SHARP ANISOTROPIC INTERPOLATION ERROR ESTIMATES
FOR RECTANGULAR RAVIART-THOMAS ELEMENTS

MARTIN STYNES

Abstract. Anisotropic $L_p$-norm error estimates are derived for the standard rectangular Raviart-Thomas elements $RT[k](\tilde{K})$ in $\mathbb{R}^d$ for $p \in [1, \infty)$, $k \geq 0$ and $d \geq 2$. Here $\tilde{K}$ is an affine image of an axi-parallel parallelotope $K$. The proofs are based on a variant of the classical Poincaré inequality. The estimates derived make full use of the asymmetric nature of the vector space components of $RT[k](\tilde{K})$; a Shishkin mesh example demonstrates their superiority over previous estimates.

1. Introduction

In this paper we develop anisotropic interpolation error estimates for the standard rectangular Raviart-Thomas elements $RT[k]$ in $\mathbb{R}^d$ for $k \geq 0$ and $d \geq 2$. See, e.g., [5,7] for a description of these spaces and their use in mixed finite element methods.

On axi-parallel parallelotopes $K$ in $\mathbb{R}^d$, and using the standard interpolation operator $\Pi_k : H^1(K) \rightarrow RT[k](K)$ that will be defined in Section 3, the isotropic estimate

\[(1.1) \quad \|u - \Pi_k u\|_{L^2(K)} \leq Ch^{k+1}\|u\|_{H^{k+1}(K)}\]

is well known [5] Chapter III; here each component of $u$ lies in the Sobolev space $H^{k+1}(K)$ and $h$ is the diameter of $K$. Example 3.3 below will show that bounds of this type are too imprecise for sharp analyses on the highly anisotropic meshes that are used to handle singularly perturbed problems whose solutions exhibit boundary layers. For such problems, instead of (1.1), one needs anisotropic interpolation error estimates that are stated explicitly in terms of the width of the parallelotope in different directions. The only such estimate in the literature seems to be [2, Remark 4.1]: for $d = 2$ and $k = 0$ one has

\[(1.2) \quad \|(u - \Pi_0 u)_1\|_{L^2(K)} \leq C[h_1\|(u_1)_x\|_{L^2(K)} + h_2\|(u_1)_y\|_{L^2(K)}]\]

with a similar bound for $\|(u - \Pi_0 u)_2\|_{L^2(K)}$, where $K$ is the rectangle $[-h_1, h_1] \times [-h_2, h_2]$ and $(u - \Pi_0 u)_i$ denotes the $i^{th}$ component of $u - \Pi_0 u$.

We shall generalize (1.2) to all $d \geq 2$ and $k \geq 0$ in the $L_p(K)$ norm for $p \in [1, \infty]$. Furthermore, in this generalization we shall remedy a weakness of (1.2): although $RT[0](K) = (\text{span}\{1, x\}, \text{span}\{1, y\})$, the bound (1.2) does not seem to use the non-constant components of $RT[0](K)$. In contrast to this, our estimates exploit fully the asymmetric nature of the vector space bases in the components of $RT[k](K)$ for each $k \geq 0$. Our results are also extended to affine images of axi-parallel $K$. 

Received by the editor August 12, 2012 and, in revised form, April 16, 2013.

2010 Mathematics Subject Classification. Primary 65N30; Secondary 65N12, 65N15.

©2014 American Mathematical Society

Reverts to public domain 28 years from publication

License or copyright restrictions may apply to redistribution; see https://www.ams.org/journal-terms-of-use

2675
The proofs that we give are self-contained apart from invoking some well-known results from functional and finite element analysis.

Finally, we mention that anisotropic interpolation error bounds for triangular Raviart-Thomas elements are proved in [125] and their references.

Notation. Throughout the paper we take $d \geq 2$ and $k \geq 0$. Boldface font is used for vector-valued functions. We often use $\mathbf{x} = (x_1, x_2, \ldots, x_d)$ to denote a generic point in $\mathbb{R}^d$. The transpose of a vector $\mathbf{v}$ is denoted by $\mathbf{v}^T$. For any suitable domain $\Omega$, the norm in $L_p(\Omega)$ is denoted by $\| \cdot \|_{L_p(\Omega)}$, and standard notation is used for Sobolev spaces, e.g., $H^1(\Omega), W^{1,p}(\Omega)$. If $\mathbf{v} = (v_1, \ldots, v_d)$ is a vector-valued function defined on $\Omega$ with each of its components in the Sobolev space $W^{m,p}(\Omega)$, we write $\mathbf{v} \in W^{m,p}(\Omega)$ and define

$$
\| \mathbf{v} \|_{W^{m,p}(\Omega)} = \left\{ \begin{array}{ll}
\left[ \sum_{i=1}^d \| v_i \|_{W^{m,p}(\Omega)}^p \right]^{1/p} & \text{for } 1 \leq p < \infty, \\
\max_{i=1,\ldots,d} \| v_i \|_{W^{m,\infty}(\Omega)} & \text{for } p = \infty.
\end{array} \right.
$$

Finally, $C$ denotes a generic constant that depends only on one or more of $d, k$ and $p$; it can take different values in different places.

2. **POINCARÉ-TYPE INEQUALITIES**

The interpolation error analysis of the paper is based on a variant of the classical Poincaré inequality that is related to the variant of this inequality used in [2].

Let $\hat{K} = \prod_{i=1}^d [-1, 1] \subset \mathbb{R}^d$ be our reference element. The classical Poincaré inequality (sometimes called the Friedrichs inequality) is the following:

**Lemma 2.1.** Let $p \in [1, \infty]$. Let $g \in W^{1,p}(\hat{K})$ satisfy $\int_{\hat{K}} \mathbf{g} = 0$. Then there exists a constant $C = C(p)$, which is independent of $g$, such that

$$
\| g \|_{L_p(\hat{K})} \leq C \| \nabla g \|_{L_p(\hat{K})}.
$$

Let $\hat{F}$ denote the set of $(d-1)$-dimensional faces of $\hat{K}$, viz.,

$$
\hat{F} = \{ \mathbf{x} = (x_1, x_2, \ldots, x_d) \in \hat{K} : x_i = t \} \text{ for } i = 1, \ldots, d, \text{ where } |t| = 1.
$$

Note that $g \in W^{1,p}(\hat{K})$ implies that its trace (which we also write as $g$) lies in $H^{1/2}(\hat{f})$ for each $\hat{f} \in \hat{F}$, so $\int_{\hat{f}} g$ is well defined. In [2] Lemma 2.2, Acosta and Durán establish a variant of Lemma 2.1

**Lemma 2.2.** Let $p \in [1, \infty]$. Let $g \in W^{1,p}(\hat{K})$ satisfy $\int_{\hat{f}} g = 0$ for some face $\hat{f} \in \hat{F}$. Then there exists a constant $C = C(p)$, which is independent of $g$, such that

$$
\| g \|_{L_p(\hat{K})} \leq C \| \nabla g \|_{L_p(\hat{K})}.
$$

We shall make use of a generalisation of both of these lemmas. Let $m_1, \ldots, m_d$ be non-negative integers. Define $\hat{w} : \hat{K} \mapsto \mathbb{R}$ by

$$
\hat{w}(\mathbf{x}) = \prod_{j=1}^d (1 - x_j^2)^{m_j}.
$$

Let $\hat{f} \in \hat{F}$ be some face of $\hat{K}$ with $x_j^2 = 1$ for all $\mathbf{x} \in \hat{f}$. Define $\hat{w}_f : \hat{f} \mapsto \mathbb{R}$ by

$$
\hat{w}_f(x_1, x_2, \ldots, x_{i-1}, x_{i+1}, \ldots, x_d) = \prod_{j=1}^d (1 - x_j^2)^{m_j}.
$$
Lemma 2.3. Let $p \in [1, \infty]$. Let $g \in W^{1,p}(\hat{K})$ satisfy $\int_{\hat{K}} \hat{w}g = 0$ or $\int_{\hat{f}} \hat{w}_jg = 0$ for some face $\hat{f} \in \hat{\mathcal{F}}$. Then there exists a constant $C = C(p)$, which is independent of $g$, such that
\[
\|g\|_{L^p(\hat{K})} \leq C\|\nabla g\|_{L^p(\hat{K})}.
\]

