

ITERATIVE TECHNIQUES FOR TIME DEPENDENT STOKES PROBLEMS

JAMES H. BRAMBLE AND JOSEPH E. PASCIAK

January 1994

ABSTRACT. In this paper, we consider solving the coupled systems of discrete equations which arise from implicit time stepping procedures for the time dependent Stokes equations using a mixed finite element spatial discretization. At each time step, a two by two block system corresponding to a perturbed Stokes problem must be solved. Although there are a number of techniques for iteratively solving this type of block system, to be effective, they require a good preconditioner for the resulting pressure system, to be effective, they require a good preconditioner for the resulting pressure operator (Schur complement). In contrast to the time independent Stokes equations where the pressure operator is well conditioned, the pressure operator for the perturbed system becomes more ill conditioned as the time step is reduced (and/or the Reynolds number is increased). In this paper, we shall describe and analyze preconditioners for the resulting pressure systems. These preconditioners give rise to iterative rates of convergence which are independent of both the mesh size h as well as the time step and Reynolds number parameter k .

1. INTRODUCTION.

In this paper we analyze efficient iterative techniques for solving the coupled systems of linear equations which arise from fully discrete approximations of time dependent Stokes equations. Such systems also arise when the Navier-Stokes equations are advanced in time by using the modified method of characteristics.

The coupled linear systems have a block matrix representation of the form

$$(1.1) \quad \begin{pmatrix} \mathbf{A}_k & \mathbf{B}^* \\ \mathbf{B} & 0 \end{pmatrix} \begin{pmatrix} X \\ Y \end{pmatrix} = \begin{pmatrix} F \\ G \end{pmatrix}.$$

Here \mathbf{A}_k is symmetric, \mathbf{B}^* is the adjoint of \mathbf{B} and the parameter k is related to the time step size and Reynolds number. There are a number of techniques which lead to efficient iterative schemes for solving (1.1) provided that effective preconditioners for \mathbf{A}_k and $\mathbf{B}\mathbf{A}_k^{-1}\mathbf{B}^*$ are available (cf. [5], [6], [21]). These methods will be reviewed

1991 *Mathematics Subject Classification.* 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. DMS-9007185 and by the U.S. Army Research Office through the Mathematical Sciences Institute, Cornell University.

in Section 3. The goal of this paper is the analysis of preconditioners for the so-called Schur complement $\mathbf{B}\mathbf{A}_k^{-1}\mathbf{B}^*$ for the time dependent Stokes application.

For the time dependent Stokes application, the problem is posed in terms of two finite dimensional spaces, \mathbf{V}_h and W_h . The space \mathbf{V}_h consists of vector valued functions defined from a mesh of size h on a connected bounded domain Ω . The space W_h consists of scalar valued functions on a mesh of size h . It is assumed that the pair of spaces satisfy the classical Babuška-Brezzi condition: For all $p \in W_h$,

$$\|p\|^2 \leq c \sup_{\mathbf{v} \in \mathbf{V}_h} \frac{(p, \nabla \cdot \mathbf{v})^2}{D(\mathbf{v}, \mathbf{v})}.$$

Here $D(\cdot, \cdot)$ denotes the sum of the componentwise Dirichlet forms. In the case of the time dependent Stokes application, the form which results from the Schur complement is given by

$$(1.2) \quad (\mathbf{B}\mathbf{A}_k^{-1}\mathbf{B}^*p, p) = \sup_{\mathbf{v} \in \mathbf{V}_h} \frac{(p, \nabla \cdot \mathbf{v})^2}{D_k(\mathbf{v}, \mathbf{v})}.$$

Here $D_k(\mathbf{w}, \mathbf{w}) = (\mathbf{w}, \mathbf{w}) + kD(\mathbf{w}, \mathbf{w})$.

The preconditioner for $\mathbf{B}\mathbf{A}_k^{-1}\mathbf{B}^*$ which we shall study is of the form

$$(1.3) \quad \mathcal{K} = kI + Q_h T_\gamma$$

where I denotes the identity on W_h , Q_h denotes the $L^2(\Omega)$ projection onto W_h and T_γ is solution operator for a finite element approximation to the Neumann problem (see Section 4). Preconditioners of this form have been used for the solution of the coupled systems resulting from fully discrete approximations to time dependent Stokes problems (see, e.g., [8], [17]). Although numerical results were reported, there has been no theoretical work explaining the success of the preconditioner. In this paper, we provide a theory which shows that this preconditioner gives rise to rates of convergence which can be bounded independently of both the mesh size h as well as the time step parameter k .

The outline of the remainder of the paper is as follows. In Section 2, we develop the coupled linear systems corresponding to fully discrete time stepping approximations to the time dependent Stokes problem. In Section 3, we survey some iterative techniques for solving block systems of the form of (1.1). These techniques give rise to rapidly converging iterative schemes provided that effective preconditioners are available for \mathbf{A}_k and $\mathbf{B}\mathbf{A}_k^{-1}\mathbf{B}^*$. The problem of preconditioning \mathbf{A}_k has been well studied and effective algorithms are available. In Section 4, we give the analysis which provides bounds on the condition number for the preconditioner (1.3) applied to $\mathbf{B}\mathbf{A}_k^{-1}\mathbf{B}^*$. Finally, we present the results of numerical experiments which illustrate the effectiveness of the proposed preconditioner in Section 5.

2. THE PROBLEM AND NOTATION.

We shall be concerned with solving the discrete systems which result from time

stepping approximation to the linearized Navier–Stokes equations:

$$(2.1) \quad \frac{\partial \mathbf{u}}{\partial t} - \nu \Delta \mathbf{u} - \mathbf{grad} P = \mathbf{f} \quad \text{in } \mathcal{W} = \Omega \times (0, T),$$

$$(2.2) \quad \nabla \cdot \mathbf{u} = 0 \quad \text{in } \mathcal{W},$$

$$(2.3) \quad \mathbf{u} = 0 \quad \text{on } \partial\Omega \times [0, T],$$

$$(2.4) \quad \mathbf{u}(x, 0) = \mathbf{u}_0(x) \quad \text{in } \Omega,$$

$$(2.5) \quad \int_{\Omega} P(x, t) dx = 0 \quad \text{for each } t \in (0, T).$$

Here Ω is a domain d dimensional Euclidean space (with $d = 2$ or $d = 3$) and Δ denotes the componentwise Laplace operator. In addition, \mathbf{u} is a vector valued function and P is a scalar valued function on \mathcal{W} . We restrict our attention to the above model problem for simplicity.

The discrete approximation to (2.1)–(2.5) is defined in terms of the discrete approximation of the stationary Stokes problem. The Stokes problem is as follows: Find \mathbf{v} and Q satisfying

$$(2.6) \quad -\Delta \mathbf{v} - \mathbf{grad} Q = \mathbf{g} \quad \text{in } \Omega,$$

$$(2.7) \quad \nabla \cdot \mathbf{v} = f \quad \text{in } \Omega,$$

$$(2.8) \quad \mathbf{v} = 0 \quad \text{on } \partial\Omega,$$

$$(2.9) \quad \int_{\Omega} Q dx = 0.$$

The regularity properties for the stationary Stokes equation will play a fundamental role in the construction and analysis of iterative methods for solving the discrete systems arising from approximations to (2.1)–(2.5). These properties are defined in terms of Sobolev spaces. The Sobolev spaces $\{H^s(\Omega)\}$ for non-negative integers s are defined to be the distributions which along with their partial derivatives of order s are in $L^2(\Omega)$. A complete development and discussion of these spaces can be found in, e.g., [13], [16], [18]. The norm on $H^s(\Omega)$ will be denoted $\|\cdot\|_s$. For negative s , the space $H^s(\Omega)$ is defined by duality and is the set of linear functionals on $H^{-s}(\Omega)$ for which the norm

$$\|v\|_s = \sup_{\phi \in C^\infty(\Omega)} \frac{(v, \phi)}{\|\phi\|_{-s}}$$

is finite. Here (\cdot, \cdot) denotes the $L^2(\Omega)$ inner product. We shall also use Sobolev spaces of vector valued functions. A vector function \mathbf{w} is in $(H^s(\Omega))^d$ if each of its components is in $H^s(\Omega)$. The norm in $(H^s(\Omega))^d$ will also be denoted by $\|\cdot\|_s$. There is no ambiguity with this definition since the specific norm used will be uniquely identified by the type of function on which it is applied. When $s = 0$, the norm will be denoted by $\|\cdot\|$ in both the vector and scalar case.

