MATHEMATICS OF COMPUTATION
VOLUME 49, NUMBER 179
JULY 1987, PAGES 1–16

The Construction of Preconditioners for Elliptic Problems by Substructuring. II

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

Abstract. We give a method for constructing preconditioners for the discrete systems arising in the approximation of solutions of elliptic boundary value problems. These preconditioners are based on domain decomposition techniques and lead to algorithms which are well suited for parallel computing environments. The method presented in this paper leads to a preconditioned system with condition number proportional to d/h where d is the subdomain size and h is the mesh size. These techniques are applied to singularly perturbed problems and problems in three dimensions. The results of numerical experiments illustrating the performance of the method on problems in two and three dimensions are given.

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.

For $N = 2$ or $N = 3$, let Ω be a bounded domain in R^N with a piecewise smooth boundary $\partial\Omega$. As a model problem for a second-order uniformly elliptic equation we shall consider the Dirichlet problem

$$\begin{aligned} Lu &= f \quad \text{in } \Omega, \\ u &= 0 \quad \text{on } \partial\Omega, \end{aligned} \tag{1.1}$$

where

$$Lv = -\sum_{i,j=1}^{N} \frac{\partial}{\partial x_i}\left(a_{ij}\frac{\partial v}{\partial x_j}\right) + av, \tag{1.2}$$

with a_{ij} symmetric, uniformly positive definite and bounded above on Ω. For ease of exposition, we assume that either $a \equiv 0$ or a is bounded above and below by

Received February 21, 1986.
1980 *Mathematics Subject Classification* (1985 *Revision*). Primary 65N30; Secondary 65F10.
*This manuscript has been authored under contract number DE-AC02-76CH00016 with the U.S. Department of Energy. Accordingly, the U.S. Government retains a non-exclusive, royalty-free license to publish or reproduce the published form of this contribution, or allow others to do so, for U.S. Government purposes. This work was also supported in part under the National Science Foundation Grant No. DMS84-05352 and under the Air Force Office of Scientific Research, Contract No. ISSA86-0026.

positive constants. The generalized Dirichlet form is given by

$$A(v,\phi) = \sum_{i,j=1}^{N} \int_\Omega a_{ij} \frac{\partial v}{\partial x_i} \frac{\partial \phi}{\partial x_j}\, dx + \int_\Omega av\phi\, dx, \tag{1.3}$$

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

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

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

$$A(u,\phi) = (f,\phi)_\Omega \tag{1.4}$$

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

$$A(U,\Phi) = (f,\Phi)_\Omega \tag{1.5}$$

for all $\Phi \in S_h^0(\Omega)$.

We shall also be interested in solving (1.5) when the form A of (1.3) corresponds to the singularly perturbed operator

$$\tilde{L}v = v + \varepsilon Lv, \tag{1.6}$$

where L was defined by (1.2) and ε is a possibly small constant which in some applications depends upon h. The A form corresponding to (1.6) is then given by

$$A(v,\phi) = \varepsilon \left\{ \sum_{i,j=1}^{N} \int_\Omega a_{ij} \frac{\partial v}{\partial x_i} \frac{\partial \phi}{\partial x_j}\, dx + \int_\Omega av\phi\, dx \right\} + (v,\phi)_\Omega. \tag{1.7}$$

Singularly perturbed problems arise, for example, in time-stepping methods for the numerical approximation of parabolic problems.

Now it is easy to see that if ε is bounded away from zero, then any preconditioner for (1.5) gives a preconditioner for (1.7). Furthermore, if ε is of order h^2, then the quadratic form $A(v,v)$ restricted to the subspace $S_h^0(\Omega)$ is equivalent to $(v,v)_\Omega$ and no preconditioner is necessary. We shall provide a preconditioner for (1.7) which has conditioning properties similar to those of the preconditioner developed for (1.5) independent of ε.

As illustrated in Part 1 [3], the preconditioning problem can be reduced to the problem of defining an appropriate form B on $S_h^0(\Omega)\times S_h^0(\Omega)$ satisfying the following criterion. Firstly, the problem of finding $W \in S_h^0(\Omega)$, given g, satisfying

$$B(W,\Phi) = (g,\Phi)_\Omega \quad \text{for all } \Phi \in S_h^0(\Omega) \tag{1.8}$$

should be easier to obtain than the solution of (1.5). Secondly, the forms B and A should be comparable in the sense that there are positive constants λ_0 and λ_1 satisfying

$$\lambda_0 B(V,V) \le A(V,V) \le \lambda_1 B(V,V) \quad \text{for all } V \in S_h^0(\Omega) \tag{1.9}$$

with λ_1/λ_0 "not too large."

It should be noted that it is generally not possible to develop an effective preconditioner for (1.7) directly from a preconditioner for (1.5). If B is a preconditioner for (1.5), then a natural choice of a preconditioner for (1.7) would be the form given by

$$\varepsilon B(u,v) + (u,v)_\Omega. \tag{1.10}$$

Unfortunately, the problem corresponding to (1.8) using the form (1.10) cannot, in general, be efficiently solved.

In this paper we shall develop a particularly simple method for defining preconditioners by domain decomposition. As is typical with domain decomposition techniques, the given domain Ω is broken into a number of subdomains $\{\Omega_i\}$. Our preconditioner is defined so that the calculation of the solution of (1.8) involves solving in parallel related Galerkin equations on the subregions and some interconnecting equations. For the method to be developed, the number of unknowns involved in the interconnecting equations will be at most equal to the number of subdomains.

Other papers providing iterative methods involving domain decomposition for the solution of elliptic problems have appeared in the literature [1]–[8]. The earliest papers involved splitting the domain into subdomains without interior corner points [1], [2], [4]–[6], [8]. These methods became inefficient when many long thin subdomains were used. Consequently, it became natural to develop decomposition methods which use quasi-uniform subregions. In Part I, we defined and analyzed such a method for two-dimensional problems. That method was shown to have a condition number for the preconditioned system which was bounded by $c(1+\log(d/h))^2$ (here d and h correspond, respectively, to the diameter of the subregions and the discretization size of the mesh).

