

The Analysis of Multigrid Algorithms for Nonsymmetric and Indefinite Elliptic Problems*

By James H. Bramble, Joseph E. Pasciak, and Jinchao Xu

Abstract. We prove some new estimates for the convergence of multigrid algorithms applied to nonsymmetric and indefinite elliptic boundary value problems. We provide results for the so-called 'symmetric' multigrid schemes. We show that for the variable \mathcal{V} -cycle and the \mathcal{W} -cycle schemes, multigrid algorithms with any amount of smoothing on the finest grid converge at a rate that is independent of the number of levels or unknowns, provided that the initial grid is sufficiently fine. We show that the \mathcal{V} -cycle algorithm also converges (under appropriate assumptions on the coarsest grid) but at a rate which may deteriorate as the number of levels increases. This deterioration for the \mathcal{V} -cycle may occur even in the case of full elliptic regularity. Finally, the results of numerical experiments are given which illustrate the convergence behavior suggested by the theory.

1. Introduction. In recent years, multigrid methods have been used extensively as tools for obtaining the solution of the discrete systems which arise in the numerical approximation of partial differential equations (cf. [6], [8]). In conjunction, there has been intensive research aimed at attaining a more thorough theoretical understanding of the multigrid technique [1]–[5], [8], [13]–[18], [21]. In this paper, we shall provide some new iterative convergence estimates for multigrid algorithms applied to nonsymmetric and indefinite problems.

The theory for the analysis of multigrid methods applied to symmetric positive definite problems is most completely developed [2], [4], [5], [13], [15], [21]. Generally, these results assume a 'regularity and approximation' hypothesis which involves a parameter $0 < \alpha \leq 1$. The results in these papers guarantee convergence rates for multigrid \mathcal{V} -cycle, the variable \mathcal{V} -cycle (cf. [5]) and the \mathcal{W} -cycle algorithms for various α . In particular, [5], [15] give iterative convergence results for the symmetric problem which are valid for any amount of smoothing and any α .

The theory for multigrid methods applied to nonsymmetric and indefinite problems is not so completely developed. Two types of algorithms are the so-called 'symmetric' and 'nonsymmetric' multigrid schemes. The nonsymmetric scheme uses a relaxation procedure based on the original equations whereas the symmetric

Received November 9, 1987.

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 non-exclusive, royalty-free license to publish or reproduce the published form of this contribution, or allow others to do so, for U.S. Government purposes. This work was also supported in part under the National Science Foundation Grant No. 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.

scheme uses a relaxation based on the symmetric positive definite system associated with the normal equations. Some results only hold under rather restrictive assumptions involving the relation between the number of smoothings m and the size of the coarsest grid h_j . For example, Bank [1] gives \mathscr{W} -cycle results for both schemes and for arbitrary α which, however, require first that m be sufficiently large, and secondly that h_j be sufficiently small (depending on m). Mandel [14] gives results for the nonsymmetric \mathscr{W} -cycle scheme and the \mathscr{V} -cycle scheme (assuming full regularity $\alpha = 1$) which are valid for any m if h_j is chosen sufficiently small (depending on m).

In this paper, we shall prove some new iterative convergence estimates for the symmetric multigrid scheme applied to nonsymmetric and indefinite problems. We give results for the \mathscr{V} -cycle, variable \mathscr{V} -cycle and \mathscr{W} -cycle algorithms for any amount of smoothing under the assumption of $\alpha > 3/4$. Our theorems for the variable \mathscr{V} -cycle and \mathscr{W} -cycle algorithms require that h_j be sufficiently small (independent of the amount of smoothing) and guarantee an iterative convergence rate which is uniformly independent of the number of levels and the mesh size on the finest grid. The assumption that h_j is sufficiently small is not very restrictive since such an assumption must be made for solvability on the coarsest grid. The results for the \mathscr{V} -cycle algorithm are somewhat weaker. We show that the \mathscr{V} -cycle converges if h_j is small enough (depending on the number of levels and α), at a rate which deteriorates as more and more levels are used. Even in the case $\alpha = 1$, the \mathscr{V} -cycle convergence estimates deteriorate like $1 - c/\ln(h^{-1})$.

We derive our iterative convergence estimates for multigrid algorithms in an abstract setting. The use of this abstract approach more clearly identifies the relevant hypotheses.

The outline of the remainder of the paper is as follows. In Section 2 we describe the abstract framework to be used in the paper. The assumptions used in our analysis and some preliminary definitions are also given there. Section 3 shows how this framework can be applied in the case of nonsymmetric and indefinite uniformly elliptic second-order boundary value problems. Section 4 defines the multigrid operator and provides a basic recurrence relation used in our subsequent analysis. The convergence estimates given in this paper are based on three technical lemmas. In Section 5 we prove our multigrid theorems, assuming the technical lemmas. Section 6 provides the proof of the lemmas and represents the core of our analysis. Finally, the results of numerical experiments illustrating the earlier derived theory are given in Section 7.

Throughout this paper, c and C , with or without subscript will denote a generic positive constant which may take on different values in different places. These constants will always be independent of the mesh parameters.

2. Abstract Framework and Assumptions. In this section, we first give an abstract framework for our nonsymmetric multigrid application. This abstract presentation more clearly identifies the relevant hypotheses used in the iterative convergence analysis to be developed. We then list the assumptions required for the multigrid analysis presented in later sections. To keep the paper from becoming too abstract, we show how a model application to a second-order problem fits into this framework in the next section.

We start with a Hilbert scale (cf. [11]) of spaces $\{H^\gamma\}$ for $\gamma \in [0, 2]$. The norm on H^γ will be denoted by $\|\cdot\|_{H^\gamma}$. We assume that $H^s \subset H^t$ whenever $t < s$. The largest space (i.e., $\gamma = 0$) will be denoted H with norm $\|\cdot\|_H$ and inner product (\cdot, \cdot) . The space H^γ is assumed to be compactly contained in H^δ whenever $\gamma > \delta$. Let \mathcal{M} be a closed subspace of H^1 . The spaces H^s for $-1 \leq s < 0$ are defined by duality and with norm given by

$$\|v\|_{H^s} \equiv \sup_{\phi \in \mathcal{M}} \frac{(v, \phi)}{\|\phi\|_{H^{-s}}}.$$

Assume that we are given a nested sequence of ‘approximation’ subspaces

$$\mathcal{M}_1 \subset \mathcal{M}_2 \subset \dots \subset \mathcal{M}_J \subset \mathcal{M}.$$

In addition, let $\hat{A}(\cdot, \cdot)$ be a positive definite symmetric quadratic form on $\mathcal{M} \times \mathcal{M}$ satisfying

$$(2.1) \quad c\|v\|_{H^1}^2 \leq \hat{A}(v, v) \leq C\|v\|_{H^1}^2 \quad \text{for all } v \in \mathcal{M}$$

and $D(\cdot, \cdot)$ be a quadratic form on $\mathcal{M} \times \mathcal{M}$. We shall be interested in approximating the solution of

$$(2.2) \quad A(u, \phi) \equiv \hat{A}(u, \phi) + D(u, \phi) = (f, \phi) \quad \text{for all } \phi \in \mathcal{M},$$

for a given function $f \in H$. We shall assume that (2.2) is uniquely solvable for any $f \in H$.

We will be interested in applying multigrid procedures to develop a rapidly converging iterative algorithm for the solution of the Galerkin approximation of (2.2) in the subspace \mathcal{M}_J . Specifically, we seek the function $U \in \mathcal{M}_J$ which satisfies

$$(2.3) \quad A(U, \chi) = (f, \chi) \quad \text{for all } \chi \in \mathcal{M}_J.$$

Our multigrid algorithms will require the use of discrete inner products $(\cdot, \cdot)_k$ on $\mathcal{M}_k \times \mathcal{M}_k$ for $k = 1, \dots, J$. The corresponding norm will be denoted $\|\cdot\|_k$. In the algorithms, these inner products are used instead of (\cdot, \cdot) to avoid the inversion of Gram matrices. This means that the problem of computing $W \in \mathcal{M}_k$ satisfying

$$(2.4) \quad (W, \theta)_k = F(\theta) \quad \text{for all } \theta \in \mathcal{M}_k$$

for a given linear functional F should be simple.

We next list the assumptions required for our multigrid analysis.

(A.1): The first assumption involves elliptic regularity for the forms $A(\cdot, \cdot)$ and $\hat{A}(\cdot, \cdot)$. We assume that solutions u of (2.2) and the corresponding equation

$$\hat{A}(u, \theta) = (f, \theta) \quad \text{for all } \theta \in \mathcal{M}$$

satisfy

$$(2.5) \quad \|u\|_{H^{1+\alpha}} \leq c\|f\|_{H^{\alpha-1}}$$

for some $\alpha \in (3/4, 1]$ independent of f .

(A.2): We assume first that D satisfies

$$(2.6) \quad |D(v, w)| \leq C\|v\|_{H^1}\|w\|_H \quad \text{for all } v, w \in \mathcal{M}.$$

It is an immediate consequence of (2.6) that the operator $D : \mathcal{M} \mapsto H$ defined by

$$(Dv, \theta) = D(v, \theta) \quad \text{for all } \theta \in H$$

is well defined and satisfies

$$(2.7) \quad \|Dv\|_H \leq C\|v\|_{H^1}.$$

We further assume that D maps $H^{1+\alpha}$ into H^α , i.e.,

$$(2.8) \quad \|Dv\|_{H^\alpha} \leq C\|v\|_{H^{1+\alpha}}.$$

Let $D^* : H \mapsto H^{-1}$ be defined by

$$(D^*w, \phi) = (w, D\phi).$$

We assume that D^* is a bounded operator from H^1 into $H^{-1/2-\varepsilon}$ for any positive ε .

(A.3): We require approximation properties for the subspaces $\{\mathcal{M}_k\}$. These are given in terms of a parameter h_k which satisfies

$$c\kappa^k \leq h_k \leq C\kappa^k$$

for constants c, C and $\kappa < 1$ independent of k . We assume that for v in H^s and $s \in [1, 1 + \alpha]$, there exists $\chi \in \mathcal{M}_k$ such that

$$\|v - \chi\|_H + h_k\|v - \chi\|_{H^1} \leq Ch_k^s\|v\|_{H^s}.$$

(A.4): We require that the inverse inequality,

$$\|W\|_{H^\beta} \leq Ch_k^{\gamma-\beta}\|W\|_{H^\gamma} \quad \text{for all } W \in \mathcal{M}_k$$

holds for all $\beta > \gamma$ with $\beta, \gamma \in [0, 1 + \alpha]$.

