

## THE ANALYSIS OF MULTIGRID ALGORITHMS WITH NONNESTED SPACES OR NONINHERITED QUADRATIC FORMS

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

**ABSTRACT.** We provide a theory for the analysis of multigrid algorithms for symmetric positive definite problems with nonnested spaces and noninherited quadratic forms. By this we mean that the form on the coarser grids need not be related to that on the finest, i.e., we do not stay within the standard variational setting. In this more general setting, we give new estimates corresponding to the  $\mathcal{V}$  cycle,  $\mathcal{W}$  cycle and a  $\mathcal{V}$  cycle algorithm with a variable number of smoothings on each level. In addition, our algorithms involve the use of nonsymmetric smoothers in a novel way.

We apply this theory to various numerical approximations of second-order elliptic boundary value problems. In our first example, we consider certain finite difference multigrid algorithms. In the second example, we consider a finite element multigrid algorithm with nested spaces, which however uses a prolongation operator that does not coincide with the natural subspace imbedding. The third example gives a multigrid algorithm derived from a loosely coupled sequence of approximation grids. Such a loosely coupled grid structure results from the most natural standard finite element application on a domain with curved boundary. The fourth example develops and analyzes a multigrid algorithm for a mixed finite element method using the so-called Raviart-Thomas elements.

### 1. INTRODUCTION

In recent years, multigrid methods have been used extensively as tools for obtaining approximations to the solutions of partial differential equations (see the references in [8, 16, 21]). In conjunction, there has been intensive research into the theoretical understanding of these methods (cf. [2, 3, 4, 5, 6, 16, 20, 21, 22, 30] and others). In this paper, we shall extend the theory for symmetric problems so that it applies in a more general framework.

The analysis of this paper can be broken down into two distinct parts. In the first part (§§2, 3, and 4), we provide a general theoretical framework for

---

Received August 29, 1988; revised June 29, 1989 and October 9, 1989.

1980 *Mathematics Subject Classification* (1985 Revision). Primary 65N30; Secondary 65F10.

This manuscript has been authored under contract number DE-AC02-76CH00016 with the U.S. Department of Energy. Accordingly, the U.S. Government retains a nonexclusive, royalty-free license to publish or reproduce the published form of this contribution, or allow others to do so, for U.S. Government purposes. This work was also supported in part under the National Science Foundation Grant No. DMS84-05352 and under the Air Force Office of Scientific Research, Contract No. ISSA86-0026 and by the U.S. Army Research Office through the Mathematical Science Institute, Cornell University.

the analysis of multigrid algorithms for symmetric problems with nonnested subspaces. Our algorithms allow the use of nonsymmetric smoothers in a novel way. In the second part (§§5–8), we use the general theory to provide iterative convergence estimates for a number of applications. The results obtained for these examples are new.

One of the most powerful tools for the development of an iterative convergence analysis for multigrid algorithms involves the use of the variational multigrid framework (see, e.g., [3, 4, 5, 21]). This is motivated from the study of the finite element multigrid technique, where the coarser multigrid spaces are nested and the discrete operators on the subspaces are given in terms of a form defined on a larger space (which contains all of the subspaces). As is well known, this analysis generalizes in a straightforward manner to the case of nonnested spaces under the constraint that the form on the coarser grid is equal to the form on the finer applied to the interpolated image (see (2.5), [13, 21, 22]).

This paper provides an analysis which allows the above-mentioned constraint to be violated. In §3, we consider the case when the equality constraint is replaced by a corresponding inequality (see (A.2)). With this weakened assumption, a “regularity and approximation” assumption (see (A.3)), and an appropriate smoother, we show that all of the results in [5] hold. This means that for the  $\mathcal{V}$  cycle, the variable  $\mathcal{V}$  cycle and the  $\mathcal{W}$  cycle algorithm, with any amount of smoothing,  $I - B_j A_j$  is a reducer. Here  $B_j$  is the corresponding multigrid operator (symmetric and positive definite) and  $A_j$  is the operator which we are trying to invert.

In §4, we consider the case when the inequality constraint (A.2) no longer holds. In this case,  $I - B_j A_j$  may no longer be a reducer. However, for the  $\mathcal{V}$  and variable  $\mathcal{V}$  cycle algorithms, the operator  $B_j$  is still symmetric and positive definite and hence can be used as a preconditioner. Section 4 provides bounds on the spectrum of  $B_j A_j$ . We prove that for the variable  $\mathcal{V}$  cycle algorithm with the additional regularity and approximation assumption, the system  $B_j A_j$  is uniformly well-conditioned (independent of the number of multigrid levels). Thus, we can construct rapidly converging iterative schemes for computing the action of  $A_j^{-1}$  using  $B_j$  (corresponding to the variable  $\mathcal{V}$  multigrid cycle) as a preconditioner. We next provide a result for the  $\mathcal{W}$  cycle algorithm without assuming (A.2). In this case,  $I - B_j A_j$  is still a reducer (uniformly, independent of  $j$ ) for the  $\mathcal{W}$  cycle provided that  $m$  (the number of smoothing steps) is chosen sufficiently large. We finally provide a result for the  $\mathcal{V}$  cycle algorithm which is valid if (A.2) holds, up to a perturbation.

Earlier papers provided a technique for proving  $\mathcal{W}$  cycle results without the variational framework under the assumption that the number of smoothings was sufficiently large [3, 16]. This approach was used extensively in [16] and, for example, [9, 32].

It is interesting to note that if (A.2) is not satisfied, the operator  $B_j$  corresponding to the  $\mathcal{W}$  cycle algorithm with a fixed number of smoothings may

be indefinite and hence of little use in an iterative algorithm for computing the action of  $A_j^{-1}$ . This is illustrated computationally in an example in §9. This indicates that the variable  $\mathcal{V}$  cycle algorithm is more robust than the  $\mathcal{W}$  cycle algorithm.

We note that many of the results given in this paper (and also [2, 4, 5, 6, 20] in the variational case) provide multigrid analysis for any number of smoothings on the finest grid. Such results are important to the code developer in that they guarantee that algorithms will work with just one smoothing. Accordingly, it is not necessary to experiment with various amounts of smoothing and one need not be concerned that the number of smoothing iterations may become so large as to make the algorithm no longer practical.

In the second part of the paper, we apply the earlier developed theory to a number of examples. The major part of the analysis necessary for the application of our theory involves the proof of the so-called “regularity and approximation” property. Its proof generally uses the elliptic regularity of the underlying problem as well as the approximation properties of the numerical method.

Sections 5 through 8 consider four different applications of the general theory. In §5, we consider a finite difference example with a lower-order term discretized by the “lumped mass” approximation. Section 6 considers a finite element example with alternative prolongation operators. Section 7 studies a finite element example where the multigrid algorithm was derived with loosely coupled grids. This example can be used to develop and analyze multigrid algorithms for problems with curved boundaries. Section 8 considers a multigrid algorithm for a mixed finite element approximation using the “Raviart-Thomas” elements. In all of these applications, the equality constraint mentioned above does not hold and hence the usual “variational” theory does not apply.

Unless otherwise stated,  $c$ ,  $C$ , and  $M$ , with or without subscript will denote generic positive constants which may take on different values in different places. These constants will always be independent of mesh parameters and the number of levels in the multigrid algorithms.

## 2. MULTIGRID ALGORITHMS

In this section, we describe the symmetric multigrid algorithm in the general setting. We also derive a basic recursion relation which plays a major role in the analysis given in §3.

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

$$\mathcal{M}_0, \mathcal{M}_1, \dots, \mathcal{M}_j$$

along with linear operators  $I_k: \mathcal{M}_{k-1} \mapsto \mathcal{M}_k$  for  $k = 1, \dots, j$ . The operators  $\{I_k\}$  will sometimes be called “prolongation” operators. In addition, we assume that we are given symmetric positive definite quadratic forms  $A_k(\cdot, \cdot)$  and  $(\cdot, \cdot)_k$  defined on  $\mathcal{M}_k \times \mathcal{M}_k$  for  $k = 0, \dots, j$ . The norm corresponding to  $(\cdot, \cdot)_k$  will be denoted by  $\|\cdot\|_k$ . Examples of families of spaces, operators, and forms will be given in later sections.

We shall develop multigrid algorithms for the solution of the problem: Given  $f \in \mathcal{M}_j$ , find  $v \in \mathcal{M}_j$  satisfying

$$(2.1) \quad A_j(v, \phi) = (f, \phi)_j \quad \text{for all } \phi \in \mathcal{M}_j.$$

To define these algorithms, we first define auxiliary operators. For  $k = 0, \dots, j$ , define the operator  $A_k: \mathcal{M}_k \mapsto \mathcal{M}_k$  by

$$(A_k w, \phi)_k = A_k(w, \phi) \quad \text{for all } \phi \in \mathcal{M}_k.$$

The operator  $A_k$  is clearly symmetric (in both the  $A_k(\cdot, \cdot)$  and  $(\cdot, \cdot)_k$  inner products) and positive definite. Also define the operators  $P_{k-1}: \mathcal{M}_k \mapsto \mathcal{M}_{k-1}$  and  $P_{k-1}^0: \mathcal{M}_k \mapsto \mathcal{M}_{k-1}$  by

$$A_{k-1}(P_{k-1} w, \phi) = A_k(w, I_k \phi) \quad \text{for all } \phi \in \mathcal{M}_{k-1},$$

and

$$(P_{k-1}^0 w, \phi)_{k-1} = (w, I_k \phi)_k \quad \text{for all } \phi \in \mathcal{M}_{k-1}.$$

It is easy to see that  $I_k P_{k-1}$  is a symmetric operator with respect to the  $A_k$  form. Note that, in general, neither  $P_k^0$  nor  $P_k$  is a projection.

To define the smoothing process, we require linear operators  $R_k: \mathcal{M}_k \mapsto \mathcal{M}_k$  for  $k = 1, \dots, j$ . This operator may be symmetric or nonsymmetric with respect to the inner product  $(\cdot, \cdot)_k$ . If  $R_k$  is nonsymmetric, then we define  $R_k^l$  to be its adjoint and set

$$R_k^{(l)} = \begin{cases} R_k & \text{if } l \text{ is odd,} \\ R_k^l & \text{if } l \text{ is even.} \end{cases}$$

The multigrid operator  $B_k: \mathcal{M}_k \mapsto \mathcal{M}_k$  is defined by induction and is given as follows.

### Multigrid Algorithm

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

- (1) Set  $x^0 = 0$  and  $q^0 = 0$ .
- (2) Define  $x^l$  for  $l = 1, \dots, m(k)$  by

$$(2.2) \quad x^l = x^{l-1} + R_k^{(l+m(k))}(g - A_k x^{l-1}).$$

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

$$(2.3) \quad q^i = q^{i-1} + B_{k-1}[P_{k-1}^0(g - A_k x^{m(k)}) - A_{k-1} q^{i-1}].$$

- (4) Define  $y^l$  for  $l = m(k) + 1, \dots, 2m(k)$  by

$$y^l = y^{l-1} + R_k^{(l+m(k))}(g - A_k y^{l-1}).$$

- (5) Set  $B_k g = y^{2m(k)}$ .

In this algorithm,  $m(k)$  is a positive integer which may vary from level to level and determines the number of smoothing iterations on that level. Because of this variable smoothing, the above algorithm is more general than that usually described [2, 3, 8, 16]. If  $R_k$  is symmetric and all of the  $m(k)$  are the same, then the algorithm is the usual symmetric multigrid cycling algorithm described in a notation which is convenient for our analysis. Note that  $B_k$  is clearly a linear operator for each  $k$ . In this algorithm,  $p$  is a positive integer. We shall study the cases  $p = 1$  and  $p = 2$ , which correspond respectively to the symmetric  $\mathcal{V}$  and  $\mathcal{W}$  cycles of multigrid.

In the above algorithm, we alternate between  $R_k$  and  $R_k^t$  in Step 2. In Step 4, we use the adjoints of the Step 2 smoothings applied in the reverse order. This results in a symmetric operator  $B_j$ . As far as we know, this form of the multigrid algorithm has not previously been suggested. The exact form of the above algorithm is motivated by the theory presented in later sections. Nonsymmetric smoothers were previously considered in [22, 23], but the theory there assumed a "variational" multigrid setup and full elliptic regularity.

Set  $K_k = I - R_k A_k$ ; then  $K_k^* = I - R_k^t A_k$  is the adjoint with respect to  $A_k(\cdot, \cdot)$ . We now make the following basic assumption:

(A.1) The spectrum of  $K_k^* K_k$  is in the interval  $[0, 1)$ .

*Remark 2.1.* We note that the Richardson iteration is an example of a symmetric  $R_k$  satisfying (A.1). In addition, one sweep of the Gauss-Seidel iteration with any ordering is an example of a nonsymmetric iteration satisfying (A.1).

Set

$$\tilde{K}_k^{(m)} = \begin{cases} (K_k^* K_k)^{m/2} & \text{if } m \text{ is even,} \\ (K_k^* K_k)^{(m-1)/2} K_k^* & \text{if } m \text{ is odd.} \end{cases}$$

Let  $I$  denote the identity operator. It is straightforward to check that

$$(2.4) \quad I - B_k A_k = (\tilde{K}_k^{(m(k))})^* [(I - I_k P_{k-1}) + I_k (I - B_{k-1} A_{k-1})^p P_{k-1}] \tilde{K}_k^{(m(k))},$$

cf. (2.7) of [5]. In (2.4),  $*$  denotes the adjoint with respect to the inner product  $A_k(\cdot, \cdot)$ .

Equation (2.4) gives a fundamental recurrence relation for the multigrid operator  $B_k$ . A straightforward argument using (2.4) and mathematical induction implies that  $I - B_k A_k$  is a symmetric operator on  $\mathcal{M}_k$  (even when  $R_k$  is nonsymmetric) with respect to the  $A_k$  form. This immediately implies that  $B_k$  is symmetric with respect to the  $(\cdot, \cdot)_k$  inner product.

