NEW ESTIMATES FOR MULTILEVEL
ALGORITHMS INCLUDING THE V-CYCLE

JAMES H. BRAMBLE AND JOSEPH E. PASCIAK

ABSTRACT. The purpose of this paper is to provide new estimates for certain multilevel algorithms. In particular, we are concerned with the simple additive multilevel algorithm discussed recently together with J. Xu and the standard V-cycle algorithm with one smoothing step per grid. We shall prove that these algorithms have a uniform reduction per iteration independent of the mesh sizes and number of levels, even on nonconvex domains which do not provide full elliptic regularity. For example, the theory applies to the standard multigrid V-cycle on the L-shaped domain, or a domain with a crack, and yields a uniform convergence rate. We also prove uniform convergence rates for the multigrid V-cycle for problems with nonuniformly refined meshes. Finally, we give a new multigrid approach for problems on domains with curved boundaries and prove a uniform rate of convergence for the corresponding multigrid V-cycle algorithms.

1. INTRODUCTION

In recent years, multigrid methods have been used extensively to efficiently solve the discrete equations which arise in the numerical approximation of partial differential equations (see the references in [13, 18, 21]). In conjunction, there has been intensive research into the theoretical understanding of the convergence properties of these methods (cf. [2, 3, 6, 7, 9–12, 18, 20, 21]). In this paper, we present a new general theory based on two assumptions which are different from those made in earlier works. By using the new theory, we are able to derive some surprising uniform convergence bounds for a number of problems. The earlier theories suggested that the rates of convergence for these applications deteriorated as the number of multigrid levels increased.

Previously, there were two general approaches for proving convergence of multigrid algorithms. The first was based on the so-called regularity and approximation assumption [7]. The verification of this hypothesis used both the approximation properties of the discrete method as well as the regularity properties of the approximated partial differential equation. The theory of [7, 16] only provides a uniform convergence rate for the V-cycle algorithm in the case

of full elliptic regularity. It gives a deteriorating estimate, for example, in the case of an L-shaped domain or a domain with a crack boundary.

The second general approach is based solely on approximation and is given in [9, 12]. The "no regularity" theory gives rise to estimates which deteriorate at least linearly with the number of levels in the multigrid scheme.

In contrast, the theory developed in this paper uses two assumptions. The first assumption replaces the regularity and approximation assumption by a much weaker inequality on the whole space (see (3.1)). As we will demonstrate, this inequality often can be verified in applications where full elliptic regularity fails to hold. In [22], this assumption was shown to hold for the standard application using Besov space arguments. In this paper, we show that the estimate can be verified in this case using regularity properties of elliptic problems which depend only on the domain.

The second assumption is that the underlying discrete operator should be appropriately small when restricted to coarser grid spaces (see (3.5)). This assumption was motivated by an inequality proved by Zhang [23] for the standard application. Zhang's proof used local arguments. In this paper, we provide a more general approach, which has been successfully applied even in cases with nonlocal operators (cf. [4, 10]).

We provide a general theory which shows how the rate of convergence of multigrid algorithms can be bounded in terms of the constants appearing in these two new assumptions. We will give three applications showing that the general theory can be used to prove stronger convergence estimates for V-cycle algorithms. The first applies the general theory to second-order uniformly elliptic problems in d-dimensional Euclidean space. We will show that the V-cycle algorithm, with only one smoothing per grid per iteration, leads to a uniformly convergent algorithm independent of the number of levels. This is true on the L-shaped domain and domains with a crack boundary. In addition, these results are valid for both the additive (multilevel) and multiplicative (standard multigrid) form of the algorithm.

The second application involves an example with a general mesh refinement. The best earlier results known for this problem were given in [9] and [12], where it was shown that the convergence rate for the additive and multiplicative algorithms could deteriorate at worst at a rate of $1 - c/J^2$ and $1 - c/J$, respectively. Here, J is the number of levels and c is a positive constant independent of the number of levels. Applying the general theory developed in this paper, we prove, for instance, that the convergence rate for the V-cycle for this example is bounded uniformly below one, independently of J.

The final example applies to domains with curved boundaries. We construct a simple set of nested multilevel spaces and show that our general theory may be used to prove uniform estimates for the additive multilevel schemes and that the V-cycle multigrid scheme (with one smoothing step per grid) has an associated contraction number bounded uniformly below one.

The outline of the remainder of the paper is as follows. Following [7], we provide a general framework for the development of multilevel algorithms in §2. Section 3 provides a general theory for the analysis of these algorithms based on the new assumptions mentioned above. Section 4 verifies the new assumptions in the case of the quasi-uniform finite element approximation. The theorems of §3 then give new convergence estimates for this application. The case of

mesh refinement is considered in §5. Application of the new theory also leads to uniform convergence estimates. Finally, we consider a multigrid algorithm for domains with curved boundaries in §6. Again, uniform rates of convergence are proven.

2. GENERAL ADDITIVE AND MULTIPLICATIVE MULTIGRID ALGORITHMS

Following [7], we first give a general framework for the development of multigrid algorithms in this section. We next define the additive and multiplicative versions of the multigrid algorithms. The additive version provides a preconditioner for the resulting operator. The multiplicative algorithm also gives rise to a preconditioner which can be used in a simple linear iteration (the standard multigrid approach) or as a convergence accelerator in a conjugate gradient iteration. For convenience, the algorithms are presented 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

$$M_1 \subset M_2 \subset \cdots \subset M_J.$$

Associated with this sequence, we assume that we are given additional subspaces, $\widetilde{M_i} \subseteq M_i$ for $i = 2, 3, \ldots, J$. The multilevel algorithms will involve smoothing only on the subspaces $\{\widetilde{M_i}\}$. In addition, let $A(\cdot, \cdot)$ and $(\cdot, \cdot)$ be symmetric positive definite bilinear forms on M_J. Let $\|\cdot\|$ denote the norm corresponding to $(\cdot, \cdot)$. We shall study multigrid algorithms for the solution of the problem: Given $f \in M_J$, find $v \in M_J$ satisfying

$$(2.1) \qquad A(v, \phi) = (f, \phi) \qquad \text{for all } \phi \in M_J.$$

The multigrid algorithms are described in terms of auxiliary operators. For $k = 1, \ldots, J$, define the operator $A_k : M_k \mapsto M_k$ by

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

The operator A_k is clearly symmetric (in both the $A(\cdot, \cdot)$ and $(\cdot, \cdot)$ inner products) and positive definite. Also define the orthogonal projectors $P_k, Q_k : M_J \mapsto M_k$ by

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

and

$$(Q_k w, \phi) = (w, \phi) \qquad \text{for all } \phi \in M_k.$$

With M_k replaced by $\widetilde{M_k}$, the operators $\widetilde{A_k}$, $\widetilde{P_k}$, and $\overline{Q_k}$ are defined analogously. It is easy to check the fundamental identity $Q_l A_k = A_l P_l$ whenever $l < k$. This and the analogous identity $\widetilde{Q_k} A_k = \widetilde{A_k} \widetilde{P_k}$ will be used in various places throughout this paper.

Equation (2.1) can then be rewritten

$$(2.3) \qquad A_J u_J = f.$$

Both the additive and multiplicative (standard) versions of the multigrid algorithms can be thought of as defining an operator $B_J : M_J \mapsto M_J$ which approximately inverts A_J. The goal of the analysis is to provide estimates for either the spectrum of $B_J A_J$ or an appropriate norm of $I - B_J A_J$.

To introduce smoothing into the multigrid algorithms, we shall use "generic" smoothing operators $R_k \colon M_k \mapsto \widetilde{M}_k$, for $k = 2, \dots, J$. Examples of these operators are given in [6]. The properties which they satisfy will be discussed in the subsequent analysis. We set $R_1 = A_1^{-1}$, i.e., we solve on the coarsest space. The additive multigrid preconditioner is then defined by

$$(2.4) \qquad\qquad B_J^a = \sum_{k=1}^{J} R_k Q_k.$$

To analyze the above preconditioner, we must provide estimates for the spectrum of the operator

$$(2.5) \qquad\qquad B_J^a A_J = \sum_{k=1}^{J} T_k,$$

where $T_k = R_k Q_k A_J = R_k A_k P_k$. Note that $T_1 = P_1$.

We shall always take R_k to be symmetric with respect to the $(\cdot, \cdot)$ inner product when used in the additive algorithm. This implies that $R_k = R_k \overline{Q}_k$ and hence $T_k = R_k \widetilde{A}_k \widetilde{P}_k$. This also results in a symmetric operator B_J^a. In general, preconditioned iterative techniques for symmetric problems are much more effective when applied with symmetric preconditioners. The use of a nonsymmetric preconditioner is inappropriate in this case.

The standard multigrid algorithm is often defined as a process which produces a function $\mathrm{MG}_k(w_k, g_k)$. Here, k is the grid level and $w_k, g_k \in M_k$. The function w_k can be thought of as a given approximation to the solution u_k of

$$(2.6) \qquad\qquad A_k u_k = g_k.$$

The result of the multigrid process is to produce $\mathrm{MG}_k(w_k, g_k) \in M_k$, an improved approximation to the solution $u_k = A_k^{-1} g_k$. A standard presentation of this algorithm is given below.

Algorithm 2.1. *For $k = 1$, define $MG_1(w_1, g_1) = A_1^{-1} g_1$. For $k > 1$, $MG_k(w_k, g_k)$ is defined in terms of $MG_{k-1}(\cdot, \cdot)$ as follows:*

    (1) *Set*

$$(2.7) \qquad\qquad x_k = w_k + R_k^t(g_k - A_k w_k).$$

    (2) *Set $y_k = x_k + q$ where*

$$(2.8) \qquad\qquad q = MG_{k-1}(0, Q_{k-1}(g_k - A_k x_k)).$$

    (3) *Set $MG_k(w_k, g_k) = y_k + R_k(g_k - A_k y_k)$.*

The first and third steps above correspond to smoothing. The second step is a correction step. The operator R_k^t is the $(\cdot, \cdot)$ adjoint of the operator R_k. Many generalizations of the above algorithm exist involving more smoothing and correction iterations [7, 11, 18, 21]. We only consider the above algorithm since the results of this paper are most interesting in this case. We note that the results immediately extend to more general algorithms (see the remark after the second theorem of §3) with more than one correction step (e.g., the W-cycle algorithm) as well as algorithms with more than one smoothing step per level.

