The Construction of Preconditioners
for Elliptic Problems by Substructuring. I

By J. H. Bramble,* J. E. Pasciak* and A. H. Schatz*

*Dedicated to Professor Joachim Nitsche on the occasion
of the sixtieth anniversary of his birthday.*

Abstract. We consider the problem of solving the algebraic system of equations which arise from the discretization of symmetric elliptic boundary value problems via finite element methods. A new class of preconditioners for these discrete systems is developed based on substructuring (also known as domain decomposition). The resulting preconditioned algorithms are well suited to emerging parallel computing architectures. The proposed methods are applicable to problems on general domains involving differential operators with rather general coefficients. A basic theory for the analysis of the condition number of the preconditioned system (which determines the iterative convergence rate of the algorithm) is given. Techniques for applying the theory and algorithms to problems with irregular geometry are discussed and the results of extensive numerical experiments are reported.

1. Introduction. The aim of this series of papers is to propose and analyze methods for efficiently solving the equations resulting from finite element discretizations of second-order elliptic boundary value problems on general domains in R^2 and R^3. In particular, we shall be concerned with constructing easily invertible and "effective" preconditioners for the resulting system of discrete equations which can be used in a preconditioned iterative algorithm to achieve a rapid solution method. The methods to be presented are well suited to parallel computing architectures.

In this paper we shall restrict ourselves to boundary value problems in R^2. Let Ω be a bounded domain in R^2 with a piecewise smooth boundary $\partial\Omega$. As a model problem for a second-order uniformly elliptic equation we shall consider the Dirichlet problem

$$(1.1) \qquad Lu = f \quad \text{in } \Omega, \qquad u = 0 \quad \text{on } \partial\Omega,$$

where

$$Lv = -\sum_{i,j=1}^{2} \frac{\partial}{\partial x_i}\left(a_{ij}\frac{\partial v}{\partial x_j}\right),$$

with a_{ij} uniformly positive definite, bounded and piecewise smooth on Ω. The generalized Dirichlet form is given by

$$(1.2) \qquad A(v,\phi) = \sum_{i,j=1}^{2} \int_{\Omega} a_{ij} \frac{\partial v}{\partial x_i} \cdot \frac{\partial \phi}{\partial x_j} \, dx,$$

which is defined for all v and ϕ in the Sobolev space $H^1(\Omega)$ (the space of distributions with square-integrable first derivatives). The $L^2(\Omega)$-inner product is denoted

$$(v,\phi) = \int_{\Omega} v\phi \, dx.$$

The subspace $H_0^1(\Omega)$ is the completion of the smooth functions with support in Ω with respect to the norm in $H^1(\Omega)$. The weak formulation of the problem defined by (1.1) is: Find $u \in H_0^1(\Omega)$ such that

$$(1.3) \qquad A(u,\phi) = (f,\phi)$$

for all $\phi \in H_0^1(\Omega)$. This leads immediately to the standard Galerkin approximation. Let $S_h^0(\Omega)$ be a finite-dimensional subspace of $H_0^1(\Omega)$. The Galerkin approximation is defined as the solution of the following problem: Find $U \in S_h^0(\Omega)$ such that

$$(1.4) \qquad A(U,\Phi) = (f,\Phi)$$

for all $\Phi \in S_h^0(\Omega)$. Once a basis $\{\chi_i\}_{i=1}^{N}$ for $S_h^0(\Omega)$ is chosen, (1.4) leads to a system of linear algebraic equations. Write $U = \sum_{i=1}^{N} \alpha_i \chi_i$. Then (1.4) becomes

$$(1.5) \qquad \sum_{i=1}^{N} \alpha_i A(\chi_i, \chi_j) = (f, \chi_j),$$

$j = 1, \ldots, N$. We shall choose $S_h^0(\Omega)$ so that firstly, the function U will be a good approximation to u and secondly, efficient algorithms for the solution of the underlying linear system (1.5) can be developed. In particular, we will consider subspaces $S_h^0(\Omega)$ of $H_0^1(\Omega)$ which are defined so that certain related subproblems can be efficiently solved. We will see that this leads to algorithms for the solution of the global linear system which is well suited to parallel processing.

The strategy of choosing $S_h^0(\Omega)$ so that efficient algorithms exist for the solution of the resulting linear system is not unusual. For example, for the Laplace operator on a rectangular region, a subspace $S_h^0(\Omega)$ of piecewise linear functions on a uniform triangulation leads to the usual 5-point approximation to the Laplacian. The resulting equations may be solved "fast" using, for example, fast Fourier transform techniques. In this case, other choices of $S_h^0(\Omega)$ may lead to good approximate solutions, but these solutions may be more difficult to obtain computationally. Another example of a special choice of $S_h^0(\Omega)$ which leads often to a fast algorithm is one which may be thought of as connected with a nested set of grids. For such spaces, a "multigrid" algorithm may be applied.

The underlying method which we will consider is a preconditioned iterative method. The choice of a particular iterative method within a certain class is not essential, but for the purpose of this exposition we may think of the well-known conjugate gradient method [12], [15] which is often used in practice. Roughly, the application of a preconditioned method may be described as follows. Let A be the $N \times N$ matrix with entries $A(\chi_i, \chi_j)$, α the column vector whose components are

as in (1.5), and F the vector with components (f, χ_j). Then (1.5) may be written as

$$(1.6) \qquad\qquad A\alpha = F.$$

Generally, the matrix A is not well-conditioned so that a direct application of the conjugate gradient method to the symmetric positive-definite system (1.6) will not be a very efficient algorithm. The preconditioned conjugate gradient method (PCG) consists of choosing a positive-definite symmetric matrix B and writing the equivalent system

$$(1.7) \qquad\qquad B^{-1}A\alpha = B^{-1}F.$$

In the present context the matrix B will be associated with another bilinear form $B(\cdot, \cdot)$ defined on $S_h^0(\Omega) \times S_h^0(\Omega)$. The system (1.7) is symmetric with respect to the inner product defined by

$$(1.8) \qquad\qquad [\alpha, \beta] \equiv \sum_{i,j=1}^{N} B_{ij}\alpha_i\beta_j.$$

Thus, the conjugate gradient method may be applied to (1.7) with respect to (1.8). The importance of making a "good" choice for B is well known. The matrix B should have two properties. First, the solution of the problem

$$(1.9) \qquad\qquad B\beta = b$$

should be easy to obtain. This is tantamount to applying the operator B^{-1} to the vector b. Secondly, B should be spectrally close to A in the sense that the condition number K of $B^{-1}A$ should not be large. Clearly, $K \le \lambda_1/\lambda_0$, where λ_0 and λ_1 are constants such that

$$\lambda_0[\beta, \beta] \le [B^{-1}A\beta, \beta] \le \lambda_1[\beta, \beta] \quad \text{for all } \beta \in R^N.$$

In terms of the form $B(\cdot, \cdot)$, the first property means that the solution W of

$$(1.10) \qquad\qquad B(W, \Phi) = (g, \Phi), \quad \text{for all } \Phi \in S_h^0(\Omega)$$

for a given function g should be easier to obtain than the solution of (1.4). The spectral condition, in terms of the forms, is

$$(1.11) \qquad \lambda_0 B(V, V) \le A(V, V) \le \lambda_1 B(V, V) \quad \text{for all } V \in S_h^0(\Omega).$$

These two properties will guarantee, firstly, that the work per iterative step in applying the preconditioned method will be small, and, secondly, that the number of steps to reduce the error to a given size will also be small so that an efficient algorithm will result.

In this paper we will describe and analyze a technique for constructing the bilinear form $B(\cdot, \cdot)$ so that the action of the corresponding matrix problem B^{-1} is easy to compute. As a preliminary step, the domain is subdivided into subdomains. Our preconditioner will be defined so that the computation of its inverse applied to a vector only involves solving in parallel related Galerkin (or matrix) equations on subregions of Ω and some interconnecting equations, which may also be solved in parallel. The preconditioner B will be the first of our domain decomposition preconditioners to be developed in this series of papers and will sometimes be denoted DD1.

In Section 2, the preconditioner B^{-1} will be defined and the essential step in the iterative algorithm of computing the action of B^{-1} will be described in detail.

The main result concerning the condition number K is also stated in this section as Theorem 1. Section 3 is devoted to a complete proof of Theorem 1. In Section 4 we show how $S_h^0(\Omega)$ can be constructed and how various coefficients introduced in the definition of the preconditioner (see Section 2) can be chosen so that the related subproblems can be efficiently solved, even in rather complex domain geometry. Section 5 contains a reexamination of the process of applying the action of B^{-1} in the context of "block Gauss elimination". Finally, in Section 6 we describe the results of numerical calculations which show that the theoretical estimates are fully realized in practice.

For other works dealing with the numerical solution of boundary value problems via substructuring we refer to [1], [2], [4]–[7], [10]. We emphasize that a novel feature of our approach is that more than two subdomains can meet at an interior point of the original domain. In addition, our results remain valid independently of the number of such points. As a simple example, our approach applies to a checkerboard subdivision of a square.

2. The Construction of $B(\cdot, \cdot)$ and the Preconditioning Algorithm. As mentioned in the introduction, the preconditioner which we will construct involves the solution of smaller related problems on subdomains and subdomain boundaries. For the sake of simplicity of exposition we shall proceed with the discussion only for the special case of polygonal domains and piecewise linear approximations.

More precisely, we shall begin with the following assumptions with regard to Ω.

A.1: Ω is a polygonal domain.

A.2: For each h, $0 < h < 1$ a parameter, Ω has been given a quasi-uniform triangulation Ω^h. By this we mean that there exists a positive constant c_1 independent of h such that each triangle $\tau^h \in \Omega^h$ contains a ball of radius $c_1 h$ and is contained in a ball of radius h.

A.3: For each triangulation Ω^h, Ω may be written as the union of disjoint regions Ω_k, which are either quadrilaterals or triangles whose sides coincide with the mesh lines of the original triangulation and which are quasi-uniform of size $d \geq h$ with constants, as above, which are independent of d and h. If Ω_k is a quadrilateral, we require additionally that the lengths of each side be bounded from below by $c_1 d$ and that any interior angle α satisfy $0 < C_0 \leq \alpha \leq C_1 < \pi$. The collection of regions Ω_k will frequently be referred to as the subdomains.

The vertices of the $\{\Omega_k\}$ will be labeled v_j (ordered in some way) and Γ_{ij} will denote the straight line segment with endpoints v_i and v_j. Throughout this paper we shall only consider Γ_{ij} when Γ_{ij} is an edge of some Ω_k. Furthermore, we associate with each Ω_k the triangulation inherited from the original triangulation Ω^h. The examples given in Figures 2.1 and 2.2 should help clarify the situation.

For each h, let $S_h(\Omega)$ be the space of continuous piecewise linear functions defined relative to the triangulation Ω^h and $S_h^0(\Omega)$ be the subspace of $S_h(\Omega)$ consisting of those functions which vanish on $\partial\Omega$. $S_h^0(\Omega_j)$ will denote the subspace of $S_h^0(\Omega)$ of functions whose supports are contained in $\bar{\Omega}_j$ (in particular, they vanish on $\partial\Omega_j$ and outside $\bar{\Omega}_j$). In addition, $S_h(\Omega_j)$ will be the set of functions which are restrictions of those in $S_h^0(\Omega)$ to $\bar{\Omega}_j$. Subspaces on the boundaries of the subdomains will be denoted as follows. $S_h(\partial\Omega_j)$ will denote the restrictions of $S_h(\Omega_j)$ to $\partial\Omega_j$ and $S_h^0(\Gamma_{ij})$, the subspace of $S_h(\partial\Omega_j)$ consisting of functions whose support is contained on the edge Γ_{ij}. In what follows, c and C (with or without subscript) will denote generic positive constants which are independent of h, d and the Ω_k.
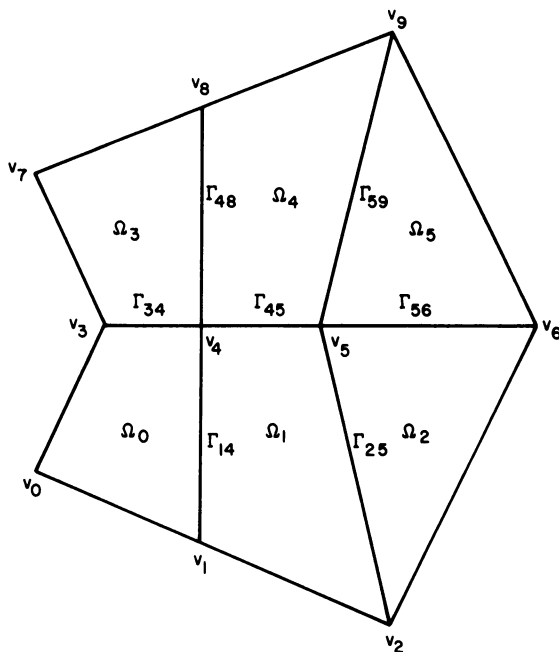
FIGURE 2.1. *The domain Ω and subdomains.*
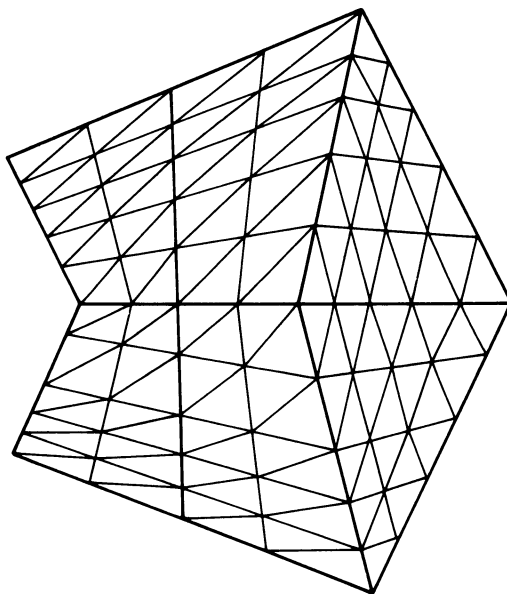
FIGURE 2.2. *The domain with mesh.*

