with vertices $p_1 := (0,0)^\top$, $p_2 := (s,0)^\top$ and $p_3 := (s^\delta, s^\varepsilon)^\top$ with $1 < \varepsilon < \delta$. We then have $h_1 = \sqrt{(s - s^\delta)^2 + s^{2\varepsilon}}$, $h_2 = \sqrt{s^{2\delta} + s^{2\varepsilon}}$ and $h_T = s$; i.e.,

$$\frac{H_T}{h_T} = \frac{\sqrt{(s - s^\delta)^2 + s^{2\varepsilon}} \sqrt{s^{2\delta} + s^{2\varepsilon}}}{\frac{1}{2} s^{1+\varepsilon}} \leq c.$$

In this case, the estimate (14.10) becomes

$$|\varphi - I_T \varphi|_{W^{1,p}(T)} \leq c \sum_{i=1}^{2} h_i \left| \frac{\partial \varphi}{\partial r_i} \right|_{W^{1,p}(T)},$$

which is the anisotropic interpolation error estimate.

Example 14.5 (Blade)**.** Let $T \subset \mathbb{R}^2$ be the simplex with vertices $p_1 := (0,0)^\top$, $p_2 := (2s,0)^\top$ and $p_3 := (s, s^\varepsilon)^\top$ with $1 < \varepsilon$. We then have $h_1 = h_2 = \sqrt{s^2 + s^{2\varepsilon}}$ and $h_T = 2s$; i.e.,

$$\frac{H_T}{h_T} = \frac{s^2 + s^{2\varepsilon}}{s^{1+\varepsilon}} \to \infty \quad \text{as } s \to 0.$$

In this case, the estimate (14.10) becomes

$$|\varphi - I_T \varphi|_{W^{1,p}(T)} \leq c s^{2-\varepsilon} |\varphi|_{W^{2,p}(T)}.$$

When $\varepsilon > 2$, this implies that the estimate diverges as $s \to 0$.

Example 14.6 (Dagger)**.** Let $T \subset \mathbb{R}^2$ be the simplex with vertices $p_1 := (0,0)^\top$, $p_2 := (s,0)^\top$ and $p_3 := (s^\delta, s^\varepsilon)^\top$ with $1 < \delta < \varepsilon$. We then have $h_1 = \sqrt{(s - s^\delta)^2 + s^{2\varepsilon}}$, $h_2 = \sqrt{s^{2\delta} + s^{2\varepsilon}}$ and $h_T = s$; i.e.,

$$\frac{H_T}{h_T} = \frac{\sqrt{(s - s^\delta)^2 + s^{2\varepsilon}} \sqrt{s^{2\delta} + s^{2\varepsilon}}}{\frac{1}{2} s^{1+\varepsilon}} \leq c s^{\delta - \varepsilon} \to \infty \quad \text{as } s \to 0.$$

In this case, the estimate (14.10) becomes

$$\begin{aligned} |\varphi - I_T \varphi|_{W^{1,p}(T)} &\leq c s^{\delta - \varepsilon} \left(s \left| \frac{\partial \varphi}{\partial r_1} \right|_{W^{1,p}(T)} + s^\delta \left| \frac{\partial \varphi}{\partial r_2} \right|_{W^{1,p}(T)} \right) \\ &\leq c s^{1 + \delta - \varepsilon} |\varphi|_{W^{2,p}(T)}. \end{aligned}$$

When $\varepsilon - \delta > 1$, this implies that the estimate diverges as $s \to 0$.

Example 14.7 (Right-angled triangle)**.** Let $T \subset \mathbb{R}^2$ be the simplex with vertices $p_1 := (0,0)^\top$, $p_2 := (s,0)^\top$ and $p_3 := (0, \delta s)^\top$ with $\delta \ll 1$. We then have $h_1 = s$, $h_2 = \delta s$ and $h_T = s\sqrt{1 + \delta^2}$; i.e.,

$$\frac{H_T}{h_T} = 2.$$

In this case, the estimate (14.10) becomes

$$|\varphi - I_T \varphi|_{W^{1,p}(T)} \leq c \sum_{i=1}^{2} h_i \left| \frac{\partial \varphi}{\partial r_i} \right|_{W^{1,p}(T)},$$

which is the anisotropic interpolation error estimate.

Example 14.8 (Blade)**.** Let $T \subset \mathbb{R}^2$ be the simplex with vertices $p_1 := (0,0)^\top$, $p_2 := (2s,0)^\top$ and $p_3 := (s, \delta s)^\top$ with $\delta \ll 1$. We then have $h_1 = h_2 = s\sqrt{1+\delta^2}$ and $h_T = 2s$; i.e.,

$$\frac{H_T}{h_T} = \frac{s^2(1+\delta^2)}{\delta s^2} \leq \frac{c}{\delta},$$

In this case, the estimate (14.10) becomes

$$|\varphi - I_T \varphi|_{W^{1,p}(T)} \leq \frac{c}{\delta} s |\varphi|_{W^{2,p}(T)}.$$

This implies that the estimate converges as $s \to 0$ and the error may be large. Thus, even if anisotropic mesh partitioning is used, it is unlikely to improve calculation efficiency.

14.4 Examples that do not satisfy conditions (12.2) in Theorem 12.1

The following lemma ([4, Lemma 4], [3, Lemma 2.3]) gives a criterion for the existence of linear functionals satisfying conditions (12.2b) and (12.2c).

Lemma 14.9. Let $\mathbb{P}$ be an arbitrary polynomial space and β be a multi-index. We set $j := \dim(\partial^\beta \mathbb{P})$. Assume that $I : \mathcal{C}^\mu(\widehat{T}) \to \mathbb{P}$, $\mu \in \mathbb{N}$, is a linear operator with $I\hat{\eta} = \hat{\eta}$ $\forall \hat{\eta} \in \mathbb{P}$. Then, there exist linear functionals $\mathscr{F}_i : \mathcal{C}^\infty(\widehat{T}) \to \mathbb{R}$, $i = 1, \ldots, j$, such that

$$\mathscr{F}_i(\partial^\beta(\hat{\varphi} - I\hat{\varphi})) = 0 \quad \forall i = 1, \ldots, j, \quad \forall \hat{\varphi} \in \mathcal{C}^\infty(\widehat{T}), \tag{14.12}$$

$$\hat{\eta} \in \mathbb{P}, \quad \mathscr{F}_i(\partial^\beta \hat{\eta}) = 0 \quad \forall i = 1, \ldots, j \quad \Rightarrow \quad \partial^\beta \hat{\eta} = 0 \tag{14.13}$$

if and only if the condition

$$\hat{\varphi} \in \mathcal{C}^\infty(\widehat{T}), \quad \partial^\beta \hat{\varphi} = 0 \quad \Rightarrow \quad \partial^\beta I \hat{\varphi} = 0 \tag{14.14}$$

holds.

Proof. A proof can be found in [4, Lemma 4]. □

If Condition (14.14) is violated, estimate (12.3) does not hold. This means that one cannot obtain the estimate (14.1), which is sharper than (11.2).

The following are examples that do not satisfy (14.14). Let $\widehat{T} \subset \mathbb{R}^2$ be the reference element with vertices $\widehat{p}_1 := (0,0)^\top$, $\widehat{p}_2 := (1,0)^\top$, $\widehat{p}_3 := (0,1)^\top$. We set $\widehat{p}_4 := (1/3, 1/3)^\top$. We define the barycentric coordinates $\lambda_i : \mathbb{R}^2 \to \mathbb{R}$, $i = 1, \ldots, 3$, on the reference element as

$$\lambda_1 := 1 - \hat{x}_1 - \hat{x}_2, \quad \lambda_2 := \hat{x}_1, \quad \lambda_3 := \hat{x}_2, \quad (\hat{x}_1, \hat{x}_2)^\top \in \widehat{T}.$$

Example 14.10 ($\mathbb{P}^1$ + bubble Finite Element)**.** As mentioned in Example 11.10, we define the local basis functions as

$$\begin{aligned} \theta_4(x) &:= 27\lambda_1(x)\lambda_2(x)\lambda_3(x), \\ \theta_i(x) &:= \lambda_i(x) - \frac{1}{3}\theta_4(x), \quad i = 1, 2, 3. \end{aligned}$$

The interpolation operator I_T^b defined by

$$I^b : \mathcal{C}(\widehat{T}) \ni \hat{\varphi} \mapsto I^b \hat{\varphi} := \sum_{i=1}^{4} \hat{\varphi}(\widehat{P}_i)\theta_i \in \operatorname{span}\{\theta_1, \theta_2, \theta_3, \theta_4\}.$$

Let $\beta = (1,0)$. Setting $\hat{\varphi}(\hat{x}_1,\hat{x}_2) := \hat{x}_2^2$, we have $\frac{\partial \hat{\varphi}}{\partial \hat{x}_1} = 0$. By simple calculation, we obtain

$$\begin{aligned}
\frac{\partial}{\partial \hat{x}_1} I^b \hat{\varphi} &= \hat{\varphi}(\widehat{p}_1)\frac{\partial \theta_1}{\partial \hat{x}_1} + \hat{\varphi}(\widehat{p}_2)\frac{\partial \theta_1}{\partial \hat{x}_1} + \hat{\varphi}(\widehat{p}_3)\frac{\partial \theta_3}{\partial \hat{x}_1} + \hat{\varphi}(\widehat{p}_4)\frac{\partial \theta_4}{\partial \hat{x}_1} \\
&= \frac{\partial \theta_3}{\partial \hat{x}_1} + \frac{1}{3^2}\frac{\partial \theta_4}{\partial \hat{x}_1} = -\frac{1}{3}\frac{\partial \theta_4}{\partial \hat{x}_1} + \frac{1}{3^2}\frac{\partial \theta_4}{\partial \hat{x}_1} \not\equiv 0.
\end{aligned}$$

Therefore, the condition (14.14) is not satisfied. This implies that the error estimate (12.3) on the reference element does not hold for the $\mathcal{P}^1$ + bubble finite element.

Example 14.11 ($\mathbb{P}^3$ Hermite Finite Element)**.** Following [17, Theorem 2.2.8], we define the Hermite interpolation operator $I^H : H^3(T) \to \mathbb{P}^3$ as

$$\begin{aligned}
I^H \hat{\varphi} &:= \sum_{i=1}^{3} \left(-2\lambda_i^3 + 3\lambda_i^2 - 7\lambda_i \sum_{1 \le j < k \le 3, j \ne i, k \ne i} \lambda_j \lambda_k \right) \hat{\varphi}(\widehat{p}_i) + 27\lambda_1\lambda_2\lambda_3 \hat{\varphi}(\widehat{p}_4) \\
&\quad + \sum_{i=1}^{3} \left(\sum_{j=1}^{3} \lambda_i \lambda_j (2\lambda_i + \lambda_j - 1)(\widehat{p}_j^{(1)} - \widehat{p}_i^{(1)}) \right) \frac{\partial \hat{\varphi}}{\partial \hat{x}_1}(\widehat{p}_i) \\
&\quad + \sum_{i=1}^{3} \left(\sum_{j=1}^{3} \lambda_i \lambda_j (2\lambda_i + \lambda_j - 1)(\widehat{p}_j^{(2)} - \widehat{p}_i^{(2)}) \right) \frac{\partial \hat{\varphi}}{\partial \hat{x}_2}(\widehat{p}_i),
\end{aligned}$$

where $\widehat{p}_i^{(k)}$, $1 \le k \le 2$, are the components of a point $\widehat{p}_i \in \mathbb{R}^2$. Let $\beta = (1,0)$. Setting $\hat{\varphi}(\hat{x}_1,\hat{x}_2) := \hat{x}_2^4$, we have $\frac{\partial \hat{\varphi}}{\partial \hat{x}_1} = 0$. Furthermore, by a simple calculation, i.e.,

$$\begin{aligned}
\frac{\partial}{\partial \hat{x}_1}(\lambda_1\lambda_2\lambda_3) &= -\hat{x}_2^2 - 2\hat{x}_1\hat{x}_2 - \hat{x}_2, \\
\frac{\partial}{\partial \hat{x}_1}\{\lambda_3\lambda_1(2\lambda_3 + \lambda_1 - 1)\} &= -\hat{x}_2 + 2\hat{x}_1\hat{x}_2, \\
\frac{\partial}{\partial \hat{x}_1}\{\lambda_3\lambda_2(2\lambda_3 + \lambda_2 - 1)\} &= -\hat{x}_2 + 2\hat{x}_1\hat{x}_2 + 2\hat{x}_2^2,
\end{aligned}$$

we obtain

$$\begin{aligned}
\frac{\partial}{\partial \hat{x}_1} I^H \hat{\varphi} &= \frac{\partial}{\partial \hat{x}_1}\left(-2\lambda_3^3 + 3\lambda_3^2 - 7\lambda_3 \sum_{1 \le j < k \le 3, j \ne 3, k \ne 3} \lambda_j \lambda_k \right) \hat{\varphi}(\widehat{p}_3) \\
&\quad + 27\frac{\partial}{\partial \hat{x}_1}(\lambda_1\lambda_2\lambda_3)\hat{\varphi}(\widehat{p}_4) \\
&\quad + \frac{\partial}{\partial \hat{x}_1}\left(\sum_{j=1}^{3} \lambda_3\lambda_j(2\lambda_3 + \lambda_j - 1)(\widehat{p}_j^{(2)} - \widehat{p}_3^{(2)}) \right) \frac{\partial \hat{\varphi}}{\partial \hat{x}_2}(\widehat{p}_3) \\
&= -7\frac{\partial}{\partial \hat{x}_1}(\lambda_1\lambda_2\lambda_3)\hat{\varphi}(\widehat{p}_3) + 27\frac{\partial}{\partial \hat{x}_1}(\lambda_1\lambda_2\lambda_3)\hat{\varphi}(\widehat{p}_4) \\
&\quad + \frac{\partial}{\partial \hat{x}_1}\left\{ \lambda_3\lambda_1(2\lambda_3 + \lambda_1 - 1)(\widehat{p}_1^{(2)} - \widehat{p}_3^{(2)}) \right\} \frac{\partial \hat{\varphi}}{\partial \hat{x}_2}(\widehat{p}_3) \\
&\quad + \frac{\partial}{\partial \hat{x}_1}\left\{ \lambda_3\lambda_2(2\lambda_3 + \lambda_2 - 1)(\widehat{p}_2^{(2)} - \widehat{p}_3^{(2)}) \right\} \frac{\partial \hat{\varphi}}{\partial \hat{x}_2}(\widehat{p}_3) \\
&= -7(-\hat{x}_2^2 - 2\hat{x}_1\hat{x}_2 - \hat{x}_2) + \frac{1}{3}(-\hat{x}_2^2 - 2\hat{x}_1\hat{x}_2 - \hat{x}_2) \\
&\quad + 8(\hat{x}_2 - 2\hat{x}_1\hat{x}_2 - \hat{x}_2^2) \not\equiv 0.
\end{aligned}$$

Here, we used

$$\hat{\varphi}(\widehat{p}_i) = 0, \quad \frac{\partial \hat{\varphi}}{\partial \hat{x}_2}(\widehat{p}_i) = 0, \quad i = 1, 2,$$
$$\hat{\varphi}(\widehat{p}_3) = 1, \quad \hat{\varphi}(\widehat{p}_4) = \frac{1}{3^4}, \quad \frac{\partial \hat{\varphi}}{\partial \hat{x}_2}(\widehat{p}_3) = 4,$$
$$\widehat{p}_1^{(2)} - \widehat{p}_3^{(2)} = -1, \quad \widehat{p}_2^{(2)} - \widehat{p}_3^{(2)} = -1.$$

Therefore, Condition (14.14) is not satisfied. This implies that error estimate (12.3) on the reference element does not hold for Hermitian finite elements.

14.5 Effect of the quantity $|T|_d^{\frac{1}{q}-\frac{1}{p}}$ in the interpolation error estimates for $d = 2, 3$

We consider the effect of the factor $|T|_d^{\frac{1}{q}-\frac{1}{p}}$.

14.5.1 Case that $q > p$

When $q > p$, the factor may affect the convergence order. In particular, the interpolation error estimate may diverge on anisotropic mesh partitions.

Let $T \subset \mathbb{R}^2$ be the triangle with vertices $p_1 := (0,0)^\top$, $p_2 := (s,0)^\top$, $p_3 := (0, s^\varepsilon)^\top$ for $0 < s \ll 1$, $\varepsilon \geq 1$, $s \in \mathbb{R}$ and $\varepsilon \in \mathbb{R}$. Then,

$$\frac{h_{\max}}{h_{\min}} = s^{1-\varepsilon}, \quad \frac{H_T}{h_T} = 2, \quad |T|_2 = \frac{1}{2} s^{1+\varepsilon}.$$

Let $k = 1$, $\ell = 2$, $m = 1$, $q = 2$, and $p \in (1, 2)$. Then, $W^{1,p}(T) \hookrightarrow L^2(T)$ and Theorem 14.1 lead to

$$|\varphi - I_T \varphi|_{H^1(T)} \leq c s^{-(1+\varepsilon)\frac{2-p}{2p}} \left(s \left| \frac{\partial \varphi}{\partial r_1} \right|_{W^{1,p}(T)} + s^\varepsilon \left| \frac{\partial \varphi}{\partial r_2} \right|_{W^{1,p}(T)} \right).$$

When $\varepsilon = 1$ (the case of the isotropic element), we get

$$|\varphi - I_T \varphi|_{H^1(T)} \leq c h_T^{\frac{2(p-1)}{p}} |\varphi|_{W^{2,p}(T)}, \quad \frac{2(p-1)}{p} > 0.$$

However, when $\varepsilon > 1$ (the case of the anisotropic element), the estimate may diverge as $s \to 0$. Therefore, if $q > p$, the convergence order of the interpolation operator may deteriorate.

We next set $m = 0$, $\ell = 2$, $q = \infty$, and $p = 2$. Let

$$\varphi(x, y) := x^2 + y^2.$$

Let $I_T^L : \mathcal{C}^0(T) \to \mathbb{P}^1$ be the local Lagrange interpolation operator. For any nodes p_i of T, because $I_T^L \varphi(p_i) = \varphi(p_i)$, we have

$$I_T^L \varphi(x, y) = s x + s^\varepsilon y.$$

It thus holds that

$$(\varphi - I_T^L \varphi)(x, y) = \left(x - \frac{s}{2} \right)^2 + \left(y - \frac{s^\varepsilon}{2} \right)^2 - \frac{1}{4}(s^2 + s^{2\varepsilon}).$$

We therefore have, because $H^2(T) \hookrightarrow L^\infty(T)$,

$$\|\varphi - I_T^L \varphi\|_{L^\infty(T)} = \frac{1}{4}(s^2 + s^{2\varepsilon}), \quad \sum_{|\gamma|=2} \widetilde{\mathscr{H}}^\gamma \|\partial_x^\gamma \varphi\|_{L^2(T)} = 2|T|_2^{\frac{1}{2}}(s^2 + s^{2\varepsilon}),$$

and thus,

$$\frac{\|\varphi - I_T^L \varphi\|_{L^\infty(T)}}{|T|_2^{-\frac{1}{2}} \sum_{|\gamma|=2} \widetilde{\mathscr{H}}^\gamma \|\partial_x^\gamma \varphi\|_{L^2(T)}} = \frac{1}{8}.$$

This example implies that the convergence order is not optimal, but the estimate converges on anisotropic meshes.

14.5.2 Case that $q < p$

We consider Theorem 14.1. Let $I_T^L : \mathcal{C}(T) \to \mathbb{P}^k$ ($k \in \mathbb{N}$) be the local Lagrange interpolation operator. Let $\varphi \in W^{\ell,\infty}(T)$ be such that $\ell \in \mathbb{N}$, $2 \le \ell \le k+1$. It then holds that, for any $m \in \{0, \dots, \ell - 1\}$ and $q \in [1, \infty]$,

$$|\varphi - I_T^L \varphi|_{W^{m,q}(T)} \le c|T|_d^{\frac{1}{q}} \left(\frac{H_T}{h_T} \right)^m \sum_{|\gamma|=\ell-m} h^\gamma |\partial_r^\gamma \varphi|_{W^{m,\infty}(T)}. \tag{14.15}$$

The convergence order is therefore improved by $|T|_d^{\frac{1}{q}}$. We do numerical tests to confirm this. Let $k = 1$ and

$$\varphi(x, y, z) := x^2 + \frac{1}{4} y^2 + z^2.$$

Let $s := \frac{1}{N}$, $N \in \mathbb{N}$ and $\varepsilon \in \mathbb{R}$, $1 < \varepsilon$. We compute the convergence order with respect to the H^1 norm defined by

$$Err_s^\varepsilon(H^1) := |\varphi - I_T^L \varphi|_{H^1(T)}.$$

The convergence indicator r is defined by

$$r = \frac{1}{\log(2)} \log \left(\frac{Err_s^\varepsilon(H^1)}{Err_{s/2}^\varepsilon(H^1)} \right).$$

(I) Let $T \subset \mathbb{R}^3$ be the simplex with vertices $p_1 := (0,0,0)^\top$, $p_2 := (s,0,0)^\top$, $p_3 := (0, s^\varepsilon, 0)^\top$, and $p_4 := (0,0,s^\delta)^\top$ ($1 < \delta \le \varepsilon$), and $0 < s \ll 1$, $s \in \mathbb{R}$. We then have $h_1 = \sqrt{s^2 + s^{2\varepsilon}}$, $h_2 = s^\varepsilon$ and $h_3 := \sqrt{s^{2\varepsilon} + s^{2\delta}}$; i.e.,

$$\frac{h_{\max}}{h_{\min}} \le c s^{1-\varepsilon}, \quad \frac{H_T}{h_T} \le c.$$

From (14.15) with $m = 1$, $\ell = 2$, and $q = 2$, because $|T|_3 \approx s^{1+\varepsilon+\delta}$, we have the estimate

$$|\varphi - I_T^L \varphi|_{H^1(T)} \le c h_T^{\frac{3+\varepsilon+\delta}{2}}.$$

Computational results are for the case that $\varepsilon = 3.0$ and $\delta = 2.0$ (Table 13).

Table 13: Error of the local interpolation operator ($\varepsilon = 3.0, \delta = 2.0$)

N	s	$Err_s^{3.0}(H^1)$	r
64	1.5625e-02	2.4336e-08	
128	7.8125e-03	1.5209e-09	4.00
256	3.9062e-03	9.5053e-11	4.00

(II) Let $T \subset \mathbb{R}^3$ be the simplex with vertices $p_1 := (0,0,0)^\top$, $p_2 := (s,0,0)^\top$, $p_3 := (\frac{s}{2}, s^\varepsilon, 0)^\top$, and $p_4 := (0,0,s)^\top$ ($1 < \varepsilon \leq 6$) and $0 < s \ll 1$, $s \in \mathbb{R}$. We then have $h_1 = s$, $h_2 = \sqrt{s^2/4 + s^{2\varepsilon}}$ and $h_3 := s$; i.e.,

$$\frac{h_{\max}}{h_{\min}} = \frac{t}{\sqrt{s^2/4 + t^{2\varepsilon}}} \leq c, \quad \frac{H_T}{h_T} \leq cs^{1-\varepsilon}.$$

From (14.15) with $m = 1$, $\ell = 2$, and $q = 2$, because $|T|_3 \approx s^{2+\varepsilon}$, we have the estimate

$$|\varphi - I_T^L \varphi|_{H^1(T)} \leq c h_T^{3-\frac{\varepsilon}{2}}.$$

Computational results are for the cases that $\varepsilon = 3.0, 6.0$ (Table 14).

Table 14: Error of the local interpolation operator ($\varepsilon = 3.0, 6.0$)

N	s	$Err_s^{3.0}(H^1)$	r	$Err_s^{6.0}(H^1)$	r
64	1.5625e-02	1.9934e-04		1.0206e-01	
128	7.8125e-03	7.0477e-05	1.50	1.0206e-01	0
256	3.9062e-03	2.4917e-05	1.50	1.0206e-01	0

14.6 What happens if violating the maximum-angle condition?

This subsection introduces two negative points by violating the maximum-angle condition. One is that it is practically disadvantageous. As an example, let $T \subset \mathbb{R}^2$ be the triangle with vertices $p_1 := (0,0)^\top$, $p_2 := (s,0)^\top$, $p_3 := (s/2, s^\varepsilon)^\top$ for $0 < s \ll 1$, $\varepsilon \geq 1$, $s \in \mathbb{R}$ and $\varepsilon \in \mathbb{R}$. From Theorem 14.1 with $k = 1$, $\ell = 2$, $m = 1$, $p = q = 2$, we have

$$|\varphi - I_T \varphi|_{H^1(T)} \leq cs^{2-\varepsilon} \left| \frac{\partial \varphi}{\partial r_1} \right|_{H^1(T)} + s \left| \frac{\partial \varphi}{\partial r_2} \right|_{H^1(T)}.$$

Even if one wants to reduce the step size in a specific direction (y-axis direction), the interpolation error may diverge as $s \to 0$ when $\varepsilon > 2$. This loses the benefits of using anisotropic meshes.

Another is that violating the condition makes it challenging to show mathematical validity in the finite element method. One of the answers can be found in [5]. That is, the maximum-angle condition is sufficient to do numerical calculations safely.

15 Lagrange Interpolation Error Estimates

15.1 One-dimensional Lagrange Interpolation

Let $\Omega := (0,1) \subset \mathbb{R}$. For $N \in \mathbb{N}$, let $\mathbb{T}_h = \{0 = x_0 < x_1 < \cdots < x_N < x_{N+1} = 1\}$ be a mesh of $\overline{\Omega}$ such as

$$\overline{\Omega} := \bigcup_{i=1}^{N} I_i, \quad \operatorname{int} I_i \cap \operatorname{int} I_j = \emptyset \quad \text{for } i \neq j,$$

where $I_i := [x_i, x_{i+1}]$ for $0 \le i \le N$. We denote $h_i := x_{i+1} - x_i$ for $0 \le i \le N$. For $\widehat{T} := [0,1] \subset \mathbb{R}$ and $\widehat{P} := \mathbb{P}^k$ with $k \in \mathbb{N}$, let $\{\widehat{T}, \widehat{P}, \widehat{\Sigma}\}$ be the reference Lagrange finite element, e.g., see [20]. The corresponding interpolation operator is defined as

$$I_{\widehat{T}}^k : \mathcal{C}(\widehat{T}) \ni \hat{v} \mapsto I_{\widehat{T}}^k(\hat{v}) := \sum_{m=0}^{k} \hat{v}(\hat{\xi}_m)\widehat{\mathcal{L}}_m^k,$$

where $\hat{\xi}_m := \frac{m}{k}$ and $\{\widehat{\mathcal{L}}_0^k, \dots, \widehat{\mathcal{L}}_k^k\}$ is the Lagrange polynomials associated with the nodes $\{\hat{\xi}_0, \dots, \hat{\xi}_k\}$. For $i \in \{0, \dots, N\}$, we consider the affine transformations

$$\Phi_i : \widehat{T} \ni t \mapsto x = x_i + t h_i \in I_i.$$

For $\hat{v} \in \mathcal{C}(\widehat{T})$, we set $\hat{v} = v \circ \Phi_i$.

Theorem 15.1. Let $1 \le p \le \infty$ and assume that there exists a nonnegative integer k such that

$$\mathcal{P}^k = \widehat{P} \subset W^{k+1,p}(\widehat{T}) \subset \mathcal{C}(\widehat{T}).$$

Let ℓ $(0 \le \ell \le k)$ be such that $W^{\ell+1,p}(\widehat{T}) \subset \mathcal{C}(\widehat{T})$ with continuous embedding. Furthermore, assume that $\ell, m \in \mathbb{N} \cup \{0\}$ and $p, q \in [1, \infty]$ such that $0 \le m \le \ell + 1$ and

$$W^{\ell+1,p}(\widehat{T}) \hookrightarrow W^{m,q}(\widehat{T}).$$

It then holds that, for any $v \in W^{\ell+1,p}(I_i)$ with $\hat{v} = v \circ \Phi_i$,

$$|v - I_{I_i}^k v|_{W^{m,q}(I_i)} \le c h_i^{\frac{1}{q} - \frac{1}{p} + \ell + 1 - m} |v|_{W^{\ell+1,p}(I_i)}. \tag{15.1}$$

Proof. We only show the outline of the proof. Scaling argument yields

$$|v - I_{I_i}^k v|_{W^{m,q}(I_i)} = h_i^{-m+\frac{1}{q}} |\hat{v} - I_{\widehat{T}}\hat{v}|_{W^{m,q}(\widehat{T})},$$
$$|\hat{v}|_{W^{\ell+1,p}(\widehat{T})} = h_i^{\ell+1-\frac{1}{p}} |v|_{W^{\ell+1,p}(I_i)}.$$

Using the Sobolev embedding theorem and the Bramble–Hilbert–type lemma, we have

$$|\hat{v} - I_{\widehat{T}}\hat{v}|_{W^{m,q}(\widehat{T})} \le c|\hat{v}|_{W^{\ell+1,p}(\widehat{T})}.$$

Therefore, we obtain the estimate (15.1). □

Remark 15.2. The assumptions of Theorem 15.1 are standard; that is, there is no need to show the existence of functionals such as Theorem 14.1. Furthermore, the quantity $h_{\max}/h_{\min}$ that deteriorates the convergent order does not appear in (15.1).

Remark 15.3. If we set $x_j := \frac{j}{N+1}$, $j = 0, 1, \dots, N, N+1$, the mesh $\mathbb{T}_h$ is said to be the uniform mesh. If we set $x_j := g\left(\frac{j}{N+1}\right)$, $j = 1, \dots, N, N+1$ with a grading function g, the mesh $\mathbb{T}_h$ is said to be the graded mesh with respect to $x = 0$, see [8]. In particular, when one sets $g(y) := y^{\varepsilon}$ $(\varepsilon > 0)$, the mesh is called the radical mesh.

Remark 15.4 (Optimal order)**.** If $p = q$, it is possible to have the optimal error estimates even if the scale is different for each element. In the one-dimensional case, when $q > p$, the convergence order of the interpolation operator may deteriorate, see Section 14.5.1.

15.2 Lagrange Finite Element

Let $\widehat{T} \subset \mathbb{R}^d$ be the reference element defined in Sections 5.1 and 5.1. Let α be a multi-index. For $k \in \mathbb{N}$, we define the set of Lagrange nodes as

$$\mathscr{P} := \{\widehat{p}_i\}_{i=1}^{N^{(2,k)}} := \left\{ \left(\frac{i_1}{k}, \frac{i_2}{k} \right)^\top \in \mathbb{R}^2 \right\}_{0 \le i_1 + i_2 \le k} = \left\{ \frac{1}{k}\alpha \in \mathbb{R}^2 \right\}_{|\alpha| \le k}, \quad \text{if } d = 2,$$

$$\mathscr{P} := \{\widehat{p}_i\}_{i=1}^{N^{(3,k)}} := \widehat{T} \cap \left\{ \left(\frac{i_1}{k}, \frac{i_2}{k}, \frac{i_3}{k} \right)^\top \in \mathbb{R}^3 \right\}_{0 \le i_1, i_2, i_3 \le k}, \quad \text{if } d = 3.$$

The Lagrange finite element on the reference element is defined by the triple $\{\widehat{T}, \widehat{P}, \widehat{\Sigma}\}$ as follows.

1. $\widehat{P} := \mathbb{P}^k(\widehat{T})$;

2. $\widehat{\Sigma}$ is a set $\{\hat{\chi}_i\}_{1 \le i \le N^{(d,k)}}$ of $N^{(d,k)}$ linear forms $\{\hat{\chi}_i\}_{1 \le i \le N^{(d,k)}}$ with its components such that, for any $\hat{q} \in \widehat{P}$,

$$\hat{\chi}_i(\hat{q}) := \hat{q}(\widehat{p}_i) \quad \forall i \in \{1, \ldots, N^{(d,k)}\}. \tag{15.2}$$

The nodal basis functions associated with the degrees of freedom by (15.2) are defined as

$$\hat{\theta}_i(\widehat{p}_j) = \delta_{ij} \quad \forall i, j \in \{1, \ldots, N^{(d,k)}\}. \tag{15.3}$$

It then holds that $\hat{\chi}_i(\hat{\theta}_j) = \delta_{ij}$ for any $i, j \in \{1, \ldots, d+1\}$. Setting $V(\widehat{T}) := \mathcal{C}(\widehat{T})$ or $V(\widehat{T}) := W^{s,p}(\widehat{T})$ with $p \in [1, \infty]$ and $ps > d$ ($s \ge d$ if $p = 1$), the local operator $I_{\widehat{T}}^L$ is defined as

$$I_{\widehat{T}}^L : V(\widehat{T}) \ni \hat{\varphi} \mapsto I_{\widehat{T}}^L \hat{\varphi} := \sum_{i=1}^{N^{(d,k)}} \hat{\varphi}(\widehat{p}_i)\hat{\theta}_i \in \widehat{P}. \tag{15.4}$$

By analogous argument in Section 9, the Lagrange finite elements $\{\widetilde{T}, \widetilde{P}, \widetilde{\Sigma}\}$ and $\{T, P, \Sigma\}$ are constructed. The local shape functions are $\tilde{\theta}_i = \psi_{\widetilde{T}}^{-1}(\hat{\theta}_i)$ and $\theta_i = \psi_{\widetilde{T}}^{-1}(\tilde{\theta}_i)$ for any $i \in \{1, \ldots, N^{(d,k)}\}$, and the associated local interpolation operators are respectively defined as

$$I_{\widetilde{T}}^L : V(\widetilde{T}) \ni \tilde{\varphi} \mapsto I_{\widetilde{T}}^L \tilde{\varphi} := \sum_{i=1}^{N^{(d,k)}} \tilde{\varphi}(\widetilde{p}_i)\tilde{\theta}_i \in \widetilde{P}, \tag{15.5}$$

$$I_T^L : V(T) \ni \varphi \mapsto I_T^L \varphi := \sum_{i=1}^{N^{(d,k)}} \varphi(p_i)\theta_i \in P, \tag{15.6}$$

where $\tilde{p}_i = \Phi_{\widetilde{T}}(\hat{p}_i)$, $p_i = \Phi_T(\tilde{p}_i)$ for $i \in \{1, \ldots, N^{(d,k)}\}$.

15.3 Local Interpolation Error Estimates

We first introduce the following lemmata.

Lemma 15.5 ($d = 2$)**.** Let β be a multi-index with $m := |\beta|$ and $\hat{\varphi} \in \mathcal{C}(\widehat{T})$ a function such that $\partial_{\hat{x}}^\beta \hat{\varphi} \in W^{\ell-m,p}(\widehat{T})$, where $\ell, m \in \mathbb{N}_0$, $p \in [1, \infty]$ are such that $0 \le m \le \ell \le k+1$ and

$$p = \infty \quad \text{if } m = 0 \text{ and } \ell = 0, \tag{15.7a}$$

$$p > 2 \quad \text{if } m = 0 \text{ and } \ell = 1, \tag{15.7b}$$

$$m < \ell \quad \text{if } \beta_1 = 0 \text{ or } \beta_2 = 0, \text{ and } m > 0. \tag{15.7c}$$

Fix $q \in [1, \infty]$ such that $W^{\ell-m,p}(\widehat{T}) \hookrightarrow L^q(\widehat{T})$. Let $I_{\widehat{T}} := I_{\widehat{T}}^L$. It then holds that

$$\|\partial_{\hat{x}}^{\beta}(\hat{\varphi} - I_{\widehat{T}}^L \hat{\varphi})\|_{L^q(\widehat{T})} \leq c|\partial_{\hat{x}}^{\beta}\hat{\varphi}|_{W^{\ell-m,p}(\widehat{T})}. \tag{15.8}$$

Proof. We follow [3, Lemma 2.4]. We first give proofs in some particular cases: $k = 1, 2$.

Let $k = 1$. Let $m = 0$, that is, $\beta = (0, 0)$. We then have $j = \dim \mathbb{P}^1 = 3$. From the Sobolev embedding theorem (Theorem 1.6), we have $W^{\ell,p}(\widehat{T}) \subset \mathcal{C}^0(\widehat{T})$ with $1 \leq p \leq \infty$ and $2 < \ell p$. Under this condition, we use

$$\mathscr{F}_i(\hat{\varphi}) := \hat{\varphi}(\widehat{p}_i), \quad \hat{\varphi} \in W^{\ell,p}(\widehat{T}), \quad i = 1, \dots, 3.$$

It then holds that

$$|\mathscr{F}_i(\hat{\varphi})| \leq \|\hat{\varphi}\|_{\mathcal{C}^0(\widehat{T})} \leq c\|\hat{\varphi}\|_{W^{\ell,p}(\widehat{T})},$$

which means $\mathscr{F}_i \in W^{\ell,p}(\widehat{T})'$ for $i = 1, \dots, 3$, that is, (12.2a) is satisfied. Furthermore, we have

$$\mathscr{F}_i(I_{\widehat{T}}^L \hat{\varphi}) = (I_{\widehat{T}}^L \hat{\varphi})(\widehat{p}_i) = \hat{\varphi}(\widehat{p}_i) = \mathscr{F}_i(\hat{\varphi}), \quad i = 1, \dots, 3,$$

which satisfies (12.2b). For all $\hat{\eta} \in \mathbb{P}^1$, if $\mathscr{F}_i(\hat{\eta}) = 0$ for $i = 1, \dots, 3$, it obviously holds $\hat{\eta} = 0$. This means that (12.2c) is satisfied.

