MATHEMATICS OF COMPUTATION
VOLUME 52, NUMBER 185
JANUARY 1989, PAGES 1–15

An Optimal-Order Multigrid Method for P1 Nonconforming Finite Elements

By Susanne C. Brenner

Abstract. An optimal-order multigrid method for solving second-order elliptic boundary value problems using P1 nonconforming finite elements is developed.

1. Introduction. Let Ω be a convex polygon in $\mathbf{R}^2$. Let $f \in L^2(\Omega)$, $\alpha \in C^1(\overline{\Omega})$ and $\beta \in C^0(\overline{\Omega})$. We assume there exist constants α_0, β_0 such that $\alpha \geq \alpha_0 > 0$ and $\beta \geq \beta_0$, where β_0 depends on the boundary condition. In this paper we develop an optimal-order multigrid method for solving the Dirichlet problem ($\beta_0 = 0$)

$$(1.1)\qquad \begin{aligned} -\nabla \cdot (\alpha \nabla u) + \beta u &= f \quad \text{in } \Omega, \\ u &= 0 \quad \text{on } \partial\Omega \end{aligned}$$

and the Neumann problem ($\beta_0 > 0$)

$$(1.2)\qquad \begin{aligned} -\nabla \cdot (\alpha \nabla u) + \beta u &= f \quad \text{in } \Omega, \\ \frac{\partial u}{\partial n} &= 0 \quad \text{on } \partial\Omega, \end{aligned}$$

using P1 nonconforming finite elements (cf. [3], [4], [10]). We refer the reader to [1], [7] and [9] for conforming multigrid methods. As in [1], our multigrid method will be described in a coordinate-free fashion.

Our method consists of smoothing on the current-grid and coarser-grid correction, as in the conforming multigrid method. The important difference in the nonconforming case is that $V_{k-1} \not\subseteq V_k$, where V_k's are the finite element spaces on mesh level k. Hence we can no longer simply use the natural injection for the intergrid transfer of grid functions. The key idea is to define an operator $I_{k-1}^k \colon V_{k-1} \to V_k$ that reduces to natural injection on continuous piecewise linear functions. By doing so, we can use the well-known analysis of the conforming multigrid method. We will show that the approximate solution satisfies the same type of error estimates as the discretization error and that it can be obtained in $\mathscr{O}(n)$ steps, where n is the dimension of the discretized finite element space. Since our intergrid transfer operator does not preserve either the energy or the L^2-norm, the standard proof of convergence (cf. [2]) for the $\mathscr{V}$-cycle does not carry over directly. We will therefore only discuss a $\mathscr{W}$-cycle method, even though the $\mathscr{V}$-cycle method may be convergent.

The paper is organized as follows. We begin with a discussion of the notation and fundamental estimates from the theory of finite elements. The intergrid transfer operator is discussed in Section 3. Section 4 contains the results on the contracting

Received July 2, 1987; revised February 2, 1988.
1980 *Mathematics Subject Classification* (1985 *Revision*). Primary 65N30, 65F10.

©1989 American Mathematical Society
0025-5718/89 $1.00 + $.25 per page

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

property of the kth level iteration followed by the convergence theorems for the nested iteration in Section 5. The singular Neumann problem ($\beta_0 = 0$ in (1.2)) will be treated in Section 6. A piecewise quadratic nonconforming finite element is discussed in the last section.

2. Preliminaries and Notation. Let $V = W_2^1(\Omega)$ for the Neumann problem (1.2) and $V = \{v \in W_2^1(\Omega)\colon v|_{\partial\Omega} = 0\}$ for the Dirichlet problem (1.1). Here, $W_2^1(\Omega)$ denotes the usual Sobolev space (cf. [3]). The variational formulation for (1.1) and (1.2) is to find $u \in V$ such that

$$a(u,v) = F(v) \quad \forall v \in V, \tag{2.1}$$

where

$$a(u,v) = \int_\Omega \alpha \nabla u \cdot \nabla v + \beta uv \quad \text{and} \quad F(v) = \int_\Omega fv. \tag{2.2}$$

Let $\{\mathscr{T}^k\}$, $k \geq 1$, be a family of triangulations of Ω, where $\mathscr{T}^{k+1}$ is obtained by connecting the midpoints of the edges of the triangles in $\mathscr{T}^k$. Let $h_k := \max_{T \in \mathscr{T}^k} \operatorname{diam} T$ (therefore $h_k = 2h_{k+1}$). Then there exist positive constants C_1, C_2, independent of k, such that

$$C_2 h_k^2 \leq |T| \leq C_1 h_k^2 \quad \forall T \in \mathscr{T}^k, \tag{2.3}$$

where $|T|$ denotes the area of the triangle T. These constants depend only on the angles that appear in $\mathscr{T}^1$. Throughout this paper we let C and C_i denote generic constants independent of k.

For the Dirichlet problem (1.1), define the finite element space

$$\begin{aligned} V_k := \{v\colon v|_T &\text{ is linear for all } T \in \mathscr{T}^k, v \text{ is continuous at the midpoints} \\ &\text{of the edges and } v = 0 \text{ at the midpoints on } \partial\Omega\}. \end{aligned} \tag{2.4}$$

For the Neumann problem (1.2), we define V_k similarly but without any restrictions on v along the boundary of Ω. Note that functions in V_k are not continuous. In other words, V_k is a nonconforming finite element space.

We also use a conforming finite element space for our analysis. Define

$$\begin{aligned} W_k := \{w\colon w|_T &\text{ is linear for all } T \in \mathscr{T}^k, w \text{ is continuous} \\ &\text{on } \Omega \text{ and } w|_{\partial\Omega} = 0\} \end{aligned} \tag{2.5}$$

for the Dirichlet problem (1.1). For the Neumann problem (1.2), we make no assumptions on w along the boundary of Ω. The space W_k will only be used to obtain our theoretical estimates. We emphasize that it will *not* play any role in the actual multigrid algorithm. Observe that $W_k = V_k \cap V = V_k \cap V_{k+1}$.

Let $\{\phi_1^k, \ldots, \phi_{n_k}^k\}$ be the basis of V_k such that each ϕ_j^k equals 1 at exactly one midpoint and equals 0 at all other midpoints. For any linear functions ψ, ϕ on a triangle K,

$$\int_K \psi\phi = \frac{1}{3}|K| \left(\sum_{i=1}^{3} \psi(m_i)\phi(m_i) \right), \tag{2.6}$$

where the m_i's are the midpoints of the sides of K (cf. [3, p. 183]). It follows that the ϕ_i's are orthogonal with respect to the L^2-inner product.