We construct our preconditioner B by constructing its corresponding bilinear form $B(\cdot, \cdot)$ defined on $S_h^0(\Omega) \times S_h^0(\Omega)$. We first introduce another form $\tilde{A}(\cdot, \cdot)$ which is defined by first setting

$$\tilde{A}_k(U, V) = \sum_{i,j=1}^{2} \int_{\Omega_k} a_{ij}^k \frac{\partial U}{\partial x_i} \frac{\partial V}{\partial x_j} \, dx,$$

and then defining

$$\tilde{A}(U,V) = \sum_k \tilde{A}_k(U,V).$$

Here for each k, a_{ij}^k is a piecewise smooth (possibly discontinuous) uniformly positive-definite matrix. The reason for the form of $\tilde{A}$ will become clear as we proceed with the development. We note, however, that

$$C_0 \tilde{A}(U,U) \le A(U,U) \le C_1 \tilde{A}(U,U)$$

for positive constants C_0 and C_1. Thus, the problem of finding a preconditioner for A is the same as finding one for $\tilde{A}$.

We next decompose functions in $S_h^0(\Omega)$ as follows: Write $W = W_P + W_H$ where $W_P \in S_h^0(\Omega_1) \oplus \cdots \oplus S_h^0(\Omega_{n_r})$ and satisfies

$$\tilde{A}_k(W_P, \Phi) = \tilde{A}_k(W, \Phi) \quad \text{for all } \Phi \in S_h^0(\Omega_k)$$

for each k. Notice that W_P is determined on Ω_k by the values of W on Ω_k and that

$$\tilde{A}_k(W_H, \Phi) = 0 \quad \text{for all } \Phi \in S_h^0(\Omega_k).$$

Thus on each Ω_k, W is decomposed into a function W_P which vanishes on $\partial\Omega_k$ and a function $W_H \in S_h(\Omega_k)$ which satisfies the above homogeneous equations and has the same boundary values as W. We shall refer to such a function W_H as "discrete $\tilde{A}_k$-harmonic."

Remark 2.1. The matrices with entries a_{ij}^k are in principle arbitrary but, as will be seen in Section 4, they may be chosen in such a way that the subproblems determining W_P and W_H may be easily solved once the values of W_H on the subdomain boundaries are known.

We note that the above decomposition is orthogonal in the $\tilde{A}$-inner product and hence,

$$\tilde{A}(W,W) = \tilde{A}(W_P, W_P) + \tilde{A}(W_H, W_H).$$

We shall define $B(\cdot, \cdot)$ by replacing the $\tilde{A}(W_H, W_H)$ term above. To do this, we decompose $W_H \in S_h(\Omega_k)$ into $W_H = W_E + W_V$, where $W_V \in S_h(\Omega_k)$ is the discrete $\tilde{A}_k$-harmonic function whose values on $\partial\Omega_k$ are the linear function along each Γ_{ij} with the same values as W at the vertices. Thus W_E is a discrete $\tilde{A}_k$-harmonic function in Ω_k for each k which vanishes at all of the vertices.

Before defining the form $B(\cdot, \cdot)$, we note that for any discrete $\tilde{A}_k$-harmonic function W with zero mean value on Ω_k,

(2.1) $$\gamma_0 \tilde{A}_k(W,W) \le |W|_{1/2, \partial\Omega_k}^2 \le \gamma_1 \tilde{A}_k(W,W),$$

where γ_0 and γ_1 are positive constants and $|\cdot|_{1/2, \partial\Omega_k}$ is the norm on the Sobolev space $H^{1/2}(\partial\Omega_k)$. This will be proved in Section 3 but is noted here to motivate our construction. Now it will also be shown in the next section that if $W = 0$ at the vertices, then the norm $|W|_{1/2, \partial\Omega_k}^2$ may be replaced in (2.1) by $\sum_{\Gamma_{ij}} \alpha_{ij} \langle a^{-1}\tilde{l}_0^{1/2} W, W \rangle_{\Gamma_{ij}}$ with new values of γ_0 and γ_1 such that $\gamma_1/\gamma_0 \le C(1 + \ln(d/h)^2)$. Here $\tilde{l}_0$ is the operator defined for each Γ_{ij} on $S_h^0(\Gamma_{ij})$ by

(2.2) $$\langle a^{-1}\tilde{l}_0 W, \Phi \rangle_{\Gamma_{ij}} = \langle aW', \Phi' \rangle_{\Gamma_{ij}} \quad \text{for all } \Phi \in S_h^0(\Gamma_{ij}).$$

The prime denotes differentiation with respect to arc length s along Γ_{ij}. In (2.2) a is, for simplicity, a positive piecewise constant function on Γ_{ij}, $0 < a_0 \leq a \leq a_1$, with a_0 and a_1 independent of Γ_{ij} and h, and

$$\langle \phi, \psi \rangle_{\Gamma_{ij}} = \int_{\Gamma_{ij}} \phi \psi \, ds.$$

Note that $\tilde{l}_0$ is symmetric and positive definite in the inner product $\langle a^{-1} \cdot, \cdot \rangle_{\Gamma_{ij}}$ and hence its square root is well-defined.

Here again we may, in principle, choose the function a quite arbitrarily, but, as will be seen in Section 4, computational considerations dictate a natural choice. α_{ij} is a positive constant which will also be chosen explicitly later. We however require that $0 < C_0 \leq \alpha_{ij} \leq C_1$ for constants C_0 and C_1 which are independent of h, d, and the Ω_k's.

Finally, as is shown in the next section, for W_V as defined above,

$$C_0 A_k(W_V, W_V) \leq \sum_{\Gamma_{ij}} \alpha_{ij}(W_V(v_i) - W_V(v_j))^2 \leq C_1 A_k(W_V, W_V)$$

holds for some positive constants C_0 and C_1.

With the above statements in mind, we now define the form $B(\cdot, \cdot)$ by

$$
\begin{aligned}
B(W, \Phi) = {} & \tilde{A}(W_P, \Phi_P) + \sum_{\Gamma_{ij}} \alpha_{ij} \langle a^{-1} \tilde{l}_0^{1/2} W_E, \Phi_E \rangle_{\Gamma_{ij}} \\
& + \sum_{\Gamma_{ij}} \alpha_{ij}(W_V(v_i) - W_V(v_j))(\Phi_V(v_i) - \Phi_V(v_j)).
\end{aligned}
$$

(2.3)

The following theorem is proved in Section 3:

THEOREM 1. *There are positive constants λ_0, λ_1 and C such that*

$$\lambda_0 B(W, W) \leq A(W, W) \leq \lambda_1 B(W, W) \quad \text{for all } W \in S_h^0(\Omega),$$

where $\lambda_1/\lambda_0 \leq C(1 + \ln(d/h)^2)$. If all of the vertices of the Ω_k lie on $\partial\Omega$, then $\lambda_1/\lambda_0 \leq C$.

Thus the condition number grows at most like $(1 + \ln(d/h)^2)$ as h tends to zero. This means that the preconditioned iteration will converge rapidly and corresponds to the second of the two desirable properties mentioned earlier.

The first property previously discussed states that problem (1.10) should be much more easily solved than the original (1.4). This means that the solution of the corresponding matrix equation (1.9) is relatively easy to obtain.

We shall demonstrate how (1.10) can be solved efficiently. In fact, we shall see that the defining equations have been chosen to conveniently lend themselves to a "block Gauss elimination" procedure. Here, we shall describe the process used to solve (1.10). The matrix interpretation is given in Section 5.

Given g, the problem of solving (1.10) reduces to finding the functions W_P and W_H. The function W_P restricted to Ω_k satisfies

(2.4) $$\tilde{A}_k(W_P, \Phi) = (g, \Phi) \quad \text{for all } \Phi \in S_h^0(\Omega_k).$$

Thus the function W_P on Ω_k can be obtained by solving the corresponding Dirichlet problem (2.4). Note that the problems on different subdomains are independent of each other so that they may be solved in parallel.

With W_P now known, we are left with the equation

$$(2.5) \quad \sum_{\Gamma_{ij}} \alpha_{ij} \langle a^{-1} \tilde{l}_0^{1/2} W_E, \Phi_E \rangle_{\Gamma_{ij}} + \sum_{\Gamma_{ij}} \alpha_{ij} (W_V(v_i) - W_V(v_j))(\Phi_V(v_i) - \Phi_V(v_j))$$

$$= (g, \Phi) - \tilde{A}(W_P, \Phi_P) = (g, \Phi) - \tilde{A}(W_P, \Phi),$$

the last equality holding since $\tilde{A}(W_P, \Phi_H) = 0$. Notice that the value of $(g, \Phi) - \tilde{A}(W_P, \Phi)$, for each Φ, depends only on the value of Φ on the Γ_{ij}'s. Thus (2.5) gives rise to a set of equations on the restriction of $S_h^0(\Omega)$ to $\bigcup \Gamma_{ij}$. To solve these equations, we proceed as follows: For each Γ_{ij} choose Φ in the subspace of $S_h^0(\Omega)$ whose elements vanish in the interior mesh points of every Ω_k and on all other Γ's and, in particular, at the endpoints of Γ_{ij}. Thus, on this subspace, (2.5) decouples into the independent problems of finding $W_E \in S_h^0(\Gamma_{ij})$ given by

$$(2.6) \qquad \alpha_{ij} \langle a^{-1} \tilde{l}_0^{1/2} W_E, \Phi \rangle_{\Gamma_{ij}} = (g, \Phi) - \tilde{A}(W_P, \Phi)$$

for each Γ_{ij}. The computational aspects of solving for W_E on each Γ_{ij} are fully discussed in Section 4; however, note that these are local problems with unknowns corresponding to the nodes on Γ_{ij} and may be solved in parallel.

Next we must solve for W_V on the edges. We consider the subspace of $S_h^0(\Omega)$ consisting of functions which are linear between the endpoints of each Γ_{ij} and vanish at mesh points in Ω^h which are interior to any Ω_j. Clearly, such a subspace has dimension equal to the number of interior vertices, i.e., vertices of the Ω_k which do not lie on $\partial \Omega$. For each Φ in this subspace, $\Phi_E = 0$ and (2.5) reduces to

$$(2.7) \qquad \sum_{\Gamma_{ij}} \alpha_{ij} (W_V(v_i) - W_V(v_j))(\Phi_V(v_i) - \Phi_V(v_j)) = (g, \Phi) - \tilde{A}(W_P, \Phi).$$

A basis for this subspace may be chosen as follows: Choose $\Phi^1, \ldots, \Phi^M$, where M is the number of vertices not on $\partial \Omega$ and $\Phi^i(v_j) = \delta_{ij}$ where $\delta_{ij} = 1$ if $i = j$ and 0 otherwise. This choice gives rise to a difference equation for the function W_V on the interior vertices which is independent of (2.6) and may be solved concurrently. The values W_V at the vertices determine W_V on the edges and hence $W_H = W_E + W_V$ is known on all of the edges Γ_{ij}.

The last step consists of determining W_H in each Ω_k so that

$$(2.8) \qquad \tilde{A}_k(W_H, \Phi) = 0 \quad \text{for all } \Phi \in S_h^0(\Omega_k).$$

The problem of finding the solution of (2.8), given the values of W_H on the boundary of the subdomains, reduces to independent Dirichlet solves on the subdomains. Hence the solution of (1.10) is determined by $W = W_P + W_H$.

We summarize the process by outlining the steps for obtaining the solution of

$$B(W, \psi) = (g, \psi) \quad \text{for all } \psi \in S_h^0(\Omega),$$

and hence for computing the action of B^{-1}.

Algorithm DD1.

1. Find W_P by solving Dirichlet problems on subregions. The solution of the individual Dirichlet problems on the subdomains may be done in parallel.

2. Find W_E on Γ_{ij} by solving one-dimensional equations on each Γ_{ij}.

3. Find W_V on $\bigcup \Gamma_{ij}$ by first finding W_V on the interior vertices of Ω_k by solving a coarse mesh difference equation and then extending piecewise linearly to the edges

Γ_{ij}. The solution on the different segments Γ_{ij} of Step 2 and Step 3 may be done in parallel.