We next consider the weak formulation of (2.6)–(2.9). Let $H_0^1(\Omega)$ be the completion of $C_0^\infty(\Omega)$ in the norm $\|\cdot\|_1$. Define $\mathbf{V} \equiv (H_0^1(\Omega))^d$, i.e., the space of vector

valued functions with each component in $H_0^1(\Omega)$. Finally, let W denote the functions in $L^2(\Omega)$ with zero mean value on Ω . Multiplying (2.6) and (2.7) by functions in \mathbf{V} and W respectively, it is easy to see that the solution (\mathbf{v}, Q) of (2.6)–(2.9) satisfies

$$(2.10) \quad \begin{aligned} D(\mathbf{v}, \mathbf{w}) + (Q, \nabla \cdot \mathbf{w}) &= (\mathbf{g}, \mathbf{w}) & \text{for all } \mathbf{w} \in \mathbf{V}, \\ (\nabla \cdot \mathbf{v}, q) &= (f, q) & \text{for all } q \in W. \end{aligned}$$

Here, $D(\cdot, \cdot)$ denotes the vector Dirichlet form. We shall let $D(\cdot, \cdot)$ denote both the vector and scalar Dirichlet forms on Ω . For scalar functions in $v, w \in H^1(\Omega)$, $D(\cdot, \cdot)$ is defined by

$$D(v, w) \equiv \int_{\Omega} \mathbf{grad} v \cdot \mathbf{grad} w \, dx.$$

For vector functions $\mathbf{w}, \tilde{\mathbf{w}}$,

$$D(\mathbf{w}, \tilde{\mathbf{w}}) \equiv \sum_{i=1}^d D(w_i, \tilde{w}_i).$$

Since Ω is bounded and connected, it follows from [12] that there is a constant C_0 satisfying

$$(2.11) \quad \|p\| \leq C_0 \sup_{\mathbf{w} \in \mathbf{V}} \frac{(p, \nabla \cdot \mathbf{w})}{\|\mathbf{w}\|_1}.$$

Here and in the remainder of the paper, C (with or without subscript) will denote a generic positive constant. These constants may take on different values in different occurrences, however, they will always be independent of mesh and time step parameters. It easily follows from (2.11) that if $\mathbf{g} = 0$ then the solution \mathbf{v} satisfies

$$(2.12) \quad \|\mathbf{v}\|_1 \leq C \|f\|.$$

In addition, we assume that the solutions of (2.6)–(2.9) satisfy regularity estimates of the form: For $f = 0$ and all $\mathbf{g} \in L^2(\Omega)^d$,

$$(2.13) \quad \|\mathbf{v}\|_2 + \|Q\|_1 \leq C \|\mathbf{g}\|.$$

This imposes some conditions on the domain. It is known that (2.13) holds in the case of convex domains in R^2 with polygonal boundaries [15] and for convex polyhedral domains in R^3 [11]. Using a duality technique and (2.13), it is shown in the appendix that solutions to (2.6)–(2.9) with $\mathbf{g} = 0$ satisfy

$$(2.14) \quad \|\mathbf{v}\| \leq C \|f\|_{-1}.$$

To approximately solve (2.10), we introduce a collection of pairs of approximation subspaces $\mathbf{V}_h \subset \mathbf{V}$ and $\bar{W}_h \subset L^2(\Omega)$ indexed by h in the interval $0 < h < 1$. We assume that the constant function is in \bar{W}_h and define W_h to be the subspace of functions in \bar{W}_h with zero mean value. We will assume that the classical L-B-B

(Ladyzhenskaya-Babuška-Brezzi) condition (cf. [12]) holds for the pair of spaces; i.e. there is a constant c_0 which does not depend upon h such that

$$(2.15) \quad \sup_{\mathbf{v} \in \mathbf{V}_h} \frac{(p, \nabla \cdot \mathbf{v})^2}{D(\mathbf{v}, \mathbf{v})} \geq c_0 \|p\|^2 \quad \text{for all } p \in W_h.$$

In addition, the subspaces are assumed to satisfy the following approximation and inverse properties.

(1) For $\mathbf{v} \in (H^2(\Omega) \cap H_0^1(\Omega))^2$,

$$(2.16) \quad \inf_{\mathbf{w} \in \mathbf{V}_h} \|\mathbf{v} - \mathbf{w}\|_1 \leq Ch \|\mathbf{v}\|_2.$$

(2) For $v \in H^1(\Omega)$,

$$(2.17) \quad \inf_{w \in \bar{W}_h} \|v - w\| \leq Ch \|v\|_1.$$

(3) For $\mathbf{v} \in \mathbf{V}_h$ and $s = 0, 1$,

$$(2.18) \quad \|\mathbf{v}\|_s \leq Ch^{-1} \|\mathbf{v}\|_{s-1}.$$

(4) For $v \in W_h$,

$$(2.19) \quad \|v\| \leq Ch^{-1} \|v\|_{-1}.$$

The constant C appearing above is independent of the approximation parameter h . Many subspace pairs satisfying (2.15)–(2.19) have been studied (see, e.g., [12],[19],[22]). A simple collection of (\mathbf{V}_h, W_h) pairs is given in the following example.

Example 1. Let Ω be the unit square $(0,1)^2$ in R^2 . We first break Ω into smaller squares with edge length of $h = 1/n$ for even integer values of n . The domain is further subdivided by breaking each smaller square into two triangles by connecting the lower left hand corner to the upper right. The subspace \mathbf{V}_h is defined to be the set of vector valued functions which are piecewise linear with respect to the above triangulation, continuous on Ω and vanish on $\partial\Omega$. We will use a pressure space of piecewise constant functions. To satisfy (2.15), this space must be taken with respect to a somewhat coarser mesh. To this end, we consider breaking up Ω into squares of size $2/n$. Let \bar{W}_h denote the set of functions which are piecewise constant on the squares of edge length $2/n$. The space W_h is then defined to be the functions in \bar{W}_h with zero mean value on Ω . The pair (\mathbf{V}_h, W_h) satisfies (2.15) with c_0 independent of $h = 1/n$ (see, [12]).

The approximation to the solution (\mathbf{v}, Q) of (2.10) is defined by replacing the spaces in (2.10) by their discrete counterparts. Specifically, the approximate functions are defined as the unique elements $\mathbf{v}_h \in \mathbf{V}_h$ and $Q_h \in W_h$ satisfying

$$(2.20) \quad \begin{aligned} D(\mathbf{v}_h, \mathbf{w}) + (Q_h, \nabla \cdot \mathbf{w}) &= (\mathbf{g}, \mathbf{w}) & \text{for all } \mathbf{w} \in \mathbf{V}_h, \\ (\nabla \cdot \mathbf{v}_h, q) &= (f, q) & \text{for all } q \in W_h. \end{aligned}$$

The unique solvability of (2.20) is a consequence of (2.15). The following lemma provides error estimates for the Stokes approximation.

Lemma 2.1. *Let (\mathbf{v}, Q) solve (2.10) and (\mathbf{v}_h, Q_h) solve (2.20). Assume that (2.15)–(2.17) hold. Then*

$$(2.21) \quad \|\mathbf{v}_h - \mathbf{v}\|_1 + \|Q_h - Q\| \leq Ch(\|\mathbf{v}\|_2 + \|Q\|_1).$$

Suppose that (2.13) holds for all $\mathbf{g} \in (L^2(\Omega))^d$. Let \mathbf{v} and \mathbf{v}_h respectively solve (2.10) and (2.20) with $\mathbf{g} = 0$. Then,

$$(2.22) \quad \|\mathbf{v}_h - \mathbf{v}\| \leq Ch\|\mathbf{v}\|_1.$$

The proof of the first inequality (2.21) is well known and can be found in [12]. For completeness, we provide a proof of (2.22) in the appendix.

