SIAM J. NUMER. ANAL.
Vol. 34, No. 1, pp. 178–200, February 1997

MULTIGRID METHODS FOR NEARLY SINGULAR LINEAR EQUATIONS AND EIGENVALUE PROBLEMS*

ZHIQIANG CAI†, JAN MANDEL‡, AND STEVE MCCORMICK§

Abstract. The purpose of this paper is to develop a convergence theory for multigrid methods applied to nearly singular linear elliptic partial differential equations of the type produced from a positive definite system by a shift with the identity. One of the important aspects of this theory is that it allows such shifts to vary anywhere in the multigrid scheme, enabling its application to a wider class of eigenproblem solvers. The theory is first applied to a method for computing eigenvalues and eigenvectors that consists of multigrid iterations with zero right-hand side and updating the shift from the Rayleigh quotient before every cycle. It is then applied to the Rayleigh quotient multigrid (RQMG) method, which is a more direct multigrid procedure for solving eigenproblems. Local convergence of the multigrid V-cycle and global convergence for a full multigrid version of both methods is obtained.

Key words. eigenvalue problem, multigrid, Rayleigh quotient, singular equations

AMS subject classifications. 65N55, 65F10, 65N30

PII. S1064827594261139

1. Introduction.

In this paper, we consider the solution of the generalized eigenvalue problem based in an abstract finite-dimensional Hilbert space V with inner product $(\cdot,\cdot)$: find $\lambda \in \mathcal{R}$ and $0 \neq u \in V$ such that

$$Au = \lambda Bu. \tag{1.1}$$

Here for simplicity A and B are assumed to be linear continuous symmetric positive definite operators defined on V.

We will consider two multigrid approaches for finding the smallest eigenvalue for (1.1) based on a sequence of subspaces. One uses multigrid as an inner loop solver for an outer loop inverse iteration type process, which has been studied by many authors (cf. the early work in [1] and [7]). The other is the Rayleigh quotient multigrid (RQMG) method [5, 8], which is a more direct approach based on minimizing the Rayleigh quotient at each stage of the multigrid processing. To our knowledge, this is the first theory for methods like RQMG where the shifts are allowed to vary within multigrid cycles and be close to eigenvalues. Filling this gap is the main purpose of this paper.

We will analyze convergence of these two multigrid methods by developing and applying a general convergence theory for singular or nearly singular linear problems: given $f \in V$ and a scalar $\mu \in \mathcal{R}$, find $0 \neq u \in V$ such that

$$(A - \mu B)u = f. \tag{1.2}$$

*Received by the editors April 14, 1994; accepted for publication (in revised form) March 27, 1995.

http://www.siam.org/journals/sinum/34-1/26113.html

†Department of Mathematics, Purdue University, West Lafayette, IN 47907-1395 (zcai@math.purdue.edu). The research of this author was supported by National Science Foundation grant DMS-9619792.

‡Center for Computational Mathematics, University of Colorado at Denver, Denver, CO 80217-3364. The research of this author was supported by National Science Foundation grant ASC-9121431.

§Program in Applied Mathematics, University of Colorado at Boulder, Boulder, CO 80309-0526. The research of this author was supported by Air Force Office of Scientific Research grant AFOSR-91-0156 and by National Science Foundation grant DMS-8704169.

Here we restrict ourselves to theoretical analyses because of the numerical results available in [5] for RQMG, the principal target of this paper.

Previous convergence results for multigrid algorithms applied to (1.2) were obtained by Bank [1]. In order to establish norm estimates for the rate of convergence, the shift μ was assumed to be bounded away from the smallest eigenvalue of (1.1) in [1]. In contrast, our analysis uses an error decomposition into the eigenspace associated with the smallest eigenvalue of (1.1) and its orthogonal complement. We will not attempt to solve (1.2) in the usual sense; instead, our aim is to preserve the approximate magnitude of the components in the eigenspace of the smallest eigenvalue of (1.1) while attenuating error components in its orthogonal complement. In our analysis, the shift μ will be allowed to vary in a small neighborhood of the smallest eigenvalue of (1.1).

The outline of the remainder of this paper is as follows. In section 2, we formulate the problems, establish notation, and define a multigrid algorithm for the nearly singular problem. In section 3, we develop a convergence theory for this multigrid algorithm. The theory is first applied in section 4 to a method for computing eigenvalues and eigenvectors that uses multigrid as an inner loop solver for an outer loop inverse iteration type process. It is then applied in section 5 to RQMG. The final section develops global convergence results for full multigrid V-cycle versions of both methods.

2. Preliminaries. Let V be a real linear space, on which are given inner products $a(\cdot,\cdot)$ and $(\cdot,\cdot)$, with corresponding induced norms denoted by $|||\cdot|||$ and $\|\cdot\|$. Let $b(\cdot,\cdot)$ be a continuous, symmetric, positive definite bilinear form on $V \times V$. Consider the eigenvalue problem: find $\lambda \in \mathcal{R}$ and $0 \neq u \in V$ such that

$$a(u,v) = \lambda b(u,v) \quad \forall\, v \in V. \tag{2.1}$$

If (2.1) corresponds to the eigenvalue problem for a self-adjoint elliptic partial differential operator, it will typically admit an infinite set of nondecreasing eigenvalues. Without loss of such generality, let the (possibly multiple) eigenvalues of (2.1) satisfy

$$0 < \lambda_1 < \lambda_2 < \cdots. \tag{2.2}$$

In particular, we note that λ_1 is the minimum of the Rayleigh quotient over V:

$$\lambda_1 = \inf_{0 \neq u \in V} RQ(u) \tag{2.3}$$

where the Rayleigh quotient is defined by

$$RQ(u) = \frac{a(u,u)}{b(u,u)}. \tag{2.4}$$

We will consider multigrid methods for finding the smallest eigenvalue for (2.1), based on a sequence of finite-dimensional subspaces. To this end, let

$$V^0 \subset \cdots \subset V^k \subset \cdots \subset V$$

be a nested family of finite-dimensional subspaces of V. Let $(\cdot,\cdot)_k$ be a given inner product on V^k, $\|\cdot\|_k$ its induced norm, and $h_k = 2^{-k} h_0$ the mesh parameter associated with V^k, $h_0 > 0$. Then the corresponding finite-dimensional problem (2.1) on V^k is as follows: find $\lambda^k \in \mathcal{R}$ and $0 \neq u^k \in V^k$ such that

$$a(u^k,v) = \lambda^k b(u^k,v) \quad \forall\, v \in V^k. \tag{2.5}$$

Let the (possibly multiple) eigenvalues of (2.5) satisfy

$$0 < \lambda_1^k < \lambda_2^k < \cdots < \lambda_{max}^k. \tag{2.6}$$

Note that λ_1^k is the minimum of the Rayleigh quotient over V^k:

$$\lambda_1^k = \inf_{0 \neq u^k \in V^k} RQ(u^k).$$

In order to analyze convergence of the multigrid algorithm for the eigenvalue problem (2.5), we will first study the behavior of multigrid applied to the following *singular* or *nearly singular* problem: given a source term f^l in the dual space $(V^l)'$ and a scalar $\lambda \in \mathcal{R}$, find $0 \neq u^l \in V^l$ such that

$$a(u^l, v) - \lambda b(u^l, v) = f^l(v) \quad \forall\, v \in V^l. \tag{2.7}$$

The shift λ is assumed to satisfy

$$\lambda_1 \leq \lambda \leq \lambda_1^l + O(h_0/l), \quad l \to \infty. \tag{2.8}$$

By this we mean that

$$\lambda_1 \leq \lambda \leq \lambda_1^l + Ch_0/l, \quad l \to \infty \tag{2.9}$$

for some constant C independent of h_0 and l. This condition, which allows for shifts that are arbitrarily close to λ_1^l, can be guaranteed for the eigenproblem solvers we consider, as we show in the last section. However, since $\lambda = \lambda_1^l$ is allowed, problem (2.7) may not have a solution. This is acceptable because our real interest is eigenproblems: we will attempt to solve (2.7) not in the strict sense, but only in that the approximation is correct up to the eigencomponents belonging to λ_1^l.

The reader is strongly advised to keep in mind that all of the following estimates allow for λ to change any time during multigrid processing, provided it continues to satisfy the bounds in (2.8). *This allowance for a floating shift is one of the key distinguishing points of the theory developed here, and it is just what enables treatment of the nonlinear scheme RQMG below. Unfortunately, it is also what dramatically complicates our exposition.*

For any $u,\, v \in V$, define the bilinear forms $c_\lambda(\cdot,\, \cdot)$ and $c(\cdot,\, \cdot)$ on $V \times V$ by

$$c_\lambda(u,\, v) = a(u,\, v) - \lambda b(u,\, v) \quad \text{and} \quad c(u,\, v) = a(u,\, v) - \lambda_1 b(u,\, v).$$

Fix $k \in \{0,\, 1,\, \cdots,\, l\}$ and let $q^k(\cdot,\, \cdot)$ be a bilinear form on $V^k \times V^k$. We will use this form to define the smoothing step in the multigrid algorithm defined below. Define the operators $A^k,\, B^k,\, C_\lambda^k,\, C^k,\, Q^k : V^k \longrightarrow V^k$ by requiring

$$\begin{aligned}
(A^k u,\, v) &= a(u,\, v),\\
(B^k u,\, v) &= b(u,\, v),\\
(C_\lambda^k u,\, v) &= c_\lambda(u,\, v),\\
(C^k u,\, v) &= c(u,\, v),\\
(Q^k u,\, v) &= q^k(u,\, v)
\end{aligned}$$

for all u and v in V^k. Note that

$$C_\lambda^k = A^k - \lambda B^k \quad \text{and} \quad C^k = A^k - \lambda_1 B^k.$$

In practice, the operator Q^k will be constructed to "approximately invert" C^k_λ in some weak sense. The basic idea is that we want Q^k to satisfy the condition that it is an adequate approximate inverse of C^k_λ on oscillatory vectors in V^k. We will be precise about this condition in section 3.3.