In the above framework, the multigrid spaces need not be related to each other. Note that in the so-called "variational" case studied in [2, 4, 5, 20], it is assumed that

$$(2.5) \quad A_k(I_k u, I_k v) = A_{k-1}(u, v) \quad \text{for all } u, v \in \mathcal{M}_{k-1}.$$

Hence, the forms on all of the coarser grids are defined in terms of, or inherited

from, the form on the finest. The purpose of this paper is to analyze more general multigrid algorithms not satisfying assumption (2.5).

### 3. GENERAL MULTIGRID THEORY ASSUMING (A.2)

We provide a general multigrid theory in this and the following section. In this section, we consider the case when (2.5) is replaced by the assumption that for  $k = 1, \dots, j$ ,

$$(A.2) \quad A_k(I_k u, I_k u) \leq A_{k-1}(u, u) \quad \text{for all } u \in \mathcal{M}_{k-1}.$$

The reason for such an assumption will become clear as the analysis develops. As illustrated in Theorem 1, this assumption along with (A.1) is sufficient to guarantee that  $I - B_j A_j$  is a reducer and that the linear multigrid algorithm converges. In §4, we consider the case when (A.2) fails to hold.

*Remark 3.1.* Note that by definition,  $P_{k-1}$  is the adjoint of  $I_k$  and hence (A.2) holds if and only if

$$(3.1) \quad A_{k-1}(P_{k-1} u, P_{k-1} u) \leq A_k(u, u) \quad \text{for all } u \in \mathcal{M}_k.$$

Inequality (3.1) is also equivalent to the nonnegativity of the operator  $I - I_k P_{k-1}$  on  $\mathcal{M}_k$ . A straightforward argument using (2.4) and mathematical induction implies that  $I - B_k A_k$  is also nonnegative. If (A.2) does not hold, it is unlikely that  $I - B_k A_k$  is nonnegative.

The goal of this section is to prove that  $I - B_k A_k$  is a reducer and to estimate its rate of reduction under the Assumption (A.2). It suffices to show that the inequality

$$(3.2) \quad |A_k((I - B_k A_k)u, u)| \leq \delta_k A_k(u, u) \quad \text{for all } u \in \mathcal{M}_k$$

holds for a constant  $\delta_k < 1$  and estimate the dependence of  $\delta_k$  on  $k$  and additional assumptions. The above inequality implies that  $I - B_k A_k$  is a contraction with contraction number  $\delta_k$ . Moreover, if (A.2) holds, then  $I - B_k A_k$  is nonnegative and (3.2) is the same as

$$(3.3) \quad A_k((I - B_k A_k)u, u) \leq \delta_k A_k(u, u) \quad \text{for all } u \in \mathcal{M}_k.$$

The first theorem guarantees convergence of the multigrid algorithm under minimal assumptions.

**Theorem 1.** *Assume that (A.1) and (A.2) hold. Then (3.2) holds for some  $\delta_k < 1$ .*

The proof of the above theorem will be given later in this section. We note that the hypotheses for the theorem are rather weak. The spaces  $\mathcal{M}_k$  need not be related except for the existence of the linear maps  $I_k$ . Moreover, the maps  $I_k$  need not be injective and the assumptions on  $R_k$  are minimal. Thus, the above theorem can be thought of as a result for “algebraic multigrid” since it requires none of the stronger “regularity and approximation” assumptions used in the convergence analysis for the partial differential equation approximation applications.

Moreover, the theorem can still be used to develop multigrid algorithms, even when the forms  $A_k$  on the spaces do not a priori satisfy (A.2). Note that (A.2) can be satisfied by simply scaling the forms, i.e.,  $A_k(\cdot, \cdot) \leftarrow \alpha_k A_k(\cdot, \cdot)$ . Clearly, there is no difficulty in applying the multigrid algorithm with the scaled forms. Theorem 1 then implies stability and convergence; however, this convergence may be unacceptably slow without further conditions being satisfied. Multigrid results without this scaling of forms (i.e., when (A.2) fails to hold) are given in §4.

For stronger convergence estimates, we shall make additional a priori assumptions. Let  $0 < \alpha \leq 1$ . The first assumption is a “regularity and approximation” assumption of the form

$$(A.3) \quad |A_k((I - I_k P_{k-1})u, u)| \leq C_\alpha^2 \left( \frac{\|A_k u\|_k^2}{\lambda_k} \right)^\alpha A_k(u, u)^{1-\alpha} \quad \text{for all } u \in \mathcal{M}_k,$$

where  $\lambda_k$  is the largest eigenvalue of  $A_k$ . More precisely, we assume that (A.3) holds with  $C_\alpha$  independent of  $k$  for  $k = 1, \dots, j$ .

Let  $R_{k,\omega}$  correspond to the Richardson smoothing iteration defined by  $R_{k,\omega} = \omega \lambda_k^{-1} I$  and  $K_{k,\omega} = (I - R_{k,\omega} A_k)$  be the corresponding reducer. We assume that there exists an  $\omega$  in  $(0, 2)$  not depending on  $k$  such that

$$(A.4) \quad A_k(K_k u, K_k u) \leq A_k(K_{k,\omega} u, K_{k,\omega} u) \quad \text{for all } u \in \mathcal{M}_k,$$

i.e., the smoothing process converges as fast as Richardson’s method for some  $\omega$ .

The above assumption was used by McCormick in [22, 23]. It is not difficult to show that (A.4) is equivalent to the existence of a positive constant  $C_R$  not depending on  $k$  and satisfying

$$(3.4) \quad \frac{\|u\|_k^2}{\lambda_k} \leq C_R (\tilde{R}_k u, u)_k \quad \text{for all } u \in \mathcal{M}_k,$$

where  $\tilde{R}_k = (I - K_k^* K_k) A_k^{-1}$ . The inequality (3.4) is convenient for our analysis. We also note that (A.1) immediately follows from (3.4) and hence (A.4) implies (A.1).

*Remark 3.2.* The Richardson method is a symmetric iteration satisfying (A.4) for  $\omega \in (0, 2)$ . In finite element and finite difference applications, under reasonable assumptions, one sweep of Gauss-Seidel iteration with any ordering gives rise to a nonsymmetric  $R_k$  which also satisfies (A.4).

*Remark 3.3.* For the theory presented in [5], it was assumed that  $K_k$  was non-negative and that

$$(3.5) \quad \frac{\|u\|_k^2}{\lambda_k} \leq C_R (R_k u, u)_k \quad \text{for all } u \in \mathcal{M}_k$$

was satisfied with the (assumed symmetric) operator  $R_k$ . It is easy to see that this implies (3.4) with the same constant  $C_R$ .

We can now state and prove the theorem for estimating  $\delta_k$  in (3.2) for the symmetric  $\mathcal{V}$  cycle.

**Theorem 2.** *Assume that (A.2–A.4) hold and define  $B_j$  with  $p = 1$  and  $m(k) = m$  for all  $k$ . Then (3.2) holds with*

$$(3.6) \quad \delta_k = \frac{Mk^{\frac{1-\alpha}{\alpha}}}{Mk^{\frac{1-\alpha}{\alpha}} + m^\alpha}.$$

*Remark 3.4.* The convergence estimates in Theorem 2 and those to be stated later in this section have exactly the same form as the theorems in [5]. The interested reader is referred to [5] for various explicit expressions for  $M$  in terms of the constants  $\alpha$ ,  $C_\alpha$  and  $C_R$ .

*Proof of Theorem 2.* We shall prove (3.3) by induction on  $k$ . For  $k = 0$ , there is nothing to prove. Assume that (3.3) holds for  $k - 1$ . By (2.4),

$$(3.7) \quad \begin{aligned} A_k((I - B_k A_k)u, u) &= A_k((I - I_k P_{k-1})\tilde{u}, \tilde{u}) \\ &\quad + A_{k-1}((I - B_{k-1} A_{k-1})P_{k-1}\tilde{u}, P_{k-1}\tilde{u}), \end{aligned}$$

where  $\tilde{u} = \tilde{K}_k^{(m)}u$ . Applying the induction hypothesis and the definition of  $P_{k-1}$  gives

$$(3.8) \quad \begin{aligned} A_k((I - B_k A_k)u, u) &\leq A_k((I - I_k P_{k-1})\tilde{u}, \tilde{u}) + \delta_{k-1} A_k(I_k P_{k-1}\tilde{u}, \tilde{u}) \\ &= (1 - \delta_{k-1}) A_k((I - I_k P_{k-1})\tilde{u}, \tilde{u}) + \delta_{k-1} A_k(\tilde{u}, \tilde{u}). \end{aligned}$$

Applying (A.3) and a generalized arithmetic-geometric mean inequality gives

$$(3.9) \quad A_k((I - I_k P_{k-1})\tilde{u}, \tilde{u}) \leq C_\alpha^2 \left\{ \alpha \gamma_k \frac{\|A_k \tilde{u}\|_k^2}{\lambda_k} + (1 - \alpha) \gamma_k^{-\alpha/(1-\alpha)} A_k(\tilde{u}, \tilde{u}) \right\}.$$

Applying (3.4) gives

$$(3.10) \quad \frac{\|A_k \tilde{u}\|_k^2}{\lambda_k} \leq C_R A_k((I - \bar{K}_k) \bar{K}_k^m u, u),$$

where

$$(3.11) \quad \bar{K}_k = \begin{cases} K_k^* K_k & \text{if } m \text{ is even,} \\ K_k K_k^* & \text{if } m \text{ is odd.} \end{cases}$$

The remainder of the proof of the theorem is exactly the same as the proof of Theorem 1 in [5].  $\square$

From the proof of Theorem 2, it is apparent that the framework for nonnested spaces and noninherited forms developed in §2 fits into the machinery of [5]. The next two theorems follow in a similar manner. The first gives a result for the  $\mathcal{W}$  cycle algorithm, while the second gives a result for the variable  $\mathcal{V}$  cycle algorithm.

**Theorem 3.** Assume that (A.2–A.4) hold and define  $B_j$  with  $p = 2$  and  $m(k) = m$  for all  $k$ . Then (3.2) holds with  $\delta_k = \delta$  (independent of  $k$ ) given by

$$(3.12) \quad \delta = \frac{M}{M + m^\alpha}.$$

**Theorem 4.** Assume that (A.2–A.4) hold and define  $B_j$  with  $p = 1$ . Assume that  $m(k)$  satisfies

$$(3.13) \quad \beta_0 m(k) \leq m(k-1) \leq \beta_1 m(k).$$

Here, we assume that  $\beta_0$  and  $\beta_1$  are constants which are greater than one and independent of  $k$ . Then (3.2) holds with

$$(3.14) \quad \delta_k = \frac{M}{M + m(k)^\alpha}.$$

*Remark 3.5.* We have only provided results for the “symmetric” multigrid cycling schemes, i.e., those in which one smooths both before and after coarse grid correction. The above analysis seems to fail for the nonsymmetric multigrid schemes (described in, for example, [5, 13, 22]) due to the fact that  $I_k P_{k-1}$  is no longer a projection and the product of the so-called slash cycles [22] is no longer the symmetric  $\mathcal{V}$  cycle.

*Proof of Theorem 1.* We now prove Theorem 1. Note that since  $R_k$  is positive definite and all spaces are finite-dimensional, (3.4) holds for some constant  $C_R(j)$  which may depend on  $\{R_k\}$ ,  $k = 1, \dots, j$ . Similarly, the definiteness of  $A_k$  implies that (A.3) holds for some constant  $C_\alpha(j)$  which may depend on  $\{A_k\}$ ,  $\{I_k\}$ , and  $\{(\cdot, \cdot)_k\}$ . Theorems 2–4 still hold with some convergence parameter  $\delta_k < 1$  depending on  $C_R(j)$  and  $C_\alpha(j)$  since, in this case, the constant  $M$  is not independent of  $k$ . This proves Theorem 1.  $\square$

#### 4. GENERAL MULTIGRID THEORY WITHOUT (A.2)

In this section, we provide an analysis for the multigrid algorithm which allows (A.2) to be violated. In this case,  $I - B_k A_k$  may no longer be a reducer. Nevertheless, the operator  $B_j$  corresponding to the variable  $\mathcal{V}$  and the  $\mathcal{V}$  cycle multigrid algorithms is positive definite and hence can be used as a preconditioner in an iterative method for solving (2.1). The  $\mathcal{W}$  cycle may, however, be indefinite without increasing the number of smoothings. We first give a theorem with minimal hypotheses, which guarantees that the operator  $B_k$  corresponding to the  $\mathcal{V}$  or variable  $\mathcal{V}$  cycle algorithm is symmetric and positive definite. We next consider additional hypotheses which are sufficient to guarantee iterative convergence rates for variable  $\mathcal{V}$ ,  $\mathcal{V}$ , and  $\mathcal{W}$  cycle multigrid algorithms.

**Theorem 5.** Assume that (A.1) holds and  $p = 1$ . Then  $B_j$  is a symmetric positive definite operator on  $\mathcal{M}_j$ .

*Proof.* As already observed in §2,  $B_j$  is a symmetric operator with respect to the  $(\cdot, \cdot)_k$  inner product. By (2.4),

$$(4.1) \quad (B_k A_k u, A_k u)_k = A_k((I - \bar{K}_k^{m(k)})u, u) \\ + (B_{k-1} A_{k-1} P_{k-1} \tilde{K}_k^{(m(k))} u, A_{k-1} P_{k-1} \tilde{K}_k^{(m(k))} u)_{k-1}$$

for all  $u \in \mathcal{M}_k$ . Since the eigenvalues of  $\bar{K}_k$  are in  $[0, 1)$ , it follows that  $I - \bar{K}_k^{m(k)}$  is a positive definite operator. Thus, by (4.1) and induction,  $B_k$  is positive definite. This completes the proof of Theorem 5.  $\square$

*Remark 4.1.* In general, the theorem does not hold for the  $\mathcal{W}$  cycle multigrid algorithm. We give a computational example in §9 where (A.2) is violated and the  $B_j$  corresponding to the  $\mathcal{W}$  cycle multigrid algorithm with  $m = 1$  has negative eigenvalues. Thus, this  $\mathcal{W}$  cycle algorithm cannot be used in a preconditioning strategy or to develop a reducer. Computational results given in §9 for the same problem indicate that the corresponding variable  $\mathcal{V}$  and  $\mathcal{V}$  cycle algorithms give rise to effective preconditioners and hence lead to rapidly converging iterative schemes.