(A.5): We require first that the discrete inner product $(\cdot, \cdot)_k$ be equivalent to (\cdot, \cdot) on \mathcal{M}_k , i.e.,

$$(2.9) \quad c\|\chi\|_H \leq \|\chi\|_k \leq C\|\chi\|_H.$$

In addition, we assume that the discrete inner products accurately approximate the inner product on H in the sense that

$$(2.10) \quad |(\psi, \chi) - (\psi, \chi)_k| \leq Ch_k\|\psi\|_{H^1}\|\chi\|_k \quad \text{for all } \psi, \chi \in \mathcal{M}_k.$$

We next introduce some discrete operators which play a fundamental role both in the analysis and the algorithms to be considered in this paper:

(O.1): The operator $A_k : \mathcal{M}_k \mapsto \mathcal{M}_k$ is defined by

$$(A_k W, \theta)_k = A(W, \theta) \quad \text{for all } \theta \in \mathcal{M}_k.$$

(O.2): The operator $P_k : \mathcal{M} \mapsto \mathcal{M}_k$ is defined by

$$(2.11) \quad A(P_k w, \theta) = A(w, \theta) \quad \text{for all } \theta \in \mathcal{M}_k.$$

(O.3): The operator $\hat{A}_k : \mathcal{M}_k \mapsto \mathcal{M}_k$ is defined by

$$(\hat{A}_k W, \theta)_k = \hat{A}(W, \theta) \quad \text{for all } \theta \in \mathcal{M}_k.$$

(O.4): The operator $\hat{P}_k : \mathcal{M} \mapsto \mathcal{M}_k$ is defined by

$$\hat{A}(\hat{P}_k w, \theta) = \hat{A}(w, \theta) \quad \text{for all } \theta \in \mathcal{M}_k.$$

(O.5): The operator $D_k : \mathcal{M}_k \mapsto \mathcal{M}_k$ is defined by

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

(O.6): The operator $I_k : \mathcal{M}_{k+1} \mapsto \mathcal{M}_k$ is defined by

$$(I_k W, \theta)_k = (W, \theta)_{k+1} \quad \text{for all } \theta \in \mathcal{M}_k.$$

(O.7): The operator $P_k^0 : H \mapsto \mathcal{M}_k$ is defined by

$$(P_k^0 v, \theta)_k = (v, \theta) \quad \text{for all } \theta \in \mathcal{M}_k.$$

All of the above operators except possibly P_k are clearly well defined. We shall assume, however, that h_k is less than some positive constant ν with ν chosen small enough so that the above assumptions imply a unique solution to (2.11) (cf. [20]). This also implies that A_k is invertible.

We note that (2.3) is equivalent to

$$A_k U = P_k^0 f.$$

We define two scales of norms on \mathcal{M}_k which we shall use in our analysis. The operator \hat{A}_k is symmetric and positive definite on \mathcal{M}_k in the $(\cdot, \cdot)_k$ inner product. We define the scale of norms $\{\|\cdot\|_{k,s}\}$ for any real s by

$$\|W\|_{k,s} = \|\hat{A}_k^{s/2} W\|_k \quad \text{for all } W \in \mathcal{M}_k.$$

Similarly, the operator $A_k^* A_k$ is also symmetric and positive definite on \mathcal{M}_k (here, $*$ denotes the adjoint with respect to $(\cdot, \cdot)_k$). We define the scale of norms $\{\|\cdot\|_{k,s}\}$ for any real s by

$$\|W\|_{k,s} = ((A_k^* A_k)^{s/2} W, W)_k^{1/2} \quad \text{for all } W \in \mathcal{M}_k.$$

Let $L_k = (A_k^* A_k)^{1/2}$; then clearly

$$\|W\|_{k,s} = \|L_k^{s/2} W\|_k \quad \text{for all } W \in \mathcal{M}_k.$$

We will often consider the norms of operators from a space into itself. If $T : S \mapsto S$ is an operator on a generic space S with norm $\|\cdot\|$, then the norm of T will be denoted by $\|T\|$ and is given by

$$\|T\| = \sup_{\phi \in S} \frac{\|T\phi\|}{\|\phi\|}.$$

3. An Application to the Second-Order Problem. We consider a model second-order problem in this section and show that the hypotheses of Section 2 are satisfied. This application involves a finite element approximation of a nonsymmetric and indefinite elliptic problem in N -dimensional Euclidean space.

Let Ω be a domain in R^N . The spaces $H^s = H^s(\Omega)$ will be the Sobolev spaces of order s on Ω [12], [19]. We shall be interested in approximating the solution of the problem

$$(3.1) \quad \mathcal{L}u = f \quad \text{in } \Omega,$$

$$(3.2) \quad \frac{\partial u}{\partial \nu} = 0 \quad \text{on } \partial\Omega,$$

or

$$(3.3) \quad u = 0 \quad \text{on } \partial\Omega,$$

where

$$\mathcal{L}u = - \sum_{i,j=1}^N \frac{\partial}{\partial x_i} a_{ij}(x) \frac{\partial u}{\partial x_j} + \sum_{i=1}^N b_i(x) \frac{\partial u}{\partial x_i} + c(x)u$$

and $\frac{\partial}{\partial \nu}$ denotes the outward co-normal derivative on $\partial\Omega$.

We assume that the matrix $\{a_{ij}(x)\}$ is symmetric and uniformly positive definite.

Under appropriate smoothness assumptions for the domain Ω and coefficients defining \mathcal{L} , it is possible to prove that the solutions of (3.1)–(3.3) satisfy estimates of the form (2.5) [7], [10]. For two-dimensional polygonal domains, with coefficients in $C^1(\Omega)$, (2.5) holds for $\alpha > 3/4$, if all interior angles of the polygon are bounded by $4\pi/3$. For more general applications, we implicitly assume the appropriate hypotheses so that (2.5) holds for $\alpha > 3/4$.

The space \mathcal{M} is a subset of $H^1(\Omega)$ satisfying appropriate boundary conditions. In the case of boundary condition (3.3), \mathcal{M} is the completion of $C_0^\infty(\Omega)$ in the $H^1(\Omega)$ -norm. For boundary condition (3.2), $\mathcal{M} = H^1(\Omega)$ unless $c(x) = 0$ for all x , in which case \mathcal{M} consists of those functions in $H^1(\Omega)$ which are orthogonal to constants.

A weak formulation of (3.1)–(3.3) is: Find $u \in \mathcal{M}$ such that

$$(3.4) \quad A(u, v) = (f, v) \quad \text{for all } v \in \mathcal{M},$$

where (\cdot, \cdot) is the usual $L^2(\Omega)$ inner product and

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

Note that, in general, $A(\cdot, \cdot)$ is nonsymmetric and indefinite. We assume that (3.4) has a unique solution.

We define $\hat{A}(\cdot, \cdot)$ by

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

Then, obviously

$$D(u, v) = \sum_{i=1}^N \int_{\Omega} b_i \frac{\partial u}{\partial x_i} v dx + \int_{\Omega} (c - 1)uv dx.$$

We next check assumption (A.2). Inequality (2.6) follows immediately from the Schwarz inequality. The operator D is given by

$$Du = \sum_{i=1}^N b_i \frac{\partial u}{\partial x_i} + (c - 1)u,$$

and hence (2.8) clearly holds. Finally, we note that for $w \in H^1(\Omega)$ and $\phi \in \mathcal{M}$,

$$(3.5) \quad (D^*w, \phi) = (w, D\phi) = -(Dw, \phi) + \sum_{i=1}^N \left\{ \int_{\partial\Omega} b_i n_i w \phi ds - \left(\frac{\partial b_i}{\partial x_i} w, \phi \right) \right\},$$

where n_i is the component of the outward normal in the i th direction. We assume that b_i is in $C^1(\Omega)$ and that c is in $L^\infty(\Omega)$. The boundary term in (3.5) vanishes in the case of boundary conditions (3.3) and hence $D^* : H^1(\Omega) \mapsto L^2(\partial\Omega)$ in this case. In the case of boundary conditions (3.2), by a well-known trace inequality,

$$\left| \sum_{i=1}^N \int_{\partial\Omega} b_i n_i w \phi \, ds \right| \leq C \|w\|_{H^{1/2+\varepsilon}(\Omega)} \|\phi\|_{H^{1/2+\varepsilon}(\Omega)},$$

from which it follows that $D^* : H^1(\Omega) \mapsto H^{-1/2-\varepsilon}(\Omega)$. Thus (A.2) holds for either application.

We next consider the finite element approximation subspaces. For simplicity, we shall only describe a piecewise linear application in two dimensions. The application to higher-dimensional problems and more general approximation subspaces is straightforward. We write $\Omega = \bigcup \tau_i^1$, where $\tau_1 \equiv \{\tau_i^1\}$ is a collection of triangles with mutually disjoint interiors. We assume that these triangles are of quasi-uniform size h_1 . This means that there are positive constants c and C such that the diameter of every triangle is bounded by Ch_1 and each triangle contains a circle of radius ch_1 . We define a sequence of triangulations by induction. Assume that the triangulation $\tau_{k-1} = \{\tau_i^{k-1}\}$ has been defined. The triangles of τ_k are formed by connecting the midpoints of the edges of the triangles in τ_{k-1} . Thus, each triangle in τ_{k-1} gives rise to four triangles in τ_k .

The approximation subspace \mathcal{M}_k consists of functions which are continuous and piecewise linear with respect to the triangulation τ_k . In the case of Dirichlet boundary conditions, we additionally require that the functions in τ_k vanish on $\partial\Omega$. In the case of boundary conditions (3.2) and $c(x) = 0$, we also require that the functions in τ_k have zero mean value. For these spaces, $h_k = 2^{-k+1}h_1$ and classical techniques in the theory of finite elements imply that (A.3) and (A.4) hold.

We finally define the discrete inner products. Let x_{ij}^k , $j = 1, 2, 3$, denote the vertices of the i th triangle of the k th grid. Define

$$(3.6) \quad (\phi, \chi)_k = 1/3 \sum_i |\tau_i^k| \sum_{j=1}^3 \phi(x_{ij}^k) \chi(x_{ij}^k).$$

Here $|\tau_i^k|$ denotes the area of the triangle τ_i^k . It is not difficult to show that (A.5) holds for this inner product. Note that (3.6) can be rewritten

$$(3.7) \quad (\phi, \chi)_k = \sum_i \omega_i^k \phi(y_i^k) \chi(y_i^k),$$

where $\{y_i^k\}$ are the nodes of the k th grid and ω_i^k is an appropriate weight function. Note that (3.7) implies that the solution of problems of the form (2.4) reduces to division by the weights $\{\omega_i^k\}$.

4. Multigrid Algorithms. We will define the multigrid algorithms in this section and develop certain recurrence relations which will be used in the iterative convergence analysis given later in the paper. The multigrid algorithm defines a linear operator B_k on \mathcal{M}_k which is an approximate inverse for A_k . We will consider the so-called ‘symmetric multigrid scheme’. Here ‘symmetric’ refers to the fact that the relaxation process used results from an iterative scheme for the symmetric operator $A_k^* A_k$.

We define the operator $B_k : \mathcal{M}_k \mapsto \mathcal{M}_k$ by induction on k . As we shall see in later sections, for stability, the coarsest grid in the multigrid process must not be too coarse. To this end, we shall define our algorithms starting from the intermediate grid level j , $1 \leq j < J$. In this algorithm, we assume that the operator B_j equals A_j^{-1} , although some results still hold when B_j is defined differently (see Remark 5.2).