Denote the eigenspace associated with λ^k_1 by

$$V^k_1 = \{u \in V^k : a(u, v) = \lambda^k_1 b(u, v) \quad \forall v \in V^k\}$$

and its *a-orthogonal* complement by

$$V^k_2 = \{u \in V^k : a(u, v) = 0 \quad \forall v \in V^k_1\}.$$

For any $u \in V^k$, define the a-orthogonal projection operators $P^k_1 : V^k \longrightarrow V^k_1$ and $P^k_2 : V^k \longrightarrow V^k_2$ by

$$a(P^k_1 u, v) = a(u, v) \quad \forall v \in V^k_1$$

and

$$a(P^k_2 u, v) = a(u, v) \quad \forall v \in V^k_2.$$

Consider the following multigrid algorithm for "solving" (2.7); more precisely, this algorithm attempts to reduce the error in V^l_2 only while keeping the V^l_1 approximation component essentially unchanged. Note that the multigrid algorithm as it is posed here is based on a direct solver for the coarsest grid problem ($l = 0$). Later we will allow for approximate solvers.

MG/ALGORITHM. *Let an initial approximation $u^{l,0} \in V^l$ and a right-hand side f^l be given. Then the new approximation $u^{l,2} = \mathbf{MG}/^l(u^{l,0}, f^l)$ is defined recursively as follows:*

(a) *If $l = 0$, then compute $u^{l,2} \in V^l$ so that*

$$P^l_1 u^{l,2} = P^l_1 u^{l,0} \quad \textit{and} \quad c_\lambda(u^{l,2}, v) = f^l(v) \quad \forall v \in V^l_2. \tag{2.10}$$

(b) *If $l > 0$, then perform the following:*

1. **Coarse-grid correction step.** *Denote the current residual by $r^{l,0}(v) = f^l(v) - c_\lambda(u^{l,0}, v)$ and define $f^{l-1} \in (V^{l-1})'$ as its restriction to V^{l-1}: $f^{l-1}(v) = r^{l,0}(v)\ \forall v \in V^{l-1}$. Then set*

$$u^{l,1} = u^{l,0} + \mathbf{MG}/^{l-1}(0, f^{l-1}).$$

2. **Smoothing step.** *Define $w^l \in V^l$ by*

$$q^l(w^l, v) = r^{l,1}(v) \quad \forall v \in V^l, \tag{2.11}$$

where $r^{l,1}(v) = f^l(v) - c_\lambda(u^{l,1}, v)$. Then set

$$u^{l,2} = u^{l,1} + w^l.$$

We have introduced this abstract form of the multigrid algorithm for two basic reasons. First, we want a general scheme that controls the potential instabilities found in many approaches to nearly singular equations, namely, those caused by shifts that get too close to coarse grid eigenvalues. Our coarsest grid solver in (2.10) and the smoothers we allow in section 3.3 prevent such difficulties. Second, this abstract algorithm applies naturally to RQMG, as we show in section 5.

3. Convergence of a multigrid method for nearly singular linear equations. In this section, we will analyze convergence of multigrid for the nearly singular linear problem (2.7). Our analysis uses an error decomposition introduced in section 3.2, which is based on an interpretation of solution and error that takes into account the objective of using (2.7) to solve the eigenvalue problem (2.5) with $k = l$. We have therefore not attempted to solve (2.7) in the usual sense; instead, our aim is to preserve the approximate magnitude of the components in the eigenspace associated with λ_1^l while attenuating error components in its a-orthogonal complement. Based on this error decomposition, we establish a smoothing property of relaxation in section 3.3 and a reduced approximation property of the discretization in section 3.4 which are combined in section 3.5. A V-cycle estimate will be developed in section 3.7 that follows from the recursive estimates developed in section 3.6. Again, the reader is strongly advised to keep in mind that the shift λ is floating in the sense that it is allowed to vary anywhere in the multigrid process, provided it remains within the bounds expressed by (2.8).

3.1. Preliminaries and assumptions. Fix $k \in \{0, 1, \cdots, l\}$. For any $u \in V^k$, the a-orthogonal projections P_1^k and P_2^k yield the unique decomposition

$$u = P_1^k u + P_2^k u. \tag{3.1}$$

Remark 3.1. Note that V_1^k and V_2^k are also b-orthogonal:

$$b(u, v) = 0 \quad \forall\, u \in V_1^k \quad \forall\, v \in V_2^k. \tag{3.2}$$

In this paper, $C > 0$ will denote a generic constant that does not depend on the number of levels l or any of the mesh sizes h_k.

Assumption 3.1. The norms $||\cdot||$ and $||\cdot||_k$ are uniformly equivalent on all V^k, that is, for any $v \in V^k$,

$$\frac{1}{C}||v|| \le ||v||_k \le C||v||. \tag{3.3}$$

Assumption 3.2. Assume that the following approximation properties hold:

(1) For any $u \in V_1^k$, there exists $v \in V_1^{k-1}$ such that

$$|||u - v||| \le C h_k |||u|||. \tag{3.4}$$

(2) For any $v \in V_1^{k-1}$, there exists $u \in V_1^k$ such that

$$|||v - u||| \le C h_k |||v|||. \tag{3.5}$$

Remark 3.2. Under the usual full regularity assumptions and subspace properties, standard finite element theory (cf. [9]) concludes that for any $0 \ne u \in V$ that minimizes $RQ(u)$ over V there exists $v \in V^l$ such that

$$|||u - v||| \le C h_l |||u|||.$$

(Note that C here depends on λ_1.) Property (3.4) then follows using the triangle inequality and this estimate with $l = k$ and $l = k - 1$. An analogous argument based on a corresponding estimate can be used to establish (3.5). The eigenvalue estimate in our first lemma below also follows from standard finite element theory, but we include its proof here to show that the estimate is actually a consequence of our general assumptions.

LEMMA 3.1 (eigenvalue approximation property).

$$\lambda_1^k \le \lambda_1^{k-1} \le \lambda_1^k + Ch_k^2.$$

Proof. The left inequality follows from noting that λ_1^i is the minimum of $RQ(u^i)$ over V^i, $i = k-1, k$, and that $V^{k-1} \subset V^k$. To prove the right inequality, note for any $u \in V_1^k$and $v \in V_1^{k-1}$ that

$$\begin{aligned} a(u-v,\, u-v) &= a(u-v,\, u) - a(u,\, v) + a(v,\, v) \\ &= \lambda_1^k b(u-v,\, u) - \lambda_1^k b(u,\, v) + \lambda_1^{k-1} b(v,\, v) \\ &= \lambda_1^k b(u-v,\, u-v) + (\lambda_1^{k-1} - \lambda_1^k) b(v,\, v) \\ &\ge (\lambda_1^{k-1} - \lambda_1^k) b(v,\, v) \\ &= \frac{\lambda_1^{k-1} - \lambda_1^k}{\lambda_1^{k-1}} a(v,\, v). \end{aligned}$$

Hence,

$$\lambda_1^{k-1} - \lambda_1^k \le \lambda_1^{k-1} \frac{|||u-v|||^2}{|||v|||^2}.$$

The lemma then follows from Assumption 3.2 and the fact that λ_1^{k-1} is bounded by a constant independent of h_k. ∎

Remark 3.3. A similar argument shows that

$$\lambda_1 \le \lambda_1^k \le \lambda_1 + Ch_k^2.$$

Together with (2.8), this implies that

$$0 \le \lambda - \lambda_1 \le Ch_0/l. \tag{3.6}$$

LEMMA 3.2 (norm equivalence). *It holds for any* $v \in V_2^k$ *that*

$$c(v,\, v) \le |||v|||^2 \le C\, c(v,\, v) \tag{3.7}$$

and for all sufficiently small h_0 *that*

$$c_\lambda(v,\, v) \le |||v|||^2 \le C\, c_\lambda(v,\, v). \tag{3.8}$$

Proof. It suffices to prove (3.8) since (3.7) may be proved in a similar way. The left inequality follows immediately from the fact that the bilinear form $b(\cdot,\, \cdot)$ is nonnegative. To prove the right inequality, note that since V_2^k is the a-orthogonal complement of V_1^k, then

$$a(v,\, v) \ge \lambda_2^k b(v,\, v).$$

Thus,

$$c_\lambda(v,\, v) \ge \left(1 - \frac{\lambda}{\lambda_2^k}\right) a(v,\, v).$$

Standard "mini-max" arguments (cf. [9]) show that $\lambda_2 \le \lambda_2^k \le \lambda_2^0$, from which follows

$$c_\lambda(v,\, v) \ge \frac{\lambda_2 - \lambda}{\lambda_2^0} a(v,\, v).$$

The proof is concluded using (2.8). ∎

We may thus define new norms on the spaces V_2^k $(k = 0, 1, \cdots, l)$, which are uniformly equivalent to the norm $|||\cdot|||$, by

$$||v||_c = \sqrt{c(v, v)} \quad \forall\, v \in V_2^k.$$

Remark 3.4. $||\cdot||_c$ is a seminorm on V^k. It is in fact a norm provided $\lambda_1^k \neq \lambda_1$.

3.2. Error decomposition. The following concept of error decomposition will be the basis of our analysis. Because we consider algorithms for the nearly singular problem (2.7) with the ultimate objective of solving (2.5) with $k = l$, the usual concepts of solution and error are not really relevant here. Indeed, we will not attempt to solve the problem per se but only to attenuate "error components" in the subspace V_2^l while preserving the approximate magnitude of the components in V_1^l. The purpose of this section is to make this precise. The main point to note here is that we cannot measure error in the usual direct sense because λ and hence what we even mean by the "solution" are floating: the error component in V_1^l is more or less well defined because $\lambda << \lambda_2^l$, so we will use a direct error norm to measure it. The error component in V_1^l is elusive, however, so instead of a direct measure we will use a residual norm that is easier to pinpoint and bound.