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

Similarity is an equivalence relation on triangles. For each equivalence class $\mathscr{R}$, there exist constants $C_{\mathscr{R}} > 0$ and $C'_{\mathscr{R}} > 0$ such that for any triangle $T \in \mathscr{R}$ and $v \in \mathscr{P}_1(T)$, the space of first-degree polynomials on T, we have

$$C_{\mathscr{R}}\Theta(v) \le \int_T |\nabla v|^2 \le C'_{\mathscr{R}}\Theta(v).$$

Here, $\Theta(v) = [v(m_1)-v(m_2)]^2+[v(m_2)-v(m_3)]^2+[v(m_3)-v(m_1)]^2$ and m_1, m_2, m_3 are the midpoints of the sides of T. Since any triangle in $\mathscr{T}^k$ $(k = 1, 2, \dots)$ is similar to a triangle in $\mathscr{T}^1$, there exist $C_3, C_4 > 0$ such that

$$C_3\Theta(v) \le \int_T |\nabla v|^2 \le C_4\Theta(v) \tag{2.7}$$

for any $v \in \mathscr{P}_1(T)$, $T \in \mathscr{T}^k$, $k = 1, 2, \dots$. Moreover, as a consequence of (2.3), (2.6) and (2.7), there exists $C > 0$ such that

$$\int_T |\nabla v|^2 \le C h_k^{-2} \|v\|^2_{L^2} \tag{2.8}$$

for $v \in \mathscr{P}_1(T)$ and $T \in \mathscr{T}^k$.

For each k, define (on $V_k + W_2^1(\Omega)$)

$$a_k(u, v) := \sum_{T \in \mathscr{T}^k} \int_T (\alpha \nabla u \cdot \nabla v + \beta u v) \tag{2.9}$$

and

$$\|u\|_k := \sqrt{a_k(u, u)}. \tag{2.10}$$

The bilinear form $a_k(\cdot, \cdot)$ is obviously symmetric and positive definite on V_k. The stiffness matrix representing $a_k(\cdot, \cdot)$ with respect to the basis $\{\phi_1^k, \dots, \phi_{n_k}^k\}$ has at most five entries per row. As a consequence of (2.8), we have

$$\|u\|_k \le C h_k^{-1} \|u\|_{L^2} \quad \forall u \in V_k. \tag{2.11}$$

We also note that if $u, v \in W_2^1(\Omega)$, then $a_k(u, v) = a(u, v)$.

We now recall some fundamental estimates from the theory of finite elements.

Let Π_k and $\tilde{\Pi}_k$ be the interpolation operators associated with V_k and W_k, respectively. If $u \in W_2^2(\Omega)$, we have the following estimates for the interpolation error:

$$\|u - \Pi_k u\|_{L^2} + h_k \|u - \Pi_k u\|_k \le C h_k^2 \|u\|_{W_2^2} \tag{2.12}$$

and

$$\|u - \tilde{\Pi}_k u\|_{L^2} + h_k \|u - \tilde{\Pi}_k u\|_k \le C h_k^2 \|u\|_{W_2^2} \tag{2.13}$$

(cf. [3]).

Since $f \in L^2(\Omega)$, elliptic regularity implies that $u \in W_2^2(\Omega)$ (cf. [6]). For the same f, let $u_k \in V_k$ satisfy

$$a_k(u_k, v) = \int_\Omega f v \quad \forall v \in V_k$$

and let $\tilde{u}_k \in W_k$ satisfy

$$a_k(\tilde{u}_k, v) = \int_\Omega f v \quad \forall v \in W_k.$$

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

Since V_k satisfies the patch test (cf. [8], [10]), we have the following estimate for the discretization error:

$$\|u - u_k\|_{L^2} + h_k\|u - u_k\|_k \le Ch_k^2\|u\|_{W_2^2} \tag{2.14}$$

(cf. [4], [10]). The estimate for the conforming discretization error is, of course, well known (cf. [3]):

$$\|u - \tilde{u}_k\|_{L^2} + h_k\|u - \tilde{u}_k\|_k \le Ch_k^2\|u\|_{W_2^2}. \tag{2.15}$$

In [1], it was shown that $\tilde{u}_k$ could be calculated by an iterative procedure to within an accuracy comparable to the error estimated by (2.15) using an amount of work that is proportional to the number of unknowns, namely the dimension of W_k. Our main goal in this paper is to prove a corresponding result for the computation of u_k.

From the spectral theorem, there exist eigenvalues $0 < \lambda_1 \le \lambda_2 \le \cdots \le \lambda_{n_k}$ and eigenfunctions $\psi_1, \psi_2, \ldots, \psi_{n_k} \in V_k$, $(\psi_i, \psi_j)_{L^2} = \delta_{ij}$ (= the Kronecker delta), such that $a_k(\psi_i, v) = \lambda_i(\psi_i, v)_{L^2}$ for all $v \in V_k$. From (2.11), there exists $C_5 > 0$ such that

$$\lambda_i \le C_5 h_k^{-2}. \tag{2.16}$$

If $v \in V_k$, we can write $v = \sum_{i=1}^{n_k} \nu_i\psi_i$. The norm $|||v|||_{s,k}$ is defined (cf. [1]) as follows:

$$|||v|||_{s,k} := \left(\sum_{i=1}^{n_k} \nu_i^2\lambda_i^s\right)^{1/2}. \tag{2.17}$$

Note that $|||v|||_{0,k} = \|v\|_{L^2}$ and $|||v|||_{1,k} = \|v\|_k$.

Finally, it follows from the Cauchy-Schwarz inequality that

$$|a_k(v, w)| \le |||v|||_{1+t,k}|||w|||_{1-t,k} \tag{2.18}$$

for any $t \in \mathbf{R}$ and $v, w \in V_k$.

