UNIFORM CONVERGENCE OF MULTIGRID V–CYCLE ITERATIONS FOR INDEFINITE AND NONSYMMETRIC PROBLEMS

James H. Bramble
Do Y. Kwak
and
Joseph E. Pasciak

To appear: SIAM Journal of Numerical Analysis
Dedicated to Professor Seymour Parter on the occasion of the sixty fifth anniversary of his birthday.

1992

Abstract. In this paper, we present an analysis of a multigrid method for nonsymmetric and/or indefinite elliptic problems. In this multigrid method various types of smoothers may be used. One type of smoother which we consider is defined in terms of an associated symmetric problem and includes point and line, Jacobi and Gauss-Seidel iterations. We also study smoothers based entirely on the original operator. One is based on the normal form, that is, the product of the operator and its transpose. Other smoothers studied include point and line, Jacobi and Gauss-Seidel (with certain orderings). We show that the uniform estimates of [6] for symmetric positive definite problems carry over to these algorithms. More precisely, the multigrid iteration for the nonsymmetric and/or indefinite problem is shown to converge at a uniform rate provided that the coarsest grid in the multilevel iteration is sufficiently fine (but not depending on the number of multigrid levels).

1. Introduction.

The purpose of this paper is to study certain multigrid methods for second order elliptic boundary value problems including problems which may be nonsymmetric and/or indefinite. Multigrid methods are among the most efficient methods available for solving the discrete equations associated with approximate solutions of elliptic partial differential equations. Since their introduction by Fedorenko [15],

---

1991 *Mathematics Subject Classification.* 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 non-exclusive, 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. DMS-9007185 and by the U.S. Army Research Office through the Mathematical Sciences Institute, Cornell University. The second author was also partially supported by the Korea Science and Engineering Foundation.

there has been intensive research toward the mathematical understanding of such methods. The reader is referred to [19], [17] and [3] and the bibliographies therein. Most of these works concern symmetric, positive definite elliptic problems although a few consider nonsymmetric and/or indefinite problems. In particular, [1],[18], [10] and [24] deal with such multigrid algorithms and are most closely related to the subject of this paper. All of these papers share the requirement that the coarse grid be sufficiently fine. We shall briefly describe their contents.

The paper of Bank [1] derives uniform convergence estimates for the W–cycle multigrid iteration with both a standard Jacobi smoother and a smoother which uses the operator times its adjoint. In each case, sufficiently many smoothings are required and a sufficiently fine coarse grid depending on the number of smoothings is needed. Some regularity for the elliptic partial differential equation was also required.

Mandel studied the V–cycle iteration and showed that it was effective with only one smoothing and a sufficiently fine coarse grid. His result requires that the underlying partial differential equation satisfy the "full elliptic regularity" hypothesis and generalizes the results of Braess and Hackbusch [2] for the symmetric positive definite problem.

Bramble, Pasciak and Xu [10] studied the symmetric smoother introduced by Bank and showed that the W–cycle and variable V–cycle worked without making the undesirable requirement of "sufficiently many smoothings". Somewhat more than minimal regularity was needed.

In [24], Wang showed that, for the standard V–cycle with one smoothing, the "reduction factor" for the iteration error was bounded by $1-C/J+C_1h_1$ where J is the number of levels, h_1 is the size of the coarsest grid and C and C_1 are constants. This estimate deteriorates with the number of levels and will be less than one only if the coarse grid is subsequently finer as the number of levels increase. Minimal elliptic regularity was assumed.

In this paper uniform iterative convergence estimates for V-cycle multigrid methods applied to nonsymmetric and/or indefinite problems are proved under rather weak assumptions (e.g., the domain need not be convex). Uniform estimates were shown to hold in [6] and [8] for the V–cycle with one smoothing step in the symmetric positive definite case under such hypotheses. We show that these results carry over to the nonsymmetric and/or indefinite case for a variety of smoothers. The coarse grid must be fine enough but need not depend on the number of levels J. Such a condition seems unavoidable since, in many cases, it is needed even for the approximate problem to make sense.

In recent years, some other techniques have been proposed to handle the nonsymmetric indefinite case. One approach in [14], [4] and [7] is to precondition with a symmetric operator and then solve certain normal equations by the conjugate gradient method. One possible advantage of such a method is that some nonsymmetric problems which are not "compact perturbations" of symmetric ones may be treated. Of course, the usual normal equations may be formed and then preconditioned (cf. [7] and [20]); this approach seems to be rather restrictive in that good preconditioners may be difficult to construct. Other recent approaches have included Schwarz type methods [12] and two–level methods in which a "coarse space" is introduced to reduce the problem to one with a positive definite symmetric part (cf. [4], [13] and [25]).

The remainder of the paper is organized as follows: In Section 2, we describe a model problem and introduce the multigrid method. In Section 3, smoothers based on the symmetric problem (and used in our nonsymmetric and/or indefinite applications) are defined and the relevant properties which they satisfy are stated. Section 4 develops smoothers based on the original problem. The main results of the paper, which provide iterative convergence rates for the multigrid algorithms with the smoothers of Sections 3 and 4, are given in Section 5.

2. The problem and multigrid algorithm.

We set up the model nonsymmetric problem and the simplest multigrid algorithm in this section. We consider, for simplicity, the Dirichlet problem in two spatial dimensions approximated by piecewise linear finite elements on a quasi-uniform mesh. The multigrid convergence results hold for many extensions and generalizations as discussed at the end of Section 5.

We consider as our model problem the following second order elliptic equation with homogeneous boundary conditions.

$$
\begin{aligned}
-\sum_{i,j=1}^{2} \frac{\partial}{\partial x_j}(a_{ij}\frac{\partial u}{\partial x_i}) + \sum_{i=1}^{2} b_i \frac{\partial u}{\partial x_i} + au &= f \quad \text{in} \quad \Omega, \\
u &= 0 \quad \text{on} \quad \partial\Omega,
\end{aligned} \tag{2.1}
$$

where Ω is a polygonal domain (possibly nonconvex) in R^2 and $\{a_{ij}(x)\}$ is bounded symmetric, and uniformly positive definite for $x \in \Omega$. We assume that a_{ij} is in the Sobolev space $W_p^{\gamma}(\Omega)$ for $p > 2/\gamma$ (see, [16] for the definition of $W_p^{\gamma}(\Omega)$). Further, we assume that b_i is continuously differentiable on $\bar{\Omega}$ and that $|a|$ is bounded. Finally, we assume that the solution of (2.1) exists.

Let $H^1(\Omega)$ denote the Sobolev space of order one on Ω (cf., [16]) and let $H_0^1(\Omega)$ denote those functions in $H^1(\Omega)$ whose trace vanish on $\partial\Omega$. For $v, w \in H_0^1(\Omega)$, define

$$
A(v,w) = \sum_{i,j=1}^{2} \int_\Omega a_{ij} \frac{\partial v}{\partial x_i}\frac{\partial w}{\partial x_j}\, dx + \sum_{i=1}^{2} \int_\Omega b_i \frac{\partial v}{\partial x_i} w\, dx + \int_\Omega avw\, dx. \tag{2.2}
$$

The solution u of (2.1) satisfies

$$
A(u,v) = (f,v) \qquad \text{for all } v \in H_0^1(\Omega), \tag{2.3}
$$

where $(\cdot,\cdot)$ denotes the inner product in $L^2(\Omega)$.

