SIAM J. NUMER. ANAL.
Vol. 30, No. 3, pp. 630–649, June 1993

FAST ITERATIVE SOLUTION OF STABILISED STOKES SYSTEMS PART I: USING SIMPLE DIAGONAL PRECONDITIONERS*

ANDREW WATHEN† AND DAVID SILVESTER‡

Abstract. Mixed finite element approximation of the classical Stokes problem describing slow viscous incompressible flow gives rise to symmetric indefinite systems for the discrete velocity and pressure variables. Iterative solution of such indefinite systems is feasible and is an attractive approach for large problems. The use of stabilisation methods for convenient (but unstable) mixed elements introduces stabilisation parameters. The authors show how these can be chosen to obtain rapid iterative convergence.

The authors propose a conjugate gradient-like method (the method of preconditioned conjugate residuals) that is applicable to symmetric indefinite problems, describe the effects of stabilisation on the algebraic structure of the discrete Stokes operator, and derive estimates of the eigenvalue spectrum of this operator on which the convergence rate of the iteration depends. The simple case of diagonal preconditioning is discussed. The results apply to both locally and globally stabilised mixed elements as well as to elements which are inherently stable. It is demonstrated that convergence rates comparable to that achieved using the diagonally scaled conjugate gradient method applied to the discrete Laplacian are approachable for the Stokes problem.

Key words. Stokes problem, stabilised mixed finite elements, preconditioning

AMS subject classifications. 65N30, 65F10

1. Introduction.

Mixed finite element solution of the Stokes equations describing slow incompressible viscous flow leads to a symmetric indefinite discrete system for the pressure and velocity components. Since such systems are usually large and their solution is frequently part of an (outer) iterative scheme to solve the Navier–Stokes equations [Gl], rapid solution methods for such indefinite systems are desirable.

Given a flow domain Ω of $\Re^d (d = 2 \text{or} 3)$ with boundary $\partial\Omega$, a function $\mathbf{f}$ and appropriate boundary conditions, the classical form of the Stokes problem is to find the velocity $\mathbf{u}$ and pressure p satisfying

$$-\nu\nabla^2\mathbf{u} + \text{grad}p = \mathbf{f} \quad \text{in } \Omega, \tag{1.1}$$

$$\text{div}\mathbf{u} = 0 \quad \text{in } \Omega. \tag{1.2}$$

For simplicity, let

$$\mathbf{u} = 0 \quad \text{on } \partial\Omega, \tag{1.3}$$

representing "no-flow" on the boundary.

A standard mixed finite element approximation of the Stokes problem is obtained by constructing finite-dimensional subspaces $\mathbf{V}_h \subset \mathbf{V}$ and $P_h \subset P$, where $\mathbf{V}$ and P are

*Received by the editors October 15, 1991; accepted for publication (in revised form) June 4, 1992. This work was partially supported by the Fulbright Commission, the U.S. National Science Foundation under grant CCR-8821078-A2, and the U.K Science and Engineering Research Council under grant GR/H20299.

†School of Mathematics, University of Bristol, Bristol, BS8 1TW, United Kingdom.

‡Department of Mathematics, University of Manchester Institute of Science and Technology, P.O Box 88, Manchester, M60 1QD, United Kingdom.

appropriate Sobolev spaces (cf. [BF, p. 202]). The discrete Stokes problem is then to find $\mathbf{u}_h \in \mathbf{V}_h$ and $p_h \in P_h$ such that

$$(1.4) \qquad \begin{aligned} (\mathrm{grad}\mathbf{u}_h, \mathrm{grad}\mathbf{v}) - (p_h, \mathrm{div}\mathbf{v}) &= (\mathbf{f}, \mathbf{v}) \quad \forall \mathbf{v} \in \mathbf{V}_h \\ -(q, \mathrm{div}\mathbf{u}_h) &= 0 \quad \forall q \in P_h. \end{aligned}$$

Independently of the choice of $\mathbf{V}_h$ and P_h, the discrete system (1.4) can be written in block matrix form as

$$(1.5) \qquad \begin{pmatrix} A & B^t \\ B & O \end{pmatrix} \begin{pmatrix} u \\ p \end{pmatrix} = \begin{pmatrix} f \\ O \end{pmatrix},$$

where u is a vector of the discrete velocity variables and p a vector of the discrete pressure variables with respect to appropriate bases for $\mathbf{V}_h$ and P_h, respectively. The coefficient matrix is symmetric, but necessarily indefinite because of the zero diagonal block. The Dirichlet boundary conditions (1.3) for the velocity ensure that A is positive definite, and by applying the congruence transform

$$(1.6) \qquad \begin{pmatrix} A & B^t \\ B & O \end{pmatrix} = \begin{pmatrix} A & O \\ B & I \end{pmatrix} \begin{pmatrix} A^{-1} & O \\ O & -BA^{-1}B^t \end{pmatrix} \begin{pmatrix} A & B^t \\ O & I \end{pmatrix}$$

it is apparent that the coefficient matrix in (1.5) is nonsingular if and only if B has full row rank.

For a mixed finite element which is stable in the LBB sense ([BF, p. 205]), there exists a constant $\gamma > 0$ independent of the mesh spacing, h, such that

$$(1.7) \qquad \sup_{\mathbf{u}_h \in \mathbf{V}_h - \{0\}} \frac{(p_h, \mathrm{div}\mathbf{u}_h)}{\|\mathbf{u}_h\|_{\mathbf{V}}} \geq \gamma \|p_h\|_P \quad \forall p_h \in P_h,$$

where $\|.\|_{\mathbf{V}}$ and $\|.\|_P$ are norms in the underlying spaces $\mathbf{V}$ and P. Making the specific (and common) choice

$$\|\mathbf{u}_h\|_{\mathbf{V}} = (\mathrm{grad}\mathbf{u}_h, \mathrm{grad}\mathbf{u}_h)^{\frac{1}{2}}, \qquad \|p_h\|_P = (p_h, p_h)^{\frac{1}{2}}$$

leads to the matrix form

$$(1.8) \qquad \max_{u \in \Re^n - \{0\}} \frac{p^t B u}{[u^t A u]^{\frac{1}{2}}} \geq \gamma [p^t M_p p]^{\frac{1}{2}} \quad \forall p \in \Re^m.$$

Here m is the total number of discrete pressure variables (the dimension of P_h), n the total number of discrete velocity variables. Here and throughout the paper, $\Re^m$ is to be interpreted as excluding vectors corresponding to the hydrostatic pressure mode, i.e., vectors representing constant functions p_h, and $\Re^n$ is to be interpreted as discrete velocities satisfying the boundary condition (1.3).

In (1.8), M_p is the pressure mass matrix, i.e., the Grammian matrix of basis functions for P_h. It is symmetric and positive definite and has condition number independent of h for any usual finite element basis for P_h ([F], [W]). Also, as A is a discrete representation of a second-order operator, the Dirichlet boundary conditions (1.3) ensure that there exists a positive constant C such that $u^t A u / u^t u \geq C h^d$ for all

$u \in \Re^n - \{0\}$ ([AB, p. 240]). Thus (1.8) implies the existence of a positive constant C such that

$$\min_{p\in\Re^m-\{0\}} \max_{u\in\Re^n-\{0\}} \frac{p^tBu}{[p^tp]^{\frac{1}{2}}\,[u^tu]^{\frac{1}{2}}} \geq Ch^{\frac{d}{2}} \min_{p\in\Re^m-\{0\}} \frac{[p^tM_pp]^{\frac{1}{2}}}{[p^tp]^{\frac{1}{2}}}. \tag{1.9}$$

The left-hand side of (1.9) is a definition of the smallest singular value of B ([GVL, p. 286]), hence B is of full rank and we see that the Stokes system (1.5) for an LBB stable element is uniquely solvable on any finite computational grid.

If we also assume boundedness of B, i.e., that there exists Γ with

$$p^tBu \leq \Gamma\,[p^tM_pp]^{\frac{1}{2}}\,[u^tAu]^{\frac{1}{2}} \quad \forall u \in \Re^n \quad \text{and} \quad \forall p \in \Re^m, \tag{1.10}$$

then it is easily shown (see for example, [BF, pp. 75–78]) that in the case of an LBB stable element we have the bounds

$$\gamma^2 \leq \frac{p^t(BA^{-1}B^t)p}{p^tM_pp} \leq \Gamma^2 \quad \forall p \in \Re^m - \{0\}. \tag{1.11}$$

Since the spectral condition number of M_p is independent of the grid size h, (1.11) simply states that the "Schur complement" matrix $BA^{-1}B^t$ has spectral condition number independent of h for any stable element. Hence any system with $BA^{-1}B^t$ as coefficient matrix may be rapidly solved by a conjugate gradient or other iterative method (see, e.g., [GVL]). This has lead a number of authors to propose nested iterative solution strategies based on the block factorisation (1.6). See [V], [BP], [BWY]. In this paper, however, we will introduce the possibility of a single nonnested iterative solution of the indefinite Stokes system. An appropriate Krylov space (preconditioned conjugate residual) method is reviewed in §2. The relevent convergence analysis reveals that a certain minimax polynomial approximation problem on the eigenvalue spectrum of the coefficient matrix describes the rate of convergence of the iteration in an analogous way to the positive definite case.

