Eigenvalues of Block Matrices Arising From Problems in Fluid Mechanics

K. A. Cliffe,
T. J. Garratt and
A. Spence

CRPC-TR93333
1993

Center for Research on Parallel Computation
Rice University
P.O. Box 1892
Houston, TX 77251-1892

EIGENVALUES OF BLOCK MATRICES ARISING FROM PROBLEMS IN FLUID MECHANICS

K. A. CLIFFE*, T. J. GARRATT†, AND A. SPENCE‡

Abstract. Block matrices with a special structure arise from mixed finite element discretizations of incompressible flow problems. This paper is concerned with an analysis of the eigenvalue problem for such matrices, and the derivation of two shifted eigenvalue problems which are more suited to numerical solution by iterative algorithms like simultaneous iteration and Arnoldi's method. The application of the shifted eigenvalue problems to the determination of the eigenvalue of smallest real part is discussed and a numerical example arising from a stability analysis of double-diffusive convection is described.

Keywords. block matrices, eigenvalues, finite elements, Navier-Stokes

AMS (MOS) Subject Classification . 15A18, 65F15, 65F50, 76M10

1. Introduction. Let A and B be $N \times N$ real matrices with the following block structure

$$A = \begin{bmatrix} K & C \\ C^T & 0 \end{bmatrix} \quad B = \begin{bmatrix} M & 0 \\ 0 & 0 \end{bmatrix} \tag{1.1}$$

where $N = n+m$, $n > m$, K is $n \times n$, C is $n \times m$ of rank m, and M is $n \times n$ symmetric positive definite. The paper is concerned with the theory of the generalised eigenvalue problem

$$Aw = \mu Bw \tag{1.2}$$

that we call EVP1, and three related eigenvalue problems called EVP2, EVP3 and EVP4 that are introduced in Section 3. Since the matrices are typically large and sparse numerical techniques based on transformation methods, like the QZ algorithm, will be very expensive, and the reason for introducing the related eigenvalue problems EVP2, EVP3 and EVP4 is that they should be amenable to iterative techniques commonly used to find selected eigenvalues of large sparse matrices (Section I-5, [2]). In the applications we have mind K, C, and M arise from mixed finite element discretizations of the "velocity-pressure" formulation of the Navier-Stokes equations for incompressible flow problems [3], [9], and the eigenvalue problem (1.2) arises in the determination of the stability of steady flows [5]. The problem is to find the eigenvalues of (1.2) with smallest real part.

As is standard, the finite values $\mu \in C$ such that $\det(A - \mu B) = 0$ are known as finite eigenvalues, though we usually drop the term "finite". Since B is singular there are also infinite eigenvalues, which are defined to be zero eigenvalues of $\nu Aw = Bw$, with corresponding eigenvectors that are null vectors of B. The theory for (1.2) is more complicated than for the standard eigenvalue problem $Aw = \mu w$ [21]. However, the assumptions made in this paper on C and M allow very precise statements to be made about the number of eigenvalues of (1.2) and make possible the introduction of related eigenvalue problems which are better suited to solution by iterative algorithms.

* Theoretical Studies Department, B424.4, Harwell Laboratory, Didcot, OX11 0RA, UK.,
† AspenTech UK Ltd, Sheraton House, Castle Park, Cambridge, CB3 0AX.
‡ School of Mathematical Sciences, University of Bath, Bath, BA2 7AY.

Eigenvalue problems of the form (1.2) with block structure (1.1) arise in applications involving constraints. For example, Malkus [14] discusses the case when K is symmetric in an analysis of the discrete LBB stability condition for incompressible finite elements, arising in linear elasticity or Stokes flow. For symmetric K, the results on the eigenvalues of EVP1 in Theorem 1 and Lemma 1 are contained in ([14], Theorem 3) though our method of proof produces the results in a more direct manner. In addition, the case of K symmetric and $M = I$ is discussed in Golub [10].

The plan of the paper is as follows. In Section 2 the basic theory for the eigenvalue problem (1.1)(1.2) is presented. Section 3 contains an analysis of some related eigenvalue problems which provide some shift strategies for the eigenvalues. In Section 4 some practical aspects are considered. First there is a discussion relating to the execution of certain matrix-vector operations, and second the estimation of the accuracy of computed eigensolutions is examined. Section 5 contains a discussion of strategies which could be used to determine the eigenvalues of smallest real part of (1.1)(1.2), and these are illustrated with reference to a matrix problem arising from a finite element discretization of 2–dimensional double–diffusive convection in a box.

