DIVERGENCE–FREE SCOTT–VOGELIUS ELEMENTS ON CURVED DOMAINS

MICHAEL NEILAN* AND M. BARIS OTUS†

Abstract. We construct and analyze an isoparametric finite element pair for the Stokes problem in two dimensions. The pair is defined by mapping the Scott-Vogelius finite element space via a Piola transform. The velocity space has the same degrees of freedom as the quadratic Lagrange finite element space, and therefore, the proposed spaces reduce to the Scott-Vogelius pair in the interior of the domain. We prove that the resulting method converges with optimal order, is divergence–free, and is pressure robust. Numerical examples are provided which support the theoretical results.

1. Introduction. Isoparametric finite element methods are a well-known and extensively studied technique to approximate PDEs on smooth domains. Such schemes use polynomial diffeomorphisms between reference and physical elements with degree dictated by the approximation properties of the underlying finite element space. The use of such mappings yield curved elements on the boundary that, while still do not conform exactly to the physical domain, generally lead to higher–order approximations and mitigate the geometric error. In particular, the resulting geometric error is generally of the same order as the discretization error, and thus, the resulting methods are potentially robust with respect to rates of convergence. The implementation and analysis of isoparametric elements for second–order, scalar elliptic problems are well–established, and classical theories exist [23, 9, 17, 8, 20]. On the other hand, isoparametric elements for mixed problems, in particular the Stokes problem, is less developed [2, 21, 11].

In this paper, we adopt and expand the isoparametric framework to construct a divergence–free method for incompressible flow, i.e., schemes that yield discrete velocity solutions that are divergence–free pointwise. The scheme is also pressure-robust, i.e., the gradient part of the source function only influences the discrete pressure solution. This feature allows a decoupling of errors between the velocity and pressure, which is beneficial for situations with fluid flow with large pressure gradient and/or small viscosity. Such divergence-free and pressure-robust finite element schemes seem to be gaining in popularity [14, 22, 13, 1, 18, 15, 3], although, as far as we are aware, the methods have only been constructed on polytopal domains. Thus, divergence–free methods are currently limited to second–order accuracy (formally) on general domains with smooth boundary.

The basis of our construction is the lowest-order two-dimensional Scott-Vogelius pair defined on Clough-Tocher refinements, i.e., simplicial triangulations obtained by connecting the vertices of each triangle in a given mesh to its barycenter. In this case, the velocity space is the space of continuous, piecewise quadratic polynomials, and the pressure space is the space of (discontinuous) piecewise linear polynomials. It is known, on affine Clough-Tocher meshes, this pair is stable, and the corresponding scheme is divergence-free and pressure-robust. However, a direct application of the isoparametric paradigm to this pair leads to a method with neither of these desirable properties. Indeed, the Scott-Vogelius pair, defined by standard isoparametric mappings, is given by

$$\breve{\boldsymbol{V}}_h = \{\boldsymbol{v} \in \boldsymbol{H}_0^1(\Omega_h) : \ \boldsymbol{v}|_K = \hat{\boldsymbol{v}} \circ F_K^{-1}, \ \exists \hat{\boldsymbol{v}} \in \boldsymbol{\mathcal{P}}_2(\hat{T}) \ \forall K \in \mathcal{T}_h^{ct}\}, \tag{1.1a}$$

$$\breve{Q}_h = \{q \in L_0^2(\Omega_h) : \ q|_K = \hat{q} \circ F_K^{-1}, \ \exists \hat{q} \in \mathcal{P}_1(\hat{T}) \ \forall K \in \mathcal{T}_h^{ct}\}, \tag{1.1b}$$

where $\hat{T}$ is a reference triangle, $\mathcal{P}_k(\hat{T})$ denotes the space of polynomials of degree $\leq k$ on $\hat{T}$, $F_K : \hat{T} \to K$ is a quadratic diffeomorphism, and $\mathcal{T}_h^{ct}$ is the Clough-Tocher refinement of a simplicial triangulation $\mathcal{T}_h$ (cf. Section 2 for a detailed explanation of the notation). Applying the chain rule shows $\operatorname{div} \boldsymbol{v}_h \notin \breve{Q}_h$ for general $\boldsymbol{v}_h \in \breve{\boldsymbol{V}}_h$ (unless F_K is affine $\forall K \in \mathcal{T}_h$), and simple calculations show the exact enforcement of the divergence–free constraint and the pressure–robustness of the scheme using $\breve{\boldsymbol{V}}_h \times \breve{Q}_h$ is lost on curved elements.

Our methodology to construct divergence–free and pressure robust schemes consists of two main ideas. First, instead of composition, we use a divergence–preserving transformation to recover the divergence–free property, i.e., we use a Piola transform in the definition of the local velocity space instead of compo-

*Department of Mathematics, University of Pittsburgh, Pittsburgh, PA 15260 (neilan@math.pitt.edu). Supported in part by the National Science Foundation grant DMS-2011733.

†Department of Mathematics, University of Pittsburgh, Pittsburgh, PA 15260 (mbo13@pitt.edu).

sition. Combining the local spaces defined through this mapping with the Lagrange degrees of freedom yields a global non-conforming (velocity) finite element space that is $\boldsymbol{H}^1$-conforming in the interior of the domain and $\boldsymbol{H}(\mathrm{div})$-conforming globally. We also show that the resulting space is "weakly continuous," and therefore suitable for second-order elliptic problems.

The second main idea in our construction is to treat the Scott-Vogelius pair as a macro-element, rather than a finite element space defined on a refined (Clough-Tocher) triangulation. In particular, local spaces are defined by mapping a macro reference local space, and therefore the corresponding finite element code does not "see" the global Clough-Tocher triangulation. This modification is motivated by the stability analysis of the Scott-Vogelius pair, which is based on Stenberg's macro-element technique [7]. Adopting this technique to the isoparametric setting, we show that the resulting pair satisfies the inf-sup condition, and therefore the finite element method for the Stokes problem is well-posed.

The rest of the paper is organized as follows. In the next section, we set the notation, state the properties of the quadratic diffeomorphisms, and provide some preliminary results. In Section 3, we define the local spaces of the velocity-pressure pair and provide a unisolvent set of degrees of freedom. Here, we also prove a local inf-sup stability result. Section 4 states the global spaces and proves a global inf-sup stability result. We also show in this section that functions in the discrete velocity space enjoy weak continuity properties. In Section 5, we state the finite element method and show that the method is optimally convergent. Section 6 gives a pressure-robust scheme through the use of commuting projections, and Section 7 provides numerical experiments which confirm the theoretical results. Some auxiliary results are given in Appendix A.

2. Preliminaries. We assume that the domain $\Omega \subset \mathbb{R}^2$ is sufficiently smooth, and the boundary $\partial\Omega$ is given by a finite number of local charts. The construction of the mesh with curved boundaries follows the standard isoparametric framework in [17, 8, 9, 5]. In particular, we start with a shape-regular and affine triangulation $\tilde{\mathcal{T}}_h$, with mesh size sufficiently small, such that the boundary vertices of $\tilde{\mathcal{T}}_h$ lie on $\partial\Omega$, and $\tilde{\Omega}_h := \mathrm{int}\Big(\cup_{\tilde{T}\in\tilde{\mathcal{T}}_h} \overline{\tilde{T}}\Big)$ is an $\mathcal{O}(h^2)$ polygonal approximation to Ω. Here, $h = \max_{\tilde{T}\in\tilde{\mathcal{T}}_h} \mathrm{diam}(\tilde{T})$. We assume each $\tilde{T} \in \tilde{\mathcal{T}}_h$ has at most two boundary vertices.

Remark 2.1. *For the continuation of the paper, we use C (with or without subscript) to denote a generic constant that is independent of any mesh size parameter. For a regular mapping $H : \mathbb{R}^2 \to \mathbb{R}^2$, we denote its Jacobian by DH.*

We let $G : \tilde{\Omega}_h \to \Omega$ be a bijective map with $\|G\|_{W^{1,\infty}(\tilde{\Omega}_h)} \le C$ such that such that $G|_{\tilde{T}}(x) = x$ at all vertices of $\tilde{T}$, in particular, G is the identity map for any triangle $\tilde{T} \in \tilde{\mathcal{T}}_h$ with three interior vertices. We denote by G_h the piecewise quadratic nodal interpolant of G satisfying $\|DG_h\|_{W^{1,\infty}(\tilde{T})} \le C$ and $\|DG_h^{-1}\|_{W^{1,\infty}(\tilde{T})} \le C$ for all $\tilde{T} \in \tilde{\mathcal{T}}_h$. We then set

$$\mathcal{T}_h = \{G_h(\tilde{T}) : \ \tilde{T} \in \tilde{\mathcal{T}}_h\}, \qquad \Omega_h := \mathrm{int}\Big(\cup_{T\in\mathcal{T}_h} \overline{T}\Big)$$

to be the isoparametric triangulation and computational domain, respectively.

Denote by $\hat{T}$ the reference triangle with vertices $(1,0),(0,1)$, and $(0,0)$. For $\tilde{T} \in \tilde{\mathcal{T}}_h$, we denote by $F_{\tilde{T}} : \hat{T} \to \tilde{T}$ an affine mapping satisfying $|F_{\tilde{T}}|_{W^{1,\infty}(\hat{T})} \le Ch_T$ and $|F_{\tilde{T}}^{-1}|_{W^{1,\infty}(\tilde{T})} \le Ch_T^{-1}$, where $h_T = \mathrm{diam}(\tilde{T})$. We define the quadratic diffeomorphism $F_T : \hat{T} \to T$ as $F_T = G_h \circ F_{\tilde{T}}$ which satisfies

$$\begin{aligned} &|F_T|_{W^{m,\infty}(\hat{T})} \le Ch_T^m \quad 0 \le m \le 2, \qquad |F_T^{-1}|_{W^{m,\infty}(T)} \le Ch_T^{-m} \quad 0 \le m \le 3, \\ &c_1 h_T^2 \le \det(DF_T) \le c_2 h_T^2, \end{aligned} \tag{2.1}$$

where $h_T = \mathrm{diam}(G_h^{-1}(T))$. Note the mappings F_T and $F_{\tilde{T}}$ (with $T = G_h(\tilde{T})$) are oriented in the same way so that $F_T = F_{\tilde{T}}$ at the vertices of $\hat{T}$. In particular, the mappings coincide if $G|_{\tilde{T}}$ is the identity operator. Furthermore, if $e \subset \partial T$ is a straight edge with $e = F_T(\hat{e})$ and $\hat{e} \subset \partial\hat{T}$, then $F_T|_{\hat{e}}$ is affine. If $T \in \mathcal{T}_h$ has all straight edges, then F_T is affine and $T = G_h(\tilde{T}) = \tilde{T}$. The conditions on F_T and the shape-regularity of $\tilde{\mathcal{T}}_h$ imply $|T|/|G_h^{-1}(T)| \le C$ and $|G_h^{-1}(T)|/|T| \le C$ for all $T \in \mathcal{T}_h$.

Denote by $\hat{T}^{ct} = \{\hat{K}_i\}_{i=1}^3$ the Clough–Tocher triangulation of the reference triangle, obtained by connecting the vertices of $\hat{T}$ with its barycenter. We then define the analogous local triangulations on $\tilde{T} \in \tilde{\mathcal{T}}_h$ and $T \in \mathcal{T}_h$, respectively, (cf. Figure 2.1)

$$\tilde{T}^{ct} = \{F_{\tilde{T}}(\hat{K}) : \hat{K} \in \hat{T}^{ct}\},\ T^{ct} = \{F_T(\hat{K}) :\ \hat{K} \in \hat{T}^{ct}\}.$$

The properties of F_T show $|T| \le C|K|$ for all $K \in T^{ct}$.

We denote by $\mathcal{E}_h^I$ the interior (straight) edges of $\mathcal{T}_h$, and by $\mathcal{E}_h^{I,\partial} \subset \mathcal{E}_h^I$ the set of interior edges that have one endpoint on $\partial\Omega_h$, i.e., the set of interior edges that "touch" the computational boundary. We use the generic $\boldsymbol{n}$ to denote the outward unit normal of a domain which is clear from its context. The tangent vector $\boldsymbol{t}$ is obtained by rotating $\boldsymbol{n}$ 90 degrees counterclockwise.

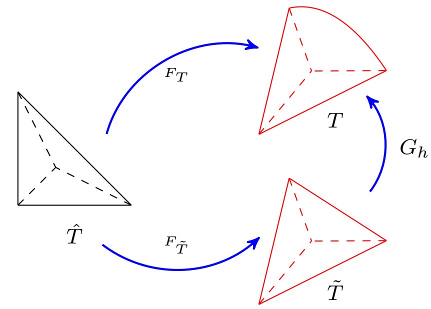

Figure 2.1. *Left: Clough-Tocher split of the reference triangle $\hat{T}$. Right: The corresponding curved and straight macro elements induced by the mappings F_T and $F_{\tilde{T}}$.*

Remark 2.2.

1. *The globally refined triangulations are given by*

$$\tilde{\mathcal{T}}_h^{ct} = \{\tilde{K} :\ \tilde{K} \in \tilde{T}^{ct},\ \exists \tilde{T} \in \tilde{\mathcal{T}}_h\}, \qquad \mathcal{T}_h^{ct} = \{K :\ K \in T^{ct},\ \exists T \in \mathcal{T}_h\}.$$

However, we emphasize that the construction of the Clough-Tocher isoparametric mesh $\mathcal{T}_h^{ct}$ is constructed by mapping the reference macro element *$\hat{T}^{ct}$. In particular, the finite element spaces, given in subsequent sections, are defined on $\mathcal{T}_h$ (not $\mathcal{T}_h^{ct}$); in fact, the corresponding finite element code does not "see" the refined triangulation $\mathcal{T}_h^{ct}$.*

2. *Note that this construction leads to curved interior edges in $\mathcal{T}_h^{ct}$, as interior edges of T^{ct} may be curved.*

The proofs of the following two lemmas are given in Appendix A.

Lemma 2.3. *For each $T \in \mathcal{T}_h$, define the matrix valued function $A_T : \hat{T} \to \mathbb{R}^{2\times 2}$ as*

$$A_T(\hat{x}) = \frac{DF_T(\hat{x})}{\det(DF_T(\hat{x}))}. \tag{2.2}$$

Then there holds

$$|A_T|_{W^{m,\infty}(\hat{T})} \le C h_T^{m-1}, \qquad \textit{and} \qquad |A_T^{-1}|_{W^{m,\infty}(\hat{T})} \le \begin{cases} C h_T^{1+m} & m = 0,1 \\ 0 & m \ge 2 \end{cases} \tag{2.3}$$

Lemma 2.4. *Let $\tilde{T} \in \tilde{\mathcal{T}}_h$ and $T \in \mathcal{T}_h$ with $T = G_h(\tilde{T})$. Let $\hat{e}$ be an edge of $\hat{T}$ with outward unit normal $\hat{\boldsymbol{n}}$, and assume that the corresponding edge $e = F_T(\hat{e})$ on T is straight. Then*

$$\det(DF_T(\hat{x}))(DF_T(\hat{x}))^{-\intercal}\hat{\boldsymbol{n}} = \det(DF_{\tilde{T}}(\hat{x}))(DF_{\tilde{T}}(\hat{x}))^{-\intercal}\hat{\boldsymbol{n}}$$

is constant on $\hat{e}$.

We also need a scaling result which is found in [5].

Lemma 2.5. *Suppose that $\boldsymbol{w}(x) = \hat{\boldsymbol{w}}(\hat{x})$ for sufficiently smooth $\boldsymbol{w} \in W^{m,p}(T)$. Then for any $K \in T^{ct}$,*

$$|\boldsymbol{w}|_{W^{m,p}(K)} \le C h_T^{2/p-m} \sum_{r=0}^m h_T^{2(m-r)} |\hat{\boldsymbol{w}}|_{W^{r,p}(\hat{K})},$$

$$|\hat{\boldsymbol{w}}|_{W^{m,p}(\hat{K})} \le C h_T^{m-2/p} \sum_{r=0}^m |\boldsymbol{w}|_{W^{r,p}(K)},$$

with $\hat{K} = F_T^{-1}(K)$.

3. Local Spaces. Recall $\hat{T} \subset \mathbb{R}^2$ is the reference triangle, and $\hat{T}^{ct} = \{\hat{K}_1, \hat{K}_2, \hat{K}_3\}$ is the Clough–Tocher triangulation, obtained by connecting the vertices of $\hat{T}$ with its barycenter. We define the polynomial spaces on $\hat{T}$ without boundary conditions:

$$\hat{\boldsymbol{V}} = \{\hat{\boldsymbol{v}} \in \boldsymbol{H}^1(\hat{T}) : \ \hat{\boldsymbol{v}}|_{\hat{K}} \in \boldsymbol{\mathcal{P}}_2(\hat{K}) \ \forall \hat{K} \in \hat{T}^{\rm ct}\}, \quad \hat{Q} = \{\hat{q} \in L^2(\hat{T}) : \ \hat{q}|_{\hat{K}} \in \mathcal{P}_1(\hat{K}) \ \forall \hat{K} \in \hat{T}^{\rm ct}\},$$

where $\mathcal{P}_k(S)$ is the space of scalar polynomials of degree $\le k$ with domain S, and $\boldsymbol{\mathcal{P}}_k(S) = [\mathcal{P}_k(S)]^2$.

