File : nonsym

January 18, 1991

Time : 15 : 2

Typeset by $\mathcal{A}\mathcal{M}\mathcal{S}$-TEX

Iterative Schemes for Non-symmetric and Indefinite Elliptic Boundary Value Problems

James H. Bramble, Zbigniew Leyk, and Joseph E. Pasciak

January 18, 1991

Abstract. The purpose of this paper is twofold. The first is to describe some simple and robust iterative schemes for non-symmetric and indefinite elliptic boundary value problems. The schemes are based in the Sobolev space $H^1(\Omega)$ and require minimal hypotheses. The second is to develop algorithms utilizing a coarse grid approximation. This leads to iteration matrices whose eigenvalues lie in the right half of the complex plane. In fact, for symmetric indefinite problems, the iteration is reduced to a well conditioned symmetric positive definite system which can be solved by conjugate gradient iteration. Application of the general theory as well as numerical examples are given.

1. Introduction.

In the first part of this paper, we shall describe methods based on the normal equations with respect to the Sobolev space $H^1(\Omega)$. Methods of this kind have been suggested in [**8**] and in [**3**]. In [**3**], a theorem providing bounds for iterative convergence rate was given. Here, we give a somewhat more general version of the above mentioned result and elaborate on its applicability and implementation. These methods are particularly robust in that preconditioners can often be developed from problems with different boundary conditions and only limited regularity is required on the solutions of the underlying partial differential equation.

In contrast, iterative schemes for non-symmetric and indefinite systems have been studied which are based on the normal equations in discrete L^2. The analysis of the resulting iterative schemes seems to require full elliptic regularity. In addition, rapid convergence of the L^2 based algorithms requires more stringent restrictions on the boundary conditions of the problem from which the preconditioner is derived [**12**].

The $H^1(\Omega)$ based schemes which we shall describe are simple to analyze and robust. Alternative schemes based on generalizations of conjugate gradient and conjugate residual methods have been proposed. Theoretically, none of the generalized schemes can be shown to be asymptotically faster than the $H^1(\Omega)$ normal equation method which we shall describe. Extensive comparisons of these methods have been made [**8**]. These experiments suggest that the methods exhibit similar performance when they converge. However, the generalized approaches may fail to converge in certain applications.

In the second part of this paper, we introduce a technique for reducing certain indefinite or non-symmetric problems to ones whose spectrum is contained in the right half of the complex plane. We then show how this may be utilized to define an easily computable

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 by the U.S. Army Research Office through the Mathematical Sciences Institute, Cornell University.

Typeset by $\mathcal{A}\mathcal{M}\mathcal{S}$-TEX

iterative procedure for solving the resulting reduced problem. In the symmetric indefinite case, the reduced problem is symmetric positive definite and hence, for example, the conjugate gradient method may be applied. In the nonsymmetric case, the reduced problem may be solved using, for example, the GMRES algorithm (cf. [**16**]).

We shall develop the iterative schemes in an abstract way. To do this, we assume that we are given a Hilbert space $\mathcal{S}$ with norm and inner product respectively denoted by $\|\cdot\|_{H^1}$ and $\hat{A}(\cdot,\cdot)$. The space $\mathcal{S}$ is compactly and densely contained in a larger space L^2 with norm and inner product denoted $\|\cdot\|$ and $(\cdot,\cdot)$. We shall be interested in approximating the solution $u \in \mathcal{S}$ of the problem

$$A(u,\theta) = (f,\theta) \qquad \text{for all } \theta \in \mathcal{S}, \tag{1.1}$$

for a given function f. The quadratic form A may be non-symmetric or indefinite but is assumed to be bounded in the norm on $\mathcal{S}$.

We shall describe the iterative schemes applied to a family of approximations to (1.1). To this end, we assume that we are given finite dimensional approximation spaces $\mathcal{S}_h \subset \mathcal{S}$ indexed by $h \in (0,1]$. The parameter h corresponds roughly to an approximation grid size and problems with larger h have fewer unknowns. The approximation $u_h \in \mathcal{S}_h$ is defined to be the Galerkin projection, i.e.

$$A(u_h,\phi) = (f,\phi) \qquad \text{for all } \phi \in \mathcal{S}_h. \tag{1.2}$$

We shall always assume that (1.1) and (1.2) are uniquely solvable. In applications, this is often the case with only minor restriction on the size of h, e.g. $h \le h_0$ for a fixed constant h_0. We shall develop various preconditioned iterative schemes for computing the solution of (1.2) in the remainder of the paper.

In Section 2, we describe the preconditioned iterative schemes based in H^1 for (1.2). These schemes have appeared in earlier literature [**3**],[**8**]. We state a couple of simple theorems which provide bounds for their convergence.

In Section 3, we describe iterative schemes in a reduced subspace utilizing a coarse grid approximation. Alternative ways of using a coarse grid approximation in these types of problems have appeared, for example, in [**5**], [**6**] as well as in the multigrid literature (see [**4**],[**11**] and the extended bibliography in [**13**]. Our approach is unique in that it leads to a problem with lesser rank and a well conditioned symmetric positive definite system in some applications.

We present applications in 4. There we consider second order equations with first order terms, oblique boundary derivative terms, as well as zero'th order terms of the Helmholz type. Finally, the results of numerical experiments illustrating the convergence of the proposed iterative schemes are given in Section 5.

2. A general iterative method with preconditioner based in H^1.

In this section, we provide iterative schemes for (1.2). These schemes will be defined in terms of a symmetric positive definite quadratic form B_h on $\mathcal{S}_h \times \mathcal{S}_h$. We suppose that there are constants C_0 and C_1 satisfying

$$C_0 \hat{A}(\phi,\phi) \le B_h(\phi,\phi) \le C_1 \hat{A}(\phi,\phi) \qquad \text{for all } \phi \in \mathcal{S}_h. \tag{2.1}$$

The form B_h will provide the preconditioner for A. For the subsequently described iterative methods to be effective, the ratio C_1/C_0 should remain relatively small even as h becomes small.

To describe the iterative methods, we first consider computer representation of the finite element problem (1.2). The computer realization of (1.2) assumes a given basis $\{\phi_h^i\}_{i=1}^{n_h}$ for $\mathcal{S}_h$. The function u_h solving (1.2) is expanded in this basis as

$$u_h = \sum_{i=1}^{n_h} U_h^i \phi_h^i.$$

The unknown coefficients satisfy the matrix equation

$$\tilde{A}_h U_h = F_h \tag{2.2}$$

where $\tilde{A}_h$ is the "stiffness matrix" with entries $(\tilde{A}_h)_{ij} = A(\phi_h^i, \phi_h^j)$ and the F_h is the vector with entries $(F_h)_i = (f, \phi_h^i)$. The preconditioning matrix $(\tilde{B}_h)_{ij} = B_h(\phi_h^i, \phi_h^j)$ is analogously defined. Finally, the preconditioner $\tilde{M}_h$ is defined to be the inverse of $\tilde{B}_h$.