The main aim of the paper is to consider the effect of stabilisation on the convergence of such a PCR iteration for unstable as well as stable elements. To this end in §3 we establish estimates for the eigenvalue spectrum of the stable (and a stabilised) Stokes operator after diagonal preconditioning, and in §4 we present some computational results and relate these to our analytic estimates.

For a review of the notion of stabilisation in the context of mixed finite elements, see [FHS]. In this paper we assume that the underlying mixed approximation is low-order so that a *stabilised* discrete Stokes formulation is a simple perturbation of (1.4) of the form: find $\mathbf{u}_h \in \mathbf{V}_h$ and $p_h \in P_h$ such that

$$\begin{aligned} (\mathrm{grad}\mathbf{u}_h, \mathrm{grad}\mathbf{v}) - (p_h, \mathrm{div}\mathbf{v}) &= (\mathbf{f}, \mathbf{v}) \quad \forall \mathbf{v}\in\mathbf{V}_h \\ -(q, \mathrm{div}\mathbf{u}_h) - \beta\mathcal{C}_h(p_h, q) &= 0 \quad \forall q\in P_h. \end{aligned} \tag{1.12}$$

Here $\beta > 0$ is a stabilisation parameter, and $\mathcal{C}_h(\cdot,\cdot)$ is a symmetric continuous bilinear form which is positive semidefinite on $P_h \times P_h$, and satisfies a *weak stabilisation condition*:

$$(p, \mathrm{div}\mathbf{v}) = 0 \quad \forall \mathbf{v}\in\mathbf{V}_h \;\Rightarrow\; \mathcal{C}_h(p,p) \neq 0. \tag{1.13}$$

Expressed in matrix form the system (1.12) is

$$\begin{pmatrix} A & B^t \\ B & -\beta C \end{pmatrix} \begin{pmatrix} u \\ p \end{pmatrix} = \begin{pmatrix} f \\ g \end{pmatrix} \tag{1.14}$$

where C is the symmetric positive semidefinite "stabilisation matrix."

For an unstable element, the upper bound (only) in (1.11) holds. Instability can arise either because the LBB constant γ is zero (corresponding to "pure" spurious pressure modes); or else it depends on the mesh parameter h, as discussed in [BF, p. 79]. The stabilisation condition (1.13) is clearly a necessary condition for a well-posed discrete problem since it ensures that all the eigenvalues of the matrix in (1.14) are nonzero, so a solution of the discrete system always exists and is unique. In practice, however, we need to ensure that the stability constant is *independent* of h. In particular, in order to derive eigenvalue bounds in §3, we need to assume that making a specific choice of C_h implies the existence of a value β_0 independent of h, such that for all $\beta > \beta_0$ the (*uniform-*) *stabilisation condition* holds:

$$\gamma^2 \leq \frac{p^t(BA^{-1}B^t + \beta C)p}{p^t M_p p} \leq \Gamma^2 \quad \forall p \in \Re^m - \{0\}, \tag{1.15}$$

with constants γ,Γ also independent of h. This condition is an obvious extension of (1.11) to the stabilised case.

Since C is always a bounded operator, the right-hand inequality in (1.15) is immediately satisfied for any stabilisation. Ensuring that the left-hand inequality holds in (1.15) turns out to be the key issue if an unstable mixed method is to be properly stabilised. One general approach which leads to uniform stability (in the sense of (1.15)) clearly arises if the pressure space P_h can be decomposed into a stable part, say Q_h in which (1.11) holds, and the orthogonal complement $Q_h^\perp$ (the "unstable" part):

$$P_h = Q_h \oplus Q_h^\perp. \tag{1.16}$$

In this case, using a minimax argument the trivial bound

$$0 \leq \frac{p^t C p}{p^t M_p p} \quad \forall p \in Q_h \tag{1.17}$$

can be combined with the bound (1.11) which holds in Q_h, to give

$$\gamma^2 \leq \frac{p^t(BA^{-1}B^t + \beta C)p}{p^t M_p p} \quad \forall p \in Q_h. \tag{1.18}$$

If a C can be constructed so that over the space $Q_h^\perp$ we have that

$$\phi^2 \leq \frac{p^t C p}{p^t M_p p} \quad \forall p \in Q_h^\perp, \tag{1.19}$$

where ϕ is independent of h, then combining with the trivial bound over $Q_h^\perp$:

$$0 \leq \frac{p^t(BA^{-1}B^t)p}{p^t M_p p} \quad \forall p \in Q_h^\perp \tag{1.20}$$

gives the bound

$$\beta\phi^2 \leq \frac{p^t(BA^{-1}B^t + \beta C)p}{p^t M_p p} \quad \forall p \in Q_h^\perp. \tag{1.21}$$

Clearly if the spaces Q_h and $Q_h^{\perp}$ are also eigenspaces of the Schur complement, then comparing (1.18) and (1.21) gives the desired result (1.15) for any value of β greater than the critical stabilisation parameter $\beta_0 = \gamma^2/\phi^2$.

Using this construction, the validity of a particular stabilisation ultimately depends on the existence of an appropriate splitting (1.16) in order that (1.19) be satisfied. Stabilisation methods based on macro-elements, so-called *locally* stabilised methods, for example the methods analysed in [KS] and [PS], may be shown to satisfy (1.15) by this approach. The uniform stability (in the sense of (1.15)) of so-called *globally* stabilised methods, sometimes referred to as Galerkin Least Squares formulations, may also be amenable to such a decomposition of the pressure space, although a more elegant theoretical framework for establishing (1.15) in this case can almost certainly be constructed using "Verfürth's trick" (cf. [FHS]).

2. Iterative methods for indefinite systems. The applicability and efficiency of Conjugate Gradient methods ([GVL]) for solving symmetric positive definite systems is widely appreciated. For symmetric indefinite problems such as the discrete Stokes problem, block elimination yields a definite Schur complement (for a stable element) and back substitution gives also a definite system, hence using nested iteration ([BWY], [BP]) is an attractive approach. However, Conjugate Gradient methods exist for indefinite systems also, and a nonnested iterative solution of the Stokes problem is perfectly feasible using, for example, the method of Preconditioned Conjugate Residuals (PCR) which is alternatively called the MINRES algorithm [AMS]. We sketch here the relevant properties of PCR which may also be worked out from more general treatments such as given in [AMS].

The PCR algorithm for solving $\mathcal{A}x = b$ with symmetric indefinite $\mathcal{A}$ and symmetric positive definite preconditioning matrix $\mathcal{M}$ is expressible in either the Orthodir or Orthomin forms ([JY], [AMS]). The robust Orthodir form, for example, is

$$r_0 = b - \mathcal{A}x_0, \qquad p_0 = \mathcal{M}^{-1}r_0$$

$$\begin{aligned}
\alpha_i &= p_i^t \mathcal{A}\mathcal{M}^{-1} r_i / p_i^t \mathcal{A}\mathcal{M}^{-1}\mathcal{A}p_i \\
x_{i+1} &= x_i + \alpha_i p_i \\
r_{i+1} &= r_i - \alpha_i \mathcal{A}p_i \\
\gamma_i &= p_i^t \mathcal{A}\mathcal{M}^{-1}\mathcal{A}\mathcal{M}^{-1}\mathcal{A}p_i / p_i^t \mathcal{A}\mathcal{M}^{-1}\mathcal{A}p_i \\
\sigma_i &= p_{i-1}^t \mathcal{A}\mathcal{M}^{-1}\mathcal{A}\mathcal{M}^{-1}\mathcal{A}p_i / p_{i-1}^t \mathcal{A}\mathcal{M}^{-1}\mathcal{A}p_{i-1} \\
p_{i+1} &= \mathcal{M}^{-1}\mathcal{A}p_i - \gamma_i p_i - \sigma_i p_{i-1}.
\end{aligned} \tag{2.1}$$

This method can be implemented with only two matrix $\times$ vector products at each iteration. Only a single matrix $\times$ vector product is needed at each iteration in general if a hybrid Orthomin/Orthodir form is used ([AMS]).

