Chapter 16

Some Domain Decomposition Algorithms for Elliptic Problems

MAKSYMILIAN DRYJA
Warsaw University

and

OLOF B. WIDLUND
New York University

Dedicated to David M. Young, Jr., on the occasion of his sixty-fifth birthday.

Abstract

We discuss domain decomposition methods with which the often very large linear systems of algebraic equations, arising when elliptic problems are discretized by finite differences or finite elements, can be solved with the aid of exact or approximate solvers for the same equations restricted to subregions. The interaction between the subregions, to enforce appropriate continuity requirements, is handled by an iterative method, often a preconditioned conjugate gradient method. Much of the work is local and can be carried out in parallel. We first explore how ideas from structural engineering computations naturally lead to certain matrix splittings. In preparation for the detailed design and analysis of related domain decomposition methods, we then consider the Schwarz alternating algorithm, discovered in 1869. That algorithm can conveniently be expressed in terms of certain projections. We develop these ideas further and discuss an interesting additive variant of the Schwarz method. This also leads to the development of a general framework, which already has proven quite useful in the study of a variety of domain decomposition methods and certain related algorithms. We demonstrate this by developing several algorithms and by showing how their rates of convergence can be estimated. One of them is a Schwarz-type method, for which the subregions overlap, while the others are so called iterative substructuring methods, where the subregions do not overlap. Compared to previous studies

of iterative substructuring methods, our proof is simpler and in one case it can be completed without using a finite element extension theorem. Such a theorem has, to our knowledge, always been used in the previous analysis in all but the very simplest cases.

1 Introduction

Domain decomposition methods have recently become an important focus of research on numerical methods for partial differential equations. They appear to offer the best promise for the parallel solution of the often very large systems of linear or nonlinear algebraic systems of equations which arise when the elliptic problems of elasticity, fluid dynamics and many other important applications are discretized by finite elements or finite differences. The rapidly growing interest in this field is reflected in a series of SIAM sponsored symposia and mini-symposia exclusively devoted to research in this area; see [9, 15].

The domain decomposition methods considered here can be regarded as *divide and conquer* algorithms. In each step of an iteration, the original discrete elliptic problem is solved on the subregions into which the original region has been divided. (Under favorable circumstances, an already existing code can be used for this purpose.) The interaction between the different parts of the region is handled by transferring suitable data across the interfaces, which are created by the subdivision. These data are generated in each iteration from the residual of the original or a derived system of equations. A conjugate gradient method is often used to accelerate the convergence. In addition to the local problems, some global, coarse problem must be incorporated into the preconditioner in order to obtain a fast rate of convergence in the case of many subregions. That is the case of primary interest in parallel computing research. We note that the need for a coarse, global part of the preconditioner is well known in multi-grid work. More formally, we have shown in our previous work that if in each iteration step of a domain decomposition algorithm information is only exchanged between neighboring substructures, the rate of convergence can be no better than if the conjugate gradient method is used, without any preconditioning, for the coarse model obtained by using the subregions as elements; cf. Widlund [29].

There has been considerable progress in this research area and a number of fast algorithms have been designed and studied for which the condition number of the iteration matrix is uniformly bounded or grows only in proportion to the expression $(1 + \log(H/h))^q, q = 2$ or 3, where H is the diameter of a typical subregion and h the diameter of a typical element into which the subregions are divided; see, e.g., Bramble, Pasciak and Schatz [5, 6, 7, 8] , Dryja [11], Dryja, Proskurowski and Widlund [13], Dryja and Widlund [14] and Widlund [29]. Most of the important results in this field have been developed in a finite element framework and we also adopt such an approach in this paper.

To assess the complexity of domain decomposition algorithms, traditional tools

must also be used to measure the cost of solving the different subproblems. If a parallel computing system is used, there are of course many additional aspects to estimating the overall cost of the computation. We provide no detailed discussion of such matters here. Instead, we focus on the design of preconditioners that give rapid rates of convergence while decreasing the amount of work per step, and on the development of a general framework inside which a variety of algorithms can be designed and analyzed. We note that several sets of careful numerical experiments with various domain decomposition algorithms have been carried out on parallel computers; cf. Keyes and Gropp [18, 19], who used a hypercube, and Greenbaum et al. [17], who used an ultra computer prototype built at the Courant Institute.

Domain decomposition methods can be classified according to whether the subregions (substructures) overlap or not. The algorithms which do not use any overlap are often called *iterative substructuring* methods while the others are called *Schwarz-type* methods. This is in recognition of the importance of ideas of structural engineering computing in the development of the former class of methods and of the fundamental contribution by H. A. Schwarz, who introduced his alternating method in 1869; cf. Schwarz [24]. The two classes of methods have much in common and there is a real promise that a unified theory can be developed. This paper is an effort in that direction. We are able to show that some interesting iterative substructuring methods can also be derived from a so called *additive variant* of the Schwarz algorithm. Technically this work also offers something new. One of the algorithms can be analyzed without using a so called *finite element extension theorem*. Such theorems have previously been central to the development of the theory; see Widlund [28] for a proof of such a theorem for general conforming finite element methods and a general discussion.

We note that in a recent paper, Bjørstad and Widlund [3] have returned to the two subregion case to show that a method, introduced by Chan and Resasco [10], is in fact the classical Schwarz's method accelerated by the conjugate gradient method. That study provides additional evidence that we need not strictly distinguish between methods which use overlapping subregions and those who do not.

In our development of the theory, a prominent role is played by subspaces and projections. Already fifty years ago Sobolev [25] showed that the classical Schwarz algorithm can be analyzed using a variational framework and more recently this approach has been further developed by Pierre-Louis Lions [20]. An additive Schwarz algorithms was first studied in Dryja and Widlund [12]. In this paper, we strengthen and simplify that result. We note that similar ideas have also been discussed by Matsokin and Nepomnyaschikh [22]. The general analytic framework that is now available for the study of these additive methods, has also proven very useful in our study of so called iterative refinement methods; cf. Widlund [32, 33].

