An Algebraic Schwarz Theory

Michael Holst

CRPC-TR94460
June, 1994

Center for Research on Parallel Computation
Rice University
P.O. Box 1892
Houston, TX 77251-1892

This work was supported in part by the NSF.

An Algebraic Schwarz Theory*

Michael Holst, Caltech
Applied Math 217-50
Pasadena, CA 91125

June 13, 1994

Abstract

In this paper, we discuss the general theory of additive and multiplicative Schwarz methods for self-adjoint positive linear operator equations, representative methods of particular interest here being multigrid and domain decomposition. We examine closely one of the most useful and elegant modern convergence theories for these methods, following closely the recent work of Dryja and Widlund, Xu, and their colleagues. Our motivation is to fully understand this theory, and then to develop a variation of the theory in a slightly more general setting, which will be useful in the analysis of algebraic multigrid and domain decomposition methods, when little or no finite element structure is available. Using this approach we can show some convergence results for a very broad class of fully algebraic domain decomposition methods, without regularity assumptions about the continuous problem. Although we cannot at this time use the theory to provide a "good" convergence theory for algebraic multigrid methods, we believe that with additional analysis it may be possible to do so using this framework, as well as to use the framework to guide the design of the coarse problems. The language we employ throughout is algebraic, and can be interpreted abstractly in terms of operators on Hilbert spaces, or in terms of matrix operators.

*This work was supported in part by the NSF under Cooperative Agreement No. CCR-9120008.
The government has certain rights in this material.

Contents

1. Introduction **1**

2. Linear operator equations **3**
2.1 Linear operators and spectral theory 3
2.2 The basic linear method 4
2.3 Properties of the error propagation operator 5
2.4 Conjugate gradient acceleration of linear methods 8

3. A theory for products and sums of operators **14**
3.1 Basic product and sum operator theory 14
3.2 The interaction hypothesis 18
3.3 Allowing for a global operator 22
3.4 Main results of the theory 24

4. Abstract Schwarz theory **26**
4.1 The Schwarz methods 26
4.2 Subspace splitting theory 28
4.3 Product and sum splitting theory for non-nested Schwarz methods 33
4.4 Product and sum splitting theory for nested Schwarz methods 35

5. Applications to domain decomposition **37**
5.1 Variational formulation and subdomain-based subspaces 37
5.2 The multiplicative and additive Schwarz methods 37
5.3 Algebraic domain decomposition methods 37
5.4 Convergence theory for the algebraic case 38
5.5 Improved results through finite element theory 39

6. Applications to multigrid **40**
6.1 Recursive multigrid and nested subspaces 40
6.2 Variational multigrid as a multiplicative Schwarz method 40
6.3 Algebraic multigrid methods 41
6.4 Convergence theory for the algebraic case 41
6.5 Improved results through finite element theory 42

Acknowledgments **43**

Bibliography **44**

1. Introduction

In this paper, we discuss the general theory of additive and multiplicative Schwarz methods for self-adjoint positive linear operator equations, representative methods of particular interest here being multigrid and domain decomposition. We examine closely one of the most useful and elegant modern convergence theories for these methods, following closely the recent work of Dryja and Widlund, Xu, and their colleagues. Our motivation is to fully understand this theory, and then to develop a variation of the theory in a slightly more general setting, which will be useful in the analysis of algebraic multigrid and domain decomposition methods, when little or no finite element structure is available. Using this approach we can show some convergence results for a very broad class of fully algebraic domain decomposition methods, without regularity assumptions about the continuous problem. Although we cannot at this time use the theory to provide a "good" convergence theory for algebraic multigrid methods, we believe that with additional analysis it may be possible to do so using this framework, as well as to use the framework to guide the design of the coarse problems. The language we employ throughout is algebraic, and can be interpreted abstractly in terms of operators on Hilbert spaces, or in terms of matrix operators.

Our approach in this paper is quite similar (and owes much) to [39], with the following exceptions. We first develop a separate and complete theory for products and sums of operators, without reference to subspaces, and then use this theory to formulate a Schwarz theory based on subspaces. In addition, we develop the Schwarz theory allowing for completely general prolongation and restriction operators, so that the theory is not restricted to the use of inclusion and projection as the transfer operators (a similar Schwarz framework with general transfer operators was constructed recently by Hackbusch [17]). The resulting theoretical framework is useful for analyzing specific algebraic methods, such as algebraic multigrid and algebraic domain decomposition, without requiring the use of finite element spaces (and their associated transfer operators of inclusions and projection). The framework may also be useful for analyzing methods based on transforms to other spaces not naturally thought of as subspaces, such as methods based on successive wavelet or other transforms. Finally, we show quite clearly how the basic product/sum and Schwarz theories must be modified and refined to analyze the effects of using a global operator, or of using additional nested spaces as in the case of multigrid-type methods. We also present (adding somewhat to the length of an already lengthy paper) a number of (albeit simple but useful) results in the product/sum and Schwarz theory frameworks which are commonly used in the literature, the proofs of which are often difficult to locate (for example, the relationship between the usual condition number of an operator and its generalized or A-condition number). The result is a consistent and self-contained theoretical framework for analyzing abstract linear methods for self-adjoint positive linear operator equations, based on subspace-decomposition ideas.

Outline

As a brief outline, we begin in §2 with a review of the basic theory of self-adjoint operators (or symmetric matrices), the idea of a linear iterative method, and some key ideas about conjugate gradient acceleration of linear methods. While most of this material is well-known, it seems to be scattered around the literature, and many of the simple proofs seem unavailable or difficult to locate. Therefore, we have chosen to present this background material in an organized way at the beginning of the paper.

In §3, we present a unified approach for bounding the norms and condition numbers of products and sums of self-adjoint operators on a Hilbert space, derived from work due to Dryja and Widlund [14], Bramble et al. [8], and Xu [39]. Our particular approach is quite general in that we establish the main norm and condition number bounds without reference to subspaces; each of the three required assumptions for the theory involve only the operators on the original Hilbert space. Therefore, this product/sum operator theory may find use in other applications without natural subspace decompositions. Later, we will apply the product and sum operator theory to the case when the operators correspond to corrections in subspaces of the original space, as in multigrid and domain decomposition methods.

In §4, we consider abstract Schwarz methods based on subspaces, and apply the general product and sum operator theory to these methods. The resulting theory, which is a variation of that presented in [39] and [14], rests on the notion of a stable subspace splitting of the original Hilbert space (cf. [31, 32]). Although our derivation here is presented in a somewhat different, algebraic language, many of the intermediate results we use have appeared previously in the literature in other forms (we provide references at the appropriate points). In contrast to earlier approaches, we develop the entire theory employing general prolongation and restriction operators; the use of inclusion and projection as prolongation and restriction are represented in our approach as a special case.

In §5 and §6, we apply the theory derived earlier to domain decomposition methods and to multigrid methods, and in particular to their algebraic forms. Since our theoretical framework allows for general prolongation and restriction operators, the theory is applicable to methods for general algebraic equations (coming from finite difference or finite volume discretization of elliptic equations) for which strong theories are currently lacking. For algebraic domain decomposition, we are able to derive useful (although not optimal) convergence estimates. Although the algebraic multigrid results are not as interesting, the theory does provide yet another proof of the robustness of the algebraic multigrid approach. We also indicate how the convergence estimates for multigrid and domain decomposition methods may be improved (giving optimal estimates), following the recent work of Dryja and Widlund, Bramble et al., and Xu, which requires some of the additional structure provided in the finite element setting.

In addition to the references cited directly in the text below, the material here owes much to the following sources: [5, 6, 7, 12, 14, 17, 26, 27, 28, 38, 39].

2. Linear operator equations

In this section, we first review the theory of self-adjoint linear operators on a Hilbert space. The results required for the analysis of linear methods, as well as conjugate gradient methods, are summarized. We then develop carefully the theory of classical linear methods for operators equations. The conjugate gradient method is then considered, and the relationship between the convergence rate of linear methods as preconditioners and the convergence rate of the resulting preconditioned conjugate gradient method is explored in some detail.

As a motivation, consider that if either the box-method or the finite element method is used to discretize the second order linear elliptic partial differential equation $\mathcal{L}u = f$, a set of linear algebraic equations results, which we denote as:

$$A_k u_k = f_k. \tag{1}$$

The subscript k denotes the discretization level, with larger k corresponding to a more refined mesh, and with an associated mesh parameter h_k representing the diameter of the largest element or volume in the mesh Ω_k. For a self-adjoint strongly elliptic partial differential operator, the matrix A_k produced by the box or finite element method is SPD. In this work, we are interested in linear iterations for solving the matrix equation (1) which have the general form:

$$u_k^{n+1} = (I - B_k A_k) u_k^n + B_k f_k, \tag{2}$$

where B_k is an SPD matrix approximating A_k^{-1} in some sense. The classical stationary linear methods fit into this framework, as well as domain decomposition methods and multigrid methods.

2.1 Linear operators and spectral theory

In this section we compile some material on self-adjoint linear operators in finite-dimensional spaces which will be used throughout the work.

Let $\mathcal{H}$ be a real finite-dimensional Hilbert space equipped with the inner-product $(\cdot,\cdot)$ inducing the norm $\|\cdot\| = (\cdot,\cdot)^{1/2}$. Since we are concerned only with finite-dimensional spaces, $\mathcal{H}$ can be thought of as the Euclidean space $\mathbb{R}^n$; however, the preliminary material below and the algorithms we develop are phrased in terms of the unspecified space $\mathcal{H}$, so that the algorithms may be interpreted directly in terms of finite element spaces as well. This is necessary to set the stage for our discussion of multigrid and domain decomposition theory later in the work.

If the operator $A : \mathcal{H} \mapsto \mathcal{H}$ is linear, we denote this as $A \in \mathbf{L}(\mathcal{H},\mathcal{H})$. If $A \in \mathbf{L}(\mathcal{H},\mathcal{H})$ is self-adjoint positive definite (SPD) with respect to $(\cdot,\cdot)$, then the bilinear form $A(u,v) = (Au,v)$ defines another inner-product on $\mathcal{H}$, which we denote as $(\cdot,\cdot)_A = A(\cdot,\cdot)$ to emphasize the fact that it is an inner-product. This second inner-product also induces a norm $\|\cdot\|_A = (\cdot,\cdot)_A^{1/2}$. For each inner-product the Cauchy-Schwarz inequality holds:

$$|(u,v)| \le (u,u)^{1/2}(v,v)^{1/2}, \qquad |(u,v)_A| \le (u,u)_A^{1/2}(v,v)_A^{1/2}, \qquad \forall u,v \in \mathcal{H}.$$

The adjoint of the operator $M \in \mathbf{L}(\mathcal{H},\mathcal{H})$ with respect to $(\cdot,\cdot)$ is denoted M^T where $(Mu,v) = (u,M^Tv)\ \forall u,v \in \mathcal{H}$, and the adjoint with respect to $(\cdot,\cdot)_A$ is called the A-adjoint and denoted M^*, where $(Mu,v)_A = (u,M^*v)_A\ \forall u,v \in \mathcal{H}$.

For the operator M we denote the eigenvalues satisfying $Mu_i = \lambda_i u_i$ for eigenfunctions $u_i \ne 0$ as $\lambda_i(M)$. The spectral theory for self-adjoint linear operators states that the eigenvalues of the self-adjoint operator M are real and lie in the closed interval $[\lambda_{\min}(M), \lambda_{\max}(M)]$ defined by the Raleigh quotients:

$$\lambda_{\min}(M) = \min_{u\ne 0} \frac{(Mu,u)}{(u,u)}, \qquad \lambda_{\max}(M) = \max_{u\ne 0} \frac{(Mu,u)}{(u,u)}.$$

Similarly, if an operator M is A-self-adjoint, then the eigenvalues are real and lie in the interval defined by the Raleigh quotients generated by the A-inner-product:

$$\lambda_{\min}(M) = \min_{u\ne 0} \frac{(Mu,u)_A}{(u,u)_A}, \qquad \lambda_{\max}(M) = \max_{u\ne 0} \frac{(Mu,u)_A}{(u,u)_A}.$$

We denote the set of eigenvalues as the spectrum $\sigma(M)$ and the largest of these in absolute value as the spectral radius as $\rho(M) = \max(|\lambda_{\min}(M)|, |\lambda_{\max}(M)|)$. For SPD (or A-SPD) operators M, the eigenvalues of M are real and positive, and the powers M^s for real s are well-defined through the spectral decomposition; see for example §79 and §82 in [18]. Finally, recall that a matrix representing the operator M with respect to any basis for $\mathcal{H}$ has the same eigenvalues as the operator M.

Linear operators on finite-dimensional spaces are always bounded, and these bounds define the operator norms induced by the norms $\|\cdot\|$ and $\|\cdot\|_A$:

$$\|M\| = \max_{u\neq 0} \frac{\|Mu\|}{\|u\|}, \qquad \|M\|_A = \max_{u\neq 0} \frac{\|Mu\|_A}{\|u\|_A}.$$

A well-known property is that if M is self-adjoint, then $\rho(M) = \|M\|$. This property can also be shown to hold for A-self-adjoint operators. The following lemma can be found in [2] (as Lemma 4.1), although the proof there is for A-normal matrices rather than A-self-adjoint operators.

Lemma 2.1 *If A is SPD and M is A-self-adjoint, then $\|M\|_A = \rho(M)$.*

Proof. We simply note that

$$\|M\|_A = \max_{u\neq 0} \frac{\|Mu\|_A}{\|u\|_A} = \max_{u\neq 0} \frac{(AMu, Mu)^{1/2}}{(Au, u)^{1/2}} = \max_{u\neq 0} \frac{(AM^*Mu, u)^{1/2}}{(Au, u)^{1/2}} = \lambda_{\max}^{1/2}(M^*M),$$

since M^*M is always A-self-adjoint. Since by assumption M itself is A-self-adjoint, we have that $M^* = M$, which yields: $\|M\|_A = \lambda_{\max}^{1/2}(M^*M) = \lambda_{\max}^{1/2}(M^2) = (\max_i[\lambda_i^2(M)])^{1/2} = \max[|\lambda_{\min}(M)|, |\lambda_{\max}(M)|] = \rho(M)$. $\square$

2.2 The basic linear method

In this section, we introduce the basic linear method which we study and use in the remainder of the work.

Assume we are faced with the operator equation $Au = f$, where $A \in \mathbf{L}(\mathcal{H}, \mathcal{H})$ is SPD, and we desire the unique solution u. Given a *preconditioner* (approximate inverse) $B \approx A^{-1}$, consider the equivalent *preconditioned system* $BAu = Bf$. The operator B is chosen so that the simple linear iteration:

$$u^1 = u^0 - BAu^0 + Bf = (I - BA)u^0 + Bf,$$

which produces an improved approximation u^1 to the true solution u given an initial approximation u^0, has some desired convergence properties. This yields the following basic linear iterative method which we study in the remainder of this work:

Algorithm 2.1 *(Basic Linear Method for solving $Au = f$)*

$$u^{n+1} = u^n + B(f - Au^n) = (I - BA)u^n + Bf.$$

Subtracting the iteration equation from the identity $u = u - BAu + Bf$ yields the equation for the error $e^n = u - u^n$ at each iteration:

$$e^{n+1} = (I - BA)e^n = (I - BA)^2 e^{n-1} = \cdots = (I - BA)^{n+1} e^0. \tag{3}$$

The convergence of Algorithm 2.1 is determined completely by the spectral radius of the error propagation operator $E = I - BA$.

Theorem 2.2 *The condition $\rho(I - BA) < 1$ is necessary and sufficient for convergence of Algorithm 2.1.*

Proof. See for example Theorem 10.11 in [23] or Theorem 7.1.1 in [30]. $\square$

Since $|\lambda|\|u\| = \|\lambda u\| = \|Mu\| \leq \|M\|\,\|u\|$ for any norm $\|\cdot\|$, it follows that $\rho(M) \leq \|M\|$ for all norms $\|\cdot\|$. Therefore, $\|I - BA\| < 1$ and $\|I - BA\|_A < 1$ are both sufficient conditions for convergence of Algorithm 2.1. In fact, it is the norm of the error propagation operator which will bound the reduction of the error at each iteration, which follows from (3):

$$\|e^{n+1}\|_A \leq \|I - BA\|_A \|e^n\|_A \leq \|I - BA\|_A^{n+1} \|e^0\|_A. \tag{4}$$

The spectral radius $\rho(E)$ of the error propagator E is called the *convergence factor* for Algorithm 2.1, whereas the norm of the error propagator $\|E\|$ is referred to as the *contraction number* (with respect to the particular choice of norm $\|\cdot\|$).

2.3 Properties of the error propagation operator

In this section, we establish some simple properties of the error propagation operator of an abstract linear method. We note that several of these properties are commonly used, especially in the multigrid literature, although the short proofs of the results seem difficult to locate. The particular framework we construct here for analyzing linear methods is based on the recent work of Xu [39], on the recent papers on multigrid and domain decomposition methods referenced therein, and on the text by Varga [34].

An alternate sufficient condition for convergence of the basic linear method is given in the following lemma, which is similar to *Stein's Theorem* (Theorem 7.1.8 in [30], or Theorem 6.1, page 80 in [40]).

Lemma 2.3 *If E^* is the A-adjoint of E, and $I - E^*E$ is A-positive, then it holds that $\rho(E) \leq \|E\|_A < 1$.*

Proof. By hypothesis, $(A(I - E^*E)u, u) > 0 \ \forall u \in \mathcal{H}$. This implies that $(AE^*Eu, u) < (Au, u) \ \forall u \in \mathcal{H}$, or $(AEu, Eu) < (Au, u) \ \forall u \in \mathcal{H}$. But this last inequality implies that

$$\rho(E) \leq \|E\|_A = \max_{u \neq 0} \frac{(AEu, Eu)}{(Au, u)} < 1.$$

□

We now state three very simple lemmas that we use repeatedly in the following sections.

Lemma 2.4 *If A is SPD, then BA is A-self-adjoint if and only if B is self-adjoint.*

Proof. Simply note that: $(ABAx, y) = (BAx, Ay) = (Ax, B^T Ay) \ \forall x, y \in \mathcal{H}$. The lemma follows since $BA = B^T A$ if and only if $B = B^T$. □

Lemma 2.5 *If A is SPD, then $I - BA$ is A-self-adjoint if and only if B is self-adjoint.*

Proof. Begin by noting that: $(A(I - BA)x, y) = (Ax, y) - (ABAx, y) = (Ax, y) - (Ax, (BA)^* y) = (Ax, (I - (BA)^*)y), \ \ \forall x, y \in \mathcal{H}$. Therefore, $E^* = I - (BA)^* = I - BA = E$ if and only if $BA = (BA)^*$. But by Lemma 2.4, this holds if and only if B is self-adjoint, so the result follows. □

Lemma 2.6 *If A and B are SPD, then BA is A-SPD.*

Proof. By Lemma 2.4, BA is A-self-adjoint. Also, we have that:

$$(ABAu, u) = (BAu, Au) = (B^{1/2}Au, B^{1/2}Au) > 0 \quad \forall u \neq 0, \ u \in \mathcal{H}.$$

Therefore, BA is also A-positive, and the result follows. □

We noted above that the property $\rho(M) = \|M\|$ holds in the case that M is self-adjoint with respect to the inner-product inducing the norm $\|\cdot\|$. If B is self-adjoint, the following theorem states that the resulting error propagator $E = I - BA$ has this property with respect to the A-norm.

Theorem 2.7 *If A is SPD and B is self-adjoint, then $\|I - BA\|_A = \rho(I - BA)$.*

Proof. By Lemma 2.5, $I - BA$ is A-self-adjoint, and by Lemma 2.1 the result follows. □

The following simple lemma, similar to Lemma 2.3, will be very useful later in the work.

Lemma 2.8 *If A and B are SPD, and $E = I - BA$ is A-non-negative, then it holds that $\rho(E) = \|E\|_A < 1$.*

Proof. By Lemma 2.5, E is A-self-adjoint, and by assumption E is A-non-negative, and so from §2.1 we see that E must have real non-negative eigenvalues. By hypothesis, $(A(I - BA)u, v) \geq 0 \ \forall u, v \in \mathcal{H}$, which implies that $(ABAu, u) \leq (Au, u) \ \forall u \in \mathcal{H}$. By Lemma 2.6, BA is A-SPD, and we have that

$$0 < (ABAu, u) \leq (Au, u) \quad \forall u \in \mathcal{H}, \quad u \neq 0,$$

which implies that $0 < \lambda_i(BA) \leq 1 \ \forall \lambda_i \in \sigma(BA)$. Thus, since $\lambda_i(E) = \lambda_i(I - BA) = 1 - \lambda_i(BA) \ \forall i$, we have that

$$\rho(E) = \max_i \lambda_i(E) = 1 - \min_i \lambda_i(BA) < 1.$$

Finally, by Theorem 2.7, we have $\|E\|_A = \rho(E) < 1$. □

The following simple lemma relates the contraction number bound to two simple inequalities; it is a standard result which follows directly from the spectral theory of self-adjoint linear operators.

Lemma 2.9 *If A is SPD and B is self-adjoint, and $E = I - BA$ is such that:*

$$-C_1(Au, u) \leq (AEu, u) \leq C_2(Au, u), \qquad \forall u \in \mathcal{H},$$

for $C_1 \geq 0$ and $C_2 \geq 0$, then $\rho(E) = \|E\|_A \leq \max\{C_1, C_2\}$.