For an affine triangle $\tilde{T} \in \tilde{\mathcal{T}}_h$ in the polygonal mesh, we define the spaces via composition

$$\tilde{\boldsymbol{V}}(\tilde{T}) = \{\tilde{\boldsymbol{v}} \in \boldsymbol{H}^1(\tilde{T}) : \ \tilde{\boldsymbol{v}}(\tilde{x}) = \hat{\boldsymbol{v}}(\hat{x}), \ \exists \hat{\boldsymbol{v}} \in \hat{\boldsymbol{V}}\}, \quad \tilde{Q}(\tilde{T}) = \{\tilde{q} \in L^2(\tilde{T}) : \ \tilde{q}(\tilde{x}) = \hat{q}(\hat{x}), \ \exists \hat{q} \in \hat{Q}\},$$

where $\tilde{x} = F_{\tilde{T}}(\hat{x})$. Thus, $\tilde{\boldsymbol{V}}(\tilde{T})$ is the local, quadratic Lagrange finite element space with respect to $\tilde{T}^{ct}$, and $\tilde{Q}(\tilde{T})$ is the space of (discontinuous) piecewise linear polynomials with respect to $\tilde{T}^{ct}$. We also define the analogous spaces with boundary conditions

$$\begin{aligned} \hat{\boldsymbol{V}}_0 &= \hat{\boldsymbol{V}} \cap \boldsymbol{H}^1_0(\hat{T}), & \hat{Q}_0 &= \hat{Q} \cap L^2_0(\hat{T}), \\ \tilde{\boldsymbol{V}}_0(\tilde{T}) &= \tilde{\boldsymbol{V}}(\tilde{T}) \cap \boldsymbol{H}^1_0(\tilde{T}), & \tilde{Q}_0(\tilde{T}) &= \tilde{Q}(\tilde{T}) \cap L^2_0(\tilde{T}). \end{aligned}$$

For $T \in \mathcal{T}_h$, possibly with curved boundary, we define the spaces with the aid of the Piola transform

$$\begin{aligned} &\boldsymbol{V}(T) = \{\boldsymbol{v} \in \boldsymbol{H}^1(T) : \ \boldsymbol{v}(x) = A_T(\hat{x})\hat{\boldsymbol{v}}(\hat{x}), \ \exists \hat{\boldsymbol{v}} \in \hat{\boldsymbol{V}}\}, \ \ \boldsymbol{V}_0(T) = \boldsymbol{V}(T) \cap \boldsymbol{H}^1_0(T), \\ &Q(T) = \{q \in L^2(T) : \ q(x) = \hat{q}(\hat{x}), \ \exists \hat{q} \in \hat{Q}\}, \qquad\qquad Q_0(T) = \{q \in L^2(T) : \ q(x) = \hat{q}(\hat{x}), \ \exists \hat{q} \in \hat{Q}_0\}. \end{aligned}$$

Here, $x = F_T(\hat{x})$ and we recall $A_T(\hat{x}) = DF_T(\hat{x})/\det(DF_T(\hat{x}))$. If F_T is affine, then $\boldsymbol{V}(T) = \tilde{\boldsymbol{V}}(\tilde{T})$ and $Q(T) = \tilde{Q}(\tilde{T})$; otherwise, both $\boldsymbol{V}(T)$ and $Q(T)$ are not necessarily piecewise polynomial spaces. Moreover, for $\boldsymbol{v} \in \boldsymbol{V}(T)$ and for a straight edge $e \subset \partial T$, the restriction of $\boldsymbol{v}$ to e is not necessarily a polynomial, even though F_T^{-1} is affine on e. Nonetheless, the next lemma shows the normal component of $\boldsymbol{v}$ is a polynomial on straight edges.

LEMMA 3.1. *Let $\boldsymbol{v} \in \boldsymbol{V}(T)$, and suppose that e is a straight edge of ∂T with unit normal $\boldsymbol{n}$. Then $\boldsymbol{v} \cdot \boldsymbol{n}|_e$ is a quadratic polynomial.*

Proof. Write $\boldsymbol{v}(x) = A_T(\hat{x})\hat{\boldsymbol{v}}(\hat{x})$ for some $\hat{\boldsymbol{v}} \in \hat{\boldsymbol{V}}$, and set $\hat{e} = F_T^{-1}(e)$ to be the corresponding edge in $\partial\hat{T}$ with outward unit normal $\hat{\boldsymbol{n}}$. We then have

$$\hat{\boldsymbol{v}} \cdot \hat{\boldsymbol{n}} = (\det(DF_T)DF_T^{-1}\boldsymbol{v}) \cdot \hat{\boldsymbol{n}} = (\det(DF_T)DF_T^{-\intercal}\hat{\boldsymbol{n}}) \cdot \boldsymbol{v}.$$

By Lemma 2.4, $(\det(DF_T)DF_T^{-\intercal}\hat{\boldsymbol{n}})$ is a constant vector. Using the identity $\boldsymbol{n} = DF^{-\intercal}\hat{\boldsymbol{n}}/|DF^{-\intercal}\hat{\boldsymbol{n}}|$ [19], we conclude $(\det(DF_T)DF_T^{-\intercal}\hat{\boldsymbol{n}})$ is a non-zero multiple of $\boldsymbol{n}$. In particular $\boldsymbol{v} \cdot \boldsymbol{n}$ is a non-zero multiple of $\hat{\boldsymbol{v}} \cdot \hat{\boldsymbol{n}}$. Because $F_T|_{\hat{e}}$ is affine and $\hat{\boldsymbol{v}} \cdot \hat{\boldsymbol{n}}$ is a quadratic polynomial on $\hat{e}$, we conclude $\boldsymbol{v} \cdot \boldsymbol{n}|_e$ is a quadratic polynomial on e. ∎

LEMMA 3.2. *Suppose $\boldsymbol{v} = A_T\hat{\boldsymbol{v}} \in \boldsymbol{V}(T)$ for some $\hat{\boldsymbol{v}} \in \hat{\boldsymbol{V}}$. There holds $\|\boldsymbol{v}\|_{H^1(T)} \le Ch_T^{-1}\|\hat{\boldsymbol{v}}\|_{H^1(\hat{T})}$.*

Proof. By a change of variables, the chain rule, Lemma 2.3, and Lemma 2.5, we have

$$\begin{aligned} \|\boldsymbol{v}\|_{H^1(T)} &\le C(|A_T\hat{\boldsymbol{v}}|_{H^1(\hat{T})} + h_T\|A_T\hat{\boldsymbol{v}}\|_{L^2(\hat{T})}) \\ &\le C(\|A_T\|_{L^\infty(\hat{T})}\|\hat{\boldsymbol{v}}\|_{H^1(\hat{T})} + \|A_T\|_{W^{1,\infty}(\hat{T})}\|\hat{\boldsymbol{v}}\|_{L^2(\hat{T})}) \le Ch_T^{-1}\|\hat{\boldsymbol{v}}\|_{H^1(\hat{T})}. \end{aligned}$$

∎

3.1. Degrees of freedom for $\boldsymbol{V}(T)$. The canonical (nodal) degrees of freedom (DOFs) of the quadratic Lagrange finite element space on T^{ct} are a given function's values at the (four) vertices in T^{ct}, and its values at the (six) edge midpoints in T^{ct}. Here, we show that these Lagrange DOFs form a unisolvent set over $\boldsymbol{V}(T)$.

Let $\mathcal{N}_{\hat{T}} := \{\hat{a}_i\}_{i=1}^{10}$ denote the set of (four) vertices and (six) edge midpoints in $\hat{T}^{ct}$. We let $\mathcal{N}_T := \{a_i\}_{i=1}^{10}$ and $\mathcal{N}_{\tilde{T}} := \{\tilde{a}_i\}_{i=1}^{10}$ be the corresponding sets on T^{ct} and $\tilde{T}^{ct}$, respectively, with $a_i = F_T(\hat{a}_i)$, and $\tilde{a}_i = F_{\tilde{T}}(\hat{a}_i)$.

LEMMA 3.3. *A function $\boldsymbol{v} \in \boldsymbol{V}(T)$ is uniquely determined by the values $\boldsymbol{v}(a)$ for all $a \in \mathcal{N}_T$.*

Proof. The number of DOFs given is 20 which matches in the dimension of $\boldsymbol{V}(T)$. Thus, it suffices to show that if $\boldsymbol{v} \in \boldsymbol{V}(T)$ vanishes on the DOFs, then $\boldsymbol{v} \equiv 0$.

Write $\boldsymbol{v}(x) = A_T(\hat{x})\hat{\boldsymbol{v}}(\hat{x})$ for some $\hat{\boldsymbol{v}} \in \hat{\boldsymbol{V}}$. We then have

$$0 = \boldsymbol{v}(a) = A_T(\hat{a})\hat{\boldsymbol{v}}(\hat{a}) \qquad \forall a \in \mathcal{N}_T.$$

Because $A_T(\hat{a})$ is invertible, we conclude $\hat{\boldsymbol{v}}(\hat{a}) = 0$ for all $\hat{a} \in \mathcal{N}_{\hat{T}}$. Since $\hat{\boldsymbol{v}}$ is uniquely determined by these values, we conclude $\hat{\boldsymbol{v}} \equiv 0$, and therefore $\boldsymbol{v} \equiv 0$. ◻

LEMMA 3.4. *There holds, for all $\boldsymbol{v} \in \boldsymbol{V}(T)$,*

$$\|\boldsymbol{v}\|^2_{H^1(T)} \le C \sum_{a \in \mathcal{N}_T} |\boldsymbol{v}(a)|^2.$$

Proof. Again, we write $\boldsymbol{v}(x) = A_T(\hat{x})\hat{\boldsymbol{v}}(\hat{x})$ with $A_T(\hat{x}) = DF_T(\hat{x})/\det(DF_T(\hat{x}))$ for some $\hat{\boldsymbol{v}} \in \hat{\boldsymbol{V}}$. By equivalence of norms in a finite dimensional setting, and the estimate $\|A_T^{-1}\|_{L^\infty(\hat{T})} \le C h_T$, we have

$$\begin{aligned}
\|\hat{\boldsymbol{v}}\|^2_{H^1(\hat{T})} &\le C \sum_{\hat{a} \in \mathcal{N}_{\hat{T}}} |\hat{\boldsymbol{v}}(\hat{a})|^2 = C \sum_{\hat{a} \in \mathcal{N}_{\hat{T}}} |A_T^{-1}(\hat{a}) A_T(\hat{a})\hat{\boldsymbol{v}}(\hat{a})|^2 \\
&\le C h_T^2 \sum_{\hat{a} \in \mathcal{N}_{\hat{T}}} |A_T(\hat{a})\hat{\boldsymbol{v}}(\hat{a})|^2 = C h_T^2 \sum_{a \in \mathcal{N}_T} |\boldsymbol{v}(a)|^2.
\end{aligned}$$

Therefore by Lemma 3.2,

$$\|\boldsymbol{v}\|^2_{H^1(T)} \le C\|A_T \hat{\boldsymbol{v}}\|^2_{H^1(\hat{T})} \le C\|A_T\|^2_{W^{1,\infty}(\hat{T})} \|\hat{\boldsymbol{v}}\|^2_{H^1(\hat{T})} \le C h_T^{-2} \|\hat{\boldsymbol{v}}\|^2_{H^1(\hat{T})} \le C \sum_{a \in \mathcal{N}_T} |\boldsymbol{v}(a)|^2.$$

◻

LEMMA 3.5. *For $T \in \mathcal{T}_h$, let $\boldsymbol{I}_T : \boldsymbol{H}^3(T) \to \boldsymbol{V}(T)$ be uniquely determined by the conditions*

$$(\boldsymbol{I}_T \boldsymbol{u})(a) = \boldsymbol{u}(a) \qquad \forall a \in \mathcal{N}_T.$$

Then there holds

$$\|\boldsymbol{u} - \boldsymbol{I}_T \boldsymbol{u}\|_{H^m(T)} \le C h_T^{3-m} \|\boldsymbol{u}\|_{H^3(T)} \qquad \forall \boldsymbol{u} \in \boldsymbol{H}^3(T), \quad m = 0, 1.$$

Proof. Let $\boldsymbol{u} \in \boldsymbol{H}^3(T)$, and for notational convenience, we set $\boldsymbol{v} = \boldsymbol{I}_T \boldsymbol{u}$.

Write

$$\boldsymbol{v}(x) = (A_T \hat{\boldsymbol{v}})(\hat{x}), \qquad \boldsymbol{u}(x) = (A_T \hat{\boldsymbol{u}})(\hat{x})$$

with $\hat{\boldsymbol{v}} \in \hat{\boldsymbol{V}}$ and $\hat{\boldsymbol{u}} \in \boldsymbol{H}^3(\hat{T})$. By definition of $\boldsymbol{I}_T \boldsymbol{u}$ and the nodal points, we find

$$(A_T \hat{\boldsymbol{v}})(\hat{a}) = (A_T \hat{\boldsymbol{u}})(\hat{a}) \qquad \forall \hat{a} \in \mathcal{N}_{\hat{T}}.$$

Therefore, because A_T is invertible, $\hat{\boldsymbol{v}}(\hat{a}) = \hat{\boldsymbol{u}}(\hat{a})$ for all $\hat{a} \in \mathcal{N}_{\hat{T}}$, i.e., $\hat{\boldsymbol{v}}$ is the quadratic Lagrange nodal interpolant of $\hat{\boldsymbol{u}}$ with respect to the local triangulation $\hat{T}^{ct}$. It then follows from standard interpolation theory that

$$\|\hat{\boldsymbol{u}} - \hat{\boldsymbol{v}}\|_{H^m(\hat{T})} \le C|\hat{\boldsymbol{u}}|_{H^3(\hat{T})}.$$

Applying Lemmas 2.5 and 2.3 then yields

$$|\boldsymbol{u} - \boldsymbol{v}|_{H^m(T)} \le C h_T^{1-m} \|A_T(\hat{\boldsymbol{u}} - \hat{\boldsymbol{v}})\|_{H^m(\hat{T})} \le C h_T^{1-m} \|A_T\|_{W^{m,\infty}(\hat{T})} \|\hat{\boldsymbol{u}} - \hat{\boldsymbol{v}}\|_{H^m(\hat{T})} \le C h_T^{-m} |\hat{\boldsymbol{u}}|_{H^3(\hat{T})}.$$

Finally, we once again use Lemmas 2.3 and 2.5 to obtain

$$\begin{aligned}
|\hat{\boldsymbol{u}}|_{H^3(\hat{T})} = |A_T^{-1} A_T \hat{\boldsymbol{u}}|_{H^3(\hat{T})} &\le C\big(\|A_T^{-1}\|_{L^\infty(\hat{T})} |A_T \hat{\boldsymbol{u}}|_{H^3(\hat{T})} + |A_T^{-1}|_{W^{1,\infty}(\hat{T})} |A_T \hat{\boldsymbol{u}}|_{H^2(\hat{T})}\big) \\
&\le C\big(h_T |A_T \hat{\boldsymbol{u}}|_{H^3(\hat{T})} + h_T^2 |A_T \hat{\boldsymbol{u}}|_{H^2(\hat{T})}\big) \le C h_T^3 \|\boldsymbol{u}\|_{H^3(T)}.
\end{aligned}$$

◻

3.2. A connection between local finite element spaces. In this section, we explicitly identify a correspondence between piecewise polynomials defined on the affine local triangulation $\tilde{T}^{ct}$ and functions on T^{ct} with $T = G_h(\tilde{T})$. This connection will be used to prove global inf-sup stability in the subsequent section.

DEFINITION 3.6. *Let $\tilde{T} \in \tilde{\mathcal{T}}_h$ and $T \in \mathcal{T}_h$ with $T = G_h(\tilde{T})$.*

1. *We define the operator $\Psi_T : \tilde{\boldsymbol{V}}(\tilde{T}) \to \boldsymbol{V}(T)$ uniquely by the conditions*

$$(\Psi_T \tilde{\boldsymbol{v}})(a) = \tilde{\boldsymbol{v}}(\tilde{a}) \qquad \forall \tilde{a} \in \mathcal{N}_{\tilde{T}}, \quad \text{where } a = G_h(\tilde{a}).$$

2. *We define the operator $\Upsilon_T : \tilde{Q}(\tilde{T}) \to Q(T)$ as*

$$(\Upsilon_T \tilde{q})(x) = \tilde{q}(F_{\tilde{T}}(\hat{x})).$$

THEOREM 3.7.

1. *If F_T is affine, then $(\Psi_T \tilde{\boldsymbol{v}})(x) = \tilde{\boldsymbol{v}}(\tilde{x})$, in particular, Ψ_T is the identity operator.*
2. *If $e \subset \partial T$ is a straight edge, so that $e \subset \partial \tilde{T}$, then*

$$(\Psi_T \tilde{\boldsymbol{v}}) \cdot \boldsymbol{n}|_e = \tilde{\boldsymbol{v}} \cdot \boldsymbol{n}|_e.$$

3. *There holds $\|\Psi_T \tilde{\boldsymbol{v}}\|_{H^1(T)} \le C \|\tilde{\boldsymbol{v}}\|_{H^1(\tilde{T})}$.*

Proof. For notational simplicity, we set $\boldsymbol{v} = \Psi_T \tilde{\boldsymbol{v}} \in \boldsymbol{V}(T)$.

1. If F_T is affine, so that DF_T is constant, we have $\boldsymbol{V}(T) = \tilde{\boldsymbol{V}}(\tilde{T})$. We then conclude that $(\Psi_T \tilde{\boldsymbol{v}}) = \tilde{\boldsymbol{v}}$ by Lemma 3.3.
2. Let $e \subset \partial T$ be a straight edge with outward unit normal $\boldsymbol{n}$, endpoints a_1 and a_2, and midpoint a_3. Then $e \subset \partial \tilde{T}$ and

$$(\boldsymbol{v} \cdot \boldsymbol{n})(a_1) = (\tilde{\boldsymbol{v}} \cdot \boldsymbol{n})(a_1), \quad (\boldsymbol{v} \cdot \boldsymbol{n})(a_2) = (\tilde{\boldsymbol{v}} \cdot \boldsymbol{n})(a_2), \quad (\boldsymbol{v} \cdot \boldsymbol{n})(a_3) = (\tilde{\boldsymbol{v}} \cdot \boldsymbol{n})(a_3).$$

   By Lemma 3.3, $\boldsymbol{v} \cdot \boldsymbol{n}|_e$ and $\tilde{\boldsymbol{v}} \cdot \boldsymbol{n}|_e$ are both quadratic polynomials, and therefore, these conditions imply $\boldsymbol{v} \cdot \boldsymbol{n}|_e = \tilde{\boldsymbol{v}} \cdot \boldsymbol{n}|_e$.
3. Set $\hat{\tilde{\boldsymbol{v}}}(\hat{x}) = \tilde{\boldsymbol{v}}(\tilde{x})$ with $\tilde{x} = F_{\tilde{T}}(\hat{x})$. Using Lemma 3.4 and a standard scaling argument, we have

$$\|\boldsymbol{v}\|^2_{H^1(T)} \le C \sum_{a \in \mathcal{N}_T} |\boldsymbol{v}(a)|^2 = C \sum_{\tilde{a} \in \mathcal{N}_{\tilde{T}}} |\tilde{\boldsymbol{v}}(\tilde{a})|^2 = C \sum_{\hat{a} \in \mathcal{N}_{\hat{T}}} |\hat{\tilde{\boldsymbol{v}}}(\hat{a})|^2 \le C \|\hat{\tilde{\boldsymbol{v}}}\|^2_{H^1(\hat{T})} \le C \|\tilde{\boldsymbol{v}}\|^2_{H^1(\tilde{T})}.$$

∎

3.3. Local Inf-sup stability. In this section, we derive an indirect local inf-sup stability result of the pair $\boldsymbol{V}_0(T) \times Q_0(T)$. As a first step, we use the stability of the analogous pair $\hat{\boldsymbol{V}}_0 \times \hat{Q}_0$ defined on the reference triangle. The proof of the following lemma is found in, e.g., [1, 13].