Proof. The argument imitates the standard proof of Lemma 2.1 which is also used in [2]. We give the proof under the hypothesis $\int_{\hat{f}} \hat{w}_jg = 0$; the case $\int_{\hat{K}} \hat{w}g = 0$ is similar. Suppose the desired inequality is false. Then there exists a sequence \{\{g_j\}\} $\subset W^{1,p}(\hat{K})$ such that $\int_{\hat{f}} \hat{w}_jg_j = 0$, $\|g_j\|_{L^p(\hat{K})} = 1$, and $\|\nabla g_j\|_{L^p(\hat{K})} < 1/j$ for $j = 1, 2, \ldots$. Now $W^{1,p}(\hat{K})$ can be compactly imbedded into $L^p(\hat{K})$ by the Rellich-Kondracov theorem [3], so a subsequence of \{\{g_j\}\} (which for convenience we also call \{\{g_j\}\}) converges in $\| \cdot \|_{p, \hat{K}}$ to some $\hat{g} \in L^2(\hat{K})$. One must then have $\|\hat{g}\|_{L^p(\hat{K})} = 1$ and $\int_{\hat{f}} \hat{w}_j\hat{g} = 0$. Furthermore, if $\hat{g}$ is any $C^\infty$ function defined on $\hat{K}$ that has compact support in the interior of $\hat{K}$, then
\[
\int_{\hat{K}} \hat{g}\nabla \phi = \lim_{j \to \infty} \int_{\hat{K}} g_j \nabla \phi = -\lim_{j \to \infty} \int_{\hat{K}} \phi \nabla g_j = 0,
\]
so $\hat{g} \in W^{1,p}(\hat{K})$ and $\|\nabla \hat{g}\|_{L^p(\hat{K})} = 0$. Hence $\hat{g}$ is constant on $\hat{K}$ and, consequently, its trace is constant on $\hat{f}$; but $\int_{\hat{f}} \hat{w}_j\hat{g} = 0$ and $\hat{w}_j > 0$ almost everywhere on $\hat{f}$ then imply that the constant value of $\hat{g}$ is zero, which contradicts the earlier assertion that $\|\hat{g}\|_{L^p(\hat{K})} = 1$ and thereby completes the proof. \qed

3. Interpolation Error Bounds for Axi-Parallel $K$

The interpolation error estimates of this section are derived on the generic axi-parallel $d$-dimensional mesh parallelepode
\[
K = \prod_{i=1}^{d} [-h_i, h_i],
\]
where the $h_i$ satisfy $0 < h_i \leq 1$ for all $i$. The mapping $T : K \mapsto \mathbb{R}^d$ defined by $(Tx)_i = x_i/h_i \forall i$ is a bijection from $K$ to $\hat{K}$. For each face $\hat{f} \in \hat{\mathcal{F}}$, it is clear that $T^{-1} \hat{f}$ is a $(d-1)$-dimensional face of $K$. Let $\mathcal{F}$ denote the set of $(d-1)$-dimensional faces of $K$.

Following standard finite element practice, we shall work on the reference parallelepode $\hat{K}$ and then transform our results to $K$.

For non-negative integers $k_1, \ldots, k_d$ and $\Omega = \hat{K}$ or $\hat{\mathcal{F}}$ for some $\hat{F} \in \hat{\mathcal{F}}$, let $Q_{k_1,k_2,\ldots,k_d}(\Omega)$ denote the set of polynomials in $d$ variables defined on $\Omega$ where, for each $i$, the degree of the variable $x_i$ is at most $k_i$. If $k_1 = k_2 = \cdots = k_d = k$, we denote this set by $Q_k(\Omega)$.

For each $k \geq 0$, the rectangular Raviart-Thomas space $RT_{[k]}(\hat{K})$ (see, e.g., [5]) is defined by
\[
RT_{[k]}(\hat{K}) = \left[ Q_k(\hat{K}) \right]^d + xQ_k(\hat{K}).
\]
One can also write this as
\[
RT_{[k]}(\hat{K}) = Q_{k+1,k,k,\ldots,k}(\hat{K}) \times Q_{k,k+1,k,k,\ldots,k}(\hat{K}) \times \cdots \times Q_{k,k,\ldots,k,k+1}(\hat{K}) \times Q_{k,k,\ldots,k,k+1}(\hat{K}).
\]
Thus, for example, in the case $d = 2$ one has $RT_{[0]}(\hat{K}) = (\text{span } \{1, x_1\}, \text{span } \{1, x_2\})$
and $RT_{[1]}(\hat{K}) = (\text{span } \{1, x_1, x_2, x_1 x_2, x_1^2, x_2^2\}, \text{span } \{1, x_1, x_2, x_1 x_2, x_2^2, x_1 x_2^2\})$.

The asymmetry between the $d$ components of $RT_{[k]}(\hat{K})$ will be exploited fully in our error bounds.

For each $\hat{u} \in [H^1(\hat{K})]^d$ the Raviart-Thomas interpolant $\Pi_k \hat{u} \in RT_{[k]}(\hat{K})$ of order $k \geq 0$ is defined by

$$\int_{\hat{K}} (\hat{\Pi}_k \hat{u} \cdot n) q_k = \int_{\hat{f}} (\hat{u} \cdot n) q_k \quad \forall q_k \in Q_k(\hat{f}), \forall \hat{f} \in \hat{F},$$

where $n$ is a unit normal to $\hat{f}$, and

$$\int_{\hat{K}} \hat{\Pi}_k \hat{u} \cdot q = \int_{\hat{K}} \hat{u} \cdot q \quad \forall q \in \Psi_k(\hat{K}) \text{ if } k > 0,$$

where $\Psi_k(\hat{K}) := Q_{k-1,k,k,\ldots,k}(\hat{K}) \times Q_{k,k-1,k,k,\ldots,k}(\hat{K}) \times \cdots \times Q_{k,k,\ldots,k,k-1}(\hat{K})$ for each $k > 0$. The Raviart-Thomas space $RT_{[k]}(\hat{K})$ and the interpolant $\Pi_k u \in RT_{[k]}(K)$, where $u = \hat{u} \circ T$, are then defined via the transformation $T : K \mapsto \hat{K}$; see [5].

Existence and uniqueness of $\hat{\Pi}_k \hat{u}$ is well known when $d = 2$ or 3; the case of general $d \geq 2$ follows from [9], or for a direct elementary argument adapt the proof for $d$-simplexes given in [7].

Note that (3.2) implies that for each $\ell \in \{1, \ldots, d\}$, the component $(\hat{\Pi}_k \hat{u})_{\ell}$ depends only on $\hat{u}_\ell$, unlike the situation when triangular Raviart-Thomas elements are employed; cf. [1].

The next result presents an error bound for each component of $\hat{u} - \hat{\Pi}_k \hat{u}$ that will transform later to yield an anisotropic error bound on $K$ in Theorem 3.2.

