MATHEMATICS OF COMPUTATION
VOLUME 55, NUMBER 192
OCTOBER 1990, PAGES 411-437

A NONCONFORMING MULTIGRID METHOD FOR THE STATIONARY STOKES EQUATIONS

SUSANNE C. BRENNER

ABSTRACT. An optimal-order W-cycle multigrid method for solving the stationary Stokes equations is developed, using P1 nonconforming divergence-free finite elements.

1. INTRODUCTION

Let Ω be a bounded convex polygonal domain in $\mathbb{R}^2$. The stationary Stokes equations for an incompressible viscous fluid are given by

$$\begin{aligned} -\Delta \mathbf{u} + \mathbf{grad}\, p &= \mathbf{f} \quad \text{in } \Omega, \\ \operatorname{div} \mathbf{u} &= 0 \quad \text{in } \Omega, \\ \mathbf{u} &= \mathbf{0} \quad \text{on } \partial\Omega. \end{aligned} \tag{1.1}$$

Here the viscosity constant is taken to be 1, p is the pressure, $\mathbf{u} = (u_1, u_2)$ is the velocity of the fluid, and $\mathbf{f} = (f_1, f_2)$ denotes the body force. In this paper, vectors are always represented by boldfaced letters. We assume $\mathbf{f} \in (L^2(\Omega))^2$. There exist a unique solution $(\mathbf{u}, p) \in ((H_0^1(\Omega))^2 \cap (H^2(\Omega))^2) \times (H^1(\Omega)/\mathbb{R})$ of (1.1) and a positive constant C_Ω such that

$$\|\mathbf{u}\|_{(H^2(\Omega))^2} + |p|_{H^1(\Omega)} \leq C_\Omega \|\mathbf{f}\|_{(L^2(\Omega))^2} \tag{1.2}$$

(cf. [11, 13]).

In this paper we will use the following notation for the Sobolev norms and seminorms:

$$\|\mathbf{v}\|_{(H^m(\Omega))^2} := \left(\int_\Omega \sum_{|\alpha| \leq m} |\partial^\alpha \mathbf{v}|^2 \, dx \right)^{1/2}$$

and

$$|\mathbf{v}|_{(H^m(\Omega))^2} := \left(\int_\Omega \sum_{|\alpha| = m} |\partial^\alpha \mathbf{v}|^2 \, dx \right)^{1/2}.$$

Similar notation is also used for scalar functions.

Received March 24, 1989.
1980 *Mathematics Subject Classification* (1985 *Revision*). Primary 65N30, 65F10.
Key words and phrases. Nonconforming multigrid method, stationary Stokes equations.
This work was supported in part by the National Science Foundation under Grant No. DMS-8904911.

A weak form of (1.1) is to find a divergence-free $\mathbf{u}$ in $(H_0^1(\Omega))^2$ such that

$$a(\mathbf{u}, \mathbf{v}) + \int_\Omega \mathbf{grad}\, p \cdot \mathbf{v} = \int_\Omega \mathbf{f} \cdot \mathbf{v}\, dx \quad \forall \mathbf{v} \in (H_0^1(\Omega))^2, \tag{1.3}$$

where

$$a(\mathbf{v}_1, \mathbf{v}_2) := \int_\Omega \nabla \mathbf{v}_1 \cdot \nabla \mathbf{v}_2\, dx, \tag{1.4}$$

and $\nabla \mathbf{v}_1 \cdot \nabla \mathbf{v}_2 = \sum_{i=1}^2 \nabla v_{1,i} \cdot \nabla v_{2,i}$ for $\mathbf{v}_1 = (v_{1,1}, v_{1,2})$ and $\mathbf{v}_2 = (v_{2,1}, v_{2,2})$ in $(H_0^1(\Omega))^2$.

Let $V = \{\mathbf{v} : \mathbf{v} \in (H_0^1(\Omega))^2,\ \operatorname{div} \mathbf{v} = 0\}$. If we restrict (1.3) to V, the pressure term disappears and the problem becomes to find $\mathbf{u} \in V$ such that

$$a(\mathbf{u}, \mathbf{v}) = \int_\Omega \mathbf{f} \cdot \mathbf{v}\, dx \quad \forall \mathbf{v} \in V. \tag{1.5}$$

The velocity $\mathbf{u}$ can be characterized as the unique solution of (1.5) (cf. [10]).

In order to apply the Ritz-Galerkin method to equation (1.5), we introduce a family of triangulations of Ω: $\{\mathscr{T}^k\}_{k=1}^\infty$, where $\mathscr{T}^{k+1}$ is obtained by connecting the midpoints of the edges of the triangles in $\mathscr{T}^k$. We will denote $\max\{\operatorname{diam} T : T \in \mathscr{T}^k\}$ by h_k.

The finite element spaces V_k are defined as follows:

$$\begin{aligned} V_k := \{\mathbf{v}|_T &\text{ is linear and divergence-free for all } T \in \mathscr{T}^k, \\ &\mathbf{v} \text{ is continuous at the midpoints of interelement boundaries}, \\ &\qquad \text{and } \mathbf{v} = \mathbf{0} \text{ at the midpoints of } \mathscr{T}^k \text{ along } \partial\Omega\}. \end{aligned} \tag{1.6}$$

Note that V_k is nonconforming because $V_k \not\subset V$.

On $V_k + V$ we define the following positive definite symmetric bilinear form,

$$a_k(\mathbf{v}_1, \mathbf{v}_2) := \sum_{T \in \mathscr{T}^k} \int_T \nabla \mathbf{v}_1 \cdot \nabla \mathbf{v}_2\, dx, \tag{1.7}$$

and its associated nonconforming energy norm

$$\|\mathbf{v}\|_k := \sqrt{a_k(\mathbf{v}, \mathbf{v})}. \tag{1.8}$$

The discretized problem for (1.5) is to find $\mathbf{u}_k \in V_k$ such that

$$a_k(\mathbf{u}_k, \mathbf{v}) = \int_\Omega \mathbf{f} \cdot \mathbf{v}\, dx \quad \forall \mathbf{v} \in V_k. \tag{1.9}$$

It is proved in [10] that there exists a positive constant C such that

$$\|\mathbf{u} - \mathbf{u}_k\|_{(L^2(\Omega))^2} + h_k \|\mathbf{u} - \mathbf{u}_k\|_k \le C h_k^2 (|\mathbf{u}|_{(H^2(\Omega))^2} + |p|_{H^1(\Omega)}). \tag{1.10}$$

Throughout this paper, C (with or without subscripts) denotes a positive constant independent of the mesh parameter k.

We will develop an optimal-order multigrid method for (1.9). Let n_k be the dimension of V_k. Our full multigrid algorithm will yield an approximate solution $\hat{\mathbf{u}}_k$ to (1.9) in $\mathscr{O}(n_k)$ steps such that

$$\|\mathbf{u}_k - \hat{\mathbf{u}}_k\|_{(L^2(\Omega))^2} + h_k\|\mathbf{u}_k - \hat{\mathbf{u}}_k\|_k \le Ch_k^2(|\mathbf{u}|_{(H^2(\Omega))^2} + |p|_{H^1(\Omega)}). \tag{1.11}$$

For background information on multigrid methods, we refer the reader to [12, 14] and the references therein.

The crucial part in the development of a nonconforming multigrid method is the correct choice of an intergrid transfer operator $I_{k-1}^k\colon V_{k-1} \to V_k$. (Since $V_{k-1} \not\subset V_k$, natural injection no longer works.) The intergrid transfer operator we use is defined by averaging and has the following three properties:

$$\|I_{k-1}^k \mathbf{v}\|_k \le C\|\mathbf{v}\|_{k-1} \quad \forall \mathbf{v} \in V_{k-1}, \tag{1.12}$$

$$\|I_{k-1}^k \mathbf{v} - \mathbf{v}\|_{(L^2(\Omega))^2} \le Ch_k\|\mathbf{v}\|_{k-1} \quad \forall \mathbf{v} \in V_{k-1}, \tag{1.13}$$

and

$$\begin{aligned} &\|I_{k-1}^k(\Pi_{k-1}\mathbf{g}) - \Pi_k\mathbf{g}\|_{(L^2(\Omega))^2} + h_k\|I_{k-1}^k(\Pi_{k-1}\mathbf{g}) - \Pi_k\mathbf{g}\|_k \\ &\quad \le Ch_k^2|\mathbf{g}|_{(H^2(\Omega))^2} \quad \forall \mathbf{g} \in (H^2(\Omega))^2 \cap (H_0^1(\Omega))^2, \end{aligned} \tag{1.14}$$

where Π_k denotes an interpolation operator from V onto V_k (cf. §2). These three estimates will play an important role in our convergence analysis. Analogous estimates have been used for other nonconforming finite elements (cf. [7, 8]). We also refer the interested reader to other related results in nonconforming and nonnested multigrid methods in [3, 5, 18].

This paper is organized as follows. We review some facts about the finite element space V_k in §2. In §3 we define the intergrid transfer operator and prove the three estimates (1.12)–(1.14). The multigrid algorithm is described in §4. In §5 we discuss the mesh-dependent norms, which is followed by the convergence analysis in §6.

2. The divergence-free P1 nonconforming finite element space

Let P be a simply connected polygonal domain and $\mathscr{T}$ be a triangulation of P. Denote $\max\{\operatorname{diam} T\colon T \in \mathscr{T}\}$ by h. Let

$$\begin{aligned} W := \{\mathbf{w} \in (L^2(\Omega))^2 \colon\ & \mathbf{w}|_T \text{ is linear and divergence-free for all } T \in \mathscr{T}, \\ & \mathbf{w} \text{ is continuous at the midpoints} \\ & \text{of interelement boundaries, and} \\ & \mathbf{w} = \mathbf{0} \text{ at the midpoints of } \mathscr{T} \text{ along } \partial P\}. \end{aligned} \tag{2.1}$$

We will describe a basis of W. First we make an observation on the divergence-free condition. Let $\mathbf{w}$ be a linear function on a triangle T with midpoints m_1, m_2, and m_3 on edges e_1, e_2, and e_3 (cf. Figure 1).

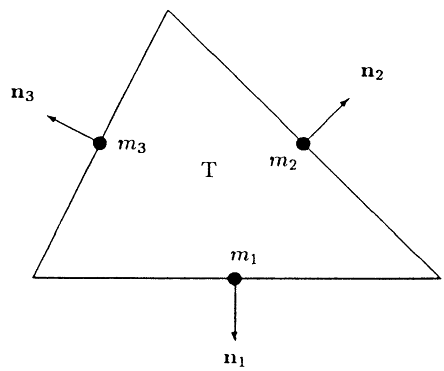

FIGURE 1

Then

$$(2.2)\qquad \begin{aligned} \operatorname{div}\mathbf{w}=0 &\Leftrightarrow \int_T \operatorname{div}\mathbf{w}\,dx=0\\ &\Leftrightarrow \int_{\partial T}\mathbf{w}\cdot\mathbf{n}\,ds=0 \Leftrightarrow \sum_{i=1}^{3}(\mathbf{w}(m_i)\cdot\mathbf{n}_i)|e_i|=0, \end{aligned}$$

where $\mathbf{n}_i$ denotes the outer normal to edge e_i.

Let e be an edge in $\mathscr{T}$. Denote by ϕ_e the piecewise linear function on P that takes the value 1 at the midpoint of the edge e and 0 at all other midpoints.

The first kind of basis functions are associated with internal edges. Let $\mathbf{w}_e := \phi_e\mathbf{t}_e$, where e is an internal edge and $\mathbf{t}_e$ is a unit vector tangential to e. Then it follows from (2.2) that $\mathbf{w}_e \in W$.