If we define the norm

$$\|\!|y|\!\| = (y^t \mathcal{A}\mathcal{M}^{-1}\mathcal{A}y)^{\frac{1}{2}},$$

then the PCR iterates have the property that

$$\|\!|x - x_k|\!\| \leq \|\!|x - y|\!\|$$

for all y in the (affine) Krylov space

$$x_0 + \operatorname{span}\{\mathcal{M}^{-1}r_0, (\mathcal{M}^{-1}\mathcal{A})\mathcal{M}^{-1}r_0, \ldots, (\mathcal{M}^{-1}\mathcal{A})^{k-1}\mathcal{M}^{-1}r_0\}.$$

Thus if Π_{k-1} is the set of real polynomials of degree $k-1$, we have

$$\|\!|x - x_k|\!\| = \min_{p \in \Pi_{k-1}} \|\!|\mathcal{A}^{-1} r_0 - p(\mathcal{M}^{-1}\mathcal{A})\mathcal{M}^{-1} r_0|\!\| \tag{2.2}$$

from which we may further deduce

LEMMA. *If $x = \mathcal{A}^{-1}b$ and x_k is the kth* PCR *iterate defined by* (2.1), *then*

$$\|\!|x - x_k|\!\| \leq \min_{p \in \Pi_k^1} \max_{i \in \{1,\ldots,N\}} |p(\lambda_i)| \, \|r_0\|_{\mathcal{M}^{-1}}, \tag{2.3}$$

where $\{\lambda_i : i = 1, \ldots, N\}$ are the eigenvalues of $\mathcal{M}^{-1}\mathcal{A}$, Π_k^1 is the set of kth degree polynomials with constant term one and $\|y\|_{\mathcal{M}^{-1}} = (y^t \mathcal{M}^{-1} y)^{\frac{1}{2}}$.

The λ_i are real since positive definiteness of $\mathcal{M}$ implies that $\mathcal{M}^{\frac{1}{2}}$ exists and so $\mathcal{M}^{-1}\mathcal{A}$ is similar to the symmetric matrix $\mathcal{M}^{-\frac{1}{2}}\mathcal{A}\mathcal{M}^{-\frac{1}{2}}$. The analysis is entirely analogous to that used in the case of positive definite $\mathcal{A}$ (see, for example, [AB, pp. 24–26]).

The convergence estimate (2.3) indicates that the convergence of the PCR iteration depends crucially on the spectrum of $\mathcal{M}^{-\frac{1}{2}}\mathcal{A}\mathcal{M}^{-\frac{1}{2}}$. The PCR method can be applied when $\mathcal{A}$ is positive definite, in which case the spectrum is contained in a real positive interval $[\lambda_{\min}, \lambda_{\max}]$ (using an obvious notation), and the estimate (2.3) can be expressed in terms of shifted Chebyshev polynomials, which are the minimax polynomials on this interval. Using this approach, one can derive the estimate

$$\|\!|x - x_k|\!\| \leq 2\left(\frac{1-\sqrt{\kappa}}{1+\sqrt{\kappa}}\right)^k \|\!|x - x_0|\!\| \tag{2.4}$$

for the PCR algorithm where $\kappa = \lambda_{\max}(\mathcal{M}^{-\frac{1}{2}}\mathcal{A}\mathcal{M}^{-\frac{1}{2}})/\lambda_{\min}(\mathcal{M}^{-\frac{1}{2}}\mathcal{A}\mathcal{M}^{-\frac{1}{2}})$ is the spectral condition number. The corresponding estimate for the common preconditioned version of the Hestenes–Stiefel CG algorithm, which is applicable when $\mathcal{A}$ and $\mathcal{M}$ are both symmetric and positive definite, is

$$\|x - x_k\|_{\mathcal{M}^{-\frac{1}{2}}\mathcal{A}\mathcal{M}^{-\frac{1}{2}}} \leq 2\left(\frac{1-\sqrt{\kappa}}{1+\sqrt{\kappa}}\right)^k \|x - x_0\|_{\mathcal{M}^{-\frac{1}{2}}\mathcal{A}\mathcal{M}^{-\frac{1}{2}}}, \tag{2.5}$$

where $\|y\|_{\mathcal{M}^{-\frac{1}{2}}\mathcal{A}\mathcal{M}^{-\frac{1}{2}}} = (y^t \mathcal{M}^{-\frac{1}{2}}\mathcal{A}\mathcal{M}^{-\frac{1}{2}} y)^{\frac{1}{2}}$. (Positive definiteness of $\mathcal{M}$ is not in fact necessary; see [AMS].) The Hestenes–Stiefel algorithm requires a single matrix × vector product at each iteration and thus is the method of choice in the positive definite case. If there are significant gaps in the spectrum of $\mathcal{A}$, then the Chebyshev estimate (2.5) is pessimistic and more accurate estimates can be derived by considering the polynomial minimax approximation problem in (2.3) on disjoint subintervals which contain the spectrum [AB], [G], [R].

When $\mathcal{A}$ is indefinite (but nonsingular), PCR is still applicable, but the Hestenes–Stiefel method can fail. In this case the spectrum is contained in $[-b, -a] \cup [a, b]$ for some $a, b > 0$ and again analytic expressions for the minimax errors on this set in terms of Chebyshev polynomials are available [L]. However, such an estimate is most likely to be pessimistic unless the eigenvalues are essentially symmetric about the origin. Other estimates are available which are applicable in situations where there are very few negative eigenvalues [SzW] but neither are these cases important here. In §4 we show minimax polynomials on the discrete eigenvalue sets which come from some of the specific problems considered in §3.

The role of preconditioning for conjugate gradient solution of symmetric matrix equations is to make the spectrum more clustered: this is often replaced by the simpler goal of reducing the condition number κ. Unless there is some symmetry in the eigenvalues about the origin which is preserved under preconditioning, then the simple reduction of κ seems a less appropriate goal in the indefinite case.

In §4 we will establish inclusion intervals for the eigenvalues of various indefinite discrete representations of the Stokes problem when only an elementary diagonal scaling is used. More sophisticated preconditioning strategies are developed in [SW]. The simple diagonal preconditioning case serves to illustrate the difference between various elements, and in particular to highlight the fact that stabilisation has an effect on iterative convergence.

3. Eigenvalue estimates. We are interested in solving systems with the symmetric coefficient matrix

$$\mathcal{A} = \begin{pmatrix} A & B^t \\ B & -\beta C \end{pmatrix}, \tag{3.1}$$

where A represents a block diagonal matrix of discrete Laplacians, β is the stabilisation parameter, $-C$ the stabilisation matrix, and B the coupling terms between velocities and pressure. A is positive definite and C is positive semidefinite. We will denote by n the order of the square matrix A and by m the order of C. (Thus, B is $m \times n$). In all practical cases $n > m$. For a stable element, we may simply take $C = 0$ throughout this section.

We choose the positive definite preconditioner

$$\mathcal{M} = \begin{pmatrix} D_A & O \\ O & \beta D_C \end{pmatrix} \tag{3.2}$$

where $D_A = \mathrm{diag}(A)$, $D_C = \mathrm{diag}(C)$ if $C \neq 0$, or else $D_C = h^d I$ where d is the spatial dimension as before. We are assuming quasi uniformity of the grid. Certainly D_A is positive definite with (diagonal) entries of $O(1)$ in $\Re^2$ and $O(h)$ in $\Re^3$. For all types of stabilisation described in the previous section, D_C is also positive definite with the diagonal entries being $O(h^2)$ in two dimensions, and $O(h^3)$ in three dimensions. The definition of D_C in the stable case is designed to satisfy a corresponding scaling with the mesh size. The important point is simply that D_C be spectrally equivalent to the pressure mass matrix, i.e., that there exist constants θ, Θ independent of h such that

$$\theta^2 \leq \frac{p^t M_p p}{p^t D_C p} \leq \Theta^2 \quad \forall p \in \Re^m - \{0\}. \tag{3.3}$$

Using the results of [W], $D_C = \mathrm{diag}(M_p)$ satisfies (3.3) and is thus another reasonable choice for both stable and unstable elements. As the parameter β does not arise in (3.1) in the stable case and the scaling of D_C with h is prescribed, it is not appropriate to take β to be other than unity in the case of an unstabilised stable element.

Recalling the convergence estimate (2.3) we are interested in the optimal minimax polynomials of increasing degree on the eigenvalue spectrum of the diagonally scaled matrix

$$\mathcal{M}^{-\frac{1}{2}} \mathcal{A} \mathcal{M}^{-\frac{1}{2}} = \begin{pmatrix} D_A^{-\frac{1}{2}} A D_A^{-\frac{1}{2}} & \frac{1}{\sqrt{\beta}} D_A^{-\frac{1}{2}} B^t D_C^{-\frac{1}{2}} \\ \frac{1}{\sqrt{\beta}} D_C^{-\frac{1}{2}} B D_A^{-\frac{1}{2}} & -D_C^{-\frac{1}{2}} C D_C^{-\frac{1}{2}} \end{pmatrix} = \begin{pmatrix} \widetilde{A} & \frac{1}{\sqrt{\beta}} \widetilde{B}^t \\ \frac{1}{\sqrt{\beta}} \widetilde{B} & -\widetilde{C} \end{pmatrix} = \widetilde{\mathcal{A}}. \tag{3.4}$$

(Note $\widetilde{C} = 0$ in the unstabilised case.)

