ON THE SPECTRA OF SUMS OF ORTHOGONAL PROJECTIONS WITH APPLICATIONS TO PARALLEL COMPUTING

PETTER E. BJØRSTAD* AND JAN MANDEL†

Abstract. Many parallel iterative algorithms for solving symmetric, positive definite problems proceed by solving in each iteration, a number of independent systems on subspaces. The convergence of such methods is determined by the spectrum of the sums of orthogonal projections on those subspaces, while the convergence of a related sequential method is determined by the spectrum of the product of complementary projections. We study spectral properties of sums of orthogonal projections and in the case of two projections, characterize the spectrum of the sum completely in terms of the spectrum of the product.

AMSMOS: 65N30 65J10 35J20 15A18

Keywords: Orthogonal Projections, Parallel Computing, Domain Decomposition, Grid Refinement, Schwarz Alternating Method.

1. Introduction. Recently there has been a strong revival of the interest in domain decomposition algorithms for elliptic problems; cf. e.g. Glowinski et al. [12], and Chan et al. [4]. A classical algorithm of this kind is the Schwarz alternating method [20]. It proceeds by computing the solution on subdomains in a sequential fashion, and is therefore not necessarily attractive in a parallel computing environment. Similarly, the FAC algorithm [18, 17], an iterative algorithm for grid refinement problems, computes the solution to subproblems on a sequence of uniform grids.

Alternative methods that may be more suitable for parallel computing, have recently been proposed. These so called additive methods [8, 19] proceed by computing the solution on all subdomains, or in the refinement case on all grids, simultaneously, thereby making the algorithms more suitable for parallel computers. The present work has been motivated by the observed success of the above mentioned algorithms despite a rather incomplete theoretical foundation.

The convergence of these methods is determined by properties of the spectrum of certain sums of orthogonal projections. However, studies of sums of orthogonal projections *per se* seem to be missing from the literature, and more such tools are needed for the analysis of parallel iterative methods. In this paper, we collect some known propositions in a unified framework and complement this by new results.

In Section 2, we recall several parallel iterative algorithms, the analysis of which leads to sums of orthogonal projections. In Section 3, we study sums of an arbitrary number of projections, and in Section 4, we completely characterize the spectrum of

* Institutt for Informatikk, University of Bergen, Thormøhlens gate 55, N-5008 Bergen, Norway. This work was supported in part by the Norwegian Research Council for Science and the Humanities under grant D.01.08.054 and by The Royal Norwegian Council for Scientific and Industrial Research under grant IT2.28.28484.

† Computational Mathematics Group, University of Colorado at Denver, 1200 Larimer Street, Denver, CO 80204. This work was supported in part by the Air Force Office of Scientific Research under grant AFOSR-86-0126 and by the National Science Foundation under grant DMS-8704169.

the sum of two projections. Finally, in Section 5, we apply the theory of Section 4 to the case of finite element spaces.

2. Additive Algorithms for Parallel Solution of Linear Systems. Let H be a Hilbert space with inner product $a(\cdot,\cdot)$, and $\langle\cdot,\cdot\rangle$ another inner product on H extended to a duality pairing $\langle f,u\rangle$, $f\in H'$, $u\in H$ in the usual way. Consider the variational problem

$$u\in H:\quad a(u,v)=\langle f,v\rangle,\quad \forall v\in H. \tag{2.1}$$

We are mainly concerned with a discrete version of this problem, that is, the case when H is finite dimensional; however, most of the results hold in the general case. Let $\mathbf{u}$ be the solution of (2.1).

Let $V=\sum V_i,\subset H$, where V_i, $i=1,\ldots,n$ are closed subspaces of H. Following P. L. Lions [15, 16], the classical Schwarz alternating method for the iterative solution of (2.1) can be written in an abstract way as iterations $u^k\mapsto u^{k+1}$ defined by

$$\left.\begin{array}{l} u_i\in V_i:\quad a(u_i,v_i)=\langle f,v_i\rangle - a(u^{k+(i-1)/n},v_i),\quad \forall v_i\in V_i \\ u^{k+i/n}=u^{k+(i-1)/n}+u_i \end{array}\right\},\quad i=1,\ldots,n. \tag{2.2}$$

It is easy to see that it holds for the transformation of error that

$$\mathbf{u}-u^{k+1}=(I-P_{V_n})\cdots(I-P_{V_1})(\mathbf{u}-u^k),$$

where P_{V_i} is the orthogonal projection on V_i. This method decomposes the problem (2.1) into a series of subproblems (2.2) on subspaces V_i.

A parallel version of this method [2, 9, 11] is defined by

$$\begin{array}{l} u_i\in V_i:\quad a(u_i,v_i)=\langle f,v_i\rangle - a(u^k,v_i),\quad \forall v_i\in V_i,\quad i=1,\ldots,n, \\ u^{k+1}=u^k+\sum_{i=1}^n u_i, \end{array} \tag{2.3}$$

and it is easy to see that the error in (2.3) is transformed by

$$\mathbf{u}-u^{k+1}=\Big(I-\sum_{i=1}^n P_{V_i}\Big)(\mathbf{u}-u^k).$$

For obvious reasons, the method (2.2) is called the multiplicative method, and the method (2.3) is called the additive method.