The above theorem can be thought of as a result for algebraic multigrid since the hypotheses are minimal. The theorem guarantees that corresponding preconditioned iterative algorithms (for example, preconditioned conjugate gradient) for the solution of (2.1) will be stable and convergent. The convergence rates of these algorithms may be unacceptably slow without further conditions being satisfied.

In the remainder of this section, we shall make additional assumptions which will lead to theorems which guarantee that the iterative algorithms converge at more reasonable rates. For the  $\mathcal{V}$  cycle algorithms, this involves deriving bounds on the largest and smallest eigenvalues of the operator  $B_j A_j$ . Equivalently, we shall provide positive constants  $\eta_0$  and  $\eta_1$  which may depend on  $k$  and satisfy the inequalities

$$(4.2) \quad \eta_0 A_k(u, u) \leq A_k(B_k A_k u, u) \leq \eta_1 A_k(u, u) \quad \text{for all } u \in \mathcal{M}_k.$$

Note that if (4.2) holds, then the preconditioned conjugate gradient method converges with an asymptotic reduction rate of

$$\frac{1 - \sqrt{\eta_0/\eta_1}}{1 + \sqrt{\eta_0/\eta_1}}$$

per iterative step.

The next theorem provides estimates for  $\eta_0$  and  $\eta_1$  for the variable  $\mathcal{V}$  cycle algorithm.

**Theorem 6.** *Assume that (A.3) and (A.4) hold and define  $B_j$  with  $p = 1$ . Assume that  $\{m(k)\}$  satisfy (3.13). Then the constants  $\eta_0$  and  $\eta_1$  in (4.2) satisfy*

$$\eta_0 \geq \frac{m(k)^\alpha}{M + m(k)^\alpha} \quad \text{and} \quad \eta_1 \leq \frac{M + m(k)^\alpha}{m(k)^\alpha},$$

*i.e., the system  $B_j A_j$  is well-conditioned independently of  $j$ .*

We shall use the following lemma in the proof of Theorem 6.

**Lemma 4.1.** *Assume that  $p = 1$  and that  $\bar{\delta}_i$  for  $i = 1, 2, \dots, k$  satisfies the inequality*

$$(4.3) \quad -A_i((I - I_i P_{i-1})\tilde{u}, \tilde{u}) \leq \bar{\delta}_i A_i(u, u) \quad \text{for all } u \in \mathcal{M}_i$$

where  $\tilde{u} = \tilde{K}_i^{(m(i))} u$ . Then

$$\eta_1 \leq \prod_{i=1}^k (1 + \bar{\delta}_i).$$

*Proof.* It suffices to show that

$$(4.4) \quad -A_k((I - B_k A_k)u, u) \leq (\tau_k - 1)A_k(u, u) \quad \text{for all } u \in \mathcal{M}_k$$

where  $\tau_0 = 1$  and  $\tau_k = \prod_{i=1}^k (1 + \bar{\delta}_i)$ . We prove (4.4) by induction. For  $k = 0$ , there is nothing to prove. Assume (4.4) holds for  $k - 1$ . Then by (2.4), the induction assumption and (4.3),

$$\begin{aligned} & -A_k((I - B_k A_k)u, u) \\ &= -A_k((I - I_k P_{k-1})\tilde{u}, \tilde{u}) - A_{k-1}((I - B_{k-1} A_{k-1})P_{k-1}\tilde{u}, P_{k-1}\tilde{u}) \\ &\leq -A_k((I - I_k P_{k-1})\tilde{u}, \tilde{u}) + (\tau_{k-1} - 1)A_{k-1}(P_{k-1}\tilde{u}, P_{k-1}\tilde{u}) \\ &\leq [\bar{\delta}_k + (\tau_{k-1} - 1)(1 + \bar{\delta}_k)]A_k(u, u) = (\tau_k - 1)A_k(u, u). \end{aligned}$$

This completes the proof of the lemma.  $\square$

*Proof of Theorem 6.* We note that Assumption (A.2) was used in the proof of Theorem 4 only to reduce to the proof of inequality (3.3). The subsequent arguments showing (3.3) remain valid without (A.2) and lead to the inequality

$$A_k((I - B_k A_k)u, u) \leq \delta_k A_k(u, u) \quad \text{for all } u \in \mathcal{M}_k,$$

where  $\delta_k$  is given by (3.14). It immediately follows that (4.2) holds with  $\eta_0 = 1 - \delta_k$ .

To estimate  $\eta_1$ , we note that by (A.3),

$$(4.5) \quad -A_k((I - I_k P_{k-1})\tilde{u}, \tilde{u}) \leq C_\alpha^2 (\lambda_k^{-1} \|A_k \tilde{u}\|_k^2)^\alpha A_k(\tilde{u}, \tilde{u})^{1-\alpha}.$$

By (3.10) and [5, (3.16)],

$$(4.6) \quad \frac{\|A_k \tilde{u}\|_k^2}{\lambda_k} \leq \frac{C}{m(k)} A_k((I - \bar{K}_k^{m(k)})u, u).$$

Since the eigenvalues of  $\bar{K}_k$  are in the interval  $[0, 1)$ ,

$$(4.7) \quad -A_k((I - I_k P_{k-1})\tilde{u}, \tilde{u}) \leq C m(k)^{-\alpha} A_k(u, u).$$

Elementary arguments imply that

$$\prod_{k=1}^j \left(1 + \frac{C}{m(k)^\alpha}\right) \leq 1 + \frac{C}{m(j)^\alpha},$$

and hence the bound for  $\eta_1$  follows from Lemma 4.1. This completes the proof of Theorem 6.  $\square$

We next give a theorem for the  $\mathscr{W}$  cycle algorithm.

**Theorem 7.** *Assume that all the hypotheses except (A.2) hold for Theorem 3. Then, for the  $\mathscr{W}$  cycle algorithm with  $m$  sufficiently large, (3.2) holds with  $\delta_k = \delta$  where  $\delta$  is given by (3.12). Furthermore, the same conclusion holds if “ $m$  sufficiently large” is replaced by the assumption*

$$(A.5) \quad A_k(I_k u, I_k u) \leq 2A_{k-1}(u, u) \quad \text{for all } u \in \mathcal{M}_{k-1}.$$

*Proof.* We first consider the case without Assumption (A.5). We first show that

$$(4.8) \quad -A_k((I - B_k A_k)u, u) \leq \delta A_k(u, u) \quad \text{for all } u \in \mathcal{M}_k.$$

By (2.4), it clearly suffices to show that

$$(4.9) \quad -A_k((I - I_k P_{k-1})\tilde{u}, \tilde{u}) \leq \delta A_k(u, u)$$

where  $\tilde{u} = \tilde{K}_k^{(m)} u$ .

Inequality (4.9) immediately follows from (4.7) if  $m$  and  $M$  are chosen sufficiently large. With (4.8) verified, the proof of the opposite inequality,

$$A_k((I - B_k A_k)u, u) \leq \delta A_k(u, u)$$

follows in the same way as Theorem 3. This completes the proof of Theorem 7 without the assumption of (A.5).

If (A.5) holds, then

$$-A_k((I - I_k P_{k-1})\tilde{u}, \tilde{u}) \leq A_k(\tilde{u}, \tilde{u}).$$

Hence,

$$(4.10) \quad -A_k((I - I_k P_{k-1})\tilde{u}, \tilde{u}) \leq (1 - \delta^2)|A_k((I - I_k P_{k-1})\tilde{u}, \tilde{u})| + \delta^2 A_k(\tilde{u}, \tilde{u}).$$

We note that the first step in the analysis of the  $\mathscr{W}$  cycle algorithm in Theorem 3 (see [5, (3.32)]) is to show that the right-hand side of (4.10) bounds  $A_k((I - B_k A_k)u, u)$ . The remainder of the proof of Theorem 3 bounds the right-hand side of (4.10) by  $\delta A_k(u, u)$  (see the proof of Theorem 3 of [5]). This completes the proof of Theorem 7.  $\square$

*Remark 4.2.* It is elementary but tedious to see from the proof of Theorem 7 (and in particular, the proof of Theorem 3 of [5]) that the constant 2 in (A.5) can always be replaced by  $2 + \varepsilon$ , where the size of  $\varepsilon$  depends upon the size of  $C_R$  and  $C_\alpha$ . Larger  $C_R$  and  $C_\alpha$  require smaller  $\varepsilon$ .

The last theorem of this section provides a result for the  $\mathscr{V}$  cycle algorithm. For this result, we assume that

$$(A.6) \quad A_k(I_k u, I_k u) \leq (1 + c\lambda_k^{-\gamma})A_{k-1}(u, u) \quad \text{for all } u \in \mathcal{M}_{k-1}$$

which holds for some  $\gamma$  in the interval  $(0, 1]$ . In many applications,  $\lambda_k$  grows like  $h_k^{-2}$ , where  $h_k$  is the mesh size. Thus, (A.6) is a perturbation of (A.2) up to some power of  $h_k$ . We have the following theorem for the  $\mathscr{V}$  cycle algorithm.

**Theorem 8.** Assume that (A.3), (A.4) and (A.6) hold, and define  $B_j$  with  $p = 1$  and  $m(k) = m$ . Assume further that the maximum eigenvalue  $\lambda_k \geq \kappa^k$  for some  $\kappa > 1$ . Then  $\eta_1$  in (4.2) can be chosen independently of  $k$  and  $\eta_0 \leq 1 - \delta_k$ , where  $\delta_k$  is given by (3.6).

*Proof.* Since  $I_k$  and  $P_{k-1}$  are adjoint operators, (A.6) implies that

$$-A_k((I - I_k P_{k-1})\tilde{u}, \tilde{u}) \leq c\lambda_k^{-\gamma} A_k(\tilde{u}, \tilde{u}) \leq c\lambda_k^{-\gamma} A_k(u, u).$$

Elementary manipulations imply

$$\prod_{k=1}^{\infty} (1 + c\lambda_k^{-\gamma})^{-1} < \infty.$$

Lemma 4.1 shows that  $\eta_1$  can be bounded independently of  $k$ . The bound for  $\eta_0$  follows from the proof of Theorem 2 (see also the proof of Theorem 6).  $\square$

*Remark 4.3.* We note that if (A.6) is satisfied, the results of Theorem 7 still hold if the the assumption “ $m$  is sufficiently large” is replaced by the assumption “the coarse grid is sufficiently fine”.

## 5. A FINITE DIFFERENCE APPLICATION

In this section, we consider a finite difference application approximating the solution of the problem

$$(5.1) \quad \begin{aligned} -\Delta u + u &= f & \text{in } \Omega, \\ u &= 0 & \text{on } \partial\Omega. \end{aligned}$$

Here, we do not assume that  $\Omega$  is a rectangle, and hence standard multigrid analysis for square or periodic domains does not apply. We define a multigrid algorithm in terms of the general approach of §2. In this as well as the remaining sections, we shall consider only simple model problems, even though the techniques obviously extend to more general applications.

We set up a sequence of nodes in the usual way. Without loss of generality, we assume that the domain  $\Omega \subset [0, 1]^2$  and  $h_k = 2^{-k}/M$  for some integer  $M > 1$ . Let  $N_k = 2^k M - 1$ . The nodes of the finite difference approximation on the  $k$ th level are given by

$$x_{ij}^k = (ih_k, jh_k) \quad \text{for } i, j = 1, \dots, N_k.$$

We assume that the boundary of the domain  $\Omega$  aligns with the mesh lines on the coarsest grid, i.e.,  $\Omega$  is the union of coarse grid rectangles. Let  $\Omega_k$  denote the nodes of the  $k$ th grid which are in the interior of  $\Omega$ . The space  $\mathcal{M}_k$  is defined to be the vector space of nodal values defined on  $\Omega_k$ . The prolongation operator  $I_k$  is defined as follows:

- (1) If  $x_{ij}^k$  is a node on the  $(k-1)$ st grid, then  $(I_k V)(x_{ij}^k) = V(x_{ij}^k)$ .

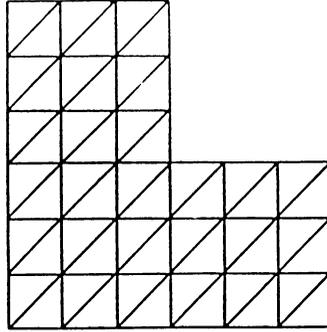


FIGURE 5.1  
*The triangular mesh*

- (2) If  $x_{ij}^k$  is a node on an edge in the  $x$  or  $y$  direction between two nodes in the  $(k-1)$ st grid, then  $(I_k V)(x_{ij}^k)$  is the average of the nodal values of  $V$  at the two nodes.
- (3) Otherwise,  $(I_k V)(x_{ij}^k)$  is the average of the nodal value of the node immediately above and to the right with that of the node immediately below and to the left.

The operator  $I_k$  corresponds to piecewise linear interpolation on the triangulation of size  $h_{k-1}$  in Figure 5.1. For a function  $V \in \mathcal{M}_k$ ,  $V_{ij}$  will denote the value of  $V$  at  $x_{ij}^k$ . The quadratic forms  $A_k(\cdot, \cdot)$  and  $(\cdot, \cdot)_k$  are defined by

$$(5.2) \quad A_k(U, V) = \sum_{x_{ij}^k \in \Omega_k} [h_k^2 U_{ij} + 4U_{i,j} - U_{i+1,j} - U_{i-1,j} - U_{i,j+1} - U_{i,j-1}] V_{ij}$$

and

$$(5.3) \quad (U, V)_k = h_k^2 \sum_{x_{ij}^k \in \Omega_k} U_{ij} V_{ij}.$$

In (5.2) and the remainder of this section, nodal values are set to zero when the node is not inside of  $\Omega$ . It is obvious that, with the above definitions, the solution of (2.1) corresponds to the standard finite difference approximation to the solution of (5.1). In fact, the problems on the coarser grids are also standard finite difference approximations. Because of the way the lower-order term of (5.1) is approximated, (2.5) does not hold and the ‘‘variational’’ theory does not apply.