The iterative schemes are based on the following simple observation. The solution U_h of (2.2) satisfies

$$\mathcal{A} U_h \equiv \tilde{M}_h \tilde{A}_h^t \tilde{M}_h \tilde{A}_h U_h = \tilde{M}_h \tilde{A}_h^t \tilde{M}_h F_h. \tag{2.3}$$

The matrix $\mathcal{A} = \tilde{M}_h \tilde{A}_h^t \tilde{M}_h \tilde{A}_h$ is symmetric on R^{n_h} equipped with the inner product

$$[V, W]_h = V^t \tilde{B}_h W. \tag{2.4}$$

It is immediate that

$$[\mathcal{A}V, W]_h = (\tilde{A}_h V)^t \tilde{M}_h (\tilde{A}_h W) = [V, \mathcal{A}W]_h. \tag{2.5}$$

Solvability of (1.2) and (2.1) imply that $\mathcal{A}$ is positive definite with respect to this inner product. The preconditioned H^1 based iterative scheme for solving (2.2) is nothing more than the conjugate gradient method applied to (2.3) in the inner product (2.5). This leads to the following algorithm.

ALGORITHM 2.1.

(1) *Let an initial approximation V_0 to U_h be given (e.g., $V_0 = 0$).*
(2) *Set $R_0 = \tilde{A}_h^t \tilde{M}_h (F_h - \tilde{A}_h V_0)$ and $P_0 = \tilde{M}_h R_0$.*
(3) *For $i \geq 0$ define:*

$$\begin{aligned}
\alpha_i &= \frac{R_i^t P_i}{(\tilde{A}_h P_i)^t \tilde{M}_h (\tilde{A}_h P_i)}, \\
V_{i+1} &= V_i + \alpha_i P_i, \\
R_{i+1} &= R_i - \alpha_i \tilde{A}_h^t \tilde{M}_h \tilde{A}_h P_i, \\
\beta_i &= \frac{(\tilde{M}_h R_{i+1})^t \tilde{A}_h^t \tilde{M}_h \tilde{A}_h P_i}{(\tilde{A}_h P_i)^t \tilde{M}_h (\tilde{A}_h P_i)} \\
P_{i+1} &= \tilde{M}_h R_{i+1} - \beta_i P_i.
\end{aligned}$$

REMARK 2.1: Note that the inner product (2.4) does not appear in the above algorithm. Thus, it is possible to use a preconditioner $\tilde{M}_h$ without explicitly knowing the corresponding form $B_h(\cdot,\cdot)$ as long as an algorithm for computing the action of $\tilde{M}_h$ is available. In addition, the above algorithm can be coded so that exactly two evaluations of $\tilde{M}_h$ and one evaluation of $\tilde{A}_h$ and $\tilde{A}_h^t$ are required per iterative step.

REMARK 2.2: The above algorithm is not new. Equivalent versions have appeared in the literature (e.g., Algorithm II of [**3**] and Algorithm 9.3 of [**8**]).

To bound the rate of convergence of the above conjugate gradient algorithm [**15**], it suffices to estimate constants satisfying inequalities of the form

$$C_2[V,V]_h \leq [\mathcal{A}V,V]_h \leq C_3[V,V]_h, \tag{2.6}$$

for all vectors $V \in R^{n_h}$. In fact, it is well know that i steps of the above algorithm reduces the initial error by a factor which is less than or equal to

$$\rho_i = 2\left(\frac{\sqrt{C_3}-\sqrt{C_2}}{\sqrt{C_3}+\sqrt{C_2}}\right)^i \tag{2.7}$$

in an appropriate norm.

For the purposes of analysis, it is more convenient to deal with operators defined on spaces. We start with the case where $B_h(\cdot,\cdot) = \hat{A}(\cdot,\cdot)$. Let V, W be vectors and set

$$v = \sum_{i=1}^{n_h} V^i \phi_h^i \text{ and } w = \sum_{i=1}^{n_h} W^i \phi_h^i.$$

Note that $\tilde{M}_h \tilde{A}_h$ provides a matrix representation of an operator on the subspace $\mathcal{S}_h$. Indeed, if $W = \tilde{M}_h \tilde{A}_h V$ then

$$\hat{A}(w,\phi) = A(v,\phi) \qquad \text{for all } \phi \in \mathcal{S}_h. \tag{2.8}$$

We note, in addition, that

$$[V,W]_h = \hat{A}(v,w).$$

Let $w \in \mathcal{S}_h$ solve (2.8). We can write $w = R_h v$ where $R_h : \mathcal{S}_h \mapsto \mathcal{S}_h$ is defined by $R_h v = \chi$ where χ is the unique function in $\mathcal{S}_h$ satisfying

$$\hat{A}(\chi,\phi) = A(v,\phi) \qquad \text{for all } \phi \in \mathcal{S}_h. \tag{2.9}$$

Note that $R_h = \hat{A}_h^{-1} A_h$ where $A_h : \mathcal{S}_h \mapsto \mathcal{S}_h$ is defined by

$$(A_h v, \theta) = A(v,\theta) \qquad \text{for all } \theta \in \mathcal{S}_h$$

and $\hat{A}_h$ is defined analogously. It immediately follows that

$$[\mathcal{A}V,V]_h = \hat{A}(R_h v, R_h v).$$

Thus, (2.6) can be rewritten

$$C_2 \|v\|_{H^1}^2 \le \|R_h v\|_{H^1}^2 \le C_3 \|v\|_{H^1}^2 \qquad \text{for all } v \in \mathcal{S}_h. \tag{2.10}$$

Of course, (2.6) is not the same as (2.10) when $B_h(\cdot,\cdot) \neq \hat{A}(\cdot,\cdot)$. In that case, (2.1) and (2.10) can be combined to show that

$$C_2 C_1^{-2}[V,V]_h \le [\mathcal{A}V,V]_h \le C_3 C_0^{-2}[V,V]_h, \tag{2.11}$$

Thus, in either case, estimates for the rate of convergence of the above conjugate gradient algorithm will follow from estimates of the form of (2.10) provided that B_h is a good preconditioner for $\hat{A}$ on $\mathcal{S}_h$.