Proof. By Lemma 2.5, $E = I - BA$ is A-self-adjoint, and by the spectral theory outlined in §2.1, the inequality above simply bounds the most negative and most positive eigenvalues of E with $-C_1$ and C_2, respectively. The result then follows by Theorem 2.7. □

Corollary 2.10 *If A and B are SPD, then Lemma 2.9 holds for some $C_2 < 1$.*

Proof. By Lemma 2.6, BA is A-SPD, which implies that the eigenvalues of BA are real and positive by the discussion in §2.1. By Lemma 2.5, $E = I - BA$ is A-self-adjoint, and therefore has real eigenvalues. The eigenvalues of E and BA are related by $\lambda_i(E) = \lambda_i(I - BA) = 1 - \lambda_i(BA)\ \forall i$, and since $\lambda_i(BA) > 0\ \forall i$, we must have that $\lambda_i(E) < 1\ \forall i$. Since C_2 in Lemma 2.9 bounds the largest positive eigenvalue of E, we have that $C_2 < 1$. □

We now define the *A-condition number* of an invertible operator M by extending the standard notion to the A-inner-product:

$$\kappa_A(M) = \|M\|_A \|M^{-1}\|_A.$$

In the next section we show (Lemma 2.12) that if M is an A-self-adjoint operator, then in fact the following simpler expression holds:

$$\kappa_A(M) = \frac{\lambda_{\max}(M)}{\lambda_{\min}(M)}.$$

The generalized condition number κ_A is employed in the following lemma, which states that there is an optimal relaxation parameter for a basic linear method, and gives the best possible convergence estimate for the method employing the optimal parameter. This lemma has appeared many times in the literature in one form or another; cf. [31].

Lemma 2.11 *If A and B are SPD, then*

$$\rho(I - \alpha BA) = \|I - \alpha BA\|_A < 1.$$

if and only if $\alpha \in (0, 2/\rho(BA))$. Convergence is optimal when $\alpha = 2/[\lambda_{\min}(BA) + \lambda_{\max}(BA)]$, giving

$$\rho(I - \alpha BA) = \|I - \alpha BA\|_A = 1 - \frac{2}{1 + \kappa_A(BA)} < 1.$$

Proof. Note that $\rho(I - \alpha BA) = \max_\lambda |1 - \alpha\lambda(BA)|$, so that $\rho(I - \alpha BA) < 1$ if and only if $\alpha \in (0, 2/\rho(BA))$, proving the first part. Taking $\alpha = 2/[\lambda_{\min}(BA) + \lambda_{\max}(BA)]$, we have

$$\rho(I - \alpha BA) = \max_\lambda |1 - \alpha\lambda(BA)| = \max_\lambda (1 - \alpha\lambda(BA))$$

$$= \max_\lambda \left(1 - \frac{2\lambda(BA)}{\lambda_{\min}(BA) + \lambda_{\max}(BA)}\right) = 1 - \frac{2\lambda_{\min}(BA)}{\lambda_{\min}(BA) + \lambda_{\max}(BA)} = 1 - \frac{2}{1 + \frac{\lambda_{\max}(BA)}{\lambda_{\min}(BA)}}.$$

Since BA is A-self-adjoint, by Lemma 2.12 we have that $\kappa_A(BA) = \lambda_{\max}(BA)/\lambda_{\min}(BA)$, so that if $\alpha = 2/[\lambda_{\min}(BA) + \lambda_{\max}(BA)]$, then

$$\rho(I - \alpha BA) = \|I - \alpha BA\|_A = 1 - \frac{2}{1 + \kappa_A(BA)}.$$

To show this is optimal, we must solve $\min_{\alpha}[\max_{\lambda} |1-\alpha\lambda|]$, where $\alpha \in (0, 2/\lambda_{\max})$. Note that each α defines a polynomial of degree zero in λ, namely $P_o(\lambda) = \alpha$. Therefore, we can rephrase the problem as

$$P_1^{\text{opt}}(\lambda) = \min_{P_o}\left[\max_{\lambda}|1-\lambda P_o(\lambda)|\right].$$

It is well-known that the scaled and shifted Chebyshev polynomials give the solution to this "mini-max" problem:

$$P_1^{\text{opt}}(\lambda) = 1 - \lambda P_o^{\text{opt}} = \frac{T_1\left(\frac{\lambda_{\max}+\lambda_{\min}-2\lambda}{\lambda_{\max}-\lambda_{\min}}\right)}{T_1\left(\frac{\lambda_{\max}+\lambda_{\min}}{\lambda_{\max}-\lambda_{\min}}\right)}.$$

Since $T_1(x) = x$, we have simply that

$$P_1^{\text{opt}}(\lambda) = \frac{\frac{\lambda_{\max}+\lambda_{\min}-2\lambda}{\lambda_{\max}-\lambda_{\min}}}{\frac{\lambda_{\max}+\lambda_{\min}}{\lambda_{\max}-\lambda_{\min}}} = 1 - \lambda\left(\frac{2}{\lambda_{\min}+\lambda_{\max}}\right),$$

showing that in fact $\alpha_{\text{opt}} = 2/[\lambda_{\min}+\lambda_{\max}]$. $\Box$

Remark 2.1. Theorem 2.7 will be exploited later since $\rho(E)$ is usually much easier to compute numerically than $\|E\|_A$, and since it is the energy norm $\|E\|_A$ of the error propagator E which is typically bounded in various convergence theories for iterative processes.

Note that if we wish to reduce the initial error $\|e^0\|_A$ by the factor ϵ, then equation (4) implies that this will be guaranteed if

$$\|E\|_A^{n+1} \le \epsilon.$$

Taking natural logarithms of both sides and solving for n, we see that the maximum number of iterations required to reach the desired tolerance, as a function of the contraction number, is given by:

$$n \le \frac{|\ln\epsilon|}{|\ln\|E\|_A|}. \tag{5}$$

If the bound on the norm is of the form in Lemma 2.11, then to achieve a tolerance of ϵ after n iterations will require:

$$n \le \frac{|\ln\epsilon|}{\left|\ln\left(1-\frac{2}{1+\kappa_A(BA)}\right)\right|} = \frac{|\ln\epsilon|}{\left|\ln\left(\frac{\kappa_A(BA)-1}{\kappa_A(BA)+1}\right)\right|}. \tag{6}$$

Using the approximation:

$$\ln\left(\frac{a-1}{a+1}\right) = \ln\left(\frac{1+(-1/a)}{1-(-1/a)}\right) = 2\left[\left(\frac{-1}{a}\right) + \frac{1}{3}\left(\frac{-1}{a}\right)^3 + \frac{1}{5}\left(\frac{-1}{a}\right)^5 + \cdots\right] < \frac{-2}{a},$$

we have that $|\ln[(\kappa_A(BA)-1)/(\kappa_A(BA)+1)]| > 2/\kappa_A(BA)$, so that:

$$n \le \frac{1}{2}\kappa_A(BA)|\ln\epsilon| + 1.$$

We then have that the maximum number of iterations required to reach an error on the order of the tolerance ϵ is:

$$n = O(\kappa_A(BA)|\ln\epsilon|).$$

If a single iteration of the method costs $O(N)$ arithmetic operators, then the overall complexity to solve the problem is $O(|\ln\|E\|_A|^{-1}N|\ln\epsilon|)$, or $O(\kappa_A(BA)N|\ln\epsilon|)$. If the quantity $\|E\|_A$ can be bounded less than one independent of N, or if $\kappa_A(BA)$ can be bounded independent of N, then the complexity is near optimal $O(N|\ln\epsilon|)$.

Note that if E is A-self-adjoint, then we can replace $\|E\|_A$ by $\rho(E)$ in the above discussion. Even when this is not the case, $\rho(E)$ is often used above in place of $\|E\|_A$ to obtain an estimate, and the quantity $R_\infty(E) = -\ln\rho(E)$ is referred to as the *asymptotic convergence rate* (see page 67 of [34], or page 88 of [40]).

In [34], the *average rate of convergence of m iterations* is defined as the quantity $R(E^m) = -\ln(\|E^m\|^{1/m})$, the meaning of which is intuitively clear from equation (4). As noted on page 95 in [34], since $\rho(E) = \lim_{m\to\infty} \|E^m\|^{1/m}$ for all bounded linear operators E and norms $\|\cdot\|$ (Theorem 7.5-5 in [24]), it follows that $\lim_{m\to\infty} R(E^m) = R_\infty(E)$.

While $R_\infty(E)$ is considered the standard measure of convergence of linear iterations (it is called the "convergence rate" in [40], page 88), this is really an asymptotic measure, and the convergence behavior for the early iterations may be better monitored by using the norm of the propagator E directly in (5); an example is given on page 67 of [34] for which $R_\infty(E)$ gives a poor estimate of the number of iterations required.

2.4 Conjugate gradient acceleration of linear methods

Consider now the linear equation $Au = f$ in the space $\mathcal{H}$. The conjugate gradient method was developed by Hestenes and Stiefel [19] for linear systems with symmetric positive definite operators A. It is common to *precondition* the linear system by the SPD *preconditioning operator* $B \approx A^{-1}$, in which case the generalized or preconditioned conjugate gradient method [10] results. Our purpose in this section is to briefly examine the algorithm, its contraction properties, and establish some simple relationships between the contraction number of a basic linear preconditioner and that of the resulting preconditioned conjugate gradient algorithm. These relationships are commonly used, but some of the short proofs seem unavailable.

In [3], a general class of conjugate gradient methods obeying three-term recursions is studied, and it is shown that each instance of the class can be characterized by three operators: an inner product operator X, a preconditioning operator Y, and the system operator Z. As such, these methods are denoted as CG(X,Y,Z). We are interested in the special case that $X = A$, $Y = B$, and $Z = A$, when both B and A are SPD. Choosing the *Omin* [3] algorithm to implement the method CG(A,B,A), the *preconditioned conjugate gradient method* results:

Algorithm 2.2 *(Preconditioned Conjugate Gradient Algorithm)*

$$
\begin{array}{l}
\textit{Let } u^0 \in \mathcal{H} \textit{ be given.} \\
r^0 = f - Au^0, \quad s^0 = Br^0, \quad p^0 = s^0. \\
\textit{Do } i = 0, 1, \ldots \textit{ until convergence:} \\
\quad \alpha_i = (r^i, s^i)/(Ap^i, p^i) \\
\quad u^{i+1} = u^i + \alpha_i p^i \\
\quad r^{i+1} = r^i - \alpha_i A p^i \\
\quad s^{i+1} = Br^{i+1} \\
\quad \beta_{i+1} = (r^{i+1}, s^{i+1})/(r^i, s^i) \\
\quad p^{i+1} = s^{i+1} + \beta_{i+1} p^i \\
\textit{End do.}
\end{array}
$$

If the dimension of $\mathcal{H}$ is n, then the algorithm can be shown to converge in n steps since the preconditioned operator BA is A-SPD [3]. Note that if $B = I$, then this algorithm is exactly the Hestenes and Stiefel algorithm.

Since we wish to understand a little about the convergence properties of the conjugate gradient method, and how these will be effected by a linear method representing the preconditioner B, we will briefly review a well-known conjugate gradient contraction bound. To begin, it is not difficult to see that the error at each iteration of Algorithm 2.2 can be written as a polynomial in BA times the initial error:

$$e^{i+1} = [I - BAp_i(BA)]e^0,$$

where $p_i \in \mathcal{P}_i$, the space of polynomials of degree i. At each step the energy norm of the error $\|e^{i+1}\|_A = \|u - u^{i+1}\|_A$ is minimized over the Krylov subspace:

$$V_{i+1}(BA, Br^0) = \text{span } \{Br^0, (BA)Br^0, (BA)^2Br^0, \ldots, (BA)^iBr^0\}.$$

Therefore, it must hold that:

$$\|e^{i+1}\|_A = \min_{p_i \in \mathcal{P}_i} \|[I - BAp_i(BA)]e^0\|_A.$$

Since BA is A-SPD, the eigenvalues $\lambda_j \in \sigma(BA)$ of BA are real and positive, and the eigenvectors v_j of BA are A-orthonormal. By expanding $e^0 = \sum_{j=1}^n \alpha_j v_j$, we have:

$$\|[I - BAp_i(BA)]e^0\|_A^2 = (A[I - BAp_i(BA)]e^0, [I - BAp_i(BA)]e^0)$$

$$= (A[I - BAp_i(BA)](\sum_{j=1}^n \alpha_j v_j), [I - BAp_i(BA)](\sum_{j=1}^n \alpha_j v_j))$$

$$= (\sum_{j=1}^n [1 - \lambda_j p_i(\lambda_j)]\alpha_j \lambda_j v_j, \sum_{j=1}^n [1 - \lambda_j p_i(\lambda_j)]\alpha_j v_j) = \sum_{j=1}^n [1 - \lambda_j p_i(\lambda_j)]^2 \alpha_j^2 \lambda_j$$

$$\leq \max_{\lambda_j \in \sigma(BA)} [1 - \lambda_j p_i(\lambda_j)]^2 \sum_{j=1}^n \alpha_j^2 \lambda_j = \max_{\lambda_j \in \sigma(BA)} [1 - \lambda_j p_i(\lambda_j)]^2 \sum_{j=1}^n (A\alpha_j v_j, \alpha_j v_j)$$

$$= \max_{\lambda_j \in \sigma(BA)} [1 - \lambda_j p_i(\lambda_j)]^2 (A \sum_{j=1}^n \alpha_j v_j, \sum_{j=1}^n \alpha_j v_j) = \max_{\lambda_j \in \sigma(BA)} [1 - \lambda_j p_i(\lambda_j)]^2 \|e^0\|_A^2.$$

Thus, we have that

$$\|e^{i+1}\|_A \leq \left(\min_{p_i \in \mathcal{P}_i} \left[\max_{\lambda_j \in \sigma(BA)} |1 - \lambda_j p_i(\lambda_j)| \right] \right) \|e^0\|_A.$$

The scaled and shifted Chebyshev polynomials $T_{i+1}(\lambda)$, extended outside the interval $[-1, 1]$ as in the Appendix A of [4], yield a solution to this *mini-max* problem. Using some simple well-known relationships valid for $T_{i+1}(\cdot)$, the following contraction bound is easily derived:

$$\|e^{i+1}\|_A \leq 2 \left(\frac{\sqrt{\frac{\lambda_{\max}(BA)}{\lambda_{\min}(BA)}} - 1}{\sqrt{\frac{\lambda_{\max}(BA)}{\lambda_{\min}(BA)}} + 1} \right)^{i+1} \|e^0\|_A = 2\ \delta_{\text{cg}}^{i+1}\ \|e^0\|_A. \tag{7}$$

The ratio of the extreme eigenvalues of BA appearing in the bound is often mistakenly called the (spectral) condition number $\kappa(BA)$; in fact, since BA is not self-adjoint (it is A-self-adjoint), this ratio is not in general equal to the condition number (this point is discussed in great detail in [2]). However, the ratio does yield a condition number in a different norm. The following lemma is a special case of Corollary 4.2 in [2].

Lemma 2.12 *If A and B are SPD, then*

$$\kappa_A(BA) = \|BA\|_A \|(BA)^{-1}\|_A = \frac{\lambda_{\max}(BA)}{\lambda_{\min}(BA)}. \tag{8}$$

Proof. For any A-SPD M, it is easy to show that M^{-1} is also A-SPD, so that from §2.1 both M and M^{-1} have real, positive eigenvalues. From Lemma 2.1 it then holds that:

$$\|M^{-1}\|_A = \rho(M^{-1}) = \max_{u \neq 0} \frac{(AM^{-1}u, u)}{(Au, u)} = \max_{u \neq 0} \frac{(AM^{-1/2}u, M^{-1/2}u)}{(AMM^{-1/2}u, M^{-1/2}u)}$$

$$= \max_{v \neq 0} \frac{(Av, v)}{(AMv, v)} = \left[\min_{v \neq 0} \frac{(AMv, v)}{(Av, v)} \right]^{-1} = \lambda_{\min}(M)^{-1}.$$

By Lemma 2.6, BA is A-SPD, which together with Lemma 2.1 implies that $\|BA\|_A = \rho(BA) = \lambda_{\max}(BA)$. From above we have that $\|(BA)^{-1}\|_A = \lambda_{\min}(BA)^{-1}$, implying that the A-condition number is given as the ratio of the extreme eigenvalues of BA as in equation (8). □

More generally, it can be shown that if the operator D is C-normal for some SPD inner-product operator C, then the generalized condition number given by $\kappa_C(D) = \|D\|_C \|D^{-1}\|_C$ is equal to the ratio of the extreme eigenvalues of the operator D. A proof of this fact is given in Corollary 4.2 of [2], along with a detailed discussion of this and other relationships for more general conjugate gradient methods. The conjugate gradient contraction number δ_{cg} can now be written as:

$$\delta_{cg} = \frac{\sqrt{\kappa_A(BA)} - 1}{\sqrt{\kappa_A(BA)} + 1} = 1 - \frac{2}{1 + \sqrt{\kappa_A(BA)}}.$$

The following lemma is used in the analysis of multigrid and other linear preconditioners (it appears for example as Proposition 5.1 in [38]) to bound the condition number of the operator BA in terms of the extreme eigenvalues of the linear preconditioner error propagator $E = I - BA$. We have given our own short proof of this result for completeness.

Lemma 2.13 *If A and B are SPD, and $E = I - BA$ is such that:*

$$-C_1(Au, u) \leq (AEu, u) \leq C_2(Au, u), \qquad \forall u \in \mathcal{H},$$

for $C_1 \geq 0$ and $C_2 \geq 0$, then the above must hold with $C_2 < 1$, and it follows that:

$$\kappa_A(BA) \leq \frac{1 + C_1}{1 - C_2}.$$

Proof. First, since A and B are SPD, by Corollary 2.10 we have that $C_2 < 1$. Since $(AEu, u) = (A(I - BA)u, u) = (Au, u) - (ABAu, u)$, $\forall u \in \mathcal{H}$, it is immediately clear that

$$-C_1(Au, u) - (Au, u) \leq -(ABAu, u) \leq C_2(Au, u) - (Au, u), \qquad \forall u \in \mathcal{H}.$$

After multiplying by -1, we have

$$(1 - C_2)(Au, u) \leq (ABAu, u) \leq (1 + C_1)(Au, u), \qquad \forall u \in \mathcal{H}.$$

By Lemma 2.6, BA is A-SPD, and it follows from §2.1 that the eigenvalues of BA are real and positive, and lie in the interval defined by the Raleigh quotients of §2.1, generated by the A-inner-product. From above, we see that the interval is given by $[(1 - C_2), (1 + C_1)]$, and by Lemma 2.12 the result follows. □

The next corollary appears for example as Theorem 5.1 in [38].

Corollary 2.14 *If A and B are SPD, and BA is such that:*

$$C_1(Au, u) \leq (ABAu, u) \leq C_2(Au, u), \qquad \forall u \in \mathcal{H},$$

for $C_1 \geq 0$ and $C_2 \geq 0$, then the above must hold with $C_1 > 0$, and it follows that:

$$\kappa_A(BA) \leq \frac{C_2}{C_1}.$$

Proof. This follows easily from the argument used in the proof of Lemma 2.13. □

The following corollary, which relates the contraction property of a linear method to the condition number of the operator BA, appears without proof as Proposition 2.2 in [39].

Corollary 2.15 *If A and B are SPD, and $\|I - BA\|_A \leq \delta < 1$, then*

$$\kappa_A(BA) \leq \frac{1 + \delta}{1 - \delta}. \qquad (9)$$

Proof. This follows immediately from Lemma 2.13 with $\delta = \max\{C_1, C_2\}$. □

We comment briefly on an interesting implication of Lemma 2.13, which was apparently first noticed in [38]. It seems that even if a linear method is not convergent, for example if $C_1 > 1$ so that $\rho(E) > 1$, it may still be a good preconditioner. For example, if A and B are SPD, then by Corollary 2.10 we always have $C_2 < 1$. If it is the case that $C_2 << 1$, and if $C_1 > 1$ does not become too large, then $\kappa_A(BA)$

will be small and the conjugate gradient method will converge rapidly. A multigrid method will often diverge when applied to a problem with discontinuous coefficients unless special care is taken. Simply using conjugate gradient acceleration in conjunction with the multigrid method often yields a convergent (even rapidly convergent) method without employing any of the special techniques that have been developed for these problems; Lemma 2.13 may be the explanation for this behavior.

The following result from [39] connects the contraction number of the linear method used as the preconditioner to the contraction number of the resulting conjugate gradient method, and it shows that the conjugate gradient method always accelerates a linear method.

Theorem 2.16 *If A and B are SPD, and $\|I - BA\|_A \leq \delta < 1$, then $\delta_{\text{cg}} < \delta$.*

Proof. An abbreviated proof appears in [39]; we fill in the details here for completeness. Assume that the given linear method has contraction number bounded as $\|I - BA\|_A < \delta$. Now, since the function:

$$\frac{\sqrt{\kappa_A(BA)} - 1}{\sqrt{\kappa_A(BA)} + 1}$$

is an increasing function of $\kappa_A(BA)$, we can use the result of Lemma 2.13, namely $\kappa_A(BA) \leq (1+\delta)/(1-\delta)$, to bound the contraction rate of preconditioned conjugate gradient method as follows:

$$\delta_{\text{cg}} \leq \left(\frac{\sqrt{\kappa_A(BA)} - 1}{\sqrt{\kappa_A(BA)} + 1} \right) \leq \frac{\sqrt{\frac{1+\delta}{1-\delta}} - 1}{\sqrt{\frac{1+\delta}{1-\delta}} + 1} \cdot \frac{\sqrt{\frac{1+\delta}{1-\delta}} - 1}{\sqrt{\frac{1+\delta}{1-\delta}} - 1} = \frac{\frac{1+\delta}{1-\delta} - 2\sqrt{\frac{1+\delta}{1-\delta}} + 1}{\frac{1+\delta}{1-\delta} - 1} = \frac{1 - \sqrt{1-\delta^2}}{\delta}.$$

Note that this last term can be rewritten as:

$$\delta_{\text{cg}} \leq \frac{1 - \sqrt{1-\delta^2}}{\delta} = \delta \left(\frac{1}{\delta^2} [1 - \sqrt{1-\delta^2}] \right).$$

Now, since $0 < \delta < 1$, clearly $1 - \delta^2 < 1$, so that $1 - \delta^2 > (1-\delta^2)^2$. Thus, $\sqrt{1-\delta^2} > 1 - \delta^2$, or $-\sqrt{1-\delta^2} < \delta^2 - 1$, or finally $1 - \sqrt{1-\delta^2} < \delta^2$. Therefore, $(1/\delta^2)\left[1 - \sqrt{1-\delta^2}\right] < 1$, or

$$\delta_{\text{cg}} \leq \delta \left(\frac{1}{\delta^2} \left[1 - \sqrt{1-\delta^2} \right] \right) < \delta.$$

A more direct proof follows by recalling from Lemma 2.11 that the *best* possible contraction of the linear method, when provided with an optimal parameter, is given by:

$$\delta_{\text{opt}} = 1 - \frac{2}{1 + \kappa_A(BA)},$$

whereas the conjugate gradient contraction is

$$\delta_{\text{cg}} = 1 - \frac{2}{1 + \sqrt{\kappa_A(BA)}}.$$

Assuming $B \neq A^{-1}$, we always have $\kappa_A(BA) > 1$, so we must have that $\delta_{\text{cg}} < \delta_{\text{opt}} \leq \delta$. □

Remark 2.2. This result implies that it always pays in terms of an improved contraction number to use the conjugate gradient method to accelerate a linear method; the question remains of course whether the additional computational labor involved will be amortized by the improvement. This is not clear from the above analysis, and seems to be problem-dependent in practice.