A few years ago, much of the work on iterative substructuring methods was focused on elliptic problems defined on regions divided into two or a few subregions with separating curves (surfaces) which do not intersect; see e.g., Bjørstad and Widlund [2] and Widlund [28]. In this paper, we show that some of those algorithms

and results can be recycled and combined with an iterative substructuring algorithm derived by using subspaces and projections to obtain previously known algorithms and results. This approach highlights how preconditioners can be built from parts which are strictly local to an individual substructure, parts which involve interaction between pairs of neighboring substructures and a coarse global model with relatively few degrees of freedom.

This paper is organized as follows. In Sec. 2, we review some of the ideas of substructuring that are very important in the development of computational methods of structural engineering. This discussion naturally leads to matrix splittings, which provide preconditioners for the large linear systems of algebraic equations, which arises in finite element work. In Sec. 3, we discuss different Schwarz methods and some general tools for estimating their rates of convergence. In the concluding sections, we show how two types of domain decomposition algorithms can be analyzed by using relatively simple tools of mathematical and finite element analysis. While we can do a lot with linear algebra, we ultimately have to resort to tools of analysis in order to complete the proofs of our main results.

2 Substructures, Subspaces and Projections

The domain decomposition methods, considered in this paper, provide preconditioners of systems of equations representing discrete elliptic problems. Preconditioners can be associated with matrix splittings but we note that those considered here differ in certain respects from most of those encountered the standard literature on iterative methods; cf Varga [27]. Domain decomposition methods can best be understood in terms of the substructuring ideas, which have provided a very successful framework for the development of very large programming systems for structural mechanics computing; cf. Bell, Hatlestad, Hansteen and Araldsen [1] or Przemieniecki [23].

We consider a linear, self adjoint, elliptic problem, which is discretized by a finite element method on a bounded Lipschitz region. The region Ω is a subset of $I\!R^n$, n=2 or 3, the differential operator is the Laplacian, and zero Dirichlet conditions and continuous, piecewise linear finite elements are used. The theory could equally well be developed for much more general linear elliptic problems, which can be formulated as minimization problems. Arbitrary conforming finite elements could also be considered without further major complications. Nonconforming finite elements, non-self adjoint problems and problems that give rise to indefinite symmetric systems of equations are also quite important. Some progress has already been made in such cases.

In the model case considered here, the continuous and discrete problems take the form

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

and

$$a(u_h, v_h) = f(v_h), \ \forall \ v_h \in \ V^h \ , \tag{2.1}$$

respectively. The bilinear form is defined by

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

This form defines a semi-norm $|u|_{H^1(\Omega)} = (a(u,u))^{1/2}$ in the Sobolev space $H^1(\Omega)$. It is a norm for $V = H_0^1(\Omega)$. Here $H_0^1(\Omega)$ is the subspace of $H^1(\Omega)$ functions with zero trace; all elements of V and its subspace V^h vanish on $\partial\Omega$, the boundary of Ω. The triangulation of Ω is introduced in the following way. The region is divided into non-overlapping substructures $\Omega_i \ , i = 1, 2, \ldots, N$. To simplify the description, our study is confined to triangular (simplicial) substructures. In such a case, the original region must of course be a polygon (polyhedron). All the substructures Ω_i are further divided into elements. The common assumption in finite element theory that all elements are shape regular is adopted and the same assumption is made concerning the substructures. On the element level this means that there is a bound on h_K/ρ_K, which is independent of the number of degrees of freedom and of K. Here h_K is the diameter of the element K and ρ_K the diameter of the largest sphere that can be inscribed in K.

Since $a(u_h, v_h) = a(u, v_h), \ \forall \ v_h \in \ V^h$, the finite element solution is the projection of the exact solution onto the finite element space with respect to the inner product defined by the bilinear form. We will see that the problems defined on the subregions, from which preconditioners for the entire problem can be assembled, can similarly be viewed in terms of orthogonal projections onto subspaces directly associated with the subregion in question.

By using ideas of structural engineering, the stiffness matrix K can be constructed in the following way. The elements of K are given by

$$k_{i,j} = a(\varphi_i, \varphi_j) \ ,$$

where φ_i and φ_j are standard finite element basis functions. Since an integral over Ω can be written as a sum of integrals over the substructures, the stiffness matrix can be assembled from the stiffness matrices $K^{(k)}$ the elements of which are defined by

$$k_{i,j}^{(k)} = a_{\Omega_k}(\varphi_i, \varphi_j) \ ,$$

with i and j corresponding to the degrees of freedom of the substructure Ω_k. The form $a_{\Omega_k}(u_h, v_h)$ represents the contribution to the integral from that substructure.

This so called subassembly process can be summarized in the formula

$$x^T K y = \sum_i {x^{(i)}}^T K^{(i)} y^{(i)} \ ,$$

where $x^{(i)}$ is the subvector of parameter values associated with the substructure Ω_i and its boundary $\partial\Omega_i$.

In engineering computing practice, the large linear system with the coefficient matrix K is solved using programs, which are elaborate implementations of block Gaussian elimination; cf., e.g., [1]. Particular attention is placed on the use of data structures, etc., which allows for efficient I/O, and on the need to integrate the factorization and solution steps as much as possible with the generation of the blocks which together define the stiffness matrix. By the so-called *subassembly process*, described above in a special case, the contributions from the individual elements are first computed and these element stiffness matrices are then merged with their neighbors, creating so-called *super elements*. This process is often used recursively a substantial number of times. On each level, the variables associated with the formation of a particular super element can be divided into two sets; those which are common to other super elements and the interior variables which are not. In our description of domain decomposition algorithms only three levels are considered, the elements, with a characteristic diameter h, the substructures with a diameter H and the entire region Ω, which, without loss of generality, is assumed to have unit diameter.

