MATHEMATICS OF COMPUTATION
VOLUME 57, NUMBER 195
JULY 1991, PAGES 23–45

CONVERGENCE ESTIMATES FOR MULTIGRID ALGORITHMS WITHOUT REGULARITY ASSUMPTIONS

JAMES H. BRAMBLE, JOSEPH E. PASCIAK, JUNPING WANG, AND JINCHAO XU

ABSTRACT. A new technique for proving rate of convergence estimates of multigrid algorithms for symmetric positive definite problems will be given in this paper. The standard multigrid theory requires a "regularity and approximation" assumption. In contrast, the new theory requires only an easily verified approximation assumption. This leads to convergence results for multigrid refinement applications, problems with irregular coefficients, and problems whose coefficients have large jumps. In addition, the new theory shows why it suffices to smooth only in the regions where new nodes are being added in multigrid refinement applications.

1. INTRODUCTION

In recent years, multigrid methods have been used extensively as tools for obtaining approximations to solutions of partial differential equations (see the references in [13], [14], [20]). In conjunction, there has been intensive research into the theoretical understanding of the convergence properties of the methods (cf. [3]–[10], [14], [17], [19]). This paper will present some new results on the convergence of multigrid algorithms. In particular, we shall prove multigrid convergence results under regularity-free assumptions.

The standard multigrid analysis uses a "regularity and approximation" assumption [6], [20]. This hypothesis is proved using elliptic regularity for the solution of the underlying partial differential equation as well as approximation and inverse properties of the discrete multilevel spaces. In contrast, two-level schemes have been proved to converge under apparently weaker assumptions [12], [14], [16], [18], [22], [23]. These arguments often only require approximation properties of the subspace. In this paper, we shall prove convergence

Received April 17, 1990; revised July 24, 1990.
1980 *Mathematics Subject Classification* (1985 *Revision*). Primary 65N30; Secondary 65F10.
This manuscript has been authored under contract number DE-AC02-76CH00016 with the U.S. Department of Energy. Accordingly, the U.S. Government retains a nonexclusive, royalty-free license to publish or reproduce the published form of this contribution, or allow others to do so, for U.S. Government purposes. This work was also supported in part under the National Science Foundation Grant No. DMS84-05352, DMS88-05311-04 and by the U.S. Army Research Office through the Mathematical Sciences Institute, Cornell University. The third author's research was supported by Office of Naval Research Contract No. 0014-88-K-0370 and by the Institute for Scientific Computation of the University of Wyoming through National Science Foundation Grant No. RII-8610680.

0025-5718/91 $1.00 + $.25 per page

estimates for general multilevel procedures under these weaker approximation assumptions. Somewhat weaker results of this form are proved in [7].

In the case of mesh refinement, the regularity and approximation constants may grow with some power of the ratio of the finest to coarsest mesh size. Our theory will show that the rate of convergence for these examples is no worse than $1 - c/j$, where j is the number of levels in the multigrid scheme. In addition, our analysis explains why it is only necessary to smooth in the region where refined triangles (or nodes) are being added.

Other applications of the theory to be provided include general H_0^1 equivalent forms and operators whose coefficients have large jumps. Both of the above-mentioned multigrid applications satisfy the approximation assumption, and hence the theory of this paper provides a fundamental analysis. In contrast, an analysis proving regularity and approximation estimates for these cases is either unavailable or leads to inequalities with large constants.

We should note that there has been some attempt to generalize the two-level analysis to the multilevel case [1], [21]. For example, [21] provides an analysis for an algorithm which requires applying a Chebyshev iterative scheme for the coarse grid correction, the number of terms depending on the convergence rate of the two-level scheme. The work estimates for this scheme are the same as those for the general multigrid algorithm with the number of correction steps (p in Algorithm S or N) equal to the number of terms in the Chebyshev series and hence is unacceptable for large p. The analysis presented here provides an estimate for the multilevel algorithm (including the V-cycle) without placing any restrictions on the convergence rate of the two-level algorithm. This is important since, as far as we know, there is no computationally effective way of improving the theoretical convergence rate of the two-level algorithm without assuming the regularity and approximation property.

The outline of the remainder of the paper is as follows. In §2, following [6], we define the multigrid algorithms and derive some fundamental recurrence relations. In §3, we give the regularity-free multigrid convergence analysis. In §4, we consider the application of the earlier theory to the case of general H_0^1 equivalent forms. Finite element and finite difference applications are considered there. An example in which the grid is locally refined and smoothing is only on the refined subregion is given in §5. In §6, we consider the applications to problems which have large jumps in coefficients. In §7, we give the results of numerically computed convergence factors which illustrate the theory developed in this paper. Finally, we include an appendix in which some issues concerning the implementation of the schemes are considered.

2. The multigrid algorithms

In this section, following [6], we describe the multigrid algorithms in both the symmetric and nonsymmetric cases. Along the way, we derive basic recurrence relations which play major roles in the analysis of the methods. For convenience, the algorithms are developed in an abstract Hilbert space setting. The

results most naturally apply to finite element multigrid algorithms but can also be applied to certain formulations of finite difference multigrid algorithms.

Let us assume that we are given a nested sequence of finite-dimensional vector spaces

$$M_0 \subset M_1 \subset \cdots \subset M_j.$$

In addition, let $A(\cdot, \cdot)$ and $(\cdot, \cdot)_k$ be symmetric positive definite bilinear forms on M_k for $k = 0, \ldots, j$. We shall develop multigrid algorithms for the solution of the problem: Given $f \in M_j$, find $v \in M_j$ satisfying

$$A(v, \phi) = (f, \phi)_j \quad \text{for all } \phi \in M_j. \tag{2.1}$$

To define the multigrid algorithms, we shall define auxiliary operators. For $k = 0, \ldots, j$, define the operator $A_k : M_k \mapsto M_k$ by

$$(A_k w, \phi)_k = A(w, \phi) \quad \text{for all } \phi \in M_k. \tag{2.2}$$

The operator A_k is clearly symmetric (in both the $A(\cdot, \cdot)$- and $(\cdot, \cdot)_k$-inner products) and positive definite. Also define the projectors $P_k : M_j \mapsto M_k$ and operators $P_k^0 : M_{k+1} \mapsto M_k$ by

$$A(P_k w, \phi) = A(w, \phi) \quad \text{for all } \phi \in M_k,$$

and

$$(P_k^0 w, \phi)_k = (w, \phi)_{k+1} \quad \text{for all } \phi \in M_k.$$

Note that P_k is symmetric in the A-inner product.

To define the smoothing process, we require a linear operator $R_k : M_k \mapsto M_k$ for $k = 1, \ldots, j$. We assume that R_k is a symmetric and positive semi-definite operator in the $(\cdot, \cdot)_k$-inner product and set $K_k = I - R_k A_k$ on M_k. Clearly, K_k is symmetric in the inner product $A(\cdot, \cdot)$, and we further assume that K_k is nonnegative in the sense that $A(K_k u, u) \geq 0$ for all $u \in M_k$. These assumptions imply that the spectrum of K_k is in $[0, 1]$.

We first define the symmetric multigrid operator $B_k^s : M_k \mapsto M_k$ by induction.

Algorithm S. Set $B_0^s = A_0^{-1}$. Assume that B_{k-1}^s has been defined and define $B_k^s g$ for $g \in M_k$ as follows:

(1) Set $x^0 = 0$ and $q^0 = 0$.

(2) Define x^l for $l = 1, \ldots, m(k)$ by

$$x^l = x^{l-1} + R_k(g - A_k x^{l-1}). \tag{2.3}$$

(3) Define $x^{m(k)+1} = x^{m(k)} + q^p$ where q^i for $i = 1, \ldots, p$ is defined by

$$q^i = q^{i-1} + B_{k-1}^s [P_{k-1}^0 (g - A_k x^{m(k)}) - A_{k-1} q^{i-1}].$$

(4) Set $B_k^s g = x^{2m(k)+1}$ where x^l is defined for $l = m(k)+2, \ldots, 2m(k)+1$ by (2.3).