The second kind of basis functions are associated with internal vertices. Let p be an internal vertex and let $e_1, e_2, \dots, e_l$ be the edges in $\mathscr{T}$ that have p as an endpoint. Let $\mathbf{w}_p := \sum_{i=1}^{l} |e_i|^{-1}\phi_{e_i}\mathbf{n}_{e_i}$, where $\mathbf{n}_{e_i}$ is a unit vector normal to e_i pointing in the counterclockwise direction (cf. Figure 2). It again follows from (2.2) that $\mathbf{w}_p \in W$.

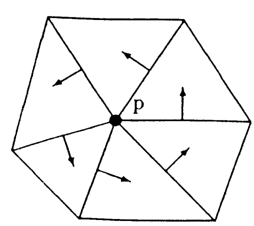

FIGURE 2

The proof of the following lemma can be found in Appendix 3 of [17].

Lemma 1. *The set of vector functions* $\{\mathbf{w}_e : e$ *is an internal edge of* $\mathcal{T}\} \cup \{\mathbf{w}_p : p$ *is an internal vertex of* $\mathcal{T}\}$ *is a basis of* W. *In particular,*

$$\dim W = e^I + v^I, \tag{2.3}$$

where e^I *denotes the number of internal edges and* v^I *denotes the number of internal vertices.*

We can apply (2.3) to derive an exact formula for the dimension n_k of the finite element space V_k in (1.6). Let e_k^I be the number of internal edges in $\mathcal{T}^k$. Denote by f_k the number of triangles in $\mathcal{T}^k$. Then e_k^I and f_k satisfy the difference equations

$$e_k^I = 2e_{k-1}^I + 3f_{k-1}, \qquad f_k = 4f_{k-1}. \tag{2.4}$$

Equation (2.3) and Euler's formula imply that

$$n_k = 2e_k^I - f_k + 1. \tag{2.5}$$

If we solve (2.4) and substitute the solution into (2.5), we obtain

$$n_k = 2^k(e_1^I - \tfrac{3}{2}f_1) + 2f_1 4^{k-1} + 1. \tag{2.6}$$

Therefore, asymptotically,

$$n_k \sim 2f_1 4^{k-1}. \tag{2.7}$$

Henceforth, we will use the following set of vector functions as the standard basis for V_k:

$$\{\mathbf{v}_e^k : e \text{ is an internal edge of } \mathcal{T}^k\} \cup \{\mathbf{v}_p^k : p \text{ is an internal vertex of } \mathcal{T}^k\}. \tag{2.8}$$

Let $Z := \{\mathbf{z} \in (L^2(\Omega))^2 : \mathbf{z}|_T$ is linear for all $T \in \mathcal{T}$, $\mathbf{z}$ is continuous at the midpoints of interelement boundaries, and $\mathbf{z} = \mathbf{0}$ at the midpoints of $\partial P\}$.

The interpolation operator Π: $(H^2(P))^2 \cap (H_0^1(P))^2 \to Z$ is defined by (cf. [10])

$$\Pi\mathbf{g} \in Z \quad \text{and} \quad \int_e \Pi\mathbf{g}\, ds = \int_e \mathbf{g}\, ds \quad \text{for all edges } e \in \mathcal{T}. \tag{2.9}$$

More explicitly, we have

$$\Pi\mathbf{g}(m_e) = \frac{1}{|e|}\int_e \mathbf{g}\, ds, \tag{2.10}$$

where m_e is the midpoint of the edge e.

The following lemma is proved in [10].

Lemma 2. *Let* $\mathbf{g} \in (H^2(P))^2 \cap (H_0^1(P))^2$. *Then*

$$\int_T \operatorname{div}(\Pi\mathbf{g}|_T)\, dx = \int_T \operatorname{div}\mathbf{g}\, dx \quad \forall T \in \mathcal{T}, \tag{2.11}$$

and there exists a positive constant C *which depends only on the angles of the triangles in* $\mathcal{T}$ *such that*

$$\|\mathbf{g}-\Pi\mathbf{g}\|_{(L^2(P))^2}+h\left(\sum_{T\in\mathcal{T}^k}|\mathbf{g}-\Pi\mathbf{g}|^2_{(H^1(T))^2}\right)^{1/2}\le Ch^2|\mathbf{g}|_{(H^2(P))^2}. \tag{2.12}$$

As a corollary to Lemma 2, Π: $\{\mathbf{g}: \mathbf{g}\in(H^2(P))^2\cap(H_0^1(P))^2$ and $\operatorname{div}\mathbf{g}=0\}\to W$. If we apply this result to Ω, V, and V_k, there exists a sequence of interpolation operators $\Pi_k: V\to V_k$ such that

$$\|\mathbf{g}-\Pi_k\mathbf{g}\|_{(L^2(\Omega))^2}+h_k\|\mathbf{g}-\Pi_k\mathbf{g}\|_k\le Ch_k^2|\mathbf{g}|_{(H^2(\Omega))^2}. \tag{2.13}$$

3. The intergrid transfer operator I_{k-1}^k

In [6, 8], we described the construction of an intergrid transfer operator for the scalar P1 nonconforming finite element. The construction here is similar, except that special care must be taken to preserve the divergence-free condition.

Let $\mathbf{v}\in V_{k-1}$. To define the piecewise linear vector function $I_{k-1}^k\mathbf{v}$, it suffices to specify its values at the midpoints of $\mathcal{T}^k$. If $m\in\partial\Omega$, then $(I_{k-1}^k\mathbf{v})(m)=0$. If m lies in the interior of Ω, then there are two cases to consider. For a midpoint m of $\mathcal{T}^k$ that lies on the common edge of two triangles T_1 and T_2 of $\mathcal{T}^{k-1}$ (e.g., $m_1,\dots,m_6$ in Figure 3), we define

$$(I_{k-1}^k\mathbf{v})(m):=\tfrac{1}{2}[\mathbf{v}|_{T_1}(m)+\mathbf{v}|_{T_2}(m)].$$

If a midpoint m lies in the interior of a triangle T in $\mathcal{T}^{k-1}$ (e.g., m_7, m_8, and m_9 in Figure 3), then the tangential component of $(I_{k-1}^k\mathbf{v})(m)$ is the same as the tangential component of $\mathbf{v}(m)$, and the normal component will be determined by the condition that $\operatorname{div}(I_{k-1}^k\mathbf{v})=0$ on the three outer triangles in the subdivision of T. In other words, if we denote by e_i the edge in Figure 3 that has m_i as its midpoint, then $(I_{k-1}^k\mathbf{v})(m_i)\cdot\mathbf{n}_i$, $i=7,8,9$, are determined by the following equations:

$$\begin{aligned}\sum_{i=6,1,7}(I_{k-1}^k\mathbf{v})(m_i)\cdot\mathbf{n}_i|e_i|&=0,\\ \sum_{i=2,3,8}(I_{k-1}^k\mathbf{v})(m_i)\cdot\mathbf{n}_i|e_i|&=0,\\ \sum_{i=4,5,9}(I_{k-1}^k\mathbf{v})(m_i)\cdot\mathbf{n}_i|e_i|&=0.\end{aligned} \tag{3.1}$$

Proposition 1. *The intergrid transfer operator* I_{k-1}^k *maps* V_{k-1} *into* V_k, *i.e.,*

$$I_{k-1}^k\mathbf{v}\in V_k\quad\forall\mathbf{v}\in V_{k-1}. \tag{3.2}$$

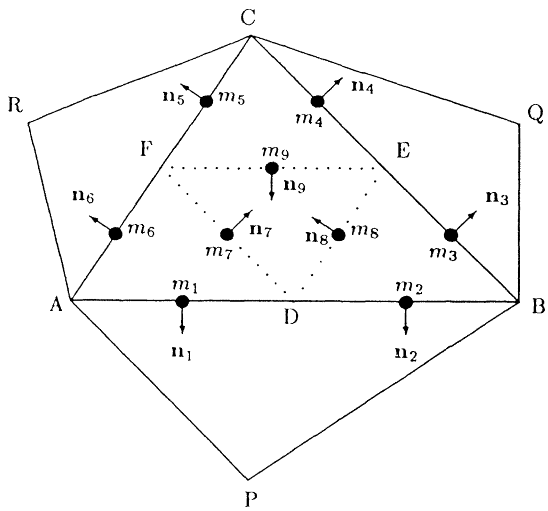

FIGURE 3

Proof. It suffices to check the divergence-free condition on $\triangle DEF$ in Figure 3. (By construction, $\operatorname{div}(I_{k-1}^k \mathbf{v}) = 0$ on $\triangle ADF$, $\triangle DBE$, and $\triangle FEC$.)

Let $\mathbf{w} = I_{k-1}^k \mathbf{v}$. We want to show that

$$\sum_{i=7}^{9} \mathbf{w}(m_i) \cdot \mathbf{n}_i |e_i| = 0. \tag{3.3}$$

In view of (3.1), equation (3.3) follows from

$$\sum_{i=1}^{6} \mathbf{w}(m_i) \cdot \mathbf{n}_i |e_i| = 0. \tag{3.4}$$

Let $\hat{\mathbf{v}} = \mathbf{v}|_{\triangle ABC}$. Since $\operatorname{div} \hat{\mathbf{v}} = 0$, we have

$$\begin{aligned} &\sum_{i=1,6,7} \hat{\mathbf{v}}(m_i) \cdot \mathbf{n}_i |e_i| = 0, \qquad &&\sum_{i=2,3,8} \hat{\mathbf{v}}(m_i) \cdot \mathbf{n}_i |e_i| = 0, \\ &\sum_{i=4,5,9} \hat{\mathbf{v}}(m_i) \cdot \mathbf{n}_i |e_i| = 0, &&\sum_{i=7,8,9} \hat{\mathbf{v}}(m_i) \cdot \mathbf{n}_i |e_i| = 0. \end{aligned} \tag{3.5}$$

By subtracting the last equation in (3.5) from the sum of the first three equations, we have

$$\sum_{i=1}^{6} \hat{\mathbf{v}}(m_i) \cdot \mathbf{n}_i |e_i| = 0. \tag{3.6}$$

Therefore, it suffices to show that

$$\sum_{i=1}^{6} \mathbf{w}(m_i) \cdot \mathbf{n}_i |e_i| = \sum_{i=1}^{6} \hat{\mathbf{v}}(m_i) \cdot \mathbf{n}_i |e_i|. \tag{3.7}$$

Let $\tilde{\mathbf{v}} = \mathbf{v}|_{\triangle ABP}$. By the definition of the intergrid transfer operator,

$$\begin{aligned} \mathbf{w}(m_i) &= \tfrac{1}{2}[\hat{\mathbf{v}}(m_i) + \tilde{\mathbf{v}}(m_i)] \\ &= \hat{\mathbf{v}}(m_i) + \tfrac{1}{2}[\tilde{\mathbf{v}}(m_i) - \hat{\mathbf{v}}(m_i)], \qquad i = 1, 2. \end{aligned} \tag{3.8}$$

The function $g = (\tilde{\mathbf{v}} - \hat{\mathbf{v}}) \cdot \mathbf{n}$, where $\mathbf{n} = \mathbf{n}_1 = \mathbf{n}_2$, is a linear function along $\overline{AB}$ which vanishes at the midpoint D. Therefore,

$$g(m_1) + g(m_2) = 0. \tag{3.9}$$

Combining (3.8) and (3.9), we have

$$\sum_{i=1}^{2} \mathbf{w}(m_i) \cdot \mathbf{n}_i = \sum_{i=1}^{2} \hat{\mathbf{v}}(m_i) \cdot \mathbf{n}_i.$$

Similarly,

$$\sum_{i=3}^{4} \mathbf{w}(m_i) \cdot \mathbf{n}_i = \sum_{i=3}^{4} \hat{\mathbf{v}}(m_i) \cdot \mathbf{n}_i$$

and

$$\sum_{i=5}^{6} \mathbf{w}(m_i) \cdot \mathbf{n}_i = \sum_{i=5}^{6} \hat{\mathbf{v}}(m_i) \cdot \mathbf{n}_i. \qquad \square$$

It is obvious that $I_{k-1}^k : V_{k-1} \to V_k$ is a linear operator. The following proposition will be useful in the work estimate of the full multigrid algorithm.