For the analysis, we introduce a symmetric positive definite form $\hat{A}(\cdot,\cdot)$ which has same second order part as $A(\cdot,\cdot)$. We define $\hat{A}(\cdot,\cdot)$ by

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

The difference is denoted by

$$D(u,v) = A(u,v) - \hat{A}(u,v).$$

The form $D(\cdot,\cdot)$ satisfies the inequalities

$$|D(u,v)| \le C \|u\|_1 \|v\| \text{ and } |D(u,v)| \le C \|u\| \|v\|_1 . \tag{2.4}$$

Here $\|\cdot\|_1$ and $\|\cdot\|$ denote the norms in $H^1(\Omega)$ and $L^2(\Omega)$ respectively. The second inequality above follows from integration by parts. Here and throughout the paper, c or C, with or without subscript, will denote a generic positive constant. These constants can take on different values in different occurrences but will always be independent of the mesh size and the number of levels in multigrid algorithms.

By the assumptions on the coefficients appearing in the definition of $\hat{A}(\cdot,\cdot)$, it follows that the norm $\hat{A}(v,v)^{1/2}$ for $v \in H^1(\Omega)$ is equivalent to the norm on $H^1(\Omega)$. Thus, we take

$$\|v\|_1 = \hat{A}(v,v)^{1/2}.$$

We develop a sequence of nested triangulations of Ω in the usual way. We assume that a coarse triangulation $\{\tau_1^i\}$ of Ω is given. Successively finer triangulations $\{\tau_m^i\}$ for $m > 1$ are defined by subdividing each triangle (in a coarser triangulation) into four by connecting the midpoints of the edges. The mesh size of $\{\tau_1^i\}$ will be denoted to be d_1 and can be taken to be the diameter of the largest triangle. By similarity, the mesh size of $\{\tau_m^i\}$ is $2^{1-m}d_1$.

For theoretical and practical purposes, the coarsest grid in the multilevel algorithms must be sufficiently fine. In practice, however, the coarse grid is still considerably coarser than the solution grid. Let L and J be greater than or equal to one and set M_k, for $k = 1, \ldots, J$, to be the functions which are piecewise linear with respect to the triangulation $\{\tau_{k+L}^i\}$, continuous on Ω and vanish on $\partial\Omega$. Since the triangulations are nested, it follows that

$$M_1 \subset M_2 \subset \ldots \subset M_J.$$

The space M_k has a mesh size of $h_k = 2^{1-L-k}d_1 = 2^{1-k}h_1$.

Fix k in $\{1, 2, \ldots\}$. Let us temporarily assume that for every $u \in M_k$,

$$A(u,v) = 0 \quad \text{for all } v \in M_k \quad \text{implies} \quad u = 0. \tag{2.5}$$

This assumption immediately implies the existence and uniqueness of solutions to problems of the form: Given a linear functional $F(\cdot)$ defined on M_k, find $u \in M_k$ satisfying

$$A(u,\phi) = F(\phi) \qquad \text{for all } \phi \in M_k.$$

In particular, the projection operator $P_k : H^1(\Omega) \mapsto M_k$ satisfying

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

is well defined.

Clearly, if (2.2) has a positive definite symmetric part then (2.5) holds. More generally, if solutions of (2.1) satisfy regularity estimates of the form

$$\|u\|_{1+\alpha} \leq C\|f\|_{-1+\alpha}, \tag{2.6}$$

then, it is well known (cf., [22]) that there exists a constant h_0 such that for $h_k \leq h_0$, (2.5) holds and furthermore

$$||(I - P_k)u|| \leq ch_k^\alpha ||(I - P_k)u||_1. \tag{2.7}$$

and finally,

$$\|P_k u\|_1 \leq C \, \|u\|_1 \, . \tag{2.8}$$

Even if regularity estimates of the form of (2.6) are not known to hold, then (2.5) is known from a recent result by Schatz and Wang [23].

Lemma 2.1 [23]**.** *There exists an h_0 such that (2.5) holds for $h_k \leq h_0$. Moreover, given $\epsilon > 0$, there exists an $h_0(\epsilon) > 0$ such that for all $h_k \in (0, h_0]$, (2.8) holds and*

$$||(I - P_k)u|| \leq \epsilon ||(I - P_k)u||_1. \tag{2.9}$$

Remark 2.1. The above ϵ will appear in our subsequent analysis. We note that ϵ can be taken arbitrarily small. However, L will be taken large enough so that (2.5), (2.8) and (2.9) hold. Thus, the coarse grid size (i.e., L) for any estimate in which ϵ appears will depend on ϵ.

In our analysis, we shall use the orthogonal projectors $\hat{P}_k : H_0^1(\Omega) \mapsto M_k$ and $Q_k : L^2(\Omega) \mapsto M_k$ which, respectively, denote the elliptic projection corresponding to $\hat{A}(\cdot, \cdot)$ and the $L^2(\Omega)$ projection. These are defined by

$$\hat{A}(\hat{P}_k u, v) = \hat{A}(u, v) \quad \text{for all } v \in M_k,$$

and

$$(Q_k u, v) = (u, v) \quad \text{for all } v \in M_k.$$

The multigrid algorithms will be defined in terms of an additional inner product $(\cdot, \cdot)_k$ on $M_k \times M_k$. Examples of this inner product in our applications will be given in the next section. Additional operators are defined in terms of this inner product as follows: For each k, define $A_k : M_k \to M_k$ and $\hat{A}_k : M_k \to M_k$ by

$$(A_k u, v)_k = A(u, v) \quad \text{for all } v \in M_k,$$

and

$$(\hat{A}_k u, v)_k = \hat{A}(u, v) \quad \text{for all } v \in M_k.$$

Finally, the restriction operator $P_{k-1}^0 : M_k \mapsto M_{k-1}$ is defined by

$$(P_{k-1}^0 u, v)_{k-1} = (u, v)_k \qquad \text{for all } v \in M_{k-1}.$$

We seek the solution of

$$A(u, v) = (f, v), \quad \text{for all } v \in M_J. \tag{2.10}$$

This can be rewritten in the above notation as

$$A_J u = Q_J f. \tag{2.11}$$

We describe the simplest V–cycle multigrid algorithm for iteratively computing the solution u of (2.3). Given an initial iterate $u_0 \in M_J$, we define a sequence approximating u by

$$u_{i+1} = \mathrm{Mg}_J(u_i, Q_J f). \tag{2.12}$$

Here $\mathrm{Mg}_J(\cdot, \cdot)$ is a map of $M_J \times M_J$ into M_J and is defined as follows.

Definition MG. *Set $Mg_1(v, w) = A_1^{-1} w$. Let $k > 1$ and v, w be in M_k. Assuming that $Mg_{k-1}(\cdot, \cdot)$ has been defined, we define $Mg_k(v, w)$ by:*

(1) $x_k = v + R_k(w - A_k v)$.