4. Find W_H by extending the values of $W_E + W_V$ on the Γ_{ij} discrete $\tilde{A}_k$-harmonically to the subregions; i.e., solve Dirichlet problems on the subregions. As in Step 1, the solutions of the individual Dirichlet problems on the various subdomains may be done in parallel.

5. Set $W = W_P + W_H$.

We shall now discuss several features of this preconditioning algorithm.

Remark 2.2. The process described above is just that which is required for applying the "action" of the matrix B^{-1} to an arbitrary vector. We again emphasize that it involves solving some local problems on subdomains which are independent of each other so that they can be solved concurrently on computers with parallel architecture.

Remark 2.3. As remarked previously in this section, the matrices of coefficients a_{ij}^k defining the forms $\tilde{A}_k$ need only, in principle, be chosen so that they are uniformly positive definite (which implies the spectral equivalence of $\tilde{A}$ and A). However, as will be seen in later sections, a judicious choice of a_{ij}^k can often be made which results in subdomain problems which can be "fast" solved. In Section 4, we shall explicitly show one method of choosing the coefficients a_{ij}^k in a simple way so that known efficient direct methods may be used to solve the problems on the subdomains.

Remark 2.4. The theoretical results for this algorithm remain valid independent of the number of subdomains and interior vertices used in the decomposition of Ω. This is important when the coefficients $a_{ij}(x)$ are rapidly varying, in which case preconditioners with smaller subdomains more closely reflect the behavior of the coefficients and give rise to more rapidly convergent algorithms. The freedom to use many subdomains may also prove to be important in developing the most efficient preconditioner for a computer with a large number of parallel processors.

Remark 2.5. For simplicity of presentation, we have assumed that Ω is a polygonal domain and the subdomains Ω_j are either quadrilaterals or triangles and that the subspaces consist of piecewise linear functions. The algorithm and theorem can be extended, under reasonable assumptions, to the case where Ω is a bounded domain with piecewise smooth boundary and the subdomains Ω_j have either piecewise smooth boundaries or are mesh domains which approximate piecewise smooth boundaries. We can also extend the above algorithms to a class of higher-order piecewise polynomial subspaces. We intend to deal with these extensions in a later paper.

Remark 2.6. We could also define another preconditioner by replacing the form on the left-hand side of (2.7) by a form, corresponding to a weighted identity operator, leading to the equations for W_V given by

$$(2.9) \qquad \sum_{i=1}^{M} \alpha_i W_V(v_i) \Phi_V(v_i) = (g, \Phi) - \tilde{A}(W_p, \Phi),$$

where the sum is taken over all interior vertices of the Ω_j's. Note that in the case $M = 1$, this formulation coincides with (2.7). The choice of basis functions Φ^i as indicated after (2.7) leads to

$$W_V(v_i) = \frac{(g, \Phi^i) - \tilde{A}(W_P, \Phi^i)}{\alpha_i}, \qquad i = 1, \ldots, M.$$

In this case one can prove, using the techniques given in Section 3, that Theorem 1 holds with

$$(2.10) \qquad\qquad \lambda_1/\lambda_0 \leq C d^{-2}(1 + \ln(d/h)^2),$$

where C is independent of d and h. The estimate indicates that this procedure may be reasonable if d is large, i.e., there are very few subdomains, but will become inefficient as d becomes small, i.e., as the number of subdomains is increased. This is illustrated by the results of Example 6 of Section 6.

Remark 2.7. In the case that the forms $\tilde{A}_k(\cdot, \cdot)$ and $A(\cdot, \cdot)$ coincide on functions in $S_h(\Omega_k)$ then the variables which are interior to Ω_k can be eliminated from the iterative process. Consequently, if the above forms coincide on every subdomain, then the iterative process can be reduced to a boundary iteration. The resulting algorithm is more efficient than the general algorithm in that each iteration does not require the solution of (2.4). However, much of the generality and flexibility of the general algorithm is lost.

3. A Proof of Theorem 1. In this section, we prove the main theorem of the paper which provides bounds on the condition number for the preconditioned system corresponding to (1.7). This, as previously noted, reduces to the estimation of the quantities λ_0 and λ_1 appearing in quadratic form inequalities (1.11). This will be done here in the special case of assumptions A.1, A.2, A.3, and where the finite element subspaces are as in Section 2.

We shall first need some preliminaries. We remind the reader that c or C, with or without subscript, will denote a generic positive constant which is independent of h, d, the subdivision Ω_k and the triangulation Ω^h.

The derivation of the estimates in this section requires the use of various norms defined on the subdomain boundaries. Let Ω_i be a subdomain of Ω^h (as defined in Section 2) and β_i be the set of indices jk with $\Gamma_{jk} \in \partial\Omega_i$, hence $\partial\Omega_i = \bigcup \Gamma_{jk}$ for $jk \in \beta_i$. The Sobolev space of order one half on $\partial\Omega_i$ will be denoted $H^{1/2}(\partial\Omega_i)$ and is defined in [11], [14], [16]. With d as in A.3 (roughly the diameter of Ω_i), we define the weighted norm on $H^{1/2}(\partial\Omega_i)$ by

$$(3.1) \qquad |w|_{1/2, \partial\Omega_i} = \left(\int_{\partial\Omega_i} \int_{\partial\Omega_i} \frac{(w(x) - w(y))^2}{|x - y|^2} \, ds(x) \, ds(y) \right.$$
$$\left. + d^{-1} |w|^2_{L^2(\partial\Omega_i)} \right)^{1/2},$$

where s is arc length along $\partial\Omega_i$. If v is a smooth function on $\partial\Omega_i$ with support contained in one of the edges $\Gamma_{jk} \subset \partial\Omega_i$, then the integral term in (3.1) reduces to

$$\int_{\Gamma_{jk}} \int_{\Gamma_{jk}} \frac{(v(x) - v(y))^2}{|x - y|^2} \, ds(x) \, ds(y) + 2 \int_{\Gamma_{jk}} \int_{\partial\Omega_i/\Gamma_{jk}} \frac{v(x)^2}{|x - y|^2} \, ds(y) \, ds(x).$$

A straightforward computation gives that

$$c \int_{\partial\Omega_i/\Gamma_{jk}} |x - y|^{-2} \, ds(y) \leq |x - v_k|^{-1} + |x - v_j|^{-1} \leq C \int_{\partial\Omega_i/\Gamma_{jk}} |x - y|^{-2} \, ds(y).$$

Thus, for smooth v with support contained on Γ_{jk}, the norm in (3.1) is equivalent to

$$(3.2) \qquad \left(\int_{\Gamma_{jk}} \int_{\Gamma_{jk}} \frac{(v(x) - v(y))^2}{|x - y|^2} \, ds(x) \, ds(y) \right.$$
$$\left. + \int_{\Gamma_{jk}} \frac{v(x)^2}{|x - v_k|} + \frac{v(x)^2}{|x - v_j|} \, ds(x) \right)^{1/2}.$$

The space $\overset{\circ\circ}{H}{}^{1/2}(\Gamma_{jk})$ is defined to be the completion of the smooth functions with compact support in Γ_{jk} with respect to the norm (3.2). We shall denote by $|\cdot|_{1/2,\Gamma_{jk}}$ the norm on $\overset{\circ\circ}{H}{}^{1/2}(\Gamma_{jk})$ given by (3.2). It is well known that the space $\overset{\circ\circ}{H}{}^{1/2}(\Gamma_{jk})$ is the interpolation space which is halfway between $H_0^1(\Gamma_{jk})$ and $L^2(\Gamma_{jk})$ [14], [16]. We note that the operator $(-\partial^2/\partial s^2)$ with domain of definition $H_0^1(\Gamma_{jk})$ is positive definite and selfadjoint on $L^2(\Gamma_{jk})$ and domain of $(-\partial^2/\partial s^2)^{1/2} = \overset{\circ\circ}{H}{}^{1/2}(\Gamma_{jk})$. Consequently, the corresponding norm given by

$$(3.3) \qquad \left(\left\langle (-\partial^2/\partial s^2)^{1/2} w, w \right\rangle_{\Gamma_{jk}} \right)^{1/2}$$

is equivalent to the norm of (3.2) on $\overset{\circ\circ}{H}{}^{1/2}(\Gamma_{jk})$. We note that the discrete operator l_0 defined on $S_h^0(\Gamma_{jk})$ by

$$\langle l_0 W, \phi \rangle_{\Gamma_{jk}} = \langle W', \phi' \rangle_{\Gamma_{jk}} \quad \text{for all } \phi \in S_h^0(\Gamma_{jk}),$$

is a finite-dimensional approximation to $(-\partial^2/\partial s^2)$. Using A.2, it can be shown by interpolation [13, Theorem 9.1] that

$$(3.4) \qquad c|W|_{1/2,\Gamma_{jk}}^2 \leq \left\langle l_0^{1/2} W, W \right\rangle_{\Gamma_{jk}} \leq C|W|_{1/2,\Gamma_{jk}}^2 \quad \text{for all } W \in S_h^0(\Gamma_{jk}).$$

We also note that by the assumptions on the coefficients defining $\tilde{l}_0$ in (2.2),

$$c \langle l_0 W, W \rangle_{\Gamma_{jk}} \leq \left\langle a^{-1} \tilde{l}_0 W, W \right\rangle_{\Gamma_{jk}} \leq C \langle l_0 W, W \rangle_{\Gamma_{jk}} \quad \text{for all } W \in S_h^0(\Gamma_{jk}).$$

By A.3 and a similar interpolation argument,

$$(3.5) \qquad \begin{aligned} c \left\langle l_0^{1/2} W, W \right\rangle_{\Gamma_{jk}} &\leq \left\langle a^{-1} \tilde{l}_0^{1/2} W, W \right\rangle_{\Gamma_{jk}} \\ &\leq C \left\langle l_0^{1/2} W, W \right\rangle_{\Gamma_{jk}} \quad \text{for all } W \in S_h^0(\Gamma_{jk}). \end{aligned}$$

We shall need several lemmas which will be used in the proof of the main theorem.

LEMMA 3.1. *For $V \in S_h^0(\Gamma_{jk})$, let $\bar{V}$ be the function which is equal to V on Γ_{jk} and is equal to zero on the remaining edges of Ω_i. Let $\tilde{v}$ denote the $\tilde{A}_i$-harmonic extension of $\bar{V}$ satisfying $\tilde{v} = \bar{V}$ on $\partial\Omega_i$ and*

$$(3.6) \qquad \tilde{A}_i(\tilde{v}, \phi) = 0 \quad \text{for all } \phi \in H_0^1(\Omega_i).$$

Then

$$c\tilde{A}_i(\tilde{v}, \tilde{v}) \leq \left\langle a^{-1} \tilde{l}_0^{1/2} V, V \right\rangle_{\Gamma_{jk}} \leq C\tilde{A}_i(\tilde{v}, \tilde{v}).$$

Proof. Let $D_i(u, \phi) = \int_{\Omega_i} \nabla u \cdot \nabla \phi \, dx$, and let $v^* \in H^1(\Omega)$ satisfy $v^* = \bar{V}$ on $\partial\Omega_i$, and $D_i(v^*, \phi) = 0$ for all $\phi \in H_0^1(\Omega_i)$. Then, using a trace inequality, A.3, and the uniform positive definiteness of the $\{a_{jk}^i\}$,

$$(3.7) \qquad \begin{aligned} |\bar{V}|_{1/2,\partial\Omega_i}^2 &\leq cD_i(\tilde{v}, \tilde{v}) \leq c\tilde{A}_i(\tilde{v}, \tilde{v}) \\ &\leq c\tilde{A}_i(v^*, v^*) \leq CD_i(v^*, v^*). \end{aligned}$$

Using A.3 and a well-known *a priori* inequality, we have $D_i(v^*, v^*) \leq C |\bar{V}|_{1/2,\partial\Omega_i}^2$, and hence

$$(3.8) \qquad c |\bar{V}|_{1/2,\partial\Omega_i}^2 \leq \tilde{A}_i(\tilde{v}, \tilde{v}) \leq C |\bar{V}|_{1/2,\partial\Omega_i}^2.$$

The lemma easily follows from the equivalence of norms (3.1) and (3.2) for functions $\bar{V}$, (3.4), (3.5) and (3.8). $\square$

We shall need some *a priori* estimates for discrete $\tilde{A}_i$-harmonic functions.