The Multigrid Algorithm.

Set $B_j = A_j^{-1}$. Assume that $B_{k-1} : \mathcal{M}_{k-1} \mapsto \mathcal{M}_{k-1}$ has been defined and define $B_k g$ for $g \in \mathcal{M}_k$ and $k = j + 1, \dots, J$ as follows:

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

$$(4.1) \quad x^l = x^{l-1} + \mu_k^{-2} A_k^* (g - A_k x^{l-1}),$$

where μ_k is the largest eigenvalue of $L_k = (A_k^* A_k)^{1/2}$.

- (3) Define $B_k g = x^{m(k)} + q^p$, where q^i , for $i = 1, 2, \dots, p$, is defined by

$$(4.2) \quad q^i = q^{i-1} + B_{k-1} [I_{k-1} (g - A_k x^{m(k)}) - A_{k-1} q^{i-1}].$$

The heuristic motivation for the above algorithm is as follows. Step (2) is a smoothing process and is designed to reduce the high-frequency components of the error. The low-frequency components of the error are then reduced by the coarser grid correction (3).

Remark 4.1. We have used μ_k^2 in (4.1) for convenience. In actual algorithms, any reasonable bound for the largest eigenvalue of the system $A_k^* A_k$ can be used.

Let $g = A_k x$ and $K_k = I - \mu_k^{-2} A_k^* A_k$. Clearly

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

It is straightforward to check that q^p satisfies

$$q^p = (I - (I - B_{k-1} A_{k-1})^p) P_{k-1} (x - x^{m(k)}).$$

Combining the above equalities gives

$$(4.3) \quad I - B_k A_k = [(I - P_{k-1}) + (I - B_{k-1} A_{k-1})^p P_{k-1}] K_k^{m(k)}.$$

The relation (4.3) provides a fundamental identity for the analysis of the multigrid algorithm.

The goal of this paper is to prove inequalities of the form

$$(4.4) \quad \|I - B_k A_k\|_{k,1}^2 \leq \delta_k.$$

Such inequalities immediately imply that the linear iteration

$$U^{n+1} = U^n + B_k (F - A_k U^n)$$

converges to the solution U of

$$A_k U = F$$

with a rate of $\sqrt{\delta_k}$ per step in the norm $\|\cdot\|_{k,1}$. Equality (4.3) gives a way of relating the reduction δ_k to that of the $(k-1)$ -grid and hence provides a key ingredient for a mathematical induction argument.

5. The Convergence Theorems and their Proofs. We give our convergence results for multigrid algorithms in this section. We first give results for the variable \mathcal{V} -cycle. Next, we consider the \mathcal{V} -cycle with constant $m(k) = m$. Finally, we consider the \mathcal{W} -cycle algorithms. The proofs of these theorems depend on three lemmas. These lemmas are central to the analysis of the paper and will be proved in the next section. In this section, we prove our multigrid theorems, assuming the lemmas.

We start by stating the lemmas. The first lemma gives a so-called ‘regularity and approximation’ estimate for the projection operator P_k .

LEMMA 5.1. *If h_j is sufficiently small, there exists a positive constant C not depending on k such that*

$$\|(I - P_{k-1})v\|_1^2 \leq C(\mu_k^{-1} \|L_k v\|_k^2)^\alpha (L_k v, v)_k^{1-\alpha} \quad \text{for all } v \in \mathcal{M}_k.$$

The next two lemmas represent an essential part of the analysis of this paper. Their proof uses the Dunford-Taylor integral formula for operators and is given in the next section.

LEMMA 5.2. *If h_j is sufficiently small, there exists a positive constant C not depending on k such that for all $v \in \mathcal{M}_k, \chi \in \mathcal{M}_{k-1}$,*

$$(5.1) \quad (L_k(I - P_{k-1})v, \chi)_k \leq Ch_k^{\alpha-1/2-\varepsilon} \|(I - P_{k-1})v\|_{k,1} \|\chi\|_{k,1}$$

holds for any positive ε .

LEMMA 5.3. *If h_j is sufficiently small, there exists a positive constant C not depending on k such that for all $\chi \in \mathcal{M}_{k-1}$,*

$$\left| \|\chi\|_{k,1}^2 - \|\chi\|_{k-1,1}^2 \right| \leq Ch_k^{4\alpha-3} (1 + |\ln h_k|) \|\chi\|_{k,1}^2.$$

We can now state and prove the convergence theorem for the variable \mathcal{V} -cycle algorithm.

THEOREM 1. *Let $p = 1$ and assume that $m(k)$ satisfies*

$$(5.2) \quad \beta_0 m(k) \leq m(k-1) \leq \beta_1 m(k)$$

where β_0 and β_1 are constants greater than one and independent of k for $k = j+2, \dots, J$. Let γ be positive and less than $\min(\alpha - 1/2, 4\alpha - 3)$. Then there exist positive constants M and ν not depending on k such that when $h_j \leq \nu$, (4.4) holds with

$$(5.3) \quad \delta_k = \frac{M}{M + m(k)^{\alpha/2}}$$

for $k = j+1, \dots, J$.

Proof. We will prove the theorem by induction. For the purpose of this proof, let $m(j) = \beta_0 m(j+1)$ (note that $m(j)$ does not appear in the definition of the multigrid process). Clearly, (4.4) holds for $k = j$ with δ_k given by (5.3). Let $k \in \{j+1, \dots, J\}$ and assume that (4.4) holds for $k-1$ with δ_{k-1} given by (5.3). It follows from the recursive relation (4.3) that

$$\begin{aligned} \|(I - B_k A_k)v\|_{k,1}^2 &= \|(I - P_{k-1})\tilde{v}\|_{k,1}^2 + \|(I - B_{k-1}A_{k-1})P_{k-1}\tilde{v}\|_{k,1}^2 \\ &\quad + 2(L_k(I - P_{k-1})\tilde{v}, (I - B_{k-1}A_{k-1})P_{k-1}\tilde{v})_k, \end{aligned}$$

where $\tilde{v} = K_k^{m(k)}v$. Applying Lemma 5.2 gives

$$\begin{aligned} & \|(I - B_k A_k)v\|_{k,1}^2 \\ & \leq (1 + Ch_k^{\alpha-1/2-\varepsilon}) (\|(I - P_{k-1})\tilde{v}\|_{k,1}^2 + \|(I - B_{k-1}A_{k-1})P_{k-1}\tilde{v}\|_{k,1}^2). \end{aligned}$$

Using Lemma 5.3 and the induction hypothesis, we deduce that

$$\begin{aligned} \|(I - B_{k-1}A_{k-1})P_{k-1}\tilde{v}\|_{k,1}^2 & \leq (1 + Ch_k^\gamma)\|(I - B_{k-1}A_{k-1})P_{k-1}\tilde{v}\|_{k-1,1}^2 \\ & \leq \delta_{k-1}(1 + Ch_k^\gamma)\|P_{k-1}\tilde{v}\|_{k-1,1}^2 \\ & \leq \delta_{k-1}(1 + Ch_k^\gamma)\|P_{k-1}\tilde{v}\|_{k,1}^2 \end{aligned}$$

holds for any fixed γ less than $4\alpha - 3$. We remind the reader that here and throughout the paper, C denotes a generic positive constant which may take on different values from line to line. It follows from Lemma 5.2 that

$$\begin{aligned} \|\tilde{v}\|_{k,1}^2 & = \|P_{k-1}\tilde{v}\|_{k,1}^2 + \|(I - P_{k-1})\tilde{v}\|_{k,1}^2 + 2(L_k(I - P_{k-1})\tilde{v}, P_{k-1}\tilde{v})_k \\ & \geq (1 - Ch_k^{\alpha-1/2-\varepsilon})(\|P_{k-1}\tilde{v}\|_{k,1}^2 + \|(I - P_{k-1})\tilde{v}\|_{k,1}^2) \end{aligned}$$

and thus for ν sufficiently small

$$\|P_{k-1}\tilde{v}\|_{k,1}^2 \leq (1 + Ch_k^{\alpha-1/2-\varepsilon})\|\tilde{v}\|_{k,1}^2 - \|(I - P_{k-1})\tilde{v}\|_{k,1}^2.$$

Requiring, in addition, that $\gamma < \alpha - 1/2$ and combining the above inequalities gives

$$\|(I - B_k A_k)v\|_{k,1}^2 \leq (1 + Ch_j^\gamma) \{ (1 - \delta_{k-1})\|(I - P_{k-1})\tilde{v}\|_{k,1}^2 + \delta_{k-1}\|\tilde{v}\|_{k,1}^2 \}.$$

By Lemma 5.1, the Schwarz inequality, and a generalized arithmetic geometric mean inequality,

$$\begin{aligned} \|(I - P_{k-1})\tilde{v}\|_{k,1}^2 & \leq C (\mu_k^{-1}(L_k^2\tilde{v}, \tilde{v})_k)^\alpha (L_k\tilde{v}, \tilde{v})_k^{1-\alpha} \\ & \leq C (\mu_k^{-2}(L_k^3\tilde{v}, \tilde{v})_k)^{\alpha/2} (L_k\tilde{v}, \tilde{v})_k^{1-\alpha/2} \\ & \leq C \left\{ \eta_k \mu_k^{-2}(L_k^3\tilde{v}, \tilde{v})_k + \eta_k^{-\alpha/(2-\alpha)}(L_k\tilde{v}, \tilde{v})_k \right\} \end{aligned}$$

holds for any positive constant η_k . Using the definition of K_k and the fact that its eigenvalues are in the interval $[0, 1)$ gives

$$\begin{aligned} \mu_k^{-2}(L_k^3\tilde{v}, \tilde{v})_k & = (L_k(I - K_k)K_k^{m(k)}v, K_k^{m(k)}v)_k \\ & \leq (2m(k))^{-1} \sum_{l=0}^{2m(k)-1} (L_k(I - K_k)K_k^l v, v)_k \\ & = (2m(k))^{-1} (L_k(I - K_k^{2m(k)})v, v)_k. \end{aligned}$$

Combining the above inequalities gives

$$\begin{aligned} & \|(I - B_k A_k)v\|_{k,1}^2 \\ (5.4) \quad & \leq (1 + C_1 h_k^\gamma) \left\{ C_0(1 - \delta_{k-1})\eta_k m(k)^{-1} (L_k(I - K_k^{2m(k)})v, v)_k \right. \\ & \quad \left. + [C_0(1 - \delta_{k-1})\eta_k^{-\alpha/(2-\alpha)} + \delta_{k-1}] (L_k K_k^{2m(k)}v, v)_k \right\}. \end{aligned}$$

Setting $C_2 = C_0(1 + C_1)$, we see that the theorem will follow if we can choose η_k , h_j and M such that

$$(5.5) \quad C_2(1 - \delta_{k-1})\eta_k m(k)^{-1} \leq \delta_k$$

and

$$(5.6) \quad C_2(1 - \delta_{k-1})\eta_k^{-\alpha/(2-\alpha)} + C_1h_j^\gamma\delta_{k-1} \leq \delta_k - \delta_{k-1}.$$