Let $m = 1$. We set $\beta = (1, 0)$. We then have $j = \dim(\partial^{\beta}\mathbb{P}^1) = 1$. We consider a functional

$$\mathscr{F}_1(\hat{\varphi}) := \int_0^1 \hat{\varphi}(\hat{x}_1, 0) d\hat{x}_1, \quad \hat{\varphi} \in W^{2,p}(\widehat{T}), \quad 1 < p.$$

We set $\widehat{I} := \{\hat{x} \in \widehat{T};\ \hat{x}_2 = 0\}$. The continuity is then shown by the trace theorem (e.g., see Theorem 1.7): if $1 = m < \ell$,

$$|\mathscr{F}_1(\hat{\varphi})| \leq \|\hat{\varphi}\|_{L^1(\widehat{I})} \leq c\|\hat{\varphi}\|_{W^{\ell-1,p}(\widehat{T})},$$

which means $\mathscr{F}_1 \in W^{\ell-1,p}(\widehat{T})'$, that is, (12.2a) is satisfied. Furthermore, it holds that

$$\begin{aligned}
\mathscr{F}_1(\partial^{(1,0)}(\hat{\varphi} - I_{\widehat{T}}^L \hat{\varphi})) &= \int_0^1 \frac{\partial}{\partial \hat{x}_1}(\hat{\varphi} - I_{\widehat{T}}^L \hat{\varphi})(\hat{x}_1, 0) d\hat{x}_1 \\
&= \left[\hat{\varphi} - I_{\widehat{T}}^L \hat{\varphi}\right]_{(0,0)}^{(1,0)} = 0,
\end{aligned}$$

which satisfy (12.2b). Let $\hat{\eta} := a\hat{x}_1 + b\hat{x}_2 + c$. We then have

$$\mathscr{F}_1(\partial^{(1,0)}\hat{\eta}) = a.$$

If $\mathscr{F}_1(\partial^{(1,0)}\hat{\eta}) = 0$, $a = 0$. This implies that $\partial^{(1,0)}\hat{\eta} = 0$. This means that (12.2c) is satisfied.

By analogous argument, the case $\beta = (0, 1)$ holds.

Let $k = 2$. Let $m = 0$, that is, $\beta = (0, 0)$. We then have $j = \dim \mathbb{P}^1 = 6$. Because $\dim \mathbb{P}^2 = 6$, we can show as in the case $k = 1$ and $\beta = (0, 0)$.

Let $\beta := (1, 0)$. We define three functionals as

$$\begin{aligned}
\mathscr{F}_1(\hat{\varphi}) &:= \int_0^{\frac{1}{2}} \hat{\varphi}(\hat{x}_1, 0) d\hat{x}_1, \\
\mathscr{F}_2(\hat{\varphi}) &:= \int_{\frac{1}{2}}^1 \hat{\varphi}(\hat{x}_1, 0) d\hat{x}_1, \\
\mathscr{F}_3(\hat{\varphi}) &:= \int_0^{\frac{1}{2}} \hat{\varphi}(\hat{x}_1, 1/2) d\hat{x}_1.
\end{aligned}$$

We then show (12.2a) and (12.2b) as above. Let $\hat{\eta} \in \mathbb{P}^2$ be such that

$$\mathscr{F}_i(\partial_{\hat{x}}^{\beta} \hat{\eta}) = 0, \quad i = 1, 2, 3. \tag{15.9}$$

We set the polynomial:

$$\begin{aligned} \hat{q} &:= \hat{\eta} - \hat{\eta}(1,0) \cdot 2 \left(\hat{x}_2 - \frac{1}{2} \right) (\hat{x}_2 - 1) - \hat{\eta} \left(\frac{1}{2}, \frac{1}{2} \right) \cdot [-4 \hat{x}_2 (\hat{x}_2 - 1)] \\ &\quad - \hat{\eta}(0,1) \cdot 2 \hat{x}_2 \left(\hat{x}_2 - \frac{1}{2} \right) \in \mathcal{P}^2. \end{aligned} \tag{15.10}$$

This has the following properties:

$$\partial_{\hat{x}} \hat{\eta} = \partial_{\hat{x}} \hat{q}, \quad \hat{q}(1,0) = \hat{q} \left(\frac{1}{2}, \frac{1}{2} \right) = \hat{q}(0,1) = 0. \tag{15.11}$$

We thus have

$$0 = \mathscr{F}_3(\partial_{\hat{x}}^{\beta} \hat{\eta}) = \mathscr{F}_3(\partial_{\hat{x}}^{\beta} \hat{q}) = \hat{q} \left(\frac{1}{2}, \frac{1}{2} \right) - \hat{q} \left(0, \frac{1}{2} \right),$$

hence, $\hat{q} \left(0, \frac{1}{2} \right) = 0$. By similar way,

$$\begin{aligned} 0 &= \mathscr{F}_2(\partial_{\hat{x}}^{\beta} \hat{\eta}) = \mathscr{F}_2(\partial_{\hat{x}}^{\beta} \hat{q}) = \hat{q}(1,0) - \hat{q} \left(\frac{1}{2}, 0 \right), \\ 0 &= \mathscr{F}_1(\partial_{\hat{x}}^{\beta} \hat{\eta}) = \mathscr{F}_1(\partial_{\hat{x}}^{\beta} \hat{q}) = \hat{q} \left(\frac{1}{2}, 0 \right) - \hat{q}(0,0), \end{aligned}$$

thus, $\hat{q} \left(\frac{1}{2}, 0 \right) = 0$ and $\hat{q}(0,0) = 0$. Therefore, $\hat{q} \equiv 0$. Together with (15.13), we have $\hat{q} = \hat{q}(\hat{x}_2)$, $\partial_{\hat{x}}^{\beta} \hat{\eta} = 0$. □

Lemma 15.6 ($d = 3$)**.** Let β be a multi-index with $m := |\beta|$ and $\hat{\varphi} \in \mathcal{C}(\widehat{T})$ a function such that $\partial_{\hat{x}}^{\beta} \hat{\varphi} \in W^{\ell - m, p}(\widehat{T})$, where $\ell, m \in \mathbb{N}_0$, $p \in [1, \infty]$ are such that $0 \le m \le \ell \le k+1$ and

$$\begin{aligned} & p = \infty \quad \text{if } m = 0 \text{ and } \ell = 0, && (15.12a) \\ & p > \frac{3}{\ell} \quad \text{if } m = 0 \text{ and } \ell = 1, 2, && (15.12b) \\ & m < \ell \quad \text{if } \beta_1 = 0, \ \beta_2 = 0, \text{ or } \beta_3 = 0, && (15.12c) \\ & p > 2 \quad \text{if } \beta \in \{(\ell - 1, 0, 0); (0, \ell - 1, 0); (0, 0, \ell - 1)\}. && (15.12d) \end{aligned}$$

Fix $q \in [1, \infty]$ such that $W^{\ell - m, p}(\widehat{T}) \hookrightarrow L^q(\widehat{T})$. Let $I_{\widehat{T}} := I_{\widehat{T}}^L$. It then holds that

$$\| \partial_{\hat{x}}^{\beta} (\hat{\varphi} - I_{\widehat{T}} \hat{\varphi}) \|_{L^q(\widehat{T})} \le c | \partial_{\hat{x}}^{\beta} \hat{\varphi} |_{W^{\ell - m, p}(\widehat{T})}. \tag{15.13}$$

Proof. A proof can be found in [3, Lemma 2.6]. □

We have the following new Lagrange interpolation error estimates.

Theorem 15.7. Let $\{\widehat{T}, \widehat{P}, \widehat{\Sigma}\}$ be the Lagrange finite element with $V(\widehat{T}) := \mathcal{C}(\widehat{T})$ and $\widehat{P} := \mathbb{P}^k(\widehat{T})$ with $k \ge 1$. Let $I_{\widehat{T}} := I_{\widehat{T}}^L$. Let $m \in \mathbb{N}_0$, $\ell \in \mathbb{N}$, and $p \in \mathbb{R}$ be such that $0 \le m \le \ell \le k+1$ and

$$d = 2: \begin{cases} p \in (2, \infty] & \text{if } m = 0, \ \ell = 1, \\ p \in [1, \infty] & \text{if } m = 0, \ \ell \ge 2 \text{ or } m \ge 1, \ \ell - m \ge 1, \end{cases}$$

$$d = 3: \begin{cases} p \in \left(\frac{3}{\ell}, \infty \right] & \text{if } m = 0, \ \ell = 1, 2, \\ p \in (2, \infty] & \text{if } m \ge 1, \ \ell - m = 1, \\ p \in [1, \infty] & \text{if } m = 0, \ \ell \ge 3 \text{ or } m \ge 1, \ \ell - m \ge 2. \end{cases}$$

Setting $q \in [1,\infty)$ be such that

$$W^{\ell-m,p}(\widehat{T}) \hookrightarrow L^q(\widehat{T}), \tag{15.14}$$

that is $(\ell - m) - \frac{d}{p} \geq -\frac{d}{q}$. Then, for all $\hat{\varphi} \in W^{\ell,p}(\widehat{T})$ with $\varphi := \hat{\varphi} \circ \Phi^{-1}$, we have

$$|\varphi - I_T^L \varphi|_{W^{m,q}(T)} \leq c|T|_d^{\frac{1}{q}-\frac{1}{p}} \left(\frac{H_T}{h_T} \right)^m \sum_{|\varepsilon|=\ell-m} h^{\varepsilon} |\partial_r^{\varepsilon} \varphi|_{W^{m,p}(T)}. \tag{15.15}$$

In particular, if Condition 10.1 is imposed, it holds that,for all $\hat{\varphi} \in W^{\ell,p}(\widehat{T})$ with $\varphi := \hat{\varphi} \circ \Phi^{-1}$,

$$|\varphi - I_T^L \varphi|_{W^{m,q}(T)} \leq C_2^I |T|_d^{\frac{1}{q}-\frac{1}{p}} \left(\frac{H_T}{h_T} \right)^m \sum_{|\varepsilon|=\ell-m} \widetilde{\mathscr{H}}^{\varepsilon} |\partial_{\tilde{x}}^{\varepsilon} (\varphi \circ \Phi_T)|_{W^{m,p}(\Phi_T^{-1}(T))}. \tag{15.16}$$

Furthermore, for any $\hat{\varphi} \in \mathcal{C}(\widehat{T})$ with $\varphi := \hat{\varphi} \circ \Phi^{-1}$, it holds that

$$\|\varphi - I_T^L \varphi\|_{L^\infty(T)} \leq c \|\varphi\|_{L^\infty(T)}.$$

Proof. Proved in a similar way to Theorem 14.1. □

15.4 Global Interpolation Error Estimates

Recall the space V_h^n with $n = 1$ (see (14.7)). We consider the space

$$V_h^L := \left\{ \varphi_h \in V_h^1 : \ [\![\varphi_h]\!]_F = 0 \ \forall F \in \mathcal{F}_h^i \right\} \subset H^1(\Omega). \tag{15.17}$$

We also define the global interpolation I_h^L to space V_h^L as

$$(I_h^L \varphi)|_T := I_T^L(\varphi|_T) \quad \forall T \in \mathbb{T}_h, \quad \forall \varphi \in \mathcal{C}(\overline{\Omega}).$$

Corollary 15.8. Suppose that the assumptions of Theorem 15.7 are satisfied. We impose Condition 6.2. Let I_h^L be the corresponding global Lagrange interpolation operator. It then holds that, for any $\varphi \in W^{\ell,p}(\Omega)$;

(I) if Condition 10.1 is not imposed,

$$|\varphi - I_h^L \varphi|_{W^{m,q}(\Omega)} \leq c \sum_{T \in \mathbb{T}_h} |T|_d^{\frac{1}{q}-\frac{1}{p}} \sum_{|\varepsilon|=\ell-m} h^{\varepsilon} |\partial_r^{\varepsilon} \varphi|_{W^{m,p}(T)}. \tag{15.18}$$

(II) if Condition 10.1 is imposed,

$$|\varphi - I_h^L \varphi|_{W^{m,q}(\Omega)} \leq c \sum_{T \in \mathbb{T}_h} |T|_d^{\frac{1}{q}-\frac{1}{p}} \sum_{|\varepsilon|=\ell-m} \widetilde{\mathscr{H}}^{\varepsilon} |\partial_{\tilde{x}}^{\varepsilon} (\varphi \circ \Phi_T)|_{W^{m,p}(\Phi_T^{-1}(T))}. \tag{15.19}$$

Proof. This corollary is proved in the same argument as Corollary 14.2. □

16 L^2-orthogonal Projection

This section considers error estimates of the L^2-orthogonal projection, e.g., for standard argument, see [20, Section 1.4.3] and [21, Section 11.5.3].

16.1 Finite Element

Let $k \in \mathbb{N}_0$. Let $\widehat{T} \subset \mathbb{R}^d$ be the reference element defined in Section 5.1. Let $\widehat{P}$ be a finite-dimensional space such that $\mathbb{P}^k \subset \widehat{P} \subset W^{k+1,\infty}(\widehat{T})$. The L^2-orthogonal projection onto $\widehat{P}$ is the linear operator $\Pi_{\widehat{T}}^k : L^1(\widehat{T}) \to \widehat{P}$ defined as

$$\int_{\widehat{T}} (\Pi_{\widehat{T}}^k \hat{\varphi} - \hat{\varphi}) \hat{q} d\hat{x} = 0 \quad \forall \hat{q} \in \widehat{P}, \quad \forall \hat{\varphi} \in L^1(\widehat{T}). \tag{16.1}$$

Because $\Pi_{\widehat{T}}^k \hat{\varphi} - \hat{\varphi}$ and $\Pi_{\widehat{T}}^k \hat{\varphi} - \hat{q}$ are L^2-orthogonal for any $\hat{q} \in \widehat{P}$, the Pythagorean identity yields

$$\|\hat{\varphi} - \hat{q}\|_{L^2(\widehat{T})}^2 = \|\hat{\varphi} - \Pi_{\widehat{T}}^k \hat{\varphi}\|_{L^2(\widehat{T})}^2 + \|\Pi_{\widehat{T}}^k \hat{\varphi} - \hat{q}\|_{L^2(\widehat{T})}^2.$$

This implies that

$$\Pi_{\widehat{T}}^k \hat{\varphi} = \arg\min_{\hat{q} \in \widehat{P}} \|\hat{\varphi} - \hat{q}\|_{L^2(\widehat{T})}.$$

Therefore, $\widehat{P}$ is pointwise invariant under $\Pi_{\widehat{T}}^k$.

Let $\Phi_{\widetilde{T}} : \widehat{T} \to \widetilde{T}$ and $\Phi_T : \widetilde{T} \to T$ be the two affine mappings defined in Section 5.2. For any $T \in \mathbb{T}_h$ with $\widetilde{T} = \Phi_{\widetilde{T}}(\widehat{T})$ and $T = \Phi_T(\widetilde{T})$, let $\hat{\varphi} := \tilde{\varphi} \circ \Phi_{\widetilde{T}}$ and $\tilde{\varphi} := \varphi \circ \Phi_T$. Furthermore, we set

$$\begin{aligned} \widetilde{P} &:= \{\psi_{\widetilde{T}}^{-1}(\hat{q});\ \hat{q} \in \widehat{P}\}, \\ P &:= \{\psi_{\widetilde{T}}^{-1}(\tilde{q});\ \tilde{q} \in \widetilde{P}\}. \end{aligned}$$

The L^2-orthognal projections onto $\widehat{P}$ and P are respectively the linear operators $\Pi_{\widetilde{T}}^k : L^1(\widetilde{T}) \to \widetilde{P}$ and $\Pi_T^k : L^1(T) \to P$ defined as

$$\begin{aligned} &\int_{\widetilde{T}} (\Pi_{\widetilde{T}}^k \tilde{\varphi} - \tilde{\varphi}) \tilde{q} d\tilde{x} = 0 \quad \forall \tilde{q} \in \widetilde{P}, \quad \forall \tilde{\varphi} \in L^1(\widetilde{T}), \\ &\int_{T} (\Pi_{T}^k \varphi - \varphi) q dx = 0 \quad \forall q \in P, \quad \forall \varphi \in L^1(T). \end{aligned}$$

Then, $\widetilde{P}$ and P are respectively pointwise invariant under $\Pi_{\widetilde{T}}^k$ and Π_T^k.

16.2 Local Interpolation Error Estimates

We have the following stability estimate of the projection $\Pi_{\widehat{T}}^k$.

Lemma 16.1. Let $q \in [1, \infty)$. It holds that

$$\|\Pi_{\widehat{T}}^k \hat{\varphi}\|_{L^q(\widehat{T})} \le c \|\hat{\varphi}\|_{L^q(\widehat{T})} \quad \forall \hat{\varphi} \in L^q(\widehat{T}). \tag{16.2}$$

Proof. Because all the norms in the finite-dimensional space $\widehat{P}$ are equivalent, there exist $\hat{c}_1$ and $\hat{c}_2$, depending on $\widehat{T}$, such that

$$\|\Pi_{\widehat{T}}^k \hat{\varphi}\|_{L^q(\widehat{T})} \le \hat{c}_1 \|\Pi_{\widehat{T}}^k \hat{\varphi}\|_{L^2(\widehat{T})}, \tag{16.3}$$

$$\|\Pi_{\widehat{T}}^k \hat{\varphi}\|_{L^{q'}(\widehat{T})} \le \hat{c}_2 \|\Pi_{\widehat{T}}^k \hat{\varphi}\|_{L^q(\widehat{T})}, \tag{16.4}$$

where $\frac{1}{q} + \frac{1}{q'} = 1$. Then,

$$
\begin{aligned}
\|\Pi_{\widehat{T}}^k \hat{\varphi}\|_{L^q(\widehat{T})}^2 &\leq c\|\Pi_{\widehat{T}}^k \hat{\varphi}\|_{L^2(\widehat{T})}^2 = c\int_{\widehat{T}} \hat{\varphi}\Pi_{\widehat{T}}^k \hat{\varphi} d\hat{x} \\
&\leq c\|\hat{\varphi}\|_{L^q(\widehat{T})}\|\Pi_{\widehat{T}}^k \hat{\varphi}\|_{L^{q'}(\widehat{T})} \\
&\leq c\|\hat{\varphi}\|_{L^q(\widehat{T})}\|\Pi_{\widehat{T}}^k \hat{\varphi}\|_{L^q(\widehat{T})},
\end{aligned}
$$

where we used (16.3), (16.4), (16.1) with $\hat{q} := \Pi_{\widehat{T}}^k \hat{\varphi}$, and the Hölder's inequality with $\frac{1}{q} + \frac{1}{q'} = 1$. This proves the target inequality. □

The following theorem gives an anisotropic error estimate of the projection Π_T^k.

Theorem 16.2. For $k \in \mathbb{N}_0$, let $\ell \in \mathbb{N}_0$ be such that $0 \leq \ell \leq k$. Let $p \in [1, \infty)$ and $q \in [1, \infty)$ be such that

$$
W^{1,p}(T) \hookrightarrow L^q(T), \tag{16.5}
$$

that is $1 - \frac{d}{p} \geq -\frac{d}{q}$. It then holds that, for any $\hat{\varphi} \in W^{\ell+1,p}(\widehat{T})$ with $\varphi := \hat{\varphi} \circ \Phi^{-1}$,

$$
\|\Pi_T^k \varphi - \varphi\|_{L^q(T)} \leq c|T|_d^{\frac{1}{q}-\frac{1}{p}} \sum_{|\epsilon|=\ell+1} h^{\varepsilon} \|\partial_r^{\varepsilon} \varphi\|_{L^p(T)}. \tag{16.6}
$$

In particular, if Condition 10.1 is imposed, it holds that, for any $\hat{\varphi} \in W^{\ell+1,p}(\widehat{T})$ with $\varphi := \hat{\varphi} \circ \Phi^{-1}$,

$$
\|\Pi_T^k \varphi - \varphi\|_{L^q(T)} \leq c|T|_d^{\frac{1}{q}-\frac{1}{p}} \sum_{|\epsilon|=\ell+1} \widetilde{\mathscr{H}}^{\varepsilon} \|\partial_{\tilde{x}}^{\varepsilon} (\varphi \circ \Phi_T)\|_{L^p(\Phi_T^{-1}(T))}. \tag{16.7}
$$

Proof. Using the scaling argument, we have

$$
\|\Pi_T^k \varphi - \varphi\|_{L^q(T)} = c|\det(A_{\widetilde{T}})|^{\frac{1}{q}} \|\Pi_{\widehat{T}}^k \hat{\varphi} - \hat{\varphi}\|_{L^q(\widehat{T})}. \tag{16.8}
$$

where we used $|\det(A_T)| = 1$. For any $\hat{\eta} \in \mathcal{P}^{\ell} \subset \widehat{P}$ with $0 \leq \ell \leq k$, from the triangle inequality and $\Pi_{\widehat{T}}^k \hat{\eta} = \hat{\eta}$, we have

$$
\|\Pi_{\widehat{T}}^k \hat{\varphi} - \hat{\varphi}\|_{L^q(\widehat{T})} \leq \|\Pi_{\widehat{T}}^k (\hat{\varphi} - \hat{\eta})\|_{L^q(\widehat{T})} + \|\hat{\eta} - \hat{\varphi}\|_{L^q(\widehat{T})}. \tag{16.9}
$$

Using (16.2) for the first term on the right-hand side of (16.9), we have

$$
\|\Pi_{\widehat{T}}^k (\hat{\varphi} - \hat{\eta})\|_{L^q(\widehat{T})} \leq c\|\hat{\varphi} - \hat{\eta}\|_{L^q(\widehat{T})}. \tag{16.10}
$$

Using the Sobolev embedding theorem for the second term on the right-hand side of (16.9) and (16.10), we obtain

$$
\|\hat{\varphi} - \hat{\eta}\|_{L^q(\widehat{T})} \leq c\|\hat{\varphi} - \hat{\eta}\|_{W^{1,p}(\widehat{T})}. \tag{16.11}
$$

Combining (16.8), (16.9), (16.10), and (16.11), we have

$$
\|\Pi_T^k \varphi - \varphi\|_{L^q(T)} \leq c|\det(A_{\widetilde{T}})|^{\frac{1}{q}} \inf_{\hat{\eta} \in \mathcal{P}^{\ell}} \|\hat{\varphi} - \hat{\eta}\|_{W^{1,p}(\widehat{T})}. \tag{16.12}
$$

From the Bramble–Hilbert-type lemma (e.g., see Subsection 1.7.4), there exists a constant $\hat{\eta}_{\beta} \in \mathbb{P}^{\ell}$ such that, for any $\hat{\varphi} \in W^{\ell+1,p}(\widehat{T})$,

$$
|\hat{\varphi} - \hat{\eta}_{\beta}|_{W^{t,p}(\widehat{T})} \leq C^{BH}(\widehat{T})|\hat{\varphi}|_{W^{\ell+1,p}(\widehat{T})}, \quad t = 0, 1. \tag{16.13}
$$

If Condition 10.1 is not imposed, using (10.2) ($m = 0$) and (17.20), we then have

$$\begin{aligned} \|\hat{\varphi} - \hat{\eta}_\beta\|_{W^{1,p}(\widehat{T})} &\leq c|\hat{\varphi}|_{W^{\ell+1,p}(\widehat{T})} \\ &\leq c|\det(A_{\widetilde{T}})|^{-\frac{1}{p}} \sum_{|\epsilon|=\ell+1} h^{\varepsilon} \|\partial_r^{\varepsilon} \varphi\|_{L^p(T)}. \end{aligned} \tag{16.14}$$

If Condition 10.1 is imposed, using (10.3) ($m = 0$) and (17.20), we then have

$$\begin{aligned} \|\hat{\varphi} - \hat{\eta}_\beta\|_{W^{1,p}(\widehat{T})} &\leq c|\hat{\varphi}|_{W^{\ell+1,p}(\widehat{T})} \\ &\leq c|\det(A_{\widetilde{T}})|^{-\frac{1}{p}} \sum_{|\epsilon|=\ell+1} \widetilde{\mathscr{H}^{\varepsilon}} \|\partial_{\tilde{x}}^{\varepsilon} \tilde{\varphi}\|_{L^p(\widetilde{T})}. \end{aligned} \tag{16.15}$$

Therefore, combining (16.12), (16.14), and (16.15) with (6.9), we have (16.6) and (16.7) using $\widetilde{T} = \Phi_T^{-1}(T)$ and $\tilde{\varphi} = \varphi \circ \Phi_T$. □

16.3 Global Interpolation Error Estimates

Recall the space V_h^n with $n = 1$ (see (14.7)). We define the standard discontinuous space as

$$P_{dc,h}^k := V_h^1 = \left\{ p_h \in L^1(\Omega);\ p_h|_T \in P \quad \forall T \in \mathbb{T}_h \right\}.$$

We also define the global interpolation Π_h^k to space $P_{dc,h}^k$ as

$$(\Pi_h^k \varphi)|_T := \Pi_T^k(\varphi|_T) \quad \forall T \in \mathbb{T}_h, \quad \forall \varphi \in L^1(\Omega).$$

Corollary 16.3. Suppose that the assumptions of Theorem 16.2 are satisfied. We impose Condition 6.2. Let Π_h^k be the corresponding global L^2-orthogonal projection. It then holds that, for any $\varphi \in W^{\ell+1,p}(\Omega)$;

(I) if Condition 10.1 is not imposed,

$$\|\Pi_h^k \varphi - \varphi\|_{L^q(\Omega)} \leq c \sum_{T \in \mathbb{T}_h} |T|_d^{\frac{1}{q} - \frac{1}{p}} \sum_{|\epsilon|=\ell+1} h^{\varepsilon} \|\partial_r^{\varepsilon} \varphi\|_{L^p(T)}. \tag{16.16}$$

(II) if Condition 10.1 is imposed,

$$\|\Pi_h^k \varphi - \varphi\|_{L^q(\Omega)} \leq c \sum_{T \in \mathbb{T}_h} |T|_d^{\frac{1}{q} - \frac{1}{p}} \sum_{|\epsilon|=\ell+1} \widetilde{\mathscr{H}^{\varepsilon}} \|\partial_{\tilde{x}}^{\varepsilon} (\varphi \circ \Phi_T)\|_{L^p(\Phi_T^{-1}(T))}. \tag{16.17}$$

Proof. This corollary is proved in the same argument as Corollary 14.2. □

16.4 Another Estimate

Theorem 16.4. Let $T \subset \mathbb{R}^d$ be a simplex. Let $\Pi_T^0 : L^2(T) \to \mathbb{P}^0(T)$ be the local L^2-projection defined by

$$\Pi_T^0 \varphi := \frac{1}{|T|_d} \int_T \varphi dx \quad \forall \varphi \in L^2(T).$$

It then holds that

$$\|\Pi_T^0 \varphi - \varphi\|_{L^2(T)} \leq \frac{h_T}{\pi} |\varphi|_{H^1(T)} \quad \forall \varphi \in H^1(T). \tag{16.18}$$

Proof. For any $\varphi \in H^1(T)$, we set $w := \Pi_T^0 \varphi - \varphi$. It then holds that

$$\int_T w dx = \int_T (\Pi_T^0 \varphi - \varphi) dx = \frac{1}{|T|_d} \int_T \varphi dx |T|_d - \int_T \varphi dx = 0.$$

Therefore, using the Poincaré inequality (1.15), we conclude (16.18). □

17 New Nonconforming FE Interpolation Error Estimates

17.1 Local Interpolation Error Estimates

We introduce the following theorem using the error estimates of the L^2-orthogonal projection.

Theorem 17.1. Let $\alpha := (\alpha_1, \ldots, \alpha_d)^\top \in \mathbb{N}_0^d$ be a multi-index and $k \in \mathbb{N}$. Let $p \in [1, \infty)$ and $q \in [1, \infty)$ be such that (16.5) holds. We define an interpolation operator $I_T : W^{k,p}(T) \to \mathbb{P}^k(T)$ that satisfies:

$$\partial_x^\alpha (I_T \varphi) = \Pi_T^0 (\partial_x^\alpha \varphi) \quad \forall \varphi \in W^{k,p}(T) \quad \forall \alpha : \ |\alpha| \leq k. \tag{17.1}$$

Then, for any $\hat{\varphi} \in W^{k+1,p}(\widehat{T})$ with $\varphi := \hat{\varphi} \circ \Phi^{-1}$ and any α with $|\alpha| \leq k$,

$$|I_T \varphi - \varphi|_{W^{k,q}(T)} \leq c |T|_d^{\frac{1}{q} - \frac{1}{p}} \sum_{i=1}^d h_i \left| \frac{\partial \varphi}{\partial r_i} \right|_{W^{k,p}(T)}. \tag{17.2}$$

If Condition 10.1 is imposed, then:

$$|I_T \varphi - \varphi|_{W^{k,q}(T)} \leq c |T|_d^{\frac{1}{q} - \frac{1}{p}} \sum_{i=1}^d \widetilde{\mathscr{H}_i} \left| \frac{\partial (\varphi \circ \Phi_T)}{\partial \tilde{x}_i} \right|_{W^{k,p}(\Phi_T^{-1}(T))}. \tag{17.3}$$

Proof. The error estimate of the L^2-orthogonal projection (16.6) with $\ell = 0$ yields

$$\begin{aligned}
|I_T \varphi - \varphi|_{W^{k,q}(T)}^q &= \sum_{|\alpha|=k} \|\partial_x^\alpha (I_T \varphi - \varphi)\|_{L^q(T)}^q \\
&= \sum_{|\alpha|=k} \left\| \Pi_T^0 (\partial_x^\alpha \varphi) - \partial_x^\alpha \varphi \right\|_{L^q(T)}^q \\
&\leq c |T|_d^{\left(\frac{1}{q} - \frac{1}{p}\right) q} \sum_{|\alpha|=k} \sum_{i=1}^d h_i^q \left\| \partial_x^\alpha \frac{\partial \varphi}{\partial r_i} \right\|_{L^p(T)}^q.
\end{aligned}$$

Using the Jensen-type inequality (1.5), we obtain

$$\begin{aligned}
|I_T \varphi - \varphi|_{W^{k,q}(T)} &\leq c |T|_d^{\frac{1}{q} - \frac{1}{p}} \left(\sum_{i=1}^d h_i^q \sum_{|\alpha|=k} \left\| \partial_x^\alpha \frac{\partial \varphi}{\partial r_i} \right\|_{L^p(T)}^q \right)^{\frac{1}{q}} \\
&\leq c |T|_d^{\frac{1}{q} - \frac{1}{p}} \sum_{i=1}^d h_i \left| \frac{\partial \varphi}{\partial r_i} \right|_{W^{k,p}(T)},
\end{aligned}$$

which is the target inequality in Eq. (17.2).

If Condition 10.1 is imposed, then (16.7) with $\ell = 0$ yields

$$|I_T \varphi - \varphi|_{W^{k,q}(T)}^q \leq c |T|_d^{\left(\frac{1}{q} - \frac{1}{p}\right) q} \sum_{|\alpha|=k} \sum_{i=1}^d \widetilde{\mathscr{H}_i^q} \left\| \partial_{\tilde{x}}^\alpha \frac{\partial (\varphi \circ \Phi_T)}{\partial \tilde{x}_i} \right\|_{L^p(\Phi_T^{-1}(T))}^q.$$

Using the Jensen-type inequality in (1.5), we obtained the target inequality in (17.3). $\square$

Note 17.2. The operators that satisfy (17.1) exist; see Theorems 17.3 and 17.9.

17.2 CR Finite Element

Let $\widehat{T} \subset \mathbb{R}^d$ be the reference element defined in Sections 5.1 and 5.1. Let $\widehat{F}_i$ be the face of $\widehat{T}$ opposite to $\widehat{p}_i$. The CR finite element on the reference element is defined by the triple $\{\widehat{T}, \widehat{P}, \widehat{\Sigma}\}$ as follows.

1. $\widehat{P} := \mathbb{P}^1(\widehat{T})$;
2. $\widehat{\Sigma}$ is a set $\{\hat{\chi}_i\}_{1 \leq i \leq N^{(d,1)}}$ of $N^{(d,1)}$ linear forms $\{\hat{\chi}_i\}_{1 \leq i \leq N^{(d,1)}}$ with its components such that, for any $\hat{q} \in \widehat{P}$,

$$\hat{\chi}_i(\hat{q}) := \frac{1}{|\widehat{F}_i|_{d-1}} \int_{\widehat{F}_i} \hat{q} d\hat{s} \quad \forall i \in \{1, \dots, d+1\}. \tag{17.4}$$

The nodal basis functions associated with the degrees of freedom by (17.4) are defined as

$$\hat{\theta}_i(\hat{x}) := d \left(\frac{1}{d} - \hat{\lambda}_i(\hat{x}) \right) \quad \forall i \in \{1, \dots, d+1\}. \tag{17.5}$$

It then holds that $\hat{\chi}_i(\hat{\theta}_j) = \delta_{ij}$ for any $i, j \in \{1, \dots, d+1\}$. Setting $V(\widehat{T}) := W^{1,1}(\widehat{T})$, the local operator $I_{\widehat{T}}^{CR}$ is defined as

$$I_{\widehat{T}}^{CR} : V(\widehat{T}) \ni \hat{\varphi} \mapsto I_{\widehat{T}}^{CR} \hat{\varphi} := \sum_{i=1}^{d+1} \left(\frac{1}{|\widehat{F}_i|_{d-1}} \int_{\widehat{F}_i} \hat{\varphi} d\hat{s} \right) \hat{\theta}_i \in \widehat{P}. \tag{17.6}$$

By analogous argument in Section 9, the CR finite elements $\{\widetilde{T}, \widetilde{P}, \widetilde{\Sigma}\}$ and $\{T, P, \Sigma\}$ are constructed. The local shape functions are $\tilde{\theta}_i = \psi_{\widehat{T}}^{-1}(\hat{\theta}_i)$ and $\theta_i = \psi_{\widetilde{T}}^{-1}(\tilde{\theta}_i)$ for any $i \in \{1, \dots, d+1\}$, and the associated local interpolation operators are respectively defined as

$$I_{\widetilde{T}}^{CR} : V(\widetilde{T}) \ni \tilde{\varphi} \mapsto I_{\widetilde{T}}^{CR} \tilde{\varphi} := \sum_{i=1}^{d+1} \left(\frac{1}{|\widetilde{F}_i|_{d-1}} \int_{\widetilde{F}_i} \tilde{\varphi} d\tilde{s} \right) \tilde{\theta}_i \in \widetilde{P}, \tag{17.7}$$

$$I_{T}^{CR} : V(T) \ni \varphi \mapsto I_{T}^{CR} \varphi := \sum_{i=1}^{d+1} \left(\frac{1}{|F_i|_{d-1}} \int_{F_i} \varphi ds \right) \theta_i \in P, \tag{17.8}$$

where $\{\widetilde{F}_i := \Phi_{\widetilde{T}}(\widehat{F}_i)\}_{i \in \{1,\dots,d+1\}}$ and $\{F_i := \Phi_T(\widetilde{F}_i)\}_{i \in \{1,\dots,d+1\}}$.

17.3 Local CR Interpolation Error Estimates

We present anisotropic CR interpolation error estimates.