3. The Intergrid Transfer Operator. For $v \in V_{k-1}$ the intergrid transfer operator $I_{k-1}^k \colon V_{k-1} \to V_k$ is defined as follows. Let p be a midpoint of a side of a triangle in $\mathscr{T}^k$. If p lies in the interior of a triangle in $\mathscr{T}^{k-1}$, then we define

$$(I_{k-1}^k v)(p) := v(p).$$

Otherwise, if p lies on the common edge of two adjacent triangles T_1 and T_2 in $\mathscr{T}^{k-1}$, then we define

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

Note that the matrix for I_{k-1}^k with respect to the bases $\{\phi_1^{k-1}, \ldots, \phi_{n_{k-1}}^{k-1}\}$ and $\{\phi_1^k, \ldots, \phi_{n_k}^k\}$ has at most five entries per row.

From the definition of I_{k-1}^k, it is clear that

$$I_{k-1}^k v = v \quad \forall v \in W_{k-1} = V_k \cap V_{k-1} \subseteq V. \tag{3.1}$$

In other words, $I_{k-1}^k|_{W_{k-1}}$ is just the natural injection.

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

LEMMA 1. *There exists $C > 0$ such that*

$$\|I_{k-1}^k v\|_k \le C\|v\|_{k-1} \tag{3.2}$$

and

$$\|I_{k-1}^k v\|_{L^2} \le C\|v\|_{L^2} \tag{3.3}$$

for all $v \in V_{k-1}$.

Proof. The inequality (3.3) follows immediately from (2.6), the definition of I_{k-1}^k and the quasi-uniformity of the triangulations.

Inequality (3.2) can be deduced from (3.3) as follows. Given $v \in V_{k-1}$, define $g \in V_{k-1}$ by

$$\int_\Omega g\phi = a_{k-1}(v,\phi) \quad \forall \phi \in V_{k-1}, \tag{3.4}$$

$w \in W_{k-1}$ by

$$a_k(w,\phi) = \int_\Omega g\phi \quad \forall \phi \in W_{k-1}, \tag{3.5}$$

and $z \in V$ by

$$a(z,\phi) = \int_\Omega g\phi \quad \forall \phi \in V. \tag{3.6}$$

Then (3.3), (2.11), (2.14), (2.15) and elliptic regularity imply that

$$\begin{aligned} \|I_{k-1}^k v\|_k &\le \|I_{k-1}^k(v-w)\|_k + \|w\|_k \\ &\le Ch_k^{-1}\|I_{k-1}^k(v-w)\|_{L^2} + \|w-v\|_{k-1} + \|v\|_{k-1} \\ &\le Ch_k^{-1}\|v-w\|_{L^2} + \|v\|_{k-1} \\ &\le Ch_k^{-1}[\|v-z\|_{L^2} + \|w-z\|_{L^2}] + \|v\|_{k-1} \\ &\le Ch_k\|z\|_{W_2^2} + \|v\|_{k-1} \le Ch_k\|g\|_{L^2} + \|v\|_{k-1}. \end{aligned} \tag{3.7}$$

But

$$\|g\|_{L^2}^2 = a_{k-1}(v,g) \le \|v\|_{k-1}\|g\|_{k-1} \le Ch_{k-1}^{-1}\|v\|_{k-1}\|g\|_{L^2}.$$

Therefore,

$$\|g\|_{L^2} \le Ch_{k-1}^{-1}\|v\|_{k-1}. \tag{3.8}$$

Combining (3.7) and (3.8), we obtain (3.2). □

4. Contracting Properties of the kth Level Iteration. The kth level iteration with initial guess z_0 yields $MG(k, z_0, G)$ as an approximate solution to the following problem.

Find $z \in V_k$ such that $a_k(z,v) = G(v)\ \forall v \in V_k$, where $G \in V_k'$.

For $k = 1, MG(1, z_0, G)$ is the solution obtained from a direct method. For $k > 1, MG(k, z_0, G) = z_m + I_{k-1}^k q_p$, where the approximation $z_m \in V_k$ is constructed recursively from the initial guess z_0 and the equations

$$(z_i - z_{i-1}, v)_{L^2} = (\Lambda_k)^{-1}(G(v) - a_k(z_{i-1}, v)), \quad \forall v \in V_k,\ 1 \le i \le m.$$

Here, $\Lambda_k = C_5 h_k^{-2}$ (cf. (2.16)), which is greater than or equal to $\max_{1\le i\le n_k}\lambda_i$, and m is an integer to be determined later. With respect to $\{\phi_1^k, \dots, \phi_{n_k}^k\}, z_m$

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

can be obtained from z_0 by iterating a sparse band matrix because the ϕ_i^k's are L^2-orthogonal. The coarser-grid correction $q_p \in V_{k-1}$ is obtained by applying the $(k-1)$-level iteration p times $(p = 2, 3)$. More precisely,

$$q_0 = 0,$$
$$q_i = MG(k-1, q_{i-1}, \overline{G}), \qquad 1 \le i \le p,$$

where $\overline{G} \in V'_{k-1}$ is defined by $\overline{G}(v) := G(I_{k-1}^k v) - a_k(z_m, I_{k-1}^k v)\ \forall v \in V_{k-1}$.

The main result in this section is the following theorem.

THEOREM 1. *If the number of smoothing steps is large enough, then the kth level iteration is a contraction for both the energy norm and the L^2-norm. Moreover, the contraction number is independent of k.*

Theorem 1 is a trivial consequence of the following lemmas. In order to simplify the notation, we define the statements

(S_k): When the kth level iteration is applied to the problem of finding $z \in V_k$ such that $a_k(z, v) = G(v)\ \forall v \in V_k$, we have $\|z - MG(k, z_0, G)\|_k \le \gamma \|z - z_0\|_k$,

and

$(\overline{S}_k)$: When the kth level iteration is applied to the problem of finding $z \in V_k$ such that $a_k(z, v) = G(v)\ \forall v \in V_k$, we have $\|z - MG(k, z_0, G)\|_{L^2} \le \overline{\gamma} \|z - z_0\|_{L^2}$.