The above algorithm results in a very simple error reduction process. For any k, let u_k solve (2.6), $v_k = \mathrm{MG}_k(w_k, g_k)$, $e_k^0 = u_k - w_k$, and $e_k^1 = u_k - v_k$. Note that the error e_k^1 is the resulting error after one application of the multigrid process on the kth subspace with an initial error e_k^0. We shall demonstrate that these errors are related by a linear operator F_k, i.e., $e_k^1 = F_k e_k^0$. This is obviously true for $k = 1$, where $F_1 = 0$. Assume that $e_{k-1}^1 = F_{k-1} e_{k-1}^0$ holds for all $w_{k-1}, g_{k-1} \in M_{k-1}$. From (2.7) we have that $u_k - x_k = K_k^* e_k^0$, where $K_k^* = I - R_k^t A_k$. The function q produced in Step 2 approximates the function

$$\tilde{q} = A_{k-1}^{-1} Q_{k-1}(g_k - A_k x_k)$$
$$= A_{k-1}^{-1} Q_{k-1} A_k(u_k - x_k) = P_{k-1}(u_k - x_k).$$

By assumption,

$$\tilde{q} - q = F_{k-1}\tilde{q},$$

or

$$u_k - x_k - q = u_k - x_k - (I - F_{k-1})P_{k-1}(u_k - x_k).$$

Thus,

$$u_k - y_k = [(I - P_{k-1}) + F_{k-1}P_{k-1}](u_k - x_k).$$

Finally, $e_k^1 = K_k(u_k - y_k)$, where $K_k = (I - R_k A_k)$. Hence,

$$e_k^1 = K_k[(I - P_{k-1}) + F_{k-1}P_{k-1}]K_k^* e_k^0.$$

Thus, we see that the errors on the kth level are related by the linear operator F_k defined by the recurrence

$$(2.9) \qquad F_k = K_k[(I - P_{k-1}) + F_{k-1}P_{k-1}]K_k^*.$$

The multigrid process is often applied repeatedly to develop an iterative method for solving problem (2.3). Given an initial approximation u^0, subsequent approximations are defined by

$$(2.10) \qquad u^{l+1} = \mathrm{MG}_J(u^l, g) \quad \text{for } l = 1, \dots.$$

From the above discussion, the error $e^l = u - u^l$ is given by $e^l = (F_J)^l e^0$. Consequently, the multigrid iterative process corresponds to a linear iterative procedure. This can be written equivalently as

$$(2.11) \qquad u^{l+1} = u^l + B_J^m(g - A_J u^l)$$

for the operator $B_J^m = (I - F_J)A_J^{-1}$. Alternatively, this operator B_J^m can be directly defined by the following algorithm.

Algorithm 2.2. *Define* $B_1^m = A_1^{-1}$. *For* $k > 1$, $B_k^m g$ *for* $g \in M_k$ *is defined as follows:*

(1) *Set*

$$(2.12) \qquad x = R_k^t(g).$$

(2) *Set* $y = x + q$, *where* q *is given by*

$$(2.13) \qquad q = B_{k-1}^m Q_{k-1}(g - A_k x).$$

(3) *Set* $B_k^m g = y + R_k(g - A_k y)$.

It is straightforward to show that B_k^m satisfies (cf. [7])

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

with B_k^m, B_{k-1}^m defined by Algorithm 2.2. This shows that $F_k = I - B_k^m A_k$, for $k = 1, \ldots, J$, i.e., the linear iteration (2.11) with B_J^m defined by Algorithm 2.2 is equivalent to the multigrid iteration (2.10). This is an important observation in that it allows the use of the multigrid process to define preconditioning operators B_J^m. For example, the operator B_J^m can be used as a preconditioner with the conjugate gradient method to develop more effective iteration procedures in many applications. It also allows us to use the operator presentation of Algorithm 2.2 for the analysis of the multigrid iteration.

It was shown in [9] that the error reduction operator associated with Algorithm 2.2 (the standard multigrid algorithm) can be written

$$(2.14) \quad \begin{aligned} (I - B_J^m A_J) &= (I - T_J)(I - T_{J-1}) \cdots (I - T_2)(I - T_1) \\ &\quad \cdot (I - T_1^*)(I - T_2^*) \cdots (I - T_{J-1}^*)(I - T_J^*). \end{aligned}$$

This identity depends upon the assumption that the subspaces are imbedded and that one form is used to define the operators on all levels (see (2.2)).

Note that T_k^* is the adjoint of T_k with respect to the $A(\cdot, \cdot)$ inner product. Comparing (2.5) and (2.14) clearly shows the relation between additive and multiplicative multilevel algorithms.

Remark 2.1. The multigrid algorithms are often defined in terms of inner products $(\cdot, \cdot)_k$ which may vary as a function of k. In this case, $(\cdot, \cdot)_k$ replaces the $(\cdot, \cdot)$ inner product in (2.1) and the operator Q_{k-1} is replaced by $Q_{k-1}' : M_k \mapsto M_{k-1}$ defined by

$$(Q_{k-1}'v, \psi)_{k-1} = (v, \psi)_k \qquad \text{for all } \psi \in M_{k-1}.$$

The reason for introducing (possibly discrete) inner products on each level is that it may appear that the projection Q_k requires the inversion of Gram matrices. In fact, for appropriately defined smoothers [6], this inversion is avoided and Q_k never explicitly appears in the computational algorithm [9].

3. A GENERAL FRAMEWORK FOR THE ANALYSIS OF MULTIGRID ALGORITHMS

We provide a general theory for multigrid algorithms in this section which is based on a number of abstract assumptions. Two of these assumptions are different from those used in earlier analyses of multigrid algorithms. In later sections, we will apply this theory to prove stronger results concerning the convergence rate of multigrid algorithms in certain applications.

We first describe the new assumptions. The first is much weaker than the full regularity and approximation assumption (cf. [7]). Let λ_k denote the largest eigenvalue of A_k. The new assumption is that there exists a constant $C_0 \geq 1$ satisfying

$$(3.1) \qquad A(v, v) \leq C_0 \left[A(P_1 v, v) + \sum_{k=2}^{J} \frac{\|\widetilde{A}_k \widetilde{P}_k v\|^2}{\lambda_k} \right] \qquad \text{for all } v \in M_J.$$

Remark 3.1. The full regularity and approximation assumption is that there is a constant C_K not depending on k such that

$$(3.2) \qquad A((I - P_{k-1})u, u) \leq C_K \frac{\|A_k u\|^2}{\lambda_k} \qquad \text{for all } u \in M_k.$$

Let v be in M_J. Taking $u = P_k v$ in (3.2) and summing over k gives

$$A(v, v) \leq A(P_1 v, v) + C_K \sum_{k=2}^{J} \frac{\|A_k P_k v\|^2}{\lambda_k} \qquad \text{for all } v \in M_J.$$

Thus, (3.2) implies (3.1) in the case when $\widetilde{M_k} = M_k$. In general, the converse is not true.

The following lemma, which will be crucial in applying the general theory, illustrates that the above assumption is much weaker than the standard full regularity and approximation assumption.

Lemma 3.1. *Let A and $\mathscr{A}$ be equivalent quadratic forms on M_J. By this we mean that there are positive constants c_0 and c_1 (not depending on J) satisfying*

$$(3.3) \qquad c_0 A(v, v) \leq \mathscr{A}(v, v) \leq c_1 A(v, v) \qquad \text{for all } v \in M_J.$$

Let $\widetilde{\mathscr{A}_k}$, $\mathscr{P}_k$, $\widetilde{\mathscr{P}_k}$, and Λ_k be the quantities defined with respect to $\mathscr{A}$ corresponding to $\widetilde{A_k}$, P_k, $\widetilde{P}_k$, and λ_k. Assume that (3.1) holds. Then

$$(3.4) \quad \mathscr{A}(v, v) \leq C_0 c_1/c_0 \left[\mathscr{A}(\mathscr{P}_1 v, v) + \sum_{k=2}^{J} \frac{\|\widetilde{\mathscr{A}_k} \widetilde{\mathscr{P}_k} v\|^2}{\Lambda_k} \right] \qquad \text{for all } v \in M_J.$$

Proof. Let v be in M_J. We note that (3.1) can be rewritten as

$$A(v, v) \leq C_0 \left[(A_1^{-1} Q_1 A_J v, A_J v) + \sum_{k=2}^{J} \frac{\|\overline{Q}_k A_J v\|^2}{\lambda_k} \right].$$

Setting $w = A_J v$ gives the equivalent inequality

$$(A_J^{-1} w, w) \leq C_0 \left[(A_1^{-1} Q_1 w, w) + \sum_{k=2}^{J} \frac{\|\overline{Q}_k w\|^2}{\lambda_k} \right].$$

It then follows from (3.3) that

$$c_0 (\mathscr{A}_J^{-1} w, w) \leq C_0 c_1 \left[(\mathscr{A}_1^{-1} Q_1 w, w) + \sum_{k=2}^{J} \frac{\|\overline{Q}_k w\|^2}{\Lambda_k} \right] \qquad \text{for all } w \in M_J.$$

This is just a restatement of (3.4) and hence the proof is complete. $\square$

Remark 3.2. We allow for the constant C_0 appearing in (3.1) to depend on J. The results of the general theorems will always depend in a simple way on this constant. We will provide applications where (3.1) can be proved with C_0 independent of J, even though it is known that, for these applications, the corresponding C_K in (3.2) must tend to infinity. It is also shown in [12] that (3.1) holds with $C_0 = CJ$ for many applications. $\square$

For $k = 2, \ldots, J$, set $\widetilde{T}_k = \lambda_k^{-1}\widetilde{A}_k\widetilde{P}_k$ and $\widetilde{T}_1 = P_1$. The second assumption is that the operator $\widetilde{T}_k$ is "small" when applied to functions in M_l with $l \leq k$. More precisely, we assume that there is a positive number $\varepsilon < 1$ and a positive constant $\widetilde{C}$ satisfying