Fully discrete time-stepping approximations to (2.1)–(2.5) using the above spaces lead to systems of equations of the form

$$(2.23) \quad \begin{aligned} D_k(\mathbf{u}_h, \mathbf{w}) + (P_h, \nabla \cdot \mathbf{w}) &= (\mathbf{f}, \mathbf{w}) && \text{for all } \mathbf{w} \in \mathbf{V}_h, \\ (\nabla \cdot \mathbf{u}_h, q) &= 0 && \text{for all } q \in W_h. \end{aligned}$$

Here

$$(2.24) \quad D_k(\mathbf{v}, \mathbf{w}) = (\mathbf{v}, \mathbf{w}) + kD(\mathbf{v}, \mathbf{w}) \quad \text{for all } \mathbf{v}, \mathbf{w} \in \mathbf{V},$$

and k is a positive number which is related to the time step size.

The above problem can be formulated in terms of operators as follows: Let $\mathbf{A}_k : \mathbf{V}_h \mapsto \mathbf{V}_h$, $\mathbf{B} : \mathbf{V}_h \mapsto W_h$ and $\mathbf{B}^* : W_h \mapsto \mathbf{V}_h$ be defined by

$$(2.25) \quad \begin{aligned} (\mathbf{A}_k \mathbf{v}, \mathbf{w}) &= D_k(\mathbf{v}, \mathbf{w}), && \text{for all } \mathbf{v}, \mathbf{w} \in \mathbf{V}_h, \\ (\mathbf{B} \mathbf{v}, w) &= (\nabla \cdot \mathbf{v}, w), && \text{for all } \mathbf{v} \in \mathbf{V}_h, w \in W_h, \\ (\mathbf{B}^* w, \mathbf{v}) &= (w, \nabla \cdot \mathbf{v}), && \text{for all } \mathbf{v} \in \mathbf{V}_h, w \in W_h. \end{aligned}$$

Note that \mathbf{B}^* is the adjoint of \mathbf{B} . Moreover, (2.23) can be rewritten as

$$(2.26) \quad \begin{pmatrix} \mathbf{A}_k & \mathbf{B}^* \\ \mathbf{B} & 0 \end{pmatrix} \begin{pmatrix} \mathbf{u}_h \\ P_h \end{pmatrix} = \begin{pmatrix} \mathbf{f}_h \\ 0 \end{pmatrix},$$

where \mathbf{f}_h denotes the $(L^2(\Omega))^d$ orthogonal projection of \mathbf{f} into \mathbf{V}_h .

3. ITERATIVE METHODS FOR SYSTEMS OF THE FORM OF (2.26).

In this section, we present some iterative methods for block systems of the form of (2.26). All of these methods involve the introduction of a preconditioner for a reduced system on W_h . Three of the methods involve the use of an additional preconditioner for the operator \mathbf{A}_k on \mathbf{V}_h . In this section, we describe these iterative techniques and discuss how their rates of convergence depend upon the condition numbers of related preconditioned sub-systems. The techniques discussed in this

section are not restricted to the specific system (2.26) but rather, they are applicable to general block systems of the form of (2.26). Block systems of this form arise in many other applications. For example, such systems must be solved for finite element Lagrange multiplier approximations to Dirichlet and interface problems [2],[3], velocity-pressure formulations of the equations of Stokes and elasticity [12] and mixed finite element methods [20].

We start by considering generic block operator equations of which (2.26) is an example. Let H^1 and H^2 be finite dimensional Hilbert spaces and consider the problem

$$(3.1) \quad M \begin{pmatrix} X \\ Y \end{pmatrix} = \begin{pmatrix} F \\ G \end{pmatrix}$$

where X, F are in H^1 and Y, G are in H^2 . We study operators M of the form,

$$(3.2) \quad M = \begin{pmatrix} A & B^* \\ B & 0 \end{pmatrix}.$$

We assume that A is a positive definite, symmetric operator on H^1 . In addition, we assume that B and B^* are adjoints with respect to the inner products in H^1 and H^2 . We shall use the notation (\cdot, \cdot) and $\|\cdot\|$ to denote the inner products and norms on H^1 and H^2 . This can be done without ambiguity since the particular inner product and norm can be identified by the type of function on which the inner product or norm operates.

Applying block Gaussian elimination to (3.1) shows that the solution of (3.1) satisfies

$$(3.3) \quad \begin{pmatrix} A & B^* \\ 0 & BA^{-1}B^* \end{pmatrix} \begin{pmatrix} X \\ Y \end{pmatrix} = \begin{pmatrix} F \\ BA^{-1}F - G \end{pmatrix}.$$

Thus, (3.1) is nonsingular if and only if $BA^{-1}B^*$ is invertible. Clearly, $BA^{-1}B^*$ is symmetric and non-negative. A straightforward computation gives that

$$(3.4) \quad (BA^{-1}B^*U, U) = \sup_{\theta \in H^1} \frac{(U, B\theta)^2}{(A\theta, \theta)} \quad \text{for all } U \in H^2,$$

and hence solvability of (3.1) will follow if

$$(3.5) \quad \sup_{\Theta \in H^1} \frac{(U, B\Theta)^2}{(A\Theta, \Theta)} \geq c \|U\|^2 \quad \text{for all } U \in H^2,$$

holds for some positive number c . Inequality (3.5) is the classical L-B-B (Ladyzhenskaya-Babuška-Brezzi) condition (cf. [12]).

The first scheme which we shall consider for solving (3.1) involves iteratively solving the equation

$$(3.6) \quad BA^{-1}B^*Y = BA^{-1}F - G$$

and subsequently back solving (3.3) for X , i.e. $X = A^{-1}(F - B^*Y)$. This is a classical technique but has two potential drawbacks. First, the operator $BA^{-1}B^*$ may be ill-conditioned and hence the iteration for (3.6) may converge slowly without preconditioning. Second, the action of the inverse of the operator A must be computed at each step of the iteration. This latter drawback is perhaps the more serious one and for this reason we will not focus our attention on this method.

As we will see, it will be essential for the efficiency of all of the other methods discussed here to be able to construct a good preconditioner for $BA^{-1}B^*$. For this purpose let \mathcal{K} be a symmetric positive definite operator on H^2 and let κ be the condition number of $\mathcal{K}BA^{-1}B^*$. Since $\mathcal{K}BA^{-1}B^*$ is symmetric with respect to the inner product defined by $(\mathcal{K}^{-1}\cdot, \cdot)$, it follows easily that $\kappa \leq c_1/c_0$ for any pair c_0 and c_1 of positive numbers satisfying

$$(3.7) \quad c_0(\mathcal{K}V, V) \leq (BA^{-1}B^*\mathcal{K}V, \mathcal{K}V) \leq c_1(\mathcal{K}V, V) \quad \text{for all } V \in H^2.$$

The construction of such operators \mathcal{K} , with κ not too large, will be important for the performance of the methods which we will now describe.

Remark 3.1. This paper is concerned with the solution of the systems resulting from time stepping procedures for the linearized Navier–Stokes problem. In the next section, we shall show how to construct efficient preconditioners for the particular operator $\mathbf{B}\mathbf{A}_k^{-1}\mathbf{B}^*$ corresponding to (2.26).

As already observed, one problem with the iterative technique just described is that it requires the evaluation of the action of A^{-1} at each step in the iteration. In the application considered in this paper, the action of $A^{-1} = \mathbf{A}_k^{-1}$ is more expensive to compute than that of a suitable preconditioner. We next consider a natural preconditioned conjugate gradient technique for solving (3.1) which does not require the evaluation of the action of A^{-1} . To this end, we assume that we are given a symmetric positive definite operator \mathcal{J} which acts as a preconditioner for A . This means that there are positive numbers c_2 and c_3 satisfying

$$(3.8) \quad c_2(\mathcal{J}V, V) \leq (A\mathcal{J}V, \mathcal{J}V) \leq c_3(\mathcal{J}V, V) \quad \text{for all } V \in H^1,$$

with c_3/c_2 not too large.

The second method which we present here is as follows. We will precondition M by the block operator M_0 defined by

$$M_0 = \begin{pmatrix} \mathcal{J} & 0 \\ 0 & \mathcal{K} \end{pmatrix}$$

and then form the normal equations corresponding to (3.1). That is, we write the equivalent system

$$(3.9) \quad M_0 M M_0 M \begin{pmatrix} X \\ Y \end{pmatrix} = M_0 M M_0 \begin{pmatrix} F \\ G \end{pmatrix}.$$

Let (\cdot, \cdot) denote the sum of the componentwise inner products on $H^1 \times H^2$. Note that $M_0 M M_0 M$ is a symmetric operator with respect to the inner product $(M_0^{-1}\cdot, \cdot)$.