Remark 2.3. Note that if a given linear method requires a parameter α as in Lemma 2.11 in order to be competitive, one can simply use the conjugate gradient method as an accelerator for the method without a parameter, avoiding the possibly costly estimation of a good parameter α. Theorem 2.16 guarantees that the resulting method will have superior contraction properties, without requiring the parameter estimation. This is exactly why additive multigrid and domain decomposition methods (which we discuss in more detail later) are used almost exclusively as preconditioners for conjugate gradient methods; in contrast to the multiplicative variants, which can be used effectively without a parameter, the additive variants always require a good parameter α to be effective, unless used as preconditioners.

To finish this section, we remark briefly on the complexity of Algorithm 2.2. If a tolerance of ϵ is required, then the computational cost to reduce the energy norm of the error below the tolerance can be determined from the expression above for δ_{cg} and from equation (7). To achieve a tolerance of ϵ after n iterations will require:

$$2\,\delta_{\text{cg}}^{n+1} = 2\left(\frac{\sqrt{\kappa_A(BA)}-1}{\sqrt{\kappa_A(BA)}+1}\right)^{n+1} < \epsilon.$$

Dividing by 2 and taking natural logarithms yields:

$$n \le \frac{\left|\ln\frac{\epsilon}{2}\right|}{\left|\ln\left(\frac{\sqrt{\kappa_A(BA)}-1}{\sqrt{\kappa_A(BA)}+1}\right)\right|}.$$

Using the approximation:

$$\ln\left(\frac{a-1}{a+1}\right) = \ln\left(\frac{1+(-1/a)}{1-(-1/a)}\right) = 2\left[\left(\frac{-1}{a}\right)+\frac{1}{3}\left(\frac{-1}{a}\right)^3+\frac{1}{5}\left(\frac{-1}{a}\right)^5+\cdots\right] < \frac{-2}{a},$$

we have that $|\ln[(\kappa_A^{1/2}(BA)-1)/(\kappa_A^{1/2}(BA)+1)]| > 2/\kappa_A^{1/2}(BA)$, so that:

$$n \le \frac{1}{2}\kappa_A^{1/2}(BA)\left|\ln\frac{\epsilon}{2}\right| + 1.$$

We then have that the maximum number of iterations required to reach an error on the order of the tolerance ϵ is:

$$n = O\left(\kappa_A^{1/2}(BA)\left|\ln\frac{\epsilon}{2}\right|\right).$$

If the cost of each iteration is $O(N)$, which will hold in the case of the sparse matrices generated by standard discretizations of elliptic partial differential equations, then the overall complexity to solve the problem is $O(\kappa_A^{1/2}(BA)N|\ln[\epsilon/2]|)$. If the preconditioner B is such that $\kappa_A^{1/2}(BA)$ can be bounded independently of the problem size N, then the complexity becomes (near) optimal order $O(N|\ln[\epsilon/2]|)$.

We make some final remarks regarding the idea of *spectral equivalence.*

Definition 2.1 *The SPD operators $B \in \mathbf{L}(\mathcal{H},\mathcal{H})$ and $A \in \mathbf{L}(\mathcal{H},\mathcal{H})$ are called spectrally equivalent if there exists constants $C_1 > 0$ and $C_2 > 0$ such that:*

$$C_1(Au,u) \le (Bu,u) \le C_2(Au,u), \qquad \forall u \in \mathcal{H}.$$

In other words, B defines an inner-product which induces a norm equivalent to the norm induced by the A-inner-product. If a given preconditioner B is spectrally equivalent to A^{-1}, then the condition number of the preconditioned operator BA is uniformly bounded.

Lemma 2.17 *If the SPD operators B and A^{-1} are spectrally equivalent, then:*

$$\kappa_A(BA) \le \frac{C_2}{C_1}.$$

Proof. By hypothesis, we have that $C_1(A^{-1}u,u) \le (Bu,u) \le C_2(A^{-1}u,u)$, $\forall u \in \mathcal{H}$. But this can be written as: $C_1(A^{-1/2}u, A^{-1/2}u) \le (A^{1/2}BA^{1/2}A^{-1/2}u, A^{-1/2}u) \le C_2(A^{-1/2}u, A^{-1/2}u)$, or:

$$C_1(\tilde{u},\tilde{u}) \le (A^{1/2}BA^{1/2}\tilde{u},\tilde{u}) \le C_2(\tilde{u},\tilde{u}), \qquad \forall \tilde{u} \in \mathcal{H}.$$

Now, since $BA = A^{-1/2}(A^{1/2}BA^{1/2})A^{1/2}$, we have that BA is similar to the SPD operator $A^{1/2}BA^{1/2}$. Therefore, the above inequality bounds the extreme eigenvalues of BA, and as a result the lemma follows by Lemma 2.12. $\square$

Remark 2.4. Of course, since all norms on finite-dimensional spaces are equivalent (which follows from the fact that all linear operators on finite-dimensional spaces are bounded), the idea of spectral equivalence is only important in the case of infinite-dimensional spaces, or when one considers how the equivalence constants behave as one increases the sizes of the spaces. This is exactly the issue in multigrid and domain decomposition theory: as one decreases the mesh size (increases the size of the spaces involved), one would like the quantity $\kappa_A(BA)$ to remain nicely bounded (in other words, one would like the equivalence constants to remain constant or grow only slowly). A discussion of these ideas appears in [31].

3. A theory for products and sums of operators

In this section, we present a unified approach for bounding the norms and condition numbers of products and sums of self-adjoint operators on a Hilbert space, derived from work due to Dryja and Widlund [14], Bramble et al. [8], and Xu [39]. Our particular approach is quite general in that we establish the main norm and condition number bounds without reference to subspaces; each of the three required assumptions for the theory involve only the operators on the original Hilbert space. Therefore, this product/sum operator theory may find use in other applications without natural subspace decompositions. Later, we will apply the product and sum operator theory to the case when the operators correspond to corrections in subspaces of the original space, as in multigrid and domain decomposition methods.

3.1 Basic product and sum operator theory

Let $\mathcal{H}$ be a real Hilbert space equipped with the inner-product $(\cdot,\cdot)$ inducing the norm $\|\cdot\| = (\cdot,\cdot)^{1/2}$. Let there be given an SPD operator $A \in \mathbf{L}(\mathcal{H},\mathcal{H})$ defining another inner-product on $\mathcal{H}$, which we denote as $(\cdot,\cdot)_A = (A\cdot,\cdot)$. This second inner-product also induces a norm $\|\cdot\|_A = (\cdot,\cdot)_A^{1/2}$. We are interested in general product and sum operators of the form

$$E = (I - T_J)(I - T_{J-1}) \cdots (I - T_1), \tag{10}$$

$$P = T_1 + T_2 + \cdots + T_J, \tag{11}$$

for some A-self-adjoint operators $T_k \in \mathbf{L}(\mathcal{H},\mathcal{H})$. If E is the error propagation operator of some linear method, then the convergence rate of this linear method will be governed by the norm of E. Similarly, if a preconditioned linear operator has the form of P, then the convergence rate of a conjugate gradient method employing this system operator will be governed by the condition number of P.

The A-norm is convenient here, as it is not difficult to see that P is A-self-adjoint, as well as $E^s = EE^*$. Therefore, we will be interested in deriving bounds of the form:

$$\|E\|_A^2 \le \delta < 1, \qquad \kappa_A(P) = \frac{\lambda_{\max}(P)}{\lambda_{\min}(P)} \le \gamma. \tag{12}$$

The remainder of this section is devoted to establishing some minimal assumptions on the operators T_k in order to derive bounds of the form in equation (12). If we define $E_k = (I-T_k)(I-T_{k-1})\cdots(I-T_1)$, and define $E_0 = I$ and $E_J = E$, then we have the following relationships.

Lemma 3.1 *The following relationships hold for $k = 1, \ldots, J$:*

(1) $E_k = (I - T_k)E_{k-1}$
(2) $E_{k-1} - E_k = T_k E_{k-1}$
(3) $I - E_k = \sum_{i=1}^{k} T_i E_{i-1}$

Proof. The first relationship is obvious from the definition of E_k, and the second follows easily from the first. Taking $E_0 = I$, and summing the second relationship from $i = 1$ to $i = k$ gives the third. □

Regarding the operators T_k, we make the following assumption:

Assumption 3.1 *The operators $T_k \in \mathbf{L}(\mathcal{H},\mathcal{H})$ are A-self-adjoint, A-non-negative, and*

$$\rho(T_k) = \|T_k\|_A \le \omega < 2, \quad k = 1, \ldots, J.$$

Note that this implies that $0 \le \lambda_i(T_k) \le \omega < 2$, $k = 1, \ldots, J$.

The following simple lemma, first appearing in [8], will often be useful at various points in the analysis of the product and sum operators.

Lemma 3.2 *Under Assumption 3.1, it holds that*

$$(AT_k u, T_k u) \le \omega(AT_k u, u), \quad \forall u \in \mathcal{H}.$$

Proof. Since T_k is A-self-adjoint, we know that $\rho(T_k) = \|T_k\|_A$, so that

$$\rho(T_k) = \max_{v \neq 0} \frac{(AT_k v, v)}{(Av, v)} \leq \omega < 2,$$

so that $(AT_k v, v) \leq \omega(Av, v)$, $\forall v \in \mathcal{H}$. But this gives $(AT_k u, T_k u) = (AT_k^{1/2} T_k u, T_k^{1/2} u) = (AT_k T_k^{1/2} u, T_k^{1/2} u)$ $= (AT_k v, v) \leq \omega(Av, v) = \omega(AT_k^{1/2} u, T_k^{1/2} u) = \omega(AT_k u, u)$, $\forall u \in \mathcal{H}$. □

The next lemma, also appearing first in [8], will be a key tool in the analysis of the product operator.

Lemma 3.3 *Under Assumption 3.1, it holds that*

$$(2-\omega) \sum_{k=1}^{J} (AT_k E_{k-1} v, E_{k-1} v) \leq \|v\|_A^2 - \|E_J v\|_A^2.$$

Proof. Employing the relationships in Lemma 3.1, we can rewrite the following difference as

$$\begin{aligned} \|E_{k-1} v\|_A^2 - \|E_k v\|_A^2 &= (AE_{k-1} v, E_{k-1} v) - (AE_k v, E_k v) \\ &= (AE_{k-1} v, E_{k-1} v) - (A[I - T_k] E_{k-1} v, [I - T_k] E_{k-1} v) \\ &= 2(AT_k E_{k-1} v, E_{k-1} v) - (AT_k E_{k-1} v, T_k E_{k-1} v) \end{aligned}$$

By Lemma 3.2 we have $(AT_k E_{k-1} v, T_k E_{k-1} v) \leq \omega(AT_k E_{k-1} v, E_{k-1} v)$, so that

$$\|E_{k-1} v\|_A^2 - \|E_k v\|_A^2 \geq (2-\omega)(AT_k E_{k-1} v, E_{k-1} v).$$

With $E_0 = I$, by summing from $k = 1$ to $k = J$ we have:

$$\|v\|_A^2 - \|E_J v\|_A^2 \geq (2-\omega) \sum_{k=1}^{J} (AT_k E_{k-1} v, E_{k-1} v).$$

□

We now state four simple assumptions which will, along with Assumption 3.1, allow us to give norm and condition number bounds by employing the previous lemmas. These four assumptions form the basis for the product and sum theory, and the remainder of our work will chiefly involve establishing conditions under which these assumptions are satisfied.

Assumption 3.2 *(Splitting assumption) There exists $C_0 > 0$ such that*

$$\|v\|_A^2 \leq C_0 \sum_{k=1}^{J} (AT_k v, v), \quad \forall v \in \mathcal{H}.$$

Assumption 3.3 *(Composite assumption) There exists $C_1 > 0$ such that*

$$\|v\|_A^2 \leq C_1 \sum_{k=1}^{J} (AT_k E_{k-1} v, E_{k-1} v), \quad \forall v \in \mathcal{H}.$$

Assumption 3.4 *(Product assumption) There exists $C_2 > 0$ such that*

$$\sum_{k=1}^{J} (AT_k v, v) \leq C_2 \sum_{k=1}^{J} (AT_k E_{k-1} v, E_{k-1} v), \quad \forall v \in \mathcal{H}.$$

Assumption 3.5 *(Sum assumption) There exists $C_3 > 0$ such that*

$$\sum_{k=1}^{J} (AT_k v, v) \leq C_3 \|v\|_A^2, \quad \forall v \in \mathcal{H}.$$

Lemma 3.4 *Under Assumptions 3.2 and 3.4, Assumption 3.3 holds with $C_1 = C_0C_2$.*

Proof. This is immediate, since

$$\|v\|_A^2 \leq C_0 \sum_{k=1}^{J} (AT_k v, v) \leq C_0C_2 \sum_{k=1}^{J} (AT_k E_{k-1}v, E_{k-1}v), \quad \forall v \in \mathcal{H}.$$

□

Remark 3.5. In what follows, it will be necessary to satisfy Assumption 3.3 for some constant C_1. Lemma 3.4 provides a technique for verifying Assumption 3.3 by verifying Assumptions 3.2 and 3.4 separately. In certain cases it will still be necessary to verify Assumption 3.3 directly.

The following theorems provide a fundamental framework for analyzing product and sum operators, employing only the five assumptions previously stated. A version of the product theorem similar to the one below first appeared in [8]. Theorems for sum operators were established early by Dryja and Widlund [12].

Theorem 3.5 *Under Assumptions 3.1 and 3.3, the product operator (10) satisfies:*

$$\|E\|_A^2 \leq 1 - \frac{2-\omega}{C_1}.$$

Proof. To prove the result, it suffices to show that

$$\|Ev\|_A^2 \leq \left(1 - \frac{2-\omega}{C_1}\right) \|v\|_A^2, \quad \forall v \in \mathcal{H},$$

or that

$$\|v\|_A^2 \leq \frac{C_1}{2-\omega} \left(\|v\|_A^2 - \|Ev\|_A^2\right), \quad \forall v \in \mathcal{H}.$$

By Lemma 3.3 (which required only Assumption 3.1), it is enough to show

$$\|v\|_A^2 \leq C_1 \sum_{k=1}^{J} (AT_k E_{k-1}v, E_{k-1}v), \quad \forall v \in \mathcal{H}.$$

But, by Assumption 3.3 this result holds, and the theorem follows. □

Corollary 3.6 *Under Assumptions 3.1, 3.2, and 3.4, the product operator (10) satisfies:*

$$\|E\|_A^2 \leq 1 - \frac{2-\omega}{C_0C_2}.$$

Proof. This follows from Theorem 3.5 and Lemma 3.4. □

Theorem 3.7 *Under Assumptions 3.1, 3.2, and 3.5, the sum operator (11) satisfies:*

$$\kappa_A(P) \leq C_0C_3.$$

Proof. This result follows immediately from Assumptions 3.2 and 3.5, since $P = \sum_{k=1}^{J} T_k$ is A-self-adjoint by Assumption 3.1, and since

$$\frac{1}{C_0}(Av, v) \leq \sum_{k=1}^{J} (AT_k v, v) = (APv, v) \leq C_3(Av, v), \quad \forall v \in \mathcal{H}.$$

This implies that $C_0^{-1} \leq \lambda_i(P) \leq C_3$, and by Lemma 2.12 it holds that $\kappa_A(P) \leq C_0C_3$. □

The constants C_0 and C_1 in Assumptions 3.2 and 3.3 will depend on the specific application; we will discuss estimates for C_0 and C_1 in the following sections. The constants C_2 and C_3 in Assumptions 3.4 and 3.5 will also depend on the specific application; however, we can derive bounds which grow with the number of operators J, which will always hold without additional assumptions. Both of these default or worst case results appear essentially in [8]. First, we recall the Cauchy-Schwarz inequality in $\mathbb{R}^n$, and state a useful corollary.

Lemma 3.8 *If $a_k, b_k \in \mathbb{R}$, $k = 1, \ldots, n$, then it holds that*

$$\left(\sum_{k=1}^{n} a_k b_k\right)^2 \leq \left(\sum_{k=1}^{n} a_k^2\right)\left(\sum_{k=1}^{n} b_k^2\right).$$

Proof. See for example [22]. □

Corollary 3.9 *If $a_k \in \mathbb{R}$, $k = 1, \ldots, n$, then it holds that*

$$\left(\sum_{k=1}^{n} a_k\right)^2 \leq n \sum_{k=1}^{n} a_k^2.$$

Proof. This follows easily from Lemma 3.8 by taking $b_k = 1$ for all k. □

Lemma 3.10 *Under only Assumption 3.1, we have that Assumption 3.4 holds, where:*

$$C_2 = 2 + \omega^2 J(J-1).$$

Proof. We must show that

$$\sum_{k=1}^{J}(AT_k v, v) \leq [2 + \omega^2 J(J-1)] \sum_{k=1}^{J}(AT_k E_{k-1}v, E_{k-1}v), \quad \forall v \in \mathcal{H}.$$

By Lemma 3.1, we have that

$$(AT_k v, v) = (AT_k v, E_{k-1}v) + (AT_k v, [I - E_{k-1}]v) = (AT_k v, E_{k-1}v) + \sum_{i=1}^{k-1}(AT_k v, T_i E_{i-1}v)$$

$$\leq (AT_k v, v)^{1/2}(AT_k E_{k-1}v, E_{k-1}v)^{1/2} + \sum_{i=1}^{k-1}(AT_k v, T_k v)^{1/2}(AT_i E_{i-1}v, T_i E_{i-1}v)^{1/2}.$$

By Lemma 3.2, we have

$$(AT_k v, v) \leq (AT_k v, v)^{1/2}(AT_k E_{k-1}v, E_{k-1}v)^{1/2} + \omega(AT_k v, v)^{1/2}\sum_{i=1}^{k-1}(AT_i E_{i-1}v, E_{i-1}v)^{1/2},$$

or finally

$$(AT_k v, v) \leq \left[(AT_k E_{k-1}v, E_{k-1}v)^{1/2} + \omega \sum_{i=1}^{k-1}(AT_i E_{i-1}v, E_{i-1}v)^{1/2}\right]^2. \tag{13}$$

Employing Corollary 3.9 for the two explicit terms in the inequality (13) yields:

$$(AT_k v, v) \leq 2\left[(AT_k E_{k-1}v, E_{k-1}v) + \omega^2 \left[\sum_{i=1}^{k-1}(AT_i E_{i-1}v, E_{i-1}v)^{1/2}\right]^2\right].$$

Employing Corollary 3.9 again for the $k-1$ terms in the sum yields

$$(AT_k v, v) \leq 2\left[(AT_k E_{k-1}v, E_{k-1}v) + \omega^2(k-1)\sum_{i=1}^{k-1}(AT_i E_{i-1}v, E_{i-1}v)\right].$$

Summing the terms, and using the fact that the T_k are A-non-negative, we have

$$\sum_{k=1}^{J}(AT_kv,v) \leq 2\left[\sum_{k=1}^{J}\left\{(AT_kE_{k-1}v,E_{k-1}v)+\omega^2(k-1)\sum_{i=1}^{k-1}(AT_iE_{i-1}v,E_{i-1}v)\right\}\right]$$

$$\leq 2\left[1+\omega^2\sum_{i=1}^{J}(i-1)\right]\sum_{k=1}^{J}(AT_kE_{k-1}v,E_{k-1}v).$$

Since $\sum_{i=1}^{J} i = (J+1)J/2$, we have that the lemma follows. □

Lemma 3.11 *Under only Assumption 3.1, we have that Assumption 3.5 holds, where:*

$$C_3 = \omega J.$$

Proof. By Assumption 3.1, we have

$$\sum_{k=1}^{J}(AT_kv,v) \leq \sum_{k=1}^{J}(AT_kv,T_kv)^{1/2}(Av,v)^{1/2} \leq \sum_{k=1}^{J}\omega(Av,v) = \omega J\|v\|_A^2,$$

so that $C_3 = \omega J$. □

Remark 3.6. Note that since Lemmas 3.10 and 3.11 provide default (worst case) estimates for C_2 and C_3 in Assumptions 3.4 and 3.5, due to Lemma 3.4 it suffices to estimate only C_0 in Assumption 3.2 in order to employ the general product and sum operator theorems (namely Corollary 3.6 and Theorem 3.7).