Define $A:H\to H'$ by $\langle Au,v\rangle=a(u,v)$ and $C:H'\to H$ by

$$CA=\sum_{i=1}^n P_{V_i}. \tag{2.4}$$

Then the additive method (2.3) can be written in the standard form

$$u^{k+1}=u^k-C(Au^k-f),$$

the mapping C being an *approximate solver* for $Au = r$, defined by

$$Cr = \sum_{i=1}^{n} u_i, \quad u_i \in V_i, \quad a(u_i, v_i) = \langle r, v_i \rangle, \quad \forall v_i \in V_i.$$

Note that by (2.4), CA is symmetric on H in the inner product $a(\cdot,\cdot)$. Since the additive method will fail to converge if the spectral radius $\rho(CA) > 2$, which may well happen in the general case, the approximate solver C is often used as a preconditioner in the conjugate gradient method whose convergence properties are then determined by the spectrum $\sigma(CA)$. In particular, the number of steps of the preconditioned conjugate gradient method required to solve the problem to a fixed precision, grows like $(\lambda_{\max}(CA)/\lambda_{\min}(CA))^{1/2}$. See [13] for more details.

The analysis of such iterative methods thus leads to the problem of localizing the spectrum of $\sum_{i=1}^{n} P_{V_i}$. We proceed to give a few examples of highly interesting algorithms that can be put into this framework.

The Schwarz' alternating method is obtained when H is a space of functions on a domain $\Omega = \bigcup_{i=1}^{n} \Omega_i$, and

$$V_i = \{v \in V : \mathrm{supp}\ v \subset \overline{\Omega}_i\}.$$

Additive FAC (AFAC). Following [18] we briefly describe an additive algorithm for computing the composite solution on a grid having one level of refinement. This algorithm generalizes in a straightforward fashion to many (nested) levels of refinement. Let $H = H_{2h} + H_h$, where H_{2h} and H_h are finite element spaces such that $H_{2h} \subset H_0^1(\Omega_1)$, $H_h \subset H_0^1(\Omega_2)$, and naturally extended by zero. Here $\Omega_2 \subset \overline{\Omega}_1$ and $H_{2h} \cap H_0^1(\Omega_2) \subset H_h$.

Given $u \in H$ as the current approximation to the composite solution, the additive algorithm proceeds as follows [18]:

$$u_{2h} \in H_{2h} \ : \ a(u - u_{2h}, v_{2h}) = f(v_{2h}), \ \ \forall v_{2h} \in H_{2h}$$

Update u by $u \leftarrow u - u_{2h}$. The next step is

$$\begin{aligned} u_h \in H_h : \qquad & a(u - u_h, v_h) = f(v_h), \forall v_h \in H_h \\ w_h \in H_{2h} \cap H_h : \qquad & a(u - w_{2h}, v_{2h}) = f(v_{2h}), \forall v_{2h} \in H_{2h} \cap H_h. \end{aligned}$$

Update u by $u \leftarrow u - u_h + w_{2h}$.

Following [18] define

$$H_{2h}^{2h-harm} = \{u_{2h} \in H_{2h} \ : \ a(u_{2h}, v_{2h}) = 0, \forall v_{2h} \in H_{2h} \cap H_h\}.$$

With this definition, one can show [18] that the error propagates according to a formula of the same structure as in the additive Schwarz case:

$$e^{k+1} = (I - (P_{H_{2h}^{2h-harm}} + P_{H_h}))e^k,$$

and the method is a particular case of (2.3) with $V_1 = H_h$ and $V_2 = H_{2h}^{2h-harm}$. For problem-specific analysis of AFAC, see [18] for the case of two levels and [10] for a general number of levels.

Douglas and Miranker [6] studied a method in which the subspaces V_i are defined as spaces of functions which satisfy suitable symmetry and antisymmetry properties and gave algebraical conditions under which the subspaces V_i are mutually orthogonal; then the method (2.3) reduces to a direct method. The "robust multigrid method" by Hackbush [14] is also of the form (2.3) with the subspaces V_i being defined as ranges of suitable stencil operators ("prolongations") on a uniform grid. Douglas and Smith [7] studied several other methods within this framework and proved a convergence bound for Hackbusch's method. For convergence estimates in the case when the subproblems in (2.3) are solved only approximately, see Douglas and Mandel [5].

3. The spectra of sums of projections. In this section, we give several results valid for arbitrary sums of projections. In the next section we give a more complete theory for the sum of only two projections. Because we will not need to refer to the problem (2.1) any more, denote the inner product in H by (u, v) rather than $a(u, v)$. The corresponding norm is $\|u\| = (u, u)^{1/2}$.

Our first theorem is an extension of Lemma 2.3 in [18] to the infinite dimensional case. It shows that the bounds on the spectrum of the sum of projections on linearly independent subspaces can be reduced to the bounds on the spectra of Gram matrices of unit vectors, one from each subspace.