Proposition 2. *The matrix representing I_{k-1}^k with respect to the standard bases of V_{k-1} and V_k (cf. (2.8)) is sparse, with the number of nonzero entries per row bounded by 9.*

Proof. First we look at the effect of I_{k-1}^k on basis functions in V_{k-1} that are associated with internal edges. Let $\hat{e} = \overline{AC}$ be an internal edge of $\mathcal{T}^{k-1}$ (cf. Figure 4) and $\mathbf{v}_{\hat{e}}^{k-1}$ be the basis function in V_{k-1} associated with $\hat{e}$. Denote by P the simply connected polygonal domain $AFBGCHDE$. By definition, $I_{k-1}^k \mathbf{v}_{\hat{e}}^{k-1}$ is supported on P and vanishes at the midpoints of $\mathcal{T}^k$ along ∂P. It follows from Lemma 1 and the definition of I_{k-1}^k that $I_{k-1}^k \mathbf{v}_{\hat{e}}^{k-1}$ is a linear combination of $\mathbf{v}_e^k$, where e ranges over all edges of $\mathcal{T}^k$ in the domain $ABCD$, and $\mathbf{v}_{p_i}^k$, $1 \le i \le 5$. In the case that one or more of the edges $\overline{AB}$, $\overline{BC}$, $\overline{CD}$, or $\overline{DA}$ are along $\partial\Omega$, the results are similar.

We now examine the effect of I_{k-1}^k on basis functions in V_{k-1} that are associated with internal vertices. Let p be a vertex in $\mathcal{T}^{k-1}$ and $\mathbf{v}_p^{k-1}$ be the basis function on V_{k-1} associated with p. We assume that there are, say, five edges in $\mathcal{T}^{k-1}$ that have p as a vertex (cf. Figure 5). Denote by $\widetilde{P}$ the simply connected polygonal domain $AGBHCIDJEF$. From the definition of I_{k-1}^k, the function $I_{k-1}^k \mathbf{v}_p$ is supported in $\widetilde{P}$ and vanishes at the midpoints of $\mathcal{T}^k$

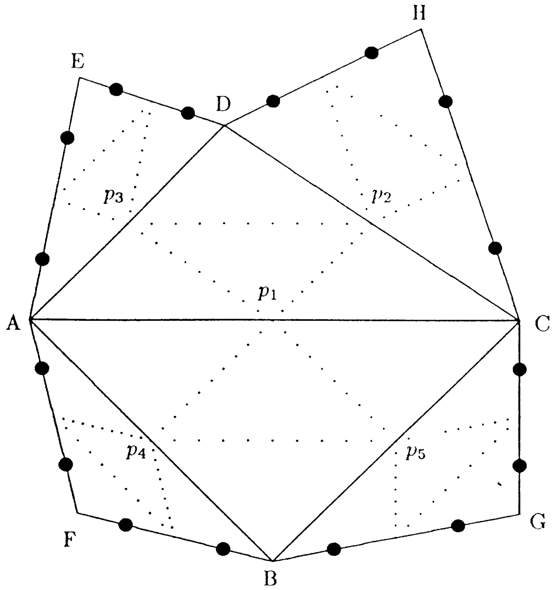

FIGURE 4

along $\partial\widetilde{P}$. Therefore, Lemma 1 and the definition of I_{k-1}^k imply that $I_{k-1}^k \mathbf{v}_p^{k-1}$ is a linear combination of $\mathbf{v}_p^k$, $\mathbf{v}_{p_i}^k$, $i = 1, \ldots, 10$, and $\mathbf{v}_e^k$, where e ranges over all edges of $\mathscr{T}^k$ in the domain $ABCDE$. Again, the results are similar if one or more of the edges $\overline{AB}$, $\overline{BC}$, $\overline{CD}$, $\overline{DE}$, or $\overline{EA}$ are along $\partial\Omega$.

The proposition now follows from the two observations above. □

The rest of this section is devoted to the proofs of (1.12)–(1.14).

We first give more explicit descriptions of $\|\cdot\|_{L^2}$ and $|\cdot|_{H^1}$ for piecewise linear vector functions on a triangle. Let T be a triangle and $\mathbf{v}$ be a piecewise linear vector function on T. We have the quadrature formula

$$\|\mathbf{v}\|^2_{(L^2(T))^2} = \frac{|T|}{3}\sum_{i=1}^{3}|\mathbf{v}(m_i)|^2, \tag{3.10}$$

where the m_i are the midpoints of the sides of T for $i = 1, 2, 3$ (cf. [9, p. 183]). Also, by a standard homogeneity argument, there exist constants C_1, C_2 which depend only on the angles in T, such that

$$C_1\Theta(\mathbf{v}) \le |\mathbf{v}|^2_{(H^1(T))^2} \le C_2\Theta(\mathbf{v}), \tag{3.11}$$

where

$$\Theta(\mathbf{v}) = [\mathbf{v}(m_1) - \mathbf{v}(m_2)]^2 + [\mathbf{v}(m_2) - \mathbf{v}(m_3)]^2 + [\mathbf{v}(m_3) - \mathbf{v}(m_1)]^2. \tag{3.12}$$

The following two lemmas prepare the way for the proofs of (1.12) and (1.13).

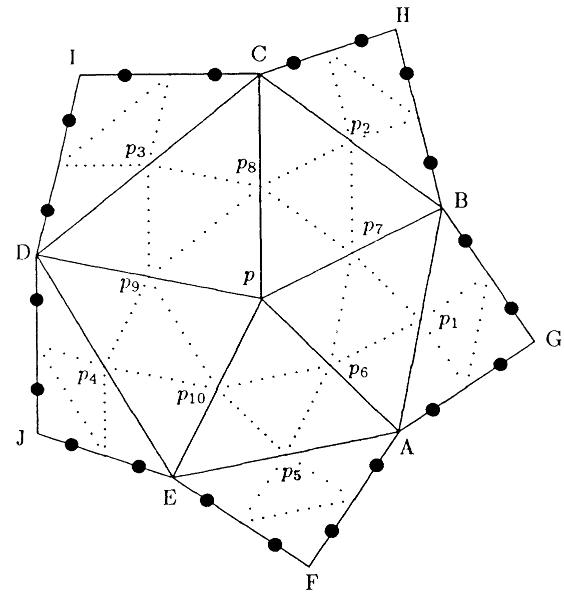

FIGURE 5

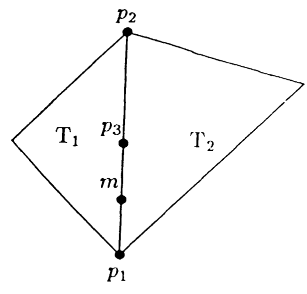

FIGURE 6

Lemma 3. *Let G be the union of two neighboring triangles T_1 and T_2. Let p_3 be the midpoint of $\overline{p_1p_2}$ and m be the midpoint of $\overline{p_1p_3}$ (cf. Figure 6). Let $Z := \{\mathbf{w}\colon \mathbf{w}|_{T_i}$ is linear for $i = 1, 2$ and $\mathbf{w}$ is continuous at $p_3\}$. Then there exists a positive constant C depending only on the angles of T_1 and T_2 such that*

$$\left|\mathbf{w}|_{T_1}(m) - \mathbf{w}|_{T_2}(m)\right| \le C(|\mathbf{w}|_{(H^1(T_1))^2} + |\mathbf{w}|_{(H^1(T_2))^2}) \tag{3.13}$$

for all $\mathbf{w} \in Z$.

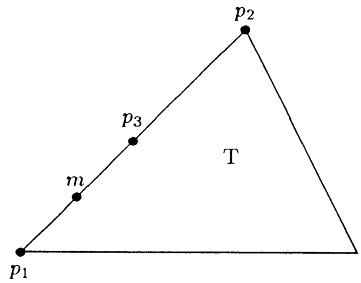

FIGURE 7

Proof. Given $\mathbf{w} \in Z$, let $\mathbf{w}_1 = \mathbf{w}|_{T_1}$ and $\mathbf{w}_2 = \mathbf{w}|_{T_2}$. Let $\mathbf{t}$ be a unit vector in the direction $\overline{p_3p_1}$. Then we have

$$\begin{aligned}|\mathbf{w}_1(m) - \mathbf{w}_2(m)| &= |\mathbf{w}_1(m) - \mathbf{w}_1(p_3) + \mathbf{w}_2(p_3) - \mathbf{w}_2(m)| \\ &\le |\mathbf{w}_1(m) - \mathbf{w}_1(p_3)| + |\mathbf{w}_2(p_3) - \mathbf{w}_2(m)| \\ &= \left|\frac{\partial \mathbf{w}_1}{\partial t}\right| |\overline{mp_3}| + \left|\frac{\partial \mathbf{w}_2}{\partial t}\right| |\overline{mp_3}| \\ &\le C[|\mathbf{w}|_{(H^1(T_1))^2} + |\mathbf{w}|_{(H^1(T_2))^2}]. \quad \square\end{aligned}$$

The next lemma is proved similarly.

Lemma 4. *Let T be a triangle. Let p_3 be the midpoint of $\overline{p_1p_2}$ and m be the midpoint of $\overline{p_1p_3}$ (cf. Figure 7). Let $Z = \{\mathbf{w}\colon \mathbf{w}$ is linear and $\mathbf{w} = \mathbf{0}$ at $p_3\}$. Then there exists a positive constant C depending only on the angles in T such that*

$$|\mathbf{w}(m)| \le C|\mathbf{w}|_{(H^1(T))^2}. \tag{3.14}$$

Theorem 1. *There exists a positive constant C such that for all $\mathbf{v} \in V_{k-1}$,*

$$\|I_{k-1}^k \mathbf{v}\|_k \le C\|\mathbf{v}\|_{k-1}$$

and

$$\|I_{k-1}^k \mathbf{v} - \mathbf{v}\|_{(L^2(\Omega))^2} \le Ch_k\|\mathbf{v}\|_{k-1},$$

i.e., (1.12) *and* (1.13) *hold.*

Proof. Given $\mathbf{v} \in V_{k-1}$, we can write

$$\sum_{T\in\mathscr{T}^k} \sum_{i=1}^{3} |(I_{k-1}^k \mathbf{v} - \mathbf{v}|_T)(m_i)|^2 = S_1 + S_2 + S_3,$$

where S_1, S_2, and S_3 are defined as follows:

$$S_1 = \sum_m [|(I_{k-1}^k \mathbf{v} - \mathbf{v}|_{\widehat{T}_1})(m)|^2 + |(I_{k-1}^k \mathbf{v} - \mathbf{v}|_{\widehat{T}_2})(m)|^2],$$

where m ranges over all the midpoints in $\mathscr{T}^k$ that belong to an internal edge in $\mathscr{T}^{k-1}$ and $\widehat{T}_1$, $\widehat{T}_2 \in \mathscr{T}^{k-1}$ are the two triangles that contain m;

$$S_2 = \sum_m |(I_{k-1}^k \mathbf{v} - \mathbf{v}|_{\widehat{T}})(m)|^2,$$

where m ranges over all midpoints of $\mathscr{T}^k$ along $\partial\Omega$ and $\widehat{T} \in \mathscr{T}^{k-1}$ is the triangle that contains m; and

$$S_3 = 2\sum_m |(I_{k-1}^k \mathbf{v} - \mathbf{v}|_{\widehat{T}})(m) \cdot \mathbf{n}_m|^2,$$

where m ranges over all the midpoints in $\mathscr{T}^k$ that are inside some triangle $\widehat{T} \in \mathscr{T}^{k-1}$ and $\mathbf{n}_m$ is a unit vector normal to the edge containing m.