(3.5) $$A(\widetilde{T}_k w, w) \leq (\widetilde{C}\varepsilon^{k-l})^2 A(w, w) \qquad \text{for all } w \in M_l.$$

Additional assumptions required for the theory are standard and will be stated when needed. However, we note that (3.1) can be rewritten as

(3.6) $$A(v, v) \leq C_0 \sum_{k=1}^{J} A(\widetilde{T}_k v, v) \qquad \text{for all } v \in M_J.$$

The first theorem of this section provides an estimate for the condition number associated with the additive multilevel method. For this result, we use the following hypothesis on the smoothing operator: For $k = 2, \ldots, J$, we assume that R_k is a symmetric operator with respect to $(\cdot, \cdot)$ and satisfies

(3.7) $$C_1 \frac{\|w\|^2}{\lambda_k} \leq (R_k w, w) \leq C_2 \frac{\|w\|^2}{\lambda_k} \qquad \text{for all } w \in \widetilde{M}_k.$$

Without loss of generality, we assume that $C_1 \leq 1 \leq C_2$. Note that for $k > 1$,

$$A(\widetilde{T}_k v, v) = \frac{\|\widetilde{A}_k\widetilde{P}_k v\|^2}{\lambda_k},$$

and hence (3.7) implies

(3.8) $$C_1 A(\widetilde{T}_k v, v) \leq A(T_k v, v) \leq C_2 A(\widetilde{T}_k v, v) \qquad \text{for all } v \in M_J.$$

Theorem 3.1. *Assume that R_k satisfies (3.7) and that (3.1) and (3.5) hold. Then the condition number $K(B_J^a A_J)$ satisfies*

$$K(B_J^a A_J) \leq \left(\frac{\widetilde{C}}{1-\varepsilon}\right)^2 \frac{C_2 C_0}{C_1}.$$

Proof. It suffices to estimate the constants c_1 and c_2 satisfying the inequalities

(3.9) $$c_1 A(\psi, \psi) \leq A(B_J^a A_J \psi, \psi) \leq c_2 A(\psi, \psi) \qquad \text{for all } \psi \in M_J.$$

We first bound the sum on the right-hand side of (3.6). Clearly, for $\psi \in M_J$,

$$\sum_{k=1}^{J} A(\widetilde{T}_k \psi, \psi) = \sum_{k=1}^{J}\sum_{l=1}^{k} A(\widetilde{T}_k \psi, (P_l - P_{l-1})\psi),$$

with $P_0 = 0$. Applying the Schwarz inequality with respect to $A(\widetilde{T}_k \cdot, \cdot)$ and (3.5) gives that

(3.10) $$\sum_{k=1}^{J} A(\widetilde{T}_k \psi, \psi) \leq \widetilde{C}\sum_{k=1}^{J}\sum_{l=1}^{k} \varepsilon^{k-l} A(\widetilde{T}_k \psi, \psi)^{1/2} A((P_l - P_{l-1})\psi, \psi)^{1/2}.$$

Let $\mathscr{E} = 1 + \varepsilon + \varepsilon^2 + \cdots = (1-\varepsilon)^{-1}$. Then for $\{\alpha_k\}, \{\beta_k\}$ arbitrary real vectors, the following inequality is elementary:

(3.11)
$$\sum_{k=1}^{J}\sum_{l=1}^{k} \varepsilon^{k-l}\alpha_k\beta_l \leq \frac{\eta}{\mathscr{E}}\sum_{k=1}^{J}\sum_{l=1}^{k} \varepsilon^{k-l}\alpha_k^2 + \frac{\mathscr{E}}{4\eta}\sum_{k=1}^{J}\sum_{l=1}^{k} \varepsilon^{k-l}\beta_l^2$$

$$\leq \eta\sum_{k=1}^{J}\alpha_k^2 + \frac{\mathscr{E}^2}{4\eta}\sum_{k=1}^{J}\beta_k^2,$$

for $\eta > 0$. Combining (3.10) and (3.11) gives

$$\sum_{k=1}^{J} A(\widetilde{T}_k \psi, \psi) \leq \widetilde{C}\eta \sum_{k=1}^{J} A(\widetilde{T}_k \psi, \psi) + \frac{\widetilde{C}\mathcal{E}^2}{4\eta} A(\psi, \psi).$$

Taking $\eta = (2\widetilde{C})^{-1}$ above and (3.6) show that

$$(3.12) \qquad C_0^{-1} A(\psi, \psi) \leq \sum_{k=1}^{J} A(\widetilde{T}_k \psi, \psi) \leq \widetilde{C}^2 \mathcal{E}^2 A(\psi, \psi).$$

Combining (3.12) with (3.8) shows that (3.9) holds with $c_1 = C_1/C_0$ and $c_2 = C_2\widetilde{C}^2\mathcal{E}^2$. This completes the proof of the theorem. □

The following corollary is an obvious consequence of the theorem. It shows that we may reduce the analysis of the additive algorithm to that for any equivalent quadratic form.

Corollary 3.1. *Assume that R_k satisfies (3.7) and that (3.1) and (3.5) hold. Let $\mathscr{A}$ be an equivalent quadratic form on M_J (satisfying (3.3)) and $\mathscr{B}_J^a$ be the additive preconditioner corresponding to $\mathscr{A}$, i.e.,*

$$\mathscr{B}_J^a = \mathscr{A}_1^{-1} Q_1 + \sum_{k=2}^{J} R_k Q_k.$$

Then the condition number $K(\mathscr{B}_J^a \mathscr{A}_J)$ is bounded independently of J.

Remark 3.3. The upper inequality of (3.7) does not hold for many smoothing operators but it does hold for the additive point smoother (cf. [12, inequality (4.3)]). The additive point smoother is defined on spaces $\widetilde{M}_k$ (of dimension $\widetilde{N}_k$) with a nodal finite element basis $\{\phi_k^i\}$ by

$$(3.13) \qquad R_k v = \sum_{i=1}^{\widetilde{N}_k} A(\phi_k^i, \phi_k^i)^{-1} (v, \phi_k^i) \phi_k^i.$$

Note that the upper inequality in (3.7) fails to hold with general R_k if R_k is too much like A_k^{-1}. In fact, the largest eigenvalue of B_J^a is J when $R_k = A_k^{-1}$, for $k = 1, \ldots, J$. This shows that the convergence of the additive multilevel algorithm may deteriorate if the smoothers do not behave like point methods. For example, the upper inequality in (3.7) holds for point and line relaxation schemes but fails to hold for some block schemes. In contrast, the analysis for the multiplicative version (standard multigrid) will not require the upper inequality of (3.7) and no convergence deterioration can be seen in the resulting algorithms. □

We next provide an analysis of the multiplicative form of the multigrid algorithm (Algorithm 2.1). For the multigrid algorithms, we shall also allow the use of nonsymmetric smoothers. In this case, the lower inequality of (3.7) is replaced by the following conditions on the smoother.

(C.1) There is a constant $C_R \geq 1$ which does not depend on k such that the smoothing procedure satisfies

$$(3.14) \qquad \frac{\|u\|^2}{\lambda_k} \leq C_R(\overline{R}_k u, u) \qquad \text{for all } u \in \widetilde{M}_k.$$

Here, $\overline{R}_k = (I - K_k^* K_k)A_k^{-1}$ (recalling that $K_k = I - R_k A_k$). Note that (3.14) holds with $C_R = 1$ for $k = 1$ since $\overline{R}_1 = A_1^{-1}$. In addition, smoothers in multigrid algorithms must be properly scaled as stated in the following condition.

(C.2) There is a constant $\theta < 2$ not depending on k such that

$$(3.15) \qquad A(T_k v, T_k v) \le \theta A(T_k v, v) \qquad \text{for all } v \in M_k.$$

Finally, $\widetilde{M}_k$ should be an invariant subspace under R_k^t. Explicitly, we require that

(C.3) $R_k = R_k \overline{Q}_k$.

The above conditions are shown to hold in [6] for the smoothing operators corresponding to many variations (including line- and point-based schemes) of Jacobi and Gauss-Seidel iterative procedures. It is easy to see that the lower inequality of (3.7) implies (3.14) (with a slightly different constant) in the case of symmetric R_k. In addition, (C.3) is automatically satisfied for symmetric smoothers since, by definition, the range of R_k is contained in $\widetilde{M}_k$.

The following theorem provides the general convergence result for the multigrid algorithm.

Theorem 3.2. *Let R_k satisfy (C.1)–(C.3) and assume that (3.1) and (3.5) hold. Then*

$$(3.16) \qquad 0 \le A((I - B_J^m A_J)v, v) \le (1 - 1/C_M)A(v, v) \qquad \text{for all } v \in M_J$$

holds for

$$(3.17) \qquad C_M = 2C_0\left(C_R + \frac{\widetilde{C}^2\varepsilon^2}{(1-\varepsilon)^2}\frac{\theta}{2-\theta}\right).$$

Proof. Set $E_0 = I$, and for $k = 1, \dots, J$, define

$$E_k = (I - T_k)(I - T_{k-1})\cdots(I - T_1).$$

Note that by (2.14),

$$A((I - B_J^m A_J)v, v) = A(E_J^* v, E_J^* v),$$

and the lower inequality of (3.16) follows. The proof of the upper inequality requires bounding the norm of E_J^* or, equivalently, the norm of its adjoint E_J.

We first derive some identities involving the above operators. Clearly, for $k = 1, \dots, J$,

$$(3.18) \qquad E_{k-1} - E_k = T_k E_{k-1},$$

from which it follows that

$$(3.19) \qquad I - E_k = \sum_{m=1}^{k} T_m E_{m-1}.$$

It is obvious from (3.18) that

$$(3.20) \quad A(E_{k-1}v, E_{k-1}v) - A(E_k v, E_k v) = A((2I - T_k)E_{k-1}v, T_k E_{k-1}v).$$

Let $\overline{T}_k = \overline{R}_k A_k P_k = (I - K_k^* K_k)P_k$. Clearly,

$$(3.21) \qquad A(\overline{T}_k E_{k-1}v, E_{k-1}v) = A((2I - T_k)E_{k-1}v, T_k E_{k-1}v).$$

Summing (3.20) gives that

$$(3.22) \qquad A(v\,,\,v) - A(E_J v\,,\,E_J v) = \sum_{k=1}^{J} A(\overline{T}_k E_{k-1} v\,,\,E_{k-1} v).$$

Note that the upper inequality of (3.16) immediately follows if we prove that

$$(3.23) \qquad A(v\,,\,v) \le C_M \left[A(v\,,\,v) - A(E_J v\,,\,E_J v) \right].$$

Thus, the proof reduces to showing that $A(v\,,\,v)$ can be bounded by C_M times the sum on the right-hand side of (3.22).