The last phases of a standard engineering calculation of this type involve a substantial fraction of the arithmetical work and require more global communication than during the previous phases, where the work can proceed in a distributed fashion, without synchronization between the substructures. The iterative substructuring methods require synchronization once every iteration, but the best of them converge quite fast. They also demand much less communication of data between the substructures than methods which use direct methods throughout. The use of inexact solvers for the subproblems can also be considered when an iterative substructuring method is used. If a substructure has a simple geometry, special fast solvers may also be used.

If we divide the subvectors $x^{(i)}$ associated with the i-th substructure into two, $x_B^{(i)}$ and $x_I^{(i)}$, corresponding to the variables shared with other substructures and those which are interior to the substructure, then the matrix $K^{(i)}$ can be written as

$$\begin{pmatrix} K_{II}^{(i)} & K_{IB}^{(i)} \\ K_{IB}^{(i)T} & K_{BB}^{(i)} \end{pmatrix}.$$

Since the interior variables are associated with only one of the substructures, they can be eliminated locally and in parallel. The reduced matrix is a so-called *Schur complement* and has the form

$$S^{(i)} = K_{BB}^{(i)} - {K_{IB}^{(i)}}^T {K_{II}^{(i)}}^{-1} K_{IB}^{(i)}.$$

It is now easy to show that if the corresponding Schur complement of the global stiffness matrix K is denoted by S, then

$$x^T S y = \sum_i {x_B^{(i)}}^T S^{(i)} y_B^{(i)}. \tag{2.2}$$

The elimination of the interior variables from the substructures can be viewed in terms of orthogonal projections, with respect to the bilinear form, of the solution u_h of Eq. (2.1) onto the subspaces $H_0^1(\Omega_i) \bigcap V^h$, $i = 1, 2, \ldots, N$. These subspaces can easily be shown to be orthogonal, in the sense of the bilinear form, to the so-called *piecewise discrete harmonic functions* which satisfy

$$K_{II}^{(i)} x_I^{(i)} + K_{IB}^{(i)} x_B^{(i)} = 0, \ \forall i.$$

If the local problems are solved exactly, what remains is to find a sufficiently accurate approximation of the part of the solution which is piecewise discrete harmonic. This is done by approximately solving the reduced linear system with the matrix S. An iterative substructuring method is obtained by selecting a preconditioner for the the matrix S. Once an approximation of the solution has been found on the boundaries of the substructures, the solution can be found everywhere by solving local Dirichlet problems on each substructure separately.

The matrices $S^{(i)}$ are full. Complete information on the sparsity and block structure of S can be obtained by using Eq. (2.2). It is also natural to partition the vector $x_B^{(i)}$ and the Schur complement $S^{(i)}$ further. We group the variables associated with the interior of each edge (face) of the substructure into separate subvectors. Each of them is associated with exactly two substructures which are neighbors of each other. The remaining degrees of freedom of $x_B^{(i)}$ form a separate subvector. In two dimensions, the components of this vector, $x_V^{(i)}$, are the values at the vertices of Ω_i, while in three dimensions they are associated with all the nodes of the edges of the simplex, including its vertices. We can say that these are the nodes of the wire basket outlining the substructure. The Schur complement $S^{(i)}$ is thus represented by a four by four (five by five) block matrix. The blocks correspond to individual edges (faces) and the set of vertices the *wire basket*. If one or several edges of Ω_i is part of $\partial\Omega$, and thus not associated with any degrees of freedom in the Dirichlet case, then the numbers of blocks of $S^{(i)}$ is reduced. In what follows, we consider splittings where the off-diagonal blocks of $S^{(i)}$ are dropped, and others where all the contributions from certain substructures are left out in the construction of the preconditioner.

In the case of two substructures, there is only one set of interface variables and S is the sum of $S^{(1)}$ and $S^{(2)}$, the Schur complements originating from the two subregions respectively. The basic *Neumann-Dirichlet algorithm* amounts to using $S^{(1)}$ or $S^{(2)}$ as a preconditioner for S; cf. Bjørstad and Widlund [2] and Widlund [28]. This algorithm can be extended directly to the case of many substructures if a red-black ordering of the substructures exists; cf. Dryja, Proskurowski and Widlund [13] and Widlund [29]. This splitting corresponds to dropping all contributions from the red substructures in the sum given in Eq. (2.2). The resulting preconditioner corresponds to a global problem since the remaining black substructures are connected at their vertices (wire baskets). If we group all edge (face) variables and all the vertex (wire basket) variables into two subvectors $x_E^{(i)}$ and $x_V^{(i)}$, then we can

write $S^{(i)}$ as

$$\begin{pmatrix} S_{EE}^{(i)} & S_{EV}^{(i)} \\ S_{EV}^{(i)T} & S_{VV}^{(i)} \end{pmatrix}.$$

In this quadratic form, which represents the sum over the black substructures only, the variables of $x_E^{(i)}$ appear only once. They are therefore coupled only to variables which are associated with the same substructure. The entire system of linear equations, which correspond to the Neumann-Dirichlet preconditioner, can, after a permutation, therefore be written in terms of a two-by-two block matrix where the leading block is the direct sum of the $S_{EE}^{(i)}$ which correspond to the black substructures. An analysis of the structure of this large matrix reveals that it is quite economical to use standard block Gaussian elimination, in the case of two dimensions. The Schur complement that remains after the elimination of the edge variables is sparse and its dimension is equal to the number of variables associated with substructure vertices which do not fall on $\partial\Omega$; see further Dryja, Proskurowski and Widlund [13]. It is shown in Widlund [29], that the condition number of the resulting preconditioner grows in proportion to $(1 + \log(H/h))^2$. The algorithm is less attractive in the three dimensional case, since the number of variables, which have to be treated in a special way, are relatively more numerous than for a problem in the plane. However, the same kind of bound can be established on the rate of convergence. There is an interesting variant of the algorithm, where the wire basket variables are handled differently and the preconditioner is cheaper to use; see Dryja [11]. We will not discuss it further in this paper.