We choose η_k by

$$(5.7) \quad C_2(1 - \delta_{k-1})\eta_k m(k)^{-1} = \delta_{k-1},$$

from which (5.5) immediately follows. Solving for η_k in (5.7) and using this result in (5.6) implies that it is sufficient to choose M and h_j so that

$$(5.8) \quad C_3(1 - \delta_{k-1})^{2/(2-\alpha)}m(k)^{-\alpha/(2-\alpha)} + C_1h_j^\gamma\delta_{k-1}^{2/(2-\alpha)} \leq (\delta_k - \delta_{k-1})\delta_{k-1}^{\alpha/(2-\alpha)}.$$

Let $\mathcal{D}(k) \equiv M + m(k)^{\alpha/2}$ and $\beta_k = m(k-1)/m(k) \in [\beta_0, \beta_1]$; then

$$(5.9) \quad 1 - \delta_{k-1} = \frac{(\beta_k m(k))^{\alpha/2}}{\mathcal{D}(k-1)}.$$

A direct computation using (5.9) and the identity $\delta_k = M/\mathcal{D}(k)$ shows that (5.8) is equivalent to

$$(5.10) \quad C_3\beta_k^{\alpha/(2-\alpha)}M^{-\alpha/(2-\alpha)} + C_1h_j^\gamma M \leq \frac{Mm(k)^{\alpha/2}}{\mathcal{D}(k)}(\beta_k^{\alpha/2} - 1).$$

Note that if $M \geq 1$ then

$$C_4 \equiv (\beta_0^{\alpha/2} - 1)/2 \leq \frac{Mm(k)^{\alpha/2}}{M + m(k)^{\alpha/2}}(\beta_k^{\alpha/2} - 1),$$

hence it suffices to have

$$C_5M^{-\alpha/(2-\alpha)} + C_1h_j^\gamma M \leq C_4,$$

where $C_5 = C_3\beta_1^{\alpha/(2-\alpha)}$. Thus, taking $M \geq 1$ large enough so that

$$C_5M^{-\alpha/(2-\alpha)} \leq C_4/2$$

and

$$h_j \leq \nu \leq C_4^{1/\gamma}(2C_1M)^{-1/\gamma}$$

completes the proof of the theorem.

We next prove a theorem for the standard \mathcal{V} -cycle algorithm.

THEOREM 2. *Consider the \mathcal{V} -cycle algorithm ($p = 1$) with $m(k) = m$ for all k . Let γ be positive and less than $\min(\alpha - 1/2, 4\alpha - 3)$. Then there exist positive constants M , c , and ν not depending on k such that when $h_j \leq \min(\nu, c(j-1)^{-2/(\alpha\gamma)})$, (4.4) holds with*

$$(5.11) \quad \delta_k = \frac{Mk^{(2-\alpha)/\alpha}}{(Mk^{(2-\alpha)/\alpha} + m^{\alpha/2})}$$

for $k = j + 1, \dots, J$.

Remark 5.1. The theorem suggests that the \mathcal{V} -cycle may be less robust than the variable \mathcal{V} -cycle. Note that the convergence estimate for the \mathcal{V} -cycle algorithm deteriorates as k becomes larger, even in the case $\alpha = 1$. Furthermore, the theorem suggests that for stability, the coarsest grid must become finer as the number of grid levels increases.

Proof. The proof of this theorem is essentially contained in the proof of Theorem 1 and the proof of Theorem 1 of [5]. Indeed, (5.4) is valid with $m(k) = m$ and hence it suffices to choose η_k , M , h_j , and c so that (5.5) and (5.6) are satisfied. We choose η_k by (5.7) and reduce (5.5)–(5.6) to (5.8). Making similar algebraic manipulations (compare with (5.10)), we see it suffices to choose the parameters so that

$$\begin{aligned} C_3 M^{-\alpha/(2-\alpha)} + C_1 h_j^\gamma (k-1)^{2/\alpha} \\ \leq \frac{M m^{\alpha/2}}{\mathcal{D}(k)} [k^{(2-\alpha)/\alpha} - (k-1)^{(2-\alpha)/\alpha}] (k-1), \end{aligned}$$

where $\mathcal{D}(k) \equiv M k^{(2-\alpha)/\alpha} + m^{\alpha/2}$. Noting that $k \geq 2$ and $(2-\alpha)/\alpha > 0$, elementary arguments imply

$$k^{(2-\alpha)/\alpha} \leq C_6 [k^{(2-\alpha)/\alpha} - (k-1)^{(2-\alpha)/\alpha}] (k-1).$$

Thus, it suffices to prove

$$C_3 M^{-\alpha/(2-\alpha)} + C_1 h_j^\gamma M (k-1)^{2/\alpha} \leq C_6 \frac{M m^{\alpha/2} k^{(2-\alpha)/\alpha}}{\mathcal{D}(k)}.$$

We set

$$M = \left(\frac{1 + \tilde{M}}{\tilde{M}} \right)^{(2-\alpha)/\alpha} \tilde{M}$$

and define \tilde{M} by

$$C_3 \tilde{M}^{-\alpha/(2-\alpha)} = C_6/2.$$

Then

$$C_3 M^{-\alpha/(2-\alpha)} \leq \frac{C_6}{2} \frac{M}{1+M} \leq \frac{C_6}{2} \frac{M m^{\alpha/2} k^{(2-\alpha)/\alpha}}{\mathcal{D}(k)}.$$

We then set

$$c = \left(\frac{C_6 M}{2 C_1 (1+M)} \right)^{1/\gamma},$$

from which it follows that $h_j \leq c(j-1)^{-2/(\gamma\alpha)}$ implies

$$C_1 h_j^\gamma M (k-1)^{2/\alpha} \leq \frac{C_6}{2} \frac{M}{1+M} \leq \frac{C_6}{2} \frac{M m^{\alpha/2} k^{(2-\alpha)/\alpha}}{\mathcal{D}(k)}.$$

Combining the above inequalities proves the theorem.

The last theorem which we shall prove is for the \mathscr{W} -cycle algorithm.

THEOREM 3. *Consider the \mathscr{W} -cycle algorithm ($p = 2$) with $m(k) \equiv m$ for all k . Let γ be positive and less than $\min(\alpha - 1/2, 4\alpha - 3)$. Then there exist positive constants M and ν such that when $h_j \leq \nu$, (4.4) holds with*

$$(5.12) \quad \delta_k \equiv \delta = (1 + m/M)^{-\alpha/2}$$

for $k = j + 1, \dots, J$.

Proof. The proof of this theorem is essentially contained in the proof of Theorem 1 and the proof of Theorem 3 of [5]. Since the term involving $(I - B_{k-1} A_{k-1})$ appears squared in (4.3), following the proof of Theorem 1, we see that (5.4) holds

with δ_{k-1} replaced by δ^2 . We see that the theorem will follow if we can choose $\eta_k = \eta$, ν and M such that

$$(5.13) \quad C_2(1 - \delta^2)\eta m^{-1} \leq \delta$$

and

$$(5.14) \quad C_2(1 - \delta^2)\eta^{-\alpha/(2-\alpha)} + C_1 h_j^\gamma \delta^2 \leq \delta - \delta^2.$$

We choose η so that (5.13) holds with equality. Solving for η and using this result in (5.14) implies that it is sufficient to choose M and ν so that

$$(5.15) \quad C_3(1 - \delta^2)^{2/(2-\alpha)} m^{-\alpha/(2-\alpha)} + C_1 h_j^\gamma \delta^{(4-\alpha)/(2-\alpha)} \leq (1 - \delta)\delta^{2/(2-\alpha)}.$$

It is elementary to see that

$$(5.16) \quad 2^{-2/(2-\alpha)} \leq (1 - \delta)^{-\alpha/(2-\alpha)} \left(\frac{m}{M}\right)^{\alpha/(2-\alpha)} \left(\frac{\delta}{1 + \delta}\right)^{2/(2-\alpha)}$$

for δ given by (5.12). Define M by

$$M^{\alpha/(2-\alpha)} 2^{-2/(2-\alpha)} = 2C_3.$$

Then

$$C_3(1 - \delta^2)^{2/(2-\alpha)} m^{-\alpha/(2-\alpha)} \leq \frac{1}{2}(1 - \delta)\delta^{2/(2-\alpha)}.$$

Choosing

$$(5.17) \quad h_j \leq \nu \leq \left(\frac{(1 - \delta)}{2C_1\delta}\right)^{1/\gamma}$$

implies

$$C_1 h_j^\gamma \delta^{(4-\alpha)/(2-\alpha)} \leq \frac{1}{2}(1 - \delta)\delta^{2/(2-\alpha)}.$$

This completes the proof of the theorem.

Remark 5.2. The multigrid process described in Section 4 requires that the problem on the coarsest grid be solved exactly, i.e., $B_j = A_j^{-1}$. It is possible to relax this restriction and still apply the results of this paper. We consider, for example, the variable \mathcal{V} -cycle multigrid algorithm. From the proof of Theorem 1 it is immediate that the theorem will still hold as long as B_j satisfies

$$(5.18) \quad \|(I - B_j A_j)\|_{k,1}^2 \leq \delta_j,$$

where

$$(5.19) \quad \delta_j = \frac{M}{M + \beta_0^{\alpha/2} m(j + 1)^{\alpha/2}}.$$

One obvious choice for an iterative definition of B_j is $B_j g = x^{m_j}$ where x^l for $l = 1, \dots, m_j$ is given by (4.1) with $k = j$. Here m_j is some integer to be specified. An iterative definition of B_j has the advantage that no additional coding is necessary (in contrast to the use of $B_j = A_j^{-1}$, where direct solvers for nonsymmetric and indefinite problems must be introduced into the code). There are two additional factors involved in the use of an iterative process for B_j . First, one would like to avoid the coarsest grids so that $h_j \leq \nu$ is satisfied. Secondly, the computational work on the coarsest grid should not increase the asymptotic work of the algorithm.

We consider the application described in Section 3. We should like the multigrid algorithm to achieve a reduction δ_J which is independent of h_J , with computational effort bounded by a constant times the number of grid points in the finest grid. Let $N(k)$ denote the number of degrees of freedom in the k th grid level. We assume $N(k)/N(k - 1) \geq c_0 > \beta_1$, and hence the amount of work on the grids $1, \dots, J$ will be bounded by $O(N(J))$ [3], [6]. It is not difficult to see that for B_j defined as above,

$$\| (I - B_j A_j) \|_{k,1}^2 \leq (1 - ch_j^4)^{m_j}.$$

On the other hand, if we take $\beta_0 = \beta_1 = 2$,

$$\frac{M}{M + 2^{\alpha/2} m(j + 1)^{\alpha/2}} \geq C(h_J/h_j)^{\alpha/2}.$$

Consequently, to satisfy (5.18)–(5.19), we need only take

$$(5.20) \quad m_j = O(h_J^{-1} h_j^{-3}).$$

The work constraint is then $h_J^{-1} h_j^{-5} \leq ch_J^{-2}$. Thus setting $h_j = h_J^{1/5}$ and defining m_j by (5.20) gives rise to a multigrid algorithm which yields a uniform reduction independent of h_J , with an operation count bounded by a constant times the number of degrees of freedom on the finest grid.