The preconditioner defined and analyzed in this paper has the following advantages over that defined in Part I. Firstly, it is somewhat simpler, both conceptually and computationally. Secondly, it extends in a straightforward manner to three-dimensional problems. Thirdly, it applies to singularly perturbed systems without deterioration in the iterative convergence rates.

On the negative side, the preconditioner defined in this paper shows a somewhat faster asymptotic growth of the condition number for the preconditioned system than that of the Part I preconditioner. We will show that the condition number for the new method is bounded by cd/h in contrast to the $(1+\log(d/h))^2$ growth for the preconditioned system of Part I. This is a reasonable growth for many rather large three-dimensional problems when d and h are judiciously chosen.

An important aspect of this paper involves the introduction of certain constants or 'average values' associated with discrete functions on the subdomains as part of the definition of the preconditioner B. A technique for computing these average values is presented. A future part in this series of papers will provide a three-dimensional preconditioner employing this averaging technique with a $(1+\log(d/h))^2$ condition number growth for the preconditioned system.

The outline of the remainder of the paper is as follows. In Section 2 we describe the domain decomposition preconditioners and prove estimates for the growth of the condition numbers for the preconditioned system. In Section 3 we show how to compute the solution to (1.8). Numerical examples of the preconditioner applied to problems in two and three dimensions are given in Section 4.

We shall also let c and C, with or without subscript, denote generic positive constants. These constants will always be independent of the mesh and subdomain parameters h and d (see Section 2).

2. The Construction and Analysis of the Preconditioner. We will describe the preconditioner in this section and prove an estimate for the condition number of the preconditioned system. We start by giving some hypothesis on the domain and subdomain partitioning and the associated finite element subspaces.

For the sake of simplicity of exposition we shall proceed with the discussion only for the special case of polyhedral domains and piecewise linear approximations. Many generalities are possible and will be discussed in later papers.

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

(A.1) Ω is a polyhedral domain in R^2 or R^3, which for each h, $0 < h < 1$, a parameter, has been given a triangulation Ω^h of maximal size h. That is, $\Omega^h = \bigcup_{j=1}^{m(h)} \tau_j^h$, where each τ_j^h is a simplex which is contained in a ball of radius h.

Any union of simplexes of Ω^h will be called a mesh subdomain, and the vertices of the simplexes in Ω^h will be denoted by x_i ordered in some fashion. We shall partition the domain Ω into a number of mesh subdomains $\{\Omega_k\}$.

(A.2) We assume that the triangulation is quasi-uniform near the boundaries of the subdomains, i.e., if $\tau^h \in \Omega^h$ is a simplex such that $\tau^h \cap \partial\Omega_k \neq \emptyset$, then τ^h contains a ball of radius ch where c is independent of h.

(A.3) The Ω_k are quasi-uniform of size d. This means there exists a positive constant c_1 which is independent of d and h such that each Ω_k contains a ball of radius $c_1 d$ and is contained in a ball of radius d. The number of domains n_d is proportional to d^{-N}.

(A.4) Each Ω_k is uniformly star-shaped with respect to a point. This means that for each Ω_k there is a point $\hat{x}_k$ and a constant $c_2 > 0$, independent of d and h, such that $(x - \hat{x}_k) \cdot n(x) \geq c_2 d$ for all $x \in \partial\Omega_k$. Here, $n(x)$ denotes the outward unit normal to $\partial\Omega_k$ at x.

(A.5) Let $\tilde{\Omega}_k$ be the scaled domain defined by

$$\tilde{\Omega}_k \equiv \{x | dx \in \Omega_k\}.$$

We assume that $\tilde{\Omega}_k$ has a Lipschitz continuous boundary with Lipschitz constants which are independent of d.

Remark 2.1. Assumption (A.5) is a weak regularity hypothesis for the boundary of Ω_k. It guarantees that a Poincaré inequality of the form

$$\|V\|_{\Omega_k}^2 \leq Cd^2 D_k(V, V) \tag{2.1}$$

holds for functions V with zero mean value on Ω_k with a constant C independent of d and k. Here $D_k(\cdot, \cdot)$ denotes the Dirichlet inner product on Ω_k.

Remark 2.2. We note that Assumption (A.4) implies the inequality

$$|u|_{\partial\Omega_k}^2 \leq c\{d^{-1} \|u\|_{\Omega_k}^2 + dD_k(u, u)\}. \tag{2.2}$$

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 $\partial\Omega$. $S_h^0(\Omega_k)$ will denote the subspace of $S_h^0(\Omega)$ of functions whose supports are contained in $\bar{\Omega}_k$ (in particular, they vanish on $\partial\Omega_k$

and outside $\bar{\Omega}_k$). Let Γ denote $\bigcup_k \partial\Omega_k$ and let $S_h(\Gamma)$ denote the functions which are restrictions to Γ of functions in $S_h^0(\Omega)$.

We shall need some additional notation. The $L^2(\partial\Omega_k)$ inner product shall be denoted

$$\langle v, w\rangle_{\partial\Omega_k} = \int_{\partial\Omega_k} vw \, ds$$

with corresponding norm

$$|v|^2_{\partial\Omega_k} = \langle v, v\rangle_{\partial\Omega_k},$$

where ds is an element of arc length or surface area of $\partial\Omega_k$. The analogous discrete inner product is given by

$$\langle v, w\rangle_{\partial\Omega_k,h} = h^{N-1} \sum_{x_i\in\partial\Omega_k} v(x_i)w(x_i)$$

with corresponding discrete norm

$$|v|^2_{\partial\Omega_k,h} = \langle v, v\rangle_{\partial\Omega_k,h}.$$

It follows from (A.2) that

$$c\,|v|_{\partial\Omega_k} \le |v|_{\partial\Omega_k,h} \le C\,|v|_{\partial\Omega_k} \tag{2.3}$$

holds for functions $v \in S_h(\Gamma)$. In addition, we have the following lemma.