LEMMA 3.8. *For any $\hat{q} \in \hat{Q}_0$, there exists $\hat{\boldsymbol{v}} \in \hat{\boldsymbol{V}}_0$ such that $\hat{\nabla} \cdot \hat{\boldsymbol{v}} = \hat{q}$ with the bound $\|\hat{\boldsymbol{v}}\|_{H^1(\hat{T})} \le C \|\hat{q}\|_{L^2(\hat{T})}$.*

THEOREM 3.9. *Given $q \in Q_0(T)$, then there exists $\boldsymbol{v} \in \boldsymbol{V}_0(T)$ such that*

$$(\nabla \cdot \boldsymbol{v})(x) = \frac{h_T^2 q(x)}{\det(DF_T(F_T^{-1}(x)))}, \quad \text{and} \quad \|\boldsymbol{v}\|_{H^1(T)} \le C \|q\|_{L^2(T)}.$$

Proof. Let $q \in Q_0(T)$. Then there exists $q \in \hat{Q}_0$ such that $q(x) = \hat{q}(\hat{x})$. Because $h_T^2 \hat{q} \in \hat{Q}_0$, by Lemma 3.8, there exists $\hat{\boldsymbol{v}} \in \hat{\boldsymbol{V}}_0$ such that $\hat{\nabla} \cdot \hat{\boldsymbol{v}} = h_T^2 \hat{q}$ and $\|\hat{\boldsymbol{v}}\|_{H^1(\hat{T})} \le C h_T^2 \|\hat{q}\|_{L^2(\hat{T})}$. Setting $\boldsymbol{v}(x) = A_T \hat{\boldsymbol{v}} \in \boldsymbol{V}_0(T)$, we compute

$$(\nabla \cdot \boldsymbol{v})(x) = \frac{(\hat{\nabla} \cdot \hat{\boldsymbol{v}})(\hat{x})}{\det(DF_T(\hat{x}))} = \frac{h_T^2 \hat{q}(\hat{x})}{\det(DF_T(\hat{x}))} = \frac{h_T^2 q(x)}{\det(DF_T(F_T^{-1}(x)))}.$$

Applying Lemma 3.2 and a change of variables yields

$$\|\boldsymbol{v}\|_{H^1(T)} \le C h_T^{-1} \|\hat{\boldsymbol{v}}\|_{H^1(\hat{T})} \le C h_T \|\hat{q}\|_{L^2(\hat{T})} \le C \|q\|_{L^2(T)}.$$

∎

4. The Global Spaces. Define the Scott–Vogelius pair with respect to the affine triangulation $\tilde{\mathcal{T}}_h$:

$$\tilde{\boldsymbol{V}}^h = \{\tilde{\boldsymbol{v}} \in \boldsymbol{H}_0^1(\tilde{\Omega}_h) :\ \tilde{\boldsymbol{v}}|_{\tilde{T}} \in \tilde{\boldsymbol{V}}(\tilde{T}),\ \forall \tilde{T} \in \tilde{\mathcal{T}}_h\}, \qquad \tilde{Q}^h = \{\tilde{q} \in L_0^2(\tilde{\Omega}_h) :\ \tilde{q}|_{\tilde{T}} \in \tilde{Q}(\tilde{T}),\ \forall \tilde{T} \in \tilde{\mathcal{T}}_h\}.$$

We construct the global spaces $\boldsymbol{V}^h \times Q^h$ defined on $\mathcal{T}_h$ using the spaces $\tilde{\boldsymbol{V}}^h \times \tilde{Q}^h$ and with the aid of the operators Ψ_T and Υ_T given in Definition 3.6. To this end, we define Ψ and Υ to be the operators given by

$$\Psi|_T = \Psi_T, \qquad \Upsilon|_T = \Upsilon_T \qquad \forall T \in \mathcal{T}_h.$$

The global spaces, defined on the isoparametric mesh $\mathcal{T}_h$, are then given by

$$\boldsymbol{V}^h := \{\boldsymbol{v} :\ \boldsymbol{v} = \Psi\tilde{\boldsymbol{v}},\ \exists \tilde{\boldsymbol{v}} \in \tilde{\boldsymbol{V}}^h\}, \qquad Q^h := \{q :\ q = \Upsilon\tilde{q},\ \exists \tilde{q} \in \tilde{Q}^h\}.$$

Remark 4.1. *It is easy to see that the space $\boldsymbol{V}^h$ is equivalently defined as functions locally in $\boldsymbol{V}(T)$ on each $T \in \mathcal{T}_h$, are continuous on the DOFs in Lemma 3.3, and vanish on $\partial\Omega_h$.*

Theorem 4.2.

1. *There holds $\boldsymbol{V}^h \subset \boldsymbol{H}_0(\mathrm{div};\Omega_h) = \{\boldsymbol{v} \in \boldsymbol{L}^2(\Omega_h) :\ \nabla\cdot\boldsymbol{v} \in L^2(\Omega_h),\ \boldsymbol{v}\cdot\boldsymbol{n}|_{\partial\Omega_h} = 0\}$.*
2. *There holds $q \in Q^h$ if and only if $q|_T \circ F_T \in \hat{Q}$ for all $T \in \mathcal{T}_h$, and*

$$\sum_{T\in\mathcal{T}_h} 2|\tilde{T}| \int_T \frac{q}{\det(DF_T \circ F_T^{-1})} = 0.$$

Proof.

1. Let $T_1, T_2 \in \mathcal{T}_h$ such that $\emptyset \neq \partial T_1 \cap \partial T_2 =: e$, and let $\boldsymbol{n}$ be a unit normal of e. Note that e is a straight edge in $\mathcal{T}_h$. Let $\boldsymbol{v} = \Psi(\tilde{\boldsymbol{v}})$ for some $\tilde{\boldsymbol{v}} \in \tilde{\boldsymbol{V}}^h$, and denote by $\boldsymbol{v}_i$ the restriction of $\boldsymbol{v}$ to T_i. Likewise, let $\tilde{\boldsymbol{v}}_i$ denote the restriction of $\tilde{\boldsymbol{v}}$ to $\tilde{T}_i$. Then by Theorem 3.7 and the continuity of $\tilde{\boldsymbol{v}}$, we have

$$\boldsymbol{v}_1\cdot\boldsymbol{n}|_e = \tilde{\boldsymbol{v}}_1\cdot\boldsymbol{n}|_e = \tilde{\boldsymbol{v}}_2\cdot\boldsymbol{n}|_e = \boldsymbol{v}_2\cdot\boldsymbol{n}|_e.$$

   Thus, the normal component of $\boldsymbol{v}$ is single-valued on interior edges. Because $\boldsymbol{v}|_{\partial T\cap\partial\Omega_h} = 0$ for all $T \in \mathcal{T}_h$, we conclude that $\boldsymbol{v} \in \boldsymbol{H}_0(\mathrm{div};\Omega_h)$.
2. Let $q \in Q^h$. Then there exists a (unique) $\tilde{q} \in \tilde{Q}^h$ such that $q = \Upsilon\tilde{q}$, with $q|_T(F_T(\hat{x})) = \tilde{q}|_{\tilde{T}}(F_{\tilde{T}}(\hat{x}))$. We then find by a change of variables

$$0 = \int_{\tilde{\Omega}_h} \tilde{q} = \sum_{\tilde{T}\in\tilde{\mathcal{T}}_h} \int_{\tilde{T}} \tilde{q} = \sum_{\tilde{T}\in\tilde{\mathcal{T}}_h} 2|\tilde{T}| \int_{\hat{T}} \tilde{q}\circ F_{\tilde{T}} = \sum_{T\in\mathcal{T}_h} 2|\tilde{T}| \int_{\hat{T}} q\circ F_T = \sum_{T\in\mathcal{T}_h} 2|\tilde{T}| \int_T \frac{q}{\det(DF_T\circ F_T^{-1})}.$$

   The converse is proved similarly.

∎

4.1. Global inf-sup stability. In this section, we show the finite element pair $\boldsymbol{V}^h \times Q^h$ is inf-sup stable. This is achieved by using the local stability result given in Theorem 3.9 combined with Stenberg's macro element technique.

We define the spaces of piecewise constants with respect to $\tilde{\mathcal{T}}_h$ and $\mathcal{T}_h$:

$$\tilde{Y}^h := \{q \in L_0^2(\tilde{\Omega}_h) :\ \tilde{q}|_T \in \mathcal{P}_0(\tilde{T})\ \forall \tilde{T} \in \tilde{\mathcal{T}}_h\} \subset \tilde{Q}^h, \qquad Y^h := \{q :\ q = \Upsilon(\tilde{q}),\ \exists \tilde{q} \in \tilde{Y}^h\} \subset Q^h.$$

We first show that the pair $\boldsymbol{V}^h \times Y^h$ is stable in the following lemma.

Lemma 4.3. *There holds*

$$\sup_{\boldsymbol{v}\in\boldsymbol{V}^h\setminus\{0\}} \frac{\int_{\Omega_h}(\nabla\cdot\boldsymbol{v})q}{\|\nabla\boldsymbol{v}\|_{L^2(\Omega_h)}} \geq \gamma_1\|q\|_{L^2(\Omega_h)} \qquad \forall q \in Y^h,$$

where the gradient of $\boldsymbol{v}$ is understood piecewise with respect to $\mathcal{T}_h$. Here, $\gamma_1 > 0$ is a constant independent of h.

Proof. Fix $q \in Y^h$, and let $\tilde{q} \in \tilde{Y}^h$ be the piecewise constant function such that $q = \Upsilon\tilde{q}$. Note that, because q and $\tilde{q}$ are both piecewise constant, there holds $q|_{\Omega_h\cap\tilde{\Omega}_h} = \tilde{q}|_{\Omega_h\cap\tilde{\Omega}_h}$. In particular, we have

$$\int_T q = \frac{|T|}{|\tilde{T}|}\int_{\tilde{T}} \tilde{q}, \quad \text{and} \quad \|q\|^2_{L^2(T)} = \frac{|T|}{|\tilde{T}|}\|\tilde{q}\|^2_{L^2(\tilde{T})} \qquad \forall \tilde{T} \in \tilde{\mathcal{T}}_h,$$

with $T = G_h(\tilde{T})$. Thus we have $\|q\|_{L^2(\Omega_h)} \le C\|\tilde{q}\|_{L^2(\tilde{\Omega}_h)}$.

Let $\tilde{\boldsymbol{w}} \in \boldsymbol{H}^1_0(\tilde{\Omega}_h)$ satisfy $\tilde{\nabla}\cdot\tilde{\boldsymbol{w}} = \tilde{q}$ and $\|\tilde{\nabla}\tilde{\boldsymbol{w}}\|_{L^2(\tilde{\Omega}_h)} \le C\|\tilde{q}\|_{L^2(\tilde{\Omega}_h)}$. The results in [6, Theorem 4.4] and the properties of G ensure that $C > 0$ is independent of h. From the stability proof of the piecewise quadratic-constant pair [4, 7], there exists $\tilde{\boldsymbol{v}} \in \tilde{\boldsymbol{V}}^h$ such that

$$\int_{\tilde{e}} \tilde{\boldsymbol{v}} = \int_{\tilde{e}} \tilde{\boldsymbol{w}}, \quad \text{and} \quad \|\tilde{\nabla}\tilde{\boldsymbol{v}}\|_{L^2(\tilde{\Omega}_h)} \le C\|\tilde{\nabla}\tilde{\boldsymbol{w}}\|_{L^2(\tilde{\Omega}_h)}.$$

Let $\boldsymbol{v} = \Psi\tilde{\boldsymbol{v}}$ and note that $\|\nabla\boldsymbol{v}\|_{L^2(\Omega_h)} \le C\|\nabla\tilde{\boldsymbol{v}}\|_{L^2(\tilde{\Omega}_h)}$ by Theorem 3.7 (item (3)). Furthermore, this theorem shows that, on each $T \in \mathcal{T}_h$,

$$\int_T \nabla\cdot\boldsymbol{v} = \int_{\partial T}(\boldsymbol{v}\cdot\boldsymbol{n}) = \int_{\partial\tilde{T}}(\tilde{\boldsymbol{v}}\cdot\tilde{\boldsymbol{n}}) = \int_{\partial\tilde{T}}(\tilde{\boldsymbol{w}}\cdot\tilde{\boldsymbol{n}}) = \int_{\tilde{T}}\tilde{\nabla}\cdot\tilde{\boldsymbol{w}} = \int_{\tilde{T}}\tilde{q} = \frac{|\tilde{T}|}{|T|}\int_T q,$$

and therefore, because q is constant on T,

$$\int_T (\nabla\cdot\boldsymbol{v})q = \frac{|\tilde{T}|}{|T|}\int_T q^2 = \|\tilde{q}\|^2_{L^2(\tilde{T})}.$$

Summing over $T \in \mathcal{T}_h$ then gets

$$\begin{aligned}\int_{\Omega_h}(\nabla\cdot\boldsymbol{v})q = \|\tilde{q}\|^2_{L^2(\tilde{\Omega}_h)} &\ge C\|\tilde{q}\|_{L^2(\tilde{\Omega}_h)}\|\tilde{\nabla}\tilde{\boldsymbol{w}}\|_{L^2(\tilde{\Omega}_h)} \ge C\|\tilde{q}\|_{L^2(\tilde{\Omega}_h)}\|\tilde{\nabla}\tilde{\boldsymbol{v}}\|_{L^2(\tilde{\Omega}_h)}\\ &\ge C\|\tilde{q}\|_{L^2(\tilde{\Omega}_h)}\|\nabla\boldsymbol{v}\|_{L^2(\Omega_h)} \ge C\|q\|_{L^2(\Omega_h)}\|\nabla\boldsymbol{v}\|_{L^2(\Omega_h)}.\end{aligned}$$

Dividing this expression by $\|\nabla\boldsymbol{v}\|_{L^2(\Omega_h)}$ gives us the desired estimate. ◻

THEOREM 4.4. *There holds*

$$\sup_{\boldsymbol{v}\in\boldsymbol{V}^h\setminus\{0\}}\frac{\int_{\Omega_h}(\nabla\cdot\boldsymbol{v})q}{\|\nabla\boldsymbol{v}\|_{L^2(\Omega_h)}} \ge C\|q\|_{L^2(\Omega_h)} \qquad \forall q \in Q^h,$$

where the gradient of $\boldsymbol{v}$ is understood piecewise with respect to $\mathcal{T}_h$.

Proof. Let $q \in Q^h$. For each $T \in \mathcal{T}_h$, we define $\bar{q}_T \in \mathcal{P}_0(T)$ such that

$$\int_T \frac{(q - \bar{q}_T)}{\det(DF_T)} = 0,$$

and set $\bar{q}$ such that $\bar{q}|_T = \bar{q}_T$ for all $T \in \mathcal{T}_h$. Then $(q-\bar{q})|_T \in Q_0(T)$ for all $T \in \mathcal{T}_h$, and $\bar{q} \in Y^h$. Consequently, by Theorem 3.9, for each $T \in \mathcal{T}_h$, there exists $\boldsymbol{v}_{1,T} \in \boldsymbol{V}_0(T)$ such that

$$\nabla\cdot\boldsymbol{v}_{1,T} = \frac{h_T^2(q-\bar{q})}{\det(DF_T)}, \qquad \|\nabla\boldsymbol{v}_{1,T}\| \le C\|q-\bar{q}\|_{L^2(T)}.$$

Set $\boldsymbol{v}_1$ such that $\boldsymbol{v}_1|_T = \boldsymbol{v}_{1,T}$ for all $T \in \mathcal{T}_h$. Then $\boldsymbol{v}_1 \in \boldsymbol{V}^h$ because $\boldsymbol{v}_{1,T}|_{\partial T} = 0$. We also have $\|\nabla\boldsymbol{v}_1\|_{L^2(\Omega_h)} \le C\|q-\bar{q}\|_{L^2(\Omega_h)}$, and

$$\int_{\Omega_h}(\nabla\cdot\boldsymbol{v}_1)(q-\bar{q}) = \sum_{T\in\mathcal{T}_h}\int_T(\nabla\cdot\boldsymbol{v}_1)(q-\bar{q})$$

$$
\begin{aligned}
&= \sum_{T\in\mathcal{T}_h} \int_T \frac{h_T^2|q-\bar{q}|^2}{\det(DF_T)} \ge C \sum_{T\in\mathcal{T}_h} \int_T |q-\bar{q}|^2 \\
&= C\|q-\bar{q}\|^2_{L^2(\Omega_h)} \\
&\ge C\|q-\bar{q}\|_{L^2(\Omega_h)} \|\nabla \boldsymbol{v}_1\|_{L^2(\Omega_h)}.
\end{aligned}
$$

Next, recall $\boldsymbol{v}_1|_{\partial T} = \boldsymbol{v}_{1,T}|_{\partial T} = 0$, and $\bar{q}$ is constant on each T. Therefore by the divergence theorem,

$$
\int_{\Omega_h} (\nabla\cdot \boldsymbol{v}_1)\bar{q} = \sum_{T\in\mathcal{T}_h} \int_T (\nabla\cdot \boldsymbol{v}_1)\bar{q} = \sum_{T\in\mathcal{T}_h} \int_{\partial T} (\boldsymbol{v}_1\cdot \boldsymbol{n})\bar{q} = 0.
$$

Thus, we conclude the existence of a constant γ_0 independent of h such that

$$
\gamma_0 \|q-\bar{q}\|_{L^2(\Omega_h)} \le \sup_{\boldsymbol{v}\in \boldsymbol{V}^h\setminus\{0\}} \frac{\int_{\Omega_h} (\nabla\cdot\boldsymbol{v})q}{\|\nabla \boldsymbol{v}\|_{L^2(\Omega_h)}}.
$$

Next, we use the stability of the $\boldsymbol{V}^h \times Y^h$ pair given in Lemma 4.3:

$$
\begin{aligned}
\gamma_1 \|\bar{q}\|_{L^2(\Omega_h)} &\le \sup_{\boldsymbol{v}\in \boldsymbol{V}^h\setminus\{0\}} \frac{\int_{\Omega_h} (\nabla\cdot\boldsymbol{v})\bar{q}}{\|\nabla \boldsymbol{v}\|_{L^2(\Omega_h)}} \\
&\le \sup_{\boldsymbol{v}\in \boldsymbol{V}^h\setminus\{0\}} \frac{\int_{\Omega_h} (\nabla\cdot\boldsymbol{v})q}{\|\nabla \boldsymbol{v}\|_{L^2(\Omega_h)}} + \|q-\bar{q}\|_{L^2(\Omega_h)} \le (1+\gamma_0^{-1}) \sup_{\boldsymbol{v}\in \boldsymbol{V}^h\setminus\{0\}} \frac{\int_{\Omega_h} (\nabla\cdot\boldsymbol{v})q}{\|\nabla \boldsymbol{v}\|_{L^2(\Omega_h)}}.
\end{aligned}
$$

Therefore,

$$
\|q\|_{L^2(\Omega_h)} \le \|q-\bar{q}\|_{L^2(\Omega_h)} + \|\bar{q}\|_{L^2(\Omega_h)} \le (\gamma_0^{-1} + \gamma_1^{-1}(1+\gamma_0^{-1})) \sup_{\boldsymbol{v}\in \boldsymbol{V}^h\setminus\{0\}} \frac{\int_{\Omega_h} (\nabla\cdot\boldsymbol{v})q}{\|\nabla \boldsymbol{v}\|_{L^2(\Omega_h)}}.
$$

This is the desired inf-sup condition. ◻

4.2. $\boldsymbol{V}^h$ as an approximate $\boldsymbol{H}^1_0(\Omega_h)$ function space. Recall from Theorem 4.2 that the discrete velocity space is $\boldsymbol{H}_0(\mathrm{div};\Omega_h)$-conforming. However, in general there holds $\boldsymbol{V}^h \not\subset \boldsymbol{H}^1(\Omega_h)$ because $\boldsymbol{v}|_e$ is not a quadratic polynomial on all edges e in $\mathcal{T}_h$. More precisely, $\boldsymbol{v}|_e$ is not a quadratic polynomial if $e\subset \partial T$, and T has a curved edge (otherwise $\boldsymbol{v}|_e$ is quadratic and is continuous across the edge). Nonetheless, the definition of $\boldsymbol{V}^h$ shows that functions in $\boldsymbol{V}^h$ are single-valued at three points on each internal edge. We use this property in the next two lemmas to show that functions in $\boldsymbol{V}^h$ are "weakly continuous."