To prove regularity and approximation for this and the remaining applications in this paper, we shall need to use various Sobolev spaces. For nonnegative integers  $m$ , the Sobolev space  $H^m(\Omega)$  is defined to be the set of functions in  $L^2(\Omega)$  whose distributional derivatives up to order  $m$  are in  $L^2(\Omega)$  (see, e.g., [19, 25]). For  $\sigma \in (0, 1)$ , the space  $H^\sigma(\Omega)$  is defined to be the set of functions

in  $L^2(\Omega)$  for which the norm

$$\|v\|_{H^\sigma(\Omega)}^2 \equiv \int_{\Omega} \int_{\Omega} \frac{(v(x) - v(y))^2}{|x - y|^{2+2\sigma}} dx dy$$

is finite. For  $s \in (m, m + 1)$  and  $m$  an integer greater than zero, we define  $H^s(\Omega)$  to be the set of functions in  $H^m(\Omega)$  for which the norm

$$\|v\|_{H^s(\Omega)}^2 \equiv \|v\|_{H^m(\Omega)}^2 + \sum_{|k|=m} \|D^k v\|_{H^{s-m}(\Omega)}^2$$

is finite. It can be shown that these spaces coincide with those defined by Hilbert scale interpolation between  $L^2(\Omega)$  and  $H^{m+1}(\Omega)$  (with equivalent norms) provided that there exists an extension operator which is simultaneously bounded from  $L^2(\Omega) \mapsto L^2(R^2)$  and  $H^{m+1}(\Omega) \mapsto H^{m+1}(R^2)$  (cf. [18, Theorem 9.3]). The existence of such operators is known in the case of domains with minimally smooth boundary [28] which includes the case of plane polygons.

For a nodal function  $V \in \mathcal{M}_k$ , let  $\tilde{V}$  denote the piecewise linear function on the triangulation depicted by Figure 5.1 which interpolates  $V$  at the nodes of  $\mathcal{M}_k$ . It can be shown that

$$(5.4) \quad |(\tilde{W}, \tilde{V}) - (W, V)_k| \leq ch_k \|\tilde{W}\|_{H^1(\Omega)} \|V\|_k \quad \text{for all } V, W \in \mathcal{M}_k.$$

Here,  $(\cdot, \cdot)$  denotes the  $L^2$  inner product on  $\Omega$  and  $\|\cdot\|_k = (\cdot, \cdot)_k^{1/2}$ .

Regularity results for problem (5.1) have been proven in [17]. The space  $H^{-1}(\Omega)$  is defined as the set of distributions on  $\Omega$  for which the norm

$$\|f\|_{H^{-1}(\Omega)} = \sup_{\phi \in H_0^1(\Omega)} \frac{(f, \phi)}{\|\phi\|_{H^1(\Omega)}}$$

is finite. For  $s \in (0, 1)$ , the spaces  $H^{-s}(\Omega)$  are defined by interpolation between  $L^2(\Omega)$  and  $H^{-1}(\Omega)$  with norms denoted by  $\|\cdot\|_{H^{-s}(\Omega)}$ . In [17], it is shown that the solution  $u$  of (5.1) satisfies inequalities of the form

$$(5.5) \quad \|u\|_{H^{1+\beta}(\Omega)} \leq c \|f\|_{H^{-1+\beta}(\Omega)},$$

where  $0 < \beta \leq 1$  is a constant which depends upon  $\partial\Omega$ . In particular, a result of [17] shows that (5.5) holds for some  $\beta > 1/2$  for any polygonal domain in  $R^2$  with interior angles less than  $2\pi$ .

For our purposes, we shall need an alternative representation of the negative Sobolev norms defined above. Consider the problem: Given  $g \in L^2(\Omega)$ , find  $w \in H_0^1(\Omega)$  satisfying

$$D(w, \chi) = (g, \chi) \quad \text{for all } \chi \in H_0^1(\Omega),$$

where  $D(\cdot, \cdot)$  denotes the Dirichlet form on  $\Omega$ . By Rellich's Lemma [25],  $H_0^1(\Omega)$  is compactly contained in  $L^2(\Omega)$ . It follows that the solution operator  $T: L^2(\Omega) \mapsto H_0^1(\Omega)$  defined by  $T(g) = w$  has a complete orthonormal basis

of eigenfunctions  $\{\psi_i\}$  [27]. The eigenvalues of  $T$  corresponding to these eigenfunctions will be denoted  $\{\sigma_i\}$ . We define the spaces

$$\dot{H}^s = \left\{ \phi = \sum c_i \psi_i \text{ such that } \sum \sigma_i^{-s} c_i^2 < \infty \right\}$$

with norm

$$(5.6) \quad \|\phi\|_s = \left( \sum \sigma_i^{-s} c_i^2 \right)^{1/2}.$$

It is not difficult to show that  $\dot{H}^1 = H_0^1(\Omega)$  with equivalent norms. By duality and interpolation, it immediately follows that  $\dot{H}^{-s} = H^{-s}(\Omega)$  for  $s \in [0, 1]$ . Thus inequality (5.5) can be rewritten

$$(5.7) \quad \|u\|_{H^{1+\beta}(\Omega)} \leq c \|f\|_{-1+\beta}.$$

We can now prove the following proposition.

**Proposition 5.1.** *Let  $\mathcal{M}_k$ ,  $A_k(\cdot, \cdot)$ ,  $(\cdot, \cdot)_k$ , and  $I_k$  be defined as above. Then (A.2) holds and (A.3) holds for  $\alpha = \beta/2$ .*

Combining the proposition with Theorems 2–4 gives results for the corresponding multigrid algorithms (with appropriately chosen  $R_k$ ). Note that we get uniform (independent of  $h_j$ ) convergence for the  $\mathcal{W}$  cycle and variable  $\mathcal{V}$  cycle algorithms. With the  $\mathcal{V}$  cycle, we may see some logarithmic ( $j \sim \ln h_j^{-1}$ ) deterioration even in the case of full elliptic regularity ( $\beta = 1$ ).

*Proof of Proposition 5.1.* We first prove (A.2). We write  $A_k(\cdot, \cdot) = (\cdot, \cdot)_k + D_k(\cdot, \cdot)$ . It is not difficult to see that

$$(5.8) \quad D_k(V, W) = D(\tilde{V}, \tilde{W}) \quad \text{for all } V, W \in \mathcal{M}_k.$$

This uses the assumption that  $\partial\Omega$  aligns with the mesh lines of the coarsest grid. Consequently, to prove (A.2), it suffices to show that

$$(5.9) \quad (I_k V, I_k V)_k \leq (V, V)_{k-1} \quad \text{for all } V \in \mathcal{M}_{k-1}.$$

Note that the form  $(\cdot, \cdot)_k$  can also be written

$$(W, W)_k = h_k^2/6 \sum_{\tau_i} \sum_{j=1}^3 W(y_{ij})^2,$$

where the first sum is taken over the triangles of the  $k$ th mesh and  $\{y_{ij}\}_{j=1}^3$  denotes the three vertices of the  $i$ th triangle. It clearly suffices to prove the inequality analogous to (5.9) on a typical triangle of the  $(k-1)$ st grid. Note that  $I_k V = \tilde{V}$  on the nodes of the  $k$ th mesh and  $\tilde{V}$  is piecewise linear on each triangle of the  $(k-1)$ st mesh. Consider a triangle of the  $(k-1)$ st grid and a linear function which takes on the values  $a$ ,  $b$ , and  $c$  at the nodes. Computing the local contributions to  $(I_k V, I_k V)_k$  and  $(V, V)_{k-1}$  corresponding to this triangle, we see that (5.9) follows from

$$a^2 + b^2 + c^2 + 3(a+b)^2/4 + 3(b+c)^2/4 + 3(a+c)^2/4 \leq 4(a^2 + b^2 + c^2).$$

This completes the proof of (A.2).

We first introduce some additional notation for the proof of (A.3). Let  $\tilde{A}$  be the form on  $H_0^1(\Omega) \times H_0^1(\Omega)$  defined by

$$\tilde{A}(w, v) = (w, v) + D(w, v).$$

In addition, let  $S_k \subset H_0^1(\Omega)$  denote the collection of piecewise linear functions on the triangulation of size  $h_k$  (see Figure 5.1). Let  $\bar{P}_k$  denote the elliptic projection onto  $S_k$  with respect to  $\tilde{A}(\cdot, \cdot)$ , i.e.,  $\bar{P}_k w$  for  $w \in H_0^1(\Omega)$  is the unique function in  $S_k$  satisfying

$$\tilde{A}(\bar{P}_k w, v) = \tilde{A}(w, v) \quad \text{for all } v \in S_k.$$

From the definitions and (5.8), we have for  $W \in \mathcal{M}_k$  and  $\Phi \in \mathcal{M}_{k-1}$ ,

$$\tilde{A}(\overline{P_{k-1}W - \bar{P}_{k-1}\tilde{W}}, \tilde{\Phi}) = (P_{k-1}W, \tilde{\Phi}) - (P_{k-1}W, \Phi)_{k-1} + (W, I_k\Phi)_k - (\tilde{W}, \tilde{\Phi}).$$

Here we used the identity  $\tilde{I}_k\tilde{\Phi} = \tilde{\Phi}$ . Thus, by (3.1) and (5.4),

$$(5.10) \quad \tilde{A}(P_{k-1}W - \bar{P}_{k-1}\tilde{W}, P_{k-1}W - \bar{P}_{k-1}\tilde{W}) \leq ch_k^2 A_k(W, W).$$

To complete the proof of (A.3), we shall need the following lemma. Its proof is the same as that given for the finite element case in [3].

**Lemma 5.1.** *Let  $0 \leq s \leq 1$ . There are constants  $c_0$  and  $c_1$  which are independent of  $j$  such that*

$$(5.11) \quad \begin{aligned} c_0 \|A_k^{s/2} V\|_k &\leq \|\tilde{V}\|_{H^s(\Omega)} \\ &\leq c_1 \|A_k^{s/2} V\|_k \quad \text{for all } V \in \mathcal{M}_k, \quad k = 0, \dots, j. \end{aligned}$$

Continuing with the proof of Proposition 5.1, we have by Lemma 5.1,

$$(5.12) \quad A_k((I - I_k P_{k-1})W, W) \leq C \|A_k^{(1+\beta)/2} W\|_k \|\tilde{W} - P_{k-1}W\|_{H^{1-\beta}(\Omega)}.$$

Now

$$\|\tilde{W} - P_{k-1}W\|_{H^{1-\beta}(\Omega)} \leq \|\tilde{W} - \bar{P}_{k-1}\tilde{W}\|_{H^{1-\beta}(\Omega)} + \|P_{k-1}W - \bar{P}_{k-1}\tilde{W}\|_{H^1(\Omega)}.$$

By finite element duality [1] (see also (6.12) and the following estimates; this makes use of the assumption (5.7)) we have that

$$\|\tilde{W} - \bar{P}_{k-1}\tilde{W}\|_{H^{1-\beta}(\Omega)} \leq ch_k^\beta \|\tilde{W}\|_{H^1(\Omega)},$$

and combining with (5.10) gives

$$(5.13) \quad \|\tilde{W} - P_{k-1}W\|_{H^{1-\beta}(\Omega)} \leq ch_k^\beta A_k(W, W)^{1/2}.$$

Clearly,

$$(5.14) \quad \|A_k^{(1+\beta)/2} W\|_k \leq \|A_k W\|_k^\beta A_k(W, W)^{(1-\beta)/2}.$$

Combining (5.10), (5.12), (5.13), and (5.14) proves (A.3) for  $\alpha = \beta/2$  and hence completes the proof of the proposition.  $\square$

*Remark 5.1.* One obvious application of the results of this section is to the case when  $\Omega$  is an L-shaped domain. As far as we know, this is the first proof which guarantees convergence for a multigrid algorithm for this five-point operator on this domain with linear interpolation as a prolongation operator.

*Remark 5.2.* It is possible to analyze the analogous multigrid algorithm in the case when the lower-order term in (5.1) has variable coefficients. In that case, it will be unlikely that (A.2) holds, and hence one should use the theory of §4.

## 6. FINITE ELEMENT EXAMPLES WITH ALTERNATIVE PROLONGATIONS

In this section, we consider two finite element examples with nested spaces where the prolongation operator does not correspond to the natural imbedding of the coarser space into the finer. An immediate consequence of the use of these prolongation operators is that (2.5) no longer holds, and hence the variational theory does not apply. The first example leads to an algorithm which is equivalent to a rather reasonable finite difference multigrid application [8], and our theory provides new estimates for its convergence. The second example provides an instance when (A.2) is violated, in fact, (A.5) is sharp as  $k \rightarrow \infty$  (see Remark 6.1).

We consider the simplest of all finite element applications. We start with a domain  $\Omega$  which is a union of rectangles and consider the problem

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

We shall provide two different finite element subspaces for this problem.

In the first case, we define a coarse grid triangulation by dividing each rectangle into two triangles, using one of the diagonals of the rectangle. Finer grids are defined by successively dividing each triangle into four by connecting the midpoints of the edges of the triangle (see Figure 6.1). The finite element subspace  $\mathcal{M}_k$  is defined to be the space of continuous piecewise linear functions on the  $k$ th grid which vanish on  $\partial\Omega$ .

In the second case, we consider the corresponding sequence of rectangular grids, i.e., the original collection of rectangles is successively refined by dividing each rectangle into four subrectangles in the obvious way. The finite element subspace  $\tilde{\mathcal{M}}_k$  is defined to be the space of continuous piecewise bilinear functions on the  $k$ th rectangular grid which vanish on  $\partial\Omega$ .