**Theorem 3.1.** Let $p \in [1, \infty]$. Let $\hat{u} = (\hat{u}_1, \ldots, \hat{u}_d) \in (W^{d+k+2,p}(K))^d$. Then for $\ell = 1, \ldots, d$ there exists a constant $C = C(d, k, p)$, which is independent of $\hat{u}$, such that

$$\| (\hat{u} - \hat{\Pi}_k \hat{u})_{\ell} \|_{L_p(K)} \leq C \left[ \sum_{i=1}^{d} \sum_{r \neq \ell}^{d} \sum_{m_{r}=0}^{k} \left\| \frac{\partial}{\partial x_i} \left( \prod_{m_{r}=0}^{m_{r}-1} \frac{\partial}{\partial x_{\ell}} \frac{m_{r}+1+\cdots+m_{r}+m_{r}+\cdots+m_{r}+\cdots+m_{r}+1}{\partial x_{r} \partial x_{\ell}} \right) \right\|_{L_p(K)} 
+ \sum_{j=1}^{d} \sum_{r \neq j}^{d} \sum_{m_{r}=0}^{k} \left\| \frac{\partial}{\partial x_j} \left( \prod_{m_{r}=0}^{m_{r}-1} \frac{\partial}{\partial x_{r}} \frac{m_{r}+1+\cdots+m_{r}+m_{r}+\cdots+m_{r}+\cdots+m_{r}+1}{\partial x_{j} \partial x_{r}} \right) \right\|_{L_p(K)} \right].$$

**Proof.** We shall bound $\| (\hat{u} - \hat{\Pi}_k \hat{u})_1 \|_{L_p(K)}$; the other components can be handled similarly. Set $z = (\hat{u} - \hat{\Pi}_k \hat{u})_1$. Specify two $(d - 1)$-dimensional faces $\hat{F}^+$ and $\hat{F}^-$ of $\hat{K}$ that are associated with $z$ through (3.2a) by setting

$$\hat{F}^+ = \{ x \in \hat{K} : x_1 = 1 \}, \quad \hat{F}^- = \{ x \in \hat{K} : x_1 = -1 \}.$$
For $m_2, m_3, \ldots, m_d \in \{0, \ldots, k\}$, integrations by parts and (3.2a) yield

$$
\int_{\hat{F}^\pm} \frac{\partial^{m_2 + \cdots + m_d} z}{\partial x_2^{m_2} \cdots \partial x_d^{m_d}} \frac{d}{dx_j} \prod_{j=2}^d (1 - x_j^2)^{m_j} = (-1)^{m_2 + \cdots + m_d} \int_{\hat{F}^\pm} \frac{d}{dx_j} \prod_{j=2}^d (1 - x_j^2)^{m_j}
$$

(3.4)

where $\hat{F}^\pm$ means that we take either $\hat{F}^+$ or $\hat{F}^-$. Consequently, Lemma 2.3 gives

$$
\left\| \frac{\partial^{m_2 + \cdots + m_d} z}{\partial x_2^{m_2} \cdots \partial x_d^{m_d}} \right\|_{L_p(K)} \leq C \sum_{i=1}^d \left\| \frac{\partial}{\partial x_i} \left( \frac{\partial^{m_2 + \cdots + m_d} z}{\partial x_2^{m_2} \cdots \partial x_d^{m_d}} \right) \right\|_{L_p(K)}
$$

(3.5)

for $m_2, m_3, \ldots, m_d \in \{0, \ldots, k\}$.

Next, for $m_2, m_3, \ldots, m_d \in \{0, \ldots, k\}$, a simple integration gives

$$
\int_{K} \frac{\partial^{1+m_2 + \cdots + m_d} z}{\partial x_1^{m_2} \cdots \partial x_d^{m_d}} \frac{d}{dx_j} \prod_{j=2}^d (1 - x_j^2)^{m_j} = \left( \int_{\hat{F}^+} - \int_{\hat{F}^-} \right) z \frac{\partial^{m_2 + \cdots + m_d} z}{\partial x_2^{m_2} \cdots \partial x_d^{m_d}} \prod_{j=2}^d (1 - x_j^2)^{m_j} = 0
$$

by (3.4). Hence Lemma 2.3 implies that

$$
\left\| \frac{\partial^{1+m_2 + \cdots + m_d} z}{\partial x_1^{m_2} \cdots \partial x_d^{m_d}} \right\|_{L_p(K)} \leq C \sum_{i=1}^d \left\| \frac{\partial}{\partial x_i} \left( \frac{\partial^{1+m_2 + \cdots + m_d} z}{\partial x_1^{m_2} \cdots \partial x_d^{m_d}} \right) \right\|_{L_p(K)}
$$

(3.6)

for $m_2, m_3, \ldots, m_d \in \{0, \ldots, k\}$.

There remains the case where more than one derivative of $z$ with respect to $x_1$ is present. For $m_1 = 2, 3, \ldots, k + 1$ and $m_2, m_3, \ldots, m_d \in \{0, \ldots, k\}$, integrations by parts yield

$$
\int_{K} \frac{\partial^{m_1+m_2 + \cdots + m_d} z}{\partial x_1^{m_1} \partial x_2^{m_2} \cdots \partial x_d^{m_d}} (1 - x_1^2)^{m_1-1} \prod_{j=2}^d (1 - x_j^2)^{m_j} = (-1)^{m_2 + \cdots + m_d} \int_{K} \frac{d}{dx_1^{m_1}} (1 - x_1^2)^{m_1-1} \prod_{j=2}^d \frac{d}{dx_j^{m_j}} \left( (1 - x_j^2)^{m_j} \right)
$$

$$
= (-1)^{m_2 + \cdots + m_d} \int_{K} z \frac{d}{dx_1^{m_1}} (1 - x_1^2)^{m_1-1} \prod_{j=2}^d \frac{d}{dx_j^{m_j}} \left( (1 - x_j^2)^{m_j} \right)
$$

(3.7)

by (3.2b), where in the penultimate step we integrated by parts $m_1$ times and invoked (3.2a) to eliminate the boundary terms in the final integration by parts. Consequently, Lemma 2.3 gives

$$
\left\| \frac{\partial^{m_1+m_2 + \cdots + m_d} z}{\partial x_1^{m_1} \partial x_2^{m_2} \cdots \partial x_d^{m_d}} \right\|_{L_p(K)} \leq C \sum_{i=1}^d \left\| \frac{\partial}{\partial x_i} \left( \frac{\partial^{m_1+m_2 + \cdots + m_d} z}{\partial x_1^{m_1} \partial x_2^{m_2} \cdots \partial x_d^{m_d}} \right) \right\|_{L_p(K)}
$$

(3.8)

for $m_1 = 2, 3, \ldots, k + 1$ and $m_2, m_3, \ldots, m_d \in \{0, \ldots, k\}$.

We can summarize (3.5)–(3.7) as

$$
\left\| \frac{\partial^{m_1+m_2 + \cdots + m_d} z}{\partial x_1^{m_1} \partial x_2^{m_2} \cdots \partial x_d^{m_d}} \right\|_{L_p(K)} \leq C \sum_{i=1}^d \left\| \frac{\partial}{\partial x_i} \left( \frac{\partial^{m_1+m_2 + \cdots + m_d} z}{\partial x_1^{m_1} \partial x_2^{m_2} \cdots \partial x_d^{m_d}} \right) \right\|_{L_p(K)}
$$

for $m_1 = 0, \ldots, k + 1$ and $m_2, m_3, \ldots, m_d \in \{0, \ldots, k\}$. 
Now (3.8) is applied iteratively as many times as possible, i.e., until the largest permissible values of the \( m_i \) in (3.8) are reached:

\[
\| (\hat{u} - \Pi_k \hat{u})_1 \|_{L_p(K)} = \| z \|_{L_p(K)}
\]

\[
\leq C \sum_{i=1}^{d} \left\| \frac{\partial z}{\partial x_i} \right\|_{L_p(K)}
\]

\[
\leq C \sum_{j=1}^{d} \sum_{i=1}^{d} \left\| \frac{\partial^2 z}{\partial x_j \partial x_i} \right\|_{L_p(K)}
\]