LEMMA 2. *There exists $\gamma \in (0, 1)$ and an integer $m \ge 1$, both independent of k, such that*

$$(S_{k-1}) \quad \textit{implies} \quad (S_k).$$

LEMMA 3. *There exists $\overline{\gamma} \in (0, 1)$ and an integer $m \ge 1$, both independent of k, such that*

$$(\overline{S}_{k-1}) \quad \textit{implies} \quad (\overline{S}_k).$$

Our analysis is based on estimates of the following errors. Let $e_0 := z - z_0$, $e_m := z - z_m$ and $e_f := z - MG(k, z_0, G)$. Also let $e \in V_{k-1}$ satisfy

$$a_{k-1}(e, v) = \overline{G}(v) = a_k(e_m, I_{k-1}^k v) \quad \forall v \in V_{k-1}, \tag{4.1}$$

and let $\tilde{e} \in W_{k-1}$ satisfy

$$a_{k-1}(\tilde{e}, v) = \overline{G}(v) = a_k(e_m, I_{k-1}^k v) = a_k(e_m, v) \quad \forall v \in W_{k-1}. \tag{4.2}$$

As in the conforming case (cf. [1]), we have the following effects of the smoothing steps.

LEMMA 4. *There exists $C > 0$ such that*

$$\|e_m\|_{L^2} \le \|e_0\|_{L^2}, \tag{4.3}$$
$$\|e_m\|_k \le \|e_0\|_k \tag{4.4}$$

and

$$|||e_m|||_{2,k} \le C h_k^{-1} m^{-1/2} |||e_0|||_{1,k} = C h_k^{-1} m^{-1/2} \|e_0\|_k. \tag{4.5}$$

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

From (4.1) and (3.2) we have

$$\|e\|_{k-1}^2 = a_k(e_m, I_{k-1}^k e) \leq \|e_m\|_k \|I_{k-1}^k e\|_k \leq C\|e_m\|_k \|e\|_{k-1}.$$

Therefore, there exists $C > 0$ such that

$$\|e\|_{k-1} \leq C\|e_m\|_k \leq C\|e_0\|_k. \tag{4.6}$$

Since e and $\tilde{e}$ are approximate solutions in different spaces to the same problem (cf. (4.1) and (4.2)), they are close to each other.

LEMMA 5. *There exists* $C > 0$ *such that*

$$\|e - \tilde{e}\|_{L^2} \leq Ch_k m^{-1/2}\|e_0\|_{L^2} \tag{4.7}$$

and

$$\|e - \tilde{e}\|_{k-1} \leq Cm^{-1/2}\|e_0\|_k. \tag{4.8}$$

Proof. Let $f_0 \in V_{k-1}$ satisfy

$$(f_0, v)_{L^2} = a_k(e_m, I_{k-1}^k v) \quad \forall v \in V_{k-1}.$$

It follows from (2.18) and (3.2) that

$$\begin{aligned}\|f_0\|_{L^2}^2 &= a_k(e_m, I_{k-1}^k f_0) \leq |||e_m|||_{2,k} |||I_{k-1}^k f_0|||_{0,k} \\ &\leq C|||e_m|||_{2,k}\|f_0\|_{L^2}.\end{aligned}$$

Therefore,

$$\|f_0\|_{L^2} \leq C|||e_m|||_{2,k}. \tag{4.9}$$

Let $v_0 \in W_2^2(\Omega)$ satisfy

$$-\nabla\cdot(\alpha\nabla v_0) + \beta v_0 = f_0 \quad \text{in } \Omega, \qquad v_0 = 0 \quad \text{on } \partial\Omega \qquad \left(\frac{\partial v_0}{\partial n} = 0 \text{ on } \partial\Omega\right)$$

in the Dirichlet (Neumann) case. Note that e and $\tilde{e}$ are the finite element (Galerkin) approximations to v_0 in V_{k-1} and W_{k-1}, respectively. By (2.14) and (2.15), we have $\|v_0 - e\|_{L^2} + h_{k-1}\|v_0 - e\|_{k-1} \leq Ch_{k-1}^2\|v_0\|_{W_2^2}$ and $\|v_0 - \tilde{e}\|_{L^2} + h_{k-1}\|v_0 - \tilde{e}\|_{k-1} \leq Ch_{k-1}^2\|v_0\|_{W_2^2}$.

Since $h_{k-1} = 2h_k$, it follows from the triangle inequality that

$$\|e - \tilde{e}\|_{L^2} + h_k\|e - \tilde{e}\|_{k-1} \leq Ch_k^2\|v_0\|_{W_2^2}.$$

By elliptic regularity, $\|v_0\|_{W_2^2} \leq C\|f_0\|_{L^2}$. Therefore, from (4.9) we obtain

$$\|e - \tilde{e}\|_{L^2} + h_k\|e - \tilde{e}\|_{k-1} \leq Ch_k^2|||e_m|||_{2,k}. \tag{4.10}$$

The inequalities (4.7) and (4.8) now follow from (4.10) and (4.5). □

Next, observe that from (4.2) and the fact that $W_{k-1} \subseteq V$ we have an orthogonality relation,

$$a_k(e_m - \tilde{e}, v) = 0 \quad \forall v \in W_{k-1}. \tag{4.11}$$

The analysis of $e_m - \tilde{e}$ is similar to the one used in conforming multigrid methods.

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

LEMMA 6. *There exists $C > 0$ such that*

$$\|e_m - \tilde{e}\|_{L^2} \leq Cm^{-1/2}\|e_0\|_{L^2} \tag{4.12}$$

and

$$\|e_m - \tilde{e}\|_k \leq Cm^{-1/2}\|e_0\|_k. \tag{4.13}$$

Proof. By (4.11), (2.18) and (4.5) we have

$$\begin{aligned} \|e_m - \tilde{e}\|_k^2 &= a_k(e_m - \tilde{e}, e_m - \tilde{e}) = a_k(e_m - \tilde{e}, e_m) \\ &\leq |||e_m - \tilde{e}|||_{0,k}|||e_m|||_{2,k} \\ &\leq Ch_k^{-1}m^{-1/2}|||e_m - \tilde{e}|||_{0,k}\|e_0\|_k. \end{aligned} \tag{4.14}$$

