SOME SCHWARZ METHODS FOR SYMMETRIC AND NONSYMMETRIC ELLIPTIC PROBLEMS

OLOF B. WIDLUND *

Abstract. This paper begins with an introduction to additive and multiplicative Schwarz methods. A two-level method is then reviewed and a new result on its rate of convergence is established for the case when the overlap is small. Recent results by Xuejun Zhang, on multi-level Schwarz methods, are formulated and discussed. The paper is concluded with a discussion of recent joint results with Xiao-Chuan Cai on nonsymmetric and indefinite problems.

Key Words. domain decomposition, Schwarz methods, finite elements, nonsymmetric and indefinite elliptic problems

AMS(MOS) subject classifications. 65F10, 65N30

1. Introduction.

Over the last few years, a general theory has been developed for the study of *additive* and *multiplicative* Schwarz methods. Many domain decomposition and certain multigrid methods can now be successfully analyzed inside this framework. Early work by P.-L. Lions [23], [24] gave an important impetus to this effort. The additive Schwarz methods were then developed by Dryja and Widlund [15], [16], [17], Matsokin and Nepomnyaschikh [25] and Nepomnyaschikh [26] and others. Recent efforts by Bramble, Pasciak, Wang and Xu [4] and Xu [36] have extended the general framework making a systematic study of multiplicative Schwarz methods possible. The multiplicative algorithms are direct generalizations of the original alternating method discovered more than 120 years ago by H. A. Schwarz [30]. We note that most of the work in recent years has focused on the positive definite, symmetric case.

While this theory is quite general, the applications so far have primarily been to the solution of the often large linear systems of algebraic equations, which arise in the finite element discretization of elliptic and parabolic boundary value problems. As shown in P.-L. Lions [24], the classical Schwarz algorithms can conveniently be described in terms of subspaces of the given space. The relevant error propagation operator of a particular Schwarz method can be written as a polynomial of orthogonal projections onto these subspaces. The use of these projections in computations involves the evaluation of the residual of the original finite element problem and the exact, or approximate, solution of several finite element problems on subregions. An additional coarse discrete model is also often used to enhance the rate of convergence. For a discussion of many applications, see Dryja and Widlund [17]. For other current projects, which also use the Schwarz framework, see Dryja and Widlund [18], [20] and Dryja, Smith and Widlund [14].

* Courant Institute of Mathematical Science, New York University, 251 Mercer Street, New York, N.Y. 10012. This work was supported in part by the National Science Foundation under Grant NSF-CCR-8903003 and in part by the U.S. Department of Energy under contract DE-FG02-88ER250553 at the Courant Mathematics and Computing Laboratory.

This paper is organized as follows. In Section 2, we first briefly review the theory for the positive definite, symmetric case and then, to provide an example, turn to a special case already discussed in our first papers on Schwarz methods, cf. Dryja and Widlund [15], [16]. This additive method uses two levels of discretization corresponding to a coarse space and a number of local finite element spaces. The former plays a role similar to the coarse mesh solver in a multigrid method; it provides global transportation of information across the region.

In our original proof that the condition number for this Schwarz method is independent of the number of subspaces and the mesh size, it is important to assume that there is a relatively generous overlap between the subregions. However, numerical experiments have shown that such methods often perform quite satisfactorily with little overlap, cf. Bjørstad, Moe and Skogen [1], Bjørstad and Skogen [2], Cai [7], [8] and Cai, Gropp and Keyes [9]. In Section 3, we prove a theorem on additive Schwarz methods with minimal overlap to help provide an explanation. This result has recently been obtained in joint work with Maksymilian Dryja.

In Section 4, we describe a recent result by Xuejun Zhang [37], a Courant Institute student, on multilevel additive Schwarz methods. He has shown that the condition number of a family of such methods grows at most linearly with the number of levels. The parallel multilevel preconditioner of Bramble, Pasciak and Xu [5] belongs to this family and, consequently, the estimate of the condition number of the BPX operator can be improved. For earlier results and a general discussion of multilevel methods, see Dryja and Widlund [19] and Xu [35].

In Section 5, we discuss recent joint work with Xiao-Chuan Cai on nonsymmetric and indefinite elliptic problems. Cai's dissertation [7] was a pioneering study on nonsymmetric problems. He discovered that the use a coarse model is even more important than in the positive definite, symmetric case. Thus, already in the nonsymmetric, positive definite case, the use of a sufficiently fine coarse mesh model is required in order for the additive Schwarz operator to have a positive definite, symmetric part. Confining the spectrum of the additive Schwarz operator to the right half plane allows us to derive strong results on the convergence rate of GMRES, a conjugate gradient-like method, which is effective for many nonsymmetric problems, cf. Saad and Schultz [28]. Generally, it is known that the performance of iterative methods of the conjugate gradient family suffers if the origin of the complex plane is surrounded by eigenvalues. The experimental evidence is also very clear; the performance is clearly enhanced in the presence of a coarse space with sufficiently many degrees of freedom, cf. Cai [7], Cai, Gropp and Keyes [9] and Cai and Widlund [11].

Our joint work concerns both additive and multiplicative algorithms, cf. Cai and Widlund [11], [10]. In the second paper, the theory, previously developed for multiplicative methods by Bramble et al. [4] for the positive definite, symmetric case, is generalized to a class of nonsymmetric and indefinite elliptic problems.