In this algorithm, $m(k)$ is a positive integer which may vary from level to level and determines the number of smoothing iterations on that level. Because of this variable smoothing, the above algorithm is more general than that usually described [3], [4], [13], [14]. If all of the $m(k)$ are the same, then this algorithm is the usual symmetric multigrid algorithm described in a notation which is convenient for our analysis. Note that B_k^s is clearly a linear operator for each k. In this algorithm, p is a positive integer. The cases $p = 1$ and $p = 2$ correspond respectively to the symmetric $\mathscr{V}$ and $\mathscr{W}$ cycles of multigrid.

The definition of the nonsymmetric multigrid operator B_k^n is similar except that the smoothings of Step 4 are excluded. More precisely, we define $B_k^n : M_k \mapsto M_k$ by induction.

Algorithm N. Set $B_0^n = A_0^{-1}$. Assume that B_{k-1}^n has been defined and define $B_k^n g$ for $g \in M_k$ as follows:

(1) Set $x^0 = 0$ and $q^0 = 0$.

(2) Define x^l for $l = 1, \ldots, m(k)$ by (2.3).

(3) Define $B_k^n g = x^{m(k)} + q^p$ where q^i for $i = 1, \ldots, p$ is defined by

$$q^i = q^{i-1} + B_{k-1}^n \left[P_{k-1}^0 (g - A_k x^{m(k)}) - A_{k-1} q^{i-1}\right]. \tag{2.4}$$

The above algorithm defines a linear operator B_k^n which is equivalent to the standard nonsymmetric multigrid algorithms described in [4], [14] when $m(k)$ is constant.

Let $g = A_k x$. It is straightforward to check that q^p defined by (2.4) satisfies

$$q^p = (I - (I - B_{k-1}^n A_{k-1})^p) A_{k-1}^{-1} P_{k-1}^0 A_k (x - x^{m(k)}). \tag{2.5}$$

A trivial computation gives that

$$x - x^{m(k)} = K_k^{m(k)} x. \tag{2.6}$$

Noting that on M_k, $P_{k-1}^0 A_k = A_{k-1} P_{k-1}$, and combining (2.5) and (2.6) gives

$$I - B_k^n A_k = \left[(I - P_{k-1}) + (I - B_{k-1}^n A_{k-1})^p P_{k-1}\right] K_k^{m(k)} \quad \text{on } M_k. \tag{2.7}$$

Equation (2.7) gives a fundamental recurrence relation for the nonsymmetric multigrid algorithms. The analogous recurrence in the symmetric case is

$$I - B_k^s A_k = K_k^{m(k)} \left[(I - P_{k-1}) + (I - B_{k-1}^s A_{k-1})^p P_{k-1}\right] K_k^{m(k)} \quad \text{on } M_k, \tag{2.8}$$

which follows from similar reasoning.

We now restrict our attention to the case of $p = 1$. Note that $K_k = K_k^*$, where K_k^* is the adjoint with respect to the inner product $A(\cdot, \cdot)$. Consequently,

$$(I - B_k^n A_k)^* = K_k^{m(k)} \left[(I - P_{k-1}) + (I - B_{k-1}^n A_{k-1})^* P_{k-1}\right].$$

Multiplying this with (2.7), we find by a straightforward induction argument

$$I - B_j^s A_j = (I - B_j^n A_j)^* (I - B_j^n A_j). \tag{2.9}$$

Let $|||\cdot|||$ denote the norm corresponding to the inner product $A(\cdot,\cdot)$ on M_j. Clearly, the convergence rate of Algorithm S is bounded by the operator norm of $I - B_j^s A_j$ with respect to $|||\cdot|||$, which is the square of the operator norm of $I - B_j^n A_j$.

For our analysis, we will write $I - B_j^n A_j$ as a product of operators. By (2.7), for $k = 1, \ldots, j$, we clearly have that, on M_j,

$$\begin{aligned} I - B_k^n A_k P_k &= (I - P_k) + (I - B_{k-1}^n A_{k-1} P_{k-1}) K_k^{m(k)} P_k \\ &= (I - B_{k-1}^n A_{k-1} P_{k-1})(I - P_k + K_k^{m(k)} P_k) \\ &= (I - B_{k-1}^n A_{k-1} P_{k-1})(I - T_k), \end{aligned} \tag{2.10}$$

where $T_k = (I - K_k^{m(k)}) P_k$. Note that we have used the fact that $P_{k-1}(I - P_k) = 0$ in the derivation of the above equality. For convenience, we define $T_0 = P_0$ and hence

$$I - B_j^n A_j = (I - T_0)(I - T_1) \cdots (I - T_j),$$

from which it follows that

$$(I - B_j^n A_j)^* = (I - T_j) \cdots (I - T_1)(I - T_0), \tag{2.11}$$

since $T_k^* = T_k$. We note that each factor of (2.11) is defined on all of M_j. To develop bounds for the convergence of either Algorithm S or N, it suffices to estimate the operator norm of $(I - B_j^n A_j)^*$.

The more general case $p \geq 1$ is quite similar. We first consider the nonsymmetric algorithm. Set

$$\varepsilon_k = (I - B_k^n A_k P_k)^* \quad \text{for } k = 0, \ldots, j.$$

A manipulation similar to that used above yields $\varepsilon_k = (I - T_k)\varepsilon_{k-1}^p$ for $k = 1, \ldots, j$. Thus,

$$\begin{aligned} \varepsilon_j &= (I - T_j)(I - T_{j-1})\varepsilon_{j-2}^p \varepsilon_{j-1}^{p-1} \\ &= (I - T_j) \cdots (I - T_1)(I - T_0)\varepsilon_1^{p-1} \varepsilon_2^{p-1} \cdots \varepsilon_{j-1}^{p-1}. \end{aligned} \tag{2.12}$$

Clearly, $|||\varepsilon_1^{p-1} \varepsilon_2^{p-1} \cdots \varepsilon_{j-1}^{p-1}||| \leq 1$. Consequently, estimates for Algorithm N with $p > 1$ follow from the estimates for Algorithm N with $p = 1$. Similar arguments applied to the symmetric algorithm imply that

$$\begin{aligned} I - B_j^s A_j = (I - T_j) \cdots (I - T_1)(I - T_0) D_j \\ \cdot (I - T_0)(I - T_1) \cdots (I - T_j), \end{aligned} \tag{2.13}$$

where the $|||\cdot|||$ norm of the operator D_j is also bounded by one. Consequently, the convergence rate for $p > 1$ is always bounded by the convergence rate for the case of $p = 1$. Our analysis will not guarantee an improved rate of convergence for $p > 1$ under the regularity-free assumption (3.3). In contrast, the proofs

assuming the regularity and approximation assumption [6], [14], [20] do suggest such improvements.

3. Multigrid analysis

In this section, we give an analysis of the multigrid algorithms described in the previous section. Our goal is to prove inequalities of the form:

$$A((I - B_j^s A_j)u, u) \leq \delta_j A(u, u) \quad \text{for all } u \in M_j, \tag{3.1}$$

or

$$A((I - B_j^n A_j)u, (I - B_j^n A_j)u) \leq \delta_j A(u, u) \quad \text{for all } u \in M_j. \tag{3.2}$$

From the previous discussion, it suffices to prove (3.2) when $p = 1$.

In contrast to the usual multigrid analysis [6], [20], we will relate δ_j in (3.2) to somewhat different a priori assumptions. Let λ_k denote the largest eigenvalue of A_k. Our main assumption is the existence of linear operators $Q_k: M_j \mapsto M_k$ for $k = 0, 1, \ldots, j$, with $Q_j = I$, satisfying the following properties. We assume that there are constants C_1 and C_2 not depending on k for which

$$\begin{aligned} \|(Q_k - Q_{k-1})u\|_k^2 &\leq C_1 \lambda_k^{-1} A(u, u) \quad \text{for } k = 1, \ldots, j, \\ A(Q_k u, Q_k u) &\leq C_2 A(u, u) \quad \text{for } k = 0, \ldots, j-1. \end{aligned} \tag{3.3}$$

The inequalities in (3.3) hold for all $u \in M_j$. As will be demonstrated in §§4–6, in various applications the derivation of inequalities of the form (3.3) only uses the approximation properties of the subspaces.