REMARK 2.3: As we have seen above, in the case when $B_h(\cdot,\cdot) = \hat{A}(\cdot,\cdot)$, $\tilde{M}_h\tilde{A}_h$ is a matrix representation of the operator R_h. Similarly, $\tilde{M}_h\tilde{A}_h^t$ is a matrix representation of the adjoint of the operator R_h with respect to the inner product $\hat{A}(\cdot,\cdot)$. Thus, the reformulation (2.3) is developed by first preconditioning the system (i.e. applying $\tilde{M}_h$) and subsequently multiplying by the $\hat{A}(\cdot,\cdot)$ adjoint of the preconditioned system. In many applications, $\hat{A}(\cdot,\cdot)$ corresponds to the inner product in $H^1(\Omega)$.

Let the operator $R : \mathcal{S} \mapsto \mathcal{S}$ be defined by $Rv = \chi$ where χ is the unique function in $\mathcal{S}$ satisfying

$$\hat{A}(\chi,\theta) = A(v,\theta) \qquad \text{for all } \theta \in \mathcal{S}. \tag{2.12}$$

The following theorem provides some hypotheses for proving the lower estimate in (2.10). Note that the upper estimate follows immediately from the boundedness assumption on the form $A(\cdot,\cdot)$. For the purposes of this theorem and elsewhere in the remainder of this manuscript, the character C will be used to denote a generic positive constant which may take on different values in different places. The constant C will always be independent of the mesh parameter h.

THEOREM 2.1. *Let R_h and R be defined respectively by (2.9) and (2.12). Assume that the following estimates hold:*

$$\|\theta\|_{H^1} \le C\left\{\|R_h\theta\|_{H^1} + \|\theta\|\right\} \qquad \text{for all } \theta \in \mathcal{S}_h. \tag{2.13}$$

There exists a fixed positive number γ with

$$\|(R - R_h)\theta\| \le Ch^\gamma \|\theta\|_{H^1} \qquad \text{for all } \theta \in \mathcal{S}. \tag{2.14}$$

For any $\epsilon > 0$ there exists a constant C_ϵ such that

$$\|\theta\| \le C_\epsilon \|R\theta\| + \epsilon \|\theta\|_{H^1} \qquad \text{for all } \theta \in \mathcal{S}. \tag{2.15}$$

Then there exists a constant $h_0 > 0$ such that for $h \le h_0$,

$$\|\theta\|_{H^1} \le C \|R_h\theta\|_{H^1} \qquad \text{for all } \theta \in \mathcal{S}_h.$$

This theorem was essentially given in [**3**]. Its proof is obvious. Applications of the theorem are given in detail in Section 4.

3. A COARSE GRID REDUCTION TECHNIQUE.

We develop a method which utilizes a "coarse grid solve" in this section. This results in a system with a reduced number of unknowns which is either positive definite in the case when A is symmetric indefinite or has a positive definite symmetric part in the general case.

To describe this technique, we assume that along with our approximation space $\mathcal{S}_h$, we are given a "coarser" subspace $\mathcal{S}_H \subset \mathcal{S}_h$ with mesh parameter $H > h$. We are interested in solving (1.2).

The algorithm is developed in terms of the projector $P_H : \mathcal{S} \mapsto \mathcal{S}_H$ which is defined by $P_H v = w$ where $w \in \mathcal{S}_H$ is the solution of

$$A(w, \phi) = A(v, \phi) \qquad \text{for all } \phi \in \mathcal{S}_H. \tag{3.1}$$

Subsequently, we shall impose sufficient conditions such that the solvability of (3.1) is guaranteed provided that H is sufficiently small. Throughout this development, we shall assume that the solution of the coarse grid problems, e.g. (3.1) is relatively inexpensive. We then write the solution of (1.2) as

$$u_h = P_H u_h + (I - P_H) u_h. \tag{3.2}$$

We next provide a technique for computing $(I - P_H)u_h$. Let Q_H denote the L^2 projection onto $\mathcal{S}_H$, i.e. $Q_H v$ is the unique function in $\mathcal{S}_H$ satisfying

$$(Q_H v, w) = (v, w) \qquad \text{for all } w \in \mathcal{S}_H.$$

Note that $(I - P_H)u_h$ satisfies the equation

$$A((I - P_H)u_h, \phi) = (f, \phi) - A(P_H u_h, \phi) \qquad \text{for all } \phi \in \mathcal{S}_h. \tag{3.3}$$

Let $\mathcal{S}_h^{\perp}$ be defined as the image of $\mathcal{S}_h$ under the operator $I - Q_H$. Then it is easy to see that $(I - P_H)u_h = (I - P_H)v$ for any function $v \in \mathcal{S}_h$ satisfying

$$A((I - P_H)v, \phi) = (f, \phi) - A(P_H u_h, \phi) \qquad \text{for all } \phi \in \mathcal{S}_h^{\perp}. \tag{3.4}$$

We will later impose conditions which guarantee the existence of a unique function $v \in \mathcal{S}_h^{\perp}$ satisfying (3.4). The algorithm for the reduced equations is, essentially, a scheme for solving (3.4).

Before stating the reduced algorithm, we provide a theorem which guarantees existence and uniqueness of solutions to (3.1) and (3.4).

THEOREM 3.1. *Assume that the form A satisfies a Gårding inequality of the form*

$$C_4 \hat{A}(v, v) - C_5 \|v\|^2 \leq A(v, v) \qquad \text{for all } v \in \mathcal{S}. \tag{3.5}$$

Assume that there is a fixed $\gamma > 0$, such that functions $w \in \mathcal{S}_H$ satisfying (3.1) also satisfy

$$\|v - w\| \leq C H^{\gamma} \|v - w\|_{H^1} \qquad \text{for all } v \in \mathcal{S}. \tag{3.6}$$

Then there exists a positive constant h_0 such that for $H < h_0$, (3.1) is uniquely solvable and (3.4) is uniquely solvable in $\mathcal{S}_h^\perp$. Moreover, there exists a positive constant C such that

$$C_6\hat{A}((I-P_H)v,(I-P_H)v) \le A((I-P_H)v,(I-P_H)v), \tag{3.7}$$

for all $v \in \mathcal{S}_h$.

Proof: The unique solvability of (3.1) under the above assumptions follows applying an argument given in [**17**]. Inequality (3.7) follows combining (3.5) and (3.6).

We need only show the unique solvability of (3.4) on $\mathcal{S}_h^\perp$. Inequality (3.7) implies that the quadratic form on the left hand side of (3.4) has a non-negative symmetric part which vanishes only on functions v with $(I-P_H)v = 0$. For $v \in \mathcal{S}_h^\perp$,

$$\|v\| = \|(I-Q_H)(I-P_H)v\| \le \|(I-P_H)v\|.$$

This completes the proof of the theorem.