Other splittings of S are known, which give as good (or better) results as the Neumann-Dirichlet method. We note that when we considered that algorithm, the values at the vertices (on the wire basket) were treated differently from those of the interior of the edges (faces). The methods considered by Bramble, Pasciak and Schatz [5, 8] also treat these sets of variables differently. As was shown in Widlund [29, 30], the part of the preconditioner introduced in [5] that relates to the vertex variables of plane problems can be viewed in terms of a projection onto a finite element subspace, where the substructures play the role of elements. The dimension of this subspace, V^H, is equal to the number of substructure vertices which belong to the open set Ω. In the splitting which corresponds to this preconditioner, the couplings between the different edges (faces) are also ignored. Thus in the four-by-four blocks representation of $S^{(i)}$ previously discussed, the off-diagonal blocks are set to zero. The diagonal blocks may also be changed; cf. Sec. 5. In Sec. 5, we will show that the variables associated with an individual edge can be naturally related to a subspace of functions which vanish outside the two substructures, which have this edge in common. The few preconditioners which have been studied in detail for the three dimensional case, cf. Bramble, Pasciak and Schatz [8] and Dryja [11], can also be described in similar general terms.

3 Schwarz Methods

We begin by briefly discussing the classical formulation of Schwarz's method in the continuous case. There are two fractional steps corresponding to two overlapping subregions, Ω_1' and Ω_2', the union of which is the region Ω. The initial guess $u^0 \in V$. The iterate u^{n+1} is determined from u^n by sequentially updating the approximate solution on the two subregions.

$$\begin{aligned} -\Delta u^{n+1/2} &= f && \text{in } \Omega_1', \\ u^{n+1/2} &= u^n && \text{on } \partial\Omega_1' \end{aligned}$$

and

$$\begin{aligned} -\Delta u^{n+1} &= f && \text{in } \Omega_2', \\ u^{n+1} &= u^{n+1/2} && \text{on } \partial\Omega_2' . \end{aligned}$$

We could equally well have written down the finite element version of the algorithm. From now on, we only consider that case. It is easy and convenient to describe this classical method in terms of two projections P_i, $i = 1, 2$, onto $V_i^h = H_0^1(\Omega_i') \bigcap V^h$; cf. Lions [20]. They are defined by

$$a(P_i v_h, \phi_h) = a(v_h, \phi_h), \ \forall \phi_h \in V_i^h . \tag{3.1}$$

It is also 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

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

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. If more than two subspaces are used, this effect is further pronounced, even if the degree of the polynomial representing the multiplicative algorithm often is lower than maximal. This is so because a product of two projections associated with subregions which do not overlap, vanishes; cf. the discussion in Widlund [31]. The basic idea behind the additive form of the algorithm is to work with the simplest possible polynomial in the projections. Therefore, the equation

$$P u_h = (P_1 + P_2 + \cdots + P_N) u_h = g_h' , \tag{3.2}$$

is solved by an iterative method. Since the operator P is symmetric, with respect to the bilinear form, and positive definite, the method of choice is the conjugate gradient method. Equation (3.2) must have the same solution as Eq. (2.1), i.e., the correct right-hand side must be found. Since by Eq. (2.1), $a(u_h, \phi_h) = f(\phi_h)$, the

right-hand side g_h' can be constructed by solving Eq. (3.1) for all values of i and adding the results. It is similarly possible to apply the operator P of Eq. (3.2) to any given element of V^h by applying each projection P_i to the element and adding the results. Most of the work, in particular that which involves the individual projections, can be carried out in parallel.

We now describe the additive Schwarz method introduced in Dryja and Widlund [14]; cf. also Dryja [12]. We start with the same triangular (simplicial) nonoverlapping substructures Ω_i, that have been considered before. Since Schwarz-type domain decomposition algorithms use overlapping subregions, we extend each substructure to a larger region Ω_i'. We assume that the overlap is generous assuming that the distance between the boundaries $\partial\Omega_i$ and $\partial\Omega_i'$ is 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.

The analysis of Schwarz methods is more complicated when the boundaries of the different subdomains Ω_i' intersect at one or several points; cf. the discussion in Lions [20]. Such a situation occurs if the region is L-shaped and is partitioned into two overlapping rectangles. We discuss such a more complicated situation in Sec. 5.

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

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

The first subspace V_0^h is equal to V^H, the same coarse global space of continuous, piecewise linear functions on the coarse mesh defined by the substructures Ω_i that we introduced in Sec. 2. The other subspaces are related to the subdomains, in the same way as in a traditional Schwarz algorithm, i.e., $V_i^h = V^h \bigcap H_0^1(\Omega_i')$. The computation of the projection of an arbitrary function onto the subspace V^H involves the solution of a standard finite element linear system of algebraic equations which is on the order of N. This coarse, global approximation of the elliptic equation is of the same type as the local problems associated with subdomains. The only real difference between the problem related to the first subspace and the others lies in the way that the right-hand-side of the linear system is generated as weighted averages with weights determined by the basis functions associated with the coarse mesh. We note that if we make the dimension of all the subspaces approximately equal, then we will have $N + 1$ linear systems, each with about than N unknowns, to solve in each step of the iterative solution of a linear system with about N^2 unknowns.

It is well-known that the number of steps required to decrease an appropriate norm of the error of a conjugate gradient iteration by a fixed factor is proportional to $\sqrt{\kappa}$, where κ is the condition number of P; see, e.g., Golub and Van Loan [16]. We therefore need to establish that the operator P of Eq. (3.2) is not only invertible but that satisfactory upper and lower bounds on its eigenvalues can be obtained. A constant upper bound can easily be obtained for P; cf. Sec. 4. For certain other additive algorithms a useful technique is based on strengthened Cauchy inequalities;

cf. Mandel and McCormick [21], Widlund [33] and Yserentant [34].