LEMMA 3.2. *Let* $W \in S_h(\Omega)$ *be discrete* $\tilde{A}_i$*-harmonic, then*

$$(3.9) \qquad \tilde{A}_i(W, W) \leq C \, |W|^2_{1/2, \partial\Omega_i} \, .$$

Furthermore, the following hold:

(i) *If* W *has mean value zero on* Ω_i, *then*

$$(3.10) \qquad c \, |W|^2_{1/2, \partial\Omega_i} \leq \tilde{A}_i(W, W).$$

(ii) *If* W *vanishes at the vertices of* Ω_i, *then*

$$(3.11) \qquad \tilde{A}_i(W, W) \leq C \sum_{jk \in \beta_i} \alpha_{jk} \left\langle a^{-1} \tilde{l}_0^{1/2} W, W \right\rangle_{\Gamma_{jk}} .$$

(iii) *If* W *is a linear function on each edge* $\Gamma_{jk} \subset \partial\Omega_i$, *then*

$$(3.12) \qquad \tilde{A}_i(W, W) \leq C \sum_{jk \in \beta_i} \alpha_{jk} (W(v_j) - W(v_k))^2.$$

Proof. The inequalities (3.9), (3.10), and (3.11) were essentially proved in [4]. For completeness we shall include a proof of (3.9) and (3.10) at the end of this section. Here we shall show how (3.11) and (3.12) follow from (3.9), (3.10) and earlier inequalities. We prove (3.11) as follows. Let Γ_{jk} be any edge of Ω_i and let W_{jk} be the discrete $\tilde{A}_i$-harmonic function which is equal to W on Γ_{jk} and vanishes on all the other edges of $\partial\Omega_i$. Clearly, $W = \sum_{jk \in \beta_i} W_{jk}$ and the triangle inequality yields

$$(3.13) \qquad \tilde{A}_i(W, W) \leq C \sum_{jk \in \beta_i} \tilde{A}_i(W_{jk}, W_{jk}).$$

Applying (3.9), the equivalence of norms (3.1) and (3.2) on the functions W_{jk}, (3.4) and (3.5) gives

$$(3.14) \qquad \begin{aligned} \tilde{A}_i(W_{jk}, W_{jk}) &\leq C \, |W_{jk}|^2_{1/2, \partial\Omega_i} \\ &\leq C \, |W_{jk}|^2_{1/2, \Gamma_{jk}} \leq C \left\langle a^{-1} \tilde{l}_0^{1/2} W, W \right\rangle_{\Gamma_{jk}} . \end{aligned}$$

Combining (3.13) and (3.14) proves (3.11).

We next prove (3.12). Applying (3.9) to the function $W - \beta$, where β is a constant to be determined later, gives

$$(3.15) \qquad \tilde{A}_i(W, W) = \tilde{A}_i(W - \beta, W - \beta) \leq C \, |W - \beta|^2_{1/2, \partial\Omega_i} \, .$$

If Ω_i is a triangle (resp. quadrilateral), let W^* be the linear (resp. bilinear) function on Ω_i which has the same boundary values as $W - \beta$. Choosing β so that the average of W^* on Ω_i is zero, applying a trace and Poincaré inequality gives

$$(3.16) \qquad |W - \beta|^2_{1/2, \partial\Omega_i} \leq C D_i(W^*, W^*).$$

An elementary calculation yields

$$(3.17) \qquad \begin{aligned} D_i(W^*, W^*) &\leq C \sum_{jk \in \beta_i} (W(v_j) - W(v_k))^2 \\ &\leq C \sum_{jk \in \beta_i} \alpha_{jk} (W(v_j) - W(v_k))^2. \end{aligned}$$

Thus (3.12) follows from (3.15), (3.16), and (3.17). $\square$

An important ingredient for our analysis is a certain type of discrete Sobolev inequality. Let $\hat{\Omega}$ be a polygonal domain which satisfies a cone condition with radius d and angle γ. Let $\tilde{S}_h$ for $0 < h < 1$ be any family of subspaces of $W_\infty^1(\hat{\Omega})$ satsifying the inverse inequality

$$(3.18) \qquad \|\nabla W\|_{L^\infty(\hat{\Omega})} \leq C_1 h^{-1} \|W\|_{L^\infty(\hat{\Omega})} \quad \text{for all } W \in \tilde{S}_h,$$

where $d \geq h$. We then have the following lemma.

LEMMA 3.3. *There exists a positive constant C independent of h and d and depending only on γ and C_1 in (3.18) such that*

$$(3.19) \quad \|W\|_{L^\infty(\hat{\Omega})}^2 \leq C \left(d^{-2} \|W\|_{L^2(\hat{\Omega})}^2 + \ln(d/h) D_{\hat{\Omega}}(W,W) \right) \quad \text{for all } W \in \tilde{S}_h,$$

where $D_{\hat{\Omega}}(\cdot,\cdot)$ denotes the Dirichlet form on $\hat{\Omega}$.

Various discrete Sobolev inequalities have appeared in the literature [3], [19]. Since the results in the literature do not correspond exactly to the given lemma, we shall include an elementary proof of the lemma after the proof of Theorem 1.

Some consequences of Lemma 3.3 which are important in our present considerations are the following discrete type Sobolev inequalities.

LEMMA 3.4. *Let W be in $S_h(\Omega_i)$.*
(i) *If $W(p) = 0$ for some point $p \in \bar{\Omega}_i$, then*

$$(3.20) \qquad \|W\|_{L^\infty(\Omega_i)}^2 \leq C(1 + \ln(d/h)) \tilde{A}_i(W,W).$$

(ii) *For any function $W \in S_h(\Omega_i)$,*

$$(3.21) \qquad \sum_{jk \in \beta_i} \alpha_{jk}(W(v_j) - W(v_k))^2 \leq C(1 + \ln(d/h)) \tilde{A}_i(W,W).$$

Proof. In order to prove (3.20) we first observe that by A.2, (3.19) is satisfied for $\tilde{S}_h = S_h(\Omega_i)$. Let α be the average value of W on Ω_i. Applying the Poincaré inequality yields

$$d^{-2} \|W - \alpha\|_{L^2(\Omega_i)}^2 \leq C D_i(W,W) \leq C \tilde{A}_i(W,W).$$

Thus, applying Lemma 3.3 to the function $W - \alpha$ gives

$$\|W - \alpha\|_{L^\infty(\Omega_i)}^2 \leq C(1 + \ln(d/h)) \tilde{A}_i(W,W).$$

Note that since $W(p) = 0$,

$$(3.22) \qquad |\alpha| \leq \|W - \alpha\|_{L^\infty(\Omega_i)},$$

and (3.20) follows by the triangle inequality.
Since

$$\sum_{jk \in \beta_i} \alpha_{jk}(W(v_j) - W(v_k))^2 \leq C \sum_{jk \in \beta_i} (W(v_j) - W(v_k))^2,$$

the inequality (3.21) follows by applying (3.20) to the function $W(x) - W(v_m)$, where v_m is a vertex of Ω_i. This completes the proof of the lemma. $\quad \square$

LEMMA 3.5. *Let* $W \in S_h(\Omega_i)$ *satisfy* $W = 0$ *on the vertices of* Ω_i *and let* $W_L \in S_h(\Omega_i)$ *be a discrete* $\tilde{A}_i$-*harmonic function which is linear on each edge* $\Gamma_{jk} \subset \partial\Omega_i$. *Then*

$$(3.23) \qquad \sum_{jk \in \beta_i} \alpha_{jk} \left\langle a^{-1} \tilde{l}_0^{1/2} W, W \right\rangle_{\Gamma_{jk}}$$
$$\leq C(1 + \ln(d/h)^2) \tilde{A}_i(W + W_L, W + W_L).$$

Proof. We shall first prove (3.23) in the case that $W_L = 0$. Let Γ_{jk} be any edge of Ω_i. It follows from (3.5), (3.4), (3.2), and (3.1) that

$$(3.24) \quad \alpha_{jk} \left\langle a^{-1} \tilde{l}_0^{1/2} W, W \right\rangle_{\Gamma_{jk}} \leq c \left\{ |W|_{1/2, \partial\Omega_i}^2 + \int_{\Gamma_{jk}} \frac{W(x)^2}{|x - v_k|} + \frac{W(x)^2}{|x - v_j|} ds(x) \right\}.$$

Let α be the average value of W on Ω_i. Applying (3.22), (3.10), and (3.20) leads to

$$|W|_{1/2, \partial\Omega_i}^2 \leq C \left(|\alpha|^2 + |W - \alpha|_{1/2, \partial\Omega_i}^2 \right)$$
$$\leq C \left(\|W - \alpha\|_{L^\infty(\Gamma_{jk})}^2 + \tilde{A}_i(W, W) \right)$$
$$\leq C(1 + \ln(d/h)) \tilde{A}_i(W, W).$$

Hence, it suffices to show that

$$(3.25) \qquad \begin{aligned} I(W) &\equiv I_1(W) + I_2(W) \\ &\equiv \int_{\Gamma_{jk}} \frac{W(x)^2}{|x - v_k|} ds(x) + \int_{\Gamma_{jk}} \frac{W(x)^2}{|x - v_j|} ds(x) \\ &\leq C(1 + \ln(d/h)^2) \tilde{A}_i(W, W). \end{aligned}$$

Without loss of generality, we assume that v_k is the origin and that Γ_{jk} is the line segment with $x_1 = 0$ and $x_2 \in [0, Y]$. Then,

$$I_1(W) = \int_0^Y \frac{W(0, y)^2}{y} dy.$$

Let y_1 be the y-value of the node on Γ_{jk} closest to zero. We bound the preceding integral by considering

$$(3.26) \qquad \int_0^Y \frac{W(0, y)^2}{y} dy = \int_0^{y_1} \frac{W(0, y)^2}{y} dy + \int_{y_1}^Y \frac{W(0, y)^2}{y} dy.$$

Note that by A.2, $ch \leq y_1 \leq Ch$. Therefore, by the mean-value theorem (using the hypothesis that $W(0, 0) = 0$),

$$(3.27) \qquad \int_0^{y_1} \frac{W(0, y)^2}{y} dy \leq Ch^2 \left\| \frac{\partial W(0, \cdot)}{\partial y} \right\|_{L^\infty([0, y_1])}^2 \leq C \|W\|_{L^\infty(\Omega_i)}^2,$$

where the second inequality used the inverse property for the subspace $S_h^0(\Omega_i)$. Hence, by (3.20) and (3.27),

$$\int_0^{y_1} \frac{W(0, y)^2}{y} dy \leq C(1 + \ln(d/h)) \tilde{A}_i(W, W).$$

For the second term in (3.26) we have

$$\int_{y_1}^Y \frac{W(0, y)^2}{y} dy \leq \|W\|_{L^\infty(\Omega_i)}^2 \int_{y_1}^Y \frac{dy}{y} \leq C(1 + \ln(d/h)^2) \tilde{A}_i(W, W).$$

Combining the above estimates gives a bound for the first term in (3.25); the second term is estimated similarly. Hence (3.23) follows in the case that $W_L = 0$.

To prove (3.23) in the general case, let $W_\perp$ be the function in $S_h(\Omega_i)$ which satisfies $W_\perp(v_m) = W_L(v_m)$ for all vertices v_m of $\partial\Omega_i$ and $\tilde{A}_i(W_\perp, \phi) = 0$ for all $\phi \in S_h(\Omega_i)$ with $\phi(v_m) = 0$ on all vertices of $\partial\Omega_i$. Notice that $W + W_L - W_\perp$ vanishes at the vertices of Ω_i. Applying the arithmetic-geometric mean inequality and the special case of (3.23) proved above gives

$$\alpha_{jk} \left\langle a^{-1}\tilde{l}_0^{1/2}W, W \right\rangle_{\Gamma_{jk}}$$
$$\leq 2\alpha_{jk} \left\langle a^{-1}\tilde{l}_0^{1/2}(W + W_L - W_\perp), (W + W_L - W_\perp) \right\rangle_{\Gamma_{jk}}$$
$$+ 2\alpha_{jk} \left\langle a^{-1}\tilde{l}_0^{1/2}(W_L - W_\perp), (W_L - W_\perp) \right\rangle_{\Gamma_{jk}}$$
$$\leq C\,(1 + \ln(d/h)^2)\tilde{A}_i((W + W_L - W_\perp), (W + W_L - W_\perp))$$
$$+ 2\alpha_{jk} \left\langle a^{-1}\tilde{l}_0^{1/2}(W_L - W_\perp), (W_L - W_\perp) \right\rangle_{\Gamma_{jk}}.$$

Since the functions $(W + W_L - W_\perp)$ and $W_\perp$ are orthogonal in the $\tilde{A}_i(\cdot, \cdot)$-inner product,

$$\tilde{A}_i((W + W_L - W_\perp), (W + W_L - W_\perp)) \leq \tilde{A}_i(W + W_L, W + W_L).$$

Thus to complete the proof of the lemma we need only show that

$$(3.28) \quad \begin{aligned} &\left\langle a^{-1}\tilde{l}_0^{1/2}(W_L - W_\perp), (W_L - W_\perp) \right\rangle_{\Gamma_{jk}} \\ &\qquad \leq C(1 + \ln(d/h)^2)\tilde{A}_i(W + W_L, W + W_L). \end{aligned}$$