2. An introduction to Schwarz methods. As usual, we write our continuous and finite element elliptic problems as

$$a(u,v) = f(v),\ \forall\ v \in\ V\ ,$$

and

$$a(u_h,v_h) = f(v_h),\ \forall\ v_h \in\ V^h\ , \tag{1}$$

respectively. We assume, in the next few sections, that the problem is selfadjoint and elliptic in a suitably chosen space V and that $a(u,v)$ is bounded in $V \times V$. In the case of Poisson's equation, the bilinear form is defined by

$$a(u,v) = \int_\Omega \nabla u \cdot \nabla v\ dx\ .$$

To avoid unnecessary complications, we first confine our discussion to this operator, to homogeneous Dirichlet conditions and to continuous, piecewise linear finite elements.

In the classical formulation of Schwarz's alternating method, two overlapping subregions, Ω_1' and Ω_2' are used; the union of the two is Ω. There are two sequential, fractional steps in which the approximate solution of the elliptic equation on Ω is updated by solving the problem restricted to one of the subregions, one at a time. The most recent values of the solution are used as boundary values on the part of $\partial\Omega_i'$, the boundary of Ω_i', which is not a part of $\partial\Omega$.

The finite element version of the algorithm can conveniently be described in terms of projections $P_i : V^h \to V_i^h$, defined by

$$a(P_i v_h, \phi_h) = a(v_h, \phi_h),\ \forall \phi_h \in V_i^h\ . \tag{2}$$

Here V^h is the finite element space on Ω and $V_i^h = H_0^1(\Omega_i') \bigcap V^h$. It is easy to show that the error propagation operator of this multiplicative Schwarz method is

$$(I - P_2)(I - P_1).$$

This algorithm can therefore be viewed as a simple iterative method for solving

$$(P_1 + P_2 - P_2P_1)u_h = g_h,$$

with an appropriate right-hand side g_h.

This operator is a polynomial of degree two and thus not ideal for parallel computing since two sequential steps are involved. This effect is further pronounced if more than two subspaces are used even if the degree of the polynomial representing the multiplicative algorithm often is lower than maximal because the product of two projections associated with subregions, which do not overlap, vanishes.

We note that it is often advantageous to group subspaces, which do not intersect, into sets. The subspaces of each set can then be merged and the value of N reduced.

The decrease in the number of fractional steps makes the algorithm easier to parallelize. Numerical experiments also show that the convergence rate often is enhanced if the number of subspaces, created in this way, is small.

In the additive form of the algorithm, we work with the simplest possible polynomial of the projections: The equation

$$Pu_h = (P_1 + P_2 + \cdots + P_N)u_h = g_h' \ , \tag{3}$$

is solved by an iterative method. The choice of this polynomial can also be motivated by first introducing a relaxation parameter τ in the multiplicative algorithm, which, for the case of $N = 2$, results in the polynomial $\tau(P_1 + P_2 - \tau P_2 P_1)$. Cancelling the factor τ and letting τ decrease to zero, we again obtain the first order polynomial of equation (3).

Since the operator P is positive definite, symmetric, with respect to the bilinear form, the iterative method of choice is the conjugate gradient method. Equation (3) must have the same solution as equation (1), i.e. the correct right-hand side must be found. Since by equation (1), $a(u_h, \phi_h) = f(\phi_h)$, the right-hand side g_h' can be constructed by solving equation (2) for all values of i and adding the results. Similarly the operator P of equation (3) can be applied to any given element of V^h by applying each projection P_i to the element. Most of the work, in particular that which involves the individual projections, can be carried out in parallel.

We now describe the special additive Schwarz method introduced in Dryja and Widlund [15], cf. also [16] and Dryja [13]. We start by introducing two triangulations of Ω into nonoverlapping *substructures* Ω_i and *elements.* We obtain the elements by subdividing the substructures. We always assume that these triangulations are *shape regular*, cf. e.g. Ciarlet [12]. In this additive method, we use overlapping subregions obtained by extending each substructure to a larger region Ω_i'. We first assume that the overlap is generous with the distance δ_i between the boundaries $\partial\Omega_i$ and $\partial\Omega_i'$ bounded from below by a fixed fraction of H_i, the diameter of Ω_i. We also assume that $\partial\Omega_i'$ does not cut through any element. We make the same construction for the substructures that meet the boundary except that we cut off the part of Ω_i' that is outside of Ω.

Our finite element space is represented as the sum of N+1 subspaces

$$V^h = V_0^h + V_1^h + ... + V_N^h \ .$$

The first subspace V_0^h is equal to V^H, the space of continuous, piecewise linear functions on the coarse mesh defined by the substructures Ω_i. The other subspaces are related to the subdomains, in the same way as in a traditional Schwarz algorithm, i.e. $V_i^h = H_0^1(\Omega_i') \bigcap V^h$.

It is often more economical to use approximate rather than exact solvers of the problems on the subspaces. We use the following notations. $b_i(u, v)$ is an inner product defined on $V_i^h \times V_i^h$. We assume that there exists a constant ω such that

$$a(u, u) \leq \omega b_i(u, u) \ , \forall u \in V_i^h \ . \tag{4}$$

An operator $T_i : V^h \to V_i^h$, which replaces P_i, is defined by

$$b_i(T_i u, \phi_i) = a(u, \phi_h),\ \forall \phi_h \in V_i^h\ . \tag{5}$$

It is easy to show that the operator T_i is positive semidefinite and symmetric with respect to $a(\cdot,\cdot)$ and that the minimal constant ω in equation (4) is $||T_i||_a$.

In order to estimate the rate of convergence of our special additive Schwarz method, or any other, we need upper and lower bounds for the spectrum of the operator relevant in the conjugate gradient iteration. The technique used to obtain lower bounds is often associated with P.-L. Lions, cf. [24]. One form of this result is given in the following lemma.