THEOREM 3.1. *Let H be a Hilbert space with inner product $(\cdot,\cdot)$, V_i closed subspaces of H, and $H = \oplus_{i=1}^n V_i$. Let P_{V_i} be the orthogonal projection onto V_i. Then*

$$\begin{aligned}\inf \sigma(\sum_{i=1}^n P_{V_i}) &= \inf_{\substack{\|v_i\|=1 \\ v_i \in V_i}} \lambda_{\min} G(v_1, \dots, v_n), \\ \sup \sigma(\sum_{i=1}^n P_{V_i}) &= \sup_{\substack{\|v_i\|=1 \\ v_i \in V_i}} \lambda_{\max} G(v_1, \dots, v_n),\end{aligned}$$

where $G(v_1, ..., v_n) = (g_{ij})$ is the Gram matrix, $g_{ij} = (v_i, v_j)$, $i, j = 1, \dots, n$.

Proof. Define the inner product in $V_1 \times \cdots \times V_n$ by

$$\Big((u_1, \dots, u_n), (v_1, \dots, v_n)\Big)_{V_1 \times \cdots \times V_n} = \sum_{i=1}^n (u_i, v_i).$$

Let $A : V_1 \times V_2 \times .... \times V_n \to H$ be given by

$$A : (v_1, v_2, \dots, v_n) \mapsto \sum_{i=1}^n v_i.$$

Then the adjoint $A^* : H \to V_1 \times V_2 \times ... \times V_n$ is

$$A^* : v \to (P_{V_1} v, \dots, P_{V_n} v)$$

because

$$\begin{aligned}\Big(A(0, \dots, 0, v_j, 0, \dots), w\Big) &= (v_j, w) = (v_j, P_{V_j} w) \\ &= \Big((0, \dots, 0, v_j, 0, \dots, 0r), (0, \dots, 0, P_{V_j} w, 0, \dots, 0)\Big)_{V_1 \times \cdots \times V_n} \\ &= \Big((0, \dots, 0, v_j, 0, \dots, 0), A^* w\Big)_{V_1 \times \cdots \times V_n}.\end{aligned}$$

Consequently,

$$AA^* = \sum_{i=1}^{n} P_{V_i}.$$

Because A is a bounded, one to one mapping of the Hilbert space $V_1 \times \cdots \times V_n$ onto H, its inverse is also bounded and thus $\sigma(A^*A) = \sigma(A^{-1}(AA^*)A) = \sigma(AA^*)$. Now write $v \in V_1 \times \cdots \times V_n$ as $v = (b_1 v_1, \ldots, b_n v_n)$ with $\|v_i\| = 1$, $v_i \in V_i$, Then the Rayleigh quotient of v is

$$\begin{aligned} RQ(v) &= \frac{(A^*Av, v)_{V_1 \times \ldots \times V_n}}{(v, v)_{V_1 \times \ldots \times V_n}} = \frac{(Av, Av)}{\sum_{i=1}^{n} b_i^2} \\ &= \frac{\sum_{i,j=1}^{n} b_i b_j (v_i, v_j)}{\sum_{i=1}^{n} b_i^2} = \frac{b^t G(v_1, \ldots, v_n) b}{b^t b}, \end{aligned}$$

where $b = (b_1, ..., b_n)^t$. ∎

The first part of the following theorem shows that for the spectrum of the sum of orthogonal projections to be bounded from below, it is sufficient that the corresponding decomposition is bounded from above. It is due to P. L. Lions [15] and the proof is given here for completeness only. The second part of the theorem provides an analogous statement for the upper bound of the spectrum.

THEOREM 3.2. *Let H be a Hilbert space with inner product $(\cdot, \cdot)$, V_i closed subspaces of H, and $H = \sum_{i=1}^{n} V_i$. Let P_{V_i} be the orthogonal projection onto V_i. Then it holds:*

(i) If there exists a constant $c_1 > 0$ such that

$$\forall v \in H \quad \exists v_i \in V_i: \ v = \sum_{i=1}^{n} v_i, \ \|v\|^2 \geq c_1 \sum_{i=1}^{n} \|v_i\|^2, \tag{3.1}$$

then

$$\inf \sigma\Big(\sum_{i=1}^{n} P_{V_i}\Big) \geq c_1.$$

(ii) If there is a constant c_2 such that

$$\forall v \in H \quad \forall v_i \in V_i, \ v = \sum_{i=1}^{n} v_i: \ \|v\|^2 \leq c_2 \sum_{i=1}^{n} \|v_i\|^2,$$

then

$$\sup \sigma\Big(\sum_{i=1}^{n} P_{V_i}\Big) \leq c_2.$$

Proof. (i) Let $v = \sum_{i=1}^{n} v_i$. Then

$$\begin{aligned} \|v\|^2 &= (v, v) = \sum_{i=1}^{n} (v, v_i) = \sum_{i=1}^{n} (v, P_{V_i} v_i) = \sum_{i=1}^{n} (P_{V_i} v, v_i) \\ &\leq \sum_{i=1}^{n} \|P_{V_i} v\| \, \|v_i\| \leq (\sum_{i=1}^{n} \|P_{V_i} v\|^2)^{1/2} (\sum_{i=1}^{n} \|v_i\|^2)^{1/2}. \end{aligned}$$

Using the assumption, we therefore get

$$\|v\|^2 \le c_1{}^{-1} \sum_{i=1}^{n} \|P_{V_i} v\|^2 = c_1^{-1}(v, \sum_{i=1}^{n} P_{V_i} v).$$

proving the first part of the theorem.