LEMMA 4.5. *There exists an operator $\boldsymbol{E}_h : \boldsymbol{V}^h \to \boldsymbol{H}^1_0(\Omega_h)$ such that for all $\boldsymbol{v}\in\boldsymbol{V}^h$,*

$$
\|\boldsymbol{v} - \boldsymbol{E}_h\boldsymbol{v}\|_{L^2(T)} + h_T\|\nabla(\boldsymbol{v}-\boldsymbol{E}_h\boldsymbol{v})\|_{L^2(T)} \le Ch_T^2\|\nabla\boldsymbol{v}\|_{L^2(T)} \qquad \forall T\in\mathcal{T}_h. \tag{4.1}
$$

Proof. For given $\boldsymbol{v}\in\boldsymbol{V}^h$, there exists $\tilde{\boldsymbol{v}}\in\tilde{\boldsymbol{V}}^h$ such that $\boldsymbol{v} = \Psi\tilde{\boldsymbol{v}}$. In particular, $\tilde{\boldsymbol{v}}$ is uniquely determined by

$$
\boldsymbol{v}|_T(a) = \tilde{\boldsymbol{v}}|_{\tilde{T}}(\tilde{a}) \qquad \forall a\in\mathcal{N}_T, \quad \forall T\in\mathcal{T}_h,
$$

with $T = G_h(\tilde{T})$. We define the function $\boldsymbol{E}_h\boldsymbol{v}$ via

$$
\boldsymbol{E}_h\boldsymbol{v}|_T = (\tilde{\boldsymbol{v}}\circ F_{\tilde{T}}\circ F_T^{-1})|_T \qquad \forall T\in\mathcal{T}_h.
$$

That is, $\boldsymbol{E}_h\boldsymbol{v}$ is the function in the standard isoparametric quadratic Lagrange finite element space associated with $\tilde{\boldsymbol{v}}$. We then have $\boldsymbol{E}_h\boldsymbol{v}\in\boldsymbol{H}^1_0(\Omega_h)$. Furthermore, since F_T^{-1} is affine on straight edges, we see $\tilde{\boldsymbol{v}} = \boldsymbol{E}_h\boldsymbol{v}$ on straight edges. In particular, we conclude

$$
\boldsymbol{E}_h\boldsymbol{v}|_T(a) = \boldsymbol{v}|_T(a) \qquad \forall a\in\mathcal{N}_T, \quad \forall T\in\mathcal{T}_h.
$$

We now estimate the difference $\boldsymbol{v} - \boldsymbol{E}_h\boldsymbol{v}$. On affine (non-curved) triangles, we easily see $\boldsymbol{v} = \boldsymbol{E}_h\boldsymbol{v}$ because both functions are piecewise quadratic polynomials. Therefore the estimate trivially holds in this case.

Next, let $T \in \mathcal{T}_h$ with curved boundary. We then have $\boldsymbol{v}|_{\partial T\cap\partial\Omega_h} = 0$. Write $\boldsymbol{v}|_T(x) = A_T(\hat{x})\hat{\boldsymbol{v}}(\hat{x})$ with $\hat{\boldsymbol{v}} \in \hat{\boldsymbol{V}}$ and $A_T = DF_T/\det(DF_T)$. We also set $\hat{\boldsymbol{w}} \in \hat{\boldsymbol{V}}$ such that $\hat{\boldsymbol{w}}(\hat{x}) = \boldsymbol{E}_h\boldsymbol{v}|_T(x)$. We then have

$$A_T(\hat{a})\hat{\boldsymbol{v}}(\hat{a}) = \hat{\boldsymbol{w}}(\hat{a}) \qquad \forall \hat{a} \in \mathcal{N}_{\hat{T}}.$$

Thus, $\hat{\boldsymbol{w}}$ is the piecewise quadratic Lagrange interpolant of $A_T\hat{\boldsymbol{v}}$ on $\hat{T}^{ct}$. It then follows from the Bramble–Hilbert lemma that

$$\|A_T\hat{\boldsymbol{v}} - \hat{\boldsymbol{w}}\|_{H^m(\hat{K})} \le C|A_T\hat{\boldsymbol{v}}|_{H^3(\hat{K})} \qquad \forall \hat{K} \in \hat{T}^{ct}, \quad m = 0, 1.$$

Expanding the right–hand side, using Lemma 2.3, and the fact that $\hat{\boldsymbol{v}}|_{\hat{K}}$ is a quadratic polynomial shows

$$\begin{aligned}|A_T\hat{\boldsymbol{v}}|_{H^3(\hat{K})} &\le C\big(|A|_{W^{3,\infty}(\hat{K})}\|\hat{\boldsymbol{v}}\|_{L^2(\hat{K})} + |A|_{W^{2,\infty}(\hat{K})}|\hat{\boldsymbol{v}}|_{H^1(\hat{K})} + |A|_{W^{1,\infty}(\hat{K})}|\hat{\boldsymbol{v}}|_{H^2(\hat{K})}\big)\\ &\le C\|\hat{\boldsymbol{v}}\|_{H^2(\hat{K})} \le C\|\hat{\boldsymbol{v}}\|_{L^2(\hat{K})},\end{aligned}$$

where we used the equivalence of norms in a finite dimensional setting in the last inequality. Using the estimate $\|A_T^{-1}\|_{L^\infty(\hat{T})} \le Ch_T$, we conclude

$$\|A_T\hat{\boldsymbol{v}} - \hat{\boldsymbol{w}}\|_{H^m(\hat{T})} \le Ch_T\|A_T\hat{\boldsymbol{v}}\|_{L^2(\hat{T})}, \qquad m = 0, 1.$$

We then use Lemma 2.5 and the Poincare inequality to get $(m = 0, 1)$

$$\begin{aligned}\|\boldsymbol{v} - \boldsymbol{E}_h\boldsymbol{v}\|_{H^m(T)} &\le Ch_T^{1-m}\|A_T\hat{\boldsymbol{v}} - \hat{\boldsymbol{w}}\|_{H^m(\hat{T})}\\ &\le Ch_T^{2-m}\|A_T\hat{\boldsymbol{v}}\|_{L^2(\hat{T})}\\ &\le Ch_T^{1-m}\|\boldsymbol{v}\|_{L^2(T)} \le Ch_T^{2-m}\|\nabla\boldsymbol{v}\|_{L^2(T)}.\end{aligned}$$

∎

Recall, $\mathcal{E}_h^I$ is the set of internal edges of $\mathcal{T}_h$. For $e = \mathcal{E}_h^I$, write $e = \partial T_+ \cap \partial T_-$ for some $T_\pm \in \mathcal{T}_h$. Let $\boldsymbol{n}_\pm$ denote the outward unit normal of $\partial T_\pm$ restricted to e, and for a piecewise smooth function $\boldsymbol{v}$, let $\boldsymbol{v}_\pm$ denote the restriction of $\boldsymbol{v}$ to $T_\pm$. We then define the jump operator

$$[\boldsymbol{v}]|_e = \boldsymbol{v}_+ \otimes \boldsymbol{n}_+ + \boldsymbol{v}_- \otimes \boldsymbol{n}_-,$$

where $(\boldsymbol{a} \otimes \boldsymbol{b})_{i,j} = a_i b_j$.

LEMMA 4.6. *Let $e \in \mathcal{E}_h^I$ with $e = \partial T_+ \cap \partial T_-$ for some $T_\pm \in \mathcal{T}_h$. Then there holds for all $\boldsymbol{v} \in \boldsymbol{V}^h$,*

$$\Big|\int_e [\boldsymbol{v}]\Big| \le Ch_T^3(\|\nabla\boldsymbol{v}\|_{L^2(T_+)} + \|\nabla\boldsymbol{v}\|_{L^2(T_-)}), \tag{4.2}$$

where $h_T = \max\{h_{T_+}, h_{T_-}\}$.

Proof. Let a_1, a_2 be the endpoints of e, and let a_3 be the midpoint of e. If both T_+ and T_- are affine, then $\boldsymbol{v}|_{T_+\cup T_-} \in \boldsymbol{H}^1(T_+ \cup T_-)$. This implies $[\boldsymbol{v}]|_e = 0$, and the estimate trivially follows.

Next suppose that at least one of $T_\pm$ has a curved edge, which implies that one of the endpoints a_1, a_2 lie on $\partial\Omega_h$. Without loss of generality, we assume that T_+ has a curved edge.

By construction of the space $\boldsymbol{V}^h$, in particular, the definition of Ψ, we have $[\boldsymbol{v}]|_e(a_i) = 0$, $i = 1, 2, 3$. It then follows from the error of Simpson's rule that

$$\Big|\int_e [\boldsymbol{v}]\Big| \le C|e|^5\big|[\boldsymbol{v}]\big|_{W^{4,\infty}(e)} \le Ch_T^5(|\boldsymbol{v}|_{W^{4,\infty}(K_+)} + |\boldsymbol{v}|_{W^{4,\infty}(K_-)}), \tag{4.3}$$

where $K_\pm \in T_\pm^{ct}$ satisfy $\partial K_+ \cap \partial K_- = e$.

Write $\boldsymbol{v}|_{K_\pm}(x) = (A_{T_\pm}\hat{\boldsymbol{v}}_\pm)|_{\hat{K}_\pm}(\hat{x})$ with $\hat{\boldsymbol{v}}_\pm \in \hat{\boldsymbol{V}}$, and $\hat{K}_\pm = F_{T_\pm}^{-1}(K_\pm)$. We apply Lemmas 2.5 and 2.3 to the right-hand side of (4.3) and use the fact that $\boldsymbol{v}_\pm$ is a quadratic polynomial:

$$\begin{aligned}
|\boldsymbol{v}|_{W^{4,\infty}(K_\pm)} &\le Ch_{T_\pm}^{-4} \sum_{r=0}^{4} h_{T_\pm}^{2(4-r)} |A_{T_\pm}\hat{\boldsymbol{v}}_\pm|_{W^{r,\infty}(\hat{K}_\pm)} \\
&\le Ch_{T_\pm}^{4} \sum_{r=0}^{4} h_{T_\pm}^{-2r} \sum_{j=0}^{r} |A_{T_\pm}|_{W^{r-j,\infty}(\hat{T}_\pm)} |\hat{\boldsymbol{v}}_\pm|_{W^{j,\infty}(\hat{K}_\pm)} \\
&\le Ch_{T_\pm}^{4} \sum_{r=0}^{4} h_{T_\pm}^{-2r} \sum_{j=0}^{2} h_{T_\pm}^{r-j-1} |\hat{\boldsymbol{v}}_\pm|_{W^{j,\infty}(\hat{K}_\pm)} \\
&\le C \sum_{j=0}^{2} h_{T_\pm}^{-j-1} |\hat{\boldsymbol{v}}_\pm|_{W^{j,\infty}(\hat{K}_\pm)} \le Ch_{T_\pm}^{-3} \|\hat{\boldsymbol{v}}_\pm\|_{L^2(\hat{K}_\pm)},
\end{aligned}$$

where we used equivalence of norms in the last inequality. Using the estimate $\|A_{T_\pm}^{-1}\|_{L^\infty(\hat{T})} \le Ch_{T_\pm}$ and Lemma 2.5 we get

$$|\boldsymbol{v}|_{W^{4,\infty}(K_\pm)} \le Ch_{T_\pm}^{-2} \|A_{T_\pm}\hat{\boldsymbol{v}}_\pm\|_{L^2(\hat{T})} \le Ch_{T_\pm}^{-3} \|\boldsymbol{v}\|_{L^2(T_\pm)}.$$

Combining this estimate with (4.3) and applying the Poincare inequality (on T_+) yields

$$\Big| \int_e [\boldsymbol{v}] \Big| \le Ch_T^2 \big(\|\boldsymbol{v}\|_{L^2(T_-)} + h_T \|\nabla \boldsymbol{v}\|_{L^2(T_+)} \big). \tag{4.4}$$

Next we show $\|\boldsymbol{v}\|_{L^2(T_-)} \le Ch_T \|\nabla \boldsymbol{v}\|_{L^2(T_-)}$. To this end, we set $\boldsymbol{w} = \boldsymbol{E}_h \boldsymbol{v}$, where $\boldsymbol{E}_h \boldsymbol{v}$ is given in Lemma 4.5. We then write

$$\|\boldsymbol{v}\|_{L^2(T_-)} \le \|\boldsymbol{v} - \boldsymbol{w}\|_{L^2(T_-)} + \|\boldsymbol{w}\|_{L^2(T_-)} \le \|\boldsymbol{w}\|_{L^2(T_-)} + Ch_{T_-}^2 \|\nabla \boldsymbol{v}\|_{L^2(T_-)}.$$

Let $\hat{\boldsymbol{w}} \in \hat{\boldsymbol{V}}$ such that $\hat{\boldsymbol{w}}(\hat{x}) = \boldsymbol{w}(x)$ with $x = F_{T_-}(\hat{x})$. Noting that $\boldsymbol{w}$ vanishes on $\partial\Omega_h$, in particular $\boldsymbol{w}$ vanishes on at least one vertex of T_-, we conclude that

$$\hat{\boldsymbol{w}} \to \|\hat{\nabla} \boldsymbol{w}\|_{L^2(\hat{T})}$$

is a norm. Therefore by Lemma 2.5 and equivalence of norms,

$$\|\boldsymbol{w}\|_{L^2(T_-)} \le Ch_T \|\hat{\boldsymbol{w}}\|_{L^2(\hat{T})} \le Ch_T \|\hat{\nabla} \hat{\boldsymbol{w}}\|_{L^2(\hat{T})} \le Ch_T \|\nabla \boldsymbol{w}\|_{L^2(T_-)}.$$

Hence, we have

$$\|\boldsymbol{v}\|_{L^2(T_-)} \le C\Big(h_T \|\nabla \boldsymbol{w}\|_{L^2(T_-)} + h_T^2 \|\nabla \boldsymbol{v}\|_{L^2(T_-)} \Big) \le Ch_T \|\nabla \boldsymbol{v}\|_{L^2(T_-)}.$$

Combining this estimate with (4.4), we obtain the desired estimate (4.2). This concludes the proof. ☐

5. Finite Element Method and Convergence Analysis. For a given function $\boldsymbol{f}$, we let $(\boldsymbol{u}, p) \in \boldsymbol{H}_0^1(\Omega) \times L_0^2(\Omega)$ be the solution to the Stokes problem

$$-\nu \Delta \boldsymbol{u} + \nabla p = \boldsymbol{f}, \qquad \nabla \cdot \boldsymbol{u} = 0 \qquad \text{in } \Omega,$$

where $\nu > 0$ is the viscosity. We assume that $\partial\Omega$ and $\boldsymbol{f}$ are sufficiently smooth such that $(\boldsymbol{u}, p) \in \boldsymbol{H}^3(\Omega) \times H^2(\Omega)$, and can be extended to $\mathbb{R}^2$ in a way such that $(\boldsymbol{u}, p) \in \boldsymbol{H}^3(\mathbb{R}^2) \times H^2(\mathbb{R}^2)$ with $\nabla \cdot \boldsymbol{u} = 0$ and (cf. [16])

$$\|\boldsymbol{u}\|_{H^3(\mathbb{R}^2)} \le C \|\boldsymbol{u}\|_{H^3(\Omega)}, \qquad \|p\|_{H^2(\mathbb{R}^2)} \le C \|p\|_{H^2(\Omega)}.$$

We then extend $\boldsymbol{f}$ by

$$\boldsymbol{f} = -\nu\Delta\boldsymbol{u} + \nabla p,$$

so that $\boldsymbol{f} \in \boldsymbol{H}^1(\mathbb{R}^2)$.

We denote by $\boldsymbol{f}_h \in \boldsymbol{L}^2(\Omega_h)$ a computable approximation of $\boldsymbol{f}|_\Omega$. For example, $\boldsymbol{f}_h$ could be the (global) quadratic Lagrange nodal interpolant of $\boldsymbol{f}$.