By applying Sylvester's Law of Inertia to the congruence transform of $\widetilde{A} - \xi I$,

(3.5)
$$\begin{pmatrix} I & O \\ \frac{1}{\sqrt{\beta}}\widetilde{B}(\widetilde{A}-\xi I)^{-1} & I \end{pmatrix}\begin{pmatrix} \widetilde{A}-\xi I & O \\ O & -\frac{1}{\beta}\widetilde{B}(\widetilde{A}-\xi I)^{-1}\widetilde{B}^t - \widetilde{C} - \xi I \end{pmatrix}\begin{pmatrix} I & \frac{1}{\sqrt{\beta}}(\widetilde{A}-\xi I)^{-1}\widetilde{B}^t \\ O & I \end{pmatrix}$$

for $\xi = 0$, we see that the inertia of $\widetilde{\mathcal{A}}$ is invariant with $h > 0$ and $\beta > 0$ provided C (equivalently $\mathcal{C}_h$) satisfies the *weak stabilisation condition* (1.13) or the element is stable. This follows since in algebraic form (1.13) is

$$p^t B v = 0 \quad \forall v \in \Re^n \Rightarrow p^t C p \neq 0 \tag{3.6}$$

or in scaled form

$$p^t \widetilde{B} v = p^t D_C^{-\frac{1}{2}} B D_A^{-\frac{1}{2}} v = 0 \quad \forall v \in \Re^n \Rightarrow p^t D_C^{-\frac{1}{2}} C D_C^{-\frac{1}{2}} p = p^t \widetilde{C} p \neq 0. \tag{3.7}$$

We thus denote the eigenvalues of $\widetilde{\mathcal{A}}$ by

$$\mu_{-m} \leq \mu_{-m+1} \leq \cdots \mu_{-1} \leq 0 < \mu_1 \leq \mu_2 \leq \cdots \leq \mu_n. \tag{3.8}$$

Correspondingly, we denote the eigenvalues of $\widetilde{A}$ by $(0<)\, \lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n$ and the eigenvalues of $-\widetilde{C}$ by $\lambda_{-m} \leq \lambda_{-m+1} \leq \cdots \leq \lambda_{-1} (\leq 0)$. Note $\lambda_{-i} = 0$, $i = 1, \ldots, m$ for an unstabilised method. We also (break the usual convention to consistently) write the singular values of $\widetilde{B}$ as $(0 \leq) \sigma_1 \leq \sigma_2 \leq \cdots \leq \sigma_m$.

The matrix A is a set of discrete Laplacians with Dirichlet boundary conditions; thus its smallest eigenvalue is $O(h^d)$ and its largest eigenvalue is $O(h^{d-2})$ for a problem on a quasi-uniform grid in $\Re^d$ ([AB, p. 240]). Since the diagonal entries of A are all $O(h^{d-2})$, the eigenvalues of the scaled discrete Laplacians $\widetilde{A}$ satisfy

$$\begin{aligned} \lambda_1 &= \min_{v \in \Re^n - \{0\}} \frac{v^t \widetilde{A} v}{v^t v} = \min_{u \in \Re^n - \{0\}} \frac{u^t A u}{u^t D_A u} = \upsilon^2 h^2, \\ \lambda_n &= \max_{v \in \Re^n - \{0\}} \frac{v^t \widetilde{A} v}{v^t v} = \max_{u \in \Re^n - \{0\}} \frac{u^t A u}{u^t D_A u} = \Upsilon^2. \end{aligned} \tag{3.9}$$

Our first task is to prove that σ_m is bounded independently of h.

From (1.11), whether we are using a stable or unstable element, we have

$$\frac{p^t (B A^{-1} B^t) p}{p^t M_p p} \leq \Gamma^2 \quad \forall p \in \Re^m - \{0\}. \tag{3.10}$$

Also, by construction, D_C satisfies (3.3), so

$$\frac{p^t (B A^{-1} B^t) p}{p^t D_C p} = \frac{q^t D_C^{-\frac{1}{2}} B A^{-1} B^t D_C^{-\frac{1}{2}} q}{q^t q} \leq \Gamma^2 \Theta^2 \tag{3.11}$$

for all $q = D_C^{\frac{1}{2}} p \, (\neq 0)$. Furthermore, from (3.9)

$$\Upsilon^{-2} \leq \frac{u^t A^{-1} u}{u^t D_A^{-1} u} \quad \forall u \in \Re^n - \{0\}. \tag{3.12}$$

Thus in (3.11), for all q with $B^t D_C^{-\frac{1}{2}} q \neq 0$, we have

$$\frac{q^t D_C^{-\frac{1}{2}} B A^{-1} B^t D_C^{-\frac{1}{2}} q}{q^t D_C^{-\frac{1}{2}} B D_A^{-1} B^t D_C^{-\frac{1}{2}} q} \quad \frac{q^t D_C^{-\frac{1}{2}} B D_A^{-1} B^t D_C^{-\frac{1}{2}} q}{q^t q} \leq \Gamma^2 \Theta^2$$

which using (3.12) and the definition (3.4) of $\widetilde{B}$ gives

$$\frac{q^t \widetilde{B} \widetilde{B}^t q}{q^t q} \leq \Gamma^2 \Theta^2 \Upsilon^2. \tag{3.13}$$

For $q \neq 0$ with $B^t D_C^{-\frac{1}{2}} q = 0$, (3.13) is trivially satisfied, so (3.13) holds for all $q \in \Re^m - \{0\}$. Thus, since the singular values of $\widetilde{B}$ are the square roots of the eigenvalues of $\widetilde{B}\widetilde{B}^t$, it follows that $\sigma_m \leq \Gamma\Theta\Upsilon$. That is, σ_m is bounded independently of h.

We now write

$$\begin{pmatrix} \widetilde{A} & \frac{1}{\sqrt{\beta}} \widetilde{B}^t \\ \frac{1}{\sqrt{\beta}} \widetilde{B} & -\widetilde{C} \end{pmatrix} = \begin{pmatrix} \widetilde{A} & 0 \\ 0 & -\widetilde{C} \end{pmatrix} + \frac{1}{\sqrt{\beta}} \begin{pmatrix} 0 & \widetilde{B}^t \\ \widetilde{B} & 0 \end{pmatrix} \tag{3.14}$$

and apply the minimax result on the eigenvalues of the sum of two symmetric matrices ([Wi, pp. 101–102]). The eigenvalues of

$$\begin{pmatrix} 0 & \widetilde{B}^t \\ \widetilde{B} & 0 \end{pmatrix}$$

are $-\sigma_m, \ldots, -\sigma_1, 0\,(n-m \text{ times}), \sigma_1, \ldots, \sigma_m$ ([GVL, p. 286]), thus the eigenvalues of $\widetilde{\mathcal{A}}$ satisfy the bounds

$$\lambda_i - \frac{\sigma_m}{\sqrt{\beta}} \leq \mu_i \leq \lambda_i + \frac{\sigma_m}{\sqrt{\beta}}, \qquad i = -m, \ldots, -1, 1, \ldots, n. \tag{3.15}$$

Since $\sigma_m = O(1)$, this gives simple β-dependent inclusion intervals for the eigenvalues of $\widetilde{\mathcal{A}}$. The intervals are centred on the eigenvalues of $\widetilde{A}$ and $-\widetilde{C}$. Note that the inclusion intervals are small for large β. For small β, we may interchange the roles of the matrices in (3.14) to obtain

$$\begin{aligned} -\frac{\sigma_{-i}}{\sqrt{\beta}} - \lambda_{\max} \leq \mu_i \leq -\frac{\sigma_{-i}}{\sqrt{\beta}} + \lambda_{\max}, \qquad & i = -m, \ldots, -1, \\ -\lambda_{\max} \leq \mu_i \leq \lambda_{\max}, \qquad & i = 1, \ldots, n-m, \\ \frac{\sigma_{i-n+m}}{\sqrt{\beta}} - \lambda_{\max} \leq \mu_i \leq \frac{\sigma_{i-n+m}}{\sqrt{\beta}} + \lambda_{\max}, \qquad & i = n-m+1, \ldots, n. \end{aligned} \tag{3.16}$$

Here, $\lambda_{\max} = \max\{-\lambda_{-m}, \lambda_n\}$ is independent of h because of (3.9) and the assumed boundedness of $\widetilde{C}$ which we express as

$$\max_{p \in \Re^m - \{0\}} \frac{p^t \widetilde{C} p}{p^t p} = \max_{q \in \Re^m - \{0\}} \frac{q^t C q}{q^t D_C q} \leq \Delta. \tag{3.17}$$

Δ is a constant independent of h: for the two stabilisation methods specifically considered in the next section, $\Delta = 2$ is an appropriate value for use in (3.17).

Certain refinement of (3.15) is possible: for any $0 < \xi < \lambda_1$, application of the Sylvester Law of Inertia to the congruence transform (3.5) shows that $\lambda_1 \leq \mu_1$ for any β. Thus the positive spectrum of $\widetilde{\mathcal{A}}$ lies in the interval $[\lambda_1, \lambda_n + \sigma_m/\sqrt{\beta}]$ for any stabilisation. In practical computations we have observed that the smallest positive eigenvalue of $\widetilde{\mathcal{A}}$ moves away from the origin as β is decreased for a fixed value of h.