The Galerkin approximation to the solution  $u$  of (6.1) is, of course, defined as the function  $u_j \in \mathcal{M}_j$  (resp.  $\tilde{\mathcal{M}}_j$ ) satisfying

$$D(u_j, \chi) = (f, \chi) \quad \text{for all } \chi \in \mathcal{M}_j \text{ (resp. } \tilde{\mathcal{M}}_j).$$

We define  $A_k(\cdot, \cdot) = D(\cdot, \cdot)$  and

$$(6.2) \quad (u, v)_k = h_k^2 \sum_{ij} u(x_{ij}^k) v(x_{ij}^k).$$

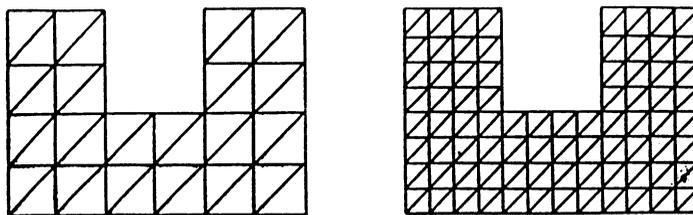


FIGURE 6.1  
The grids for  $\mathcal{M}_0$  and  $\mathcal{M}_1$

The sum in (6.2) is taken over the nodal points  $x_{ij}^k$  of the  $k$ th mesh, and  $h_k = 2^{-k} h_0$ , where  $h_0$  corresponds to the size of the rectangles of the coarsest mesh.

Note that we get the standard variational finite element multigrid algorithms if we define  $I_k$  to be the imbedding of  $\mathcal{M}_{k-1}$  into  $\mathcal{M}_k$  (resp.  $\tilde{\mathcal{M}}_{k-1}$  into  $\tilde{\mathcal{M}}_k$ ). Instead, in the first case, for  $u \in \mathcal{M}_{k-1}$ , we define the values of  $I_k u$  at the nodes of  $\mathcal{M}_k$  by first interpolating  $u$  into  $\tilde{\mathcal{M}}_{k-1}$  and subsequently interpolating the result into  $\mathcal{M}_k$ . Note that the natural imbedding uses linear interpolation on the  $(k-1)$ st triangulation and differs from  $I_k$  only in that it assigns  $(b+c)/2$  to the center node in Figure 6.2 instead of  $(a+b+c+d)/4$ . Analogously, in the second case, we define  $\tilde{I}_k u$  at the nodes of  $\tilde{\mathcal{M}}_k$  by interpolation into the subspace  $\mathcal{M}_{k-1}$  followed by interpolation into  $\tilde{\mathcal{M}}_k$ . Thus at the fine grid nodes, the interpolation operator for the first problem corresponds to the natural imbedding for the second, and vice versa.

The multigrid algorithm in the first case can be thought of as a finite difference multigrid application. Indeed, the stiffness matrix is the standard five-point difference stencil. Moreover, from the finite difference point of view, the prolongation  $I_k$  is as reasonable as any other [8].

We can now give the following proposition.

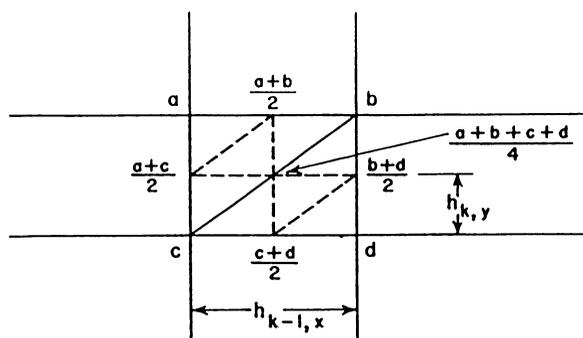


FIGURE 6.2  
Nodal values for the  $I_k$  interpolation

**Proposition 6.1.** *Let  $\mathcal{M}_k$ ,  $A_k(\cdot, \cdot)$ ,  $(\cdot, \cdot)_k$ , and  $I_k$  be defined as above and assume that (5.7) holds for the solution  $u$  of (6.1). Then (A.2) holds, and (A.3) holds for  $\alpha \leq \beta/2$ . If  $\mathcal{M}_k$  and  $I_k$  are replaced by  $\widetilde{\mathcal{M}}_k$  and  $\widetilde{I}_k$ , then (A.2) may no longer hold but (A.3) still holds with  $\alpha \leq \beta/2$ .*

Combining the proposition with Theorems 2–4 gives results for the corresponding multigrid algorithms using  $\mathcal{M}_k$  and  $I_k$  and an appropriate smoothing process. Note that we get uniform (independent of  $h_j$ ) convergence for the  $\mathcal{W}$  cycle and variable  $\mathcal{V}$  cycle algorithms. With the  $\mathcal{V}$  cycle, we may see deterioration in the convergence rate like  $1 - c/\ln(h_j^{-1})$  even in the case of full elliptic regularity ( $\beta = 1$ ).

In the case of  $\widetilde{\mathcal{M}}_k$  and  $\widetilde{I}_k$ , (A.2) will not hold in general. Hence, the multigrid operator  $I - B_k A_k$  need not be a reducer. In contrast, the  $\mathcal{V}$  cycle multigrid strategies employing the multigrid operator as a preconditioner will always be stable and convergent.

*Proof of Proposition 6.1.* We first prove (A.2) in the case of  $\mathcal{M}_k$ ,  $I_k$ , i.e.,

$$(6.3) \quad D(I_k W, I_k W) \leq D(W, W) \quad \text{for all } W \in \mathcal{M}_{k-1}.$$

It clearly suffices to prove the corresponding local inequality

$$(6.4) \quad D_R(I_k W, I_k W) \leq D_R(W, W) \quad \text{for all } W \in \mathcal{M}_{k-1},$$

where the Dirichlet form is over the domain  $R$  and  $R$  is a typical rectangle of the  $(k-1)$ st grid. Let  $R$  be the rectangle pictured in Figure 6.2 and assume that  $W$  takes on the values indicated on the corners of the rectangle. Then (6.4) is equivalent to ( $r = h_{ky}/h_{kx}$ )

$$(6.5) \quad \begin{aligned} & \frac{1}{8}[(a+b-c-d)^2/r + r(b+d-c-a)^2] \\ & + \frac{1}{4}[r(a-b)^2 + r(c-d)^2 + (d-b)^2/r + (c-a)^2/r] \\ & \leq \frac{1}{2}[r(a-b)^2 + r(c-d)^2 + (d-b)^2/r + (c-a)^2/r]. \end{aligned}$$

Inequality (6.5) clearly holds and hence (A.2) follows.

We next prove (A.3) in the case of  $\mathcal{M}_k$ ,  $I_k$ . Lemma 5.1 with  $v$  replacing  $V$  and  $\widetilde{V}$  in (5.11) and  $A_k(\cdot, \cdot)$  as defined above was proved in [3]. Consequently,

$$(6.6) \quad A_k((I - I_k P_{k-1})w, w) \leq \|A_k^{(1+\beta)/2} w\|_k \|(I - I_k P_{k-1})w\|_{H^{1-\beta}(\Omega)}$$

holds for all  $w \in \mathcal{M}_k$ . Let  $\overline{P}_k$  denote the elliptic projection onto  $\mathcal{M}_k$ , i.e.,

$$D(\overline{P}_k v, \phi) = D(v, \phi) \quad \text{for all } \phi \in \mathcal{M}_k.$$

By standard finite element techniques (the duality argument) [1, 12],

$$(6.7) \quad \|(I - \overline{P}_{k-1})w\|_{H^{1-\beta}(\Omega)} \leq ch_k^\beta \|w\|_{H^1(\Omega)}.$$

Here we have used hypothesis (5.7).

Let  $\bar{I}_k$  denote the standard interpolation operator onto the subspace  $\mathcal{M}_k$ . Applying the Bramble-Hilbert Lemma directly in the fractional-order spaces and [14], noting that on each rectangle of the  $(k-1)$ st grid,  $I_k \bar{I}_{k-1} - I$  annihilates linear functions, we conclude that

$$(6.8) \quad \|(I_k \bar{I}_{k-1} - I)v\|_{H^{1-\beta}(\Omega)} \leq ch_k^{\beta+\delta} \|v\|_{H^{1+\delta}(\Omega)},$$

and

$$(6.9) \quad \|(\bar{I}_k - I)v\|_{H^{1-\beta}(\Omega)} \leq ch_k^{\beta+\delta} \|v\|_{H^{1+\delta}(\Omega)}$$

holds for  $0 \leq \beta \leq 1$  and  $0 < \delta \leq 1$ . We note that by the inverse property

$$(6.10) \quad \|v\|_{H^{1+\delta}(\Omega)} \leq ch_{k-1}^{-\delta} \|v\|_{H^1(\Omega)};$$

(6.8) and (6.9) hold for  $\delta = 0$  when  $v \in \mathcal{M}_{k-1}$ . The inequality (6.10) will be proved in the appendix.

By the triangle inequality,

$$\begin{aligned} \|(I - I_k P_{k-1})w\|_{H^{1-\beta}(\Omega)} &\leq \|(I - \bar{P}_{k-1})w\|_{H^{1-\beta}(\Omega)} + \|(I - I_k \bar{I}_{k-1})P_{k-1}w\|_{H^{1-\beta}(\Omega)} \\ &\quad + \|(P_{k-1} - \bar{P}_{k-1})w\|_{H^{1-\beta}(\Omega)}. \end{aligned}$$

By (6.6), (6.7), (6.8) and (3.1), (A.3) will follow if we can show

$$(6.11) \quad \|(P_{k-1} - \bar{P}_{k-1})w\|_{H^{1-\beta}(\Omega)} \leq ch_k^\beta \|w\|_{H^1(\Omega)}.$$

We use a duality argument to derive (6.11). Since  $\dot{H}^1 = H_0^1(\Omega)$  is contained in  $H^1(\Omega)$ , by interpolation

$$(6.12) \quad \begin{aligned} \|(P_{k-1} - \bar{P}_{k-1})w\|_{H^{1-\beta}(\Omega)} &\leq C \| (P_{k-1} - \bar{P}_{k-1})w \|_{1-\beta} \\ &= C \sup_{\phi \in C_0^\infty(\Omega)} \frac{((P_{k-1} - \bar{P}_{k-1})w, T^{(\beta-1)/2}\phi)}{\|\phi\|_{L^2(\Omega)}}. \end{aligned}$$

The power of  $T$  is, of course, defined in terms of its eigenfunction expansion and the equality above is an immediate consequence of the definition of the norm in (5.6). Let  $\zeta$  solve

$$\begin{aligned} -\Delta \zeta &= T^{(\beta-1)/2} \phi \quad \text{in } \Omega, \\ \zeta &= 0 \quad \text{on } \partial\Omega. \end{aligned}$$

Then

$$(6.13) \quad \begin{aligned} ((P_{k-1} - \bar{P}_{k-1})w, T^{(\beta-1)/2}\phi) &= D((P_{k-1} - \bar{P}_{k-1})w, \zeta - \bar{I}_{k-1}\zeta) \\ &\quad + D(w, (I_k - I)\bar{I}_{k-1}\zeta). \end{aligned}$$

By (6.9) and (3.1), the first term in (6.13) can be bounded in absolute value by

$$D((P_{k-1} - \bar{P}_{k-1})w, \zeta - \bar{I}_{k-1}\zeta) \leq Ch_k^\beta \|w\|_{H^1(\Omega)} \|\zeta\|_{H^{1+\beta}(\Omega)}.$$

For the second term,

$$\|(I_k - I)\bar{I}_{k-1}\zeta\|_{H^1(\Omega)} \leq \|(I_k \bar{I}_{k-1} - I)\zeta\|_{H^1(\Omega)} + \|(I - \bar{I}_{k-1})\zeta\|_{H^1(\Omega)}.$$

Applying (6.8) and (6.9) shows that the second term in (6.13) can be bounded similarly. Thus, by (5.7) we have

$$\begin{aligned} ((P_{k-1} - \bar{P}_{k-1})w, T^{(\beta-1)/2}\phi) &\leq ch_k^\beta \|w\|_{H^1(\Omega)} \|\zeta\|_{H^{1+\beta}(\Omega)} \\ &\leq ch_k^\beta \|\phi\|_{L^2(\Omega)} \|w\|_{H^1(\Omega)}. \end{aligned}$$

Combining the above inequalities proves (6.11) and hence completes the proof of the proposition in the case of  $\mathcal{M}_k, I_k$ .

The proof in the case  $\tilde{\mathcal{M}}_k$  and  $\tilde{I}_k$  is similar except that we use the inequality corresponding to (6.8) to deduce the boundedness of  $\cdot|_{k-1}$  in  $H^1(\Omega)$ . This completes the proof of Proposition 6.1.  $\square$

*Remark 6.1.* In general, (A.2) does not hold in the second case. In fact, when  $h_x = h_y$ , there is a local function defined on the four nodes of size  $h_{k-1}$  in Figure 6.2 such that

$$D_R(\tilde{I}_k W, \tilde{I}_k W) = 2D_R(W, W).$$

In addition, we have computed the minimal constant  $\mu_k$  satisfying

$$(6.14) \quad A_k(I_k u, I_k u) \leq \mu_k A_{k-1}(u, u) \quad \text{for all } u \in \mathcal{M}_{k-1},$$

for the slit domain (see Example 9.1) and found that, for this example,  $\mu_k \rightarrow 2$  as  $k \rightarrow \infty$ .

## 7. A FINITE ELEMENT EXAMPLE WITH LOOSELY COUPLED GRIDS

In this section, we consider a finite element example using a sequence of loosely coupled grids. By loosely coupled, we mean that the triangulation on the  $k$ th grid is quasi-uniform of size  $h_k$ . In general, the grids and their corresponding finite element subspaces are nonnested. Our results apply to the natural finite element method applied to a problem with curved boundaries where a sequence of grids are generated which successively more closely approximate the boundary of the original domain. In general, (A.2) will not hold. We will show that (A.3) holds with appropriate  $\alpha$  and  $C_\alpha$  independent of the number of levels. Thus, the preconditioning results of Theorem 6 hold.