Theorem 17.3. Let $p \in [1, \infty)$ and $q \in [1, \infty)$ be such that (16.5) holds. Then,

$$|I_T^{CR} \varphi - \varphi|_{W^{1,q}(T)} \leq c |T|_d^{\frac{1}{q} - \frac{1}{p}} \sum_{i=1}^{d} h_i \left| \frac{\partial \varphi}{\partial r_i} \right|_{W^{1,p}(T)} \quad \forall \varphi \in W^{2,p}(T), \tag{17.9}$$

$$\|I_T^{CR} \varphi - \varphi\|_{L^q(T)} \leq c |T|_d^{\frac{1}{q} - \frac{1}{p}} \sum_{i=1}^{d} h_i \left\| \frac{\partial \varphi}{\partial r_i} \right\|_{L^p(T)} \quad \forall \varphi \in W^{1,p}(T), \tag{17.10}$$

$$\|I_T^{CR} \varphi - \varphi\|_{L^q(T)} \leq c |T|_d^{\frac{1}{q} - \frac{1}{p}} \sum_{|\varepsilon|=2} h^{\varepsilon} \, \|\partial_r^{\varepsilon} \varphi\|_{L^p(T)} \quad \forall \varphi \in W^{2,p}(T). \tag{17.11}$$

If Condition 10.1 is imposed, then:

$$|I_T^{CR}\varphi - \varphi|_{W^{1,q}(T)} \leq c|T|_d^{\frac{1}{q}-\frac{1}{p}} \sum_{i=1}^{d} \widetilde{\mathscr{H}_i} \left| \frac{\partial(\varphi \circ \Phi_T)}{\partial \tilde{x}_i} \right|_{W^{1,p}(\Phi_T^{-1}(T))} \quad \forall \varphi \in W^{2,p}(T), \tag{17.12}$$

$$\|I_T^{CR}\varphi - \varphi\|_{L^q(T)} \leq c|T|_d^{\frac{1}{q}-\frac{1}{p}} \sum_{i=1}^{d} \widetilde{\mathscr{H}_i} \left\| \frac{\partial}{\partial \tilde{x}_i}(\varphi \circ \Phi_T) \right\|_{L^p(\Phi_T^{-1}(T))} \quad \forall \varphi \in W^{1,p}(T), \tag{17.13}$$

$$\|I_T^{CR}\varphi - \varphi\|_{L^q(T)} \leq c|T|_d^{\frac{1}{q}-\frac{1}{p}} \sum_{|\varepsilon|=2} \widetilde{\mathscr{H}^{\varepsilon}} \|\partial_{\tilde{x}}^{\varepsilon}(\varphi \circ \Phi_T)\|_{L^p(\Phi_T^{-1}(T))} \quad \forall \varphi \in W^{2,p}(T), \tag{17.14}$$

Proof. Only CR interpolation satisfies the condition (17.1) to prove the estimate (17.9) and (17.12).

For $\varphi \in W^{2,p}(T)$, Green's formula and the definition of the CR interpolation imply that because $I_T^{CR}\varphi \in \mathbb{P}^1$,

$$\begin{aligned}
\frac{\partial}{\partial x_j}(I_T^{CR}\varphi) &= \frac{1}{|T|}\int_T \frac{\partial}{\partial x_j}(I_T^{CR}\varphi)dx = \frac{1}{|T|}\sum_{i=1}^{d+1} n_T^{(j)} \int_{F_i} I_T^{CR}\varphi ds \\
&= \frac{1}{|T|}\sum_{i=1}^{d+1} n_T^{(j)} \int_{F_i} \varphi ds = \frac{1}{|T|}\int_T \frac{\partial \varphi}{\partial x_j} dx = \Pi_T^0 \left(\frac{\partial \varphi}{\partial x_j} \right)
\end{aligned}$$

for $j = 1, \ldots, d$, where n_T denotes the outer unit normal vector to T and $n_T^{(j)}$ denotes the jth component of n_T. Therefore, from Theorem 17.1, the target inequalities (17.9) and (17.12) hold.

The (standard) scaling argument with $|\det(A_T)| = 1$ yields

$$\|I_T^{CR}\varphi - \varphi\|_{L^q(T)} = |\det(A_{\widetilde{T}})|^{\frac{1}{q}} \|I_{\widehat{T}}^{CR}\hat{\varphi} - \hat{\varphi}\|_{L^q(\widehat{T})}. \tag{17.15}$$

For any $\hat{\eta} \in \mathbb{P}^\ell$ with $\ell \in \{0, 1\}$, from the triangle inequality and $I_{\widehat{T}}^{CR}\hat{\eta} = \hat{\eta}$, we have

$$\|I_{\widehat{T}}^{CR}\hat{\varphi} - \hat{\varphi}\|_{L^q(\widehat{T})} \leq \|I_{\widehat{T}}^{CR}(\hat{\varphi} - \hat{\eta})\|_{L^q(\widehat{T})} + \|\hat{\eta} - \hat{\varphi}\|_{L^q(\widehat{T})}. \tag{17.16}$$

Using the definition of the CR interpolation (17.6) and the trace theorem, we have

$$\|I_{\widehat{T}}^{CR}(\hat{\varphi} - \hat{\eta})\|_{L^q(\widehat{T})} \leq \sum_{i=1}^{d+1} \frac{1}{|\widehat{F}_i|_{d-1}} \int_{\widehat{F}_i} |\hat{\varphi} - \hat{\eta}| d\hat{s} \|\hat{\theta}_i\|_{L^q(\widehat{T})} \leq c\|\hat{\varphi} - \hat{\eta}\|_{W^{1,p}(\widehat{T})}. \tag{17.17}$$

Using the Sobolev embedding theorem for the second term on the right-hand side of (17.16), we obtain

$$\|\hat{\varphi} - \hat{\eta}\|_{L^q(\widehat{T})} \leq c\|\hat{\varphi} - \hat{\eta}\|_{W^{1,p}(\widehat{T})}. \tag{17.18}$$

Combining (17.15), (17.16), (17.17) and (17.18), we have

$$\|I_T^{CR}\varphi - \varphi\|_{L^q(T)} \leq C(\widehat{T})|\det(A_{\widetilde{T}})|^{\frac{1}{q}} \inf_{\hat{\eta} \in \mathbb{P}^\ell} \|\hat{\varphi} - \hat{\eta}\|_{W^{1,p}(\widehat{T})}. \tag{17.19}$$

From the Bramble–Hilbert-type lemma (e.g., see Subsection 1.7.4), there exists a constant $\hat{\eta}_\beta \in \mathbb{P}^\ell$ with $\ell \in \{0, 1\}$ such that, for any $\hat{\varphi} \in W^{\ell+1,p}(\widehat{T})$,

$$|\hat{\varphi} - \hat{\eta}_\beta|_{W^{t,p}(\widehat{T})} \leq C^{BH}(\widehat{T})|\hat{\varphi}|_{W^{\ell+1,p}(\widehat{T})}, \quad t = 0, 1. \tag{17.20}$$

Thus, from (17.19) and (17.20),

$$\begin{aligned}\|I_T^{CR}\varphi - \varphi\|_{L^q(T)} &\le C(\widehat{T})|\det(A_{\widetilde{T}})|^{\frac{1}{q}} \inf_{\hat{\eta}\in\mathbb{P}^\ell} \|\hat{\varphi} - \hat{\eta}\|_{W^{1,p}(\widehat{T})} \\ &\le C(\widehat{T})|\det(A_{\widetilde{T}})|^{\frac{1}{q}} \|\hat{\varphi} - \hat{\eta}_\beta\|_{W^{1,p}(\widehat{T})} \\ &\le C_1(\widehat{T})|\det(A_{\widetilde{T}})|^{\frac{1}{q}} |\hat{\varphi}|_{W^{\ell+1,p}(\widehat{T})}, \quad \ell = 0, 1. \end{aligned} \tag{17.21}$$

Therefore, using (6.9), (17.21) and (10.2),

$$\|I_T^{CR}\varphi - \varphi\|_{L^q(T)} \le c|T|_d^{\frac{1}{q}-\frac{1}{p}} \sum_{|\epsilon|=\ell+1} h^\varepsilon \|\partial_r^\epsilon \varphi\|_{L^p(T)}, \quad \ell = 0, 1,$$

which leads to (17.10) and (17.11). In particular, if Condition 10.1 is imposed, using (6.9), (17.21) and (10.3),

$$\|I_T^{CR}\varphi - \varphi\|_{L^q(T)} \le c|T|_d^{\frac{1}{q}-\frac{1}{p}} \sum_{|\epsilon|=\ell+1} \widetilde{\mathscr{H}^\varepsilon} \|\partial_{\tilde{x}}^\varepsilon \tilde{\varphi}\|_{L^p(\widetilde{T})}, \quad \ell = 0, 1,$$

which leads to (17.13) and (17.14). $\square$

17.4 Global CR Interpolation Error Estimates

Recall the space V_h^n with $n = 1$ (see (14.7)). We define the CR finite element space as

$$V_h^{CR} := \left\{ \varphi_h \in V_h^1; \ \int_F [\![\varphi_h]\!]_F ds = 0 \ \forall F \in \mathcal{F}_h^i \right\}.$$

We also define the global interpolation $I_h^{CR} : W^{1,1}(\Omega) \to V_h^{CR}$ as follows.

$$(I_h^{CR}\varphi)|_T := I_T^{CR}(\varphi|_T) = \sum_{i=1}^{d+1} \left(\frac{1}{|F_i|_{d-1}} \int_{F_i} \varphi|_T ds \right) \theta_i \quad \forall T \in \mathbb{T}_h, \ \forall \varphi \in W^{1,1}(\Omega).$$

Corollary 17.4. Suppose that the assumptions of Theorem 17.3 are satisfied. Let I_h^{CR} be the corresponding global CR interpolation operator. Then,

$$|\varphi - I_h^{CR}\varphi|_{W^{1,q}(\mathbb{T}_h)} \le c \sum_{T\in\mathbb{T}_h} |T|_d^{\frac{1}{q}-\frac{1}{p}} \sum_{i=1}^d h_i \left| \frac{\partial\varphi}{\partial r_i} \right|_{W^{1,p}(T)} \quad \forall \varphi \in W^{2,p}(\Omega), \tag{17.22}$$

$$\|\varphi - I_h^{CR}\varphi\|_{L^q(\Omega)} \le c \sum_{T\in\mathbb{T}_h} |T|_d^{\frac{1}{q}-\frac{1}{p}} \sum_{i=1}^d h_i \left\| \frac{\partial\varphi}{\partial r_i} \right\|_{L^p(T)} \quad \forall \varphi \in W^{1,p}(\Omega), \tag{17.23}$$

$$\|\varphi - I_h^{CR}\varphi\|_{L^q(\Omega)} \le c \sum_{T\in\mathbb{T}_h} |T|_d^{\frac{1}{q}-\frac{1}{p}} \sum_{|\varepsilon|=2} h^\varepsilon \, \|\partial_r^\varepsilon \varphi\|_{L^p(T)} \quad \forall \varphi \in W^{2,p}(\Omega). \tag{17.24}$$

If Condition 10.1 is imposed, then:

$$|\varphi - I_h^{CR}\varphi|_{W^{1,q}(\mathbb{T}_h)} \le c \sum_{T\in\mathbb{T}_h} |T|_d^{\frac{1}{q}-\frac{1}{p}} \sum_{i=1}^d \widetilde{\mathscr{H}_i} \left| \frac{\partial(\varphi\circ\Phi_T)}{\partial\tilde{x}_i} \right|_{W^{1,p}(\Phi_T^{-1}(T))} \quad \forall \varphi \in W^{2,p}(\Omega), \tag{17.25}$$

$$\|\varphi - I_h^{CR}\varphi\|_{L^q(\Omega)} \le c \sum_{T\in\mathbb{T}_h} |T|_d^{\frac{1}{q}-\frac{1}{p}} \sum_{i=1}^d \widetilde{\mathscr{H}_i} \left\| \frac{\partial}{\partial\tilde{x}_i}(\varphi\circ\Phi_T) \right\|_{L^p(\Phi_T^{-1}(T))} \quad \forall \varphi \in W^{1,p}(\Omega), \tag{17.26}$$

$$\|\varphi - I_h^{CR}\varphi\|_{L^q(\Omega)} \le c \sum_{T\in\mathbb{T}_h} |T|_d^{\frac{1}{q}-\frac{1}{p}} \sum_{|\varepsilon|=2} \widetilde{\mathscr{H}^\varepsilon} \|\partial_{\tilde{x}}^\varepsilon(\varphi\circ\Phi_T)\|_{L^p(\Phi_T^{-1}(T))} \quad \forall \varphi \in W^{2,p}(\Omega). \tag{17.27}$$

Proof. This corollary is proved in the same argument as Corollary 14.2. $\square$

17.5 Another Estimate

Theorem 17.5. Let $T \subset \mathbb{R}^d$ be a simplex. Let $I_T^{CR} : H^1(T) \to \mathbb{P}^1(T)$ be the local CR interpolation operator defined as

$$I_T^{CR} : H^1(T) \ni \varphi \mapsto I_T^{CR}\varphi := \sum_{i=1}^{d+1} \left(\frac{1}{|F_i|_{d-1}} \int_{F_i} \varphi ds \right) \theta_i \in \mathbb{P}^1.$$

It then holds that

$$|I_T^{CR}\varphi - \varphi|_{H^1(T)} \leq \frac{h_T}{\pi} |\varphi|_{H^2(T)} \quad \forall \varphi \in H^2(T). \tag{17.28}$$

Proof. Using (16.18).

$$\begin{aligned}
|I_T^{CR}\varphi - \varphi|_{H^1(T)}^2 &= \sum_{j=1}^{d} \left\| \frac{\partial}{\partial x_j} (I_T^{CR}\varphi - \varphi) \right\|_{L^2(T)}^2 \\
&= \sum_{j=1}^{d} \left\| \Pi_T^0 \left(\frac{\partial \varphi}{\partial x_j} \right) - \left(\frac{\partial \varphi}{\partial x_j} \right) \right\|_{L^2(T)}^2 \\
&\leq \left(\frac{h_T}{\pi} \right)^2 \sum_{i,j=1}^{d} \left\| \frac{\partial^2 \varphi}{\partial x_i \partial x_j} \right\|_{L^2(T)}^2 \\
&= \left(\frac{h_T}{\pi} \right)^2 |\varphi|_{H^2(T)}^2,
\end{aligned}$$

which conclude (17.28). □

17.6 Nodal CR Interpolation Error Estimates

Let $\widehat{T} \subset \mathbb{R}^d$ be the reference element defined in Sections 5.1 and 5.1. Let $\widehat{F}_i$ be the face of $\widehat{T}$ opposite to $\widehat{p}_i$ and let $\hat{x}_{\widehat{F}_i}$ the barycentre of the face $\widehat{F}_i$. The (nodal) CR finite element on the reference element is defined by the triple $\{\widehat{T}, \widehat{P}, \widehat{\Sigma}\}$ as follows.

1. $\widehat{P} := \mathbb{P}^1(\widehat{T})$;

2. $\widehat{\Sigma}$ is a set $\{\hat{\chi}_i\}_{1 \leq i \leq N^{(d,1)}}$ of $N^{(d,1)}$ linear forms $\{\hat{\chi}_i\}_{1 \leq i \leq N^{(d,1)}}$ with its components such that, for any $\hat{q} \in \widehat{P}$,

$$\hat{\chi}_i(\hat{p}) := \hat{q}(\hat{x}_{\widehat{F}_i}) \quad \forall i \in \{1, \ldots, d+1\}. \tag{17.29}$$

The nodal basis functions associated with the degrees of freedom by (17.29) are defined as

$$\hat{\theta}_i(\hat{x}) := d \left(\frac{1}{d} - \hat{\lambda}_i(\hat{x}) \right) \quad \forall i \in \{1, \ldots, d+1\}. \tag{17.30}$$

It then holds that $\hat{\chi}_i(\hat{\theta}_j) = \delta_{ij}$ for any $i, j \in \{1, \ldots, d+1\}$. Setting $V(\widehat{T}) := \mathcal{C}(\widehat{T})$ or $V(\widehat{T}) := W^{s,p}(\widehat{T})$ with $p \in [1, \infty]$ and $ps > d$ ($s \geq d$ if $p = 1$), the local operator $I_{\widehat{T}}^{nCR}$ is defined as

$$I_{\widehat{T}}^{nCR} : V(\widehat{T}) \ni \hat{\varphi} \mapsto I_{\widehat{T}}^{nCR}\hat{\varphi} := \sum_{i=1}^{d+1} \hat{\varphi}(\hat{x}_{\widehat{F}_i}) \hat{\theta}_i \in \widehat{P}. \tag{17.31}$$

By analogous argument in Section 9, the nodal CR finite elements $\{\widetilde{T}, \widetilde{P}, \widetilde{\Sigma}\}$ and $\{T, P, \Sigma\}$ are constructed. The local shape functions are $\tilde{\theta}_i = \psi_{\widetilde{T}}^{-1}(\hat{\theta}_i)$ and $\theta_i = \psi_{\widetilde{T}}^{-1}(\tilde{\theta}_i)$ for any $i \in \{1, \ldots, d+1\}$, and the associated local interpolation operators are respectively defined as

$$I_{\widetilde{T}}^{nCR} : V(\widetilde{T}) \ni \tilde{\varphi} \mapsto I_{\widetilde{T}}^{nCR} \tilde{\varphi} := \sum_{i=1}^{d+1} \tilde{\varphi}(\tilde{x}_{\widetilde{F}_i}) \tilde{\theta}_i \in \widetilde{P}, \tag{17.32}$$

$$I_{T}^{nCR} : V(T) \ni \varphi \mapsto I_{T}^{nCR} \varphi := \sum_{i=1}^{d+1} \varphi(x_{F_i}) \theta_i \in P, \tag{17.33}$$

where $\{\widetilde{F}_i := \Phi_{\widetilde{T}}(\widehat{F}_i)\}_{i \in \{1,\ldots,d+1\}}$, $\{F_i := \Phi_T(\widetilde{F}_i)\}_{i \in \{1,\ldots,d+1\}}$, $\tilde{x}_{\widetilde{F}_i} = \Phi_{\widetilde{T}}(\hat{x}_{\widehat{F}_i})$, $x_{F_i} = \Phi_T(\tilde{x}_{\widetilde{F}_i})$ for $i \in \{1, \ldots, d+1\}$.

Corollary 17.6. Let $\{\widehat{T}, \widehat{P}, \widehat{\Sigma}\}$ be the Crouzeix–Raviart finite element with $V(\widehat{T}) := \mathcal{C}(\widehat{T})$ and $\widehat{P} := \mathbb{P}^1(\widehat{T})$. Set $I_{\widehat{T}}; = I_{\widehat{T}}^{nCR}$. Let $m \in \mathbb{N}_0$, $\ell \in \mathbb{N}$, and $p \in \mathbb{R}$ be such that

$$d = 2 : \begin{cases} p \in (2, \infty] & \text{if } m = 0,\ \ell = 1, \\ p \in [1, \infty] & \text{if } m = 0,\ \ell = 2 \text{ or } m = 1,\ \ell = 2, \end{cases}$$

$$d = 3 : \begin{cases} p \in \left(\frac{3}{\ell}, \infty\right] & \text{if } m = 0,\ \ell = 1, 2, \\ p \in (2, \infty] & \text{if } m = 1,\ \ell = 2. \end{cases}$$

Setting $q \in [1, \infty]$ such that $W^{\ell-m,p}(\widehat{T}) \hookrightarrow L^q(\widehat{T})$. Then, for all $\hat{\varphi} \in W^{\ell,p}(\widehat{T})$ with $\varphi := \hat{\varphi} \circ \Phi^{-1}$, we have

$$|\varphi - I_T^{nCR} \varphi|_{W^{m,q}(T)} \leq c |T|_d^{\frac{1}{q} - \frac{1}{p}} \left(\frac{H_T}{h_T} \right)^m \sum_{|\varepsilon| = \ell - m} h^{\varepsilon} |\partial_r^{\varepsilon} \varphi|_{W^{m,p}(T)}, \tag{17.34}$$

In particular, if Condition 10.1 is imposed, it holds that, for all $\hat{\varphi} \in W^{\ell,p}(\widehat{T})$ with $\varphi := \hat{\varphi} \circ \Phi^{-1}$,

$$|\varphi - I_T \varphi|_{W^{m,q}(T)} \leq c |T|_d^{\frac{1}{q} - \frac{1}{p}} \left(\frac{H_T}{h_T} \right)^m \sum_{|\varepsilon| = \ell - m} \widetilde{\mathscr{H}}^{\varepsilon} |\partial_{\tilde{x}}^{\varepsilon} (\varphi \circ \Phi_{\widetilde{T}})|_{W^{m,p}(\Phi_{\widetilde{T}}^{-1}(T))}. \tag{17.35}$$

Furthermore, for any $\hat{\varphi} \in \mathcal{C}(\widehat{T})$ with $\varphi := \hat{\varphi} \circ \Phi^{-1}$, it holds that

$$\|\varphi - I_T \varphi\|_{L^\infty(T)} \leq c \|\varphi\|_{L^\infty(T)}.$$

Proof. For $k = 1$, we only introduce functionals $\mathscr{F}_i$ satisfying (12.2) in Theorem 14.1 (or Theorem 12.1) for each ℓ and m.

Let $m = 0$, that is, $\beta = (0, \cdots, 0) \in \mathbb{N}_0^d$. We then have $j = \dim \mathbb{P}^1 = d + 1$. From the Sobolev embedding theorem (Theorem 1.6), we have $W^{\ell,p}(\widehat{T}) \subset \mathcal{C}^0(\widehat{T})$ with $1 < p \leq \infty$, $d < \ell p$ or $p = 1$, $d \leq \ell$. Under this condition, we use

$$\mathscr{F}_i(\hat{\varphi}) := \hat{\varphi}(\hat{x}_{\widehat{F}_i}), \quad \hat{\varphi} \in W^{\ell,p}(\widehat{T}), \quad i = 1, \ldots, d+1.$$

It then holds that

$$|\mathscr{F}_i(\hat{\varphi})| \leq \|\hat{\varphi}\|_{\mathcal{C}^0(\widehat{T})} \leq c \|\hat{\varphi}\|_{W^{\ell,p}(\widehat{T})},$$

which means $\mathscr{F}_i \in W^{\ell,p}(\widehat{T})'$ for $i = 1, \ldots, d+1$, that is, (12.2a) is satisfied. Furthermore, we have

$$\mathscr{F}_i(I_{\widehat{T}}^{nCR} \hat{\varphi}) = (I_{\widehat{T}}^{nCR} \hat{\varphi})(\hat{x}_{\widehat{F}_i}) = \hat{\varphi}(\hat{x}_{\widehat{F}_i}) = \mathscr{F}_i(\hat{\varphi}), \quad i = 1, \ldots, d+1,$$

which satisfies (12.2b). For all $\hat{\eta} \in \mathbb{P}^1$, if $\mathscr{F}_i(\hat{\eta}) = 0$ for $i = 1, \ldots, d+1$, it obviously holds $\hat{\eta} = 0$. This means that (12.2c) is satisfied.

Let $d = 2$ and $m = 1$ ($\ell = 2$). We set $\beta = (1,0)$. We then have $j = \dim(\partial^\beta \mathbb{P}^1) = 1$. We consider a functional

$$\mathscr{F}_1(\hat{\varphi}) := \int_0^{\frac{1}{2}} \hat{\varphi}(\hat{x}_1, 1/2) d\hat{x}_1, \quad \hat{\varphi} \in W^{2,p}(\widehat{T}), \quad 1 < p.$$

We set $\widehat{I} := \{\hat{x} \in \widehat{T};\ \hat{x}_2 = \frac{1}{2}\}$. The continuity is then shown by the trace theorem (e.g., see Theorem 1.7):

$$|\mathscr{F}_1(\hat{\varphi})| \le \|\hat{\varphi}\|_{L^1(\widehat{I})} \le c \|\hat{\varphi}\|_{W^{1,p}(\widehat{T})},$$

which means $\mathscr{F}_1 \in W^{2,p}(\widehat{T})'$, that is, (12.2a) is satisfied. Furthermore, it holds that

$$\begin{aligned} \mathscr{F}_1(\partial^{(1,0)}(\hat{\varphi} - I_{\widehat{T}}^{nCR}\hat{\varphi})) &= \int_0^{\frac{1}{2}} \frac{\partial}{\partial \hat{x}_1} (\hat{\varphi} - I_{\widehat{T}}^{nCR}\hat{\varphi})(\hat{x}_1, 1/2) d\hat{x}_1 \\ &= \left[\hat{\varphi} - I_{\widehat{T}}^{nCR}\hat{\varphi} \right]_{(0,1/2)}^{(1/2,1/2)} = 0, \end{aligned}$$

which satisfy (12.2b). Let $\hat{\eta} := a\hat{x}_1 + b\hat{x}_2 + c$. We then have

$$\mathscr{F}_1(\partial^{(1,0)}\hat{\eta}) = \frac{1}{2} a.$$

If $\mathscr{F}_1(\partial^{(1,0)}\hat{\eta}) = 0$, $a = 0$. This implies that $\partial^{(1,0)}\hat{\eta} = 0$. This means that (12.2c) is satisfied.

By analogous argument, the case $\beta = (0,1)$ holds.

Let $d = 3$ and $m = 1$ ($\ell = 2$). We consider Type (i) in Section 5.1 in detail. That is, the reference element is $\widehat{T} = \mathrm{conv}\{0, e_1, e_2, e_3\}$. Here, $e_1, \ldots, e_3 \in \mathbb{R}^3$ are the canonical basis. We set $\beta = (1,0,0)$. We then have $j = \dim(\partial^\beta \mathcal{P}^1) = 1$. We consider a functional

$$\mathscr{F}_1(\hat{\varphi}) := \int_0^{\frac{1}{3}} \hat{\varphi}(\hat{x}_1, 1/3, 1/3) d\hat{x}_1, \quad \hat{\varphi} \in W^{2,p}(\widehat{T}), \quad \frac{3}{2} < p.$$

We set $\widehat{I} := \{\hat{x} \in \widehat{T};\ \hat{x}_2 = \frac{1}{3},\ \ \hat{x}_3 = \frac{1}{3}\}$. The continuity is then shown by the trace theorem:

$$|\mathscr{F}_1(\hat{\varphi})| \le \|\hat{\varphi}\|_{L^1(\widehat{I})} \le c \|\hat{\varphi}\|_{W^{2,p}(\widehat{T})} \quad \text{if } p > 2,$$

which means $\mathscr{F}_1 \in W^{2,p}(\widehat{T})'$, that is, (12.2a) is satisfied. Furthermore, it holds that

$$\mathscr{F}_1(\partial^{(1,0,0)}(\hat{\varphi} - I_{\widehat{T}}^{nCR}\hat{\varphi})) = \left[\hat{\varphi} - I_{\widehat{T}}^{nCR}\hat{\varphi} \right]_{(0,1/3,1/3)}^{(1/3,1/3,1/3)} = 0,$$

which satisfy (12.2b). Let $\hat{\eta} := a\hat{x}_1 + b\hat{x}_2 + c\hat{x}_3 + d$. We then have

$$\mathscr{F}_1(\partial^{(1,0,0)}\hat{\eta}) = \frac{1}{3} a.$$

If $\mathscr{F}_1(\partial^{(1,0,0)}\hat{\eta}) = 0$, $a = 0$. This implies that $\partial^{(1,0,0)}\hat{\eta} = 0$. This means that (12.2c) is satisfied.

By analogous argument, it holds the cases $\beta = (0,1,0), (0,0,1)$.

We consider Type (ii) in Section 5.1. That is, the reference element is $\widehat{T} = \mathrm{conv}\{0, e_1, e_1 + e_2, e_3\}$. We set $\beta = (1,0,0)$. We then have $j = \dim(\partial^\beta \mathbb{P}^1) = 1$. We consider a functional

$$\mathscr{F}_1(\hat{\varphi}) := \int_{\frac{1}{3}}^{\frac{2}{3}} \hat{\varphi}(\hat{x}_1, 1/3, 1/3) d\hat{x}_1, \quad \hat{\varphi} \in W^{2,p}(\widehat{T}).$$

We can deduce the result by the similar argument with Type (i).

When $m = \ell = 0$, $p = \infty$ and $q \in [1, \infty]$, it holds that

$$\|\hat{\varphi} - I_{\widehat{T}}^{nCR}\hat{\varphi}\|_{L^q(\widehat{T})} \leq c\|\hat{\varphi}\|_{L^\infty(\widehat{T})},$$

because we have

$$|(I_{\widehat{T}}^{nCR}\hat{\varphi})(\hat{x})| \leq \sum_{i=1}^{d+1} |\hat{\varphi}(\hat{x}_{\widehat{F}_i})||\hat{\theta}_i(\hat{x})| \leq (d+1)\left(\max_{1\leq i\leq d+1} \|\hat{\theta}_i\|_{L^\infty(\widehat{T})}\right)\|\hat{\varphi}\|_{L^\infty(\widehat{T})}.$$

□

17.7 Morley Finite Element

Any dimensional Morley finite element is introduced in [51].

Let $\widehat{T} \subset \mathbb{R}^d$, $d \in \{2, 3\}$, be the reference element defined in Sections 5.1 and 5.1. Let $\widehat{F}_i$, $1 \leq i \leq d+1$, be the $(d-1)$-dimensional subsimplex of $\widehat{T}$ without $\widehat{P}_i$ and $\widehat{S}_{i,j}$, $1 \leq i < j \leq d+1$, the $(d-2)$-dimensional subsimplex of $\widehat{T}$ without $\widehat{P}_i$ and $\widehat{P}_j$. The d-dimensional Morley finite element on the reference element is defined by the triple $\{\widehat{T}, \widehat{P}, \widehat{\Sigma}\}$ as

1. $\widehat{P} := \mathbb{P}^2(\widehat{T})$;
2. $\widehat{\Sigma}$ is a set $\{\hat{\chi}_i\}_{1\leq i\leq N^{(d,2)}}$ of $N^{(d,2)}$ linear forms $\{\hat{\chi}_{i,j}^{(1)}\}_{1\leq i<j\leq d+1} \cup \{\hat{\chi}_i^{(2)}\}_{1\leq i\leq d+1}$ with its components such that, for any $\hat{q} \in \widehat{P}$,

$$\hat{\chi}_{i,j}^{(1)}(\hat{q}) := \frac{1}{|\widehat{S}_{i,j}|}\int_{\widehat{S}_{i,j}} \hat{q}\, d\hat{s}, \quad 1 \leq i < j \leq d+1, \tag{17.36a}$$

$$\hat{\chi}_i^{(2)}(\hat{q}) := \frac{1}{|\widehat{F}_i|}\int_{\widehat{F}_i} \frac{\partial \hat{q}}{\partial \hat{n}_i} d\hat{s}, \quad 1 \leq i \leq d+1, \tag{17.36b}$$

   where $\frac{\partial}{\partial \hat{n}_i} = n_{\widehat{T},i} \cdot \nabla$, and $n_{\widehat{T},i}$ is the unit outer normal to $\widehat{F}_i \subset \partial\widehat{T}$. For $d = 2$, $\hat{\chi}_{i,j}^{(1)}(\hat{p})$ is interpreted as

$$\hat{\chi}_{i,j}^{(1)}(\hat{q}) = \hat{q}(\widehat{p}_k), \quad k = 1, 2, 3, \quad k \neq i, j.$$

For a Morley finite element, $\widehat{\Sigma}$ is unisolvent (see [51, Lemma 2]). The nodal basis functions associated with the degrees of freedom provided by (17.36) are defined as follows:

$$\begin{aligned}\hat{\theta}_{i,j}^{(1)} &:= 1 - (d-1)(\hat{\lambda}_i + \hat{\lambda}_j) + d(d-1)\hat{\lambda}_i\hat{\lambda}_j \\ &\quad - (d-1)(\nabla\hat{\lambda}_i)^T\nabla\hat{\lambda}_j \sum_{k=i,j} \frac{\hat{\lambda}_k(d\hat{\lambda}_k - 2)}{2|\nabla\hat{\lambda}_k|_E^2}, \quad 1 \leq i < j \leq d+1, \end{aligned} \tag{17.37a}$$

$$\hat{\theta}_i^{(2)} := \frac{\hat{\lambda}_i(d\hat{\lambda}_i - 2)}{2|\nabla\hat{\lambda}_i|_E}, \quad 1 \leq i \leq d+1, \tag{17.37b}$$

where $|\nabla\hat{\lambda}_i|_E$ denotes the Euclidean norm in $\mathbb{R}^d$. Subsequently, [51, Theorem 1] proved that, for $1 \leq i < j \leq d+1$,

$$\hat{\chi}_{k,\ell}^{(1)}(\hat{\theta}_{i,j}^{(1)}) = \delta_{ik}\delta_{j\ell}, \quad 1 \leq k < \ell \leq d+1, \quad \hat{\chi}_k^{(2)}(\hat{\theta}_{i,j}^{(1)}) = 0, \quad 1 \leq k \leq d+1, \tag{17.38}$$

and, for $1 \le i \le d+1$,

$$\hat{\chi}_{k,\ell}^{(1)}(\hat{\theta}_i^{(2)}) = 0, \quad 1 \le k < \ell \le d+1, \quad \hat{\chi}_k^{(2)}(\hat{\theta}_i^{(2)}) = \delta_{ik}, \quad 1 \le k \le d+1. \tag{17.39}$$

The local interpolation operator $I_{\widehat{T}}^M$ is defined by

$$I_{\widehat{T}}^M : W^{2,1}(\widehat{T}) \ni \hat{\varphi} \mapsto I_{\widehat{T}}^M \hat{\varphi} \in \widehat{P}, \tag{17.40}$$

with

$$I_{\widehat{T}}^M \hat{\varphi} := \sum_{1 \le i < j \le d+1} \hat{\chi}_{i,j}^{(1)}(\hat{\varphi}) \hat{\theta}_{i,j}^{(1)} + \sum_{1 \le i \le d+1} \hat{\chi}_i^{(2)}(\hat{\varphi}) \hat{\theta}_i^{(2)}. \tag{17.41}$$

Then, it holds that $I_{\widehat{T}}^M \hat{q} = \hat{q}$ for any $\hat{q} \in \widehat{P}$ and, for any $\hat{\varphi} \in W^{2,1}(\widehat{T})$,

$$\hat{\chi}_{i,j}^{(1)}(I_{\widehat{T}}^M \hat{\varphi}) = \hat{\chi}_{i,j}^{(1)}(\hat{\varphi}), \quad 1 \le i < j \le d+1, \tag{17.42a}$$

$$\hat{\chi}_i^{(2)}(I_{\widehat{T}}^M \hat{\varphi}) = \hat{\chi}_i^{(2)}(\hat{\varphi}), \quad 1 \le i \le d+1. \tag{17.42b}$$

By analogous argument in Section 9, the Morley finite elements $\{\widetilde{T}, \widetilde{P}, \widetilde{\Sigma}\}$ and $\{T, P, \Sigma\}$ are constructed. The local shape functions are

$$\begin{aligned} \tilde{\theta}_{i,j}^{(1)} &= \psi_{\widehat{T}}^{-1}(\hat{\theta}_{i,j}^{(1)}), \quad 1 \le i < j \le d+1, \quad \tilde{\theta}_i^{(2)} = \psi_{\widehat{T}}^{-1}(\hat{\theta}_i^{(2)}), \quad 1 \le i \le d+1, \\ \theta_{i,j}^{(1)} &= \psi_{\widetilde{T}}^{-1}(\tilde{\theta}_{i,j}^{(1)}), \quad 1 \le i < j \le d+1, \quad \theta_i^{(2)} = \psi_{\widetilde{T}}^{-1}(\tilde{\theta}_i^{(2)}), \quad 1 \le i \le d+1. \end{aligned}$$

The associated local Morley interpolation operators are defined as

$$I_{\widetilde{T}}^M : W^{2,1}(\widetilde{T}) \ni \tilde{\varphi} \mapsto I_{\widetilde{T}}^M \tilde{\varphi} \in \widetilde{P}, \tag{17.43}$$

with, for any $\tilde{\varphi} \in W^{2,1}(\widetilde{T})$,

$$\tilde{\chi}_{i,j}^{(1)}(I_{\widetilde{T}}^M \tilde{\varphi}) = \tilde{\chi}_{i,j}^{(1)}(\tilde{\varphi}), \quad 1 \le i < j \le d+1, \tag{17.44a}$$

$$\tilde{\chi}_i^{(2)}(I_{\widetilde{T}}^M \tilde{\varphi}) = \tilde{\chi}_i^{(2)}(\tilde{\varphi}), \quad 1 \le i \le d+1, \tag{17.44b}$$

and

$$I_T^M : W^{2,1}(T) \ni \varphi \mapsto I_T^M \varphi \in P, \tag{17.45}$$

with, for any $\varphi \in W^{2,1}(T)$,

$$\chi_{i,j}^{(1)}(I_T^M \varphi) = \chi_{i,j}^{(1)}(\varphi), \quad 1 \le i < j \le d+1, \tag{17.46a}$$

$$\chi_i^{(2)}(I_T^M \varphi) = \chi_i^{(2)}(\varphi), \quad 1 \le i \le d+1. \tag{17.46b}$$

Remark 17.7. The Morley FEM has not been defined uniquely. There are two versions: one defined in [46], which is the original paper, and the other in [6, 44, 51]. In original Morley FEM, by normal derivatives on faces, the spans of the nodes are not preserved under push-forward. To overcome this difficulty, the mean value of the first normal derivative is used [6, 44, 51]. The original Morley interpolation error estimates are obtained using the modified Morley interpolation error estimates (see [44]). In this study, we used the Morley FEM introduced in [51].

17.8 Local Morley Interpolation Error Estimates

Using the idea of [51, Lemma 1], the following lemma holds.