\[
\leq \ldots
\]

\[
\leq C \left[ \sum_{j=1}^{d} \sum_{r=1}^{d} \sum_{m_r=0}^{k} \left\| \frac{\partial^{m_1 + \cdots + m_{j-1} + k+1 + m_{j+1} + \cdots + m_d} z}{\partial x_1^{m_1} \cdots \partial x_{j-1}^{m_{j-1}} \partial x_j^{k+1} \partial x_{j+1}^{m_{j+1}} \cdots \partial x_d^{m_d}} \right\|_{L_p(K)} \right]
\]

\[
\leq C \left[ \sum_{j=2}^{d} \sum_{r=1}^{d} \sum_{m_r=0}^{k} \left\| \frac{\partial^{m_1 + \cdots + m_{j-1} + k+1 + m_{j+1} + \cdots + m_d} z}{\partial x_1^{m_1} \cdots \partial x_{j-1}^{m_{j-1}} \partial x_j^{k+1} \partial x_{j+1}^{m_{j+1}} \cdots \partial x_d^{m_d}} \right\|_{L_p(K)} \right]
\]

where in the final inequality we used the exceptional upper limit on the number of derivatives with respect to \( x_1 \). The orders of derivatives of \( z \) that appear here annihilate \( (\Pi_k \hat{u})_1 \in Q_{k+1,k,\ldots,k} \), leaving the desired bound (3.3) with \( \ell = 1 \). \( \square \)

Now using standard finite element techniques to transform Theorem 3.1 via the mapping \( T \), with \( \hat{u} = T u \), yields the desired anisotropic error bound:

**Theorem 3.2.** Let \( p \in [1, \infty] \). Let \( u = (u_1, \ldots, u_d) \in (W^{d+2,p}(K))^d \), where \( K \) is defined in (3.1). Then for \( \ell = 1, \ldots, d \) there exists a constant \( C = C(d, k, p) \), which is independent of \( u \) and of \( h_i \), such that

\[
\| (u - \Pi_k u) \|_{L_p(K)} \leq C \left[ \sum_{i=1}^{d} h_i^{k+1} \sum_{r=1}^{d} \sum_{m_r=0}^{k} h_1^{m_1} \cdots h_{r-1}^{m_{r-1}} h_r^{m_r+1} \cdots h_d^{m_d} \right.
\]

\[
\times \left\| \frac{\partial}{\partial x_i} \left( \frac{\partial^{m_1 + \cdots + m_{r-1} + k+1 + m_{r+1} + \cdots + m_d} u_\ell}{\partial x_1^{m_1} \cdots \partial x_{r-1}^{m_{r-1}} \partial x_r^{k+1} \partial x_{r+1}^{m_{r+1}} \cdots \partial x_d^{m_d}} \right) \right\|_{L_p(K)}
\]

\[
\left. + \sum_{j=1}^{d} h_j^{k+1} \sum_{r=1}^{d} \sum_{m_r=0}^{k} h_1^{m_1} \cdots h_{j-1}^{m_{j-1}} h_j^{m_j+1} \cdots h_d^{m_d} \right]
\]

\[
\times \left\| \frac{\partial^{m_1 + \cdots + m_{j-1} + k+1 + m_{j+1} + \cdots + m_d} u_\ell}{\partial x_1^{m_1} \cdots \partial x_{j-1}^{m_{j-1}} \partial x_j^{k+1} \partial x_{j+1}^{m_{j+1}} \cdots \partial x_d^{m_d}} \right\|_{L_p(K)} \right] \quad \text{(3.10)}
\]
The sharp error bounds for each component of \( u - \Pi_{\ell} u \) given in Theorem 3.2 are suited to anisotropic elements \( K \) where the \( h_{\ell} \) can differ greatly in relative magnitude.

### 3.1. Shishkin mesh example.

To illustrate the superiority of Theorem 3.2 over isotropic error bounds and even over the anisotropic error bound of (1.2), we now consider the interpolation error when the simplest Raviart-Thomas elements \( RT_0(K) \) are used in two dimensions to approximate the gradient of the solution of a singularly perturbed reaction-diffusion problem on a Shishkin mesh.

When \( d = 2 \) and \( k = 0 \), Theorem 3.2 yields

\[
\|(u - \Pi_{0} u)_{j}\|_{L_{2}(K)} \leq C \left[ h_{j}^{2} \left\| \frac{\partial^{2} u_{j}}{\partial x_{j}^{2}} \right\|_{L_{2}(K)} + h_{j} h_{j'} \left\| \frac{\partial^{2} u_{j}}{\partial x_{j} \partial x_{j'}} \right\|_{L_{2}(K)} \right] + h_{j'} \left\| \frac{\partial u_{j}}{\partial x_{j'}} \right\|_{L_{2}(K)}
\]

for \( j = 1, 2 \), where \( j' = 2/j \).

Note that the bound (3.11) makes full use of the vector spaces in \( RT_{0}(K) = (\text{span } \{1, x_1\}, \text{span } \{1, x_2\}) \). In contrast, the anisotropic bound of [2, Remark 4.1], which we met already in (1.2), is

\[
\|(u - \Pi_{0} u)_{j}\|_{L_{2}(K)} \leq C \sum_{i=1}^{2} h_{i} \left\| \frac{\partial u_{j}}{\partial x_{i}} \right\|_{L_{2}(K)}
\]

for \( j = 1, 2 \).

One might suspect that (3.11), because its sum contains a first-order term, is no better than the first-order estimate given by (3.12), but Example 3.3 will show that (3.11) does indeed yield sharper results.

The classical isotropic bound for \( d = 2 \) and \( k = 0 \) from [5] is

\[
\|(u - \Pi_{0} u)_{j}\|_{L_{2}(K)} \leq C \left( \max_{i=1,2} h_{i} \right) \sum_{i=1}^{2} \left\| \frac{\partial u_{j}}{\partial x_{i}} \right\|_{L_{2}(K)}
\]

for \( j = 1, 2 \).

### Example 3.3.

To keep this example as short as possible, we discuss only the approximation of the gradient of a boundary layer on the fine part of a Shishkin mesh. An analysis of the \( RT_{0} \) interpolation error of the gradient of the entire solution of a singularly perturbed reaction-diffusion problem on a complete Shishkin mesh, using (3.12), can be found in [10].

Let the singularly perturbed reaction-diffusion problem be

\[
-\varepsilon^{2} \Delta v + bv = f \quad \text{in } \Omega := [0, 1]^{2}, \quad v = 0 \quad \text{on } \partial \Omega,
\]

where \( b > 0 \) and the parameter \( \varepsilon \) typically satisfies \( 0 < \varepsilon \ll 1 \). From [6][11], assuming sufficient smoothness and compatibility of the data, the boundary layer component \( w(x_1, x_2) \) of the solution of (3.14) that is associated with the edge \( x_2 = 0 \) of \( \Omega \) satisfies

\[
\frac{\partial^{m+n} w(x_1, x_2)}{\partial x_1^m \partial x_2^n} \leq C(1 + \varepsilon^{2-m}) \varepsilon^{-n} e^{-\beta x_2/\varepsilon} \quad \text{for } 0 \leq m + n \leq 3,
\]

(3.15a)

\[
\frac{\partial^{3} w(x_1, x_2)}{\partial x_1^3} \leq C e^{-\beta x_2/\varepsilon}.
\]

(3.15b)

(The above papers prove only (3.15a), but from an inspection of the construction of \( w \) in [11] one can also prove (3.15b) using a maximum principle.) Here and in
the rest of this example, the generic constants $C$ are independent of $\varepsilon$ and of the mesh parameter $N$ that is now introduced.