By (3.6),

$$
\begin{aligned}
A(v\,,\,v) &\le C_0 \sum_{k=1}^{J} A(\widetilde{T}_k v\,,\,v) \\
(3.24) \qquad &\le 2C_0 \left(\sum_{k=1}^{J} A(\widetilde{T}_k E_{k-1} v\,,\,E_{k-1} v) \right. \\
&\qquad\qquad \left. + \sum_{k=2}^{J} A(\widetilde{T}_k (I - E_{k-1}) v\,,\,(I - E_{k-1}) v) \right).
\end{aligned}
$$

For $k = 2\,,\,\ldots\,,\,J$, condition (C.1) implies

$$A(\widetilde{T}_k E_{k-1} v\,,\,E_{k-1} v) = \frac{\|\widetilde{A}_k \widetilde{P}_k E_{k-1} v\|^2}{\lambda_k} \le C_R (\overline{R}_k \widetilde{A}_k \widetilde{P}_k E_{k-1} v\,,\,\widetilde{A}_k \widetilde{P}_k E_{k-1} v).$$

By (C.3), the image of $\overline{R}_k$ is in $\widetilde{M}_k$ and $\overline{R}_k \overline{Q}_k = \overline{R}_k$. Consequently,

$$(\overline{R}_k \widetilde{A}_k \widetilde{P}_k E_{k-1} v\,,\,\widetilde{A}_k \widetilde{P}_k E_{k-1} v) = A(\overline{T}_k E_{k-1} v\,,\,E_{k-1} v).$$

Moreover, since $\widetilde{T}_1 = \overline{T}_1$,

$$(3.25) \qquad \sum_{k=1}^{J} A(\widetilde{T}_k E_{k-1} v\,,\,E_{k-1} v) \le C_R \sum_{k=1}^{J} A(\overline{T}_k E_{k-1} v\,,\,E_{k-1} v).$$

For the second sum on the right-hand side of (3.24), we use (3.19) and (3.5) to get

$$
\begin{aligned}
\sum_{k=2}^{J} A(\widetilde{T}_k (v - E_{k-1} v)\,,\,v - E_{k-1} v) &= \sum_{k=2}^{J} \sum_{l=1}^{k-1} \sum_{m=1}^{k-1} A(\widetilde{T}_k T_l E_{l-1} v\,,\,T_m E_{m-1} v) \\
&\le \widetilde{C}^2 \sum_{k=2}^{J} \sum_{l=1}^{k-1} \sum_{m=1}^{k-1} \varepsilon^{2k-l-m} A(T_l E_{l-1} v\,,\,T_l E_{l-1} v)^{1/2} A(T_m E_{m-1} v\,,\,T_m E_{m-1} v)^{1/2} \\
&\le \frac{\widetilde{C}^2}{2} \sum_{k=2}^{J} \sum_{l=1}^{k-1} \sum_{m=1}^{k-1} \varepsilon^{2k-l-m} [A(T_l E_{L-1} v\,,\,T_l E_{l-1} v) + A(T_m E_{m-1} v\,,\,T_m E_{m-1} v)].
\end{aligned}
$$

It suffices to bound either of the two terms on the right-hand side above. By

(C.2),

$$\sum_{k=2}^{J}\sum_{l=1}^{k-1}\sum_{m=1}^{k-1}\varepsilon^{2k-l-m}A(T_lE_{l-1}v\,,\,T_lE_{l-1}v)$$

$$\leq \theta\varepsilon\mathscr{E}\sum_{k=2}^{J}\sum_{l=1}^{k-1}\varepsilon^{k-l}A(T_lE_{l-1}v\,,\,E_{l-1}v)$$

$$= \theta\varepsilon\mathscr{E}\sum_{l=1}^{J-1}\sum_{k=l+1}^{J}\varepsilon^{k-l}A(T_lE_{l-1}v\,,\,E_{l-1}v)$$

$$\leq \theta\varepsilon^2\mathscr{E}^2\sum_{l=1}^{J-1}A(T_lE_{l-1}v\,,\,E_{l-1}v).$$

Using (C.2) and (3.21) gives that

$$A(T_lE_{l-1}v\,,\,E_{l-1}v) \leq (2-\theta)^{-1}A(\overline{T}_lE_{l-1}v\,,\,E_{l-1}v).$$

Thus,

(3.26)
$$\sum_{k=2}^{J}A(\widetilde{T}_k(v-E_{k-1}v)\,,\,v-E_{k-1}v)$$
$$\leq \frac{\widetilde{C}^2\varepsilon^2}{(1-\varepsilon)^2}\frac{\theta}{2-\theta}\sum_{k=1}^{J-1}A(\overline{T}_kE_{k-1}v\,,\,E_{k-1}v).$$

Combining (3.24), (3.25), and (3.26) shows that (3.23) holds with C_M given by (3.17). This completes the proof of the theorem. □

Remark 3.4. The above theorem holds for many generalizations of the multigrid algorithm given in Algorithm 2.1. For example, it holds for W-cycle and other algorithms which use more than one iteration for the coarse-grid correction (Step 2). The proof follows from an argument given in §2 of [9]. The result also holds for algorithms which use more smoothings per grid level as long as one alternates between R_k and R_k^t (see, [6, 11]). The modification to the proof is minor and illustrated in the proof of Theorem 4.3 of [6]. Finally, an analogous contraction result holds for nonsymmetric cycling algorithms where smoothing is only done either before or after the correction step, i.e., Step 1 or Step 3 is skipped.

4. THE QUASI-UNIFORM FINITE ELEMENT APPROXIMATION

In this section, we verify the hypotheses for the general multigrid theory in the case of a model second-order elliptic problem. We first describe the model problem and its finite element approximation. In particular, a nested sequence of quasi-uniform approximation spaces are defined in a standard fashion. Next, some notation concerning Sobolev spaces and the corresponding norms is provided. Finally, the conditions (3.1) and (3.5) are shown to hold with constants that are independent of the mesh parameters. Application of the general theory of the previous section then implies that the multilevel algorithms converge with rates that are independent of the number of levels, even in many examples which do not satisfy full elliptic regularity.

Let Ω be a bounded domain in R^d with polygonal boundary $\partial\Omega$. We will include the case when $\Omega \subset R^2$ is a domain with a crack. We consider the Dirichlet problem

(4.1)
$$\begin{aligned} Lu &= f &&\text{in } \Omega, \\ u &= 0 &&\text{on } \partial\Omega, \end{aligned}$$

where

$$Lv = -\sum_{i,j=1}^{d} \frac{\partial}{\partial x_i}\left(a_{ij}\frac{\partial v}{\partial x_j}\right) + av.$$

Let $\overline{\Omega} = \bigcup \overline{\Omega}_l$, for a fixed number of subdomains $\{\Omega_l\}$, where each Ω_l has a Lipschitz continuous boundary. For each i, j, l assume that a_{ij} is in $W_p^\gamma(\Omega_l)$ for some $\gamma \in (0, 1/2)$ and $p > d/\gamma$. Here, $W_p^\gamma(\Omega)$ is the Sobolev space of order γ defined in terms of the norm $L^p(\Omega)$ (cf. [17]). This condition implies that the coefficients are continuous on Ω_l but may jump across the boundaries. In the case of the additive multilevel algorithm, this assumption can be replaced by the assumption that the functions a_{ij} are in $L^\infty(\Omega)$. We further assume that the matrix $\{a_{ij}(x)\}$ is uniformly positive definite almost everywhere. In addition, we assume, for simplicity, that $a(x) \in L^\infty(\Omega)$ is nonnegative.

Remark 4.1. We impose Dirichlet boundary conditions in the above problem for simplicity. The techniques provided in this section can also be applied to problems with mixed boundary conditions. An example of such an application is given at the end of this section. □

The form $A(\cdot, \cdot)$ is defined for this example by the generalized Dirichlet form corresponding to (4.1), i.e.,

(4.2)
$$A(v, w) = \sum_{i,j=1}^{d} \int_\Omega a_{ij}\frac{\partial v}{\partial x_i}\frac{\partial w}{\partial x_j}\, dx + \int_\Omega a\, vw\, dx.$$

This is defined for all v and w in the Sobolev space $H^1(\Omega)$ (the space of distributions with square-integrable first derivatives).

Note that by the assumptions on the coefficients in (4.1), the quadratic form $A(\cdot, \cdot)$ is uniformly equivalent to the form corresponding to the constant-coefficient operator $-\Delta$. Thus, for the purpose of proving (3.1), it suffices to consider the case of the Laplacian. We will assume that there is an α in $(0, 1]$ such that solutions u of (4.1) with $L = -\Delta$ satisfy the following regularity estimate :

(4.3)
$$\|u\|_{1+\alpha} \le C\,\|f\|_{-1+\alpha}.$$

Here, $\|\cdot\|_{-1+\alpha}$ is the interpolated norm between $L^2(\Omega)$ and $H^{-1}(\Omega)$ (the dual of $H_0^1(\Omega)$). Thus, we assume that the domain results in some elliptic regularity for smooth coefficient problems (but not necessarily full elliptic regularity). This assumption is weak, since (4.3) may not hold for any $\alpha > 0$ for equations with bad coefficients. Such an inequality holds for plane domains with polygonal boundaries, including domains with cracks (cf. [15]).

We shall consider the case of quasi-uniform finite element approximation of the solution of (4.1). To define the approximation spaces, we will first define the

underlying mesh partitioning. We assume that a unit-size coarse finite element partitioning of the original domain Ω is given ($\overline{\Omega} = \bigcup \bar{\tau}_1^i$). For example, we take this partitioning in terms of triangles in the case of two spatial dimensions. For examples of such constructions see [14].