A lower bound can often conveniently be obtained by using a lemma, given by Lions [20] for the case of $N = 2$; a proof is also given in Widlund [33].

Lemma 3.1 *Let $u_h = \sum_{i=1}^{N} u_{h,i}$, where $u_{h,i} \in V_i$, be a representation of an element of $V^h = V_1 + \cdots + V_N$. If the representation can be chosen so that it follows that $\sum_{i=1}^{N} a(u_{h,i}, u_{h,i}) \leq C_0^2 a(u_h, u_h), \forall\ u_h \in V^h$, then $\lambda_{min}(P) \geq C_0^{-2}$.*

We remark that it is clear that C_0 decreases if we expand the subspaces. This follows from the fact that there is a larger choice in selecting $u_{h,i} \in V_i$. If we can expand the subspaces without worsing the upper bound, and that is often possible, our estimate of $\kappa(P)$ improves. On the other hand a larger subspace also means that the subproblems have more variables and that they are worse conditioned. For the special case of the classical Schwarz method on two regions, this tradeoff is well understood; for a discussion of precise estimates of the rate of convergence cf. Bjørstad and Widlund [3].

As was previously pointed out, the framework with subspaces and projections has already proven quite useful not only for the study of Schwarz-type methods, but also for iterative refinement methods. Similarly, it can be shown, that the bound for the condition number of the iteration operator of the hierarchical basis multigrid method, introduced by Yserentant [34] can be derived by using these techniques. Yserentant's algorithm involves the use of a direct sum of cleverly chosen subspaces, which are quite different from those considered here. All except one of the subproblems that result are very well conditioned. However, since a direct sum of subspaces is used, there is no flexibility in the representation of the elements of V^h. This fact adds to the understanding why Yserentant's method is much less attractive in the case of three dimensions since the best possible C_0 and the condition number grows rather rapidly in that case, when the mesh is refined.

4 Analysis of an Additive Schwarz Method

In this section, we will study the method introduced in the previous section and give a proof of the following result.

Theorem 4.1 *The operator P of the additive algorithm defined by the spaces V^H and V_i^h satisfies the estimate $\kappa(P) \leq$ const.*

Here as elsewhere in this paper, the constants in our estimates, are independent of h and H.

An upper bound for the spectrum of P is quite easy to obtain. We only have to note that for $i \geq 1$,

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

We recall that the subscript indicates the domain of integration of the bilinear form. The basic observation here is that $P_i u_h$ can be regarded as a projection of

$H^1(\Omega_i') \bigcap V^h$ onto $H_0^1(\Omega_i') \bigcap V^h$. In the partion of the region considered here, each point is covered by subregions Ω_i' a finite number of times. A constant upper bound of the eigenvalues of P is therefore obtained by noting that, additionally, the norm of P_0 is equal to one.

The lower bound is obtained by using Lemma 3.1. We partition the finite element function u_h as follows. We first choose $u_{h,0} \in V^H$. By using smoothing and interpolation, cf., e.g., Strang [26], we can find a linear map $\hat{I}_H$ into V^H, which is bounded in $H_0^1(\Omega)$ and which satisfies

$$\|u_h - \hat{I}_H u_h\|_{L_2(\Omega)} \leq (\text{const.}) H |u_h|_{H^1(\Omega)}. \tag{4.1}$$

Let $w_h = u_h - \hat{I}_H u_h$. The other terms in the representation of u_h are defined by $u_{h,i} = I_h(\theta_i w_h)$, $i = 1, 2, \ldots, N$. Here I_h is the interpolation operator into the space V^h and the θ_i define a partition of unity with $\theta_i \in C_0^\infty(\Omega_i')$ and $\sum \theta_i(x) = 1$. Because of the relative generous overlap of the subregions, introduced in Sec. 3, these functions can be chosen so that $\nabla \theta_i$ is bounded by const. $/H_i$. By using the linearity of I_h , we can easily show that we obtain a correct partitioning of u_h . In order to estimate the semi-norm of $u_{h,i}$, we work on one element K at a time. We obtain

$$|u_{h,i}|^2_{H^1(K)} \leq 2|\overline{\theta_i} w_h|^2_{H^1(K)} + 2|I_h((\theta_i - \overline{\theta_i}) w_h)|^2_{H^1(K)} .$$

Here $\overline{\theta_i}$ is the average value of θ_i over K. It is easy to see, by using an inverse inequality, that

$$|I_h((\theta_i - \overline{\theta_i}) w_h)|_{H^1(K)} \leq (\text{const.})\, h^{-1} \|I_h((\theta_i - \overline{\theta_i}) w_h)\|_{L_2(K)} .$$

We can now use the fact that on K, θ_i differs from its average by at most const.h/H_i. After summing over all elements of Ω_i' , we arrive at the inequality

$$|u_{h,i}|^2_{H^1(\Omega_i')} \leq (\text{const.}) \left(|w_h|^2_{H^1(\Omega_i')} + H_i^{-2} \|w_h\|^2_{L_2(\Omega_i')}\right) .$$

We now sum over all i and use that each point in Ω is covered only a fixed number of times. We then obtain a uniform bound on C_0^2, and conclude the proof of Theorem 4.1, by estimating the two terms of

$$|w_h|^2_{H^1(\Omega)} + H_i^{-2} \|w_h\|^2_{L_2(\Omega)}$$

by $|u_h|^2_{H_0^1(\Omega)}$. The bounds follow by using the boundedness of $\hat{I}_H$ in H^1 and inequality 4.1, respectively.