Theorem 3.1 justifies the following three step algorithm for computing the solution u_h of (1.2).

Algorithm 3.1.

(1) *Compute P_Hu_h and the data on right hand side of (3.3).*
(2) *Find the unique function $v \in \mathcal{S}_h^\perp$ satisfying (3.4).*
(3) *Compute $(I-P_H)v$ and set $u_h = P_Hu_h + (I-P_H)v$.*

We propose to solve problem (3.4) by preconditioned iteration where the preconditioner is defined in all of $\mathcal{S}_h$. We shall illustrate this by first considering the case when the form A is symmetric (but indefinite). To this end, we note that the function $v \in \mathcal{S}_h^\perp$ satisfying (3.4) is the solution of the operator equation

$$A_h(I-P_H)v = Q_hf - A_hP_Hu_h \tag{3.8}$$

where Q_h is the L^2 projection onto $\mathcal{S}_h$. By (3.7), $A_h^\perp = A_h(I-P_H)$ is a (symmetric) positive definite operator on $\mathcal{S}_h^\perp$ under the symmetry assumption on A. As in Section 2, let B_h be a symmetric positive definite preconditioning form and define the corresponding preconditioning operator $M_h : \mathcal{S}_h \mapsto \mathcal{S}_h$ by

$$B_h(M_h\chi,\theta) = (\chi,\theta) \qquad \text{for all } \theta \in \mathcal{S}_h.$$

Then, $M_h^\perp = (I-Q_H)M_h$ is a symmetric positive definite operator on $\mathcal{S}_h^\perp$

A conjugate gradient algorithm for solving the equation

$$\begin{aligned} M_h^\perp A_h^\perp v &= M_h^\perp(Q_hf - A_hP_Hu_h) \\ &= M_h^\perp(I-Q_H)(Q_hf - A_hP_Hu_h) \end{aligned} \tag{3.9}$$

with inner product $((M_h^\perp)^{-1}\cdot,\cdot)$ follows. It is assumed that an initial iterate v_0 is provided. For example, one could take $v_0 = 0$.

ALGORITHM 3.2.

(1) *Let* v_0 *be given.*
(2) *Set* $r_0 = (Q_h f - A_h P_H u_h - A_h^\perp v_0)$ *and* $p_0 = M_h^\perp r_0$.
(3) *For* $i \geq 0$ *define:*

$$\begin{aligned}
\alpha_i &= \frac{(r_i, p_i)}{(A_h^\perp p_i, p_i)}, \\
v_{i+1} &= v_i + \alpha_i p_i, \\
r_{i+1} &= r_i - \alpha_i A_h^\perp p_i, \\
\beta_i &= \frac{(r_{i+1}, M_h^\perp A_h^\perp p_i)}{(A_h^\perp p_i, p_i)} \\
p_{i+1} &= M_h^\perp r_{i+1} - \beta_i p_i.
\end{aligned}$$

REMARK 3.1: We note some of the computational properties of the above algorithm. Notice that the inner product $((M_h^\perp)^{-1}\cdot, \cdot)$ never enters into the computation. Thus, it is possible to use a preconditioner M_h without explicitly knowing $(M_h^\perp)^{-1}$ as long as an algorithm for computing the action of M_h is available. In addition, the above algorithm can be coded so that exactly one evaluation of $M_h^\perp$ and $A_h^\perp$ is required per iterative step. Also, the algorithm can be implemented using the basis for $\mathcal{S}_h$ explicitly avoiding the use of a computational basis for $\mathcal{S}_h^\perp$. Finally, the algorithm requires the evaluation of $(I - P_H)$ and $(I - Q_H)$ each time $A_h^\perp$ and $M_h^\perp$ are evaluated. An algorithm which avoids the $(I - Q_H)$ evaluations will be described later. First, we consider the rate of convergence of the algorithm just described.

To analyze the rate of convergence for the above, we must provide some estimates for the eigenvalues of the operator $M_h^\perp A_h^\perp$ (see (2.7)). To this end, we let $\|\cdot\|_{B_h^\perp} = ((M_h^\perp)^{-1}\cdot, \cdot)^{1/2}$ and prove the following lemma.

LEMMA 3.1. *Assume that (2.1) and (3.7) hold. Then for any* $u, v \in \mathcal{S}_h^\perp$,

$$|A((I - P_H)u, v)| \leq C_U \|u\|_{B_h^\perp} \|v\|_{B_h^\perp} \tag{3.10}$$

and

$$C_L \|u\|_{B_h^\perp}^2 \leq A((I - P_H)u, u). \tag{3.11}$$

PROOF: Let $P_B : \mathcal{S}_h \mapsto \mathcal{S}_H$ be defined by

$$B_h(P_B v, \theta) = B_h(v, \theta) \qquad \text{for all } \theta \in \mathcal{S}_H.$$

It is easy to check that

$$((M_h^\perp)^{-1} w, w) = B_h((I - P_B)w, w) \qquad \text{for all } w \in \mathcal{S}_h^\perp.$$

First we prove (3.10). By the boundedness of A,

$$\begin{aligned}
|A((I - P_H)u, v)| &= |A((I - P_H)u, (I - P_B)v)| \\
&\leq C_7 \|(I - P_H)u\|_{H^1} \|(I - P_B)v\|_{H^1}.
\end{aligned}$$

Using (3.7), we get

$$\begin{aligned}\|(I-P_H)u\|_{H^1}^2 &\le CA((I-P_H)u,(I-P_H)u) = CA((I-P_H)u,(I-P_B)u)\\ &\le C\,\|(I-P_H)u\|_{H^1}\,\|(I-P_B)u\|_{H^1}\,.\end{aligned}$$

Dividing by $\|(I-P_H)u\|_{H^1}$, using (2.1) and the identity

$$A((I-P_H)u,v) = A((I-P_H)u,(I-P_B)v)$$

proves inequality (3.10).

We next prove (3.11). By (2.1), (3.7) and obvious properties of P_B,

$$\begin{aligned}\|u\|_{B_h^\perp}^2 &\le B_h((I-P_H)u,(I-P_H)u)\\ &\le CA((I-P_H)u,u).\end{aligned}$$

This completes the proof of the lemma.

In the case when A_h is symmetric, Theorem 3.1 and Lemma 3.1 give

$$C_L((M_h^\perp)^{-1}u,u) \le (A_h^\perp u,u) \le C_U((M_h^\perp)^{-1}u,u) \qquad \text{for all } u\in \mathcal{S}_h^\perp.$$

Hence, the condition number of the operator $M_h^\perp A_h^\perp$ is bounded by C_U/C_L.