Let  $\Omega$  be a domain in  $R^2$  with piecewise smooth boundary  $\partial\Omega$ . We consider the numerical approximation to the solution  $u$  of the problem

$$(7.1) \quad \begin{aligned} Lu &= f \quad \text{in } \Omega, \\ u &= 0 \quad \text{on } \partial\Omega, \end{aligned}$$

where

$$Lv = - \sum_{i,j=1}^2 \frac{\partial}{\partial x_i} a_{ij} \frac{\partial v}{\partial x_j}$$

with  $\{a_{ij}(x)\}$  smooth, symmetric, and uniformly positive definite.

We assume that we have defined a sequence of grids  $\{T_k\}$  for  $k = 0, \dots, j$  approximating  $\Omega$  such that the  $k$ th grid consists of triangles of quasi-uniform

size  $h_k$ . In most applications,  $h_k$  is roughly twice the size of  $h_{k+1}$ , although for our theory we need only assume that there are positive constants  $c_0$  and  $c_1$  not depending on  $k$  satisfying

$$(7.2) \quad c_0 h_k \leq h_{k+1} \leq c_1 h_k.$$

We define  $\mathcal{M}_k$  to be the set of functions which are piecewise linear on  $T_k$  and vanish on the nodes of  $T_k$  on  $\partial\Omega$ . For good approximation, the boundary nodes of the triangulation  $T_k$  should lie on  $\partial\Omega$ . Note that we have not assumed that the nodes of the triangles of the  $k$ th grid are related in any way to the nodes of the triangles of the  $(k-1)$ st.

For convenience of exposition, we shall only consider the case where every triangle of each  $T_k$  lies in  $\bar{\Omega}$ . We consider the functions in  $\mathcal{M}_k$  to be extended by zero to  $\Omega$  and thus can think of  $\mathcal{M}_k$  as being contained in  $H^{1+\beta}(\Omega)$  for  $\beta < 1/2$ . The forms on  $\mathcal{M}_k$  are defined by

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

where

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

The prolongation operators  $I_k$  are defined by the natural interpolation operator associated with the subspace  $\mathcal{M}_k$ , i.e.,

$$I_k w(x_i) = w(x_i)$$

for nodes  $x_i$  of  $T_k$  and functions  $w \in \mathcal{M}_{k-1}$ . The discrete inner products are defined by

$$(u, v)_k = h_k^2 \sum u(x_i) v(x_i)$$

where the sum is taken over all the nodes of  $T_k$ .

The following proposition shows that (A.3) holds with appropriate  $\alpha$ . Combining this result with Theorems 6 and 7 implies conditioning results for the variable  $\mathcal{V}$  cycle and  $\mathcal{W}$  cycle multigrid algorithms. The variable  $\mathcal{V}$  cycle results hold with  $m(j) = 1$  while the  $\mathcal{W}$  cycle results hold only assuming that  $m$  is sufficiently large.

**Proposition 7.1.** *Let  $\mathcal{M}_k$ ,  $A_k(\cdot, \cdot)$ ,  $(\cdot, \cdot)_k$ , and  $I_k$  be defined as above. Furthermore, assume that (5.7) holds for the solution  $u$  of (7.1). Then (A.3) holds for  $\alpha < \min(\beta/2, 1/4)$ .*

*Proof.* Without loss of generality, we assume  $\beta < 1/2$ . The argument given in [3] can be used to show that

$$c_0 \|A_k^{(1-\beta)/2} W\|_k \leq \|W\|_{H^{1-\beta}(\Omega)} \leq C_1 \|A_k^{(1-\beta)/2} W\|_k \quad \text{for all } W \in \mathcal{M}_k.$$

The proof proceeds as the proof of Proposition 6.1. By (6.6), it suffices to estimate the norm in  $H^{1-\beta}(\Omega)$  of  $(I - I_k P_{k-1})w$ . Let  $\bar{P}_{k-1}$  denote the elliptic projection into  $\mathcal{M}_{k-1}$  defined by

$$A(\bar{P}_{k-1} v, \phi) = A(v, \phi) \quad \text{for all } \phi \in \mathcal{M}_{k-1}.$$

The triangle inequality gives

$$(7.3) \quad \begin{aligned} & \|(I - I_k P_{k-1})w\|_{H^{1-\beta}(\Omega)} \\ & \leq \|(I - \bar{P}_{k-1})w\|_{H^{1-\beta}(\Omega)} + \|(I - I_k)P_{k-1}w\|_{H^{1-\beta}(\Omega)} \\ & \quad + \|(P_{k-1} - \bar{P}_{k-1})w\|_{H^{1-\beta}(\Omega)}. \end{aligned}$$

As in (6.9), the estimate

$$(7.4) \quad \|(I - I_k)v\|_{H^{1-\beta}(\Omega)} \leq ch_k^{\beta+\delta} \|v\|_{H^{1+\delta}(\Omega)}$$

holds for  $0 \leq \beta \leq 1$  and  $0 < \delta \leq 1$ . Again, by (6.10), (7.4) holds for  $\delta = 0$  when  $v$  is in  $\mathcal{M}_{k-1}$ . By (7.4) with  $\delta = \beta = 0$ ,  $P_{k-1}$  is a bounded operator with respect to the norm induced by the  $A(\cdot, \cdot)$  inner product. Hence, we have that

$$\|(I - I_k)P_{k-1}w\|_{H^{1-\beta}(\Omega)} \leq Ch_k^\beta \|P_{k-1}w\|_{H^1(\Omega)} \leq Ch_k^\beta \|w\|_{H^1(\Omega)}.$$

The estimate corresponding to (6.7) is well known under the assumption (5.7). Thus, we are left to bound the third term of (7.3).

As in the proof of Proposition 6.1, we use a duality argument. Replacing  $T^{-1}$  by  $L$  in (6.12)–(6.13) gives

$$\|(P_{k-1} - \bar{P}_{k-1})w\|_{H^{1-\beta}(\Omega)} \leq C \sup_{\phi \in C_0^\infty(\Omega)} \frac{A((P_{k-1} - \bar{P}_{k-1})w, \zeta)}{\|\phi\|_{L^2(\Omega)}},$$

where  $\zeta$  is defined by

$$\begin{aligned} L\zeta &= L^{(1-\beta)/2} \phi && \text{in } \Omega, \\ \zeta &= 0 && \text{on } \partial\Omega. \end{aligned}$$

But

$$A((P_{k-1} - \bar{P}_{k-1})w, \zeta) = A(w, (I_k - I)\bar{P}_{k-1}\zeta).$$

In the appendix, we show that for  $\beta < 1/2$ ,

$$(7.5) \quad \|\bar{P}_{k-1}v\|_{H^{1+\beta}(\Omega)} \leq C\|v\|_{H^{1+\beta}(\Omega)}$$

and hence (7.4) implies

$$A(w, (I_k - I)\bar{P}_{k-1}\zeta) \leq Ch_k^\beta \|\zeta\|_{H^{1+\beta}(\Omega)} \|w\|_{H^1(\Omega)}.$$

Combining the above estimates with (5.7) proves

$$\|(I - I_k P_{k-1})w\|_{H^{1-\beta}(\Omega)} \leq Ch_k^\beta \|w\|_{H^1(\Omega)}.$$

This completes the proof of the proposition.  $\square$

*Remark 7.1.* Slightly stronger results can be obtained with further assumptions on the relationships between grids. For example, one natural way of developing a sequence of grids is as follows. Each coarse grid triangle gives rise to four triangles in the finer grid connecting the midpoints of interior edges and the midpoint (with respect to arc-length) along the boundary curve for boundary edges (see Figure 7.1). In this case, under reasonable smoothness assumptions

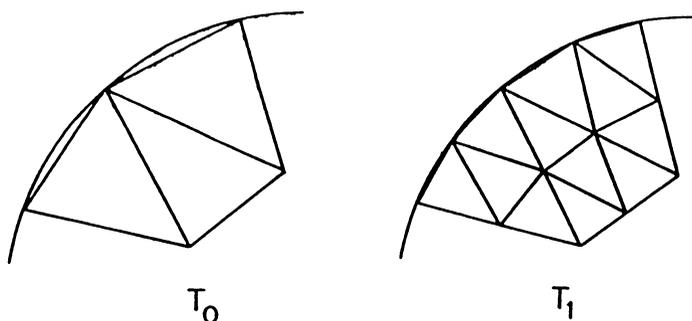


FIGURE 7.1  
The mesh of Remark 7.1

on the boundary, it is possible to prove by a perturbation argument that (A.6) holds with  $\gamma = 1/4$  [31]. Hence, Theorems 7 and 8, and Remark 4.3, provide results for the corresponding  $\mathcal{V}$  and  $\mathcal{W}$  cycle algorithms.

#### 8. A MIXED FINITE ELEMENT EXAMPLE

In this section, we develop a multigrid technique for a mixed method finite element approximation of a second-order elliptic problem. We consider the so-called “Raviart-Thomas” elements on triangles and the analogous elements on rectangles [26]. In this example, assumption (A.2) will be satisfied. A similar treatment of the “Brezzi-Douglas-Marini” elements [10] may also be carried out.

Let  $\Omega$  be a bounded domain in  $R^N$  for  $N = 2$  or  $N = 3$ . We consider the problem

$$(8.1) \quad \begin{aligned} -\nabla \cdot (\kappa \nabla w) &= f \quad \text{in } \Omega, \\ w &= 0 \quad \text{on } \partial\Omega. \end{aligned}$$

We assume that  $\kappa = \kappa(x)$  is smooth and bounded from below by some constant  $\kappa_0 > 0$ .

The mixed approximation to (8.1) can be developed as follows. We define  $\mathbf{P} = \kappa \nabla w$  and note that the pair  $(\mathbf{P}, w)$  satisfies

$$(8.2) \quad \begin{aligned} (\kappa^{-1} \mathbf{P}, \mathbf{Q}) + (w, \nabla \cdot \mathbf{Q}) &= 0 \quad \text{for all } \mathbf{Q} \in H(\text{div}; \Omega), \\ (\nabla \cdot \mathbf{P}, v) &= -(f, v) \quad \text{for all } v \in L^2(\Omega), \end{aligned}$$

where

$$H(\text{div}; \Omega) = \{\Phi \in [L^2(\Omega)]^N \text{ such that } \nabla \cdot \Phi \in L^2(\Omega)\}.$$

$H(\text{div}; \Omega)$  is a Hilbert space [29] with norm given by

$$(\|\Phi\|_{L^2(\Omega)}^2 + \|\nabla \cdot \Phi\|_{L^2(\Omega)}^2)^{1/2}.$$

The pair  $(\mathbf{P}, w)$  is approximated in mixed finite element subspace pairs  $\mathcal{E}_h, V_h$  contained respectively in  $H(\text{div}; \Omega)$  and  $L^2(\Omega)$ . Associated with these pairs is an integer  $r$  which is related to the approximation order. We assume that the

reader is familiar with the construction of these spaces as described in [26]. The mixed approximation is defined to be the pair  $(\mathbf{P}_h, w_h) \in \mathcal{Q}_h \times V_h$  satisfying

$$(8.3) \quad \begin{aligned} (\kappa^{-1} \mathbf{P}_h, Q_h) + (w_h, \nabla \cdot Q_h) &= 0 \quad \text{for all } Q_h \in \mathcal{Q}_h, \\ (\nabla \cdot \mathbf{P}_h, v) &= -(f, v) \quad \text{for all } v \in V_h. \end{aligned}$$

Techniques for solving systems of the form (8.3) have been considered (e.g., [7, 15]). We believe the multigrid technique to be described is new.

To describe our multigrid algorithm, we shall need some additional operator notation. Define the operators  $A: \mathcal{Q}_h \mapsto \mathcal{Q}_h$ ,  $B: V_h \mapsto \mathcal{Q}_h$ , and  $B^*: \mathcal{Q}_h \mapsto V_h$  by

$$\begin{aligned} (Ap, q) &= (\kappa^{-1} p, q) \quad \text{for all } q \in \mathcal{Q}_h, \\ (Bv, q) &= (v, \nabla \cdot q) \quad \text{for all } q \in \mathcal{Q}_h, \\ (B^* p, v) &= (\nabla \cdot p, v) \quad \text{for all } v \in V_h. \end{aligned}$$

Clearly,  $Ap$  is the  $L^2(\Omega)$  orthogonal projection of  $\kappa^{-1} p$  onto  $\mathcal{Q}_h$ ,  $B^* p$  is the  $L^2(\Omega)$  orthogonal projection of  $\nabla \cdot p$  onto  $V_h$ , and  $B$  is the adjoint of  $B^*$ .

With the above notation, (8.3) can be rewritten

$$\begin{pmatrix} A & B \\ B^* & 0 \end{pmatrix} \begin{pmatrix} \mathbf{P}_h \\ w_h \end{pmatrix} = \begin{pmatrix} 0 \\ -P^0 f \end{pmatrix},$$

where  $P^0$  denotes the  $L^2(\Omega)$  orthogonal projection onto  $V_h$ . Thus, the solution  $w_h$  of (8.3) satisfies

$$(8.4) \quad B^* A^{-1} B w_h = P^0 f.$$

We shall develop a multigrid algorithm for (8.4). In the remainder of this section, we restrict ourselves to the case of  $R^2$ .

For the multigrid algorithm to be developed, we assume that the cost of evaluating  $A^{-1}$  applied to a vector in  $\mathcal{Q}_h$  is not too expensive. This is true in the case of tensor product elements on a regular rectangular grid, where the evaluation of the action of  $A^{-1}$  involves banded solves (of bandwidth proportional to  $r$ ) along lines of constant  $y$  for the  $v_x$ -component and along lines of constant  $x$  for the  $v_y$ -component. The overall cost is thus proportional to the total number of unknowns. In the case of triangles, one might consider iterative evaluation of  $A^{-1}$ , a well-conditioned system. To highlight the ideas, from here on, we limit our discussion to the case of tensor product elements on rectangles and assume that  $\Omega$  is the union of such rectangles.