6. The Proof of Lemmas 5.1, 5.2 and 5.3. This section will provide the proofs of Lemmas 5.1–5.3. Before proceeding, let us state two propositions and two preliminary lemmas.

PROPOSITION 6.1. *There are positive constants c and C not depending on $v \in \mathcal{M}$ such that*

$$\|v\|_{H^1}^2 \leq C\{A(v, v) + c\|v\|_H^2\}.$$

PROPOSITION 6.2. *For $v \in H^{1+\alpha}$ and $0 \leq \delta \leq \alpha$,*

$$(6.1) \quad \|(I - \hat{P}_k)v\|_{H^{1-\delta}} \leq Ch_k^\delta \|(I - \hat{P}_k)v\|_{H^1}.$$

If h_j is sufficiently small, then P_k is well defined and

$$(6.2) \quad \|(I - P_k)v\|_{H^1} \leq C \inf_{\chi \in \mathcal{M}_k} \|v - \chi\|_{H^1},$$

for all $v \in \mathcal{M}$.

Proposition 6.1 follows immediately from (2.6). (6.1) follows from a standard duality argument and (6.2) can be proved by using the techniques given in [20].

We next introduce the preliminary lemmas. The first lemma was essentially proved in [1].

LEMMA 6.1. *Let $0 \leq s \leq 1$. There exist positive constants c_1, c_2 and c_3 such that*

$$\|\chi\|_{H^s} \leq c_1 \|\chi\|_{k,s} \leq c_2 \|\chi\|_{k,s} \leq c_3 \|\chi\|_{H^s} \quad \text{for all } \chi \in \mathcal{M}_k.$$

In addition, there are constants c and C satisfying

$$c \|\chi\|_{k,2} \leq \|\chi\|_{k,2} \leq C \|\chi\|_{k,2} \quad \text{for all } \chi \in \mathcal{M}_k.$$

LEMMA 6.2. *There exists a positive constant C which does not depend upon $v \in H^1$ such that*

$$(6.3) \quad \|(I - P_k^0)v\|_H \leq Ch_k \|v\|_{H^1},$$

$$(6.4) \quad \|P_k^0 v\|_{H^s} \leq C \|v\|_{H^s} \quad \text{for all } 0 \leq s \leq 1.$$

Proof. Let πv denote the H projection of v into \mathcal{M}_k . Using (A.3), (A.4) and standard techniques of finite element analysis gives

$$\begin{aligned} \|\pi v\|_{H^1} &\leq C \|v\|_{H^1}, \\ \|(I - \pi)v\|_H &\leq Ch_k \|v\|_{H^1}. \end{aligned}$$

For $\chi \in \mathcal{M}_k$, by (2.10),

$$((P_k^0 - \pi)v, \chi)_k = (\pi v, \chi) - (\pi v, \chi)_k \leq Ch_k \|\pi v\|_{H^1} \|\chi\|_k,$$

hence

$$\|(P_k^0 - \pi)v\|_k \leq Ch_k \|\pi v\|_{H^1}.$$

Estimate (6.3) follows from the triangle inequality.

For (6.4), by interpolation, it suffices to verify the cases $s = 0$ and $s = 1$. The case for $s = 0$ follows immediately from the definition of P_k^0 and (2.9). For $s = 1$, the argument is standard and proceeds as follows:

$$\begin{aligned} \|P_k^0 v\|_{H^1} &\leq \|(P_k^0 - \pi)v\|_{H^1} + \|\pi v\|_{H^1} \\ &\leq Ch_k^{-1} \|(P_k^0 - \pi)v\|_k + C \|v\|_{H^1} \leq C \|v\|_{H^1}. \end{aligned}$$

This completes the proof of the lemma.

We can now prove Lemma 5.1.

Proof of Lemma 5.1. Following the argument in [5], we can easily show (using our assumptions and definitions) that

$$\|(I - \hat{P}_{k-1})v\|_1^2 \leq C(h_k^2 \|\hat{A}_k v\|_k^2)^\alpha \hat{A}(v, v)^{1-\alpha} \quad \text{for all } v \in \mathcal{M}_k.$$

We note that (A.4) and Lemma 6.1 imply that $h_k^2 \leq C\mu_k^{-1}$. The lemma now follows from (6.2) and Lemma 6.1.

The proofs of Lemmas 5.2 and 5.3 require some technical perturbation estimates. We consider the term on the left-hand side of (5.1). Let $G_k = L_k - A_k$; then since

$$(A_k(I - P_{k-1})v, \chi)_k = 0,$$

we have

$$(6.5) \quad \begin{aligned} (L_k(I - P_{k-1})v, \chi)_k &= (G_k(I - P_{k-1})v, \chi)_k = ((I - P_{k-1})v, G_k^* \chi)_k \\ &\leq \|(I - P_{k-1})v\|_k \|G_k^* \chi\|_k. \end{aligned}$$

Thus, we must estimate $G_k^* = L_k - \hat{A}_k - D_k^*$.

In light of (6.5), we see that it would be useful to estimate the difference $L_k - \hat{A}_k$. Note that L_k is defined as the positive square root of the discrete operator $L_k^2 \equiv A_k^* A_k$. An alternative expression for L_k is given by the Dunford-Taylor integral representation (cf. [9]):

$$(6.6) \quad L_k = (2\pi i)^{-1} \int_{\Gamma} z^{1/2} \mathcal{R}_z(L_k^2) dz,$$

where $\mathcal{R}_z(L_k^2) \equiv (z - L_k^2)^{-1}$ and Γ is a simple closed curve in the right half (complex) plane which encloses the spectrum of L_k^2 . Let $\kappa_1, \kappa_2 > 0$ be such that the eigenvalues of L_k^2 and \hat{A}_k^2 are in the interval $[2\kappa_1, \kappa_2]$. In this paper, we will take Γ as illustrated in Figure 6.1, i.e.,

$$\Gamma = \{(\kappa_1, y) \mid y \in [-\kappa_1, \kappa_1]\} \cup \{(t, t) \mid t \in [\kappa_1, 2\kappa_2]\} \\ \cup \{(t, -t) \mid t \in [\kappa_1, 2\kappa_2]\} \cup \{(2\kappa_2, y) \mid y \in [-2\kappa_2, 2\kappa_2]\}.$$

Using an expression similar to (6.6) for \hat{A}_k gives

$$(6.7) \quad L_k - \hat{A}_k = (2\pi i)^{-1} \int_{\Gamma} z^{1/2} \mathcal{R}_z(L_k^2) (L_k^2 - \hat{A}_k^2) \mathcal{R}_z(\hat{A}_k^2) dz.$$

To estimate (6.7) we shall use the bounds given in the following lemma.

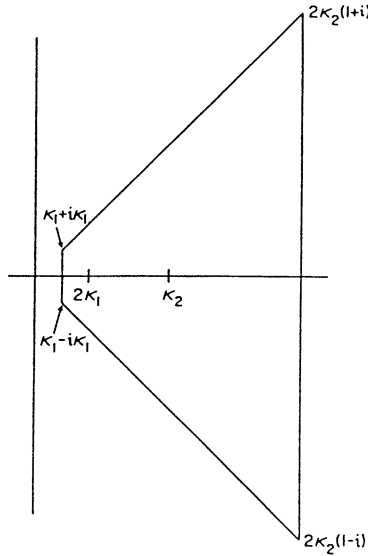


FIGURE 6.1
The curve Γ used in (6.6).

LEMMA 6.3. Let S and T be symmetric positive definite operators on \mathcal{M}_k satisfying

$$2\kappa_1 \|\chi\|_k^2 \leq (S^2 \chi, \chi)_k \leq \kappa_2 \|\chi\|_k^2, \\ 2\kappa_1 \|\chi\|_k^2 \leq (T^2 \chi, \chi)_k \leq \kappa_2 \|\chi\|_k^2,$$

for all $\chi \in \mathcal{M}_k$. Assume that $\kappa_1 \geq c$ independently of k . We allow S, T and κ_2 to depend on k . Then

$$(6.8) \quad \int_{\Gamma} |z|^{1/2} \|S \cdot \mathcal{R}_z(S^2)\|_k \|\mathcal{R}_z(T^2)\|_k d|z| \leq C(1 + \ln(\kappa_2/\kappa_1)),$$

and for any $\chi \in \mathcal{M}_k$,

$$(6.9) \quad \int_{\Gamma} |z|^{1/2} \|S^{1/2} \cdot \mathcal{R}_z(S^2)\chi\|_k^2 d|z| \leq C\|\chi\|_k^2.$$

Proof. By symmetry, it suffices to derive the above bounds for the curve $\Gamma_+ \equiv \Gamma_1 \cup \Gamma_2 \cup \Gamma_3$, where $\Gamma_1 \equiv \{(\kappa_1, y) \mid y \in [0, \kappa_1]\}$, $\Gamma_2 \equiv \{(t, t) \mid t \in [\kappa_1, 2\kappa_2]\}$ and $\Gamma_3 \equiv \{(2\kappa_2, y) \mid y \in [0, 2\kappa_2]\}$. By expansion in terms of eigenvectors, it is easy to see that

$$(6.10) \quad \|S^\beta \mathcal{R}_z(S^2)\|_k \leq \max_{\lambda \in [\sqrt{2\kappa_1}, \sqrt{\kappa_2}]} \lambda^\beta |\lambda^2 - z|^{-1}, \quad \beta = 0, 1/2, 1.$$

A similar inequality obviously holds for T .

Let

$$\mathcal{F}_1(\gamma) \equiv \int_\gamma |z|^{1/2} \|S \cdot \mathcal{R}_z(S^2)\|_k \|\mathcal{R}_z(T^2)\|_k d|z|$$

and

$$\mathcal{F}_2(\gamma) \equiv \int_\gamma |z|^{1/2} \|S^{1/2} \cdot \mathcal{R}_z(S^2)\chi\|_k^2 d|z|.$$

Then by (6.10) and elementary estimates,

$$\begin{aligned} \mathcal{F}_1(\Gamma_1) &\leq C \int_{\Gamma_1} |z|^{1/2} \kappa_1^{-3/2} d|z| \leq C, \\ \mathcal{F}_1(\Gamma_2) &\leq C \int_{\Gamma_2} |z|^{-1} d|z| \leq C \ln(2\kappa_2/\kappa_1), \\ \mathcal{F}_1(\Gamma_3) &\leq C \int_0^\infty \frac{\kappa_2}{\kappa_2^2 + y^2} dy \leq C. \end{aligned}$$

This verifies (6.8). Similar arguments give

$$\begin{aligned} \mathcal{F}_2(\Gamma_1) &\leq C \left(\int_{\Gamma_1} |z|^{1/2} \kappa_1^{-3/2} d|z| \right) \|\chi\|_k^2 \leq C \|\chi\|_k^2, \\ \mathcal{F}_2(\Gamma_3) &\leq C \left(\int_0^\infty \frac{\kappa_2}{\kappa_2^2 + y^2} dy \right) \|\chi\|_k^2 \leq C \|\chi\|_k^2. \end{aligned}$$

To bound $\mathcal{F}_2(\Gamma_2)$, we expand in terms of the eigenvectors of S . Let $\{\lambda_i, \theta_i\}$ denote the eigenvalue-eigenvector pairs for the operator S . Without loss of generality, we may assume that $\{\theta_i\}$ form an orthonormal basis for \mathcal{M}_j . Clearly, $2\kappa_1 \leq \lambda_i^2 \leq \kappa_2$ holds for each i . Decomposing

$$\chi = \sum_i c_i \theta_i$$

gives

$$\|S^{1/2} \cdot \mathcal{R}_z(S^2)\chi\|_k^2 = \sum_i \frac{\lambda_i c_i^2}{|\lambda_i^2 - z|^2}.$$

Integrating term by term yields

$$(6.11) \quad \mathcal{F}_2(\Gamma_2) = \sum_i 2^{3/4} c_i^2 \int_{\kappa_1}^{2\kappa_2} \frac{\lambda_i t^{1/2}}{(\lambda_i^2 - t)^2 + t^2} dt.$$

Elementary manipulations show that the integrals in (6.11) are bounded uniformly in κ_1, κ_2 , and λ_i . Hence $\mathcal{F}_2(\Gamma_2) \leq C \|\chi\|_k^2$. This completes the proof of the lemma.