We note that in Algorithm 3.1, it suffices to compute any function $\tilde v$ which differs from v by something in $\mathcal{S}_H$ since only $(I-P_H)v$ is required in Step 3. This observation can be used to develop a more efficient algorithm which avoids the computation of $(I-Q_H)$.

We shall describe this algorithm in terms of the computational basis for $\mathcal{S}_h$ already introduced in Section 2 as well as a computational basis $\{\phi_H^i\}$ for $\mathcal{S}_H$. Let $\tilde A_H$ denote the stiffness matrix for the form A with respect to this basis and $\tilde I_H$ denote the matrix operator which takes coefficients corresponding to a function in $\mathcal{S}_H$ into the coefficients which represent the function in the basis for $\mathcal{S}_h$. We note that $\tilde I_H\tilde A_H^{-1}\tilde I_H^t\tilde A_h$ is the matrix representation for the operator P_H. Similarly, $\tilde M_h\tilde A_h$ is the matrix representation of the product M_hA_h. For convenience, we shall denote by $\tilde P_H^\perp$ the matrix operator $I-\tilde I_H\tilde A_H^{-1}\tilde I_H^t\tilde A_h$ and by F_h the vector with entries $\{(f,\phi_h^i)\}$. We consider the following algorithm.

Algorithm 3.3.

(1) *Let W_0 be given.*
(2) *Set $R_0 = F_h - \tilde A_h\tilde I_H\tilde A_H^{-1}\tilde I_H^tF_h - \tilde A_h\tilde P_H^\perp W_0$ and $P_0=\tilde M_hR_0$.*
(3) *For $i\ge 0$ define:*

$$\begin{aligned}\alpha_i &= \frac{R_i^tP_i}{P_i^t\tilde A_h\tilde P_H^\perp P_i},\\ W_{i+1} &= W_i+\alpha_iP_i,\\ R_{i+1} &= R_i-\alpha_i\tilde A_h\tilde P_H^\perp P_i,\\ \beta_i &= \frac{(\tilde M_hR_{i+1})^t\tilde A_h\tilde P_H^\perp P_i}{P_i^t\tilde A_h\tilde P_H^\perp P_i}\\ P_{i+1} &= \tilde M_hR_{i+1}-\beta_iP_i.\end{aligned}$$

REMARK 3.2: The vectors generated in Algorithm 3.3 are related to the functions generated in Algorithm 3.2. Let w_i denote the function in $\mathcal{S}_h$ corresponding to the coefficient vector W_i and set $v_0 = (I - Q_H)w_0$. In general, w_i is not in $\mathcal{S}_h^\perp$ however the above algorithm is constructed so that $(I - Q_H)w_i = v_i$ where v_i is generated by Algorithm 3.2, i.e. w_i differs from v_i by a function in $\mathcal{S}_H$. Thus, after i steps of Algorithm 3.3, $\tilde{P}_H^\perp W_i$ will be equal to the coefficients of the function $(I - P_H)v_i$. Thus, $\tilde{P}_H^\perp W_i$ need be computed only after a practical convergence criterion is met. Note that Algorithm 3.3 avoids explicit computation of the action of the operator $(I - Q_H)$.

REMARK 3.3: Algorithm 3.3 can be programmed in such a way as to only require one evaluation of each of the operators $\tilde{A}_h$, $\tilde{P}_H^\perp$, and $\tilde{M}_h$ per iterative step.

In the general nonsymmetric case, inequality (3.7) only guarantees that the operator $A_h^\perp$ has a positive definite symmetric part. It is then possible to use, for example, GMRES [**16**] to develop an algorithm for computing a function v such that $(I - P_H)u_h = (I - P_H)v$. Specifically, we apply GMRES to the preconditioned equation (3.9) in the inner product $((M_h^\perp)^{-1}\cdot, \cdot)$. We subsequently provide a modification of this algorithm which avoids evaluation of $(M_h^\perp)^{-1}$ as well as Q_H.

Mathematically, GMRES provides an algorithm for computing a best approximation in a certain Krylov space. In our application, one assumes an initial approximation $v_0 \in \mathcal{S}_h^\perp$ and defines the Krylov space $\mathcal{K}_m$ to be the span of the vectors $r_0, M_h^\perp A_h^\perp r_0, \ldots, (M_h^\perp A_h^\perp)^{m-1} r_0$ where $r_0 = M_h^\perp(Q_h f - A_h P_H u_h - A_h v_0)$. The improved approximation v_m is equal to $v_0 + \chi$ where χ is the unique function in $\mathcal{S}_h^\perp$ which minimizes the residual error

$$\left\| r_0 - M_h^\perp A_h^\perp \theta \right\|_{B_h^\perp}$$

over all functions $\theta \in \mathcal{K}_m$.

Arguments analogous to those used to develop Algorithm 3.3 lead to the following GMRES like algorithm for computing for a vector of coefficients W_m corresponding to a function $\tilde{v}_m$ which differs from v_m by a function in $\mathcal{S}_H$.

ALGORITHM 3.4.

(1) *Let* W_0 *be given.*
(2) *Set* $R_0 = F_h - \tilde{A}_h \tilde{I}_H \tilde{A}_H^{-1} \tilde{I}_H^t F_h - \tilde{A}_h \tilde{P}_H^\perp W_0$ *and* $V_1 = \tilde{M}_h R_0 / (R_0^t \tilde{M}_h R_0)^{1/2}$.
(3) *For* $j = 1, 2, \ldots, m$ *define:*

$$\begin{aligned} &\textit{for } i = 1, 2, \ldots, j \textit{ define} \\ h_{ij} &= V_i^t \tilde{A}_h \tilde{P}_H^\perp V_j; \\ \hat{V}_{j+1} &= \tilde{M}_h \tilde{A}_h \tilde{P}_H^\perp V_j - \sum_{i=1}^{j} h_{ij} V_i; \\ h_{j+1,j} &= (\hat{V}_{j+1}^t \tilde{A}_h \tilde{P}_H^\perp V_j)^{1/2}; \\ V_{j+1} &= \hat{V}_{j+1} / h_{j+1,j}. \end{aligned}$$

(4) *Define* $W_m = W_0 + \chi$ *where*

$$\chi = \sum_{i=1}^{m} y_{mi} V_i,$$

and the coefficients y_{mi} are chosen to minimize the quantity

$$(R_0 - \tilde{A}_h \tilde{P}_H^\perp \chi)^t \tilde{M}_h (R_0 - \tilde{A}_h \tilde{P}_H^\perp \chi).$$

REMARK 3.4: Steps (2) and (3) above are essentially the Arnoldi algorithm for implementing a Gram-Schmidt orthorgonalization of the Krylov space $\mathcal{K}_m$. In general, the use such an orthogonal basis results in a more stable numerical algorithm.