The finite element method seeks $(\boldsymbol{u}_h, p_h) \in \boldsymbol{V}^h \times Q^h$ such that

$$\int_{\Omega_h} \nu\nabla\boldsymbol{u}_h : \nabla\boldsymbol{v} - \int_{\Omega_h} (\nabla\cdot\boldsymbol{v})p_h = \int_{\Omega_h} \boldsymbol{f}_h \cdot \boldsymbol{v} \qquad \forall \boldsymbol{v} \in \boldsymbol{V}^h, \tag{5.1a}$$

$$\int_{\Omega_h} (\nabla\cdot\boldsymbol{u}_h)q = 0 \qquad \forall q \in Q^h, \tag{5.1b}$$

where the gradient is understood piecewise with respect to the triangulation.

By the inf-sup condition established in Lemma 4.4 and standard theory of mixed finite element methods, problem (5.1) is well-posed.

Theorem 5.1. *There exists a unique solution* $(\boldsymbol{u}_h, p_h) \in \boldsymbol{V}^h \times Q^h$ *satisfying* (5.1).

Next, we show that, despite the non-inclusion $\nabla\cdot\boldsymbol{V}^h \not\subset Q^h$, the finite element method yields exactly divergence–free velocity approximations.

Lemma 5.2. *Suppose* $\boldsymbol{u}_h \in \boldsymbol{V}^h$ *satisfies* (5.1b). *Then* $\nabla\cdot\boldsymbol{u}_h \equiv 0$ *in* Ω_h.

Proof. For each $T \in \mathfrak{T}_h$, write $\boldsymbol{u}_h|_T = A_T\hat{\boldsymbol{u}}_T$, with $\hat{\boldsymbol{u}}_T \in \hat{\boldsymbol{V}}$. Define q to be the piecewise function

$$q|_T(x) = \frac{1}{2|\tilde{T}|}(\hat{\nabla}\cdot\hat{\boldsymbol{u}}_T)(\hat{x}), \qquad x = F_T(\hat{x}), \qquad T = G_h(\tilde{T})$$

for all $T \in \mathfrak{T}_h$. Well-known properties of the Piola transform show

$$q|_T = \frac{\det(DF_T \circ F_T^{-1})}{2|\tilde{T}|}(\nabla\cdot\boldsymbol{u}_h|_T) \qquad \forall T \in \mathfrak{T}_h,$$

and therefore

$$\sum_{T\in\mathfrak{T}_h} 2|\tilde{T}| \int_T \frac{q}{\det(DF_T \circ F_T^{-1})} = \sum_{T\in\mathfrak{T}_h} \int_T \nabla\cdot\boldsymbol{u}_h = \int_{\partial\Omega_h} \boldsymbol{u}_h\cdot\boldsymbol{n} = 0.$$

Thus we conclude $q \in Q^h$ by Theorem 4.2.

Next, by (5.1b),

$$\begin{aligned} 0 = \int_{\Omega_h} (\nabla\cdot\boldsymbol{u}_h)q = \sum_{T\in\mathfrak{T}_h} \int_T (\nabla\cdot\boldsymbol{u}_h)q &= \sum_{T\in\mathfrak{T}_h} \frac{1}{2|\tilde{T}|} \int_{\hat{T}} \frac{\hat{\nabla}\cdot\hat{\boldsymbol{u}}_T}{\det(DF_T)}(\hat{\nabla}\cdot\hat{\boldsymbol{u}}_T)\det(DF_T) \\ &= \sum_{T\in\mathfrak{T}_h} \frac{1}{2|\tilde{T}|} \int_{\hat{T}} |\hat{\nabla}\cdot\hat{\boldsymbol{u}}_T|^2. \end{aligned}$$

Thus, $\hat{\nabla}\cdot\hat{\boldsymbol{u}}_T = 0$ for all $T \in \mathfrak{T}_h$, and therefore $\nabla\cdot\boldsymbol{u}_h = 0$. ∎

5.1. Convergence Analysis. Define the subspace of divergence–free functions:

$$\boldsymbol{X}^h := \{\boldsymbol{v} \in \boldsymbol{V}^h : \ \nabla\cdot\boldsymbol{v} = 0\} \not\subset \boldsymbol{X} := \{\boldsymbol{v} \in \boldsymbol{H}_0^1(\Omega) : \ \nabla\cdot\boldsymbol{v} = 0\}.$$

Then by Lemma 5.2 and (5.1), the discrete velocity solution is uniquely determined by the problem: Find $\boldsymbol{u}_h \in \boldsymbol{X}^h$ such that

$$a_h(\boldsymbol{u}_h, \boldsymbol{v}) := \int_{\Omega_h} \nu\nabla\boldsymbol{u}_h : \nabla\boldsymbol{v} = \int_{\Omega_h} \boldsymbol{f}_h\cdot\boldsymbol{v} \qquad \forall \boldsymbol{v} \in \boldsymbol{X}^h.$$

Standard theory of non–conforming and mixed finite element methods (e.g., [8]), along with Lemma 3.5 and Theorem 4.4 yields

$$\begin{aligned}\nu\|\nabla(\boldsymbol{u}-\boldsymbol{u}_h)\|_{L^2(\Omega_h)} &\le \inf_{\boldsymbol{w}\in\boldsymbol{X}^h}\nu\|\nabla(\boldsymbol{u}-\boldsymbol{w})\|_{L^2(\Omega_h)}+\sup_{\boldsymbol{v}\in\boldsymbol{X}^h\backslash\{0\}}\frac{a_h(\boldsymbol{u}_h-\boldsymbol{u},\boldsymbol{v})}{\|\nabla\boldsymbol{v}\|_{L^2(\Omega_h)}} \\ &\le C\inf_{\boldsymbol{w}\in\boldsymbol{V}^h}\nu\|\nabla(\boldsymbol{u}-\boldsymbol{w})\|_{L^2(\Omega_h)}+\sup_{\boldsymbol{v}\in\boldsymbol{X}^h\backslash\{0\}}\frac{a_h(\boldsymbol{u}_h-\boldsymbol{u},\boldsymbol{v})}{\|\nabla\boldsymbol{v}\|_{L^2(\Omega_h)}} \\ &\le Ch^2\nu\|\boldsymbol{u}\|_{H^3(\Omega)}+\sup_{\boldsymbol{v}\in\boldsymbol{X}^h\backslash\{0\}}\frac{a_h(\boldsymbol{u}_h-\boldsymbol{u},\boldsymbol{v})}{\|\nabla\boldsymbol{v}\|_{L^2(\Omega_h)}}.\end{aligned} \tag{5.2}$$

Recalling $-\nu\Delta\boldsymbol{u}+\nabla p=\boldsymbol{f}$ in $\mathbb{R}^2$ and $\boldsymbol{X}^h\subset\boldsymbol{H}_0(\mathrm{div};\Omega)$, we have

$$\begin{aligned}a_h(\boldsymbol{u}_h-\boldsymbol{u},\boldsymbol{v}) &= \int_{\Omega_h}\boldsymbol{f}\cdot\boldsymbol{v}-a_h(\boldsymbol{u},\boldsymbol{v})+\int_{\Omega_h}(\boldsymbol{f}_h-\boldsymbol{f})\cdot\boldsymbol{v} \\ &= -\int_{\Omega_h}\nu\Delta\boldsymbol{u}\cdot\boldsymbol{v}+\int_{\Omega_h}\nabla p\cdot\boldsymbol{v}-a_h(\boldsymbol{u},\boldsymbol{v})+\int_{\Omega_h}(\boldsymbol{f}_h-\boldsymbol{f})\cdot\boldsymbol{v} \\ &= -\int_{\Omega_h}\nu\Delta\boldsymbol{u}\cdot\boldsymbol{v}-a_h(\boldsymbol{u},\boldsymbol{v})+\int_{\Omega_h}(\boldsymbol{f}_h-\boldsymbol{f})\cdot\boldsymbol{v}\qquad\forall\boldsymbol{v}\in\boldsymbol{X}^h.\end{aligned}$$

We then integrate by parts to conclude

$$-\int_{\Omega_h}\nu\Delta\boldsymbol{u}\cdot\boldsymbol{v}-a_h(\boldsymbol{u},\boldsymbol{v})=\nu\sum_{e\in\mathcal{E}_h^I}\int_e\nabla\boldsymbol{u}:[\boldsymbol{v}],$$

where we used $\boldsymbol{v}$ is zero on $\partial\Omega_h$.

Recall $\mathcal{E}_h^{I,\partial}$ is the set of edges in $\mathcal{E}_h^I$ that have one endpoint on $\partial\Omega_h$. Then by properties of $\boldsymbol{V}^h$, there holds $[\boldsymbol{v}]|_e=0$ for all $e\in\mathcal{E}_h^I\backslash\mathcal{E}_h^{I,\partial}$, in particular,

$$a_h(\boldsymbol{u}_h-\boldsymbol{u},\boldsymbol{v})=\nu\sum_{e\in\mathcal{E}_h^{I,\partial}}\int_e\nabla\boldsymbol{u}:[\boldsymbol{v}]+\int_{\Omega_h}(\boldsymbol{f}_h-\boldsymbol{f})\cdot\boldsymbol{v}. \tag{5.3}$$

Lemma 5.3. *There holds*

$$\nu\sum_{e\in\mathcal{E}_h^{I,\partial}}\int_e\nabla\boldsymbol{u}:[\boldsymbol{v}]\le C\nu h^2\|\boldsymbol{u}\|_{H^3(\Omega)}\|\nabla\boldsymbol{v}\|_{L^2(\Omega_h)}\qquad\forall\boldsymbol{v}\in\boldsymbol{V}^h.$$

Proof. For each $e\in\mathcal{E}_h^{I,\partial}$, let $G_e\in\mathbb{R}^{2\times2}$ be the average of $\nabla\boldsymbol{u}$ on e. Standard interpolation estimates show

$$h_e^{-1}\|\nabla\boldsymbol{u}-G_e\|_{L^2(e)}^2\le C|\boldsymbol{u}|_{H^2(T)}\qquad h_e=\mathrm{diam}(e), \tag{5.4}$$

for T satisfying $e\subset\partial T$. Moreover, we clearly have $|G_e|\le C|\boldsymbol{u}|_{W^{1,\infty}(\Omega)}$.

Let $\boldsymbol{E}_h\boldsymbol{v}\in\boldsymbol{H}_0^1(\Omega_h)$ satisfy (4.1). Then $[\boldsymbol{E}_h\boldsymbol{v}]|_e=0$ for all $e\in\mathcal{E}_h^I$, and so,

$$\begin{aligned}\nu\sum_{e\in\mathcal{E}_h^{I,\partial}}\int_e\nabla\boldsymbol{u}:[\boldsymbol{v}] &= \nu\sum_{e\in\mathcal{E}_h^{I,\partial}}\Big(\int_e(\nabla\boldsymbol{u}-G_e):[\boldsymbol{v}-\boldsymbol{E}_h\boldsymbol{v}]+\int_e G_e:[\boldsymbol{v}]\Big) \\ &=: I_1+I_2.\end{aligned} \tag{5.5}$$

To bound I_1 we use the Cauchy–Schwarz inequality, (5.4), a trace inequality, and Lemma 4.5:

$$I_1\le\nu\Big(\sum_{e\in\mathcal{E}_h^{I,\partial}}h_e^{-1}\|\nabla\boldsymbol{u}-G_e\|_{L^2(e)}^2\Big)^{1/2}\Big(\sum_{e\in\mathcal{E}_h^{I,\partial}}h_e\|[\boldsymbol{v}-\boldsymbol{E}_h\boldsymbol{v}]\|_{L^2(e)}^2\Big)^{1/2} \tag{5.6}$$

$$\leq C\nu h^2 |\boldsymbol{u}|_{H^2(\Omega)} \|\nabla \boldsymbol{v}\|_{L^2(\Omega_h)}.$$

To bound I_2, we apply Lemma 4.6.

$$\begin{aligned} I_2 &= \nu \sum_{e\in\mathcal{E}_h^{I,\partial}} G_e : \int_e [\boldsymbol{v}] \qquad (5.7)\\ &\leq C\nu |\boldsymbol{u}|_{W^{1,\infty}(\Omega)} \Big(\sum_{e\in\mathcal{E}_h^{I,\partial}} h_e\Big)^{1/2} \Big(\sum_{e\in\mathcal{E}_h^{I,\partial}} h_e^{-1} \Big|\int_e [\boldsymbol{v}]\Big|^2\Big)^{1/2} \\ &\leq C\nu h^{5/2} \|\boldsymbol{u}\|_{W^{1,\infty}(\Omega)} \Big(\sum_{e\in\mathcal{E}_h^{I,\partial}} h_e\Big)^{1/2} \|\nabla \boldsymbol{v}\|_{L^2(\Omega_h)} \leq C\nu h^{5/2} \|\boldsymbol{u}\|_{H^3(\Omega)} \|\nabla \boldsymbol{v}\|_{L^2(\Omega_h)}. \end{aligned}$$

Combining (5.5)–(5.7) yields the desired result. □

Finally, we combine (5.2), (5.3), Lemma 5.3 to obtain the main result of the section.

THEOREM 5.4. *There holds*

$$\|\nabla(\boldsymbol{u}-\boldsymbol{u}_h)\|_{L^2(\Omega_h)} \leq C\big(h^2 \|\boldsymbol{u}\|_{H^3(\Omega)} + \nu^{-1} |\boldsymbol{f}-\boldsymbol{f}_h|_{X_h^*}\big), \qquad (5.8)$$

where

$$|\boldsymbol{f}-\boldsymbol{f}_h|_{X_h^*} = \sup_{\boldsymbol{v}\in \boldsymbol{X}_h\setminus\{0\}} \frac{\int_{\Omega_h} (\boldsymbol{f}-\boldsymbol{f}_h)\cdot \boldsymbol{v}}{\|\nabla \boldsymbol{v}\|_{L^2(\Omega_h)}}.$$

Therefore if, for example, $\boldsymbol{f}_h$ is the nodal quadratic interpolant of $\boldsymbol{f}$, and if $\boldsymbol{f}$ is sufficiently smooth, there holds

$$\|\nabla(\boldsymbol{u}-\boldsymbol{u}_h)\|_{L^2(\Omega_h)} \leq C\big(h^2 \|\boldsymbol{u}\|_{H^3(\Omega)} + \nu^{-1} h^3 \|\boldsymbol{f}\|_{H^3(\Omega)}\big).$$

The pressure approximation satisfies

$$\|p-p_h\|_{L^2(\Omega_h)} \leq C\big(\nu \|\nabla(\boldsymbol{u}-\boldsymbol{u}_h)\|_{L^2(\Omega_h)} + \nu h^2 \|\boldsymbol{u}\|_{H^3(\Omega)} + \inf_{q\in Q^h} \|p-q\|_{L^2(\Omega_h)} + \|\boldsymbol{f}-\boldsymbol{f}_h\|_{L^2(\Omega_h)}\big). \qquad (5.9)$$

Proof. The error estimate of the velocity follows from (5.2), (5.3), and Lemma 5.3; thus, it remains to prove (5.9).

For any $q \in Q^h$ and $\boldsymbol{v} \in \boldsymbol{V}^h$, we have by (5.1) and Lemma 5.3,

$$\begin{aligned} \int_{\Omega_h} (\nabla\cdot \boldsymbol{v}_h)(p_h - q) &= a_h(\boldsymbol{u}_h, \boldsymbol{v}) - \int_{\Omega_h} (\nabla\cdot\boldsymbol{v}) q - \int_{\Omega_h} \boldsymbol{f}_h \cdot \boldsymbol{v} \\ &= a_h(\boldsymbol{u}_h - \boldsymbol{u}, \boldsymbol{v}) - \int_{\Omega_h} (\nabla\cdot\boldsymbol{v})(q-p) - \int_{\Omega_h} (\boldsymbol{f}_h - \boldsymbol{f})\cdot \boldsymbol{v} + \nu \sum_{e\in\mathcal{E}_h^{I,\partial}} \int_e \nabla \boldsymbol{u} : [\boldsymbol{v}] \\ &\leq C\big(\nu \|\nabla(\boldsymbol{u}-\boldsymbol{u}_h)\|_{L^2(\Omega_h)} + \nu h^2 \|\boldsymbol{u}\|_{H^3(\Omega)} + \|p-q\|_{L^2(\Omega_h)}\big) \|\nabla \boldsymbol{v}\|_{L^2(\Omega_h)} + \|\boldsymbol{f}-\boldsymbol{f}_h\|_{L^2(\Omega_h)} \|\boldsymbol{v}\|_{L^2(\Omega_h)}. \end{aligned}$$

Using the estimate (4.1) and the Poincare inequality, we have

$$\|\boldsymbol{v}\|_{L^2(\Omega_h)} \leq \|\boldsymbol{E}_h \boldsymbol{v}\|_{L^2(\Omega_h)} + \|\boldsymbol{v} - \boldsymbol{E}_h \boldsymbol{v}\|_{L^2(\Omega_h)} \leq C \|\nabla \boldsymbol{v}\|_{L^2(\Omega_h)}.$$

Therefore

$$\begin{aligned} \int_{\Omega_h} (\nabla\cdot \boldsymbol{v}_h)(p_h - q) \leq C\big(&\nu \|\nabla(\boldsymbol{u}-\boldsymbol{u}_h)\|_{L^2(\Omega_h)} + \nu h^2 \|\boldsymbol{u}\|_{H^3(\Omega)} \\ &+ \|p-q\|_{L^2(\Omega_h)} + \|\boldsymbol{f}-\boldsymbol{f}_h\|_{L^2(\Omega_h)}\big) \|\nabla \boldsymbol{v}\|_{L^2(\Omega_h)}. \end{aligned}$$

We then use the inf-sup condition given in Theorem 4.4 to obtain

$$
\begin{aligned}
C\|p_h - q\|_{L^2(\Omega_h)} &\le \sup_{\boldsymbol{v}\in \boldsymbol{V}^h\setminus\{0\}} \frac{\int_{\Omega_h} (\nabla\cdot \boldsymbol{v})(p_h - q)}{\|\nabla \boldsymbol{v}\|_{L^2(\Omega_h)}}\\
&\le C\big(\nu\|\nabla(\boldsymbol{u}-\boldsymbol{u}_h)\|_{L^2(\Omega_h)} + \nu h^2\|\boldsymbol{u}\|_{H^3(\Omega)} + \|p-q\|_{L^2(\Omega_h)} + \|\boldsymbol{f}-\boldsymbol{f}_h\|_{L^2(\Omega_h)}\big).
\end{aligned}
$$

Applying the triangle inequality and taking the infimum over $q\in Q^h$, we obtain (5.9). ∎

6. A Pressure Robust Scheme. In this section, we construct a computable approximation $\boldsymbol{f}_h$ such that the term $\nu^{-1}|\boldsymbol{f}-\boldsymbol{f}_h|_{\boldsymbol{X}_h^*}$ appearing in estimate (5.8) is independent of the viscosity, in particular, such that the method is pressure robust. Essentially, this construction is done by applying a commuting operator to the function $\boldsymbol{f}|_\Omega$. In particular, we adopt and modify the recent results in [12] for Scott–Vogelius elements to construct commuting operators on meshes with curved boundary.