(2) *$Mg_k(v, w) = x_k + q$, where q is defined by*

$$q = Mg_{k-1}(0, P^0_{k-1}(w - A_k x_k)).$$

Here $R_k : M_k \mapsto M_k$ is a linear smoothing operator. Note that in this V–cycle, we smooth only as we proceed to coarser grids.

In Section 3, we define R_k in terms of smoothing operators defined for the form $\hat{A}(\cdot, \cdot)$. Specifically, the smoothing procedure for the symmetric problem will be denoted $\hat{R}_k : M_k \mapsto M_k$ and we set $R_k = \hat{R}_k$. In Section 4, we consider smoothers which are directly defined in terms of the original operator A_k.

A straightforward mathematical induction argument shows that $\mathrm{Mg}_J(\cdot, \cdot)$ is a linear map from $M_J \times M_J$ into M_J. Moreover, the scheme is consistent in the sense that $v = \mathrm{Mg}_J(v, A_J v)$ for all $v \in M_J$. It easily follows that the linear operator $E = \mathrm{Mg}_J(\cdot, 0)$ is the error reduction operator for (2.12), that is

$$u - u_{i+1} = E(u - u_i).$$

Let $T_k = R_k A_k P_k$ for $k > 1$ and set $T_1 = P_1$. Using the facts that $P^0_{k-1} A_k = A_{k-1} P_{k-1}$ and $P_{k-1} P_k = P_{k-1}$ and Definition MG, a straightforward manipulation gives that for $k > 1$ and any $u \in M_J$,

$$u - \mathrm{Mg}_k(0, A_k P_k u) = (I - T_k)u - \mathrm{Mg}_{k-1}(0, A_{k-1} P_{k-1}(I - T_k)u).$$

Let $E_k u = u - \mathrm{Mg}_k(0, A_k P_k u)$. In terms of E_k, the above identity is the same as

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

Moreover, by consistency, $E = E_J$ and hence

$$E = (I - T_1)(I - T_2) \cdots (I - T_J). \tag{2.13}$$

The product representation of the error operator given above will be a fundamental ingredient in the convergence analysis presented in Section 4. Similar representations in the case of multigrid algorithms for symmetric problems were given in [9].

The above algorithm is a special case of more general multigrid algorithms in that we only use pre-smoothing. Alternatively, we could define an algorithm with just post-smoothing or both pre- and post-smoothing. The analysis of these algorithms is similar to that above and will not be presented.

Often algorithms with more than one smoothing are considered [3], [17], [19]. This is not advised in the above algorithm since the smoothing iteration is generally unstable.

3. Smoothers based on the symmetric problem.

In this section, we consider smoothers which are based on the symmetric problem. The symmetric smoother will be denoted by $\hat{R}_k$. We state a number of abstract conditions concerning these smoothing operators. We then give three examples of smoothing procedures which satisfy these assumptions. In Section 5, we provide convergence estimates for multigrid algorithms with $R_k = \hat{R}_k$ in Definition MG.

The first two conditions are standard assumptions used in earlier multigrid analyses. For $k > 1$, let $\hat{K}_k = I - \hat{R}_k\hat{A}_k$ (defined on M_k) and $\hat{T}_k = \hat{R}_k\hat{A}_k\hat{P}_k$ (defined on M_J). We assume that:

(1) There is a constant C_R such that

$$\text{(C.1)} \qquad \frac{(u,u)_k}{\lambda_k} \le C_R(\bar{R}_k u, u)_k, \quad \text{for all } u \in M_k,$$

where $\bar{R}_k = (I - \hat{K}_k^*\hat{K}_k)\hat{A}_k^{-1}$ and λ_k is the largest eigenvalue of $\hat{A}_k$. Here and in the remainder of this paper, $*$ denotes the adjoint with respect to the inner product $\hat{A}(\cdot,\cdot)$.

(2) There is a constant $\theta < 2$ not depending on k satisfying

$$\text{(C.2)} \qquad \hat{A}(\hat{T}_k v, \hat{T}_k v) \le \theta \hat{A}(\hat{T}_k v, v) \qquad \text{for all } v \in M_k.$$

Provided that (C.2) holds, (C.1) is equivalent to

$$\text{(3.1)} \qquad \frac{(u,u)_k}{\lambda_k} \le C(\hat{R}_k u, u)_k, \quad \text{for all } u \in M_k.$$

When $\hat{R}_k$ is symmetric with respect to $(\cdot,\cdot)_k$, (C.2) states that the norm of $\hat{T}_k$ is less than or equal to θ. Even in the case of non-symmetric $\hat{R}_k$, (C.2) implies stability of $(I - \hat{T}_k)$. In fact, for any $w \in M_J$, (C.2) implies that

$$\text{(3.2)} \qquad \begin{aligned} \hat{A}((I-\hat{T}_k)w, (I-\hat{T}_k)w) &= \hat{A}(w,w) - 2\hat{A}(\hat{T}_k w, w) + \hat{A}(\hat{T}_k w, \hat{T}_k w) \\ &\le \hat{A}(w,w) - (2-\theta)\hat{A}(\hat{T}_k w, w) \le \hat{A}(w,w). \end{aligned}$$

The final condition is that for $k > 1$, there exists a constant C satisfying

$$\text{(C.3)} \qquad (\hat{T}_k u, \hat{T}_k u)_k \le C\lambda_k^{-1}\hat{A}(\hat{T}_k u, u) \qquad \text{for all } u \in M_k.$$

A simple change of variable shows that (C.3) is the same as

$$(\hat{R}_k v, \hat{R}_k v)_k \le C\lambda_k^{-1}(\hat{R}_k v, v)_k \qquad \text{for all } v \in M_k.$$

In the case when $\hat{R}_k$ is symmetric, this is equivalent to

$$(\hat{R}_k v, v)_k \le C\lambda_k^{-1}(v, v)_k \qquad \text{for all } v \in M_k \tag{3.3}$$

and is the opposite inequality of (3.1). Note that both (C.2) and (C.3) hold on M_J.

Remark 3.1. If Conditions (C.1)–(C.3) hold for a smoother R_k then they hold for its adjoint R_k^t with respect to the inner product $(\cdot,\cdot)_k$. This means that (C.1) holds for $\bar{R}_k = (I - \hat{K}_k\hat{K}_k^*)\hat{A}_k^{-1}$ and that (C.2) and (C.3) hold with $\hat{T}_k^*$ replacing $\hat{T}_k$. In the case of (C.2) and (C.3), the corresponding inequalities hold with the same constants as those appearing in the original inequalities.

Example 1. The first example of a smoother is the operator

$$\hat{R}_k = \bar{\lambda}_k^{-1} I$$

where I denotes the identity operator on M_k and $\lambda_k \le \bar{\lambda}_k \le C\lambda_k$. In this case, (3.1) holds with $C = \bar{\lambda}_k/\lambda_k$, (C.2) holds with $\theta = 1$ and (3.3) holds with $C = \lambda_k/\bar{\lambda}_k$. To avoid the inversion of L^2 Gram matrices in the multigrid algorithm, we use the inner product

$$(u, v)_k = h_k^2 \sum_i u(x_i)v(x_i). \tag{3.4}$$

Here the sum is taken over all nodes x_i of the subspace M_k. Note that $(\cdot,\cdot)_k$ is uniformly (independent of k) equivalent to $(\cdot,\cdot)$ on M_k.