The negative part of the spectrum of $\widetilde{\mathcal{A}}$ will depend on the form of stabilisation $\mathcal{C}_h$ which gives rise to the stabilisation matrix C. For a stable element (without stabilisation), $\widetilde{C} = 0$, so (3.15) says how far the negative eigenvalues can move away from the origin.

For stabilisation methods (with stabilisations such as (4.1) and (4.2) given in the next section) a key issue is how the zero eigenvalues of $\widetilde{C}$ move away from the origin for finite values of β. Certainly $\mu_{-j} < 0, j = 1, \ldots, m$ for any finite β since the inertia is invariant with any bounded value of β and any h. Some further analysis, which also applies in the case of stable elements, helps answer this question.

Once again we make use of the Sylvester Law of Inertia. From (3.5), if we can show the existence of $\xi < 0$ with

$$\frac{1}{\beta}\frac{p^t\widetilde{B}(\widetilde{A}-\xi I)^{-1}\widetilde{B}^t p}{p^t p} + \frac{p^t\widetilde{C}p}{p^t p} > -\xi \quad \forall p \in \Re^m - \{0\} \tag{3.18}$$

then it follows that $\mu_{-1} < \xi$. (Note $(\widetilde{A}-\xi I)^{-1}$ is positive definite for all $\xi \leq 0$.) First, using the definition (3.4) of $\widetilde{A}$, $\widetilde{B}$, and $\widetilde{C}$ we see that the left-hand side of (3.18) is

$$\frac{1}{\beta}\frac{q^t B(A-\xi D_A)^{-1}B^t q}{q^t D_C q} + \frac{q^t C q}{q^t D_C q} \tag{3.19}$$

$$\geq \frac{\theta^2}{\beta}\frac{q^t B(A-\xi D_A)^{-1}B^t q}{q^t M_p q} + \theta^2\frac{q^t C q}{q^t M_p q} \tag{3.20}$$

from (3.3) with $q = D_C^{-\frac{1}{2}}p$. For an unstable element, if $B^t q = 0$ then from (1.15)

$$\frac{q^t C q}{q^t M_p q} \geq \frac{\gamma^2}{\beta}$$

thus (3.18) holds for any such $q = D_C^{-\frac{1}{2}}p$ with $-\xi = O(1)$. For any other q satisfying $B^t q \neq 0$, then (3.20) is

$$\theta^2\frac{z^t(I-\xi A^{-\frac{1}{2}}D_A A^{-\frac{1}{2}})^{-1}z}{z^t z}\,\frac{1}{\beta}\,\frac{q^t B A^{-1}B^t q}{q^t M_p q} + \theta^2\frac{q^t C q}{q^t M_p q} \tag{3.21}$$

with $z = A^{-\frac{1}{2}}B^t q\,(\neq 0)$. Now writing $y = (I - \xi A^{-\frac{1}{2}}D_A A^{-\frac{1}{2}})^{-\frac{1}{2}}z\,(\neq 0)$, we have

$$\frac{z^t(I-\xi A^{-\frac{1}{2}}D_A A^{-\frac{1}{2}})^{-1}z}{z^t z} = \left[\frac{y^t(I-\xi A^{-\frac{1}{2}}D_A A^{-\frac{1}{2}})y}{y^t y}\right]^{-1} \tag{3.22}$$

$$= \left[1 - \xi\frac{w^t D_A w}{w^t A w}\right]^{-1}, \qquad w = A^{-\frac{1}{2}}y \neq 0.$$

Now from (3.9), the expression (3.22) is as small as possible for w satisfying $v^2h^2 = w^tAw/w^tD_Aw$ for any $\xi < 0$. In (3.21) we thus have

$$\begin{aligned}
&\theta^2 \frac{z^t(I-\xi A^{-\frac{1}{2}} D_A A^{-\frac{1}{2}})^{-1} z}{z^t z} \frac{1}{\beta} \frac{q^t B A^{-1} B^t q}{q^t M_p q} + \theta^2 \frac{q^t C q}{q^t M_p q} \\
&\geq \theta^2 \left[1 - \xi v^{-2} h^{-2}\right]^{-1} \frac{1}{\beta} \frac{q^t B A^{-1} B^t q}{q^t M_p q} + \theta^2 \frac{q^t C q}{q^t M_p q} \\
&\geq \theta^2 \left[1 - \xi v^{-2} h^{-2}\right]^{-1} \left[\frac{1}{\beta} \frac{q^t B A^{-1} B^t q}{q^t M_p q} + \frac{q^t C q}{q^t M_p q}\right]
\end{aligned} \tag{3.23}$$

as $\xi < 0 \quad \Rightarrow \quad [1 - \xi v^{-2}h^{-2}] \geq 1$. But now we may use (1.15) (which is (1.11) in the stable case, $C = 0$) giving

$$\frac{1}{\beta} \frac{p^t \widetilde{B}(\widetilde{A}-\xi I)^{-1} \widetilde{B}^t p}{p^t p} + \frac{p^t \widetilde{C} p}{p^t p} \geq \frac{\theta^2 \gamma^2}{\beta} \left[1 - \xi v^{-2} h^{-2}\right]^{-1} \tag{3.24}$$

for any $\xi < 0$. Finally,

$$\frac{\theta^2 \gamma^2}{\beta} \left[1 - \xi v^{-2} h^{-2}\right]^{-1} \geq -\xi \tag{3.25}$$

is certainly satisfied for

$$\xi = -\frac{h\, \theta s \gamma v}{\sqrt{\beta}} + \frac{h^2\, v^2}{2}. \tag{3.26}$$

Thus μ_{-1} is not closer to the origin than $O(h/\sqrt{\beta})$.

In the case of a large stabilisation parameter, a more refined estimate of μ_{-1} can be derived by noting that for equality in (3.25), we have that

$$\xi = \frac{v^2 h^2}{2} - \frac{v^2 h^2}{2} \sqrt{1 + \frac{4\theta^2 \gamma^2}{\beta v^2 h^2}}, \tag{3.27}$$

so that if $\beta > 4\theta^2\gamma^2/v^2h^2 = O(h^{-2})$, we can expand in terms of the binomial series to give

$$\xi = -\frac{v^2 h^2}{2} \left(\frac{2\theta^2 \gamma^2}{\beta v^2 h^2} - \frac{2\theta^4 \gamma^4}{\beta^2 v^4 h^4} + \cdots\right) \approx -\frac{\theta^2 \gamma^2}{\beta} \tag{3.28}$$

for β sufficiently large. In this case μ_{-1} is not closer to the origin than $O(1/\beta)$, i.e., it is independent of h.

We summarise our results for the stabilised and stable cases separately, in the following.

THEOREM 1. *For a stabilised mixed finite element approximation of the Stokes problem* (1.12) *satisfying the uniform stabilisation condition* (1.15) *on a quasi-uniform grid, the eigenvalues of the diagonally preconditioned Stokes matrix* (3.4) *lie in the union of intervals*

$$\left[-\Delta - \frac{\Gamma\Theta\Upsilon}{\sqrt{\beta}},\ -\frac{\theta\gamma v}{\sqrt{\beta}} h + \frac{v^2}{2} h^2\right] \cup \left[v^2 h^2,\ \Upsilon^2 + \frac{\Gamma\Theta\Upsilon}{\sqrt{\beta}}\right],$$

where the constants γ, Γ *from* (1.15), θ, Θ *from* (3.3), υ, Υ *from* (3.9), *and* Δ *from* (3.17) *are independent of the grid size* h *and* β *is the stabilisation parameter.*

For a stable method, (1.11) necessarily holds. The corresponding result follows.

THEOREM 2. *For a stable mixed finite element approximation of the Stokes problem on a quasi-uniform grid, the eigenvalues of the diagonally preconditioned Stokes matrix* (3.4) (*with* $\beta = 1$ *and* $C = \widetilde{C} = 0$) *lie in the union of intervals*

$$\left[-\Gamma\Theta\Upsilon\ ,\ -\theta\gamma\upsilon\, h + \frac{\upsilon^2 h^2}{2}\right] \cup \left[\upsilon^2 h^2\ ,\ \Upsilon^2 + \Gamma\Theta\Upsilon\right],$$

where the constants γ, Γ *from* (1.11), θ, Θ *from* (3.3) *and* υ, Υ *from* (3.9) *are independent of the grid size* h.

The choice of $\beta\,(> 0)$ in the preconditioner (3.2) is apparently unconstrained, however choosing other than $\beta = 1$ affects the bounds (3.15), (3.16) on μ_{-m} and (3.26) on μ_{-1} in a compensating manner, i.e., making one better makes the other worse. For this reason we have taken $\beta = 1$ in Theorem 2. Our computations in the next section in any case indicate that (3.26) is pessimistic at least for the particular stable element considered there. The method behaves more like (3.28) for that element.