LEMMA 1. *Let T_i be the operator defined in equation (5) and let $T = T_0 + T_1 + \cdots + T_N$. Then*

$$a(T^{-1}u, u) = \min_{u=\sum u_i} \sum b_i(u_i, u_i),\ u_i \in V_i\ .$$

Therefore, if a representation, $u = \sum u_i$, can be found, such that

$$\sum b_i(u_i, u_i) \le C_0^2 a(u,u), \quad \forall u \in V^h\ ,$$

then

$$\lambda_{min}(T) \ge C_0^{-2}.$$

The upper bound for the spectrum of T can often conveniently be obtained in terms of strengthened Cauchy-Schwarz inequalities between the different subspaces. Note that we now exclude the index 0; the coarse subspace is treated separately.

DEFINITION 1. *The matrix $\mathcal{E} = \{\varepsilon_{ij}\}$ is the matrix of strengthened Cauchy-Schwarz constants, i.e. ε_{ij} is the smallest constant for which*

$$|a(v_i, v_j)| \le \varepsilon_{ij} \|v_i\|_a \|v_j\|_a\ ,\forall v_i \in V_i\ ,\forall v_j \in V_j\ ,i,j \ge 1\ , \tag{6}$$

holds.

The following lemma is easy to prove.

LEMMA 2. *Let $\rho(\mathcal{E})$ be the spectral radius of the matrix $\mathcal{E}$. The operator T satisfies the following upper bound,*

$$T \le \omega(\rho(\mathcal{E}) + 1) I\ .$$

Combining these two lemmas, we obtain

THEOREM 1. *The condition number of the operator of the additive Schwarz method satisfies*

$$\kappa(T) \le \omega(\rho(\mathcal{E}) + 1) C_0^2\ .$$

In the special case considered in this section, it is easy to show that there is a uniform upper bound; see Section 3, where an estimate of C_0 is also given.

It is clear that C_0 does not increase if we expand (or add) individual subspaces. This follows from the fact that there is a larger choice in selecting $u_i \in V_i$. If we can expand the subspaces without worsing the upper bound, and that is often possible, our estimate of the condition number $\kappa(T)$ can improve. On the other hand, a larger subspace also means that the subproblems have more variables and that they are worse conditioned.

In the multiplicative case, we need to provide an upper bound for the spectral radius of the error propagation operator

$$E_J = (I - T_J) \cdots (I - T_0) \ . \tag{7}$$

The following theorem is a variant of the important results of Bramble, Pasciak, Wang and Xu [4] and Xu [36].

THEOREM 2. *In the symmetric, positive definite case,*

$$\rho(E_J) \leq \sqrt{1 - \frac{(2-\omega)}{(2\omega^2\rho(\mathcal{E})^2 + 1)C_0^2}} \ .$$

We note that this formula is useless if $\omega \geq 2$. Since $\|I - T_i\|_a > 1$ if $\|T_i\|_a > 2$, the assumption that $\omega < 2$ is most natural.

3. The case of a small overlap. As we have already pointed out, the arithmetic work grows with the overlap while, at the same time, the smallest eigenvalue of the operator P, or T, can increase, often leading to a decrease of the condition number. Numerical experiments suggest that it is often advantageous to use a minimal overlap, with a single layer of interior mesh points common to neighboring subregions, if we wish to minimize the running time of the algorithm rather than the number of iterations, cf. Bjørstad, Moe and Skogen [1], Bjørstad and Skogen [2], Cai [7], [8] and Cai, Gropp and Keyes [9]. Some experiments also indicate that for large values of H/h, the condition number decreases by about a factor of two when the overlap is doubled. All these results are for problems in the plane. For a discussion of a similar phenomenon with a related algorithm, see also Smith [33] or Chapter 4 of Smith [32]. It is important to note that in these experiments, and in the proof given below, the special coarse, global subspace is very important. Without it, these algorithms can be very slow.

We will give an estimate which helps explain the results of these experiments. Our proof is valid in two as well as three dimensions and this suggests that additive Schwarz methods with small overlap might produce satisfactory results even in three dimensions, in particular, if the value of H/h, and thus the number of degrees of freedom for each subproblem, is not enormous. We also note that parallel computing practitioners have expressed a clear preference for domain decomposition methods where the subregions only share one layer of points, cf. e.g. Smith [31], [34]. Iterative substructuring methods have that property but so do the additive Schwarz methods with minimal overlap. From now on, C will denote a generic constant. To simplify

the proof, we assume that the maximum width of $\Gamma_{\delta,i}$, the neighborhood inside the boundary $\partial\Omega_i$ that is in common with other subregions, is bounded by $C\delta_i$.

THEOREM 3. *In the case of exact solvers for the subproblems, the condition number of the additive Schwarz method satisfies*

$$\kappa(P) \leq C(1 + C \max_i \frac{H_i}{\delta_i}).$$

Here H_i is the diameter of the substructure Ω_i and δ_i the distance between the boundaries of Ω_i and Ω_i'. The constants are independent of the parameters H_i, h and δ_i.

Proof. The proof is a refinement of a result first given in Dryja and Widlund [15]; cf. [16] for a better discussion. We first note that a constant upper bound for the spectrum of P can be obtained as follows. For $i \geq 1$, P_i is also an orthogonal projection of $H^1(\Omega_i')$ onto V_i. Therefore,

$$a(P_i u_h, u_h) \leq a_{\Omega_i'}(u_h, u_h).$$

Since there is an upper bound, m, on the number of subregions to which any $x \in \Omega$ can belong, we have

$$\sum_{i=1}^{N} a_{\Omega_i'}(u_h, u_h) \leq m\, a(u_h, u_h).$$

In addition, we use the fact that the norm of P_0 is equal to one and obtain

$$\lambda_{max}(P) \leq (m+1)\ .$$

The lower bound is obtained by using Lemma 1. A natural choice of u_0 is the L_2-projection $Q_H u_h$ of u_h onto V^H. It is well known, cf. e.g. Bramble and Xu [6], that Q_H is bounded not only in L_2 but also in H^1 and that there exists a constant, independent of h and H, such that