REMARK 3.5: A particularly efficient technique for implementing the minimization process in Step (4) is given in [**16**].

REMARK 3.6: The above algorithm becomes somewhat inefficient if m becomes too large. Let n denote the number of unknowns. Then implementation of the above algorithm requires storage on the order of mn and operations on the order of nm^2. Accordingly, it is often convenient to fix m and repetitively restart the algorithm.

It is known (see [**gmres**]) that the rate of convergence of GMRES when applied to a problem with a positive definite symmetric part can be bounded in terms of the smallest eigenvalue of the symmetric part and the norm of the original operator. Let $e_m = (I - P_H)(v - \tilde{v}_m)$. Applying the above mentioned analysis shows that

$$(M_h^\perp A_h^\perp e_m, A_h^\perp e_m) \le (1 - \alpha^2/\beta^2)^{m/2} (M_h^\perp A_h^\perp e_0, A_h^\perp e_0).$$

The constant α above is the smallest eigenvalue of the symmetric part of the operator $M_h^\perp A_h^\perp$. By Theorem 3.1 and Lemma 3.1, $\alpha \ge C_L$. The constant β above is the norm of the operator $M_h^\perp A_h^\perp$. By Theorem 3.1 and Lemma 3.1, $\beta \le C_U$.

4. APPLICATIONS.

In this section, we consider applications of the results of the previous to second order elliptic boundary value problems. Let Ω be a domain in d dimensional Euclidean space and consider the problem

$$\mathcal{L}u = f \qquad \text{in } \Omega, \tag{4.1}$$

$$u = 0 \qquad \text{on } \Gamma_D \tag{4.2}$$

$$\frac{\partial u}{\partial v} + \beta(x) u = 0 \qquad \text{on } \Gamma_N. \tag{4.3}$$

Here $\partial\Omega = \Gamma_D \cup \Gamma_N$ and $\frac{\partial}{\partial v}$ denotes the outward co-normal derivative on $\partial\Omega$. The operator $\mathcal{L}$ is given by

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

We assume that the matrix $\{a_{ij}(x)\}$ is symmetric for each x and uniformly positive definite and bounded. We also assume that β is in $L^\infty(\partial\Omega)$.

We shall also consider oblique derivative problems when $d = 2$. In this case, we require that $\partial\Omega$ be piecewise smooth and let t denote the (positively oriented) tangential direction along $\partial\Omega$. (4.3) is replaced by

$$\frac{\partial u}{\partial v} + \beta(x)u + \sigma(x)\frac{\partial u}{\partial t} = 0 \qquad \text{on } \Gamma_N. \tag{4.4}$$

We assume that σ is in the Sobolev space $W^{1,\infty}(\partial\Omega)$ (c.f. [**1**],[**10**]).

To develop and analyze this example, we shall need to use Sobolev spaces on Ω and $\partial\Omega$. For non-negative l, the Sobolev space of order l on Ω will be denoted $H^l(\Omega)$ with norm $\|\cdot\|_l$. On the boundary, the space of order l will be denoted $H^l(\partial\Omega)$ with norm $|\cdot|_l$.

We next provide the weak formulation of (?). Define $\mathcal{S}$ to be the functions in $H^1(\Omega)$ whose trace vanish on Γ_D. Let $(\cdot,\cdot)$ denote the $L^2(\Omega)$ inner product and $< \cdot,\cdot >$ denote the $L^2(\partial\Omega)$ inner product. The weak formulation of (4.1)-(4.4) is: Find $u \in \mathcal{S}$ such that

$$A(u,\psi) = (f,\psi) \qquad \text{for all } \psi \in \mathcal{S}.$$

Here, the form A is defined by

$$\begin{aligned} A(\psi,\theta) = \sum_{i,j=1}^{d} (a_{ij}\frac{\partial \psi}{\partial x_j}, \frac{\partial \theta}{\partial x_i}) + \sum_{i=1}^{d}(b_i \frac{\partial \psi}{\partial x_i}, \theta) \\ + (c\psi,\theta) + < \beta\psi + \sigma\frac{\partial \psi}{\partial t}, \theta > . \end{aligned} \tag{4.5}$$

By convention, we take $\sigma = 0$ when $d \neq 2$. Clearly, A is a bounded bilinear form with respect to the norm in $H^1(\Omega)$.

To prove the hypotheses of the theorems in the previous section, we shall use the equivalent inner product

$$\hat{A}(\psi,\theta) = \sum_{i,j=1}^{d} (a_{ij}\frac{\partial \psi}{\partial x_j}, \frac{\partial \theta}{\partial x_i}) + (\psi,\theta)$$

on $H^1(\Omega)$.

We note that the Gårding inequality (3.5) holds for the form A. In fact, the term associated with the tangential derivative results in no difficulty since

$$< \sigma\frac{\partial \psi}{\partial t}, \psi > = - < \frac{\partial \sigma}{\partial t}\psi, \psi > - < \sigma\psi, \frac{\partial \psi}{\partial t} >$$

and hence

$$< \sigma\frac{\partial \psi}{\partial t}, \psi > = -1/2 < \frac{\partial \sigma}{\partial t}\psi, \psi > . \tag{4.6}$$

The remaining terms are handled by standard perturbation arguments.

Remark 4.1: Many approaches for providing estimates for non-symmetric and indefinite problems require that the form differs from a "nice" form (e.g., $\hat{A}(\cdot,\cdot)$) by a lower order perturbation. Our theory only requires that the symmetric part of the form differs from a nice form by a lower order perturbation. Note that the oblique derivative term in (4.4) is the same strength as the second order derivative terms in (4.1). However, (4.6) shows that term $< \sigma \frac{\partial \psi}{\partial t}, \psi >$ is weaker than $\hat{A}(\psi, \psi)$.

We shall assume that (?) and its adjoint have unique solutions. In addition, we assume that solutions of adjoint problem satisfy a modest amount of elliptic regularity. Specifically, given $g \in L^2(\Omega)$, let v solve

$$A(\phi, v) = (\phi, g) \qquad \text{for all } \phi \in \mathcal{S}.$$

We assume that for some $\gamma \in (0, 1]$, there is a constant C not depending on $g \in L^2(\Omega)$ such that

$$\|v\|_{1+\gamma} \leq C \|g\| . \tag{4.7}$$

Such estimates for problems with boundary conditions (4.3) can be found in [**9**],[**14**]. In the case of (4.4), we refer to [**18**].