LEMMA 2.1. *If $v \in S_h^0(\Omega)$ and vanishes at all interior nodes of Ω_k then*

$$c_1 h\,|v|^2_{\partial\Omega_k,h} \le \|v\|^2_{\Omega_k} \le C_1 h\,|v|^2_{\partial\Omega_k,h} \tag{2.4}$$

and

$$c_1 h^{-1}\,|v|^2_{\partial\Omega_k,h} \le D_k(v,v) \le C_1 h^{-1}\,|v|^2_{\partial\Omega_k,h}. \tag{2.5}$$

We next construct the bilinear form B corresponding to our preconditioner. We first introduce another form $\tilde{A}(\cdot,\cdot)$ on $S_h^0(\Omega)$. If A is given by (1.3), we define

$$\tilde{A}_k(v,\phi) = \sum_{i,j=1}^{N} \int_{\Omega_k} a_{ij}^k \frac{\partial v}{\partial x_i}\frac{\partial \phi}{\partial x_j}\, dx + \int_{\Omega_k} a^k v\phi \, dx.$$

Alternatively, if A is given by (1.7) then we define

$$\tilde{A}_k(v,\phi) = \varepsilon\left\{\sum_{i,j=1}^{N} \int_{\Omega_k} a_{ij}^k \frac{\partial v}{\partial x_i}\frac{\partial \phi}{\partial x_j}\, dx + \int_{\Omega_k} a^k v\phi \, dx\right\} + \int_{\Omega_k} b^k v\phi \, dx. \tag{2.6}$$

We then define

$$\tilde{A}(U,V) = \sum_k \tilde{A}_k(U,V). \tag{2.7}$$

Here $a^k = 0$ if $a = 0$ or $a^k \ge c > 0$. Furthermore $b^k \ge c > 0$. These functions are piecewise smooth (possibly discontinuous) for each k. Finally, $a_{ij}^k(x)$ for $i,j = 1,\dots,N$ is a piecewise smooth (possibly discontinuous) uniformly positive definite matrix. The reason for the form of $\tilde{A}$ was discussed in Part I ([3], Section 4). Basically, it allows for greater flexibility in the definition of the preconditioner and, for example, the use of constant coefficient fast solvers (even when L has variable coefficients).

We note that

$$C_0\tilde{A}(U,U) \le A(U,U) \le C_1\tilde{A}(U,U) \quad \text{for all } U \in S_h^0(\Omega) \tag{2.8}$$

holds. 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 $W_P \in S_h^0(\Omega_1) \oplus \cdots \oplus S_h^0(\Omega_{n_d})$ and W_P restricted to Ω_k 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

$$\tilde{A}_k(W_H,\Phi) = 0 \quad \text{for all } \Phi \in S_h^0(\Omega_k). \tag{2.9}$$

Thus on each Ω_k, W is decomposed into a function W_P which vanishes on $\partial\Omega_k$ and a function W_H which satisfies the above homogeneous equation and has the same boundary values as W on $\partial\Omega_k$. We shall refer to such a function W_H as "discrete $\tilde{A}_k$-harmonic". The subspace of discrete $\tilde{A}_k$-harmonic functions shall be denoted by $H(\Omega_k)$.

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

$$\tilde{A}(W,W) = \tilde{A}(W_P,W_P) + \tilde{A}(W_H,W_H). \tag{2.10}$$

We shall define the preconditioning form B by replacing the $\tilde{A}(W_H,W_H)$ term in (2.10).

Note that a discrete $\tilde{A}_k$-harmonic function is completely determined by its values on the boundary. Accordingly, the form $\tilde{A}(W_H,W_H)$ can be replaced by a form which only involves the boundary values. The particular choice of the boundary form will depend on whether we are considering (1.3) or the singularly perturbed case (1.7).

Remark 2.3. It seems reasonable to consider replacing the $\tilde{A}(W_H,W_H)$ by the identity (or a weighted identity) on the subdomain boundary. This works reasonably well if A is given by (1.3), d is not too large, and the coefficients of A are smooth. The replacement forms to be described work better in more general situations.

We first consider the case when A is given by (1.3). To understand the motivation for the form to be defined, it is instructive to consider the case when $a = 0$. We would like to replace the form $\tilde{A}_k$ (restricted to discrete harmonic functions) and define the replacement for $\tilde{A}(W_H,W_H)$ by summation. Note that if $a^k = 0$ then $\tilde{A}_k$ is indefinite, and so its replacement should also be indefinite. Let α_k be a constant which will be chosen later; then

$$\tilde{A}_k(W_H,W_H) = \tilde{A}_k(W_H - \alpha_k, W_H - \alpha_k) \le \tilde{A}_k(W,W), \tag{2.11}$$

where $W \in S_h(\Omega)|_{\Omega_k}$ is defined by the function which equals $W_H - \alpha_k$ on $\partial\Omega_k$ and vanishes on all interior nodes of Ω_k. Now

$$\tilde{A}_k(W,W) \le c\tilde{a}_k D_k(W,W), \tag{2.12}$$

where $\tilde{a}_k$ is (for example) the smallest eigenvalue of the matrix $\{a_{ij}^k(x)\}_{i,j=1}^N$ for some point $x \in \Omega_k$. It follows from (2.5), (2.11), and (2.12) that

$$\tilde{A}_k(W_H,W_H) \le c\tilde{a}_k h^{-1} |W_H - \alpha_k|^2_{\partial\Omega_k,h}. \tag{2.13}$$

To make the right-hand side of (2.13) correspond to an indefinite form, we choose α_k to be the discrete mean value of W_H on $\partial\Omega_k$, i.e.,

$$\alpha_k = \bar{W}_H \equiv \frac{\langle W_H, 1\rangle_{\partial\Omega_k,h}}{\langle 1,1\rangle_{\partial\Omega_k,h}}.$$