Since it is also positive definite we can apply the conjugate gradient method (in the inner product $(M_0^{-1}\cdot, \cdot)$) to solve (3.9). Note that the asymptotic rate of convergence per step of the conjugate gradient method can be bounded by

$$(3.10) \quad \rho = \frac{\sqrt{\kappa} - 1}{\sqrt{\kappa} + 1}$$

where κ is the condition number of M_0MM_0M . It is essentially shown in [6] that

$$(3.11) \quad \kappa(M_0MM_0M) \leq \left(\frac{3 + \sqrt{5}}{3 - \sqrt{5}}\right)^2 \frac{\alpha_1^2}{\alpha_0^2},$$

where $\alpha_0 = \min(c_0, c_2)$ and $\alpha_1 = \max(c_1, c_3)$. We omit the proof here.

The importance of an estimate of the type (3.11) is that it shows that the convergence rate bound (combining (3.11) and (3.10)) is improved when good preconditioners \mathcal{J} and \mathcal{K} are used.

Remark 3.2. Even though we are applying the conjugate gradient method in the inner product $(M_0^{-1}\cdot, \cdot)$, the algorithm can be implemented in such a way as to avoid the explicit evaluation of M_0^{-1} . This is because of the special form of the equations being solved. Every instance of the inner product which appears in the conjugate gradient algorithm involves variables, e.g., $(M_0^{-1}x_1, x_2)$, where $x_i = M_0w_i$ for $i = 1, 2$ with either w_1 or w_2 known. Thus, M_0^{-1} can be avoided in the implementation. This is important in that there are many preconditioners whose evaluation is implemented as a process. Efficient computational procedures for computing the action of the inverse of such processes may not be known or available.

A third method, which we describe now was studied by Rustin and Winther [21]. It is related to the above ‘‘normal equations method’’ but seems to perform quite a bit better when used in the application studied in this paper (see Section 5).

This time we write (3.1) as

$$(3.12) \quad M_0M \begin{pmatrix} X \\ Y \end{pmatrix} = M_0 \begin{pmatrix} F \\ G \end{pmatrix}.$$

and note that M_0M is a symmetric operator with respect to the inner product $(M_0^{-1}\cdot, \cdot)$. Even though it is not positive definite we can apply the minimal residual method to (3.12). This is computationally no more involved than the conjugate gradient method in that it also involves only a three term recurrence relation for its implementation. This method can be characterized as follows. Let E_i be the error after i steps of the iteration and define the Krylov space

$$\mathbf{K}_i = \text{span}_{l=1}^i \{(M_0M)^l E_0\}.$$

Then, $E_i = E_0 + \chi$ for the unique $\chi \in \mathbf{K}_i$ which minimizes

$$(M_0M(E_0 + \theta), M(E_0 + \theta)) \quad \text{for all } \theta \in \mathbf{K}_i.$$

$2i$ steps of this method is necessarily no worse than i steps of the normal equation method since the error after i steps of the normal equation method satisfies the above minimization but with \mathbf{K}_i replaced by

$$\mathbf{K}'_i = \underset{l=1}{\text{span}}^i \{(M_0 M)^{2l} E_0\}.$$

Thus the rate of convergence for the minimal residual method applied to (3.12) can be bounded by the normal method rate (see, (3.10) and (3.11)). Alternatively, bounds for the rate of convergence of the minimal residual method may be inferred from the $M_0 M$ eigenvalue estimates given in [21]. Note that the computational cost of two steps of the minimal residual method is approximately that of one step of the normal equation method.

Remark 3.3. In both of the preceding methods, spectral bounds may not really predict the convergence behavior. We consider the case $\mathcal{J}^{-1} = A$ and $\mathcal{K}^{-1} = BA^{-1}B^*$. Clearly,

$$(3.13) \quad M_0 M = \begin{pmatrix} I & A^{-1}B^* \\ \mathcal{K}B & 0 \end{pmatrix}.$$

It is straightforward to check that if the null space of B is nonempty, then (3.13) has exactly the three eigenvalues 1 , $(1 + \sqrt{5})/2$ and $(1 - \sqrt{5})/2$. It follows that both the minimal residual method and the normal method converge in three iterations which is considerably better than the rate predicted by the spectral estimates. For preconditioners close to these limiting cases, it is clear that the eigenvalues must cluster near the three values 1 , $(1 + \sqrt{5})/2$ and $(1 - \sqrt{5})/2$.

Remark 3.4. The minimal residual method applied to a nonsingular symmetric operator N is, in general, no better than the conjugate gradient method applied to the positive definite symmetric operator N^2 . An example which illustrates this is as follows. Suppose that N has a spectrum which is symmetric about 0 and that each eigenvalue is simple. Any vector x has an eigenvector expansion of the form

$$x = \sum_{\lambda_i > 0} (c_i \phi_i + c_{-i} \phi_{-i}).$$

Here $\{\lambda_i, \phi_i\}$ and $\{-\lambda_i, \phi_{-i}\}$ are the eigenpairs for the operator N . We say that a vector x is even (respectively, odd) if $c_i = c_{-i}$ (respectively, $c_i = -c_{-i}$). Consider solving the problem $Nx = y$ with an initial iterate x_0 chosen such that the initial error $e_0 = x - x_0$ is even. It is easy to see that $N^k e_0$ is even for k even and odd for k odd. Thus the terms with odd powers of the operator in \mathbf{K}_i are already orthogonal to the error e_0 . It follows that the minimization over \mathbf{K}_i is identical to the minimization over \mathbf{K}'_i , i.e., $2l$ steps of minimal residual method produce the same result as l steps of the normal equation iteration. However, this is not the situation in our application where the minimal residual method appears to perform significantly better than the normal equation method (see the numerical examples in Section 5).

The last method which we will consider involves a somewhat different reformulation of (3.1). Assume that c_2 in (3.8) is greater than one. Let A_0 be defined to be \mathcal{J}^{-1} . Clearly, $A - A_0$ satisfies

$$(3.14) \quad 0 < (1 - 1/c_2)(AV, V) \leq ((A - A_0)V, V) \leq (1 - 1/c_3)(AV, V)$$

for all nonzero V in H^1 .

Using straightforward manipulations we obtain the system

$$(3.15) \quad \tilde{M} \begin{pmatrix} X \\ Y \end{pmatrix} \equiv \begin{pmatrix} A_0^{-1}A & A_0^{-1}B^* \\ BA_0^{-1}(A - A_0) & BA_0^{-1}B^* \end{pmatrix} \begin{pmatrix} X \\ Y \end{pmatrix} = \begin{pmatrix} A_0^{-1}F \\ BA_0^{-1}F - G \end{pmatrix}.$$

As a consequence of (3.14) we can define an inner product on $H^1 \times H^2$ by

$$(3.16) \quad \left[\begin{pmatrix} U \\ V \end{pmatrix}, \begin{pmatrix} X \\ Y \end{pmatrix} \right] \equiv (AU, X) - (A_0U, X) + (V, Y).$$

We note that

$$\begin{aligned} \left[\tilde{M} \begin{pmatrix} U \\ V \end{pmatrix}, \begin{pmatrix} X \\ Y \end{pmatrix} \right] &= ((AA_0^{-1}A - A)U, X) + ((A - A_0)A_0^{-1}B^*V, X) \\ &\quad + (BA_0^{-1}(A - A_0)U, Y) + (BA_0^{-1}B^*V, Y) \end{aligned}$$

and hence the operator \tilde{M} is symmetric with respect to the inner product defined by (3.16). It was shown in [5] that \tilde{M} is also positive definite. Let

$$\tilde{M}_0 \equiv \begin{pmatrix} I & 0 \\ 0 & \mathcal{K} \end{pmatrix}.$$

It was also shown in [5] that the condition number of $\tilde{M}_0\tilde{M}$ is uniformly bounded (a bound was given in terms of c_0 , c_1 and c_3). It follows that the conjugate gradient method for solving (3.15) preconditioned by \tilde{M}_0 converges rapidly. In fact, as A_0 tends to A and \mathcal{K}^{-1} tends to $BA^{-1}B^*$, the condition number for the preconditioned system tends to one.