We now state and prove a lemma for estimating $L_k - \hat{A}_k$.

LEMMA 6.4. *Let h_j be sufficiently small. Then there exists a constant C such that for all $\chi \in \mathcal{M}_k$,*

$$\|L_k \chi - \hat{A}_k \chi\|_{H^1} \leq C h_k^{\alpha-1} (1 + |\ln h_k|) \|\hat{A}_k \chi\|_k$$

and

$$\|L_k \chi - \hat{A}_k \chi\|_k \leq C h_k^{\alpha-1} \|\chi\|_{H^1}.$$

Proof. By Lemma 6.1,

$$\|L_k \chi - \hat{A}_k \chi\|_{H^1} \leq C \|L_k^{1/2} (L_k - \hat{A}_k) \chi\|_k.$$

By (6.7), for any $\chi, \theta \in \mathcal{M}_k$,

$$(L_k^{1/2} (L_k - \hat{A}_k) \chi, \theta)_k = (2\pi i)^{-1} \int_{\Gamma} z^{1/2} (E_k \mathcal{R}_z(\hat{A}_k^2) \hat{A}_k \chi, L_k \mathcal{R}_z(L_k^2) \theta)_k dz,$$

where

$$E_k = L_k^{-1/2} (L_k^2 - \hat{A}_k^2) \hat{A}_k^{-1}.$$

By the Schwarz inequality and (6.8), with $S = L_k$ and $T = \hat{A}_k$,

$$(L_k^{1/2} (L_k - \hat{A}_k) \chi, \theta)_k \leq C (1 + |\ln h_k|) \|E_k\|_k \|\hat{A}_k \chi\|_k \|\theta\|_k.$$

Note that we have used the fact that κ_1 is bounded uniformly from below and by (A.3), we can take $\kappa_2 \leq C h_k^{-4}$. Similarly, by (6.7),

$$(L_k \chi - \hat{A}_k \chi, \theta)_k = (2\pi i)^{-1} \int_{\Gamma} z^{1/2} (E_k \hat{A}_k^{1/2} \mathcal{R}_z(\hat{A}_k^2) \hat{A}_k^{1/2} \chi, L_k^{1/2} \mathcal{R}_z(L_k^2) \theta)_k dz.$$

By the Schwarz inequality, Lemma 6.1 and (6.9),

$$|(L_k \chi - \hat{A}_k \chi, \theta)_k| \leq C \|E_k\|_k \|\chi\|_{H^1} \|\theta\|_k.$$

Thus, the proof of the lemma will be complete if we can show that

$$(6.12) \quad \|E_k\|_k \leq C h_k^{\alpha-1}.$$

Obviously,

$$(6.13) \quad L_k^2 - \hat{A}_k^2 = \hat{A}_k D_k + D_k^* \hat{A}_k + D_k^* D_k$$

and hence

$$(6.14) \quad \|E_k\|_k \leq \|L_k^{-1/2} \hat{A}_k D_k \hat{A}_k^{-1}\|_k + \|L_k^{-1/2} D_k^* \hat{A}_k \hat{A}_k^{-1}\|_k + \|L_k^{-1/2} D_k^* D_k \hat{A}_k^{-1}\|_k.$$

Using Lemmas 6.1 and 6.2 and (2.7) gives

$$(6.15) \quad \|L_k^{-1/2} D_k^* \hat{A}_k \hat{A}_k^{-1}\|_k = \|D_k L_k^{-1/2}\|_k = \|P_k^0 D L_k^{-1/2}\|_k \leq C.$$

Similarly,

$$\|L_k^{-1/2} D_k^* D_k \hat{A}_k^{-1}\|_k \leq C \|D_k L_k^{-1/2}\|_k \|D_k \hat{A}_k^{-1/2}\|_k \leq C.$$

For the first term of (6.14), using Lemma 6.1 gives

$$\|L_k^{-1/2} \hat{A}_k D_k \hat{A}_k^{-1}\|_k \leq \|L_k^{-1/2} \hat{A}_k^{1/2}\|_k \|\hat{A}_k^{1/2} D_k \hat{A}_k^{-1}\|_k \leq C \|\hat{A}_k^{1/2} D_k \hat{A}_k^{-1}\|_k.$$

Combining the above estimates, making an obvious change of variable, and applying Lemma 6.2 implies that the proof of the lemma will be complete if we show

$$(6.16) \quad \|D_k \chi\|_{H^1} \leq C h_k^{\alpha-1} \|\hat{A}_k \chi\|_k \quad \text{for all } \chi \in \mathcal{M}_k.$$

Fix $\chi \in \mathcal{M}_k$ and let $w \in \mathcal{M}$ be the solution to

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

Clearly $\chi = \hat{P}_k w$. Now

$$\|D_k \chi\|_{H^1} \leq \|P_k^0 D(\chi - w)\|_{H^1} + \|P_k^0 D w\|_{H^1}.$$

Applying (2.7), (A.3), (A.4), and Lemma 6.2 gives

$$\|P_k^0 D(\chi - w)\|_{H^1} \leq C h_k^{-1} \|\chi - w\|_{H^1} \leq h_k^{\alpha-1} \|w\|_{H^{1+\alpha}}.$$

Finally, by (A.4), Lemma 6.2 and (2.8),

$$\|P_k^0 D w\|_{H^1} \leq C h_k^{\alpha-1} \|P_k^0 D w\|_{H^\alpha} \leq C h_k^{\alpha-1} \|D w\|_{H^\alpha} \leq C h_k^{\alpha-1} \|w\|_{1+\alpha}.$$

Inequality (6.16) now follows combining the above estimates with (A.1). This completes the proof of Lemma 6.4.

We can now prove Lemma 5.2.

Proof of Lemma 5.2. By (6.5), Lemma 6.1 and Proposition 6.2, it suffices to show that

$$\|G_k^* \chi\|_k \leq C h^{-1/2-\varepsilon} \|\chi\|_{H^1}.$$

In turn, by Lemma 6.4 and the triangle inequality, noting that $\alpha > 1/2$, it suffices to show

$$(6.17) \quad \|D_k^* \chi\|_k \leq C h^{-1/2-\varepsilon} \|\chi\|_{H^1}.$$

Let $\theta \in \mathcal{M}_k$; then by (A.2), (A.4) and Lemma 6.1,

$$(D_k^* \chi, \theta)_k = (D^* \chi, \theta) \leq C \|\chi\|_{H^1} \|\theta\|_{H^{1/2+\varepsilon}} \leq C h_k^{-1/2-\varepsilon} \|\chi\|_{H^1} \|\theta\|_k.$$

Inequality (6.17) immediately follows. This completes the proof of the lemma.

We shall need two additional lemmas for the proof of Lemma 5.3. The first involves stability and approximation for the operator I_k .

LEMMA 6.5. *There exists a positive constant C such that for all $\chi \in \mathcal{M}_k$*

$$(6.18) \quad \|(I - I_{k-1})\chi\|_H \leq C h_k \|\chi\|_{H^1}$$

and

$$(6.19) \quad \|I_{k-1}\chi\|_{H^1} \leq C \|\chi\|_{H^1}.$$

Proof. Note that by (2.9) and (2.10), for $\varphi \in \mathcal{M}_{k-1}$,

$$((I_{k-1} - P_{k-1}^0)\chi, \varphi)_{k-1} = (\chi, \varphi)_k - (\chi, \varphi) \leq C h_k \|\chi\|_{H^1} \|\varphi\|_{k-1}.$$

This implies that

$$\|(I_{k-1} - P_{k-1}^0)\chi\|_{k-1} \leq C h_k \|\chi\|_{H^1}.$$

The lemma then follows from Lemmas 6.1 and 6.2 and (A.4).

LEMMA 6.6. *There exists a positive constant C such that for all $\chi \in \mathcal{M}_{k-1}$*

$$(6.20) \quad \|\hat{A}_k \chi\|_k \leq C h_k^{\alpha-1} \|\hat{A}_{k-1} \chi\|_{k-1},$$

$$(6.21) \quad \|L_k \chi\|_k \leq C h_k^{\alpha-1} \|L_{k-1} \chi\|_{k-1},$$

$$(6.22) \quad \|\hat{A}_{k-1} \chi\|_{k-1} \leq C h_k^{\alpha-1} \|I_{k-1} L_k \chi\|_{k-1}$$

and

$$(6.23) \quad \|L_k \chi\|_k \leq Ch_k^{2\alpha-2} \|I_{k-1} L_k \chi\|_{k-1}.$$

Proof. By Proposition 2, Lemma 6.1, (A.3) and (A.4), for all $\varphi \in \mathcal{M}_k$,

$$\|\varphi - \hat{P}_{k-1} \varphi\|_k \leq Ch_k^\alpha \|\hat{\varphi}\|_{H^1} \leq Ch_k^{\alpha-1} \|\varphi\|_k,$$

hence

$$\|\hat{P}_{k-1} \varphi\|_k \leq Ch_k^{\alpha-1} \|\varphi\|_k.$$

Therefore, for $\chi \in \mathcal{M}_{k-1}$,

$$\begin{aligned} (\hat{A}_k \chi, \varphi)_k &= \hat{A}(\chi, \varphi) = \hat{A}(\chi, \hat{P}_{k-1} \varphi) \\ &= (\hat{A}_{k-1} \chi, \hat{P}_{k-1} \varphi)_{k-1} \leq Ch_k^{\alpha-1} \|\hat{A}_{k-1} \chi\|_{k-1} \|\varphi\|_k. \end{aligned}$$

This proves (6.20). Inequality (6.21) then follows from (6.20) and Lemma 6.1.