We replace $\tilde{A}(W_H, W_H)$ with

$$Q(W_H,W_H) = h^{-1}\sum_k \tilde{a}_k \left|W_H - \bar{W}_H\right|^2_{\partial\Omega_k,h}. \tag{2.14}$$

For more general A given by (1.3) with $a^k \neq 0$ we use

$$\begin{aligned} Q(W_H,W_H) &= \sum_k Q_k(W_H,W_H) \\ &\equiv \sum_k h^{-1}\left\{(\tilde{a}_k + \bar{a}_k h^2)\left|W_H - (\bar{W}_H)_k\right|^2_{\partial\Omega_k,h} + \bar{a}_k h d^N (\bar{W}_H)_k^2\right\}. \end{aligned} \tag{2.15}$$

Finally, when A is given by (1.7) we use

$$\begin{aligned} Q(W_H,W_H) &= \sum_k Q_k(W_H,W_H) \\ &\equiv \sum_k h^{-1}\Big\{(\varepsilon\tilde{a}_k + (\bar{b}_k + \varepsilon\bar{a}_k)h^2)\left|W_H - (\bar{W}_H)_k\right|^2_{\partial\Omega_k,h} \\ &\qquad + (\bar{b}_k + \varepsilon\bar{a}_k) h d^N (\bar{W}_H)_k^2\Big\}. \end{aligned} \tag{2.16}$$

The constants $\bar{b}_k$ and $\bar{a}_k$ are defined as the average values of b^k and a^k over Ω_k. As before, it suffices to take $\tilde{a}_k$ to be the minimal eigenvalue of the matrix $\{a^k_{ij}(x)\}$ for some point $x \in \Omega_k$.

We have the following theorem.

THEOREM 1. *Let A be given by* (1.3) *or* (1.7), *respectively. Let B be defined on $S^0_h(\Omega) \times S^0_h(\Omega)$ by*

$$B(W,W) \equiv \tilde{A}(W_P,W_P) + Q(W_H,W_H), \tag{2.17}$$

where Q is given by (2.15) *or* (2.16) *respectively. We then have*

$$\frac{ch}{d}B(W,W) \leq A(W,W) \leq CB(W,W) \quad \textit{for all } W \in S^0_h(\Omega), \tag{2.18}$$

with c and C independent of h and d.

Proof. By (2.7), (2.8) and (2.10) it suffices to prove that

$$\frac{ch}{d}Q_k(V,V) \leq \tilde{A}_k(V,V) \leq CQ_k(V,V) \quad \text{for all } V \in H(\Omega_k). \tag{2.19}$$

We shall first prove the theorem when A is given by (1.3) and $a^k = 0$. In this case, (2.19) reduces to

$$\frac{c\tilde{a}_k}{d}\left|V - \bar{V}_k\right|^2_{\partial\Omega_k,h} \leq \tilde{A}_k(V,V) \leq \frac{C\tilde{a}_k}{h}\left|V - \bar{V}_k\right|^2_{\partial\Omega_k,h} \quad \text{for all } V \in H(\Omega_k). \tag{2.20}$$

The second inequality is just (2.13).

To prove the first inequality, let β_k denote the mean value of V on Ω_k. By (2.3) and the definition of $\bar{V}_k$, we have

$$|V - \bar{V}_k|^2_{\partial\Omega_k,h} \le |V - \beta_k|^2_{\partial\Omega_k,h} \le C\,|V - \beta_k|^2_{\partial\Omega_k}.$$

Applying (2.2), (2.1) and (2.3) gives

$$\begin{aligned} |V - \bar{V}_k|^2_{\partial\Omega_k,h} &\le c(d^{-1}\,\|V - \beta_k\|^2_{\Omega_k} + dD_k(V - \beta_k, V - \beta_k)) \\ &\le CdD_k(V - \beta_k, V - \beta_k) = CdD_k(V,V). \end{aligned}$$

Then, by the ellipticity assumptions on the coefficients defining $\tilde{A}_k$ and A, we have

$$\tilde{a}_k\,|V - \bar{V}_k|^2_{\partial\Omega_k,h} \le Cd\tilde{A}_k(V,V). \tag{2.21}$$

Thus we have shown that the theorem holds for the case $a^k = 0$.

We next prove the theorem for the remaining cases. When $a \neq 0$ and A given by (1.3), we set $\bar{b}_k = 0$ and $\varepsilon = 1$. Hence (2.16) defines Q in either case. We first prove the second inequality of (2.19). Let $\tilde{V}_k$ denote the discrete $\tilde{A}_k$-harmonic extension of $\bar{V}_k$. Note that in general, $\tilde{V}_k$ is nonconstant even though $\bar{V}_k$ is constant on $\partial\Omega_k$. Evidently,

$$\tilde{A}_k(V,V) \le 2(\tilde{A}_k(V - \tilde{V}_k, V - \tilde{V}_k) + \tilde{A}_k(\tilde{V}_k, \tilde{V}_k)). \tag{2.22}$$

By the harmonicity of $\tilde{V}_k$,

$$\tilde{A}_k(\tilde{V}_k, \tilde{V}_k) \le \tilde{A}_k(\bar{V}_k, \bar{V}_k) \le c(\bar{b}_k + \varepsilon\bar{a}_k)\,\|\bar{V}_k\|^2_{\Omega_k} \le C(\bar{b}_k + \varepsilon\bar{a}_k)d^N\bar{V}_k^2. \tag{2.23}$$

By the harmonicity of $V - \tilde{V}_k$,

$$\tilde{A}_k(V - \tilde{V}_k, V - \tilde{V}_k) \le \tilde{A}_k(W,W), \tag{2.24}$$

where W is the function which equals $V - \tilde{V}_k$ on $\partial\Omega_k$ and vanishes on the interior nodes of Ω_k. The assumptions on the coefficients of the operator and the preconditioner and Lemma 2.1 give

$$\begin{aligned} \tilde{A}_k(W,W) &\le c\{(\bar{b}_k + \varepsilon\bar{a}_k)\,\|W\|^2_{\Omega_k} + \varepsilon\tilde{a}_k D_k(W,W)\} \\ &\le C\left\{\left(\frac{\varepsilon\tilde{a}_k}{h} + (\bar{b}_k + \varepsilon\bar{a}_k)h\right)|V - \bar{V}_k|^2_{\partial\Omega_k,h}\right\}. \end{aligned} \tag{2.25}$$

Combining (2.22) through (2.25) proves the second inequality of (2.19).