To construct a Shishkin mesh for this problem, let $N$ be a positive integer that is divisible by 4; it is the number of mesh intervals in each coordinate direction. We make a reasonable practical assumption that $\varepsilon \leq N^{-1}$. The parameter $\lambda$ that specifies where the mesh changes from coarse to fine is defined by

$$\lambda = 2\varepsilon \beta^{-1} \ln N.$$ 

(3.16)

Divide each of the $x$-intervals $[0, \lambda]$ and $[1 - \lambda, 1]$ into $N/4$ equidistant subintervals and divide $[\lambda, 1 - \lambda]$ into $N/2$ equidistant subintervals. This gives a coarse mesh on $[\lambda, 1 - \lambda]$ and a fine mesh elsewhere. Treat the $y$-interval $[0, 1]$ in the same way. The final 2-dimensional mesh is a tensor product of these 1-dimensional piecewise equidistant Shishkin meshes; see Figure 1, where $N = 8$.

![Figure 1. Shishkin mesh for reaction-diffusion](image)

This mesh divides $\Omega$ into a set $T_h$ of mesh rectangles $K$ whose sides are parallel to the axes. The coarse mesh width is $O(N^{-1})$, while the fine mesh width is $O(\varepsilon N^{-1} \ln N)$.

We shall use each one of (3.11)–(3.13) to derive three different bounds for the weighted $L_2$ error $\varepsilon^{1/2} \| \nabla w - \Pi_0(\nabla w) \|_{L_2(\Omega_1)}$, where $\Omega_1 := [\lambda, 1 - \lambda] \times [0, \lambda]$ is the region where the boundary layer $w$ is most active. The weighting $\varepsilon^{1/2}$ is appropriate for problems such as this since typically $\varepsilon^{1/2} \| \nabla w \|_{L_2(\Omega_1)} = O(1)$; see (3.15a).

Set $\phi(x_1, x_2) = e^{-\beta x_2/\varepsilon}$. Summing (3.13) over all $K \subset \Omega_1$ and invoking (3.15) yields

$$\varepsilon \| (\nabla w - \Pi_0 \nabla w) \|_{L_2(\Omega_1)}^2 \leq C \varepsilon N^{-2} \left[ \| \phi \|_{L_2(\Omega_1)}^2 + \| \varepsilon^{-1} \phi \|_{L_2(\Omega_1)}^2 \right] \leq CN^{-2},$$

(3.17)

$$\varepsilon \| (\nabla w - \Pi_0 \nabla w) \|_{L_2(\Omega_1)}^2 \leq C \varepsilon N^{-2} \left[ \| \varepsilon^{-1} \phi \|_{L_2(\Omega_1)}^2 + \| \varepsilon^{-2} \phi \|_{L_2(\Omega_1)}^2 \right] \leq C \varepsilon^{-2} N^{-2}.$$

(3.18)
Alternatively, summing (3.12) over all \( K \subset \Omega_1 \) and invoking (3.15) yields
\[
\varepsilon \|(\nabla w - \Pi_0 \nabla w)_1\|_{L_2(\Omega_1)}^2 \leq C\varepsilon \left[ N^{-2} \|\phi\|^2_{L_2(\Omega_1)} + \varepsilon^2 N^{-2}(\ln N)^2 \|\varepsilon^{-1} \phi\|^2_{L_2(\Omega_1)} \right]
\]
(3.19)
\[
\leq C\varepsilon^2 N^{-2}(\ln N)^2,
\]
\[
\varepsilon \|(\nabla w - \Pi_0 \nabla w)_2\|_{L_2(\Omega_1)}^2 \leq C\varepsilon \left[ N^{-2} \|\varepsilon^{-1} \phi\|^2_{L_2(\Omega_1)} + \varepsilon^2 N^{-2}(\ln N)^2 \|\varepsilon^{-2} \phi\|^2_{L_2(\Omega_1)} \right]
\]
(3.20)
\[
\leq CN^{-2}(\ln N)^2.
\]
Finally, summing (3.11) over all \( K \subset \Omega_1 \) and invoking (3.15) yields
\[
\varepsilon \|(\nabla w - \Pi_0 \nabla w)_1\|_{L_2(\Omega_1)}^2 \leq C\varepsilon \left[ N^{-4} \|\phi\|^2_{L_2(\Omega_1)} + \varepsilon^2 N^{-4}(\ln N)^2 \|\varepsilon^{-1} \phi\|^2_{L_2(\Omega_1)} + \varepsilon^2 N^{-2}(\ln N)^2 \|\varepsilon^{-1} \phi\|^2_{L_2(\Omega_1)} \right]
\]
(3.21)
\[
\leq C\varepsilon^2 N^{-2}(\ln N)^2
\]
and
\[
\varepsilon \|(\nabla w - \Pi_0 \nabla w)_2\|_{L_2(\Omega_1)}^2 \leq C\varepsilon \left[ \varepsilon^4 N^{-4}(\ln N)^4 \|\varepsilon^{-3} \phi\|^2_{L_2(\Omega_1)} + \varepsilon^2 N^{-4}(\ln N)^2 \|\varepsilon^{-2} \phi\|^2_{L_2(\Omega_1)} + N^{-2} \|\varepsilon^{-1} \phi\|^2_{L_2(\Omega_1)} \right]
\]
(3.22)
\[
\leq CN^{-2}.
\]
The assumption that \( \varepsilon \leq N^{-1} \) implies that (3.18) dominates (3.17), (3.20) dominates (3.19), and (3.22) dominates (3.21). That is, the classical bound (3.13) yields \( \varepsilon^{1/2} \|\nabla w - \Pi_0 \nabla w\|_{L_2(\Omega_1)} \leq C\varepsilon^{-1} N^{-1} \) which is clearly unsatisfactory when \( \varepsilon \ll N^{-1} \); the estimate (3.12) of (1) yields the improved bound
\[
\varepsilon^{1/2} \|\nabla w - \Pi_0 \nabla w\|_{L_2(\Omega_1)} \leq CN^{-1} \ln N,
\]
but the sharpest bound is given by (3.11), i.e., by Theorem 3.2
\[
\varepsilon^{1/2} \|\nabla w - \Pi_0 \nabla w\|_{L_2(\Omega_1)} \leq CN^{-1}.
\]
\[\square\]

4. Reduced regularity of \( u \)

If \( \hat{u} \) is less regular than is required in Theorem 3.1 i.e., if \( \hat{u} \notin (W^{d+2,\ell}(K))^d \), then one must halt the calculation leading to (3.9) at the stage when the maximum regularity of \( \hat{u} \) is reached. This will lead to a bound on \( \|\hat{u} - \Pi_k \hat{u}\|_{L_2(\Omega_1)} \) that involves lower-order derivatives of \( \hat{u} - \Pi_k \hat{u} \) so we need to eliminate the derivatives of \( \Pi_k \hat{u} \) from this bound (in (3.9) they were equal to zero so no difficulty arose there). For this we have the following result:

**Lemma 4.1.** Let \( p \in [2, \infty] \). Let \( k \) be a non-negative integer. Let \( \ell \in \{1, 2, \ldots, d\} \). Let \( m_\ell \) be a positive integer and \( m_1, \ldots, m_{\ell-1}, m_{\ell+1}, \ldots, m_d \) non-negative integers. Assume that \( \hat{u} = (\hat{u}_1, \ldots, \hat{u}_d) \in (H^1(K))^d \) so that \( \Pi_k \hat{u} \in RT_{[k]}(K) \) is defined.
Assume also that $\partial^{m_1+\cdots+m_d}\hat{u}_\ell/\partial x_1^{m_1}\cdots\partial x_d^{m_d} \in L_p(\hat{K})$. Then there is a constant $C = C(d,k,p)$ such that

\begin{equation}
(4.1) \quad \left\| \frac{\partial^{m_1+\cdots+m_d}(\hat{\Pi}_k\hat{u})_\ell}{\partial x_1^{m_1}\cdots\partial x_d^{m_d}} \right\|_{L_p(\hat{K})} \leq C \left\| \frac{\partial^{m_1+\cdots+m_d}\hat{u}_\ell}{\partial x_1^{m_1}\cdots\partial x_d^{m_d}} \right\|_{L_p(\hat{K})}.
\end{equation}