Lemma 17.8. Let $T \subset \mathbb{R}^d$ be a simplex. $n_{T,k}$ denotes the unit outer normal to the face F_k, $k = 1, \dots, d+1$ of T, $S_1, \dots, S_d$ are all $(d-2)$-dimensional subsimplexes of F_k. Let $v \in \mathcal{C}^1(T)$ be such that

$$\int_{S_\ell} v = 0, \quad \int_{F_k} \frac{\partial v}{\partial n_k} = 0, \tag{17.47}$$

for any $\ell = 1, \dots, d$ and $k = 1, \dots, d+1$. It then holds that

$$\int_{F_k} \frac{\partial v}{\partial x_i} = 0, \quad i = 1, \dots, d, \quad k = 1, \dots, d+1. \tag{17.48}$$

Proof. Let $v \in \mathcal{C}^1(T)$. Let $\xi \in \mathbb{R}^d$ be a constant vector, and let $\tau := \xi - (\xi \cdot n_{T,k}) n_{T,k}$. We have

$$\tau \cdot n_{T,k} = \xi \cdot n_{T,k} - (\xi \cdot n_{T,k}) n_{T,k} \cdot n_{T,k} = 0,$$

that is, τ is the tangent vector of F_k. Subsequently, from (17.47) we obtain

$$\int_{F_k} (\xi \cdot \nabla) v = \int_{F_k} \frac{\partial v}{\partial \tau} + (\xi \cdot n_{T,k}) \int_{F_k} \frac{\partial v}{\partial n_k} = \int_{F_k} \frac{\partial v}{\partial \tau}.$$

Let $d = 2$. Let p_{k1} and p_{k2} be the endpoints of the edge F_k, that is, $F_k = \overline{p_{k1} p_{k2}}$. Subsequently, from (17.47) we obtain

$$\int_{F_k} \frac{\partial v}{\partial \tau} = \int_{s=0}^{s=|F_k|} \frac{dv}{ds} \left(\frac{|F_k| - s}{|F_k|} p_{k1} + \frac{s}{|F_k|} p_{k2} \right) = v(p_{k2}) - v(p_{k1}) = 0. \tag{17.49}$$

Let $d = 3$. Let $\zeta^{(\ell)}$ be the unit outer normal of S_ℓ for $\ell = 1, 2, 3$, From (17.47), the Gauss–Green formula yields

$$\int_{F_k} \frac{\partial v}{\partial \tau} = \sum_{\ell=1}^{3} \tau \cdot \zeta^{(\ell)} \int_{S_\ell} v = 0. \tag{17.50}$$

From (17.49) and (17.50), it holds that for $d = 2, 3$

$$\int_{F_k} (\xi \cdot \nabla) v = 0. \tag{17.51}$$

Let $e_1, \dots, e_d \in \mathbb{R}^d$ be a canonical basis. By setting $\xi := e_i$ in (17.51), we obtain the desired result in (17.48) under Assumption (17.47). □

The anisotropic Morley interpolation error estimate is expressed as

Theorem 17.9. Let $p \in [1, \infty)$ and $q \in [1, \infty)$ be such that (16.5) holds true. Subsequently, for any $\varphi \in W^{3,p}(T) \cap \mathcal{C}^1(T)$,we have

$$|I_T^M \varphi - \varphi|_{W^{2,q}(T)} \le c |T|_d^{\frac{1}{q} - \frac{1}{p}} \sum_{i=1}^{d} h_i \left| \frac{\partial \varphi}{\partial r_i} \right|_{W^{2,p}(T)}. \tag{17.52}$$

If Condition 10.1 is imposed, then:

$$|I_T^M \varphi - \varphi|_{W^{2,q}(T)} \le c |T|_d^{\frac{1}{q} - \frac{1}{p}} \sum_{i=1}^{d} \widetilde{\mathscr{H}_i} \left| \frac{\partial (\varphi \circ \Phi_T)}{\partial \tilde{x}_i} \right|_{W^{2,p}(\Phi_T^{-1}(T))}. \tag{17.53}$$

Proof. Only Morley interpolation satisfies the condition (17.1).

Let $\varphi \in W^{3,p}(T) \cap \mathcal{C}^1(T)$ and set $v := I_T^M \varphi - \varphi$. Using the definition of the Morley interpolation operator (17.46), we obtain

$$\int_{S_{i,j}} v ds = 0, \quad 1 \le i < j \le d+1, \quad \int_{F_i} \frac{\partial v}{\partial n_i} ds = 0, \quad 1 \le i \le d+1.$$

Therefore, from Lemma 17.8, we have

$$\int_{F_i} \frac{\partial v}{\partial x_k} = 0, \quad i = 1, \dots, d+1, \quad k = 1, \dots, d. \tag{17.54}$$

From Green's formula and (17.54), it follows that, for $1 \le j, k \le d$,

$$\int_T \frac{\partial^2 v}{\partial x_j \partial x_k} dx = \sum_{i=1}^{d+1} n_{T,j} \int_{F_i} \frac{\partial v}{\partial x_k} ds = 0,$$

which leads to

$$\frac{\partial^2}{\partial x_j \partial x_k}(I_T^M \varphi) = \frac{1}{|T|_d} \int_T \frac{\partial^2}{\partial x_j \partial x_k}(I_T^M \varphi) dx = \frac{1}{|T|_d} \int_T \frac{\partial^2 \varphi}{\partial x_j \partial x_k} dx = \Pi_T^0 \left(\frac{\partial^2 \varphi}{\partial x_j \partial x_k} \right),$$

because $\frac{\partial^2}{\partial x_j \partial x_k}(I_T^M \varphi) \in \mathbb{P}^0(T)$, Therefore, by Theorem 17.1, the target inequalities (17.52) and (17.53) hold. □

17.9 Global Morley Interpolation Error Estimates

Recall the space V_h^n with $n = 2$ (see (14.7)). the Morley finite element space is as follows:

$$\begin{aligned} V_h^M := \Bigg\{ &\varphi_h \in V_h^2 : \int_F \left[\!\left[\frac{\partial \varphi_h}{\partial n} \right]\!\right] ds = 0 \ \forall F \in \mathcal{F}_h^i, \\ &\text{the integral average of } \varphi_h \text{ over each } (d-2)\text{-dimensional} \\ &\text{subsimplex of } T \in \mathbb{T}_h \text{ is continuous} \Bigg\}, \\ V_{h0}^M := \big\{ &\varphi_h \in V_h^M ; \text{ degrees of freedom of } \varphi_h \text{ in (17.36) vanish on } \partial\Omega \big\}. \end{aligned}$$

In particular, for $d = 2$, the space V_{h0}^M is described as

$$\begin{aligned} V_{h0}^M := \Bigg\{ &\varphi_h \in V_h^2 : \int_F \left[\!\left[\frac{\partial \varphi_h}{\partial n} \right]\!\right] ds = 0 \ \forall F \in \mathcal{F}_h, \\ &\varphi_h \text{ is continuous at each vertex in } \Omega, \ \varphi_h(p) = 0, \ p \in \partial\Omega \Bigg\}. \end{aligned}$$

We also define the global interpolation $I_h^M : W^{2,1}(\Omega) \to V_h^M$ (or $I_h^M : W_0^{2,1}(\Omega) \to V_{h0}^M$) as follows.

$$(I_h^M \varphi)|_T := I_T^M (\varphi|_T) \quad \forall T \in \mathbb{T}_h, \ \forall \varphi \in W^{2,1}(\Omega).$$

Corollary 17.10. Suppose that the assumptions of Theorem 17.9 are satisfied. Let I_h^M be the corresponding global Morley interpolation operator. It then holds that, for any $\varphi \in W^{3,p}(\Omega) \cap \mathcal{C}^1(\overline{\Omega})$,

(I) if Condition 10.1 is not imposed,

$$|I_T^M \varphi - \varphi|_{W^{2,q}(\mathbb{T}_h)} \leq c \sum_{T \in \mathbb{T}_h} |T|_d^{\frac{1}{q}-\frac{1}{p}} \sum_{i=1}^{d} h_i \left| \frac{\partial \varphi}{\partial r_i} \right|_{W^{2,p}(T)}, \tag{17.55}$$

(II) if Condition 10.1 is imposed,

$$|I_T^M \varphi - \varphi|_{W^{2,q}(\mathbb{T}_h)} \leq c \sum_{T \in \mathbb{T}_h} |T|_d^{\frac{1}{q}-\frac{1}{p}} \sum_{i=1}^{d} \widetilde{\mathscr{H}_i} \left| \frac{\partial (\varphi \circ \Phi_T)}{\partial \tilde{x}_i} \right|_{W^{2,p}(\Phi_T^{-1}(T))}. \tag{17.56}$$

Proof. This corollary is proved in the same argument as Corollary 14.2. □

17.10 Another Estimate

Theorem 17.11. Let $T \subset \mathbb{R}^d$ be a simplex. Let $I_T^M : H^2(T) \to \mathbb{P}^2(T)$ be the local Morley interpolation operator defined as

$$I_T^M : H^2(T) \ni \varphi \mapsto I_T^M \varphi \in \mathbb{P}^2(T),$$

with

$$\begin{aligned} \chi_{i,j}^{(1)}(I_T^M \varphi) &= \chi_{i,j}^{(1)}(\varphi), \quad 1 \leq i < j \leq d+1, \\ \chi_i^{(2)}(I_T^M \varphi) &= \chi_i^{(2)}(\varphi), \quad 1 \leq i \leq d+1. \end{aligned}$$

for any $\varphi \in H^2(T)$. It then holds that

$$|I_T^M \varphi - \varphi|_{H^2(T)} \leq \frac{h_T}{\pi} |\varphi|_{H^3(T)} \quad \forall \varphi \in H^3(T). \tag{17.58}$$

Proof. Using (16.18).

$$\begin{aligned} |I_T^M \varphi - \varphi|_{H^2(T)}^2 &= \sum_{j,k=1}^{d} \left\| \frac{\partial^2}{\partial x_j \partial x_k} (I_T^M \varphi - \varphi) \right\|_{L^2(T)}^2 \\ &= \sum_{j,k=1}^{d} \left\| \Pi_T^0 \left(\frac{\partial^2 \varphi}{\partial x_j \partial x_k} \right) - \left(\frac{\partial^2 \varphi}{\partial x_j \partial x_k} \right) \right\|_{L^2(T)}^2 \\ &\leq \left(\frac{h_T}{\pi} \right)^2 \sum_{j,k=1}^{d} \left| \frac{\partial^2 \varphi}{\partial x_j \partial x_k} \right|_{H^1(T)}^2 = \left(\frac{h_T}{\pi} \right)^2 |\varphi|_{H^3(T)}^2, \end{aligned}$$

which conclude (17.58). □

18 New Scaling Argument: Part 2

18.1 Two-step Piola Transforms

We adopt the following two-step Piola transformations.

Definition 18.1 (Two-step Piola transforms)**.** Let $V(\widehat{T}) := \mathcal{C}(\widehat{T})^d$. The Piola transformation $\Psi := \Psi_{\widetilde{T}} \circ \Psi_{\widehat{T}} : V(\widehat{T}) \to V(T)$ is defined as

$$\begin{aligned} \Psi : V(\widehat{T}) &\to V(T) \\ \hat{v} \mapsto v(x) &:= \Psi(\hat{v})(x) = \frac{1}{\det(A)} A\hat{v}(\hat{x}), \end{aligned} \tag{18.1}$$

with two Piola transformations:

$$\begin{aligned} \Psi_{\widehat{T}} : V(\widehat{T}) &\to V(\widetilde{T}) \\ \hat{v} \mapsto \tilde{v}(\tilde{x}) &:= \Psi_{\widehat{T}}(\hat{v})(\tilde{x}) := \frac{1}{\det(A_{\widetilde{T}})} A_{\widetilde{T}}\hat{v}(\hat{x}), \\ \Psi_{\widetilde{T}} : V(\widetilde{T}) &\to V(T) \\ \tilde{v} \mapsto v(x) &:= \Psi_{\widetilde{T}}(\tilde{v})(x) := \frac{1}{\det(A_T)} A_T\tilde{v}(\tilde{x}). \end{aligned}$$

18.2 Property of the Piola Transformations

Lemma 18.2. If $\hat{v} \in \mathcal{C}^1(\widehat{T})^d$, then $v := \Psi\hat{v} \in \mathcal{C}^1(T)^d$ and it holds

$$\begin{aligned} J_x v &= \frac{1}{\det(A)} A\widehat{J}_{\hat{x}}\hat{v}A^{-1}, \\ \operatorname{div} v &= \frac{1}{\det(A)} \widehat{\operatorname{div}}\hat{v}, \end{aligned}$$

where $J_x v$ and $\widehat{J}_{\hat{x}}\hat{v}$ denote the Jacobian matrixes of v and $\hat{v}$, respectively.

Proof. From the definition of the Piola transformation (18.1), we have

$$\begin{aligned} J_x v(x) &= \frac{1}{\det(A)} AJ_x(\hat{v} \circ \Phi^{-1})(x) = \frac{1}{\det(A)} A\widehat{J}_{\hat{x}}\hat{v}(\hat{x})J_x\Phi^{-1}(x) \\ &= \frac{1}{\det(A)} A\widehat{J}_{\hat{x}}\hat{v}(\hat{x})A^{-1}. \end{aligned}$$

Due to the property of the trace, we get

$$\begin{aligned} \operatorname{div} v = \operatorname{Tr}(J_x v) &= \frac{1}{\det(A)} \operatorname{Tr}(A\widehat{J}_{\hat{x}}\hat{v}A^{-1}) \\ &= \frac{1}{\det(A)} \operatorname{Tr}(\widehat{J}_{\hat{x}}\hat{v}) = \frac{1}{\det(A)} \widehat{\operatorname{div}}\hat{v}. \end{aligned}$$

□

Lemma 18.3. For $\hat{\varphi} \in \mathcal{C}^1(\widehat{T})$, $\hat{v} \in \mathcal{C}^1(\widehat{T})^d$ with $\varphi := \hat{\varphi} \circ \Phi^{-1}$ and $v := \Psi(\hat{v})$, it holds that

$$\int_T \operatorname{div} v\varphi dx = \int_{\widehat{T}} \widehat{\operatorname{div}}\hat{v}\hat{\varphi} d\hat{x}, \tag{18.2}$$

$$\int_T (v \cdot \nabla_x)\varphi dx = \int_{\widehat{T}} (\hat{v} \cdot \widehat{\nabla}_{\hat{x}})\hat{\varphi} d\hat{x}, \tag{18.3}$$

$$\int_{\partial T} (v \cdot n)\varphi ds = \int_{\partial\widehat{T}} (\hat{v} \cdot \hat{n})\hat{\varphi} d\hat{s}. \tag{18.4}$$

Proof. Because $\det(A)$ is positive, by a change a variable,

$$\int_T \operatorname{div} v\varphi dx = \frac{1}{\det(A)}\int_{\widehat{T}} \widehat{\operatorname{div}}\hat{v}\hat{\varphi}|\det(A)|d\hat{x},$$

which leads to (18.2). Because

$$\nabla_x\varphi = A^{-T}\widehat{\nabla}_{\hat{x}}\hat{\varphi},$$

we have

$$\begin{aligned}\int_T (v\cdot\nabla_x)\varphi dx &= \frac{1}{\det(A)}\int_{\widehat{T}}(A\hat{v}\cdot A^{-T}\widehat{\nabla}_{\hat{x}})\hat{\varphi}|\det(A)|d\hat{x}\\ &= \int_{\widehat{T}}[(A\hat{v})^T A^{-T}\widehat{\nabla}_{\hat{x}}]\hat{\varphi}d\hat{x} = \int_{\widehat{T}}(\hat{v}\cdot\widehat{\nabla}_{\hat{x}})\hat{\varphi}d\hat{x},\end{aligned}$$

which is (18.3).

From (18.2) and (18.3), applying the Gauss–Green formula yields

$$\begin{aligned}\int_{\partial T}(v\cdot n)\varphi ds &= \int_T(\operatorname{div} v)\varphi dx + \int_T (v\cdot\nabla_x)\varphi dx\\ &= \int_{\widehat{T}}(\widehat{\operatorname{div}}\hat{v})\hat{\varphi}d\hat{x} + \int_{\widehat{T}}(\hat{v}\cdot\widehat{\nabla}_{\hat{x}})\hat{\varphi}d\hat{x} = \int_{\partial\widehat{T}}(\hat{v}\cdot\hat{n})\hat{\varphi}d\hat{s},\end{aligned}$$

which is (18.4). $\square$

18.3 Preliminalies

18.3.1 Calculations 1

We use the following calculations in (18.6). Let $\hat{v}\in\mathcal{C}^2(\widehat{T})^d$ with $\tilde{v} = \Psi_{\widetilde{T}}\hat{v}$ and $v = \Psi_{T}\tilde{v}$. Using the definition of Piola transformations (Definition 18.1) yields, for $1\le i,k\le d$,

$$\begin{aligned}\frac{\partial\hat{v}_k}{\partial\hat{x}_i} &= \det(A_{\widetilde{T}})\sum_{\eta=1}^d [A_{\widetilde{T}}^{-1}]_{k\eta}\sum_{i_1^{(1)}=1}^d \frac{\partial\tilde{v}_\eta}{\partial\tilde{x}_{i_1^{(1)}}}\frac{\partial\tilde{x}_{i_1^{(1)}}}{\partial\hat{x}_i}\\ &= \det(A_{\widetilde{T}})\sum_{\eta=1}^d [A_{\widetilde{T}}^{-1}]_{k\eta}\sum_{i_1^{(1)}=1}^d \frac{\partial\tilde{v}_\eta}{\partial\tilde{x}_{i_1^{(1)}}}[A_{\widetilde{T}}]_{i_1^{(1)}i}\\ &= \det(A_{\widetilde{T}})h_i h_k^{-1}\sum_{\eta=1}^d[\widetilde{A}^{-1}]_{k\eta}\sum_{i_1^{(1)}=1}^d \frac{\partial\tilde{v}_\eta}{\partial\tilde{x}_{i_1^{(1)}}}[\widetilde{A}]_{i_1^{(1)}i},\\ \frac{\partial\tilde{v}_\eta}{\partial\tilde{x}_{i_1^{(1)}}} &= \det(A_T)\sum_{\nu=1}^d[A_T^{-1}]_{\eta\nu}\sum_{i_1^{(0,1)}=1}^d\frac{\partial v_\nu}{\partial x_{i_1^{(0,1)}}}[A_T]_{i_1^{(0,1)}i_1^{(1)}},\end{aligned}$$

which leads to

$$\frac{\partial\hat{v}_k}{\partial\hat{x}_i} = \det(A_{\widetilde{T}})\det(A_T)h_k^{-1}\sum_{\eta,\nu=1}^d[\widetilde{A}^{-1}]_{k\eta}[A_T^{-1}]_{\eta\nu}\sum_{i_1^{(1)},i_1^{(0,1)}=1}^d h_i[\widetilde{A}]_{i_1^{(1)}i}[A_T]_{i_1^{(0,1)}i_1^{(1)}}\frac{\partial v_\nu}{\partial x_{i_1^{(0,1)}}}.$$

By an analogous calculation, for $1 \le i, j, k \le d$,

$$\frac{\partial^2 \hat{v}_k}{\partial \hat{x}_i \partial \hat{x}_j} = \det(A_{\widetilde{T}}) h_i h_j h_k^{-1} \sum_{\eta=1}^{d} [\widetilde{A}^{-1}]_{k\eta} \sum_{i_1^{(1)}, j_1^{(1)}=1}^{d} \frac{\partial^2 \tilde{v}_\eta}{\partial \tilde{x}_{i_1^{(1)}} \partial \tilde{x}_{j_1^{(1)}}} [\widetilde{A}]_{i_1^{(1)} i} [\widetilde{A}]_{j_1^{(1)} j},$$

$$\frac{\partial^2 \tilde{v}_\eta}{\partial \tilde{x}_{i_1^{(1)}} \partial \tilde{x}_{j_1^{(1)}}} = \det(A_T) \sum_{\nu=1}^{d} [A_T^{-1}]_{\eta\nu} \sum_{i_1^{(0,1)}, j_1^{(0,1)}=1}^{d} \frac{\partial^2 v_\nu}{\partial x_{i_1^{(0,1)}} \partial x_{j_1^{(0,1)}}} [A_T]_{i_1^{(0,1)} i_1^{(1)}} [A_T]_{j_1^{(0,1)} j_1^{(1)}},$$

which leads to

$$\frac{\partial^2 \hat{v}_k}{\partial \hat{x}_i \partial \hat{x}_j} = \det(A_{\widetilde{T}}) \det(A_T) h_k^{-1} \sum_{\eta,\nu=1}^{d} [\widetilde{A}^{-1}]_{k\eta} [A_T^{-1}]_{\eta\nu}$$
$$\sum_{i_1^{(1)}, i_1^{(0,1)}=1}^{d} h_i [\widetilde{A}]_{i_1^{(1)} i} [A_T]_{i_1^{(0,1)} i_1^{(1)}} \sum_{j_1^{(1)}, j_1^{(0,1)}=1}^{d} h_j [\widetilde{A}]_{j_1^{(1)} j} [A_T]_{j_1^{(0,1)} j_1^{(1)}} \frac{\partial^2 v_\nu}{\partial x_{i_1^{(0,1)}} \partial x_{j_1^{(0,1)}}}.$$

For any multi-indices β and γ, for $1 \le k \le d$, Let $\hat{v} \in \mathcal{C}^{|\beta|+|\gamma|}(\widehat{T})^d$ with $\tilde{v} = \Psi_{\widehat{T}} \hat{v}$ and $v = \Psi_{\widetilde{T}} \tilde{v}$. Then,

$$\partial_{\hat{x}}^{\beta+\gamma} \hat{v}_k = \frac{\partial^{|\beta|+|\gamma|}}{\partial \hat{x}_1^{\beta_1} \cdots \partial \hat{x}_d^{\beta_d} \partial \hat{x}_1^{\gamma_1} \cdots \partial \hat{x}_d^{\gamma_d}} \hat{v}_k$$
$$= \det(A_{\widetilde{T}}) \det(A_T) h_k^{-1} \sum_{\eta,\nu=1}^{d} [\widetilde{A}^{-1}]_{k\eta} [A_T^{-1}]_{\eta\nu}$$
$$\underbrace{\sum_{i_1^{(1)}, i_1^{(0,1)}=1}^{d} h_1 [\widetilde{A}]_{i_1^{(1)} 1} [A_T]_{i_1^{(0,1)} i_1^{(1)}} \cdots \sum_{i_{\beta_1}^{(1)}, i_{\beta_1}^{(0,1)}=1}^{d} h_1 [\widetilde{A}]_{i_{\beta_1}^{(1)} 1} [A_T]_{i_{\beta_1}^{(0,1)} i_{\beta_1}^{(1)}} \cdots}_{\beta_1 \text{times}}$$
$$\underbrace{\sum_{i_1^{(d)}, i_1^{(0,d)}=1}^{d} h_d [\widetilde{A}]_{i_1^{(d)} d} [A_T]_{i_1^{(0,d)} i_1^{(d)}} \cdots \sum_{i_{\beta_d}^{(d)}, i_{\beta_d}^{(0,d)}=1}^{d} h_d [\widetilde{A}]_{i_{\beta_d}^{(d)} d} [A_T]_{i_{\beta_d}^{(0,d)} i_{\beta_d}^{(d)}}}_{\beta_d \text{times}}$$
$$\underbrace{\sum_{j_1^{(1)}, j_1^{(0,1)}=1}^{d} h_1 [\widetilde{A}]_{j_1^{(1)} 1} [A_T]_{j_1^{(0,1)} j_1^{(1)}} \cdots \sum_{j_{\gamma_1}^{(1)}, j_{\gamma_1}^{(0,1)}=1}^{d} h_1 [\widetilde{A}]_{j_{\gamma_1}^{(1)} 1} [A_T]_{j_{\gamma_1}^{(0,1)} j_{\gamma_1}^{(1)}} \cdots}_{\gamma_1 \text{times}}$$
$$\underbrace{\sum_{j_1^{(d)}, j_1^{(0,d)}=1}^{d} h_d [\widetilde{A}]_{j_1^{(d)} d} [A_T]_{j_1^{(0,d)} j_1^{(d)}} \cdots \sum_{j_{\gamma_d}^{(d)}, j_{\gamma_d}^{(0,d)}=1}^{d} h_d [\widetilde{A}]_{j_{\gamma_d}^{(d)} d} [A_T]_{j_{\gamma_d}^{(0,d)} j_{\gamma_d}^{(d)}}}_{\gamma_d \text{times}}$$
$$\underbrace{\frac{\partial^{\beta_1}}{\partial x_{i_1^{(0,1)}} \cdots \partial x_{i_{\beta_1}^{(0,1)}}}}_{\beta_1 \text{times}} \cdots \underbrace{\frac{\partial^{\beta_d}}{\partial x_{i_1^{(0,d)}} \cdots \partial x_{i_{\beta_d}^{(0,d)}}}}_{\beta_d \text{times}} \underbrace{\frac{\partial^{\gamma_1}}{\partial x_{j_1^{(0,1)}} \cdots \partial x_{j_{\gamma_1}^{(0,1)}}}}_{\gamma_1 \text{times}} \cdots \underbrace{\frac{\partial^{\gamma_d}}{\partial x_{j_1^{(0,d)}} \cdots \partial x_{j_{\gamma_d}^{(0,d)}}}}_{\gamma_d \text{times}} v_\nu.$$

Using (1.1), (5.7), (6.8c) and (6.9), we have, for $1 \le i,k \le d$,

$$\left|\frac{\partial \hat{v}_k}{\partial \hat{x}_i}\right| \le |\det(A_{\widetilde{T}})||\det(A_T)|h_k^{-1} \sum_{\eta,\nu=1}^{d} |[\widetilde{A}^{-1}]_{k\eta}||[A_T^{-1}]_{\eta\nu}| \left| h_i \sum_{i_1^{(1)}, i_1^{(0,1)}=1}^{d} [A_T]_{i_1^{(0,1)} i_1^{(1)}} (\tilde{r}_i)_{i_1^{(1)}} \frac{\partial v_\nu}{\partial x_{i_1^{(0,1)}}} \right|$$

$$\le c |\det(A_{\widetilde{T}})| h_k^{-1} \|\widetilde{A}^{-1}\|_2 \sum_{\nu=1}^{d} h_i \left| \frac{\partial v_\nu}{\partial r_i} \right|,$$

and, for $1 \le i,j,k \le d$,

$$\left|\frac{\partial^2 \hat{v}_k}{\partial \hat{x}_i \partial \hat{x}_j}\right| \le |\det(A_{\widetilde{T}})||\det(A_T)|h_k^{-1} \sum_{\eta,\nu=1}^{d} |[\widetilde{A}^{-1}]_{k\eta}||[A_T^{-1}]_{\eta\nu}|$$

$$\left| \sum_{i_1^{(1)}, i_1^{(0,1)}=1}^{d} h_i [\widetilde{A}]_{i_1^{(1)} i} [A_T]_{i_1^{(0,1)} i_1^{(1)}} \sum_{j_1^{(1)}, j_1^{(0,1)}=1}^{d} h_j [\widetilde{A}]_{j_1^{(1)} j} [A_T]_{j_1^{(0,1)} j_1^{(1)}} \frac{\partial^2 v_\nu}{\partial x_{i_1^{(0,1)}} \partial x_{j_1^{(0,1)}}} \right|$$

$$\le c |\det(A_{\widetilde{T}})| h_k^{-1} \|\widetilde{A}^{-1}\|_2 \sum_{\nu=1}^{d} h_i h_j \left| \frac{\partial^2 v_\nu}{\partial r_i \partial r_j} \right|.$$

18.3.2 Calculations 2

We use the following calculations in (18.7). Let $\hat{v} \in \mathcal{C}^2(\widehat{T})^d$ with $\tilde{v} = \Psi_{\widehat{T}} \hat{v}$ and $v = \Psi_{\widetilde{T}} \tilde{v}$. Using the definition of Piola transformations (Definition 18.1) yields, for $1 \le i,k \le d$,

$$\frac{\partial \hat{v}_k}{\partial \hat{x}_i} = \det(A_{\widetilde{T}}) h_k^{-1} \sum_{\eta=1}^{d} [\widetilde{A}^{-1}]_{k\eta} \sum_{i_1^{(1)}=1}^{d} h_i [\widetilde{A}]_{i_1^{(1)} i} \frac{\partial \tilde{v}_\eta}{\partial \tilde{x}_{i_1^{(1)}}},$$

and, for $1 \le i,j,k \le d$,

$$\frac{\partial^2 \hat{v}_k}{\partial \hat{x}_i \partial \hat{x}_j} = \det(A_{\widetilde{T}}) h_k^{-1} \sum_{\eta=1}^{d} [\widetilde{A}^{-1}]_{k\eta} \sum_{i_1^{(1)}=1}^{d} h_i [\widetilde{A}]_{i_1^{(1)} i} \sum_{j_1^{(1)}=1}^{d} h_j [\widetilde{A}]_{j_1^{(1)} j} \frac{\partial^2 \tilde{v}_\eta}{\partial \tilde{x}_{i_1^{(1)}} \partial \tilde{x}_{j_1^{(1)}}}.$$

For any multi-indices β and γ, for $1 \le k \le d$, Let $\hat{v} \in \mathcal{C}^{|\beta|+|\gamma|}(\widehat{T})^d$ with $\tilde{v} = \Psi_{\widehat{T}} \hat{v}$ and $v = \Psi_{\widetilde{T}} \tilde{v}$. Then,

$$\partial_{\hat{x}}^{\beta+\gamma} \hat{v}_k = \frac{\partial^{|\beta|+|\gamma|}}{\partial \hat{x}_1^{\beta_1} \cdots \partial \hat{x}_d^{\beta_d} \partial \hat{x}_1^{\gamma_1} \cdots \partial \hat{x}_d^{\gamma_d}} \hat{v}_k$$

$$= \det(A_{\widetilde{T}}) h_k^{-1} \sum_{\eta=1}^{d} [\widetilde{A}^{-1}]_{k\eta}$$

$$\underbrace{\sum_{i_1^{(1)}=1}^{d} h_1 [\widetilde{A}]_{i_1^{(1)} 1} \cdots \sum_{i_{\beta_1}^{(1)}=1}^{d} h_1 [\widetilde{A}]_{i_{\beta_1}^{(1)} 1}}_{\beta_1 \text{times}} \cdots \underbrace{\sum_{i_1^{(d)}=1}^{d} h_d [\widetilde{A}]_{i_1^{(d)} d} \cdots \sum_{i_{\beta_d}^{(d)}=1}^{d} h_d [\widetilde{A}]_{i_{\beta_d}^{(d)} d}}_{\beta_d \text{times}}$$

$$\underbrace{\sum_{j_1^{(1)}=1}^{d} h_1 [\widetilde{A}]_{j_1^{(1)} 1} \cdots \sum_{j_{\gamma_1}^{(1)}=1}^{d} h_1 [\widetilde{A}]_{j_{\gamma_1}^{(1)} 1}}_{\gamma_1 \text{times}} \cdots \underbrace{\sum_{j_1^{(d)}=1}^{d} h_d [\widetilde{A}]_{j_1^{(d)} d} \cdots \sum_{j_{\gamma_d}^{(d)}=1}^{d} h_d [\widetilde{A}]_{j_{\gamma_d}^{(d)} d}}_{\gamma_d \text{times}}$$

$$\underbrace{\frac{\partial^{\beta_1}}{\partial \tilde{x}_{i_1^{(1)}} \cdots \partial \tilde{x}_{i_{\beta_1}^{(1)}}}}_{\beta_1 \text{times}} \cdots \underbrace{\frac{\partial^{\beta_d}}{\partial \tilde{x}_{i_1^{(d)}} \cdots \partial \tilde{x}_{i_{\beta_d}^{(d)}}}}_{\beta_d \text{times}} \underbrace{\frac{\partial^{\gamma_1}}{\partial \tilde{x}_{j_1^{(1)}} \cdots \partial \tilde{x}_{j_{\gamma_1}^{(1)}}}}_{\gamma_1 \text{times}} \cdots \underbrace{\frac{\partial^{\gamma_d}}{\partial \tilde{x}_{j_1^{(d)}} \cdots \partial \tilde{x}_{j_{\gamma_d}^{(d)}}}}_{\gamma_d \text{times}} \tilde{v}_\eta.$$

Using Section 10.1.1 and (1.1), we have, for $1 \leq i, k \leq d$,

$$\begin{aligned}
\left| \frac{\partial \hat{v}_k}{\partial \hat{x}_i} \right| &\leq |\det(A_{\widetilde{T}})| h_k^{-1} \sum_{\eta=1}^{d} |[\widetilde{A}^{-1}]_{k\eta}| \sum_{i_1^{(1)}=1}^{d} h_i |[\widetilde{A}]_{i_1^{(1)} i}| \left| \frac{\partial \tilde{v}_\eta}{\partial \tilde{x}_{i_1^{(1)}}} \right| \\
&\leq c |\det(A_{\widetilde{T}})| h_k^{-1} \|\widetilde{A}^{-1}\|_2 \sum_{\eta=1}^{d} \sum_{i_1^{(1)}=1}^{d} \widetilde{\mathscr{H}}_{i_1^{(1)}} \left| \frac{\partial \tilde{v}_\eta}{\partial \tilde{x}_{i_1^{(1)}}} \right|,
\end{aligned}$$

and, for $1 \leq i, j, k \leq d$,

$$\begin{aligned}
\left| \frac{\partial^2 \hat{v}_k}{\partial \hat{x}_i \partial \hat{x}_j} \right| &\leq |\det(A_{\widetilde{T}})| h_k^{-1} \sum_{\eta=1}^{d} |[\widetilde{A}^{-1}]_{k\eta}| \sum_{i_1^{(1)}=1}^{d} h_i |[\widetilde{A}]_{i_1^{(1)} i}| \sum_{j_1^{(1)}=1}^{d} h_j |[\widetilde{A}]_{j_1^{(1)} j}| \left| \frac{\partial^2 \tilde{v}_\eta}{\partial \tilde{x}_{i_1^{(1)}} \partial \tilde{x}_{j_1^{(1)}}} \right| \\
&\leq c |\det(A_{\widetilde{T}})| h_k^{-1} \|\widetilde{A}^{-1}\|_2 \sum_{\eta=1}^{d} \sum_{i_1^{(1)}, j_1^{(1)}=1}^{d} \widetilde{\mathscr{H}}_{i_1^{(1)}} \widetilde{\mathscr{H}}_{j_1^{(1)}} \left| \frac{\partial^2 \tilde{v}_\eta}{\partial \tilde{x}_{i_1^{(1)}} \partial \tilde{x}_{j_1^{(1)}}} \right|.
\end{aligned}$$

18.4 Main Results

Lemma 18.4. Let $p \in [1, \infty)$. Let $T \in \mathbb{T}_h$ satisfy Condition 5.1 or Condition 5.2 with $T = \Phi_T(\widetilde{T})$ and $\widetilde{T} = \Phi_{\widetilde{T}}(\widehat{T})$. Then, for any $\hat{v} = (\hat{v}_1, \ldots, \hat{v}_d)^\top \in L^p(\widehat{T})^d$ with $\tilde{v} = (\tilde{v}_1, \ldots, \tilde{v}_d)^\top := \Psi_{\widehat{T}} \hat{v}$ and $v = (v_1, \ldots, v_d)^\top := \Psi_{\widetilde{T}} \tilde{v}$,

$$\|v\|_{L^p(T)^d} \leq c |\det(A_{\widetilde{T}})|^{\frac{1-p}{p}} \|\widetilde{A}\|_2 \left(\sum_{j=1}^{d} h_j^p \|\hat{v}_j\|_{L^p(\widehat{T})}^p \right)^{1/p}. \tag{18.5}$$

Proof. Because the space $\mathcal{C}(\widehat{T})^d$ is dense in the space $L^p(\widehat{T})^d$, we show (18.5) for $\hat{v} \in \mathcal{C}(\widehat{T})^d$ with $\tilde{v} = \Psi_{\widehat{T}} \hat{v}$ and $v = \Psi_{\widetilde{T}} \tilde{v}$. From the definition of the Piola transformation, for $i = 1, \ldots, d$,

$$v_i(x) = \frac{1}{\det(A_T)} \sum_{k=1}^{d} [A_T]_{ik} \tilde{v}_k(\tilde{x}), \quad \tilde{v}_i(\tilde{x}) = \frac{1}{\det(A_{\widetilde{T}})} \sum_{j=1}^{d} [\widetilde{A}]_{ij} h_j \hat{v}_j(\hat{x}),$$

which leads to

$$v_i(x) = \frac{1}{\det(A_T) \det(A_{\widetilde{T}})} \sum_{j,k=1}^{d} [A_T]_{ik} [\widetilde{A}]_{kj} h_j \hat{v}_j(\hat{x}).$$

If $1 \leq p < \infty$, using (1.1), (6.8c) and (6.9),

$$\|v\|_{L^p(T)^d}^p = \sum_{i=1}^{d} \|v_i\|_{L^p(T)}^p \leq c |\det(A_{\widetilde{T}})|^{1-p} \|\widetilde{A}\|_2^p \sum_{j=1}^{d} h_j^p \|\hat{v}_j\|_{L^p(\widehat{T})}^p,$$

which leads to (18.5). □

Lemma 18.5. Let $p \in [1, \infty)$. Let $T \in \mathbb{T}_h$ satisfy Condition 5.1 or Condition 5.2 with $T = \Phi_T(\widetilde{T})$ and $\widetilde{T} = \Phi_{\widetilde{T}}(\widehat{T})$. Let $\ell, m \in \mathbb{N}_0$ and $k \in \mathbb{N}$ with $1 \le k \le d$. Let $\beta := (\beta_1, \dots, \beta_d) \in \mathbb{N}_0^d$ and $\gamma := (\gamma_1, \dots, \gamma_d) \in \mathbb{N}_0^d$ be multi-indices with $|\beta| = \ell$ and $|\gamma| = m$, respectively. Then, for any $\hat{v} = (\hat{v}_1, \dots, \hat{v}_d)^\top \in W^{|\beta|+|\gamma|,p}(\widehat{T})^d$ with $\tilde{v} = (\tilde{v}_1, \dots, \tilde{v}_d)^\top := \Psi_{\widehat{T}} \hat{v}$ and $v = (v_1, \dots, v_d)^\top := \Psi_{\widetilde{T}} \tilde{v}$,

$$\left\| \partial_{\hat{x}}^\beta \partial_{\hat{x}}^\gamma \hat{v}_k \right\|_{L^p(\widehat{T})} \le c |\det(A_{\widetilde{T}})|^{\frac{p-1}{p}} h_k^{-1} \|\widetilde{A}^{-1}\|_2 \sum_{|\varepsilon| = \ell + m} h^\varepsilon \left\| \partial_r^\varepsilon v \right\|_{L^p(T)^d}. \tag{18.6}$$

If Condition 10.1 is imposed, then

$$\left\| \partial_{\hat{x}}^\beta \partial_{\hat{x}}^\gamma \hat{v}_k \right\|_{L^p(\widehat{T})} \le c |\det(A_{\widetilde{T}})|^{\frac{p-1}{p}} h_k^{-1} \|\widetilde{A}^{-1}\|_2 \sum_{|\varepsilon| = \ell + m} \widetilde{\mathscr{H}^\varepsilon} \| \partial_{\tilde{x}}^\varepsilon (\Psi_{\widetilde{T}}^{-1} v) \|_{L^p(\Phi_T^{-1}(T))^d}. \tag{18.7}$$

Proof. Because the space $\mathcal{C}^{\ell+m}(\widehat{T})^d$ is dense in the space $W^{\ell+m,p}(\widehat{T})^d$, we show (18.6) and (18.7) for $\hat{v} \in \mathcal{C}^{\ell+m}(\widehat{T})^d$ with $\tilde{v} = \Psi_{\widehat{T}} \hat{v}$ and $v = \Psi_{\widetilde{T}} \tilde{v}$.