We finally prove the first inequality of (2.19). Noting that (2.21) is also valid in the present case, it suffices to show that

$$h(\bar{b}_k + \varepsilon\bar{a}_k)\left\{h\,|V - \bar{V}_k|^2_{\partial\Omega_k,h} + d^N\bar{V}_k^2\right\} \le Cd\tilde{A}_k(V,V) \tag{2.26}$$

for all $V \in H(\Omega_k)$. By the arithmetic-geometric mean inequality,

$$h\,|V - \bar{V}_k|^2_{\partial\Omega_k,h} \le 2h(|V|^2_{\partial\Omega_k,h} + |\bar{V}_k|^2_{\partial\Omega_k,h}). \tag{2.27}$$

Using (2.4), it follows trivially that

$$h\,|V|^2_{\partial\Omega_k,h} \le C\,\|V\|^2_{\Omega_k} \le \frac{C}{(\bar{b}_k + \varepsilon\bar{a}_k)}\tilde{A}_k(V,V). \tag{2.28}$$

Since d is larger than h, a straightforward computation using (A.3) gives

$$h\,|\bar{V}_k|^2_{\partial\Omega_k,h} \le d\,|\bar{V}_k|^2_{\partial\Omega_k,h} \le Cd^N\bar{V}_k^2.$$

Hence it remains to bound $d^N \bar{V}_k^2$. By the definition of $\bar{V}_k$ and the Schwarz inequality,

$$d^N \bar{V}_k^2 \leq c d^N (h/d)^{2N-2} \left(\sum_i V(x_i) \right)^2$$
$$\leq C d h^{N-1} \sum_i V(x_i)^2 = C d |V|^2_{\partial\Omega_k, h},$$

where the sum over i is taken over the set of nodes x_i on $\partial\Omega_k$. Hence by (2.4),

$$d^N \bar{V}_k^2 \leq C \frac{d}{h} \|V\|^2_{\Omega_k}. \tag{2.29}$$

Combining (2.28) and (2.29) proves (2.26) and hence completes the proof of the theorem. □

Remark 2.4. The coefficients c and C appearing in the theorem depend on the local (with respect to the subdomains) behavior of the operator and preconditioner. Accordingly, the preconditioner will work well even in situations where there are large jumps in the coefficients defining L as long as these jumps only occur across the subdomain boundaries.

Remark 2.5. There is a fair amount of freedom in weighting the boundary form. For example, the Q form in the case $a = 0$ could have been defined by

$$Q(W_H, W_H) = \gamma^{-1} \sum_k \tilde{a}_k |W_H - \alpha_k|^2_{\partial\Omega_k, h}.$$

Then the condition number for the preconditioned system would remain unchanged as long as $h \leq \gamma \leq d$. The forms (2.15) and (2.16) could be similarly weighted.

3. The Solution of the Preconditioning Problem. In this section we describe an efficient algorithm for solving (1.8). In general, when B is of the form (2.17), we solve first for W_P, then for the values of W_H on Γ, and finally extend W_H to all of Ω.

We now give the details of a three-step algorithm for the solution of (1.8). As already mentioned, the problem of finding the solution W to (1.8) reduces to that of computing W_P and W_H. The first step is to compute W_P. By taking $\Phi \in S_h^0(\Omega_k)$ in (1.8) and using (2.9), we note that

$$\tilde{A}_k(W_P, \Phi) = (g, \Phi) \quad \text{for all } \Phi \in S_h^0(\Omega_k). \tag{3.1}$$

Equation (3.1) shows that W_P can be determined by solving independent discrete Dirichlet problems on the subregions. The second step involves the computation of the values of W_H on Γ. These values are determined as the solution of the following problem:

$$Q(W_H, \theta) = (g, \tilde{\theta})_\Omega - \tilde{A}(W_P, \tilde{\theta}) \quad \text{for all } \theta \in S_h(\Gamma). \tag{3.2}$$

Here $\tilde{\theta}$ denotes any extension of θ in $S_h^0(\Omega)$ and we note that by (3.1), the right-hand side of (3.2) is independent of the extension chosen. The development of an algorithm for solving (3.2) is an important part of this section and will be considered shortly. The third step is to compute the discrete $\tilde{A}_k$-harmonic extension of the boundary values of W_H computed in the previous step. This is done as follows: Let $\tilde{W}_H$ be any extension of the boundary values of W_H in $S_h^0(\Omega)$, e.g., the extension

which is zero at all of the nodes not on Γ. Then $W_H = Y + \tilde{W}_H$, where Y vanishes on Γ and is the solution of

$$\tilde{A}_k(Y, \Phi) = -\tilde{A}_k(\tilde{W}_H, \Phi) \quad \text{for all } \Phi \in S_h^0(\Omega_k), \tag{3.3}$$

for $k = 1, \ldots, n_d$. Equation (3.3), as in (3.1), requires independent discrete Dirichlet solves on the subdomains.