$$\|Q_H u_h - u_h\|_{L_2} \leq C\, H \|u_h\|_a. \tag{8}$$

Let $w_h = u_h - Q_H u_h$ and let $u_i = I_h(\theta_i w_h)\,, i = 1, \cdots, N$. Here I_h is the interpolation operator onto the space V^h and the θ_i define a partition of unity. These functions can be chosen in many ways, in particular, as nonnegative elements of V^h. Then, all the values of θ_i are determined by those at the mesh points. In the interior part of Ω_i', which does not belong to any other subregion, $\theta_i \equiv 1$. This function must also drop to 0 over a distance on the order of δ_i. We denote the region, where there is overlap, by Γ_δ and we also use the notation $\Gamma_{\delta,i} = \Gamma_\delta \cap \Omega_i$. It is easy to construct a partition of unity with $0 \leq \theta_i \leq 1$ and

$$|\nabla \theta_i| \leq \frac{C}{\delta_i}\ .$$

In order to use Lemma 1, we now estimate $a(u_i, u_i)$ in terms of $a(w_h, w_h)$. We consider the contribution from one substructure at a time and note that, trivially,

$$a_{\Omega_i \setminus \Gamma_{\delta,i}}(u_i, u_i) = a_{\Omega_i \setminus \Gamma_{\delta,i}}(w_h, w_h)\ .$$

Let K be an element in $\Gamma_{\delta,i}$. Then, using the definition of u_i,

$$a_K(u_i,u_i) \leq 2a_K(\bar{\theta}_i w_h, \bar{\theta}_i w_h) + 2a_K(I_h((\theta_i - \bar{\theta}_i)w_h), I_h((\theta_i - \bar{\theta}_i)w_h)) \ ,$$

where $\bar{\theta}_i$ is the average of θ_i over the element K. Using the fact that the diameter of K is on the order of h and the bound on $\nabla\theta_i$, we obtain

$$a_K(u_i,u_i) \leq 2a_K(w_h,w_h) + \frac{C}{\delta_i^2}\|w_h\|^2_{L^2(K)} \ .$$

To complete the proof, we need to estimate $\|w_h\|^2_{L^2(\Gamma_{\delta,i})}$. This estimate is based on the following lemma.

LEMMA 3.

$$\|u\|^2_{L^2(\Gamma_{\delta,i})} \leq C\,\delta_i^2((1+\frac{H_i}{\delta_i})a_{\Omega_i}(u,u) + \frac{1}{H_i\delta_i}\|u\|^2_{L^2(\Omega_i)}) \ .$$

Proof. All the essential ideas of the proof can be illustrated by considering a square region $(0,H)\times(0,H)$. Then, since

$$u(x,0) = u(x,y) - \int_0^y \frac{\partial u(x,\tau)}{\partial y} d\tau \ ,$$

we find, after some manipulations, that

$$H\int_0^H |u(x,0)|^2 dx \leq 2\int_0^H\int_0^H |u(x,y)|^2 dxdy + H^2\int_0^H\int_0^H |\frac{\partial u}{\partial y}|^2 dxdy \ .$$

Therefore,

$$H\int_0^H |u(x,0)|^2 dx \leq 2\|u\|^2_{L^2(\Omega_i)} + H^2 a_{\Omega_i}(u,u) \ .$$

Now consider the integral over a narrow subregion next to one of the sides of the square. Using similar arguments, we obtain

$$\int_0^H\int_0^\delta |u(x,y)|^2 dxdy \leq \delta^2 a_{\Omega_i}(u,u) + 2\delta\int_0^H |u(x,0)|^2 dx \ .$$

By combining this with the previous inequality, we obtain

$$\int_0^H\int_0^\delta |u(x,y)|^2 dxdy \leq \delta^2 a_{\Omega_i}(u,u) + 2\delta(\frac{2}{H}\|u\|^2_{L^2(\Omega_i)} + H a_{\Omega_i}(u,u)) \ ,$$

as required.

The modifications necessary to cover the case of an arbitrary, shape regular substructure are routine. ∎

We now apply Lemma 3 to the function w_h and use inequality (8) to complete the proof of the Theorem 1. We also note that there are no difficulties extending the arguments to three dimensions. ∎

We note that for the case of two subregions, it is easy to show that the result of this theorem cannot be improved.

4. Multi-level additive Schwarz methods. The method, described in the two previous sections, uses two levels corresponding to a coarse and a fine mesh and a decomposition of Ω into overlapping subregions. If H is small, the local problems are small but we then have to contend with a relatively large linear system corresponding to $V_0^h = V^H$. It is natural to try to decrease the cost by replacing the exact solver on the coarse mesh by using the algorithm recursively. This introduces additional levels of triangulation.

Let us consider $\ell+1$ rather than two levels of triangulation of Ω with substructures Ω_{ki} and discretization parameters $h_k, k = 0, \cdots, \ell$. We assume that the triangulation on level k is a refinement of that of level $k-1$. On each level, we define a finite element space V^{h_k} with $V^{h_\ell} = V^h$. Each V^{h_k} is represented, as in the two-level case, by

$$V^{h_k} = V_{k1} + \cdots + V_{kN_k}.$$

The subspace V_{ki} is associated with a subregion Ω'_{ki} obtained by extending Ω_{ki} as in Section 2. Thus, $V_{ki} = H_0^1(\Omega'_{ki}) \bigcap V^{h_k}$, extended by zero outside of $\partial\Omega'_{ki}$. The entire space can then be represented as

$$V^h = V^{h_0} + \sum_{k=1}^{\ell} \sum_{i=1}^{N_k} V_{ki} \ . \tag{9}$$

The projections $P_{ki} : V^h \to V_{ki}$, are defined as before by

$$a(P_{ki}u_h, \phi_h) = a(u_h, \phi_h), \forall \phi_h \in V_{ki} \ ,$$

and the original problem is replaced by

$$Pu_h = (P_0 + \sum_{k=1}^{\ell} \sum_{i=1}^{N_k} P_{ki})u_h = g_h \ .$$

The right hand side is constructed as in Section 2.