We will also make a second assumption concerning the symmetric smoothing operator R_k. Assume that $\widetilde{M}_k \equiv$ "the range of R_k" contains the range of $Q_k - Q_{k-1}$. We further assume that there exists a constant C_R independent of k which satisfies

$$\frac{\|u\|_k^2}{\lambda_k} \leq C_R (R_k u, u)_k \quad \text{for all } u \in \widetilde{M}_k. \tag{3.4}$$

Let $\widetilde{P}_k^0$ denote the orthogonal projection of M_k onto $\widetilde{M}_k$ with respect to $(\cdot, \cdot)_k$. We note that for $v \in M_k$ and $\phi \in \widetilde{M}_k$,

$$(R_k \widetilde{P}_k^0 v, \phi)_k = (v, R_k \phi)_k = (R_k v, \phi)_k,$$

i.e., $R_k \widetilde{P}_k^0 = R_k$.

Remark 3.1. The above assumption on R_k represents a minimal requirement for a symmetric operator to be considered a smoother and is a natural condition to be satisfied. The fact that R_k need not be definite on M_k plays an important role in certain applications. In effect, the above assumption only requires that R_k smooth on the range of $Q_k - Q_{k-1}$. In a refinement application described in §5, $\widetilde{M}_k$ will only contain degrees of freedom in a refinement subdomain and hence smoothing need only be done in that subdomain. Algorithms which only

smooth in the refinement regions have been suggested in [2]. An example of an R_k satisfying (3.4) is $R_k = \bar{\lambda}_k^{-1}\tilde{P}_k^0$, where $\bar{\lambda}_k$ is any upper bound for λ_k (we need K_k nonnegative) and $\tilde{P}_k^0$ is the orthogonal projection of M_k onto $\widetilde{M}_k$ with respect to $(\cdot, \cdot)_k$.

We can now state and prove the theorem for estimating δ_j in (3.1) and (3.2).

Theorem 1. *Assume that* (3.3) *and* (3.4) *hold, and define* B_j^s *and* B_j^n *by Algorithm* S *and* N, *respectively. Then* (3.1) *and* (3.2) *hold with*

$$\delta_j = 1 - \frac{1}{Cj}, \tag{3.5}$$

where $C = [1 + C_2^{1/2} + (C_R C_1)^{1/2}]^2$.

Proof. From the earlier discussion, we may assume, without loss of generality, that $p = 1$. In this case, (3.1) and (3.2) are equivalent. Let $E_k = (I - B_k^n A_k P_k)^*$. By (2.10), we have for $k = 1, \ldots, j$

$$E_k = (I - T_k)E_{k-1}. \tag{3.6}$$

Equality (3.6) also holds for $k = 0$ provided that we define $E_{-1} = I$.

We note that to prove (3.2) for δ_j given by (3.5), it suffices to show that

$$A(u, u) \le Cj[A(u, u) - A(E_j u, E_j u)]. \tag{3.7}$$

We start from (3.6) and compute

$$\begin{aligned} A(E_{k-1}u, E_{k-1}u) &= A(E_k u, E_k u) + 2A(T_k(I - T_k)E_{k-1}u, E_{k-1}u) \\ &\quad + A(T_k E_{k-1}u, T_k E_{k-1}u) \\ &= A(E_k u, E_k u) + A(T_k(I - T_k)E_{k-1}u, E_{k-1}u) \\ &\quad + A(T_k E_{k-1}u, E_{k-1}u). \end{aligned}$$

From the definition of T_k, the operator $T_k(I - T_k)$ is symmetric with respect to $A(\cdot, \cdot)$ and positive semidefinite, and hence

$$A(T_k E_{k-1}u, E_{k-1}u) \le A(E_{k-1}u, E_{k-1}u) - A(E_k u, E_k u). \tag{3.8}$$

Summing from 0 to j gives

$$\sum_{k=0}^{j} A(T_k E_{k-1}u, E_{k-1}u) \le A(u, u) - A(E_j u, E_j u). \tag{3.9}$$

By (3.7) and (3.9), (3.5) is reduced to proving that

$$A(u, u) \le Cj \sum_{k=0}^{j} A(T_k E_{k-1}u, E_{k-1}u). \tag{3.10}$$

To this end, we use the fact that $Q_j = I$ to write

$$u = \sum_{k=1}^{j} (Q_k - Q_{k-1})u + Q_0 u.$$

Thus,

$$
\begin{aligned}
A(u, u) &= \sum_{k=1}^{j} A(u, (Q_k - Q_{k-1})u) + A(u, Q_0 u) \\
&= \sum_{k=1}^{j} A(E_{k-1}u, (Q_k - Q_{k-1})u) + A(u, Q_0 u) \\
&\quad + \sum_{k=1}^{j} A((I - E_{k-1})u, (Q_k - Q_{k-1})u).
\end{aligned} \tag{3.11}
$$

We now estimate the first sum on the right of (3.11). Using Schwarz's inequality, we see that

$$
\begin{aligned}
\sum_{k=1}^{j} A(E_{k-1}u, (Q_k - Q_{k-1})u) &= \sum_{k=1}^{j} A(P_k E_{k-1}u, (Q_k - Q_{k-1})u) \\
&= \sum_{k=1}^{j} (\tilde{P}_k^0 A_k P_k E_{k-1}u, (Q_k - Q_{k-1})u)_k \\
&\le \sum_{k=1}^{j} \|\tilde{P}_k^0 A_k P_k E_{k-1}u\|_k \|(Q_k - Q_{k-1})u\|_k.
\end{aligned}
$$

Applying (3.3) and (3.4) gives

$$
\begin{aligned}
&\sum_{k=1}^{j} A(E_{k-1}u, (Q_k - Q_{k-1})u) \\
&\quad \le (C_1)^{1/2} A^{1/2}(u, u) \sum_{k=1}^{j} \lambda_k^{-1/2} \|\tilde{P}_k^0 A_k P_k E_{k-1}u\|_k \\
&\quad \le (C_R C_1 j)^{1/2} A^{1/2}(u, u) \left(\sum_{k=1}^{j} A(R_k A_k P_k E_{k-1}u, P_k E_{k-1}u) \right)^{1/2}.
\end{aligned}
$$

We used the fact that $R_k \tilde{P}_k^0 = R_k$ in the last inequality above. The spectrum of K_k is in $[0, 1]$, and hence

$$
\begin{aligned}
A(R_k A_k P_k E_{k-1}u, P_k E_{k-1}u) &= A((I - K_k)P_k E_{k-1}u, P_k E_{k-1}u) \\
&\le A(T_k E_{k-1}u, E_{k-1}u).
\end{aligned} \tag{3.12}
$$

Thus, we obtain

$$
\begin{aligned}
&\sum_{k=1}^{j} A(E_{k-1}u, (Q_k - Q_{k-1})u) \\
&\quad \le (C_1 C_R j)^{1/2} A^{1/2}(u, u) \left(\sum_{k=1}^{j} A(T_k E_{k-1}u, E_{k-1}u) \right)^{1/2}.
\end{aligned} \tag{3.13}
$$

For the remaining terms in (3.11), we have

$$\begin{aligned}(3.14)\qquad &\sum_{k=1}^{j} A((I-E_{k-1})u,\,(Q_k-Q_{k-1})u)+A(u,\,Q_0u)\\ &\quad=\sum_{k=1}^{j-1} A((E_k-E_{k-1})u,\,Q_ku)+A((I-E_{j-1})u,\,u).\end{aligned}$$

From (3.6), $E_{k-1}-E_k=T_kE_{k-1}$, and it follows that

$$I-E_{j-1}=\sum_{k=0}^{j-1} T_kE_{k-1}.$$

From this and (3.14),

$$\begin{aligned}(3.15)\qquad &\sum_{k=1}^{j} A((I-E_{k-1})u,\,(Q_k-Q_{k-1})u)+A(u,\,Q_0u)\\ &\quad=\sum_{k=1}^{j-1} A(T_kE_{k-1}u,\,(I-Q_k)u)+A(T_0u,\,u).\end{aligned}$$

Note that the spectrum of T_k is contained in $[0,\,1]$, and hence the right-hand side of (3.15) is bounded by

$$(3.16)\qquad j^{1/2}(1+C_2^{1/2})\left(\sum_{k=0}^{j-1} A(T_kE_{k-1}u,\,E_{k-1}u)\right)^{1/2} A^{1/2}(u,\,u).$$

Combining (3.11), (3.13), (3.15), and (3.16) gives

$$A(u,\,u)\le j[1+C_2^{1/2}+(C_RC_1)^{1/2}]^2\left(\sum_{k=0}^{j} A(T_kE_{k-1}u,\,E_{k-1}u)\right),$$

i.e., (3.10) holds with $C=[1+C_2^{1/2}+(C_RC_1)^{1/2}]^2$. This completes the proof of the theorem. □

Remark 3.2. Theorem 1 gives bounds for the convergence factor δ_k in (3.1) and (3.2) in terms of j. In finite element applications, j (the number of levels) is proportional to a logarithm of the mesh size. This theorem shows that the multigrid algorithms converge rapidly under the regularity-free assumption (3.3). We note, however, that the result does not show any improvement in the rate of convergence when the number of smoothings $m(k)$ is greater than one. In contrast, the multigrid results of [5], [6] imply improved convergence with larger $m(k)$ provided that the regularity and approximation assumption holds.