LEMMA 3.3 (an invariant subspace property). *Let f_2^k be a continuous linear functional on V^k that vanishes on V_1^k. Then for sufficiently small h_0 there exists a unique $\mathcal{U} \in V_2^k$ such that*

$$c_\lambda(\mathcal{U}, v) = f_2^k(v) \quad \forall\, v \in V^k. \tag{3.9}$$

Proof. It follows from the definition of the subspace V_2^k and from (2.8) that the bilinear form $c_\lambda(\cdot, \cdot)$ is positive definite on V_2^k. By the Lax–Milgram lemma there exists a unique solution $\mathcal{U} \in V_2^k$ of the variational problem

$$c_\lambda(\mathcal{U}, v) = f_2^k(v) \quad \forall\, v \in V_2^k.$$

The proof is completed by noting that $\mathcal{U}$ is orthogonal to V_1^k with respect to the bilinear form $c_\lambda(\cdot, \cdot)$. ∎

Let u be an approximation to the solution of the following problem: find $u^k \in V^k$ such that

$$c_\lambda(u^k, v) = f^k(v) \quad \forall\, v \in V^k \tag{3.10}$$

for a given functional $f^k : V^k \longrightarrow \mathcal{R}$. We then define the *residual* r as the functional on V^k given by

$$r(v) = f^k(v) - c_\lambda(u, v). \tag{3.11}$$

Suppose that we are given a decomposition

$$f^k = f_1^k + f_2^k, \tag{3.12}$$

where f_2^k vanishes on V_1^k. By Lemma 3.3, there exists a $\mathcal{U} \in V_2^k$ such that (3.9) holds. Write

$$u = u_1 + u_2, \qquad u_1 = P_1^k u, \qquad u_2 = P_2^k u.$$

Then

$$r(v) = f_1(v) - c_\lambda(u_1, v) + f_2(v) - c_\lambda(u_2, v) = r_1(v) + c_\lambda(e_2, v),$$

where

$$e_2 = \mathcal{U} - u_2 \tag{3.13}$$

and

$$r = r_1 + r_2 \tag{3.14}$$

with

$$r_1(v) = f_1^k(v) - c_\lambda(u_1, v), \tag{3.15}$$

$$r_2(v) = f_2^k(v) - c_\lambda(u_2, v) = c_\lambda(e_2, v). \tag{3.16}$$

We define the *size* of the error to be the pair

$$\begin{pmatrix} ||r_1|| \\ ||e_2||_c \end{pmatrix},$$

where $\|r_1\|$ is defined as the functional norm

$$||r_1|| = \sup_{v \in V^k} \frac{|r_1(v)|}{||v||}. \tag{3.17}$$

Remark 3.5. Note that the decomposition (3.12), and therefore the definition of the size of the error, is not unique. We allow this freedom to accommodate a general theory but will specify the decomposition later to suit our purposes.

3.3. Smoothing properties of relaxation. Recall that the smoothing step in the linear multigrid algorithm is defined as the replacement of the current approximation $u \in V^k$ by

$$\hat{u} = u + w,$$

with the correction $w \in V^k$ defined using q^k as a preconditioner:

$$q^k(w, v) = r(v) \quad \forall\, v \in V^k, \tag{3.18}$$

where r is the residual functional defined in (3.11). Also recall that $Q^k : V^k \longrightarrow V^k$ is the operator induced by the form q^k. Below we use $(Q^k)^T$ to denote its adjoint in $(\cdot, \cdot)$. Here we assume that r is decomposed according to (3.14) for some given decomposition of f according to (3.12). To analyze the smoother, we will need two additional assumptions.

Assumption 3.3 (properties of the forms a and b). Assume for all $u, v \in V^k$ that

$$a(u, v) \le Ch_k^{-2}||u||\; ||v||, \tag{3.19}$$

$$a(u, u) \ge \frac{1}{C}||u||^2, \tag{3.20}$$

$$b(u, v) \le C||u||\; ||v||, \tag{3.21}$$

$$b(u, u) \ge \frac{1}{C}||u||^2. \tag{3.22}$$

Remark 3.6. Let $\rho(\cdot)$ denote spectral radius; then (3.19)–(3.22) are equivalent to the respective inequalities

$$\rho(A^k) \le Ch_k^{-2}, \quad \rho\left((A^k)^{-1}\right) \le C, \quad \rho(B^k) \le C. \tag{3.23}$$

In view of Assumption 3.1, we also have the equivalent respective assumptions

$$||A^k||_k \le Ch_k^{-2}, \quad ||(A^k)^{-1}||_k \le C, \quad ||B^k||_k \le C, \tag{3.24}$$

which are naturally satisfied by a wide class of discrete elliptic problems.

Assumption 3.4 (smoother). Assume that $w \in V^k$ is defined uniquely by (3.18) and that

$$||w|| \le Ch_k^2||r||. \tag{3.25}$$

Further, suppose there exists a constant $\sigma > 0$ such that

$$Q^k + (Q^k)^T - C_\lambda^k \ge \frac{\sigma}{\rho(A^k)}(Q^k)^T Q^k. \tag{3.26}$$

Remark 3.7. Property (3.25) is equivalent to the inequality

$$||(Q^k)^{-1}||_k \le Ch_k^2. \tag{3.27}$$

Let $Q^k = D^k - L^k$, where D^k is self-adjoint in the inner product $(\cdot, \cdot)$ and D^k and L^k decompose C_λ^k according to $C_\lambda^k = D^k - L^k - (L^k)^T$. Assume that

$$\frac{1}{C}h_k^{-2}I \le D^k \le \sigma\rho(A^k)I \quad \text{and} \quad (L^k)^T L^k \le (1-\eta)(D^k)^2 \tag{3.28}$$

for some appropriate constant $\eta > 0$. (Self-adjoint operators E and F are said to satisfy the relation $E \ge F$ if $E - F$ is nonnegative definite.) Then (3.26) and (3.27) are easily verified for this choice of Q^k. This shows that Assumption 3.4 is natural for Gauss–Seidel relaxation applied to linear equations that satisfy (3.28), which is easily verified for Poisson's equation on a uniform grid, for example. For further discussion, see [6]. Note also that the choice $Q^k = \frac{1}{\omega}D^k$ satisfies (3.26) and (3.27) under these assumptions provided $\omega < 1$, which corresponds to damped Jacobi relaxation.

LEMMA 3.4 (properties of the smoother).

(i) *Define $z \in V^k$ as the solution of the problem*

$$q^k(z, v) = r_1^k(v) \quad \forall\, v \in V^k. \tag{3.29}$$

Then

$$|||z||| \le Ch_k||r_1^k||. \tag{3.30}$$

(ii) *Define $y \in V^k$ as the solution of the problem*

$$q^k(y, v) = c_\lambda(e_2^k, v) \quad \forall\, v \in V^k. \tag{3.31}$$

Then

$$|||P_1^k y||| \le Ch_k||e_2^k||_c \tag{3.32}$$

and, for sufficiently small h_0,

$$\begin{aligned} ||e_2^k - P_2^k y||_c^2 \le{}& \left(1 + C(\lambda - \lambda_1) + Ch_k^2|\lambda - \lambda_1^k|\right)||e_2^k||_c^2 \\ &-(1 + C(\lambda - \lambda_1))\frac{\sigma}{\rho(A^k)}||C_\lambda^k e_2^k||_k^2. \end{aligned} \tag{3.33}$$

Proof. (i) Bound (3.30) immediately follows from (3.19) and (3.25).

(ii) We first prove (3.32). From (3.16) we have $r_2^k(v) = c_\lambda(e_2^k, v)$. It therefore follows from (3.7) and (3.19) that

$$|r_2^k(v)| \le \|e_2^k\|_c \|v\|_c \le \|e_2^k\|_c \, |||v||| \le Ch_k^{-1} \|e_2^k\|_c \|v\|.$$

Hence,

$$||r_2^k|| \le Ch_k^{-1} ||e_2^k||_c. \tag{3.34}$$

Write $y_1 = P_1^k y$ and $y_2 = P_2^k y$. Using the fact that $y_1 \in V_1^k$ and relations (3.2), (3.21), (3.25), and (3.34), we get

$$|||y_1||| = \sqrt{\lambda_1^k b(y_1, y_1)} \le \sqrt{\lambda_1^k b(y, y)} \le C||y|| \le Ch_k^2 ||r_2^k|| \le Ch_k ||e_2^k||_c,$$

which is just (3.32).

To prove (3.33), note that

$$e_2^k - y = (I - (Q^k)^{-1} C_\lambda^k) e_2^k.$$

First note that by (3.26) we have

$$\begin{aligned}
&(I - (Q^k)^{-1} C_\lambda^k)^T C_\lambda^k (I - (Q^k)^{-1} C_\lambda^k) \\
&\quad = C_\lambda^k - C_\lambda^k \left((Q^k)^{-T} + (Q^k)^{-1} \right) C_\lambda^k + C_\lambda^k (Q^k)^{-T} C_\lambda^k (Q^k)^{-1} C_\lambda^k \\
&\quad = C_\lambda^k - C_\lambda^k (Q^k)^{-T} \left((Q^k)^T + Q^k - C_\lambda^k \right) (Q^k)^{-1} C_\lambda^k \\
&\quad \le C_\lambda^k - \frac{\sigma}{\rho(A^k)} (C_\lambda^k)^2.
\end{aligned} \tag{3.35}$$

This bound and the inequality $c_\lambda(e_2^k, e_2^k) \le c(e_2^k, e_2^k)$ imply that

$$\begin{aligned}
c_\lambda(e_2^k - y, e_2^k - y) &\le c_\lambda(e_2^k, e_2^k) - \frac{\sigma}{\rho(A^k)} ||C_\lambda^k e_2^k||_k^2 \\
&\le \|e_2^k\|_c^2 - \frac{\sigma}{\rho(A^k)} ||C_\lambda^k e_2^k||_k^2.
\end{aligned}$$