4. THE PRECONDITIONER AND ITS ANALYSIS.

We develop the preconditioner for the perturbed pressure system in this section. It should be noted that, in contrast to the case of stationary Stokes, the pressure operator $\mathbf{B}\mathbf{A}_k^{-1}\mathbf{B}^*$ corresponding to the time stepping equation (2.26) becomes ill conditioned as k becomes small. In this section, we will develop preconditioners for $\mathbf{B}\mathbf{A}_k^{-1}\mathbf{B}^*$ which lead to preconditioned systems with condition number bounds which are independent of k .

To define the preconditioner \mathcal{K} , we use a conforming scalar valued finite element space $\bar{S}_h \subset H^1(\Omega)$. We assume that \bar{S}_h satisfies the standard approximation property: For each $v \in H^2(\Omega)$,

$$(4.1) \quad \inf_{\theta \in \bar{S}_h} \|v - \theta\|_1 \leq Ch \|v\|_2.$$

Finally, we assume that the constant function is in \bar{S}_h and define S_h to be the set of functions in \bar{S}_h which have zero mean value on Ω . Many examples of the construction of subspaces satisfying the above hypotheses can be found in [1], [9], [10].

Let T denote the solution operator for the continuous Neumann problem, i.e., given $f \in H^{-1}(\Omega)$, $w = Tf$ is the unique function (with zero mean value on Ω) in $H^1(\Omega)$ satisfying

$$(4.2) \quad D(w, \phi) = (f, \phi)$$

for all $\phi \in H^1(\Omega)/R$, the functions in $H^1(\Omega)$ with zero mean value on Ω . The unique solvability of (4.2) follows from the Poincaré inequality. We shall assume full elliptic regularity for solutions of (4.2), i.e.,

$$(4.3) \quad \|w\|_2 \leq C \|f\|.$$

Note that (4.3) holds for convex domains in R^2 and R^3 [13]. Let T_h be the solution operator for the Galerkin approximation to the second order Neumann problem. Specifically, for a function $f \in L^2(\Omega)$, $w = T_h f$ is the unique function in S_h satisfying

$$(4.4) \quad D(w, \theta) = (f, \theta) \quad \text{for all } \theta \in S_h.$$

Note that the functions in S_h do not satisfy boundary conditions but have zero mean value. Under the assumptions (4.1) and (4.3), it is well known that

$$(4.5) \quad \|(T - T_h)v\| \leq Ch^2 \|v\|.$$

Let Q_h denote the $L^2(\Omega)$ orthogonal projection onto the subspace W_h . The major result of this paper is given in the following theorem and its corollary.

Theorem 4.1. *Assume that (2.13), (2.15)–(2.19), (4.1) and (4.3) hold and that $k \geq h^2$. Let γ be in the interval $[h, k^{1/2}]$ and define the operator $\mathcal{A}_\gamma^k : W_h \mapsto W_h$ by*

$$\mathcal{A}_\gamma^k = Q_h T_\gamma + kI.$$

Let $\mathbf{BA}_k^{-1}\mathbf{B}^$ be defined by (2.25). Then there are constants C_0 and C_1 not depending on h or k such that*

$$(4.6) \quad C_0(\mathcal{A}_\gamma^k q, q) \leq (\mathbf{BA}_k^{-1}\mathbf{B}^* \mathcal{A}_\gamma^k q, \mathcal{A}_\gamma^k q) \leq C_1(\mathcal{A}_\gamma^k q, q) \quad \text{for all } q \in W_h.$$

Corollary 4.1. *Assume that (2.13), (2.15)–(2.19), (4.1) and (4.3) hold and that $k \leq h^2$. Then there are constants C_0 and C_1 not depending on h or k such that*

$$C_0(\mathcal{A}_h^{h^2} q, q) \leq (\mathbf{BA}_k^{-1}\mathbf{B}^* \mathcal{A}_h^{h^2} q, \mathcal{A}_h^{h^2} q) \leq C_1(\mathcal{A}_h^{h^2} q, q) \quad \text{for all } q \in W_h.$$

Remark 4.1. The above theorem shows that one can use $\mathcal{K} = \mathcal{A}_\gamma^k$ as a preconditioner for $\mathbf{BA}_k^{-1}\mathbf{B}^*$. Each evaluation of \mathcal{K} then requires solving a discrete Neumann

problem (4.4) on a mesh of size γ . If γ is relatively small, it may be more economical to replace the operator T_γ by a preconditioner \tilde{T}_γ . Such an operator is constructed so that

$$C_2(T_\gamma\theta, \theta) \leq (\tilde{T}_\gamma\theta, \theta) \leq C_3(T_\gamma\theta, \theta) \quad \text{for all } \theta \in S_\gamma,$$

with C_2 and C_3 not depending on h . Suppose that (4.6) holds. If \tilde{T}_γ is scaled so that $C_2 \leq 1 \leq C_3$ then $\mathcal{K} = Q_h\tilde{T}_\gamma + kI$ satisfies

$$C_0C_2(\mathcal{K}q, q) \leq (\mathbf{B}\mathbf{A}_k^{-1}\mathbf{B}^*\mathcal{K}q, \mathcal{K}q) \leq C_1C_3(\mathcal{K}q, q) \quad \text{for all } q \in W_h.$$

This remark holds for $\gamma = h$ which can be applied in the case of the corollary.

Proof of Theorem 4.1. Let $\mathbf{M}_k = \mathbf{B}\mathbf{A}_k^{-1}\mathbf{B}^*$. It is easy to see that the upper inequality in (4.6) is equivalent to

$$(4.7) \quad (\mathcal{A}_\gamma^k\mathbf{M}_kq, \mathbf{M}_kq) \leq C_1(\mathbf{M}_kq, q) \quad \text{for all } q \in W_h.$$

Let q be in W_h . Using (4.5), we have that

$$(4.8) \quad \begin{aligned} (\mathcal{A}_\gamma^k\mathbf{M}_kq, \mathbf{M}_kq) &= ((T_\gamma - T)\mathbf{M}_kq, \mathbf{M}_kq) + (T\mathbf{M}_kq, \mathbf{M}_kq) + k\|\mathbf{M}_kq\|^2 \\ &\leq (T\mathbf{M}_kq, \mathbf{M}_kq) + Ck\|\mathbf{M}_kq\|^2. \end{aligned}$$

As in (3.4), a straightforward computation gives that

$$(4.9) \quad (\mathbf{M}_kq, q) = \sup_{\mathbf{v} \in \mathbf{V}_h} \frac{(q, \mathbf{B}\mathbf{v})^2}{(\mathbf{A}_k\mathbf{v}, \mathbf{v})} = \sup_{\mathbf{v} \in \mathbf{V}_h} \frac{(q, \nabla \cdot \mathbf{v})^2}{D_k(\mathbf{v}, \mathbf{v})}.$$

We first bound the last term in (4.8). We clearly have that

$$\|\nabla \cdot \mathbf{v}\|^2 \leq D(\mathbf{v}, \mathbf{v})$$

and hence (4.9) and the Schwarz inequality imply that

$$(4.10) \quad k(\mathbf{M}_kq, q) \leq \|q\|^2.$$

Using the fact that \mathbf{M}_k is a symmetric positive definite operator, it follows that

$$(4.11) \quad k\|\mathbf{M}_kq\|^2 \leq (\mathbf{M}_kq, q).$$

We note that for any w in $H^{-1}(\Omega)$,

$$(4.12) \quad \begin{aligned} (Tw, w) &= D(Tw, Tw) = \sup_{\phi \in H^1(\Omega)/R} \frac{D(Tw, \phi)^2}{D(\phi, \phi)} \\ &= \sup_{\phi \in H^1(\Omega)/R} \frac{(w, \phi)^2}{D(\phi, \phi)} \leq C\|w\|_{-1}^2. \end{aligned}$$

We used the Poincaré inequality for the last inequality above. If, in addition, w has zero mean value on Ω , then

$$(4.13) \quad \|w\|_{-1}^2 = \sup_{\phi \in H^1(\Omega)/R} \frac{(w, \phi)^2}{\|\phi\|_1^2} \leq (Tw, w).$$

We used the representation for (Tw, w) given in (4.12) for the last inequality above.