Remark 3.3. The nonnegativity assumption on K_k can be relaxed to the assumption that $\sigma(K_k)\subseteq[1-\omega,\,1]$ for some ω with $1\le\omega<2$. Here, $\sigma(K_k)$ denotes the spectrum of K_k. The value of C in Theorem 1 can be taken to

be $C = [1 + C_2^{1/2} + (C_R C_1 \omega/(2-\omega))^{1/2}]^2/(2-\omega)$. The assumption on $\sigma(K_k)$ implies that $\sigma(T_k) \subseteq [0, \omega]$. In this case, (3.8) is replaced by

$$(2-\omega)A(T_k E_{k-1}u, E_{k-1}u) \leq A(E_{k-1}u, E_{k-1}u) - A(E_k u, E_k u)$$

and (3.12) is replaced by

$$A(R_k A_k P_k E_{k-1}u, P_k E_{k-1}u) \leq \frac{\omega}{2-\omega} A(T_k E_{k-1}u, E_{k-1}u).$$

The remainder of the proof is similar.

4. Application to an H_0^1 equivalent form

In this section, we consider applying the multilevel results of the previous section to forms $A(\cdot, \cdot)$ defined on a subspace $M_j \subseteq H_0^1(\Omega)$. We will only assume that the quadratic form is equivalent to the square of the $H^1(\Omega)$ norm on M_j, i.e.,

$$c_0 \|v\|_1^2 \leq A(v, v) \leq c_1 \|v\|_1^2 \quad \text{for all } v \in M_j \tag{4.1}$$

holds for positive constants c_0, c_1. Here, $\|\cdot\|_1^2$ denotes the norm in $H^1(\Omega)$. As far as we know, it is not possible to prove the regularity and approximation assumption ((3.3) of [6]) under such a weak assumption on the form. In contrast, we shall see that it is possible to apply the theory of the preceding section to natural multigrid algorithms and guarantee rapid convergence. Examples for which (4.1) is known, yet regularity and approximation results are not available, include elliptic problems with bounded measurable coefficients and various finite difference applications.

We start with a nested sequence of multigrid spaces

$$M_0 \subset M_1 \subset \cdots \subset M_j.$$

The most natural examples result from finite element subspaces defined on a sequence of nested triangulations. For example, in the two-dimensional case, when Ω is a polygonal domain, we can define spaces of piecewise linear functions as follows. Since Ω is polygonal, start with a coarse triangulation $\tau_0 = \bigcup_l \tau_0^l$ of quasi-uniform size h_0, where τ_0^l represents an individual triangle and τ_0 denotes the triangulation. Successively finer triangulations $\{\tau_k, k = 1, \ldots, j\}$ of size $h_k = 2^{-k} h_0$ are defined by breaking each triangle of a coarser triangulation into four triangles by connecting the midpoints of the edges. The subspace M_k is defined to be the continuous functions defined on Ω which are piecewise linear with respect to τ_k and vanish on $\partial\Omega$.

Let Q_k denote the $L^2(\Omega)$ projection onto M_k. It is known that, since the triangulations are quasi-uniform,

$$\|(I - Q_k)v\| \leq c h_k \|v\|_1 \tag{4.2}$$

and

$$\|Q_k v\|_1 \leq C \|v\|_1, \tag{4.3}$$

for all $v \in H_0^1(\Omega)$ (see, e.g., [11], [23]). Here, $\|\cdot\|$ denotes the norm in $L^2(\Omega)$. Let $(\cdot, \cdot)_k$ be a discrete inner product on $M_k \times M_k$ which satisfies

$$c\|v\|^2 \leq (v, v)_k \leq C\|v\|^2 \quad \text{for all } v \in M_k. \tag{4.4}$$

The constants c and C in (4.4) are assumed to be independent of $k = 0, \ldots, j$. For example, we can take

$$(u, v)_k = h_k^2 \sum u(x_i)v(x_i), \tag{4.5}$$

where $\{x_i\}$ is the collection of nodes for M_k and the sum in (4.5) is taken over this set. Combining (4.1)–(4.4) shows that (3.3) holds. Since, in this case, we use discrete inner products, we can take R_k to be $\bar{\lambda}_k^{-1} I$, where $\bar{\lambda}_k = O(h_k^{-2})$ is a bound for the largest eigenvalue of A_k. Consequently, we can apply Theorem 1 to provide the following proposition.

Proposition 4.1. *Assume that* (4.1)–(4.4) *hold. Then the convergence rates of Algorithms* S *and* N *are bounded by*

$$\delta_j = 1 - \frac{1}{Cj}.$$

As the first application of this result, we will consider the following elliptic boundary value problem. Let Ω be a polygonal domain in R^2 and consider the problem

$$\begin{aligned} -\sum_{i,j=1}^{2} \frac{\partial}{\partial x_i} a_{ij} \frac{\partial u}{\partial x_j} + au &= f \quad \text{in } \Omega, \\ u &= 0 \quad \text{on } \partial\Omega. \end{aligned} \tag{4.6}$$

We assume that the matrix $\{a_{ij}\}$ is symmetric for each $x \in \Omega$. The corresponding form $A(\cdot, \cdot)$ is defined by

$$A(u, v) = \sum_{i,j=1}^{2} \int_\Omega a_{ij} \frac{\partial u}{\partial x_j} \frac{\partial v}{\partial x_i} \, dx + \int_\Omega auv \, dx. \tag{4.7}$$

For this application, the only additional requirement is that the coefficients are such that (4.1) holds. This means that the coefficients can have rather complex jumps provided that they remain between fixed upper and lower bounds. Proposition 4.1 applies and gives a convergence bound for the corresponding multigrid algorithms. As far as we know, appropriate elliptic regularity estimates (cf. [4], [6]) for the weak solutions of problem (4.6) are not known. Without such estimates, neither the theory of [4] nor [6] would yield useful convergence bounds.

For our second application of Proposition 4.1, we consider a finite difference operator applied to the problem

$$\begin{aligned} -\nabla \cdot a \nabla u &= f \quad \text{in } \Omega, \\ u &= 0 \quad \text{on } \partial\Omega. \end{aligned} \tag{4.8}$$

For this application, we assume that the domain is the union of squares in a coarse square mesh of size h_0 and that the coefficient a in (4.8) is continuous on $\overline{\Omega}$ and bounded from below by a positive constant c. The coarse finite difference mesh is formed by the above-mentioned squares. Successively finer grids are defined by breaking each square into four subsquares in the obvious way. We seek the solution of the finite difference equations on the jth grid.

The finite difference stencil at the grid point i, k of the jth grid is given by

$$(L_h\bar{x})_{i,k} = a_{i,k+1/2}(\bar{x}_{i,k} - \bar{x}_{i,k+1}) + a_{i,k-1/2}(\bar{x}_{i,k} - \bar{x}_{i,k-1}) + a_{i+1/2,k}(\bar{x}_{i,k} - \bar{x}_{i+1,k}) + a_{i-1/2,k}(\bar{x}_{i,k} - \bar{x}_{i-1,k}).$$

We also consider finite element spaces of piecewise linear functions as defined above. The coarse triangulation of Ω is formed by dividing each of the coarse squares mentioned above into two triangles by one of the diagonals. Nodal interpolation $\bar{x} \to x$ provides a natural correspondence between the nodal vectors in the finite difference approximation and functions in M_j, i.e., x is the function in M_j which takes on the values of $\bar{x}$ on the nodes of the grid. It is easy to show that the norm induced by the form

$$A(x, y) = \sum (L_h\bar{x})_{i,k}\bar{y}_{i,k} \tag{4.9}$$

is equivalent to the H^1 norm on M_j. The sum in (4.9) is over all nodes of the jth mesh. Consequently, Proposition 4.1 guarantees rates of convergence for the corresponding multigrid algorithms (with discrete inner products given by, for example, (4.5)).