On the other hand, by (3.32) and (3.7) we have

$$\begin{aligned}
&c(e_2^k - y_2, e_2^k - y_2) \\
&\quad = c_\lambda(e_2^k - y_2, e_2^k - y_2) + (\lambda - \lambda_1) b(e_2^k - y_2, e_2^k - y_2) \\
&\quad = c_\lambda(e_2^k - y, e_2^k - y) - c_\lambda(y_1, y_1) + (\lambda - \lambda_1) b(e_2^k - y_2, e_2^k - y_2) \\
&\quad \le c_\lambda(e_2^k - y, e_2^k - y) + \frac{\lambda - \lambda_1^k}{\lambda_1^k} a(y_1, y_1) + \frac{\lambda - \lambda_1}{\lambda_2^k} a(e_2^k - y_2, e_2^k - y_2) \\
&\quad \le c_\lambda(e_2^k - y, e_2^k - y) + Ch_k^2 |\lambda - \lambda_1^k| \|e_2^k\|_c^2 + C(\lambda - \lambda_1) \|e_2^k - y_2\|_c^2.
\end{aligned}$$

We can thus conclude from the two inequalities above that

$$||e_2^k - y_2||_c^2 \le \frac{(1 + Ch_k^2 |\lambda - \lambda_1^k|) ||e_2^k||_c^2 - \frac{\sigma}{\rho(A^k)} ||C_\lambda^k e_2^k||_k^2}{1 - C(\lambda - \lambda_1)},$$

which with (2.8) gives (3.33) for all sufficiently small h_0. ∎

3.4. Properties of coarse grid correction. Define the projection operator $S^k : V^k \to V_2^{k-1}$ by

$$c(S^k u,\, v) = c(u,\, v) \quad \forall\, u \in V^k \quad \forall\, v \in V_2^{k-1}.$$

Further, define the projection onto the *c-orthogonal* complement of V_2^{k-1} by

$$T^k = I - S^k. \tag{3.36}$$

Remark 3.8. For any $v \in V^k$, we have

$$||v||_c^2 = ||S^k v||_c^2 + ||T^k v||_c^2 \tag{3.37}$$

and, consequently,

$$||S^k v||_c \le ||v||_c \quad \text{and} \quad \|T^k v\|_c \le \|v\|_c. \tag{3.38}$$

LEMMA 3.5 (stability of S^k).

$$|||S^k v||| \le C|||v||| \quad \forall\, v \in V^k. \tag{3.39}$$

Proof. This bound follows immediately from (3.7) and (3.38). □

LEMMA 3.6 (eigenvector approximation properties).

$$|||P_1^k v||| \le Ch_k|||v||| \quad \forall\, v \in V_2^{k-1} \tag{3.40}$$

and

$$|||P_2^k v||| \le Ch_k|||v||| \quad \forall\, v \in V_1^{k-1}. \tag{3.41}$$

Proof. Let $v \in V_2^{k-1}$ and choose $w \in V_1^{k-1}$, guaranteed by (3.4) to satisfy

$$|||P_1^k v - w||| \le Ch_k|||P_1^k v|||.$$

Then

$$\begin{aligned} a(P_1^k v, P_1^k v) &= a(v,\, P_1^k v) = a(v,\, P_1^k v - w) \\ &\le |||v|||\,|||P_1^k v - w||| \le Ch_k|||v|||\,|||P_1^k v|||, \end{aligned}$$

which establishes (3.40). A similar argument proves (3.41). □

LEMMA 3.7 (eigenvector approximation property).

$$|||P_1^k S^k v||| \le Ch_k\|v\|_c \quad \forall\, v \in V_2^k.$$

Proof. This is an immediate consequence of (3.40), (3.39), and (3.7). □

The following standard approximation assumption follows from H^2-elliptic regularity of the bilinear form $a(\cdot,\, \cdot)$. See, for example, [6].

Assumption 3.5 (standard approximation property). There exists a constant $\delta > 0$ such that for any $v \in V^k$, there exists $w \in V^{k-1}$ satisfying

$$|||v - w|||^2 \le \frac{\delta}{\rho(A^k)}\|A^k v\|_k^2.$$

THEOREM 3.1 (reduced approximation property). *There exists a constant $\delta' > 0$ such that for sufficiently small h_0,*

$$||T^k v||_c^2 \le \frac{\delta'}{\rho(A^k)}||C_\lambda^k v||_k^2 \quad \forall\, v \in V_2^k. \tag{3.42}$$

Proof. Let $v \in V_2^k$. According to the Cauchy–Schwarz inequality, Assumption 3.1, (3.20), and (3.8), we have

$$(C_\lambda^k v, v) \le \|C_\lambda^k v\|_k \|v\|_k \le C\|C_\lambda^k v\|_k |||v||| \le C\|C_\lambda^k v\|_k (C_\lambda^k v, v)^{\frac{1}{2}}.$$

Hence

$$(C_\lambda^k v, v)^{\frac{1}{2}} \le C\|C_\lambda^k v\|_k,$$

which together with (3.21), (3.20), and (3.8) implies that

$$\begin{aligned} \|A^k v\|_k &\le ||C_\lambda^k v||_k + \lambda||B^k v||_k \\ &\le ||C_\lambda^k v||_k + C\lambda(C_\lambda^k v, v)^{\frac{1}{2}} \\ &\le C||C_\lambda^k v||_k. \end{aligned} \tag{3.43}$$

Let $\tilde{P} : V^k \to V^{k-1}$ be the a-orthogonal projection operator. Then, by Assumption 3.5 and (3.43), there exists a constant $\delta > 0$ such that

$$|||v - \tilde{P}v|||^2 \le \frac{\delta}{\rho(A^k)} \|A^k v\|_k^2 \le \frac{C}{\rho(A^k)} \|C_\lambda^k v\|_k^2. \tag{3.44}$$

Let $u \in V_1^k$ be such that (3.5) holds for v replaced by $P_1^{k-1}\tilde{P}v$. Then

$$\begin{aligned} |||P_1^{k-1}\tilde{P}v|||^2 &= a(P_1^{k-1}\tilde{P}v, \tilde{P}v) = a(P_1^{k-1}\tilde{P}v, v) = a(P_1^{k-1}\tilde{P}v - u, v) \\ &\le |||P_1^{k-1}\tilde{P}v - u|||\,|||v||| \le Ch_k|||P_1^{k-1}\tilde{P}v|||\,|||v|||. \end{aligned}$$

Hence

$$|||P_1^{k-1}\tilde{P}v||| \le Ch_k|||v||| \le Ch_k\|A^k v\|_k \le Ch_k\|C_\lambda^k v\|_k,$$

where we used (3.43). Now from this, (3.7), (3.44), and (3.23) we have

$$\begin{aligned} \|T^k v\|_c^2 &\le \|v - P_2^{k-1}\tilde{P}v\|_c^2 \\ &\le 2(|||v - \tilde{P}v|||^2 + |||P_1^{k-1}\tilde{P}v|||^2) \\ &\le \frac{C}{\rho(A^k)}\|C_\lambda^k v\|_k^2. \end{aligned}$$

This completes the proof. ∎

3.5. Combined smoothing and approximation properties.

LEMMA 3.8 (reduced approximation property). *Let $w \in V^k$ be the solution of the problem*

$$q^k(w, v) = r_1^k(v) + c_\lambda(e_2^k, v) \quad \forall\, v \in V^k.$$

Then for sufficiently small h_0,

$$\begin{aligned} ||e_2^k - P_2^k w||_c \le \{&(1 + C(\lambda - \lambda_1))(||S^k e_2^k||_c^2 + \beta^2||T^k e_2^k||^2) \\ &+ Ch_k^2|\lambda - \lambda_1^k|\,\|e_2^k\|_c^2\}^{\frac{1}{2}} + Ch_k\|r_1^k\| \end{aligned}$$

with

$$\beta^2 = 1 - \frac{\sigma}{\delta'} < 1,$$

where σ is given in (3.26) *and δ' is given in* (3.42).

Proof. Let z and y be the solutions of the problems (3.29) and (3.31), respectively, so that $w = z + y$. The triangle inequality and (3.7) then imply that

$$\|e_2^k - P_2^k w\|_c \le \|e_2^k - P_2^k y\|_c + |||z|||.$$

The proof of the lemma now follows from (3.33), (3.30), Theorem 3.1, and (3.37). ∎

Before we estimate the error in the multigrid algorithm, we need to estimate the error quantities just before the coarse grid correction. To this end, recall in the correction step of **MG**/ that the functional $f^{k-1} \in (V^{k-1})'$ is the restriction of the residual functional $r^{k,0}$ to V^{k-1}:

$$f^{k-1}(v) = r^{k,0}(v) = f^k(v) - c_\lambda(u^{k,0}, v) \quad \forall\, v \in V^{k-1}.$$

Now

$$u_1^{k,0} = P_1^k u^{k,0} \quad \text{and} \quad u_2^{k,0} = P_2^k u^{k,0}.$$

A given decomposition of f^k according to (3.12) yields a decomposition of $r^{k,0}$ according to (3.14)–(3.16):

$$r^{k,0}(v) = r_1^{k,0}(v) + r_2^{k,0}(v) \quad \forall\, v \in V^k,$$

where

$$r_1^{k,0}(v) = f_1^k(v) - c_\lambda(u_1^{k,0}, v), \quad r_2^{k,0}(v) = c_\lambda(e_2^{k,0}, v), \quad e_2^{k,0} = \mathcal{U}_2^k - u_2^{k,0},$$

and $\mathcal{U}_2^k$ is the solution of the problem (3.9). For any $v \in V^{k-1}$,

$$\begin{aligned} r_2^{k,0}(v) &= c(e_2^{k,0}, v) - (\lambda - \lambda_1) b(e_2^{k,0}, v) \\ &= c(e_2^{k,0}, P_2^{k-1} v) - (\lambda - \lambda_1) b(e_2^{k,0}, v) + c(e_2^{k,0}, P_1^{k-1} v) \\ &= c(S^k e_2^{k,0}, P_2^{k-1} v) - (\lambda - \lambda_1) b(e_2^{k,0}, v) + c(e_2^{k,0}, P_1^{k-1} v). \end{aligned}$$