(ii) Let $v \in H$ and take $w_i = 0$ if $P_{V_i} v = 0$, and $w_i = P_{V_i} v / \|P_{V_i} v\|$ when $P_{V_i} v \neq 0$. Then

$$\|P_{V_i} v\|^2 = (P_{V_i} v, P_{V_i} v) = (P_{V_i} v, v) = (w_i \|P_{V_i} v\|, v) = (w_i, v)\|P_{V_i} v\|,$$

and, consequently, $\|P_{V_i} v\| = (w_i, v)$. Define $X : \mathbf{R}^n \to H$ by $X : d \mapsto \sum_{i=1}^n d_i w_i$. Then $X^* : H \to \mathbf{R}^n$, $X^* : u \mapsto \{(w_i, u)\}_{i=1}^n$. We now have

$$\Big(v, \sum_{i=1}^{n} P_{V_i} v\Big) = \sum_{i=1}^{n} \|P_{V_i} v\|^2 = \sum_{i=1}^{n} (w_i, v)^2 = \|X^* v\|^2 \le \|X^*\|^2 \ \|v\|^2.$$

But it holds for all $d \in R^n$ that

$$\|Xd\|^2 = \|\sum_{i=1}^{n} w_i d_i\|^2 \le c_2 \sum_{i=1}^{n} \|d_i w_i\|^2 = c_2 \sum_{i=1}^{n} d_i^2 = c_2 \|d\|^2,$$

so $\|X\| \le c_2^{1/2}$. Since $\|X\| = \|X^*\|$, it follows that

$$\Big(v, \sum_{i=1}^{n} P_{V_i} v\Big) \le \|X^*\|^2 \|v\|^2 \le c_2 \|v\|^2,$$

which concludes the proof. ∎

We should note that there is always the trivial upper bound $\sup \sigma\Big(\sum_{i=1}^n P_{V_i}\Big) \le n$, because all projections are orthogonal. Nontrivial upper bounds can often be obtained by different means; for example, for the additive Schwarz' method, it is easy to see that the upper bound can be taken to be the maximum number of subdomains having a common nonempty intersection [11].

In the case when the subspaces V_i are linearly independent, the question arises, if the Lions' assumption (3.1) implies also a nontrivial upper bound, perhaps one independent of the number of subspaces n. In the case of two linearly independent subspaces, we show in the next section that the extreme points of the spectrum of the sum of the projections are symmetrical around one; however, in the general case the problem is open.

4. The case of two subspaces. In this section, let H be a Hilbert space, which is the sum of two closed subspaces, $H = U + V$, where possibly $U \cap V \neq \{0\}$. Since all propositions hold when exchanging the roles of U and V, we may state and prove only one variant in such cases. The following lemma summarizes a number of properties we need in order to prepare for a decomposition of the space H. We note that (4.2) below is an abstract version of the specific result stated as Lemma 3.4 in [18]. Define

$$\tilde{U} = U \cap (U \cap V)^{\perp}, \qquad \tilde{V} = V \cap (U \cap V)^{\perp}.$$

LEMMA 4.1. *It holds that*

$$U = \tilde{U} \oplus (U \cap V), \tag{4.1}$$
$$P_{\tilde{U}^\perp} P_{V^\perp} = P_{U^\perp} P_{V^\perp}, \tag{4.2}$$
$$P_U = P_{U\cap V} + P_{\tilde{U}}, \tag{4.3}$$
$$P_U + P_V = 2P_{U\cap V} + P_{\tilde{U}} + P_{\tilde{V}}, \tag{4.4}$$
$$H = \tilde{U} \oplus \tilde{V} \oplus (U \cap V), \qquad \tilde{U} \oplus \tilde{V} \perp (U \cap V), \tag{4.5}$$
$$P_{\tilde{U}} P_{\tilde{V}} = P_U P_{\tilde{V}}. \tag{4.6}$$

Proof. For (4.1), note that uniqueness of the decomposition follows from the fact that $\tilde{U} \cap (U \cap V) = U \cap (U \cap V)^\perp \cap (U \cap V) = \{0\}$. In order to prove existence, let $u \in U$, $\bar{u} = P_{U\cap V} u$, and $u = \bar{u} + \tilde{u}$. But $\tilde{u} \in U$, since $U \cap V \subset U$, so $\tilde{u} \in (U \cap V)^\perp \cap U = \tilde{U}$.

Now we prove (4.2). Let $v \in V^\perp$ and $w = P_{\tilde{U}} v$. For any $z \in U$, write $z = \tilde{z} + \bar{z}$, with $\tilde{z} \in \tilde{U}$ and $\bar{z} \in U \cap V$. Then from $\bar{z} \in U \cap V$ and $w \in \tilde{U} \subset (U \cap V)^\perp$, it follows that

$$(v, \bar{z}) = (w, \bar{z}) = 0.$$

Because $(w, \tilde{z}) = (v, \tilde{z})$ by definition of a projection, it follows that $(w, z) = (v, z)$ for all $z \in U$, which implies that $w = P_U v$. Consequently, $P_{\tilde{U}^\perp} v = P_{U^\perp} v$.

Equation (4.3) follows from (4.1) and $\tilde{U} \perp (U \cap V)$.

Equation (4.4) follows immediately from (4.3).