We will use a duality argument to estimate $|||e_m - \tilde{e}|||_{0,k} = \|e_m - \tilde{e}\|_{L^2}$. Let $w \in W_2^2(\Omega)$ satisfy

$$-\nabla \cdot (\alpha \nabla w) + \beta w = e_m - \tilde{e} \quad \text{in } \Omega, \qquad w = 0 \quad \text{on } \partial\Omega \qquad \left(\frac{\partial w}{\partial n} = 0 \text{ on } \partial\Omega\right)$$

in the Dirichlet (Neumann) case. Let $w_k \in V_k$ satisfy

$$a_k(w_k, v) = \int_\Omega (e_m - \tilde{e})v \quad \forall v \in V_k.$$

Then

$$\begin{aligned} \|e_m - \tilde{e}\|_{L^2}^2 &= \int_\Omega (e_m - \tilde{e})(e_m - \tilde{e}) = a_k(w_k, e_m - \tilde{e}) \\ &= a_k(w - v, e_m - \tilde{e}) + a_k(w_k - w, e_m - \tilde{e}) \end{aligned}$$

for any $v \in W_{k-1}$ by the orthogonality relation (4.11). Therefore, by the approximation property (2.13), the discretization error estimate (2.14) and elliptic regularity we have

$$\begin{aligned} \|e_m - \tilde{e}\|_{L^2}^2 &\leq \inf_{v \in W_{k-1}} \|w - v\|_k \|e_m - \tilde{e}\|_k + \|w_k - w\|_k \|e_m - \tilde{e}\|_k \\ &\leq Ch_k \|w\|_{W_2^2} \|e_m - \tilde{e}\|_k \leq Ch_k \|e_m - \tilde{e}\|_{L^2} \|e_m - \tilde{e}\|_k. \end{aligned}$$

Thus,

$$\|e_m - \tilde{e}\|_{L^2} \leq Ch_k \|e_m - \tilde{e}\|_k. \tag{4.15}$$

The inequality (4.13) now follows by combining (4.14) and (4.15).

From (4.15), (4.13) and (2.11) it follows that

$$\begin{aligned} \|e_m - \tilde{e}\|_{L^2} &\leq Ch_k \|e_m - \tilde{e}\|_k \\ &\leq Cm^{-1/2} h_k \|e_0\|_k \leq Cm^{-1/2} \|e_0\|_{L^2}. \quad \square \end{aligned}$$

Proof of Lemma 2. Recall that $e_f = z - MG(k, z_0, G)$ and $e_0 = z - z_0$. We have by (3.1)

$$\begin{aligned} \|e_f\|_k &= \|e_m - I_{k-1}^k q_p\|_k \\ &\leq \|e_m - \tilde{e}\|_k + \|I_{k-1}^k(\tilde{e} - e)\|_k + \|I_{k-1}^k(e - q_p)\|_k. \end{aligned}$$

From (4.13), (4.8), (3.2) and (S_{k-1}), it follows that

$$\|e_f\|_k \leq Cm^{-1/2}\|e_0\|_k + C\gamma^p \|e\|_{k-1}.$$

But (4.6) yields

$$\|e_f\|_k \leq (Cm^{-1/2} + C\gamma^p)\|e_0\|_k.$$

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

If $\gamma \in (0,1)$ is small enough, then $C\gamma^p < \gamma/2$ (since $p > 1$). If m is large enough, then $Cm^{-1/2} < \gamma/2$. For such choices, we have

$$\|e_f\|_k \le \gamma \|e_0\|_k. \quad \square$$

Proof of Lemma 3. We have by (3.1)

$$\|e_f\|_{L^2} \le \|e_m - \tilde{e}\|_{L^2} + \|I_{k-1}^k(\tilde{e} - e)\|_{L^2} + \|I_{k-1}^k(e - q_p)\|_{L^2}.$$

By (4.12), (4.7), (3.3) and $(\overline{S}_{k-1})$ it follows that

$$\|e_f\|_{L^2} \le Cm^{-1/2}\|e_0\|_{L^2} + C\overline{\gamma}^p \|e\|_{L^2}. \tag{4.16}$$

But

$$\begin{aligned}\|e\|_{L^2} &\le \|e - \tilde{e}\|_{L^2} + \|\tilde{e} - e_m\|_{L^2} + \|e_m\|_{L^2} \\ &\le Cm^{-1/2}\|e_0\|_{L^2} + \|e_0\|_{L^2}\end{aligned}$$

by (4.7), (4.12) and (4.3). Therefore, $\|e\|_{L^2} \le C\|e_0\|_{L^2}$. Hence from (4.16),

$$\|e_f\|_{L^2} \le (Cm^{-1/2} + C\overline{\gamma}^p)\|e_0\|_{L^2}.$$

If $\overline{\gamma} \in (0,1)$ is small enough, then $C\overline{\gamma}^p < \overline{\gamma}/2$. If m is large enough, then $Cm^{-1/2} < \overline{\gamma}/2$. For such choices we have

$$\|e_f\|_{L^2} \le \overline{\gamma}\|e_0\|_{L^2}. \quad \square$$

5. Nested Iteration. We have a sequence of discretizations for the problem (1.1) or (1.2). For each k, we want to find an approximate solution $\hat{u}_k$ to the problem of finding $u_k \in V_k$ such that

$$a_k(u_k, v) = \int_\Omega fv \quad \forall v \in V_k.$$

In the overall multigrid strategy, $\hat{u}_1$ is obtained by a direct method. The approximations $\hat{u}_k$ $(k \ge 2)$ are obtained recursively by

$$\begin{aligned} u_0^j &= I_{j-1}^j \hat{u}_{j-1}, \\ u_l^j &= MG(j, u_{l-1}^j, F), \qquad 1 \le l \le r, \qquad F(v) = \int_\Omega fv, \\ \hat{u}_j &= u_r^j. \end{aligned}$$

Here r is an integer to be determined.

Define $\hat{e}_k := u_k - \hat{u}_k$. In particular, $\hat{e}_1 = 0$. Lemma 2, (2.14), (2.13), (3.1) and (3.2) imply that

$$\begin{aligned}\|\hat{e}_k\|_k &\le \gamma^r \|u_k - I_{k-1}^k \hat{u}_{k-1}\|_k \\ &\le \gamma^r\{\|u_k - u\|_k + \|u - \tilde{\Pi}_{k-1}u\|_k + \|I_{k-1}^k(\tilde{\Pi}_{k-1}u - \hat{u}_{k-1})\|_k\} \\ &\le C\gamma^r\{h_k\|u\|_{W_2^2} + \|\tilde{\Pi}_{k-1}u - \hat{u}_{k-1}\|_{k-1}\} \\ &\le C\gamma^r\{h_k\|u\|_{W_2^2} + \|\tilde{\Pi}_{k-1}u - u\|_{k-1} + \|u - u_{k-1}\|_{k-1} + \|u_{k-1} - \hat{u}_{k-1}\|_{k-1}\} \\ &\le C\gamma^r\{h_k\|u\|_{W_2^2} + h_{k-1}\|u\|_{W_2^2} + \|\hat{e}_{k-1}\|_{k-1}\}.\end{aligned}$$