Proof of (18.6). Using (1.1), (6.8c) and (6.9), through a simple calculation, we have, for $1 \le k \le d$,

$$\begin{aligned}
&|\partial_{\hat{x}}^{\beta+\gamma} \hat{v}_k| = \left| \frac{\partial^{|\beta|+|\gamma|}}{\partial \hat{x}_1^{\beta_1} \cdots \partial \hat{x}_d^{\beta_d} \partial \hat{x}_1^{\gamma_1} \cdots \partial \hat{x}_d^{\gamma_d}} \hat{v}_k \right| \\
&\le c |\det(A_{\widetilde{T}})| \|\widetilde{A}^{-1}\|_2 h_k^{-1} \sum_{\nu=1}^d h^\beta h^\gamma \\
&\Bigg| \underbrace{\sum_{i_1^{(1)}, i_1^{(0,1)}=1}^d [A_T]_{i_1^{(0,1)} i_1^{(1)}} (\tilde{r}_1)_{i_1^{(1)}} \cdots \sum_{i_{\beta_1}^{(1)}, i_{\beta_1}^{(0,1)}=1}^d [A_T]_{i_{\beta_1}^{(0,1)} i_{\beta_1}^{(1)}} (\tilde{r}_1)_{i_{\beta_1}^{(1)}}}_{\beta_1 \text{times}} \cdots \\
&\underbrace{\sum_{i_1^{(d)}, i_1^{(0,d)}=1}^d [A_T]_{i_1^{(0,d)} i_1^{(d)}} (\tilde{r}_d)_{i_1^{(d)}} \cdots \sum_{i_{\beta_d}^{(d)}, i_{\beta_d}^{(0,d)}=1}^d [A_T]_{i_{\beta_d}^{(0,d)} i_{\beta_d}^{(d)}} (\tilde{r}_d)_{i_{\beta_d}^{(d)}}}_{\beta_d \text{times}} \\
&\underbrace{\sum_{j_1^{(1)}, j_1^{(0,1)}=1}^d [A_T]_{j_1^{(0,1)} j_1^{(1)}} (\tilde{r}_1)_{j_1^{(1)}} \cdots \sum_{j_{\gamma_1}^{(1)}, j_{\gamma_1}^{(0,1)}=1}^d [A_T]_{j_{\gamma_1}^{(0,1)} j_{\gamma_1}^{(1)}} (\tilde{r}_1)_{j_{\gamma_1}^{(1)}}}_{\gamma_1 \text{times}} \cdots \\
&\underbrace{\sum_{j_1^{(d)}, j_1^{(0,d)}=1}^d [A_T]_{j_1^{(0,d)} j_1^{(d)}} (\tilde{r}_d)_{j_1^{(d)}} \cdots \sum_{j_{\gamma_d}^{(d)}, j_{\gamma_d}^{(0,d)}=1}^d [A_T]_{j_{\gamma_d}^{(0,d)} j_{\gamma_d}^{(d)}} (\tilde{r}_d)_{j_{\gamma_d}^{(d)}}}_{\gamma_d \text{times}} \\
&\underbrace{\frac{\partial^{\beta_1}}{\partial x_{i_1^{(0,1)}} \cdots \partial x_{i_{\beta_1}^{(0,1)}}}}_{\beta_1 \text{times}} \cdots \underbrace{\frac{\partial^{\beta_d}}{\partial x_{i_1^{(0,d)}} \cdots \partial x_{i_{\beta_d}^{(0,d)}}}}_{\beta_d \text{times}} \underbrace{\frac{\partial^{\gamma_1}}{\partial x_{j_1^{(0,1)}} \cdots \partial x_{j_{\gamma_1}^{(0,1)}}}}_{\gamma_1 \text{times}} \cdots \underbrace{\frac{\partial^{\gamma_d}}{\partial x_{j_1^{(0,d)}} \cdots \partial x_{j_{\gamma_d}^{(0,d)}}}}_{\gamma_d \text{times}} v_\nu \Bigg| \\
&\le c |\det(A_{\widetilde{T}})| \|\widetilde{A}^{-1}\|_2 h_k^{-1} \sum_{\nu=1}^d \sum_{|\varepsilon| = |\beta| + |\gamma|} h^\varepsilon |\partial_r^\varepsilon v_\nu|.
\end{aligned}$$

Because $1 \le p < \infty$, it holds that, for $1 \le k \le d$,

$$\left\| \partial_{\hat{x}}^\beta \partial_{\hat{x}}^\gamma \hat{v}_k \right\|_{L^p(\widehat{T})}^p \le c |\det(A_{\widetilde{T}})|^{p-1} \|\widetilde{A}^{-1}\|_2^p h_k^{-p} \sum_{|\varepsilon| = |\beta| + |\gamma|} h^{\varepsilon p} \int_T |\partial_r^\varepsilon v|^p dx,$$

which leads to (18.6) together with (1.5).

Proof of (18.7)**.** Using Section 10.1.1 and (1.1), through a simple calculation, we have, for $1 \le k \le d$,

$$
\begin{aligned}
|\partial_{\hat{x}}^{\beta+\gamma} \hat{v}_k| &= \left| \frac{\partial^{|\beta|+|\gamma|}}{\partial \hat{x}_1^{\beta_1} \cdots \partial \hat{x}_d^{\beta_d} \partial \hat{x}_1^{\gamma_1} \cdots \partial \hat{x}_d^{\gamma_d}} \hat{v}_k \right| \\
&\le c |\det(A_{\widetilde{T}})| \|\widetilde{A}^{-1}\|_2 h_k^{-1} \\
&\quad \sum_{\eta=1}^{d} \underbrace{\sum_{i_1^{(1)}=1}^{d} \cdots \sum_{i_{\beta_1}^{(1)}=1}^{d}}_{\beta_1 \text{times}} \cdots \underbrace{\sum_{i_1^{(d)}=1}^{d} \cdots \sum_{i_{\beta_d}^{(d)}=1}^{d}}_{\beta_d \text{times}} \underbrace{\sum_{j_1^{(1)}=1}^{d} \cdots \sum_{j_{\gamma_1}^{(1)}=1}^{d}}_{\gamma_1 \text{times}} \cdots \underbrace{\sum_{j_1^{(d)}=1}^{d} \cdots \sum_{j_{\gamma_d}^{(d)}=1}^{d}}_{\gamma_d \text{times}} \\
&\quad \underbrace{\widetilde{\mathscr{H}}_{i_1^{(1)}} \cdots \widetilde{\mathscr{H}}_{i_{\varepsilon_1}^{(1)}}}_{\beta_1 \text{times}} \cdots \underbrace{\widetilde{\mathscr{H}}_{i_1^{(d)}} \cdots \widetilde{\mathscr{H}}_{i_{\varepsilon_d}^{(d)}}}_{\beta_d \text{times}} \underbrace{\widetilde{\mathscr{H}}_{j_1^{(1)}} \cdots \widetilde{\mathscr{H}}_{j_{\varepsilon_1}^{(1)}}}_{\gamma_1 \text{times}} \cdots \underbrace{\widetilde{\mathscr{H}}_{j_1^{(d)}} \cdots \widetilde{\mathscr{H}}_{j_{\varepsilon_d}^{(d)}}}_{\gamma_d \text{times}} \\
&\quad \left| \underbrace{\frac{\partial^{\beta_1}}{\partial \tilde{x}_{i_1^{(1)}} \cdots \partial \tilde{x}_{i_{\beta_1}^{(1)}}}}_{\beta_1 \text{times}} \cdots \underbrace{\frac{\partial^{\beta_d}}{\partial \tilde{x}_{i_1^{(d)}} \cdots \partial \tilde{x}_{i_{\beta_d}^{(d)}}}}_{\beta_d \text{times}} \underbrace{\frac{\partial^{\gamma_1}}{\partial \tilde{x}_{j_1^{(1)}} \cdots \partial \tilde{x}_{j_{\gamma_1}^{(1)}}}}_{\gamma_1 \text{times}} \cdots \underbrace{\frac{\partial^{\gamma_d}}{\partial \tilde{x}_{j_1^{(d)}} \cdots \partial \tilde{x}_{j_{\gamma_d}^{(d)}}}}_{\gamma_d \text{times}} \tilde{v}_\eta \right| \\
&\le c |\det(A_{\widetilde{T}})| \|\widetilde{A}^{-1}\|_2 h_k^{-1} \sum_{\eta=1}^{d} \sum_{|\varepsilon|=|\beta|+|\gamma|} \widetilde{\mathscr{H}}^{\varepsilon} |\partial_{\tilde{x}}^{\varepsilon} \tilde{v}_\eta|.
\end{aligned}
$$

Because $1 \le p < \infty$, it holds that, for $1 \le k \le d$,

$$
\left\| \partial_{\hat{x}}^{\beta} \partial_{\hat{x}}^{\gamma} \hat{v}_k \right\|_{L^p(\widehat{T})}^{p} \le c |\det(A_{\widetilde{T}})|^{p-1} \|\widetilde{A}^{-1}\|_2^p h_k^{-p} \sum_{|\varepsilon|=|\beta|+|\gamma|} \widetilde{\mathscr{H}}^{\varepsilon p} \int_{\widetilde{T}} |\partial_{\tilde{x}}^{\varepsilon} \tilde{v}|^p d\tilde{x},
$$

which leads to (18.7) together with (1.5). □

Remark 18.6. In inequality (18.7), it is possible to obtain the estimates in T by specifically determining the matrix A_T.

Let $\hat{v} \in \mathcal{C}^1(\widehat{T})^d$ with $\tilde{v} = \Psi_{\widehat{T}} \hat{v}$ and $v = \Psi_{\widetilde{T}} \tilde{v}$. Using (1.1), (6.8c), (6.9) and the definition of Piola transformations (Definition 18.1), we have, for $1 \le i, k \le d$,

$$
\left| \frac{\partial \hat{v}_k}{\partial \hat{x}_i} \right| \le c |\det(A_{\widetilde{T}})| \|\widetilde{A}^{-1}\|_2 h_k^{-1} \sum_{\nu=1}^{d} \sum_{i_1^{(1)}, i_1^{(0,1)}=1}^{d} \widetilde{\mathscr{H}}_{i_1^{(1)}} |[A_T]_{i_1^{(0,1)} i_1^{(1)}}| \left| \frac{\partial v_\nu}{\partial x_{i_1^{(0,1)}}} \right|.
$$

Let $d = 3$. We define the matrix A_T as

$$
\mathcal{A}_T := \begin{pmatrix} \cos \frac{\pi}{2} & -\sin \frac{\pi}{2} & 0 \\ \sin \frac{\pi}{2} & \cos \frac{\pi}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix}.
$$

We then have

$$
\left| \frac{\partial \hat{v}_k}{\partial \hat{x}_i} \right| \le c |\det(A_{\widetilde{T}})| \|\widetilde{A}^{-1}\|_2 h_k^{-1} \sum_{\nu=1}^{3} \left(\widetilde{\mathscr{H}_1} \left| \frac{\partial v_\nu}{\partial x_2} \right| + \widetilde{\mathscr{H}_2} \left| \frac{\partial v_\nu}{\partial x_1} \right| + \widetilde{\mathscr{H}_3} \left| \frac{\partial v_\nu}{\partial x_3} \right| \right).
$$

Because $1 \le p < \infty$, it holds that, for $1 \le i, k \le 3$,

$$
\begin{aligned}
\left\| \frac{\partial \hat{v}_k}{\partial \hat{x}_i} \right\|_{L^p(\widehat{T})}^{p} &\le c |\det(A_{\widetilde{T}})|^{p-1} \|\widetilde{A}^{-1}\|_2^p h_k^{-p} \\
&\quad \times \left(\widetilde{\mathscr{H}_1^p} \left\| \frac{\partial v}{\partial x_2} \right\|_{L^p(T)}^{p} + \widetilde{\mathscr{H}_2^p} \left\| \frac{\partial v}{\partial x_1} \right\|_{L^p(T)}^{p} + \widetilde{\mathscr{H}_3^p} \left\| \frac{\partial v}{\partial x_3} \right\|_{L^p(T)}^{p} \right).
\end{aligned}
$$

The following two lemmata are divided into the element on $\mathfrak{T}^{(2)}$ or $\mathfrak{T}_1^{(3)}$ and the element on $\mathfrak{T}_2^{(3)}$.

Lemma 18.7. Let $p \in [1,\infty)$. Let $T \in \mathbb{T}_h$ satisfy Condition 5.1 or Condition 5.2 with $T = \Phi_T(\widetilde{T})$ and $\widetilde{T} = \Phi_{\widetilde{T}}(\widehat{T})$, where $\widetilde{T} \in \mathfrak{T}^{(2)}$ or $\widetilde{T} \in \mathfrak{T}_1^{(3)}$. Let $\beta := (\beta_1,\dots,\beta_d) \in \mathbb{N}_0^d$ be a multi-index with $|\beta| = \ell$. Then, for any $\hat{v} \in W^{\ell+1,p}(\widehat{T})^d$ with $\tilde{v} = \Psi_{\widehat{T}}\hat{v}$ and $v = \Psi_{\widetilde{T}}\tilde{v}$,

$$\left\| \partial_{\hat{x}}^{\beta} \nabla_{\hat{x}} \cdot \hat{v} \right\|_{L^p(\widehat{T}_1)} \leq c |\det(A_{\widetilde{T}})|^{\frac{p-1}{p}} \sum_{|\varepsilon|=\ell} h^{\varepsilon} \left\| \partial_r^{\varepsilon} \nabla \cdot v \right\|_{L^p(T)}. \tag{18.8}$$

If Condition 10.1 is imposed, it holds that

$$\left\| \partial_{\hat{x}}^{\beta} \nabla_{\hat{x}} \cdot \hat{v} \right\|_{L^p(\widehat{T}_1)} \leq c |\det(A_{\widetilde{T}})|^{\frac{p-1}{p}} \sum_{|\varepsilon|=\ell} \widetilde{\mathscr{H}}^{\varepsilon} \left\| \partial_{\tilde{x}}^{\varepsilon} \nabla_{\tilde{x}} \cdot (\Psi_{\widetilde{T}}^{-1} v) \right\|_{L^p(\Phi_T^{-1}(T))}. \tag{18.9}$$

Proof. Because the space $\mathcal{C}^{\ell+1}(\widehat{T})^d$ is dense in the space $W^{\ell+1,p}(\widehat{T})^d$, we show (18.8) and (18.9) for $\hat{v} \in \mathcal{C}^{\ell+1}(\widehat{T})^d$ with $\tilde{v} = \Psi_{\widehat{T}}\hat{v}$ and $v = \Psi_{\widetilde{T}}\tilde{v}$.

By a simple calculation from Sections 18.3.1 and 18.3.2,

$$\begin{aligned}
\nabla_{\hat{x}} \cdot \hat{v} &= \sum_{k=1}^{d} \frac{\partial \hat{v}_k}{\partial \hat{x}_k} \\
&= \det(A_{\widetilde{T}}) \det(A_T) \sum_{k,\eta,\nu,i_1^{(1)},i_1^{(0,1)}=1}^{d} [\widetilde{A}^{-1}]_{k\eta} [\widetilde{A}]_{i_1^{(1)} k} [A_T^{-1}]_{\eta\nu} [A_T]_{i_1^{(0,1)} i_1^{(1)}} \frac{\partial v_\nu}{\partial x_{i_1^{(0,1)}}} \\
&= \det(A_{\widetilde{T}}) \det(A_T) \nabla \cdot v, \\
\frac{\partial}{\partial \hat{x}_i} \nabla_{\hat{x}} \cdot \hat{v} &= \sum_{k=1}^{d} \frac{\partial^2 \hat{v}_k}{\partial \hat{x}_i \partial \hat{x}_k} \\
&= \det(A_{\widetilde{T}}) \det(A_T) h_i \sum_{i_1^{(1)},i_1^{(0,1)}=1}^{d} [\widetilde{A}]_{i_1^{(1)} i} [A_T]_{i_1^{(0,1)} i_1^{(1)}} \\
&\qquad \sum_{k,\eta,\nu,j_1^{(1)},j_1^{(0,1)}=1}^{d} [\widetilde{A}^{-1}]_{k\eta} [\widetilde{A}]_{j_1^{(1)} k} [A_T^{-1}]_{\eta\nu} [A_T]_{j_1^{(0,1)} j_1^{(1)}} \frac{\partial^2 v_\nu}{\partial x_{i_1^{(0,1)}} \partial x_{j_1^{(0,1)}}} \\
&= \det(A_{\widetilde{T}}) \det(A_T) h_i \sum_{i_1^{(1)},i_1^{(0,1)}=1}^{d} [\widetilde{A}]_{i_1^{(1)} i} [A_T]_{i_1^{(0,1)} i_1^{(1)}} \frac{\partial (\nabla \cdot v)}{\partial x_{i_1^{(0,1)}}}.
\end{aligned}$$

For a general derivative $\partial_{\hat{x}}^{\beta} \nabla_{\hat{x}} \cdot \hat{v}$ with order $|\beta| = \ell$, we obtain

$$\begin{aligned}
&\partial_{\hat{x}}^{\beta} \nabla_{\hat{x}} \cdot \hat{v} = \frac{\partial^{|\beta|}}{\partial \hat{x}_1^{\beta_1} \cdots \partial \hat{x}_d^{\beta_d}} \nabla_{\hat{x}} \cdot \hat{v} \\
&= \det(A_{\widetilde{T}}) \det(A_T) \\
&\quad \underbrace{\sum_{i_1^{(1)},i_1^{(0,1)}=1}^{d} h_1 [\widetilde{A}]_{i_1^{(1)} 1} [A_T]_{i_1^{(0,1)} i_1^{(1)}} \cdots \sum_{i_{\beta_1}^{(1)},i_{\beta_1}^{(0,1)}=1}^{d} h_1 [\widetilde{A}]_{i_{\beta_1}^{(1)} 1} [A_T]_{i_{\beta_1}^{(0,1)} i_{\beta_1}^{(1)}} \cdots}_{\beta_1 \text{times}} \\
&\quad \underbrace{\sum_{i_1^{(d)},i_1^{(0,d)}=1}^{d} h_d [\widetilde{A}]_{i_1^{(d)} d} [A_T]_{i_1^{(0,d)} i_1^{(d)}} \cdots \sum_{i_{\beta_d}^{(d)},i_{\beta_d}^{(0,d)}=1}^{d} h_d [\widetilde{A}]_{i_{\beta_d}^{(d)} d} [A_T]_{i_{\beta_d}^{(0,d)} i_{\beta_d}^{(d)}}}_{\beta_d \text{times}}
\end{aligned}$$

$$\underbrace{\frac{\partial^{\beta_1}}{\partial x_{i_1^{(0,1)}} \cdots \partial x_{i_{\beta_1}^{(0,1)}}}}_{\beta_1 \text{times}} \cdots \underbrace{\frac{\partial^{\beta_d}}{\partial x_{i_1^{(0,d)}} \cdots \partial x_{i_{\beta_d}^{(0,d)}}}}_{\beta_d \text{times}} \nabla \cdot v$$

$$= \det(A_{\widetilde{T}}) \det(A_T)$$

$$\underbrace{\sum_{i_1^{(1)}, i_1^{(0,1)}=1}^{d} h_1 [A_T]_{i_1^{(0,1)} i_1^{(1)}} (\tilde{r}_1)_{i_1^{(1)}} \cdots \sum_{i_{\beta_1}^{(1)}, i_{\beta_1}^{(0,1)}=1}^{d} h_1 [A_T]_{i_{\beta_1}^{(0,1)} i_{\beta_1}^{(1)}} (\tilde{r}_1)_{i_{\beta_1}^{(1)}}}_{\beta_1 \text{times}} \cdots$$

$$\underbrace{\sum_{i_1^{(d)}, i_1^{(0,d)}=1}^{d} h_d [A_T]_{i_1^{(0,d)} i_1^{(d)}} (\tilde{r}_d)_{i_1^{(d)}} \cdots \sum_{i_{\beta_d}^{(d)}, i_{\beta_d}^{(0,d)}=1}^{d} h_d [A_T]_{i_{\beta_d}^{(0,d)} i_{\beta_d}^{(d)}} (\tilde{r}_d)_{i_{\beta_d}^{(d)}}}_{\beta_d \text{times}}$$

$$\underbrace{\frac{\partial^{\beta_1}}{\partial x_{i_1^{(0,1)}} \cdots \partial x_{i_{\beta_1}^{(0,1)}}}}_{\beta_1 \text{times}} \cdots \underbrace{\frac{\partial^{\beta_d}}{\partial x_{i_1^{(0,d)}} \cdots \partial x_{i_{\beta_d}^{(0,d)}}}}_{\beta_d \text{times}} \nabla \cdot v.$$

It then holds that, using (6.9) and (1.5),

$$|\partial_{\hat{x}}^{\beta} \nabla_{\hat{x}} \cdot \hat{v}| \leq c |\det(A_{\widetilde{T}})| \sum_{|\varepsilon|=\ell} h^{\varepsilon} |\partial_r^{\varepsilon} \nabla \cdot v|,$$

which leads to

$$\|\partial_{\hat{x}}^{\beta} \nabla_{\hat{x}} \cdot \hat{v}\|_{L^p(\widehat{T})} \leq c |\det(A_{\widetilde{T}})|^{\frac{p-1}{p}} \sum_{|\varepsilon|=\ell} h^{\varepsilon} \|\partial_r^{\varepsilon} \nabla \cdot v\|_{L^p(T)}.$$

Using an analogous argument, if Condition 10.1 is imposed, for a general derivative $\partial_{\hat{x}}^{\beta} \nabla_{\hat{x}} \cdot \hat{v}$ with order $|\beta| = \ell$, we obtain

$$\partial_{\hat{x}}^{\beta} \nabla_{\hat{x}} \cdot \hat{v} = \frac{\partial^{|\beta|}}{\partial \hat{x}_1^{\beta_1} \cdots \partial \hat{x}_d^{\beta_d}} \nabla_{\hat{x}} \cdot \hat{v}$$

$$= \det(A_{\widetilde{T}})$$

$$\underbrace{\sum_{i_1^{(1)}=1}^{d} h_1 [\widetilde{A}]_{i_1^{(1)} 1} \cdots \sum_{i_{\beta_1}^{(1)}=1}^{d} h_1 [\widetilde{A}]_{i_{\beta_1}^{(1)} 1}}_{\beta_1 \text{times}} \cdots \underbrace{\sum_{i_1^{(d)}=1}^{d} h_d [\widetilde{A}]_{i_1^{(d)} d} \cdots \sum_{i_{\beta_d}^{(d)}=1}^{d} h_d [\widetilde{A}]_{i_{\beta_d}^{(d)} d}}_{\beta_d \text{times}}$$

$$\underbrace{\frac{\partial^{\beta_1}}{\partial \tilde{x}_{i_1^{(1)}} \cdots \partial \tilde{x}_{i_{\beta_1}^{(1)}}}}_{\beta_1 \text{times}} \cdots \underbrace{\frac{\partial^{\beta_d}}{\partial \tilde{x}_{i_1^{(d)}} \cdots \partial \tilde{x}_{i_{\beta_d}^{(d)}}}}_{\beta_d \text{times}} \nabla_{\tilde{x}} \cdot \tilde{v}.$$

It then holds that

$$\left| \partial_{\hat{x}}^{\beta} \nabla_{\hat{x}} \cdot \hat{v} \right|$$

$$\leq c |\det(A_{\widetilde{T}})|$$

$$\underbrace{\sum_{i_1^{(1)}=1}^{d} h_1 |[\widetilde{A}]_{i_1^{(1)} 1}| \cdots \sum_{i_{\beta_1}^{(1)}=1}^{d} h_1 |[\widetilde{A}]_{i_{\beta_1}^{(1)} 1}|}_{\beta_1 \text{times}} \cdots \underbrace{\sum_{i_1^{(d)}=1}^{d} h_d |[\widetilde{A}]_{i_1^{(d)} d}| \cdots \sum_{i_{\beta_d}^{(d)}=1}^{d} h_d |[\widetilde{A}]_{i_{\beta_d}^{(d)} d}|}_{\beta_d \text{times}}$$

$$
\left| \underbrace{\frac{\partial^{\beta_1}}{\partial \tilde{x}_{i_1^{(1)}} \cdots \partial \tilde{x}_{i_{\beta_1}^{(1)}}}}_{\beta_1 \text{times}} \cdots \underbrace{\frac{\partial^{\beta_d}}{\partial \tilde{x}_{i_1^{(d)}} \cdots \partial \tilde{x}_{i_{\beta_d}^{(d)}}}}_{\beta_d \text{times}} \nabla_{\tilde{x}} \cdot \tilde{v} \right|
$$
$$
\leq c |\det(A_{\widetilde{T}})| \sum_{|\varepsilon|=\ell} \widetilde{\mathscr{H}^{\varepsilon}} |\partial_{\tilde{x}}^{\varepsilon} \nabla_{\tilde{x}} \cdot \tilde{v}|,
$$

which leads to

$$
\|\partial_{\hat{x}}^{\beta} \nabla_{\hat{x}} \cdot \hat{v}\|_{L^p(\widehat{T}_1)} \leq c |\det(A_{\widetilde{T}})|^{\frac{p-1}{p}} \sum_{|\varepsilon|=\ell} \widetilde{\mathscr{H}^{\varepsilon}} \|\partial_{\tilde{x}}^{\varepsilon} \nabla_{\tilde{x}} \cdot \tilde{v}\|_{L^p(\widetilde{T})}.
$$

$\square$

Lemma 18.8. Let $p \in [1,\infty)$ and $d = 3$. Let $T \in \mathbb{T}_h$ satisfy Condition 5.2 with $T = \Phi_T(\widetilde{T})$ and $\widetilde{T} = \Phi_{\widetilde{T}}(\widehat{T})$, where $\widehat{T} \in \mathfrak{T}_2^{(3)}$. Let $\ell \in \mathbb{N}_0$ and $k \in \mathbb{N}$ with $1 \leq k \leq 3$. Let $\beta := (\beta_1, \beta_2, \beta_3) \in \mathbb{N}_0^3$ be a multi-index with $|\beta| = \ell$. Then, for any $\hat{v} \in W^{\ell+1,p}(\widehat{T})^3$ with $\tilde{v} = \Psi_{\widetilde{T}} \hat{v}$ and $v = \Psi_{T} \tilde{v}$,

$$
\left\| \partial_{\hat{x}}^{\beta} \frac{\partial \hat{v}_k}{\partial \hat{x}_k} \right\|_{L^p(\widehat{T})} \leq c |\det(A_{\widetilde{T}})|^{\frac{p-1}{p}} \|\widetilde{A}^{-1}\|_2 \sum_{|\varepsilon|=\ell} h^{\varepsilon} \left\| \partial_r^{\varepsilon} \frac{\partial v}{\partial r_k} \right\|_{L^p(T)^3}. \tag{18.10}
$$

If Condition 10.1 is imposed, it holds that

$$
\left\| \partial_{\hat{x}}^{\beta} \frac{\partial \hat{v}_k}{\partial \hat{x}_k} \right\|_{L^p(\widehat{T})} \leq c |\det(A_{\widetilde{T}})|^{\frac{p-1}{p}} \|\widetilde{A}^{-1}\|_2 \sum_{|\varepsilon|=\ell} \widetilde{\mathscr{H}^{\varepsilon}} \left\| \partial_{\tilde{x}}^{\varepsilon} \frac{\partial (\Psi_{\widetilde{T}}^{-1} v)}{\partial \tilde{r}_k} \right\|_{L^p(\Phi_T^{-1}(T))^3}. \tag{18.11}
$$

Proof. Because the space $\mathcal{C}^{\ell+1}(\widehat{T})^3$ is dense in the space $W^{\ell+1,p}(\widehat{T})^3$, we show (18.10) and (18.11) for $\hat{v} \in \mathcal{C}^{\ell+1}(\widehat{T})^3$ with $\tilde{v} = \Psi_{\widetilde{T}} \hat{v}$ and $v = \Psi_{T} \tilde{v}$.

By a simple calculation from Section 18.3.1, for $1 \leq i, k \leq 3$,

$$
\begin{aligned}
\frac{\partial \hat{v}_k}{\partial \hat{x}_k} &= \det(A_{\widetilde{T}}) \det(A_T) \sum_{\eta,\nu,i_1^{(1)},i_1^{(0,1)}=1}^{3} [\widetilde{A}^{-1}]_{k\eta} [\widetilde{A}]_{i_1^{(1)} k} [A_T^{-1}]_{\eta\nu} [A_T]_{i_1^{(0,1)} i_1^{(1)}} \frac{\partial v_{\nu}}{\partial x_{i_1^{(0,1)}}} \\
&= \det(A_{\widetilde{T}}) \det(A_T) \sum_{\eta,\nu,i_1^{(1)},i_1^{(0,1)}=1}^{3} [\widetilde{A}^{-1}]_{k\eta} [A_T^{-1}]_{\eta\nu} [A_T]_{i_1^{(0,1)} i_1^{(1)}} (\tilde{r}_k)_{i_1^{(1)}} \frac{\partial v_{\nu}}{\partial x_{i_1^{(0,1)}}} \\
&= \det(A_{\widetilde{T}}) \det(A_T) \sum_{\eta,\nu=1}^{3} [\widetilde{A}^{-1}]_{k\eta} [A_T^{-1}]_{\eta\nu} \frac{\partial v_{\nu}}{\partial r_k}, \\
\frac{\partial^2 \hat{v}_k}{\partial \hat{x}_i \partial \hat{x}_k} &= \det(A_{\widetilde{T}}) \det(A_T) \sum_{\eta,\nu,i_1^{(1)},i_1^{(0,1)},j_1^{(1)},j_1^{(0,1)}=1}^{3} [\widetilde{A}^{-1}]_{k\eta} [A_T^{-1}]_{\eta\nu} \\
&\qquad h_i [\widetilde{A}]_{i_1^{(1)} i} [\widetilde{A}]_{j_1^{(1)} k} [A_T]_{i_1^{(0,1)} i_1^{(1)}} [A_T]_{j_1^{(0,1)} j_1^{(1)}} \frac{\partial^2 v_{\nu}}{\partial x_{i_1^{(0,1)}} \partial x_{j_1^{(0,1)}}} \\
&= \det(A_{\widetilde{T}}) \det(A_T) \sum_{\eta,\nu,i_1^{(1)},i_1^{(0,1)},j_1^{(1)},j_1^{(0,1)}=1}^{3} [\widetilde{A}^{-1}]_{k\eta} [A_T^{-1}]_{\eta\nu} \\
&\qquad h_i [A_T]_{i_1^{(0,1)} i_1^{(1)}} (\tilde{r}_i)_{i_1^{(1)}} [A_T]_{j_1^{(0,1)} j_1^{(1)}} (\tilde{r}_k)_{j_1^{(1)}} \frac{\partial^2 v_{\nu}}{\partial x_{i_1^{(0,1)}} \partial x_{j_1^{(0,1)}}} \\
&= \det(A_{\widetilde{T}}) \det(A_T) \sum_{\eta,\nu=1}^{3} [\widetilde{A}^{-1}]_{k\eta} [A_T^{-1}]_{\eta\nu} h_i \frac{\partial^2 v_{\nu}}{\partial r_i \partial r_k}.
\end{aligned}
$$

For a general derivative $\partial_{\hat{x}}^{\beta} \frac{\partial \hat{v}_k}{\partial \hat{x}_k}$ $(1 \leq k \leq 3)$ with order $|\beta| = \ell$, we obtain

$$
\begin{aligned}
\partial_{\hat{x}}^{\beta} \frac{\partial \hat{v}_k}{\partial \hat{x}_k} &= \frac{\partial^{|\beta|}}{\partial \hat{x}_1^{\beta_1} \partial \hat{x}_2^{\beta_2} \partial \hat{x}_3^{\beta_3}} \frac{\partial \hat{v}_k}{\partial \hat{x}_k} \\
&= \det(A_{\widetilde{T}}) \det(A_T) \sum_{\eta,\nu=1}^{3} [\widetilde{A}^{-1}]_{k\eta} [A_T^{-1}]_{\eta\nu} \\
&\quad \underbrace{\sum_{i_1^{(1)}, i_1^{(0,1)}=1}^{3} h_1 [A_T]_{i_1^{(0,1)} i_1^{(1)}} (\tilde{r}_1)_{i_1^{(1)}} \cdots \sum_{i_{\beta_1}^{(1)}, i_{\beta_1}^{(0,1)}=1}^{3} h_1 [A_T]_{i_{\beta_1}^{(0,1)} i_{\beta_1}^{(1)}} (\tilde{r}_1)_{i_{\beta_1}^{(1)}}}_{\beta_1 \text{times}} \cdots \\
&\quad \underbrace{\sum_{i_1^{(3)}, i_1^{(0,3)}=1}^{3} h_3 [A_T]_{i_1^{(0,3)} i_1^{(3)}} (\tilde{r}_3)_{i_1^{(3)}} \cdots \sum_{i_{\beta_3}^{(3)}, i_{\beta_3}^{(0,3)}=1}^{3} h_3 [A_T]_{i_{\beta_3}^{(0,3)} i_{\beta_3}^{(3)}} (\tilde{r}_3)_{i_{\beta_3}^{(3)}}}_{\beta_d \text{times}} \\
&\quad \underbrace{\frac{\partial^{\beta_1}}{\partial x_{i_1^{(0,1)}} \cdots \partial x_{i_{\beta_1}^{(0,1)}}}}_{\beta_1 \text{times}} \cdots \underbrace{\frac{\partial^{\beta_3}}{\partial x_{i_1^{(0,3)}} \cdots \partial x_{i_{\beta_d}^{(0,d)}}}}_{\beta_d \text{times}} \frac{\partial v_\nu}{\partial r_k} \\
&= \det(A_{\widetilde{T}}) \det(A_T) \sum_{\eta,\nu=1}^{3} [\widetilde{A}^{-1}]_{k\eta} [A_T^{-1}]_{\eta\nu} h^{\beta} \underbrace{\frac{\partial^{\beta_1}}{\partial r_1 \cdots \partial r_1}}_{\beta_1 \text{times}} \cdots \underbrace{\frac{\partial^{\beta_3}}{\partial r_3 \cdots \partial r_3}}_{\beta_3 \text{times}} \frac{\partial v_\nu}{\partial r_k}.
\end{aligned}
$$