The second statement in (4.5) follows trivially from the definitions of $\tilde{U}$ and $\tilde{V}$. To prove uniqueness in the first statement in (4.5), note that

$$\tilde{U} \cap \tilde{V} \cap (U \cap V)^\perp = U \cap V \cap (U \cap V)^\perp = \{0\}.$$

To prove existence, let $w \in H$, $w = u + v$, $u \in U$, $v \in V$. Now by (4.1), $u = \tilde{u} + \bar{u}$, $\tilde{u} \in \tilde{U}$, $\bar{u} \in U \cap V$ and in the same way, $v = \tilde{v} + \bar{v}$, $\tilde{v} \in \tilde{V}$, $\bar{v} \in U \cap V$. The proof of (4.5) is concluded by noting that $w = \tilde{u} + (\bar{u} + \bar{v}) + \tilde{v}$. The proof of (4.6) is completely analogous to that of (4.2). ∎

We now turn to measuring the angle between spaces and spectral radii of associated products of projections. Recall that the cosine of two vectors $u, v \in H$ is defined by

$$\cos(u, v) = \frac{(u, v)}{\|u\| \|v\|}$$

and the cosine of two subspaces $X, Y \subset H$ by

$$\cos(X, Y) = \sup_{\substack{x \in X \\ y \in Y}} |\cos(x, y)|.$$

We have the following simple result, which was stated and proved, e.g., by Bank and Dupont [1] and Braess [3].

Lemma 4.2. *If $H = X \oplus Y$, where X and Y are closed subspaces of H, then*

$$\rho(P_{X^\perp} P_{Y^\perp}) \le \cos^2(X, Y) \tag{4.7}$$

We can further relate projections and the cosine of subspaces as follows.

Lemma 4.3. *If $X, Y \subset H$ are closed subspaces of H, then*

$$\rho(P_X P_Y) = \cos^2(X, Y). \tag{4.8}$$

Proof. We have

$$\begin{aligned}
\rho(P_X P_Y) &= \|P_X P_Y P_X\| = \sup_{u \in H} \frac{(P_X P_Y u, P_Y u)}{\|u\|^2} \\
&= \sup_{y \in Y} \frac{(P_X y, P_X y)}{\|y\|^2} = \sup_{y \in Y} \sup_{x \in X} \frac{(x, y)^2}{\|x\|^2 \|y\|^2} \\
&= \cos^2(X, Y).
\end{aligned}$$

∎

The following statement extends Lemma 2.2 in [18] to the general (infinite dimensional) case.

Lemma 4.4. *If $X, Y \subset H$ are closed subspaces of H and $H = X \oplus Y$, then*

$$\cos(X^\perp, Y^\perp) = \cos(X, Y). \tag{4.9}$$

Proof. From Lemmas 4.2 and 4.3, we get the inequality

$$\cos(X^\perp, Y^\perp) \le \cos(X, Y).$$

Because orthogonal complements are closed, it will suffice to show for the converse inequality that

$$H = X^\perp \oplus Y^\perp.$$

Since $H = X \oplus Y$, we have $X^\perp \cap Y^\perp = \{0\}$ and $H = \overline{X^\perp + Y^\perp}$; therefore, we need only to show that $X^\perp + Y^\perp$ is closed. Let $w \in \overline{X^\perp + Y^\perp}$, that is,

$$w_n = u_n + v_n, \qquad u_n \in X^\perp,\ v_n \in Y^\perp, \qquad \|w_n - w\| \to 0, \quad n \to \infty.$$

By [15, Theorem I.1], we have from $H = X + Y$ that $\|P_{X^\perp} P_{Y^\perp}\| < 1$, so by Lemma 4.3, $\cos(X^\perp, Y^\perp) < 1$. It follows that the sequences u_n and v_n are bounded and we can thus extract weakly convergent subsequences

$$u_{n_k} \rightharpoonup u \in X^\perp, \quad v_{n_k} \rightharpoonup v \in Y^\perp, \quad k \to \infty.$$

Consequently,

$$u_{n_k} + v_{n_k} \rightharpoonup u + v = w \in X^{\perp} + Y^{\perp}, \quad k \to \infty,$$

which concludes the proof. ∎

The following theorem shows that there is a *single* number characterizing the relation of U and V.

THEOREM 4.5. *It holds that*

$$\cos(U^{\perp}, V^{\perp}) = \cos(\tilde{U}, V) = \cos(U, \tilde{V}) = \cos(\tilde{U}, \tilde{V}).$$

Proof. We have

$$\begin{aligned} \cos^2(U^{\perp}, V^{\perp}) &= \rho(P_{U^{\perp}} P_{V^{\perp}}) = \rho(P_{\tilde{U}^{\perp}} P_{V^{\perp}}) = \cos^2(\tilde{U}^{\perp}, V^{\perp}) = \cos^2(\tilde{U}, V) \\ &= \rho(P_V P_{\tilde{U}}) = \rho(P_{\tilde{V}} P_{\tilde{U}}) = \cos^2(\tilde{U}, \tilde{V}), \end{aligned}$$

using (4.8), (4.2), (4.8), (4.9), (4.7), (4.6) and (4.7) in this order. ∎

The following theorem is our first localization of the spectrum of two projections.