This algorithm corresponds to a natural finite difference multigrid algorithm. In fact, the finite element spaces can be used to provide interpolation and restriction operators connecting the original sequence of finite difference spaces $\overline{M}_0, \ldots, \overline{M}_j$. Define the kth grid operator L_h^k by ([20])

$$L_h^k = I_k^t L_h I_k, \tag{4.10}$$

where I_k is the interpolation matrix (corresponding to the imbedding of M_k in M_j) and I_k^t is its transpose. Then the multigrid algorithm using the form of (4.9) is exactly the same as the finite difference multigrid algorithms using (4.10). Proposition 4.1 provides an estimate for its rate of convergence.

The discussion given in this section extends without difficulty to problems in more than two dimensions and other boundary conditions as well as more general finite element spaces.

5. A REFINEMENT APPLICATION

In this section, we shall apply the results of §3 to the case of finite element approximation with locally refined grids. Such mesh refinements are convenient for accurate modeling of problems with various types of singular behavior. Here, we shall make use of the fact that the operator R_k need not satisfy

smoothing conditions on the whole space but rather only on a subspace containing the range of $Q_k - Q_{k-1}$. This will allow an analysis of algorithms which only smooth on the nodes which are interior to the subdomain of refinement.

We consider the boundary value problem (4.6), but for convenience, we take $a = 0$. As in §4, we assume that the coefficient matrix is symmetric and that the resulting form (4.7) satisfies (4.1). It is well known that jumps in coefficients, nonconvex domains, or singular data results in solutions with various types of singular behavior [15]. To accurately solve such problems, it is convenient to use an appropriate mesh refinement. We shall present a flexible mesh refinement scheme in this section which leads to multigrid algorithms whose convergence rate can be bounded by applying Theorem 1.

For convenience, we will only consider the case of piecewise linear finite elements. Applications with other finite element spaces are possible. Also the case of more than two dimensions presents no inherent difficulty.

The first step in the definition of the algorithms involves the development of sequences of nested refined grids. These grids are defined in terms of a given sequence of nested subdomains

$$\Omega_j \subseteq \Omega_{j-1} \subseteq \cdots \subseteq \Omega_0 = \Omega.$$

We assume that we are given a coarse triangulation of $\Omega = \bigcup_l \tau_0^l$. This coarse triangulation provides the first grid $\{\tau_0^l\}$. Given that a grid $\{\tau_{k-1}^l\}$ has been defined, the grid $\{\tau_k^l\}$ is defined by refining those triangles of $\{\tau_{k-1}^l\}$ which are in Ω_k. This refinement is done by breaking each triangle of the mesh $\{\tau_{k-1}^l\}$ in Ω_k into four triangles by connecting the midpoints of the edges. We assume $\partial\Omega_k$ aligns with the mesh $\{\tau_{k-1}^l\}$.

The space M_k is defined to be the set of continuous functions on Ω which are piecewise linear with respect to the grid $\{\tau_k^l\}$ and vanish on $\partial\Omega$. We note that the continuity constraint implies that there are no new degrees of freedom corresponding to nodes on $\partial\Omega_k$ (see Figure 5.1). These new nodes on $\partial\Omega_k$ will be called slave nodes since, by continuity, their values are determined by the values of their neighboring nodes (which were already in the previous grids). It is easy to see that the space M_k has a nodal basis consisting of the vertices of $\{\tau_k\}$ excluding the slave nodes.

To fit this application into the abstract framework of §2 and §3, we need to define discrete inner products and smoothing operators R_k. We shall use the $L^2(\Omega)$-inner product for the discrete inner product on each level. The use of $L^2(\Omega)$-inner products in the multigrid algorithms often leads to algorithms which require the solution of Gram matrix problems in the implementation. We shall avoid this (see the Appendix) by a judicious choice of the smoothing operator R_k. Let

$$\widetilde{M}_k = \{\phi \in M_k \mid \operatorname{supp}\phi \subseteq \Omega_k\}. \tag{5.1}$$

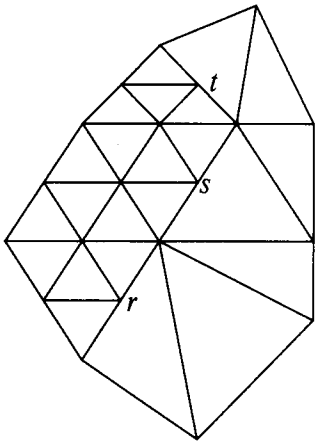

FIGURE 5.1
A mesh transition region

In addition, let $(\cdot, \cdot)$ denote the $L^2(\Omega)$-inner product and $\{\theta_k^l\}$ denote the usual nodal basis for $\widetilde{M}_k$. Define

$$R_k u = \alpha \sum_l (d_k^l)^{-1}(u, \theta_k^l)\, \theta_k^l \quad \text{for all } u \in M_k. \tag{5.2}$$

The sum in (5.2) is taken over the nodes of $\widetilde{M}_k$, $d_k^i = A(\theta_k^i, \theta_k^i)$ and α is a constant which is chosen to ensure that K_k is nonnegative.

In general, it suffices to take $\alpha = 1/3$. The nonnegativity of K_k can be seen as follows. For $u \in M_k$,

$$A(R_k A_k u, u) = 1/3 \sum_i A(u, \theta_k^i)^2 A(\theta_k^i, \theta_k^i)^{-1} \le 1/3 \sum_i A_{\Omega_k^i}(u, u). \tag{5.3}$$

Here Ω_k^i denotes the support of θ_k^i. Each triangle of the mesh $\{\tau_k^l\}$ appears at most three times on the right-hand sum in (5.3). The nonnegativity of K_k immediately follows.

The operator R_k is clearly a symmetric semidefinite operator on M_k with range contained in $\widetilde{M}_k$. We note that the space $\widetilde{M}_k$ is a finite element space developed from a mesh with elements of quasi-uniform size. It is shown in [9] that (3.4) holds with the discrete norms defined to be the $L^2(\Omega)$ norm. Note that this definition of R_k provides a smoothing operator which only involves smoothing on the nodes of M_k in the interior of the refinement subregion Ω_k. We will prove the following proposition.

Proposition 5.1. *There exists a sequence of operators* $Q_k : M_j \mapsto M_k$ *for* $k = 0, \ldots, j$, *with* $Q_j = I$, *satisfying* (3.3). *Moreover,* $(Q_k - Q_{k-1})v \in \widetilde{M}_k$ *for all*

$v \in M_j$. *Consequently, Algorithms* S *and* N *using* $A(\cdot, \cdot)$, $L^2(\Omega)$*-inner products, and* R_k *by* (5.2), *converge with rates bounded by*

$$\delta_j = 1 - \frac{1}{Cj}.$$

Proof. We need only construct a sequence of operators Q_k, $k = 0, \ldots, j-1$, satisfying (3.3) with $(Q_k - Q_{k-1})v \in \widetilde{M}_k$ for all $v \in M_k$. To do this, we use a technique given in [9]. Let h_0 denote the mesh size of the original coarse mesh and define $h_k = 2^{-k}h_0$. We clearly have that $\lambda_k \leq ch_k^{-2}$. Let $\overline{M}_k$ be the space defined from the full mesh of size h_k, i.e., $\overline{M}_k$ is the space obtained above when taking $\Omega_k = \Omega_{k-1} = \cdots = \Omega_0 = \Omega$. Clearly, $\overline{M}_k$ satisfies the usual approximation and inverse properties, which imply that the $L^2(\Omega)$ projection operator $\overline{Q}_k$ onto $\overline{M}_k$ satisfies

$$\|(I - \overline{Q}_k)v\| \leq c\lambda_k^{-1/2} A^{1/2}(v, v) \tag{5.4}$$

and

$$A(\overline{Q}_k v, \overline{Q}_k v) \leq CA(v, v), \tag{5.5}$$

for all $v \in H_0^1(\Omega)$. Moreover, both M_k and $\overline{M}_k$ have the same mesh restricted to Ω_k and $M_k \subseteq \overline{M}_k$. Let $v \in M$. For $k = 0, \ldots, j-1$, define $Q_k v = w$, where w is the unique function in M_k satisfying

$$w = \begin{cases} \overline{Q}_k v & \text{at the nodes of } M_k \text{ in the interior of } \Omega_{k+1}, \\ v & \text{at the remaining nodes of } M_k. \end{cases}$$

Note that from the construction, $(Q_k - Q_{k-1})v$ is a function in M_k with support in Ω_k, i.e., it is in $\widetilde{M}_k$. We are left to verify (3.3).