3.2 The interaction hypothesis

We now consider an additional assumption, which will be natural in multigrid and domain decomposition applications, regarding the "interaction" of the operators T_k. This assumption brings together more closely the theory for the product and sum operators. The constants C_2 and C_3 in Assumptions 3.4 and 3.5 can both be estimated in terms of the constants C_4 and C_5 appearing below, which will be determined by the interaction properties of the operators T_k. We will further investigate the interaction properties more precisely in a moment. This approach to quantifying the interaction of the operators T_k is similar to that taken in [39].

Assumption 3.6 *(Interaction assumption - weak) There exists $C_4 > 0$ such that*

$$\sum_{k=1}^{J}\sum_{i=1}^{k-1}(AT_ku_k,T_iv_i) \leq C_4\left(\sum_{k=1}^{J}(AT_ku_k,u_k)\right)^{1/2}\left(\sum_{i=1}^{J}(AT_iv_i,v_i)\right)^{1/2}, \quad \forall u_k,v_i \in \mathcal{H}.$$

Assumption 3.7 *(Interaction assumption - strong) There exists $C_5 > 0$ such that*

$$\sum_{k=1}^{J}\sum_{i=1}^{J}(AT_ku_k,T_iv_i) \leq C_5\left(\sum_{k=1}^{J}(AT_ku_k,u_k)\right)^{1/2}\left(\sum_{i=1}^{J}(AT_iv_i,v_i)\right)^{1/2}, \quad \forall u_k,v_i \in \mathcal{H}.$$

Remark 3.7. We introduce the terminology "weak" and "strong" because in the *weak* interaction assumption above, the interaction constant C_4 is defined by considering the interaction of a particular operator T_k *only* with operators T_i with $i < k$; note that this implies an ordering of the operators T_k, and different orderings may produce different values for C_4. In the *strong* interaction assumption above, the interaction constant C_5 is defined by considering the interaction of a particular operator T_k with *all* operators T_i (the ordering of the operators T_k is now unimportant).

The interaction assumptions can be used to bound the constants C_2 and C_3 in Assumptions 3.4 and 3.5.

Lemma 3.12 *Under Assumptions 3.1 and 3.6, we have that Assumption 3.4 holds, where:*

$$C_2 = (1 + C_4)^2.$$

Proof. Consider

$$\sum_{k=1}^{J}(AT_k v, v) = \sum_{k=1}^{J}\{(AT_k v, E_{k-1}v) + (AT_k v, [I - E_{k-1}]v)\} \tag{14}$$

$$= \sum_{k=1}^{J}(AT_k v, E_{k-1}v) + \sum_{k=1}^{J}\sum_{i=1}^{k-1}(AT_k v, T_i E_{i-1}v).$$

For the first term, the Cauchy-Schwarz inequalities give

$$\sum_{k=1}^{J}(AT_k v, E_{k-1}v) \leq \sum_{k=1}^{J}(AT_k v, v)^{1/2}(AT_k E_{k-1}v, E_{k-1}v)^{1/2}$$

$$\leq \left(\sum_{k=1}^{J}(AT_k v, v)\right)^{1/2}\left(\sum_{k=1}^{J}(AT_k E_{k-1}v, E_{k-1}v)\right)^{1/2}.$$

For the second term, we have by Assumption 3.6 that

$$\sum_{k=1}^{J}\sum_{i=1}^{k-1}(AT_k v, T_i E_{i-1}v) \leq C_4\left(\sum_{k=1}^{J}(AT_k v, v)\right)^{1/2}\left(\sum_{k=1}^{J}(AT_k E_{k-1}v, E_{k-1}v)\right)^{1/2}.$$

Thus, together we have

$$\sum_{k=1}^{J}(AT_k v, v) \leq (1 + C_4)\left(\sum_{k=1}^{J}(AT_k v, v)\right)^{1/2}\left(\sum_{k=1}^{J}(AT_k E_{k-1}v, E_{k-1}v)\right)^{1/2},$$

which yields

$$\sum_{k=1}^{J}(AT_k v, v) \leq (1 + C_4)^2\sum_{k=1}^{J}(AT_k E_{k-1}v, E_{k-1}v).$$

□

Lemma 3.13 *Under Assumptions 3.1 and 3.7, we have that Assumption 3.5 holds, where:*

$$C_3 = C_5.$$

Proof. Consider first that $\forall v \in \mathcal{H}$, Assumption 3.7 implies

$$\|\sum_{k=1}^{J}T_k v\|_A^2 = \sum_{k=1}^{J}\sum_{i=1}^{J}(AT_k v, T_i v) \leq C_5\left(\sum_{k=1}^{J}(AT_k v, v)\right)^{1/2}\left(\sum_{i=1}^{J}(AT_i v, v)\right)^{1/2}$$

$$= C_5\sum_{k=1}^{J}(AT_k v, v).$$

If $P = \sum_{k=1}^{J}T_k$, then we have shown that $(APv, Pv) \leq C_5(APv, v)$, $\forall v \in \mathcal{H}$, so that

$$(APv, v) \leq (APv, Pv)^{1/2}(Av, v)^{1/2} \leq C_5^{1/2}(APv, v)^{1/2}(Av, v)^{1/2}, \forall v \in \mathcal{H}.$$

This implies that $(APv, v) \leq C_5\|v\|_A^2, \forall v \in \mathcal{H}$, which proves the lemma. □

The constants C_4 and C_5 can be further estimated, in terms of the following two *interaction matrices.* An early approach employing an interaction matrix appears in [8]; the form appearing below is most closely related to that used in [17] and [39]. The idea of employing a strictly upper-triangular interaction matrix to improve the bound for the weak interaction property is due to Hackbusch [17]. The default bound for the strictly upper-triangular matrix is also due to Hackbusch [17].

Definition 3.1 *Let Ξ be the strictly upper-triangular part of the interaction matrix $\Theta \in \mathbf{L}(\mathbb{R}^J, \mathbb{R}^J)$, which is defined to have as entries Θ_{ij} the smallest constants satisfying:*

$$|(AT_iu, T_jv)| \le \Theta_{ij}(AT_iu, T_iu)^{1/2}(AT_jv, T_jv)^{1/2}, \quad 1 \le i, j \le J, \quad \forall u, v \in \mathcal{H}.$$

The matrix Θ is symmetric, and $0 \le \Theta_{ij} \le 1, \forall i, j$. Also, we have that $\Theta = I + \Xi + \Xi^T$.

Lemma 3.14 *It holds that $\|\Xi\|_2 \le \rho(\Theta)$. Also, $\|\Xi\|_2 \le \sqrt{J(J-1)/2}$ and $1 \le \rho(\Theta) \le J$.*

Proof. Since Θ is symmetric, we know that $\rho(\Theta) = \|\Theta\|_2 = \max_{\mathbf{x}\neq 0} \|\Theta\mathbf{x}\|_2/\|\mathbf{x}\|_2$. Now, given any $\mathbf{x} \in \mathbb{R}^J$, define $\bar{\mathbf{x}} \in \mathbb{R}^J$ such that $\bar{x}_i = |x_i|$. Note that $\|\mathbf{x}\|_2^2 = \sum_{i=1}^J |x_i|^2 = \|\bar{\mathbf{x}}\|_2^2$, and since $0 \le \Theta_{ij} \le 1$, we have that

$$\|\Theta\mathbf{x}\|_2^2 = \sum_{i=1}^J \left(\sum_{j=1}^J \Theta_{ij}x_j\right)^2 \le \sum_{i=1}^J \left(\sum_{j=1}^J \Theta_{ij}|x_j|\right)^2 = \|\Theta\bar{\mathbf{x}}\|_2^2.$$

Therefore, it suffices to consider only $\mathbf{x} \in \mathbb{R}^J$ with $x_i \ge 0$. For such an $\mathbf{x} \in \mathbb{R}^J$, it is clear that $\|\Xi\mathbf{x}\|_2 \le \|\Theta\mathbf{x}\|_2$, so we must have that

$$\|\Xi\|_2 = \max_{\mathbf{x}\neq 0} \frac{\|\Xi\mathbf{x}\|_2}{\|\mathbf{x}\|_2} \le \max_{\mathbf{x}\neq 0} \frac{\|\Theta\mathbf{x}\|_2}{\|\mathbf{x}\|_2} = \|\Theta\|_2 = \rho(\Theta).$$

The worst case estimate $\|\Xi\|_2 \le \sqrt{J(J-1)/2}$ follows easily, since $0 \le \Xi_{ij} \le 1$, and since:

$$[\Xi^T\Xi]_{ij} = \sum_{k=1}^J [\Xi^T]_{ik}\Xi_{kj} = \sum_{k=1}^J \Xi_{ki}\Xi_{kj} = \sum_{k=1}^{\min\{i-1,j-1\}} \Xi_{ki}\Xi_{kj} \le \min\{i-1, j-1\}.$$

Thus, we have that

$$\|\Xi\|_2^2 = \rho(\Xi^T\Xi) \le \|\Xi^T\Xi\|_1 = \max_j \left\{\sum_{i=1}^J |\ [\Xi^T\Xi]_{ij}\ |\right\} \le \sum_{i=1}^J (i-1) = \frac{J(J-1)}{2}.$$

It remains to show that $1 \le \rho(\Theta) \le J$. The upper bound follows easily since we know that $0 \le \Theta_{ij} \le 1$, and so that $\rho(\Theta) \le \|\Theta\|_1 = \max_j\{\sum_i |\Theta_{ij}|\} \le J$. Regarding the lower bound, recall that the trace of a matrix is equal to the sum of it's eigenvalues. Since all diagonal entries of Θ are unity, the trace is simply equal to J. If all the eigenvalues of Θ are unity, we are done. If we suppose there is at least one eigenvalue $\lambda_i < 1$ (possibly negative), then in order for the J eigenvalues of Θ to sum to J, there must be a corresponding eigenvalue $\lambda_j > 1$. Therefore, $\rho(\Theta) \ge 1$. □

We now have the following lemmas.

Lemma 3.15 *Under Assumption 3.1 we have that Assumption 3.6 holds, where:*

$$C_4 \le \omega\|\Xi\|_2.$$

Proof. Consider

$$\sum_{k=1}^J \sum_{i=1}^{k-1} (AT_ku_k, T_iv_i) \le \sum_{k=1}^J \sum_{i=1}^J \Xi_{ik}\|T_ku_k\|_A\|T_iv_i\|_A = (\Xi\mathbf{x}, \mathbf{y})_2,$$

where $\mathbf{x}, \mathbf{y} \in \mathbb{R}^J$, $x_k = \|T_ku_k\|_A$, $y_i = \|T_iv_i\|_A$, and $(\cdot,\cdot)_2$ is the usual Euclidean inner-product in $\mathbb{R}^J$. Now, we have that

$$(\Xi\mathbf{x}, \mathbf{y})_2 \le \|\Xi\|_2\|\mathbf{x}\|_2\|\mathbf{y}\|_2 = \|\Xi\|_2 \left(\sum_{k=1}^J (AT_ku_k, T_ku_k)\right)^{1/2} \left(\sum_{i=1}^J (AT_iv_i, T_iv_i)\right)^{1/2}$$

$$\leq \omega \|\Xi\|_2 \left(\sum_{k=1}^{J} (AT_k u_k, u_k) \right)^{1/2} \left(\sum_{i=1}^{J} (AT_i v_i, v_i) \right)^{1/2}.$$

Finally, this gives

$$\sum_{k=1}^{J} \sum_{i=1}^{k-1} (AT_k u_k, T_i v_i) \leq \omega \|\Xi\|_2 \left(\sum_{k=1}^{J} (AT_k u_k, u_k) \right)^{1/2} \left(\sum_{i=1}^{J} (AT_i v_i, v_i) \right)^{1/2}, \quad \forall u_k, v_i \in \mathcal{H}.$$

□

Lemma 3.16 *Under Assumption 3.1 we have that Assumption 3.7 holds, where:*

$$C_5 \leq \omega \rho(\Theta).$$

Proof. Consider

$$\sum_{k=1}^{J} \sum_{i=1}^{J} (AT_k u_k, T_i v_i) \leq \sum_{k=1}^{J} \sum_{i=1}^{J} \Theta_{ik} \|T_k u_k\|_A \|T_i v_i\|_A = (\Theta \mathbf{x}, \mathbf{y})_2,$$

where $\mathbf{x}, \mathbf{y} \in \mathbb{R}^J$, $x_k = \|T_k u_k\|_A$, $y_i = \|T_i v_i\|_A$, and $(\cdot,\cdot)_2$ is the usual Euclidean inner-product in $\mathbb{R}^J$. Now, since Θ is symmetric, we have that

$$(\Theta \mathbf{x}, \mathbf{y})_2 \leq \rho(\Theta) \|\mathbf{x}\|_2 \|\mathbf{y}\|_2 = \rho(\Theta) \left(\sum_{k=1}^{J} (AT_k u_k, T_k u_k) \right)^{1/2} \left(\sum_{i=1}^{J} (AT_i v_i, T_i v_i) \right)^{1/2}$$

$$\leq \omega \rho(\Theta) \left(\sum_{k=1}^{J} (AT_k u_k, u_k) \right)^{1/2} \left(\sum_{i=1}^{J} (AT_i v_i, v_i) \right)^{1/2}.$$

Finally, this gives

$$\sum_{k=1}^{J} \sum_{i=1}^{J} (AT_k u_k, T_i v_i) \leq \omega \rho(\Theta) \left(\sum_{k=1}^{J} (AT_k u_k, u_k) \right)^{1/2} \left(\sum_{i=1}^{J} (AT_i v_i, v_i) \right)^{1/2}, \quad \forall u_k, v_i \in \mathcal{H}.$$

□

This leads us finally to

Lemma 3.17 *Under Assumption 3.1 we have that Assumption 3.4 holds, where:*

$$C_2 = (1 + \omega \|\Xi\|_2)^2.$$

Proof. This follows from Lemmas 3.12 and 3.15. □

Lemma 3.18 *Under Assumption 3.1 we have that Assumption 3.5 holds, where:*

$$C_3 = \omega \rho(\Theta).$$

Proof. This follows from Lemmas 3.13 and 3.16. □

Remark 3.8. Note that Lemmas 3.17 and 3.14 reproduce the worst case estimate for C_2 given in Lemma 3.10, since:

$$C_2 = (1 + \omega \|\Xi\|_2)^2 \leq 2(1 + \omega^2 \|\Xi\|_2^2) \leq 2 + \omega^2 J(J-1).$$

In addition, Lemmas 3.18 and 3.14 reproduce the worst case estimate of $C_3 = \omega \rho(\Theta) \leq \omega J$ given in Lemma 3.11.

3.3 Allowing for a global operator

Consider the product and sum operators

$$E = (I - T_J)(I - T_{J-1}) \cdots (I - T_0), \tag{15}$$

$$P = T_0 + T_1 + \cdots + T_J, \tag{16}$$

where we now include a special operator T_0, which we assume may interact with *all* of the other operators. For example, T_0 might later represent some "global" coarse space operator in a domain decomposition method. Note that if such a global operator is included directly in the analysis of the previous section, then the bounds on $\|\Xi\|_2$ and $\rho(\Theta)$ necessarily depend on the number of operators; thus, to develop an optimal theory, we must exclude T_0 from the interaction hypothesis. This was recognized early in the domain decomposition community, and the modification of the theory in the previous sections to allow for such a global operator has been achieved mainly by Widlund and his co-workers. We will follow essentially their approach in this section.

In the following, we will use many of the results and assumptions from the previous section, where we now explicitly require that the $k = 0$ term *always* be included; the only exception to this will be the interaction assumption, which will still involve only the $k \neq 0$ terms. Regarding the minor changes to the results of the previous sections, note that we must now define $E_{-1} = I$, which modifies Lemma 3.1 in that

$$I - E_k = \sum_{i=0}^{k} T_i E_{i-1},$$

the sum beginning at $k = 0$. We make the usual Assumption 3.1 on the operators T_k (now including T_0 also), and we then have the results from Lemmas 3.2 and 3.3. The main assumptions for the theory are as in Assumptions 3.2, 3.4, and 3.5, with the additional term $k = 0$ included in each assumption. The two main results in Theorems 3.5 and 3.7 are unchanged. The default bounds for C_2 and C_3 given in Lemmas 3.10 and 3.11 now must take into account the additional operator T_0:

$$C_2 = 2 + \omega^2 J(J+1), \qquad C_3 = \omega(J+1).$$

The remaining analysis becomes now somewhat different from the case when T_0 is not present. First, we will quantify the interaction properties of the remaining operators T_k for $k \neq 0$ exactly as was done earlier, except that we must now employ the strong interaction assumption (Assumption 3.7) for both the product and sum theories. (In the previous section, we were able to use only the weak interaction assumption for the product operator.) This leads us to the following two lemmas.

Lemma 3.19 *Under Assumptions 3.1 (including T_0), 3.6 (excluding T_0), and 3.7 (excluding T_0), we have that Assumption 3.4 (including T_0) holds, where:*

$$C_2 = [1 + \omega^{1/2} C_5^{1/2} + C_4]^2.$$

Proof. Beginning with Lemma 3.1 we have that

$$\sum_{k=0}^{J} (AT_k v, v) = (AT_0 v, v) + \sum_{k=1}^{J} \{(AT_k v, E_{k-1} v) + (AT_k v, [I - E_{k-1}]v)\}$$

$$= \sum_{k=0}^{J} (AT_k v, E_{k-1} v) + \sum_{k=1}^{J} \sum_{i=0}^{k-1} (AT_k v, T_i E_{i-1} v)$$

$$= \sum_{k=0}^{J} (AT_k v, E_{k-1} v) + \sum_{k=1}^{J} (AT_k v, T_0 v) + \sum_{k=1}^{J} \sum_{i=1}^{k-1} (AT_k v, T_i E_{i-1} v) = \mathbf{S}_1 + \mathbf{S}_2 + \mathbf{S}_3. \tag{17}$$

We now estimate $\mathbf{S}_1$, $\mathbf{S}_2$, and $\mathbf{S}_3$ separately. For $\mathbf{S}_1$, we employ the Cauchy-Schwarz inequality to obtain

$$\mathbf{S}_1 = \sum_{k=0}^{J} (AT_k v, E_{k-1} v) \leq \sum_{k=0}^{J} (AT_k v, v)^{1/2} (AT_k E_{k-1} v, E_{k-1} v)^{1/2}$$

$$\leq \left(\sum_{k=0}^{J}(AT_kv,v)\right)^{1/2}\left(\sum_{k=0}^{J}(AT_kE_{k-1}v,E_{k-1}v)\right)^{1/2}.$$

To bound $\mathbf{S}_2$, we employ Assumption 3.7 as follows:

$$\mathbf{S}_2=\sum_{k=1}^{J}(AT_kv,T_0v)\leq \|\sum_{k=1}^{J}T_kv\|_A\|T_0v\|_A=\left(\sum_{k=1}^{J}\sum_{i=1}^{J}(AT_kv,T_iv)\right)^{1/2}(AT_0v,T_0v)^{1/2}$$

$$\leq \omega^{1/2}C_5^{1/2}\left(\sum_{k=1}^{J}(AT_kv,v)\right)^{1/2}(AT_0v,v)^{1/2}$$

$$\leq \omega^{1/2}C_5^{1/2}\left(\sum_{k=0}^{J}(AT_kv,v)\right)^{1/2}\left(\sum_{k=0}^{J}(AT_kE_{k-1}v,E_{k-1}v)\right)^{1/2}.$$

We now bound $\mathbf{S}_3$, employing Assumption 3.6 as

$$\mathbf{S}_3=\sum_{k=1}^{J}\sum_{i=1}^{k-1}(AT_kv,T_iE_{i-1}v)\leq C_4\left(\sum_{k=1}^{J}(AT_kv,v)\right)^{1/2}\left(\sum_{k=1}^{J}(AT_kE_{k-1}v,E_{k-1}v)\right)^{1/2}$$

$$\leq C_4\left(\sum_{k=0}^{J}(AT_kv,v)\right)^{1/2}\left(\sum_{k=0}^{J}(AT_kE_{k-1}v,E_{k-1}v)\right)^{1/2}.$$

Putting the bounds for $\mathbf{S}_1$, $\mathbf{S}_2$, and $\mathbf{S}_3$ together, dividing (17) by $\sum_{k=1}^{J}(AT_kv,v)$ and squaring, yields

$$\sum_{k=0}^{J}(AT_kv,v)\leq [1+\omega^{1/2}C_5^{1/2}+C_4]^2\sum_{k=0}^{J}(AT_kE_{k-1}v,E_{k-1}v).$$

Therefore, Assumption 3.4 holds, where:

$$C_2=[1+\omega^{1/2}C_5^{1/2}+C_4]^2.$$

□

Results similar to the next lemma are used in several recent papers on domain decomposition [14]; the proof is quite simple once the proof of Lemma 3.13 is available.

Lemma 3.20 *Under Assumptions 3.1 (including T_0) and 3.7 (excluding T_0), we have that Assumption 3.5 (including T_0) holds, where:*

$$C_3=\omega+C_5.$$

Proof. The proof of Lemma 3.13 gives immediately $\sum_{k=1}^{J}(AT_kv,v)\leq C_5\|v\|_A^2$. Now, since $(AT_0v,v)\leq \omega\|v\|_A^2$, we simply add in the $k=0$ term, yielding

$$\sum_{k=0}^{J}(AT_kv,v)\leq(\omega+C_5)\|v\|_A^2.$$

□

We finish the section by relating the constants C_2 and C_3 (required for Corollary 3.6 and Theorem 3.7) to the interaction matrices. The constants C_4 and C_5 are estimated by using the interaction matrices exactly as before, since the interaction conditions still involve only the operators T_k for $k\neq 0$.