THEOREM 4.6. *Decompose*

$$H = (U \cap V) \oplus (U \cap V)^{\perp}.$$

Then in the block notation corresponding to this decomposition,

$$P_U + P_V = \begin{pmatrix} 2I & 0 \\ 0 & P_{\tilde{U}} + P_{\tilde{V}} \end{pmatrix},$$

with the first block void if $U \cap V = \{0\}$, and it holds

$$\begin{aligned} \inf \sigma(P_{\tilde{U}} + P_{\tilde{V}}) &= 1 - \cos(\tilde{U}, \tilde{V}) \\ \sup \sigma(P_{\tilde{U}} + P_{\tilde{V}}) &= 1 + \cos(\tilde{U}, \tilde{V}) \end{aligned}$$

Proof. From (4.4), we know that $P_U + P_V = 2P_{U \cap V} + P_{\tilde{U}} + P_{\tilde{V}}$. It remains to show that we have the stated bounds of the spectrum of $P_{\tilde{U}} + P_{\tilde{V}}$. But by Theorem 3.1, these bounds are the infimum and supremum of the eigenvalues of the 2 by 2 matrices

$$\begin{pmatrix} 1 & a \\ a & 1 \end{pmatrix}$$

with $-\cos(\tilde{U}, \tilde{V}) \le a \le \cos(\tilde{U}, \tilde{V})$, which are $1 \pm a$. ∎

In the case when the subspaces U and V are linearly independent, we recover from Theorems 4.5 and 4.6 a result of [18].

COROLLARY 4.7. *If H is finite dimensional and $H = U \oplus V$, then*

$$\rho(P_{U^{\perp}} P_{V^{\perp}}) = \big(\rho(I - P_U - P_V)\big)^2.$$

In order to obtain more detailed information of the spectrum, we need to decompose our spaces $\tilde{U}$ and $\tilde{V}$ further. This will provide us with an abstract version of the decomposition given in [2] in the context of finite element spaces.

Write $\tilde{U} = U_p \oplus U_b$ where $U_p = \tilde{U} \cap \tilde{V}^\perp$ and $U_b = \tilde{U} \cap {U_p}^\perp$. Similarly, $\tilde{V} = V_p \oplus V_b$, with the subspaces V_p and V_b defined analogously. Reordering the subspaces, we can now write a decomposition of H,

$$H = (U \cap V) \oplus U_p \oplus V_p \oplus U_b \oplus V_b. \tag{4.10}$$

Note that all subspaces in this decomposition are pairwise orthogonal except for the pair U_b and V_b.

Let us use the notation $P_{X|Y}$ to mean the orthogonal projection of the (sub)space Y to X, or, equivalently, the orthogonal projection operator onto X with the domain restricted to Y. With this notation we can state our complete decomposition.

THEOREM 4.8. *In the block notation corresponding to the decomposition (4.10), it holds that*

$$P_U + P_V = \begin{pmatrix} 2I & 0 & 0 & 0 & 0 \\ 0 & I & 0 & 0 & 0 \\ 0 & 0 & I & 0 & 0 \\ 0 & 0 & 0 & I & P_{U_b|V_b} \\ 0 & 0 & 0 & P_{V_b|U_b} & I \end{pmatrix} \tag{4.11}$$

Proof. We use Theorem 4.6 and further decompose $P_{\tilde{U}} + P_{\tilde{V}}$ on $(U \cap V)^\perp = \tilde{U} \oplus \tilde{V}$:

$$(P_{\tilde{U}} + P_{\tilde{V}})|_{\tilde{U} \oplus \tilde{V}} = \begin{pmatrix} I_{\tilde{U}} & P_{\tilde{U}|\tilde{V}} \\ P_{\tilde{V}|\tilde{U}} & I_{\tilde{V}} \end{pmatrix}. \tag{4.12}$$

Now $P_{\tilde{V}|\tilde{U}} : U_p \oplus U_b \to V_p \oplus V_b$ and

$$P_{\tilde{V}|\tilde{U}} = \begin{pmatrix} P_{V_p|U_p} & P_{V_p|U_b} \\ P_{V_b|U_p} & P_{V_b|U_b} \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & P_{V_b|U_b} \end{pmatrix},$$

because $P_{V_p|U_p} = 0$, $P_{V_p|U_b} = 0$, and $P_{V_b|U_p} = 0$ from $U_p \perp \tilde{V}$ and $V_p \perp \tilde{U}$. Substituting into (4.12) along with the block form of identity the $I_{\tilde{U}}$ and analogous expressions with U and V interchanged gives (4.11). ∎

From Theorem 4.8, we may deduce the complete structure of the spectrum.

COROLLARY 4.9. *The operator $P_U + P_V$ has the following eigenvalues and invariant subspaces:*

Eigenvalue 2 with the invariant subspace $U \cap V$.

Eigenvalue 1 with the invariant subspace $U_p \oplus V_p$.

The rest of the spectrum is of the form $1 \pm \mu$ where

$$\mu^2 \in \sigma(P_{U_b|V_b} P_{V_b|U_b}).$$

If H is finite dimensional, then the number of such eigenvalues different from 1 is at most $2 \min\{\dim U_b, \dim V_b\}$. Consequently, the conjugate gradient method for the problem (2.1) preconditioned by the approximate solver (2.3) converges in at most $1 + 2\min\{\dim U_b, \dim V_b\}$ steps.