Lemmas 3 and 4 and the definition of I_{k-1}^k imply that $S_1 + S_2 \leq C\|\mathbf{v}\|_{k-1}^2$. On the other hand, S_3 can be estimated in terms of S_1 and S_2. Referring back to Figure 3, let m_7 be a typical midpoint in S_3 and $\widehat{T} = \triangle ABC \in \mathscr{T}^{k-1}$. Since $\mathbf{v}$ and $I_{k-1}^k \mathbf{v}$ are both divergence-free on $\triangle ADF$, (2.2) implies that

$$\begin{aligned}((I_{k-1}^k \mathbf{v} - \mathbf{v}|_{\widehat{T}})(m_7) \cdot \mathbf{n}_7)|\overline{DF}| = &-((I_{k-1}^k \mathbf{v} - \mathbf{v}|_{\widehat{T}})(m_6) \cdot \mathbf{n}_6)|\overline{AF}| \\ &-((I_{k-1}^k \mathbf{v} - \mathbf{v}|_{\widehat{T}})(m_1) \cdot \mathbf{n}_1)|\overline{AD}|.\end{aligned}$$

Hence,

$$|(I_{k-1}^k \mathbf{v} - \mathbf{v}|_{\widehat{T}})(m_7) \cdot \mathbf{n}_7|^2 \leq C\{|I_{k-1}^k \mathbf{v} - \mathbf{v}|_{\widehat{T}}(m_6)|^2 + |I_{k-1}^k \mathbf{v} - \mathbf{v}|_{\widehat{T}}(m_1)|^2\}.$$

Therefore, $S_3 \leq C(S_1 + S_2)$, and we have

$$\sum_{T\in\mathscr{T}^k} \sum_{i=1}^3 |(I_{k-1}^k \mathbf{v} - \mathbf{v}|_T)(m_i)|^2 \leq C\|\mathbf{v}\|_{k-1}^2. \tag{3.15}$$

From (3.11) and (3.15),

$$\begin{aligned}\|I_{k-1}^k \mathbf{v}\|_k^2 &= \sum_{T\in\mathscr{T}^k} |I_{k-1}^k \mathbf{v}|_{H^1(T)}^2 \leq C \sum_{T\in\mathscr{T}^k} \Theta((I_{k-1}^k \mathbf{v})|_T) \\ &\leq C\left[\sum_{T\in\mathscr{T}^k} \Theta(\mathbf{v}|_T) + \sum_{T\in\mathscr{T}^k} \Theta((I_{k-1}^k \mathbf{v})|_T - \mathbf{v}|_T)\right] \\ &\leq C\left[\|\mathbf{v}\|_{k-1}^2 + \sum_{T\in\mathscr{T}^k}\sum_{i=1}^3 |(I_{k-1}^k \mathbf{v} - \mathbf{v}|_T)(m_i)|^2\right] \leq C\|\mathbf{v}\|_{k-1}^2.\end{aligned}$$

This completes the proof of the first inequality.

From (3.10) and (3.15), we have

$$\begin{aligned}\|I_{k-1}^k\mathbf{v}-\mathbf{v}\|^2_{(L^2(\Omega))^2} &= \sum_{T\in\mathcal{T}^k} |I_{k-1}^k\mathbf{v}-\mathbf{v}|^2_{(L^2(T))^2}\\ &\le Ch_k^2\sum_{T\in\mathcal{T}^k}\sum_{i=1}^3 |(I_{k-1}^k\mathbf{v}-\mathbf{v}|_T)(m_i)|^2\\ &\le Ch_k^2\|\mathbf{v}\|^2_{k-1}. \quad \square\end{aligned}$$

Corollary 1. *There exists a positive constant C such that*

$$\|I_{k-1}^k\mathbf{v}\|_{(L^2(\Omega))^2} \le C\|\mathbf{v}\|_{(L^2(\Omega))^2} \quad \forall \mathbf{v}\in V_{k-1}. \tag{3.16}$$

Proof. From Theorem 1 and a standard inverse estimate (cf. [9, p. 140]), we have

$$\begin{aligned}\|I_{k-1}^k\mathbf{v}\|_{(L^2(\Omega))^2} &\le \|I_{k-1}^k\mathbf{v}-\mathbf{v}\|_{(L^2(\Omega))^2} + \|\mathbf{v}\|_{(L^2(\Omega))^2}\\ &\le Ch_k\|\mathbf{v}\|_{k-1} + \|\mathbf{v}\|_{(L^2(\Omega))^2} \le C\|\mathbf{v}\|_{(L^2(\Omega))^2}. \quad \square\end{aligned}$$

Inequality (1.14) will be proved by a homogeneity argument. We will therefore first prove some estimates on reference domains.

Lemma 5. *Let G be the union of two neighboring triangles T_1 and T_2 such that* $\operatorname{diam} G = 1$. *Let m_i $(1\le i\le 5)$ be the midpoints of the edges e_i $(1\le i\le 5)$ of T_1 and T_2 and let $\tilde{m}$ be the midpoint of $e=\overline{p_1m_1}$ (cf. Figure 8). Then*

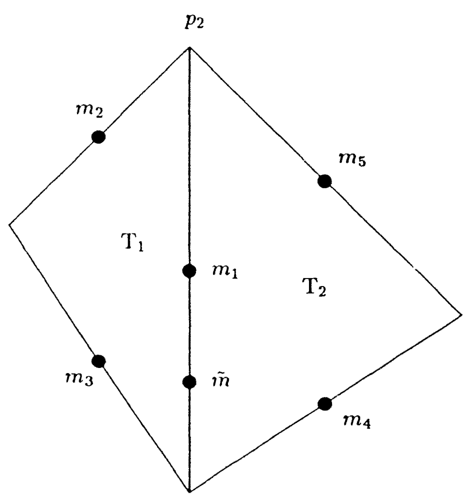

FIGURE 8

there exists a positive constant C *depending only on the angles in* T_1 *and* T_2 *such that*

$$\begin{aligned}(3.17)\qquad &\left|\frac{1}{|e_1|}\int_{e_1}\mathbf{h}\,ds+\frac{1}{4}\frac{1}{|e_3|}\int_{e_3}\mathbf{h}\,ds-\frac{1}{4}\frac{1}{|e_2|}\int_{e_2}\mathbf{h}\,ds\right.\\ &\left.+\frac{1}{4}\frac{1}{|e_4|}\int_{e_4}\mathbf{h}\,ds-\frac{1}{4}\frac{1}{|e_5|}\int_{e_5}\mathbf{h}\,ds-\frac{1}{|e|}\int_{e}\mathbf{h}\,ds\right|\leq C|\mathbf{h}|_{(H^2(G))^2},\end{aligned}$$

for all $\mathbf{h}\in(H^2(G))^2$.

Proof. It suffices to prove (3.17) for scalar functions in $H^2(G)$. Define a linear functional l on $H^2(G)$ by

$$\begin{aligned}(3.18)\qquad l(\eta)=&\frac{1}{|e_1|}\int_{e_1}\eta\,ds+\frac{1}{4}\frac{1}{|e_3|}\int_{e_3}\eta\,ds-\frac{1}{4}\frac{1}{|e_2|}\int_{e_2}\eta\,ds\\ &+\frac{1}{4}\frac{1}{|e_4|}\int_{e_4}\eta\,ds-\frac{1}{4}\frac{1}{|e_5|}\int_{e_5}\eta\,ds-\frac{1}{|e|}\int_{e}\eta\,ds.\end{aligned}$$

Observe that if $g\in\mathscr{P}_1(G)$ (i.e., g is linear), then $l(g)=\eta(m_1)+\frac{1}{4}\eta(m_3)-\frac{1}{4}\eta(m_2)+\frac{1}{4}\eta(m_4)-\frac{1}{4}\eta(m_5)-\eta(\tilde{m})=0$.

By the trace theorem (cf. [1, p. 114]), for any $g\in\mathscr{P}_1(G)$,

$$|l(\eta)|=|l(\eta+g)|\leq C\|\eta+g\|_{H^2(G)}.$$

Therefore, by the Bramble-Hilbert lemma (cf. [4]),

$$|l(\eta)|\leq C\inf_{g\in\mathscr{P}_1(G)}\|\eta+g\|_{H^2(G)}\leq C|\eta|_{H^2(G)}.\quad\square$$

The proof of the next lemma is similar.

Lemma 6. *Let* G *be a triangle such that* $\operatorname{diam}G=1$. *Let* m_i $(1\leq i\leq 3)$ *be the midpoints of edges* e_i $(1\leq i\leq 3)$ *of* G *and* $e=\overline{m_2m_3}$. *Then there exists a positive constant* C *depending only on the angles in* G *such that*

$$\begin{aligned}(3.19)\qquad &\left|\frac{1}{2}\frac{1}{|e_2|}\int_{e_2}\mathbf{h}\,ds+\frac{1}{2}\frac{1}{|e_3|}\int_{e_3}\mathbf{h}\,ds-\frac{1}{|e|}\int_{e}\mathbf{h}\,ds\right|\\ &\leq C|\mathbf{h}|_{(H^2(G))^2}\quad\forall\mathbf{h}\in(H^2(G))^2.\end{aligned}$$

Finally, we are ready to prove inequality (1.14).

Theorem 2. *There exists a positive constant* C *such that*

$$\|I_{k-1}^k(\Pi_{k-1}\mathbf{g})-\Pi_k\mathbf{g}\|_k\leq Ch_k|\mathbf{g}|_{(H^2(\Omega))^2}$$

and

$$\|I_{k-1}^k(\Pi_{k-1}\mathbf{g})-\Pi_k\mathbf{g}\|_{(L^2(\Omega))^2}\leq Ch_k^2|\mathbf{g}|_{(H^2(\Omega))^2}\quad\forall\mathbf{g}\in(H^2(\Omega))^2\cap(H_0^1(\Omega))^2.$$

Proof. From (3.11),

$$\begin{aligned}\|I_{k-1}^{k}(\Pi_{k-1}\mathbf{g})-\Pi_{k}\mathbf{g}\|_{k}^{2} &= \sum_{T\in\mathcal{T}^{k}}|I_{k-1}^{k}(\Pi_{k-1}\mathbf{g})-\Pi_{k}\mathbf{g}|_{(H^{1}(T))^{2}}^{2}\\ &\le C\sum_{T\in\mathcal{T}^{k}}\Theta((I_{k-1}^{k}(\Pi_{k-1}\mathbf{g})-\Pi_{k}\mathbf{g})|_{T})\\ &\le C\sum_{T\in\mathcal{T}^{k}}\sum_{i=1}^{3}|I_{k-1}^{k}(\Pi_{k-1}\mathbf{g})-\Pi_{k}\mathbf{g}|^{2}(m_{i}),\end{aligned}$$

where the m_i are the midpoints of T.