We develop a sequence of rectangular subgrids  $\{R_k\}$ ,  $k = 0, \dots, j-1$ , in the usual way. We start by a coarse partitioning of  $\Omega$  into rectangular elements. Each successively finer grid is defined by partitioning coarser grid rectangles into four equal size subrectangles. The mixed element approximation subspaces are defined with respect to the finest grid  $R_{j-1}$  only, and we define  $\mathcal{M}_j = V_h$ . The space  $\mathcal{M}_k$  for  $k < j$  is defined to be the set of continuous piecewise bilinear functions with respect to the  $k$ th grid which vanish on  $\partial\Omega$ .

Note that  $\mathcal{M}_j$  and  $\mathcal{M}_{j-1}$  are defined with respect to the same grid. The spaces  $\mathcal{M}_0 \subset \mathcal{M}_1 \subset \cdots \subset \mathcal{M}_{j-1}$  form a nested sequence of spaces, but  $\mathcal{M}_{j-1} \subset \mathcal{M}_j$  only if  $r > 0$ .

We next define the multigrid forms. The form  $A_j$  is defined directly from the mixed method by

$$A_j(\theta, \chi) = (B^* A^{-1} B \theta, \chi) \quad \text{for all } \theta, \chi \in \mathcal{M}_j.$$

The forms on the spaces  $\mathcal{M}_k$  for  $k < j$  are defined by

$$A_k(\theta, \chi) = \int_{\Omega} \kappa \nabla \theta \cdot \nabla \chi \, dx \quad \text{for all } \theta, \chi \in \mathcal{M}_k.$$

We assume that “discrete” inner products  $(\cdot, \cdot)_k$  are defined satisfying

$$(8.5) \quad c \|\theta\|_{L^2(\Omega)}^2 \leq (\theta, \theta)_k \leq C \|\theta\|_{L^2(\Omega)}^2 \quad \text{for all } \theta \in \mathcal{M}_k,$$

for  $k = 0, \dots, j$ , with  $c, C$  independent of  $k$ . Note that, for  $k < j$  we can use (6.2) to define  $(\cdot, \cdot)_k$ . In addition,  $(\cdot, \cdot)_j$  can be defined to be the  $L^2(\Omega)$  inner product.

To complete the definition of the multigrid algorithm, we need only define the operators  $I_k$ ,  $k = 1, \dots, j$ . Except in the case  $r = 0$ , all spaces are nested, and hence  $I_k$  can be defined by the natural injection. For  $r = 0$ , only  $\mathcal{M}_{j-1}$  is not contained in  $\mathcal{M}_j$ . In this case, we define  $I_j \theta$  to be the function in  $\mathcal{M}_j$  whose value on a grid rectangle is the mean value of  $\theta$  on that rectangle.

For sufficiently smooth  $\kappa$ , regularity results of the form (5.7) hold for the solution of (8.1) as discussed earlier.

We now give the proposition which shows that the hypotheses (A.2) and (A.3) hold with the above operator definitions. Combining these results with the theorems of §§3 and 4 implies iterative convergence estimates for the corresponding multigrid algorithms (with appropriate smoothers).

**Proposition 8.1.** *Let  $\mathcal{M}_k$ ,  $A_k(\cdot, \cdot)$ ,  $(\cdot, \cdot)_k$ , and  $I_k$  be defined as above. Then (A.2) holds. Furthermore, if the solution  $u$  of (8.1) satisfies (5.7) for some  $\beta \in (1/2, 1]$ , then (A.3) holds for  $\alpha = \beta/2$ .*

*Proof.* We first show that (A.2) holds. Since the above setup corresponds to the usual finite element multigrid for  $k < j$ , the stronger result (2.5) holds for  $k \neq j$ . Thus, we need only verify (A.2) for  $k = j$ . It is easy to see that for  $\theta \in \mathcal{M}_j$ ,

$$(8.6) \quad A_j(\theta, \theta) = \sup_{\chi \in \mathcal{E}_h} \frac{(\theta, \nabla \cdot \chi)^2}{(\kappa^{-1} \chi, \chi)}.$$

If  $r > 0$ , then for  $\theta \in \mathcal{M}_{j-1}$ ,

$$A_j(\theta, \theta) = \sup_{\chi \in \mathcal{E}_h} \frac{(\theta, \nabla \cdot \chi)^2}{(\kappa^{-1} \chi, \chi)} = \sup_{\chi \in \mathcal{E}_h} \frac{(\kappa^{1/2} \nabla \theta, \kappa^{-1/2} \chi)^2}{(\kappa^{-1} \chi, \chi)} \leq A_{j-1}(\theta, \theta).$$

For  $r = 0$ , since  $\nabla \cdot \chi$  is constant on the rectangles of size  $h_j$ ,

$$\begin{aligned} A_j(I_j\theta, I_j\theta) &= \sup_{\chi \in \mathcal{E}_h} \frac{(I_j\theta, \nabla \cdot \chi)^2}{(\kappa^{-1}\chi, \chi)} = \sup_{\chi \in \mathcal{E}_h} \frac{(\theta, \nabla \cdot \chi)^2}{(\kappa^{-1}\chi, \chi)} \\ &= \sup_{\chi \in \mathcal{E}_h} \frac{(\kappa^{1/2}\nabla\theta, \kappa^{-1/2}\chi)^2}{(\kappa^{-1}\chi, \chi)} \leq A_{j-1}(\theta, \theta). \end{aligned}$$

Combining the above inequalities verifies (A.2).

We next prove (A.3). Again, since for  $k < j$ , the above setup corresponds to the usual finite element multigrid, we need only consider  $k = j$ . Fix  $u \in \mathcal{M}_j$  and define  $f \in \mathcal{M}_j$  to be the solution of

$$(8.7) \quad (f, \theta) = A_j(u, \theta) \quad \text{for all } \theta \in \mathcal{M}_j.$$

Clearly,  $f$  is well defined. Moreover,  $u$  is the mixed approximation to the function  $W \in H_0^1(\Omega)$  satisfying

$$\int_{\Omega} \kappa \nabla W \cdot \nabla \phi \, dx = (f, \phi) \quad \text{for all } \phi \in H_0^1(\Omega).$$

Note that  $P_{j-1}u$  satisfies

$$A_{j-1}(P_{j-1}u, \chi) = (f, \chi) \quad \text{for all } \chi \in \mathcal{M}_{j-1}$$

and hence  $P_{j-1}u$  is the standard (conforming) finite element approximation to  $W$  in  $\mathcal{M}_{j-1}$ .

By the Schwarz inequality,

$$A_j((I - I_j P_{j-1})u, u) \leq A_j((I - I_j P_{j-1})u, (I - I_j P_{j-1})u)^{1/2} A_j(u, u)^{1/2}.$$

Consequently, (A.3) will follow if we can show

$$(8.8) \quad \begin{aligned} &A_j((I - I_j P_{j-1})u, (I - I_j P_{j-1})u) \\ &\leq Ch_j^{2\beta} \|A_j u\|_j^{2\beta} A_j(u, u)^{1-\beta} \quad \text{for all } u \in \mathcal{M}_j. \end{aligned}$$

Fix  $u \in \mathcal{M}_j$  and let  $q \in \mathcal{E}_h$  satisfy  $Aq + Bu = 0$ . Applying known error estimates for the mixed finite element method with  $\beta > 1/2$  [24] and the standard finite element method [1, 12] gives

$$(8.9) \quad \begin{aligned} &A_j((I - I_j P_{j-1})u, (I - I_j P_{j-1})u)^{1/2} = \sup_{\chi \in \mathcal{E}_h} \frac{(u - P_{j-1}u, \nabla \cdot \chi)}{(\kappa^{-1}\chi, \chi)^{1/2}} \\ &= \sup_{\chi \in \mathcal{E}_h} \frac{(\nabla W - \kappa^{-1}q, \chi)}{(\kappa^{-1}\chi, \chi)^{1/2}} + \sup_{\chi \in \mathcal{E}_h} \frac{(\nabla(P_{j-1}u - W), \chi)}{(\kappa^{-1}\chi, \chi)^{1/2}} \\ &\leq Ch_j^\beta \|W\|_{H^{1+\beta}(\Omega)} \leq Ch_j^\beta \|f\|_{H^{-1+\beta}(\Omega)}. \end{aligned}$$

The last inequality of (8.9) used (5.7).

Clearly, by (8.5) and (8.7),

$$\|f\|_{H^{-1+\beta}(\Omega)} \leq C \|f\|_{H^{-1}(\Omega)}^{1-\beta} \|f\|_{L^2(\Omega)}^\beta \leq C \|f\|_{H^{-1}(\Omega)}^{1-\beta} \|A_j u\|_j^\beta.$$

In addition,

$$\begin{aligned} \|f\|_{H^{-1}(\Omega)} &= \sup_{\phi \in H_0^1(\Omega)} \frac{(f, P^0 \phi)}{\|\phi\|_{H^1(\Omega)}} = \sup_{\phi \in H_0^1(\Omega)} \frac{(A_j u, P^0 \phi)_j}{\|\phi\|_{H^1(\Omega)}} \\ &\leq A_j(u, u)^{1/2} \sup_{\phi \in H_0^1(\Omega)} \frac{A_j(P^0 \phi, P^0 \phi)^{1/2}}{\|\phi\|_{H^1(\Omega)}}. \end{aligned}$$

But the mixed element spaces satisfy  $\nabla \cdot \mathcal{E}_h \subset V_h$ , and hence

$$\begin{aligned} A_j(P^0 \phi, P^0 \phi) &= \sup_{\chi \in \mathcal{E}_h} \frac{(P^0 \phi, \nabla \cdot \chi)^2}{(\kappa^{-1} \chi, \chi)} = \sup_{\chi \in \mathcal{E}_h} \frac{(\phi, \nabla \cdot \chi)^2}{(\kappa^{-1} \chi, \chi)} \\ &= \sup_{\chi \in \mathcal{E}_h} \frac{(\nabla \phi, \chi)^2}{(\kappa^{-1} \chi, \chi)} \leq C \|\phi\|_{H^1(\Omega)}^2. \end{aligned}$$

Combining the above results gives

$$(8.10) \quad \|f\|_{H^{-1+\beta}(\Omega)} \leq C A_j(u, u)^{(1-\beta)/2} \|A_j u\|_j^\beta.$$

Combining (8.9)–(8.10) verifies (8.8). This completes the proof of the proposition.  $\square$

## 9. NUMERICAL RESULTS

We provide the results of a few numerical experiments to illustrate the theory developed in the earlier sections. We have made no attempt to provide numerical results for all of the examples. Instead, we provide examples only to illustrate the theorems in §§3 and 4.

**Example 9.1.** We consider the Laplace equation on a slit domain. Specifically, we define  $\Omega$  to be the points interior to the unit square which are not on the line  $(1/2, y)$  for  $y \in [1/2, 1)$ , and we approximate the solution to (6.1). We define  $\mathcal{M}_k$  to be the space  $\tilde{\mathcal{M}}_k$  of piecewise bilinear functions on the  $k$ th rectangular grid as developed in §6. The prolongation operator  $I_k$  corresponds to the linear interpolant  $\tilde{I}_k$  with respect to the triangular mesh defined in §6. For this example, the form  $A_k$  corresponds to the Dirichlet form on the subspace, i.e.,

$$A_k(u, u) = D(u, u) \quad \text{for all } u \in \mathcal{M}_k.$$

In this application, (A.5) is satisfied, but (A.2) is not.

Table 9.1 gives the condition number  $K$  for the system  $B_j A_j$  and the reduction factor  $\delta$  ( $\delta = \delta_j$  in (3.2)) for the system  $I - B_j A_j$  as a function of the mesh size on the finest grid. We compare the  $\mathcal{V}$  cycle  $(K_v, \delta_v)$ , the variable  $\mathcal{V}$  cycle  $(K_{vv}, \delta_{vv})$  and the  $\mathcal{W}$  cycle  $(K_w, \delta_w)$  multigrid algorithm. We use Richardson smoothing, and hence (A.4) is satisfied. The variable  $\mathcal{V}$  cycle used twice the number of smoothings on each consecutively coarser grid (i.e.,  $\beta_0 = \beta_1 = 2$ ) and one smoothing on the finest grid. The  $\mathcal{V}$  and  $\mathcal{W}$  cycle

TABLE 9.1  
Convergence results for Example 9.1

$h_j$	$K_{vv}(\delta_{vv})$	$K_w(\delta_w)$	$K_v(\delta_v)$
1/8	2.1 (.45)	2.1 (.45)	2.1 (.45)
1/16	2.2 (.45)	2.2 (.45)	2.4 (.50)
1/32	2.2 (.46)	2.2 (.46)	2.7 (.54)
1/64	2.2 (.46)	2.2 (.46)	3.0 (.57)
1/128	2.2 (.46)	2.2 (.46)	3.2 (.59)

algorithms used  $m(k) = m = 1$ . For all of the runs, the coarse grid corresponded to a mesh of size  $1/4$  and the coarse grid problems were essentially solved by applying 40 smoothing steps. Note that for this example, the computational results for the variable  $\mathcal{V}$  and the  $\mathcal{W}$  cycle multigrid algorithms are essentially identical. This is reasonable since both algorithms have exactly the same number of total smoothings on each grid in the multi-level iteration. This example satisfies the hypotheses of Theorems 6 and 7, and the observed behavior of the variable  $\mathcal{V}$  and  $\mathcal{W}$  cycle algorithms agree with the theory. However, the behavior of the  $\mathcal{V}$  cycle algorithm is perhaps better than one would expect from the theory of §§3 and 4.

**Example 9.2.** This example illustrates what can happen to the multigrid algorithms when the minimal constant  $\mu_k$  satisfying (6.14) is greater than 2. We consider the same setup as in Example 9.1 except that we define  $A_k$  by

$$(9.1) \quad A_k(u, u) = \tau_k D(u, u) \quad \text{for all } u \in \mathcal{M}_k,$$

where  $\tau_j = 1$  and for  $k < j$ ,

$$\tau_k = \prod_{i=k}^{j-1} (1 + 6h_i).$$

A result of this scaling is that (A.5) no longer holds. Clearly,  $\tau_k = 1 + O(h_k)$  and it is not difficult to show that (A.3) still holds.