Associated with the mesh partitioning, we are given a rule for refinement. For example, in two dimensions, a triangle can be refined into four by connecting the midpoints of the edges.

The mesh triangulations can now be defined by mathematical induction. The coarse triangulation defined above provides the first grid $\{\tau_1^m\}$. Given that a grid $\{\tau_{k-1}^m\}$ has been defined, the grid $\{\tau_k^m\}$ is defined by refining $\{\tau_{k-1}^m\}$ using the refinement rule. We assume that the mesh size of the kth triangulation is on the order of δ^k for fixed $\delta < 1$.

The finite element space M_k is defined to be a space of piecewise polynomial functions with respect to the mesh $\{\tau_k^i\}$ which are continuous on Ω and vanish on $\partial\Omega$ (cf. [1, 14]). Note that a nodal finite element basis is not required for the application of this section. In the present example, $\widetilde{M}_k = M_k$, i.e., we smooth on all functions of M_k. The simplest two-dimensional case is when M_k consists of functions which are piecewise linear with respect to the triangulation and the refinements are defined by breaking triangles into four by connecting the midpoints of the edges.

We next define the Galerkin approximation to the solution of (4.1). Multiplying (4.1) by a smooth function ϕ which vanishes on $\partial\Omega$, and integrating by parts, gives that u satisfies

$$A(u, \phi) = (f, \phi).$$

Here, $(\cdot, \cdot)$ denotes the $L^2(\Omega)$ inner product. The Galerkin approximation to u is the unique function $u_J \in M_J$ satisfying

(4.4) $A(u_J, \psi) = (f, \psi)$ for all $\psi \in M_J$.

We shall analyze the multilevel iterative methods of §2 for solving (4.4).

To verify the hypotheses of §3, we require some notation for norms in Sobolev spaces. Let s be a nonnegative real number and $\|\cdot\|_{s, \widetilde{\Omega}}$ denote the Sobolev norm of order s on a domain $\widetilde{\Omega}$ (cf. [17, 19]). The $\widetilde{\Omega}$ will be left out of this notation when $\widetilde{\Omega} = \Omega$, and the s will be left out in the case of $L^2(\widetilde{\Omega})$ ($s = 0$).

Remark 4.2. The above assumptions on the mesh sizes imply that λ_k (where A_k is defined by (2.2)) is on the order of δ^{-2k}.

Remark 4.3. For this example, it was shown in [12] that (3.1) holds with $C_0 = CJ$. The next lemma improves this bound and shows that (3.1) holds with $C_0 = C$ (independent of J). This inequality was proved earlier in [22] using Besov space techniques. We base our proof on the well-known fact that a wide class of domains give rise to elliptic regularity.

Lemma 4.1. *Assume that (4.3) holds for $L = -\Delta$ and some $\alpha \in (0, 1]$. Then, for general L, there exists a positive constant C_0 not depending on the number of mesh levels J such that (3.1) holds.*

Proof. Here and in the remainder of this paper, we shall use c with or without subscript to denote a generic positive constant. Such constants will always be

independent of the number of levels in the multilevel algorithm. As already observed, by Lemma 3.1, it suffices to prove (3.1) with $L = -\Delta$. Clearly, for $w \in M_J$,

$$A(w, w) = A(w, Q_1 w) + \sum_{k=2}^{J} (A_k P_k w, (Q_k - Q_{k-1})w)$$

$$\leq \left(A(P_1 w, w) + \sum_{k=2}^{J} \frac{\|A_k P_k w\|^2}{\lambda_k} \right)^{1/2}$$

$$\left(A(Q_1 w, Q_1 w) + \sum_{k=2}^{J} \lambda_k \|(Q_k - Q_{k-1})w\|^2 \right)^{1/2}.$$

Consequently, it suffices to show that

$$(4.5) \qquad A(Q_1 w, Q_1 w) + \sum_{k=2}^{J} \lambda_k \|(Q_k - Q_{k-1})w\|^2 \leq cA(w, w).$$

The following approximation and boundedness properties for the operators $\{Q_k\}$ are well known: For $v \in H_0^1(\Omega)$,

$$(4.6) \qquad \begin{aligned} \|(Q_k - Q_{k-1})v\|^2 &\leq c\lambda_k^{-1} A(v, v) \quad \text{for } k = 2, \ldots, J, \\ A(Q_k v, Q_k v) &\leq cA(v, v) \quad \text{for } k = 1, \ldots, J. \end{aligned}$$

In addition, it is well known that (4.3) and the approximation properties of the finite element spaces imply that

$$(4.7) \qquad \|(I - P_{k-1})v\|_{1-\alpha}^2 \leq c\lambda_k^{-\alpha} A((I - P_{k-1})v, v).$$

Here we have used the fact that $\lambda_k \leq ch_k^{-2}$ for this application. From (4.6), it suffices to bound the terms appearing in the sum on the left-hand side of (4.5). Let β satisfy $0 < \beta < \alpha$. Then

$$(4.8) \qquad \begin{aligned} \sum_{k=2}^{J} \lambda_k \|(Q_k - Q_{k-1})w\|^2 &= \sum_{k=2}^{J} \lambda_k \left\| \sum_{l=k}^{J} (Q_k - Q_{k-1})(P_l - P_{l-1})w \right\|^2 \\ &\leq \sum_{k=2}^{J} \lambda_k \left(\sum_{l=k}^{J} \lambda_l^{-\beta} \right) \left(\sum_{l=k}^{J} \lambda_l^{\beta} \|(Q_k - Q_{k-1})(P_l - P_{l-1})w\|^2 \right). \end{aligned}$$

By Remark 4.2, the first sum over l above is bounded by $c\lambda_k^{-\beta}$. Thus, the boundedness of Q_k on $L^2(\Omega)$, the first inequality of (4.6), and interpolation gives

$$\sum_{k=2}^{J} \lambda_k \|(Q_k - Q_{k-1})w\|^2 \leq c \sum_{k=2}^{J} \lambda_k^{\alpha-\beta} \sum_{l=k}^{J} \lambda_l^{\beta} \|(P_l - P_{l-1})w\|_{1-\alpha}^2.$$

Applying (4.7) and changing the order of summation gives

$$\sum_{k=2}^{J} \lambda_k \|(Q_k - Q_{k-1})w\|^2 \leq c \sum_{k=2}^{J} \lambda_k^{\alpha-\beta} \sum_{l=k}^{J} \lambda_l^{\beta-\alpha} A((P_l - P_{l-1})w, w)$$

$$= c \sum_{l=2}^{J} \lambda_l^{\beta-\alpha} A((P_l - P_{l-1})w, w) \sum_{k=2}^{l} \lambda_k^{\alpha-\beta}.$$

Since α is greater than β, Remark 4.2 implies that the sum over k above is bounded by a constant times $\lambda_l^{\alpha-\beta}$. Thus,

$$\sum_{k=2}^{J} \lambda_k \|(Q_k - Q_{k-1})w\|^2 \leq c \sum_{l=2}^{J} A((P_l - P_{l-1})w, w) \leq cA(w, w).$$

This completes the proof of the lemma. □

Remark 4.4. We note that replacing $(Q_k - Q_{k-1})$ by $(I - Q_{k-1})$ in (4.8) and following the proof of the lemma gives that

$$\sum_{k=2}^{J} \lambda_k \|(I - Q_{k-1})w\|^2 \leq cA(w, w).$$

This inequality will be used in §5. □

The next lemma provides a proof of (3.5) for the application described above.

Lemma 4.2. *Let l be less than or equal to k. Then there is a constant $\widetilde{C}$ not depending on the mesh parameters satisfying*

(4.9) $A(\widetilde{T}_k v, v) \leq \widetilde{C}(h_k/h_l)^{2\gamma} A(v, v)$ *for all $v \in M_l$.*

Proof. The proof of Lemma 4.2 is based on the following lemma. Its proof will be given after the proof of Lemma 4.2.

Lemma 4.3. *There exists a constant c such that for all $\eta > 0$, $\phi \in H^1(\Omega)$ and $\psi \in H^{1+\gamma}(\Omega)$,*

(4.10) $|A(\phi, \psi)| \leq c(\eta^{-1}\|\phi\|^2 + \eta^{\gamma/(1-\gamma)}\|\phi\|_1^2)^{1/2}\|\psi\|_{1+\gamma}.$

Assuming Lemma 4.3, we now complete the proof of Lemma 4.2. Let $w \in M_l$. We clearly have that for $k > 2$,

(4.11) $A(\widetilde{T}_k w, w) = \dfrac{\|A_k w\|^2}{\lambda_k} = \lambda_k^{-1}\left(\sup\limits_{\phi \in M_k} \dfrac{A(w, \phi)}{\|\phi\|}\right)^2.$

By Lemma 4.3,

$$|A(w, \phi)| \leq c(\eta^{-1}\|\phi\|^2 + \eta^{\gamma/(1-\gamma)}\|\phi\|_1^2)^{1/2}\|w\|_{1+\gamma}$$
$$\leq c(\eta^{-1} + \eta^{\gamma/(1-\gamma)}h_k^{-2})^{1/2}\|\phi\|\|w\|_{1+\gamma}.$$

Taking $\eta = h_k^{2(1-\gamma)}$ and using the inverse property (see the Appendix of [11] for a proof)

$$\|w\|_{1+\gamma} \leq ch_l^{-\gamma}\|w\|_1$$

gives

(4.12) $|A(w, \phi)| \leq ch_k^{-1+\gamma}h_l^{-\gamma}\|w\|_1\|\phi\|.$

Inequality (4.9) follows combining (4.11), (4.12) and the fact that $\lambda_k \geq ch_k^{-2}$. This completes the proof of Lemma 4.2. □

Proof of Lemma 4.3. This lemma was essentially given in [5] and we will follow its proof. However, the version stated here gives a somewhat more explicit form

of the bounds and requires less regularity on the coefficients than that of Lemma 4.1 of [5].