It then holds that, using (1.1), (6.8c) and (6.9),

$$
\left| \partial_{\hat{x}}^{\beta} \frac{\partial \hat{v}_k}{\partial \hat{x}_k} \right| \leq c |\det(A_{\widetilde{T}})| \|\widetilde{A}^{-1}\|_2 \sum_{\nu=1}^{3} \sum_{|\varepsilon|=|\beta|} h^{\varepsilon} \left| \partial_r^{\varepsilon} \frac{\partial v_\nu}{\partial r_k} \right|,
$$

which leads to, using (1.5),

$$
\left\| \partial_{\hat{x}}^{\beta} \frac{\partial \hat{v}_k}{\partial \hat{x}_k} \right\|_{L^p(\widehat{T}_2)} \leq c |\det(A_{\widetilde{T}})|^{\frac{p-1}{p}} \|\widetilde{A}^{-1}\|_2 \sum_{|\varepsilon|=|\beta|} h^{\varepsilon} \left\| \partial_r^{\varepsilon} \frac{\partial v}{\partial r_k} \right\|_{L^p(T)^3}.
$$

If Condition 10.1 is imposed, by a simple calculation from Section 18.3.2, for $1 \leq i, k \leq 3$,

$$
\begin{aligned}
\frac{\partial \hat{v}_k}{\partial \hat{x}_k} &= \det(A_{\widetilde{T}}) \sum_{\eta=1}^{3} [\widetilde{A}^{-1}]_{k\eta} \sum_{i_1^{(1)}=1}^{3} [\widetilde{A}]_{i_1^{(1)} k} \frac{\partial \tilde{v}_\eta}{\partial \tilde{x}_{i_1^{(1)}}} \\
&= \det(A_{\widetilde{T}}) \sum_{\eta=1}^{3} [\widetilde{A}^{-1}]_{k\eta} \sum_{i_1^{(1)}=1}^{3} (\tilde{r}_k)_{i_1^{(1)}} \frac{\partial \tilde{v}_\eta}{\partial \tilde{x}_{i_1^{(1)}}} \\
&= \det(A_{\widetilde{T}}) \sum_{\eta=1}^{3} [\widetilde{A}^{-1}]_{k\eta} \frac{\partial \tilde{v}_\eta}{\partial \tilde{r}_k}, \\
\frac{\partial^2 \hat{v}_k}{\partial \hat{x}_i \partial \hat{x}_k} &= \det(A_{\widetilde{T}}) \sum_{\eta, i_1^{(1)}=1}^{3} [\widetilde{A}^{-1}]_{k\eta} h_i [\widetilde{A}]_{i_1^{(1)} i} \sum_{j_1^{(1)}=1}^{3} [\widetilde{A}]_{j_1^{(1)} k} \frac{\partial^2 \tilde{v}_\eta}{\partial \tilde{x}_{i_1^{(1)}} \partial \tilde{x}_{j_1^{(1)}}} \\
&= \det(A_{\widetilde{T}}) \sum_{\eta, i_1^{(1)}=1}^{3} [\widetilde{A}^{-1}]_{k\eta} h_i [\widetilde{A}]_{i_1^{(1)} i} \sum_{j_1^{(1)}=1}^{3} (\tilde{r}_k)_{j_1^{(1)}} \frac{\partial^2 \tilde{v}_\eta}{\partial \tilde{x}_{i_1^{(1)}} \partial \tilde{x}_{j_1^{(1)}}}
\end{aligned}
$$

$$= \det(A_{\widetilde{T}}) \sum_{\eta, i_1^{(1)}=1}^{3} [\widetilde{A}^{-1}]_{k\eta} h_i [\widetilde{A}]_{i_1^{(1)} i} \frac{\partial^2 \tilde{v}_\eta}{\partial \tilde{x}_{i_1^{(1)}} \partial \tilde{r}_k^s}.$$

For a general derivative $\partial_{\hat{x}}^{\beta} \frac{\partial \hat{v}_k}{\partial \hat{x}_k}$ ($1 \leq k \leq 3$) with order $|\beta| = \ell$, we obtain

$$\partial_{\hat{x}}^{\beta} \frac{\partial \hat{v}_k}{\partial \hat{x}_k} = \frac{\partial^{|\beta|}}{\partial \hat{x}_1^{\beta_1} \partial \hat{x}_2^{\beta_2} \partial \hat{x}_3^{\beta_3}} \frac{\partial \hat{v}_k}{\partial \hat{x}_k}$$

$$= \det(A_{\widetilde{T}}) \sum_{\eta=1}^{3} [\widetilde{A}^{-1}]_{k\eta}$$

$$\underbrace{\sum_{i_1^{(1)}=1}^{3} h_1 [\widetilde{A}]_{i_1^{(1)} 1} \cdots \sum_{i_{\beta_1}^{(1)}=1}^{3} h_1 [\widetilde{A}]_{i_{\beta_1}^{(1)} 1}}_{\beta_1 \text{times}} \cdots \underbrace{\sum_{i_1^{(3)}=1}^{3} h_3 [\widetilde{A}]_{i_1^{(3)} 3} \cdots \sum_{i_{\beta_3}^{(3)}=1}^{3} h_3 [\widetilde{A}]_{i_{\beta_3}^{(3)} 3}}_{\beta_3 \text{times}}$$

$$\underbrace{\frac{\partial^{\beta_1}}{\partial \tilde{x}_{i_1^{(1)}} \cdots \partial \tilde{x}_{i_{\beta_1}^{(1)}}}}_{\beta_1 \text{times}} \cdots \underbrace{\frac{\partial^{\beta_3}}{\partial \tilde{x}_{i_1^{(3)}} \cdots \partial \tilde{x}_{i_{\beta_3}^{(3)}}}}_{\beta_3 \text{times}} \frac{\partial \tilde{v}_\eta}{\partial \tilde{r}_k}.$$

It then holds that, using (1.1),

$$\left| \partial_{\hat{x}}^{\beta} \frac{\partial \hat{v}_k}{\partial \hat{x}_k} \right|$$

$$\leq |\det(A_{\widetilde{T}})| \|\widetilde{A}^{-1}\|_2 \sum_{\eta=1}^{3}$$

$$\underbrace{\sum_{i_1^{(1)}=1}^{3} h_1 |[\widetilde{A}]_{i_1^{(1)} 1}| \cdots \sum_{i_{\beta_1}^{(1)}=1}^{3} h_1 |[\widetilde{A}]_{i_{\beta_1}^{(1)} 1}|}_{\beta_1 \text{times}} \cdots \underbrace{\sum_{i_1^{(3)}=1}^{3} h_3 |[\widetilde{A}]_{i_1^{(3)} 3}| \cdots \sum_{i_{\beta_3}^{(3)}=1}^{3} h_3 |[\widetilde{A}]_{i_{\beta_d}^{(3)} 3}|}_{\beta_3 \text{times}}$$

$$\left| \underbrace{\frac{\partial^{\beta_1}}{\partial \tilde{x}_{i_1^{(1)}} \cdots \partial \tilde{x}_{i_{\beta_1}^{(1)}}}}_{\beta_1 \text{times}} \cdots \underbrace{\frac{\partial^{\beta_3}}{\partial \tilde{x}_{i_1^{(3)}} \cdots \partial \tilde{x}_{i_{\beta_3}^{(3)}}}}_{\beta_3 \text{times}} \frac{\partial \tilde{v}_\eta}{\partial \tilde{r}_k} \right|$$

$$\leq c |\det(A_{\widetilde{T}})| \|\widetilde{A}^{-1}\|_2 \sum_{\eta=1}^{3} \sum_{|\varepsilon|=|\beta|} \widetilde{\mathscr{H}}^{\varepsilon} \left| \partial_{\tilde{x}}^{\varepsilon} \frac{\partial \tilde{v}_\eta}{\partial \tilde{r}_k} \right|,$$

which leads to, using (1.5),

$$\left\| \partial_{\hat{x}}^{\beta} \frac{\partial \hat{v}_k}{\partial \hat{x}_k} \right\|_{L^p(\widehat{T})} \leq c |\det(A_{\widetilde{T}})|^{\frac{p-1}{p}} \|\widetilde{A}^{-1}\|_2 \sum_{|\varepsilon|=|\beta|} \widetilde{\mathscr{H}}^{\varepsilon} \left\| \partial_{\tilde{x}}^{\varepsilon} \frac{\partial \tilde{v}}{\partial \tilde{r}_k} \right\|_{L^p(\widetilde{T})^3}.$$

□

19 New RT Interpolation Error Estimates

19.1 RT Finite Element

Let $T \subset \mathbb{R}^d$ be a simplex. We define a space as

$$\mathbb{R}^k(\partial T) := \{\varphi_h \in L^2(\partial T) : \ \varphi_h|_F \in \mathbb{P}^k(F) \ \forall F \in \mathcal{F}_T\}.$$

Let $\widehat{T} \subset \mathbb{R}^d$ be the reference element defined in Sections 5.1 and 5.1. Let $\widehat{F}_i$ be the face of $\widehat{T}$ opposite to $\widehat{p}_i$. The RT finite element on the reference element is defined by the triple $\{\widehat{T}, \widehat{P}, \widehat{\Sigma}\}$ as follows:

1. $\widehat{P} := \mathbb{RT}^k(\widehat{T})$;

2. $\widehat{\Sigma}$ is a set $\{\hat{\chi}_i\}_{1 \le i \le N^{(RT)}}$ of $N^{(RT)}$ linear forms with its components such that, for any $\hat{r} \in \widehat{P}$,

$$\int_{\widehat{F}} \hat{r} \cdot \hat{n}_{\widehat{F}} \hat{q}_k d\hat{s}, \quad \forall \hat{q}_k \in \mathbb{R}^k(\partial \widehat{T}), \tag{19.1}$$

$$\int_{\widehat{T}} \hat{r} \cdot \hat{q}_{k-1} d\hat{x}, \quad \forall \hat{q}_{k-1} \in \mathbb{P}^{k-1}(\widehat{T})^d, \tag{19.2}$$

   shere $\hat{n}_{\widehat{F}}$ denotes the outer unit normal vector of $\widehat{T}$ on the face $\widehat{F}$. Note that for $k = 0$, the local degrees of freedom of type (19.2) are violated.

For the simplicial RT element in $\mathbb{R}^d$, it holds that

$$\dim \mathbb{RT}^k(\widehat{T}) = \begin{cases} (k+1)(k+3) & \text{if } d = 2, \\ \frac{1}{2}(k+1)(k+2)(k+4) & \text{if } d = 3. \end{cases} \tag{19.3}$$

The RT finite element with the local degrees of freedom with respect to (19.1) and (19.2) is unisolvent; for example, see [12, Proposition 2.3.4].

We set the domain of the local RT interpolation to $V(\widehat{T}) := W^{1,1}(\widehat{T})^d$; for example, see Theorem 1.9. The local RT interpolation $I_{\widehat{T}}^{RT^k} : V(\widehat{T}) \to \widehat{P}$ is then defined as follows: For any $\hat{v} \in V(\widehat{T})$,

$$\int_{\widehat{F}} (I_{\widehat{T}}^{RT^k} \hat{v} - \hat{v}) \cdot \hat{n}_{\widehat{F}} \hat{q}_k d\hat{s} = 0 \quad \forall \hat{q}_k \in \mathbb{R}^k(\partial \widehat{T}), \tag{19.4}$$

and if $k \ge 1$,

$$\int_{\widehat{T}} (I_{\widehat{T}}^{RT^k} \hat{v} - \hat{v}) \cdot \hat{q}_{k-1} d\hat{x} = 0 \quad \forall \hat{q}_{k-1} \in \mathbb{P}^{k-1}(\widehat{T})^d. \tag{19.5}$$

In particular, when $k = 0$, the degrees of freedom by (19.1) are describe as

$$\hat{\chi}_i(\hat{r}) := \int_{\widehat{F}_i} \hat{r} \cdot \hat{n}_{\widehat{F}_i} d\hat{s} \quad \forall \hat{r} \in \mathbb{RT}^0(\widehat{T}), \quad \forall i \in \{1, \ldots, d+1\}. \tag{19.6}$$

The local shape functions are as follows.

$$\hat{\theta}_i^{RT^0}(x) := \frac{\iota_{\widehat{F}_i, \widehat{T}}}{d |\widehat{T}|_d} (\hat{x} - \hat{p}_i) \quad \forall i \in \{1, \ldots, d+1\},$$

where $\iota_{\widehat{F}_i, \widehat{T}} := 1$ if $\hat{n}_{\widehat{F}_i}$ points outwards, and -1 otherwise [21, Chapter 14]. Indeed, $\hat{\theta}_i^{RT^0} \in \mathbb{RT}^0(\widehat{T})$ and $\hat{\chi}_i(\hat{\theta}_j^{RT^0}) = \delta_{ij}$ for any $i, j \in \{1, \ldots, d+1\}$. The local RT interpolation $I_{\widehat{T}}^{RT^0} : V(\widehat{T}) \to \mathbb{RT}^0(\widehat{T})$ is then described as

$$I_{\widehat{T}}^{RT^0} : V(\widehat{T}) \ni \hat{v} \mapsto I_{\widehat{T}}^{RT^0} \hat{v} := \sum_{i=1}^{d+1} \left(\int_{\widehat{F}_i} \hat{v} \cdot \hat{n}_{\widehat{F}_i} d\hat{s} \right) \hat{\theta}_i^{RT^0} \in \mathbb{RT}^0(\widehat{T}). \tag{19.7}$$

Let $\Phi_{\widetilde{T}} : \widehat{T} \to \widetilde{T}$ and $\Phi_T : \widetilde{T} \to T$ be the affine mappings defined in Section 5.2. Let $\Psi_{\widehat{T}} : V(\widehat{T}) \to V(\widetilde{T})$ and $\Psi_{\widetilde{T}} : V(\widetilde{T}) \to V(T)$ be the Piola transformations defined in Definition 18.1. The triples $\{\widetilde{T}, \widetilde{P}, \widetilde{\Sigma}\}$ and $\{T, P, \Sigma\}$ are defined as

$$\begin{cases} \widetilde{T} = \Phi_{\widetilde{T}}(\widehat{T}); \\ \widetilde{P} = \{\Psi_{\widehat{T}}(\hat{q});\ \hat{q} \in \widehat{P}\}; \\ \widetilde{\Sigma} = \{\{\tilde{\chi}_i\}_{1 \le i \le N^{(RT)}};\ \tilde{\chi}_i = \hat{\chi}_i(\widehat{\Psi}^{-1}(\tilde{q})), \forall \tilde{q} \in \widetilde{P}, \hat{\chi}_i \in \widehat{\Sigma}\}; \end{cases}$$

and

$$\begin{cases} T = \Phi_T(\widetilde{T}); \\ P = \{\Psi_{\widetilde{T}}(\tilde{q});\ \tilde{q} \in \widetilde{P}\}; \\ \Sigma = \{\{\chi_i\}_{1 \le i \le N^{(RT)}};\ \chi_i = \tilde{\chi}_i(\Psi_{\widetilde{T}}^{-1}(q)), \forall q \in P, \tilde{\chi}_i \in \widetilde{\Sigma}\}. \end{cases}$$

The triples $\{\widetilde{T}, \widetilde{P}, \widetilde{\Sigma}\}$ and $\{T, P, \Sigma\}$ are then the RT finite elements. Furthermore, let

$$I_{\widetilde{T}}^{RT^k} : V(\widetilde{T}) \to \mathbb{RT}^k(\widetilde{T}) \tag{19.8}$$

and

$$I_T^{RT^k} : V(T) \to \mathbb{RT}^k(T) \tag{19.9}$$

be the associated local RT interpolation defined in (19.4) and (19.5), respectively.

Remark 19.1. Let $T \subset \mathbb{R}^d$ be a simplex. Let $v \in H(\mathrm{div}; T)$ and $q \in \mathbb{R}^k(\partial T)$. Let $\gamma^d : H(\mathrm{div}; T) \to H^{-\frac{1}{2}}(\partial T)$ be the trace operator. Then, $\gamma^d(v) \cdot n_T \in H^{-\frac{1}{2}}(\partial T)$, see Theorem 1.2. We consider

$$\int_{\partial T} (\gamma^d(v) \cdot n_T) q ds.$$

Then, we cannot take the integral over an edge F of ∂T. Because functions $q \in \mathbb{R}^k(\partial T)$ do not belong in $H^{\frac{1}{2}}(\partial T)$.

As an example, we introduce the following remark ([12, Remark 2.5.1]). Given a function $\chi \in H^{-\frac{1}{2}}(\partial T)$, even if we are allowed to take

$$\int_{\partial T} \chi ds := \langle \chi, \psi \rangle \quad \text{with } \psi \equiv 1,$$

we cannot take the integral over an edge F of ∂T. Because the function identically equal to 1 on the whole boundary ∂T belongs to $H^{\frac{1}{2}}(\partial T)$, while the function that is equal to 1 on the edge F and 0 on the rest of ∂T does not belong to $H^{\frac{1}{2}}(\partial T)$.

Proposition 19.2. For any $\hat{v} \in H^1(\widehat{T})^d$ with $v := \Psi(\hat{v})$, it holds that

$$\Psi(I_{\widehat{T}}^{RT^k} \hat{v}) = I_T^{RT^k}(\Psi \hat{v}),$$

that is, the diagrams

$$\begin{tikzcd} V(T) \arrow[r, "\Psi_{\widetilde{T}}^{-1}"] \arrow[d, "I_T^{RT^k}"'] & V(\widetilde{T}) \arrow[r, "\Psi_{\widehat{T}}^{-1}"] \arrow[d, "I_{\widetilde{T}}^{RT^k}"'] & V(\widehat{T}) \arrow[d, "I_{\widehat{T}}^{RT^k}"] \\ P \arrow[r, "\Psi_{\widetilde{T}}^{-1}"'] & \widetilde{P} \arrow[r, "\Psi_{\widehat{T}}^{-1}"'] & \widehat{P} \end{tikzcd}$$

commute.

Proof. A proof can be found in [11, Lemma 3.4]. $\square$

Lemma 19.3. Let $T \in \mathbb{T}_h$. Let $V^{RT}(T) := W^{1,1}(T)^d$ and $V^{L^2}(T) := L^1(T)$. For $k \in \mathbb{N}_0$, let $I_T^{RT^k} : V^{RT}(T) \to \mathbb{RT}^k(T)$ and $\Pi_T^k : V^{L^2}(T) \to \mathbb{P}^k(T)$ be the RT interpolation operator and the L^2-orthogonal projection, respectively. Then, the following diagram commutes:

$$\begin{array}{ccc} V^{RT}(T) & \xrightarrow{\nabla\cdot} & V^{L^2}(T) \\ {\scriptstyle I_T^{RT^k}}\downarrow & & \downarrow{\scriptstyle \Pi_T^k} \\ \mathbb{RT}^k(T) & \xrightarrow[\nabla\cdot]{} & \mathbb{P}^k(T) \end{array}$$

In other words,

$$\nabla \cdot (I_T^{RT^k} v) = \Pi_T^k(\nabla \cdot v) \quad \forall v \in V^{RT}(T). \tag{19.10}$$

Proof. A proof can be found in [21, Lemma 16.2]. $\square$

Lemma 19.4. Let $T \in \mathbb{T}_h$ and $q \in \mathbb{RT}^k(T)$. Then,

$$\operatorname{div} q \in \mathbb{P}^k(T), \tag{19.11}$$

$$q \cdot n|_{\partial T} \in \mathbb{R}^k(\partial T). \tag{19.12}$$

Proof. A proof can be found in [12, Proposition 2.3.3]. $\square$

Lemma 19.5. The RT finite element with the nodal values in (19.1) and (19.2) is unisolvent.

Proof. A proof can be found in [12, Proposition 2.3.4]. $\square$

19.2 Remarks on the Anisotropic RT Interpolation Error Estimate

We consider the simplex $\widehat{T} \subset \mathbb{R}^2$ with vertices $\widehat{p}_1 := (0,0)^\top$, $\widehat{p}_2 := (1,0)^\top$ and $\widehat{p}_3 := (0,1)^\top$. For $1 \le i \le 3$, let $\widehat{F}_i$ be the face of $\widehat{T}$ opposite to $\widehat{p}_i$. The RT interpolation of $\hat{v}$ is defined as

$$I_{\widehat{T}}^{RT^0} \hat{v} = \sum_{i=1}^{3} \left(\int_{\widehat{F}_i} \hat{v} \cdot \hat{n}_i d\hat{s} \right) \hat{\theta}_i \in \mathbb{RT}^0,$$

where

$$\hat{\theta}_i := \frac{1}{2|\widehat{T}|}(\hat{x} - \widehat{p}_i), \quad \hat{x} = (\hat{x}_1, \hat{x}_2)^\top.$$

Setting $\hat{v} := (0, \hat{x}_2^2)^\top$ yields

$$\begin{aligned} I_{\widehat{T}}^{RT^0} \hat{v} &= \frac{1}{\sqrt{2}} \left(\int_{\widehat{F}_1} \hat{x}_2^2 d\hat{s} \right) (\hat{x}_1, \hat{x}_2)^\top - \left(\int_{\widehat{F}_3} \hat{x}_2^2 d\hat{s} \right) (\hat{x}_1, \hat{x}_2 - 1)^\top \\ &= \frac{1}{3}(\hat{x}_1, \hat{x}_2)^\top. \end{aligned}$$

This implies that $(I_{\widehat{T}}^{RT^0} \hat{v})_1 - \hat{v}_1 \not\equiv 0$ for any $\hat{x} \in \mathbb{R}^2$ and the following component-wise stability does not hold:

$$\|(I_{\widehat{T}}^{RT^0} \hat{v})_1\|_{L^2(\widehat{T})} \le c|\hat{v}_1|_{H^1(\widehat{T})}.$$

In other words, $(I_{\widehat{T}}^{RT^0} \hat{v})_1$ depends on both $\hat{v}_1$ and $\hat{v}_2$. Meanwhile, setting $\hat{v} := (0, \hat{x}_1^2)^\top$ yields $I_{\widehat{T}}^{RT^0} \hat{v} = \frac{1}{3}(0,1)^\top$. A key observation is that if $\hat{r} := (0, g(\hat{x}_1))^\top$, then $(I_{\widehat{T}}^{RT^0} \hat{r})_1 = 0$. In the next section, we introduce component-wise stabilities of the RT interpolation on the reference element by [1].

19.3 Component-wise Stability of the RT interpolation on the Reference Element

We will use symbols only used in this subsection.

19.3.1 Two-dimensional case

Let $\widehat{T} \subset \mathbb{R}^2$ be the reference triangle with vertices $\widehat{A}_1 := (1,0)^\top$, $\widehat{A}_2 := (0,1)^\top$, and $\widehat{A}_3 := (0,0)^\top$ with $\hat{N}_1 := (-1,0)^\top$, $\hat{N}_2 := (0,-1)^\top$, and $\hat{N}_3 := \frac{1}{\sqrt{2}}(1,1)^\top$. For $1 \le i \le 3$, let $\widehat{E}_i$ be the edge of $\widehat{T}$ opposite to $\widehat{A}_i$.

We use the same notation for a function of some variable as for its extension to $\widehat{T}$ as a function independent of the other variable. For example, $f(\hat{x}_2)$ denotes a function defined on $\widehat{E}_1$ as well as one is defined in $\widehat{T}$. Furthermore, the same notation is used to denote a polynomial $\hat{p}_k$ on an edge and a polynomial in two variables such that its restriction to that edge agrees with $\hat{q}_k$. For example, for $\hat{q}_k \in \mathbb{P}^k(\widehat{E}_3)$, we write $\hat{q}_k(1-\hat{x}_2, \hat{x}_2)$.

Lemma 19.6. Let $\hat{f}_i \in L^p(\widehat{E}_i)$, $i \in \{1,2\}$. If

$$\hat{u}(\hat{x}) = (\hat{f}_1(\hat{x}_2), 0)^\top, \quad \hat{v}(\hat{x}) = (0, \hat{f}_2(\hat{x}_1))^\top,$$

then there exist polynomials $\hat{q}_i \in \mathcal{P}^k(\widehat{E}_i)$, $i \in \{1,2\}$, such that

$$I_{\widehat{T}}^{RT^k} \hat{u} = (\hat{q}_1(\hat{x}_2), 0)^\top, \quad I_{\widehat{T}}^{RT^k} \hat{v} = (0, \hat{q}_2(\hat{x}_1))^\top.$$

Proof. A proof is provided in [1, Lemma 3.2] (also see Lemma 19.8) for the case $d = 3$. The estimate in the case $d = 2$ can be proved analogously. $\square$

Lemma 19.7. For $k \in \mathbb{N}_0$, there exists a constant c such that, for all $\hat{u} = (\hat{u}_1, \hat{u}_2)^\top \in W^{1,p}(\widehat{T})^2$,

$$\|(I_{\widehat{T}}^{RT^k} \hat{u})_i\|_{L^p(\widehat{T})} \le c \left(\|\hat{u}_i\|_{W^{1,p}(\widehat{T})} + \|\widehat{\operatorname{div}} \hat{u}\|_{L^p(\widehat{T})} \right), \quad i = 1, 2. \tag{19.13}$$

Proof. The proof is provided in [1, Lemma 3.3] (also see Lemma 19.9) for the case $d = 3$. The estimate in the case $d = 2$ can be proved analogously. $\square$

19.3.2 Three-dimensional case: Type i

Let $\widehat{T} \subset \mathbb{R}^3$ be the reference triangle with vertices $\widehat{A}_1 := (1,0,0)^\top$, $\widehat{A}_2 := (0,1,0)^\top$, $\widehat{A}_3 := (0,0,1)^\top$, and $\widehat{A}_4 := (0,0,0)^\top$ with $\widehat{N}_1 := (-1,0,0)^\top$, $\widehat{N}_2 := (0,-1,0)^\top$, $\widehat{N}_3 := (0,0,-1)^\top$, and $\widehat{N}_4 := \frac{1}{\sqrt{3}}(1,1,1)^\top$. For $1 \le i \le 4$, let $\widehat{E}_i$ be the edge of $\widehat{T}$ opposite to $\widehat{A}_i$.

In the two-dimensional case, we use the same notation for a function of some variable as for its extension to $\widehat{T}$ as a function independent of the other variable. For example, $f(\hat{x}_2, \hat{x}_3)$ denotes a function define on $\widehat{E}_1$ as well as one is defined in $\widehat{T}$. Furthermore, the same notation is used to denote a polynomial $\hat{p}_k$ on an edge and a polynomial in two variables such that its restriction to that edge agrees with $\hat{p}_k$. For example, for $\hat{p}_k \in \mathbb{P}^k(\widehat{E}_4)$, we write $\hat{p}_k(1-\hat{x}_2-\hat{x}_3, \hat{x}_2, \hat{x}_3)$.

Lemma 19.8. Let $k \in \mathbb{N}_0$. Let $\hat{f}_i \in L^p(\widehat{E}_i)$, $i \in \{1,2,3\}$. If

$$\begin{aligned} &\hat{u}(\hat{x}) = (\hat{f}_1(\hat{x}_2, \hat{x}_3), 0, 0)^\top, \quad \hat{v}(\hat{x}) = (0, \hat{f}_2(\hat{x}_1, \hat{x}_3), 0)^\top, \\ &\hat{w}(\hat{x}) = (0, 0, \hat{f}_3(\hat{x}_1, \hat{x}_2))^\top, \end{aligned}$$

then there exist polynomials $\hat{q}_i \in \mathbb{P}^k(\widehat{E}_i)$, $i \in \{1,2,3\}$, such that

$$\begin{aligned} &I_{\widehat{T}}^{RT^k} \hat{u} = (\hat{q}_1(\hat{x}_2, \hat{x}_3), 0, 0)^\top, \quad I_{\widehat{T}}^{RT^k} \hat{v} = (0, \hat{q}_2(\hat{x}_1, \hat{x}_3), 0)^\top, \\ &I_{\widehat{T}}^{RT^k} \hat{w} = (0, 0, \hat{q}_3(\hat{x}_1, \hat{x}_2))^\top. \end{aligned}$$

Proof. We follow [1, Lemma 3.2]. Because $\widehat{\mathrm{div}}\hat{u} = 0$, from the definition of the RT interpolation and the Green's formula, we have, for any $\hat{r}_k \in \mathbb{P}^k(\widehat{T})$,

$$
\begin{aligned}
0 &= \int_{\widehat{T}} \hat{r}_k \widehat{\mathrm{div}}\hat{u} d\hat{x} \\
&= \sum_{i=1}^{4} \int_{\widehat{E}_i} (\hat{r}_k \widehat{N}_i) \cdot \hat{u} d\hat{s} - \int_{\widehat{T}} (\hat{u} \cdot \widehat{\nabla}) \hat{r}_k d\hat{x} \\
&= \sum_{i=1}^{4} \int_{\widehat{E}_i} (\hat{r}_k \widehat{N}_i) \cdot (I_{\widehat{T}}^{RT^k} \hat{u}) d\hat{s} - \int_{\widehat{T}} ((I_{\widehat{T}}^{RT^k} \hat{u}) \cdot \widehat{\nabla}) \hat{r}_k d\hat{x} \\
&= \int_{\widehat{T}} \hat{r}_k \widehat{\mathrm{div}} (I_{\widehat{T}}^{RT^k} \hat{u}) d\hat{x},
\end{aligned}
$$

which leads to $\widehat{\mathrm{div}}(I_{\widehat{T}}^{RT^k} \hat{u}) = 0$. Therefore, form the property of the RT interpolation, $I_{\widehat{T}}^{RT^k} \hat{u} \in \mathbb{P}^k(\widehat{T})^3$, e.g. see [11, Lemma 3.1].

Using (19.4) for $i = 2, 3$, and $\hat{u}_2 = \hat{u}_3 = 0$, we have

$$
\int_{\widehat{E}_i} (I_{\widehat{T}}^{RT^k} \hat{u})_i \hat{r}_k d\hat{s} = 0 \quad \forall \hat{r}_k \in \mathbb{P}^k(\widehat{E}_i), \quad i = 2, 3.
$$

Setting $\hat{r}_k := (I_{\widehat{T}}^{RT^k} \hat{u})_i$, we obtain that $(I_{\widehat{T}}^{RT^k} \hat{u})_i|_{\widehat{E}_i} = 0$ for $i = 2, 3$.

For $k = 0$, because $I_{\widehat{T}}^{RT^0} \hat{u} \in \mathbb{P}^0(\widehat{T})^3$ and $(I_{\widehat{T}}^{RT^0} \hat{u})_i|_{\widehat{E}_i} = 0$ for $i = 2, 3$, it holds that $(I_{\widehat{T}}^{RT^0} \hat{u})_i = 0$ in $\widehat{T}$ for $i = 2, 3$. This implies that the first result holds.

For $k \geq 1$, there exists a polynomial $\hat{r}_i \in \mathbb{P}^{k-1}(\widehat{T})$, $i = 2, 3$, such that $(I_{\widehat{T}}^{RT^k} \hat{u})_i = \hat{x}_i \hat{r}_i$. Using (19.5) for $i = 2, 3$, and $\hat{u}_2 = \hat{u}_3 = 0$, we have, for $i = 2, 3$,

$$
\int_{\widehat{T}} (I_{\widehat{T}}^{RT^k} \hat{u})_i \hat{r}_i d\hat{x} = 0. \quad \text{as } \hat{q}_{k-1} := (0, \hat{r}_i, 0)^\top \text{ in (19.5)},
$$

which leads to

$$
\int_{\widehat{T}} (I_{\widehat{T}}^{RT^k} \hat{u})_i^2 d\hat{x} = \int_{\widehat{T}} \hat{x}_i \hat{x}_i \hat{r}_i^2 d\hat{x} \leq \|\hat{x}_i\|_{L^\infty(\widehat{T})} \int_{\widehat{T}} \hat{x}_i \hat{r}_i^2 d\hat{x} = 0.
$$

Note that $\hat{x}_i \geq 0$ in $\widehat{T}$ for $i = 2, 3$. We hence conclude that $(I_{\widehat{T}}^{RT^k} \hat{u})_i = 0$ in $\widehat{T}$ for $i = 2, 3$.

Because $\widehat{\mathrm{div}}(I_{\widehat{T}}^{RT} \hat{u}) = 0$, it follows that

$$
\frac{\partial (I_{\widehat{T}}^{RT^k} \hat{u})_1}{\partial \hat{x}_1} = 0.
$$

This means that $(I_{\widehat{T}}^{RT^k} \hat{u})_1$ is independent of $\hat{x}_1$.

The other two results are analogous. $\square$

Lemma 19.9. For $k \in \mathbb{N}_0$, there exists a constant c such that, for all $\hat{u} = (\hat{u}_1, \hat{u}_2, \hat{u}_3)^\top \in W^{1,p}(\widehat{T})^3$,

$$
\|(I_{\widehat{T}}^{RT^k} \hat{u})_i\|_{L^p(\widehat{T})} \leq c \left(\|\hat{u}_i\|_{W^{1,p}(\widehat{T})} + \|\widehat{\mathrm{div}}\hat{u}\|_{L^p(\widehat{T})} \right), \quad i = 1, 2, 3. \tag{19.14}
$$

Proof. Only shown when $k = 0$. When $k \geq 1$. see [1, Lemma 3.3].

From Lemma 19.8, if

$$
\hat{v} := (\hat{u}_1, \hat{u}_2 - \hat{u}_2(\hat{x}_1, 0, \hat{x}_3), \hat{u}_3 - \hat{u}_3(\hat{x}_1, \hat{x}_2, 0))^\top,
$$

it holds that

$$I_{\widehat{T}}^{RT^k}\hat{v} = I_{\widehat{T}}^{RT^k}\hat{u} - I_{\widehat{T}}^{RT^k}(0, \hat{u}_2(\hat{x}_1, 0, \hat{x}_3), 0)^\top - I_{\widehat{T}}^{RT^k}(0, 0, \hat{u}_3(\hat{x}_1, \hat{x}_2, 0))^\top,$$

and thus, $(I_{\widehat{T}}^{RT^k}\hat{v})_1 = (I_{\widehat{T}}^{RT^k}\hat{u})_1$.

Let $k = 0$. Because $\hat{v}_2|_{\widehat{E}_2} = 0$ and $\hat{v}_3|_{\widehat{E}_3} = 0$, $I_{\widehat{T}}^{RT^0}\hat{v}$ is determined by the equations

$$\int_{\widehat{E}_1} (I_{\widehat{T}}^{RT^0}\hat{v})_1 d\hat{s} = \int_{\widehat{E}_1} \hat{v}_1 d\hat{s}, \tag{19.15a}$$

$$\int_{\widehat{E}_2} (I_{\widehat{T}}^{RT^0}\hat{v})_2 d\hat{s} = 0, \tag{19.15b}$$

$$\int_{\widehat{E}_3} (I_{\widehat{T}}^{RT^0}\hat{v})_3 d\hat{s} = 0, \tag{19.15c}$$

$$\int_{\widehat{E}_4} \{(I_{\widehat{T}}^{RT^0}\hat{v})_1 + (I_{\widehat{T}}^{RT^0}\hat{v})_2 + (I_{\widehat{T}}^{RT^0}\hat{v})_3\} d\hat{s} = \int_{\widehat{E}_4} (\hat{v}_1 + \hat{v}_2 + \hat{v}_3) d\hat{s}. \tag{19.15d}$$

From the divergence formula and the definition of $\hat{v}$, we have

$$\begin{aligned}\int_{\widehat{T}} \widehat{\operatorname{div}}\hat{v} d\hat{x} &= \int_{\partial\widehat{T}} \hat{v}\cdot\hat{n} d\hat{s} = \frac{1}{\sqrt{3}}\int_{\widehat{E}_4} (\hat{v}_1 + \hat{v}_2 + \hat{v}_3) d\hat{s} + \int_{\partial\widehat{T}\setminus\widehat{E}_4} \hat{v}\cdot\hat{n} d\hat{s} \\ &= \frac{1}{\sqrt{3}}\int_{\widehat{E}_4} (\hat{v}_1 + \hat{v}_2 + \hat{v}_3) d\hat{s} + \int_{\widehat{E}_1} \hat{v}_1 d\hat{s}. \end{aligned} \tag{19.16}$$

Because $\hat{u}_1 = \hat{v}_1$, $\widehat{\operatorname{div}}\hat{u} = \widehat{\operatorname{div}}\hat{v}$, $(I_{\widehat{T}}^{RT^0}\hat{u})_1 = (I_{\widehat{T}}^{RT^0}\hat{v})_1$, (19.15), (19.16), the definition of the Raviart–Thomas interpolation, and the trace theorem, we have

$$\begin{aligned}\|(I_{\widehat{T}}^{RT^0}\hat{u})_1\|_{L^p(\widehat{T})} &= \|(I_{\widehat{T}}^{RT^0}\hat{v})_1\|_{L^p(\widehat{T})} \\ &\le \sum_{i=1}^{4} \left| \int_{\widehat{E}_i} \hat{v}\cdot\widehat{N}_i d\hat{s} \right| \|(\hat{\theta}_i)_1\|_{L^p(\widehat{T})} \\ &\le c \left| \int_{\widehat{E}_1} \hat{v}_1 d\hat{s} + \frac{1}{\sqrt{3}}\int_{\widehat{E}_4} (\hat{v}_1 + \hat{v}_2 + \hat{v}_3) d\hat{s} \right| \\ &\le c \left(\|\hat{u}_1\|_{W^{1,p}(\widehat{T})} + \|\widehat{\operatorname{div}}\hat{u}\|_{L^p(\widehat{T})} \right),\end{aligned}$$

which is the desired result for $k = 0$. By analogous argument, the estimates for $(I_{\widehat{T}}^{RT^0}\hat{u})_i$, $i = 2, 3$, can be proved.