Since $W_L - W_\perp$ vanishes at the vertices of Ω_i, applying inequality (3.24) and the subsequent arguments give

$$\left\langle a^{-1}\tilde{l}_0^{1/2}(W_L - W_\perp), (W_L - W_\perp) \right\rangle_{\Gamma_{jk}}$$
$$\leq C(1 + \ln(d/h))\tilde{A}_i(W_L - W_\perp, W_L - W_\perp) + I(W_L - W_\perp),$$

where I is defined in (3.25). Since $W_\perp$ is orthogonal to $W_L - W_\perp$ in the $\tilde{A}_i(\cdot, \cdot)$-inner product, we have in view of (3.12) and (3.21) that

$$\tilde{A}_i(W_L - W_\perp, W_L - W_\perp) \leq \tilde{A}_i(W_L, W_L)$$
$$\leq c \sum_{jk \in \beta_i} \alpha_{jk} \left[(W(v_j) + W_L(v_j)) - (W(v_k) + W_L(v_k))\right]^2$$
$$\leq C(1 + \ln(d/h))\tilde{A}_i(W + W_L, W + W_L).$$

Hence, in order to complete the proof of (3.28), it suffices to show that

$$(3.29) \qquad I(W_L - W_\perp) \leq C(1 + \ln(d/h)^2)\tilde{A}_i(W + W_L, W + W_L).$$

Now with I_1 and I_2 as in (3.25) we have by the arithmetic-geometric mean inequality,

$$(3.30) \qquad I_1(W_L - W_\perp) \leq 2I_1(W_\perp - W_\perp(v_k)) + 2I_1(W_L - W_L(v_k)).$$

For the first term on the right we apply (3.25) and then use the fact that $\tilde{A}_i(W_\perp, W_\perp)$ $\leq \tilde{A}_i(W + W_L, W + W_L)$ to obtain

$$I_1(W_\perp - W_\perp(v_k)) \leq C(1 + \ln(d/h)^2)\tilde{A}_i(W + W_L, W + W_L).$$

A simple calculation using the linearity of W_L on Γ_{jk} and (3.21) yields

$$I_1(W_L - W_L(v_k)) \leq C\alpha_{jk}(W_L(v_j) - W_L(v_k))^2$$
$$\leq C(1 + \ln(d/h))\tilde{A}_i(W + W_L, W + W_L).$$

Thus

$$I_1(W_L - W_\perp) \leq C(1 + \ln(d/h)^2)\tilde{A}_i(W + W_L, W + W_L).$$

Obviously, the same bound holds for $I_2(W_L - W_\perp)$, which completes the proof of (3.29) and hence the lemma. □

We are now in a position to prove Theorem 1.

Proof of Theorem 1. By the uniform positive definiteness of the matrices $\{a_{jk}\}$ and $\{a^i_{jk}\}$,

$$c\tilde{A}(W,W) \leq A(W,W) \leq C\tilde{A}(W,W) \quad \text{for all } W \in S^0_h(\Omega).$$

Hence, it suffices to compare the quadratic forms $\tilde{A}(\cdot, \cdot)$ with $B(\cdot, \cdot)$. As in Section 2, we decompose $W \in S^0_h(\Omega)$ into $W = W_P + W_E + W_V$. With $W_H = W_E + W_V$, we have (as noted in Section 2)

$$\tilde{A}(W,W) = \tilde{A}(W_P, W_P) + \tilde{A}(W_H, W_H)$$

and

$$B(W,W) = \tilde{A}(W_P, W_P) + B(W_H, W_H).$$

Hence, it suffices to compare $\tilde{A}(W_H, W_H)$ with $B(W_H, W_H)$. More specifically, the proof will be complete when we have shown that

(3.31) $$\tilde{A}(W_H, W_H) \leq CB(W_H, W_H),$$

and

(3.32) $$B(W_H, W_H) \leq C(1 + \ln(d/h)^2)\tilde{A}(W_H, W_H).$$

Consider a subdomain Ω_i. Using the arithmetic-geometric mean inequality, (3.11) and (3.12) yield

$$\tilde{A}_i(W_H, W_H) \leq 2(\tilde{A}_i(W_E, W_E) + \tilde{A}_i(W_V, W_V))$$
$$\leq C \sum_{jk \in \beta_i} \alpha_{jk} \left(\left\langle a^{-1}\tilde{l}_0^{1/2}W_E, W_E \right\rangle_{\Gamma_{jk}} + (W_V(v_j) - W_V(v_k))^2 \right).$$

Summing with respect to i gives (3.31). In view of (3.21) applied to W_V, and (3.23) applied to W_E and W_V (replacing W and W_L, respectively, in (3.23)), we have on each Ω_i,

$$\sum_{jk \in \beta_i} \alpha_{jk} \left(\left\langle a^{-1}\tilde{l}_0^{1/2}W_E, W_E \right\rangle_{\Gamma_{jk}} + (W_V(v_j) - W_V(v_k))^2 \right)$$
$$\leq C(1 + \ln(d/h)^2)\tilde{A}_i(W_H, W_H),$$

and summing with respect to i gives (3.32) which completes the proof of the theorem in the case where interior vertices are present.

We now turn to the case where all the vertices of the Ω_i lie on $\partial\Omega$. Since the function W_V vanishes at all the vertices, it follows that $W_V \equiv 0$ on Ω so that $W_H = W_E$ and hence the preconditioning form B simplifies to

$$B(W,W) = \tilde{A}(W_P, W_P) + \sum_{\Gamma_{jk}} \alpha_{jk} \left\langle a^{-1}\tilde{l}_0^{1/2}W_E, W_E \right\rangle_{\Gamma_{jk}}.$$

Reasoning as in the proof of the first part of this theorem, we have $\tilde{A}(W_E, W_E) \leq CB(W_E, W_E)$ which is a special case of (3.31). Hence, we need only show the sharper version of (3.32),

$$(3.33) \qquad\qquad B(W_E, W_E) \leq C\tilde{A}(W_E, W_E).$$

Let us first note that $W_E \equiv 0$ on $\partial\Omega$ and by our assumption that the vertices of the Ω_i are in this case fixed independent of h, there are a fixed number of interior edges Γ_{jk} on which possibly $W_E \neq 0$. Consider any interior edge Γ_{jk}. Identifying Γ_{jk} as two segments, say Γ_{jk}^1 and Γ_{jk}^2 with opposite orientations, it is not difficult to prove that either

(i) Γ_{jk} separates Ω into two polygonal domains, one of which (say $\tilde{\Omega}_1$) has a boundary $\partial\tilde{\Omega}_1$ consisting of Γ_{jk}^1 and nonempty parts of $\partial\Omega$ meeting each vertex of Γ_{jk}^1, or

(ii) Γ_{jk}^1 and Γ_{jk}^2 may be considered to be part of the boundary $\partial\tilde{\Omega}_1$ of a polygonal domain bounding a subdomain, say $\tilde{\Omega}_1$ of Ω, where Γ_{jk}^1 and Γ_{jk}^2 are separated (there is a positive distance with respect to arc length along $\partial\tilde{\Omega}_1$ between them) by components of $\partial\Omega$.

In either case, by equivalence of norms,

$$\alpha_{jk} \left\langle a^{-1}\tilde{l}_0^{1/2}W_E, W_E \right\rangle_{\Gamma_{jk}} \equiv \alpha_{jk} \left\langle a^{-1}\tilde{l}_0^{1/2}W_E, W_E \right\rangle_{\Gamma_{jk}^1} \leq c\,|W_E|_{1/2,\partial\tilde{\Omega}_1}^2.$$

Using a trace theorem for $\tilde{\Omega}_1$ we have

$$\alpha_{jk} \left\langle a^{-1}\tilde{l}_0^{1/2}W_E, W_E \right\rangle_{\Gamma_{jk}} \leq C\tilde{A}(W_E, W_E).$$

The inequality (3.33) follows after summing over all interior edges Γ_{jk}. This completes the proof of Theorem 1.

In the remainder of this section, we shall give the proof of Lemma 3.3 and the inequalities (3.9) and (3.10) of Lemma 3.2.

Proof of (3.9) and (3.10). It is not difficult to see (by scaling Ω_i to unit size and using A.3) that it suffices to prove (3.9) and (3.10) under the assumption that $d = 1$.

We first prove (3.9). Let $W \in S_h(\Omega_i)$ be $\tilde{A}_i$-discrete harmonic and w be the $\tilde{A}_i$-harmonic function defined by

$$\tilde{A}_i(w, \phi) = 0 \quad \text{for all } \phi \in H_0^1(\Omega_i), \qquad w = W \quad \text{on } \partial\Omega_i.$$

Using the well-known *a priori* inequality for harmonic functions, $\tilde{A}_i(w,w) \leq c\,|W|_{1/2,\partial\Omega_i}^2$, and the triangle inequality, it suffices to prove

$$(3.34) \qquad\qquad \tilde{A}_i(w - W, w - W) \leq c\,|W|_{1/2,\partial\Omega_i}^2.$$

Now from the definition of w and W it follows easily that

$$\tilde{A}_i(w - W, w - W) \leq \inf \tilde{A}_i(\Phi - w, \Phi - w),$$

with the infimum taken over functions $\Phi \in S_h(\Omega_i)$ with $\Phi = W$ on $\partial\Omega_i$. By well-known properties of $S_h(\Omega_i)$, we see that for $0 < \varepsilon < 1/2$,

$$\inf \tilde{A}_i(\Phi - w, \Phi - w) \leq Ch^{2\varepsilon} \|w\|^2_{H^{1+\varepsilon}(\Omega_i)}.$$

Now using a well-known *a priori* inequality (cf. [14], [16]) and an "inverse property" implied by A.2, we see that

$$h^{2\varepsilon} \|w\|^2_{H^{1+\varepsilon}(\Omega_i)} \leq Ch^{2\varepsilon} |W|^2_{H^{1/2+\varepsilon}(\partial\Omega_i)} \leq C |W|^2_{1/2,\partial\Omega_i},$$

which proves (3.34) and hence also (3.9).

We next prove (3.10). Let $W \in S_h(\Omega)$ have mean value zero on Ω_i. Applying a trace inequality gives $|W|^2_{1/2,\partial\Omega_i} \leq c\{D_i(W,W) + \|W\|^2_{L^2(\Omega)}\}$. Using a Poincaré inequality and the assumptions on the coefficients a^i_{jk} defining $\tilde{A}_i$ yields

$$|W|^2_{1/2,\partial\Omega_i} \leq CD_i(W,W) \leq C\tilde{A}_i(W,W),$$

which proves (3.10). $\square$

Proof of Lemma 3.3. Let W be in $\tilde{S}_h$ and x be a point of $\hat{\Omega}$ where $|W(x)| = \|W\|_{L^\infty(\hat{\Omega})}$. Let $\Lambda \subset \hat{\Omega}$ be a cone of radius d, angle γ, and vertex x. Without loss of generality assume that $x = 0$. For any point y in Λ, the Fundamental Theorem of Calculus gives that

$$W(0) = W(y) - \int_0^{|y|} \nabla W\left(\frac{ty}{|y|}\right) \cdot \frac{y}{|y|} \, dt.$$

Breaking up the integral into two regions and integrating over θ gives

$$\gamma |W(0)| \leq \left|\int_0^\gamma W(y(\theta)) \, d\theta\right| + \left|\int_0^\gamma \int_0^{\delta h} \nabla W\left(\frac{ty(\theta)}{|y(\theta)|}\right) \cdot \frac{y(\theta)}{|y(\theta)|} \, dt \, d\theta\right|$$
$$+ \left|\int_{\Lambda/\Lambda_{\delta h}} \nabla W(\varsigma) \cdot \frac{\varsigma}{|\varsigma|^2} \, d\varsigma\right|,$$

where $\delta > 0$ is to be chosen and $\Lambda_{\delta h}$ is the cone contained in Λ of radius δh, angle γ and vertex $(0,0)$. The second term above is estimated by

$$\left|\int_0^\gamma \int_0^{\delta h} \nabla W\left(\frac{ty(\theta)}{|y(\theta)|}\right) \cdot \frac{y(\theta)}{|y(\theta)|} \, dt \, d\theta\right| \leq \gamma\delta h \|\nabla W\|_{L^\infty(\hat{\Omega})} \leq c\delta |W(0)|$$

and then kicked back. Applying the Schwarz inequality to the first and the third terms gives

$$(3.35) \quad |W(0)| \leq C\left\{\left(\int_0^\gamma |W(y(\theta))|^2 \, d\theta\right)^{1/2} + \|\nabla W\|_{L^2(\hat{\Omega})} \left(\int_{\Lambda/\Lambda_{\delta h}} \frac{d\varsigma}{|\varsigma|^2}\right)^{1/2}\right\}$$
$$\leq C\left\{\left(\int_0^\gamma |W(y(\theta))|^2 \, d\theta\right)^{1/2} + \ln\left(\frac{d}{h}\right)^{1/2} \|\nabla W\|_{L^2(\hat{\Omega})}\right\}.$$

Squaring Eq. (3.35), applying the arithmetic-geometric mean inequality, multiplying by $|y|$ and integrating from 0 to d with respect to $|y|$ gives

$$\frac{d^2}{2} W(0)^2 \leq C\left(\|W\|^2_{L^2(\hat{\Omega})} + \frac{d^2}{2} \ln\left(\frac{d}{h}\right) D(W,W)\right)$$

which completes the proof of the lemma. $\square$

4. The Choice of the Coefficients Defining $\tilde{A}$; the "Fast Solvability" of the Subproblems.

As stated in Section 2, the application of our preconditioning algorithm involves solving the subproblems defined in (2.4), (2.6), (2.7) and (2.8). We have shown that Theorem 1 holds when the coefficients defining these problems are chosen among a rather large class of functions. Once a specific choice of coefficients has been made, one is then faced with the problem of choosing a method for solving the resulting equations. There are many ways in which this can be done. In general, a good choice may depend on the particular problem and the architecture of the particular computing machine. It is for this reason that we have stated Theorem 1 in a somewhat general form. In a forthcoming paper we shall give a detailed discussion of some possible approaches which are applicable to a large variety of boundary value problems and which have the additional feature that the subproblems (2.4), (2.6) and (2.8) may be solved by known "fast" methods. Our aim in this section is to give a very brief discussion of a very special case of one of these methods, which will be used in the calculation of the examples given in Section 6.
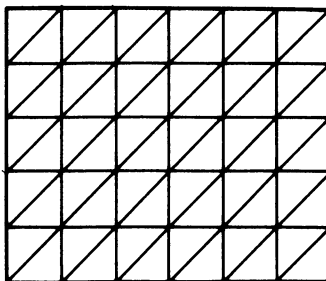

FIGURE 4.1. *A regular mesh on a rectangle.*

A globally efficient algorithm results when the original domain is split into subregions whose subproblems can be efficiently solved. The possibility of "fast solvability" for a subproblem is inherently linked to the coefficients defining $\tilde{A}_k$ and the geometry of the mesh on the subdomain. We shall first consider some simple subdomain problems where "fast solvers" are available. We will then show how these solvers can be used with Algorithm DD1 to solve much more complex problems on the original domain.