We note that we are now able to design methods for which there is fixed upper bound on the dimension of the subspaces; when h decreases, we can gradually increase ℓ keeping the size of the subproblems uniformly bounded. In this case, the work per iteration step grows only linearly with the number of degrees of freedom of the whole problem. The condition number of the stiffness matrices of the small problems, corresponding to individual subspaces, are uniformly bounded and these matrices can be replaced by diagonal matrices without affecting the condition number of the iteration operator by more than a constant factor.

Dryja and Widlund [19], have previously established that the condition number of P grows at most quadratically with ℓ, the number of levels. Recently, Xuejun Zhang [37] has established a constant upper bound for the eigenvalues of P and obtained:

THEOREM 4. *There exist constants γ_0 and γ_1, independent of the mesh sizes and ℓ, the number of levels, such that the operator P defined by the decomposition (9), satisfies*

$$\gamma_0(\ell+1)^{-1}a(u_h, u_h) \leq a(Pu_h, u_h) \leq \gamma_1 a(u_h, u_h) \ .$$

In the case of H^2-regularity, e.g. if the region is convex, there is also a constant lower bound.

Zhang also notes that the case, where the refinement from each level to the next is by a factor two and all the subspaces $V_{ki}, k \geq 1$, are of dimension one, is covered by his theory. The resulting algorithm can also be derived from the *BPX algorithm*, see Bramble, Pasciak and Xu [5], by using a simple diagonal scaling. As a consequence, Zhang has improved the bound for the BPX algorithm. Previously only a $C\ell^2$ bound was known, in the absence of an extra regularity assumption.

5. Nonsymmetric and indefinite problems. Let Ω be an open, bounded polygonal region in $R^d, d = 2$ or 3, with boundary $\partial\Omega$. Consider the homogeneous Dirichlet boundary value problem:

$$\left\{ \begin{array}{rcll} Lu & = & f & \text{in} \quad \Omega, \\ u & = & 0 & \text{on} \quad \partial\Omega. \end{array} \right. \tag{10}$$

The elliptic operator L has the form

$$Lu(x) = -\sum_{i,j=1}^{d} \frac{\partial}{\partial x_i}(a_{ij}(x)\frac{\partial u(x)}{\partial x_j}) + 2\sum_{i=1}^{d} b_i(x)\frac{\partial u(x)}{\partial x_i} + c(x)u(x).$$

All the coefficients are, by assumption, sufficiently smooth and the matrix $\{a_{ij}(x)\}$ is symmetric and uniformly positive definite for $x \in \Omega$. The right hand side $f \in L^2(\Omega)$. We also assume that the equation has a unique solution in $H_0^1(\Omega)$.

The weak form of equation (10) is: Find $u \in H_0^1(\Omega)$ such that

$$b(u,v) = f(v), \qquad \forall v \in H_0^1(\Omega). \tag{11}$$

The bilinear form $b(u,v)$ is defined by

$$b(u,v) \;=\; \sum_{i,j=1}^{d} \int_\Omega a_{ij}\frac{\partial u}{\partial x_j}\frac{\partial v}{\partial x_i}dx + \sum_{i=1}^{d} 2\int_\Omega b_i\frac{\partial u}{\partial x_i}vdx + \int_\Omega cuvdx$$

or

$$b(u,v) \;=\; \sum_{i,j=1}^{d} \int_\Omega a_{ij}\frac{\partial u}{\partial x_j}\frac{\partial v}{\partial x_i}dx + \sum_{i=1}^{d} \int_\Omega b_i\frac{\partial u}{\partial x_i}v + \frac{\partial(b_i u)}{\partial x_i}vdx \;+ \int_\Omega \tilde{c}uvdx.$$

Here, $\tilde{c}(x) = c(x) - \sum_{i=1}^{d} \partial b_i(x)/\partial x_i$.

We also use two other bilinear forms

$$a(u,v) \;=\; \sum_{i,j=1}^{d} \int_\Omega a_{ij}\frac{\partial u}{\partial x_j}\frac{\partial v}{\partial x_i}dx$$

and

$$s(u,v) \;=\; \sum_{i=1}^{d} \int_\Omega b_i\frac{\partial u}{\partial x_i}v + \frac{\partial(b_i u)}{\partial x_i}vdx \;,$$

which correspond to the second order terms and the skew-symmetric part of L, respectively. The bilinear form $a(\cdot,\cdot)$ defines a norm, which we denote by $\|\cdot\|_a$. Under the given assumptions on the coefficients a_{ij}, this norm is equivalent to the H_0^1 norm. Symmetry is always with respect to $a(\cdot,\cdot)$ and the adjoint S^T of an operator S is defined by $a(S^T u, v) = a(u, Sv)$. It is also easy to verify that

$$s(u,v) \;=\; -s(v,u), \quad \forall u, v \in H_0^1(\Omega).$$

Using elementary, standard tools, it is easy to establish the following inequalities:

(i) $\mid b(u,v) \mid \leq C\|u\|_a\|v\|_a, \quad \forall u, v \in H_0^1(\Omega).$

(ii) Gårding's inequality: There exists a constant C, such that

$$\|u\|_a^2 - C\|u\|_{L^2(\Omega)}^2 \leq b(u,u), \quad \forall u \in H_0^1(\Omega)\ .$$

(iii) There exists a constant C, such that

$$\mid s(u,v) \mid \leq C\|u\|_a\|v\|_{L^2(\Omega)}, \quad \forall u, v \in H_0^1(\Omega),$$

$$\mid s(u,v) \mid \leq C\|v\|_a\|u\|_{L^2(\Omega)}, \quad \forall u, v \in H_0^1(\Omega).$$

We note that the bounds for $b(\cdot,\cdot)$ and $s(\cdot,\cdot)$ are different, since each term in $s(\cdot,\cdot)$ contains a factor, which is of zero order. This enables us to control the skew-symmetric term and makes our analysis possible; see further discussion at the end of this section.