Proof. For convenience of notation take $\ell = 1$ as the case $\ell \neq 1$ is similar. We abuse the notation slightly by writing $\hat{\Pi}_k \hat{u}_1$ instead of $(\hat{\Pi}_k \hat{u})_1$. For brevity set

$$\tilde{\partial} \equiv \frac{\partial^{m_1+\cdots+m_d}}{\partial x_1^{m_1}\cdots\partial x_d^{m_d}} \quad \text{and} \quad |\tilde{\partial}| = m_1 + \cdots + m_d.$$ 

We first prove the result for the case $p = 2$. Integrating by parts $|\tilde{\partial}| - 1$ times and observing that the boundary terms are eliminated by $(1 - x_1^2)^{m_1-1}\prod_{i=2}^d (1 - x_i^2)^{m_i}$, one gets

$$\int_{\hat{K}} [\hat{\partial}(\hat{\Pi}_k \hat{u}_1)]^2(1 - x_1^2)^{m_1-1}\prod_{i=2}^d (1 - x_i^2)^{m_i} = (-1)^{|\tilde{\partial}| - 1} \int_{\hat{K}} \hat{\partial}(\hat{\Pi}_k \hat{u}_1)$$

$$\frac{\partial^{m_1+1+m_2+\cdots+m_d}}{\partial x_1^{m_1-1}\partial x_2^{m_2} \cdots \partial x_d^{m_d}} \left\{ \hat{\partial}(\hat{\Pi}_k \hat{u}_1)(1 - x_1^2)^{m_1-1}\prod_{i=2}^d (1 - x_i^2)^{m_i} \right\}$$

$$= (-1)^{|\tilde{\partial}|} \int_{\hat{K}} (\hat{\Pi}_k \hat{u}_1) \tilde{\partial} \left\{ \hat{\partial}(\hat{\Pi}_k \hat{u}_1)(1 - x_1^2)^{m_1-1}\prod_{i=2}^d (1 - x_i^2)^{m_i} \right\}$$

$$+ (-1)^{|\tilde{\partial}| - 1} \int_{\hat{F}^+} (\hat{\Pi}_k \hat{u}_1) \frac{\partial^{m_1+1+m_2+\cdots+m_d}}{\partial x_1^{m_1-1}\partial x_2^{m_2} \cdots \partial x_d^{m_d}} \left\{ \hat{\partial}(\hat{\Pi}_k \hat{u}_1)(1 - x_1^2)^{m_1-1}\prod_{i=2}^d (1 - x_i^2)^{m_i} \right\}$$

$$- (-1)^{|\tilde{\partial}| - 1} \int_{\hat{F}^-} (\hat{\Pi}_k \hat{u}_1) \frac{\partial^{m_1+1+m_2+\cdots+m_d}}{\partial x_1^{m_1-1}\partial x_2^{m_2} \cdots \partial x_d^{m_d}} \left\{ \hat{\partial}(\hat{\Pi}_k \hat{u}_1)(1 - x_1^2)^{m_1-1}\prod_{i=2}^d (1 - x_i^2)^{m_i} \right\},$$

after another integration by parts, where the faces $\hat{F}^\pm$ were defined in the proof of Theorem 3.1. But

$$\frac{\partial^{m_1+1+m_2+\cdots+m_d}}{\partial x_1^{m_1-1}\partial x_2^{m_2} \cdots \partial x_d^{m_d}} \left\{ \hat{\partial}(\hat{\Pi}_k \hat{u}_1)(1 - x_1^2)^{m_1-1}\prod_{i=2}^d (1 - x_i^2)^{m_i} \right\} \in Q_k(\hat{\hat{F}}^\pm)$$

and

$$\tilde{\partial} \left\{ \hat{\partial}(\hat{\Pi}_k \hat{u}_1)(1 - x_1^2)^{m_1-1}\prod_{i=2}^d (1 - x_i^2)^{m_i} \right\} \in Q_{k-1,k-1,k}(\hat{K}).$$
Consequently, the definition of $\hat{\Pi}_k$ in (3.2) enables us to replace $\hat{\Pi}_k \hat{u}_1$ by $\hat{u}_1$ in the above integrals:

$$\int_K [\tilde{\partial}(\hat{\Pi}_k \hat{u}_1)]^2 (1 - x_1^2)^{m_1-1} \prod_{i=2}^d (1 - x_i^2)^{m_i}$$

$$= (-1)^{\lvert \partial \rvert} \int_K (\hat{u}_1) \tilde{\partial} \left\{ \tilde{\partial}(\hat{\Pi}_k \hat{u}_1)(1 - x_1^2)^{m_1-1} \prod_{i=2}^d (1 - x_i^2)^{m_i} \right\}$$

$$+ (-1)^{\lvert \partial \rvert-1} \int_{\partial K} (\hat{u}_1) \frac{\partial}{\partial x_1} (\hat{\Pi}_k \hat{u}_1)(1 - x_1^2)^{m_1-1} \prod_{i=2}^d (1 - x_i^2)^{m_i}$$

$$- (-1)^{\lvert \partial \rvert-1} \int_{\partial K} (\hat{u}_1) \frac{\partial}{\partial x_1} \left( \tilde{\partial}(\hat{\Pi}_k \hat{u}_1)(1 - x_1^2)^{m_1-1} \prod_{i=2}^d (1 - x_i^2)^{m_i} \right)$$

$$= \int_K (\tilde{\partial} \hat{u}_1) [\tilde{\partial}(\hat{\Pi}_k \hat{u}_1)] (1 - x_1^2)^{m_1-1} \prod_{i=2}^d (1 - x_i^2)^{m_i},$$

after $\lvert \partial \rvert$ integrations by parts that reverse the earlier calculation. Hence

$$\int_K [\tilde{\partial}(\hat{\Pi}_k \hat{u}_1)]^2 (1 - x_1^2)^{m_1-1} \prod_{i=2}^d (1 - x_i^2)^{m_i} \leq \|\tilde{\partial} \hat{u}_1\|_{L_2(K)} \|\tilde{\partial}(\hat{\Pi}_k \hat{u}_1)\|_{L_2(K)}.$$

The conclusion of the lemma (in the case $p = 2$) will now follow if we can show that there exists a positive constant $C$ such that

$$\int_K [\tilde{\partial} q]^2 (1 - x_1^2)^{m_1-1} \prod_{i=2}^d (1 - x_i^2)^{m_i} \leq \int_K [\tilde{\partial} q]^2 (1 - x_1^2)^{m_1-1} \prod_{i=2}^d (1 - x_i^2)^{m_i}$$

for all $q \in Q_{k+1,k}, \ldots, k(K)$.

Suppose that (4.3) is false. Then there exists a sequence $\{p_n\}_{n=1}^\infty \subset Q_{k+1,k}, \ldots, k(K)$ such that

$$\int_K [\tilde{\partial} p_n]^2 = 1 \quad \text{and} \quad \int_K [\tilde{\partial} p_n]^2 (1 - x_1^2)^{m_1-1} \prod_{i=2}^d (1 - x_i^2)^{m_i} \leq \frac{1}{n} \quad \text{for} \quad n = 1, 2, \ldots.$$

The compactness of $\{q \in Q_{k+1,k}, \ldots, k(K) : \int_K [\tilde{\partial} q]^2 = 1\}$ implies the existence of a subsequence, which we also call $\{p_n\}_{n=1}^\infty$, such that $\lim_{n \to \infty} \int_K [\tilde{\partial}(p_n - p)]^2 = 0$ for some $p \in Q_{k+1,k}, \ldots, k$. It follows that $\int_K [\tilde{\partial} p]^2 = 1$ and $\int_K [\tilde{\partial} p]^2 (1 - x_1^2)^{m_1-1} \prod_{i=2}^d (1 - x_i^2)^{m_i} = 0$, which contradict each other. Thus (4.3) is true, and the case $p = 2$ of the lemma is proved.