We first prove that

$$\|(I - Q_k)v\|^2 \leq c\lambda_k^{-1} A(v, v), \tag{5.6}$$

from which the first inequality of (3.3) follows. By the definition of Q_k and the triangle inequality,

$$\begin{aligned} \|(I - Q_k)v\| &= \|v - w\| = \|v - w\|_{\Omega_{k+1}} \\ &\leq \|(I - \overline{Q}_k)v\| + \|\overline{Q}_k v - w\|_{\Omega_{k+1}}, \end{aligned} \tag{5.7}$$

where $\|\cdot\|_{\Omega_{k+1}}$ denotes the L^2 norm on Ω_{k+1}. By (5.4), it suffices to estimate the second term on the right-hand side of (5.7) by the first. But by the definition of w,

$$\|\overline{Q}_k v - w\|_{\Omega_{k+1}}^2 \leq Ch_k^2 \sum_i (\overline{Q}_k v(x_k^i) - v(x_k^i))^2,$$

where the sum is taken over the nodes x_k^i of M_k on $\partial\Omega_{k+1}$. Clearly,

$$h_k^2 \sum_i (\overline{Q}_k v(x_k^i) - v(x_k^i))^2 \leq C\|(I - \overline{Q}_k)v\|.$$

This proves (5.6) and completes the proof of the first inequality of (3.3).

To prove the second inequality of (3.3), it obviously suffices to prove

$$A((I - Q_k)v, (I - Q_k)v) \leq cA(v, v). \tag{5.8}$$

As above, we write

$$\begin{aligned} A^{1/2}((I - Q_k)v, (I - Q_k)v) &= A_{\Omega_{k+1}}^{1/2}(v - w, v - w) \\ &\leq A^{1/2}((I - \overline{Q}_k)v, (I - \overline{Q}_k)v) + A_{\Omega_{k+1}}^{1/2}(\overline{Q}_k v - w, \overline{Q}_k v - w). \end{aligned} \tag{5.9}$$

Here, $A_{\Omega_{k+1}}(\cdot, \cdot)$ is defined as in (4.7) but with integration taken only over the subdomain Ω_{k+1}. We apply (5.5) to get

$$A((I - \overline{Q}_k)v, (I - \overline{Q}_k)v) \leq CA(v, v).$$

For the remaining term in (5.9), as above

$$\begin{aligned} A_{\Omega_{k+1}}(\overline{Q}_k v - w, \overline{Q}_k v - w) &\leq C \sum_i (\overline{Q}_k v(x_k^i) - v(x_k^i))^2 \\ &\leq C h_k^{-2} \|(I - \overline{Q}_k)v\| \leq CA(v, v). \end{aligned}$$

This proves the second inequality of (3.3) and hence completes the proof of the proposition. □

Remark 5.1. The use of $L^2(\Omega)$-inner products for discrete inner products sometimes leads to multigrid algorithms which require the solution of Gram matrix problems. However, the special form of the operators R_k enables us to avoid such complications. This is discussed in detail in the appendix.

6. Interface problems with large jumps in coefficients

As the final example of the applications of the theory in §3, in this section we will present some multigrid estimates for second-order problems which may have large jumps in coefficients. We consider the following problem defined on a domain $\Omega \subset R^2$:

$$\begin{aligned} -\nabla(a\nabla)u &= f \quad \text{in } \Omega, \\ u &= 0 \quad \text{on } \partial\Omega. \end{aligned} \tag{6.1}$$

More precisely, we assume

$$\overline{\Omega} = \bigcup_{i=1}^{J} \overline{\Omega}_i, \tag{6.2}$$

where $\{\Omega_i\}$ are mutually disjoint open polygons or tetrahedrons, and set ω_i to be the average value of a over Ω_i. In particular, we shall be interested in applications when there is large variation in $\{\omega_i\}$ but little variation of a over the sets Ω_i. The coefficient a may not be smooth, e.g., it might have jump discontinuities, in which case (6.1) is understood in the weak sense. Consequently,

we assume that $\omega_i > 0$ for each i, and that there are constants c_0 and c_1 not depending on $i = 1, \ldots, J$ satisfying

$$c_0\omega_i D_i(v, v) \leq A_i(v, v) \leq c_1\omega_i D_i(v, v) \quad \text{for all } v \in H^1(\Omega_i). \tag{6.3}$$

Here, $A_i(\cdot, \cdot)$ is defined by

$$A_i(u, v) = \int_{\Omega_i} a\nabla u \cdot \nabla v \, dx.$$

Similarly, $D_i(\cdot, \cdot)$ denotes the Dirichlet form on Ω_i and

$$A(u, v) = \sum_i A_i(u, v). \tag{6.4}$$

The purpose of this section is to apply the results of §3 to provide estimates for appropriately defined multigrid algorithms for (6.1) which depend only on the constants c_0, c_1 of (6.3) but not on the values of ω_i. This means that $\max_i \omega_i / \min_i \omega_i$ can be very large without significantly reducing the rate of convergence for these multigrid algorithms. To achieve this rate of convergence, we use a discrete inner product which is weighted by the coefficients $\{\omega_i\}$ and a coarse triangulation which aligns with the boundaries of $\{\Omega_i\}$. By this we mean that $\bigcup_i \partial\Omega_i$ is a subset of $\bigcup_l \partial\tau_0^l$, where $\{\tau_0^l\}$ is the coarse triangulation of Ω (see §4).

As far as we know, the dependence of the elliptic regularity estimates (in the weighted norms) is not known for this type of problem. Consequently, the regularity and approximation assumptions necessary for the standard theory are not available. Some analysis for a fixed number of levels has been given in [23] although the estimates given there tend rapidly to one as the number of levels is increased. In contrast, the bounds we shall derive here only deteriorate quadratically with the number of levels (see Proposition 6.1).

To begin our analysis, we introduce the following weighted inner products:

$$(u, v)_{L^2_\omega(\Omega)} = \sum_{i=1}^{J} \omega_i (u, v)_{L^2(\Omega_i)}, \tag{6.5}$$

and

$$(u, v)_{H^1_\omega(\Omega)} = \sum_{i=1}^{J} \omega_i D_i(u, v), \tag{6.6}$$

with the induced norms denoted by $\|\cdot\|_{L^2_\omega(\Omega)}$ and $\|\cdot\|_{H^1_\omega(\Omega)}$, respectively. Notice that by (6.3), $A^{1/2}(\cdot, \cdot)$ is equivalent to $\|\cdot\|_{H^1_\omega(\Omega)}$. As is done in §4, we assume that Ω is triangulated by a nested sequence of quasi-uniform meshes $\{\tau_k : k = 0, \ldots, j\}$ with $\{\partial\Omega_i\}$ being a subgrid of $\{\tau_0^l\}$. Corresponding to these triangulations, as in §4, we have the multilevel spaces

$$M_0 \subset M_1 \subset \cdots \subset M_j.$$

The operators Q_k needed in the analysis of §3 can be taken to be the weighted L^2 projections $Q_k^\omega : L^2(\Omega) \mapsto M_k$ defined by

$$(Q_k^\omega u, v)_{L^2_\omega(\Omega)} = (u, v)_{L^2_\omega(\Omega)} \quad \text{for all } u \in L^2(\Omega), v \in M_k. \tag{6.7}$$

The following result is taken from [23] for the verification of (3.3).