Lemma 3.21 *Under Assumption 3.1 we have that Assumption 3.4 holds, where:*

$$C_2 \leq 6[1 + \omega^2 \rho(\Theta)^2].$$

Proof. From Lemma 3.19 we have that

$$C_2 = [1 + \omega^{1/2} C_5^{1/2} + C_4]^2.$$

Now, from Lemmas 3.15 and 3.16, and since $\omega < 2$, it follows that

$$C_2 = [1 + \omega^{1/2} C_5^{1/2} + C_4]^2 \leq [1 + \sqrt{2}(\omega\rho(\Theta))^{1/2} + \omega\|\Xi\|_2]^2.$$

Employing first Lemma 3.14 and then Corollary 3.9 twice, we have

$$\begin{aligned} C_2 &\leq [1 + \sqrt{2}(\omega\rho(\Theta))^{1/2} + \omega\rho(\Theta)]^2 \leq 3[1 + 2\omega\rho(\Theta) + \omega^2\rho(\Theta)^2] \\ &= 3[1 + \omega\rho(\Theta)]^2 \leq 6[1 + \omega^2\rho(\Theta)^2]. \end{aligned}$$

□

Lemma 3.22 *Under Assumption 3.1 we have that Assumption 3.5 holds, where:*

$$C_3 \leq \omega(\rho(\Theta) + 1).$$

Proof. From Lemmas 3.20 and 3.16 it follows that

$$C_3 = \omega + C_5 \leq \omega + \omega\rho(\Theta) = \omega(\rho(\Theta) + 1).$$

□

Remark 3.9. It is apparently possible to establish a sharper bound [9, 14] than the one given above in Lemma 3.21, the improved bound having the form

$$C_2 = 1 + 2\omega^2\rho(\Theta)^2.$$

This result is stated and used in several recent papers on domain decomposition, e.g., in [14], but the proof of the result has apparently not been published. A proof of a similar result is established for some related nonsymmetric problems in [9].

3.4 Main results of the theory

The main theory may be summarized in the following way. We are interested in norm and condition number bounds of the product and sum operators:

$$E = (I - T_J)(I - T_{J-1}) \cdots (I - T_0), \tag{18}$$

$$P = T_0 + T_1 + \cdots + T_J. \tag{19}$$

The necessary assumptions for the theory are as follows.

Assumption 3.8 *(Operator norms) The operators $T_k \in \mathbf{L}(\mathcal{H}, \mathcal{H})$ are A-self-adjoint, A-non-negative, and*

$$\rho(T_k) = \|T_k\|_A \leq \omega < 2, \quad k = 0, \ldots, J.$$

Assumption 3.9 *(Splitting constant) There exists $C_0 > 0$ such that*

$$\|v\|_A^2 \leq C_0 \sum_{k=0}^{J} (AT_k v, v), \quad \forall v \in \mathcal{H}.$$

Definition 3.2 *(Interaction matrices) Let Ξ be the strictly upper-triangular part of the interaction matrix $\Theta \in \mathbf{L}(\mathbb{R}^J, \mathbb{R}^J)$, which is defined to have as entries Θ_{ij} the smallest constants satisfying:*

$$|(AT_iu, T_jv)| \leq \Theta_{ij}(AT_iu, T_iu)^{1/2}(AT_jv, T_jv)^{1/2}, \quad 1 \leq i, j \leq J.$$

The main theorems are as follows.

Theorem 3.23 *(Product operator) Under Assumptions 3.8 and 3.9, the product operator (18) satisfies:*

$$||E||_A^2 \leq 1 - \frac{2-\omega}{C_0(6 + 6\omega^2\rho(\Theta)^2)}.$$

Proof. Assumptions 3.8 and 3.9 are clearly equivalent to Assumptions 3.1 and 3.2, and by Lemma 3.21 we know that Assumption 3.4 must hold with $C_2 = [6 + 6\omega^2\rho(\Theta)^2]$. The theorem then follows by application of Corollary 3.6. □

Theorem 3.24 *(Sum operator) Under Assumptions 3.8 and 3.9, the sum operator (19) satisfies:*

$$\kappa_A(P) \leq C_0\omega(\rho(\Theta) + 1).$$

Proof. Assumptions 3.8 and 3.9 are clearly equivalent to Assumptions 3.1 and 3.2, and by Lemma 3.22 we know that Assumption 3.5 must hold with $C_3 = \omega(1 + \rho(\Theta))$. The theorem then follows by application of Theorem 3.7. □

For the case when there is *not* a global operator T_0 present, set $T_0 \equiv 0$ in the above definitions and assumptions. Note that this implies that all $k = 0$ terms in the assumptions and definitions are ignored. The main theorems are now modified as follows.

Theorem 3.25 *(Product operator) If $T_0 \equiv 0$, then under Assumptions 3.8 and 3.9, the product operator (18) satisfies:*

$$||E||_A^2 \leq 1 - \frac{2-\omega}{C_0(1 + \omega||\Xi||_2)^2}.$$

Proof. Assumptions 3.8 and 3.9 are clearly equivalent to Assumptions 3.1 and 3.2, and by Lemma 3.17 we know that Assumption 3.4 must hold with $C_2 = (1 + \omega||\Xi||)^2$. The theorem then follows by application of Corollary 3.6. □

Theorem 3.26 *(Sum operator) If $T_0 \equiv 0$, then under Assumptions 3.8 and 3.9, the sum operator (19) satisfies:*

$$\kappa_A(P) \leq C_0\omega\rho(\Theta).$$

Proof. Assumptions 3.8 and 3.9 are clearly equivalent to Assumptions 3.1 and 3.2, and by Lemma 3.18 we know that Assumption 3.5 must hold with $C_3 = \omega\rho(\Theta)$. The theorem then follows by application of Theorem 3.7. □

Remark 3.10. We see that the product and sum operator theory now rests completely on the estimation of the constant C_0 in Assumption 3.9 and the bounds on the interaction matrices. (The bound involving ω in Assumption 3.8 always holds for any reasonable method based on product and sum operators.) We will further reduce the estimate of C_0 to simply the estimate of a "splitting" constant, depending on the particular splitting of the main space $\mathcal{H}$ into subspaces $\mathcal{H}_k$, and to an estimate of the effectiveness of the approximate solver in the subspaces.

Remark 3.11. Note that if we cannot estimate $||\Xi||_2$ or $\rho(\Theta)$, then we can still use the above theory since we have worst case estimates from Lemmas 3.15 and 3.16, namely:

$$||\Xi||_2 \leq \sqrt{J(J-1)/2} < J, \qquad \rho(\Theta) \leq J.$$

In the case of the nested spaces in multigrid methods, it may be possible to analyze $||\Xi||_2$ through the use of *strengthened Cauchy-Schwarz inequalities*, showing in fact that $||\Xi||_2 = O(1)$. In the case of domain decomposition methods, it will *always* be possible to show that $||\Xi||_2 = O(1)$ and $\rho(\Theta) = O(1)$, due to the local nature of the domain decomposition projection operators.

4. Abstract Schwarz theory

In this section, we consider abstract Schwarz methods based on subspaces, and apply the general product and sum operator theory to these methods. The resulting theory, which is a variation of that presented in [39] and [14], rests on the notion of a stable subspace splitting of the original Hilbert space (cf. [31, 32]). Although our derivation here is presented in a somewhat different, algebraic language, many of the intermediate results we use have appeared previously in the literature in other forms (we provide references at the appropriate points). In contrast to earlier approaches, we develop the entire theory employing general prolongation and restriction operators; the use of inclusion and projection as prolongation and restriction are represented in our approach as a special case.

4.1 The Schwarz methods

Consider now a Hilbert space $\mathcal{H}$, equipped with an inner-product $(\cdot,\cdot)$ inducing a norm $\|\cdot\| = (\cdot,\cdot)^{1/2}$. Let there be given an SPD operator $A \in \mathbf{L}(\mathcal{H},\mathcal{H})$ defining another inner-product on $\mathcal{H}$, which we denote as $(\cdot,\cdot)_A = (A\cdot,\cdot)$. This second inner-product also induces a norm $\|\cdot\|_A = (\cdot,\cdot)_A^{1/2}$. We are also given an associated set of spaces

$$\mathcal{H}_1, \mathcal{H}_2, \ldots, \mathcal{H}_J, \qquad \dim(\mathcal{H}_k) \le \dim(\mathcal{H}), \qquad I_k\mathcal{H}_k \subseteq \mathcal{H}, \qquad \mathcal{H} = \sum_{k=1}^{J} I_k\mathcal{H}_k,$$

for some operators $I_k : \mathcal{H}_k \mapsto \mathcal{H}$, where we assume that $\mathrm{null}(I_k) = \{0\}$. This defines a splitting of $\mathcal{H}$ into the subspaces $I_k\mathcal{H}_k$, although the spaces $\mathcal{H}_k$ alone may not relate to the largest space $\mathcal{H}$ in any natural way without the operator I_k. No requirements are made on the associated spaces $\mathcal{H}_k$ beyond the above, so that they are not necessarily nested, disjoint, or overlapping.

Associated with each space $\mathcal{H}_k$ is an inner-product $(\cdot,\cdot)_k$ inducing a norm $\|\cdot\|_k = (\cdot,\cdot)_k^{1/2}$, and an SPD operator $A_k \in \mathbf{L}(\mathcal{H}_k,\mathcal{H}_k)$, defining a second inner-product $(\cdot,\cdot)_{A_k} = (A_k\cdot,\cdot)_k$ and norm $\|\cdot\|_{A_k} = (\cdot,\cdot)_{A_k}^{1/2}$. The spaces $\mathcal{H}_k$ are related to the finest space $\mathcal{H}$ through the *prolongation* I_k defined above, and also through the *restriction* operator, defined as the adjoint of I_k relating the inner-products in $\mathcal{H}$ and $\mathcal{H}_k$:

$$(I_kv_k, v) = (v_k, I_k^Tv)_k, \qquad I_k^T : \mathcal{H} \mapsto \mathcal{H}_k.$$

It will always be completely clear from the arguments of the inner-product (or norm) which particular inner-product (or norm) is implied; i.e., if the arguments lie in $\mathcal{H}$ then either $(\cdot,\cdot)$ or $(A\cdot,\cdot)$ is to be used, whereas if the arguments lie in $\mathcal{H}_k$, then either $(\cdot,\cdot)_k$ or $(A_k\cdot,\cdot)_k$ is to be used. Therefore, we will leave off the implied subscript k from the inner-products and norms in all of the following discussions, without danger of confusion. Finally, we assume the existence of SPD linear operators $R_k \in \mathbf{L}(\mathcal{H}_k,\mathcal{H}_k)$, such that $R_k \approx A_k^{-1}$.

Definition 4.1 *The operator $A_k \in \mathbf{L}(\mathcal{H}_k,\mathcal{H}_k)$ is called variational with respect to $A \in \mathbf{L}(\mathcal{H},\mathcal{H})$ if, for a fixed operator $I_k \in \mathbf{L}(\mathcal{H}_k,\mathcal{H})$, it holds that:*

$$A_k = I_k^TAI_k.$$

If the operators A_k are each variational with A, then the operator A_k in space $\mathcal{H}_k$ is in some sense a representation of the operator A in the space $\mathcal{H}_k$. For example, in a multigrid or domain decomposition algorithm, the operator I_k^T may correspond to an orthogonal projector, and I_k to the natural inclusion of a subspace into the whole space.

Regarding the operators R_k, a natural condition to impose is that they correspond to some convergent linear methods in the associated spaces, the necessary and sufficient condition for which would be (by Theorem 2.7):

$$\rho(I - R_kA_k) = \|I - R_kA_k\|_A < 1, \qquad k = 1,\cdots,J.$$

Note that if $R_k = A_k^{-1}$, this is trivially satisfied. More generally, $R_k \approx A_k^{-1}$, corresponding to some classical linear smoothing method (in the case of multigrid), or some other linear solver.

An abstract multiplicative Schwarz method, employing associated space corrections in the spaces $\mathcal{H}_k$, has the form:

Algorithm 4.1 *(Abstract Multiplicative Schwarz Method – Implementation Form)*

$$u^{n+1} = MS(u^n, f)$$

where the operation $u^{\mathrm{NEW}} = MS(u^{\mathrm{OLD}}, f)$ *is defined as:*

$$\begin{array}{l} \textit{Do } k = 1, \ldots, J \\ \quad r_k = I_k^T(f - Au^{\mathrm{OLD}}) \\ \quad e_k = R_k r_k \\ \quad u^{\mathrm{NEW}} = u^{\mathrm{OLD}} + I_k e_k \\ \quad u^{\mathrm{OLD}} = u^{\mathrm{NEW}} \\ \textit{End do.} \end{array}$$

Note that the first step through the loop in $MS(\cdot,\cdot)$ gives:

$$u^{\mathrm{NEW}} = u^{\mathrm{OLD}} + I_1 e_1 = u^{\mathrm{OLD}} + I_1 R_1 I_1^T (f - Au^{\mathrm{OLD}}) = (I - I_1 R_1 I_1^T A) u^{\mathrm{OLD}} + I_1 R_1 I_1^T f.$$

Continuing in this fashion, and by defining $T_k = I_k R_k I_k^T A$, we see that after the full loop in $MS(\cdot,\cdot)$ the solution transforms according to:

$$u^{n+1} = (I - T_J)(I - T_{J-1}) \cdots (I - T_1) u^n + Bf,$$

where B is a quite complicated combination of the operators R_k, I_k, I_k^T, and A. By defining $E_k = (I - T_k)(I - T_{k-1}) \cdots (I - T_1)$, we see that $E_k = (I - T_k)E_{k-1}$. Therefore, since $E_{k-1} = I - B_{k-1}A$ for some (implicitly defined) B_{k-1}, we can identify the operators B_k through the recursion $E_k = I - B_k A = (I - T_k)E_{k-1}$, giving

$$B_k A = I - (I - T_k)E_{k-1} = I - (I - B_{k-1}A) + T_k(I - B_{k-1}A) = B_{k-1}A + T_k - T_k B_{k-1} A$$

$$= B_{k-1}A + I_k R_k I_k^T A - I_k R_k I_k^T A B_{k-1} A = \left[B_{k-1} + I_k R_k I_k^T - I_k R_k I_k^T A B_{k-1}\right] A,$$

so that $B_k = B_{k-1} + I_k R_k I_k^T - I_k R_k I_k^T A B_{k-1}$. But this means the above algorithm is equivalent to:

Algorithm 4.2 *(Abstract Multiplicative Schwarz Method – Operator Form)*

$$u^{n+1} = u^n + B(f - Au^n) = (I - BA)u^n + Bf$$

where the multiplicative Schwarz error propagator E *is defined by:*

$$E = I - BA = (I - T_J)(I - T_{J-1}) \cdots (I - T_1), \qquad T_k = I_k R_k I_k^T A, \quad k = 1, \ldots, J.$$

The operator $B \equiv B_J$ *is defined implicitly, and obeys the recursion:*

$$B_1 = I_1 R_1 I_1^T, \quad B_k = B_{k-1} + I_k R_k I_k^T - I_k R_k I_k^T A B_{k-1}, \quad k = 2, \ldots, J.$$

An abstract additive Schwarz method, employing corrections in the spaces $\mathcal{H}_k$, has the form:

Algorithm 4.3 *(Abstract Additive Schwarz Method – Implementation Form)*

$$u^{n+1} = MS(u^n, f)$$

where the operation $u^{\mathrm{NEW}} = MS(u^{\mathrm{OLD}}, f)$ *is defined as:*

$$\begin{array}{l} r = f - Au^{\mathrm{OLD}} \\ \textit{Do } k = 1, \ldots, J \\ \quad r_k = I_k^T r \\ \quad e_k = R_k r_k \\ \quad u^{\mathrm{NEW}} = u^{\mathrm{OLD}} + I_k e_k \\ \textit{End do.} \end{array}$$

Since each loop iteration depends only on the original approximation u^{OLD}, we see that the full correction to the solution can be written as the sum:

$$u^{n+1} = u^n + B(f - Au^n) = u^n + \sum_{k=1}^{J} I_k R_k I_k^T (f - Au^n),$$

where the preconditioner B has the form $B = \sum_{k=1}^{J} I_k R_k I_k^T$, and the error propagator is $E = I - BA$. Therefore, the above algorithm is equivalent to:

Algorithm 4.4 *(Abstract Additive Schwarz Method - Operator Form)*

$$u^{n+1} = u^n + B(f - Au^n) = (I - BA)u^n + Bf$$

where the additive Schwarz error propagator E is defined by:

$$E = I - BA = I - \sum_{k=1}^{J} T_k, \qquad T_k = I_k R_k I_k^T A, \quad k = 1, \ldots, J.$$

The operator B is defined explicitly as $B = \sum_{k=1}^{J} I_k R_k I_k^T$.

4.2 Subspace splitting theory

We now consider the framework of §4.1, employing the abstract results of §3.4. First, we prove some simple results about projectors, and the relationships between the operators R_k on the spaces $\mathcal{H}_k$ and the resulting operators $T_k = I_k R_k I_k^T A$ on the space $\mathcal{H}$. We then consider the "splitting" of the space $\mathcal{H}$ into subspaces $I_k \mathcal{H}_k$, and the verification of the assumptions required to apply the abstract theory of §3.4 is reduced to deriving an estimate of the "splitting constant".

Recall that an orthogonal projector is an operator $P \in \mathbf{L}(\mathcal{H}, \mathcal{H})$ having a closed subspace $\mathcal{V} \subseteq \mathcal{H}$ as its range (on which P acts as the identity), and having the orthogonal complement of $\mathcal{V}$, denoted as $\mathcal{V}^\perp \subseteq \mathcal{H}$, as its null space. By this definition, the operator $I - P$ is also clearly a projector, but having the subspace $\mathcal{V}^\perp$ as range and $\mathcal{V}$ as null space. In other words, a projector P splits a Hilbert space $\mathcal{H}$ into a direct sum of a closed subspace and its orthogonal complement as follows:

$$\mathcal{H} = \mathcal{V} \oplus \mathcal{V}^\perp = P\mathcal{H} \oplus (I - P)\mathcal{H}.$$

The following lemma gives a useful characterization of a projection operator; note that this characterization is often used as an equivalent alternative definition of a projection operator.

Lemma 4.1 *Let $A \in \mathbf{L}(\mathcal{H}, \mathcal{H})$ be SPD. Then the operator $P \in \mathbf{L}(\mathcal{H}, \mathcal{H})$ is an A-orthogonal projector if and only if P is A-self-adjoint and idempotent ($P^2 = P$).*

Proof. See [24], Theorem 9.5-1, page 481. □

Lemma 4.2 *Assume* $\dim(\mathcal{H}_k) \le \dim(\mathcal{H})$, $I_k : \mathcal{H}_k \mapsto \mathcal{H}$, $\text{null}(I_k) = \{0\}$, *and that A is SPD. Then*

$$Q_k = I_k (I_k^T I_k)^{-1} I_k^T, \qquad P_k = I_k (I_k^T A I_k)^{-1} I_k^T A,$$

are the unique orthogonal and A-orthogonal projectors onto $I_k \mathcal{H}_k$.

Proof. By assuming that $\text{null}(I_k) = \{0\}$, we guarantee that both $\text{null}(I_k^T I_k) = \{0\}$ and $\text{null}(I_k^T A I_k) = \{0\}$, so that both Q_k and P_k are well-defined. It is easily verified that Q_k is self-adjoint and P_k is A-self-adjoint, and it is immediate that $Q_k^2 = Q_k$ and that $P_k^2 = P_k$. Clearly, $Q_k : \mathcal{H} \mapsto I_k \mathcal{H}_k$, and $P_k : \mathcal{H} \mapsto I_k \mathcal{H}_k$, so that by Lemma 4.1 these operators are orthogonal and A-orthogonal projectors onto $I_k \mathcal{H}_k$. All that remains is to show that these operators are unique. By definition, a projector onto a subspace $I_k \mathcal{H}_k$ acts as the identity on $I_k \mathcal{H}_k$, and as the zero operator on $(I_k \mathcal{H}_k)^\perp$. Therefore, any two projectors P_k and $\tilde{P}_k$ onto $I_k \mathcal{H}_k$ must act identically on the entire space $\mathcal{H} = I_k \mathcal{H}_k \oplus (I_k \mathcal{H}_k)^\perp$, and therefore $P_k = \tilde{P}_k$. Similarly, Q_k is unique. □

We now make the following natural assumption regarding the operators $R_k \approx A_k^{-1}$.

Assumption 4.1 *The operators* $R_k \in \mathbf{L}(\mathcal{H}_k, \mathcal{H}_k)$ *are SPD. Further, there exists a subspace* $\mathcal{V}_k \subseteq \mathcal{H}_k$*, and parameters* $0 < \omega_0 \leq \omega_1 < 2$*, such that*

$$\begin{array}{llll} (a) & \omega_0(A_k v_k, v_k) \leq (A_k R_k A_k v_k, v_k), & \forall v_k \in \mathcal{V}_k \subseteq \mathcal{H}_k, & k = 1, \ldots, J, \\ (b) & (A_k R_k A_k v_k, v_k) \leq \omega_1(A_k v_k, v_k), & \forall v_k \in \mathcal{H}_k, & k = 1, \ldots, J. \end{array}$$

This implies that on the subspace $\mathcal{V}_k \subseteq \mathcal{H}_k$, it holds that $0 < \omega_0 \leq \lambda_i(R_k A_k)$, $k = 1, \ldots, J$, whereas on the entire space $\mathcal{H}_k$, it holds that $\lambda_i(R_k A_k) \leq \omega_1 < 2$, $k = 1, \ldots, J$.

There are several consequences of the above assumption which will be useful later.

Lemma 4.3 *Assumption 4.1(b) implies that* $0 < \lambda_i(R_k A_k) \leq \omega_1$*, and* $\rho(I - R_k A_k) = \|I - R_k A_k\|_{A_k} < 1$.

Proof. Since R and A are SPD by assumption, we have by Lemma 2.6 that RA is A-SPD. By Assumption 4.1(b), the Rayleigh quotients are bounded above by ω_1, so that

$$0 < \lambda_i(RA) \leq \omega_1.$$

Thus,

$$\rho(I - RA) = \max_i |\lambda_i(I - RA)| = \max_i |1 - \lambda_i(RA)|.$$

Clearly then $\rho(I - RA) < 1$ since $0 < \omega_1 < 2$. $\Box$

Lemma 4.4 *Assumption 4.1(b) implies that* $(A_k v_k, v_k) \leq \omega_1(R_k^{-1} v_k, v_k), \forall v_k \in \mathcal{H}_k$.