We approximate the solution of (?) by using the finite element method. This involves the use of a sequence of approximations spaces $\{\mathcal{S}_h\}$ which are subspaces of $\mathcal{S}$ and indexed by $h \in (0, 1]$. Many examples of such constructions can be found in [**2**],[**7**]. We put very little restriction on these spaces other than the requirement that they satisfy standard approximation properties. In particular, we allow mesh refinement in all of our applications. In this case, the parameter h corresponds to the size of the largest triangle or finite element.

For the purposes of developing iterative algorithms, we only require that the subspaces satisfy

$$\inf_{\phi \in \mathcal{S}_h} \|v - \phi\|_1 \leq Ch \|v\|_2 . \tag{4.4}$$

The inequality (?) holds with fixed C and for all $v \in H^2(\Omega) \cap \mathcal{S}$. This does not exclude the use of higher order spaces and refinement in order to obtain better solution accuracy.

The Galerkin approximation $u_h \in \mathcal{S}_h$ to the solution u of (?) is defined by (1.2). Inequality (3.6) follows applying the standard finite element duality technique [**2**],[**7**], (4.7) and (?). Thus we have shown that the hypotheses of Theorem 3.1 hold for this application. The following theorem show that the hypothesis of Theorem 2.1 also hold for this application.

THEOREM 4.1. *Under the above assumptions and definitions of $\mathcal{S}$, $\mathcal{S}_h$, $\hat{A}_h$, and A, (?)-(2.15) hold, i.e., the hypotheses of Theorem 2.1 hold.*

We shall use the following lemma in the proof of Theorem 4.1. Its proof will be given in the appendix.

LEMMA 4.1. *Let $0 < \epsilon < 1/2$ be fixed. For all $\phi \in \mathcal{S}$ and $\psi \in H^{1+\epsilon}(\Omega) \cap \mathcal{S}$,*

$$\hat{A}(\phi, \psi) \leq C \|\phi\|_{1-\epsilon} \|\psi\|_{1+\epsilon} . \tag{4.8}$$

Proof of Theorem 3: We first prove (2.13). By (3.5), for $\theta \in \mathcal{S}_h$,

$$\begin{aligned} \hat{A}(\theta,\theta) &\le C_4^{-1}(C_4\hat{A}(\theta,\theta) - C_5\|\theta\|^2 + C_5\|\theta\|^2) \\ &\le C_4^{-1}(A(\theta,\theta) + C_5\|\theta\|^2) \\ &= C_4^{-1}(\hat{A}(R_h\theta,\theta) + C_5\|\theta\|^2). \end{aligned} \tag{4.9}$$

Inequality (2.13) follows from (4.9) and obvious manipulations.

We next prove (2.14). Let P_h denote the elliptic projection onto the space $\mathcal{S}_h$, i.e. $P_h : \mathcal{S} \mapsto \mathcal{S}_h$ be defined analogously to P_H in (3.1). From the definition, it is immediate that $R_h = P_h R$. Inequality (3.6) with h replacing H gives

$$\|(R - R_h)\phi\| \le Ch^\gamma \|R\phi\|_1 .$$

Note that R is defined by the relation

$$\hat{A}(Rv,\phi) = A(v,\phi) \qquad \text{for all } \phi \in \mathcal{S}.$$

Taking $\phi = Rv$ and using the boundedness of $A(\cdot,\cdot)$ gives

$$\|Rv\|_1 \le C\|v\|_1 . \tag{4.10}$$

Inequality (2.14) follows.

Before proving (2.15), we first define some additional notation. Let $\mathcal{S}^{-1}$ denote the dual of $\mathcal{S}$, i.e., $\mathcal{S}^{-1}$ is the set of distributions on Ω for which the norm

$$\|u\|_{-1} = \sup_{\phi\in\mathcal{S}} \frac{(u,\phi)}{\|\phi\|_1}$$

is finite. In addition, we define the following operators:

(1) $A : \mathcal{S} \mapsto \mathcal{S}^{-1}$ by $Aw = v$ where v is the unique function in $\mathcal{S}^{-1}$ satisfying

$$(v,\phi) = A(w,\phi) \qquad \text{for all } \phi \in \mathcal{S}.$$

(2) $\hat{A} : \mathcal{S} \mapsto \mathcal{S}^{-1}$ by $\hat{A}w = v$ where v is the unique function in $\mathcal{S}^{-1}$ satisfying

$$(v,\phi) = \hat{A}(w,\phi) \qquad \text{for all } \phi \in \mathcal{S}.$$

(3) $T : \mathcal{S}^{-1} \mapsto \mathcal{S}$ by $Tw = v$ where v is the unique function in $\mathcal{S}$ satisfying

$$A(v,\phi) = (w,\phi) \qquad \text{for all } \phi \in \mathcal{S}.$$

(4) $T^* : \mathcal{S}^{-1} \mapsto \mathcal{S}$ by $T^*w = v$ where v is the unique function in $\mathcal{S}$ satisfying

$$A(\phi,v) = (w,\phi) \qquad \text{for all } \phi \in \mathcal{S}.$$

(5) $\hat{T} : \mathcal{S}^{-1} \mapsto \mathcal{S}$ by $\hat{T}w = v$ where v is the unique function in $\mathcal{S}$ satisfying

$$\hat{A}(v,\phi) = (w,\phi) \qquad \text{for all } \phi \in \mathcal{S}.$$

Note that $R = \hat{T}A$.

We now prove (2.15). For $u \in \mathcal{S}$,

$$\|u\|^2 = A(u, T^*u) = (Au, T^*u) = \hat{A}(\hat{T}Au, T^*u).$$

Applying Lemma 4.1 gives

$$\|u\|^2 \le C \left\|\hat{T}Au\right\|_{1-\epsilon} \|T^*u\|_{1+\epsilon}.$$

Inequality (4.7) and convexity give

$$\begin{aligned} \|u\| &\le C \|Ru\|_{1-\epsilon} \le C \|Ru\|^{\epsilon} \|Ru\|_1^{1-\epsilon} \\ &\le C_\delta \|Ru\| + \delta \|Ru\|_1 . \end{aligned} \tag{4.11}$$

Inequality (2.15) follows immediately follows combining (4.10) and (4.11). This completes the proof of the theorem.

5. Numerical experiments.

In this section, we provide the results of numerical examples illustrating the theory developed in earlier sections. We shall consider a model problem in two dimensional space. Specifically, we consider problem (4.1)-(4.4) where

$$\begin{aligned} \Omega &= [0,1] \times [0,1] \\ \mathcal{L}u &= -\Delta u + au_x + bu_y - du. \end{aligned}$$

The constants a, b, d are parameters in the examples to be presented later. Examples will be given when $\Gamma_1 = \emptyset$ and when $\Gamma_1 \neq \emptyset$ and $\sigma \neq 0$ on Γ_1. Here Δ denotes the Laplacian.