Lemma 6.1. *Assume the decomposition* (6.2) *has no cross points, namely there is no point on the interface that belongs to more than two* $\overline{\Omega}_i$*'s. Then, for all* $u \in H_0^1(\Omega)$,

$$\|(I - Q_k^\omega)u\|_{L^2_\omega(\Omega)} \leq Ch_k \|u\|_{H^1_\omega(\Omega)}$$

and

$$\|Q_k^\omega u\|_{H^1_\omega(\Omega)} \leq C\|u\|_{H^1_\omega(\Omega)}.$$

More generally, in the presence of cross points, we have, for all $u \in M_j$,

$$\|(I - Q_k^\omega)u\|_{L^2_\omega(\Omega)} \leq Ch_k \left(\log \frac{h_k}{h_j}\right)^{1/2} \|u\|_{H^1_\omega(\Omega)}$$

and

$$\|Q_k^\omega u\|_{H^1_\omega(\Omega)} \leq C \left(\log \frac{h_k}{h_j}\right)^{1/2} \|u\|_{H^1_\omega(\Omega)}.$$

Remark 6.1. The first part of the above result holds in three dimensions but, in general, the second does not [23].

To completely define the multigrid algorithms, we need only define the discrete inner products and the smoothing operators $\{R_k\}$. The discrete inner products are defined by the weighted inner product (6.5). We define the operator R_k by

$$R_k v = \alpha \sum_i (d_k^i)^{-1} (v, \theta_k^i)_{L^2_\omega(\Omega)} \theta_k^i, \tag{6.8}$$

Here, $d_k^i = A(\theta_k^i, \theta_k^i)$ and α is a constant chosen so that K_k is nonnegative. From the discussion in §5, it suffices to take $\alpha = 1/3$.

By the inverse inequality, we have that the largest eigenvalue of A_k (defined by (2.2)) is bounded by Ch_k^{-2}. As a consequence, Lemma 6.1 implies that the assumption (3.3) holds with C_1 and C_2 satisfying

$$C_1 \leq \hat{C}_1 j^\gamma, \qquad C_2 \leq \hat{C}_2 j^\gamma,$$

where $\hat{C}_1$ and $\hat{C}_2$ are constants independent of k and j. The constant γ is equal to zero if the interface has no cross points and otherwise γ equals one. Applying Theorem 1 gives the following proposition.

Proposition 6.1. *For the operators and inner products defined above, the convergence rate of Algorithm* S *and* N *are bounded by*

$$\delta_j \leq 1 - \frac{1}{Cj^{1+\gamma}},$$

where $\gamma = 0$ *or* 1 *as explained above.*

Remark 6.2. The above proposition holds for three-dimensional applications provided that, for example, there are no internal cross points.

Remark 6.3. The special form of the smoothing operators R_k defined in (6.8) enables us to avoid the solution of Gram matrix problems associated with the inner products appearing in the multigrid algorithms. This is discussed in detail in the appendix.

7. Numerical results

In this section, we provide the results of numerical examples illustrating the theory developed in the earlier sections. Specifically, the actual reduction factor δ_j satisfying (3.1) is numerically computed in several examples. Note that δ_j is the largest eigenvalue of the operator $I - B_j A_j$ and can be computed numerically. We shall provide results in the case of local refinement (see §5) as well as quasi-uniform meshes applied to problems where the coefficients defining the differential operator have jumps (see §§4 and 6). The results of the refinement calculations appear to be independent of the number of levels. In contrast, the examples in which the coefficients have jumps show a slow deterioration in the rate of convergence as the number of levels increases. Both cases show asymptotic convergence behavior which is somewhat better than the worst-case analysis provided by the earlier theory. In all of the reported results, we use a symmetric V-cycle algorithm with $m(k) = 1$ for all k.

We report the numerically computed value of δ_j as a function of the mesh parameters. We note however that, since the operator B_j is symmetric, it can be used as a preconditioner and the overall convergence of the algorithm can be accelerated by preconditioned conjugate gradient iteration. In this case, the condition number of the preconditioned system is bounded by $(1 - \delta_j)^{-1}$.

For the first example, we consider the application of the multigrid algorithm to the finite element equations corresponding to a problem with mesh refinement. For this example, the domain Ω will be the unit square, and we shall approximate the solution to

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

The corresponding form $A(\cdot, \cdot)$ is the Dirichlet form on Ω.

The sequence of grids which we shall consider will be progressively more refined as we approach the corner $(1, 1)$. We will use the scheme described in §5 for generating the mesh. We start by breaking the square into sixteen smaller

squares of side length $1/4$. The coarse triangulation is defined by splitting each of these smaller squares into two triangles, for example, along the diagonal between the bottom left to the upper right corner. For integers $0 < J < j$, we define $\Omega_k = \Omega$ for $k = 0, \ldots, J$ and $\Omega_k = [1 - 2^{J-k}, 1] \times [1 - 2^{J-k}, 1]$ for $k > J$. This generates a sequence of meshes with geometrically decreasing mesh size, with local refinement (for $k > J$) on domains of geometrically decreasing size. Such a mesh would be effective if, for example, the function f in (7.1) behaved like a δ function distribution at the point (1,1). The finite element spaces $M_0 \subset \cdots \subset M_j$ are defined as in §5.

We use (5.2) to define R_k. A more careful analysis shows that, for this example, $I - R_k A_k$ will be a nonnegative operator if α is taken to be $1/2$. We define $(\cdot, \cdot)_k$ to be the $L^2(\Omega)$-inner product.

Table 7.1 gives numerically computed values of δ_j as a function of j and J. For example, $J = 4$ and $j = 8$ corresponds to a uniform grid of size $h_J = 1/64$ with four refinement levels in the corner and $h_j = 1/1024$. For all practical purposes, these values are independent of both j and J. In contrast, Proposition 5.1 suggests that they may deteriorate like $1 - c/j$. In this example, all of the domains $\Omega_0, \ldots, \Omega_j$ are squares and the worst-case deterioration does not appear to be occurring.

TABLE 7.1
Values of δ_j for the refinement example

$j - J$	$J = 1$	$J = 2$	$J = 3$	$J = 4$
1	.670	.668	.668	.668
2	.669	.668	.668	.668
3	.669	.668	.668	.668
4	.669	.668	.668	.668

The remaining example considers multigrid applied to a problem with jumps in the coefficients. Specifically, we consider finite element approximation of the solution of the problem (6.1) where Ω is the unit square and the coefficient $a(x)$ is piecewise constant on the coarse grid triangles. Actually, $a(x) = 1$ except in the set $[1/4, 1/2] \times [1/4, 1/2] \cup [1/2, 3/4] \times [1/2, 3/4]$, where it is equal to μ. We shall present numerically computed values of δ_j for $\mu = 1, 2, 1000, 10000$.

For this example, we use quasi-uniform grids and subspaces corresponding to $k = 0, \ldots, J$ described in the previous example. The form $A(\cdot, \cdot)$ is given by (6.4) and R_k is defined by (6.8). As in the previous example, we take $\alpha = 1/2$.

Table 7.2 gives the numerically computed values of δ_j as a function of j and μ. Note that the results for $\mu = 1, 2$ appear to be bounded independently of the number of levels. For $\mu = 2$, this is somewhat better than the bound of Proposition 4.1. In addition, little change is seen in the reduction rates

of $\mu = 1{,}000$ and $\mu = 10{,}000$. Both $\mu = 1{,}000$ and $\mu = 10{,}000$ show a deterioration in the rate of convergence which is better than the theoretical bound of $1 - c/j^2$ provided by Proposition 6.1; in fact, δ_j is approximately $1 - .8/j$ (see the last column of Table 7.2).

TABLE 7.2
Values of δ_j for the discontinuous coefficient example

$j(h_j)$	$\mu = 1$	$\mu = 2$	$\mu = 1,000$	$\mu = 10,000$	$1 - .8/j$
2(1/16)	.57	.59	.62	.62	.60
3(1/32)	.59	.61	.72	.73	.73
4(1/64)	.59	.61	.80	.80	.80
5(1/128)	.59	.61	.84	.85	.84

8. APPENDIX

In this section, we shall consider the implementation of Algorithms S and N. In particular, we shall show that it is sometimes possible to implement these algorithms in a way which is independent of the discrete inner products. This is possible for the applications described in §5 and §6.