The case $p > 2$ now follows quickly: the equivalence of norms on the finite-dimensional space $\{\tilde{\partial} q : q \in Q_{k+1,k}, \ldots, k(K)\}$ gives

$$\|\tilde{\partial}(\hat{\Pi}_k \hat{u}_1)\|_{L_p(K)} \leq C \|\tilde{\partial}(\hat{\Pi}_k \hat{u}_1)\|_{L_2(K)}$$

with $C = C(k, p)$, while by Hölder’s inequality

$$\|\tilde{\partial} \hat{u}_1\|_{L_2(K)}^2 \leq \left( \int_K |\tilde{\partial} \hat{u}_1|^p \right)^{2/p} \left( \int_K 1^{p/(p-2)} \right)^{1-\frac{2}{p}} = 2^{d\left(1-\frac{2}{p}\right)} \|\tilde{\partial} \hat{u}_1\|_{L_p(K)}^2.$$
Remark 4.2. The restriction \( m_\ell \geq 1 \) of Lemma 3.1 cannot be discarded. For example, suppose that \( d = 2, k = 1, \ell = 1, m_1 = 0, m_2 = 1 \) and \( \hat{u}_1(x_1, x_2) = x_2 e^{(x_1 - 1)/\varepsilon} \) for \((x_1, x_2) \in \tilde{K}\), where the parameter \( \varepsilon \) satisfies \( 0 < \varepsilon \ll 1 \). A calculation shows that \( \| \partial \hat{u}_1 / \partial x_2 \|_{L_2(K)} = O(\varepsilon^{1/2}) \) but \( \| \partial (\hat{\Pi}_1 \hat{u}_1) / \partial x_2 \|_{L_2(K)} = O(1) \), so as \( \varepsilon \) varies one cannot have \( \| \partial (\hat{\Pi}_1 \hat{u}_1) / \partial x_2 \|_{L_2(K)} \leq C \| \partial \hat{u}_1 / \partial x_2 \|_{L_2(K)} \) for some fixed constant \( C \).

The requirement that \( p \in [2, \infty] \) is imposed by the final paragraph of the proof. It is unclear whether or not the lemma remains true when \( 1 \leq p < 2 \).

Now we can halt the calculation leading to (3.9) at any intermediate stage then appeal to Lemma 3.1 to replace derivatives of \( \hat{\Pi}_k \hat{u}_\ell \) by the same derivatives of \( \hat{u}_\ell \) for \( \ell = 1, \ldots, d \) and \( p \in [2, \infty] \). In this way one can obtain an analogue of Theorem 3.1 when \( u \) has any specified degree of regularity. To write down a formula that encompasses all cases where \( u \) has reduced regularity would be very complicated, so we give only the following sample result.

Theorem 4.3. Let \( d = 2, p \in [2, \infty] \) and \( \hat{u} = (\hat{u}_1, \ldots, \hat{u}_d) \in W^{m,p}(\tilde{K}) \) for some \( m \) satisfying \( k + 1 \leq m < 2k + 2 \). Then there is a constant \( C = C(d, k, p) \) such that

\[
\| (\hat{u} - \Pi_k \hat{u})_1 \|_{L_p(K)} \leq C \left[ \sum_{j=0}^{m-k-2} \left\| \frac{\partial^{k+j+1} \hat{u}_1}{\partial x_1^{k+1} \partial x_2^j} \right\|_{L_p(K)} \right] \]

\[
+ \sum_{\ell=m-k}^{k+1} \left\| \frac{\partial^m \hat{u}_1}{\partial x_1^\ell \partial x_2^{m-\ell}} \right\|_{L_p(K)} + \sum_{i=0}^{m-k-1} \left\| \frac{\partial^{i+k+1} \hat{u}_1}{\partial x_1^i \partial x_2^{k+1}} \right\|_{L_p(K)} \]  

(4.4)

Estimates like (4.4) can be easily transformed to the parallelootope \( K \) of (3.1), just as Theorem 3.1 was transformed to Theorem 3.2.

5. Interpolation error bounds for affine images of axi-parallel \( K \)

In this section we extend Theorem 3.2 to elements \( \tilde{K} \) that are non-degenerate affine images of the axi-parallel element \( K \) of (3.1). For simplicity of notation we ignore translations of \( K \) since interpolation errors are unaffected by these.

Thus let \( \tilde{K} \) be a non-degenerate \( d \)-dimensional parallelootope, e.g., when \( d = 2 \) it is a parallelogram and when \( d = 3 \) it is a parallelepiped. Let the one-dimensional edges emanating from some vertex of \( \tilde{K} \) have lengths \( 2h_1, \ldots, 2h_d \) and directions given by the unit vectors \( g_1, \ldots, g_d \) respectively; these data will be the same for each vertex of \( \tilde{K} \) up to rearrangements and multiplications of some \( g_i \) by \(-1\).

Construct the matrix \( M \in \mathbb{R}^{d \times d} \) by setting its \( i \)-th column equal to \( g_i \) for \( i = 1, \ldots, d \). As \( \tilde{K} \) is non-degenerate the vectors \( g_i \) are linearly independent and consequently \( M \) is invertible. Let \( e_1 = (1, 0, \ldots, 0) \), \( e_2 = (0, 1, 0, \ldots, 0) \), \ldots, \( e_d = (0, \ldots, 0, 1) \) denote the standard orthonormal basis of \( \mathbb{R}^d \). Then \( g_i = Me_i \) for \( i = 1, \ldots, d \) and, ignoring any translation, \( \tilde{K} = M(K) \).

Given \( \hat{v} \in H^1(\tilde{K}) \) and \( i \in \{1, \ldots, d\} \), we shall write \( \partial \hat{v} / \partial g_i \) for the directional derivative \( \nabla \hat{v} \cdot g_i \).

For each \( u \in [H^1(\tilde{K})]^d \) the Raviart-Thomas interpolant \( \hat{\Pi}_k u \in RT_{[k]}(\tilde{K}) \) of order \( k \geq 0 \) is defined simply by replacing \( K \) and its faces in the definition (3.2) of \( \Pi_k \) by \( \tilde{K} \) and its faces. To analyse the interpolation error on \( \tilde{K} \) we introduce the standard Piola transformation: given any \( \hat{v} \in [L_2(\tilde{K})]^d \), define \( v \in [L_2(K)]^d \) by

\[
\hat{v}(\hat{x}) = \frac{1}{\det M} M v(x), \quad \text{where } \hat{x} = M x.

(5.1)
Then, for instance, it follows that
\begin{equation}
(5.2) \quad \tilde{\Pi}_k \hat{\mathbf{u}}(\tilde{x}) = \frac{1}{\det M} M \Pi_k \mathbf{u}(x) \ \forall \ x \in K;
\end{equation}
see [5,7] for further properties of this transformation.

Let \( \| \cdot \|_p \) denote the norm induced on \( d \times d \) matrices by the vector norm in \( L_p(\mathbb{R}^d) \) for \( p \in [1, \infty] \).