The sequence of subspaces are the usual finite element approximation spaces. Specifically, the domain Ω is first partitioned into $N \times N$ square subdomains of side length $1/N$. Each smaller square is then divided into two triangles by one of the diagonals. The approximation space $\mathcal{S}_h$ is defined to be the set of functions which are continuous on Ω, piecewise linear with respect to the triangulation, and vanish on Γ_0.

We seek the Galerkin solution $u_h \in \mathcal{S}_h$ satisfying (1.2) where $A(\cdot,\cdot)$ is defined by (4.5).

A preconditioning form for the equation (1.2) is as follows

$$B_h(u, \phi) = \int_\Omega u_x\phi_x + u_y\phi_y \, dxdy.$$

For our first example, we consider $\Gamma_1 = \emptyset$ and $a = 0$, $b = 0$. By **ite** we denote the number of iterative steps for an iterative method to get the relative accuracy 10^{-8}, i.e. $\|W_i\| \le 10^{-8}\|W_0\|$, where W_i is the approximation of U_h in i^{th} iterative step. Let κ denote the condition number of the system (2.3) and N_0 be the partition parameter of "coarse mesh" analogous to N.

Table 5.1
Test results for Algorithm 2.1

N	σ	κ	**ite**
32	115	1136.8	61
64	115	1030.0	62
32	150	1953.7	75
64	150	2234.8	77

Table 5.2
Test results for Algorithm 3.3

N	N_0	σ	κ	**ite**
32	8	115	≤ 0	no convergence
32	16	115	1.84	11
64	16	115	2.21	12
32	16	150	2.26	11
64	16	150	2.79	13

Table 5.3
Test results for Algorithm 3.4

N	N_0	σ	κ	**ite**
32	8	115	??	??
32	16	115	??	??
64	16	115	??	??
32	16	150	??	??
64	16	150	??	??

Consider Algorithm 2.1. Table 5.1 gives us the condition numbers of the system (2.3) and the number of iterative steps for different values of N, N_0 and σ.

We compare these results with the results given by Algorithm 3.3 (see Table 5.2) and Algorithm 3.4 (see Table 5.2)

It is evident that Algorithms 3.3 and 3.4 outperform Algorithm 2.1. The Algorithm 3.4 is more robust than Algorithm 3.3, since it can perform computations for any parameter N_0.

We next consider the following parameters

$$\Gamma_1 = \{(x, y)\colon x = 0, 0 \leq y \leq 1\} \tag{5.1}$$

and $a = 1$, $b = 2$. Since the problem is not symmetric we compare only Algorithms 2.1 and

3.4 (for nonsymmetric problems we do not compute the condition number). The results are given in Table 5.4 and Table 5.5.

the code is not ready yet

Table 5.4
Test results for Algorithm 2.1

N	σ	$\kappa(?)$	**ite**
32	115	??	??
32	115	??	??
64	115	??	??
32	150	??	??
64	150	??	??

Table 5.5
Test results for Algorithm 3.4

N	N_0	σ	**ite**
32	8	115	??
32	16	115	??
64	16	115	??
32	16	150	??
64	16	150	??

References

1. J. P. Aubin, "Approximation of Elliptic Boundary-Value Problems," Wiley-Interscience, New York, 1972.
2. A.K. Aziz and I. Babuška, *Part I, survey lectures on the mathematical foundations of the finite element method*, in "The Mathematical Foundations of the Finite Element Method with Applications to Partial Differential Equations," (Ed. A.K. Aziz), Academic Press, New York, NY, 1972, pp. 1–362.
3. J.H. Bramble and J.E. Pasciak, *Preconditioned iterative methods for nonselfadjoint or indefinite elliptic boundary value problems*, in "Unification of finite element methods (ed. H. Kardestuncer)," Elsevier Science Publ. (North-Holland), New York, 1984, pp. 167 – 184.
4. J.H. Bramble, J.E. Pasciak and J. Xu, *The analysis of multigrid algorithms for nonsymmetric and indefinite elliptic problems*, Math. Comp. **51** (1988), 389–414.
5. X.-C. Cai and O. Widlund, *Domain decomposition algorithms for indefinite elliptic problems*, SIAM J. Sci. Stat. Comp. (to appear).
6. X.-C. Cai and J. Xu, *A preconditioned GMRES method for nonsymmetric and indefinite problems*, (Preprint).
7. P.G. Ciarlet, "The Finite Element Method for Elliptic Problems," North-Holland, New York, 1978.
8. H.C. Elman, *Iterative methods for large, sparse, nonsymmetric systems of linear equations*, Yale Univ. Dept. of Comp. Sci. Rep. 229, (1982).

9. R.B. Kellogg, in "Interpolation between subspaces of a Hilbert space," Univ. of Maryland,, Inst. Fluid Dynamics and App. Math., Tech. Note BN-719, 1971.
10. J.L. Lions and E. Magenes, "Problèmes aux Limites non Homogènes et Applications," Dunod, Paris, 1968.
11. Mandel, J., *Multigrid convergence for nonsymmetric, indefinite variational problems and one smoothing step*, in "Proc. Copper Mtn. Conf. Multigrid Methods," Applied Math. Comput., 1986, pp. 201–216.
12. T. Manteuffel and S. Parter.
13. S. McCormick (ed.), in "Multigrid Methods," SIAM, Philadelphia, Penn., 1987.
14. J. Nečas, "Les Méthodes Directes en Théorie des Équations Elliptiques," Academia, Prague, 1967.
15. W.M. Patterson, 3rd, "Iterative Methods for the Solution of a Linear Operator Equation in Hilbert Space - A Survey," Lecture Notes in Mathematics 394, Springer-Verlag, New York, 1974.
16. Y. Saad and M.H. Schultz, *GMRES: A generalized minimal residual algorithm for solving nonsymmetric linear systems*, SIAM J. Sci. Stat. Comput. **7** (1986), 856 – 869.
17. A.H. Schatz, *An observation concerning Ritz-Galerkin methods with indefinite bilinear forms*, Math. Comp. **28** (1974), 959–962.
18. ?????.

1980 *Mathematics subject classifications*: Primary 65N30; Secondary 65F10

Department of Mathematics
Cornell University
Ithaca, NY 14853
E-mail : bramble@mssun7.msi.cornell.edu

Mathematical Sciences Institute
Cornell University
Ithaca, NY 14853
E-mail : leyk@macomb.tn.cornell.edu

Department of Applied Science
Brookhaven National Laboratory
Upton, NY 11973
E-mail : pasciak@bnl.gov