We can write

$$\sum_{T\in\mathcal{T}^{k}}\sum_{i=1}^{3}|I_{k-1}^{k}(\Pi_{k-1}\mathbf{g})-\Pi_{k}\mathbf{g}|^{2}(m_{i})=S_{1}+S_{2}+S_{3},$$

where S_1, S_2, and S_3 are defined as follows:

$$S_{1}=\sum_{m}\alpha(m)|I_{k-1}^{k}(\Pi_{k-1}\mathbf{g})-\Pi_{k}\mathbf{g}|^{2}(m),$$

where m ranges over all the midpoints in $\mathcal{T}^k$ that belong to an edge in $\mathcal{T}^{k-1}$, $\alpha(m)=1$ if $m\in\partial\Omega$, otherwise $\alpha(m)=2$;

$$S_{2}=2\sum_{m}[(I_{k-1}^{k}(\Pi_{k-1}\mathbf{g})-\Pi_{k}\mathbf{g})(m)\cdot\mathbf{t}_{m}]^{2},$$

where m ranges over all the midpoints in $\mathcal{T}^k$ that are inside some triangle in $\mathcal{T}^{k-1}$ and $\mathbf{t}_m$ is a unit vector tangential to the edge that contains m as its midpoint; and

$$S_{3}=2\sum_{m}[(I_{k-1}^{k}(\Pi_{k-1}\mathbf{g})-\Pi_{k}\mathbf{g})(m)\cdot\mathbf{n}_{m}]^{2},$$

where m ranges over all the midpoints in $\mathcal{T}^k$ that are inside some triangle in $\mathcal{T}^{k-1}$ and $\mathbf{n}_m$ is a unit vector normal to the edge containing m.

The definition of I_{k-1}^k, Lemma 5, and a homogeneity argument imply that $S_1\le Ch_k^2|\mathbf{g}|^2_{(H^2(\Omega))^2}$. Similarly, $S_2\le Ch_k^2|\mathbf{g}|^2_{(H^2(\Omega))^2}$ follows from the definition of I_{k-1}^k, Lemma 6, and a homogeneity argument. On the other hand, $S_3\le CS_1$ by the divergence-free condition. Therefore, we have established the first inequality.

The second inequality follows from the observation that (3.10) implies

$$\begin{aligned}\|I_{k-1}^{k}(\Pi_{k-1}\mathbf{g})-\Pi_{k}\mathbf{g}\|_{(L^{2}(\Omega))^{2}} &\le Ch_{k}^{2}\sum_{T\in\mathcal{T}^{k}}\sum_{i=1}^{3}|I_{k-1}^{k}(\Pi_{k}\mathbf{g})-\Pi_{k}\mathbf{g}|^{2}(m_{i})\\ &\le Ch_{k}^{2}(S_{1}+S_{2}+S_{3}).\quad\square\end{aligned}$$

4. The multigrid algorithm

Given $\mathbf{v} \in V_k$, we can write $\mathbf{v} = \sum a_i \mathbf{v}^k_{e_i} + \sum b_j \mathbf{v}^k_{p_j}$, where the e_i ranges over all internal edges of $\mathscr{T}^k$ and p_j ranges over all internal vertices of $\mathscr{T}^k$ (cf. (2.8)). The inner product $(\cdot, \cdot)_k$ on V_k is defined by

$$(\mathbf{v}_1, \mathbf{v}_2)_k := h_k^4 \sum a_{1,i} a_{2,i} + h_k^2 \sum b_{1,j} b_{2,j}, \tag{4.1}$$

where $\mathbf{v}_1 = \sum a_{1,i} \mathbf{v}^k_{e_i} + \sum b_{1,j} \mathbf{v}^k_{p_j}$ and $\mathbf{v}_2 = \sum a_{2,i} \mathbf{v}^k_{e_i} + \sum b_{2,j} \mathbf{v}^k_{p_j}$ belong to V_k.

Using the quadrature formula (3.10), it is easy to see that

$$(\mathbf{v}, \mathbf{v})_{(L^2(\Omega))^2} \leq C h_k^{-2} (\mathbf{v}, \mathbf{v})_k \quad \forall \mathbf{v} \in V_k. \tag{4.2}$$

The fine-to-coarse intergrid transfer operator $I_k^{k-1}: V_k \to V_{k-1}$ is defined by

$$(I_{k-1}^k \mathbf{v}, \mathbf{w})_k = (\mathbf{v}, I_k^{k-1} \mathbf{w})_{k-1} \quad \forall \mathbf{v} \in V_{k-1}, \mathbf{w} \in V_k. \tag{4.3}$$

The symmetric positive definite operator $A_k: V_k \to V_k$ is defined by

$$(A_k \mathbf{v}, \mathbf{w})_k = a_k(\mathbf{v}, \mathbf{w}) \quad \forall \mathbf{v}, \mathbf{w} \in V_k, \tag{4.4}$$

where $a_k(\cdot, \cdot)$ is defined in (1.7).

Remark 1. With respect to the standard basis, A_k is represented by a sparse matrix. The number of nonzero entries per row is bounded by $\max(6, N)$, where N represents the maximum number of edges in $\mathscr{T}^1$ that have a common vertex inside Ω.

By a standard inverse estimate (cf. [9, p. 140]),

$$a_k(\mathbf{v}, \mathbf{v}) \leq C h_k^{-2} (\mathbf{v}, \mathbf{v})_{(L^2(\Omega))^2} \quad \forall \mathbf{v} \in V_k. \tag{4.5}$$

Then (4.2) and (4.5) imply that the largest eigenvalue of A_k is bounded by

$$\Lambda_k := C h_k^{-4}. \tag{4.6}$$

The W-cycle multigrid algorithm can now be described. We first describe the kth-level iteration scheme. The full multigrid algorithm consists of a nested iteration of these schemes.

The kth-level iteration. The kth-level iteration with initial guess z_0 yields $MG(k, \mathbf{z}_0, \mathbf{g})$ as an approximate solution to the equation

$$A_k \mathbf{z} = \mathbf{g}.$$

For $k = 1$, $MG(1, \mathbf{z}_0, \mathbf{g})$ is the solution obtained from a direct method. In other words,

$$MG(1, \mathbf{z}_0, \mathbf{g}) = A_1^{-1} \mathbf{g}.$$

For $k > 1$, there are two steps:

Smoothing step. Let $\mathbf{z}_l \in V_k$ $(1 \leq l \leq m)$ be defined recursively by the equations

$$\mathbf{z}_l = \mathbf{z}_{l-1} + \frac{1}{\Lambda_k} (\mathbf{g} - A_k \mathbf{z}_{l-1}), \qquad 1 \leq l \leq m, \tag{4.7}$$

where m is a positive integer independent of k.

Correction step. Let $\bar{\mathbf{g}} := I_k^{k-1}(\mathbf{g} - A_k\mathbf{z}_m)$. Let $\mathbf{q}_i \in V_{k-1}$ $(0 \le i \le p,\ p = 2$ or $3)$ be defined recursively by

$$\mathbf{q}_0 = 0$$

and

$$\mathbf{q}_i = MG(k-1, \mathbf{q}_{i-1}, \bar{\mathbf{g}}), \qquad 1 \le \quad \le p.$$

Then $MG(k, \mathbf{z}_0, \mathbf{g})$ is defined to be $\mathbf{z}_m + I_{k-1}^k \mathbf{q}_p$.

The full multigrid algorithm. In the case $k = 1$, the approximate solution $\hat{\mathbf{u}}_1$ of (1.9) is obtained by a direct method. The approximate solutions $\hat{\mathbf{u}}_k$ $(k \ge 2)$ of (1.9) are obtained recursively from

$$\begin{aligned} \mathbf{u}_0^k &= I_{k-1}^k \hat{\mathbf{u}}_{k-1}, \\ \mathbf{u}_l^k &= MG(k, \mathbf{u}_{l-1}^k, \mathbf{f}_k), \qquad 1 \le l \le r, \\ &\qquad\qquad (\mathbf{f}_k, \mathbf{v})_k := \int_\Omega \mathbf{f}\mathbf{v}\, dx \ \forall \mathbf{v} \in V_k, \end{aligned}$$

and

$$\hat{\mathbf{u}}_k = \mathbf{u}_r^k,$$

where r is a positive integer independent of k.

Remark 2. By Proposition 2 and Remark 1, relative to the standard basis of V_k, the operators I_{k-1}^k, I_k^{k-1}, and A_k are represented by matrices with $\mathscr{O}(n_k)$ nonzero entries. Along with the asymptotic formula (2.7) and the fact that the number of corrections p is less than four in the kth-level iteration, the total work of the full multigrid algorithm is therefore $\mathscr{O}(n_k)$. The proof is the same as the one in [2].

5. Mesh-dependent norms

The mesh-dependent norm $|||\cdot|||_{s,k}$ on V_k is defined by

$$|||\mathbf{v}|||_{s,k}^2 := (A_k^{s/2}\mathbf{v}, \mathbf{v})_k. \tag{5.1}$$

Therefore,

$$|||\mathbf{v}|||_{0,k} = \sqrt{(\mathbf{v}, \mathbf{v})_k} \quad \text{and} \quad |||\mathbf{v}|||_{2,k} = \sqrt{(A_k\mathbf{v}, \mathbf{v})_k} = \sqrt{a_k(\mathbf{v}, \mathbf{v})} = \|\mathbf{v}\|_k.$$

From definition (5.1), it is easy to deduce the following inequality:

$$|a_k(\mathbf{v}, \mathbf{w})| \le |||\mathbf{v}|||_{2+t,k}|||\mathbf{w}|||_{2-t,k}. \tag{5.2}$$

The rest of this section will be devoted to the proof of the following proposition, which is needed for the proof of the approximation property in §6.

Proposition 3. *We have* $|||\mathbf{v}|||_{1,k} \leq C\|\mathbf{v}\|_{(L^2(\Omega))^2}$.

The proof of this proposition is based on the relationship between the divergence-free P1 nonconforming space and the Morley finite element space (cf. [15]). Let M_k be the Morley finite element space associated with $\mathscr{T}^k$. Then $\phi \in M_k$ if and only if it has the following three properties:

(i) $\phi|_T$ is quadratic for all $T \in \mathscr{T}^k$,
(ii) ϕ is continuous at the vertices and vanishes at the vertices along $\partial\Omega$, and
(iii) $\partial\phi/\partial n$ is continuous at the midpoints of interelement boundaries and vanishes at the midpoints along $\partial\Omega$.

The Morley finite element space can be used to construct a nonconforming multigrid method for the biharmonic equation (cf. [7, 16]), which is closely related to the stationary Stokes equations (cf. [9, p. 280]). We can define two mesh-dependent inner products on M_k.

For ϕ and ψ in M_k,

$$b_k(\phi, \psi) := \sum_{T\in\mathscr{T}^k} \int_T D^2\phi : D^2\psi \, dx, \tag{5.3}$$

where

$$D^2\phi : D^2\psi := \sum_{i,j} \frac{\partial^2\phi}{\partial x_i \partial x_j} \cdot \frac{\partial^2\psi}{\partial x_i \partial x_j},$$

and

$$\langle \phi, \psi\rangle_k := h_k^2 \left[\sum_p \phi(p)\psi(p) + h_k^2 \sum_m \frac{\partial\phi}{\partial n}(m) \frac{\partial\psi}{\partial n}(m) \right], \tag{5.4}$$

where p ranges over all internal vertices and m ranges over all internal midpoints of $\mathscr{T}^k$.

Let $B_k : M_k \to M_k$ be a symmetric positive definite operator defined by

$$\langle B_k\phi, \psi\rangle_k := b_k(\phi, \psi). \tag{5.5}$$

We can define the mesh-dependent norms $||||\cdot||||_{s,k}$ on M_k by

$$||||\phi||||_{s,k}^2 := \langle B_k^{s/2}\phi, \phi\rangle_k. \tag{5.6}$$

Given $\phi \in M_k$, we denote by ϕ^I the continuous piecewise linear function that has the same value as ϕ at the vertices of $\mathscr{T}^k$. The following lemma is proved in Proposition 8.1 of [16].