We next bound the first term on the right hand side of (4.8). By (4.12),

$$(T\mathbf{M}_k q, \mathbf{M}_k q) \leq C \|\mathbf{M}_k q\|_{-1}^2.$$

Note that $\mathbf{M}_k q = \mathcal{Q}_h \nabla \cdot \mathbf{w}$ where $\mathbf{w} = \mathbf{A}_k^{-1} \mathbf{B}^* q$. Consequently,

$$(4.14) \quad (T\mathbf{M}_k q, \mathbf{M}_k q) \leq 2C (\|(I - \mathcal{Q}_h) \nabla \cdot \mathbf{w}\|_{-1}^2 + \|\nabla \cdot \mathbf{w}\|_{-1}^2).$$

It follows from (2.17) and a straightforward duality argument that

$$\|(I - \mathcal{Q}_h) \nabla \cdot \mathbf{w}\|_{-1} \leq Ch \|\nabla \cdot \mathbf{w}\|.$$

Thus, by the inverse inequality (2.18),

$$(4.15) \quad \begin{aligned} (T\mathbf{M}_k q, \mathbf{M}_k q) &\leq C(h^2 \|\nabla \cdot \mathbf{w}\|^2 + \|\mathbf{w}\|^2) \\ &\leq C \|\mathbf{w}\|^2 = C \|\mathbf{A}_k^{-1} \mathbf{B}^* q\|^2 \\ &\leq C(\mathbf{A}_k^{-1} \mathbf{B}^* q, \mathbf{B}^* q) = C(\mathbf{M}_k q, q). \end{aligned}$$

We used the fact that the spectrum of \mathbf{A}_k^{-1} is in the interval $(0, 1)$ for the last inequality above. Combining (4.8), (4.11) and (4.15) completes the proof of (4.7).

We now prove the lower inequality of (4.6). Evidently this is the same as

$$(4.16) \quad C_0((\mathcal{A}_\gamma^k)^{-1} q, q) \leq (\mathbf{M}_k q, q) \quad \text{for all } q \in W_h.$$

Let q be an arbitrary function in W_h . We prove (4.16) by constructing a function $\mathbf{v}_h \in \mathbf{V}_h$ such that $\mathbf{B} \mathbf{v}_h = (\mathcal{A}_\gamma^k)^{-1} q$ and

$$(4.17) \quad D_k(\mathbf{v}_h, \mathbf{v}_h) \leq C((\mathcal{A}_\gamma^k)^{-1} q, q).$$

If we have such a function, then by (4.9) and (4.17),

$$(\mathbf{M}_k q, q) \geq \frac{(q, \mathbf{B} \mathbf{v}_h)^2}{D_k(\mathbf{v}_h, \mathbf{v}_h)} \geq C^{-1}((\mathcal{A}_\gamma^k)^{-1} q, q).$$

We start by considering the pair (\mathbf{v}, Q) solving the Stokes equation:

$$\begin{aligned} -\Delta \mathbf{v} - \nabla Q &= 0 && \text{in } \Omega, \\ \nabla \cdot \mathbf{v} &= (\mathcal{A}_\gamma^k)^{-1} q && \text{in } \Omega, \\ \mathbf{v} &= 0 && \text{on } \partial\Omega. \end{aligned}$$

By (2.12),

$$(4.18) \quad \|\mathbf{v}\|_1 \leq C \|(\mathcal{A}_\gamma^k)^{-1} q\|$$

and by (2.14),

$$(4.19) \quad \|\mathbf{v}\| \leq C \|(\mathcal{A}_\gamma^k)^{-1} q\|_{-1}.$$

Combining the above two inequalities and using (4.13) gives

$$(4.20) \quad D_k(\mathbf{v}, \mathbf{v}) \leq C((T + kI)(\mathcal{A}_\gamma^k)^{-1} q, (\mathcal{A}_\gamma^k)^{-1} q).$$

Let p be in W_h . By (4.5),

$$(4.21) \quad |((T - T_\gamma)p, p)| \leq C\gamma^2 \|p\|^2 \leq Ck \|p\|^2.$$

Thus, it follows from the triangle inequality that

$$(4.22) \quad ((T + kI)p, p) \leq C((T_\gamma + kI)p, p).$$

Combining (4.22) and (4.20) gives

$$(4.23) \quad D_k(\mathbf{v}, \mathbf{v}) \leq C((T_\gamma + kI)(\mathcal{A}_\gamma^k)^{-1} q, (\mathcal{A}_\gamma^k)^{-1} q) = C((\mathcal{A}_\gamma^k)^{-1} q, q).$$

We define \mathbf{v}_h to be the mixed approximation to \mathbf{v} . Specifically, \mathbf{v}_h satisfies

$$\begin{aligned} D(\mathbf{v}_h, \mathbf{w}) + (Q_h, \nabla \cdot \mathbf{w}) &= 0 \quad \text{for all } \mathbf{w} \in \mathbf{V}_h, \\ (\nabla \cdot \mathbf{v}_h, s) &= ((\mathcal{A}_\gamma^k)^{-1} q, s) \quad \text{for all } s \in W_h. \end{aligned}$$

Note that $\mathbf{B}\mathbf{v}_h = (\mathcal{A}_\gamma^k)^{-1} q$ follows from the second equation above. The proof of the theorem will be complete once we verify (4.17) for this choice of \mathbf{v}_h . Clearly,

$$(4.24) \quad \begin{aligned} D(\mathbf{v}_h, \mathbf{v}_h) &= -(Q_h, \nabla \cdot \mathbf{v}_h) = -(Q_h, \nabla \cdot \mathbf{v}) \\ &\leq \|Q_h\| \|\nabla \cdot \mathbf{v}\| \leq \|Q_h\| D(\mathbf{v}, \mathbf{v})^{1/2}. \end{aligned}$$

Now, by (2.15)

$$(4.25) \quad \begin{aligned} \|Q_h\| &\leq C \sup_{\mathbf{w} \in \mathbf{V}_h} \frac{(Q_h, \nabla \cdot \mathbf{w})}{D(\mathbf{w}, \mathbf{w})^{1/2}} \\ &= C \sup_{\mathbf{v} \in \mathbf{V}_h} \frac{D(\mathbf{v}_h, \mathbf{w})}{D(\mathbf{w}, \mathbf{w})^{1/2}} \leq C D(\mathbf{v}_h, \mathbf{v}_h)^{1/2}. \end{aligned}$$

Combining (4.23), (4.24) and (4.25) gives

$$(4.26) \quad kD(\mathbf{v}_h, \mathbf{v}_h) \leq CkD(\mathbf{v}, \mathbf{v}) \leq C((\mathcal{A}_\gamma^k)^{-1} q, q).$$

To verify (4.17), we need only estimate $\|\mathbf{v}_h\|$. By Lemma 2.1 and (4.18),

$$\|\mathbf{v} - \mathbf{v}_h\| \leq Ch \|\mathbf{v}\|_1 \leq Ch \|(\mathcal{A}_\gamma^k)^{-1} q\|.$$

Thus, by the inverse inequality (2.19), (4.13) and the triangle inequality,

$$\|\mathbf{v}_h\| \leq \|\mathbf{v}\| + C(T(\mathcal{A}_\gamma^k)^{-1} q, (\mathcal{A}_\gamma^k)^{-1} q)^{1/2}.$$

Applying (4.22) gives

$$(T(\mathcal{A}_\gamma^k)^{-1} q, (\mathcal{A}_\gamma^k)^{-1} q) \leq C((T_\gamma + kI)(\mathcal{A}_\gamma^k)^{-1} q, (\mathcal{A}_\gamma^k)^{-1} q) = C((\mathcal{A}_\gamma^k)^{-1} q, q).$$

Combining (4.23) and the above two inequalities gives

$$(4.27) \quad \|\mathbf{v}_h\|^2 \leq C((\mathcal{A}_\gamma^k)^{-1} q, q).$$

Finally, (4.17) follows from (4.26) and (4.27). This completes the proof of the theorem.

Proof of Corollary 4.1. It immediately follows from the inverse inequality (2.18) that for $k \leq h^2$,

$$\|\mathbf{v}\|^2 \leq D_k(\mathbf{v}, \mathbf{v}) \leq C \|\mathbf{v}\|^2 \quad \text{for all } \mathbf{v} \in \mathbf{V}_h,$$

holds for C independent of k and h . Using this with (4.9) shows that all of the forms $(\mathbf{M}_k q, q)$ are equivalent, provided that $k \leq h^2$. The corollary then follows from Theorem 4.1 with $k = h^2$.

5. NUMERICAL EXAMPLES.

In this section, we present the results of numerical computations which illustrate the conditioning guaranteed by Theorem 4.1. We also report the number of iterations required to achieve a given accuracy using the preconditioned block methods discussed in Section 3.