4. Discussion. In this section, the question of how to choose the stabilisation parameter so as to enhance the rate of convergence of the PCR algorithm of §2 is addressed. We compare the performance of two representative examples of stabilised mixed methods with that of an a priori stable method. For convenience we restrict attention to two-dimensional elements and consider only the lowest order continuous velocity approximation, i.e., based on linear triangle or bilinear square elements.

As a test example we solved the "leaky" two-dimensional lid-driven cavity problem in a unit square domain with a flow solution, calculated using the stable method below, as illustrated in Fig. 1. This problem was also discussed by Pierre [P], wherein he showed the sensitivity of the solution accuracy to the choice of the stabilisation parameter using (globally-) stabilised P_1–P_1 and Q_1–Q_1 methods. In this work, only half the domain was modelled exploiting the natural symmetry of the solution about the line $x = \frac{1}{2}$. Rectangular and triangular element grids were both used; starting from a uniform subdivision of $N \times 2N$ square elements, the triangular grids were constructing by dividing each square into two. In all cases an initial solution guess of zero was used, and the tolerance for convergence was a reduction of 10^{-6} in the L_2-norm of the residual. All computations were done using Pro–MATLAB on an SGI $4D/35$ Iris workstation. The stable case is discussed first.

4.1. Using an LBB stable method. Restricting ourselves to linear triangular or bilinear rectangular elements necessarily implies that the pressure approximation must be defined on a coarser grid if the element is to satisfy the LBB condition (1.7). Using a standard four-triangle macro-element (with internal edges connecting the midpoints of the macro-element edges), we can construct a stable P_1–P_1 method by using a continuous pressure approximation defined by the macro-element vertices. The method is commonly referred to as the $P_1 \text{iso} P_2$ method; cf. [BF, pp. 255]. For the uniform grids we used, the asymptotic ratio of velocity to pressure degrees of freedom is 8:1 as $h \to 0$ which is somewhat high. Thus from an approximation point of view this method is probably too under-constrained to be the "best" first-order method (i.e., with an $O(h)$ error for velocity in the H^1-norm). Note that the tetrahedral analogue of this method is also LBB stable.

Solving the test problem using the PCR algorithm gave the iteration counts in Table 1. Results for $\beta = 1$ with two choices of the matrix D_C (in (3.2)) are presented.

Fig 1a: Stokes Problem Velocity Plot

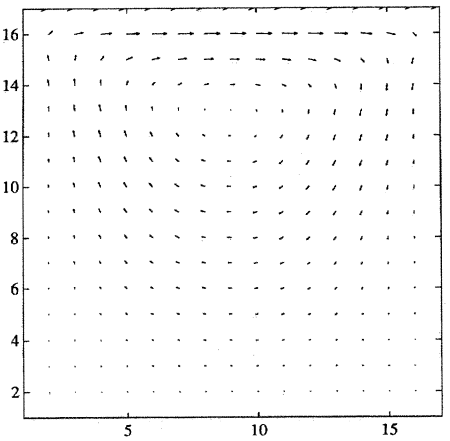

Fig 1b: Stokes Problem Pressure Distribution

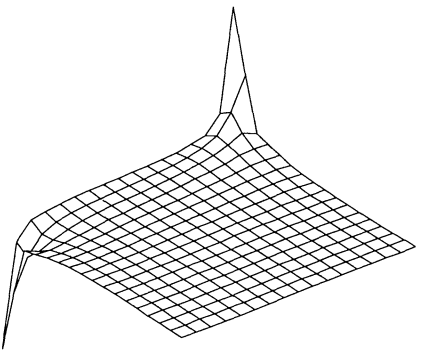

FIG. 1

Here D_{M_p} is the diagonal of M_p. These results vividly illustrate the importance of having the right scaling for the "pressure part" of the Stokes operator.

TABLE 1

Number of PCR *iterations in the stable case.*

Grid	$D_C = I$	$D_C = D_{M_p}$
2×4	17	16
4×8	88	56
8×16	261	130
16×32	*	276

A nice feature of these results is the fact that in the case where the "correct" scaling is used, the iteration count behaves like $O(h^{-1})$, as would be expected using diagonally scaled CG to solve Laplace's equation on a uniform sequence of grids. This behaviour can be explained by considering the actual eigenvalue distribution of the preconditioned Stokes operator (3.4). To get a flavour of this, the eigenvalue distribution for the 4×8 grid is illustrated in Fig. 2. Also plotted in Fig. 2 is the optimal polynomial approximation of degree 11 on the discrete set of values, i.e., the polynomial is constructed such that the PCR error contraction estimate (2.3) is minimised. Each dotted vertical bar represents an eigenvalue of (3.4) and is of height equal to twice the minimax error.

To see the way the eigenvalue distribution changes as $h \to 0$, and to compare with Theorem 2, the extremal eigenvalues $\mu_{-m}, \mu_{-1}, \mu_1, \mu_n$ are listed in Table 2. The key point is that the extremal positive eigenvalues indeed behave like those of a scaled Laplacian, whilst the negative eigenvalues remain in a fairly tight cluster which is bounded away from the origin independently of h. The behaviour of the positive eigenvalues is consistent with our perturbation analysis, but the fact that the eigenvalue μ_{-1} appears to be independent of h is slightly surprising. Our estimate (3.26) appears to be pessimistic in this case.

To conclude our discussion of Table 1, when we solved the test problem on the finest grid using the naive preconditioner $D_C = I$ (and $\beta = 1$), the PCR algorithm broke down after 576 iterations; after a residual reduction of about 10^{-5}, the IEEE

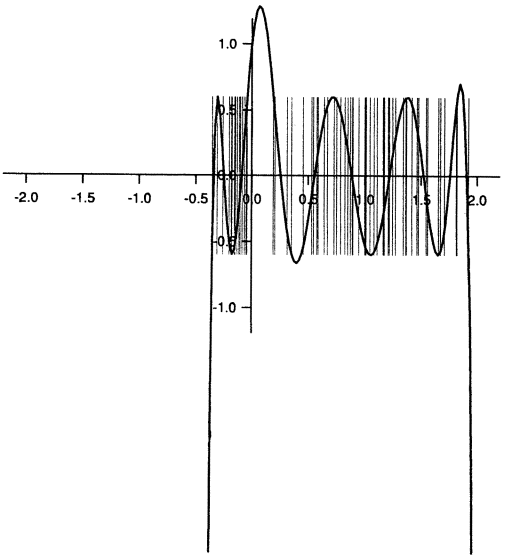

FIG. 2

TABLE 2
Eigenvalues of $\widetilde{\mathcal{A}}$ in the stable case.

Grid	μ_{-m}	μ_{-1}	μ_1	μ_n
2×4	-0.2842	-0.0286	0.4822	1.7307
4×8	-0.3538	-0.0405	0.1925	1.9255
8×16	-0.3688	-0.0407	0.0505	1.9809
16×32	-0.3721	-0.0404	0.0127	1.9952

number NaN was calculated. This erratic behaviour seems to emphasise the importance of preconditioning the Stokes operator so as to ensure the right scaling of the pressure and velocity with respect to h.

4.2. Using a globally stabilised method. Starting from a mixed approximation based on a *continuous* pressure, the obvious route to stabilisation via (1.12) is by having a stabilisation term $\mathcal{C}_h(\cdot,\cdot)$ which controls gradients in pressure:

$$\mathcal{C}_h(p_h, q_h) = \beta \sum_{K \in \tau_h} h_K^2 \int\int_K \mathrm{grad} p_h \cdot \mathrm{grad} q_h \mathrm{d}K. \tag{4.1}$$

This type of regularisation was first suggested by Brezzi and Pitkäranta [BPi] in the context of the P_1–P_1 triangular element. Continuous pressure elements which depend on internal velocity bubble functions for their stability, for example the popular "mini element" introduced by Arnold, Brezzi, and Fortin in [ABF], can also be expressed as a stabilised method of the form (1.12) with $\mathcal{C}_h(\cdot,\cdot)$ given by (4.1) (after static condensation of the bubble terms). Note that in this case the magnitude of β cannot be arbitrarily chosen; it is fixed by the underlying mixed approximation. See [P] for details in the mini-element case. An important feature of (4.1) is the "global" nature of

the stabilisation. The point is that (4.1) represents an approximation to the Laplacian of the pressure, defined over the entire domain Ω.

We will use the continuous pressure P_1–P_1 triangle stabilised using (4.1) as a representative globally stabilised method here. A very attractive feature of this method is its inherent simplicity, both the velocity and the pressure being defined by the same piecewise polynomial basis set. Another important feature, is that the approximating power of the method is much better than that of the stable method above. If the effect of the stabilisation term is ignored, then the asymptotic velocity to pressure constraint count is 2:1, which is "optimal" in two dimensions. Perhaps the only negative feature of this method is the fact that solution accuracy is known to deteriorate if the stabilisation parameter is not chosen correctly; see [P] for details. On the one hand, if the parameter is too small then the method might not be sufficiently stable to give good results. For example, solving our test problem with $\beta < 10^{-2}$ gave rise to oscillatory pressure solutions. On the other hand, it is easily seen that in the limiting case of $\beta \to \infty$ the pressure solution tends to a constant, which implies that the corresponding velocity field is nowhere near divergence-free. Solving the test problem with $\beta > 10$ gave poor solutions on all of the grids we considered.