Lemma 7. *For any* $\phi \in M_k$, *we have*

$$||||\phi||||_{1,k} \leq C(|\phi^I|_{H^1(\Omega)} + h_k ||||\phi||||_{2,k}). \tag{5.7}$$

There is an isomorphism between M_k and V_k given by the operator **curl**, where

$$(5.8)\qquad (\mathbf{curl}\,\phi)|_T = \left(\frac{\partial \phi|_T}{\partial x_2}, -\frac{\partial \phi|_T}{\partial x_1}\right) \quad \forall T \in \mathcal{T}^k.$$

More explicitly, if $\mathbf{curl}\,\phi$ is represented in terms of the standard basis of V_k, say $\mathbf{curl}\,\phi = \sum a_i \mathbf{v}^k_{e_i} + \sum b_j \mathbf{v}^k_{p_j}$, then

$$(5.9)\qquad a_i = \pm \frac{\partial \phi}{\partial n}(m_i)$$

and

$$(5.10)\qquad b_j = \phi(p_j),$$

where m_i is the midpoint of edge e_i and the sign in (5.9) depends on the choice of $\mathbf{t}_{e_i}$ and $\mathbf{n}_{e_i}$.

It follows from (4.1), (5.4), (5.9), and (5.10) that

$$(5.11)\qquad \langle \phi, \psi \rangle_k = (\mathbf{curl}\,\phi, \mathbf{curl}\,\psi)_k.$$

An easy computation also shows that

$$(5.12)\qquad a_k(\mathbf{curl}\,\phi, \mathbf{curl}\,\psi) = b_k(\phi, \psi).$$

It therefore follows from (5.11), (5.12), and the definition of the mesh-dependent norms that

$$(5.13)\qquad |||\mathbf{curl}\,\phi|||_{s,k} = ||||\phi||||_{s,k}.$$

Let $\mathbf{curl}\,\phi = \mathbf{v} = \sum a_i \mathbf{v}_{e_i} + \sum b_j \mathbf{v}_{p_j}$. The quadrature formula (3.10), (5.10), the definition of $\mathbf{v}_{p_j}$, and a straightforward computation show that there exist positive constants C_1, C_2, C_3, and C_4 such that

$$(5.14)\qquad C_1 \sum [\phi(p) - \phi(p')]^2 \le |\phi^I|^2_{H^1(\Omega)} \le C_2 \sum [\phi(p) - \phi(p')]^2$$

and

$$(5.15)\qquad C_3 \sum [\phi(p) - \phi(p')]^2 \le \left\| \sum b_j \mathbf{v}_{p_j} \right\|^2_{(L^2(\Omega))^2} \le C_4 \sum [\phi(p) - \phi(p')]^2,$$

where p and p' range over any two vertices of any triangle in $\mathcal{T}^k$.

Proof of Proposition 3. Given $\mathbf{v} = \sum a_i \mathbf{v}_{e_i} + \sum b_j \mathbf{v}_{p_j}$, there exists a unique $\phi \in M_k$ such that $\mathbf{v} = \mathbf{curl}\,\phi$. In view of (5.13), (5.14), and (5.15), the inequality (5.7) is translated into

$$(5.16)\qquad |||\mathbf{v}|||_{1,k} \le C\left(\left\| \sum b_j \mathbf{v}_{p_j} \right\|_{(L^2(\Omega))^2} + h_k |||\mathbf{v}|||_{2,k} \right).$$

From the inverse estimate (4.5),

$$(5.17)\qquad h_k |||\mathbf{v}|||_{2,k} \le C \|\mathbf{v}\|_{(L^2(\Omega))^2} \quad \forall \mathbf{v} \in V_k.$$

From the definitions of $\mathbf{v}_{e_i}$, $\mathbf{v}_{p_j}$, and the polarized form of the quadrature formula (3.10), we have $(\sum a_i \mathbf{v}_{e_i}, \sum b_j \mathbf{v}_{p_j})_{(L^2(\Omega))^2} = 0$. Therefore,

$$\left\| \sum b_j \mathbf{v}_{p_j} \right\|_{(L^2(\Omega))^2} \leq \|\mathbf{v}\|_{(L^2(\Omega))^2}. \quad \square \tag{5.18}$$

6. Convergence analysis

We will first discuss the convergence of the kth-level iteration and then the convergence of the full multigrid algorithm. Following [2], we will use a perturbation argument for the convergence proof of the kth-level iteration. In other words, we begin with a two-grid analysis.

Define the operator $P_k^{k-1}: V_k \to V_{k-1}$ by

$$a_k(\mathbf{v}, I_{k-1}^k \mathbf{w}) = a_{k-1}(P_k^{k-1}\mathbf{v}, \mathbf{w}) \quad \forall \mathbf{v} \in V_k, \mathbf{w} \in V_{k-1}. \tag{6.1}$$

In other words, P_k^{k-1} is the adjoint operator of I_{k-1}^k relative to the inner products that define the energy norms on V_k and V_{k-1}. Therefore, the following lemma is a direct consequence of (1.12).

Lemma 8. *There exists a positive constant C such that*

$$\|P_k^{k-1}\mathbf{v}\|_{k-1} \leq C\|\mathbf{v}\|_k \quad \forall \mathbf{v} \in V_k. \tag{6.2}$$

In the two-grid algorithm, we assume that the residual equation is solved exactly on the coarser grid. The final output of the kth-level iteration is therefore $\mathbf{z}_m + I_{k-1}^k \mathbf{q}$, where

$$\mathbf{q} = A_{k-1}^{-1}\overline{\mathbf{g}} = A_{k-1}^{-1}(I_k^{k-1}(\mathbf{g} - A_k\mathbf{z}_m)) = A_{k-1}^{-1}(I_k^{k-1}A_k(\mathbf{z} - \mathbf{z}_m)).$$

We denote the final error $\mathbf{z} - (\mathbf{z}_m + I_{k-1}^k\mathbf{q})$ of the two-grid algorithm by $\mathbf{e}$ and the intermediate errors $\mathbf{z} - \mathbf{z}_i$ by $\mathbf{e}_i$, for $i = 0, 1, \ldots, m$.

Lemma 9. *We have* $\mathbf{q} = P_k^{k-1}\mathbf{e}_m$.

Proof. Given any $\mathbf{w} \in V_{k-1}$,

$$\begin{aligned} a_{k-1}(\mathbf{q}, \mathbf{w}) = (A_{k-1}\mathbf{q}, \mathbf{w})_{k-1} &= (I_k^{k-1}A_k\mathbf{e}_m, \mathbf{w})_{k-1} = (A_k\mathbf{e}_m, I_{k-1}^k\mathbf{w})_k \\ &= a_k(\mathbf{e}_m, I_{k-1}^k\mathbf{w}) = a_{k-1}(P_k^{k-1}\mathbf{e}_m, \mathbf{w}). \quad \square \end{aligned}$$

From the smoothing step (4.7), we obtain

$$\mathbf{e}_l = R_k\mathbf{e}_{l-1}, \qquad l = 1, 2, \ldots, m, \tag{6.3}$$

where the relaxation operator R_k is defined by

$$R_k = I - \frac{1}{\Lambda_k}A_k. \tag{6.4}$$

Since Λ_k dominates the largest eigenvalues of A_k, it is obvious that $|||R_k\mathbf{v}|||_{s,k} \le |||\mathbf{v}|||_{s,k}$ for all $\mathbf{v} \in V_k$. Lemma 9 and (6.3) imply that

$$\begin{aligned} \mathbf{e} &= \mathbf{e}_m - I_{k-1}^k\mathbf{q} = \mathbf{e}_m - I_{k-1}^k P_k^{k-1}\mathbf{e}_m \\ &= (I - I_{k-1}^k P_k^{k-1})\mathbf{e}_m = (I - I_{k-1}^k P_k^{k-1})R_k^m\mathbf{e}_0. \end{aligned} \tag{6.5}$$

The two-grid analysis will be complete once we estimate $I - I_{k-1}^k P_k^{k-1}$ (the approximation property) and R_k^m (the smoothing property).

Lemma 10 (Smoothing property). *There exists a positive constant C such that*

$$|||R_k^m\mathbf{v}|||_{\beta,k} \le Ch_k^{-1}(4m+1)^{-1/4}|||\mathbf{v}|||_{\beta-1,k} \quad \forall \mathbf{v} \in V_k \text{ and } \beta \in \mathbb{R}. \tag{6.6}$$

Proof. Let $\lambda_1 < \lambda_2 < \cdots \le \lambda_{n_k}$ be the eigenvalues of A_k and $\tilde{\mathbf{v}}_1, \tilde{\mathbf{v}}_2, \dots, \tilde{\mathbf{v}}_{n_k}$ be the corresponding eigenvectors such that $(\tilde{\mathbf{v}}_i, \tilde{\mathbf{v}}_j)_k = \delta_{ij}$. Recall that $\lambda_{n_k} \le \Lambda_k \le Ch_k^{-4}$ (cf. (4.6)). Let $\mathbf{v} = \sum_{i=1}^{n_k} \alpha_i\tilde{\mathbf{v}}_i$. Then

$$R_k^m\mathbf{v} = \sum_{i=1}^{n_k} \alpha_i \left(1 - \frac{\lambda_i}{\Lambda_k}\right)^m \tilde{\mathbf{v}}_i.$$

From the definition of the mesh-dependent norms (5.1), we have

$$\begin{aligned} |||R_k^m\mathbf{v}|||_{\beta,k}^2 &= \sum_{i=1}^{n_k} \alpha_i^2 \left(1 - \frac{\lambda_i}{\Lambda_k}\right)^{2m} \lambda_i^{\beta/2} \\ &= \Lambda_k^{1/2} \sum_{i=1}^{n_k} \alpha_i^2\lambda_i^{(\beta-1)/2} \left[\left(1 - \frac{\lambda_i}{\Lambda_k}\right)^{2m} \left(\frac{\lambda_i}{\Lambda_k}\right)^{1/2}\right] \\ &\le Ch_k^{-2} \sup_{0\le x\le 1} [(1-x)^{2m}x^{1/2}] \sum_{i=1}^{n_k} \alpha_i^2\lambda_i^{(\beta-1)/2} \\ &\le Ch_k^{-2}(4m+1)^{-1/2}|||\mathbf{v}|||_{\beta-1,k}^2. \quad \square \end{aligned}$$

Lemma 11 (Approximation property I). *There exists a positive constant C such that*

$$|||(I - I_{k-1}^k P_k^{k-1})\mathbf{v}|||_{1,k} \le Ch_k|||\mathbf{v}|||_{2,k} \quad \forall \mathbf{v} \in V_k. \tag{6.7}$$

Proof. By Proposition 3, it suffices to show that

$$\|(I - I_{k-1}^k P_k^{k-1})\mathbf{v}\|_{(L^2(\Omega))^2} \le Ch_k|||\mathbf{v}|||_{2,k} \quad \forall \mathbf{v} \in V_k. \tag{6.8}$$

The proof of (6.8) is based on a duality argument. Given $\mathbf{v} \in V_k$, let $\hat{\mathbf{v}} = (I - I_{k-1}^k P_k^{k-1})\mathbf{v}$ and let $(\mathbf{r}, \tilde{p}) \in ((H_0^1(\Omega))^2 \cap (H^2(\Omega))^2) \times (H^1(\Omega)/\mathbb{R})$ solve the continuous problem

$$\begin{aligned} -\Delta\mathbf{r} + \mathbf{grad}\,\tilde{p} &= \hat{\mathbf{v}} \quad \text{in } \Omega, \\ \operatorname{div}\mathbf{r} &= 0 \quad \text{in } \Omega, \\ \mathbf{r} &= \mathbf{0} \quad \text{on } \partial\Omega. \end{aligned} \tag{6.9}$$

Elliptic regularity (cf. (1.2)) implies that

$$\|\mathbf{r}\|_{(H^2(\Omega))^2} + |\tilde{p}|_{H^1(\Omega)} \le C\|\hat{\mathbf{v}}\|_{(L^2(\Omega))^2}. \tag{6.10}$$