The remaining smoothers correspond to Jacobi and Gauss-Seidel, point and line iteration methods. We shall present these smoothers in terms of subspace decompositions. Specifically, we write

$$M_k = \sum_{i=1}^{l} M_k^i \tag{3.5}$$

where M_k^i is the one dimensional subspace spanned by the nodal basis function ϕ_k^i or the subspace spanned by the nodal basis functions along a line. The number of such spaces $l = l(k)$ will often depend on k. These spaces satisfy the following inequality.

$$\|v\| \le Ch_k \|v\|_1 \qquad \text{for all } v \in M_k^i. \tag{3.6}$$

Example 2. For the second example, we consider the additive smoother defined by

$$\hat{R}_k = \gamma \sum_{i=1}^{l} \hat{A}_{k,i}^{-1} Q_{k,i}. \tag{3.7}$$

Here $\hat{A}_{k,i} : M_k^i \to M_k^i$ is the defined by

$$(\hat{A}_{k,i} v, \chi)_k = \hat{A}(v, \chi) \quad \text{for all } \chi \in M_k^i$$

and $Q_{k,i} : M_k \to M_k^i$ is the projection onto M_k^i with respect to the inner product $(\cdot,\cdot)_k$. The constant γ is a scaling factor which is chosen to ensure that (C.2) is satisfied (see, e.g., [11],[5]). Note that $\hat{R}_k$ is symmetric with respect to the inner product $(\cdot,\cdot)_k$. In addition, (3.1) and (3.3) are shown to hold in [11] with point Jacobi. When the subspaces M_k^i are defined in terms of lines, (3.1) was proved in [5]. The estimate (3.3) easily follows in the line case using the support properties of the basis functions and (3.6). For this example, we take $(\cdot,\cdot)_k = (\cdot,\cdot)$ for all k.

Example 3. We next consider the multiplicative smoother. Given $f \in M_k$, we define $\hat{R}_k$ by

(1) Set $v_0 = 0 \in M_k$.
(2) Define v_i, for $i = 1, \ldots, l$, by

$$v_i = v_{i-1} + \hat{A}_{k,i}^{-1} Q_{k,i} (f - \hat{A}_k v_{i-1}).$$

(3) Set $\hat{R}_k f = v_l$.

Conditions (C.1) and (C.2) are known for this operator (see, e.g., [5]). The next lemma shows that (C.3) holds for this choice of $\hat{R}_k$. For this case, we also take $(\cdot,\cdot)_k = (\cdot,\cdot)$ for all k.

Lemma 3.1. *(C.3) holds when $\hat{R}_k$ is defined to be the multiplicative smoother of Example 3.*

Proof. The proof uses the techniques for analyzing smoothers presented in [5]. Fix $k > 1$ and let

$$\hat{\mathcal{E}}_i = (I - \hat{P}_k^i)(I - \hat{P}_k^{i-1}) \cdots (I - \hat{P}_k^1) \tag{3.8}$$

where $\hat{P}_k^i$ denotes the $\hat{A}(\cdot,\cdot)$ projection onto M_k^i. Note that $(I - \hat{T}_k) = \hat{\mathcal{E}}_l$ and $\hat{\mathcal{E}}_{i-1} = \hat{\mathcal{E}}_i + \hat{P}_k^i \hat{\mathcal{E}}_{i-1}$. Hence

$$\hat{T}_k = I - \hat{\mathcal{E}}_l = \sum_{i=1}^{l} \hat{P}_k^i \hat{\mathcal{E}}_{i-1}$$

and for every $u \in M_k$, (cf., [5])

$$\begin{aligned}\hat{A}((2I - \hat{T}_k)u, \hat{T}_k u) &= \hat{A}(u,u) - A(\hat{\mathcal{E}}_l u, \hat{\mathcal{E}}_l u) \\ &= \sum_{i=1}^{l} \hat{A}(\hat{P}_k^i \hat{\mathcal{E}}_{i-1} u, \hat{\mathcal{E}}_{i-1} u).\end{aligned}$$

Since $h_k^2 \le c \lambda_k^{-1}$, the proof of the lemma will be complete if we can show that

$$(\hat{T}_k u, \hat{T}_k u) \le c h_k^2 \sum_{i=1}^{l} \hat{A}(\hat{P}_k^i \hat{\mathcal{E}}_{i-1} u, \hat{\mathcal{E}}_{i-1} u). \tag{3.9}$$

Expanding the left hand side of (3.9) gives

$$(\hat{T}_k u, \hat{T}_k u) = \sum_{i=1}^{l} \sum_{j=1}^{l} (\hat{P}_k^i \hat{\mathcal{E}}_{i-1} u, \hat{P}_k^j \hat{\mathcal{E}}_{j-1} u). \tag{3.10}$$

Because of the support properties of $\{\phi_k^i\}$, the subspaces $\{M_k^i\}$ satisfy a limited interaction property in that for every i, the number of subspaces j for which $(v^i, v^j) \neq 0$, with $v^i \in M_k^i$ and $v^j \in M_k^j$ is bounded by a fixed constant n_0 not depending on k or l. Lemma 3.1 of [5] implies that the double sum of (3.10) can be bounded by n_0 times its diagonal, i.e.

$$(\hat{T}_k u, \hat{T}_k u) \leq n_0 \sum_{i=1}^{l} (\hat{P}_k^i \hat{\mathcal{E}}_{i-1} u, \hat{P}_k^i \hat{\mathcal{E}}_{i-1} u). \tag{3.11}$$

Applying (3.6) gives

$$(\hat{P}_k^i \hat{\mathcal{E}}_{i-1} u, \hat{P}_k^i \hat{\mathcal{E}}_{i-1} u) \leq C h_k^2 \hat{A}(\hat{P}_k^i \hat{\mathcal{E}}_{i-1} u, \hat{\mathcal{E}}_{i-1} u). \tag{3.12}$$

Combining (3.11) and (3.12) proves (3.9). This completes the proof of the lemma.

Remark 3.2. The same analysis could be used for successive overrelaxation type iteration. In that case,

$$\hat{\mathcal{E}}_l = (I - \beta \hat{P}_k^l)(I - \beta \hat{P}_k^{l-1}) \cdots (I - \beta \hat{P}_k^1)$$

where $\beta \in (0, 2)$ is the relaxation parameter.

4. Smoothers based on A_k.

In this section, we consider smoothing operators R_k which are defined directly in terms of the nonsymmetric and/or indefinite operator A_k. The first smoother is one that was originally analyzed in [1] and subsequently studied in [10].