Since $h_{k-1} = 2h_k$,

$$\|\hat{e}_k\|_k \le Ch_k\gamma^r\|u\|_{W_2^2} + C\gamma^r\|\hat{e}_{k-1}\|_{k-1}. \tag{5.1}$$

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

By iterating (5.1) we have

$$\begin{aligned}\|\hat{e}_k\|_k &\le Ch_k\gamma^r\|u\|_{W_2^2} + C^2h_{k-1}\gamma^{2r}\|u\|_{W_2^2} + \cdots + C^k h_1\gamma^{kr}\|u\|_{W_2^2}\\ &\le \|u\|_{W_2^2}\frac{h_kC\gamma^r}{1-2C\gamma^r}\end{aligned}$$

if $2C\gamma^r < 1$. Therefore,

$$\|\hat{e}_k\|_k \le Ch_k\|u\|_{W_2^2}. \tag{5.2}$$

In summary, we have proved the following theorem.

THEOREM 2. *If r is large enough, then there exists a constant $C > 0$ such that*

$$\|u-\hat{u}_k\|_k \le Ch_k\|u\|_{W_2^2}.$$

Similarly, we can prove the following theorem for the L^2-error.

THEOREM 3. *If r is large enough, then there exists a constant $C > 0$ such that*

$$\|u-\hat{u}_k\|_k \le Ch_k^2\|u\|_{W_2^2}.$$

THEOREM 4. *The cost for obtaining $\hat{u}_k$ is $\mathscr{O}(n_k)$, where n_k is the dimension of V_k.*

Proof. This is a consequence of the fact that $p = 2, 3$ and that the number of nonzero entries in the stiffness matrices, the smoothing iteration matrices and the intergrid transfer matrices are all proportional to n_k. The proof is the same as the one in [1]. □

6. The Singular Neumann Problem. When $\beta \equiv 0$ in (1.2), the necessary and sufficient condition for the existence of a solution is

$$\int_\Omega f = 0. \tag{6.1}$$

If (6.1) is satisfied, there exists a unique solution u in $\hat{V} = \{v \in W_2^1(\Omega)\colon \int_\Omega v = 0\}$. The multigrid method developed in earlier sections can be modified to yield approximate solutions of u. Let $\hat{V}_k = \{v \in V_k\colon \int_\Omega v = 0\}$ and $\hat{W}_k = \{w \in W_k\colon \int_\Omega w = 0\}$. Let $u_k \in \hat{V}_k$ satisfy

$$a_k(u_k, v) = \int_\Omega fv \quad \forall v \in \hat{V}_k \tag{6.2}$$

and $\tilde{u}_k \in \hat{W}_k$ satisfy

$$a_k(\tilde{u}_k, v) = \int_\Omega fv \quad \forall v \in \hat{W}_k. \tag{6.3}$$

Then

$$\begin{aligned}\|u-\tilde{u}_k\|_k &= \inf_{v\in\hat{W}_k}\|u-v\|_k\\ &\le \left\|u - \left(\tilde{\Pi}_k u - \frac{1}{|\Omega|}\int_\Omega \tilde{\Pi}_k u\right)\right\|_k \qquad (6.4)\\ &= \|u-\tilde{\Pi}_k u\|_k \le Ch_k\|u\|_{W_2^2}.\end{aligned}$$

A duality argument shows that

$$\|u-\tilde{u}_k\|_{L^2} \le Ch_k^2\|u\|_{W_2^2}. \tag{6.5}$$

The analog of (2.15) therefore holds for the space $\hat{W}_k$.

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

To estimate $\|u - u_k\|_k$, we start with the formula (cf. [8])

$$\|u - u_k\|_k \leq \inf_{v \in \hat{V}_k} \|u - v\|_k + \sup_{v \in \hat{V}_k} \frac{|a_k(u - u_k, v)|}{\|v\|_k}. \tag{6.6}$$

By arguments similar to those that led to (6.4), we see that the first term is bounded by $Ch_k\|u\|_{W_2^2}$. The second term is also bounded by $Ch_k\|u\|_{W_2^2}$; the analysis is similar to that in [4]. Therefore,

$$\|u - u_k\|_k \leq Ch_k\|u\|_{W_2^2}. \tag{6.7}$$

Again, a duality argument (cf. the proof of (4.15)) shows that

$$\|u - u_k\|_{L^2} \leq Ch_k^2\|u\|_{W_2^2}. \tag{6.8}$$

Thus the analog of (2.14) holds for $\hat{V}_k$.

The operator I_{k-1}^k defined in Section 3 must be modified so that it maps $\hat{V}_{k-1}$ into $\hat{V}_k$. We define $\hat{I}_{k-1}^k : \hat{V}_{k-1} \to \hat{V}_k$ by

$$\hat{I}_{k-1}^k v := I_{k-1}^k v - \frac{1}{|\Omega|} \int_\Omega I_{k-1}^k v. \tag{6.9}$$

The computation of the integral involves only $\mathcal{O}(n_k)$ operations (using the quadrature formula (2.6)). Note that $\hat{I}_{k-1}^k w = w$ for all $w \in \hat{W}_{k-1}$.

We have, by (3.2),

$$\|\hat{I}_{k-1}^k v\|_k = \|I_{k-1}^k v\|_k \leq C\|v\|_{k-1} \tag{6.10}$$

and, by the Cauchy-Schwarz inequality and (3.3),

$$\begin{aligned} \|\hat{I}_{k-1}^k v\|_{L^2} &= \left\| I_{k-1}^k v - \frac{1}{|\Omega|} \int_\Omega I_{k-1}^k v \right\|_{L^2} \\ &\leq \|I_{k-1}^k v\|_{L^2} + \left\| \frac{1}{|\Omega|} \int_\Omega I_{k-1}^k v \right\|_{L^2} \\ &\leq 2\|I_{k-1}^k v\|_{L^2} \leq C\|v\|_{L^2}. \end{aligned} \tag{6.11}$$

Because of (6.4), (6.5), (6.7), (6.8), (6.10) and (6.11), the theory developed in earlier sections carries over to this case if we replace V, V_k, W_k and I_{k-1}^k by $\hat{V}, \hat{V}_k, \hat{W}_k$ and $\hat{I}_{k-1}^k$, respectively.