We shall begin by discussing "fast solvability" in the special case where a subdomain is a rectangle which has been triangulated with a regular mesh and then indicate how this may be extended to more general subregions. For m and n positive integers and $0 < h < 1$, let $\Omega_k = R$ denote the rectangle $R = \{(x_1, x_2) \mid 0 < x_1 < mh,\ 0 < x_2 < nh\}$. Assume that R has been triangulated with a uniform mesh of size h as indicated in Figure 4.1. Here, for simplicity of notation, the dependence of R on m, n, h and k has been omitted. The choice of m, n, and h will be clear in the context in which they are used. As before, let $S_h(R)$ be the space of piecewise linear functions defined relative to the given triangulation and $S_h^0(R)$ the subspace of $S_h(R)$ whose functions vanish on ∂R. We shall restrict ourselves to a particularly simple choice of the form $\tilde{A}_k$, namely a constant times the Dirichlet form for the Laplace operator,

$$(4.1) \qquad\qquad \tilde{A}_k(u, v) = q_k D_k(u, v),$$

where q_k is a constant which is to be chosen below.

Let us first consider the problem of inverting (2.4). Let β denote the vector whose components $\{\beta_j\}$ are the values of W_P at the interior nodal points of the triangulation, ordered say successively along columns. Then the corresponding set of linear equations can be written as

$$(4.2) \qquad\qquad\qquad M\beta = b,$$

where M is the $(m-1) \times (m-1)$ block tridiagonal matrix with block order $n-1$ given by

$$M = q_k \begin{pmatrix} T & -I & & & & \\ -I & T & -I & & & \\ & \cdot & \cdot & \cdot & & \\ & & \cdot & \cdot & \cdot & \\ & & & \cdot & \cdot & \cdot \\ & & & -I & T & -I \\ & & & & -I & T \end{pmatrix}.$$

Here I is the $(n-1) \times (n-1)$ identity matrix and T is the $(n-1) \times (n-1)$ matrix

$$T = \begin{pmatrix} 4 & -1 & & & & \\ -1 & 4 & -1 & & & \\ & \cdot & \cdot & \cdot & & \\ & & \cdot & \cdot & \cdot & \\ & & & \cdot & \cdot & \cdot \\ & & & -1 & 4 & -1 \\ & & & & -1 & 4 \end{pmatrix}.$$

The matrix M of course corresponds to the usual five-point centered difference approximation to $-\Delta$. It is well known that (4.2) may be solved by, for example, "fast" direct methods. An excellent discussion of some of these methods, which are also applicable to problems more general than (4.1), (4.2) on rectangles, may be found in Swarztrauber [18]. Let us just mention that using a discrete Fourier method or a cyclic reduction algorithm, the computational complexity of solving (4.2) on a serial machine is $O(mn \log(n))$. The Facr(l) algorithm which uses a combination of both of these is shown in [18] to have a computational complexity of $O(mn \log \log(n))$, where for convenience we have taken $n \le m$.

Let us now turn to the choice of the constant q_k. As noted in Remark 2.3, any $q_k > 0$ will satisfy the hypothesis of Theorem 1. It is obvious, however, that $\tilde{A}$ should model the original form A as closely as possible (with this simple choice of q_k). One prescription is as follows: Choose any point, say $\bar{x} \in R$, and let q_0 and q_1 be the smallest and largest eigenvalue of the matrix $\{a_{ij}(\bar{x})\}$ (which may be trivially calculated). We can choose q_k to be any number satisfying

$$(4.3) \qquad\qquad\qquad q_0 \le q_k \le q_1.$$

Obviously, the problem (2.8) may be handled in exactly the same fashion since it also can be reduced to solving (4.2) with appropriate b.

We now turn to problem (2.6). Let Γ_{ij} be an edge which is on the common boundary segment for the subdomains say Ω_k and Ω_l. Since the mesh is equally spaced on Γ_{ij}, we take $a = 1$ (see (2.2)) on Γ_{ij} and hence, $\tilde{l}_0^{1/2} = l_0^{1/2}$. Without loss of generality, assume that there are $n-1$ (as opposed to $(m-1)$) interior nodes

on Γ_{ij}. Let β be the vector of length $n-1$ whose components $\{\beta_p\}$ are the nodal values of W_E (the solution of (2.6)) on Γ_{ij} and let $\{\Phi_p\}$ be the nodal basis for $S_h^0(\Gamma_{ij})$. The problem of computing β is the same as the matrix problem $N\beta = \gamma$, where N is given by

$$(4.4) \qquad N_{pq} = \alpha_{ij} \left\langle a^{-1} \tilde{l}_0^{1/2} \Phi_p, \Phi_q \right\rangle_{\Gamma_{ij}}.$$

Since the nodes are equally spaced, the eigenvectors of N are given by

$$(4.5) \qquad \Psi_p = \begin{pmatrix} \sin(\pi p/n) \\ \sin(2\pi p/n) \\ \vdots \\ \sin((n-1)\pi p/n) \end{pmatrix}.$$

The eigenvalues for N are then given by

$$(4.6) \qquad \lambda_p = \alpha_{ij} \sqrt{\frac{(2 - 2\cos(\pi p/n))(4 + 2\cos(\pi p/n))}{6}}.$$

Thus the computation of nodal values of W_E reduces to the expansion of γ in terms of the eigenvectors (4.5), the division of the resulting coefficients by the eigenvalues (4.6), and the evaluation of the resulting expansion (with the divided coefficients) at the nodal values. Both the expansion of γ in terms of the eigenvectors and the subsequent evaluation at the nodes reduce to discrete sine transforms. Using the Fast Fourier Transform, the sine transforms and hence the solution of (2.6) on each edge Γ_{ij} can be computed in computational work on the order of $O(n \log(n))$.

Finally, we come to the choice of the coefficients α_{ij} and solution of the difference equation (2.7). Again, let Γ_{ij} be an interior edge which is a common boundary segment of the two subdomains Ω_k and Ω_l. We take

$$(4.7) \qquad \alpha_{ij} = q_k + q_l.$$

We solve the difference equations (2.7) by applying some standard method, for example sparse Gaussian elimination techniques [8], [9]. We emphasize that (2.7) may be solved in parallel with (2.6) and if the number of internal nodes is reasonable, the cost of solving (2.7) will be negligible.

We now turn our attention to the more general situation where the subregions may be quadrilaterals. For simplicity of presentation, let us indicate by an example how the subproblems (2.4), (2.6) and (2.8) may be set up so as to utilize the fast solution methods previously mentioned. Consider the region given in Figure 4.2. The key idea here is to set up a mesh on each subdomain Ω_k which is topologically equivalent to a regular mesh on R. Then, the coefficients a_{ij}^k defining $\tilde{A}_k$ can be chosen so that the matrix problem for (2.4) and (2.8) can be solved by fast direct methods. To make this example nontrivial we shall consider subspaces which are somewhat refined near the nonconvex corners of the domain.

We start by imposing a "radial like" rectangular refinement around these corners which is then extended to the rest of the domain as illustrated in Figure 4.3. The final triangular grid is shown in Figure 4.4 and is formed by subdividing each quadrilateral in Figure 4.3 into two triangles.

Obviously, the triangulation of each subdomain Ω_k in Figure 4.4 has the same topological nodal structure as that of R in Figure 4.1, and we can order the nodal points in precisely the same fashion. In fact, the triangulations are equivalent in
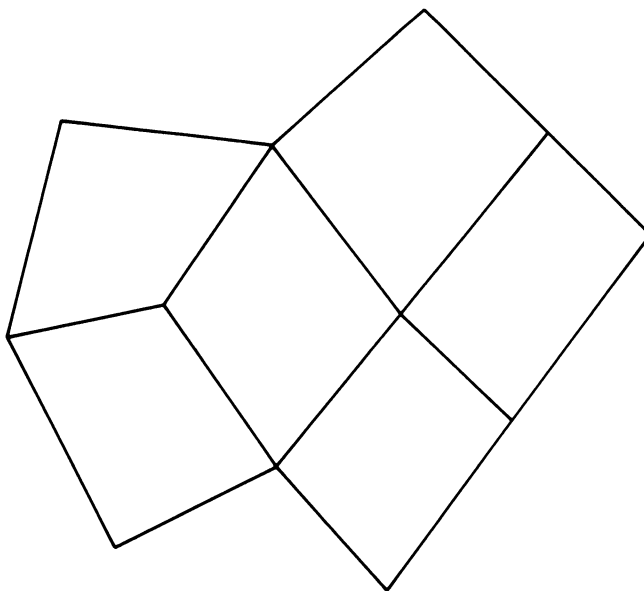
FIGURE 4.2. *A more general domain.*
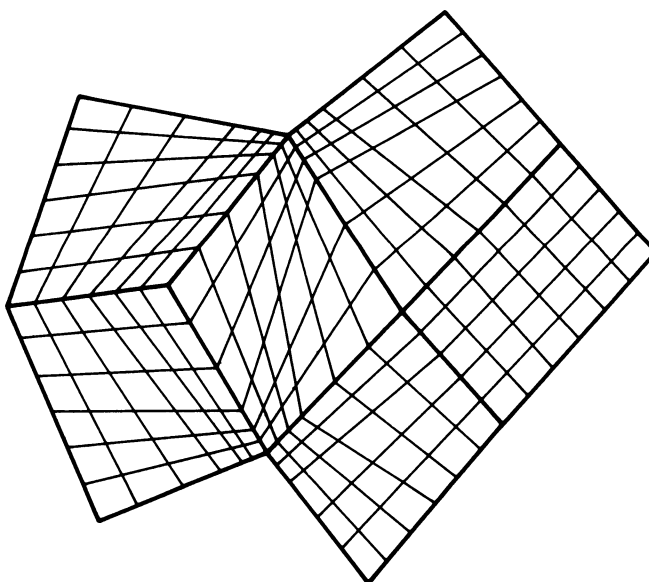
FIGURE 4.3. *The domain with rectangular subdivision.*

the sense that there exists a nondegenerate piecewise linear mapping of the triangulation of Figure 4.1 onto any of the subdomain triangulations of Figure 4.4. It can be shown that any simply connected piecewise smooth domain may be triangulated in an analogous way and hence our assumption that Ω_k is a quadrilateral is just for convenience of presentation.