Let ϕ be in $H^1(\Omega)$ and ψ be in $H^{1+\gamma}(\Omega)$. There is no problem bounding the lowest-order term of (4.2). We need only consider the derivative terms in (4.2). Fix l, and let $\mathbf{E}_l$ denote the extension operator defined on $L^2(\Omega_l)$ given by Theorem 1.4.3.1 of [17]. For a function v defined on Ω_l, let $\tilde{v}$ denote the extension of v by zero to R^d. Since $\gamma < 1/2$, Corollary 1.4.4.5 of [17] gives that the norm $\|\tilde{v}\|_{W_2^\gamma(R^d)}$ is equivalent to the norm $\|v\|_{\gamma,\Omega_l}$ for all $v \in H^\gamma(\Omega_l)$. Thus,

$$\int_{\Omega_l} a_{ij} \frac{\partial \phi}{\partial x_i} \frac{\partial \psi}{\partial x_j}\, dx = \left(\mathscr{F}\left(\frac{\partial(\mathbf{E}_l \phi)}{\partial x_i} \right), \mathscr{F}\left(\widetilde{a_{ij} \frac{\partial \psi}{\partial x_j}} \right) \right),$$

where $\mathscr{F}$ denotes the Fourier transform. By the Schwarz inequality,

$$\int_{\Omega_l} a_{ij} \frac{\partial \phi}{\partial x_i} \frac{\partial \psi}{\partial x_j}\, dx$$

(4.13)
$$\leq c \left(\int_{R^d} \frac{|\zeta|^2}{(1+|\zeta|^2)^\gamma} |\mathscr{F}(\mathbf{E}_l \phi)(\zeta)|^2\, d\zeta \right)^{1/2} \left\| a_{ij} \frac{\partial \psi}{\partial x_j} \right\|_{\gamma,\Omega_l}$$

$$\leq c(\eta^{-1} \|\phi\|^2 + \eta^{\gamma/(1-\gamma)} \|\phi\|_1^2)^{1/2} \left\| \frac{\partial \psi}{\partial x_j} \right\|_\gamma,$$

where $\eta > 0$ is arbitrary. For the last inequality, we used Theorem 1.4.4.2 of [17], which states that multiplication by $a_{ij} \in W_p^\gamma(\Omega_l)$ for $p > d/\gamma$ is a bounded operator on $H^\gamma(\Omega_l)$. The lemma immediately follows by summing over l. □

Remark 4.5. A more constructive proof of Lemma 4.2 is possible in the case of smooth coefficients and nodal finite element approximation spaces. In this case, one proves directly that

$$A(T_k w, w) \leq c(h_k/h_l) A(w, w)$$

holds for $T_k = R_k A_k P_k$, with R_k defined by (3.13). The lemma then follows with $\gamma = 1/2$ from

$$\frac{\|v\|^2}{\lambda_k} \leq c(R_k v, v) \qquad \text{for all } v \in M_k,$$

which is a general smoothing property proved in [6].

Combining Lemmas 4.1 and 4.2 with Theorems 3.1, 3.2, and Corollary 3.1 gives the following theorems.

Theorem 4.1. *Let B_J^a be defined by (2.4) with $A(\cdot, \cdot)$, $(\cdot, \cdot)$, and $\{M_k, \widetilde{M}_k\}$ as described in this section. Assume that R_k (which is symmetric) satisfies (3.7). Then the condition number $K(B_J^a A_J)$ is bounded by a constant which is independent of J.*

Theorem 4.2. *Let B_J^m be defined by Algorithm 2.1 with $A(\cdot, \cdot)$, $(\cdot, \cdot)$, and $\{M_k, \widetilde{M}_k\}$ as described in this section. Assume that R_k satisfies (C.1)–(C.3). Then*

(4.14) $0 \leq A((I - B_J^m A_J)v, v) \leq (1 - 1/C_M)A(v, v)$ *for all $v \in M_J$.*

The constant C_M in (4.14) is independent of J.

Remark 4.6. A uniform result similar to the upper estimate in Theorem 4.1 in the case of R_k defined by (3.13) was announced by Zhang at the Fifth International Symposium on Domain Decomposition Methods [23]. In addition, a version of Theorem 4.1 is proved in [22] using Besov-space equivalences.

Remark 4.7. The results of this section hold for many applications with mixed boundary conditions. We illustrate this by considering a simple example. Specifically, we consider (4.1) with $\Omega = (0, 1) \times (0, 2)$ but with the boundary conditions

$$2\frac{\partial u}{\partial v} = 0 \qquad \text{on } \Gamma_N,$$
$$u = 0 \qquad \text{on } \partial\Omega/\Gamma_N.$$

Here, Γ_N is the line segment $\{(1, y) \mid y \in (1, 2)\}$ and $\frac{\partial}{\partial v}$ denotes the outward conormal derivative on Γ_N. There is no problem with the proof of Lemma 4.2 in this case since the basic ingredient in the proof, Lemma 4.3 does not depend on boundary conditions. Furthermore, the resulting form $A(\cdot, \cdot)$ is equivalent to that corresponding to the Laplacian (with boundary condition $\partial u/\partial n = 0$ on Γ_N). The Laplacian with this boundary condition satisfies (4.3) for any α in $(0, 1/2)$. Thus, Lemma 4.1 holds.

5. General mesh refinement

In this section, we apply the general theory to an approximation which utilizes a locally refined mesh. Such mesh refinements are convenient for accurate modeling of problems with various types of singular behavior. For simplicity, we will consider the piecewise linear finite element approximation, although we will allow a very general form of refinement.

As in the previous section, we consider problem (4.1) and start with a coarse triangulation. The refinement triangulation $\{\tau_k^i\}$ is defined in terms of a sequence of (open) mesh domains

$$\Omega_J \subseteq \Omega_{J-1} \subseteq \cdots \subseteq \Omega_1 = \Omega.$$

The only restrictions on the mesh domains $\{\Omega_k\}$ are that the boundary of Ω_k, for $k > 1$, consists of edges of mesh triangles in the mesh $\{\tau_{k-1}^i\}$, and that there is at least one edge of $\{\tau_{k-1}^i\}$ contained in Ω_k. These mesh domains control the region of refinement. If τ_{k-1}^i is a triangle contained in Ω_k, then it is broken into four smaller triangles (in the triangulation $\{\tau_k^i\}$) by the lines connecting the midpoints of the edges. Alternatively, if τ_{k-1}^i is in the complement of Ω_k, then it is not subdivided but is directly included into the kth triangulation. A simple example of this construction is the case of a unit square with local refinement near the corner $(1, 1)$. In this case, we take $\Omega_k = \Omega$ for $k = 1, \ldots, j$ and $\Omega_k = [1 - 2^{k-j}, 1] \times [1 - 2^{k-j}, 1]$ for $k = j + 1, \ldots, J$.

We consider the piecewise linear finite element approximation, although many extensions are possible. The space M_k is defined to be the set of piecewise linear functions with respect to the mesh $\{\tau_k^i\}$ which are continuous on Ω and vanish on $\partial\Omega$. The continuity condition implies that the finer grid nodes on a coarse-fine boundary are slave nodes in the sense that the values of the

function there are completely determined by the values of the function on the nearby coarse grid points.

For this application, if $\Omega_{k-1} \neq \Omega_k$, then the subspace $\widetilde{M}_k$ on which we smooth is a proper nonzero subspace of M_k. In fact, we define $\widetilde{M}_k$ to be the functions in M_k which are zero outside of Ω_k. Thus, we smooth on a given level just in the region where new nodes are being added in the refinement scheme.

For this application, there is no difficulty in proving (3.5). The largest eigenvalue of M_k is on the order of h_k^{-2}, where h_k is the size of the smallest triangle defining the mesh of M_k. The argument given in the proof of Lemma 4.2 applies with little change and gives that (3.5) holds with a uniform constant $\widetilde{C}$.

To apply Theorems 3.1 and 3.2, we need to prove (3.1). By Lemma 3.1, we need only consider the case when the coefficients defining L are smooth. Let $\overline{M}_k$ denote the quasi-uniform finite element space obtained from refining over the entire domain at each level starting with $\{\tau_1^i\}$, that is, the quasi-uniform space resulting from the above construction with $\Omega = \Omega_1 = \Omega_2 = \cdots = \Omega_J$. Let $\overline{Q}_k$ denote the $(\cdot, \cdot)$ orthogonal projection onto $\overline{M}_k$. A sequence of operators $\mathcal{Q}_k : M_J \mapsto M_k$ was constructed in [9] (see the sequence $\{Q_k\}$ defined in §5 of [9]) which satisfy

(1) $\mathcal{Q}_J$ is the identity.
(2) The range of $\mathcal{Q}_k - \mathcal{Q}_{k-1}$ is contained in $\widetilde{M}_k$.
(3) The inequalities

$$\|(I - \mathcal{Q}_k)v\| \leq c \left\| (I - \overline{Q}_k)v \right\|,$$

$$A(\mathcal{Q}_1 v, \mathcal{Q}_1 v) \leq cA(v, v)$$

hold for all $v \in M_J$ and with constant c independent of k and J.

Following the proof of Lemma 4.1, for $w \in M_J$, we have

$$A(w, w) = A(w, \mathcal{Q}_1 w) + \sum_{k=2}^{J} (\widetilde{A}_k \widetilde{P}_k w, (\mathcal{Q}_k - \mathcal{Q}_{k-1})w)$$

$$\leq \left(A(P_1 w, w) + \sum_{k=2}^{J} \frac{\|\widetilde{A}_k \widetilde{P}_k w\|^2}{\lambda_k} \right)^{1/2}$$

$$\left(A(\mathcal{Q}_1 w, \mathcal{Q}_1 w) + \sum_{k=2}^{J} \lambda_k \|(\mathcal{Q}_k - \mathcal{Q}_{k-1})w\|^2 \right)^{1/2}.$$

Consequently, it suffices to show that

(5.1) $$A(\mathcal{Q}_1 w, \mathcal{Q}_1 w) + \sum_{k=2}^{J} \lambda_k \|(\mathcal{Q}_k - \mathcal{Q}_{k-1})w\|^2 \leq cA(w, w).$$

Clearly, by (3),

$$\|(\mathcal{Q}_k - \mathcal{Q}_{k-1})w\|^2 \leq 2 \left[\|(I - \mathcal{Q}_k)w\|^2 + \|(I - \mathcal{Q}_{k-1})w\|^2 \right]$$

$$\leq c \left\| (I - \bar{Q}_{k-1})w \right\|^2.$$

Thus,

$$A(\mathcal{Q}_1 w , \mathcal{Q}_1 w) + \sum_{k=2}^{J} \lambda_k \|(\mathcal{Q}_k - \mathcal{Q}_{k-1})w\|^2$$

$$\leq c \left(A(w , w) + \sum_{k=2}^{J} \lambda_k \|(I - \overline{Q}_{k-1})w\|^2 \right).$$

Inequality (5.1) follows from Remark 4.4. This shows that (3.1) holds for the refinement application of this section.