To discuss the main objections of this section further, we define the rot operator

$$
\operatorname{rot} \boldsymbol{v} = \frac{\partial v_2}{\partial x_1} - \frac{\partial v_1}{\partial x_2},
$$

and the corresponding Hilbert space

$$
\boldsymbol{H}(\mathrm{rot};\Omega_h) := \{\boldsymbol{v}\in \boldsymbol{L}^2(\Omega_h):\ \operatorname{rot}\boldsymbol{v}\in L^2(\Omega_h)\}.
$$

The main goal of this section is to prove the following result.

THEOREM 6.1. *There exists finite element spaces $\boldsymbol{W}_h\subset \boldsymbol{H}(\mathrm{rot};\Omega_h)$, $\Sigma_h\subset H_0^1(\Omega)$ with respect to the partition $\mathcal{T}_h$, and operators $\boldsymbol{\Pi}_W:\boldsymbol{H}^2(\Omega)\to \boldsymbol{W}_h$ and $\Pi_\Sigma: H^3(\Omega)\to\Sigma_h$ such that*

$$
\boldsymbol{\Pi}_W\nabla p = \nabla \Pi_\Sigma p \qquad \forall p\in H^3(\Omega). \tag{6.1}
$$

Moreover, there holds for any $\boldsymbol{f}\in \boldsymbol{H}^3(\Omega)$,

$$
\|\boldsymbol{f}-\boldsymbol{\Pi}_W\boldsymbol{f}\|_{L^2(\Omega_h)} \le Ch^2\|\boldsymbol{f}\|_{H^3(\Omega)}, \tag{6.2}
$$

where $\boldsymbol{f}$ in the left-hand side of the above inequality is an H^3 extension of $\boldsymbol{f}|_\Omega$.

COROLLARY 6.2. *Let $(\boldsymbol{u}_h,p_h)\in \boldsymbol{V}_h\times Q_h$ be the solution of the finite element method (5.1) with $\boldsymbol{f}_h = \boldsymbol{\Pi}_W\boldsymbol{f}$. Then there holds*

$$
\|\nabla(\boldsymbol{u}-\boldsymbol{u}_h)\|_{L^2(\Omega_h)} \le Ch^2\|\boldsymbol{u}\|_{H^5(\Omega)}.
$$

Proof. In light of estimate (5.8), it suffices to show $|\boldsymbol{f}-\boldsymbol{f}_h|_{X_h^*}\le C\nu h^2\|\boldsymbol{u}\|_{H^5(\Omega)}$.

Recall that the extension of $\boldsymbol{f}|_\Omega$ is given by $\boldsymbol{f} = -\nu\Delta\boldsymbol{u}+\nabla p$. Therefore by Theorem 6.1, for all $\boldsymbol{v}\in\boldsymbol{X}_h$,

$$
\begin{aligned}
\int_{\Omega_h} (\boldsymbol{f}-\boldsymbol{f}_h)\cdot\boldsymbol{v} &= \int_{\Omega_h} \big(-\nu(\Delta\boldsymbol{u}-\boldsymbol{\Pi}_W\Delta\boldsymbol{u}) + (\nabla p - \boldsymbol{\Pi}_W\nabla p)\big)\cdot\boldsymbol{v}\\
&= \int_{\Omega_h} \big(-\nu(\Delta\boldsymbol{u}-\boldsymbol{\Pi}_W\Delta\boldsymbol{u}) + \nabla(p - \Pi_\Sigma p)\big)\cdot\boldsymbol{v}\\
&= -\nu\int_{\Omega_h} (\Delta\boldsymbol{u}-\boldsymbol{\Pi}_W\Delta\boldsymbol{u})\cdot\boldsymbol{v},
\end{aligned}
$$

where we used that $\nabla\cdot\boldsymbol{v}=0$ and $\boldsymbol{v}\cdot\boldsymbol{n}|_{\partial\Omega_h}=0$. Consequently,

$$
|\boldsymbol{f}-\boldsymbol{f}_h|_{X_h^*} \le C\nu\|\Delta\boldsymbol{u}-\boldsymbol{\Pi}_W\Delta\boldsymbol{u}\|_{L^2(\Omega_h)} \le Ch^2\nu\|\Delta\boldsymbol{u}\|_{H^3(\Omega)} \le C\nu h^2\|\boldsymbol{u}\|_{H^5(\Omega)}.
$$

∎

6.1. Proof of Theorem 6.1: Preliminaries. As a first step of the proof of Theorem 6.1, we "rotate" the space $\boldsymbol{V}(T)$.

Definition 6.3. *We define*

$$\begin{aligned}\boldsymbol{W}(T) &= \{\boldsymbol{v}\in \boldsymbol{H}^1(T):\ \boldsymbol{v}(x) = (DF_T(\hat{x}))^{-\intercal}\hat{\boldsymbol{v}}(\hat{x}),\ \exists \hat{\boldsymbol{v}}\in\hat{\boldsymbol{V}}\},\\ \boldsymbol{W}_0(T) &= \boldsymbol{W}(T)\cap \boldsymbol{H}_0^1(T).\end{aligned}$$

Remark 6.4. *Define*

$$S = \begin{pmatrix} 0 & -1\\ 1 & 0\end{pmatrix},$$

so that $\operatorname{rot}(S\boldsymbol{v}) = \nabla\cdot\boldsymbol{v}$, *and* $SDF_TS^{-1} = \det(DF_T)(DF_T)^{-\intercal}$. *Therefore, if* $\boldsymbol{v}(x) = (DF_T(\hat{x}))^{-\intercal}\hat{\boldsymbol{v}}(\hat{x})$, *we have*

$$\operatorname{rot}\boldsymbol{v}(x) = \operatorname{rot}\Big(S\frac{DF_T(\hat{x})S^{-1}\hat{\boldsymbol{v}}(\hat{x})}{\det(DF_T(\hat{x}))}\Big) = \nabla\cdot\Big(\frac{DF_T(\hat{x})S^{-1}\hat{\boldsymbol{v}}(\hat{x})}{\det(DF_T(\hat{x}))}\Big) = \frac{\hat{\operatorname{rot}}\,\hat{\boldsymbol{v}}(\hat{x})}{\det(DF_T(\hat{x}))}.$$

Remark 6.5. *Note that* $\hat{\operatorname{rot}}:\hat{\boldsymbol{V}}\to\hat{Q}$ *is a surjection. Indeed, let* $\hat{q}\in\hat{Q}$. *Then there exists* $\hat{\boldsymbol{v}}\in\boldsymbol{V}$ *such that* $\hat{\nabla}\cdot\hat{\boldsymbol{v}} = \hat{q}$. *Then set* $\hat{\boldsymbol{w}} = S\hat{\boldsymbol{v}}$ *so that* $\hat{q} = \hat{\nabla}\cdot\hat{\boldsymbol{v}} = \hat{\operatorname{rot}}\,\hat{\boldsymbol{w}}$. *Similar arguments show* $\hat{\operatorname{rot}}:\hat{\boldsymbol{V}}_0\to\hat{Q}_0$ *is a bijection.*

Lemma 6.6. *Let* $\{\hat{\alpha}_i\}_{i=1}^3, \{\hat{m}_i\}_{i=1}^3\subset\mathcal{N}_{\hat{T}}$ *be, respectively, the vertices and edge midpoints of* $\hat{T}$. *Set* $\alpha_i = F_T(\hat{\alpha}_i)$ *and* $m_i = F_T(\hat{m}_i)$ *to be the corresponding points on* T. *Any* $\boldsymbol{v}\in\boldsymbol{W}(T)$ *is uniquely determined by the values*

$$\boldsymbol{v}(\alpha_i),\ (\boldsymbol{v}\cdot\boldsymbol{n})(m_i)\qquad i=1,2,3, \tag{6.3a}$$

$$\int_e \boldsymbol{v}\cdot\boldsymbol{t}\qquad \forall\ \text{edges of } T, \tag{6.3b}$$

$$\int_T(\operatorname{rot}\boldsymbol{v})q\qquad \forall q\in Q_0(T). \tag{6.3c}$$

Proof. Write $\boldsymbol{v}(x) = DF_T^{-\intercal}\hat{\boldsymbol{v}}$ for some $\hat{\boldsymbol{v}}\in\hat{\boldsymbol{V}}$, and suppose that $\boldsymbol{v}$ vanishes on the DOFs. We show $\hat{\boldsymbol{v}}\equiv 0$.

We clearly have $\hat{\boldsymbol{v}}(\hat{\alpha}_i) = 0$ for $i=1,2,3$, and by using the relation $\boldsymbol{t} = DF_T\hat{\boldsymbol{t}}/|DF_T\hat{\boldsymbol{t}}|$ [19], and a change of variables, we have

$$0 = \int_e\boldsymbol{v}\cdot\boldsymbol{t} = \int_{\hat{e}}\frac{(DF_T^{-\intercal}\hat{\boldsymbol{v}})\cdot(DF_T\hat{\boldsymbol{t}})}{|DF_T\hat{\boldsymbol{t}}|}|\det(DF_T)||DF^{-\intercal}\hat{\boldsymbol{n}}| = \int_{\hat{e}}\hat{\boldsymbol{v}}\cdot\hat{\boldsymbol{t}},$$

where we used the identity $|\det(DF_T)||DF^{-\intercal}\hat{\boldsymbol{n}}| = |DF_T\hat{\boldsymbol{t}}|$. Thus, we conclude $\hat{\boldsymbol{v}}\cdot\hat{\boldsymbol{t}}|_{\partial\hat{T}} = 0$.

Similarly, using the relation $\boldsymbol{n} = DF_T^{-\intercal}\hat{\boldsymbol{n}}/|DF_T^{-\intercal}\hat{\boldsymbol{n}}|$, we compute

$$0 = (\boldsymbol{v}\cdot\boldsymbol{n})(m_i) = \frac{\hat{\boldsymbol{v}}\cdot(DF_T^{-1}DF_T^{-\intercal}\hat{\boldsymbol{n}})}{|DF_T^{-\intercal}\hat{\boldsymbol{n}}|}(\hat{m}_i).$$

Because $(DF_T^{-1}DF_T^{-\intercal}\hat{\boldsymbol{n}})\cdot\hat{\boldsymbol{n}} = |DF_T^{-\intercal}\hat{\boldsymbol{n}}|^2\neq 0$, we conclude $(DF_T^{-1}DF_T^{-\intercal}\hat{\boldsymbol{n}})$ is not tangent to $\hat{\boldsymbol{t}}$. Thus, since $\hat{\boldsymbol{v}}\cdot\hat{\boldsymbol{t}}|_{\partial\hat{T}} = 0$, we get $\hat{\boldsymbol{v}}|_{\partial\hat{T}} = 0$, i.e., $\hat{\boldsymbol{v}}\in\hat{\boldsymbol{V}}_0$.

Now let $\hat{q}\in\hat{Q}_0$, and set $q(x) = \hat{q}(\hat{x})$ so that $q\in Q_0(T)$. Using $\operatorname{rot}\boldsymbol{v} = \hat{\operatorname{rot}}\,\hat{\boldsymbol{v}}/\det(DF_T)$, we have by a change of variables,

$$\int_T(\operatorname{rot}\boldsymbol{v})q = \int_{\hat{T}}(\hat{\operatorname{rot}}\,\hat{\boldsymbol{v}})\hat{q}.$$

Taking $\hat{q} = \hat{\text{rot}}\, \hat{\boldsymbol{v}}$, we conclude $\hat{\text{rot}}\, \hat{\boldsymbol{v}} = 0$. This implies $\hat{\boldsymbol{v}} \equiv 0$, and therefore $\boldsymbol{v} \equiv 0$. ◻

Next, we define the local Clough-Tocher space on the reference element

$$\hat{\Sigma} = \{\hat{\sigma} \in H^2(\hat{T}) :\ \hat{\sigma}|_{\hat{K}} \in \mathcal{P}_3(\hat{K})\ \forall \hat{K} \in \hat{T}^{ct}\}.$$

It is known that the dimension of $\hat{\Sigma}$ is 12 [10], and any $\hat{\sigma} \in \hat{\Sigma}$ is uniquely determined by the valuess

$$\hat{\nabla}\hat{\sigma}(\hat{\alpha}_i),\ \hat{\sigma}(\hat{\alpha}_i),\ (\hat{\nabla}\hat{\sigma} \cdot \hat{\boldsymbol{n}})(\hat{m}_i) \qquad i = 1, 2, 3. \tag{6.4}$$

We define the Clough–Tocher space on T via composition

$$\Sigma(T) = \{\sigma :\ \sigma(x) = \hat{\sigma}(\hat{x}),\ \exists \hat{\sigma} \in \hat{\Sigma}\}.$$

It is easy to see $\Sigma(T) \subset H^2(T)$. In the following lemma, we extend the above DOFs to $\Sigma(T)$.

Lemma 6.7. *A function $\sigma \in \Sigma(T)$ is uniquely determined by the values*

$$\nabla\sigma(\alpha_i),\ \sigma(\alpha_i),\ (\nabla\sigma \cdot \boldsymbol{n})(m_i) \qquad i = 1, 2, 3. \tag{6.5a}$$

Proof. Write $\sigma(x) = \hat{\sigma}(\hat{x})$ with $\hat{\sigma} \in \hat{\Sigma}$. It suffices to show that if σ vanishes at the above DOFs, then $\hat{\sigma}$ vanishes on (6.4).

If σ vanishes at the above DOFs, then clearly

$$\hat{\nabla}\hat{\sigma}(\hat{\alpha}_i) = 0,\ \hat{\sigma}(\hat{\alpha}_i) = 0 \qquad i = 1, 2, 3.$$

This implies $\hat{\sigma}|_{\partial\hat{T}} = 0$, and therefore $\hat{\nabla}\hat{\sigma} \cdot \hat{\boldsymbol{t}}|_{\partial\hat{T}} = 0$.

Next, by the chain rule and the relation $\boldsymbol{n} = DF^{-\intercal}\hat{\boldsymbol{n}}/|DF^{-\intercal}\hat{\boldsymbol{n}}|$,

$$0 = (\nabla\sigma \cdot \boldsymbol{n})(m_i) = \Big(\frac{1}{|DF_T^{-\intercal}\hat{\boldsymbol{n}}|}\hat{\nabla}\hat{\sigma} \cdot (DF_T^{-1}DF_T^{-\intercal}\hat{\boldsymbol{n}})\Big)(\hat{m}_i).$$

Thus, we have $\big(\hat{\nabla}\hat{\sigma} \cdot (DF_T^{-1}DF_T^{-\intercal}\hat{\boldsymbol{n}})\big)(\hat{m}_i) = 0$. Since

$$\big((DF_T^{-1}DF_T^{-\intercal}\hat{\boldsymbol{n}}) \cdot \hat{\boldsymbol{n}}\big)(\hat{m}_i) = |(DF_T\hat{\boldsymbol{n}})(\hat{m}_i)|^2 \neq 0,$$

the vector $(DF_T^{-1}DF_T^{-\intercal}\hat{\boldsymbol{n}})(\hat{m}_i)$ is not tangent to $\hat{e}$. Because the tangental derivative of $\hat{\sigma}$ vanishes at $\hat{m}_i$, we conclude $\hat{\nabla}\hat{\sigma}(\hat{m}_i) = 0$. Thus, $\hat{\sigma} \equiv 0$ and $\sigma \equiv 0$. ◻

Remark 6.8. *Note that if $\sigma \in \Sigma(T)$ with $\sigma(x) = \hat{\sigma}(\hat{x})$, then $\nabla\sigma(x) = (DF_T(\hat{x}))^{-\intercal}\hat{\nabla}\hat{\sigma}(\hat{x})$. We conclude $\nabla\sigma \in \boldsymbol{W}(T)$.*

As a next step, we use the DOFs stated in Lemmas 6.6–6.7 to construct commuting operators with properties stated in Theorem 6.1. Note that an added difficulty of the construction is that the operators are defined for functions with domain Ω, but map to functions with domain Ω_h. To mitigate this mismatch, we employ the mapping $G : \tilde{\Omega}_h \to \Omega$ given in Section 2.

For each $T \in \mathcal{T}_h$ and edge e in $\mathcal{T}_h$, we set

$$T_R := G(G_h^{-1}(T)) \subset \Omega, \qquad e_R := G(G_h^{-1}(e)) \subset \bar{\Omega},$$

where we recall G_h is the quadratic interpolant of G. That is, T_R is obtained by first mapping T to its associated affine element $\tilde{T} = G_h^{-1}(T) \in \tilde{\mathcal{T}}_h$, and then mapping $\tilde{T}$ to $G(\tilde{T}) \subset \Omega$. By properties of the quadratic interpolant G_h, we have $G(G_h^{-1}(\alpha_i)) = \alpha_i$ and $G(G_h^{-1}(m_i)) = m_i$ for all vertices and edge midpoints of T.