We also use the following regularity result, cf. Grisvard [22] or Nečas [27].

(iv) The solution w of the adjoint equation

$$b(\phi, w) \;=\; g(\phi), \quad \forall \phi \in H_0^1(\Omega)$$

satisfies

$$\|w\|_{H^{1+\gamma}(\Omega)} \leq C\|g\|_{L^2(\Omega)}\ ,$$

where γ depends on the interior angles of $\partial\Omega$, is independent of g and is at least $1/2$.

We approximate equation (11) by a Galerkin conforming finite element method as in (1). For simplicity, we consider only continuous, piecewise linear, triangular elements in R^2 and tetrahedral elements in R^3.

We consider the same additive Schwarz algorithm, which were introduced previously in this paper, and also its multiplicative counterpart. Our results have also been established for several other classes of algorithms, cf. Cai and Widlund [11]. There

is also a more abstract theory, which we will describe briefly without providing any proofs.

We introduce operators which can replace the projections P_i and the operators T_i used previously. They are no longer symmetric.

Definition: *For $i = 0, \cdots, N$:*

For any $w^h \in V^h$, $P_i w^h \in V_i^h$ is the solution of the finite element equation

$$b(P_i w^h, v_i^h) \;=\; b(w^h, v_i^h), \quad \forall v_i^h \in V_i^h.$$

For any $w^h \in V^h$, $T_i w^h \in V_i^h$ is the solution of the finite element equation

$$a(T_i w^h, v_i^h) \;=\; b(w^h, v_i^h), \quad \forall v_i^h \in V_i^h.$$

Other choices for the operators T_i are also possible. We introduce operators which define two possible additive Schwarz methods

$$P \;=\; P_0 \;+\; P_1 \;+ \cdots +\; P_N$$

and

$$T \;=\; P_0 \;+\; T_1 \;+ \cdots +\; T_N.$$

The only difference between P and T is that, for $i > 0$, we replace the operator P_i, corresponding to Ω_i', by T_i. The coarse mesh operator is not changed; it should always correspond to the exact solution of the original problem in the coarse space.

The problems that we are considering are always nonsymmetric and we therefore have to replace the conjugate gradient method by a more appropriate iterative method. In our experiments with additive methods, we have always used the GMRES method, cf. Saad and Schultz [28], and Eisenstat, Elman and Schultz [21]. This is a generalized minimum residual method that in practice has proven quite powerful for a large class of nonsymmetric problems. The method converges satisfactorily for positive definite problems and a bound on the decay of the norm of the residual in n iterations is given by $(1 - c_p^2/C_p^2)^{n/2}$. Here $c_p = \inf_{x \neq 0} a(x, Tx)/a(x, x)$ and $C_p = \sup_{x \neq 0} \|Tx\|_a / \|x\|_a$.

In our discussion of the symmetric, positive definite case, we have found that the analysis of both additive and multiplicative cases requires estimates of the constant C_0, the spectral radius of the matrix $\mathcal{E}$ and the norms of the operators T_i. In the nonsymmetric case, we must do additional work.

For the multiplicative algorithms, we wish to estimate the spectral radius of the error propagation operator E_J given in formula (7). Following Bramble et al. [4], we begin by observing that, with $E_j = (I - T_j) \cdots (I - T_1)(I - T_0)$, $E_{-1} = I$, and $R_j = T_j + T_j^T - T_j^T T_j$,

$$E_j^T E_j - E_{j+1}^T E_{j+1} = E_j^T R_{j+1} E_j.$$

This leads to the identity

$$I - E_J^T E_J = R_0 + \sum_{j=1}^{J} E_{j-1}^T R_j E_{j-1} \ . \tag{12}$$

To simplify the notations, we here use $T_0 = P_0$.

A satisfactory upper bound for $\rho(E_J)$ could therefore be obtained by showing that the operator on the right hand side of (12) is sufficiently positive definite. It therefore seems natural to assume that the operators R_i are positive semidefinite. It is easy to see that this is so if $T_i^T = T_i \geq 0$, but such an assumption on R_i can often not be established in our nonsymmetric applications. In the general case, we therefore make a different

ASSUMPTION 1. *There exist parameters δ_i, which can be made sufficiently small, and a constant $\gamma > 0$, such that*

$$R_i = T_i + T_i^T - T_i^T T_i \geq \gamma T_i^T T_i - \delta_i I \ . \tag{13}$$

We note that if we can bound $T_i + T_i^T + \delta_i I$ from below by some positive constant multiple of $T_i^T T_i$, then Assumption 1 is satisfied for αT_i for a sufficiently small α. It is well known that such a rescaling often is necessary to obtain convergence for nonsymmetric problems.