We are therefore led to choose

$$f_1^{k-1}(v) = r_1^{k,0}(v) - (\lambda - \lambda_1) b(e_2^{k,0}, v) + c(e_2^{k,0}, P_1^{k-1} v) \tag{3.45}$$

and

$$f_2^{k-1}(v) = c(S^k e_2^{k,0}, P_2^{k-1} v). \tag{3.46}$$

This provides a legitimate coarse grid decomposition according to (3.12) because the functional f_2^{k-1} vanishes on V_1^{k-1} and

$$f^{k-1} = f_1^{k-1} + f_2^{k-1}.$$

Let $\mathcal{U}_2^{k-1} \in V_2^{k-1}$ be the solution of the problem

$$c_\lambda(\mathcal{U}_2^{k-1}, v) = f_2^{k-1}(v) \quad \forall\, v \in V_2^{k-1}. \tag{3.47}$$

Then $r^{k-1,0}(v) = f^{k-1}(v) - c_\lambda(u^{k-1,0}, v)$ has a corresponding decomposition according to (3.14)–(3.16):

$$r_1^{k-1,0}(v) = f_1^{k-1}(v) - c_\lambda(u_1^{k-1,0}, v) \quad \text{and} \quad r_2^{k-1,0}(v) = c_\lambda(\mathcal{U}_2^{k-1} - u_2^{k-1,0}, v).$$

LEMMA 3.9 (initial coarse grid error estimates). *The initial coarse grid error satisfies*

$$u^{k-1,\,0} = 0, \mathcal{U}_2^{k-1} = e_2^{k-1,\,0}, r_i^{k-1,\,0} = f_i^{k-1},\ i = 1, 2, \tag{3.48}$$

$$\|\mathcal{U}_2^{k-1}\|_c \le (1 + C(\lambda - \lambda_1))\, \|S^k e_2^{k,\,0}\|_c \ \textit{ for sufficiently small } h_0, \tag{3.49}$$

$$\|\mathcal{U}_2^{k-1} - S^k e_2^{k,\,0}\|_c \le C(\lambda - \lambda_1)\|e_2^{k,\,0}\|_c, \tag{3.50}$$

and

$$||r_1^{k-1,\,0}|| = ||f_1^{k-1}|| \le ||r_1^{k,\,0}|| + (C(\lambda - \lambda_1) + Ch_k)||e_2^{k,\,0}||_c. \tag{3.51}$$

Proof. Relations (3.48) follow immediately from the definitions of $\mathbf{MG}/^k$, $e_i^{k-1,\,0}$, and $r_i^{k-1,\,0}$, $i = 1, 2$.

Bound (3.49) follows from using (3.47), (3.46), the Cauchy–Schwarz inequality, (3.38), (3.21), (3.20), and (3.7), to conclude that

$$\begin{aligned}\|\mathcal{U}_2^{k-1}\|_c^2 &= c_\lambda(\mathcal{U}_2^{k-1}, \mathcal{U}_2^{k-1}) + (\lambda - \lambda_1) b(\mathcal{U}_2^{k-1}, \mathcal{U}_2^{k-1})\\ &= c(S^k e_2^{k,\,0}, \mathcal{U}_2^{k-1}) + (\lambda - \lambda_1) b(\mathcal{U}_2^{k-1}, \mathcal{U}_2^{k-1})\\ &\le \|S^k e_2^{k,\,0}\|_c \|\mathcal{U}_2^{k-1}\|_c + C(\lambda - \lambda_1)\|\mathcal{U}_2^{k-1}\|_c^2.\end{aligned}$$

By (3.46) and (3.47), we have

$$c(\mathcal{U}_2^{k-1} - S^k e_2^{k,\,0},\, v) = (\lambda - \lambda_1) b(S^k e_2^{k,\,0},\, v) \quad \forall\, v \in V_2^{k-1}.$$

Choosing $v = \mathcal{U}_2^{k-1} - S^k e_2^{k,\,0}$, then (3.50) follows from (3.21), (3.20), (3.7), and (3.38).

To prove (3.51), let $v \in V^{k-1}$ and $u \in V_1^k$ be arbitrary. Then from (3.45), the Cauchy–Schwarz inequality, (3.21), and (3.7), we have

$$\begin{aligned}|f_1^{k-1}(v)| &\le |r_1^{k,\,0}(v)| + (\lambda - \lambda_1)|b(e_2^{k,0},\, v)| + |c(e_2^{k,0},\, P_1^{k-1} v)|\\ &\le |r_1^{k,\,0}(v)| + C(\lambda - \lambda_1)\|e_2^{k,0}\|\,\|v\| + |c(e_2^{k,0},\, P_1^{k-1}v - u)|\\ &\le |r_1^{k,\,0}(v)| + C(\lambda - \lambda_1)|||e_2^{k,0}|||\,\|v\| + C|||e_2^{k,0}|||\,|||P_1^{k-1}v - u|||.\end{aligned}$$

Choosing $u \in V_1^k$ according to (3.5) with $P_1^{k-1} v$ in place of v, it follows from (3.8) that

$$|f_1^{k-1}(v)| \le |r_1^{k,\,0}(v)| + C(\lambda - \lambda_1)|||e_2^{k,0}|||\,\|v\| + Ch_k\|e_2^{k,0}\|_c|||P_1^{k-1}v|||.$$

Now since $\lambda_1^{k-1} \le \lambda_1^0 = C$, then (3.21) implies that

$$|||P_1^{k-1}v||| = \sqrt{\lambda_1^{k-1} b(P_1^{k-1}v,\, P_1^{k-1}v)} \le \sqrt{\lambda_1^{k-1} b(v,\, v)} \le C\|v\|.$$

The above two inequalities imply (3.51). ∎

3.6. Recursive estimate. For convenience in what follows, we define the relation

$$\begin{pmatrix} p \\ q \end{pmatrix} \le \begin{pmatrix} r \\ s \end{pmatrix}, \quad p,\, q,\, r,\, s \in \mathcal{R},$$

to mean that $p \le r$ and $q \le s$. For $k = 0, 1, \ldots, l$, with β_k and ε_i^k, $i = 1, 2, 3$ given, let

$$\mathcal{E}^k = \begin{pmatrix} \varepsilon_1^k & \varepsilon_2^k \\ \varepsilon_3^k & \beta_k \end{pmatrix}. \tag{3.52}$$

THEOREM 3.2 (recursive estimate). *Assume that there is a constant C_0 such that $\mathbf{MG}/^{k-1}$ satisfies*

$$\begin{pmatrix} |||u_1^{k-1,2} - u_1^{k-1,0}||| \\ ||e_2^{k-1,2}||_c \end{pmatrix} \le \mathcal{E}^{k-1} \begin{pmatrix} ||r_1^{k-1,0}|| \\ ||e_2^{k-1,0}||_c \end{pmatrix} \tag{3.53}$$

with

$$0 \le \varepsilon_i^{k-1} \le C_0, \;\; i = 1, 2, 3 \tag{3.54}$$

and

$$1 > \beta_{k-1} \ge \beta > 0. \tag{3.55}$$

Then $\mathbf{MG}/^k$ satisfies

$$\begin{pmatrix} |||u_1^{k,2} - u_1^{k,0}||| \\ ||e_2^{k,2}||_c \end{pmatrix} \le \mathcal{E}^k \begin{pmatrix} ||r_1^{k,0}|| \\ ||e_2^{k,0}||_c \end{pmatrix}, \tag{3.56}$$

with

$$\varepsilon_i^{k-1} \le \varepsilon_i^k \le \varepsilon_i^{k-1} + Ch_k + C(\lambda - \lambda_1), i = 1, 2, 3 \tag{3.57}$$

and

$$\beta_{k-1} \le \beta_k \le \beta_{k-1} + Ch_k + C(\lambda - \lambda_1). \tag{3.58}$$

Proof. From (3.53), (3.49), (3.51), and (3.38) we have

$$|||u_1^{k-1,2} - u_1^{k-1,0}||| \le \varepsilon_1^{k-1} ||r_1^{k,0}|| + \varepsilon_2^{k-1} \left(1 + C(\lambda - \lambda_1)\right) ||e_2^{k,0}||_c \tag{3.59}$$

and

$$||e_2^{k-1,2}||_c \le \varepsilon_3^{k-1} ||r_1^{k,0}|| + \beta_{k-1} \left(1 + C(\lambda - \lambda_1)\right) ||S^k e_2^{k,0}||_c. \tag{3.60}$$

Letting

$$u_1^{k-1,2} = P_1^{k-1} u^{k-1,2} \quad \text{and} \quad u_2^{k-1,2} = P_2^{k-1} u^{k-1,2},$$

it then follows from (3.48) and (3.41) that

$$|||P_1^k u_1^{k-1,2}||| \le |||u_1^{k-1,2} - u_1^{k-1,0}||| \tag{3.61}$$

and

$$|||P_2^k u_1^{k-1,2}||| \le Ch_k |||u_1^{k-1,2} - u_1^{k-1,0}|||. \tag{3.62}$$

Now (3.13), (3.7), and (3.49) (noting that $\lambda - \lambda_1 \le C$) imply that

$$\begin{aligned} |||u_2^{k-1,2}||| &= |||\mathcal{U}_2^{k-1} - e_2^{k-1,2}||| \\ &\le C(\|\mathcal{U}_2^{k-1}\|_c + \|e_2^{k-1,2}\|_c) \le C(\|e_2^{k,0}\|_c + \|e_2^{k-1,2}\|_c), \end{aligned}$$

which together with (3.40) gives

$$|||P_1^k u_2^{k-1,2}||| \le Ch_k(\|e_2^{k,0}\|_c + \|e_2^{k-1,2}\|_c). \tag{3.63}$$