We next prove (6.22). Noting that $\hat{A}_{k-1} = I_{k-1} \hat{A}_k$, the triangle inequality and Lemma 6.4 give

$$\begin{aligned} \|\hat{A}_{k-1} \chi\|_{k-1} &= \|I_{k-1} \hat{A}_k \chi\|_{k-1} \leq (\|I_{k-1} L_k \chi\|_{k-1} + \|(\hat{A}_k - L_k) \chi\|_k) \\ &\leq Ch_k^{\alpha-1} (\|I_{k-1} L_k \chi\|_{k-1} + \|\chi\|_{H^1}). \end{aligned}$$

Finally, we note that by Lemma 6.1 and (2.9),

$$\|\chi\|_{H^1}^2 \leq C(L_k \chi, \chi)_k \leq C \|I_{k-1} L_k \chi\|_{k-1} \|\chi\|_{H^1},$$

and hence

$$\|\chi\|_{H^1} \leq C \|I_{k-1} L_k \chi\|_{k-1}.$$

Combining the above inequalities completes the proof of (6.22). Inequality (6.23) follows immediately from (6.22), (6.20) and Lemma 6.1.

We are now ready to prove Lemma 5.3. However, before doing so, we note a few properties of our operators which are immediate consequences of the defining relations. As noted earlier, $\hat{A}_{k-1} = I_{k-1} \hat{A}_k$. Similarly, $D_{k-1} = I_{k-1} D_k$. In addition, the operator I_{k-1} is symmetric on both \mathcal{M}_k with the $(\cdot, \cdot)_k$ inner product as well as \mathcal{M}_{k-1} with the $(\cdot, \cdot)_{k-1}$ inner product.

Proof of Lemma 5.3. For $\chi \in \mathcal{M}_{k-1}$,

$$|\|\chi\|_{k,1}^2 - \|\chi\|_{k-1,1}^2| = ((\tilde{L}_{k-1} - L_{k-1})\chi, \chi)_{k-1},$$

where $\tilde{L}_{k-1} \equiv I_{k-1} L_k$. Note that the operator $\tilde{L}_{k-1} : \mathcal{M}_{k-1} \mapsto \mathcal{M}_{k-1}$ is symmetric and the eigenvalues of \tilde{L}_{k-1}^2 are in the interval $[c, Ch_k^{-4}]$ for appropriate constants c and C (independent of k). Applying an expression analogous to (6.7) gives

$$\begin{aligned} & \left((L_{k-1} - \tilde{L}_{k-1}) \chi, \chi \right)_{k-1} \\ &= (2\pi i)^{-1} \int_{\Gamma} z^{1/2} (F_k L_{k-1} \mathcal{R}_z(L_{k-1}^2) \chi, \tilde{L}_{k-1} \mathcal{R}_z(\tilde{L}_{k-1}^2) \chi)_{k-1} dz, \end{aligned}$$

where

$$F_k = \tilde{L}_{k-1}^{-1} (L_{k-1}^2 - \tilde{L}_{k-1}^2) L_{k-1}^{-1}.$$

By the Schwarz inequality and (6.9),

$$\begin{aligned} \left| \left((L_{k-1} - \tilde{L}_{k-1}) \chi, \chi \right)_{k-1} \right| &\leq C \|F_k\|_{k-1} \|L_{k-1}^{1/2} \chi\|_{k-1} \|\tilde{L}_{k-1}^{1/2} \chi\|_{k-1} \\ &\leq C \|F_k\|_{k-1} \|\chi\|_{k,1}^2, \end{aligned}$$

where the second inequality follows from Lemma 6.1 and the identity $(\tilde{L}_{k-1}\chi, \chi)_{k-1} = (L_k\chi, \chi)_k$. To complete the proof of the lemma, we need only bound $\|F_k\|_{k-1}$.

We start first with the identity

$$F_k = Q_1 + Q_2 + Q_3,$$

where

$$\begin{aligned} Q_1 &= (I - I_{k-1})(L_k - \hat{A}_k)L_{k-1}^{-1}, \\ Q_2 &= \tilde{L}_{k-1}^{-1}I_{k-1}(L_k - \hat{A}_k)(I - I_{k-1})\hat{A}_kL_{k-1}^{-1}, \\ Q_3 &= \tilde{L}_{k-1}^{-1}(L_{k-1}^2 - I_{k-1}L_k^2 - \hat{A}_{k-1}^2 + I_{k-1}\hat{A}_k^2)L_{k-1}^{-1}. \end{aligned}$$

Obviously, it suffices to bound the norms $\|Q_i\|_{k-1}$ for $i = 1, 2, 3$.

Let $\chi, \theta \in \mathcal{M}_{k-1}$. For Q_1 , by Lemmas 6.1, 6.4, 6.5 and 6.6, we have

$$\begin{aligned} \|Q_1\chi\|_{k-1} &\leq Ch_k\|(L_k - \hat{A}_k)L_{k-1}^{-1}\chi\|_{H^1} \leq Ch_k^\alpha(1 + |\ln h_k|)\|L_kL_{k-1}^{-1}\chi\|_k \\ &\leq Ch_k^{2\alpha-1}(1 + |\ln h_k|)\|\chi\|_{k-1}. \end{aligned}$$

For Q_2 , we have

$$\begin{aligned} |(Q_2\chi, \theta)_{k-1}| &= |(\hat{A}_kL_{k-1}^{-1}\chi, (I - I_{k-1})(L_k - \hat{A}_k)\tilde{L}_{k-1}^{-1}\theta)_k| \\ &\leq \|\hat{A}_kL_{k-1}^{-1}\chi\|_k\|(I - I_{k-1})(L_k - \hat{A}_k)\hat{A}_k^{-1}\|_k\|\hat{A}_k\tilde{L}_{k-1}^{-1}\theta\|_k. \end{aligned}$$

Thus, applying Lemmas 6.1, 6.4, 6.5, and 6.6 gives

$$\|Q_2\|_{k-1} \leq Ch_k^{4\alpha-3}(1 + |\ln h_k|).$$

For Q_3 , we obviously have

$$\|Q_3\|_{k-1} \leq \|Q_{3,1}\|_{k-1} + \|Q_{3,2}\|_{k-1} + \|Q_{3,3}\|_{k-1},$$

where

$$\begin{aligned} Q_{3,1} &= \tilde{L}_{k-1}^{-1}\hat{A}_{k-1}(I_{k-1} - I)D_kL_{k-1}^{-1}, \\ Q_{3,2} &= \tilde{L}_{k-1}^{-1}(D_{k-1}^*\hat{A}_{k-1} - I_{k-1}D_k^*\hat{A}_k)L_{k-1}^{-1}, \\ Q_{3,3} &= \tilde{L}_{k-1}^{-1}(D_{k-1}^*D_{k-1} - I_{k-1}D_k^*D_k)L_{k-1}^{-1}. \end{aligned}$$

For $Q_{3,1}$, by (6.16), (6.22), and Lemma 6.5,

$$\begin{aligned} \|Q_{3,1}\chi\|_{k-1} &\leq Ch_k\|\tilde{L}_{k-1}^{-1}\hat{A}_{k-1}\|_{k-1}\|D_kL_{k-1}^{-1}\chi\|_{H^1} \\ &\leq Ch^{3\alpha-2}\|\chi\|_{k-1}. \end{aligned}$$

For $Q_{3,2}$,

$$\begin{aligned} |(Q_{3,2}\chi, \theta)_{k-1}| &= |(\hat{A}_kL_{k-1}^{-1}\chi, (I - I_{k-1})D_k\tilde{L}_{k-1}^{-1}\theta)_k| \\ &\leq \|\hat{A}_kL_{k-1}^{-1}\chi\|_k\|(I - I_{k-1})D_k\hat{A}_k^{-1}\|_k\|\hat{A}_k\tilde{L}_{k-1}^{-1}\theta\|_k. \end{aligned}$$

Applying (6.16) and Lemmas 6.5 and 6.6 gives

$$\|Q_{3,2}\|_{k-1} \leq Ch_k^{4\alpha-3}.$$

Finally, for $Q_{3,3}$,

$$\begin{aligned} |(Q_{3,3}\chi, \theta)_{k-1}| &= |(D_kL_{k-1}^{-1}\chi, (I - I_{k-1})D_k\tilde{L}_{k-1}^{-1}\theta)_k| \\ &\leq \|D_kL_{k-1}^{-1}\|_k\|L_kL_{k-1}^{-1}\chi\|_k\|(I - I_{k-1})D_k\hat{A}_k^{-1}\|_k\|\hat{A}_k\tilde{L}_{k-1}^{-1}\theta\|_k. \end{aligned}$$

Applying (6.15), (6.16), and Lemmas 6.1, 6.5 and 6.6 gives

$$\|Q_{3,3}\|_{k-1} \leq Ch_k^{4\alpha-3}.$$

Combining the above inequalities proves Lemma 5.3.

7. Numerical Results. In this section, we give the results of numerical experiments involving the multigrid algorithms. These model computations show that the assumption $h_j \leq \nu$ is necessary for convergence in practice. In contrast, the degradation of the convergence rate for $\alpha = 1$ suggested by Theorem 2 was not observed in the reported computations.

In our numerical examples, we consider the symmetric and indefinite problem

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

where Ω is the unit square in R^2 . No examples for the nonsymmetric problem are given.

The eigenvalues for the operator of (7.1) are $(j^2 + k^2)\pi^2 - \mu$ where j, k are positive integers. We will consider the cases $\mu = 30$ and $\mu = 65$. The case $\mu = 30$ has only one negative eigenvalue. When $\mu = 65$, there are two negative eigenvalues, one of which is of multiplicity two.

To triangulate Ω , we first partition it into a regular rectangular mesh and then split each rectangle into two triangles (see Figure 7.1). We use the continuous piecewise linear finite element subspace on the resulting triangulations described in Section 3 and use the discrete inner products given by (3.6). For the purpose of this computation, we deviate from the finite element approximation in that the lower-order term in (7.1) is approximated by an appropriately weighted diagonal term. This is the so-called ‘lumped mass’ finite difference operator. With this discretization it is possible to actually compute the action of L_k and its inverse.

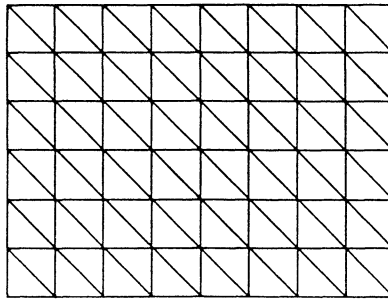


FIGURE 7.1

The regular triangular mesh defining \mathcal{M}_1 .

For these examples, it is computationally feasible to actually compute the best possible δ_k satisfying (4.4). Note that (4.4) is equivalent to the inequality,

$$((I - A_k B_k^*) L_k (I - B_k A_k) v, v)_k \leq \delta_k (L_k v, v)_k \quad \text{for all } v \in \mathcal{M}_k.$$

Thus, the best value of δ_k equals the largest eigenvalue of the operator

$$(7.2) \quad \mathcal{E} = L_k^{-1} (I - A_k B_k^*) L_k (I - B_k A_k).$$

The largest eigenvalue of \mathcal{E} can then be computed by, for example, the power method, if routines for computing the action of \mathcal{E} are available. We obviously know how to compute A_k and B_k . For the constant coefficient problem on a rectangular domain with a regular mesh, the operator L_k and its inverse can be efficiently computed by use of the Fast Fourier Transform. We are left to compute B_k^* .