5. An application to finite element spaces. Here we briefly explain the application of the theory to finite element spaces as used in domain decomposition algorithms [2], which motivated the general results above. Let H be a space of finite element functions with support on two overlapping subregions $\Omega^{(1)}$ and $\Omega^{(2)}$. We define the subspaces U and V to be the corresponding spaces of finite element functions (with zero traces on the boundaries) defined on $\Omega^{(1)}$ and $\Omega^{(2)}$. Following [2] we use the notations $\Omega_1 = \Omega \setminus \overline{\Omega}^{(2)}$, where $\overline{\Omega}^{(2)}$ is the closure of $\Omega^{(2)}$, $\Omega_2 = \Omega \setminus \overline{\Omega}^{(1)}$ and $\Omega_3 = \Omega^{(1)} \cap \Omega^{(2)}$. The region Ω is thus also divided into three nonoverlapping subregions Ω_1, Ω_2, and Ω_3 which are separated from each other by the curves (or surfaces) $\Gamma_4 = \overline{\Omega}_1 \cap \overline{\Omega}_3$ and $\Gamma_5 = \overline{\Omega}_2 \cap \overline{\Omega}_3$.

With subvectors and subscripts corresponding to the degrees of freedom associated with the open sets Ω_1, Ω_2 and Ω_3 and the curves (surfaces) Γ_4 and Γ_5, the entire discrete problem can be written as

$$Kx = \begin{pmatrix} K_{11} & 0 & 0 & K_{14} & 0 \\ 0 & K_{22} & 0 & 0 & K_{25} \\ 0 & 0 & K_{33} & K_{34} & K_{35} \\ K_{14}^T & 0 & K_{34}^T & K_{44} & K_{45} \\ 0 & K_{25}^T & K_{35}^T & K_{45}^T & K_{55} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \\ b_5 \end{pmatrix} \tag{5.1}$$

The stiffness matrix K has been generated in the usual way by the bilinear form $a(u, v)$. We make the unique correspondence between finite element functions u, v and the corresponding nodal values x and y. Similarly, $y^T Kx$ corresponds to $a_\Omega(u, v)$. We write any function $u \in H$ as the sum of 5 components u_1 to u_5 corresponding to the 5 components x_1 to x_5 of the vector x.

We can then identify the spaces defined in Section 4:

U: Functions u corresponding to x_1, x_3 and x_4, that is, all functions such that $x_2 = 0$ and $x_5 = 0$.

V: Functions v corresponding to x_2, x_3 and x_5.

$U \cap V$: Functions u_3 with support in $\overline{\Omega}_3$, that is, corresponding to x_3.

$(U \cap V)^\perp$: Functions which are discrete harmonic on Ω_3, corresponding to all vectors x such that $K_{33}x_3 + K_{34}x_4 + K_{35}x_5 = 0$.

$\tilde{U}$: Vectors x with possibly only x_1, x_3, and x_4 nonzero such that $K_{33}x_3 + K_{34}x_4 = 0$.

$\tilde{V}$: Only x_2, x_3, and x_5 may be nonzero and $K_{33}x_3 + K_{35}x_5 = 0$.

$\tilde{V}^\perp$: Vectors x such that $K_{22}x_2 + K_{25}x_5 = 0$ and for all y_3, y_5 such that $K_{33}y_3 + K_{35}y_5 = 0$, it holds

$$y_3^T(K_{33}x_3 + K_{34}x_4 + K_{35}x_5) + y_5^T(K_{25}^T x_2 + K_{35}^T x_3 + K_{45}^T x_4 + K_{55}x_5) = 0. \tag{5.2}$$

U_p: Since $u \in U_p$ implies that $u \in (U \cap V)^\perp$ the first term in 5.2 is zero. In particular,

$$K_{33}x_3 + K_{34}x_4 = 0. \tag{5.3}$$

Assuming that K_{33} is invertible, it follows from (5.2) that y_5 is arbitrary and we get the necessary and sufficient conditions for $u \in U_p$ are $x_2 = 0$ and $x_5 = 0$ (since $u \in U$), and $K_{35}^T x_3 + K_{45}^T x_4 = 0$. Using (5.3), we may conclude that

$$u \in U_p \quad \Leftrightarrow \quad x_2 = 0,\ x_5 = 0,\ \begin{pmatrix} K_{33} & K_{34} \\ K_{35}^T & K_{45}^T \end{pmatrix} \begin{pmatrix} x_3 \\ x_4 \end{pmatrix} = 0.$$

Thus if the matrix

$$\begin{pmatrix} K_{33} & K_{34} \\ K_{35}^T & K_{45}^T \end{pmatrix} \tag{5.4}$$

has a trivial null space, then $x_3 = 0$ and $x_4 = 0$. This is the case if it is of full rank and the dimension of x_5 is at least as large as the dimension of x_4. In this case, functions from U_p simply correspond to x such that only the subvector x_1 may be nonzero.

U_b: Here we restrict ourselves to the case when the matrix (5.4) has a trivial null space and K_{33} and K_{11} are nonsingular. Then functions in U_b are given by an arbitrary component x_4, the component x_3 is determined from (5.3), x_1 from $K_{11}x_1 + K_{14}x_4 = 0$, and $x_2 = 0$ and $x_5 = 0$. In other words, a function from U_b is then given by its values on Γ_4 and extended as a discrete harmonic function onto Ω_1 and Ω_3.