□

19.3.3 Three-dimensional case: Type ii

Let $\widehat{T} \subset \mathbb{R}^3$ be the reference triangle with vertices $\widehat{A}_1 := (1,0,0)^\top$, $\widehat{A}_2 := (1,1,0)^\top$, $\widehat{A}_3 := (0,0,1)^\top$, and $\widehat{A}_4 := (0,0,0)^\top$ with $\widehat{N}_1 := \frac{1}{\sqrt{2}}(-1,1,0)^\top$, $\widehat{N}_2 := (0,-1,0)^\top$, $\widehat{N}_3 := (0,0,-1)^\top$, and $\widehat{N}_4 := \frac{1}{\sqrt{2}}(1,0,1)^\top$. For $1 \le i \le 4$, let $\widehat{E}_i$ be the edge of $\widehat{T}$ opposite to $\widehat{A}_i$ and with $\overline{E}_1$ the projection of $\widehat{E}_1$ onto the plane given by $\hat{x}_1 = 0$.

Lemma 19.10. Let $k \in \mathbb{N}_0$. Let $\hat{f}_1 \in L^p(\overline{E}_1)$, and $\hat{f}_i \in L^p(\widehat{E}_i)$, $i \in \{2, 3\}$. If

$$\begin{aligned}&\hat{u}(\hat{x}) = (\hat{f}_1(\hat{x}_2, \hat{x}_3), 0, 0)^\top, \quad \hat{v}(\hat{x}) = (0, \hat{f}_2(\hat{x}_1, \hat{x}_3), 0)^\top, \\ &\hat{w}(\hat{x}) = (0, 0, \hat{f}_3(\hat{x}_1, \hat{x}_2))^\top,\end{aligned}$$

then there exist polynomials $\hat{q}_1 \in \mathbb{P}^k(\overline{E}_1)$, and $\hat{q}_i \in \mathbb{P}^k(\widehat{E}_i)$, $i \in \{2,3\}$, such that

$$
\begin{aligned}
&I_{\widehat{T}}^{RT^k} \hat{u} = (\hat{q}_1(\hat{x}_2, \hat{x}_3), 0, 0)^\top, \quad I_{\widehat{T}}^{RT^k} \hat{v} = (0, \hat{q}_2(\hat{x}_1, \hat{x}_3), 0)^\top, \\
&I_{\widehat{T}}^{RT^k} \hat{w} = (0, 0, \hat{q}_3(\hat{x}_1, \hat{x}_2))^\top.
\end{aligned}
$$

Proof. We follow [1, Lemma 4.2]. The proof is similar to that of Lemma 19.8. We prove the second equality. The other two follow in an analogous argument.

Because $\widehat{\operatorname{div}}\hat{v} = 0$, from the definition of the RT interpolation and the Green's formula, we have $\widehat{\operatorname{div}}(I_{\widehat{T}}^{RT^k} \hat{v}) = 0$. Therefore, form the property of the RT interpolation, $I_{\widehat{T}}^{RT^k} \hat{v} \in \mathbb{P}^k(\widehat{T})^3$. Using (19.4) for $i = 3$, and $\hat{v}_3 = 0$, we have

$$
\int_{\widehat{E}_3} (I_{\widehat{T}}^{RT^k} \hat{v})_3 \hat{p}_k d\hat{s} = 0 \quad \forall \hat{p}_k \in \mathbb{P}^k(\widehat{E}_3).
$$

Setting $\hat{p}_k := (I_{\widehat{T}}^{RT^k} \hat{v})_3$, we obtain that $(I_{\widehat{T}}^{RT^k} \hat{v})_3|_{\widehat{E}_3} = 0$.

Let $k = 0$. Because $I_{\widehat{T}}^{RT^0} \hat{v} \in \mathbb{P}^0(\widehat{T})^3$ and $(I_{\widehat{T}}^{RT^0} \hat{v})_3|_{\widehat{E}_3} = 0$, it holds that $(I_{\widehat{T}}^{RT^0} \hat{v})_3 = 0$ in $\widehat{T}$. Using (19.4) for $i = 4$, and $\hat{v}_1 = \hat{v}_3 = 0$, we have

$$
\int_{\widehat{E}_4} \{ (I_{\widehat{T}}^{RT^0} \hat{v})_1 + (I_{\widehat{T}}^{RT^0} \hat{v})_3 \} d\hat{s} = 0,
$$

which leads to $(I_{\widehat{T}}^{RT^0} \hat{v})_1|_{\widehat{E}_4} = 0$. It then holds that that $(I_{\widehat{T}}^{RT^0} \hat{v})_1 = 0$ in $\widehat{T}$. This implies that the second result holds.

Let $k \geq 1$. As in the proof of Lemma 19.8, we obtain that $(I_{\widehat{T}}^{RT^k} \hat{v})_3 = 0$ in $\widehat{T}$. Using (19.4) for $i = 4$, and $\hat{v}_1 = \hat{v}_3 = 0$, we have

$$
\int_{\widehat{E}_4} \{ (I_{\widehat{T}}^{RT^k} \hat{v})_1 + (I_{\widehat{T}}^{RT^k} \hat{v})_3 \} \hat{p}_k d\hat{s} = 0 \quad \forall \hat{p}_k \in \mathbb{P}^k(\widehat{E}_4),
$$

which implies that $\{ (I_{\widehat{T}}^{RT^k} \hat{v})_1 + (I_{\widehat{T}}^{RT^k} \hat{v})_3 \}|_{\widehat{E}_4} = 0$, and hence

$$
(I_{\widehat{T}}^{RT^k} \hat{v})_1 + (I_{\widehat{T}}^{RT^k} \hat{v})_3 = (1 - \hat{x}_1 - \hat{x}_3) \hat{r}
$$

for some $\hat{r} \in \mathbb{P}^{k-1}(\widehat{T})$. Using (19.5) and $\hat{v}_1 = \hat{v}_3 = 0$, we have

$$
\int_{\widehat{T}} \{ (I_{\widehat{T}}^{RT^k} \hat{v})_1 + (I_{\widehat{T}}^{RT^k} \hat{v})_3 \} \hat{r} d\hat{x} = 0. \quad \text{as } \hat{q}_{k-1} := (\hat{r}, 0, \hat{r})^\top \text{ in (19.5)},
$$

which leads to

$$
\begin{aligned}
\int_{\widehat{T}} \{ (I_{\widehat{T}}^{RT^k} \hat{v})_1 + (I_{\widehat{T}}^{RT^k} \hat{v})_3 \}^2 d\hat{x} &= \int_{\widehat{T}} (1 - \hat{x}_1 - \hat{x}_3)^2 \hat{r}^2 d\hat{x} \\
&\leq \| 1 - \hat{x}_1 - \hat{x}_3 \|_{L^\infty(\widehat{T})} \int_{\widehat{T}} (1 - \hat{x}_1 - \hat{x}_3) \hat{r}^2 d\hat{x} = 0.
\end{aligned}
$$

Note that $1 - \hat{x}_1 - \hat{x}_3 \geq 0$ in $\widehat{T}$. We hence have $(I_{\widehat{T}}^{RT^k} \hat{v})_1 + (I_{\widehat{T}}^{RT^k} \hat{v})_3 = 0$ in $\widehat{T}$. Because we know $(I_{\widehat{T}}^{RT^k} \hat{v})_3 = 0$ in $\widehat{T}$, we conclude that $(I_{\widehat{T}}^{RT^k} \hat{v})_1 = 0$ in $\widehat{T}$.

Because $\widehat{\operatorname{div}}(I_{\widehat{T}}^{RT^k} \hat{v}) = 0$, it follows that

$$
\frac{\partial (I_{\widehat{T}}^{RT^k} \hat{v})_2}{\partial \hat{x}_2} = 0.
$$

This means that $(I_{\widehat{T}}^{RT^k} \hat{v})_2$ is independent of $\hat{x}_2$. $\square$

Lemma 19.11. For $k \in \mathbb{N}_0$, there exists a constant c such that, for all $\hat{u} = (\hat{u}_1, \hat{u}_2, \hat{u}_3)^\top \in W^{1,p}(\widehat{T})^3$,

$$\|(I_{\widehat{T}}^{R^kT} \hat{u})_i\|_{L^p(\widehat{T})} \leq c \left(\|\hat{u}_1\|_{W^{1,p}(\widehat{T})} + \left\| \frac{\partial \hat{u}_2}{\partial \hat{x}_2} \right\|_{L^p(\widehat{T})} + \left\| \frac{\partial \hat{u}_3}{\partial \hat{x}_3} \right\|_{L^p(\widehat{T})} \right), \tag{19.17a}$$

$$\|(I_{\widehat{T}}^{RT^k} \hat{u})_i\|_{L^p(\widehat{T})} \leq c \left(\|\hat{u}_i\|_{W^{1,p}(\widehat{T})} + \|\mathrm{div}\, \hat{u}\|_{L^p(\widehat{T})} \right), \quad i = 2, 3. \tag{19.17b}$$

In particular,

$$\|(I_{\widehat{T}}^{RT^k} \hat{u})_i\|_{L^p(\widehat{T})} \leq c \left(\|\hat{u}_i\|_{W^{1,p}(\widehat{T})} + \sum_{j=1, j\neq i}^{3} \left\| \frac{\partial \hat{u}_j}{\partial \hat{x}_j} \right\|_{L^p(\widehat{T})} \right), \quad i = 1, 2, 3. \tag{19.18}$$

Proof. Only shown when $k = 0$. When $k \geq 1$. see [1, Lemma 4.3]. We prove the estimates (19.17a) and (19.17b) with $i = 2$. The other one follows in an analogous argument.

Case for (19.17a)**.** From Lemma 19.10, if

$$\hat{v} := (\hat{u}_1, \hat{u}_2 - \hat{u}_2(\hat{x}_1, 0, \hat{x}_3), \hat{u}_3 - \hat{u}_3(\hat{x}_1, \hat{x}_2, 0))^\top,$$

it holds that

$$I_{\widehat{T}}^{RT^k} \hat{v} = I_{\widehat{T}}^{RT^k} \hat{u} - I_{\widehat{T}}^{RT^k} (0, \hat{u}_2(\hat{x}_1, 0, \hat{x}_3), 0)^\top - I_{\widehat{T}}^{RT^k} (0, 0, \hat{u}_3(\hat{x}_1, \hat{x}_2, 0))^\top,$$

and thus, $(I_{\widehat{T}}^{RT^k} \hat{v})_1 = (I_{\widehat{T}}^{RT^k} \hat{u})_1$.

Let $k = 0$. Because $\hat{v}_2|_{\widehat{E}_2} = 0$ and $\hat{v}_3|_{\widehat{E}_3} = 0$, $I_{\widehat{T}}^{RT^0} \hat{v}$ is determined by the equations

$$\int_{\widehat{E}_1} \{-(I_{\widehat{T}}^{RT^0} \hat{v})_1 + (I_{\widehat{T}}^{RT^0} \hat{v})_2\} d\hat{s} = \int_{\widehat{E}_1} (-\hat{v}_1 + \hat{v}_2) d\hat{s}, \tag{19.19a}$$

$$\int_{\widehat{E}_2} (I_{\widehat{T}}^{RT^0} \hat{v})_2 d\hat{s} = 0, \tag{19.19b}$$

$$\int_{\widehat{E}_3} (I_{\widehat{T}}^{RT^0} \hat{v})_3 d\hat{s} = 0, \tag{19.19c}$$

$$\int_{\widehat{E}_4} \{(I_{\widehat{T}}^{RT^0} \hat{v})_1 + (I_{\widehat{T}}^{RT^0} \hat{v})_3\} d\hat{s} = \int_{\widehat{E}_4} (\hat{v}_1 + \hat{v}_3) d\hat{s}. \tag{19.19d}$$

From the divergence formula and the definition of $\hat{v}$, we have

$$\begin{aligned} \int_{\widehat{T}} \widehat{\mathrm{div}} (\hat{v}_1, \hat{v}_2, 0)^\top d\hat{x} &= \int_{\partial \widehat{T}} (\hat{v}_1, \hat{v}_2, 0)^\top \cdot \hat{n} d\hat{s} \\ &= \frac{1}{\sqrt{2}} \int_{\widehat{E}_1} (-\hat{v}_1 + \hat{v}_2) d\hat{s} + \frac{1}{\sqrt{2}} \int_{\widehat{E}_4} \hat{v}_1 d\hat{s}, \end{aligned} \tag{19.20}$$

and

$$\begin{aligned} \int_{\widehat{T}} \widehat{\mathrm{div}} (\hat{v}_1, 0, \hat{v}_3)^\top d\hat{x} &= \int_{\partial \widehat{T}} (\hat{v}_1, 0, \hat{v}_3)^\top \cdot \hat{n} d\hat{s} \\ &= \frac{1}{\sqrt{2}} \int_{\widehat{E}_1} (-\hat{v}_1) d\hat{s} + \frac{1}{\sqrt{2}} \int_{\widehat{E}_4} (\hat{v}_1 + \hat{v}_3) d\hat{s}, \end{aligned} \tag{19.21}$$

Because $\hat{u}_1 = \hat{v}_1$, $\frac{\partial \hat{u}_j}{\partial \hat{x}_j} = \frac{\partial \hat{v}_j}{\partial \hat{x}_j}$ for $j = 2, 3$, $(I_{\widehat{T}}^{RT^0} \hat{u})_1 = (I_{\widehat{T}}^{RT^0} \hat{v})_1$, (19.19), (19.20), (19.21), the definition of the RT interpolation, and the trace theorem, we have

$$\begin{aligned} \|(I_{\widehat{T}}^{RT^0} \hat{u})_1\|_{L^p(\widehat{T})} &= \|(I_{\widehat{T}}^{RT^0} \hat{v})_1\|_{L^p(\widehat{T})} \\ &\le \sum_{j=1}^{4} \left| \int_{\widehat{E}_j} \hat{v} \cdot \widehat{N}_j d\hat{s} \right| \|(\hat{\theta}_j)_1\|_{L^p(\widehat{T})} \\ &\le c \left| \frac{1}{\sqrt{2}} \int_{\widehat{E}_1} (-\hat{v}_1 + \hat{v}_2) d\hat{s} + \frac{1}{\sqrt{2}} \int_{\widehat{E}_4} (\hat{v}_1 + \hat{v}_3) d\hat{s} \right| \\ &\le c \left(\|\hat{u}_1\|_{W^{1,p}(\widehat{T})} + \left\| \frac{\partial \hat{u}_2}{\partial \hat{x}_2} \right\|_{L^p(\widehat{T})} + \left\| \frac{\partial \hat{u}_3}{\partial \hat{x}_3} \right\|_{L^p(\widehat{T})} \right), \end{aligned}$$

which is the desired result for $k = 0$.

Case for (19.17b) **with** $i = 2$**.** From Lemma 19.10, if

$$\hat{v} := (\hat{u}_1 - \hat{u}_1(\hat{x}_2, \hat{x}_2, \hat{x}_3), \hat{u}_2, \hat{u}_3 - \hat{u}_3(\hat{x}_1, \hat{x}_2, 0))^\top,$$

it holds that

$$I_{\widehat{T}}^{RT^k} \hat{v} = I_{\widehat{T}}^{RT^k} \hat{u} - I_{\widehat{T}}^{RT^k} (\hat{u}_1(\hat{x}_2, \hat{x}_2, \hat{x}_3), 0, 0)^\top - I_{\widehat{T}}^{RT^k} (0, 0, \hat{u}_3(\hat{x}_1, \hat{x}_2, 0))^\top,$$

and thus, $(I_{\widehat{T}}^{RT^k} \hat{v})_2 = (I_{\widehat{T}}^{RT^k} \hat{u})_2$.

Let $k = 0$. Because $\hat{v}_1|_{\widehat{E}_1} = 0$ and $\hat{v}_3|_{\widehat{E}_3} = 0$, $I_{\widehat{T}}^{RT^0} \hat{v}$ is determined by the equations

$$\int_{\widehat{E}_1} \{-(I_{\widehat{T}}^{RT^0} \hat{v})_1 + (I_{\widehat{T}}^{RT^0} \hat{v})_2\} d\hat{s} = \int_{\widehat{E}_1} \hat{v}_2 d\hat{s}, \tag{19.22a}$$

$$\int_{\widehat{E}_2} (I_{\widehat{T}}^{RT^0} \hat{v})_2 d\hat{s} = \int_{\widehat{E}_2} \hat{v}_2 d\hat{s}, \tag{19.22b}$$

$$\int_{\widehat{E}_3} (I_{\widehat{T}}^{RT^0} \hat{v})_3 d\hat{s} = 0, \tag{19.22c}$$

$$\int_{\widehat{E}_4} \{(I_{\widehat{T}}^{RT^0} \hat{v})_1 + (I_{\widehat{T}}^{RT^0} \hat{v})_3\} d\hat{s} = \int_{\widehat{E}_4} (\hat{v}_1 + \hat{v}_3) d\hat{s}. \tag{19.22d}$$

From the divergence formula and the definition of $\hat{v}$, we have

$$\begin{aligned} \int_{\widehat{T}} \widehat{\operatorname{div}} (\hat{v}_1, \hat{v}_2, \hat{v}_3)^\top d\hat{x} &= \int_{\partial \widehat{T}} (\hat{v}_1, \hat{v}_2, \hat{v}_3)^\top \cdot \hat{n} d\hat{s} \\ &= \frac{1}{\sqrt{2}} \int_{\widehat{E}_1} \hat{v}_2 d\hat{s} - \int_{\widehat{E}_2} \hat{v}_2 d\hat{s} + \frac{1}{\sqrt{2}} \int_{\widehat{E}_4} (\hat{v}_1 + \hat{v}_3) d\hat{s}, \end{aligned} \tag{19.23}$$

Because $\hat{u}_2 = \hat{v}_2$, $\frac{\partial \hat{u}_j}{\partial \hat{x}_j} = \frac{\partial \hat{v}_j}{\partial \hat{x}_j}$ for $j = 1, 3$, $(I_{\widehat{T}}^{RT} \hat{u})_2 = (I_{\widehat{T}}^{RT} \hat{v})_2$, (19.22), (19.23), the definition of the RT interpolation, and the trace theorem, we have

$$\begin{aligned} \|(I_{\widehat{T}}^{RT^0} \hat{u})_2\|_{L^p(\widehat{T})} &= \|(I_{\widehat{T}}^{RT^0} \hat{v})_2\|_{L^p(\widehat{T})} \\ &\le \sum_{j=1}^{4} \left| \int_{\widehat{E}_j} \hat{v} \cdot \widehat{N}_j d\hat{s} \right| \|(\hat{\theta}_j)_2\|_{L^p(\widehat{T})} \\ &\le c \left| \frac{1}{\sqrt{2}} \int_{\widehat{E}_1} \hat{v}_2 d\hat{s} + \int_{\widehat{E}_2} \hat{v}_2 d\hat{s} + \frac{1}{\sqrt{2}} \int_{\widehat{E}_4} (\hat{v}_1 + \hat{v}_3) d\hat{s} \right| \\ &\le c \left(\|\hat{u}_2\|_{W^{1,p}(\widehat{T})} + \left\| \widehat{\operatorname{div}} \hat{u} \right\|_{L^p(\widehat{T})} \right), \end{aligned}$$

which is the desired result for $k = 0$.

□

19.4 Stability of the local RT interpolation

The following two lemmata are divided into the element on $\mathfrak{T}^{(2)}$ or $\mathfrak{T}_1^{(3)}$ and the element on $\mathfrak{T}_2^{(3)}$.

Lemma 19.12. Let $p \in [1,\infty)$. Let $T \in \mathbb{T}_h$ satisfy Condition 5.1 or Condition 5.2 with $T = \Phi_T(\widetilde{T})$ and $\widetilde{T} = \Phi_{\widetilde{T}}(\widehat{T})$, where $\widetilde{T} \in \mathfrak{T}^{(2)}$ or $\widetilde{T} \in \mathfrak{T}_1^{(3)}$. Then, for any $\hat{v} \in W^{1,p}(\widehat{T})^d$ with $\tilde{v} = \Psi_{\widehat{T}}\hat{v}$ and $v = \Psi_{\widetilde{T}}\tilde{v}$,

$$\|I_T^{RT^k} v\|_{L^p(T)^d} \le c \left[\frac{H_T}{h_T} \left(\|v\|_{L^p(T)^d} + \sum_{|\varepsilon|=1} h^{\varepsilon} \|\partial_r^{\varepsilon} v\|_{L^p(T)^d} \right) + h_T \|\nabla \cdot v\|_{L^p(T)} \right]. \tag{19.24}$$

Proof. From (18.5),

$$\|I_T^{RT^k} v\|_{L^p(T)^d} \le c|\det(A_{\widetilde{T}})|^{-\frac{p-1}{p}} \|\widetilde{A}\|_2 \left(\sum_{j=1}^d h_j^p \|(I_{\widehat{T}}^{RT^k} \hat{v})_j\|^p_{L^p(\widehat{T})} \right)^{1/p}. \tag{19.25}$$

The component-wise stability (19.13) for $2d$ or (19.14) for $3d$ yields

$$\sum_{j=1}^d h_j^p \|(I_{\widehat{T}}^{RT^k} \hat{v})_j\|^p_{L^p(\widehat{T})} \le c \sum_{j=1}^d h_j^p \left(\|\hat{v}_j\|^p_{W^{1,p}(\widehat{T})} + \|\nabla_{\hat{x}} \cdot \hat{v}\|^p_{L^p(\widehat{T})} \right). \tag{19.26}$$

From (18.6) with $\ell = 0$ and $m \in \{0,1\}$,

$$\begin{aligned} \|\hat{v}_j\|^p_{W^{1,p}(\widehat{T})} &= \|\hat{v}_j\|^p_{L^p(\widehat{T})} + \sum_{k=1}^d \left\| \frac{\partial \hat{v}_j}{\partial \hat{x}_k} \right\|^p_{L^p(\widehat{T})} \\ &\le c|\det(A_{\widetilde{T}})|^{p-1} \|\widetilde{A}^{-1}\|_2^p h_j^{-p} \left[\|v\|^p_{L^p(T)^d} + \left(\sum_{|\varepsilon|=1} h^{\varepsilon} \|\partial_r^{\varepsilon} v\|_{L^p(T)^d} \right)^p \right]. \end{aligned} \tag{19.27}$$

From (18.8) with $\ell = 0$,

$$\|\nabla_{\hat{x}} \cdot \hat{v}\|_{L^p(\widehat{T}_1)} \le c|\det(A_{\widetilde{T}})|^{\frac{p-1}{p}} \|\nabla \cdot v\|_{L^p(T)}. \tag{19.28}$$

Combining the above inequalities (19.25), (19.26), (19.27), and (19.28) with (6.8b) and (1.5) yields

$$\|I_T^{RT^k} v\|_{L^p(T)^d} \le c \left[\frac{H_T}{h_T} \left(\|v\|_{L^p(T)^d} + \sum_{|\varepsilon|=1} h^{\varepsilon} \|\partial_r^{\varepsilon} v\|_{L^p(T)^d} \right) + h_T \|\nabla \cdot v\|_{L^p(T)} \right],$$

which is the desired estimate. □

Lemma 19.13. Let $p \in [1,\infty)$ and $d = 3$. Let $T \in \mathbb{T}_h$ satisfy Condition 5.2 with $T = \Phi_T(\widetilde{T})$ and $\widetilde{T} = \Phi_{\widetilde{T}}(\widehat{T})$, where $\widetilde{T} \in \mathfrak{T}_2^{(3)}$. Then, for any $\hat{v} \in W^{\ell+1,p}(\widehat{T})^3$ with $\tilde{v} = \Psi_{\widehat{T}}\hat{v}$ and $v = \Psi_{\widetilde{T}}\tilde{v}$,

$$\|I_T^{RT^k} v\|_{L^p(T)^3} \le c \frac{H_T}{h_T} \left[\|v\|_{L^p(T)^3} + h_T \sum_{k=1}^3 \left\| \frac{\partial v}{\partial r_k} \right\|_{L^p(T)^3} \right]. \tag{19.29}$$

Proof. The component-wise stability (19.17) yields

$$\sum_{j=1}^{3} h_j^p \|(I_{\widehat{T}}^{RT^k}\hat{v})_j\|_{L^p(\widehat{T})}^p \leq c \sum_{j=1}^{3} h_j^p \left(\|\hat{v}_j\|_{W^{1,p}(\widehat{T})}^p + \sum_{k=1,k\neq j}^{3} \left\| \frac{\partial \hat{v}_k}{\partial \hat{x}_k} \right\|_{L^p(\widehat{T})}^p \right). \tag{19.30}$$

From (18.10) with $\ell = 0$,

$$\left\| \frac{\partial \hat{v}_k}{\partial \hat{x}_k} \right\|_{L^p(\widehat{T})} \leq c |\det(A_{\widetilde{T}})|^{\frac{p-1}{p}} \|\widetilde{A}^{-1}\|_2 \left\| \frac{\partial v}{\partial r_k} \right\|_{L^p(T)^3}. \tag{19.31}$$

By analogous argument in Lemma 19.12,

$$\begin{aligned} \|\hat{v}_j\|_{W^{1,p}(\widehat{T})}^p &= \|\hat{v}_j\|_{L^p(\widehat{T})}^p + \sum_{k=1}^{3} \left\| \frac{\partial \hat{v}_j}{\partial \hat{x}_k} \right\|_{L^p(\widehat{T})}^p \\ &\leq c |\det(A_{\widetilde{T}})|^{p-1} \|\widetilde{A}^{-1}\|_2^p h_j^{-p} \left[\|v\|_{L^p(T)^3}^p + \left(\sum_{|\varepsilon|=1} h^{\varepsilon} \|\partial_r^{\varepsilon} v\|_{L^p(T)^3} \right)^p \right]. \end{aligned} \tag{19.32}$$

Combining the above inequalities (19.25), (19.30), (19.31), and (19.32) with (6.8b) and (1.5) yields

$$\begin{aligned} \|I_T^{RT^k} v\|_{L^p(T)^3} &\leq c |\det(A_{\widetilde{T}})|^{-\frac{p-1}{p}} \|\widetilde{A}\|_2 \left(\sum_{j=1}^{d} h_j^p \|(I_{\widehat{T}}^{RT^k}\hat{v})_j\|_{L^p(\widehat{T})}^p \right)^{1/p} \\ &\leq c \frac{H_T}{h_T} \left[\|v\|_{L^p(T)^3} + \sum_{|\varepsilon|=1} h^{\varepsilon} \|\partial_r^{\varepsilon} v\|_{L^p(T)^3} + \sum_{j=1}^{3} h_j \sum_{k=1,k\neq j}^{3} \left\| \frac{\partial v}{\partial r_k} \right\|_{L^p(T)^3} \right], \end{aligned}$$

which is the desired result. □

19.5 Local RT Interpolation Error Estimates

The following two theorems are divided into the element on $\mathfrak{T}^{(2)}$ or $\mathfrak{T}_1^{(3)}$ and the element on $\mathfrak{T}_2^{(3)}$.

Theorem 19.14. Let $p \in [1, \infty)$. Let $T \in \mathbb{T}_h$ satisfy Condition 5.1 or Condition 5.2 with $T = \Phi_T(\widetilde{T})$ and $\widetilde{T} = \Phi_{\widetilde{T}}(\widehat{T})$, where $\widetilde{T} \in \mathfrak{T}^{(2)}$ or $\widetilde{T} \in \mathfrak{T}_1^{(3)}$. For $k \in \mathbb{N}_0$, let $\{T, \mathbb{RT}^k(T), \Sigma\}$ be the RT finite element and $I_T^{RT^k}$ the local interpolation operator defined in (19.9). Let ℓ be such that $0 \leq \ell \leq k$. Then, for any $\hat{v} \in W^{\ell+1,p}(\widehat{T})^d$ with $\tilde{v} = \Psi_{\widetilde{T}}\hat{v}$ and $v = \Psi_{\widetilde{T}}\tilde{v}$,

$$\|I_T^{RT^k} v - v\|_{L^p(T)^d} \leq c \left(\frac{H_T}{h_T} \sum_{|\varepsilon|=\ell+1} h^{\varepsilon} \|\partial_r^{\varepsilon} v\|_{L^p(T)^d} + h_T \sum_{|\beta|=\ell} h^{\beta} \|\partial_r^{\beta} \nabla \cdot v\|_{L^p(T)} \right). \tag{19.33}$$

If Condition 10.1 is imposed, it holds that

$$\begin{aligned} \|I_T^{RT^k} v - v\|_{L^p(T)^d} \leq c \Bigg(&\frac{H_T}{h_T} \sum_{|\varepsilon|=\ell+1} \widetilde{\mathscr{H}}^{\varepsilon} \|\partial_{\tilde{x}}^{\varepsilon} (\Psi_{\widetilde{T}}^{-1} v)\|_{L^p(\Phi_T^{-1}(T))^d} \\ &+ h_T \sum_{|\beta|=\ell} \widetilde{\mathscr{H}}^{\beta} \|\partial_{\tilde{x}}^{\beta} \nabla_{\tilde{x}} \cdot (\Psi_{\widetilde{T}}^{-1} v)\|_{L^p(\Phi_T^{-1}(T))} \Bigg). \end{aligned} \tag{19.34}$$

Proof. Let $\hat{v} \in W^{\ell+1,p}(\widehat{T})^d$. Let $I_{\widehat{T}}^{RT^k}$ be the local interpolation operators on $\widehat{T}$ defined by (19.4) and (19.5). If $q \in \mathbb{P}^\ell(T)^d \subset \mathbb{RT}^k(T)$, then $I_T^{RT} q = q$.

We set $\mathfrak{Q}^{(\ell+1)} v := (Q^{(\ell+1)} v_1, \ldots, Q^{(\ell+1)} v_d)^\top \in \mathbb{P}^\ell(T)^d$, where $Q^{(\ell+1)} v_j$ is defined by (1.11) for any j. We then obtain

$$\|I_T^{RT^k} v - v\|_{L^p(T)^d} \leq \|I_T^{RT^k}(v - \mathfrak{Q}^{(\ell+1)} v)\|_{L^p(T)^d} + \|\mathfrak{Q}^{(\ell+1)} v - v\|_{L^p(T)^d}. \tag{19.35}$$

The inequality (18.5) for the first term on the right-hand side of (19.35) yield

$$\|I_T^{RT^k}(v - \mathfrak{Q}^{(\ell+1)} v)\|_{L^p(T)^d} \leq c|\det(A_{\widetilde{T}})|^{\frac{1-p}{p}} \|\widetilde{A}\|_2 \left(\sum_{j=1}^d h_j^p \|\{I_{\widehat{T}}^{RT^k}(\hat{v} - \widehat{\mathfrak{Q}}^{(\ell+1)} \hat{v})\}_j\|_{L^p(\widehat{T})}^p \right)^{1/p}. \tag{19.36}$$

The component-wise stability (19.13) for $2d$ or (19.14) for $3d$ yields

$$\begin{aligned} &\sum_{j=1}^d h_j^p \|\{I_{\widehat{T}_1}^{RT^k}(\hat{v} - \widehat{\mathfrak{Q}}^{(\ell+1)} \hat{v})\}_j\|_{L^p(\widehat{T})}^p \\ &\qquad \leq c \sum_{j=1}^d h_j^p \left(\|\hat{v}_j - \widehat{Q}^{(\ell+1)} \hat{v}_j\|_{W^{1,p}(\widehat{T})}^p + \|\nabla_{\hat{x}} \cdot (\hat{v} - \widehat{\mathfrak{Q}}^{(\ell+1)} \hat{v})\|_{L^p(\widehat{T})}^p \right). \end{aligned} \tag{19.37}$$

The inequality (18.5) for the second term on the right-hand side of (19.35) yields

$$\|\mathfrak{Q}^{(\ell+1)} v - v\|_{L^p(T)^d} \leq c|\det(A_{\widetilde{T}})|^{\frac{1-p}{p}} \|\widetilde{A}\|_2 \left(\sum_{j=1}^d h_j^p \|\widehat{Q}^{(\ell+1)} \hat{v}_j - \hat{v}_j\|_{L^p(\widehat{T})}^p \right)^{1/p}. \tag{19.38}$$

The Bramble–Hilbert-type lemma (Lemma 1.10) and (18.6),

$$\begin{aligned} \|\hat{v}_j - \widehat{Q}^{(\ell+1)} \hat{v}_j\|_{W^{1,p}(\widehat{T})}^p &= \|\hat{v}_j - \widehat{Q}^{(\ell+1)} \hat{v}_j\|_{L^p(\widehat{T})}^p + \sum_{k=1}^d \left\| \frac{\partial}{\partial \hat{x}_k} (\hat{v}_j - \widehat{Q}^{(\ell+1)} \hat{v}_j) \right\|_{L^p(\widehat{T})}^p \\ &\leq c \left(\sum_{|\gamma| = \ell+1} \|\partial_{\hat{x}}^\gamma \hat{v}_j\|_{L^p(\widehat{T})}^p + \sum_{k=1}^d \sum_{|\beta| = \ell} \left\| \partial_{\hat{x}}^\beta \frac{\partial \hat{v}_j}{\partial \hat{x}_k} \right\|_{L^p(\widehat{T})}^p \right) \\ &\leq c|\det(A_{\widetilde{T}})|^{p-1} h_j^{-p} \|\widetilde{A}^{-1}\|_2^p \left(\sum_{|\varepsilon| = \ell+1} h^\varepsilon \|\partial_r^\varepsilon v\|_{L^p(T)^d} \right)^p. \end{aligned} \tag{19.39}$$

Because from [14, Proposition 4.1.17] it holds that

$$\widehat{\operatorname{div}}(\widehat{\mathfrak{Q}}^{(\ell+1)} \hat{v}) = \widehat{Q}^\ell(\widehat{\operatorname{div}} \hat{v}). \tag{19.40}$$

From the Bramble–Hilbert-type lemma (Lemma 1.10) and (18.8),

$$\begin{aligned} \|\nabla_{\hat{x}} \cdot (\hat{v} - \widehat{\mathfrak{Q}}^{(\ell+1)} \hat{v}))\|_{L^p(\widehat{T})}^p &= \|\nabla_{\hat{x}} \cdot \hat{v} - \widehat{Q}^\ell(\nabla_{\hat{x}} \cdot \hat{v})\|_{L^p(\widehat{T})}^p \\ &\leq \|\nabla_{\hat{x}} \cdot \hat{v} - \widehat{Q}^\ell(\nabla_{\hat{x}} \cdot \hat{v})\|_{W^{\ell,p}(\widehat{T})}^p \\ &\leq c|\nabla_{\hat{x}} \cdot \hat{v}|_{W^{\ell,p}(\widehat{T})}^p = c \sum_{|\beta| = \ell} \|\partial^\beta \nabla_{\hat{x}} \cdot \hat{v}\|_{L^p(\widehat{T})}^p \\ &\leq c|\det(A_{\widetilde{T}})|^{p-1} \left(\sum_{|\varepsilon| = \ell} h^\varepsilon \|\partial_r^\varepsilon \nabla \cdot v\|_{L^p(T)} \right)^p. \end{aligned} \tag{19.41}$$

Combining (19.36), (19.37), (19.39), and (19.41) with (6.8b) yields

$$\|I_T^{RT^k}(v - \mathfrak{Q}^{(\ell+1)}v)\|_{L^p(T)^d} \leq c\left(\frac{H_T}{h_T}\sum_{|\varepsilon|=\ell+1} h^{\varepsilon}\,\|\partial_r^{\varepsilon} v\|_{L^p(T)^d} + h_T\sum_{|\beta|=\ell} h^{\beta}\|\partial_r^{\beta}\nabla\cdot v\|_{L^p(T)}\right). \tag{19.42}$$

Furthermore, using a similar argument, from the Bramble–Hilbert-type lemma (Lemma 1.10), (18.6), and (19.38) together with (6.8b),

$$\|\mathfrak{Q}^{(\ell+1)}v - v\|_{L^p(T)^d} \leq c\frac{H_T}{h_T}\sum_{|\varepsilon|=\ell+1} h^{\varepsilon}\,\|\partial_r^{\varepsilon} v\|_{L^p(T)^d}. \tag{19.43}$$

Therefore, from (19.35), (19.42), and (19.43), we have (19.33).