**Theorem 5.1.** Let \( \tilde{K} = M(K) \), where \( K \) is defined in \( (3.1) \) and \( M \) is an affine mapping. Let \( p \in [1, \infty] \). Let \( \hat{\mathbf{u}} = (\hat{u}_1, \ldots, \hat{u}_d) \in (W^{d k + 2, p}(\tilde{K}))^d \). Then there exists a constant \( C = C(d, k, p) \), which is independent of \( \hat{\mathbf{u}} \) and of \( h_i \), such that

\[
\| \hat{\mathbf{u}} - \tilde{\Pi}_k \hat{\mathbf{u}} \|_{L_p(\tilde{K})} \leq C \| M \|_p \left[ \sum_{i=1}^{d} h_i \sum_{\ell=1}^{d} \sum_{r \neq \ell}^{d} \sum_{m_\ell = 0}^{k} \tilde{u}_{\ell}^{m_\ell} \cdots \tilde{u}_{\ell-1}^{m_{\ell-1} + k +1 + m_{\ell+1} + \cdots + m_d} \det M \right]^{q - 1} \| M^{-1} \|_p^q \times \left\| \frac{\partial}{\partial x} \left( \frac{\partial m_1 + \cdots + m_{\ell-1} + k +1 + m_{\ell+1} + \cdots + m_d}{\partial g_1 \cdots \partial g_{\ell-1} \partial g_{\ell+1} \cdots \partial g_d} \hat{\mathbf{u}} \right) \right\|_{L_p(\tilde{K})}.
\]

\[
(5.3)
\]

where \( q = m_1 + \cdots + m_{\ell-1} + k +2 + m_{\ell+1} + \cdots + m_d \) and \( n = m_1 + \cdots + m_{j-1} + k +1 + m_{j+1} + \cdots + m_d \).

**Proof.** Suppose first that \( 1 \leq p < \infty \). By \( (5.1) \) and \( (5.2) \), the change of variables \( \tilde{x} \mapsto x \) yields

\[
\| \hat{\mathbf{u}} - \tilde{\Pi}_k \hat{\mathbf{u}} \|_{L_p(\tilde{K})}^p = \int_{\tilde{K}} \left( \sum_{i=1}^{d} \left| (\hat{\mathbf{u}} - \tilde{\Pi}_k \hat{\mathbf{u}})_i(\tilde{x}) \right| \right)^p d\tilde{x}
\]

\[
= \int_{K} \sum_{i=1}^{d} \left| \frac{1}{\det M} (M(\mathbf{u} - \Pi_k \mathbf{u}))_i(x) \right|^p |\det M| dx
\]

\[
\leq \left( \frac{\| M \|_p^p}{|\det M|^{p-1}} \| \mathbf{u} - \Pi_k \mathbf{u} \|_{L_p(K)}^p \right).
\]

Now invoke Theorem 3.2 to bound \( \| \mathbf{u} - \Pi_k \mathbf{u} \|_{L_p(K)}^p \). The derivatives of \( \mathbf{u} \) produced here are transformed to derivatives of \( \hat{\mathbf{u}} \) by using both equations in \( (5.1) \) and the definition of \( \partial \hat{\mathbf{u}} / \partial g_j \) as follows:

\[
\frac{\partial \mathbf{u}(x)}{\partial x_j} = (\det M) M^{-1} \frac{\partial \hat{\mathbf{u}}(\tilde{x})}{\partial \tilde{x}_j} = (\det M) M^{-1} \sum_{\ell=1}^{d} \frac{\partial \hat{\mathbf{u}}(\tilde{x})}{\partial \tilde{x}_\ell} \cdot \frac{\partial \tilde{x}_\ell}{\partial x_j}
\]

\[
= (\det M) M^{-1} \nabla \hat{\mathbf{u}} \cdot g_j = (\det M) M^{-1} \frac{\partial \hat{\mathbf{u}}(\tilde{x})}{\partial g_j}.
\]
for $j = 1, \ldots, d$, where $\partial \mathbf{u} / \partial x_j := (\partial u_1 / \partial x_j, \ldots, \partial u_d / \partial x_j)^T$. Consequently,

$$\left\| \frac{\partial}{\partial x_i} \left( \frac{\partial^m_{\ell-1} + k+1 + m_{\ell+1} + \cdots + m_d \mathbf{u}}{\partial x_1^{m_{\ell-1}} \cdots \partial x_{\ell-1}^{m_{\ell-1}} \partial x_{\ell+1}^{k+1} \partial x_{\ell+1}^{m_{\ell+1}} \cdots \partial x_d^{m_d}} \right) \right\|_{L_p(K)}$$

$$= \left[ \int_K \sum_{\ell=1}^d \left| \frac{\partial}{\partial x_i} \left( \frac{\partial^{m_{\ell-1}+k+1+m_{\ell+1}+\cdots+m_d u_\ell(x)}{\partial x_1^{m_{\ell-1}} \cdots \partial x_{\ell-1}^{m_{\ell-1}} \partial x_{\ell+1}^{k+1} \partial x_{\ell+1}^{m_{\ell+1}} \cdots \partial x_d^{m_d}} \right) \right|^p \, dx \right]^{1/p}$$

$$\leq \left[ \det M \cdot ||M^{-1}||_p \right] \sum_{\ell=1}^d \left| \frac{\partial}{\partial g_\ell} \left( \frac{\partial^{m_{\ell-1}+k+1+m_{\ell+1}+\cdots+m_d \tilde{u}(\tilde{x})}{\partial g_1^{m_{\ell-1}} \cdots \partial g_{\ell-1}^{m_{\ell-1}} \partial g_{\ell+1}^{k+2} \partial g_{\ell+1}^{m_{\ell+1}} \cdots \partial g_d^{m_d}} \right) \right|^p \left| \det M^{-1} \right|^{1/p}$$

One can prove a similar inequality to bound the other terms from Theorem 3.2. Substituting these inequalities into (5.4) yields the statement of the theorem.

The case $p = \infty$ is handled by a similar but slightly simpler calculation. 

The entries of the $d \times d$ matrix $M = (m_{ij})$ satisfy $|m_{ij}| \leq 1 \forall i, j$. Consequently, $||M||_p \leq C$ and $|\det M| \leq C$ for some $C = C(d, p)$. Theorem 5.1 is particularly useful when for all elements $K$ in a partition of a domain one also has $||M^{-1}||_p \leq C$ for some fixed constant $C$. This is the rectangular analogue of the “regular vertex property” for tetrahedral elements that is discussed in [1] Theorem 2.3.

Unlike Theorem 3.2, Theorem 5.1 does not give information about individual components of the interpolation error, but this is not necessarily a drawback. When we applied Theorem 3.2 in Example 3.3, the details of the calculation show that this theorem was used in a form that resembled Theorem 5.1 i.e., one could instead invoke Theorem 5.1 in this example and reach precisely the same conclusion.

**Remark 5.2.** When $\tilde{u}$ has reduced regularity so that Theorem 5.1 cannot be invoked, it is straightforward to take an applicable result such as Theorem 4.3 and transform it to $\tilde{K}$ by imitating the proof of Theorem 5.1.

**Remark 5.3.** A satisfactory extension of Theorem 5.1 to $\tilde{K} = M(K)$ for non-linear mappings $M$ and $d \geq 2$ seems to be difficult. The classical isotropic interpolation error estimates of [12] for such $\tilde{K}$ include a term $\text{div} \mathbf{u}$ that does not appear in our estimates. For $d = 2$ and shape-regular convex quadrilateral elements, in [4] an isotropic interpolation error bound is derived that does not include this $\text{div} \mathbf{u}$ term, but the authors state [4, p. 2431] that “We restrict our presentation to two-dimensional domains, the three-dimensional case being considerably more complicated”.
References


Department of Mathematics, National University of Ireland, Cork, Ireland

E-mail address: m.stynes@ucc.ie

Current address: Applied Mathematics Division, Beijing Computational Science Research Center, No. 3 He-Qing Road, Hai-Dian District, Beijing 100084, China