2. Theory for the eigenvalue problem. This section contains some results about the eigenvalue problem $Aw = \mu Bw$ which for convenience we rewrite (and rename) as

$$\begin{bmatrix} K & C \\ C^T & 0 \end{bmatrix} \begin{bmatrix} u \\ p \end{bmatrix} = \mu \begin{bmatrix} M & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} u \\ p \end{bmatrix} \qquad \text{(EVP1)}$$

where, in anology with our applications arising from the discretization of the Navier-Stokes equations, we use the notation $w = [u, p]$, $u \in \boldsymbol{R}^n$, $p \in \boldsymbol{R}^m$, where u and p correspond to velocity and pressure degrees of freedom respectively.

First we note that since C is full rank, the QR factorisation of C has the form (ignoring possible permutations, which play no role here)

$$C = QR = [Q_1\ Q_2] \begin{bmatrix} R_1 \\ 0 \end{bmatrix} \ (= Q_1 R_1) \qquad (2.1)$$

where R is $n \times m$, R_1 is $m \times m$ nonsingular and upper triangular, Q is $n \times n$ orthogonal, Q_1 is $n \times m$ and provides an orthonormal basis for range(C), and Q_2 is $n \times (n-m)$ and provides an orthonormal basis for $C^\perp$. For future use note that C^TC is $m \times m$ positive definite and has the Cholesky decomposition $C^TC = R_1^T R_1$ ([11] p. 217). Also, the matrix

$$\pi := I - C(C^TC)^{-1}C^T \qquad (2.2)$$

is a projection from $\boldsymbol{R}^n$ onto $C^\perp$ along range(C).

Now we state a fundamental result on the number of eigenvalues of EVP1.

THEOREM 2.1. *a) The eigenvalue problem EVP1 has precisely $n-m$ eigenvalues, that are those of the reduced eigenvalue problem of dimension $(n-m)$*

$$Q_2^T(K - \mu M)Q_2 z = 0. \qquad \text{(REVP1)}$$

b) If (μ, z), $z \in \boldsymbol{R}^{n-m}$ is an eigensolution of REVP1 then (μ, u, p) is a corresponding eigensolution of EVP1 where

$$u = Q_2 z, \quad p = -R_1^{-1}Q_1^T(K - \mu M)Q_2 z.$$

c) If (μ, u, p), $u \in \boldsymbol{R}^n$, $p \in \boldsymbol{R}^m$ is an eigensolution of EVP1 then (μ, z) is a corresponding eigensolution of REVP1 where $z = Q_2^T u$.

Proof. a) Introduce $Z = \begin{bmatrix} Q & 0 \\ 0 & I_m \end{bmatrix}$ where Q is defined in (2.1), and $y := Z^T w = (Q^T u, p) = (Q_1^T u, Q_2^T u, p) =: (u_1, u_2, p)$. Now EVP1 is equivalent to $Z^T A Z y = \mu Z^T B Z y$, that in block form becomes

$$\begin{bmatrix} K_{11} & K_{12} & R_1 \\ K_{21} & K_{22} & 0 \\ R_1^T & 0 & 0 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ p \end{bmatrix} = \mu \begin{bmatrix} M_{11} & M_{12} & 0 \\ M_{21} & M_{22} & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ p \end{bmatrix} \tag{2.3}$$

where $K_{ij} = Q_i^T K Q_j$, $M_{ij} = Q_i^T M Q_j$, $i, j = 1, 2$. Simple manipulation shows that $u_1 = 0$, $(K_{22} - \mu M_{22}) u_2 = 0$, $p = -R_1^{-1}(K_{12} - \mu M_{12}) u_2$. Since M and M_{22} are symmetric positive definite a) is immediate. Results b) and c) now follow. □

REMARK 1. An equivalent proof in the style of Golub [10] uses the projection π defined by (2.2). Since $\pi u = u$ for $u \in C^{\perp}$ we may write the first row of EVP1 as $K \pi u + C p = \mu M \pi u$. Premultiplication by π gives

$$\pi K \pi u = \mu \pi M \pi u \tag{2.4}$$

that has the same eigenvalues as (REVP1) plus m zero eigenvalues corresponding to eigenvectors lying in range(C) which have no relevance for EVP1.