Via Lemmas 6.6–6.7 we introduce the operator $\boldsymbol{\Pi}_W^T : \boldsymbol{H}^2(T_R) \to \boldsymbol{W}(T)$ uniquely determined by the conditions

$$(\boldsymbol{\Pi}_W^T\boldsymbol{v})(\alpha_i) = \boldsymbol{v}(\alpha_i) \qquad i = 1, 2, 3, \tag{6.6a}$$

$$(\mathbf{\Pi}_W^T \boldsymbol{v}\cdot\boldsymbol{n})(m_i) = (\boldsymbol{v}\cdot\boldsymbol{n})(m_i) \qquad i=1,2,3, \tag{6.6b}$$

$$\int_e (\mathbf{\Pi}_W^T\boldsymbol{v})\cdot\boldsymbol{t} = \int_{e_R}\boldsymbol{v}\cdot\boldsymbol{t}_{e_R} \qquad \forall \text{ edges of } T, \tag{6.6c}$$

$$\int_T (\operatorname{rot}\mathbf{\Pi}_W^T\boldsymbol{v})q = \int_{T\cap T_R}(\operatorname{rot}\boldsymbol{v})q \qquad \forall q\in Q_0(T), \tag{6.6d}$$

where $\boldsymbol{n}$ is the outward unit normal with respect to $e\subset\partial T$, $\boldsymbol{t}$ is the unit tangent of $e\subset\partial T$, and $\boldsymbol{t}_{e_R}$ is the unit tangent of $e_R\subset\partial T_R$. We also set $\Pi_\Sigma^T : H^3(T_S)\to\Sigma(T)$ uniquely determined by

$$\Pi_\Sigma^T\sigma(\alpha_i) = \sigma(\alpha_i), \quad \nabla(\Pi_\Sigma\sigma)(\alpha_i) = \nabla\sigma(a), \qquad i=1,2,3, \tag{6.7a}$$

$$\nabla(\Pi_\Sigma^T\sigma)(m_i)\cdot\boldsymbol{n}(m_i) = \nabla\sigma(m_i)\cdot\boldsymbol{n}(m_i) \qquad i=1,2,3. \tag{6.7b}$$

We define the global spaces

$$\boldsymbol{W}^h = \{\boldsymbol{v}\in\boldsymbol{H}(\operatorname{rot};\Omega_h) : \ \boldsymbol{v}|_T\in\boldsymbol{W}(T)\ \forall T\in\mathfrak{T}_h,\ \boldsymbol{v} \text{ is continuous on (6.3)}\},$$
$$\Sigma^h = \{\sigma\in H^1(\Omega_h) : \ \sigma|_T\in\Sigma(T)\ \forall T\in\mathfrak{T}_h,\ \sigma \text{ is continuous on (6.5)}\},$$

and the operators $\mathbf{\Pi}_W : H^2(\Omega)\to\boldsymbol{W}^h$, $\Pi_\Sigma : H^3(\Omega)\to\Sigma^h$ by

$$\mathbf{\Pi}_W\boldsymbol{v}|_T = \mathbf{\Pi}_W^T\boldsymbol{v}, \qquad \Pi_\Sigma\sigma|_T = \Pi_\Sigma^T\sigma, \qquad \forall T\in\mathfrak{T}_h.$$

We now prove that these operators satisfy (6.1)–(6.2).

6.2. Proof of (6.1)**.** For given $p\in H^3(\Omega)$, set $\boldsymbol{\rho} = \mathbf{\Pi}_W\nabla p - \nabla\Pi_\Sigma p\in\boldsymbol{W}(T)$. We wish to show $\boldsymbol{\rho}\equiv 0$. This this end, it suffices to show $\boldsymbol{\rho}$ vanishes at the DOFs in Lemma 6.6 for each $T\in\mathfrak{T}_h$.

First, we consider the interior DOFs of $\boldsymbol{W}(T)$. Using (6.6d) and the identity $\operatorname{rot}\nabla p = 0$, we have

$$\int_T(\operatorname{rot}\boldsymbol{\rho})q = \int_T(\operatorname{rot}(\mathbf{\Pi}_W\nabla p))q = \int_{T\cap T_R}(\operatorname{rot}(\nabla p))q = 0 \qquad \forall q\in Q_0(T).$$

Let α_i be a vertex of T. We then have by (6.6a) and (6.7a),

$$\boldsymbol{\rho}(\alpha_i) = \mathbf{\Pi}_W\nabla p(\alpha_i) - \nabla\Pi_\Sigma p(\alpha_i) = 0.$$

Next, let m_i be an edge midpoint of T and let $\boldsymbol{n}$ be the outward unit normal at m_i. Then by (6.6b) and (6.7b),

$$\boldsymbol{\rho}(m_i)\cdot\boldsymbol{n} = \mathbf{\Pi}_W\nabla p(m_i)\cdot\boldsymbol{n} - \nabla\Pi_\Sigma p(m_i)\cdot\boldsymbol{n} = 0.$$

Finally, let $e\subset\partial T$ be an edge of T with endpoints α_2 and α_1. Recalling that e_R also has endpoints α_2 and α_1, we use (6.7a) and (6.6c) to obtain

$$\begin{aligned}\int_e\boldsymbol{\rho}\cdot\boldsymbol{t} = \int_e\big(\mathbf{\Pi}_W\nabla p - \nabla\Pi_\Sigma p\big)\cdot\boldsymbol{t} &= \int_{e_R}\nabla p\cdot\boldsymbol{t}_{e_R} - \int_e(\nabla\Pi_\Sigma p)\cdot\boldsymbol{t}\\ &= p(\alpha_2)-p(\alpha_1) - \big((\Pi_\Sigma p)(\alpha_2) - (\Pi_\Sigma p)(\alpha_1)\big) = 0.\end{aligned}$$

Thus, $\boldsymbol{\rho}$ vanishes at all the DOFs in Lemma 6.6, and we conclude $\boldsymbol{\rho}\equiv 0$.

6.3. Proof of (6.2)**.** We break up the proof of estimate (6.2) into three parts.

(i) We extend $\boldsymbol{f}$ to $\mathbb{R}^2$ such that $\|\boldsymbol{f}\|_{H^3(\mathbb{R})}\le C\|\boldsymbol{f}\|_{H^3(\Omega)}$ With this extension, we define $\boldsymbol{I}_W\boldsymbol{f}\in\boldsymbol{W}(T)$ uniquely by the conditions

$$(\boldsymbol{I}_W\boldsymbol{f})(\alpha_i) = \boldsymbol{f}(\alpha_i), \quad (\boldsymbol{I}_W\boldsymbol{f}\cdot\boldsymbol{n})(m_i) = (\boldsymbol{f}\cdot\boldsymbol{n})(m_i) \qquad i=1,2,3,$$

$$\int_e(\boldsymbol{I}_W\boldsymbol{f})\cdot\boldsymbol{t} = \int_e\boldsymbol{f}\cdot\boldsymbol{t} \qquad \forall \text{ edges of } T,$$

$$\int_T(\operatorname{rot}\boldsymbol{I}_W\boldsymbol{f})q = \int_T(\operatorname{rot}\boldsymbol{f})q \qquad \forall q\in Q_0(T).$$

(ii) We now estimate $\|\boldsymbol{f} - \boldsymbol{I}_W \boldsymbol{f}\|_{L^2(T)}$. For notational convenience, we write $\boldsymbol{v} = \boldsymbol{I}_W \boldsymbol{f}$, and set

$$\boldsymbol{v}(x) = R_T(\hat{x})\hat{\boldsymbol{v}}(\hat{x}), \qquad \boldsymbol{f}(x) = R_T(\hat{x})\hat{\boldsymbol{f}}(\hat{x}),$$

with $R_T(\hat{x}) = (DF_T(\hat{x}))^{-\intercal}$. We then have

$$\hat{\boldsymbol{v}}(\hat{\alpha}_i) = \hat{\boldsymbol{f}}(\hat{\alpha}_i), \quad (\hat{\boldsymbol{v}} \cdot (R_T^\intercal \boldsymbol{n}))(\hat{m}_i) = (\hat{\boldsymbol{f}} \cdot (R_T^\intercal \boldsymbol{n}))(\hat{m}_i) \qquad i = 1,2,3. \tag{6.8}$$

We also have, by a change of variables (cf. proof of Lemma 6.6)

$$\int_{\hat{e}} \hat{\boldsymbol{v}} \cdot \hat{\boldsymbol{t}} = \int_e \boldsymbol{v} \cdot \boldsymbol{t} = \int_e \boldsymbol{f} \cdot \boldsymbol{t} = \int_{\hat{e}} \hat{\boldsymbol{f}} \cdot \hat{\boldsymbol{t}}. \tag{6.9}$$

Next, for $q \in Q_0(T)$, write $q(x) = \hat{q}(\hat{x})$ with $\hat{q} \in \hat{Q}$. We then have

$$\int_{\hat{T}} (\hat{\text{rot}}\hat{\boldsymbol{v}})\hat{q} = \int_{\hat{T}} (\det(DF_T)\text{rot}\boldsymbol{v}) \circ F_T \hat{q} = \int_T \text{rot}\boldsymbol{v} q = \int_T \text{rot}\boldsymbol{f} q = \int_{\hat{T}} \hat{\text{rot}}\hat{\boldsymbol{f}}\hat{q}. \tag{6.10}$$

It follows from (6.8)–(6.10) and a slight generalization of the Bramble–Hilbert lemma that

$$\|\hat{\boldsymbol{f}} - \hat{\boldsymbol{v}}\|_{L^2(\hat{T})} \le C|\hat{\boldsymbol{f}}|_{H^3(\hat{T})}. \tag{6.11}$$

Therefore by (2.1), Lemma 2.5 and (6.11) (and noting $R_T^{-1} = DF_T^\intercal$),

$$\begin{aligned} \|\boldsymbol{f} - \boldsymbol{I}_W \boldsymbol{f}\|_{L^2(T)} &\le C h_T \|R_T(\hat{\boldsymbol{f}} - \hat{\boldsymbol{v}})\|_{L^2(\hat{T})} \\ &\le C|\hat{\boldsymbol{f}}|_{H^3(\hat{T})} = C|R_T^{-1} R_T \hat{\boldsymbol{f}}|_{H^3(\hat{T})} \\ &\le C\big(\|R_T^{-1}\|_{L^\infty(\hat{T})} |R_T \hat{\boldsymbol{f}}|_{H^3(\hat{T})} + |R_T^{-1}|_{W^{1,\infty}(\hat{T})} |R_T \hat{\boldsymbol{f}}|_{H^2(\hat{T})}\big) \\ &\le C h_T^3 \|\boldsymbol{f}\|_{H^3(T)}. \end{aligned} \tag{6.12}$$

(iii) We now estimate $(\boldsymbol{\Pi}_W \boldsymbol{f} - \boldsymbol{I}_W \boldsymbol{f})|_T \in \boldsymbol{W}(T)$. Set $\boldsymbol{w} = \boldsymbol{\Pi}_W \boldsymbol{f} - \boldsymbol{I}_W \boldsymbol{f} \in \boldsymbol{W}(T)$. Then

$$\begin{aligned} &\boldsymbol{w}(\alpha_i) = 0, \quad (\boldsymbol{w} \cdot \boldsymbol{n})(m_i) = 0 && i = 1,2,3, \\ &\int_e \boldsymbol{w} \cdot \boldsymbol{t} = \int_{e_R} \boldsymbol{f} \cdot \boldsymbol{t}_{e_R} - \int_e \boldsymbol{f} \cdot \boldsymbol{t} && \forall \text{ edges of } T, \\ &\int_T (\text{rot}\, \boldsymbol{w}) q = \int_{T \cap T_R} (\text{rot}\, \boldsymbol{f}) q - \int_T (\text{rot}\, \boldsymbol{f}) q && \forall q \in Q_0(T). \end{aligned}$$

Write $\boldsymbol{w}(x) = R_T(\hat{x})\hat{\boldsymbol{w}}(\hat{x})$. By equivalence of norms, we have

$$\begin{aligned} \|\hat{\boldsymbol{w}}\|^2_{H^m(\hat{T})} &\le C\Big(\sum_{i=1}^3 (|\hat{\boldsymbol{w}}(\hat{\alpha}_i)|^2 + |\hat{\boldsymbol{w}}(\hat{m}_i)|^2) + \sup_{\substack{\hat{q} \in \hat{Q}_0 \\ \|\hat{q}\|_{L^2(\hat{T})} = 1}} \Big|\int_{\hat{T}} (\hat{\text{rot}}\, \hat{\boldsymbol{w}})\hat{q}\Big|^2\Big) \\ &= C\Big(\sum_{i=1}^3 |\hat{\boldsymbol{w}}(\hat{m}_i)|^2 + \sup_{\substack{\hat{q} \in \hat{Q}_0 \\ \|\hat{q}\|_{L^2(\hat{T})} = 1}} \Big|\int_{\hat{T}} (\hat{\text{rot}}\, \hat{\boldsymbol{w}})\hat{q}\Big|^2\Big). \end{aligned} \tag{6.13}$$

Next we use the algebraic identity

$$\hat{\boldsymbol{w}}(\hat{m}_i) = \frac{1}{\boldsymbol{\alpha}^\perp \cdot \boldsymbol{\beta}} \Big(((\hat{\boldsymbol{w}}(\hat{m}_i)) \cdot \boldsymbol{\beta})\boldsymbol{\alpha}^\perp - ((\hat{\boldsymbol{w}}(\hat{m}_i)) \cdot \boldsymbol{\alpha})\boldsymbol{\beta}^\perp \Big) \tag{6.14}$$

for any linearly independent vectors $\boldsymbol{\alpha}, \boldsymbol{\beta} \in \mathbb{R}^2$. Here, $\boldsymbol{\alpha}^\perp = S\boldsymbol{\alpha}$. We take $\boldsymbol{\alpha} = -\hat{\boldsymbol{t}}(\hat{m}_i)$ and $\boldsymbol{\beta} = R_T^\intercal(\hat{m}_i)\boldsymbol{n}(m_i)$, so that

$$|\boldsymbol{\alpha}^\perp \cdot \boldsymbol{\beta}| = |S\hat{\boldsymbol{t}}(\hat{m}_i) \cdot (R_T^\intercal(\hat{m}_i)\boldsymbol{n}(m_i))| = |(R_T(\hat{m}_i)\hat{\boldsymbol{n}}(\hat{m}_i)) \cdot \boldsymbol{n}(m_i)| = |(R_T \hat{\boldsymbol{n}})(\hat{m}_i)|,$$

where we used the relation $\boldsymbol{n} = R_T \hat{\boldsymbol{n}}/|R_T\hat{\boldsymbol{n}}|$ in the last equality. We use (6.14) and the identity $\hat{\boldsymbol{w}}(\hat{m}_i)\cdot\boldsymbol{\beta} = (\boldsymbol{w}\cdot\boldsymbol{n})(m_i) = 0$ to conclude

$$|\hat{\boldsymbol{w}}(\hat{m}_i)| = \frac{1}{|R_T\hat{\boldsymbol{n}}(\hat{m}_i)|}\big|(\hat{\boldsymbol{w}}\cdot\hat{\boldsymbol{t}})(\hat{m}_i)SR_T^{\intercal}(\hat{m}_i)\boldsymbol{n}(m_i)\big| \le \frac{|R_T(\hat{m}_i)|}{|R_T\hat{\boldsymbol{n}}(\hat{m}_i)|}|(\hat{\boldsymbol{w}}\cdot\boldsymbol{t})(\hat{m}_i)| \le C|(\hat{\boldsymbol{w}}\cdot\boldsymbol{t})(\hat{m}_i)|,$$

where $C>0$ is the condition number of R_T (which is independent of h).
We use this estimate in (6.13) to conclude

$$\|\hat{\boldsymbol{w}}\|^2_{H^m(\hat{T})} \le C\Big(\sum_{i=1}^3 |(\hat{\boldsymbol{w}}\cdot\hat{\boldsymbol{t}})(\hat{m}_i)|^2 + \sup_{\substack{\hat{q}\in\hat{Q}_0\\ \|\hat{q}\|_{L^2(\hat{T})}=1}} \Big|\int_{\hat{T}}(\hat{\operatorname{rot}}\,\hat{\boldsymbol{w}})\hat{q}\Big|^2\Big).$$

Using Simpson's rule, noting that $\hat{\boldsymbol{w}}$ vanishes on the vertices of $\hat{T}$, we obtain

$$\|\hat{\boldsymbol{w}}\|^2_{H^m(\hat{T})} \le C\Big(\sum_{\hat{e}\subset\partial\hat{T}} \Big|\int_{\hat{e}}\hat{\boldsymbol{w}}\cdot\hat{\boldsymbol{t}}\Big|^2 + \sup_{\substack{\hat{q}\in\hat{Q}_0\\ \|\hat{q}\|_{L^2(\hat{T})}=1}} \Big|\int_{\hat{T}}(\hat{\operatorname{rot}}\,\hat{\boldsymbol{w}})\hat{q}\Big|^2\Big). \tag{6.15}$$

We now estimate the two terms on the right-hand side of (6.15) separately.
First by a change of variables, we have

$$\int_{\hat{e}}\hat{\boldsymbol{w}}\cdot\hat{\boldsymbol{t}} = \int_e \boldsymbol{w}\cdot\boldsymbol{t} = \int_{e_R}\boldsymbol{f}\cdot\boldsymbol{t}_{e_R} - \int_e \boldsymbol{f}\cdot\boldsymbol{t}.$$

Set $\Theta := G_h\circ G^{-1}$ so that $e = \Theta(e_R)$ and $T = \Theta(T_R)$. There holds [17, Proposition 3]

$$|\Theta(x)-x| = \mathcal{O}(h_T^3),\quad |D\Theta - I_2| = \mathcal{O}(h_T^2),\quad \boldsymbol{t}(\Theta(x)) = \frac{D\Theta\boldsymbol{t}_{e_R}}{|D\Theta\boldsymbol{t}_{e_R}|}(x),\qquad x\in\bar{T}_R, \tag{6.16}$$

and therefore by a change of variables,

$$\int_e \boldsymbol{f}\cdot\boldsymbol{t} = \int_{e_R}|(D\Theta)\boldsymbol{t}_{e_R}|(\boldsymbol{f}\cdot\boldsymbol{t})\circ\Theta = \int_{e_R}(\boldsymbol{f}\circ\Theta)\cdot(D\Theta\boldsymbol{t}_{e_R}).$$

Thus,

$$\begin{aligned}\int_{\hat{e}}\hat{\boldsymbol{w}}\cdot\hat{\boldsymbol{t}} &= \int_{e_R}\big(\boldsymbol{f}\cdot\boldsymbol{t}_{e_R} - (\boldsymbol{f}\circ\Theta)\cdot(D\Theta\boldsymbol{t}_{e_R})\big)\\ &= \int_{e_R}\big(\boldsymbol{f} - (\boldsymbol{f}\circ\Theta)\big)\cdot\boldsymbol{t}_{e_R} - (\boldsymbol{f}\circ\Theta)\cdot(D\Theta\boldsymbol{t}_{e_R} - \boldsymbol{t}_{e_R})\big),\end{aligned}$$

and therefore by (6.16), Taylor's Theorem, and a Sobolev embedding,

$$\Big|\int_{\hat{e}}\hat{\boldsymbol{w}}\cdot\hat{\boldsymbol{t}}\Big| \le C\big(h_T^4|\boldsymbol{f}|_{W^{1,\infty}(\mathbb{R}^2)} + h_T^3\|\boldsymbol{f}\|_{L^\infty(\mathbb{R}^2)}\big) \le Ch_T^3\|\boldsymbol{f}\|_{H^3(\Omega)}. \tag{6.17}$$