Proof. We drop the subscripts for ease of exposition. By Assumption 4.1(b), $(ARAv, v) \leq \omega_1(Av, v)$, so that ω_1 bounds the Raleigh quotients generated by RA. Since RA is similar to $R^{1/2}AR^{1/2}$, we must also have that

$$(R^{1/2}AR^{1/2}v, v) \leq \omega_1(v, v).$$

But this implies

$$(AR^{1/2}v, R^{1/2}v) \leq \omega_1(R^{-1}R^{1/2}v, R^{1/2}v),$$

or $(Aw, w) \leq \omega_1(R^{-1}w, w), \forall w \in \mathcal{H}$. $\Box$

Lemma 4.5 *Assumption 4.1(b) implies that* $T_k = I_k R_k I_k^T A$ *is A-self-adjoint and A-non-negative, and*

$$\rho(T_k) = \|T_k\|_A \leq \omega_1 < 2.$$

Proof. That $T_k = I_k R_k I_k^T A$ is A-self-adjoint and A-non-negative follows immediately from the symmetry of R_k and A_k. To show the last result, we employ Lemma 4.4 to obtain

$$(AT_k v, T_k v) = (AI_k R_k I_k^T Av, I_k R_k I_k^T Av) = (I_k^T A I_k R_k I_k^T Av, R_k I_k^T Av)$$

$$= (A_k R_k I_k^T Av, R_k I_k^T Av) \leq \omega_1(R_k^{-1} R_k I_k^T Av, R_k I_k^T Av) = \omega_1(I_k^T Av, R_k I_k^T Av)$$

$$= \omega_1(AI_k R_k I_k^T Av, v) = \omega_1(AT_k v, v).$$

Now, from the Schwarz inequality, we have

$$(AT_k v, T_k v) \leq \omega_1(AT_k v, v) \leq \omega_1(AT_k v, T_k v)^{1/2}(Av, v)^{1/2},$$

or that

$$(AT_k v, T_k v)^{1/2} \leq \omega_1(Av, v)^{1/2},$$

which implies that $\|T_k\|_A \leq \omega_1 < 2$. $\Box$

The key idea in all of the following theory involves the splitting of the original Hilbert space $\mathcal{H}$ into a collection of subspaces $I_k \mathcal{V}_k \subseteq I_k \mathcal{H}_k \subseteq \mathcal{H}$. It will be important for the splitting to be *stable* in a certain sense, which we state as the following assumption.

Assumption 4.2 *Given any $v \in \mathcal{H} = \sum_{k=1}^J I_k\mathcal{H}_k$, $I_k\mathcal{H}_k \subseteq \mathcal{H}$, there exists subspaces $I_k\mathcal{V}_k \subseteq I_k\mathcal{H}_k \subseteq \mathcal{H} = \sum_{k=1}^J I_k\mathcal{V}_k$, and a particular splitting $v = \sum_{k=1}^J I_k v_k$, $v_k \in \mathcal{V}_k$, such that*

$$\sum_{k=1}^J \|I_k v_k\|_A^2 \leq S_0 \|v\|_A^2,$$

for some splitting constant $S_0 > 0$.

The following key lemma (in the case of inclusion and projection as prolongation and restriction) is sometimes referred to as *Lions' Lemma* [25], although the multiple-subspace case is essentially due to Widlund [36].

Lemma 4.6 *Under Assumption 4.2 it holds that*

$$\left(\frac{1}{S_0}\right)\|v\|_A^2 \leq \sum_{k=1}^J (AP_k v, v), \quad \forall v \in \mathcal{H}.$$

Proof. Given any $v \in \mathcal{H}$, we employ the splitting of Assumption 4.2 to obtain

$$\|v\|_A^2 = \sum_{k=1}^J (Av, I_k v_k) = \sum_{k=1}^J (I_k^T Av, v_k) = \sum_{k=1}^J (I_k^T A(I_k(I_k^T AI_k)^{-1}I_k^T A)v, v_k) = \sum_{k=1}^J (AP_k v, I_k v_k).$$

Now, let $\tilde{P}_k = (I_k^T AI_k)^{-1}I_k^T A$, so that $P_k = I_k\tilde{P}_k$. Then

$$\|v\|_A^2 = \sum_{k=1}^J (I_k^T AI_k\tilde{P}_k v, v_k) = \sum_{k=1}^J (A_k\tilde{P}_k v, v_k) \leq \sum_{k=1}^J (A_k v_k, v_k)^{1/2}(A_k\tilde{P}_k v, \tilde{P}_k v)^{1/2}$$

$$\leq \left(\sum_{k=1}^J (A_k v_k, v_k)\right)^{1/2}\left(\sum_{k=1}^J (A_k\tilde{P}_k v, \tilde{P}_k v)\right)^{1/2} = \left(\sum_{k=1}^J (AI_k v_k, I_k v_k)\right)^{1/2}\left(\sum_{k=1}^J (A_k\tilde{P}_k v, \tilde{P}_k v)\right)^{1/2}$$

$$= \left(\sum_{k=1}^J \|I_k v_k\|_A^2\right)^{1/2}\left(\sum_{k=1}^J (A_k\tilde{P}_k v, \tilde{P}_k v)\right)^{1/2} \leq S_0^{1/2}\|v\|_A\left(\sum_{k=1}^J (AI_k\tilde{P}_k v, I_k\tilde{P}_k v)\right)^{1/2}$$

$$= S_0^{1/2}\|v\|_A\left(\sum_{k=1}^J (AP_k v, P_k v)\right)^{1/2}, \quad \forall v \in \mathcal{H}.$$

Since $(AP_k v, P_k v) = (AP_k v, v)$, dividing the above by $\|v\|_A$ and squaring yields the result. □

The next intermediate result will be useful in the case that the subspace solver R_k is effective on only the part of the subspace $\mathcal{H}_k$, namely $\mathcal{V}_k \subseteq \mathcal{H}_k$.

Lemma 4.7 *Under Assumptions 4.1(a) and 4.2 (for the same subspaces $I_k\mathcal{V}_k \subseteq I_k\mathcal{H}_k$) it holds that*

$$\sum_{k=1}^J (R_k^{-1}v_k, v_k) \leq \left(\frac{S_0}{\omega_0}\right)\|v\|_A^2, \quad \forall v = \sum_{k=1}^J I_k v_k \in \mathcal{H}, \ v_k \in \mathcal{V}_k \subseteq \mathcal{H}_k.$$

Proof. With $v = \sum_{k=1}^J I_k v_k$, where we employ the splitting in Assumption 4.2, we have

$$\sum_{k=1}^J (R_k^{-1}v_k, v_k) = \sum_{k=1}^J (A_k A_k^{-1}R_k^{-1}v_k, v_k) = \sum_{k=1}^J (A_k v_k, v_k)\frac{(A_k A_k^{-1}R_k^{-1}v_k, v_k)}{(A_k v_k, v_k)}$$

$$\leq \sum_{k=1}^J (A_k v_k, v_k)\max_{v_k \neq 0}\frac{(A_k A_k^{-1}R_k^{-1}v_k, v_k)}{(A_k v_k, v_k)} \leq \sum_{k=1}^J \omega_0^{-1}(A_k v_k, v_k)$$

$$= \sum_{k=1}^J \omega_0^{-1}(AI_k v_k, I_k v_k) = \sum_{k=1}^J \omega_0^{-1}\|I_k v_k\|_A^2 \leq \left(\frac{S_0}{\omega_0}\right)\|v\|_A^2,$$

which proves the lemma. □

The following lemma relates the constant appearing in the "splitting" Assumption 3.9 of the product and sum operator theory to the subspace splitting constant appearing in Assumption 4.2 above.

Lemma 4.8 *Under Assumptions 4.1(a) and 4.2 (for the same subspaces $I_k\mathcal{V}_k \subseteq I_k\mathcal{H}_k$) it holds that*

$$\|v\|_A^2 \leq \left(\frac{S_0}{\omega_0}\right)\sum_{k=1}^{J}(AT_kv, v), \quad \forall v \in \mathcal{H}.$$

Proof. Given any $v \in \mathcal{H}$, we begin with the splitting in Assumption 4.2 as follows

$$\|v\|_A^2 = (Av, v) = \sum_{k=1}^{J}(Av, I_kv_k) = \sum_{k=1}^{J}(I_k^TA(I_k(I_k^TAI_k)^{-1}I_k^TA)v, v_k) = \sum_{k=1}^{J}(I_k^TAI_k\tilde{P}_kv, v_k),$$

where $P_k = I_k\tilde{P}_k = I_k(I_k^TAI_k)^{-1}I_k^TA$. We employ now the Cauchy-Schwarz inequality in the R_k inner-product, yielding

$$\|v\|_A^2 = \sum_{k=1}^{J}(A_k\tilde{P}_kv, v_k) = \sum_{k=1}^{J}(R_kA_k\tilde{P}_kv, R_k^{-1}v_k)$$

$$\leq \left(\sum_{k=1}^{J}(R_k^{-1}v_k, v_k)\right)^{1/2}\left(\sum_{k=1}^{J}(R_kA_k\tilde{P}_kv, A_k\tilde{P}_kv)\right)^{1/2} \leq \left(\frac{S_0}{\omega_0}\right)^{1/2}\|v\|_A\left(\sum_{k=1}^{J}(AT_kv, v)\right)^{1/2},$$

where we have employed for the last inequality both Lemma 4.7, and the fact that $A_k\tilde{P} = I_k^TA$, which implies that $(R_kA_k\tilde{P}_kv, A_k\tilde{P}_kv) = (R_kI_k^TAv, I_k^TAv) = (AT_kv, v)$. Dividing the inequality above by $\|v\|_A$ and squaring yields the lemma. □

In order to employ the product and sum theory, we must quantify the interaction of the operators T_k. As the T_k involve corrections in subspaces, we will see that the operator interaction properties will be determined completely by the interaction of the subspaces. Therefore, we introduce the following notions to quantify the interaction of the subspaces involved.

Definition 4.2 *(Strong interaction matrix) The interaction matrix $\Theta \in \mathbf{L}(\mathbb{R}^J, \mathbb{R}^J)$ is defined to have as entries Θ_{ij} the smallest constants satisfying:*

$$|(AI_iu_i, I_jv_j)| \leq \Theta_{ij}(AI_iu_i, I_iu_i)^{1/2}(AI_jv_j, I_jv_j)^{1/2}, \ 1 \leq i,j \leq J, \ u_i \in \mathcal{H}_i, v_j \in \mathcal{H}_j.$$

Definition 4.3 *(Weak interaction matrix) The strictly upper-triangular interaction matrix $\Xi \in \mathbf{L}(\mathbb{R}^J, \mathbb{R}^J)$ is defined to have as entries Ξ_{ij} the smallest constants satisfying:*

$$|(AI_iu_i, I_jv_j)| \leq \Xi_{ij}(AI_iu_i, I_iu_i)^{1/2}(AI_jv_j, I_jv_j)^{1/2}, \ 1 \leq i < j \leq J, \ u_i \in \mathcal{H}_i, v_j \in \mathcal{V}_j \subseteq \mathcal{H}_j.$$

The following lemma relates the interaction properties of the subspaces specified by the strong interaction matrix to the interaction properties of the associated subspace correction operators $T_k = I_kR_kI_k^TA$.

Lemma 4.9 *For the strong interaction matrix Θ given in Definition 4.2, it holds that*

$$|(AT_iu, T_jv)| \leq \Theta_{ij}(AT_iu, T_iu)^{1/2}(AT_jv, T_jv)^{1/2}, \quad 1 \leq i,j \leq J, \quad \forall u, v \in \mathcal{H}.$$

Proof. Since $T_ku = I_kR_kI_k^TAu = I_ku_k$, where $u_k = R_kI_k^TAu$, the lemma follows simply from the definition of Θ in Definition 4.2 above. □

Remark 4.12. Note that the weak interaction matrix in Definition 4.3 involves a subspace $\mathcal{V}_k \subseteq \mathcal{H}_k$, which will be necessary in the analysis of multigrid-like methods. Unfortunately, this will preclude the simple application of the product operator theory of the previous sections. In particular, we cannot estimate the constant C_2 required for the use of Corollary 3.6, because we cannot show Lemma 3.15 for arbitrary T_k. In order to prove Lemma 3.15, we would need to employ the upper-triangular portion of the strong interaction matrix Θ in Definition 4.2, involving the entire space $\mathcal{H}_k$, which is now different from the upper-triangular weak interaction matrix Ξ (employing only the subspace $\mathcal{V}_k$) defined as above in Definition 4.3. There was no such distinction between the weak and strong interaction matrices in the product and sum operator theory of the previous sections; the weak interaction matrix was defined simply as the strictly upper-triangular portion of the strong interaction matrix.

We can, however, employ the original Theorem 3.5 by attempting to estimate C_1 directly, rather than employing Corollary 3.6 and estimating C_1 indirectly through C_0 and C_2. The following result will allow us to do this, and still employ the weak interaction property above in Definition 4.3.

Lemma 4.10 *Under Assumptions 4.1 and 4.2 (for the same subspaces $I_k\mathcal{V}_k \subseteq I_k\mathcal{H}_k$), it holds that*

$$\|v\|_A^2 \leq \left(\frac{S_0}{\omega_0}\right)[1+\omega_1\|\Xi\|_2]^2 \sum_{k=1}^J (AT_kE_{k-1}v, E_{k-1}v), \qquad \forall v \in \mathcal{H},$$

where Ξ is the weak interaction matrix of Definition 4.3.

Proof. We employ the splitting of Assumption 4.2, namely $v = \sum_{k=1}^J I_k v_k$, $v_k \in \mathcal{V}_k \subseteq \mathcal{H}_k$, as follows:

$$\|v\|_A^2 = \sum_{k=1}^J (Av, I_kv_k) = \sum_{k=1}^J (AE_{k-1}v, I_kv_k) + \sum_{k=1}^J (A[I-E_{k-1}]v, I_kv_k)$$

$$= \sum_{k=1}^J (AE_{k-1}v, I_kv_k) + \sum_{k=1}^J\sum_{i=1}^{k-1} (AT_iE_{i-1}v, I_kv_k) = \mathbf{S}_1 + \mathbf{S}_2.$$

We now estimate $\mathbf{S}_1$ and $\mathbf{S}_2$ separately. For the first term, we have:

$$\mathbf{S}_1 = \sum_{k=1}^J (AE_{k-1}v, I_kv_k) = \sum_{k=1}^J (I_k^TAE_{k-1}v, v_k) = \sum_{k=1}^J (R_kI_k^TAE_{k-1}v, R_k^{-1}v_k)$$

$$\leq \sum_{k=1}^J (R_kI_k^TAE_{k-1}v, I_k^TAE_{k-1}v)^{1/2}(R_k^{-1}v_k, v_k)^{1/2} = \sum_{k=1}^J (AT_kE_{k-1}v, E_{k-1}v)^{1/2}(R_k^{-1}v_k, v_k)^{1/2}$$

$$\leq \left(\sum_{k=1}^J (AT_kE_{k-1}v, E_{k-1}v)\right)^{1/2} \left(\sum_{k=1}^J (R_k^{-1}v_k, v_k)\right)^{1/2}.$$

where we have employed the Cauchy-Schwarz inequality in the R_k inner-product for the first inequality and in $\mathbb{R}^J$ for the second. Employing now Lemma 4.7 (requiring Assumptions 4.1 and 4.2) to bound the right-most term, we have

$$\mathbf{S}_1 \leq \left(\frac{S_0}{\omega_0}\right)^{1/2} \|v\|_A \left(\sum_{k=1}^J (AT_kE_{k-1}v, E_{k-1}v)\right)^{1/2}.$$

We now bound the term $\mathbf{S}_2$, employing the weak interaction matrix given in Definition 4.3 above, as follows:

$$\mathbf{S}_2 = \sum_{k=1}^J\sum_{i=1}^{k-1} (AT_iE_{i-1}v, I_kv_k) = \sum_{k=1}^J\sum_{i=1}^{k-1} (AI_i[R_iI_i^TAE_{i-1}v], I_kv_k)$$

$$\leq \sum_{k=1}^J\sum_{i=1}^J \Xi_{ik}\|I_i[R_iI_i^TAE_{i-1}v]\|_A\|I_kv_k\|_A = \sum_{k=1}^J\sum_{i=1}^J \Xi_{ik}\|T_iE_{i-1}v\|_A\|I_kv_k\|_A = (\Xi\mathbf{x}, \mathbf{y})_2,$$

where $\mathbf{x}, \mathbf{y} \in \mathbb{R}^J$, $x_k = \|I_kv_k\|_A$, $y_i = \|T_iE_{i-1}v\|_A$, and $(\cdot,\cdot)_2$ is the usual Euclidean inner-product in $\mathbb{R}^J$. Now, we have that

$$\mathbf{S}_2 \leq (\Xi\mathbf{x}, \mathbf{y})_2 \leq \|\Xi\|_2\|\mathbf{x}\|_2\|\mathbf{y}\|_2 = \|\Xi\|_2 \left(\sum_{k=1}^J (AT_kE_{k-1}v, T_kE_{k-1}v)\right)^{1/2} \left(\sum_{k=1}^J (AI_kv_k, I_kv_k)\right)^{1/2}$$

$$\leq \omega_1^{1/2}\|\Xi\|_2 \left(\sum_{k=1}^J (AT_kE_{k-1}v, E_{k-1}v)\right)^{1/2} \left(\sum_{k=1}^J (A_kv_k, v_k)\right)^{1/2},$$

since $A_k = I_k^T A I_k$, and by Lemma 3.2, which may be applied because of Lemma 4.5. By Lemma 4.4, we have $(A_k v_k, v_k) \leq \omega_1 (R_k^{-1} v_k, v_k)$, and employing this result along with Lemma 4.7 gives

$$\mathbf{S}_2 \leq \omega_1 \|\Xi\|_2 \left(\sum_{k=1}^{J} (AT_k E_{k-1} v, E_{k-1} v) \right)^{1/2} \left(\sum_{k=1}^{J} (R_k^{-1} v_k, v_k) \right)^{1/2}$$

$$\leq \left(\frac{S_0}{\omega_0} \right)^{1/2} \|v\|_A \omega_1 \|\Xi\|_2 \left(\sum_{k=1}^{J} (AT_k E_{k-1} v, E_{k-1} v) \right)^{1/2}.$$

Combining the two results gives finally

$$\|v\|_A^2 \leq \mathbf{S}_1 + \mathbf{S}_2 \leq \left(\frac{S_0}{\omega_0} \right)^{1/2} \|v\|_A \left[1 + \omega_1 \|\Xi\|_2 \right] \left(\sum_{k=1}^{J} (AT_k E_{k-1} v, E_{k-1} v) \right)^{1/2}, \qquad \forall v \in \mathcal{H}.$$

Dividing by $\|v\|_A$ and squaring yieldings the result. □

Remark 4.13. Although our language and notation is quite different, the proof we have given above for Lemma 4.10 is similar to results in [41] and [17]. Similar ideas and results appear [35]. The main ideas and techniques underlying proofs of this type were originally developed in [7, 8, 39].

4.3 Product and sum splitting theory for non-nested Schwarz methods

The main theory for Schwarz methods based on non-nested subspaces, as in the case of overlapping domain decomposition-like methods, may be summarized in the following way. We still consider an abstract method, but we assume it satisfies certain assumptions common to real overlapping Schwarz domain decomposition methods. In particular, due to the local nature of the operators T_k for $k \neq 0$ arising from subspaces associated with overlapping subdomains, it will be important to allow for a special global operator T_0 for global communication of information (the need for T_0 will be demonstrated later). Therefore, we use the analysis framework of the previous sections which includes the use of a special global operator T_0. Note that the local nature of the remaining T_k will imply that $\rho(\Theta) \leq N_c$, where N_c is the number of maximum number of subdomains which overlap any subdomain in the region.

The analysis of domain decomposition-type algorithms is in most respects a straightforward application of the theory of products and sums of operators, as presented earlier. The theory for multigrid-type algorithms is more subtle; we will discuss this in the next section.

Let the operators E and P be defined as:

$$E = (I - T_J)(I - T_{J-1}) \cdots (I - T_0), \tag{20}$$

$$P = T_0 + T_1 + \cdots + T_J, \tag{21}$$

where the operators $T_k \in \mathbf{L}(\mathcal{H}, \mathcal{H})$ are defined in terms of the approximate corrections in the spaces $\mathcal{H}_k$ as:

$$T_k = I_k R_k I_k^T A, \qquad k = 0, \ldots, J, \tag{22}$$

where

$$I_k : \mathcal{H}_k \mapsto \mathcal{H}, \qquad \mathrm{null}(I_k) = \{0\}, \qquad I_k \mathcal{H}_k \subseteq \mathcal{H}, \qquad \mathcal{H} = \sum_{k=1}^{J} I_k \mathcal{H}_k.$$

The following assumptions are required; note that the following theory employs many of the assumptions and lemmas of the previous sections, for the case that $\mathcal{V}_k \equiv \mathcal{H}_k$.