We now develop an algorithm for solving (3.2), that is, computing the values of W_H on Γ. For notational convenience we set $V = W_H|_\Gamma$. Then (3.2) reduces to

$$h^{-1} \sum_k \left\{ (\varepsilon \tilde{a}_k + h^2(\bar{b}_k + \varepsilon \bar{a}_k)) \langle V - \bar{V}_k, \chi \rangle_{\partial\Omega_k, h} + (\bar{b}_k + \varepsilon \bar{a}_k) h d^2 \bar{V}_k \bar{\chi}_k \right\} = F(\chi) \tag{3.4}$$

for all $\chi \in S_h(\Gamma)$, where F is a known linear functional on $S_h(\Gamma)$. We shall see that the problem of computing the solution V of (3.4) is straightforward if the values of $\bar{V}_k$ are known. Indeed, V satisfies

$$h^{-1} \sum_k \left(\varepsilon \tilde{a}_k + h^2(\bar{b}_k + \varepsilon \bar{a}_k) \right) \langle V, \chi \rangle_{\partial\Omega_k, h} = G(\chi), \tag{3.5}$$

where the linear functional G in (3.5) depends upon F and the average values $\bar{V}_k$. Note that if we use the usual nodal basis for functions in $S_h(\Gamma)$, then the matrix corresponding to Problem (3.5) is diagonal and hence its solution can be trivially computed. Thus to solve (3.2), we need only demonstrate a technique for computing the average values $\bar{V}_k$.

We shall define another function $\tilde{V} \in S_h(\Gamma)$ which has the same average values as the solution V of (3.4). Obviously, the average values of V can then be computed by calculating the average values of $\tilde{V}$. Let $S_h^0(\Gamma)$ denote the collection of functions in $S_h(\Gamma)$ which have zero average value on every $\partial\Omega_k$. Clearly, $(V - \tilde{V}) \in S_h^0(\Gamma)$. To make the system of equations determining $\tilde{V}$ of minimal size, we choose $\tilde{V}$ in an orthogonal complement of $S_h^0(\Gamma)$. Specifically, let $S_h^\perp(\Gamma)$ be defined by

$$S_h^\perp(\Gamma) \equiv \{\theta \in S_h(\Gamma) \mid Q(\theta, \omega) = 0 \text{ for all } \omega \in S_h^0(\Gamma)\},$$

and define $\tilde{V}$ to be the unique function in $S_h^\perp(\Gamma)$ satisfying

$$Q(\tilde{V}, \theta) = F(\theta) \quad \text{for all } \theta \in S_h^\perp(\Gamma). \tag{3.6}$$

Note that $\tilde{V}$ is the orthogonal projection of V into $S_h^\perp(\Gamma)$ and hence $\tilde{V}$ has the same average values as V. In what follows, we shall derive a basis for $S_h^\perp(\Gamma)$. This basis will consist of functions with local (with respect to d) support, and hence $\tilde{V}$ can be computed as the solution to a sparse, positive definite and symmetric "stiffness" matrix corresponding to (3.6). The number of unknowns in this system will always be less than or equal to n_d.

Before proceeding, we shall introduce some additional notation. Let $\nu_k \equiv h^{-1}(\varepsilon \tilde{a}_k + h^2(\bar{b}_k + \varepsilon \bar{a}_k))$ and define

$$Q_0(W, W) \equiv \sum_k \nu_k \left| W - \bar{W} \right|^2_{\partial\Omega_k, h}.$$

Note that Q and Q_0 only differ by terms involving the average values squared. Hence,

$$S_h^\perp(\Gamma) = \{\theta \in S_h(\Gamma) \mid Q_0(\theta, \omega) = 0 \text{ for all } \omega \in S_h^0(\Gamma)\}. \tag{3.7}$$

We will first define functions $\phi_k \in S_h^{\perp}(\Gamma)$, for $k = 1, \ldots, n_d$. Consider a fixed subregion with boundary $\partial\Omega_k$. The function ϕ_k is defined to be zero on all of the nodes on $\Gamma/\partial\Omega_k$, its values on $\partial\Omega_k$ are to be determined. Let W be in $S_h(\Gamma)$; then

$$Q_0(W, \phi_k) = \sum_i \gamma_i W(x_i)\phi_k(x_i) + \sum_j \kappa_{k,j}\bar{W}_j, \tag{3.8}$$

where

$$\gamma_i = h^{N-1} \sum_{\{j | x_i \in \partial\Omega_j \cap \partial\Omega_k\}} \nu_j \tag{3.9}$$

and

$$\kappa_{k,j} = -h^{N-1}\nu_j \sum_{x_l \in \partial\Omega_k \cap \partial\Omega_j} \phi_k(x_l). \tag{3.10}$$

The sum over i in (3.8) is taken over the nodes x_i on $\partial\Omega_k$ and the sum over j in (3.8) is taken over the subregions Ω_j with $\partial\Omega_j \cap \partial\Omega_k \neq \varnothing$. We define the nodal values of ϕ_k on $\partial\Omega_k$ by

$$\phi_k(x_i) = \frac{\nu_k}{\gamma_i}. \tag{3.11}$$

With the above choice for ϕ_k, it is evident that the first sum in (3.8) equals $\nu_k N_k \bar{W}_k$, where N_k is defined to be the number of nodes on $\partial\Omega_k$. Hence (3.8) becomes

$$Q_0(W, \phi_k) = \nu_k N_k \bar{W}_k + \sum_j \kappa_{k,j}\bar{W}_j. \tag{3.12}$$

By (3.7) and (3.12), $\phi_k \in S_h^{\perp}(\Gamma)$.

We will next show that

$$S_h^{\perp}(\Gamma) = \operatorname*{span}_{k=1,\ldots,n_d} \phi_k. \tag{3.13}$$

It suffices to show that if

$$\theta \in S_h^{\perp}(\Gamma) \quad \text{and} \quad Q_0(\theta, \phi_k) = 0 \tag{3.14}$$

for $k = 1, \ldots, n_d$, then $\theta = 0$. Consider the matrix M defined by the right-hand side of (3.12), i.e.,

$$M_{i,j} = \delta_{i,j}\nu_i N_i + \kappa_{i,j},$$

where $\delta_{i,j}$ is the Kronecker Delta Function. Let θ satisfy (3.14) and $\bar{\theta}$ be the vector with components $\bar{\theta}_k$. Then by (3.12), (3.14) and the definition of M,

$$M\bar{\theta} = 0. \tag{3.15}$$

To show that (3.13) holds, it suffices to show that M is invertible. Indeed, if M is invertible, then (3.15) implies that θ is also in $S_h^0(\Gamma)$, i.e., $\theta = 0$.