Even though the scaling introduced in (9.1) is purely artificial, it is not unreasonable to expect similar differences in forms in actual applications. Such differences might be observed if the operator involved had variable coefficients and the forms on the individual grids were computed by numerical integration.

Table 9.2 gives computational results for this example. The condition number for the variable  $\mathcal{V}$  cycle algorithm ( $K_{vv}$ ) and the  $\mathcal{V}$  cycle ( $K_v$ ) algorithm as a function of  $h_j$  is reported. In addition, the largest ( $\eta_1^w$ ) and smallest ( $\eta_0^w$ ) eigenvalue of the operator  $B_j A_j$  is given in the case of the  $\mathcal{W}$  cycle algorithm with  $m = 1$ . Finally, the minimum value of  $\mu_k$  satisfying (6.14) is also given. Note that for these computations, the  $\mathcal{W}$  cycle algorithm leads to an indefinite operator  $B_j$  for more than two levels. Thus, the extra smoothing requirement

TABLE 9.2  
Convergence results for Example 9.2

$h_j$	$K_{vv}$	$K_v$	$\eta_0^w$ ( $\eta_1^w$ )	$\mu_k$
1/8	3.8	3.8	.57 (2.2)	3.7
1/16	5.2	6.1	-.4 (1.6)	3.2
1/32	5.7	8.5	-1.2 (1.4)	2.7
1/64	5.2	10.6	-4.4 (1.3)	2.4
1/128	4.6	12.5	-30 (1.2)	2.2

in Theorem 7 is needed to produce a stable  $\mathscr{W}$  cycle multigrid algorithm. In contrast, the hypotheses for Theorem 6 are satisfied and the computational results for the variable  $\mathscr{V}$  cycle algorithm illustrate the uniform conditioning of the  $B_j A_j$  guaranteed by the theory. As in Example 9.1, the behavior of the  $\mathscr{V}$  cycle seems better than that predicted by the theory. The  $\mathscr{V}$  cycle does show more deterioration in the condition number compared to Example 9.1, but nevertheless would lead to a reasonable preconditioned strategy for solving (2.1).

*Remark 9.1.* Although not reported, the largest eigenvalues of  $B_j A_j$  for variable  $\mathscr{V}$  and  $\mathscr{V}$  cycle computations in Table 9.2 were always greater than 2. Accordingly,  $I - B_j A_j$  is not a reducer. Obviously, there exists a constant  $\gamma < 1$  so that  $I - \gamma B_j A_j$  is a reducer with a good reduction rate. An iterative algorithm with reduction matrix  $I - \gamma B_j A_j$  can be trivially constructed and is, equivalently, a linear preconditioned iteration for the computation of the action of  $A_j^{-1}$  applied to a function in  $\mathscr{M}_j$ . Note that for the  $\mathscr{W}$  cycle algorithm with one smoothing and more than two levels, there does not exist a constant  $\gamma$  so that  $I - \gamma B_j A_j$  is a reducer. For a stable iterative technique utilizing the  $\mathscr{W}$  cycle algorithm, one would have to increase  $m$ .

## 10. APPENDIX

We give a proof of (6.10) and (7.5) in this section. We will prove the results for piecewise linear functions on triangles. The proof for bilinear functions is similar. Let  $m$  be a nonnegative integer. Assume that we are given a quasi-uniform triangulation  $\{\cup \tau_i\}$  of size  $h$  on  $\Omega$  and consider the space  $S_h$  of discontinuous piecewise polynomials up to degree  $m$  on this triangulation. We will prove that the inverse inequality

$$(10.1) \quad \|v\|_{H^\alpha(\Omega)} \leq Ch^{-\alpha} \|v\|_{L^2(\Omega)} \quad \text{for all } v \in S_h$$

holds for  $\alpha < 1/2$ .

Assuming that (10.1) holds, we can prove (6.10) and (7.5) as follows. Clearly,

$$(10.2) \quad \|v\|_{H^{1+\alpha}(\Omega)} \leq C \left( \|v\|_{H^1(\Omega)} + \sum_{i=1}^2 \left\| \frac{\partial v}{\partial x_i} \right\|_{H^\alpha(\Omega)} \right).$$

Inequality (6.10) follows applying (10.1) to  $\frac{\partial v}{\partial x_i}$ . To prove (7.5), we let the above triangulation correspond to the triangulation defining  $\mathcal{M}_{k-1}$  and let  $\pi$  denote the  $L^2(\Omega)$  orthogonal projection operator onto  $\mathcal{M}_{k-1}$ . Obviously, it suffices to prove that

$$\|\bar{P}_{k-1}u - u\|_{H^{1+\beta}(\Omega)} \leq C\|u\|_{H^{1+\beta}(\Omega)}.$$

Clearly,

$$\begin{aligned} & \|\bar{P}_{k-1}u - u\|_{H^{1+\beta}(\Omega)} \\ & \leq \|u\|_{H^1(\Omega)} + \sum_{i=1}^2 \left( \left\| (I - \pi) \frac{\partial u}{\partial x_i} \right\|_{H^\beta(\Omega)} + \left\| \pi \frac{\partial u}{\partial x_i} - \frac{\partial}{\partial x_i} (\bar{P}_{k-1}u) \right\|_{H^\beta(\Omega)} \right). \end{aligned}$$

By (10.1) and approximation properties,

$$\begin{aligned} \left\| \pi \frac{\partial u}{\partial x_i} - \frac{\partial}{\partial x_i} (\bar{P}_{k-1}u) \right\|_{H^\beta(\Omega)} & \leq Ch_k^{-\beta} \left\| \pi \frac{\partial u}{\partial x_i} - \frac{\partial}{\partial x_i} (\bar{P}_{k-1}u) \right\|_{L^2(\Omega)} \\ & \leq C\|u\|_{H^{1+\beta}(\Omega)}. \end{aligned}$$

Inequality (7.5) then follows from the fact that  $\pi$  is a bounded operator on  $H^\beta(\Omega)$ .

We provide a proof of (10.1) in the remainder of this section. To do this, we use the real method of interpolation of Lions and Peetre (see [11]) which asserts that we may take

$$\|u\|_{H^\beta(\Omega)}^2 = \int_0^\infty K(u, t)^2 t^{-2\beta-1} dt,$$

where

$$K(u, t)^2 = \inf_{v \in H^1(\Omega)} (\|u - v\|_{L^2(\Omega)}^2 + t^2 \|v\|_{H^1(\Omega)}^2).$$

In fact, a direct computation shows that the norm above is equal to a constant (depending on  $\beta$ ) multiple of the Hilbert scale norm. Taking  $v = 0$  in the definition of  $K(u, t)^2$  gives

$$(10.3) \quad \int_h^\infty K(u, t)^2 t^{-2\beta-1} dt \leq (2\beta)^{-1} h^{-2\beta} \|u\|_{L^2(\Omega)}^2.$$

Thus, we are left to estimate the integral from 0 to  $h$ .

For  $t \leq h$ , we define  $v$  (depending on  $t$ ) as follows. Let  $\phi_i$  be a smooth function defined on the  $i$ th triangle  $\tau_i$  of the triangulation defining  $S_h$  satisfying

$$\phi_i(x) = \begin{cases} 0 & \text{if } x \text{ is not in } \tau_i, \\ 1 & \text{if } x \text{ is in } \tau_i \text{ and the distance from } x \text{ to } \partial\tau_i \text{ is } \geq t. \end{cases}$$

In addition, assume that  $\phi_i$  satisfies

$$\phi_i'(x) \leq Ct^{-1} \quad \text{for all } x \in \tau_i.$$

We then define  $v \in H^1(\Omega)$  by

$$v = \sum_i \phi_i u.$$

By the quasi-uniformity of the triangulation and the smoothness of  $\phi_i$ ,

- (1)  $\|u - v\|_{L^2(\tau_i)}^2 \leq Ch t \|u\|_{L^\infty(\tau_i)}^2$  and  
 (2)  $\|v\|_{H^1(\tau_i)}^2 \leq C(\|u\|_{H^1(\tau_i)}^2 + ht^{-1} \|u\|_{L^\infty(\tau_i)}^2)$ .

We clearly have

$$\|u\|_{L^\infty(\tau_i)}^2 \leq ch^{-2} \|u\|_{L^2(\tau_i)}^2$$

and

$$\|u\|_{H^1(\tau_i)}^2 \leq ch^{-2} \|u\|_{L^2(\tau_i)}^2.$$

Consequently,

$$K(u, t)^2 \leq cth^{-1} \|u\|_{L^2(\Omega)}^2.$$

It follows that

$$(10.4) \quad \int_0^h K(u, t)^2 t^{-2\beta-1} dt \leq Ch^{-2\beta} \|u\|_{L^2(\Omega)}^2.$$

Combining (10.3) and (10.4) completes the proof of (10.1).

#### 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 C. C. Douglas, *Sharp estimates for multigrid rates of convergence with general smoothing and acceleration*, SIAM J. Numer. Anal. **22** (1985), 617–633.
3. R. E. Bank and T. Dupont, *An optimal order process for solving finite element equations*, Math. Comp. **36** (1981), 35–51.
4. 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.
5. J. H. Bramble and J. E. Pasciak, *New convergence estimates for multigrid algorithms*, Math. Comp. **49** (1987), 311–329.
6. 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.
7. J. H. Bramble and J. E. Pasciak, *A preconditioning technique for indefinite systems resulting from mixed approximations of elliptic problems*, Math. Comp. **50** (1988), 1–18.
8. A. Brandt, *Multi-level adaptive solutions to boundary-value problems*, Math. Comp. **31** (1977), 333–390.
9. S. C. Brenner, *An optimal-order multigrid method for P1 nonconforming finite elements*, Math. Comp. **52** (1989), 1–16.
10. F. Brezzi, J. Douglas, Jr., and L. D. Marini, *Two families of mixed finite elements for second order elliptic problems*, Numer. Math. **47** (1985), 217–235.
11. P. L. Butzer and H. Berens, *Semi-groups of operators and approximation*, Springer-Verlag, Berlin and New York, 1967.
12. P. G. Ciarlet, *The finite element method for elliptic problems*, North-Holland, New York, 1978.

13. C. C. Douglas, *Multi-grid algorithms with applications to elliptic boundary-value problems*, SIAM J. Numer. Anal. **21** (1984), 236–254.
14. T. Dupont and R. Scott, *Polynomial approximation of functions in Sobolev spaces*, Math. Comp. **34** (1980), 441–463.
15. R. Glowinski and M. F. Wheeler, *Domain decomposition and mixed methods for elliptic problems*, Proc. 1st Internat. Conf. on Domain Decomposition Methods, SIAM, Philadelphia, PA, 1988, pp. 144–172.
16. W. Hackbusch, *Multi-grid methods and applications*, Springer-Verlag, New York, 1985.
17. R. B. Kellogg, *Interpolation between subspaces of a Hilbert space*, Technical Note BN-719, Univ. of Maryland, Inst. of Fluid Dynamics and Appl. Math., 1971.
18. S. G. Krein and Y. I. Petunin, *Scales of Banach spaces*, Russian Math. Surveys **21** (1966), 85–160.
19. J. L. Lions and E. Magenes, *Problèmes aux limites non homogènes et applications*, vol. 1, Dunod, Paris, 1968.
20. J. Mandel, S. F. McCormick, and J. Ruge, *An algebraic theory for multigrid methods for variational problems*, preprint.
21. J. Mandel, S. McCormick, and R. Bank, *Variational multigrid theory*, Multigrid Methods, (S. McCormick, ed.), SIAM, Philadelphia, PA, pp. 131–178.
22. S. F. McCormick, *Multigrid methods for variational problems: Further results*, SIAM J. Numer. Anal. **21** (1984), 255–263.
23. S. F. McCormick, *Multigrid methods for variational problems: General theory for the V-cycle*, SIAM J. Numer. Anal. **22** (1985), 634–643.
24. F. A. Milner, *Mixed finite element methods for quasilinear second-order elliptic problems*, Math. Comp. **44** (1985), 303–320.
25. J. Nečas, *Les méthodes directes en théorie des équations elliptiques*, Academia, Prague, 1967.
26. P. A. Raviart and J. M. Thomas, *A mixed finite element method for 2-nd order elliptic problems*, Mathematical Aspects of Finite Element Methods, Lecture Notes in Math., vol. 606 (I. Galligani and E. Magenes, eds.), Springer-Verlag, New York, 1977, pp. 292–315.
27. F. Riesz and B. Sz.-Nagy, *Functional analysis*, Ungar, New York, 1955.
28. E. M. Stein, *Singular integrals and differentiability properties of functions*, Princeton Univ. Press, Princeton, NJ, 1970.
29. R. Temam, *Navier-Stokes equations*, North-Holland, New York, 1977.
30. R. Verfürth, *A multilevel algorithm for mixed problems*, SIAM J. Numer. Anal. **21** (1984), 264–271.
31. J. Xu, *Theory of multilevel methods*, Thesis, Cornell University, Ithaca, NY, 1989.
32. S. Zhang, *Multi-level iterative techniques*, Thesis, Math. Res. Rep. 88020, Penn. State Univ., 1988.

DEPARTMENT OF MATHEMATICS, CORNELL UNIVERSITY, ITHACA, NEW YORK 14853  
*E-mail address*: [bramble@mathvax.msi.cornell.edu](mailto:bramble@mathvax.msi.cornell.edu)

BROOKHAVEN NATIONAL LABORATORY, UPTON, NEW YORK 11973  
*E-mail address*: [pasciak@bnl.gov](mailto:pasciak@bnl.gov)

DEPARTMENT OF MATHEMATICS, PENNSYLVANIA STATE UNIVERSITY, UNIVERSITY PARK, PENNSYLVANIA 16802  
*E-mail address*: [xu@rayleigh.psu.edu](mailto:xu@rayleigh.psu.edu)