Case in which Condition 10.1 is imposed. From the Bramble–Hilbert-type lemma (Lemma 1.10) and (18.7),

$$\begin{aligned}
\|\hat{v}_j - \widehat{Q}^{(\ell+1)}\hat{v}_j\|^p_{W^{1,p}(\widehat{T})} &\leq c|\hat{v}_j|^p_{W^{\ell+1,p}(\widehat{T}_1)} + c\sum_{k=1}^{d}\left|\frac{\partial\hat{v}_j}{\partial\hat{x}_k}\right|^p_{W^{\ell,p}(\widehat{T})} \\
&= c\left(\sum_{|\gamma|=\ell+1}\|\partial_{\hat{x}}^{\gamma}\hat{v}_j\|^p_{L^p(\widehat{T})} + \sum_{k=1}^{d}\sum_{|\beta|=\ell}\left\|\partial_{\hat{x}}^{\beta}\frac{\partial\hat{v}_j}{\partial\hat{x}_k}\right\|^p_{L^p(\widehat{T})}\right) \\
&\leq c|\det(A_{\widetilde{T}})|^{p-1}h_j^{-p}\|\widetilde{A}^{-1}\|_2^p\left(\sum_{|\varepsilon|=\ell+1}\widetilde{\mathscr{H}^{\varepsilon}}\|\partial_{\tilde{x}}^{\varepsilon}\tilde{v}\|_{L^p(\widetilde{T})^d}\right)^p.
\end{aligned} \tag{19.44}$$

Because (19.40), from the Bramble–Hilbert-type lemma (Lemma 1.10) and (18.9),

$$\|\nabla_{\hat{x}}\cdot(\hat{v} - \widehat{\mathfrak{Q}}^{(\ell+1)}\hat{v}))\|^p_{L^p(\widehat{T})} \leq c|\det(A_{\widetilde{T}})|^{p-1}\left(\sum_{|\varepsilon|=\ell}\widetilde{\mathscr{H}^{\varepsilon}}\,\|\partial_{\tilde{x}}^{\varepsilon}\nabla_{\tilde{x}}\cdot\tilde{v}\|_{L^p(\widetilde{T})}\right)^p. \tag{19.45}$$

Combining (19.36), (19.37), (19.44), and (19.45) with (6.8b) yields

$$\|I_T^{RT^k}(v - \mathfrak{Q}^{(\ell+1)}v)\|_{L^p(T)^d} \leq c\left(\frac{H_T}{h_T}\sum_{|\varepsilon|=\ell+1}\widetilde{\mathscr{H}^{\varepsilon}}\|\partial_{\tilde{x}}^{\varepsilon}\tilde{v}\|_{L^p(\widetilde{T})^d} + h_T\sum_{|\beta|=\ell}\widetilde{\mathscr{H}^{\beta}}\|\partial_{\tilde{x}}^{\beta}\nabla_{\tilde{x}}\cdot\tilde{v}\|_{L^p(\widetilde{T})}\right). \tag{19.46}$$

Furthermore, using a similar argument, from the Bramble–Hilbert-type lemma (Lemma 1.10), (18.7), and (19.38) together with (6.8b),

$$\|\mathfrak{Q}^{(\ell+1)}v - v\|_{L^p(T)^d} \leq c\frac{H_T}{h_T}\sum_{|\varepsilon|=\ell+1}\widetilde{\mathscr{H}^{\varepsilon}}\|\partial_{\tilde{x}}^{\varepsilon}\tilde{v}\|_{L^p(\widetilde{T})^d}, \tag{19.47}$$

Therefore, from (19.35), (19.46), and (19.47), we have (19.34). □

Theorem 19.15. Let $p \in [1,\infty)$ and $d = 3$. Let $T \in \mathbb{T}_h$ satisfy Condition 5.2 with $T = \Phi_T(\widetilde{T})$ and $\widetilde{T} = \Phi_{\widetilde{T}}(\widehat{T})$, where $\widetilde{T} \in \mathfrak{T}_2^{(3)}$. For $k \in \mathbb{N}_0$, let $\{T, \mathbb{RT}^k(T), \Sigma\}$ be the RT finite element and $I_T^{RT^k}$ the local interpolation operator defined in (19.9). Let ℓ be such that $0 \leq \ell \leq k$. Then, for any $\hat{v} \in W^{\ell+1,p}(\widehat{T})^d$ with $\tilde{v} = \Psi_{\widehat{T}}\hat{v}$ and $v = \Psi_{\widetilde{T}}\tilde{v}$,

$$\|I_T^{RT^k}v - v\|_{L^p(T)^3} \leq c\frac{H_T}{h_T}\left(h_T\sum_{k=1}^{3}\sum_{|\varepsilon|=\ell} h^{\varepsilon}\left\|\partial_r^{\varepsilon}\frac{\partial v}{\partial r_k}\right\|_{L^p(T)^3}\right). \tag{19.48}$$

If Condition 10.1 is imposed, it holds that

$$\|I_T^{RT^k} v - v\|_{L^p(T)^3} \leq c \frac{H_T}{h_T} \left(\sum_{|\varepsilon|=\ell+1} \widetilde{\mathscr{H}^{\varepsilon}} \|\partial_{\tilde{x}}^{\varepsilon} (\Psi_{\widetilde{T}}^{-1} v)\|_{L^p(\Phi_T^{-1}(T))^3} + h_T \sum_{k=1}^{3} \sum_{|\varepsilon|=\ell} \widetilde{\mathscr{H}^{\varepsilon}} \left\| \partial_{\tilde{x}}^{\varepsilon} \frac{\partial (\Psi_{\widetilde{T}}^{-1} v)}{\partial \tilde{r}_k^s} \right\|_{L^p(\Phi_T^{-1}(T))^3} \right). \tag{19.49}$$

Proof. An analogous proof of Theorem 19.14 yields the desired result (19.48), where we use Lemma 19.11 instead of Lemma 19.9, and Lemma 18.8 instead of Lemma 18.7.

Let $\hat{v} \in W^{\ell+1,p}(\widehat{T}_2)^3$. Let $I_{\widehat{T}}^{RT^k}$ be the local interpolation operators on $\widehat{T}$ defined by (19.4) and (19.5). If $q \in \mathbb{P}^{\ell}(T)^3 \subset \mathbb{RT}^k(T)$, then $I_T^{RT^k} q = q$.

We set $\mathfrak{Q}^{(\ell+1)} v := (Q^{(\ell+1)} v_1, Q^{(\ell+1)} v_2, Q^{(\ell+1)} v_3)^{\top} \in \mathbb{P}^{\ell}(T)^3$, where $Q^{(\ell+1)} v_j^s$ is defined by (1.11) for any j. We then obtain

$$\|I_T^{RT^k} v - v\|_{L^p(T)^3} \leq \|I_T^{RT^k} (v - \mathfrak{Q}^{(\ell+1)} v)\|_{L^p(T)^3} + \|\mathfrak{Q}^{(\ell+1)} v - v\|_{L^p(T)^3}. \tag{19.50}$$

The inequality (18.5) for the first term on the right-hand side of (19.50) yields

$$\|I_T^{RT^k} (v - \mathfrak{Q}^{(\ell+1)} v)\|_{L^p(T)^3} \leq c |\det(A_{\widetilde{T}})|^{\frac{1-p}{p}} \|\widetilde{A}\|_2 \left(\sum_{j=1}^{3} h_j^p \|\{I_{\widehat{T}}^{RT^k} (\hat{v} - \widehat{\mathfrak{Q}}^{(\ell+1)} \hat{v})\}_j\|_{L^p(\widehat{T})}^p \right)^{1/p}. \tag{19.51}$$

The component-wise stability (19.18) for $3d$ yields

$$\begin{aligned} &\sum_{j=1}^{3} h_j^p \|\{I_{\widehat{T}}^{RT^k} (\hat{v} - \widehat{\mathfrak{Q}}^{(\ell+1)} \hat{v})\}_j\|_{L^p(\widehat{T})}^p \\ &\quad \leq c \sum_{j=1}^{3} h_j^p \left(\|\hat{v}_j - \widehat{Q}^{(\ell+1)} \hat{v}_j\|_{W^{1,p}(\widehat{T})}^p + \sum_{k=1, k \neq j}^{3} \left\| \frac{\partial}{\partial \hat{x}_k} (\hat{v} - \widehat{\mathfrak{Q}}^{(\ell+1)} \hat{v})_k \right\|_{L^p(\widehat{T})}^p \right). \end{aligned} \tag{19.52}$$

The inequality (18.5) for the second term on the right-hand side of (19.50) yields

$$\|\mathfrak{Q}^{(\ell+1)} v - v\|_{L^p(T)^3} \leq c |\det(A_{\widetilde{T}})|^{\frac{1-p}{p}} \|\widetilde{A}\|_2 \left(\sum_{j=1}^{d} h_j^p \|\widehat{Q}^{(\ell+1)} \hat{v}_j - \hat{v}_j\|_{L^p(\widehat{T})}^p \right)^{1/p}. \tag{19.53}$$

From the Bramble–Hilbert-type lemma (Lemma 1.10) and (18.6), we have

$$\begin{aligned} \|\hat{v}_j - \widehat{Q}^{(\ell+1)} \hat{v}_j\|_{W^{1,p}(\widehat{T})}^p &= \|\hat{v}_j - \widehat{Q}^{(\ell+1)} \hat{v}_j\|_{L^p(\widehat{T})}^p + \sum_{k=1}^{d} \left\| \frac{\partial}{\partial \hat{x}_k} (\hat{v}_j - \widehat{Q}^{(\ell+1)} \hat{v}_j) \right\|_{L^p(\widehat{T})}^p \\ &\leq c \left(\sum_{|\gamma|=\ell+1} \|\partial_{\hat{x}}^{\gamma} \hat{v}_j\|_{L^p(\widehat{T})}^p + \sum_{k=1}^{3} \sum_{|\beta|=\ell} \left\| \partial_{\hat{x}}^{\beta} \frac{\partial \hat{v}_j}{\partial \hat{x}_k} \right\|_{L^p(\widehat{T})}^p \right) \\ &\leq c |\det(A_{\widetilde{T}})|^{p-1} h_j^{-p} \|\widetilde{A}^{-1}\|_2^p \left(\sum_{|\varepsilon|=\ell+1} h^{\varepsilon} \|\partial_r^{\varepsilon} v\|_{L^p(T)^3} \right)^p. \end{aligned} \tag{19.54}$$

From the Bramble–Hilbert-type lemma (Lemma 1.10) and (18.10), we have

$$
\begin{aligned}
\sum_{k=1,k\neq j}^{3} \left\| \frac{\partial}{\partial \hat{x}_k} (\hat{v}_k - \widehat{Q}^{(\ell+1)} \hat{v}_k)) \right\|_{L^p(\widehat{T})}^p &= \sum_{k=1,k\neq j}^{3} \left\| \frac{\partial \hat{v}_k}{\partial \hat{x}_k} - \widehat{Q}^{(\ell+1)} \left(\frac{\partial \hat{v}_k}{\partial \hat{x}_k} \right) \right\|_{L^p(\widehat{T})}^p \\
&\leq c \sum_{k=1,k\neq j}^{3} \sum_{|\beta|=\ell} \left\| \partial_{\hat{x}}^{\beta} \frac{\partial \hat{v}_k}{\partial \hat{x}_k} \right\|_{L^p(\widehat{T})}^p \\
&\leq c |\det(A_{\widetilde{T}})|^{p-1} \|\widetilde{A}^{-1}\|_2^p \sum_{k=1,k\neq j}^{3} \sum_{|\varepsilon|=\ell} h^{\varepsilon p} \left\| \partial_r^{\varepsilon} \frac{\partial v}{\partial r_k} \right\|_{L^p(T)^3}^p . \qquad (19.55)
\end{aligned}
$$

Gathering (19.51), (19.52), (19.54) and (19.55) together with (1.5) and (6.8b) yields

$$
\begin{aligned}
&\|I_T^{RT^k} (v - \mathfrak{Q}^{(\ell+1)} v)\|_{L^p(T)^3} \\
&\quad \leq c \frac{H_T}{h_T} \left(\sum_{|\varepsilon|=\ell+1} h^{\varepsilon} \|\partial_r^{\varepsilon} v\|_{L^p(T)^3} + \sum_{j=1}^{3} h_j \sum_{k=1,k\neq j}^{3} \sum_{|\varepsilon|=\ell} h^{\varepsilon} \left\| \partial_r^{\varepsilon} \frac{\partial v}{\partial r_k} \right\|_{L^p(T)^3} \right) \\
&\quad \leq c \frac{H_T}{h_T} \left(\sum_{|\varepsilon|=\ell+1} h^{\varepsilon} \|\partial_r^{\varepsilon} v\|_{L^p(T)^3} + h_T \sum_{k=1}^{3} \sum_{|\varepsilon|=\ell} h^{\varepsilon} \left\| \partial_r^{\varepsilon} \frac{\partial v}{\partial r_k} \right\|_{L^p(T)^3} \right). \qquad (19.56)
\end{aligned}
$$

From the Bramble–Hilbert-type lemma (Lemma 1.10) and (18.6),

$$
\begin{aligned}
\|\hat{v}_j - \widehat{Q}^{(\ell+1)} \hat{v}_j\|_{L^p(\widehat{T})}^p &\leq c \sum_{|\gamma|=\ell+1} \|\partial_{\hat{x}}^{\gamma} \hat{v}_j\|_{L^p(\widehat{T})}^p \\
&\leq c |\det(A_{\widetilde{T}})|^{p-1} h_j^{-p} \|\widetilde{A}^{-1}\|_2^p \left(\sum_{|\varepsilon|=\ell+1} h^{\varepsilon} \|\partial_r^{\varepsilon} v\|_{L^p(T)^3} \right)^p . \qquad (19.57)
\end{aligned}
$$

From (19.53) and (19.57) together with (6.8b), we obtain

$$
\|\mathfrak{Q}^{(\ell+1)} v - v\|_{L^p(T)^3} \leq c \frac{H_T}{h_T} \sum_{|\varepsilon|=\ell+1} h^{\varepsilon} \|\partial_r^{\varepsilon} v\|_{L^p(T)^3} . \qquad (19.58)
$$

Therefore, from (19.50) and (19.56), (19.58), we have (19.48).

Case in which Condition 10.1 is imposed. From the Bramble–Hilbert-type lemma (Lemma 1.10) and (18.7), we have

$$
\begin{aligned}
\|\hat{v}_j - \widehat{Q}^{(\ell+1)} \hat{v}_j\|_{W^{1,p}(\widehat{T})}^p &= \|\hat{v}_j - \widehat{Q}^{(\ell+1)} \hat{v}_j\|_{L^p(\widehat{T})}^p + \sum_{k=1}^{d} \left\| \frac{\partial}{\partial \hat{x}_k} (\hat{v}_j - \widehat{Q}^{(\ell+1)} \hat{v}_j) \right\|_{L^p(\widehat{T})}^p \\
&\leq c \left(\sum_{|\gamma|=\ell+1} \|\partial_{\hat{x}}^{\gamma} \hat{v}_j\|_{L^p(\widehat{T})}^p + \sum_{k=1}^{3} \sum_{|\beta|=\ell} \left\| \partial_{\hat{x}}^{\beta} \frac{\partial \hat{v}_j}{\partial \hat{x}_k} \right\|_{L^p(\widehat{T})}^p \right) \\
&\leq c |\det(A_{\widetilde{T}})|^{p-1} h_j^{-p} \|\widetilde{A}^{-1}\|_2^p \left(\sum_{|\varepsilon|=\ell+1} \mathscr{H}^{\varepsilon} \|\partial_{\tilde{x}}^{\varepsilon} (\Psi_{\widetilde{T}}^{-1} v)\|_{L^p(\Phi_T^{-1}(T))^3} \right)^p . \qquad (19.59)
\end{aligned}
$$

From the Bramble–Hilbert-type lemma (Lemma 1.10) and (18.10), we have

$$\begin{aligned}\sum_{k=1,k\neq j}^{3}\left\|\frac{\partial}{\partial \hat{x}_k}(\hat{v}_k-\widehat{Q}^{(\ell+1)}\hat{v}_k))\right\|^p_{L^p(\widehat{T})} &= \sum_{k=1,k\neq j}^{3}\left\|\frac{\partial \hat{v}_k}{\partial \hat{x}_k}-\widehat{Q}^{(\ell+1)}\left(\frac{\partial \hat{v}_k}{\partial \hat{x}_k}\right)\right\|^p_{L^p(\widehat{T})}\\
&\leq c\sum_{k=1,k\neq j}^{3}\sum_{|\beta|=\ell}\left\|\partial^{\beta}_{\hat{x}}\frac{\partial \hat{v}_k}{\partial \hat{x}_k}\right\|^p_{L^p(\widehat{T})}\\
&\leq c|\det(A_{\widetilde{T}})|^{p-1}\|\widetilde{A}^{-1}\|_2^p\sum_{k=1,k\neq j}^{3}\sum_{|\varepsilon|=\ell}\widetilde{\mathscr{H}^{\varepsilon}}\left\|\partial^{\varepsilon}_{\tilde{x}}\frac{\partial(\Psi_{\widetilde{T}}^{-1}v)}{\partial \tilde{r}_k^s}\right\|_{L^p(\Phi_T^{-1}(T))^3}. \qquad (19.60)\end{aligned}$$

Gathering (19.51), (19.52), (19.59) and (19.60) together with (1.5) and (6.8b) yields

$$\begin{aligned}\|I_T^{RT^k}(v-\mathfrak{Q}^{(\ell+1)}v)\|_{L^p(T)^3} &\leq c\frac{H_T}{h_T}\Bigg(\sum_{|\varepsilon|=\ell+1}\widetilde{\mathscr{H}^{\varepsilon}}\|\partial^{\varepsilon}_{\tilde{x}}(\Psi_{\widetilde{T}}^{-1}v)\|_{L^p(\Phi_T^{-1}(T))^3}\\
&\qquad+\sum_{j=1}^{3}h_j\sum_{k=1,k\neq j}^{3}\sum_{|\varepsilon|=\ell}\widetilde{\mathscr{H}^{\varepsilon}}\left\|\partial^{\varepsilon}_{\tilde{x}}\frac{\partial(\Psi_{\widetilde{T}}^{-1}v)}{\partial \tilde{r}_k}\right\|_{L^p(\Phi_T^{-1}(T))^3}\Bigg)\\
&\leq c\frac{H_T}{h_T}\Bigg(\sum_{|\varepsilon|=\ell+1}\widetilde{\mathscr{H}^{\varepsilon}}\|\partial^{\varepsilon}_{\tilde{x}}(\Psi_{\widetilde{T}}^{-1}v)\|_{L^p(\Phi_T^{-1}(T))^3}\\
&\qquad+h_T\sum_{k=1}^{3}\sum_{|\varepsilon|=\ell}\widetilde{\mathscr{H}^{\varepsilon}}\left\|\partial^{\varepsilon}_{\tilde{x}}\frac{\partial(\Psi_{\widetilde{T}}^{-1}v)}{\partial \tilde{r}_k}\right\|_{L^p(\Phi_T^{-1}(T))^3}\Bigg). \qquad (19.61)\end{aligned}$$

From the Bramble–Hilbert-type lemma (Lemma 1.10) and (18.7),

$$\begin{aligned}\|\hat{v}_j-\widehat{Q}^{(\ell+1)}\hat{v}_j\|^p_{L^p(\widehat{T})} &\leq c\sum_{|\gamma|=\ell+1}\|\partial^{\gamma}_{\hat{x}}\hat{v}_j\|^p_{L^p(\widehat{T})}\\
&\leq c|\det(A_{\widetilde{T}})|^{p-1}h_j^{-p}\|\widetilde{A}^{-1}\|_2^p\left(\sum_{|\varepsilon|=\ell+1}\widetilde{\mathscr{H}^{\varepsilon}}\|\partial^{\varepsilon}_{\tilde{x}}(\Psi_{\widetilde{T}}^{-1}v)\|_{L^p(\Phi_T^{-1}(T))^3}\right)^p.\end{aligned} \qquad (19.62)$$

From (19.53) and (19.62) together with (6.8b), we obtain

$$\|\mathfrak{Q}^{(\ell+1)}v-v\|_{L^p(T)^3}\leq c\frac{H_T}{h_T}\sum_{|\varepsilon|=\ell+1}\widetilde{\mathscr{H}^{\varepsilon}}\|\partial^{\varepsilon}_{\tilde{x}}(\Psi_{\widetilde{T}}^{-1}v)\|_{L^p(\Phi_T^{-1}(T))^3}. \qquad (19.63)$$

Therefore, from (19.50) and (19.61), (19.63), we have (19.49). □

19.6 Global RT Interpolation Error Estimates

We define a broken finite element space as

$$RT^k(\mathbb{T}_h):=\left\{v_h\in L^1(\Omega)^d;\ v_h|_T\in\mathbb{RT}^k(T)\ \forall T\in\mathbb{T}_h\right\}.$$

The corresponding (global) RT finite element space is defined as

$$V_h^{RT^k}:=\{v_h\in RT^k(\mathbb{T}_h);\ [\![v_h\cdot n]\!]_F=0,\ \forall F\in\mathcal{F}_h^i\}.$$

Lemma 19.16. It holds that

$$V_h^{RT^k} \subset H(\mathrm{div};\Omega).$$

Proof. Let $v_h \in V_h^{RT^k}$. Because its restriction to every $T \in \mathbb{T}_h$ is a polynomial, it is differentiable in the classical sense. Let us consider the function $w_h \in L^2(\Omega)$ defined on T by $w_h|_T = \mathrm{div}(v_h)|_T$. Let $\varphi \in C_0^\infty(\Omega)$. Then, using the Green formula yields

$$\begin{aligned}\int_\Omega w_h \varphi dx &= \sum_{T\in\mathbb{T}_h} \int_T w_h \varphi dx \\ &= -\sum_{T\in\mathbb{T}_h} \int_T (v_h)|_T \cdot \nabla\varphi dx + \sum_{F\in\mathcal{F}_h^i} \int_F [\![v_h \cdot n]\!]_F \varphi ds.\end{aligned}$$

Because $[\![v_h \cdot n]\!]_F = 0$,

$$\int_\Omega w_h \varphi dx = -\int_\Omega (v_h \cdot \nabla)\varphi dx.$$

Therefore, the distributional divergence of v_h is w_h. Because $w_h \in L^2(\Omega)$, $\mathrm{div}\, v_h \in L^2(\Omega)$. □

We define the global RT interpolation $I_h^{RT^k} : W^{1,1}(\Omega) \to V_h^{RT^k}$ as

$$(I_h^{RT^k} v)|_T = I_T^{RT^k}(v|_T) \quad \forall T \in \mathbb{T}_h, \quad \forall v \in W^{1,1}(\Omega).$$

Corollary 19.17 (de Rham complex)**.** The following diagram commutes:

$$\begin{array}{ccc} W^{1,1}(\Omega)^d & \xrightarrow{\nabla\cdot} & L^1(\Omega) \\ {\scriptstyle I_h^{RT^k}}\downarrow & & \downarrow{\scriptstyle \Pi_h^k} \\ V_h^{RT^k} & \xrightarrow[\nabla\cdot]{} & P_{dc,h}^k \end{array}$$

In other words, it holds that

$$\mathrm{div}(I_h^{RT^k} v) = \Pi_h^k(\mathrm{div}\, v) \quad \forall v \in W^{1,1}(\Omega)^d. \tag{19.64}$$

Proof. Combine Lemma 19.3. □

Corollary 19.18 (Stability)**.** Let $p \in [1,\infty)$. We impose Condition 6.2 with $h \leq 1$. Then,

$$\|I_h^{RT^k} v\|_{L^p(\Omega)^d} \leq c\|v\|_{W^{1,p}(\Omega)^d} \quad \forall v \in W^{1,p}(\Omega)^d.$$

Proof. Lemmata 19.12 and 19.13 yield

$$\|I_h^{RT^k} v\|_{L^p(\Omega)^d}^p = \sum_{T\in\mathbb{T}_h} \|I_T^{RT^k} v\|_{L^p(T)^d}^p \leq c \sum_{T\in\mathbb{T}_h} \|v\|_{W^{1,p}(T)^d}^p = c\|v\|_{W^{1,p}(\Omega)^d}^p,$$

which leads to the desired result. □

Corollary 19.19. Let $p \in [1,\infty)$. We impose Condition 6.2 with $h \leq 1$. Let ℓ be such that $0 \leq \ell \leq k$. Then, for any $v \in W^{\ell+1,p}(\Omega)^d$, if all mesh elements $T \in \mathbb{T}_h$ satisfy Condition 5.1 or Condition 5.2 with $T = \Phi_T(\widetilde{T})$ and $\widetilde{T} = \Phi_{\widetilde{T}}(\widehat{T})$, where $\widetilde{T} \in \mathfrak{T}^{(2)}$ or $\widetilde{T} \in \mathfrak{T}_1^{(3)}$,

$$\|I_h^{RT^k} v - v\|_{L^p(\Omega)^d} \leq c \sum_{T\in\mathbb{T}_h} \sum_{|\varepsilon|=\ell+1} h^\varepsilon \|\partial_r^\varepsilon v\|_{L^p(T)^d} + h \left(\sum_{T\in\mathbb{T}_h} \sum_{|\beta|=\ell} h^{\beta p} \|\partial_r^\beta \nabla\cdot v\|_{L^p(T)}^p \right)^{1/p}. \tag{19.65}$$

Furthermore, if Condition 10.1 is imposed,

$$\|I_h^{RT^k} v - v\|_{L^p(\Omega)^d} \leq c \sum_{T \in \mathbb{T}_h} \Bigg(\sum_{|\varepsilon| = \ell+1} \widetilde{\mathscr{H}^{\varepsilon}} \|\partial_{\tilde{x}}^{\varepsilon} (\Psi_{\widetilde{T}}^{-1} v)\|_{L^p(\Phi_T^{-1}(T))^d} + h_T \sum_{|\beta| = \ell} \widetilde{\mathscr{H}^{\beta}} \|\partial_{\tilde{x}}^{\beta} \nabla_{\tilde{x}} \cdot (\Psi_{\widetilde{T}}^{-1} v)\|_{L^p(\Phi_T^{-1}(T))} \Bigg). \tag{19.66}$$

Let $d = 3$. For any $v \in W^{\ell+1,p}(\Omega)^d$, if all mesh elements $T \in \mathbb{T}_h$ satisfy Condition 5.2 with $T = \Phi_T(\widetilde{T})$ and $\widetilde{T} = \Phi_{\widetilde{T}}(\widehat{T})$, where $\widetilde{T} \in \mathfrak{T}_2^{(3)}$,

$$\|I_h^{RT^k} v - v\|_{L^p(\Omega)^3} \leq ch \left(\sum_{T \in \mathbb{T}_h} \sum_{k=1}^{3} \sum_{|\varepsilon| = \ell} h^{\varepsilon p} \left\| \partial_r^{\varepsilon} \frac{\partial v}{\partial r_k} \right\|_{L^p(T)^3}^p \right)^{1/p}. \tag{19.67}$$

Furthermore, if Condition 10.1 is imposed,

$$\|I_h^{RT^k} v - v\|_{L^p(\Omega)^3} \leq c \sum_{T \in \mathbb{T}_h} \Bigg(\sum_{|\varepsilon| = \ell+1} \widetilde{\mathscr{H}^{\varepsilon}} \|\partial_{\tilde{x}}^{\varepsilon} (\Psi_{\widetilde{T}}^{-1} v)\|_{L^p(\Phi_T^{-1}(T))^3} + h_T \sum_{k=1}^{3} \sum_{|\varepsilon| = \ell} \widetilde{\mathscr{H}^{\varepsilon}} \left\| \partial_{\tilde{x}}^{\varepsilon} \frac{\partial (\Psi_{\widetilde{T}}^{-1} v)}{\partial \tilde{r}_k} \right\|_{L^p(\Phi_T^{-1}(T))^3} \Bigg). \tag{19.68}$$

Proof. This corollary is proved in the same argument as Corollary 14.2. □

20 Inverse Inequalities on Anisotropic Meshes

References

[1] Acosta, G., Apel, Th., Durán, R.G., Lombardi, A.L.: Error estimates for Raviart–Thomas interpolation of any order on anisotropic tetrahedra, Mathematics of Computation **80** No. 273, 141-163 (2010)

[2] G. Acosta, R.G. Durán: The maximum angle condition for mixed and nonconforming elements: Application to the Stokes equations, SIAM J. Numer. Anal. 37 (1999) 18-36. Zbl 0948.65115.

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

[4] Apel, Th, Dobrowolski, M.: Anisotropic interpolation with applications to the finite element method. Computing **47**, 277-293 (1992)

[5] Apel, T., Eckardt, L., Hauhner, C., Kempf, V.: The maximum angle condition on finite elements: useful or not? , PAMM (2021)

[6] Arnord, D.T., Brezzi, F.: Mixed and nonconforming finite element methods: implementation, postprocessing and error estimates. RAIRO Modélisation mathématique et analyse numérique **19**, 7-32 (1985)

[7] Babuška, I., Aziz, A.K.: On the angle condition in the finite element method. SIAM J. Numer. Anal. **13**, 214-226 (1976)

[8] Babuška, I., Suri, M.: The *p* and *h-p* versions of the finite elements method, basic principles and properties. SIAM Review. **36**, 578-632 (1994)

[9] Barnhill, R.E., Gregory, J.A.: Sard kernel theorems on triangular domains with applications to finite element error bounds. Numer. Math. **25**, 215-229 (1975/1976)

[10] Bernardi, C., Girault, V., Hecht, F., Raviart, P.-A., Rivière, B.: Mathematics and Finite Element Discretizations of Incompressible Navier–Stokes Flows. SIAM, (2024)

[11] Boffi, D., Brezzi, F., Demkowicz, L.F., Durén, R.G., Falk, R.S., Fortin, M.: Mixed Finite Elements, Compatibility Conditions, and Applications : Lectures Given at the C.I.M.E. Summer School, Italy, 2006. Lecture Notes in Mathematics **1939**, Springer, (2008)

[12] Boffi, D., Brezzi, F., Fortin, M.: Mixed Finite Element Methods and Applications. Springer Verlag, New York (2013)

[13] Brandts, J., Korotov, S., Křížek, M.: On the equivalence of regularity criteria for triangular and tetrahedral finite element partitions. Comput. Math, Appl. **55**, 2227-2233 (2008)

[14] Brenner, S.C., Scott, L.R.: The Mathematical Theory of Finite Element Methods, Third Edition. Springer Verlag, New York (2008)

[15] Chen, S., Shi, D., Zhao, Y.: Anisotropic interpolation and quasi-Wilson element for narrow quadrilateral meshes. IMA Journal of Numerical Analysis, **24**, 77-95 (2004)

[16] Cheng, S. -W., Dey, T. K., Edelsbrunner, H., Facello, M. A., Teng, S. -H.: Sliver Exudation. J. ACM, **47**, 883-904 (2000)

[17] Ciarlet, P. G.: The Finite Element Method for Elliptic problems. SIAM, New York (2002)

[18] Dekel, S., Leviatan, D.: The Bramble–Hilbert Lemma for Convex Domains, SIAM Journal on Mathematical Analysis **35** No. 5, 1203-1212 (2004)

[19] Dupont, T., Scott, R.: Polynomial approximation of functions in Sobolev spaces. Math. Comp. **34**, 441-463 (1980)

[20] Ern, A., Guermond, J.L.: Theory and Practice of Finite Elements. Springer Verlag, New York (2004)

[21] Ern, A., Guermond, J. L.: Finite Elements I: Approximation and Interpolation. Springer-Verlag, New York (2021)

[22] Ern, A., Guermond, J. L.: Finite elements II: Galerkin Approximation, Elliptic and Mixed PDEs. Springer-Verlag, New York (2021)

[23] Ern, A. and Guermond, J. L.: Finite elements III: First-Order and Time-Dependent PDEs. Springer-Verlag, New York (2021)

[24] Gellert, W., Gottwald, S., Hellwich, M., Kästner, H., Küstner, H.: The VNR Concise Encyclopedia of Mathematics, Springer, (1975)

[25] Golub, G. H., Loan, C. F. V.: Matrix Computations 3rd edition. The Johas Hopkins University Press, (1996)

[26] Gregory, J.A.: Error bounds for linear interpolation on triangles: in: J.R. Whiteman (Ed.), Proc. MAFELAP II, Academic Press, London, (1976)

[27] Girault, V., Raviart, P.A.: Finite Element Methods for Navier-Stokes Equations. Springer-Verlag, (1986)

[28] Hitotsumatsu, S., Kuroyanagi, K.: Geometry by barycentric coordinates, Gendaishugakusha (2014) (in Japanese)

[29] Ishizaka H. Anisotropic interpolation error analysis using a new geometric parameter and its applications. Ehime University, Ph. D. thesis (2022)

[30] Ishizaka, H.: Anisotropic Raviart–Thomas interpolation error estimates using a new geometric parameter. Calcolo 59 (4), (2022)

[31] Ishizaka, H.: Anisotropic weakly over-penalised symmetric interior penalty method for the Stokes equation. Journal of Scientific Computing 100, (2024)

[32] Ishizaka, H.: Morley finite element analysis for fourth-order elliptic equations under a semi-regular mesh condition. Applications of Mathematics **69** (6), 769–805 (2024)

[33] Ishizaka, H.: Hybrid weakly over-penalised symmetric interior penalty method on anisotropic meshes. Calcolo **61**, (2024)

[34] Ishizaka, H.: Anisotropic modified Crouzeix-Raviart finite element method for the stationary Navier-Stokes equation. Numeriche Mathmatik, (2025)

[35] Ishizaka, H.: Nitsche method under a semi-regular mesh condition. preprint (2025) `https://arxiv.org/abs/2501.06824`

[36] Ishizaka, H., Kobayashi, K., Suzuki, R., Tsuchiya, T.: A new geometric condition equivalent to the maximum angle condition for tetrahedrons. Comput. Math. Appl. 99, 323–328 (2021)

[37] Ishizaka, H., Kobayashi, K., Tsuchiya, T.: General theory of interpolation error estimates on anisotropic meshes. Jpn. J. Ind. Appl. Math. 38 (1), 163–191 (2021)

[38] Ishizaka, H., Kobayashi, K., Tsuchiya, T.: Crouzeix–Raviart and Raviart–Thomas finite element error analysis on anisotropic meshes violating the maximum-angle condition. Jpn. J. Ind. Appl. Math. 38 (2), 645-675 (2021)

[39] Ishizaka, H., Kobayashi, K., Tsuchiya, T.: Anisotropic interpolation error estimates using a new geometric parameter. Jpn. J. Ind. Appl. Math. 40 (1), 475–512 (2023)

[40] Jamet, P.: Estimation de I'erreur pour des éléments finis droits presque dégénérés. RAIRO Anal. Numér **10**, 43-60 (1976)

[41] Křížek, M.: On semiregular families of triangulations and linear interpolation. Appl. Math. Praha **36**, 223-232 (1991)

[42] Křížek, M.: On the maximum angle condition for linear tetrahedral elements. SIAM J. Numer. Anal. **29**, 513-520 (1992)

[43] Ladyženskaja, O. A., Solonnikov, V. A., Ural'ceva, N. N.: Linear and Quasi-linear Equations of Parabolic Type. Translations of Mathematical Monographs **23**, AMS, Providence, (1968)

[44] P. Lascaux, P. Lesaint: Some nonconforming finite elements for the plate bending problem. RAIRO Anal. Numer. R-1 (1975) L9-53. Zbl 0319.73042.

[45] Mario, B.: A note on the Poincaré inequality for convex domains. Z. Anal. ihre. Anwend., **22**, 751-756 (2003)

[46] L.S.D. Morley: The Triangular Equilibrium Element in the Solution of Plate Bending Problems. Aero Quart 19 (1968) 149-169. DOI https://doi.org/10.1017/S0001925900004546.

[47] Payne, L.E., Weinberger, H.F.: An optimal Poincaré-inequality for convex domains, Arch. Rational Mech. Anal. **5**, 286-292 (1960)

[48] Synge, J.L.: The Hypercircle in Mathematical Physics. Cambridge Univ. Press, Cambridge (1957)

[49] Todhunter, I.: Spherical Trigonometry: For the Use of Colleges and Schools (5th ed.), MacMillan, (1886)

[50] Verfürth, R.: A note on polynomial approximation in Sobolev spaces, Math. Modelling and Numer. Anal. **33**, 715-719 (1999)

[51] Ming, W., Xu, J.: The Morley element for fourth order elliptic equations in any dimensions. Numer. Math. **103**, 155-169 (2006)

[52] Zlámak, M.: On the finite element method, Numer. Math. **12**, 394-409 (1968)