Let $\mathbf{r}_k \in V_k$ and $\mathbf{r}_{k-1} \in V_{k-1}$ solve

$$\begin{aligned} a_k(\mathbf{r}_k, \mathbf{w}) &= \int_\Omega \hat{\mathbf{v}} \cdot \mathbf{w}\, dx \quad \forall \mathbf{w} \in V_k, \\ a_{k-1}(\mathbf{r}_{k-1}, \mathbf{w}) &= \int_\Omega \hat{\mathbf{v}} \cdot \mathbf{w}\, dx \quad \forall \mathbf{w} \in V_{k-1}, \end{aligned} \tag{6.11}$$

respectively. The discretization error estimate (1.10), elliptic regularity estimate (6.10), and the fact that $h_{k-1} = 2h_k$ imply that

$$\begin{aligned} \|\mathbf{r} - \mathbf{r}_k\|_k &\le Ch_k\|\hat{\mathbf{v}}\|_{(L^2(\Omega))^2}, \\ \|\mathbf{r} - \mathbf{r}_{k-1}\|_{k-1} &\le Ch_k\|\hat{\mathbf{v}}\|_{(L^2(\Omega))^2}. \end{aligned} \tag{6.12}$$

Denote $P_k^{k-1}\mathbf{v}$ by $\mathbf{z}$. (Therefore, $\hat{\mathbf{v}} = \mathbf{v} - I_{k-1}^k\mathbf{z}$.) We have

$$\begin{aligned} \|\hat{\mathbf{v}}\|^2_{(L^2(\Omega))^2} = &\left\{ (\hat{\mathbf{v}}, \hat{\mathbf{v}})_{(L^2(\Omega))^2} - \sum_{T\in\mathcal{T}^k} \int_T \nabla\mathbf{r} \cdot \nabla(\mathbf{v} - \mathbf{z})\, dx \right\} \\ &+ \sum_{T\in\mathcal{T}^k} \int_T \nabla\mathbf{r} \cdot \nabla(\mathbf{v} - \mathbf{z})\, dx. \end{aligned} \tag{6.13}$$

Since $\mathbf{z} \in V_{k-1}$, by using the definitions of $\mathbf{r}_k$, $\mathbf{r}_{k-1}$, $a_k(\cdot,\cdot)$, and $a_{k-1}(\cdot,\cdot)$, we can rewrite the first term on the right-hand side of (6.13) as follows:

$$\begin{aligned} &(\hat{\mathbf{v}}, \hat{\mathbf{v}})_{(L^2(\Omega))^2} - \sum_{T\in\mathcal{T}^k} \int_T \nabla\mathbf{r} \cdot \nabla(\mathbf{v} - \mathbf{z})\, dx \\ &= (\hat{\mathbf{v}}, \mathbf{v} - I_{k-1}^k\mathbf{z})_{(L^2(\Omega))^2} - \sum_{T\in\mathcal{T}^k} \int_T \nabla\mathbf{r} \cdot \nabla\mathbf{v}\, dx \\ &\quad + \sum_{T\in\mathcal{T}^{k-1}} \int_T \nabla\mathbf{r} \cdot \nabla\mathbf{z}\, dx \\ &= a_k(\mathbf{r}_k, \mathbf{v}) - (\hat{\mathbf{v}}, I_{k-1}^k\mathbf{z})_{(L^2(\Omega))^2} - a_k(\mathbf{r}, \mathbf{v}) + a_{k-1}(\mathbf{r}, \mathbf{z}) \\ &= a_k(\mathbf{r}_k - \mathbf{r}, \mathbf{v}) - (\hat{\mathbf{v}}, \mathbf{z})_{(L^2(\Omega))^2} + (\hat{\mathbf{v}}, \mathbf{z} - I_{k-1}^k\mathbf{z})_{(L^2(\Omega))^2} + a_{k-1}(\mathbf{r}, \mathbf{z}) \\ &= a_k(\mathbf{r}_k - \mathbf{r}, \mathbf{v}) + a_{k-1}(\mathbf{r} - \mathbf{r}_{k-1}, \mathbf{z}) + (\hat{\mathbf{v}}, \mathbf{z} - I_{k-1}^k\mathbf{z})_{(L^2(\Omega))^2}. \end{aligned} \tag{6.14}$$

Using the Cauchy-Schwarz inequality, (6.12), the definition of $\mathbf{z}$, Lemma 8, and (1.13), we have

$$\begin{aligned} &\left| (\hat{\mathbf{v}}, \hat{\mathbf{v}})_{(L^2(\Omega))^2} - \sum_{T\in\mathcal{T}^k} \int_T \nabla\mathbf{r} \cdot \nabla(\mathbf{v} - \mathbf{z})\, dx \right| \\ &\le \|\mathbf{r}_k - \mathbf{r}\|_k \|\mathbf{v}\|_k + \|\mathbf{r} - \mathbf{r}_{k-1}\|_{k-1} \|P_k^{k-1}\mathbf{v}\|_{k-1} \\ &\quad + \|\hat{\mathbf{v}}\|_{(L^2(\Omega))^2} \|\mathbf{z} - I_{k-1}^k\mathbf{z}\|_{(L^2(\Omega))^2} \\ &\le Ch_k\|\hat{\mathbf{v}}\|_{(L^2(\Omega))^2} \|\mathbf{v}\|_k. \end{aligned} \tag{6.15}$$

By using the definitions of $a_k(\cdot,\cdot)$, $a_{k-1}(\cdot,\cdot)$, $\mathbf{z}$, and P_{k-1}^k, the remaining term on the right-hand side of (6.13) can be rewritten as follows:

$$\begin{aligned}\sum_{T\in\mathscr{T}^k}\int_T \nabla\mathbf{r}\cdot\nabla(\mathbf{v}-\mathbf{z})\,dx &= a_k(\mathbf{r},\mathbf{v})-a_{k-1}(\mathbf{r},\mathbf{z})\\ &= a_k(\mathbf{r}-\Pi_k\mathbf{r},\mathbf{v})+a_k(\Pi_k\mathbf{r}-I_{k-1}^k\Pi_{k-1}\mathbf{r},\mathbf{v})\\ &\quad + a_k(I_{k-1}^k\Pi_{k-1}\mathbf{r},\mathbf{v})-a_{k-1}(\mathbf{r},\mathbf{z})\\ &= a_k(\mathbf{r}-\Pi_k\mathbf{r},\mathbf{v})+a_k(\Pi_k\mathbf{r}-I_{k-1}^k\Pi_{k-1}\mathbf{r},\mathbf{v})\\ &\quad + a_{k-1}(\Pi_{k-1}\mathbf{r}-\mathbf{r},P_k^{k-1}\mathbf{v}).\end{aligned}\tag{6.16}$$

The interpolation error estimate (2.13), (1.14), (6.2), and (6.10) imply that

$$\begin{aligned}\left|\sum_{T\in\mathscr{T}^k}\int_T \nabla\mathbf{r}\cdot\nabla(\mathbf{v}-\mathbf{z})\,dx\right| &\le Ch_k|\mathbf{r}|_{(H^2(\Omega))^2}\|\mathbf{v}\|_k\\ &\le Ch_k\|\hat{\mathbf{v}}\|_{(L^2(\Omega))^2}\|\mathbf{v}\|_k.\end{aligned}\tag{6.17}$$

Inequality (6.8) now follows from (6.13), (6.15), and (6.17). □

Corollary 2 (Approximation property II). *There exists a positive constant C such that*

$$|||(I-I_{k-1}^kP_k^{k-1})\mathbf{v}|||_{2,k}\le Ch_k|||\mathbf{v}|||_{3,k}\quad\forall\mathbf{v}\in V_k.\tag{6.18}$$

Proof. From (6.1), (5.2), and Lemma 11, we have

$$\begin{aligned}|||(I-I_{k-1}^kP_k^{k-1})\mathbf{v}|||_{2,k} &= \sup_{\mathbf{w}\in V_k\setminus\{0\}}\frac{|a_k((I-I_{k-1}^kP_k^{k-1})\mathbf{v},\mathbf{w})|}{|||\mathbf{w}|||_{2,k}}\\ &= \sup_{\mathbf{w}\in V_k\setminus\{0\}}\frac{|a_k(\mathbf{v},(I-I_{k-1}^kP_k^{k-1})\mathbf{w})|}{|||\mathbf{w}|||_{2,k}}\\ &\le \sup_{\mathbf{w}\in V_k\setminus\{0\}}\frac{|||\mathbf{v}|||_{3,k}|||(I-I_{k-1}^kP_k^{k-1})\mathbf{w}|||_{1,k}}{|||\mathbf{w}|||_{2,k}}\\ &\le Ch_k|||\mathbf{v}|||_{3,k}.\quad\square\end{aligned}$$

Corollary 3. *There exists a positive constant C such that*

$$|||P_k^{k-1}\mathbf{v}|||_{1,k}\le C|||\mathbf{v}|||_{1,k}\quad\forall\mathbf{v}\in V_k.\tag{6.19}$$

Proof. By Proposition 3, (6.7), (1.13), and (6.2), we have

$$\begin{aligned}|||P_k^{k-1}\mathbf{v}|||_{1,k} &\le |||P_k^{k-1}\mathbf{v}-I_{k-1}^kP_k^{k-1}\mathbf{v}|||_{1,k}\\ &\quad + |||I_{k-1}^kP_k^{k-1}\mathbf{v}-\mathbf{v}|||_{1,k}+|||\mathbf{v}|||_{1,k}\\ &\le \|P_k^{k-1}\mathbf{v}-I_{k-1}^kP_k^{k-1}\mathbf{v}\|_{(L^2(\Omega))^2}+Ch_k\|\mathbf{v}\|_k+|||\mathbf{v}|||_{1,k}\\ &\le Ch_k\|P_k^{k-1}\mathbf{v}\|_{k-1}+Ch_k\|\mathbf{v}\|_k+|||\mathbf{v}|||_{1,k}\\ &\le Ch_k\|\mathbf{v}\|_k+|||\mathbf{v}|||_{1,k}.\end{aligned}$$

On the other hand, from the definition of $|||\cdot|||_{s,k}$ and the fact that the spectral radius of A_k is bounded by Ch_k^{-4} (cf. (4.6)), we have $h_k\|\mathbf{v}\|_k \le C|||\mathbf{v}|||_{1,k}$ for all $\mathbf{v} \in V_k$. $\square$

Theorem 3 (Convergence of the two-grid algorithm). *There exists a positive constant C such that*

$$\|\mathbf{e}\|_k \le Cm^{-1/4}\|\mathbf{e}_0\|_k \tag{6.20}$$

and

$$\|\mathbf{e}\|_{(L^2(\Omega))^2} \le Cm^{-1/4}\|\mathbf{e}_0\|_{(L^2(\Omega))^2}. \tag{6.21}$$

Therefore, the two-grid algorithm is a contraction if m is large enough.

Proof. By (6.5), (6.18), and (6.6) with $\beta = 3$, we have

$$\begin{aligned}\|\mathbf{e}\|_k &= \|(I - I_{k-1}^k P_k^{k-1})R_k^m \mathbf{e}_0\|_k \\ &\le Ch_k|||R_k^m\mathbf{e}_0|||_{3,k} \le Cm^{-1/4}\|\mathbf{e}_0\|_k.\end{aligned}$$

By (6.5), (6.8), and (6.6) with $\beta = 2$, we have

$$\begin{aligned}\|\mathbf{e}\|_{(L^2(\Omega))^2} &= \|(I - I_{k-1}^k P_k^{k-1})R_k^m \mathbf{e}_0\|_{(L^2(\Omega))^2} \\ &\le Ch_k|||R_k^m\mathbf{e}_0|||_{2,k} \le Cm^{-1/4}|||\mathbf{e}_0|||_{1,k}.\end{aligned} \tag{6.22}$$

Inequality (6.21) now follows from (6.22) by Proposition 3. $\square$

Theorem 4 (Convergence of the kth-level iteration). *There exists a positive constant C such that when the kth-level iteration is applied to $A_k\mathbf{z} = \mathbf{g}$, we have*

$$\|\mathbf{z} - MG(k, \mathbf{z}_0, \mathbf{g})\|_k \le Cm^{-1/4}\|\mathbf{z} - \mathbf{z}_0\|_k \tag{6.23}$$

and

$$\|\mathbf{z} - MG(k, \mathbf{z}_0, \mathbf{g})\|_{(L^2(\Omega))^2} \le Cm^{-1/4}\|\mathbf{z} - \mathbf{z}_0\|_{(L^2(\Omega))^2}, \tag{6.24}$$

provided that m is large enough.