It follows easily from Assumption 1 that

$$\|T_i\|_a \leq \omega_i \equiv (1 + \sqrt{1 + \delta_i(1+\gamma)})/(1+\gamma) \leq \frac{2}{1+\gamma} + \frac{\delta_i}{2} \ . \tag{14}$$

To simplify our calculations and formulas, we always assume that $\omega = \max_i \omega_i \geq 1$. It is easy to show that $\|I - T_i\|_a \leq 1 + \frac{\delta_i}{2}$ and that for a small enough δ_i, $\|T_i\|_a < 2$. In the case when T_i is positive definite, symmetric, $\delta_i = 0$,

$$0 \leq T_i \leq \omega I \equiv 2/(1+\gamma) I \tag{15}$$

and

$$R_i \geq (2 - \omega) T_i \ . \tag{16}$$

In our discussion of the symmetric case and the additive algorithms, we have seen that we can often establish a lower bound on the operator $T = T_0 + \cdots T_J$ in terms of the constant C_0 of Lemma 1. By using similar tools, we can often establish

ASSUMPTION 2. *There exists a constant $\tilde{C}_0 > 0$, such that*

$$\sum_{i=0}^{J} T_i^T T_i \geq \tilde{C}_0^{-2} I \ . \tag{17}$$

We note that in the symmetric case, $P_i^2 = P_i$ and generally $T_i \geq T_i^2/2$, if $\|T_i\|_a < 2$. In this case, Assumption 2 is therefore closely related to the bound given in Lemma 1.

We can establish Assumption 2, for a sufficiently fine coarse mesh, by using only Lemma 1, Gårding's inequality, cf. inequality (ii) above, and a result of Schatz's [29]. We note that Schatz proved the existence of finite elment solutions for elliptic problems, such as those introduced in the beginning of this section, for a sufficiently small mesh size. (In the variant that we need here, we replace the approximate solution by the coarse mesh solution and the exact solution of the continuous problem by the finite element solution in V^h.) In the proof of the variant of Schatz's result, Gårding's inequality, (ii), and the regularity result, (iv), but not (iii) are used. Details are given for the special Schwarz method, considered in this paper, in Lemma 5 of Cai and Widlund [11]. That paper also contains results for Yserentant's method and iterative substructuring methods for problems in the plane.

If we combine Assumptions 1 and 2, we obtain the required positive lower bound for T that is required for the estimate of the rate of convergence of the GMRES algorithm provided that $\sum \delta_i$ is made sufficiently small, cf. Cai and Widlund [10] for details. In the applications this requirement is met by decreasing the coarse mesh size. The upper bound is given by

$$\| \sum_{i=0}^{J} T_i \|_a \leq 2(\rho(\mathcal{E}) + 1)/(1+\gamma) + 1/2 \sum_{i=0}^{J} \delta_i \ .$$

This bound is obtained quite straightforwardly.

For the multiplicative case the following result holds. A proof will be given in Cai and Widlund [10] together with several applications.

THEOREM 5. *In the general case, the multiplicative algorithm is convergent if*

$$\frac{\gamma}{(\omega^2 |\mathcal{E}|_{l_1}^2 + (\sum \delta_i)^2)\tilde{C}_0^2}$$

dominates

$$\sum_{j=0}^{J} \delta_j exp(\sum_{i=0}^{j-1} \delta_i)$$

by a sufficiently large constant factor. Under this assumption, there exists a constant $C > 0$, *such that*

$$\rho(E_J) \leq \sqrt{1 - \frac{C}{(\omega^2 |\mathcal{E}|_{l_1}^2 + (\sum \delta_i)^2)\tilde{C}_0^2}} \ .$$

The main effort in the application of this theory goes into the verification of Assumption 1. So far we have have only been able to do so by using the inequality (iii) and a perturbation argument. However, there are elliptic problems which satisfy Gårding's inequality but for which the skew-symmetric part is not a relatively compact perturbation. In such a case an inequality serving the same purpose as (iii) will not hold. For an example of elliptic problems of this type, see Bramble, Leyk and Pasciak [3] in which several interesting algorithms, which differ from ours, are considered.

Acknowledgement. The author is indebted to Maksymilian Dryja for many discussions and suggestions concerning this work.

REFERENCES

[1] P. E. Bjørstad, R. Moe, and M. Skogen, *Parallel domain decomposition and iterative refinement algorithms*, in Parallel Algorithms for PDEs, Proceedings of the 6th GAMM-Seminar held in Kiel, Germany, January 19–21, 1990, W. Hackbusch, ed., Braunschweig, Wiesbaden, 1990, Vieweg-Verlag.

[2] P. E. Bjørstad and M. Skogen, *Domain decomposition algorithms of Schwarz type, designed for massively parallel computers*, in Fifth International Symposium on Domain Decomposition Methods for Partial Differential Equations, T. F. Chan, D. E. Keyes, G. A. Meurant, J. S. Scroggs, and R. G. Voigt, eds., Philadelphia, PA, 1991, SIAM. To appear.

[3] J. H. Bramble, Z. Leyk, and J. E. Pasciak, *Iterative schemes for non-symmetric and indefinite elliptic boundary value problems*, tech. rep., Cornell University, 1991.

[4] J. H. Bramble, J. E. Pasciak, J. Wang, and J. Xu, *Convergence estimates for product iterative methods with applications to domain decomposition and multigrid*, tech. rep., Cornell University, 1990. Submitted to Math. Comp.

[5] J. H. Bramble, J. E. Pasciak, and J. Xu, *Parallel multilevel preconditioners*, Math. Comp., 55 (1990), pp. 1–22.

[6] J. H. Bramble and J. Xu, *Some estimates for a weighted L^2 projection*, tech. rep., Cornell University, 1990. Submitted to Math. Comp.

[7] X.-C. Cai, *Some domain decomposition algorithms for nonselfadjoint elliptic and parabolic partial differential equations*, Tech. Rep. 461, Computer Science Department, Courant Institute of Mathematical Sciences, September 1989. Courant Institute doctoral dissertation.