As a model problem, we consider systems which arise from time stepping (2.1)–(2.5) for Ω equal to the unit square $(0,1)^2$ in R^2 . The space \mathbf{V}_h is defined as in Example 1. To define W_h , we start with the space \tilde{W}_h of piecewise constant functions with respect to the square mesh with edge lengths $1/n$ (see Example 1). Label the smaller squares in the x and y directions and let $\Phi_{i,j}$ denote the characteristic function which is one on the i, j 'th square. For $l, m \in \{1, 2, \dots, n/2\}$ let

$$\theta_{l,m} = \Phi_{2l,2m} - \Phi_{2l-1,2m} - \Phi_{2l,2m-1} + \Phi_{2l-1,2m-1}.$$

We then define

$$W_h = \left\{ \phi \in \tilde{W}_h \mid \int_{\Omega} \phi \, dx = 0 \text{ and } (\phi, \theta_{l,m}) = 0 \text{ for } 1 \leq l, m \leq n/2 \right\}.$$

It is shown in [12],[14] that the resulting pairs \mathbf{V}_h and W_h satisfy (2.15) with c_0 independent of n .

For convenience, we discretize the first term in (2.24) by a lumped mass approximation. This means that we use

$$(\mathbf{v}, \mathbf{w}) = h^2 \sum_{ij} \mathbf{v}_{ij} \cdot \mathbf{w}_{ij} \quad \text{for all } \mathbf{v}, \mathbf{w} \in \mathbf{V}_h.$$

Here the sum is taken over the nodes of the mesh and the subscript on the functions \mathbf{v} and \mathbf{w} denotes evaluation at the grid point. We will use the operator \mathcal{A}_h^k analyzed in Section 4 as a preconditioner for the Schur complement $\mathbf{B}\mathbf{A}_k^{-1}\mathbf{B}^*$ corresponding to the coupled system (2.23).

We let S_h denote the space of functions with zero mean value which are continuous on Ω and piecewise linear with respect to the triangulation defining \mathbf{V}_h . Because \mathbf{A}_k comes from a regular mesh and constant coefficient operator, its inverse can be computed by using the discrete sine transform. Similarly, the action of T_h can be computed by using the discrete cosine transform. By utilizing odd and even extensions, the discrete sine and cosine transforms can be reduced to the discrete Fourier transforms which can be efficiently computed in on the order of $n^2 \log(n)$ operations.

One can compute the action of \mathcal{A}_h^k and $\mathbf{B}\mathbf{A}_k^{-1}\mathbf{B}^*$ efficiently using the above techniques. The largest and smallest eigenvalues of $\mathcal{A}_h^k\mathbf{B}\mathbf{A}_k^{-1}\mathbf{B}^*$ were then computed using the power method. In the case of $k = 0$, we used $\mathcal{A}_h^{h^2}$ as suggested by Corollary 1. The resulting condition numbers as a function of h and k are reported in Table 1. For the reported range of h and k , the condition numbers of the system $\mathcal{A}_h^k\mathbf{B}\mathbf{A}_k^{-1}\mathbf{B}^*$ were bounded by 5.3.

Table 1.
Condition number of $\mathcal{A}_h^k\mathbf{B}\mathbf{A}_k^{-1}\mathbf{B}^$.*

h	$k = 1$	$k = h$	$k = h^2$	$k = 0$
1/8	4.6	4.3	3.8	4.3
1/16	4.9	4.6	4.3	4.5
1/32	5.2	4.7	4.5	4.6
1/64	5.3	5.0	4.5	4.6

We next illustrate the behavior of the last three iterative algorithms discussed in Section 3. For these examples, we use a V-cycle multigrid preconditioner for each component of the operator \mathbf{A}_k . Specifically, we consider the V-cycle algorithm which uses one pre and post sweep of point-Gauss-Seidel as a smoother (see [4]). For the multigrid algorithm to give rise to a symmetric preconditioning operator, the order of the points in the post Gauss-Seidel smoother is the reverse of that used in the pre Gauss-Seidel smoother. Because of the lumped mass term, the resulting multigrid scheme is not of variational type (see, Section 5 of [7]). By direct computation, the resulting preconditioner applied to \mathbf{A}_k has eigenvalues in the interval $[\frac{1}{74}, 1]$.

We applied the last three algorithms of Section 3 utilizing the multigrid preconditioner just discussed. For the normal and minimal residual algorithms, no scaling

was required. In the case of (3.15) preconditioned by \tilde{M}_0 , we defined \mathcal{J} to be the multigrid preconditioner scaled by 2. Since the original preconditioner applied to \mathbf{A}_k had eigenvalues in the interval $[\frac{1}{2}, 1]$, it follows that the scaled operator \mathcal{J} satisfies

$$.32 (\mathbf{A}_k \mathbf{v}, \mathbf{v}) \leq ((\mathbf{A}_k - \mathcal{J}^{-1}) \mathbf{v}, \mathbf{v}) \leq \frac{1}{2} (\mathbf{A}_k \mathbf{v}, \mathbf{v}) \quad \text{for all } \mathbf{v} \in \mathbf{V}_h.$$

As before, we used $\mathcal{K} = \mathcal{A}_h^k$.

In Tables 2,3 and 4, we report the number of iterations required to reduce the discrete L^2 norm error by a factor of 10^{-4} for the various preconditioned block iterative techniques discussed in Section 3. We choose the discrete L^2 norm as an unbiased measure of the solution error. All three methods are minimization methods with respect to different norms and Krylov subspaces. None of these norms correspond to the discrete L^2 norm. All methods are applied to calculate the same fixed known solution.

Table 2.
Iteration numbers for block normal method.

h	$k = 1$	$k = h$	$k = h^2$
1/8	40	32	25
1/16	49	36	30
1/32	52	35	29
1/64	53	34	27

By far, the worst method was the block normal method. It required significantly more iterations for the same accuracy when compared to the other two techniques. Moreover, each iteration of the normal equation method requires two operator and preconditioner evaluations and thus the cost per iteration is roughly double that of the other two methods.

Table 3.
Iteration numbers for the minimal residual method.

h	$k = 1$	$k = h$	$k = h^2$
1/8	24	19	15
1/16	29	23	18
1/32	31	23	18
1/64	31	22	18

Of the remaining two techniques, the reformulation (3.15) preconditioned by \tilde{M}_0 (cf. [5]) converged somewhat faster than the minimal residual technique (cf. [21]). Both of these methods involve roughly the same amount of computational

work since they only require one evaluation of the operator and preconditioner per iterative step. One must choose the scaling factor for the preconditioner in the case of reformulation (3.15) whereas the minimal residual technique does not require any parameter selection. Actually, somewhat better performance was observed using the reformulation (3.15) and A_0^{-1} defined by scaling the multigrid preconditioner by a factor of 1.5 (instead of 2). For comparison, the iteration results for for all three methods and $k = h^2$ were combined in Table 5.

Table 4.
Iteration numbers for reformulation (3.15) preconditioned by \tilde{M}_0 .

h	$k = 1$	$k = h$	$k = h^2$
1/8	18	14	13
1/16	23	18	14
1/32	25	18	14
1/64	25	17	14

Table 5.
Iteration numbers for all methods and $k = h^2$.

h	Normal	Min. Res.	Reformulation (3.15)
1/8	25	15	13
1/16	30	18	14
1/32	29	18	14
1/64	27	18	14

6. APPENDIX.

We provide a proof of (2.14) and (2.22) in this section. We include this section to show that the desired conclusions result from our assumptions. The techniques for proving these estimates are well known.

We start by proving (2.14). Let (\mathbf{v}, Q) be the solution to (2.6)–(2.9) with $\mathbf{g} = 0$. Consider the solution to the following Stokes problem in weak form: Find (Ψ, θ) in $\mathbf{V} \times W$ satisfying

$$(7.1) \quad \begin{aligned} D(\Psi, \mathbf{w}) + (\theta, \nabla \cdot \mathbf{w}) &= (\mathbf{v}, \mathbf{w}) \quad \text{for all } \mathbf{w} \in \mathbf{V}, \\ (\nabla \cdot \Psi, q) &= 0 \quad \text{for all } q \in W. \end{aligned}$$

By the regularity assumption (2.13),

$$(7.2) \quad \|\Psi\|_2 + \|\theta\|_1 \leq C \|\mathbf{v}\|.$$

It immediately follows from (2.10) and (7.1) that

$$(7.3) \quad \begin{aligned} \|\mathbf{v}\|^2 &= D(\Psi, \mathbf{v}) + (\theta, \nabla \cdot \mathbf{v}) \\ &= (f, \theta) \leq \|f\|_{-1} \|\theta\|_1. \end{aligned}$$

Inequality (2.14) then follows combining (7.2) and (7.3).