Example 4. For our first example of a smoother based on A_k, we consider R_k defined by

$$R_k = \bar{\lambda}_k^{-2} A_k^t.$$

Here, A_k^t is the adjoint of A_k with respect to the inner product $(\cdot, \cdot)_k$ and $\bar{\lambda}_k$ is as in Example 1. A possible motivation for such a choice is that, on M_k, the iteration

$$v^i = v^{i-1} + \bar{\lambda}_k^{-2} A_k^t (f - A_k v^{i-1})$$

is stable in the norm $(\cdot, \cdot)_k^{1/2}$ provided that $\bar{\lambda}_k^2$ is greater than or equal to half the largest eigenvalue of $A_k^t A_k$.

Example 5. This example is closely related to the second example of the previous section. As in that example, we define the line or point subspaces $\{M_k^i\}$ for $i = 1, \ldots, l$. Note that the form $A(\cdot, \cdot)$ satisfies a Gårding inequality

$$c_1 \hat{A}(u, u) - c \|u\|^2 \leq A(u, u) \qquad \text{for all } u \in H_0^1(\Omega).$$

Consequently, by (3.6),

$$(c_1 - Ch_k^2)\hat{A}(u,u) \le A(u,u) \qquad \text{for all } u \in M_k^i.$$

We will assume that h_2 is sufficiently small so that

$$Ch_2^2 \le c_1/2. \tag{4.1}$$

This means that $A(\cdot,\cdot)$ restricted to M_k^i has a positive definite symmetric part. Hence, the projector $P_k^i : M_k \mapsto M_k^i$ satisfying

$$A(P_k^i v, w) = A(v,w) \qquad \text{for all } w \in M_k^i$$

is well defined and satisfies

$$\left\| P_k^i u \right\|_1 \le C \left\| u \right\|_{1,\Omega_k^i}. \tag{4.2}$$

The second norm is taken only over the subdomain Ω_k^i which is the set of points of Ω where the functions in M_k^i are nonzero. In addition, the operator $A_{k,i} : M_k^i \mapsto M_k^i$ defined by

$$(A_{k,i} v, w)_k = A(v,w) \qquad \text{for all } v, w \in M_k^i,$$

is invertible. We set R_k by

$$R_k = \gamma \sum_{i=1}^{l} A_{k,i}^{-1} Q_{k,i}.$$

We choose γ as in Example 2 so that the symmetric smoother defined by (3.7) satisfies (C.2).

Example 6. Our final example is that of Gauss-Seidel directly applied to the non-symmetric/indefinite equations. We assume that the subspaces $\{M_k^i\}$ satisfy the conditions of the previous example and in addition, that l is bounded independently of k. This is possible by doing what is commonly referred to as a coloring scheme. Starting with subspaces satisfying (3.6), we group together those whose supports overlap at most on sets of measure zero and thus reducing the size of l. For a regular mesh on the square, we could group together all of the subspaces associated with the odd lines (similarly, those associated with the even lines). Since we are grouping subspaces with essentially disjoint supports, (3.6) holds on the larger subspaces. The block Gauss-Seidel algorithm (based on A_k) is given as follows:

(1) Set $v_0 = 0 \in M_k$.
(2) Define v_i, for $i = 1, \ldots, l$, by

$$v_i = v_{i-1} + A_{k,i}^{-1} Q_{k,i}(f - A_k v_{i-1}).$$

(3) Set $R_k f = v_l$.

5. Analysis of the multigrid iteration (2.12).

We provide an analysis of the multigrid iteration (2.12) in this section. This analysis is based on the product representation of the error operator (2.13). All of the analysis of this section is based on perturbation from the uniform convergence estimates for multigrid applied to symmetric problems.

We start by stating a result from [6] estimating the rate of convergence for the multigrid algorithm applied to the symmetric problem. Specifically, we replace A_k by $\hat{A}_k$ and R_k by $\hat{R}_k$ in Definition MG. Set $\hat{T}_1 = \hat{P}_1$. From the earlier discussion, the error operator associated with this iteration applied to finding solution of the symmetric problem

$$\hat{A}_J u = Q_J f$$

is given by $\hat{E} = \hat{E}_J$ where

$$\hat{E}_k = (I - \hat{T}_1)(I - \hat{T}_2) \cdots (I - \hat{T}_k). \tag{5.1}$$

We then have the following theorem.

Theorem 5.1 [6]**.** *For $k > 1$, let $\hat{R}_k$ satisfy (C.1) and (C.2). Under the assumptions on the domain Ω and the coefficients of (2.1) given in Section 2, there exists a positive constant $\hat{\delta} < 1$ not depending on J such that*

$$\hat{A}(\hat{E}_J u, \hat{E}_J u) \le \hat{\delta}^2 A(u, u) \qquad \text{for all } u \in M_J.$$

To analyze the multigrid algorithms using the smoothers of Section 3, we use the perturbation operator

$$Z_k = T_k - \hat{T}_k.$$

We note that for any $u, v \in M_J$, for $k > 1$,

$$\hat{A}(Z_k u, v) = D(u, \hat{T}_k^* v). \tag{5.2}$$

Indeed, by definition,

$$\begin{aligned}
\hat{A}(T_k u, v) &= (T_k u, \hat{A}_k \hat{P}_k v)_k = (A_k P_k u, \hat{R}_k^t \hat{A}_k \hat{P}_k v)_k \\
&= (A_k P_k u, \hat{T}_k^* v)_k = A(P_k u, \hat{T}_k^* v) \\
&= A(u, \hat{T}_k^* v) = \hat{A}(u, \hat{T}_k^* v) + D(u, \hat{T}_k^* v).
\end{aligned}$$

The equality (5.2) immediately follows.

To handle the case of $k = 1$, we have

$$\hat{A}(Z_1 u, v) = D((I - P_1)u, \hat{P}_1 v). \tag{5.3}$$

In fact, by definition,

$$\begin{aligned}
\hat{A}(P_1 u, v) &= \hat{A}(P_1 u, \hat{P}_1 v) \\
&= A(u, \hat{P}_1 v) - D(P_1 u, \hat{P}_1 v) \\
&= \hat{A}(\hat{P}_1 u, v) + D((I - P_1)u, \hat{P}_1 v).
\end{aligned}$$

The following theorem provides an estimate for the multigrid algorithm when the smoothers of Section 3 are used.

Theorem 5.2. *Let $R_k = \hat{R}_k$ and assume that (C.1)–(C.3) hold. Given $\epsilon > 0$, there exists an $h_0 > 0$ such that for $h_1 \leq h_0$,*

$$\hat{A}(Eu, Eu) \leq \delta^2 \hat{A}(u,u) \qquad \text{for all } u \in M_J,$$

for $\delta = \hat{\delta} + c(h_1 + \epsilon)$. Here $\hat{\delta}$ is less than one (independently of J) and is given by Theorem 5.1.