Assumption 4.3 *(Subspace solvers) The operators $R_k \in \mathbf{L}(\mathcal{H}_k, \mathcal{H}_k)$ are SPD. Further, there exists parameters $0 < \omega_0 \leq \omega_1 < 2$, such that*

$$\omega_0 (A_k v_k, v_k) \leq (A_k R_k A_k v_k, v_k) \leq \omega_1 (A_k v_k, v_k), \quad \forall v_k \in \mathcal{H}_k, \quad k = 0, \ldots, J.$$

Assumption 4.4 *(Splitting constant) Given any $v \in \mathcal{H}$, there exists $S_0 > 0$ and a particular splitting $v = \sum_{k=0}^{J} I_k v_k$, $v_k \in \mathcal{H}_k$, such that*

$$\sum_{k=0}^{J} \|I_k v_k\|_A^2 \le S_0 \|v\|_A^2.$$

Definition 4.4 *(Interaction matrix) The interaction matrix $\Theta \in \mathbf{L}(\mathbb{R}^J, \mathbb{R}^J)$ is defined to have as entries Θ_{ij} the smallest constants satisfying:*

$$|(AI_i u_i, I_j v_j)| \le \Theta_{ij} (AI_i u_i, I_i u_i)^{1/2} (AI_j v_j, I_j v_j)^{1/2}, \ 1 \le i, j \le J, \ u_i \in \mathcal{H}_i, v_j \in \mathcal{H}_j.$$

Theorem 4.11 *(Multiplicative method) Under Assumptions 4.3 and 4.4, it holds that*

$$\|E\|_A^2 \le 1 - \frac{\omega_0 (2 - \omega_1)}{S_0 (6 + 6\omega_1^2 \rho(\Theta)^2)}.$$

Proof. By Lemma 4.5, Assumption 4.3 implies that Assumption 3.8 holds, with $\omega = \omega_1$. By Lemma 4.8, we know that Assumptions 4.3 and 4.4 imply that Assumption 3.9 holds, with $C_0 = S_0/\omega_0$. By Lemma 4.9, we know that Definition 4.4 is equivalent to Definition 3.2 for Θ. Therefore, the theorem follows by application of Theorem 3.23. □

Theorem 4.12 *(Additive method) Under Assumptions 4.3 and 4.4, it holds that*

$$\kappa_A(P) \le \frac{S_0(\rho(\Theta) + 1)\omega_1}{\omega_0}.$$

Proof. By Lemma 4.5, Assumption 4.3 implies that Assumption 3.8 holds, with $\omega = \omega_1$. By Lemma 4.8, we know that Assumptions 4.3 and 4.4 imply that Assumption 3.9 holds, with $C_0 = S_0/\omega_0$. By Lemma 4.9, we know that Definition 4.4 is equivalent to Definition 3.2 for Θ. Therefore, the theorem follows by application of Theorem 3.24. □

Remark 4.14. Note that Assumption 4.3 is equivalent to

$$\kappa_A(R_k A_k) \le \frac{\omega_1}{\omega_0}, \qquad k = 0, \ldots, J,$$

or $\max_k\{\kappa_A(R_k A_k)\} \le \omega_1/\omega_0$. Thus, the result in Theorem 4.12 can be written as:

$$\kappa_A(P) \le S_0(\rho(\Theta) + 1) \max_k \{\kappa_A(R_k A_k)\}.$$

Therefore, the *global* condition number is completely determined by the *local* condition numbers, the splitting constant, and the interaction property.

Remark 4.15. We have the default estimate for $\rho(\Theta)$:

$$\rho(\Theta) \le J.$$

For use of the theory above, we must also estimate the splitting constant S_0, and the subspace solver spectral bounds ω_0 and ω_1, for each particular application.

Remark 4.16. Note that if a coarse space operator T_0 is not present, then the alternate bounds from the previous sections could have been employed. However, the advantage of the above approach is that the additional space $\mathcal{H}_0$ does not adversely effect the bounds, while it provides an additional space to help satisfy the splitting assumption. In fact, in the finite element case, it is exactly this coarse space which allows one to show that S_0 does not depend on the number of subspaces, yielding optimal algorithms when a coarse space is involved.

Remark 4.17. The theory in this section was derived mainly from work in the domain decomposition community, due chiefly to Widlund and his co-workers. In particular, our presentation owes much to [39] and [14].

4.4 Product and sum splitting theory for nested Schwarz methods

The main theory for Schwarz methods based on nested subspaces, as in the case of multigrid-like methods, is summarized in this section. By "nested" subspaces, we mean here that there are additional subspaces $\mathcal{V}_k \subseteq \mathcal{H}_k$ of importance, and we refine the analysis to consider these addition nested subspaces $\mathcal{V}_k$. Of course, we must still assume that $\sum_{k=1}^J I_k \mathcal{V}_k = \mathcal{H}$. Later, when analyzing multigrid methods, we will consider in fact a nested sequence $I_1\mathcal{H}_1 \subseteq I_2\mathcal{H}_2 \subseteq \cdots \subseteq \mathcal{H}_J \equiv \mathcal{H}$, with $\mathcal{V}_k \subseteq \mathcal{H}_k$, although this assumption is not necessary here. We will however assume here that one space $\mathcal{H}_1$ automatically performs the role of a "global" space, and hence it will not be necessary to include a special global space $\mathcal{H}_0$ as in the non-nested case. Therefore, we will employ the analysis framework of the previous sections which does not specifically include a special global operator T_0. (By working with the subspaces $\mathcal{V}_k$ rather than the $\mathcal{H}_k$ we will be able to avoid the problems encountered with a global operator interacting with all other operators, as in the previous sections.)

The analysis of multigrid-type algorithms is more subtle than analysis for overlapping domain decomposition methods, in that the efficiency of the method comes from the effectiveness of simple linear methods (e.g., Gauss-Seidel iteration) at reducing the error in a certain sub-subspace $\mathcal{V}_k$ of the "current" space $\mathcal{H}_k$. The overall effect on the error is not important; just the effectiveness of the linear method on error subspace $\mathcal{V}_k$. The error in the remaining space $\mathcal{H}_k \backslash \mathcal{V}_k$ is handled by subspace solvers in the other subspaces, since we assume that $\mathcal{H} = \sum_{k=1}^J I_k \mathcal{V}_k$. Therefore, in the analysis of the nested space methods to follow, the spaces $\mathcal{V}_k \subseteq \mathcal{H}_k$ introduced earlier will play a key role. This is in contrast to the non-nested theory of the previous section, where it was taken to be the case that $\mathcal{V}_k \equiv \mathcal{H}_k$. Roughly speaking, nested space algorithms "split" the error into components in $\mathcal{V}_k$, and if the subspace solvers in each space $\mathcal{H}_k$ are good at reducing the error in $\mathcal{V}_k$, then the overall method will be good.

Let the operators E and P be defined as:

$$E = (I - T_J)(I - T_{J-1}) \cdots (I - T_1), \tag{23}$$

$$P = T_1 + T_2 + \cdots + T_J, \tag{24}$$

where the operators $T_k \in \mathbf{L}(\mathcal{H}, \mathcal{H})$ are defined in terms of the approximate corrections in the spaces $\mathcal{H}_k$ as:

$$T_k = I_k R_k I_k^T A, \qquad k = 1, \ldots, J, \tag{25}$$

where

$$I_k : \mathcal{H}_k \mapsto \mathcal{H}, \qquad \text{null}(I_k) = \{0\}, \qquad I_k\mathcal{H}_k \subseteq \mathcal{H}, \qquad \mathcal{H} = \sum_{k=1}^J I_k \mathcal{H}_k.$$

The following assumptions are required.

Assumption 4.5 *(Subspace solvers) The operators $R_k \in \mathbf{L}(\mathcal{H}_k, \mathcal{H}_k)$ are SPD. Further, there exists subspaces $I_k\mathcal{V}_k \subseteq I_k\mathcal{H}_k \subseteq \mathcal{H} = \sum_{k=1}^J I_k\mathcal{V}_k$, and parameters $0 < \omega_0 \le \omega_1 < 2$, such that*

$$\omega_0(A_k v_k, v_k) \le (A_k R_k A_k v_k, v_k), \quad \forall v_k \in \mathcal{V}_k \subseteq \mathcal{H}_k, \quad k = 1, \ldots, J,$$

$$(A_k R_k A_k v_k, v_k) \le \omega_1 (A_k v_k, v_k), \quad \forall v_k \in \mathcal{H}_k, \quad k = 1, \ldots, J.$$

Assumption 4.6 *(Splitting constant) Given any $v \in \mathcal{H}$, there exists subspaces $I_k\mathcal{V}_k \subseteq I_k\mathcal{H}_k \subseteq \mathcal{H} = \sum_{k=1}^J I_k\mathcal{V}_k$ (the same subspaces $\mathcal{V}_k$ as in Assumption 4.5 above) and a particular splitting $v = \sum_{k=1}^J I_k v_k$, $v_k \in \mathcal{V}_k$, such that*

$$\sum_{k=1}^J \|I_k v_k\|_A^2 \le S_0 \|v\|_A^2, \quad \forall v \in \mathcal{H},$$

for some splitting constant $S_0 > 0$.

Definition 4.5 *(Strong interaction matrix) The interaction matrix $\Theta \in \mathbf{L}(\mathbb{R}^J, \mathbb{R}^J)$ is defined to have as entries Θ_{ij} the smallest constants satisfying:*

$$|(AI_i u_i, I_j v_j)| \le \Theta_{ij} (AI_i u_i, I_i u_i)^{1/2} (AI_j v_j, I_j v_j)^{1/2}, \; 1 \le i,j \le J, \; u_i \in \mathcal{H}_i, v_j \in \mathcal{H}_j.$$

Definition 4.6 *(Weak interaction matrix) The strictly upper-triangular interaction matrix $\Xi \in \mathbf{L}(\mathbb{R}^J, \mathbb{R}^J)$ is defined to have as entries Ξ_{ij} the smallest constants satisfying:*

$$|(AI_iu_i, I_jv_j)| \leq \Xi_{ij}(AI_iu_i, I_iu_i)^{1/2}(AI_jv_j, I_jv_j)^{1/2}, \ 1 \leq i < j \leq J, \ u_i \in \mathcal{H}_i, v_j \in \mathcal{V}_j \subseteq \mathcal{H}_j.$$

Theorem 4.13 *(Multiplicative method) Under Assumptions 4.5 and 4.6, it holds that*

$$\|E\|_A^2 \leq 1 - \frac{\omega_0(2-\omega_1)}{S_0(1+\omega_1\|\Xi\|_2)^2}.$$

Proof. The proof of this result is more subtle than the additive method, and requires more work than a simple application of the product operator theory. This is due to the fact that the weak interaction matrix of Definition 4.6 specifically involves the subspace $\mathcal{V}_k \subseteq \mathcal{H}_k$. Therefore, rather than employing Theorem 3.25, which employs Corollary 3.6 indirectly, we must do a more detailed analysis, and employ the original Theorem 3.5 directly. (See the remarks preceding Lemma 4.10.)

By Lemma 4.5, Assumption 4.5 implies that Assumption 3.1 holds, with $\omega = \omega_1$. Now, to employ Theorem 3.5, it suffices to realize that Assumption 3.3 holds with with $C_1 = S_0(1+\omega_1\|\Xi\|_2)^2/\omega_0$. This follows from Lemma 4.10. □

Theorem 4.14 *(Additive method) Under Assumptions 4.5 and 4.6, it holds that*

$$\kappa_A(P) \leq \frac{S_0\rho(\Theta)\omega_1}{\omega_0}.$$

Proof. By Lemma 4.5, Assumption 4.5 implies that Assumption 3.8 holds, with $\omega = \omega_1$. By Lemma 4.8, we know that Assumptions 4.5 and 4.6 imply that Assumption 3.9 holds, with $C_0 = S_0/\omega_0$. By Lemma 4.9, we know that Definition 4.5 is equivalent to Definition 3.2 for Θ. Therefore, the theorem follows by application of Theorem 3.26. □

Remark 4.18. We have the default estimates for $\|\Xi\|_2$ and $\rho(\Theta)$:

$$\|\Xi\|_2 \leq \sqrt{J(J-1)/2} < J, \qquad \rho(\Theta) \leq J.$$

For use of the theory above, we must also estimate the splitting constant S_0, and the subspace solver spectral bounds ω_0 and ω_1, for each particular application.

Remark 4.19. The theory in this section was derived from several sources; in particular, our presentation owes much to [39], [17], and to [41].

5. Applications to domain decomposition

Domain decomposition methods were first proposed by H.A. Schwarz as a theoretical tool for studying elliptic problems on complicated domains, constructed as the union of simple domains. An interesting early reference not often mentioned is [21], containing both analysis and numerical examples, and references to the original work by Schwarz. In this section, we briefly describe the fundamental overlapping domain decomposition methods, and apply the theory of the previous sections to give convergence rate bounds.

5.1 Variational formulation and subdomain-based subspaces

Given a domain Ω and coarse triangulation by J regions $\{\Omega_k\}$ of mesh size H_k, we refine (several times) to obtain a fine mesh of size h_k. The regions defined by the initial triangulation Ω_k are then extended by δ_k to form the "overlapping subdomains" Ω'_k. Now, let V and V_0 denote the finite element spaces associated with the h_k and H_k triangulation of Ω, respectively. The variational problem in V has the form:

$$\text{Find } u \in V \text{ such that } a(u,v) = f(v), \quad \forall v \in V.$$

The form $a(\cdot,\cdot)$ is bilinear, symmetric, coercive, and bounded, whereas $f(\cdot)$ is linear and bounded. Therefore, through the Riesz representation theorem we can associate with the above problem an abstract operator equation $Au = f$, where A is SPD.

Domain decomposition methods can be seen as iterative methods for solving the above operator equation, involving approximate projections of the error onto subspaces of V associated with the overlapping subdomains Ω'_k. To be more specific, let $V_k = H_0^1(\Omega'_k) \cap V$, $k = 1, \ldots, J$; it is not difficult to show that $V = V_1 + \cdots + V_J$, where the coarse space V_0 may also be included in the sum.

5.2 The multiplicative and additive Schwarz methods

We denote as A_k the restriction of the operator A to the space V_k, corresponding to (any) discretization of the original problem restricted to the subdomain Ω'_k. Algebraically, it can be shown that $A_k = I_k^T A I_k$, where I_k is the natural inclusion in $\mathcal{H}$ and I_k^T is the corresponding projection. The property that I_k is the natural inclusion and I_k^T is the corresponding projection holds if either $\mathcal{V}_k$ is a finite element space or the Euclidean space $\mathbb{R}^{n_k}$ (in the case of multigrid, I_k and I_k^T are inclusion and projection only in the finite element space case). In other words, domain decomposition methods automatically satisfy the variational condition, Definition 4.1, in the subspaces V_k, $k \neq 0$, for *any* discretization method.

Now, if $R_k \approx A_k^{-1}$, we can define the approximate A-orthogonal projector from V onto V_k as $T_k = I_k R_k I_k^T A$. An overlapping domain decomposition method can be written as the basic linear method, Algorithm 2.1, where the *multiplicative Schwarz* error propagator E is:

$$E = (I - T_J)(I - T_{J-1}) \cdots (I - T_0).$$

The *additive Schwarz* preconditioned system operator P is:

$$P = T_0 + T_1 + \cdots + T_J.$$

Therefore, the overlapping multiplicative and additive domain decomposition methods fit exactly into the framework of abstract multiplicative and additive Schwarz methods discussed in the previous sections.

5.3 Algebraic domain decomposition methods

As remarked above, for domain decomposition methods it automatically holds that $A_k = I_k^T A I_k$, where I_k is the natural inclusion, I_k^T is the corresponding projection, and V_k is either a finite element space or $\mathbb{R}^{n_k}$. While this *variational condition* holds for multigrid methods only in the case of finite element discretizations, or when directly enforced as in algebraic multigrid methods (see the next section), the condition holds naturally and automatically for domain decomposition methods employing any discretization technique.

We see that the Schwarz method framework then applies equally well to domain decomposition methods based on other discretization techniques (box-method or finite differences), or to algebraic equations having a block-structure which can be viewed as being associated with the discretization of an elliptic equation over a domain. The Schwarz framework can be used to provide a convergence analysis even in the algebraic case, although the results may be suboptimal compared to the finite element case when more information is available about the continuous problem.

5.4 Convergence theory for the algebraic case

For domain decomposition methods, the local nature of the projection operators will allow for a simple analysis of the interaction properties required for the Schwarz theory. To quantify the local nature of the projection operators, assume that we are given $\mathcal{H} = \sum_{k=0}^{J} I_k\mathcal{H}_k$ along with the subspaces $I_k\mathcal{H}_k \subseteq \mathcal{H}$, and denote as P_k the A-orthogonal projector onto $I_k\mathcal{H}_k$. We now make the following definition.

Definition 5.1 *For each operator P_k, $1 \le k \le J$, define $N_c^{(k)}$ to be the number of operators P_i such that $P_kP_i \neq 0$, $1 \le i \le J$, and let $N_c = \max_{1\le k\le J}\{N_c^{(k)}\}$.*

Remark 5.20. This is a natural condition for domain decomposition methods, where $N_c^{(k)}$ represents the number of subdomains which overlap a given domain associated with P_k, excluding a possible coarse space $I_0\mathcal{H}_0$. By treating the projector P_0 separately in the analysis, we allow for a global space $\mathcal{H}_0$ which may in fact interact with all of the other spaces. Note that $N_c \le J$ in general with Schwarz methods; with domain decomposition, we can show that $N_c = O(1)$. Our use of the notation N_c comes from the idea that N_c represents essentially the minimum number of colors required to color the subdomains so that no two subdomains sharing interior mesh points have the same color. (If the domains were non-overlapping, then this would be a case of the four-color problem, so that in two dimensions it would always hold that $N_c \le 4$.)

The following splitting is the basis for applying the theory of the previous sections. Note that this splitting is well-defined in a completely algebraic setting without further assumptions.

Lemma 5.1 *Given any $v \in \mathcal{H} = \sum_{k=0}^{J} I_k\mathcal{H}_k$, $I_k\mathcal{H}_k \subseteq \mathcal{H}$, there exists a particular splitting $v = \sum_{k=0}^{J} I_kv_k$, $v_k \in \mathcal{H}_k$, such that*

$$\sum_{k=0}^{J} \|I_kv_k\|_A^2 \le S_0\|v\|_A^2,$$

for the splitting constant $S_0 = 1 + N_cJ$.

Proof. We have the projectors $P_k : \mathcal{H} \mapsto I_k\mathcal{H}_k$ as defined in Lemma 4.2. Define $\hat{P}_k = P_k, k = 1,\ldots,J$, and $\hat{P}_0 = I - \sum_{k=1}^{J} P_k$. Now, let $P = \sum_{k=0}^{J} \hat{P}_k$. Clearly, $P = I$ by definition, so that P defines a splitting of $v \in \mathcal{H}$ as:

$$Pv = \sum_{k=0}^{J} \hat{P}_kv = \sum_{k=0}^{J} I_kv_k = v, \qquad v_k \in \mathcal{H}_k.$$

Now, consider

$$\sum_{k=0}^{J} \|I_kv_k\|_A^2 = \sum_{k=0}^{J} \|\hat{P}_kv\|_A^2 = \sum_{k=0}^{J}(A\hat{P}_kv, \hat{P}_kv) = \sum_{k=1}^{J}(A\hat{P}_kv, \hat{P}_kv) + (A\hat{P}_0v, \hat{P}_0v)$$

$$= \sum_{k=1}^{J}(AP_kv, P_kv) + (A[I - \sum_{k=1}^{J} P_k]v, [I - \sum_{j=1}^{J} P_j]v)$$

$$= \sum_{k=1}^{J}(AP_kv, P_kv) + (Av, v) - 2\sum_{k=1}^{J}(AP_kv, v) + \sum_{k=1}^{J}\sum_{j=1}^{J}(AP_kv, P_jv)$$

$$= (Av, v) - \sum_{k=1}^{J}(AP_kv, P_kv) + \sum_{k=1}^{J}\sum_{j=1}^{J}(AP_kv, P_jv) = (Av, v) + \sum_{k=1}^{J}\sum_{\substack{j=1 \\ j\neq k}}^{J}(AP_kv, P_jv).$$

Now, since $(AP_kv, P_jv) \leq (AP_kv, P_kv)^{1/2}(AP_jv, P_jv)^{1/2} \leq (Av, v)^{1/2}(Av, v)^{1/2} = (Av, v)$, we have that

$$\sum_{k=0}^{J} \|I_kv_k\|_A^2 = \|v\|_A^2 + \sum_{k=1}^{J}\sum_{\substack{j=1\\ j\neq k}}^{J}(AP_kv, P_jv) \leq \|v\|_A^2 + \sum_{k=1}^{J} N_c\|v\|_A^2 = (1+N_cJ)\|v\|_A^2.$$

□

Lemma 5.2 *It holds that* $\rho(\Theta) \leq N_c$.

Proof. This follows easily, since $\rho(\Theta) \leq \|\Theta\|_1 = \max_j\{\sum_i |\Theta_{ij}|\} \leq N_c$. □

We make the following assumption on the subspace solvers.

Assumption 5.1 *Assume there exists SPD operators* $R_k \in \mathbf{L}(\mathcal{H}_k, \mathcal{H}_k)$ *and parameters* $0 < \omega_0 \leq \omega_1 < 2$, *such that*

$$\omega_0(A_kv_k, v_k) \leq (A_kR_kA_kv_k, v_k) \leq \omega_1(A_kv_k, v_k), \quad \forall v_k \in \mathcal{H}_k, \quad k = 1, \ldots, J.$$

Theorem 5.3 *Under Assumption 5.1, the multiplicative Schwarz domain decomposition method has an error propagator which satisfies:*

$$\|E\|_A^2 \leq 1 - \frac{\omega_0(2-\omega_1)}{(1+N_cJ)(6+6\omega_1^2N_c^2)}.$$

Proof. By Assumption 5.1, we have that Assumption 4.3 holds. By Lemma 5.1, we have that Assumption 4.4 holds, with $S_0 = 1 + N_cJ$. By Lemma 5.2, we have that for Θ as in Definition 4.4, it holds that $\rho(\Theta) \leq N_c$. The proof now follows from Theorem 4.11. □

Theorem 5.4 *Under Assumption 5.1, the additive Schwarz domain decomposition method as a preconditioner gives a condition number bounded by:*

$$\kappa_A(P) \leq (1+N_cJ)(1+N_c)\frac{\omega_1}{\omega_0}.$$

Proof. By Assumption 5.1, we have that Assumption 4.3 holds. By Lemma 5.1, we have that Assumption 4.4 holds, with $S_0 = 1 + N_cJ$. By Lemma 5.2, we have that for Θ as in Definition 4.4, it holds that $\rho(\Theta) \leq N_c$. The proof now follows from Theorem 4.12. □

5.5 Improved results through finite element theory

If a coarse space is employed, and the overlap of the subdomains δ_k is on the order of the subdomain size H_k, i.e., $\delta_k = cH_k$, then one can show the following result. Also required is some elliptic regularity, e.g., $u \in H^{1+\alpha}(\Omega)$ along with existence of C such that $\|u\|_{H^{1+\alpha}(\Omega)} \leq C\|f\|_{H^{\alpha-1}(\Omega)}$.