Returning now to our main concern; that of finding the "optimal" choice of stabilisation parameter in the sense that the contraction factor in (2.3) is minimised. PCR iteration counts using preconditioner (3.2) for a range of values of β are listed in Table 3. These results illustrate that the efficiency of the PCR solution method is also crucially dependent on the choice of stabilisation parameter.

TABLE 3

Number of PCR *iterations in the globally stabilised case.*

Grid	$\beta = 0.01$	$\beta = 0.025$	$\beta = 0.1$	$\beta = 1.0$	$\beta = 10.0$
2×4	22	21	22	23	23
4×8	71	61	65	91	93
8×16	169	135	150	285	359
16×32	400	307	369	741	1279

The characteristic feature of globally stabilised methods is the fact that the stabilisation matrix C represents some discrete approximation to the Laplacian (with Neumann boundary conditions). This means that the eigenvalues $\lambda_{-m}, \ldots, \lambda_{-1}$ of $-\widetilde{C}$ are fairly evenly spread within the interval $[-2, 0]$. For comparison with the a priori stable case above, we plot the eigenvalue distribution corresponding to the case of $\beta = 0.025$ in Fig. 3, together with the optimal minimax polynomial of degree 11.

The results in Table 3 closely agree with our perturbation analysis. For a small value of β the estimate (3.16) applies, and the extremal eigenvalues μ_{-m} and μ_n are forced to move out towards $\pm\infty$ (by an amount proportional to $1/\sqrt{\beta}$) independently of h. Applying (3.16) in the case of the smallest positive eigenvalue μ_1, we see that its movement away from the origin is bounded independently of β by $\lambda_{\max}(= 2)$. Hence the condition of the preconditioned system might be expected to deteriorate as $\beta \to 0$, as reflected by the iteration counts in the case of $\beta = 0.01$. For a large value of β the estimate (3.15) applies, showing that the spectrum of the preconditioned system is "close" to that of $\widetilde{A}$ on the positive side, and to that of $-\widetilde{C}$ on the negative side. In particular as $\beta \to \infty$ the spectrum becomes essentially symmetric about the origin.

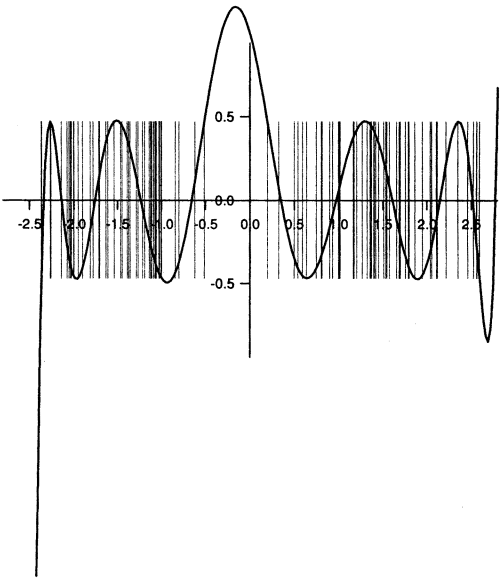

FIG. 3

In this situation, it is well known that the application of Conjugate Residuals gives essentially the same rate of convergence as would be obtained using CG to solve the "normal equations," i.e., the condition number of the system is effectively squared ([Fr]). Thus for large values of β, we could expect the number of iterations to increase by a factor of four for each successive refinement of the grid, as can be seen in the case of $\beta = 10$. Interestingly, the "optimal" choice of $\beta = 0.025$ from the table turns out to be a very natural choice. It is precisely the value (see [P]) which generates the system that would result using the analogous subdivision of P_1–P_1 mini-elements, after elimination of the internal velocity bubble terms.

TABLE 4

Eigenvalues of $\widetilde{\mathcal{A}}$ in the globally stabilised case with parameter $\beta = 0.025$.

Grid	μ_{-m}	μ_{-1}	μ_1	μ_n
2×4	-2.2251	-0.6888	0.6986	2.5778
4×8	-2.3667	-0.5122	0.1974	2.5916
8×16	-2.4553	-0.2846	0.0507	2.5931
16×32	-2.4865	-0.1476	0.0128	2.5931

In Table 4 we illustrate the variation of the extremal eigenvalues with grid refinement in the mini-element case (i.e., with $\beta = 0.025$). Comparing these values with those in Table 2, note that there is a fundamental difference in the behaviour of the largest negative eigenvalue μ_{-1}, our estimate (3.26) of $O(h)$ movement of the eigenvalue μ_{-1} appears to be tight in this case.

4.3. Using a locally stabilised method. Finally we discuss the performance of a representative locally stabilised method, namely the Q_1–P_0 quadrilateral, locally

stabilised (see [KS]) over 2×2 macro-elements via the "local" jump operator:

$$C_h(p_h, q_h) = \beta \sum_{M \in \mathcal{M}_h} \sum_{e \in \Gamma_M} h_e \int_e [\![p_h]\!]_e [\![q_h]\!]_e \mathrm{d}s. \tag{4.2}$$

Note that in the limit of $\beta \to \infty$, the use of (4.2) in formulation (1.12) leads to a *locally continuous* (i.e., within macro-elements) pressure solution. Indeed, the method is equivalent to a stable mixed method as $\beta \to \infty$. A proof that this approach leads to a uniformly stabilised method (in the sense of (1.15)) is given in [A].

The attractive feature of this method (apart from its simplicity) is the fact that it also has the "optimal" approximation property of having an asymptotic velocity to pressure constraint count of 2:1. The use of a discontinuous pressure is especially alluring since it gives the method a "local bias," for example, it leads to conservation of mass at an element level. In spite of its instability the raw Q_1–P_0 method is often used in practical computations without any stabilisation. Indeed, in terms of accuracy the velocity solutions are usually reasonable, and pressures often appear to be realistic after post-processing. The lack of inherent stability has to be overcome in our case; the crucial point is that if the method is not stabilised then the eigenvalues of the Schur complement $BA^{-1}B^t$ in (1.11) are *not* independent of h, (the LBB constant γ is $O(h)$), hence the performance of the PCR solver rapidly deteriorates as $h \to 0$. Such deterioration in convergence must also be expected with a nested iterative strategy based on the Schur complement.

As in the globally stabilised case, solution accuracy tends to deteriorate if the stabilisation parameter is not sufficiently large. For example, solving our test problem, oscillatory pressure solutions are evident if $\beta < 10^{-2}$; see [SK] for some related results. In contrast to the globally stabilised case, however, solution accuracy is retained in the limit of an arbitrarily large stabilisation parameter. This implies that there is more freedom when seeking to optimise the choice of stabilisation parameter to speed up the rate of convergence of the PCR solver. Of course, if the effect of the stabilisation is localised, then we might expect that varying β might have less effect on the iteration counts. This expectation is borne out by the results in Table 5.

TABLE 5

Number of PCR *iterations in the locally stabilised case.*

Grid	$\beta = 0.01$	$\beta = 0.1$	$\beta = 1$	$\beta = 10$	$\beta = 100$
2×4	17	17	17	17	16
4×8	73	55	55	59	65
8×16	202	142	156	152	147
16×32	*	347	375	329	263

The characteristic feature of locally stabilised mixed methods having discontinuous pressure is that the stabilisation matrix C is block diagonal. In the two-dimensional Q_1–P_0 case above the blocks are 4×4 matrices, each having eigenvalues 0,1,1,2 after diagonal scaling. The repeated eigenvalue structure of the operator C_h (after scaling) means that for large β the preconditioned Stokes matrix will have well clustered negative eigenvalues as illustrated by the spectrum plotted in Fig. 4.