In practice we can use the same scheme with I_{k-1}^k replaced by $\hat{I}_{k-1}^k$ and V_1 replaced by $\hat{V}_1$. The solution obtained is in $\hat{V}_k$ since the zero mean value is preserved by the intergrid transfer operator, the smoothing steps and the coarser-grid correction.

7. Extension to a Quadratic Nonconforming Finite Element. The principles which led to our optimal-order multigrid method for P1 nonconforming elements can also be applied to higher-order nonconforming finite elements. In this section we will indicate how this can be done for a quadratic nonconforming finite element (cf. [5]). For simplicity, we restrict our discussion to the Dirichlet problem for the Laplace equation

$$\begin{aligned} -\Delta u &= f \quad \text{in } \Omega, \\ u &= 0 \quad \text{on } \partial\Omega, \end{aligned} \tag{7.1}$$

where Ω is a convex polygon.

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

Let $V = \{v \in W_2^1(\Omega)\colon v|_{\partial\Omega} = 0\}$ and $a(u,v) = \int_\Omega \nabla u \cdot \nabla v$ for $u, v \in V$. Then the variational formulation for (7.1) is to find $u \in V$ such that

$$a(u,v) = \int fv \quad \forall v \in V. \tag{7.2}$$

We assume that $u \in W_2^3(\Omega) \cap V$ in order to fully exploit the properties of the quadratic element.

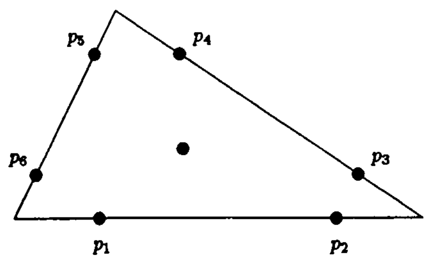

FIGURE 1

Let K be a triangle. The barycentric coordinates of the six Gauss-Legendre points $p_1, p_2, p_3, p_4, p_5, p_6$ along the sides of K (cf. Figure 1) are obtained by permuting

$$\left(\frac{1}{2}\left(1+\frac{\sqrt{3}}{3}\right), \frac{1}{2}\left(1-\frac{\sqrt{3}}{3}\right), 0\right).$$

If g is a quadratic polynomial on K, then

$$g(p_6) - g(p_5) + g(p_4) - g(p_3) + g(p_2) - g(p_1) = 0 \tag{7.3}$$

(cf. [5]).

Let $\beta, \alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5, \alpha_6 \in \mathbf{R}$ such that $\sum_{i=1}^6 (-1)^i \alpha_i = 0$. Then there is exactly one quadratic polynomial g such that $g(p_i) = \alpha_i$ and $g(\text{centroid}) = \beta$.

Let $\mathscr{T}^k$ be the family of triangulations described in Section 2. Define

$$\begin{aligned} V_k := \{v\colon\ & v|_T \text{ is a quadratic polynomial for each } T \in \mathscr{T}^k, v \text{ is continuous} \\ & \text{at the Gauss-Legendre points of interelement boundaries and } v \\ & \text{vanishes at the Gauss-Legendre points on } \partial\Omega\} \end{aligned} \tag{7.4}$$

and

$$\begin{aligned} W_k := \{w\colon\ & w|_T \text{ is a quadratic polynomial for each } T \in \mathscr{T}^k, w \text{ is} \\ & \text{continuous on } \Omega \text{ and } w \equiv 0 \text{ on } \partial\Omega\}. \end{aligned} \tag{7.5}$$

Again, note that $W_k = V_k \cap V = V_k \cap V_{k+1}$.

For each k, define

$$a_k(v_1, v_2) = \sum_{T \in \mathscr{T}^k} \int_T \nabla v_1 \cdot \nabla v_2 \quad \text{for } v_1, v_2 \in V + V_k \tag{7.6}$$

and

$$\|v\|_k = a_k(v,v)^{1/2}. \tag{7.7}$$

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

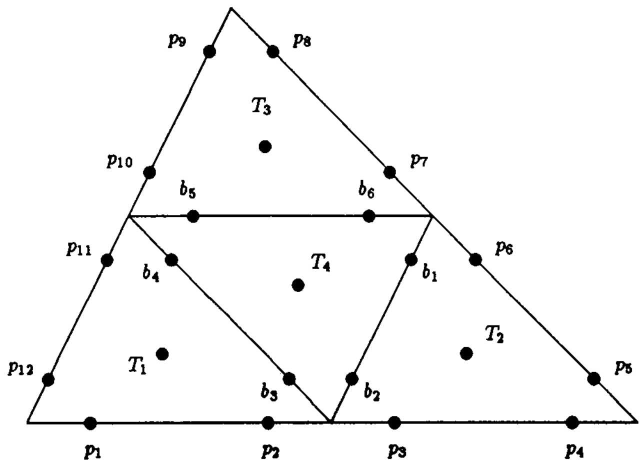

FIGURE 2

Let $u_k \in V_k$ satisfy

$$a_k(u_k, v) = \int fv \quad \forall v \in V_k,$$

and $\tilde{u}_k \in W_k$ satisfy

$$a_k(\tilde{u}_k, v) = \int fv \quad \forall v \in W_k.$$

Then we have the estimates

$$\|u - u_k\|_{L^2} + h_k\|u - u_k\|_k \leq Ch_k^3\|u\|_{W_2^3} \tag{7.8}$$

and

$$\|u - \tilde{u}_k\|_{L^2} + h_k\|u - \tilde{u}_k\|_k \leq Ch_k^3\|u\|_{W_2^3}. \tag{7.9}$$

Inequality (7.9) is well known. Inequality (7.8) holds because the nonconforming element satisfies the patch test (cf. [4], [5]). The theory for P1 nonconforming finite elements can be extended to this case in a straightforward manner if we can find an operator $I_{k-1}^k\colon V_{k-1} \to V_k$ such that (i) it is a bounded linear operator with respect to both the L^2 norm and the energy norm and (ii) it reduces to the natural injection on W_{k-1}.

For $v \in V_{k-1}$, the operator $I_{k-1}^k\colon V_{k-1} \to V_k$ is defined as follows.