[8] ——, *An additive Schwarz algorithm for nonselfadjoint elliptic equations*, in Third International Symposium on Domain Decomposition Methods for Partial Differential Equations, T. Chan, R. Glowinski, , J. Périaux, and O. Widlund, eds., SIAM, Philadelphia, PA, 1990.

[9] X.-C. Cai, W. D. Gropp, and D. E. Keyes, *A comparison of some domain decomposition algorithms for nonsymmetric elliptic problems*, in Fifth International Symposium on Domain Decomposition Methods for Partial Differential Equations, T. F. Chan, D. E. Keyes, G. A. Meurant, J. S. Scroggs, and R. G. Voigt, eds., Philadelphia, PA, 1991, SIAM. To appear.

[10] X.-C. Cai and O. Widlund, *Multiplicative Schwarz algorithms for some nonsymmetric and indefinite problems*, tech. rep., Computer Science Department, Courant Institute of Mathematical Sciences, 1991. To appear.

[11] ——, *Domain decomposition algorithms for indefinite elliptic problems*, SIAM J. Sci. Statist. Comput., 13 (1992). To appear.

[12] P. G. Ciarlet, *The Finite Element Method for Elliptic Problems*, North-Holland, 1978.

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

[14] M. Dryja, B. F. Smith, and O. B. Widlund, *Schwarz analysis of iterative substructuring algorithms for problems in three dimensions*, tech. rep., Department of Computer Science, Courant Institute, 1991. To appear.

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

[16] ——, *Some domain decomposition algorithms for elliptic problems*, in Iterative Methods for Large Linear Systems, San Diego, California, 1989, Academic Press, pp. 273–291. Proceeding 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.

[17] ——, *Towards a unified theory of domain decomposition algorithms for elliptic problems*, in

Third International Symposium on Domain Decomposition Methods for Partial Differential Equations, held in Houston, Texas, March 20-22, 1989, T. Chan, R. Glowinski, , J. Périaux, and O. Widlund, eds., SIAM, Philadelphia, PA, 1990.

[18] ——, *Additive Schwarz methods for elliptic finite element problems in three dimensions*, in Fifth Conference on Domain Decomposition Methods for Partial Differential Equations, T. F. Chan, D. E. Keyes, G. A. Meurant, J. S. Scroggs, and R. G. Voigt, eds., Philadelphia, PA, 1991, SIAM. To appear.

[19] ——, *Multilevel additive methods for elliptic finite element problems*, in Parallel Algorithms for Partial Differential Equations, Proceedings of the Sixth GAMM-Seminar, Kiel, January 19–21, 1990, W. Hackbusch, ed., Braunschweig, Germany, 1991, Vieweg & Son.

[20] ——, *The Neumann-Neumann method as an additive Schwarz method for finite element elliptic problems*, tech. rep., Department of Computer Science, Courant Institute, 1991. In preparation.

[21] S. C. Eisenstat, H. C. Elman, and M. H. Schultz, *Variational iterative methods for nonsymmetric systems of linear equations*, SIAM J. Numer. Anal., 20 (2) (1983), pp. 345–357.

[22] P. Grisvard, *Elliptic problems in nonsmooth domains*, Pitman Publishing, Boston, 1985.

[23] P. L. Lions, *Interprétation stochastique de la méthode alternée de Schwarz*, C. R. Acad. Sci. Paris, 268 (1978), pp. 325–328.

[24] ——, *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.

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

[26] S. V. Nepomnyaschikh, *Domain Decomposition and Schwarz Methods in a Subspace for the Approximate Solution of Elliptic Boundary Value Problems*, PhD thesis, Computing Center of the Siberian Branch of the USSR Academy of Sciences, Novosibirsk, USSR, 1986.

[27] J. Nečas, *Les méthodes directes en théorie des équations elliptiques*, Academia, Prague, 1967.

[28] Y. Saad and M. H. Schultz, *GMRES: A generalized minimum residual algorithm for solving nonsymmetric linear systems*, SIAM J. Sci. Stat. Comp., 7 (1986), pp. 856–869.

[29] A. H. Schatz, *An observation concerning Ritz-Galerkin methods with indefinite bilinear forms*, Math. Comp., 28 (1974), pp. 959–962.

[30] 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.

[31] B. F. Smith, *A domain decomposition algorithm for elliptic problems in three dimensions*, Tech. Rep. 519, Department of Computer Science, Courant Institute, October 1990. To appear in Numer. Math.

[32] ——, *Domain Decomposition Algorithms for the Partial Differential Equations of Linear Elasticity*, PhD thesis, Courant Institute of Mathematical Sciences, September 1990. Tech. Rep. 517, Department of Computer Science, Courant Institute.

[33] ——, *An optimal domain decomposition preconditioner for the finite element solution of linear elasticity problems*, SIAM J. Sci. Stat. Comput., (1990). To appear.

[34] ——, *An iterative substructuring algorithm for problems in three dimensions*, in Fifth International Symposium on Domain Decomposition Methods for Partial Differential Equations, T. F. Chan, D. E. Keyes, G. A. Meurant, J. S. Scroggs, and R. G. Voigt, eds., Philadelphia, PA, 1991, SIAM. To appear.

[35] J. Xu, *Theory of Multilevel Methods*, PhD thesis, Cornell University, May 1989.

[36] ——, *Iterative methods by space decomposition and subspace correction*, tech. rep., Penn State University, University Park, PA, 1990. Submitted to SIAM Review.

[37] X. Zhang, *Multilevel additive Schwarz methods*, tech. rep., Courant Institute of Mathematical Sciences, Department of Computer Science, 1991. To appear.