Proof. Let C^* be a positive constant which dominates all of the constants in (1.12), (3.16), (6.2), (6.19), (6.20), and (6.22). Assume that m satisfies

$$(2C^*/m^{1/4})^{p-1} < (2C^{*2})^{-1}, \tag{6.25}$$

and let $\gamma = 2C^*/m^{1/4}$. (Recall that $p = 2$ or 3 in the algorithm.) We shall prove the following inequalities by induction:

$$\|\mathbf{z} - MG(k, \mathbf{z}_0, \mathbf{g})\|_k \le \gamma\|\mathbf{z} - \mathbf{z}_0\|_k \tag{6.26}$$

and

$$\|\mathbf{z} - MG(k, \mathbf{z}_0, \mathbf{g})\|_{(L^2(\Omega))^2} \le \gamma|||\mathbf{z} - \mathbf{z}_0|||_{1,k}. \tag{6.27}$$

Note that (6.24) follows from (6.27) by Proposition 3.

For $k = 1$, (6.26) and (6.27) hold because $MG(1, \mathbf{z}_0, \mathbf{g}) = A_1^{-1}\mathbf{g} = \mathbf{z}$. Assume that (6.26) and (6.27) hold for $k \le n - 1$.

Let $\mathbf{e}_i = \mathbf{z} - \mathbf{z}_i$, $0 \le i \le m$. Then $\mathbf{e}_m = R_n^m \mathbf{e}_0$. We have

$$\begin{aligned} \mathbf{z} - MG(n, \mathbf{z}_0, \mathbf{g}) &= \mathbf{z} - (\mathbf{z}_m + I_{n-1}^n \mathbf{q}_p) \\ &= \mathbf{z} - (\mathbf{z}_m + I_{n-1}^n \mathbf{q}) + I_{n-1}^n (\mathbf{q} - \mathbf{q}_p), \end{aligned} \tag{6.28}$$

where $\mathbf{q} = P_n^{n-1} \mathbf{e}_m$ (cf. Lemma 9) satisfies $A_{n-1}\mathbf{q} = \bar{\mathbf{g}}$ and $\mathbf{q}_p$ is the approximation of $\mathbf{q}$ obtained by applying the $(n-1)$-level iteration p times. From (1.12), the induction hypothesis, Lemma 8, and (6.24), it follows that

$$\begin{aligned} \|I_{n-1}^n (\mathbf{q} - \mathbf{q}_p)\|_n &\le C^* \gamma^p \|\mathbf{q}\|_{n-1} = C^* \gamma^p \|P_n^{n-1} \mathbf{e}_m\|_{n-1} \\ &\le (C^*)^2 \gamma^p \|R_n^m \mathbf{e}_0\|_n \le \tfrac{\gamma}{2} \|\mathbf{e}_0\|_n. \end{aligned} \tag{6.29}$$

Since $\mathbf{z} - (\mathbf{z}_m + I_{n-1}^n \mathbf{q})$ is the final error of the two-grid algorithm, it follows from (6.20) and the choice of γ that

$$\|\mathbf{z} - (\mathbf{z}_m + I_{n-1}^n \mathbf{q})\|_n \le C^* m^{-1/4} \|\mathbf{e}_0\|_n = \tfrac{\gamma}{2} \|\mathbf{e}_0\|_n. \tag{6.30}$$

Combining (6.28), (6.29), and (6.30), we see that (6.26) holds for $k = n$.

On the other hand, by (3.16), (6.19), and the induction hypothesis, we have

$$\begin{aligned} \|I_{n-1}^n (\mathbf{q} - \mathbf{q}_p)\|_{(L^2(\Omega))^2} &\le C^* \gamma^p |||\mathbf{q}|||_{1,k-1} = C^* \gamma^p |||P_n^{n-1} \mathbf{e}_m|||_{1,k-1} \\ &\le (C^*)^2 \gamma^p |||R_n^m \mathbf{e}_0|||_{1,k} \\ &\le (C^*)^2 \gamma^p |||\mathbf{e}_0|||_{1,k} \le \tfrac{\gamma}{2} |||\mathbf{e}_0|||_{1,k}. \end{aligned} \tag{6.31}$$

It also follows from (6.22) and the choice of γ that

$$\|\mathbf{z} - (\mathbf{z}_m + I_{n-1}^n \mathbf{q})\|_{(L^2(\Omega))^2} \le C^* m^{-1/4} |||\mathbf{e}_0|||_{1,k} \le \tfrac{\gamma}{2} |||\mathbf{e}_0|||_{1,k}. \tag{6.32}$$

Therefore, (6.27) holds for $k = n$ by combining (6.28), (6.31), and (6.32). □

Theorem 5 (Full multigrid convergence). *If m is chosen so that the kth-level iteration is a contraction for $k = 1, 2, \ldots$ and the parameter r in the full multigrid algorithm is chosen large enough, then*

$$\begin{aligned} &\|\mathbf{u}_k - \hat{\mathbf{u}}_k\|_{(L^2(\Omega))^2} + h_k \|\mathbf{u}_k - \hat{\mathbf{u}}_k\|_k \\ &\quad \le C h_k^2 (|\mathbf{u}|_{(H^2(\Omega))^2} + |p|_{H^1(\Omega)}) \quad \text{for } k \ge 1, \end{aligned} \tag{6.33}$$

where $(\mathbf{u}, p)$ is the solution of (1.1), $\mathbf{u}_k$ is the exact solution of the discretized problem (1.9), and $\hat{\mathbf{u}}_k$ is the approximate solution of (1.9) obtained from the full multigrid algorithm.

Proof. It suffices to prove that

$$\begin{aligned} &\|\mathbf{u}_k - I_{k-1}^k \mathbf{u}_{k-1}\|_{(L^2(\Omega))^2} + h_k \|\mathbf{u}_k - I_{k-1}^k \mathbf{u}_{k-1}\|_k \\ &\quad \le C h_k^2 (|\mathbf{u}|_{(H^2(\Omega))^2} + |p|_{H^1(\Omega)}). \end{aligned} \tag{6.34}$$

Theorem 4 and a standard argument (cf. [14, Theorem 7.1, p. 162]) will then prove (6.33).

The discretization error estimate (1.10), the interpolation error estimate (2.13), properties (1.12), (1.14), and (3.16) imply that

$$\begin{aligned}
&\|\mathbf{u}_k - I_{k-1}^k \mathbf{u}_{k-1}\|_{(L^2(\Omega))^2} + h_k \|\mathbf{u}_k - I_{k-1}^k \mathbf{u}_{k-1}\|_k \\
&\quad \le \|\mathbf{u}_k - \Pi_k \mathbf{u}\|_{(L^2(\Omega))^2} + h_k \|\mathbf{u}_k - \Pi_k \mathbf{u}\|_k) \\
&\qquad + (\|\Pi_k \mathbf{u} - I_{k-1}^k(\Pi_{k-1}\mathbf{u})\|_{(L^2(\Omega))^2} + h_k \|\Pi_k \mathbf{u} - I_{k-1}^k(\Pi_{k-1}\mathbf{u})\|_k) \\
&\qquad + (\|I_{k-1}^k(\Pi_{k-1}\mathbf{u} - \mathbf{u}_{k-1})\|_{(L^2(\Omega))^2} + h_k \|I_{k-1}^k(\Pi_{k-1}\mathbf{u} - \mathbf{u}_{k-1})\|_k) \\
&\quad \le C h_k^2 (|\mathbf{u}|_{(H^2(\Omega))^2} + |p|_{H^1(\Omega)}) \\
&\qquad + C(\|\Pi_{k-1}\mathbf{u} - \mathbf{u}_{k-1}\|_{(L^2(\Omega))^2} + h_k \|\Pi_{k-1}\mathbf{u} - \mathbf{u}_{k-1}\|_k) \\
&\quad \le C h_k^2 (|\mathbf{u}|_{(H^2(\Omega))^2} + |p|_{H^1(\Omega)}). \quad \square
\end{aligned}$$

BIBLIOGRAPHY

1. R. A. Adams, *Sobolev spaces*, Academic Press, New York, 1975.
2. R. E. Bank and T. Dupont, *An optimal order process for solving finite element equations*, Math. Comp. **36** (1981), 35–51.
3. D. Braess and R. Verfürth, *Multi-grid methods for non-conforming finite element methods*, preprint number 453, Universität Heidelberg, March 1988.
4. J. H. Bramble and S. R. Hilbert, *Estimation of linear functionals on Sobolev spaces with application to Fourier transforms and spline interpolation*, SIAM J. Numer. Anal. **7** (1970), 113–124.
5. J. H. Bramble, J. E. Pasciak, and J. Xu, *The analysis of multigrid algorithms with nonnested spaces or noninherited quadratic forms*, Math. Comp. **56** (1991), to appear.
6. S. C. Brenner, *An optimal-order multigrid method for* P1 *nonconforming finite elements*, Math. Comp. **52** (1989), 1–15.
7. ——, *An optimal-order nonconforming multigrid method for the biharmonic equation*, SIAM J. Numer. Anal. **26** (1989), 1124–1138.
8. ——, *Multigrid methods for nonconforming finite elements*, Dissertation, Univ. of Michigan, 1988.
9. P. G. Ciarlet, *The finite element method for elliptic problems*, North-Holland, Amsterdam, New York, and Oxford, 1978.
10. M. Crouzeix and P.-A. Raviart, *Conforming and nonconforming finite element methods for solving the stationary Stokes equations*. I, RAIRO **R-3** (1973), 33–75.
11. V. Girault and P.-A. Raviart, *Finite element methods for Navier-Stokes equations*, Springer-Verlag, Berlin and Heidelberg, 1986.
12. W. Hackbusch, *Multi-grid methods and applications*, Springer-Verlag, Berlin and Heidelberg, 1985.

R. B. Kellogg and J. E. Osborn, *A regularity result for the Stokes problem on a convex polygon*, J. Funct. Anal. **21** (1976), 397–431.

14. S. McCormick (ed.), *Multigrid methods*, Frontiers in Applied Math., vol. 3, SIAM, Philadelphia, 1987.
15. L. S. D. Morley, *The triangular equilibrium problem in the solution of plate bending problems*, Aero. Quart. **19** (1968), 149–169.

P. Peisker and D. Braess, *A conjugate gradient method and a multigrid algorithm for Morley's finite element approximation of the biharmonic equation*, Numer. Math. **50** (1987), 567–586.

17. F. Thomasset, *Implementation of finite element methods for Navier-Stokes equations*, Springer-Verlag, New York, 1981.

18. S. Zhang, *Multi-level iterative techniques*, Dissertation, Pennsylvania State Univ., 1988.

DEPARTMENT OF MATHEMATICS, SYRACUSE UNIVERSITY, SYRACUSE, NEW YORK 13244

Current address: Department of Mathematics and Computer Science, Clarkson University, Potsdam, New York 13699

E-mail address: brenner@sun.mcs.clarkson.edu