5 Iterative Substructuring Methods

In this section, we show how we can obtain an iterative substructuring method by using the framework with subspaces and projections developed in Sec. 3. We primarily consider problems in the plane. We give an estimate of the condition

number of the operator P, which corresponds to a specific choice of subspaces. This proof is carried out without using an extension theorem. We then indicate how results for the the special case of two substructures can be used to derive different algorithms, among them one due to Bramble, Pasciak and Schatz [5]. We also briefly discuss the three dimensional case.

We assume that the region is divided into substructures as in Sec. 2. An additive Schwarz method is introduced by using the coarse space V^H and subspaces V_{ij}^h, which are related to an edge of a substructure. Thus, let Γ_{ij} be the edge which is common to two adjacent substructures Ω_i and Ω_j. Then, $V_{ij}^h = H_0^1(\Omega_{ij}) \bigcap V^h$, where $\Omega_{ij} = \Omega_i \bigcup \Gamma_{ij} \bigcup \Omega_j$. The region Ω, with the exception of the vertices of the substructures, is therefore covered by the subregions Ω_{ij} constructed from all pairs of adjacent substructures.

Compared with the subspaces used in the previous section, we use less overlap in the sense that only the elements of V^H differ from zero at the vertices of the substructures. This is reflected in a poorer bound on the condition number.

Theorem 5.1 *The operator P of the additive algorithm defined by the spaces V^H and V_{ij}^h satisfies the estimate $\kappa(P) \leq$ (const.)$(1 + log(H/h))^2$.*

By using the same method as in Sec. 4, we can easily show that the eigenvalues of $P \leq 4$. In the proof of the lower bound of the spectrum of P, we use Lemma 3.1 and the following lemma, which plays an important role in the more traditional theory for iterative substructuring algorithms. Variations of this result, which dates back at least to 1966, are given in a number of papers; see e.g., Bramble [4] , Bramble, Pasciak and Schatz [5] or Yserentant [34].

Lemma 5.1 *Let α be any value of $u_h(x)$, with $x \in \Omega_i$. Then*

$$\|u_h - \alpha\|^2_{L^\infty(\Omega_i)} \leq \text{(const.)}\ (1 + \log(H/h))|u_h|^2_{H^1(\Omega_i)} .$$

We note that this result holds only for regions in two dimensions and that it resembles a Sobolev inequality. Since all the elements of V_{ij}^h vanish at the vertices of the substructures, and we must pick the interpolant $I_H u_h$ as the element in V^H in the representation of u_h . It is easy to show that $|I_H u_h|^2_{H^1(\Omega_i)}$ can be estimated by

$$\sum_{k,l=1}^{3} (u_h(\tilde{V}_k) - u_h(\tilde{V}_l))^2 ,$$

where the $\tilde{V}_k's$ are the vertices of the substructure Ω_i . It then follows from Lemma 5.1 that

$$|I_H u_h|^2_{H^1(\Omega_i)} \leq \text{(const.)}(1 + \log(H/h))|u_h|^2_{H^1(\Omega_i)} . \tag{5.1}$$

Let $w_h = u_h - I_H u_h$. As in the previous section, we use a partion of unity. The argument is now more complicated since there is less overlap. The elements of V_{ij}^h, used in the representation of u_h, are given by the formula

$$u_{h,ij} = I_h(\theta_{ij} w_h) .$$

Since w_h vanishes at the vertices of the substructures and we only use values of θ_{ij} at nodal points, we only need a partion of unity at the nodal points which are not substructure vertices. The function θ_{ij} must be equal to 1 at all nodal points in the interior of Γ_{ij}, since all the other cut-off functions vanish there, and it must vanish at the corresponding nodes on the other edges of Ω_i and Ω_j. It is easy to see that $\nabla\theta_{ij}$ therefore must grow as fast as const. $/r$, where r is the distance to the closest endpoint of Γ_{ij}. Cut-off functions, for which this is also an upper bound, can indeed be constructed.

The necessary bound is obtained, one substructure at a time. In order to estimate $|u_{h,ij}|_{H^1(\Omega_i)}$ in terms of $|w_h|_{H^1(\Omega_i)}$, we first consider the few elements of Ω_i for which an end point of Γ_{ij} is a vertex. $u_{h,ij}$ and w_h vanish at such a vertex, and at the other nodes of these special elements the absolute value of $u_{h,ij}$ is no larger than that of w_h. A straightforward calculation shows that the contributions to $|u_{h,ij}|^2_{H^1(\Omega_i)}$ from these triangles are smaller than those to $|w_h|^2_{H^1(\Omega_i)}$. We now use arguments similar to those of the previous section to obtain an estimate of the integral over the rest of the substructure. Working with one element at a time, we obtain

$$\begin{aligned} |u_{h,ij}|^2_{H^1(K)} &\leq 2|w_h|^2_{H^1(K)} + (\text{const.})r^{-2}\|w_h\|^2_{L_2(K)} \\ &\leq 2|w_h|^2_{H^1(K)} + (\text{const.})(h/r)^2\|w_h\|^2_{L_\infty(K)}\,. \end{aligned}$$

By using the fact that w_h does not change if a constant is added to u_h, the formula for $u_{h,ij}$ and that $\|w_h\|_{L_\infty(K)} \leq 2\|u_h\|_{L_\infty(\Omega_i)}$, we obtain

$$|u_{h,ij}|^2_{H^1(K)} \leq 2|w_h|^2_{H^1(K)} + (\text{const.})(h/r)^2\|u_h - \alpha\|^2_{L_\infty(\Omega_i)}\,, \forall\,\alpha\,.$$

The sum of the first term over the elements can be estimated by using inequality (5.1). Since the number of elements decreases to zero linearly with the distance r, the sum of the second expression over all the elements of the substructure, except those at the endpoints of Γ_{ij}, can be estimated by

$$(\text{const.})\|u_h - \alpha\|^2_{L_\infty(\Omega_i)} \int_h^H r^{-1}dr$$

Here H represents the diameter of Ω_i and h the minimum distance of any other nodal point to the end points of Γ_{ij}. The estimate on C_0^2 and the proof of the whole theorem is now concluded by using Lemma 5.1.