Comparing the results in Table 5 with those of Table 3, we see that in both cases, the performance is poor if β is too small. In the locally stabilised case with $\beta = 0.01$

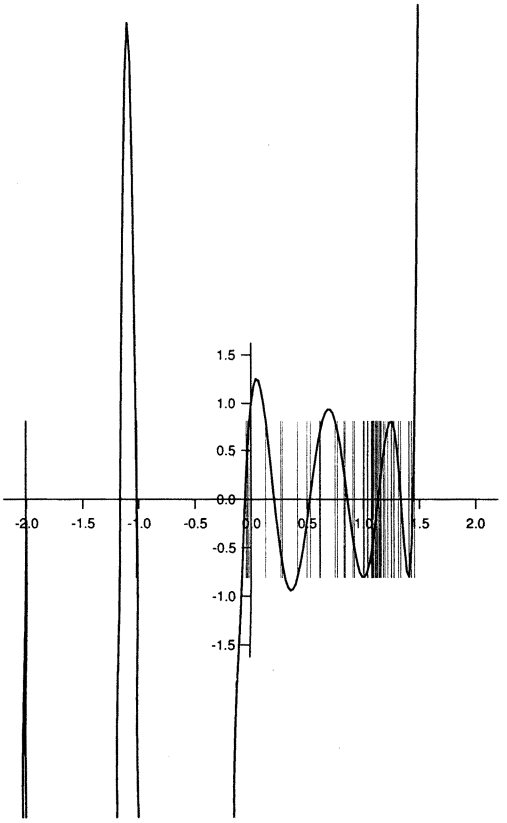

FIG. 4

the PCR algorithm breaks down on the finest grid (in the same way as when using the stable method with the "wrong" scaling). As discussed above this poor behaviour is explained by our eigenvalue theory of the last section. On the other hand, for a large stabilisation parameter the eigen-spectrum of the locally stabilised method is fundamentally different to that in the globally stabilised case, and this is reflected in the iteration counts. In the locally stabilised case, the clustering on the negative side of the spectrum can be exploited by the PCR algorithm and the nice behaviour of the iteration count growing like $O(h^{-1})$ is retained. This is in stark contrast to the globally stabilised case with large values of β. Note also that our analysis clearly shows that β must not be too large, otherwise the smallest negative eigenvalue μ_{-1} (i.e., the right-most cluster) will be close to the origin, which will certainly slow down convergence.

In Table 6 we show the variation of the extremal eigenvalues for the optimal choice of parameter from the table above, (i.e., with $\beta = 100$). Comparing with the values in Table 4, we see that the fundamental difference is the fact that with this value of β the eigenvalue μ_{-1} seems to be insensitive to h, that is, consistent with (3.28). For more moderate values of β (for example, $\beta = 1$) the eigenvalue μ_{-1} moves towards the origin like $O(h)$, so that the convergence behaviour is more like the mini-element case above.

Our conclusions from this are as follows: using an a priori stable method, convergence rates analogous to those which would be expected solving the diagonally scaled

TABLE 6
Eigenvalues of $\widetilde{A}$ in the locally stabilised case with parameter $\beta = 100$.

Grid	μ_{-m}	μ_{-1}	μ_1	μ_n
2×4	-2.0004	-0.0019	0.3983	1.2526
4×8	-2.0011	-0.0016	0.1135	1.4297
8×16	-2.0012	-0.0013	0.0309	1.4839
16×32	-2.0012	-0.0012	0.0091	1.4982

Laplacian can be obtained in the indefinite case, as long as the preconditioner (3.2) enforces the "correct" scaling. Using a globally stabilised method, good convergence rates can only be achieved by making the correct choice of stabilisation parameter. However, even when chosen optimally the iteration counts are likely to be asymptotically inferior to those which would be obtained using an a priori stable method as above. Using a locally stabilised method leads to reasonable rates of convergence and solution accuracy as long as the stabilisation parameter is not too small. Making the "optimal choice" in the locally stabilised case again gives the convergence behaviour obtained in the a priori stable case.

In part II of this work ([SW]), we extend our analysis to cover the case of more sophisticated preconditioners, for example, based on the Laplacian part of the Stokes operator.

Acknowledgments. This research was completed whilst the authors were visiting the Computer Science Department at Stanford University. We would like to thank Gene Golub for his continued encouragement and support. We are grateful to Alison Ramage for Figs. 2, 3, and 4.

REFERENCES

[A] J. ATANGA, *Iterative methods for locally stabilised mixed finite element equations*, Ph.D. thesis, University of Manchester Institute of Science and Technology, UK, 1992.

[ABF] D. ARNOLD, F. BREZZI, AND M. FORTIN, *A stable finite element for the Stokes equations*, Calcolo, 21 (1984), pp. 337–344.

[AMS] S. F. ASHBY, T. A. MANTEUFFEL, AND P. E. SAYLOR, *A taxonomy for conjugate gradient methods*, SIAM J. Numer. Anal., 27 (1990), pp. 1542–1568.

[AB] O. AXELSSON AND V. A. BARKER, *Finite Element Solution of Boundary Value Problems: Theory and Computation*, Academic Press, New York, 1984.

[BWY] R. E. BANK, B. D. WELFERT, AND H. YSERENTANT, *A class of iterative methods for solving saddle point problems*, Numer. Math., 56 (1990), pp. 645–666.

[BP] J. BRAMBLE AND J. PASCIAK, *A preconditioning technique for indefinite systems resulting from mixed approximations of elliptic problems*, Math. Comput., 50 (1988), pp. 1–17.

[BF] F. BREZZI AND M. FORTIN, *Mixed and Hybrid Finite Element Methods*, Springer-Verlag, New York, 1991.

[BPi] F. BREZZI AND J. PITKÄRANTA, *On the stabilisation of finite element approximations of the Stokes problem*, in Efficient Solutions of Elliptic Systems, Notes on Numerical Fluid Mechanics Vol. 10, W. Hackbusch, ed., Vieweg, Braunschweig, 1984, pp. 11–19.

[FHS] L. P. FRANCA, T. J. R. HUGHES, AND R. STENBERG, *Stabilised finite element methods for the Stokes problem*, Incompressible Computational Fluid Dynamics—Trends and Advances, R. A. Nicolaides and M. D. Gunzburger, eds., Cambridge University Press, Cambridge, 1993.

[Fr] R. FREUND, *On polynomial preconditioning for indefinite Hermitian matrices*, RIACS Tech. Report 89.32, RIACS, NASA Ames Research Center, Moffett Field, CA, 1989.

[F] I. FRIED, *The l_2 and l_∞ condition numbers of the finite element stiffness and mass matrices and the pointwise convergence of the method*, in The Mathematics of Finite Elements and Applications, J. R. Whiteman, ed., Academic Press, London, 1973, pp. 163–174.

[Gl] R. L. GLOWINSKI, *Numerical Methods for Nonlinear Variational Principles*, Springer-Verlag, New York, 1984.

[GVL] G. H. GOLUB AND C. F. VAN LOAN, *Matrix Computations*, John Hopkins University Press, Baltimore, MD, 1983.

[G] A. GREENBAUM, *Comparison of splittings used with the conjugate gradient algorithm*, Numer. Math., 33 (1979), pp. 181–193.

[HF] T. J. R. HUGHES AND L. P. FRANCA, *A new finite element formulation for* CFD: VII. *The Stokes problem with various well-posed boundary conditions: Symmetric formulations that converge for all velocity/pressure spaces*, Comput. Methods Appl. Mech. Engrg., 65 (1987), pp. 85–96.

[JY] W. D. JOUBERT AND D. M. YOUNG, *Necessary and sufficient conditions for the simplification of generalised conjugate gradient algorithms*, Linear Algebra Appl., 88/89 (1987), pp. 449–485.

[KS] N. KECHKAR AND D. J. SILVESTER, *Analysis of locally stabilised mixed finite element methods for the Stokes problem*, Math. Comp., 58 (1992), pp. 1–10.

[L] V. I. LEBEDEV, *Iterative methods for solving operator equations with a spectrum contained in several intervals*, USSR Comput. Math. and Math. Phys., 9 (1969), pp. 17–24.

[P] R. PIERRE, *Simple C^0 approximations for the computation of incompressible flows*, Comput. Methods Appl. Mech. Engrg., 68 (1988), pp. 205–227.

[PS] J. PITKÄRANTA AND T. SAARINEN, *A multigrid version of a simple finite element method for the Stokes problem*, Math. Comp., 45 (1985), pp. 1–14.

[R] A. RAMAGE, *Preconditioned conjugate gradient methods for Galerkin finite element equations*, Ph.D. thesis, University of Bristol, UK, 1991.

[SK] D. J. SILVESTER AND N. KECHKAR, *Stabilised bilinear-constant velocity-pressure finite elements for the conjugate gradient solution of the Stokes problem*, Comput. Methods Appl. Mech. Engrg., 79 (1990), pp. 71–86.

[SW] D. J. SILVESTER AND A. J. WATHEN, *Fast iterative solution of stabilised Stokes systems Part* II: *Using block preconditioners*, Report NA-92-06, Computer Science Department, Stanford University, Stanford, CA, 1992.

[SzW] D. B. SZYLD AND O. B. WIDLUND, *Variational analysis of some conjugate gradient methods*, East-West J. Numer. Math., 1 (1991), pp. 1–25.

[V] R. VERFURTH, *A combined conjugate gradient-multigrid algorithm for the numerical solution of the Stokes problem*, IMA J. Numer. Anal., 4 (1984), pp. 441–455.

[W] A. J. WATHEN, *Realistic eigenvalue bounds for the Galerkin Mass Matrix*, IMA J. Numer. Anal., 7 (1987), pp. 449–457.

[Wi] J.H. WILKINSON, *The Algebraic Eigenvalue Problem*, Oxford University Press, Oxford, 1965.