We shall use the piecewise linear mapping mentioned above to define the coefficients a_{ij}^k . Fix k and let T be the corresponding piecewise linear mapping of R
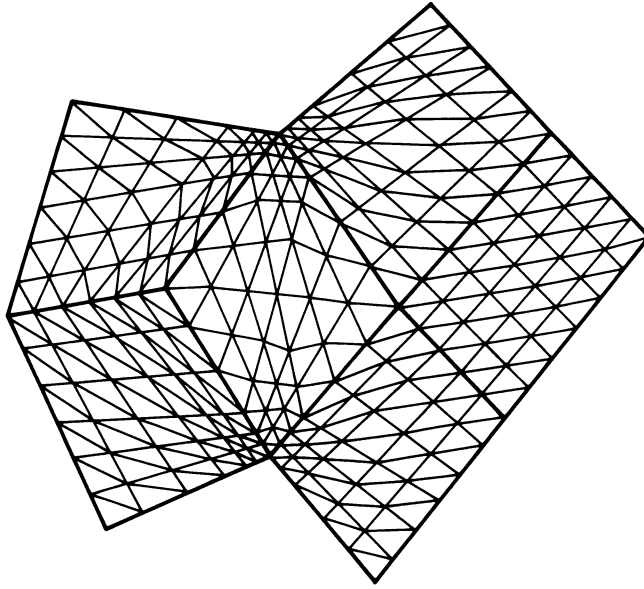
FIGURE 4.4. *The domain with triangular subdivision.*

onto Ω_k. We clearly have, by a change of variable, that

$$(4.8) \qquad q_k \int_R |\nabla v|^2 \, dx = \sum_{i,j=1}^{2} \int_{\Omega_k} a_{ij}^k \frac{\partial v(T^{-1}(x))}{\partial x_i} \frac{\partial v(T^{-1}(x))}{\partial x_j} \, dx$$

$$\equiv \tilde{A}_k(v,v),$$

where $\{a_{ij}^k\}$ is a piecewise constant 2×2 matrix (here q_k is a positive constant to be defined later). We use (4.8) to define a_{ij}^k. If $\{\Phi_i\}$ denotes the usual nodal basis for $S_h(\Omega)$ restricted to Ω_k, then $\{\Psi_i \equiv \Phi_i \circ T\}$ is the usual nodal basis for the subspace $S_h(R)$. Thus, the form $\tilde{A}$ applied to basis functions is given by

$$(4.9) \qquad \tilde{A}_k(\Phi_i, \Phi_j) = q_k \int_R \nabla \Psi_i \cdot \nabla \Psi_j \, dx.$$

In light of (4.9), it is clear that the coefficients a_{ij}^k need never be computed. Furthermore, the matrix problem for the solution of (2.4) and (2.8) is given by (4.2) and hence can be fast solved.

The constant q_k may be chosen in the following manner: Let $\bar{x}$ be any fixed point of Ω_k. By a change of variable, we clearly have

$$\int_{\Omega_k} a_{ij} \frac{\partial v}{\partial x_i} \frac{\partial v}{\partial x_j} \, dx = \int_R \bar{a}_{ij} \frac{\partial v(T(x))}{\partial x_i} \frac{\partial v(T(x))}{\partial x_j} \, dx.$$

Let q_1 and q_0 denote the largest and smallest eigenvalue of the matrix $\{\bar{a}_{ij}(\bar{x})\}$. Then we may choose q_k to be any number

$$(4.10) \qquad q_0 \leq q_k \leq q_1.$$

This again is a trivial calculation. We remark that it is easily seen that $\tilde{A}_k$ corresponds, in this case, to a form whose coefficients a_{ij}^k are piecewise constant functions on the triangles of the triangulation of Ω_k.

We now turn to defining $\tilde{l}_0$ and solving the problem (2.6). As in the definition of the coefficients a_{ij}^k above, we shall use a "mapping" technique to define the coefficient a in (2.2). Fix ij and assume that Γ_{ij} has $n-1$ interior nodes and is a common boundary segment between subdomains Ω_k and Ω_l. Let L denote the line segment $[0, nh]$ which has an associated equally-spaced mesh with n equal segments. There is a piecewise linear mapping T which maps the mesh of L onto the nodes of Γ_{ij}. We clearly have, by a variable change,

$$\langle v \circ T, v \circ T \rangle_L = \langle a^{-1} v, v \rangle_{\Gamma_{ij}}$$

and

$$\left\langle \frac{dv(T(x))}{dx}, \frac{dv(T(x))}{dx} \right\rangle_L = \langle av', v' \rangle_{\Gamma_{ij}} = \left\langle a^{-1} \tilde{l}_0 v, v \right\rangle_{\Gamma_{ij}}$$

for appropriate piecewise constant a. Let β be the vector whose components are the nodal values of W_E on Γ_{ij}. Then, β is the solution of

(4.11) $N\beta = \gamma$

for an appropriate right-hand side vector γ, where the matrix N is given by (4.4). The discussion following (4.4) describes an efficient procedure for solving (4.11). Again, the coefficient a is a theoretical device and need never be actually computed in the implementation of Algorithm DD1.

Finally, the coefficients in (2.7) are chosen in exactly the same way as for the case of a rectangle, i.e., they are defined by (4.7) and (4.10).

5. Matrix Representation of the Operators. In this section we will describe the action of inverting the preconditioner B (given by Algorithm DD1) in terms of block matrices. It will be shown that B has a special structure and that the process for solving $B\alpha = \beta$ previously described may also be seen to be a block Gauss elimination process with an appropriate basis.

We consider the inversion of B in terms of basis functions of the following form:

1. $\{\Phi_P^i\}$ is the set of basis functions for $\bigcup S_h^0(\Omega_j)$. These functions correspond to the usual nodal basis of functions which are one on one of the nodes interior to some subdomain and zero on all of the remaining nodes.

2. $\{\Phi_E^i\}$ is the set of basis functions corresponding to the variables which lie on the edges of the subregions (excluding the corners). This basis consists of the usual nodal basis functions which are one on one of the edge nodes and zero on all of the remaining nodes of Ω^h.

3. $\{\Phi_V^i\}$ is a basis for the functions which are linear on the edges of the subregions. The function Φ_V^i is one on v_i, zero on all other vertices v_j with $j \neq i$, zero on all of the nodes which are interior to any of the subregions, and extended linearly along the edges of the subdomains.

It is easily seen that the above collection of functions give rise to a basis for $S_h^0(\Omega)$. As usual, we decompose functions in $S_h^0(\Omega)$ in terms of linear combinations of these basis functions and an arbitrary function in the subspace is represented by a vector of its coefficients. We order these vectors as follows:

$$\alpha = \begin{pmatrix} v_V \\ v_E \\ v_P \end{pmatrix}$$

where v_P, v_E, and v_V represent coefficients for basis functions of type 1, 2, and 3 respectively. In terms of block matrices the system corresponding to B is then

$$(5.1) \qquad \begin{pmatrix} B_{VV} & B_{VE} & B_{VP} \\ B_{VE}^t & B_{EE} & B_{EP} \\ B_{VP}^t & B_{EP}^t & B_{PP} \end{pmatrix} \begin{pmatrix} v_V \\ v_E \\ v_P \end{pmatrix} = \begin{pmatrix} b_V \\ b_E \\ b_P \end{pmatrix}.$$

The first step of Algorithm DD1 corresponds to computing the solution of $B_{PP}^{-1} b_P$. Using this solution to calculate the data for the inversions of Steps 2 and 3 of Algorithm DD1 corresponds to eliminating two blocks in the third column of (5.1) to obtain

$$(5.2) \qquad \begin{pmatrix} B_{VV} - B_{VP}B_{PP}^{-1}B_{VP}^t & B_{VE} - B_{VP}B_{PP}^{-1}B_{EP}^t & 0 \\ B_{VE}^t - B_{EP}B_{PP}^{-1}B_{VP}^t & B_{EE} - B_{EP}B_{PP}^{-1}B_{EP}^t & 0 \\ B_{VP}^t & B_{EP}^t & B_{PP} \end{pmatrix} \begin{pmatrix} v_V \\ v_E \\ v_P \end{pmatrix}$$
$$= \begin{pmatrix} b_V - B_{VP}B_{PP}^{-1}b_P \\ b_E - B_{EP}B_{PP}^{-1}b_P \\ b_P \end{pmatrix}.$$

We note that the inversions of Steps 2 and 3 are problems which can be solved independently and involve the edge and vertex basis functions respectively. This fact means that the two blocks $B_{VE} - B_{VP}B_{PP}^{-1}B_{EP}^t$ and $B_{VE}^t - B_{EP}B_{PP}^{-1}B_{VP}^t$ in matrix (5.2) must be identically zero. Furthermore, the upper two diagonal blocks of (5.2) must correspond to

$$M_{kl} = \sum_{\Gamma_{ij}} \alpha_{ij} (\Phi_V^k(v_i) - \Phi_V^k(v_j))(\Phi_V^l(v_i) - \Phi_V^l(v_j))$$

and

$$N_{kl} = \sum_{\Gamma_{ij}} \alpha_{ij} \left\langle \tilde{l}_0^{1/2} \Phi_E^k, \Phi_E^l \right\rangle_{\Gamma_{ij}},$$

respectively. Steps 4 and 5 of the algorithm correspond to backsolving (5.2) once the values of v_V and v_E are known.

6. Numerical Experiments.

In this section, we shall present some results of numerical experiments which illustrate the convergence properties of the preconditioning algorithm using DD1 as a preconditioner discussed in Section 2, when used in conjunction with the conjugate gradient method. To this end we shall report a number of parameters which measure or effect the convergence of the scheme. We shall, for example, compute the condition number K of the preconditioned system. In some examples, we shall also report n, the number of iterations required to reduce the matrix norm $(Ax \cdot x)^{1/2}$ of the error $E_n = U - U_n$ below an indicated tolerance. Here U is a randomly generated solution of the matrix equations normalized so that $-1 \leq U \leq 1$ and U_n is the approximation to U obtained using n steps of the iterative algorithm.

The examples were chosen to illustrate the effectiveness of the algorithm on problems with both smooth and discontinuous coefficients on domains with different geometries. In all of these examples subspaces $S_h^0(\Omega)$ of piecewise linear functions defined on a quasi-uniform mesh of size h were used and the algorithm was applied to solve the finite element equations approximating the solution of an elliptic problem of the form

$$Lu = f \quad \text{in } \Omega, \qquad u = 0 \quad \text{on } \partial\Omega.$$
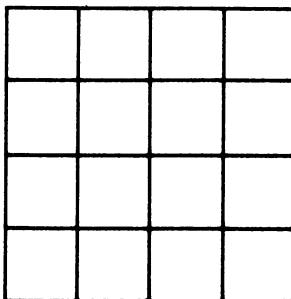
FIGURE 6.1. *Subdivision of the square.*

The procedure discussed in Section 4 for choosing the coefficients of the preconditioning form and solving the related subproblems (in particular problems (2.4), (2.6), and (2.8) by fast methods) was used throughout this section.

Example 1. For our first example we take $L = -\Delta$, the Laplace operator (i.e., $a_{11} = a_{22} = 1$ and $a_{12} = a_{21} = 0$), Ω the unit square and $S_h^0(\Omega)$ the piecewise linear functions on a regular mesh of size h which vanish on $\partial\Omega$. Note that although, in this very simple case, the resulting equations may be fast solved on a serial machine by any of the methods discussed in Section 4, the algorithm used would be particularly appealing for a machine with parallel architecture. We will also use this example as a benchmark for the more complicated examples to follow. We subdivide the domain Ω into sixteen subregions as indicated in Figure 6.1.

TABLE 6.1. *Iterative convergence for Example* 1.

Iteration	A-error	A-error Average Reduction	Max-error	Max-error Average Reduction
1	9.5×10^{-2}	.095	6.6×10^{-1}	.66
2	5.5×10^{-2}	.23	5.4×10^{-1}	.74
3	2.4×10^{-2}	.29	1.8×10^{-1}	.56
4	4.8×10^{-3}	.26	4.2×10^{-2}	.45
5	1.2×10^{-3}	.26	9.9×10^{-3}	.40
6	6.7×10^{-4}	.30	9.6×10^{-3}	.46
7	3.6×10^{-4}	.32	3.2×10^{-3}	.44
8	9.5×10^{-5}	.31	7.5×10^{-4}	.41
9	1.6×10^{-5}	.29	1.2×10^{-4}	.37
10	5.0×10^{-6}	.30	5.6×10^{-5}	.38
11	3.3×10^{-6}	.32	4.2×10^{-5}	.40

Table 6.1 illustrates the iterative reduction rates for Example 1 when $h = 1/32$. The table lists the total reduction and average reduction rate as a function of the number of iterations in the matrix norm $(Ax \cdot x)^{1/2}$ and the maximum norm. These reductions are normalized so that the initial error is unity. We see, for example, that a reduction of .0001 in the A norm (resp. maximum norm) requires only 8 (resp. 10) iterations.

$\mu = 300$	$\mu = 0.0001$	$\mu = 31400$	$\mu = 5$
$\mu = 0.05$	$\mu = 8$	$\mu = 0.07$	$\mu = 2700$
$\mu = 10^6$	$\mu = 0.1$	$\mu = 200$	$\mu = 9$
$\mu = 1$	$\mu = 6000$	$\mu = 4$	$\mu = 140000$

FIGURE 6.2. *The coefficients for Example* 2.

To more fully illustrate the convergence behavior of the method on this problem we consider Table 6.2 which gives the condition number and theoretical reduction ρ^{**} for Example 1 as a function of the mesh size h. We note that the theoretical reduction gives a pessimistic bound on the worst-case convergence in the A norm. For example, the actual reduction rate given in Table 6.1 for 11 iterations was .32 which is considerably better than the theoretical rate of .45 given in Table 6.2 for $h = 1/32$. We also compare the condition number to the function $(\log_2 1/h)^2/3.5$ and hence demonstrate the log-squared growth in the condition number which suggests that Theorem 1 is sharp.