For a symmetric problem, the operator B_k^* is also a multigrid operator and is given by the following algorithm [18]:

Algorithm for computing B_k^* . Let $B_j^* : \mathcal{M}_j \mapsto \mathcal{M}_j$ denote the adjoint of A_j^{-1} . Assume that $B_{k-1}^* : \mathcal{M}_{k-1} \mapsto \mathcal{M}_{k-1}$ has been defined and define B_k^*g for $g \in \mathcal{M}_k$ as follows:

- (I) Set $q^0 = 0$.
- (II) Define $x^0 = q^p$ where q^i , for $i = 1, 2, \dots, p$, is defined by

$$q^i = q^{i-1} + B_{k-1}^*[I_{k-1}g - A_{k-1}q^{i-1}].$$

- (III) For $l = 1, \dots, m(k)$, define

$$x^l = x^{l-1} + \mu_k^{-2}A_k^*(g - A_kx^{l-1}).$$

Arguments similar to those leading to (4.3) imply that the operator B_k^* defined by (I)–(III) satisfies the equation

$$(7.3) \quad I - B_k^*A_k = K_k^{m(k)}[(I - P_{k-1}) + (I - B_{k-1}^*A_{k-1})^pP_{k-1}].$$

A straightforward mathematical induction argument using (7.3) and the symmetry of A_k implies that the operator defined by (I)–(III) is the adjoint of B_k .

Table 7.1 gives the computed largest eigenvalue for the discrete system (7.2). We vary the mesh size on the finest grid and use $h_j = 1/8$ for the coarsest grid. We give the convergence parameter δ_j as a function of $h_J = 2^{-2-J}$ for $J = 2, 3, 4, 5$ for the variable \mathcal{V} -cycle algorithm ($\beta_0 = \beta_1 = 2$), the standard \mathcal{V} -cycle algorithm ($m(k) = 1$) and the \mathcal{W} -cycle algorithm ($m(k) = 1$). In the variable \mathcal{V} -cycle case, we use $m(J) = 1$.

The results of Table 7.1 illustrate that the multigrid process can be used to develop convergent iterative algorithms for the solution of the equations on the finest grid level. The rate of iterative convergence for these algorithms appears to be bounded independently of the number of grid levels as suggested by the theory.

TABLE 7.1
 δ_k for 'symmetric' multigrid schemes with $h_j = 1/8$
 applied to (7.1) with $\mu = 30$.

h_J	Var \mathcal{V} -cycle	\mathcal{V} -cycle	\mathcal{W} -cycle
1/16	.88	.88	.88
1/32	.88	.90	.88
1/64	.88	.90	.88
1/128	.88	.90	.88

The next table illustrates the importance of satisfying the assumption $h_j \leq \nu$. For this example, we again consider (7.1) with $\mu = 30$ but use $h_j = 1/4$. Values of δ_J greater than one indicate instability of the multigrid scheme. Note that only the \mathcal{W} -cycle examples with $h_J < 1/16$ and the variable \mathcal{V} -cycle example with $h_J = 1/16$ were stable. It should not be inferred from these results that the \mathcal{W} -cycle is generally more stable than the \mathcal{V} -cycle algorithms. Later examples will show that it shares the same type of stability problems.

TABLE 7.2
 δ_k for ‘symmetric’ multigrid schemes with $h_j = 1/4$
 applied to (7.1) with $\mu = 30$.

h_J	Var \mathcal{V} -cycle	\mathcal{V} -cycle	\mathcal{W} -cycle
1/16	.93	1.06	1.02
1/32	1.09	1.07	.88
1/64	1.09	1.07	.88
1/128	1.08	1.07	.88

The next table illustrates how the convergence rate of the multigrid schemes depends on the number of smoothings used. Table 7.3 gives δ_J as a function of $m(J)$, the number of smoothings on the finest grid level. In this example, $h_J = 1/128$ and $h_j = 1/8$ and we again use $\beta_0 = \beta_1 = 2$ in the variable \mathcal{V} -cycle scheme. The theory developed earlier indicates that, for stability, ν can be chosen independently of $m(J)$. This is consistent with the numerical results which remain stable without the use of smaller h_j as $m(J)$ increases. In contrast, the ‘nonsymmetric’ scheme requires the use of smaller h_j as the number of smoothings increases [1], [14].

TABLE 7.3
 δ_k for ‘symmetric’ multigrid schemes applied to (7.1)
 with $\mu = 30$, $h_j = 1/8$ and $h_J = 1/128$.

$m(J)$	Var \mathcal{V} -cycle	\mathcal{V} -cycle	\mathcal{W} -cycle
1	.88	.90	.88
3	.68	.74	.68
5	.52	.64	.52
7	.43	.56	.43
9	.39	.54	.39

Table 7.3 also shows that the rate of convergence δ_J decreases with larger $m(J)$ as theoretically predicted.

For the final example, we consider $\mu = 65$. In this case, we had to use $h_j = 1/16$ to get a stable algorithm. The computed values of δ_J for $1/32 \leq h_j \leq 1/128$ for the variable \mathcal{V} -cycle, the \mathcal{V} -cycle, and \mathcal{W} -cycle algorithms were approximately .88, .9 and .88, respectively. These results, as well as those given in Table 7.1, do not exhibit the convergence degradation for the \mathcal{V} -cycle algorithm suggested by Theorem 2.

Table 7.4 gives computed values of δ_J when $h_j = 1/4$ was used. These results again illustrate the importance of the theoretical assumption $h_j \leq \nu$. Note that a value of δ_j of a thousand implies that two steps of multigrid will amplify certain frequencies of the error by a factor of a thousand. Such an amplification leads to a rapidly divergent numerical scheme. This example also illustrates that the \mathcal{W} -cycle algorithm displays the same type of stability problems as the \mathcal{V} -cycle algorithms. In fact, the \mathcal{W} -cycle schemes were so unstable at smaller h_J , that it was impossible to compute the corresponding values of δ_J due to computer exponential overflow.

TABLE 7.4
 δ_k for 'symmetric' multigrid schemes applied to (7.1)
 with $\mu = 65$ and $h_j = 1/4$.

h_j	Var \mathcal{V} -cycle	\mathcal{V} -cycle	\mathcal{W} -cycle
1/16	956	1060	8.0×10^5
1/32	826	1115	6.5×10^{11}
1/64	634	1121	*
1/128	484	1120	*

When the above examples converge, we observe almost identical results for the variable \mathcal{V} -cycle and the \mathcal{W} -cycle algorithms. Note that both algorithms have the same number of smoothing iterations on the various grid levels. For these examples, the extra grid transfer involved in the \mathcal{W} -cycle algorithm does not seem to yield a faster convergence rate.

Department of Mathematics
 Cornell University
 Ithaca, New York 14853
E-mail: bramble@mathvax.msi.cornell.edu

Brookhaven National Laboratory
 Upton, New York 11973
E-mail: pasciak@bnlux0.bnl.gov

Department of Mathematics
 Cornell University
 Ithaca, New York 14853

1. R. E. BANK, "A comparison of two multilevel iterative methods for nonsymmetric and indefinite elliptic finite element equations," *SIAM J. Numer. Anal.*, v. 18, 1981, pp. 724–743.
2. R. E. BANK & C. C. DOUGLAS, "Sharp estimates for multigrid rates of convergence with general smoothing and acceleration," *SIAM J. Numer. Anal.*, v. 22, 1985, pp. 617–633.
3. R. E. BANK & T. DUPONT, "An optimal order process for solving elliptic finite element equations," *Math. Comp.*, v. 36, 1981, pp. 35–51.
4. D. BRAESS & W. HACKBUSCH, "A new convergence proof for the multigrid method including the V -cycle," *SIAM J. Numer. Anal.*, v. 20, 1983, pp. 967–975.
5. J. H. BRAMBLE & J. E. PASCIAK, "New convergence estimates for multigrid algorithms," *Math. Comp.*, v. 49, 1987, pp. 311–329.
6. A. BRANDT, "Multi-level adaptive solutions to boundary-value problems," *Math. Comp.*, v. 31, 1977, pp. 333–390.
7. P. GRISVARD, "Behavior of the solutions of an elliptic boundary value problem in a polygonal or polyhedral domain," in *Numerical Solution of Partial Differential Equations*, III (B. Hubbard, ed.), Academic Press, New York, 1976, pp. 207–274.
8. W. HACKBUSCH, *Multi-Grid Methods and Applications*, Springer-Verlag, New York, 1985.
9. T. KATO, *Perturbation Theory for Linear Operators*, Springer-Verlag, New York, 1976.
10. V. A. KONDRAT'EV, "Boundary problems for elliptic equations with conical or angular points," *Trans. Moscow Math. Soc.*, v. 16, 1967, pp. 227–313.
11. S. G. KREIN & Y. I. PETUNIN, *Scales of Banach spaces*, Russian Math. Surveys, vol. 21, 1966, pp. 85–160.
12. J. L. LIONS & E. MAGENES, *Problèmes aux Limites non Homogènes et Applications*, Dunod, Paris, 1968.

13. J. F. MAITRE & F. MUSY, "Algebraic formalization of the multigrid method in the symmetric and positive definite case—A convergence estimation for the V -cycle," in *Multigrid Methods for Integral and Differential Equations* (D. J. Paddon and H. Holstein, eds.), Clarendon Press, Oxford, 1985.
14. J. MANDEL, "Multigrid convergence for nonsymmetric, indefinite variational problems and one smoothing step," in *Proc. Copper Mtn. Conf. Multigrid Methods*, Appl. Math. Comput., 1986, pp. 201–216.
15. J. MANDEL, *Algebraic Study of Multigrid Methods for Symmetric, Definite Problems*. (Preprint.)
16. J. MANDEL, S. F. MCCORMICK & J. RUGE, *An Algebraic Theory for Multigrid Methods for Variational Problems*. (Preprint.)
17. S. F. MCCORMICK, "Multigrid methods for variational problems: Further results," *SIAM J. Numer. Anal.*, v. 21, 1984, pp. 255–263.
18. S. F. MCCORMICK, "Multigrid methods for variational problems: General theory for the V -cycle," *SIAM J. Numer. Anal.*, v. 22, 1985, pp. 634–643.
19. J. NEČAS, *Les Méthodes Directes en Théorie des Équations Elliptiques*, Academia, Prague, 1967.
20. A. H. SCHATZ, "An observation concerning Ritz-Galerkin methods with indefinite bilinear forms," *Math. Comp.*, v. 28, 1974, pp. 959–962.
21. H. YSERENTANT, "The convergence of multi-level methods for solving finite-element equations in the presence of singularities," *Math. Comp.*, v. 47, 1986, pp. 399–409.