By definition, $u^{k,1} = u^{k,0} + u^{k-1,2}$. For any $v \in V^k$, decomposing $r^{k,1}(v) = f^k(v) - c_\lambda(u^{k,1}, v)$ according to (3.14)–(3.16) yields

$$\begin{aligned} r_1^{k,1}(v) &= f_1^k(v) - c_\lambda(u_1^{k,1}, v) \\ &= r_1^{k,0}(v) - c_\lambda(P_1^k u^{k-1,2}, v) \\ &= r_1^{k,0}(v) + (\lambda - \lambda_1^k) b(P_1^k u^{k-1,2}, v). \end{aligned}$$

Hence by the Cauchy–Schwarz inequality, (3.21), and (3.20) we get

$$\|r_1^{k,1}\| \le \|r_1^{k,0}\| + C|\lambda - \lambda_1^k| |||P_1^k u^{k-1,2}|||. \tag{3.64}$$

Letting $\tilde{u} = P_1^k u_2^{k-1,2} - P_2^k u_1^{k-1,2}$, then

$$|||\tilde{u}||| \le |||P_1^k u_2^{k-1,2}||| + |||P_2^k u_1^{k-1,2}||| \tag{3.65}$$

and because $u_2^{k-1,2} - P_2^k u_2^{k-1,2} = P_1^k u_2^{k-1,2}$ and $P_2^k u_2^{k-1,2} - P_2^k u^{k-1,2} = -P_2^k u_1^{k-1,2}$, we obtain

$$\begin{aligned} e_2^{k,1} &= e_2^{k,0} - P_2^k u^{k-1,2} \\ &= (e_2^{k,0} - \mathcal{U}_2^{k-1}) + (\mathcal{U}_2^{k-1} - u_2^{k-1,2}) + (u_2^{k-1,2} - P_2^k u_2^{k-1,2}) \\ &\quad + (P_2^k u_2^{k-1,2} - P_2^k u^{k-1,2}) \\ &= (e_2^{k,0} - \mathcal{U}_2^{k-1}) + e_2^{k-1,2} + \tilde{u}. \end{aligned} \tag{3.66}$$

From this, (3.38), and (3.50) we get

$$\begin{aligned} \|e_2^{k,1}\|_c &\le \|T^k e_2^{k,0}\|_c + \|S^k e_2^{k,0} - \mathcal{U}_2^{k-1}\|_c + \|e_2^{k-1,2}\|_c + \|\tilde{u}\|_c \\ &\le (1 + C(\lambda - \lambda_1)) \|e_2^{k,0}\|_c + \|e_2^{k-1,2}\|_c + |||\tilde{u}|||. \end{aligned} \tag{3.67}$$

We can similarly conclude that

$$\|S^k e_2^{k,1}\|_c \le C(\lambda - \lambda_1)\|e_2^{k,0}\|_c + \|e_2^{k-1,2}\|_c + |||\tilde{u}||| \tag{3.68}$$

and

$$\|T^k e_2^{k,1}\|_c \le \|T^k e_2^{k,0}\|_c + |||\tilde{u}|||. \tag{3.69}$$

To estimate the size of $e_2^{k,2} = \mathcal{U}_2^k - P_2^k u^{k,2} = e_2^{k,1} - P_2^k w$, where $w \in V^k$ solves $q^k(w, v) = r_1^k(v) + c_\lambda(e_2^{k,1}, v) \ \forall v \in V^k$, note that Lemma 3.8 implies that

$$\begin{aligned} \|e_2^{k,2}\|_c \le (1 + C(\lambda - \lambda_1))^{\frac{1}{2}} &\left(\|S^k e_2^{k,1}\|_c^2 + \beta^2 \|T^k e_2^{k,1}\|_c^2\right)^{\frac{1}{2}} \\ &+ Ch_k|\lambda - \lambda_1^k|^{\frac{1}{2}} \|e_2^{k,1}\|_c + Ch_k\|r_1^{k,1}\|. \end{aligned}$$

By using (3.59)–(3.69), (3.54), and (3.55), a lengthy computation shows that

$$\|e_2^{k,2}\|_c \le \varepsilon_3^k \|r_1^{k,0}\| + \beta_k \|e_2^{k,0}\|_c,$$

with ε_3^k and β_k satisfying (3.57).

It remains to estimate $|||u_1^{k,2} - u_1^{k,0}|||$. We have from (3.30) and (3.32) that

$$\begin{aligned}|||u_1^{k,2} - u_1^{k,0}||| &\le |||u_1^{k,2} - u_1^{k,1}||| + |||u_1^{k,1} - u_1^{k,0}||| \\ &\le Ch_k\left(\|r_1^{k,1}\| + \|e_2^{k,1}\|_c\right) + |||P_1^k u_1^{k-1,2}||| + |||P_1^k u_2^{k-1,2}|||.\end{aligned}$$

It thus follows from (3.64), (3.67), (3.61), (3.63), (3.65), (3.59), and (3.60) that

$$|||u_1^{k,2} - u_1^{k,0}||| \le (\varepsilon_1^{k-1} + Ch_k)||r_1^{k,0}|| + (\varepsilon_2^{k-1} + C(\lambda - \lambda_1) + Ch_k)||e_2^{k,0}||_c,$$

which completes the proof. ☐

3.7. Linear V-cycle estimate.

THEOREM 3.3 (linear V-cycle estimate). *Suppose that* $\mathbf{MG}/^0$ *satisfies*

$$\begin{pmatrix} |||u_1^{0,2} - u_1^{0,0}||| \\ ||e_2^{0,2}||_c \end{pmatrix} \le \mathcal{E}^0 \begin{pmatrix} ||r_1^{0,0}|| \\ ||e_2^{0,0}||_c \end{pmatrix} \tag{3.70}$$

independently of h_0 *and of the choice of the decomposition* $f^0 = f_1^0 + f_2^0$ *for some positive constants* ε_i^0, $1 > \beta_0 \ge \beta > 0$. *Then for sufficiently small* h_0 *and* $\lambda - \lambda_1$, $\mathbf{MG}/^l$ *satisfies*

$$\begin{pmatrix} |||u_1^{l,2} - u_1^{l,0}||| \\ ||e_2^{l,2}||_c \end{pmatrix} \le \mathcal{E}^l \begin{pmatrix} ||r_1^{l,0}|| \\ ||e_2^{l,0}||_c \end{pmatrix}$$

with

$$\varepsilon_i^0 \le \varepsilon_i^l \le \varepsilon_i^0 + Ch_0 + C(\lambda - \lambda_1)l, \quad i = 1, 2, 3, \tag{3.71}$$

and

$$\beta_0 \le \beta_l \le \beta_0 + Ch_0 + C(\lambda - \lambda_1)l. \tag{3.72}$$

Proof. This follows directly from recursion on Theorem 3.2. ☐

LEMMA 3.10 (direct solver). *Let* $\mathbf{MG}/^0$ *be defined by* (2.10). *Then* (3.70) *holds with* $\varepsilon_1^0 = \varepsilon_2^0 = \beta_0 = 0$ *and* ε_3^0 *a constant independent of* h_0.

Proof. Relations (2.10) are equivalent to assuming that $u^{0,2} = u_1^{0,0} + u_2^{0,2}$, where $u_2^{0,2} \in V_2^0$ is defined by

$$c_\lambda(u_2^{0,2}, v) = f^0(v) \quad \forall\, v \in V_2^0. \tag{3.73}$$

Clearly then,

$$u_1^{0,2} - u_1^{0,0} = 0,$$

that is, $\varepsilon_1^0 = \varepsilon_2^0 = 0$. Let $f^0 = f_1^0 + f_2^0$ be a decomposition according to (3.12) and $\mathcal{U}^0 \in V_2^0$ be the solution of (3.50) with $k = 1$. Then $e_2^{0,2} = \mathcal{U}^0 - u_2^{0,2}$ satisfies

$$c_\lambda(e_2^{0,2}, e_2^{0,2}) = f_1^0(e_2^{0,2}) = r_1^0(e_2^{0,2}),$$

which together with (3.7), (3.8), and (3.20) implies that

$$\|e_2^{0,2}\|_c^2 \le C\, c_\lambda(e_2^{0,2},\, e_2^{0,2}) = C\, r_1^0(e_2^{0,2}) \le C\, ||r_1^0||\, \|e_2^{0,2}\|_c.$$

Hence, $\beta_0 = 0$ and ε_3^0 is a constant independent of h_0. This completes the proof. ∎

Remark 3.9. Since (3.70) holds with $\beta_0 = 0$, it certainly holds with $\beta_0 = \beta$.

LEMMA 3.11 (approximate solver). *Let* $\mathbf{MG}/^0$ *in the definition of* $\mathbf{MG}/$ *be replaced by a mapping*

$$M^0 : f^0 \to u^{0,2}$$

such that for any f^0,

$$|||u^{0,2}||| \le C \sup_{0 \ne v \in V^0} \frac{|f^0(v)|}{|||v|||} \tag{3.74}$$

and for any f_2^0 *that vanishes on* V_1^0,

$$||P_2^0 M^0 f_2^0 - \mathcal{U}^0||_c \le \beta_0 ||\mathcal{U}^0||_c, \tag{3.75}$$

where $\mathcal{U}^0$ *is defined by* (3.9) *with* $k = 0$. *Then* (3.70) *holds with* $\varepsilon_i \le C$, $i = 1, 2, 3$, *and* β_0 *from* (3.75).