Let p be a Gauss-Legendre point of a side of a triangle in $\mathscr{T}^k$. If p lies on the common edge of two adjacent triangles Δ_1 and Δ_2 in $\mathscr{T}^{k-1}$, then

$$(I_{k-1}^k v)(p) := \frac{1}{2}[v|_{\Delta_1}(p) + v|_{\Delta_2}(p)].$$

If p lies inside a triangle T in $\mathscr{T}^{k-1}$, then there are two cases to consider. $I_{k-1}^k v$ assumes the same values as v at the points b_2, b_4 and b_6; the value of $I_{k-1}^k v$ at the points b_1, b_3 and b_5 is determined by condition (7.3) applied to the three outer triangles T_1, T_2, T_3 (cf. Figure 2). To complete the definition of the intergrid transfer operator we must verify that the values of $I_{k-1}^k v$ at the six Gauss-Legendre points of T_4 satisfy (7.3). Then we let $I_{k-1}^k v$ take the same value as v at the centroids of T_1, T_2, T_3 and T_4.

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

LEMMA 7. *Let $p_1, \ldots, p_{12}$ be the Gauss-Legendre points along the boundary of $T \in \mathscr{T}^k$ (cf. Figure 2). Then*

$$\sum_{i=1}^{12} (-1)^i (I_{k-1}^k v)(p_i) = \sum_{i=1}^{12} (-1)^i v|_T(p_i). \tag{7.10}$$

Proof. Let e denote the edge that contains p_1, p_2, p_3 and p_4. We have

$$\sum_{i=1}^{4} (-1)^i (I_{k-1}^k v)(p_i) = \sum_{i=1}^{4} (-1)^i v|_T(p_i) + \sum_{i=1}^{4} \frac{(-1)^i}{2} [v|_{T'}(p_i) - v|_T(p_i)], \tag{7.11}$$

where T' is the neighboring triangle in $\mathscr{T}^{k-1}$ that also contains e. On e, $q := v|_{T'} - v|_T$ is a quadratic polynomial that vanishes at the two Gauss-Legendre points of the original triangle T, which are symmetric with respect to the midpoint. Therefore q is symmetric with respect to the midpoint of e. Thus $q(p_4) = q(p_1)$ and $q(p_3) = q(p_2)$. Hence the second sum on the right-hand side of (7.11) adds up to zero.

Since a similar equality holds for each edge, Lemma 7 is now proved. □

Because v satisfies (7.3) on T_1, T_2 and T_3,

$$\sum_{i=1}^{12} (-1)^i v(p_i) + \sum_{j=1}^{6} (-1)^j v(b_j) = 0. \tag{7.12}$$

But v also satisfies (7.3) on T_4, hence

$$\sum_{j=1}^{6} (-1)^j v(b_j) = 0. \tag{7.13}$$

Equalities (7.12), (7.13) and Lemma 7 then imply that

$$\sum_{i=1}^{12} (-1)^i (I_{k-1}^k v)(p_i) = 0.$$

But by construction,

$$\sum_{i=1}^{12} (-1)^i (I_{k-1}^k v)(p_i) + \sum_{j=1}^{6} (-1)^j (I_{k-1}^k v)(b_j) = 0.$$

Therefore,

$$\sum_{j=1}^{6} (-1)^j (I_{k-1}^k v)(b_j) = 0,$$

and the compatibility condition on T_4 is satisfied. Hence $I_{k-1}^k v$ is well defined.

By our definition of $I_{k-1}^k v$, it is obvious that $I_{k-1}^k w = w$ for $w \in W_{k-1}$.

The proof of Lemma 1 is still applicable for this intergrid transfer operator. Therefore, I_{k-1}^k is bounded with respect to both the L^2 and energy norms.

As indicated earlier, once I_{k-1}^k has these properties we can obtain a multigrid method for finding approximate solutions $\hat{u}_k$ of u_k such that

$$\|u - \hat{u}_k\|_{L^2} + h_k \|u - \hat{u}_k\|_k \le C h_k^3 \|u\|_{W_2^3}.$$

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

Acknowledgments. Special thanks go to the author's advisor, Professor Ridgway Scott, for his advice and encouragement. I would also like to thank the referee for helpful suggestions and for the simple proof of Lemma 1.

Department of Mathematics
University of Michigan
Ann Arbor, Michigan 48109

1. R. E. BANK & T. DUPONT "An optimal order process for solving finite element equations," *Math. Comp.*, v. 36, 1981, pp. 35–51.

2. D. BRAESS & W. HACKBUSCH, "A new convergence proof for the multigrid method including the V-cycle," *SIAM J. Numer. Anal.*, v. 20, 1983, pp. 967–975.

3. P. CIARLET, *The Finite Element Method for Elliptic Problems*, North-Holland, Amsterdam, 1978.

4. M. CROUZEIX & P.-A. RAVIART, "Conforming and nonconforming finite element methods for solving the stationary Stokes equations I," *RAIRO Anal. Numér. Sér. Rouge*, v. 7, no. R-3, 1973, pp. 33–75.

5. M. FORTIN & M. SOULIE, "A non-conforming piecewise quadratic element on triangles," *Internat. J. Numer. Methods Engrg.*, v. 19, 1983, pp. 505–520.

6. P. GRISVARD, "Behavior of solutions of an elliptic boundary value problem in polygonal or polyhedral domains," *Numerical Solution of Partial Differential Equations* - III (Synspade 1975) (B. Hubbard, ed.), Academic Press, New York, 1976, pp. 207–274.

7. W. HACKBUSCH, *Multi-Grid Methods and Applications*, Springer-Verlag, Berlin and New York, 1985.

8. G. STRANG & G. FIX, *An Analysis of the Finite Element Method*, Prentice-Hall, Englewood Cliffs, N.J., 1973.

9. K. STÜBEN & U. TROTTENBERG, "Multigrid methods: fundamental algorithms, model problem analysis and applications," *Multigrid Methods* (Cologne, 1981), Lecture Notes in Math., vol. 960 (W. Hackbusch & U. Trottenberg, eds.), Springer-Verlag, Berlin and New York, 1982, pp. 1–176.

10. F. THOMASSET, *Implementation of Finite Element Methods for Navier-Stokes Equations*, Springer-Verlag, New York, 1981.

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