We will see that M is symmetric and positive definite and hence invertible. Indeed, the quantity γ_i in (3.9) depends upon the point x_i but not the subregion Ω_k. Consequently, $\kappa_{j,k} = \kappa_{k,j}$, i.e., M is symmetric. Furthermore, this system is sparse with positive diagonal entries and nonpositive off-diagonal entries. Also,

$$\sum_j \kappa_{k,j} = -\tilde{N}_k \nu_k,$$

where $\tilde{N}_k$ is defined to be the number of nodes on $\partial\Omega_k$ which do not lie on $\partial\Omega$. Consequently, the matrix M is irreducibly diagonally dominant and hence positive definite; cf. [9].

In general, the functions $\phi_1, \ldots, \phi_{n_d}$ may not be linearly independent. For example, in the case of the unit square with the checkerboard subdivision, the function

$$\sum_{\text{red squares}} \phi_k - \sum_{\text{black squares}} \phi_l = 0.$$

In this case it is easy to check that $\{\phi_1, \ldots, \phi_{n_d-1}\}$ is linearly independent and hence forms a basis for $S_h^{\perp}(\Gamma)$. Bases for more complicated domains and subdivisions are also straightforward to derive.

For completeness, we restate the algorithm developed in this section for computing the solution W of (1.8).

Algorithm DD2.

(1) Compute W_P by solving (3.1). This involves Dirichlet solves on the subdomains which can be done independently and in parallel.

(2) Compute the values of W_H on Γ. First we compute the function $\tilde{V}$ by solving (3.6) using the finite element basis $\{\phi_k\}$ described above. The average values of V are computed by calculating the average values of $\tilde{V}$. The values of W_H on Γ are then computed by solving the trivial equation (3.5).

(3) Extend the boundary values of W_H by solving (3.3). As in Step 1, this involves Dirichlet solves on the subdomains which can be done independently and/or in parallel.

(4) Set $W = W_P + W_H$.

Remark 3.1. When $\bar{b}_k = \bar{a}_k = 0$, the matrix M can be directly used to compute the average values of V. Indeed, if $\bar{V}$ denotes the vector of average values of V, then by (3.12) and the definition of M,

$$Q(V, \phi_k) = (M\bar{V})_k = F(\phi_k), \tag{3.16}$$

and hence the average values of V are given by

$$\bar{V} = M^{-1} \begin{pmatrix} F(\phi_1) \\ \vdots \\ F(\phi_{n_d}) \end{pmatrix}.$$

Remark 3.2. The method for computing the average values of V described in Remark 3.1 may not work well when either $\bar{b}_k$ or $\bar{a}_k$ are nonzero. In the general case, a matrix M satisfying (3.16) can be derived using similar techniques. In such cases, M may no longer be symmetric or diagonally dominant. Consequently, it may be difficult to obtain good numerical solutions for (3.16) when low-order terms are present. Note, however, that the algorithm described earlier for computing the average values by solving (3.6) always leads to numerically stable, sparse and symmetric positive definite systems.

4. Numerical Experiments. In this section we shall present some results of numerical experiments which illustrate the convergence properties of the preconditioned conjugate gradient algorithm using DD2 as a preconditioner. Two-dimensional examples where L is given by (1.2) will be considered first. These examples are taken from [3], so that a direct comparison between the preconditioners DD1 and DD2 can be made. Next, two-dimensional singularly perturbed

problems of the form (1.5), (1.7) will be studied. Finally, a model three-dimensional problem will be considered.

To avoid making this section too long, we shall not attempt to illustrate the full power and flexibility of the algorithm. Accordingly, other numerical examples on which we have tested the algorithm will not be included. These examples include applications to problems with discontinuous coefficients, problems with smoothly varying coefficients, and problems on domains with irregular geometry (see Examples 2, 3, 4, and 5 of Part I).

We shall define a number of parameters which will be introduced to study the convergence properties of the proposed preconditioning algorithm. The condition number of the preconditioned system is denoted by K. The integer n is defined to be the number of iterations required to reduce the A-norm (defined by $A(\cdot,\cdot)^{1/2}$) of the error $E_n = U - U_n$ by a factor of .0001. Here, U is a randomly generated solution of (1.5), normalized so that $-1 \leq U \leq 1$, and U_n is the approximation to U obtained using n steps of a conjugate gradient algorithm preconditioned by DD2. It is well known that the A-norm of the iteration error satisfies the bound

$$A(E_j, E_j) \leq 4\rho^{2j} A(E_0, E_0),$$

where

$$\rho \equiv \frac{\sqrt{K} - 1}{\sqrt{K} + 1}. \tag{4.1}$$

We shall sometimes compare ρ with the average observed reduction ρ_0 defined by

$$\rho_0 = \left(\frac{A(E_n, E_n)}{A(E_0, E_0)} \right)^{1/2n}.$$

Example 1. The first set of numerical experiments is applied to the standard model problem given by

$$Lu = f \quad \text{on } \Omega \qquad \text{and} \qquad u = 0 \quad \text{on } \partial\Omega, \tag{4.2}$$

where L is taken to be the Laplace operator $-\Delta$ and Ω is the unit square. There are many techniques available for solving this problem. However, this problem is interesting in that it illustrates many of the convergence properties of the proposed preconditioner. The square is partitioned into m^2 equal subsquares and hence $d = 1/m$.

The first table gives some indication of a typical run on an iteration-by-iteration basis. Here we break the square into sixteen subsquares and hence $d = 1/4$ and set $h = 1/32$. Table 4.1 gives the normalized error reduction as a function of the number of preconditioned steps in the A-norm and the maximum norm. Note that it takes 14 iterations to reduce the A-norm error by .0001. For this example, the average error reduction over 14 steps in the A-norm (resp. maximum norm) was .52 (resp. .60).