Proof. We have

$$|||u_1^{0,2} - u_1^{0,0}||| = |||u_1^{0,2}||| = |||P_1^0 u^{0,2}||| \le |||P_1^0 M^0 f_1^0||| + |||P_1^0 M^0 f_2^0|||.$$

From (3.74) and the fact that $f_1^0 = r_1^{0,0}$ we conclude that

$$|||P_1^0 M^0 f_1^0||| \le C||r_1^{0,0}||.$$

It follows from (3.74) and the definition of $\mathcal{U}^0$ that

$$|||P_1^0 M^0 f_2^0||| \le C \sup_{0 \ne v \in V^0} \frac{|f_2^0(v)|}{|||v|||} \le C \sup_{0 \ne v \in V_2^0} \frac{|f_2^0(v)|}{|||v|||} \le C\|\mathcal{U}^0\|_c.$$

The above inequalities imply that $\varepsilon_i^0 \le C$ for $i = 1, 2$. Since $e_2^{0,2} = \mathcal{U}^0 - u_2^{0,2} = \mathcal{U}^0 - P_2^0 M^0 f^0$, from (3.75) and (3.74) we have

$$\|e_2^{0,2}\|_c \le \|\mathcal{U}^0 - P_2^0 M^0 f_2^0\|_c + \|P_2^0 M^0 f_1^0\|_c \le \beta_0\|\mathcal{U}^0\|_c + C\,\|r_1^{0,0}\|.$$

This completes the proof. ∎

4. Convergence of a linearized multigrid method for eigenvalue problems. Here we apply the theory developed in the previous section to a linearized multigrid method for solving the eigenvalue problem (2.5) with $k = l$. This method uses an outer loop iterative process, which replaces the given approximate eigenfunction $u_{\text{old}}^l \in V^l$ by the new approximation $u_{\text{new}}^l \in V^l$ that is computed by applying $\mathbf{MG}/$ to the problem

$$\lambda = RQ(u^l)\,, \tag{4.1}$$

$$a(u^l,\, v) - \lambda b(u^l,\, v) = 0 \quad \forall v \in V^l\,. \tag{4.2}$$

Assume for simplicity that $\mathbf{MG}/$ uses an exact coarsest grid solver.

Here we are considering a conventional linearized multigrid method applied to the eigenvalue problem, so the shift λ is now considered to be fixed throughout each multigrid cycle, allowed to change only *between* cycles via (4.1). However, the reader should note that changes could be allowed anywhere within the multigrid cycles, and the proof of our next theorem would still apply. We will use this observation in proving our theorem on RQMG. The reader should also note that this linearized algorithm relies on the fact that no multigrid cycle will produce the exact solution of (4.2) (i.e., we will show that $u^l_{\text{new}} \neq 0$).

LMG/ALGORITHM. *Given an initial approximation* $u^l_{\text{old}} \in V^l$ *to an eigenvector for* (2.5) *belonging to* λ^l *such that* $|||u^l_{\text{old}}||| = 1$*, then the new approximation* $u^l_{\text{new}} = \mathbf{LMG}/^l(u^l_{\text{old}})$ *is defined as follows:*

1. **MG inner solver**. *Perform one step of* $\mathbf{MG}/^l$ *applied to* (4.2) *using the initial approximation* $u^{l,0} = u^l_{\text{old}}$ *and the shift* $\lambda = RQ(u^l_{old})$*:*

$$u^{l,\,2} = \mathbf{MG}/^l(u^{l,\,0},\, 0)\ .$$

2. **Normalization**. *Set*

$$u^l_{\text{new}} = \frac{u^{l,\,2}}{|||u^{l,\,2}|||}\ .$$

To analyze the convergence of **LMG**/, we define the error in an approximation $u^l = u^l_1 + u^l_2, u^l_i = P^l_i u^l$, $i = 1,\, 2$, to an eigenvector of (2.5) belonging to λ^l_1 by

$$e^l = \frac{u^l_2}{|||u^l_1|||}\ . \tag{4.3}$$

Note that $\lambda^l_1 \leq \lambda = RQ(u^{l,\,0})$.

THEOREM 4.1 (convergence of linearized multigrid). $\mathbf{LMG}/^l$ *converges according to the estimate*

$$\|e^l_{\text{new}}\|_c \leq (Ch_0 + \beta)\,\|e^l_{\text{old}}\|_c \tag{4.4}$$

for sufficiently small h_0.

Proof. First note that the definition of λ yields

$$c_\lambda(u^{l,\,0},\, u^{l,\,0}) = a(u^{l,\,0},\, u^{l,\,0}) - \lambda b(u^{l,\,0},\, u^{l,\,0}) = 0.$$

Thus writing $u^{l,\,0} = u^{l,\,0}_1 + u^{l,\,0}_2$, where $u^{l,\,0}_i = P^l_i u^{l,\,0}$, $i = 1,\, 2$, then

$$c_\lambda(u^{l,\,0}_1,\, u^{l,\,0}_1) + c_\lambda(u^{l,\,0}_2,\, u^{l,\,0}_2) = 0\ ,$$

which can be rewritten

$$\|u^{l,\,0}_2\|^2_c = \eta\,|||u^{l,\,0}_1|||^2 + (\lambda - \lambda_1)\,b(u^{l,\,0}_2,\, u^{l,\,0}_2)\ .$$

Letting $\eta = (\lambda - \lambda^l_1)/\lambda^l_1$, then from this relation and (2.8) we can conclude for sufficiently small h_0 that

$$\|u^{l,\,0}_2\|^2_c \leq C\eta|||u^{l,\,0}_1|||^2 \tag{4.5}$$

and that

$$|||u^{l,\,0}_1|||^2 \leq \frac{1}{\eta}\|u^{l,\,0}_2\|^2_c\ . \tag{4.6}$$

Now let $f^l(v) = f_1^l(v) = f_2^l(v) = 0 \ \ \forall v \in V^l$. Then f_2^l certainly vanishes on V_1^l, and noting that $\mathcal{U} = 0$ in (3.9) with $k = l$, we apply Theorem 3.3, the Cauchy–Schwarz inequality, and (4.5) to conclude that

$$\begin{aligned}
|||u_1^{l,\,2} - u_1^{l,\,0}||| &\le \varepsilon_1^l \sup_{0\neq v\in V^l} \frac{|f_1^k(v) - c_\lambda(u_1^{l,\,0}, v)|}{\|v\|} + \varepsilon_2^l \|\mathcal{U} - u_2^{l,\,0}\|_c \\
&= (\lambda - \lambda_1^l)\, \varepsilon_1^l \sup_{0\neq v\in V^l} \frac{|b(u_1^{l,\,0}, v)|}{\|v\|} + \varepsilon_2^l \|u_2^{l,\,0}\|_c \\
&\le \varepsilon^l \left(C(\lambda - \lambda_1^l)\|u_1^{l,\,0}\| + \|u_2^{l,\,0}\|_c \right) \\
&\le C\sqrt{\eta}\varepsilon^l |||u_1^{l,\,0}||| \,. \qquad (4.7)
\end{aligned}$$

Similarly, we apply Theorem 3.3 and (4.6) to obtain

$$\begin{aligned}
\|u_2^{l,\,2}\|_c &= \|\mathcal{U} - u_2^{l,\,2}\|_c \\
&\le C\varepsilon^l \eta\, |||u_1^{l,\,0}||| + \beta_l \|u_2^{l,\,0}\|_c \\
&\le \left(C\sqrt{\eta}\varepsilon^l + \beta^l \right) \|u_2^{l,\,0}\|_c \,. \qquad (4.8)
\end{aligned}$$

From (4.7) we can conclude that

$$\begin{aligned}
|||u_1^{l,\,2}||| &\ge |||u_1^{l,\,0}||| - |||u_1^{l,\,2} - u_1^{l,\,0}||| \\
&\ge \left(1 - C\sqrt{\eta}\varepsilon^l\right) |||u_1^{l,\,0}||| \,. \qquad (4.9)
\end{aligned}$$

Bounds (4.8) and (4.9) now yield

$$\|e_{\rm new}^l\|_c \le \left(\frac{C\sqrt{\eta}\varepsilon^l + \beta^l}{1 - C\sqrt{\eta}\varepsilon^l} \right) \|e_{\rm old}^l\|_c. \qquad (4.10)$$

The proof now follows from (3.6). ∎

Remark 4.1. The theory here for **MG**/ assumes an exact coarsest grid solver. In fact, the theory holds if we assume only that the coarsest grid solver is effective enough that it does not contaminate the estimates in (3.71) and (3.72); see Lemma 3.11. In any event, this theorem shows that the **LMG**/ worst case convergence factor is arbitrarily close to the multigrid factor for standard well-posed linear elliptic equations provided (2.8) holds. We will show in section 6 that (2.8) can be guaranteed by the use of a *full multigrid* or *nested iteration* process.

5. Convergence of RQMG for eigenvalue problems. In this section, we apply the theory of section 3 to RQMG, which is a direct multigrid method for (2.5) involving minimization of the Rayleigh quotient over level k corrections to the fine-grid eigenvector approximation. The approximate eigenvalue is the Rayleigh quotient of the eigenvector approximation incorporating all current corrections, so it changes during each level k correction. The main point in establishing RQMG convergence is simply to recognize that the proof of Theorem 4.1 made no use of the fact that the shift λ was assumed to be fixed in the multigrid cycles.

Following is the definition of **RQMG**/, which (as MG was) is posed with an exact coarsest grid solver and a general smoother. (Again, approximate coarsest grid solvers can be treated as in Lemma 3.11.) To accommodate recursion, we have defined **RQMG**/ in terms of a given vector $u \in V$, which should be interpreted as the finest level eigenvector approximation.

RQMG/ALGORITHM. *Given an initial correction* $u^{l,0} \in V^l$ *for the approximate eigenvector* $u \in V^l$*, then the new correction* $u^{l,2} = \mathbf{RQMG}/^l(u^{l,0}, u)$ *is defined as follows:*

(a) *If* $l = 0$*, then compute* $u^{l,2} \in V^l$ *so that*

$$u_1^{l,2} = u_1^{l,0} \quad \textit{and} \quad RQ(u + u^{l,2}) = \min_{v \in V_2^l} RQ(u + u_1^{l,0} + v).$$

(b) *If* $l > 0$*, then perform the following:*

1. **Coarse-grid correction step**. *Set* $u^{l,1} = u^{l,0} + \mathbf{RQMG}/^{l-1}(0, u + u^{l,0})$.
2. **Smoothing step**. *Let* $\lambda = RQ(u + u^{l,1})$ *and define* $w^l \in V^l$ *by*

$$q^l(w^l, v) = -c_\lambda(u + u^{l,1}, v), \quad \forall\, v \in V^l.$$

Then set $u^{l,2} = u^{l,1} + w^l$.