We can combine the above observations with Theorems 3.1, 3.2 and Corollary 3.1 to get:

Theorem 5.1. *Let* B_J^q *be defined by* (2.4) *with* $A(\cdot , \cdot)$, $(\cdot , \cdot)$, *and* $\{M_k , \widetilde{M}_k\}$ *as described in this section. Assume that* R_k *(which is symmetric) satisfies* (3.7). *Then the condition number* $K(B_J^q A_J)$ *is bounded by a constant which is independent of* J.

Theorem 5.2. *Let* B_J^m *be defined by Algorithm* 2.1 *with* $A(\cdot , \cdot)$, $(\cdot , \cdot)$, *and* $\{M_k , \widetilde{M}_k\}$ *as described in this section. Assume that* R_k *satisfies* (C.1)–(C.3). *Then*

$$(5.2) \qquad 0 \leq A((I - B_J^m A_J)v , v) \leq (1 - 1/C_M)A(v , v) \qquad \text{for all } v \in M_J.$$

The constant C_M *in* (5.2) *is independent of* J.

6. A CURVED BOUNDARY APPLICATION

In this section, we consider applying the theory developed earlier to a finite element approximation of a boundary value problem with a curved boundary. To remain in the framework of nested spaces, we consider coarser-grid multi-grid spaces M_k which vanish in a neighborhood of order h_k of the domain boundary. Even though these spaces provide a poor approximation, we will show that they lead to multigrid algorithms which converge with uniform rates of reduction.

For convenience, we shall consider a convex domain in R^2 with smooth boundary. Many extensions are possible. We will consider problem (4.1) with the same assumptions on the coefficients as made in §4. The form $A(\cdot , \cdot)$ is defined by (4.2). We also assume that (4.3) holds for $L = -\Delta$ and some α in $(0, 1]$.

We start with a coarse approximate triangulation $\{\tau_1^i\}$ of Ω. By construction, a node will be either in the interior of Ω or on $\partial\Omega$. Without loss of generality, we assume that no triangle of $\{\tau_1^i\}$ has all three vertices on $\partial\Omega$. The triangulation $\{\tau_k^i\}$ will be defined from $\{\tau_{k-1}^i\}$ as follows:

(1) If τ_{k-1}^i is a triangle with two vertices in Ω then τ_{k-1}^i is broken into four finer-grid triangles by the lines connecting the centers of the edges.

(2) A triangle τ_{k-1}^i with two vertices on $\partial\Omega$ results in four finer-grid triangles as illustrated in Figure 6.1. The new boundary point is the midpoint along the boundary arc between the two boundary vertices of τ_{k-1}^i.

Note that not all triangles in $\{\tau_{k-1}^i\}$ can be written as the union of the triangles in $\{\tau_k^i\}$. As a consequence, we will define the coarser multilevel spaces $k < J$ in a different manner than that used for the finest space.
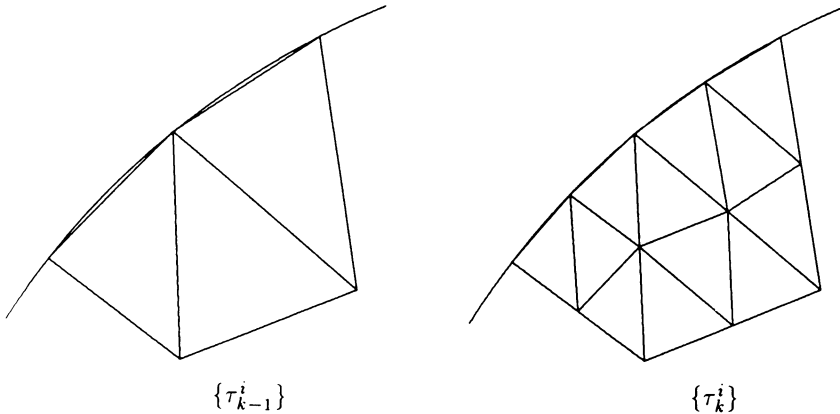
$\{\tau_{k-1}^i\}$ $\{\tau_k^i\}$

FIGURE 6.1. The mesh refinement near the boundary

We use the above grid strategy to define the sequence of approximation meshes up to J and define Ω_k to be the interior of $\bigcup_i \overline{\tau}_k^i$. The fine-grid approximation space M_J is a space of piecewise linear functions with respect to $\{\tau_j^i\}$ which are continuous on Ω_J and vanish on $\partial\Omega_J$.

For $k < J$, let $\overline{M}_k$ denote the analogous space of piecewise linear functions which are defined in terms of the triangulation $\{\tau_k^i\}$. Let Ω_k^0, for $k < J$, be the interior of the union of the closures of triangles of $\{\tau_k^i\}$ which do not have vertices on $\partial\Omega$. The spaces M_k for $k < J$, to be used in the multigrid algorithm, are defined by

$$(6.1) \qquad M_k = \{\phi \in \overline{M}_k \mid \operatorname{supp}(\phi) \in \overline{\Omega}_k^0\}.$$

Because of the convexity of Ω, it is easy to see that $\Omega_k \subseteq \Omega$ for $k = 1, \ldots, J$. Consequently, functions in $\overline{M}_k$ extended by zero are in $H_0^1(\Omega)$. Thus, we shall consider the spaces $\overline{M}_k$ and M_k as being contained in $H_0^1(\Omega)$. Moreover, the triangles of $\{\tau_{k-1}^i\}$ in Ω_k^0 can be written as the union of triangles in $\{\tau_k^i\}$. This implies that the multigrid spaces defined by (6.1) are nested, i.e.,

$$M_1 \subset M_2 \subset \cdots \subset M_J \subset H_0^1(\Omega).$$

The argument of Lemma 4.2 goes through without change for this application. To apply the theory of §4, we need only prove (3.1). From the proof of Lemma 4.1, it clearly suffices to prove the inequalities (4.6) and (4.7) for this application. To this end, we introduce the following lemma.

Lemma 6.1. *Let* Ω^η *denote the strip* $\{x \in \Omega \mid \operatorname{dist}(x, \partial\Omega) < \eta\}$ *and* $0 < s < 1/2$. *Then for all* $v \in H^{1+s}(\Omega)$,

$$(6.2) \qquad \|v\|_{1, \Omega^\eta} \leq c\eta^s \|v\|_{1+s}.$$

In addition, for $v \in H_0^1(\Omega)$,

$$(6.3) \qquad \|v\|_{\Omega^\eta} \leq c\eta \|v\|_1.$$

Proof. Note that if ω^η is a reference square of side length η, then

$$\|w\|_{\omega^\eta}^2 \leq \eta^2/2 \int_{\omega^\eta} |\nabla w|^2 \, dx$$

holds for all functions w vanishing on one edge of ω^η. Inequality (6.3) then easily follows by the use of local maps. A relatively straightforward interpolation argument (see Chapter 1, Section 15 of [19]) using the real method [19] between (6.3) and the trivial inequality $\|w\|_{\Omega^\eta} \leq \|w\|$ gives that (for $s \neq 1/2$)

$$(6.4) \qquad \|w\|_{\Omega^\eta} \leq c(s)\eta^s \|w\|_s \quad \text{for all } w \in H_0^s(\Omega).$$

By Corollary 1.4.4.5 of [17], $H_0^s(\Omega) = H^s(\Omega)$ for $s < 1/2$. Let $v \in H^{1+s}(\Omega)$. Applying (6.4) to v and the first derivatives of v implies (6.2). This completes the proof of the lemma. $\square$

We next prove (4.7). Note that by Lemma 3.1, we need only prove (4.7) for $A(\cdot, \cdot)$ corresponding to the Laplacian. Let $\alpha \in (0, 1/2)$ be such that (4.3) holds. Based on the standard finite element duality argument (cf., [1, 14]), (4.7) will follow if we show that for all $w \in H^{1+\alpha}(\Omega) \cap H_0^1(\Omega)$,

$$(6.5) \qquad \|(I - P_{k-1})w\|_1 \leq c\lambda_k^{-\alpha/2} \|w\|_{1+\alpha}.$$

Fix w in $H^{1+\alpha}(\Omega) \cap H_0^1(\Omega)$. Let χ denote the function in M_{k-1} which interpolates w on the nodes of Ω_{k-1}^0 and $\overline{\chi}$ be the function in $\overline{M}_{K-1}$ which interpolates w at the nodes of Ω_{k-1}. From the definition of P_{k-1}, we immediately have that

$$(6.6) \quad \begin{aligned} A((I - P_{k-1})w, (I - P_{k-1})w) &\leq A(w - \chi, w - \chi) \\ &\leq c(\|w - \overline{\chi}\|_{1,\Omega_{k-1}^0}^2 + \|\overline{\chi} - \chi\|_{1,\Omega_{k-1}^0}^2 + \|w\|_{1,\Omega\backslash\Omega_{k-1}^0}^2). \end{aligned}$$

Applying the Bramble-Hilbert Lemma and well-known techniques, we conclude that

$$(6.7) \qquad \|w - \overline{\chi}\|_{1,\Omega_{k-1}}^2 \leq ch_k^{2\alpha} \|w\|_{1+\alpha}^2.$$

Note that $\overline{\chi} - \chi$ is a mesh function (in $\overline{M}_{k-1}$) which vanishes on all nodes except those on $\partial\Omega_{k-1}^0$. Consequently,

$$(6.8) \qquad \|\overline{\chi} - \chi\|_{1,\Omega_{k-1}^0}^2 \leq c\sum_{x_i} \overline{\chi}(x_i)^2,$$

where the sum is taken over the nodes x_i of $\overline{M}_{k-1}$ on $\partial\Omega_{k-1}^0$. Let $\widetilde{\Omega}_{k-1} = \Omega_{k-1}\backslash\Omega_{k-1}^0$. Then (6.7) and (6.8) imply that

$$(6.9) \quad \begin{aligned} \|\overline{\chi} - \chi\|_{1,\Omega_{k-1}^0}^2 &\leq c \|\overline{\chi}\|_{1,\widetilde{\Omega}_{k-1}}^2 \leq c(\|w - \overline{\chi}\|_{1,\widetilde{\Omega}_{k-1}}^2 + \|w\|_{1,\widetilde{\Omega}_{k-1}}^2) \\ &\leq c(h_k^{2\alpha} \|w\|_{1+\alpha}^2 + \|w\|_{1,\Omega\backslash\Omega_{k-1}^0}^2). \end{aligned}$$

Applying Lemma 6.1 gives

$$(6.10) \qquad \|w\|_{1,\Omega\backslash\Omega_{k-1}^0}^2 \leq ch_k^{2\alpha} \|w\|_{1+\alpha}^2.$$

Combining (6.6), (6.7), (6.9) and (6.10) proves (6.5), i.e., (4.7) holds for this application.