This algorithm, introduced and analyzed as a method based on the subspaces V^H and V_{ij}^h and the related projections, can equally well be understood in terms of a splitting. Let us consider the case where the components of the right hand side of the system of equations corresponding to the interiors have been set to zero in a preliminary step. The right hand side of a linear system related to Ω_{ij} and V_{ij}^h then differs from zero only on Γ_{ij}. The Schur complement associated with this edge can be shown, straightforwardly, to be the sum of the corresponding blocks in the four

by four block representation of the Schur complements $S^{(i)}$ and $S^{(j)}$, which were introduced in Sec. 2.

One of the attractive features of the framework first introduced in Sec. 3 is the ease by which the subproblems can be replaced by preconditioners. Let us consider the additive Schwarz method for the problem discussed in the beginning of Sec. 3. We can write the projection P_1 in matrix terms. After a suitable permutation of the variables, it is seen to correspond to

$$y = P_1 x = \begin{pmatrix} {K^{(1)}}^{-1} & 0 \\ 0 & 0 \end{pmatrix} Kx.$$

It is easy to see that this matrix is symmetric in the $K-$ inner product which corresponds to the bilinear form. If ${K^{(1)}}^{-1}$, and the other matrices which play similar roles, are replaced by inverses of preconditioners for the subproblems, then it is easy to see that the resulting algorithm converges and that its condition number can be estimated immediately in terms of $\kappa(P)$ and bounds for the local preconditioners. Using this method, a number of algorithms can be derived from a basic method based on projections and subspaces. Thus, if the Schur complement corresponding to the problem defined on Ω_{ij} is replaced by the square root of the discrete, one-dimensional Laplacian, denoted by ${l_0}^{1/2}$ in Bramble, Pasciak and Schatz [5], we obtain the main algorithm of their paper. The estimate of the condition number of the resulting method can be obtained, by using the argument just given, combining Theorem 5.1 and a bound for the condition number of problems defined on the union of two subregions.

We conclude this paper by a remark on the part of preconditioner which corresponds to the coarse global problem. In two dimensions, we have used the subspace V^H for this purpose. The related quadratic form has, in a special case, the form

$$\sum_i \sum_{k,l=1}^{3} (u_h(\tilde{V}_k^{(i)}) - u_h(\tilde{V}_l^{(i)}))^2 .$$

It is easy to show that this form equally well can be written as a double sum over $(u_h(\tilde{V}_k^{(i)}) - \overline{u}_h^{(i)})^2$, where $\overline{u}_h^{(i)}$ is the average value of the three values associated with the vertices of Ω_i. In three dimensions, the corresponding quadratic form, with the sums and averages calculated with respect to the variables associated with the wire baskets of the individual substructures, provide an important, global part of the powerful iterative substructuring methods, which have been developed by Bramble, Pasciak and Schatz [8] and Dryja [11].

Acknowledgements

This work was supported in part by the National Science Foundation under Grant NSF-CCR-8703768 and, in part, by the U.S. Department of Energy under contract

DE-AC02-76ER03077-V at the Courant Mathematics and Computing Laboratory. This paper was originally typed by the second author in LaTeX and reformatted by Lisa Laguna.

References

[1] Bell, K., B. Hatlestad, O. E. Hansteen, and Per O. Araldsen. NORSAM, *a programming system for the finite element method. Users manual, Part 1, General description.* NTH, Trondheim, 1973.

[2] Bjørstad, Petter E., and Olof B. Widlund. "Iterative methods for the solution of elliptic problems on regions partitioned into substructures," *SIAM J. Numer. Anal.* **23** (1986) 1093–1120.

[3] Bjørstad, Petter E., and Olof B. Widlund. "To overlap or not to overlap: a note on a domain decomposition method for elliptic problems," *SIAM J. Sci. Stat. Comput.* **10** (1989) 1053–1061.

[4] Bramble, James H. "A second order finite difference analogue of the first biharmonic boundary value problem," *Numer. Math.* **9** (1966) 236–249.

[5] Bramble, James H., Joseph E. Pasciak, and Alfred H. Schatz. "The construction of preconditioners for elliptic problems by substructuring, I," *Math. Comp.* **47** (1986) 103–134.

[6] Bramble, James H., Joseph E. Pasciak, and Alfred H. Schatz. "The construction of preconditioners for elliptic problems by substructuring, II," *Math. Comp.* **49** (1987) 1–16.

[7] Bramble, James H., Joseph E. Pasciak, and Alfred H. Schatz. *The Construction of Preconditioners for Elliptic Problems by Substructuring, III.* Technical report, Cornell University, 1987.

[8] Bramble, James H., Joseph E. Pasciak, and Alfred H. Schatz. *The Construction of Preconditioners for Elliptic Problems by Substructuring, IV.* Technical report, Cornell University, 1988. To appear in Math. Comp.

[9] Chan, Tony F., Roland Glowinski, Gérard A. Meurant, Jacques Périaux, and Olof Widlund, eds. *Domain Decomposition Methods*, SIAM, Philadelphia, 1989. Proceedings of the Second International Symposium on Domain Decomposition Methods , Los Angeles, California , January 14–16, 1988.

[10] Chan, Tony F., and Diana C. Resasco. "Analysis of domain decomposition preconditioners on irregular regions." In R. Vichnevetsky and R. Stepleman, eds., *Advances in Computer Methods for Partial Differential Equations.* IMACS, 1987.

[11] Maksymilian Dryja. "A method of domain decomposition for 3-D finite element problems." In Roland Glowinski, Gene H. Golub, Gérard A. Meurant, and Jacques Périaux, eds., *Domain Decomposition Methods for Partial Differential Equations*, SIAM, Philadelphia, 1988.