TABLE 4.1
Step by Step Iterative Convergence for Example 1.

Iteration	A-error	Max-error
1	8.0×10^{-1}	1.0
2	4.0×10^{-1}	1.4
3	1.4×10^{-1}	6.7×10^{-1}
4	5.7×10^{-2}	3.5×10^{-1}
5	2.3×10^{-2}	1.7×10^{-1}
6	1.1×10^{-2}	9.5×10^{-2}
7	5.4×10^{-3}	4.5×10^{-2}
8	2.3×10^{-3}	3.5×10^{-2}
9	1.8×10^{-3}	2.6×10^{-2}
10	1.2×10^{-3}	1.5×10^{-2}
11	6.8×10^{-4}	6.9×10^{-3}
12	3.5×10^{-4}	2.9×10^{-3}
13	1.9×10^{-4}	1.6×10^{-3}
14	9.3×10^{-5}	7.7×10^{-4}

The next two tables show that, in practice, the condition number of the preconditioned systems exhibit the growth rates predicted by the theory. In Table 4.2, we fix $d = 1/4$ and vary h. As predicted by Theorem 1, K grows like d/h. For Table 4.3, we fix $d/h = 4$ and vary h. In this case, the condition number for the precondition system remains bounded independent of h. Tables 4.2 and 4.3 also give the observed average reduction ρ_0 and n, the number of iterations required to reduce the A-norm error by a factor of .0001. The number of subregions m is also included in Table 4.3.

TABLE 4.2
Iterative Convergence Results for Example 1 *when* $d = 1/4$.

h	K	$\frac{15d}{8h}$	ρ_0	n
1/8	3.4	3.8	.21	7
1/16	7.2	7.5	.39	10
1/32	14	15	.52	14
1/64	30	30	.60	19
1/128	61	60	.68	24

TABLE 4.3
Iterative Convergence Results for Example 1 *when* $d/h = 4$.

h	K	ρ_0	n	m
1/8	6.6	.20	6	4
1/16	7.2	.40	10	16
1/32	7.5	.42	11	64
1/64	7.6	.42	11	256

Example 2. The next example illustrates the algorithm applied to singularly perturbed problems. We again consider the unit square with the same subdomain

subdivisions as in Example 1. For this problem, the form A is given by

$$A(v,\phi) = \varepsilon D(v,\phi) + (v,\phi). \tag{4.3}$$

Table 4.4 gives iterative convergence results when $h = 1/32$, $d = 1/4$, $\varepsilon = h^p$, and p varies between 0 and 2. This range of ε is typically that which occurs when time-stepping procedures are applied to parabolic problems and ε is essentially the size of the time step. Table 4.4 shows that the condition numbers for the preconditioned systems, as p varies, remain bounded by the condition number corresponding to the case $p = 0$. Similar results were obtained when h and d were varied.

TABLE 4.4
Iterative Convergence Results for Example 2 as p Varies.

p	K	ρ_0	n
0	15.1	.51	14
0.5	14.7	.51	14
1.0	12.4	.49	14
1.5	9.7	.45	12
2	6.6	.33	9

Example 3. For our final example we consider a model three-dimensional problem. Here we set Ω to be the unit cube and define the subregions by breaking Ω into 27 subcubes of equal size. We let L be an elliptic operator of the form

$$Lu = -\nabla \cdot \mu \nabla u \quad \text{in } \Omega, \tag{4.4}$$

where μ is a piecewise constant function on Ω and constant on the subdomains. Figure 4.1 gives the values of μ as a function of the x, y, z coordinates of the center of the subregions. These values were chosen to exhibit relatively large jumps across subregions, but are otherwise arbitrary. Table 4.5 gives iterative convergence results for the conjugate gradient method preconditioned by DD2 applied to the finite element equations corresponding to (4.4). Note that even though the coefficients of the operator have large jumps, the condition number K of the preconditioned system remains relatively small. In fact, the results reported do not differ significantly from results (not presented) for the case $\mu \equiv 1$. This is in agreement with Remark 2.4.

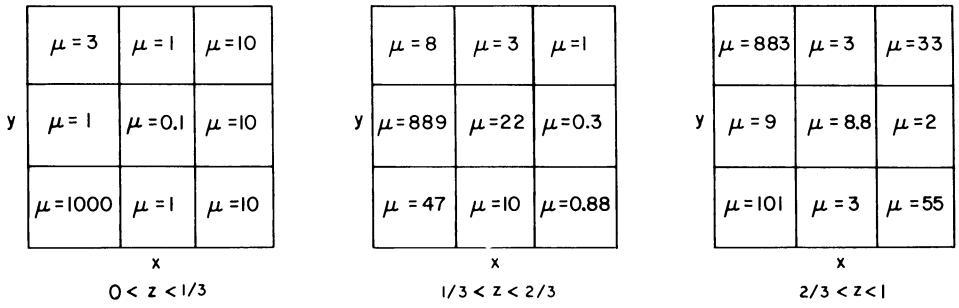

FIGURE 4.1
Coefficients for Example 3.

TABLE 4.5
Iterative Convergence Results for Example 3.

h	K	ρ_0	n
1/6	6.8	.39	11
1/12	17.4	.55	16
1/24	38	.64	21

Department of Mathematics
Cornell University
Ithaca, New York 14853

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," *SIAM J. Numer. Anal.*, v. 23, 1986, pp. 1097–1120.

3. J. H. BRAMBLE, J. E. PASCIAK & A. H. SCHATZ, "The construction of preconditioners for elliptic problems by substructuring. I," *Math. Comp.*, v. 47, 1986, pp. 103–134.

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. G. H. GOLUB & D. MEYERS, "The use of preconditioning over irregular regions," *Proc. 6th Internat. Conf. Comput. Methods in Sci. and Engng.*, Versailles, France, 1983.

9. R. S. VARGA, *Matrix Iterative Analysis*, Prentice-Hall, Englewood Cliffs, N.J., 1962.