Before proceeding, we briefly examine the way finite element equations are described in computer codes. Typically, one represents unknown functions in an approximation space (e.g., M_k) by a vector of nodal coefficients. The computer problem is to find the coefficients of the function $u \in M_k$ satisfying

$$A(u, \theta_k^i) = F(\theta_k^i), \tag{8.1}$$

for all basis function $\{\theta_k^i\}$. The functional F is represented by its action on the basis vectors. One defines the stiffness matrix M by $\tilde{A}_{ij} = A(\theta_k^i, \theta_k^j)$. Note that, in terms of these representations, the stiffness matrix corresponds to a linear operator $\tilde{A}_k$ from M_k onto M_k' (the space of linear functionals on M_k).

As we shall see, it is possible to implement a multigrid algorithm with inner products which are defined from a fixed inner product $(\cdot, \cdot)$ provided that the smoothing operator R_k has the form

$$R_k u = \sum_i r_k^i(u, \theta_k^i)\theta_k^i. \tag{8.2}$$

Here, $\{r_k^i\}$ is a vector of coefficients and $(\cdot, \cdot)$ is a fixed inner product independent of k. We will consider the case of Algorithm S; the case of Algorithm N is similar.

Corresponding to R_k, we define the linear operator $\tilde{R}_k : M_k' \mapsto M_k$ by

$$\tilde{R}_k F = \sum_i r_k^i F(\theta_k^i)\theta_k^i \quad \text{for all } F \in M_k'.$$

Instead of defining a sequence of operators $B_k: M_k \mapsto M_k$, $k = 0, \ldots, j$, we define, by induction, a sequence of operators $\tilde{B}_k: M_k' \mapsto M_k$, $k = 0, \ldots, j$, as follows:

Set $\tilde{B}_0^s = \tilde{A}_0^{-1}$. Assume that $\tilde{B}_{k-1}^s$ has been defined and define $\tilde{B}_k^s g$ for $g \in M_k'$ as follows:

(1) Set $x^0 = 0$ and $q^0 = 0$.

(2) Define x^l for $l = 1, \ldots, m(k)$ by

$$x^l = x^{l-1} + \tilde{R}_k(g - \tilde{A}_k x^{l-1}). \tag{8.3}$$

(3) Define $x^{m(k)+1} = x^{m(k)} + q^p$, where q^i for $i = 1, \ldots, p$ is defined by

$$q^i = q^{i-1} + \tilde{B}_{k-1}^s[(g - \tilde{A}_k x^{m(k)}) - \tilde{A}_{k-1} q^{i-1}].$$

(4) Set $\tilde{B}_k^s g = x^{2m(k)+1}$, where x^l is defined for $l = m(k)+2, \ldots, 2m(k)+1$ by (8.3).

It is not difficult to check that $B_j^s A_j = \tilde{B}_j^s \tilde{A}_j$. In addition, $\tilde{B}_j^s F = B_j^s f$ as long as

$$(f, \theta) = F(\theta) \quad \text{for all } \theta \in M_j.$$

This means that the multigrid algorithm using $\tilde{B}_j^s$ as a preconditioner for (8.1) leads to the same set of iterates as the multigrid algorithm using B_j^s as a preconditioner for (2.1). Thus, the above algorithm is an implementation of Algorithm S provided that we use a fixed inner product $(\cdot, \cdot)$ and R_k has the form of (8.2). Note that the use of P_{k-1}^0 in Step 3 of Algorithm S has been avoided in the above implementation. This is because the natural imbedding of M_k' into M_{k-1}' is used. In addition, the inner products do not appear in the implementation.

Remark 8.1. The above implementation is somewhat simpler than a direct implementation of Algorithm S. This is because the prolongation and restriction parts of the above implementation only involve the relations between finite element spaces. The coefficients of the operator only affect the computation of $\tilde{R}_k$ (see (5.2) and (6.8)) and $\tilde{A}_k$.

BIBLIOGRAPHY

1. O. Axelsson and P. S. Vassilevski, *Algebraic multilevel preconditioning methods,* II, (preprint).
2. D. Bai and A. Brandt, *Local mesh refinement multilevel techniques*, SIAM J. Sci. Statist. Comput. **8** (1987), 109–134.
3. R. E. Bank, and C. C. Douglas, *Sharp estimates for multigrid rates of convergence with general smoothing and acceleration*, SIAM J. Numer. Anal. **22** (1985), 617–633.
4. R. E. Bank and T. Dupont, *An optimal order process for solving finite element equations*, Math. Comp. **36** (1981), 35–51.
5. D. Braess and W. Hackbusch, *A new convergence proof for the multigrid method including the V-cycle*, SIAM J. Numer. Anal. **20** (1983), 967–975.

6. J. H. Bramble and J. E. Pasciak, *New convergence estimates for multigrid algorithms*, Math. Comp. **49** (1987), 311–329.
7. J. H. Bramble, J. E. Pasciak, J. Wang, and J. Xu, *Convergence estimates for product iterative methods with applications to domain decomposition and multigrid*, (preprint).
8. J. H. Bramble, J. E. Pasciak, and J. Xu, *The analysis of multigrid algorithms for nonsymmetric and indefinite elliptic problems*, Math. Comp. **51** (1988), 389–414.
9. J. H. Bramble, J. E. Pasciak, and J. Xu, *Parallel multilevel preconditioners*, Math. Comp. **55** (1990), 1–22.
10. 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), 1–34.
11. J. H. Bramble and J. Xu, *Some estimates for a weighted L^2 projection*, Math. Comp. **56** (1991), 463–476.
12. A. Brandt, *Algebraic multigrid theory: the symmetric case*, Appl. Math. Comput. **19** (1986), 23–56.
13. A. Brandt, *Multi-level adaptive solutions to boundary-value problems*, Math. Comp. **31** (1977), 333–390.
14. W. Hackbusch, *Multi-grid methods and applications*, Springer-Verlag, New York, 1985.
15. R. B. Kellogg, *Interpolation between subspaces of a Hilbert space*, Tech. Note BN-719, Univ. of Maryland, Inst. Fluid Dynamics and Appl. Math., 1971.
16. M. Kočvara and J. Mandel, *A multigrid method for three-dimensional elasticity and algebraic convergence estimates*, Appl. Math. Comput. **23** (1987), 121–135.
17. J. F. Maitre and F. Musy, *Algebraic formalization of the multigrid method in the symmetric and positive definite case—a convergence estimation for the V-cycle*, Multigrid Methods for Integral and Differential Equations (D. J. Paddon and H. Holstein, eds.), Clarendon Press, Oxford, 1985.
18. J. Mandel, *Étude algébrique d'une méthode multigrille pour quelques problèmes de frontière libre*, C.R. Acad. Sci. Paris Sér. I. Math. **298** (1984), 469–472.
19. J. Mandel, *Multigrid convergence for nonsymmetric, indefinite variational problems and one smoothing step*, (Proc. Copper Mtn. Conf. Multigrid Methods), Applied Math. Comput. **19** (1986), 201–216.
20. J. Mandel, S. F. McCormick, and R. Bank, *Variational multigrid theory*, Multigrid Methods (S. McCormick, ed.), SIAM, Philadelphia, PA, 1987, pp. 131–178.
21. P. Vassilevski, *Iterative methods for solving finite element equations based on multilevel splitting of the matrix*, Bulgarian Academy of Sciences, Sofia, Bulgaria, (preprint).
22. J. Xu, *Convergence estimates for some multigrid algorithms*, Proc. 1989 Houston Domain Decomp. Methods Conf. Third Internat. Sympos. Domain Decomposition Methods for Partial Differential Equations (T. Chan, R. Glowinski, J. Periaux, and O. Widlund, eds.), SIAM, Philadelphia, PA, 1990, pp. 174–187.
23. J. Xu, *Theory of multilevel methods*, Ph.D. Thesis, Cornell University and Penn State University, Dept. Math. Rep AM-48, 1989.

DEPARTMENT OF MATHEMATICS, CORNELL UNIVERSITY, ITHACA, NEW YORK 14853
E-mail address: bramble@mathvax.msi.cornell.edu

DEPARTMENT OF APPLIED SCIENCE, BROOKHAVEN NATIONAL LABORATORY, UPTON, NEW YORK 11973
E-mail address: pasciak@bnl.gov

MATHEMATICS DEPARTMENT, THE UNIVERSITY OF WYOMING, LARAMIE, WYOMING 82071

DEPARTMENT OF MATHEMATICS, PENNSYLVANIA STATE UNIVERSITY, UNIVERSITY PARK, PENNSYLVANIA 16802