[12] Maksymilian Dryja. "An additive Schwarz algorithm for two- and three-dimensional finite element elliptic problems." In Tony Chan, Roland Glowinski, Gérard A. Meurant, Jacques Périaux, and Olof Widlund, eds., *Domain Decomposition Methods*, SIAM, Philadelphia, 1989.

[13] Dryja, Maksymilian, Wlodek Proskurowski, and Olof Widlund. "A method of domain decomposition with crosspoints for elliptic finite element problems." In Bl. Sendov, ed., *Optimal Algorithms*, 97–111, Sofia, Bulgaria, 1986. Bulgarian Academy of Sciences.

[14] Dryja, Maksymilian, and Olof B. Widlund. *An Additive Variant of the Schwarz Alternating Method for the Case of Many Subregions.* Technical Report 339, also Ultracomputer Note 131, Department of Computer Science, Courant Institute, 1987.

[15] Glowinski, Roland, Gene H. Golub, Gérard A. Meurant, and Jacques Périaux, eds. *Domain Decomposition Methods for Partial Differential Equations*, SIAM, Philadelphia, 1988. Proceedings of the First International Symposium on Domain Decomposition Methods for Partial Differential Equations, Paris, France, January 1987.

[16] Golub, Gene H., and Charles F. Van Loan. *Matrix Computations.* Johns Hopkins Univ. Press, 1983.

[17] Greenbaum, Anne, Congming Li, and Han Zheng Chao. *Parallelizing Preconditioned Conjugate Gradient Algorithms.* Technical report, Courant Institute, 1988. To appear in *Computer Physics Communications.*

[18] Keyes, David E., and William D. Gropp. "A comparison of domain decomposition techniques for elliptic partial differential equations and their parallel implementation," *SIAM J. Sci. Stat. Comput.* **8** (1987) 166–202.

[19] Keyes, David E., and William D. Gropp. "Domain decomposition techniques for the parallel solution of nonsymmetric systems of elliptic bvps." In Tony Chan, Roland Glowinski, Gérard A. Meurant, Jacques Périaux, and Olof Widlund, eds., *Domain Decomposition Methods*, SIAM, Philadelphia, 1989.

[20] Lions, Pierre Louis. "On the Schwarz alternating method. I." In Roland Glowinski, Gene H. Golub, Gérard A. Meurant, and Jacques Périaux, eds., *First International Symposium on Domain Decomposition Methods for Partial Differential Equations*, SIAM, Philadelphia, 1988.

[21] Mandel, Jan, and Steve McCormick. "Iterative solution of elliptic equations with refinement: The two-level case." In Tony Chan, Roland Glowinski, Gérard A. Meurant, Jacques Périaux, and Olof Widlund, eds., *Domain Decomposition Methods*, SIAM, Philadelphia, 1989.

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

[23] Przemieniecki, J.S. "Matrix structural analysis of substructures," *Am. Inst. Aero. Astro. J.* **1** (1963) 138–147.

[24] Schwarz, H. A. *Gesammelete Mathematische Abhandlungen*, Vol. 2, 133–143. Springer, Berlin, 1890. First published in Vierteljahrsschrift der Naturforschenden Gesellschaft in Zürich **15** (1870) 272–286.

[25] Sobolev, S. L. "The Schwarz algorithm in the theory of elasticity," *Dokl. Acad. N. USSR*, **IV(XIII)** (1936) 236–238. (in Russian).

[26] Strang, Gilbert. "Approximation in the finite element method," *Numer. Math.* **19** (1972) 81–98.

[27] Varga, Richard S. *Matrix Iterative Analysis*. Prentice-Hall, 1962.

[28] Widlund, Olof B. "An extension theorem for finite element spaces with three applications." In Wolfgang Hackbusch and Kristian Witsch, eds., *Numerical Techniques in Continuum Mechanics*, 110–122, Braunschweig/Wiesbaden, 1987. Notes on Numerical Fluid Mechanics, Vol. 16, Friedr. Vieweg und Sohn. Proceedings of the Second GAMM-Seminar, Kiel, January , 1986.

[29] Widlund, Olof B. "Iterative substructuring methods: Algorithms and theory for problems in the plane." In Roland Glowinski, Gene H. Golub, Gérard A. Meurant, and Jacques Périaux, eds., *First International Symposium on Domain Decomposition Methods for Partial Differential Equations*, SIAM, Philadelphia, 1988.

[30] Widlund, Olof B. "Iterative substructuring methods:The general elliptic case." In *Computational Processes and Systems, 6*, Moscow, 1988. Nauka. Proceedings of Modern Problems in Numerical Analysis, a conference held in Moscow, USSR, September , 1986. (In Russian, also available from the author, in English, as a technical report.)

[31] Widlund, Olof B. *On the Rate of Convergence of the Classical Schwarz Alternating Method in the Case of More Than Two Subregions*. Technical report, Department of Computer Science, Courant Institute, 1988.

[32] Widlund, Olof B. *Some Domain Decomposition and Iterative Refinement Algorithms for Elliptic Finite Element Problems*. Technical Report 386, Department of Computer Science, Courant Institute, 1988. To appear in *J. Comp. Math.*

as the proceedings of the China–U.S. Seminar on Boundary Integral Equations and Boundary Element Methods in Physics and Engineering, held at the Xi'an Jiatong University, Xi'an, The People's Republic of China, December 27, 1987–January 1, 1988.

[33] Widlund, Olof B. "Optimal iterative refinement methods." In Tony Chan, Roland Glowinski, Gérard A. Meurant, Jacques Périaux, and Olof Widlund, eds., *Domain Decomposition Methods*, SIAM, Philadelphia, 1989.

[34] Yserentant, Harry. "On the multi-level splitting of finite element spaces," *Numer. Math.* **49** (1986) 379–412.