Thus in this case, the decomposition is in complete agreement with the conclusions in [2].

REFERENCES

[1] R. E. Bank and T. F. Dupont, *Analysis of a two-level scheme for solving finite element equations*, Tech. Report CNA-159, Center for Numerical Analysis, University of Texas at Austin, 1980.

[2] P. E. Bjørstad, *Multiplicative and Additive Schwarz Methods: Convergence in the 2 domain case*, in Domain Decomposition Methods, T. F. Chan, R. Glowinski, J. Périaux, and O. Widlund, eds., Philadelphia, 1989, SIAM.

[3] D. Braess, *The contraction number of a multigrid method for solving the Poisson equation*, Numerische Mathematik, 37 (1981), pp. 387–404.

[4] T. F. Chan, R. Glowinski, J. Périaux, and O. Widlund, eds., *Domain Decomposition Methods for Partial Differential Equations II*, Philadelphia, 1989, SIAM. Proceedings of the Second International Symposium on Domain Decomposition Methods , Los Angeles, California , January 14 - 16, 1988.

[5] C. C. Douglas and J. Mandel, *The domain reduction method: High way reduction in three dimensions and convergence with inexact solvers*, in Multigrid Methods, J. Mandel, S. F. McCormick, J. J. E. Dendy, C. Farhat, G. Lonsdale, S. V. Parter, J. W. Ruge, and K. Stüben, eds., Philadelphia, in print, SIAM. (Proceedings of the Fourth Copper Mountain Conference on Multigrid Methods, Copper Mountain, April 1989.).

[6] C. C. Douglas and W. L. Miranker, *Constructive interference in parallel algorithms*, SIAM J. Numer. Anal., 25 (1988), pp. 376–398.

[7] C. C. Douglas and B. F. Smith, *Using symmetries and antisymmetries to analyze a parallel multigrid algorithm: the elliptic boundary value problem case*, tech. report, IBM, 1987. Submitted to SIAM J. Num. Anal.

[8] M. Dryja, *An additive Schwarz algorithm for two- and three- dimensional finite element elliptic problems*, in Domain Decomposition Methods, T. F. Chan, R. Glowinski, J. Périaux, and O. Widlund, eds., Philadelphia, 1989, SIAM.

[9] M. Dryja and O. B. Widlund, *An additive variant of the Schwarz alternating method for the case of many subregions*, Tech. Report 339, also Ultracomputer Note 131, Department of Computer Science, Courant Institute, 1987.

[10] ——, *On the optimality of an additive iterative refinement method*, in Multigrid Methods, J. Mandel, S. McCormick, J. J.E. Dendy, C. Farhat, G. Lonsdale, S. V. Parter, J. W. Ruge,

and K. Stüben, eds., Philadelphia, 1989, SIAM. Proceedings of the Fourth Copper Mountain Conference on Multigrid Methods, April 9-14, 1989.

[11] ——, *Some domain decomposition algorithms for elliptic problems*, Tech. Report 438 also Ultracomputer Note 155, Department of Computer Science, Courant Institute, 1989. Proceedings of the Conference on Iterative Methods for Large Linear Systems held in Austin, Texas, October 19 - 21, 1988, to celebrate the sixty-fifth birthday of David M. Young, Jr.

[12] R. Glowinski, G. H. Golub, G. A. Meurant, and J. Périaux, eds., *Domain Decomposition Methods for Partial Differential Equations*, Philadelphia, 1988, SIAM. Proceedings of the First International Symposium on Domain Decomposition Methods for Partial Differential Equations, Paris, France, January 1987.

[13] G. H. Golub and C. F.Van Loan, *Matrix Computations*, Johns Hopkins Univ. Press, 1989. Second Edition.

[14] W. Hackbusch, *A new approach to robust multi-grid solvers*, in ICIAM '87, Philadelphia, 1988, SIAM, pp. 114–126. Proceedings of the First International Conference on Industrial and Applied Mathematics (Paris, 1987).

[15] 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, 1988, SIAM.

[16] ——, *On the Schwarz alternating method. II.*, in Domain Decomposition Methods, T. F. Chan, R. Glowinski, J. Périaux, and O. Widlund, eds., Philadelphia, 1989, SIAM.

[17] J. Mandel and S. McCormick, *Iterative solution of elliptic equations with refinement: The model multi-level case*, in Domain Decomposition Methods, T. F. Chan, R. Glowinski, J. Périaux, and O. Widlund, eds., Philadelphia, 1989, SIAM.

[18] ——, *Iterative solution of elliptic equations with refinement: The two-level case*, in Domain Decomposition Methods, T. F. Chan, R. Glowinski, J. Périaux, and O. Widlund, eds., Philadelphia, 1989, SIAM.

[19] A. M. Matsokin and S. V. Nepomnyaschikh, *A Schwarz alternating method in a subspace*, Soviet Mathematics, 29(10) (1985), pp. 78 – 84.

[20] H. A. Schwarz, *Gesammelte Mathematische Abhandlungen*, vol. 2, Springer, Berlin, 1890, pp. 133–143. First published in Vierteljahrsschrift der Naturforschenden Gesellschaft in Zürich, volume 15, 1870, pp.272–286.