Lemma 5.5 *There exists a splitting* $v = \sum_{k=0}^{J} I_kv_k$, $v_k \in \mathcal{H}_k$ *such that*

$$\sum_{k=0}^{J} \|I_kv_k\|_A^2 \leq S_0\|v\|_A^2, \quad \forall v \in \mathcal{H},$$

where S_0 *is independent of* J *(and* h_k *and* H_k*).*

Proof. Refer for example to the proof in [39] and the references therein to related results. □

6. Applications to multigrid

Multigrid methods were first developed by Federenko in the early 1960's, and have been extensively studied and developed since they became widely known in the late 1970's. In this section, we briefly describe the linear multigrid method as a Schwarz method, and apply the theory of the previous sections to give convergence rate bounds.

6.1 Recursive multigrid and nested subspaces

Consider a set of finite-dimensional Hilbert spaces $\mathcal{H}_k$ of increasing dimension:

$$\dim(\mathcal{H}_1) < \dim(\mathcal{H}_2) < \cdots < \dim(\mathcal{H}_J).$$

The spaces $\mathcal{H}_k$, which may for example be finite element function spaces, or simply $\mathbf{R}^{n_k}$ (where $n_k = \dim(\mathcal{H}_k)$), are assumed to be connected by prolongation operators $I_{k-1}^k \in \mathbf{L}(\mathcal{H}_{k-1}, \mathcal{H}_k)$, and restriction operators $I_k^{k-1} \in \mathbf{L}(\mathcal{H}_k, \mathcal{H}_{k-1})$. We can use these various operators to define mappings I_k that provide a nesting structure for the set of spaces $\mathcal{H}_k$ as follows:

$$I_1\mathcal{H}_1 \subset I_2\mathcal{H}_2 \subset \cdots \subset I_J\mathcal{H}_J \equiv \mathcal{H},$$

where

$$I_J = I, \quad I_k = I_{J-1}^J I_{J-2}^{J-1} \cdots I_{k+1}^{k+2} I_k^{k+1}, \quad k = 1, \ldots, J-1.$$

We assume that each space $\mathcal{H}_k$ is equipped with an inner-product $(\cdot,\cdot)_k$ inducing the norm $\|\cdot\|_k = (\cdot,\cdot)_k^{1/2}$. Also associated with each $\mathcal{H}_k$ is an operator A_k, assumed to be SPD with respect to $(\cdot,\cdot)_k$. It is assumed that the operators satisfy *variational conditions*:

$$A_{k-1} = I_k^{k-1} A_k I_{k-1}^k, \qquad I_k^{k-1} = (I_{k-1}^k)^T. \tag{26}$$

These conditions hold naturally in the finite element setting, and are imposed directly in algebraic multigrid methods.

Given $B \approx A^{-1}$ in the space $\mathcal{H}$, the *basic linear method* constructed from the preconditioned system $BAu = Bf$ has the form:

$$u^{n+1} = u^n - BAu^n + Bf = (I - BA)u^n + Bf. \tag{27}$$

Now, given some B, or some procedure for applying B, we can either formulate a linear method using $E = I - BA$, or employ a CG method for $BAu = Bf$ if B is SPD.

6.2 Variational multigrid as a multiplicative Schwarz method

The recursive formulation of multigrid methods has been well-known for more than fifteen years; mathematically equivalent forms of the method involving product error propagators have been recognized and exploited theoretically only very recently. In particular, it can be shown [7, 20, 29] that if the variational conditions (26) hold, then the multigrid error propagator can be factored as:

$$E = I - BA = (I - T_J)(I - T_{J-1}) \cdots (I - T_1),$$

where:

$$I_J = I, \quad I_k = I_{J-1}^J I_{J-2}^{J-1} \cdots I_{k+1}^{k+2} I_k^{k+1}, \quad k = 1, \ldots, J-1,$$

$$T_1 = I_1 A_1^{-1} I_1^T A, \quad T_k = I_k R_k I_k^T A, \quad k = 2, \ldots, J,$$

where $R_k \approx A_k^{-1}$ is the "smoothing" operator employed in each space $\mathcal{H}_k$.

6.3 Algebraic multigrid methods

Equations arising in various application areas often contain complicated discontinuous coefficients, the shapes of which may not be resolvable on all coarse mesh element boundaries as required for accurate finite element approximation (and as required for validity of finite element error estimates). Multigrid methods typically perform badly, and even the regularity-free multigrid convergence theory [7] is invalid.

Possible approaches include coefficient averaging methods (cf. [1]) and the explicit enforcement of the conditions (26) (cf. [1, 11, 33]). By introducing a symbolic stencil calculus and employing Maple or Mathematica, the conditions (26) can be enforced algebraically in an efficient way for certain types of sparse matrices; details may be found for example in the appendix of [20].

If one imposes the variational conditions (26) algebraically, then from our comments in the previous section we know that algebraic multigrid methods can be viewed as multiplicative Schwarz methods, and we can attempt to analyze the convergence rate of algebraic multigrid methods using the Schwarz theory framework.

6.4 Convergence theory for the algebraic case

The following splitting is the basis for applying the theory of the previous sections. Note that this splitting is well-defined in a completely algebraic setting without further assumptions.

Lemma 6.1 *Given any $v \in \mathcal{H} = \sum_{k=0}^{J} I_k\mathcal{H}_k$, $I_{k-1}\mathcal{H}_{k-1} \subseteq I_k\mathcal{H}_k \subseteq \mathcal{H}$, there exists subspaces $I_k\mathcal{V}_k \subseteq I_k\mathcal{H}_k \subseteq \mathcal{H} = \sum_{k=1}^{J} I_k\mathcal{V}_k$, and a particular splitting $v = \sum_{k=0}^{J} I_k v_k$, $v_k \in \mathcal{V}_k$, such that*

$$\sum_{k=0}^{J} \|I_k v_k\|_A^2 \equiv \|v\|_A^2.$$

The subspaces are $\mathcal{V}_k = (P_k - P_{k-1})\mathcal{H}$, and the splitting is $v = \sum_{k=1}^{J}(P_k - P_{k-1})v$.

Proof. We have the projectors $P_k : \mathcal{H} \mapsto I_k\mathcal{H}_k$ as defined in Lemma 4.2, where we take the convention that $P_J = I$, and that $P_0 = 0$. Since $I_{k-1}\mathcal{H}_{k-1} \subset I_k\mathcal{H}_k$, we know that $P_kP_{k-1} = P_{k-1}P_k = P_{k-1}$. Now, let us define:

$$\hat{P}_1 = P_1, \qquad \hat{P}_k = P_k - P_{k-1}, \qquad k = 2, \ldots, J.$$

By Theorem 9.6-2 in [24] we have that each $\hat{P}_k$ is a projection. (It is easily verified that $\hat{P}_k$ is idempotent and A-self-adjoint.) Define now $\mathcal{V}_k = \hat{P}_k\mathcal{H}, k = 1, \ldots, J$.

Note that

$$\hat{P}_k\hat{P}_j = (P_k - P_{k-1})(P_j - P_{j-1}) = P_kP_j - P_kP_{j-1} - P_{k-1}P_j + P_{k-1}P_{j-1}.$$

Thus, if $k > j$, then

$$\hat{P}_k\hat{P}_j = P_j - P_{j-1} - P_j + P_{j-1} = 0.$$

Similarly, if $k < j$, then

$$\hat{P}_k\hat{P}_j = P_k - P_k - P_{k-1} + P_{k-1} = 0.$$

Thus,

$$\mathcal{H} = \mathcal{V}_1 \oplus \mathcal{V}_2 \oplus \cdots \oplus \mathcal{V}_J = \hat{P}_1\mathcal{H} \oplus \hat{P}_2\mathcal{H} \oplus \cdots \oplus \hat{P}_J\mathcal{H},$$

and $P = \sum_{k=1}^{J} \hat{P}_k = I$ defines a splitting (an A-orthogonal splitting) of $\mathcal{H}$. We then have that

$$\|v\|_A^2 = (APv, v) = \sum_{k=1}^{J}(A\hat{P}_kv, v) = \sum_{k=1}^{J}(A\hat{P}_kv, \hat{P}_kv) = \sum_{k=1}^{J} \|\hat{P}_kv\|_A^2 = \sum_{k=1}^{J} \|I_kv_k\|_A^2.$$

□

For the particular splitting employed above, the weak interaction property is quite simple.

Lemma 6.2 *The (strictly upper-triangular) interaction matrix $\Xi \in \mathbf{L}(\mathbb{R}^J, \mathbb{R}^J)$, having entries Ξ_{ij} as the smallest constants satisfying:*

$$|(AI_iu_i, I_jv_j)| \leq \Xi_{ij}(AI_iu_i, I_iu_i)^{1/2}(AI_jv_j, I_jv_j)^{1/2}, \ 1 \leq i < j \leq J, \ u_i \in \mathcal{H}_i, v_j \in \mathcal{V}_j \subseteq \mathcal{H}_j,$$

satisfies $\Xi \equiv 0$ for the subspace splitting $\mathcal{V}_k = \hat{P}_k\mathcal{H} = (P_k - P_{k-1})\mathcal{H}$.

Proof. Since $\hat{P}_jP_i = (P_j - P_{j-1})P_i = P_jP_i - P_{j-1}P_i = P_i - P_i = 0$ for $i < j$, we have that $\mathcal{V}_j = \hat{P}_j\mathcal{H}$ is orthogonal to $\mathcal{H}_i = P_i\mathcal{H}$, for $i < j$. Thus, it holds that

$$(AI_iu_i, I_jv_j) = 0, \ 1 \leq i < j \leq J, \ u_i \in \mathcal{H}_i, v_j \in \mathcal{V}_j \subseteq \mathcal{H}_j.$$

□

The most difficult assumption to verify will be the following one.

Assumption 6.1 *There exists SPD operators R_k and parameters $0 < \omega_0 \leq \omega_1 < 2$ such that*

$$\omega_0(A_kv_k, v_k) \leq (A_kR_kA_kv_k, v_k), \quad \forall v_k \in \mathcal{V}_k = (P_k - P_{k-1})\mathcal{H} \subseteq \mathcal{H}_k, \quad k = 1, \ldots, J,$$

$$(A_kR_kA_kv_k, v_k) \leq \omega_1(A_kv_k, v_k), \quad \forall v_k \in \mathcal{H}_k, \quad k = 1, \ldots, J.$$

With this single assumption, we can state the main theorem.

Theorem 6.3 *Under Assumption 6.1, the multigrid method has an error propagator which satisfies:*

$$\|E\|_A^2 \leq 1 - \omega_0(2 - \omega_1).$$

Proof. By Assumption 6.1, Assumption 4.5 holds. The splitting in Lemma 6.1 shows that Assumption 4.6 holds, with $S_0 = 1$. Lemma 6.2 shows that for Ξ as in Definition 4.6, it holds that $\Xi \equiv 0$. The theorem now follows by Theorem 4.13. □

Remark 6.21. In order to analyze the convergence rate of an algebraic multigrid method, we now see that we must be able to estimate the two parameters ω_0 and ω_1 in Assumption 6.1. However, in an algebraic multigrid method, we are free to choose the prolongation operator I_k, which of course also influences $A_k = I_k^TAI_k$. Thus, we can attempt to select the prolongation operator I_k and the subspace solver R_k together, so that Assumption 6.1 will hold, independent of the number of levels J employed. In other words, the Schwarz theory framework can be used to help design an effective algebraic multigrid method. Whether it will be possible to select R_k and I_k satisfying the above requirements is the subject of future work.

6.5 Improved results through finite element theory

If one assumes that $u \in H^2(\Omega)$, then the following result can be shown for certain smoothing operators R_k.

Lemma 6.4 *There exists SPD operators R_k and parameters $0 < \omega_0 \leq \omega_1 < 2$ such that*

$$\omega_0(A_kv_k, v_k) \leq (A_kR_kA_kv_k, v_k), \quad \forall v_k \in \mathcal{V}_k = (P_k - P_{k-1})\mathcal{H} \subseteq \mathcal{H}_k, \quad k = 1, \ldots, J,$$

$$(A_kR_kA_kv_k, v_k) \leq \omega_1(A_kv_k, v_k), \quad \forall v_k \in \mathcal{H}_k, \quad k = 1, \ldots, J.$$

Proof. See for example the proof in [41]. □

More generally, assume only that $u \in H^1(\Omega)$, and that there exists $L^2(\Omega)$-like orthogonal projectors Q_k onto the finite element spaces $\mathcal{M}_k$, where we take the convention that $Q_J = I$ and $Q_0 = 0$. This defines the splitting

$$v = \sum_{k=1}^{J}(Q_k - Q_{k-1})v,$$

which is central to the BPWX theory [7]. Employing this splitting along with results from finite element approximation theory, it is shown in [7], using a similar Schwarz theory framework, that

$$\|E\|_A^2 \leq 1 - \frac{C}{J^{1+\nu}}, \quad \nu \in \{0, 1\}.$$

This result holds even in the presence of coefficient discontinuities (the constants being independent of the jumps in the coefficients). The restriction is that all discontinuities lie along all element boundaries on all levels. The constant ν depends on whether coefficient discontinuity "cross-points" are present.

Acknowledgments

Thanks to Dr. Stefan Vandewalle at Caltech who read the paper quite carefully, suggesting simplifications in some of the proofs and also establishing an improved bound for Lemma 3.19. The author also thanks Prof. Olof Widlund for useful discussions during his visit to Caltech in February 1994, as well as Dr. Michael Fey, Dr. Eric Van de Velde, and Prof. Dan Meiron and his applied mathematics research group at Caltech for several helpful comments.

Bibliography

[1] R. E. Alcouffe, A. Brandt, J. E. Dendy, Jr., and J. W. Painter, *The multi-grid method for the diffusion equation with strongly discontinuous coefficients*, SIAM J. Sci. Statist. Comput., 2 (1981), pp. 430–454.

[2] S. Ashby, M. Holst, T. Manteuffel, and P. Saylor, *The role of the inner product in stopping criteria for conjugate gradient iterations*, Tech. Rep. UCRL-JC-112586, Lawrence Livermore National Laboratory, 1992.

[3] S. F. Ashby, T. A. Manteuffel, and P. E. Saylor, *A taxonomy for conjugate gradient methods*, Tech. Rep. UCRL-98508, Lawrence Livermore National Laboratory, March 1988. To appear in SIAM J. Numer. Anal.

[4] O. Axelsson and V. Barker, *Finite Element Solution of Boundary Value Problems*, Academic Press, Orlando, FL, 1984.

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

[6] J. H. Bramble and J. E. Pasciak, *New convergence estimates for multigrid algorithms*, Math. Comp., 49 (1987), pp. 311–329.

[7] J. H. Bramble, J. E. Pasciak, J. Wang, and J. Xu, *Convergence estimates for multigrid algorithms without regularity assumptions*, Math. Comp., 57 (1991), pp. 23–45.

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

[9] X.-C. Cai and O. B. Widlund, *Multiplicative Schwarz algorithms for some nonsymmetric and indefinite problems*, Tech. Rep. 595, Courant Institute of Mathematical Science, New York University, New York, NY, 1992.

[10] P. Concus, G. H. Golub, and D. P. O'Leary, *A generalized conjugate gradient method for the numerical solution of elliptic partial differential equations*, in Sparse Matrix Computations, J. R. Bunch and D. J. Rose, eds., Academic Press, New York, NY, 1976, pp. 309–332.

[11] J. E. Dendy, Jr., *Two multigrid methods for three-dimensional problems with discontinuous and anisotropic coefficients*, SIAM J. Sci. Statist. Comput., 8 (1987), pp. 673–685.

[12] M. Dryja and O. B. Widlund, *Towards a unified theory of domain decomposition algorithms for elliptic problems*, in Third International Symposium on Domain Decomposition Methods for Partial Differential Equations, T. F. Chan, R. Glowinski, J. Périaux, and O. B. Widlund, eds., Philadelphia, PA, 1989, SIAM, pp. 3–21.

[13] M. Dryja and O. B. Widlund, *Multilevel additive methods for elliptic finite element problems*, Tech. Rep. 507, Courant Institute of Mathematical Science, New York University, New York, NY, 1990.

[14] M. Dryja and O. B. Widlund, *Domain decomposition algorithms with small overlap*, Tech. Rep. 606, Courant Institute of Mathematical Science, New York University, New York, NY, 1992.

[15] M. Dryja and O. B. Widlund, *Some recent results on schwarz type domain decomposition algorithms*, Tech. Rep. 615, Courant Institute of Mathematical Science, New York University, New York, NY, 1992.

[16] M. Dryja and O. B. Widlund, *Schwarz methods of neumann-neumann type for three-dimensional elliptic finite element problems*, tech. rep., Courant Institute of Mathematical Science, New York University, New York, NY, 1993. (To appear).

[17] W. Hackbusch, *Iterative Solution of Large Sparse Systems of Equations*, Springer-Verlag, Berlin, Germany, 1994.

[18] P. R. Halmos, *Finite-Dimensional Vector Spaces*, Springer-Verlag, Berlin, Germany, 1958.

[19] M. R. Hestenes and E. Stiefel, *Methods of conjugate gradients for solving linear systems*, J. Research of NBS, 49 (1952), pp. 409–435.

[20] M. Holst, *The Poisson-Boltzmann equation: An analysis and multilevel numerical solution*, tech. rep., Department of Applied Mathematics and CRPC, California Institute of Technology, 1994.

[21] L. V. Kantorovich and V. I. Krylov, *Approximate Methods of Higher Analysis*, P. Noordhoff, Ltd, Groningen, The Netherlands, 1958.

[22] A. N. Kolmogorov and S. V. Fomin, *Introductory Real Analysis*, Dover Publications, New York, NY, 1970.

[23] R. Kress, *Linear Integral Equations*, Springer-Verlag, Berlin, Germany, 1989.

[24] E. Kreyszig, *Introductory Functional Analysis with Applications*, John Wiley & Sons, Inc., New York, NY, 1990.

[25] P. L. Lions, *On the Schwarz Alternating Method. I*, in First International Symposium on Domain Decomposition Methods for Partial Differential Equations, R. Glowinski, G. H. Golub, G. A. Meurant, and J. Périaux, eds., Philadelphia, PA, 1988, SIAM, pp. 1–42.

[26] J. Mandel, *Some recent advances in multigrid methods*, Advances in Electronics and Electron Physics, 82 (1991), pp. 327–377.

[27] J. Mandel, S. McCormick, and R. Bank, *Variational multigrid theory*, in Multigrid Methods, S. McCormick, ed., SIAM, 1987, pp. 131–177.

[28] S. F. McCormick, *An algebraic interpretation of multigrid methods*, SIAM J. Numer. Anal., 19 (1982), pp. 548–560.

[29] S. F. McCormick and J. W. Ruge, *Unigrid for multigrid simulation*, Math. Comp., 41 (1983), pp. 43–62.

[30] J. M. Ortega, *Numerical Analysis: A Second Course*, Academic Press, New York, NY, 1972.

[31] P. Oswald, *Stable subspace splittings for Sobolev spaces and their applications*, Tech. Rep. MATH-93-7, Institut für Angewandte Mathematik, Friedrich-Schiller-Universitatät Jena, D-07740 Jena, FRG, September 1993.

[32] U. Rüde, *Mathematical and Computational Techniques for Multilevel Adaptive Methods*, vol. 13 of SIAM Frontiers Series, SIAM, Philadelphia, PA, 1993.

[33] J. W. Ruge and K. Stüben, *Algebraic multigrid*, in Multigrid Methods, S. McCormick, ed., SIAM, 1987, pp. 73–130.

[34] R. S. Varga, *Matrix Iterative Analysis*, Prentice-Hall, Englewood Cliffs, NJ, 1962.

[35] J. Wang, *Convergence analysis without regularity assumptions for multigrid algorithms based on SOR smoothing*, SIAM J. Numer. Anal., 29 (1992), pp. 987–1001.

[36] O. B. Widlund, *Optimal iterative refinement methods*, in Third International Symposium on Domain Decomposition Methods for Partial Differential Equations, T. F. Chan, R. Glowinski, J. Périaux, and O. B. Widlund, eds., Philadelphia, PA, 1989, SIAM, pp. 114–125.

[37] O. B. Widlund, *Some schwarz methods for symmetric and nonsymmetric elliptic problems*, Tech. Rep. 581, Courant Institute of Mathematical Science, New York University, New York, NY, 1991.

[38] J. Xu, *Theory of Multilevel Methods*, PhD thesis, Department of Mathematics, Penn State University, University Park, PA, July 1989. Technical Report AM 48.

[39] J. Xu, *Iterative methods by space decomposition and subspace correction*, SIAM Review, 34 (1992), pp. 581–613.

[40] D. M. Young, *Iterative Solution of Large Linear Systems*, Academic Press, New York, NY, 1971.

[41] H. Yserentant, *Old and new convergence proofs for multigrid methods*, Acta Numerica, (1993), pp. 285–326.

[12, 13, 14, 15, 16, 37]