Proof. For an arbitrary operator $\mathcal{O} : M_J \mapsto M_J$, let $||\mathcal{O}||_{\hat{A}}$ denote its operator norm, i.e.,

$$||\mathcal{O}||_{\hat{A}} = \sup_{u,v\in M_J} \frac{\hat{A}(\mathcal{O}u, v)}{\hat{A}(u,u)^{1/2}\hat{A}(v,v)^{1/2}}.$$

Applying (2.4), (2.9) and (2.8) to (5.3) gives

$$|\hat{A}(Z_1 u, v)| \leq C\epsilon \left\|(I - P_1)u\right\|_1 \left\|v\right\|_1 \leq C\epsilon \left\|u\right\|_1 \left\|v\right\|_1 .$$

This means that the operator norm of Z_1 is bounded by $C\epsilon$. Since the operator norm of $(I - \hat{P}_1)$ is less than or equal to one, the triangle inequality implies that the operator norm of $(I - P_1) = (I - \hat{P}_1 - Z_1)$ is bounded by $1 + C\epsilon$.

For $k > 1$, applying (2.4), (C.3), Remark 3.1, and (3.2) to (5.2) gives

$$\begin{aligned} |\hat{A}(Z_k u, v)| &\leq c h_k \left\|u\right\|_1 \hat{A}(\hat{T}_k v, v)^{1/2} \\ &\leq c h_k \left\|u\right\|_1 \left\|v\right\|_1 , \end{aligned}$$

i.e., the operator norm of Z_k is bounded by ch_k. Since, by (3.2), the operator norm of $(I - \hat{T}_k)$ is less than or equal to one, the triangle inequality implies that the operator norm of $(I - T_k) = (I - \hat{T}_k - Z_k)$ is less than or equal to $1 + ch_k$. Hence, it follows that

$$||E_k||_{\hat{A}} \leq (1 + C\epsilon) \prod_{i=2}^{k} (1 + ch_i) \leq C.$$

It is immediate from the definitions that

$$E_k - \hat{E}_k = (E_{k-1} - \hat{E}_{k-1})(I - \hat{T}_k) - E_{k-1} Z_k. \tag{5.4}$$

By (3.2) and the above estimates, for $k > 1$,

$$\begin{aligned} ||E_k - \hat{E}_k||_{\hat{A}} &\leq ||E_{k-1} - \hat{E}_{k-1}||_{\hat{A}} ||I - \hat{T}_k||_{\hat{A}} + ||E_{k-1}||_{\hat{A}} ||Z_k||_{\hat{A}} \\ &\leq ||E_{k-1} - \hat{E}_{k-1}||_{\hat{A}} + C h_k. \end{aligned} \tag{5.5}$$

Repetitively applying (5.5) and using

$$||E_1 - \hat{E}_1||_{\hat{A}} = ||Z_1||_{\hat{A}} \leq C\epsilon$$

gives that

$$||E_J - \hat{E}_J||_{\hat{A}} \leq C\epsilon + C \sum_{k=2}^{J} h_k \leq c(h_1 + \epsilon).$$

The theorem follows from the triangle inequality and Theorem 5.1.

Remark 5.1. Note that ϵ can be made arbitrarily small by taking h_1 small enough. Consequently, Theorem 5.2 shows that the multigrid iteration converges with a rate which is independent of J provided that the coarse grid is fine enough. The coarse grid mesh size can also be taken to be independent of J.

We next consider the case of Example 4. For this example, we consider first the multigrid algorithm for the symmetric problem which uses

$$\hat{R}_k = \bar{\lambda}_k^{-2}\hat{A}_k \tag{5.6}$$

as a smoother. From the discussion in Section 2, the iteration (2.12) with $\hat{R}_k$ (given by (5.6)) and $\hat{A}_k$ replacing, respectively, R_k and A_k in Definition MG, gives rise to the error operator given by (5.1) where, as above, for $k > 1$, $\hat{T}_k = \hat{R}_k\hat{A}_k\hat{P}_k$. The smoother (5.6) does not satisfy (C.1) and so the first step in the analysis of the nonsymmetric and/or indefinite example is to provide a uniform estimate for $\hat{E}_J$ given by (5.1). Such an estimate is provided in the following theorem. Its proof is given in the appendix.

Theorem 5.3. *Let $\hat{E}_J$ be given by (5.1) where $\hat{T}_k = \hat{R}_k\hat{A}_k\hat{P}_k$ and $\hat{R}_k$ is defined by (5.6). Then,*

$$\hat{A}(\hat{E}_J u, \hat{E}_J u) \le \hat{\delta}^2 A(u,u) \qquad \text{for all } u \in M_J.$$

Here $\hat{\delta}$ is less that one and independent of J.

We can now prove the convergence estimate for multigrid applied to (2.1) using the smoother of Example 4.

Theorem 5.4. *Let R_k be defined by Example 4. Given $\epsilon > 0$, there exists an $h_0 > 0$ such that for $h_1 \le h_0$,*

$$\hat{A}(Eu, Eu) \le \delta^2 \hat{A}(u,u) \qquad \text{for all } u \in M_J,$$

for $\delta = \hat{\delta} + c(h_1 + \epsilon)$. Here $\hat{\delta}$ is less than one (independently of J) and is given by Theorem 5.3.

Proof. For $k > 1$, we consider the perturbation operator

$$Z_k = T_k - \hat{T}_k = \bar{\lambda}_k^{-2}(A_k^t A_k P_k - \hat{A}_k^2\hat{P}_k).$$

Clearly,

$$Z_k = \bar{\lambda}_k^{-2}[A_k^t(A_k P_k - \hat{A}_k\hat{P}_k) + (A_k^t - \hat{A}_k)\hat{A}_k\hat{P}_k]. \tag{5.7}$$

As in (5.2),

$$\bar{\lambda}_k^{-1}\hat{A}((A_k P_k - \hat{A}_k\hat{P}_k)u, v) = \bar{\lambda}_k^{-1}D(u, \hat{A}_k\hat{P}_k v)$$

from which it follows using (2.4) that

$$||\bar{\lambda}_k^{-1}(A_k P_k - \hat{A}_k\hat{P}_k)||_{\hat{A}} \le ch_k.$$

A similar argument shows that

$$\|\bar{\lambda}_k^{-1}(A_k^t - \hat{A}_k)\hat{P}_k\|_{\hat{A}} \le c h_k.$$

It is not difficult to show that

$$\|A_k^t\|_{\hat{A}} \le C\bar{\lambda}_k.$$

Combining the above estimates with (5.7) gives

$$\begin{aligned} \|Z_k\|_{\hat{A}} \le \|\bar{\lambda}_k^{-1}A_k^t\|_{\hat{A}}\|\bar{\lambda}_k^{-1}(A_kP_k - \hat{A}_k\hat{P}_k)\|_{\hat{A}} \\ + \|\bar{\lambda}_k^{-1}(A_k^t - \hat{A}_k)\hat{P}_k\|_{\hat{A}}\|\bar{\lambda}_k^{-1}\hat{A}_k\hat{P}_k\|_{\hat{A}} \le c h_k. \end{aligned}$$

The remainder the proof is exactly the same as that of Theorem 5.2. This completes the proof the theorem.

We next consider the case of Example 5. We use perturbation from the multigrid algorithm for $\hat{A}$ which uses the smoother $\hat{R}_k$ defined by Example 2. Theorem 5.1 provides a uniform estimate for the operator norm of $\hat{E}_J$.