We next prove (2.22) by applying the finite element duality argument. Let $\mathbf{E} = \mathbf{v} - \mathbf{v}_h$ and $e = Q - Q_h$ where (\mathbf{v}, Q) and (\mathbf{v}_h, Q_h) are respectively the solutions of (2.10) and (2.20). Note that

$$\begin{aligned} D(\mathbf{E}, \mathbf{w}) + (e, \nabla \cdot \mathbf{w}) &= 0 & \text{for all } \mathbf{w} \in \mathbf{V}_h, \\ (\nabla \cdot \mathbf{E}, q) &= 0 & \text{for all } q \in W_h. \end{aligned}$$

Consider the solution to the following Stokes problem in weak form: Find (Ψ, θ) in $\mathbf{V} \times W$ satisfying

$$(7.4) \quad \begin{aligned} D(\Psi, \mathbf{w}) + (\theta, \nabla \cdot \mathbf{w}) &= (\mathbf{E}, \mathbf{w}) & \text{for all } \mathbf{w} \in \mathbf{V}, \\ (\nabla \cdot \Psi, q) &= 0 & \text{for all } q \in W. \end{aligned}$$

By the regularity assumption (2.13),

$$(7.5) \quad \|\Psi\|_2 + \|\theta\|_1 \leq C \|\mathbf{E}\|.$$

We clearly have that for any $\mathbf{w} \in \mathbf{V}_h$ and $q \in W_h$,

$$\begin{aligned} \|\mathbf{E}\|^2 &= D(\Psi, \mathbf{E}) + (\theta, \nabla \cdot \mathbf{E}) \\ &= D(\Psi - \mathbf{w}, \mathbf{E}) - (e, \nabla \cdot (\mathbf{w} - \Psi)) + (\theta - q, \nabla \cdot \mathbf{E}) \\ &\leq Ch(\|\mathbf{E}\|_1 + \|e\|)(\|\Psi\|_2 + \|\theta\|_1) \\ &\leq Ch(\|\mathbf{E}\|_1 + \|e\|) \|\mathbf{E}\|. \end{aligned}$$

We used (2.16), (2.17) and (7.5) to get the last two inequalities above. It immediately follows that

$$(7.6) \quad \|\mathbf{E}\| \leq Ch(\|\mathbf{E}\|_1 + \|e\|).$$

To complete the proof of (2.22), we need only bound the two norms on the right hand side of (7.6) by $\|v\|_1$. By (2.11) and the fact that $\mathbf{g} = 0$,

$$\begin{aligned} \|Q\|^2 &\leq C_0^2 \sup_{\mathbf{w} \in \mathbf{V}} \frac{(Q, \nabla \cdot \mathbf{w})^2}{\|\mathbf{w}\|_1^2} \\ &= C_0^2 \sup_{\mathbf{w} \in \mathbf{V}} \frac{D(\mathbf{v}, \mathbf{w})^2}{\|\mathbf{w}\|_1^2} \leq CD(\mathbf{v}, \mathbf{v}). \end{aligned}$$

Similarly, (2.15) implies that

$$\|Q_h\|^2 \leq CD(\mathbf{v}_h, \mathbf{v}_h).$$

Finally,

$$\begin{aligned} D(\mathbf{v}_h, \mathbf{v}_h) &= |(Q_h, \nabla \cdot \mathbf{v}_h)| \\ &= |(Q_h, \nabla \cdot \mathbf{v})| \leq D(\mathbf{v}, \mathbf{v})^{1/2} \|Q_h\| \end{aligned}$$

It follows from the above three inequalities that

$$\|\mathbf{E}\|_1 + \|e\| \leq CD(\mathbf{v}, \mathbf{v})^{1/2}.$$

This completes the proof (2.22).

REFERENCES

1. A.K. Aziz and I. Babuška, *Part I, survey lectures on the mathematical foundations of the finite element method*, The Mathematical Foundations of the Finite Element Method with Applications to Partial Differential Equations (A.K. Aziz, ed.), Academic Press, New York, NY, 1972, pp. 1–362.
2. J.H. Bramble, *The Lagrange multiplier method for Dirichlet's problem*, Math. Comp. **37** (1981), 1–12.
3. J.H. Bramble and J.E. Pasciak, *A boundary parametric approximation to the linearized scalar potential magnetostatic field problem*, Appl. Numer. Math. **1** (1985), 493–514.
4. J.H. Bramble and J.E. Pasciak, *The analysis of smoothers for multigrid algorithms*, Math. Comp. **58** (1992), 467–488.
5. J.H. Bramble and J.E. Pasciak, *A preconditioning technique for indefinite systems resulting from mixed approximations of elliptic problems*, Math. Comp. **50** (1988), 1–18.
6. J.H. Bramble and J.E. Pasciak, *A domain decomposition technique for Stokes problems*, App. Num. Math. **6** (1990), 251–261.
7. J.H. Bramble, J.E. Pasciak and J. Xu, *The analysis of multigrid algorithms with non-nested spaces or non-inherited quadratic forms*, Math. Comp. **56** (1991), 1–34.
8. Cahouet and Chabard, *Some fast 3D finite element solvers for the generalized Stokes problem*, Int. J. Numer. Meth. in Fluid **8** (1988), 869–895.
9. P.G. Ciarlet, *The Finite Element Method for Elliptic Problems*, North-Holland, New York, 1978.
10. P.G. Ciarlet, *Basic error estimates for elliptic problems*, Finite Element Methods : Handbook of Numerical Analysis, (P.G. Ciarlet and J.L. Lions, eds.), vol. II, North-Holland, New York, 1991, pp. 18–352.
11. M. Dauge, *Elliptic Boundary Value Problems on Corner Domains*, Lecture Notes in Mathematics, 1341, Springer-Verlag, 1988.
12. V. Girault and P.A. Raviart, *Finite Element Approximation of the Navier-Stokes Equations*, Lecture Notes in Math. # 749, Springer-Verlag, New York, 1981.
13. P. Grisvard, *Elliptic Problems in Nonsmooth Domains*, Pitman, Boston, 1985.
14. C. Johnson and J. Pitkäranta, *Analysis of some mixed finite element methods related to reduced integration*, Math. Comp. **38** (1982), 375–400.
15. R.B. Kellogg and J.E. Osborn, *A regularity result for the Stokes problem in a convex polygon*, Jour. Funct. Anal. **21** (1976), 397–431.
16. J.L. Lions and E. Magenes, *Problèmes aux Limites non Homogènes et Applications*, vol. 1, Dunod, Paris, 1968.
17. Y. Maday, D. Meiron, A.T. Patera, and E.M. Ronquist, *Analysis of iterative methods for the steady and unsteady Stokes problem: Application to spectral element discretizations*, SIAM J. Stat. Comp. (1992).
18. J. Nečas, *Les Méthodes Directes en Théorie des Équations Elliptiques*, Academia, Prague, 1967.
19. J.C. Nedelec, *Elements finis mixtes incompressibles pour l'équation de Stokes dans R^3 .*, Numer. Math. **39** (1982), 97–112.
20. P.A. Raviart and J.M. Thomas, *A mixed finite element method for 2-nd order elliptic problems*, Mathematical Aspects of Finite Element Methods, Lecture Notes in Mathematics, #606 (Eds. I. Galligani and E. Magenes), Springer-Verlag, New York, 1977, pp. 292–315.
21. T. Rusten and R. Winther, *A preconditioned iterative method for saddle point problems*, SIAM J. Matrix Anal. Appl. **13** (1992), 887–904.
22. L.R. Scott and M. Vogelius, *Conforming finite element methods for incompressible and nearly incompressible continua*, Lectures in Applied Math (B. E. Engquist, et al., eds.), vol. 22, AMS, Providence 221-244, 1985.

DEPARTMENT OF APPLIED SCIENCE
BROOKHAVEN NATIONAL LABORATORY
UPTON, NY 11973
E-MAIL : PASCIAK@BNL.GOV