We now prove (4.6). By the triangle inequality, the first inequality will follow if we can show that for all $v \in M_J$,

$$(6.11) \qquad \|(I - Q_k)v\|^2 \leq c\lambda_k^{-1} A(v, v) \quad \text{for } k = 1, \ldots, J - 1.$$

Let $\overline{Q}_k$ denote the $L^2(\Omega_k)$ orthogonal projector onto the space $\overline{M}_k$. By using the interpolant, it is easy to prove that

$$\|(I - \overline{Q}_k)v\|^2_{\Omega_k} \le ch_k^{1+\alpha} \|v\|^2_{1+\alpha}$$

for all $v \in H_0^1(\Omega) \cap H^{1+\alpha}(\Omega)$. Interpolating (using the real method [19]) between this inequality and the trivial inequality

$$\|(I - \overline{Q}_k)v\|^2_{\Omega_k} \le \|v\|^2$$

gives

$$\|(I - \overline{Q}_k)v\|^2_{\Omega_k} \le c\lambda_k^{-1} \|v\|^2_1 \quad \text{for all } v \in H_0^1(\Omega).$$

Fix $v \in H_0^1(\Omega)$ and let θ_k denote the function in M_k which interpolates $\overline{Q}_k v$ at the nodes in Ω_k^0. Then, Lemma 6.1 and the triangle inequality give

$$
\begin{aligned}
\|(I - Q_k)v\|^2 &\le \|v - \theta_k\|^2 \\
&\le c(\|(I - \overline{Q}_k)v\|^2_{\Omega_k^0} + \|\overline{Q}_k v - \theta_k\|^2_{\Omega_k^0} + \|v\|^2_{\Omega \setminus \Omega_k^0}) \\
&\le c(\lambda_k^{-1} \|v\|^2_1 + \|\overline{Q}_k v - \theta_k\|^2_{\Omega_k^0} + \lambda_k^{-1} \|v\|^2_1).
\end{aligned}
$$

The difference $\overline{Q}_k v - \theta_k$ is a mesh function which vanishes on the nodes of Ω_k^0 and hence

$$
\begin{aligned}
\|\overline{Q}_k v - \theta_k\|^2_{\Omega_k^0} &\le ch_k^2 \sum_{x_i} (\overline{Q}_k v(x_i))^2 \le c\|\overline{Q}_k v\|^2_{\Omega_k \setminus \Omega_k^0} \\
&\le c(\|\overline{Q}_k v - v\|^2_{\Omega_k \setminus \Omega_k^0} + \|v\|^2_{\Omega_k \setminus \Omega_k^0}) \le c\lambda_k^{-1} \|v\|^2_1.
\end{aligned}
$$

The sum on x_i above is over the nodes on $\partial \Omega_k^0$. Combining the above inequalities proves (6.11).

We finally prove the second inequality of (4.6). Fix $v \in H_0^1(\Omega)$ and let $\overline{\theta}_k$ denote the $L^2(\Omega_k)$ orthogonal projection of v onto the space of discontinuous piecewise linear functions with respect to the triangulation $\{\tau_k^i\}$. Note that

$$(6.12) \qquad
\begin{aligned}
\|v - \overline{\theta}_k\|^2_{\Omega_k} &\le c\lambda_k^{-1} \|v\|^2_{1, \Omega_k}, \\
\sum_{\tau_k^i \subset \Omega_k} \|\overline{\theta}_k\|^2_{1, \tau_k^i} &\le c \|v\|^2_{1, \Omega_k}.
\end{aligned}
$$

Consequently, by (6.11) and (6.12),

$$\sum_{\tau_k^i \subset \Omega_k} \|\overline{\theta}_k - Q_k v\|^2_{1, \tau_k^i} \le c\lambda_k \|\overline{\theta}_k - Q_k v\|^2_{\Omega_k} \le cA(v, v).$$

The second inequality of (4.6) follows by the triangle inequality and the second inequality of (6.12).

Since (4.6) and (4.7) are valid, the argument proving Lemma 4.1 implies that (3.1) holds for this application. Applying the theory of §3 gives the following theorems.

Theorem 6.1. *Let B_J^q be defined by (2.4) with $A(\cdot, \cdot)$, $(\cdot, \cdot)$, and $\{M_k\}$ as described in this section. Set $\widetilde{M}_k = M_k$ for $k = 2, \ldots, J$. Assume that R_k*

(*which is symmetric*) *satisfies* (3.7). *Then the condition number* $K(B_J^q A_J)$ *is bounded by a constant which is independent of* J.

Theorem 6.2. *Let* B_J^m *be defined by Algorithm* 2.1 *with* $A(\cdot,\cdot)$, $(\cdot,\cdot)$, *and* $\{M_k\}$ *as described in this section. Set* $\widetilde{M}_k = M_k$ *for* $k = 2, \ldots, J$. *Assume that* R_k *satisfies* (C.1)–(C.3). *Then*

$$(6.13) \qquad 0 \le A((I - B_J^m A_J)v, v) \le (1 - 1/C_M)A(v, v) \quad \text{for all } v \in M_J.$$

The constant C_M *in* (6.13) *is independent of* J.

Remark 6.1. Clearly, the techniques of §§5 and 6 could be combined to yield similar results for problems with curved boundaries and nonuniform mesh refinements.

BIBLIOGRAPHY

1. A. K. Aziz and I. Babuška, *Survey lectures on the mathematical foundations of the finite element method*, Part I, The Mathematical Foundations of the Finite Element Method with Applications to Partial Differential Equations (A. K. Aziz, ed.), Academic Press, New York, 1972, pp. 1–362.

2. R. E. Bank and T. Dupont, *An optimal order process for solving finite element equations*, Math. Comp. **36** (1981), 35–51.

3. D. Braess and W. Hackbusch, *A new convergence proof for the multigrid method including the V-cycle*, SIAM J. Numer. Anal. **20** (1983), 967–975.

4. J. H. Bramble, Z. Leyk, and J. E. Pasciak, *The analysis of multigrid algorithms for pseudo-differential operators of order minus one*, preprint.

5. ____, *Iterative schemes for non-symmetric and indefinite elliptic boundary value problems*, BNL Rep. 45870.

6. J. H. Bramble and J. E. Pasciak, *The analysis of smoothers for multigrid algorithms*, Math. Comp. **58** (1992), 467–488.

7. ____, *New convergence estimates for multigrid algorithms*, Math. Comp. **49** (1987), 311–329.

8. J. H. Bramble, J. E. Pasciak, J. Wang, and J. Xu, *Convergence estimates for product iterative methods with applications to domain decomposition*, Math. Comp. **57** (1991), 1–21.

9. ____, *Convergence estimates for multigrid algorithms without regularity assumptions*, Math. Comp. **57** (1991), 23–45.

10. J. H. Bramble, J. E. Pasciak and J. Xu, *A multilevel preconditioner for domain decomposition boundary systems*, Proceedings of the 10th Internat. Conf. on Comput. Methods. in Appl. Sci. and Engr., Nova Sciences, New York, 1992.

11. ____, *The analysis of multigrid algorithms with nonnested spaces or noninherited quadratic forms*, Math. Comp. **56** (1991), 1–34.

12. ____, *Parallel multilevel preconditioners*, Math. Comp. **55** (1990), 1–22.

13. A. Brandt, *Multi-level adaptive solutions to boundary-value problems*, Math. Comp. **31** (1977), 333–390.

14. P. G. Ciarlet, *The finite element method for elliptic problems*, North-Holland, New York, 1978.

15. M. Dauge, *Elliptic boundary value problems on corner domains*, Lecture Notes in Math., vol. 1341, Springer-Verlag, Berlin and New York,, 1988.

16. N. H. Decker, S. V. Parter, and J. Mandel, *On the role of regularity in multigrid methods*, Multigrid Methods, Proceedings of the Third Copper Mountain Conference (S. McCormick, ed.), Marcel Dekker, New York, 1988, pp. 143–156.

17. P. Grisvard, *Elliptic problems in non smooth domains*, Pitman, Boston, 1985.

18. W. Hackbusch, *Multi-grid methods and applications*, Springer-Verlag, New York, 1985.

19. J. L. Lions and E. Magenes, *Problèmes aux limites non homogènes et applications,* vol. 1, Dunod, Paris, 1968.

20. J.-F. Maitre and F. Musy, *Algebraic formalisation 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. Holstien, eds), Clarendon Press, Oxford, 1985, pp. 213–223.

21. J. Mandel, S. McCormick, and R. Bank, *Variational multigrid theory*, Multigrid Methods, (S. McCormick, ed.), SIAM, Philadelphia, PA, 1987, pp. 131–178.

22. P. Oswald, *On discrete norm estimates related to multilevel preconditioners in the finite element method*, preprint.

23. X. Zhang, *Multi-level additive Schwarz methods*, Courant Inst. Math. Sci., Dept. Comp. Sci. Rep. (August, 1991).

DEPARTMENT OF MATHEMATICS, CORNELL UNIVERSITY, ITHACA, NEW YORK 14853
E-mail address: bramble@mssun7.msi.cornell.edu

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