Theorem 5.5. *Let R_k be defined by Example 5. Given $\epsilon > 0$, there exists an $h_0 > 0$ such that for $h_1 \le h_0$,*

$$\hat{A}(Eu, Eu) \le \delta^2 \hat{A}(u,u) \qquad \text{for all } u \in M_J,$$

for $\delta = \hat{\delta} + c(h_1 + \epsilon)$. Here $\hat{\delta}$ is less than one (independently of J) and is given by Theorem 5.1 applied to $\hat{R}_k$ defined in Example 2.

Proof. For this case, the perturbation operator Z_k is given by

$$Z_k = \gamma \sum_{i=1}^{l} (P_k^i - \hat{P}_k^i).$$

As in (5.3),

$$\hat{A}((P_k^i - \hat{P}_k^i)u, v) = D((I - P_k^i)u, \hat{P}_k^i v).$$

Applying (2.4), (3.6) and (4.2) gives

$$\hat{A}((P_k^i - \hat{P}_k^i)u, v) \le c h_k \|u\|_{1,\Omega_k^i} \|v\|_{1,\Omega_k^i} \tag{5.8}$$

and hence

$$\hat{A}(Z_k u, v) \le c h_k \sum_{i=1}^{l} \|u\|_{1,\Omega_k^i} \|v\|_{1,\Omega_k^i}.$$

Using the limited overlap properties of the domains Ω_k^i gives

$$\|Z_k\|_{\hat{A}} \le c h_k.$$

The remainder of the proof of the theorem is exactly the same as that given in the proof of Theorem 5.2.

We finally consider the case of Example 6. We use perturbation from the multigrid algorithm for $\hat{A}$ which uses the smoother $\hat{R}_k$ defined by Example 3. Theorem 5.1 provides a uniform estimate for the operator norm of $\hat{E}_J$.

Theorem 5.6. *Let R_k be defined by Example 6. Given $\epsilon > 0$, there exists an $h_0 > 0$ such that for $h_1 \leq h_0$,*

$$\hat{A}(Eu, Eu) \leq \delta^2 \hat{A}(u, u) \qquad \text{for all } u \in M_J,$$

for $\delta = \hat{\delta} + c(h_1 + \epsilon)$. Here $\hat{\delta}$ is less than one (independently of J) and is given by Theorem 5.1 applied with $\hat{R}_k$ defined as in Example 3.

Proof. The perturbation operator for this example is

$$Z_k = T_k - \hat{T}_k = \hat{\mathcal{E}}_l - \mathcal{E}_l$$

where $\hat{\mathcal{E}}_l$ is given by (3.8) and

$$\mathcal{E}_i = (I - P_k^i)(I - P_k^{i-1}) \cdots (I - P_k^1).$$

As in (5.4),

$$\hat{\mathcal{E}}_i - \mathcal{E}_i = (I - \hat{P}_k^i)(\hat{\mathcal{E}}_{i-1} - \mathcal{E}_{i-1}) - (\hat{P}_k^i - P_k^i)\mathcal{E}_{i-1}.$$

Clearly, the operator norm of $(I - \hat{P}_k^i)$ is bounded by one. Moreover, by (5.8) the operator norm of $(\hat{P}_k^i - P_k^i)$ is bounded by ch_k. It follows that

$$||I - P_k^i||_{\hat{A}} \leq 1 + ch_k$$

and, for $j = 1, \ldots, l$,

$$||\mathcal{E}_j||_{\hat{A}} \leq C.$$

Thus,

$$\begin{aligned} ||\hat{\mathcal{E}}_i - \mathcal{E}_i||_{\hat{A}} &\leq ||\hat{\mathcal{E}}_{i-1} - \mathcal{E}_{i-1}||_{\hat{A}} + ||\hat{P}_k^i - P_k^i||_{\hat{A}} ||\mathcal{E}_{i-1}||_{\hat{A}} \\ &\leq ||\hat{\mathcal{E}}_{i-1} - \mathcal{E}_{i-1}||_{\hat{A}} + Ch_k. \end{aligned} \tag{5.9}$$

Repetitively applying (5.9) and using the fact that for this example, l is bounded independently of k gives that for $k > 1$,

$$||Z_k||_{\hat{A}} \leq Ch_k.$$

The remainder of the proof of this theorem is the same as that of Theorem 5.2.

Remark 5.1. Many extensions and generalizations of the techniques given above are possible. These techniques lead to uniform estimates for multigrid iteration methods for solving nonsymmetric and/or indefinite problems for the following applications.

(1) Approximations using higher order nodal finite element spaces.
(2) Three dimensional problems.
(3) Problems with discontinuous coefficients as discussed in [6].
(4) More general boundary conditions.
(5) Problems with local mesh refinement as described in [11].
(6) Finite element approximation of problems on domains with nonpolygonal boundaries as discussed in [6].

In addition, the perturbation analysis given above can be combined with results for additive multilevel algorithms, for example, Theorem 3.1 of [6]. This leads to new estimates for additive multilevel preconditioning iterations applied to indefinite and nonsymmetric problems. Provided that the coarse grid is sufficiently fine, the operator

$$P = \sum_{k=1}^{J} T_k$$

has a uniformly (independent of J) positive definite symmetric part with respect to the inner product $\hat{A}(\cdot, \cdot)$ and has a uniformly bounded operator norm. These results extend to all of the applications discussed in Remark 5.1.

6. Appendix

We provide a proof of Theorem 5.3 in this appendix. We will apply the analysis given in the proof of Theorem 3.2 of [6]. Note that we cannot directly apply Theorem 3.2 of [6] since the smoother $\hat{R}_k = \bar{\lambda}_k^{-2} \hat{A}_k$ does not satisfy (C.1). We note, however, that Theorem 5.3 will follow from the proof of Theorem 3.2 of [6] if we show that (C.2) holds as well as (3.5) and (3.6) of [6] with $\tilde{T}_k$ replaced by $\hat{T}_k$ defined above. Clearly, (C.2) holds with $\theta = 1$. The remaining two inequalities corresponding to (3.5) and (3.6) of [6] are

$$\hat{A}(\hat{T}_k v, v) \leq (\tilde{C}\eta^{k-l})^2 \hat{A}(v, v) \qquad \text{for all } v \in M_l, \ l < k \tag{6.1}$$

and

$$\hat{A}(v, v) \leq C \sum_{k=1}^{J} \hat{A}(\hat{T}_k v, v) \qquad \text{for all } v \in M_J. \tag{6.2}$$

Here η is less than one and independent of k and l.

From the definition of $\bar{\lambda}_k$, we obviously have

$$\hat{A}(\hat{T}_k v, v) \leq \bar{\lambda}_k^{-1} \hat{A}(\hat{A}_k v, v) = \hat{A}(\tilde{T}_k v, v).$$

As in [6], we have set $\tilde{T}_k = \bar{\lambda}_k^{-1} \hat{A}_k$. Inequality (6.1) follows from Lemma 4.2 of [6].