Remark 5.1. The RQMG coarsest grid correction step posed here is in the spirit of our analysis in the sense that the V_1^0 error component is guaranteed to be unchanged. This is in contrast to the significant changes that can happen in the more natural RQMG scheme of determining an optimal correction from the coarsest grid, which would be much more difficult to analyze. However, the coarsest grid solver used here is no less practical because it can be implemented using projections onto V_1^0 and V_2^0, which are obtained naturally in the full multigrid process. That is, the eigenvector approximation computed initially on the coarse grid can be used to implement the coarsest grid RQMG solver efficiently.

THEOREM 5.1 (RQMG convergence). *Suppose that the current eigenvector approximation* u_{old}^l *is in* $V^l \setminus \{0\}$*. Let* $u^{l,2} = \mathbf{RQMG}/^l(0, u_{\text{old}}^l)$*. (Here we assume that any nonzero initial correction has already been added into* $u = u_{\text{old}}^l$*.) Define the new eigenvector approximation by* $u_{\text{new}}^l = u_{\text{old}}^l + u^{l,2}$*. Define the errors* e_{old}^l *and* e_{new}^l *in the respective approximations* u_{old}^l *and* u_{new}^l *as in* (4.3)*. Then* $\mathbf{RQMG}/^l$ *converges according to the estimate*

$$\|e_{\text{new}}^l\|_c \le (Ch_0 + \beta)\|e_{\text{old}}^l\|_c. \tag{5.1}$$

Proof. The coarsest grid solver satisfies the assumptions of Lemma 3.11. Indeed, setting to zero the Fréchet derivative of $RQ(u_{\text{old}}^l + u_1^{0,0} + w)$ with respect to $v \in V_2^0$, we get

$$c_\lambda(w, v) = c_\lambda(u_{\text{old}}^0, v) \qquad \forall v \in V_2^0.$$

Hence

$$\begin{pmatrix} |||u_1^{0,2} - u_1^{0,0}||| \\ \|\mathcal{U}^0 - u_2^{0,2}\|_c \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \le \mathcal{E}^0 \begin{pmatrix} \|r_1^{0,0}\| \\ \|\mathcal{U}^0 - u_2^{0,0}\|_c \end{pmatrix},$$

where $\varepsilon_i^0 = 0$, $i = 1, 2, 3$, and $\beta_0 = \beta$. This proof is now completed by noting that the proof of Theorem 4.1 did not rely on fixing the shift λ, so it applies here without modification. ∎

This theorem confirms that RQMG converges linearly with a per cycle error reduction factor that is bounded below one uniformly in l, which is the usual sense of multigrid optimality. Of course, Rayleigh quotient iteration is cubically convergent when exact linear solvers are used. It should therefore come as no surprise that the overall efficiency of a Rayleigh quotient iterative method is dictated by the linear solver efficiency. For the elliptic case considered here, RQMG is optimal in this sense, as the numerical results in [5] demonstrate.

6. Full multigrid for eigenvalue problems. The theory developed in this paper applies to the eigenvalue problem (2.1) but only in a local sense: the multigrid convergence factors are guaranteed to be optimal only in a relatively small region about the eigenspace. More precisely, if $\lambda\,(\geq \lambda_1)$ is the current approximation to the eigenvalue on level l, then to obtain this optimality we need an estimate like (2.8).

Our purpose in this section is to show that this localness can be ensured by a full multigrid procedure, which is an outer loop iteration that uses the basic multigrid cycle first on coarser grids to obtain good initial guesses to the fine grid problem. We will use this scheme to guarantee that the error on each level satisfies

$$\|e^l\|_c \leq Ch_l, \tag{6.1}$$

which will suffice to prove (2.8) and, therefore, the optimality of the relevant multigrid solver.

The full multigrid procedure will be posed in terms of a generic multigrid method for solving the level l eigenvalue problem defined by (2.5). We write this generic scheme in the form

$$u^l_{new} = \mathbf{EMG}/^l(u^l_{\text{old}}),$$

which we assume converges in the sense of Theorems 4.1 and 5.1 with a factor bounded above by the constant $\gamma \equiv Ch_0 + \beta < 1$, provided (2.8) is satisfied and h_0 is sufficiently small. The following definition of the full multigrid process uses $\nu \geq 1$ cycles of EMG on progressively finer levels but makes no use of any initial approximation.

FMG/ALGORITHM. *The final approximation* $u^l = \mathbf{FMG}/^l$ *is defined as follows:*

(a) *If* $l = 0$*, then compute* $u^l \in V^l$ *so that the error satisfies* (2.8) *and* (6.1) *for* $l = 0$.

(b) *If* $l > 0$*, then perform the following:*

1. **Coarse grid solution.** *Set* $u^{l-1} = \mathbf{FMG}/^{l-1}$.
2. **Multigrid step.** *Perform* ν *steps of* **EMG**/ *applied to* (2.5) *using the initial approximation* $u^{l-1,\,0} = u^{l-1}$*:*

$$u^{l,\,\mu} = \mathbf{EMG}/^l(u^{l,\,\mu-1}), \quad \mu = 1,\, 2,\, \cdots,\, \nu.$$

Note that the coarsest grid ($l = 0$) step can be done by computing an accurate approximation to u^0_1. This follows from an estimate for the error measure $\|e^0\|_c$ that we obtain in the proof of our last theorem. (See (6.2) and (6.3) below.)

THEOREM 6.1 (global convergence). *Let* $\lambda = RQ(u^l)$*. Then, for sufficiently small* h_0 *and large* ν*, bounds* (2.8) *and* (6.1) *are satisfied for all* $l \geq 0$.

Proof. The case $l = 0$ is true by definition. Suppose for induction purposes that (2.8) and (6.1) are true for $l - 1$. Writing $u^l = u^l_1 + u^l_2$, $u^l_i \in V^l_i$, $i = 1,\, 2$, first note that

$$\begin{aligned} RQ(u^l) - \lambda_1 &= \frac{c(u^l_1,\, u^l_1) + c(u^l_2,\, u^l_2)}{b(u^l_1,\, u^l_1) + b(u^l_2,\, u^l_2)} \\ &\leq (\lambda^l_1 - \lambda_1) + \lambda^l_1 \frac{\|u^l_2\|^2_c}{|||u^l_1|||^2} \\ &\leq Ch^2_l + \lambda^l_1 \|e^l\|^2_c, \end{aligned} \tag{6.2}$$

where $e^l = \frac{u^l_2}{|||u^l_1|||}$ as in (4.3). With the analogous decomposition $u^{l-1} = u^{l-1}_1 + u^{l-1}_2$,

$u_i^{l-1} \in V_i^{l-1}$, $i = 1, 2$ and defining $u_1 = P_1^l u_1^{l-1}$, then by (3.8) and (3.5) it follows that

$$\begin{aligned}
\frac{\|P_2^l u^{l-1}\|_c}{|||P_1^l u^{l-1}|||} &\le \frac{\inf_{u \in V_1^l} |||u^{l-1} - u|||}{|||P_1^l u_1^{l-1}||| - |||P_1^l u_2^{l-1}|||} \\
&\le \frac{|||u_1^{l-1} - u_1 + u_2^{l-1}|||}{|||u_1||| - |||u_2^{l-1}|||} \\
&\le \frac{|||u_1^{l-1} - u_1||| + C\,\|u_2^{l-1}\|_c}{|||u_1^{l-1}||| - |||u_1^{l-1} - u_1||| - C\|u_2^{l-1}\|_c} \\
&= \frac{\frac{|||u_1^{l-1} - u_1|||}{|||u_1^{l-1}|||} + C\|e^{l-1}\|_c}{1 - \frac{|||u_1^{l-1} - u_1|||}{|||u_1^{l-1}|||} - C\|e^{l-1}\|_c} \\
&\le \frac{Ch_{l-1}}{1 - Ch_{l-1}}.
\end{aligned}$$

For sufficiently small h_0, we therefore have that

$$\frac{\|P_2^l u^{l-1}\|_c}{|||P_1^l u^{l-1}|||} \le Ch_l. \tag{6.3}$$

Bounds (2.8) and (6.1) for l now follow from (6.2), (6.3), and the definition of γ that implies

$$\|e^l\|_c \le \gamma^\nu \frac{\|P_2^l u^{l-1}\|_c}{|||P_1^l u^{l-1}|||}. \qquad \square$$

REFERENCES

[1] R. Bank, *Analysis of a multilevel inverse iteration procedure for eigenvalue problems*, SIAM J. Numer. Anal., 19 (1982), pp. 886–898.

[2] A. Brandt, S. McCormick, and J. Ruge, *Multigrid methods for differential eigenproblems*, SIAM J. Sci. Statist. Comput., 4 (1983), pp. 244-260.

[3] W. Hackbusch, *Multigrid solutions to linear and nonlinear eigenvalue problems for integral and differential equations*, Rostock. Math. Kolloq., 25 (1984), pp. 79–98.

[4] W. Hackbusch, *On the computation of approximate eigenvalues and eigenfunctions by means of a multigrid method*, SIAM J. Numer. Anal., 16 (1979), pp. 201–215.

[5] J. Mandel and S. McCormick, *A multilevel variational method for $Au = \lambda Bu$ on composite grids*, J. Comput. Phys., 80 (1989), pp. 442–450.

[6] J. Mandel, S. McCormick, and R. Bank, *Variational multigrid theory*, in Multigrid Methods, S. McCormick, ed., SIAM Frontiers in Applied Math., 3 (1987), pp. 131–178.

[7] S. McCormick, *A mesh refinement method for $Ax = \lambda Bx$*, Math. Comp., 36 (1981), pp. 485–498.

[8] S. McCormick, *Multilevel adaptive methods for elliptic eigenproblems: A two-level convergence theory*, SIAM J. Numer. Anal., 31 (1994), pp. 1731–1745.

[9] G. Strang and G. J. Fix, *An Analysis of the Finite Element Method*, Prentice–Hall, Englewood Cliffs, NJ, 1973.