Next we let $\hat{q}\in\hat{Q}_0$ with $\|\hat{q}\|_{L^2(\hat{T})} = 1$ and compute

$$\int_{\hat{T}}(\hat{\operatorname{rot}}\,\hat{\boldsymbol{w}})\hat{q} = \int_T \operatorname{rot}\boldsymbol{w}q = \int_{T\cap T_R}(\operatorname{rot}\boldsymbol{f})q - \int_T(\operatorname{rot}\boldsymbol{f})q = \int_{T\setminus T_R}(\operatorname{rot}\boldsymbol{f})q,$$

where $q\in Q_0(T)$ with $q(x) = \hat{q}(\hat{x})$. Using $\|q\|_{L^2(T)} \le Ch_T\|\hat{q}\|_{L^2(\hat{T})} \le Ch_T$, we obtain

$$\int_{\hat{T}}(\hat{\operatorname{rot}}\,\hat{\boldsymbol{w}})\hat{q} \le |T\setminus T_R|\|\operatorname{rot}\boldsymbol{f}\|_{L^\infty(\mathbb{R}^2)}\|q\|_{L^2(T)} \le Ch_T^3\|\boldsymbol{f}\|_{H^3(\Omega)}\|q\|_{L^2(T)} \le Ch_T^4\|\boldsymbol{f}\|_{H^3(\Omega)}. \tag{6.18}$$

Applying estimates (6.17)–(6.18) to (6.15) yields

$$\|\hat{\boldsymbol{w}}\|_{H^m(\hat{T})} \le C h_T^3 \|\boldsymbol{f}\|_{H^3(\Omega)}.$$

Therefore

$$\|\boldsymbol{\Pi}_W \boldsymbol{f} - \boldsymbol{I}_W \boldsymbol{f}\|_{L^2(T)} = \|\boldsymbol{w}\|_{L^2(T)} \le C h_T \|R_T \hat{\boldsymbol{w}}\|_{L^2(\hat{T})} \le C \|\hat{\boldsymbol{w}}\|_{L^2(\hat{T})} \le C h_T^3 \|\boldsymbol{f}\|_{H^3(\Omega)}.$$

Finally by (6.12) and the triangle inequality,

$$\|\boldsymbol{f} - \boldsymbol{\Pi}_W \boldsymbol{f}\|_{L^2(T)} \le C h_T^3 \|\boldsymbol{f}\|_{H^3(\Omega)}.$$

Summing over $T \in \mathcal{T}_h$ yields the estimate (6.2):

$$\|\boldsymbol{f} - \boldsymbol{\Pi}_W \boldsymbol{f}\|_{L^2(\Omega_h)} \le C \Big(\sum_{T \in \mathcal{T}_h} h_T^6 \|\boldsymbol{f}\|_{H^3(\Omega)}^2 \Big)^{1/2} \le C h^2 \|\boldsymbol{f}\|_{H^3(\Omega)} \Big(\sum_{T \in \mathcal{T}_h} h_T^2 \Big)^{1/2} \le C h^2 \|\boldsymbol{f}\|_{H^3(\Omega)}.$$

7. Numerical Experiments. In this section we perform a simple set of numerical experiments and compare the results with the theory established in the previous section. We let $\Omega = B_1(0) \subset \mathbb{R}^2$ be the unit ball, and take the data such that the exact solution is given by

$$\boldsymbol{u} = \begin{pmatrix} (x_1^2 + x_2^2 - 1)(8x_1^2 x_2 + x_1^2 + 5x_2^2 - 1) \\ -4x_1(x_1^2 + x_2^2 - 1)(3x_1^2 + x_2^2 + x_2 - 1) \end{pmatrix}, \qquad p = 10(x_1^2 + x_2^2 - \frac{1}{2}).$$

We compute the finite element method (5.1), taking the source approximation $\boldsymbol{f}_h$ to be the quadratic (nodal) Lagrange interpolant of $\boldsymbol{f}$, and the viscosity $\nu = 10^{-1}$. The errors for a decreasing sequence of mesh parameters h are depicted in Figure 7.1–7.2. For comparison, we also plot the errors of the analogous Scott-Vogelius finite element method using affine approximations, i.e., method (5.1) with $\boldsymbol{V}^h \times Q^h$ replaced by $\tilde{\boldsymbol{V}}^h \times \tilde{Q}^h$. The Figure shows the asymptotic convergence rates

$$\|\boldsymbol{u} - \boldsymbol{u}_h\|_{L^2(\Omega_h)} = \mathcal{O}(h^3), \quad \|\nabla(\boldsymbol{u} - \boldsymbol{u}_h)\|_{L^2(\Omega_h)} = \mathcal{O}(h^2), \quad \|p - p_h\|_{L^2(\Omega_h)} = \mathcal{O}(h^2),$$

for the isoparametric approximations. These results agree with the theoretical results stated in Theorem 5.4. In contrast, the numerics indicate the solution of the affine approximation, denoted by $(\boldsymbol{u}^{aff}, p_h^{aff}) \in \tilde{\boldsymbol{V}}^h \times \tilde{Q}^h$ satisfies the sub-optimal convergence rates

$$\|\boldsymbol{u} - \boldsymbol{u}_h^{aff}\|_{L^2(\tilde{\Omega}_h)} = \mathcal{O}(h^2), \quad \|\nabla(\boldsymbol{u} - \boldsymbol{u}_h^{aff})\|_{L^2(\tilde{\Omega}_h)} = \mathcal{O}(h^{3/2}), \quad \|p - p_h^{aff}\|_{L^2(\tilde{\Omega}_h)} = \mathcal{O}(h^{3/2}).$$

We also solve the finite element method (5.1) but with isoparametric spaces defined via the usual composition, i.e., with velocity-pressure pair (1.1). Numerical experiments indicate the method is stable and converges with optimal order. However, as Figure 7.2 shows, the method is not divergence–free (nor pressure robust).

REFERENCES

[1] D. N. Arnold and J. Qin, *Quadratic velocity/linear pressure Stokes elements*, In R. Vichnevetsky, D. Knight, and G. Richter, editors, Advances in Computer Methods for Partial Differential Equations–VII, pages 28–34. IMACS, 1992.

[2] H.-O. Bae and D. W. Kim, *Finite element approximations for the Stokes equations on curved domains, and their errors*, Appl. Math. Comput., 148(3):823–847, 2004.

[3] L. Beirão da Veiga, C. Lovadina, and G. Vacca, *Divergence free virtual elements for the Stokes problem on polygonal meshes*, ESAIM Math. Model. Numer. Anal., 51(2):509–535, 2017.

[4] C. Bernardi and G. Raugel, *Analysis of Some Finite Elements for the Stokes Problem*, Math. Comp., 44(169):71–79, 1985.

[5] C. Bernardi, *Optimal finite-element interpolation on curved domains*, SIAM J. Numer. Anal., 26(5):1212–1240, 1989.

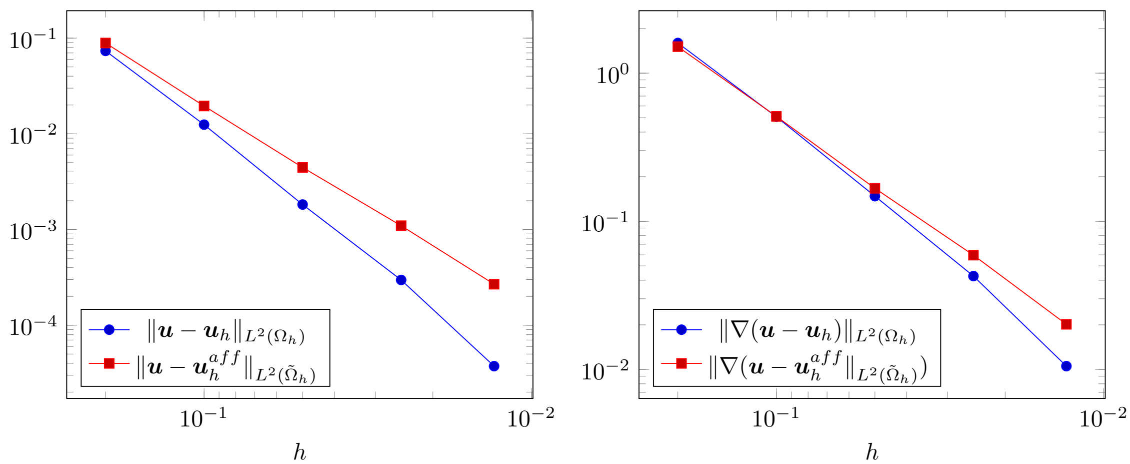

FIGURE 7.1. *Velocity errors of the isoparametric Scott-Vogelius finite element method* (5.1) *(blue) and the affine Scott-Vogelius method (red) for decreasing values of mesh parameter h.*

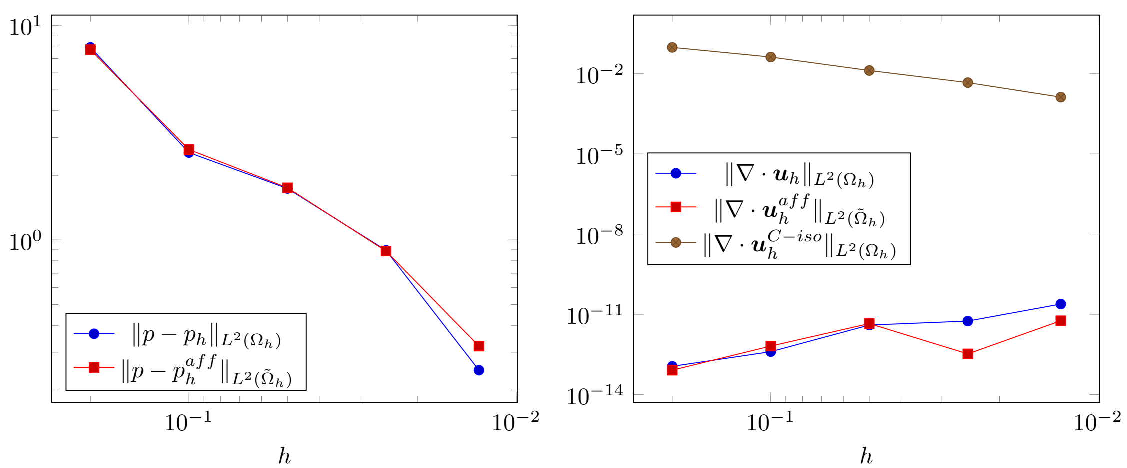

FIGURE 7.2. *Left: Pressure errors of the isoparametric Scott-Vogelius finite element method* (5.1) *(blue) and the affine Scott-Vogelius method (red) for decreasing values of mesh parameter h. Right: Divergence errors of the isoparametric Scott-Vogelius finite element method (blue), the affine Scott-Vogelius method (red), and the isoparametric Scott-Vogelius using the standard composition of isoparametric mappings (brown).*

[6] C. Bernardi, M. Costabel, M. Dauge, Monique, and V. Girault, *Continuity properties of the inf-sup constant for the divergence*, SIAM J. Math. Anal., 48(2):1250–1271, 2016.

[7] D. Boffi, F. Brezzi, L. F. Demkowicz, R. G. Durán, R. S. Falk, and M. Fortin, *Mixed finite elements, compatibility conditions, and applications,* Lectures given at the C.I.M.E. Summer School held in Cetraro, June 26–July 1, 2006. Edited by Boffi and Lucia Gastaldi. Lecture Notes in Mathematics, 1939. Springer-Verlag, Berlin; Fondazione C.I.M.E., Florence, 2008.

[8] S.C. Brenner and L.R. Scott, *The Mathematical Theory of Finite Element Methods (Third edition)*, Springer, 2008.

[9] P.G. Ciarlet and P.-A. Raviart, *Interpolation theory over curved elements, with applications to finite element methods*, Comput. Methods Appl. Mech. Engrg., 1:217–249, 1972.

[10] P. G. Ciarlet, *The Finite Element Method for Elliptic Problems*, North-Holland, Amsterdam, 1978.

[11] I. Dione and J.M. Urquiza, *Penalty: finite element approximation of Stokes equations with slip boundary conditions*, Numer. Math., 129(3):587–610, 2015.

[12] G. Fu, J. Guzmán, and M. Neilan, *Exact smooth piecewise polynomial sequences on Alfeld splits*, Math. Comp., 89(323):1059–1091, 2020.

[13] J. Guzmán, and M. Neilan, *inf-sup stable finite elements on barycentric refinements producing divergence-free approximations in arbitrary dimensions*, SIAM J. Numer. Anal., 56(5):2826–2844, 2018.

[14] V. John, Volker, A. Linke, C. Merdon, M. Neilan, and L.G. Rebholz, *On the divergence constraint in mixed finite*

element methods for incompressible flows, SIAM Rev., 59(3):492–544, 2017.
[15] G. Kanschat and N. Sharma *Divergence-conforming discontinuous Galerkin methods and C^0 interior penalty methods*, SIAM J. Numer. Anal., 52(4):1822–1842, 2014.
[16] T. Kato, M. Mitrea, G. Ponce, and M. Taylor, *Extension and representation of divergence–free vector fields on bounded domains*, Math. Research Letters, 7:643–650, 2000.
[17] M. Lenoir, *Optimal isoparametric finite elements and error estimates for domains involving curved boundaries*, SIAM J. Numer. Anal., 23(3):562–580, 1986.
[18] A. Linke, *On the role of the Helmholtz decomposition in mixed methods for incompressible flows and a new variational crime*, Comput. Methods Appl. Mech. Engrg., 268:782–800, 2014.
[19] P. Monk, *Finite element methods for Maxwell's equations*, Numerical Mathematics and Scientific Computation, Oxford University Press, New York, 2003.
[20] L. R. Scott, *Finite–element techniques for curved boundaries*, Ph.D. thesis, Massachusetts Institute of Technology, 1973.
[21] R. Verfürth, *Finite element approximation of incompressible Navier-Stokes equations with slip boundary condition*, Numer. Math., 50(6):697–721, 1987.
[22] S. Zhang, *A new family of stable mixed finite elements for the 3D Stokes equations*, Math. Comp., 74(250):543–554, 2005.
[23] Zlámal, Miloš, *Curved elements in the finite element method. I*, SIAM J. Numer. Anal., 10:229–240, 1973.

Appendix A. Proofs of Preliminary Results.

A.1. Proof of Lemma 2.3. *Proof.* For notational simplicity, we set $\hat{g}(\hat{x}) = \det(DF_T(\hat{x}))$. Using the fact that $DF_T \to \det(DF_T)$ is quadratic in two dimensions and the estimates (2.1), a simple calculation shows $|\hat{g}|_{W^{m,\infty}(\hat{T})} \le Ch_T^{2+m}$. Consequently, by the quotient rule, for any multi-index α with $|\alpha| = m$,

$$\begin{aligned}\Big|\frac{\partial^m}{\partial\hat{x}^\alpha}\frac{1}{\hat{g}}\Big| &\le C \sum_{|\beta^{(1)}|+|\beta^{(2)}|+\cdots+|\beta^{(m)}|=m} \frac{|\partial^{|\beta^{(1)}|}\hat{g}/\partial\hat{x}^{\beta^{(1)}}|\cdots|\partial^{|\beta^{(m)}|}\hat{g}/\partial\hat{x}^{\beta^{(m)}}|}{|\hat{g}^{m+1}|}\\ &\le C \sum_{|\beta^{(1)}|+|\beta^{(2)}|+\cdots+|\beta^{(m)}|=m} \frac{(h_T^{2+|\beta^{(1)}|})\cdots(h_T^{2+|\beta^{(m)}|})}{|\hat{g}^{m+1}|} \le C\frac{h_T^{3m}}{|\hat{g}^{m+1}|} \le Ch_T^{m-2},\end{aligned}$$

where we used (2.1) in the last inequality.

We then use the product rule and (2.1) to find, for any $i,j \in \{1,2\}$ and multi-index α with $|\alpha| = m$,

$$\begin{aligned}\Big|\frac{\partial^m (A_T)_{i,j}}{\partial\hat{x}^\alpha}\Big| &= \Big|\frac{\partial^m}{\partial\hat{x}^\alpha}\big((DF_T)_{i,j}/\hat{g}\big)\Big|\\ &\le C\sum_{|\beta|+|\gamma|=m}\big|\partial^\beta (DF_T)_{i,j}/\partial^{|\beta|}\hat{x}\big|\big|\partial^\gamma \hat{g}^{-1}/\partial^{|\gamma|}\hat{x}\big|\\ &\le C\sum_{|\beta|+|\gamma|=m}\big(h_T^{1+|\beta|}\big)\big(h_T^{|\gamma|-2}\big) \le Ch_T^{m-1}.\end{aligned}$$

This establishes the first inequality in (2.3).

Next, we use the identity $A_T^{-1} = \det(DF_T)(DF_T)^{-1} = \operatorname{adj}(DF_T)$, the adjugate matrix of DF_T. Because the entries of DF_T and $\operatorname{adj}(DF_T)$ are the same up to permutation and sign in two dimensions, we have by (2.1),

$$|A_T^{-1}|_{W^{m,\infty}(\hat{T})} = |DF_T|_{W^{m+1,\infty}(\hat{T})} \le \begin{cases} Ch_T^{1+m} & m = 0,1\\ 0 & m \ge 2\end{cases}$$

∎

A.2. Proof of Lemma 2.4. *Proof.* Let $\hat{\boldsymbol{t}}$ be the unit tangent vector of $\hat{e}$ obtained by rotating $\hat{\boldsymbol{n}}$ 90 degrees clockwise. Then a calculation shows

$$\det(DF_T(\hat{x}))(DF_T(\hat{x}))^{-\intercal}\hat{\boldsymbol{n}} = \begin{pmatrix} -(DF_T(\hat{x})\hat{\boldsymbol{t}})_2\\ (DF_T(\hat{x})\hat{\boldsymbol{t}})_1\end{pmatrix}.$$

Because F_T restricted to $\hat{e}$ is affine, $(DF_T(\hat{x})\hat{\boldsymbol{t}})$ is constant on $\hat{e}$. This proves the lemma. ∎