TABLE 6.2. *Condition number and theoretical reduction for Example* 1.

h	K	$(\log_2 1/h)^2/3.5$	ρ^{**}
1/8	3.0	2.6	.27
1/16	4.5	4.6	.36
1/32	7.0	7.1	.45
1/64	10.3	10.3	.52
1/128	14.0	14.0	.58
1/256	18.6	18.3	.62

Example 2. In this example, Ω is the unit square and the subdomains were taken as in Example 1 (see Figure 6.1). The operator L is taken to have coefficients which have jump discontinuities across the subdomain boundaries. More specifically, we take $a_{11} = a_{22} = \mu$ and $a_{12} = a_{21} = 0$, where μ is the randomly chosen piecewise constant function on the subdomains as indicated in Figure 6.2. Table 6.3 gives the results for the condition number of the preconditioned system and the theoretical reduction factors for this example as a function of h. Note that the results differ negligibly from those given for the Laplacian in Table 6.2. We remark that similar results were obtained in tests with other randomly chosen coefficients. This indicates that the iterative method DD1 will be extremely effective on interface

---

**It is well known (cf. [17]) that the error for preconditioned conjugate gradient iteration satisfies $(AE_n \cdot E_n) \leq 4\rho^{2n}(AE_0 \cdot E_0)$, where the reduction factor ρ is given by $\rho \equiv (\sqrt{K} - 1)/(\sqrt{K} + 1)$.

TABLE 6.3. *Condition number and theoretical reduction for Example 2.*

h	K	$(\log_2 1/h)^2/3.2$	$\rho^\dagger$
1/8	3.0	2.8	.27
1/16	5.0	5.0	.38
1/32	7.7	7.8	.47
1/64	11.2	11.3	.54
1/128	15.2	15.3	.59

problems, even when the coefficients change drastically across interfaces, as long as the subdomain boundaries align with the interface boundaries.

Example 3. Here we take L to be an operator with smoothly varying coefficients. The region Ω and the subdomains are taken exactly as in Example 1. The coefficients are defined by

$$(6.1) \qquad \begin{aligned} a_{11} &= 1 + 4(x^2 + y^2), & a_{12} &= 3xy, \\ a_{21} &= 3xy, & a_{22} &= 1 + 11(x^2 + y^2). \end{aligned}$$

This example illustrates that the introduction of more subdomains allows the preconditioner to more closely model the differential operator. Table 6.4 gives convergence results for the above problem as a function of n_r, the number of subdomains used. The coefficients defining the preconditioner were chosen as in Section 4. More precisely, we set $q_k = \sqrt{q_0 q_1}$, where q_0 and q_1 are as in (4.3) and the point $\bar{x}$ (see (4.3)) is chosen as the center of the subdomain. All computations in Table 6.4 were made for $h = 1/64$.

Note that if the Laplace operator on the original domain was used as a preconditioner for the variable coefficient problem (6.1), then the condition number would be larger than 55. In contrast, the results given in Table 6.4 show considerable improvement even when relatively few subdomains are used.

Table 6.4 also illustrates the fact that the theoretical reduction (computed from the eigenvalues of the discrete system) provides a useful bound for the actual rate of convergence. We finally included n, the number of iterations required to reduce the matrix norm $(Ax \cdot x)^{1/2}$ of the error $E_n = U - U_n$ by a factor of .0001.

TABLE 6.4. *Convergence results for Example* 3.

n_r	K	$\rho^\dagger$	Observed Reduction	n
4	42.3	.73	.56	17
16	17.5	.61	.51	14
64	11.1	.54	.45	12
256	7.4	.46	.43	11

Example 4. In this example, we consider an interface problem where the interface separates two domains with irregular geometries. The domain Ω is again the unit square subdivided into sixteen subdomains as illustrated in Figure 6.3. The space $S_h^0(\Omega)$ is taken to be piecewise linear functions defined on an irregular mesh. A

---

† See footnote **.
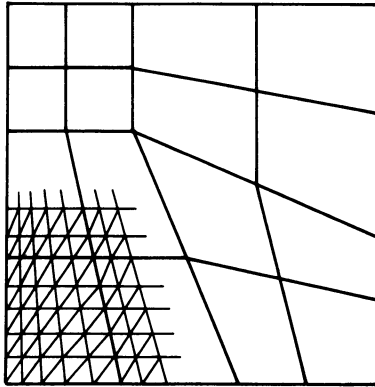
FIGURE 6.3. *The irregular geometry of Example* 4.
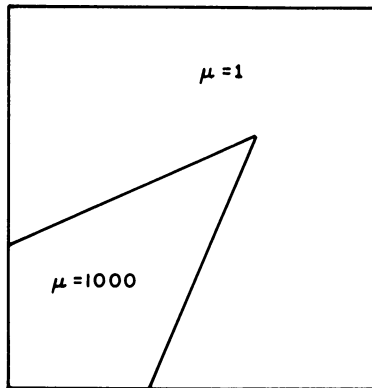
FIGURE 6.4. *The coefficients of Example* 4.

portion of this mesh is illustrated by the fine triangulation in Figure 6.3. Again the coefficients of L are piecewise constant functions defined by $a_{11} = a_{22} = \mu$ and $a_{12} = a_{21} = 0$ where μ is given by Figure 6.4.

Results for this problem are given in Table 6.5. A comparison with Table 6.2 indicates that the irregular geometry of this example only increased the condition number by about 2.5. This results in less than a factor of two increase in the number of iterations required for a given accuracy. We again remark that fast methods were used to solve the subproblems (2.4), (2.6) and (2.8) required for the preconditioner.

TABLE 6.5. *Convergence results for Example* 4.

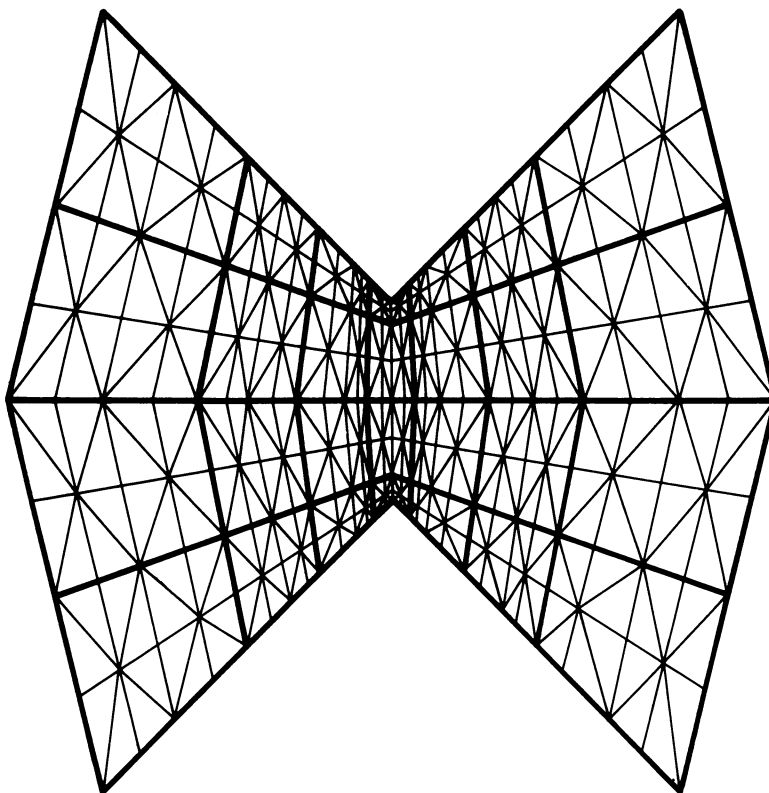| h	K	$\rho^{\dagger}$	Observed Reduction	n
1/8	5.6	.41	.32	9
1/16	10.8	.53	.45	13
1/32	17.6	.62	.51	15
1/64	25.4	.67	.55	16

† See footnote **.

FIGURE 6.5. *The mesh and subdomain structure for Example* 5.

Example 5. In this example, we illustrate Algorithm DD1 applied to the solution of a problem on a polygonal domain with nonconvex corners. The mesh and subdomain structure were chosen as illustrated in Figure 6.5. Note the mild refinement near the nonconvex corners of the domain. For the operator L we use the Laplacian as in Example 1. The results for this case are given in Table 6.6. This example illustrates some of the power and flexibility of this algorithm which will be more fully developed in later papers.

TABLE 6.6. *Convergence results for Example* 5.

Number of Unknowns	K	$\rho^\dagger$	Observed Reduction	n
405	8.54	.49	.45	12
1705	14.4	.58	.50	14
6993	20.6	.64	.55	15

Example 6. As a final example, we compare the preconditioner DD1 with the somewhat simpler preconditioner discussed in Remark 2.6. In particular, we replace the global difference equation (2.7) by a weighted identity (2.9) on the coarse mesh points. For this example, we use the Laplace operator on the square as in Example 1. This example illustrates the effect that the diameter of the subdomains has on

---

† See footnote **.

the actual condition number observed in practice. Table 6.7 compares the condition numbers for DD1 and the preconditioner of Remark 2.6 as a function of d. Observe the clear superiority of DD1 in applications with many subdivisions.

TABLE 6.7. *Comparison of* DD1 *and the preconditioner of Remark* 2.6.

d	K(DD1)	K (Remark 2.6)	h
1/2	6.3	6.3	1/16
1/4	7.0	10.5	1/32
1/8	7.5	26.6	1/64
1/16	7.5	96.9	1/128

Department of Mathematics
Cornell University
Ithaca, New York 14853

Department of Applied Mathematics
Brookhaven National Laboratory
Upton, New York 11973

Department of Mathematics
Cornell University
Ithaca, New York 14853

1. P. E. BJØRSTAD & O. B. WIDLUND, "Solving elliptic problems on regions partitioned into substructures," *Elliptic Problem Solvers II* (G. Birkhoff and A. Schoenstadt, eds.), Academic Press, New York, 1984, pp. 245–256.

2. P. E. BJØRSTAD & O. B. WIDLUND, "Iterative methods for the solution of elliptic problems on regions partitioned into substructures." (Preprint.)

3. J. H. BRAMBLE, "A second order finite difference analogue of the first biharmonic boundary value problem," *Numer. Math.*, v. 9, 1966, pp. 236–249.

4. J. H. BRAMBLE, J. E. PASCIAK & A. H. SCHATZ, "An iterative method for elliptic problems on regions partitioned into substructures," *Math. Comp.*, v. 46, 1986, pp. 361–369.

5. B. L. BUZBEE & F. W. DORR, "The direct solution of the biharmonic equation on rectangular regions and the Poisson equation on irregular regions," *SIAM J. Numer. Anal.*, v. 11, 1974, pp. 753–763.

6. B. L. BUZBEE, F. W. DORR, J. A. GEORGE & G. H. GOLUB, "The direct solution of the discrete Poisson equation on irregular regions," *SIAM J. Numer. Anal.*, v. 8, 1971, pp. 722–736.

7. Q. V. DIHN, R. GLOWINSKI & J. PÉRIAUX, "Solving elliptic problems by domain decomposition methods," *Elliptic Problem Solvers II* (G. Birkhoff and A. Schoenstadt, eds.), Academic Press, New York, 1984, pp. 395–426.

8. S. C. EISENSTAT, M. C. GURSKY, M. H. SCHULTZ & A. H. SHERMAN, *Yale Sparse Matrix Package, I. The Symmetric Codes*, Dept. of Computer Science Report No. 112, Yale University.

9. A. GEORGE & J. W. H. LIU, *User Guide for SPARSPAK*, Dept. of Computer Science Report No. CS-78-30, Waterloo University.

10. G. H. GOLUB & D. MEYERS, "The use of preconditioning over irregular regions," *Proc. Sixth Internat. Conf. on Computing Methods in Science and Engineering.* (Preprint.)

11. P. GRISVARD, *Elliptic Problems in Non Smooth Domains*, Pitman, Boston, 1985.

12. M. R. HESTENES, *The Conjugate Gradient Method for Solving Linear Systems*, Proc. Sympos. Appl. Math., vol. 6, Amer. Math. Soc., McGraw-Hill, 1956, pp. 83–102.

13. S. G. KREIN & Y. I. PETUNIN, "Scales of Banach spaces," *Russian Math. Surveys*, v. 21, 1966, pp. 85–160.

14. J. L. LIONS & E. MAGENES, *Problèmes aux Limites non Homogènes et Applications*, Dunod, Paris, 1968.

15. D. G. LUENBERGER, *Introduction to Linear and Nonlinear Programming*, Addison-Wesley, Reading, Mass., 1973.

16. J. NEČAS, *Les Méthodes Directes en Théorie des Équations Elliptiques*, Academia, Prague, 1967.

17. W. M. PATTERSON, 3RD, *Iterative Methods for the Solution of a Linear Operator Equation in Hilbert Space—A Survey*, Lecture Notes in Math., Vol. 394, Springer-Verlag, New York, 1974.

18. P. N. SWARZTRAUBER, "The methods of cyclic reduction, Fourier analysis and the FACR algorithm for the discrete solution of Poisson's equation on a rectangle," *SIAM Rev.*, v. 19, 1977, pp. 490–501.

19. V. THOMÉE, *Galerkin Finite Element Methods for Parabolic Problems*, Lecture Notes in Math., Vol. 1054, Springer-Verlag, New York, 1984.