Inequality (6.2) can be rewritten,

$$\hat{A}(u, u) \leq C \left(\hat{A}(\hat{P}_1 u, u) + \sum_{k=2}^{J} \bar{\lambda}_k^{-2} \left\| \hat{A}_k \hat{P}_k u \right\|_1^2 \right). \tag{6.3}$$

To prove this we proceed as follows. Let $u \in M_J$ and $Q_0 = 0$. Then

$$\begin{aligned} \hat{A}(u, u) &= \sum_{k=1}^{J} \hat{A}(u, (Q_k - Q_{k-1})u) \\ &\leq \left(\hat{A}(\hat{P}_1 u, u) + \sum_{k=2}^{J} \bar{\lambda}_k^{-2} \left\| \hat{A}_k \hat{P}_k u \right\|_1^2 \right)^{1/2} \left(\hat{A}(Q_1 u, Q_1 u) \right. \\ &\qquad \left. + \sum_{k=2}^{J} \bar{\lambda}_k^2 (\hat{A}_k^{-1} (Q_k - Q_{k-1})u, (Q_k - Q_{k-1})u)_k \right)^{1/2}. \end{aligned} \tag{6.4}$$

Now, for $k > 1$,

$$
\begin{aligned}
(\hat{A}_k^{-1}(Q_k - Q_{k-1})u, (Q_k - Q_{k-1})u)_k& \\
&= \sup_{\phi \in M_k} \frac{(\hat{A}_k^{-1/2}(Q_k - Q_{k-1})u, \phi)_k^2}{(\phi, \phi)_k} \\
&= \sup_{\psi \in M_k} \frac{((Q_k - Q_{k-1})u, (Q_k - Q_{k-1})\psi)_k^2}{\|\psi\|_1^2}.
\end{aligned}
$$

By well know approximation properties,

$$((Q_k - Q_{k-1})\psi, (Q_k - Q_{k-1})\psi)_k^{1/2} \leq C \, \|(Q_k - Q_{k-1})\psi\| \leq C h_k \, \|\psi\|_1 .$$

Combining the above estimates gives

$$
\begin{aligned}
\hat{A}(Q_1 u, Q_1 u) + \sum_{k=2}^{J} \bar{\lambda}_k^2 (\hat{A}_k^{-1}(Q_k - Q_{k-1})u, (Q_k - Q_{k-1})u)_k& \\
\leq C \Big(\hat{A}(Q_1 u, Q_1 u) + \sum_{k=2}^{J} \bar{\lambda}_k \, \|(Q_k - Q_{k-1})u\|^2 \Big)& \qquad (6.5) \\
\leq C \hat{A}(u, u).&
\end{aligned}
$$

The last inequality of (6.5) is (4.5) of [6] and also can be found in [21]. Combining (6.4) and (6.5) proves (6.3) and hence completes the proof of the theorem.

References

1. R. Bank, *A comparison of two multilevel iterative methods for nonsymmetric and indefinite elliptic finite element equations*, SIAM J. Numer. Anal. **18** (1981), 724–743.
2. Braess, D. and Hackbusch, W., *A new convergence proof for the multigrid method including the V-cycle*, SIAM J. Numer. Anal. **20** (1983), 967–975.
3. J.H. Bramble, *Multigrid Methods*, Cornell Mathematics Department Lecture Notes, 1992.
4. J.H. Bramble, Z. Leyk, and J.E. Pasciak, *Iterative schemes for non-symmetric and indefinite elliptic boundary value problems*, Math. Comp. (to appear).
5. J.H. Bramble and J.E. Pasciak, *The analysis of smoothers for multigrid algorithms*, Math. Comp. **58** (1992), 467–488.
6. J.H. Bramble and J.E. Pasciak, *New estimates for multigrid algorithms including the V-cycle*, Math. Comp. (to appear).
7. J.H. Bramble and J.E. Pasciak, *Preconditioned iterative methods for nonselfadjoint or indefinite elliptic boundary value problems*, Unification of finite element methods, (Ed. H. Kardestuncer), Elsevier Science Publ. (North-Holland), New York, 1984, pp. 167 – 184.
8. J.H. Bramble and J.E. Pasciak, *Uniform convergence estimates for multigrid V-cycle algorithms with less than full elliptic regularity* (1992), Brookhaven Nat. Lab. #BNL-47892.
9. J.H. Bramble, J.E. Pasciak, J. Wang, and J. Xu, *Convergence estimates for multigrid algorithms without regularity assumptions*, Math. Comp. **57** (1991), 23–45.
10. 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.
11. J.H. Bramble, J.E. Pasciak and J. Xu, *Parallel multilevel preconditioners*, Math. Comp. **55** (1990), 1–22.
12. X.-C. Cai and O. Widlund, *Domain decomposition algorithms for indefinite elliptic problems*, SIAM J. Sci. Stat. Comp. (to appear).

13. X.-C. Cai and J. Xu, *A preconditioned GMRES method for nonsymmetric and indefinite problems*, (Preprint).
14. H.C. Elman, *Iterative methods for large, sparse, nonsymmetric systems of linear equations*, Yale Univ. Dept. of Comp. Sci. Rep. 229, (1982).
15. Fedorenko, R.P., *The speed of convergence of one iterative process*, USSR Comput. Math. and Math. Phys. (1961), 1092–1096.
16. P. Grisvard, *Elliptic Problems in Nonsmooth Domains*, Pitman, Boston, 1985.
17. Hackbusch, W., *Multi-Grid Methods and Applications*, Springer-Verlag, New York, 1985.
18. Mandel, J., *Multigrid convergence for nonsymmetric, indefinite variational problems and one smoothing step*, Proc. Copper Mtn. Conf. Multigrid Methods, vol. 19, Applied Math. Comput., 1986, pp. 201–216.
19. J. Mandel, S. McCormick and R. Bank, *Variational multigrid theory*, Multigrid Methods, Ed. S. McCormick, SIAM, Philadelphia, Penn., 1987, pp. 131–178.
20. T.A. Manteuffel and S.V. Parter, *Preconditioning and boundary conditions*, SIAM J. Numer. Anal. **27** (1990), 656–694.
21. P. Oswald, *On discrete norm estimates related to multilevel preconditioners in the finite element method*, (Preprint).
22. A.H. Schatz, *An observation concerning Ritz-Galerkin methods with indefinite bilinear forms*, Math. Comp. **28** (1974), 959–962.
23. A.H. Schatz and J. Wang.
24. J. Wang, *Convergence analysis of multigrid algorithms for non-selfadjoint and indefinite elliptic problems* (1991), Proceedings of the 5'th Copper Mountain Conference on Multigrid Methods.
25. J. Xu, *A new class of iterative methods for nonsymmetric boundary value problems*, (preprint).

Department of Mathematics
White Hall, Cornell University
Ithaca, NY 14853-7901
E-mail: bramble@math.msi.cornell.edu

Department of Mathematics
Korea Advanced Institute of Science and Technology
Taejon, Korea 305-701
E-Mail: dykwak%math1.kaist.ac.kr

Department of Applied Science
Brookhaven National Laboratory
Upton, NY 11973
E-mail: pasciak@bnl.gov