For the future analysis it is convenient to exclude the possibility that $\mu = 0$ is an eigenvalue and so we assume the following:

$$\mu = 0 \text{ is } \textit{not} \text{ an eigenvalue of EVP1.} \tag{2.5}$$

This assumption is not a severe restriction in the applications we have in mind, since though zero eigenvalues are important, corresponding to "steady-state" bifurcations in some nonlinear problem, they are usually detected readily. For example, if a direct solver for A is feasible then one can check the determinant of A. If a zero eigenvalue of EVP1 has been found, an eigenvalue problem satisfying (2.5) may be obtained by considering a "shifted" eigenvalue problem with the same structure as EVP1 but with K replaced by $K - \gamma M$ for some appropriately chosen shift $\gamma \in \boldsymbol{R}$.

Under assumption (2.5), K_{22} is nonsingular and so (2.3) may be rewritten as

$$\nu w = \begin{bmatrix} 0 & 0 & R_1^{-T} \\ 0 & K_{22}^{-1} & * \\ R_1^{-1} & * & * \end{bmatrix} \begin{bmatrix} M_{11} & M_{12} & 0 \\ M_{21} & M_{22} & 0 \\ 0 & 0 & 0 \end{bmatrix} w$$

or

$$\nu w = \begin{bmatrix} 0 & 0 & 0 \\ * & K_{22}^{-1} M_{22} & 0 \\ * & * & 0 \end{bmatrix} w \tag{2.6}$$

where $*$ denotes unique submatrices which do not effect the analysis. Clearly (2.6) has a zero eigenvalue of algebraic multiplicity $2m$. Thus we have proved

LEMMA 2.2. *Under (2.5), EVP1 has an infinite eigenvalue of multiplicity $2m$.*

Similar results apply to the following generalisation of EVP1

$$\begin{bmatrix} K & C_1 \\ C_2^T & 0 \end{bmatrix} \begin{bmatrix} u \\ p \end{bmatrix} = \mu \begin{bmatrix} M & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} u \\ p \end{bmatrix} \tag{2.7}$$

where C_1 and C_2 are $n \times m$ matrices of rank m. Such a problem arises after the discretization of the Navier-Stokes equations by a spectral method [6]. Provided C_1 and C_2 satisfy the nondegeneracy condition

$$C_2^T C_1 \text{ is invertible}$$

(which is a natural condition in the context of discretizations of the Navier-Stokes equations) then results analogous to Theorem 1 hold. To see this we introduce the projection operator

$$\pi_{12} := I - C_1(C_2^T C_1)^{-1} C_2^T \tag{2.8}$$

(cf (2.2)): Clearly if $u \in \text{range}(C_1)$ then $\pi_{12}u = 0$, and if $u \in C_2^{\perp}$, then $\pi_{12}u = u$.

Now follow the approach of Remark 1 to obtain

$$\pi_{12} K \pi_{12} u = \mu \pi_{12} M \pi_{12} u \tag{2.9}$$

(cf. (2.4)) which has $n - m$ eigenvalues for $u \in C_2^{\perp}$. A corresponding reduced eigenvalue problem can be derived (cf. Theorem 1 a)).

REMARK 2. It is important to note that if K and M are large and sparse (as is the case in the applications we have in mind) then typically one would not explicitly form the matrices in (REVP1). This is because Q_2 is full and hence $Q_2^T K Q_2$ and $Q_2^T M Q_2$ are full. Rather one would employ iterative techniques to find selected eigenvalues as discussed in Section 5.

3. Some Shifted Eigenvalue Problems. It is a common technique to shift the eigenvalues of an eigenvalue problem, so that if $Aw = \mu Bw$ then $(A - \gamma B)w = (\mu - \gamma)Bw$ and all the eigenvalues μ are shifted by γ, with the corresponding eigenvector remaining unchanged. In this section we look at some generalised eigenvalue problems that are closely related to EVP1 in that the new eigenvalue problems allow us to shift both the finite and infinite eigenvalues. First consider the eigenvalue problem

$$\begin{bmatrix} K - \gamma M & \delta_1 C \\ \delta_1 C^T & 0 \end{bmatrix} v = \sigma \begin{bmatrix} M & \delta_2 C \\ \delta_2 C^T & 0 \end{bmatrix} v \tag{EVP2}$$

for some $\delta_1, \delta_2, \gamma \in \boldsymbol{R}$. Note that $\gamma = \delta_2 = 0, \delta_1 = 1$ recovers EVP1. We have the following theorem about the eigenvalues of EVP2.

THEOREM 3.1. *Denote the finite eigenvalues of REVP1 by $\mu_i, i = 1, ..., n - m$. Assume (2.5) and*

$$i)\ \delta_2 \neq 0, \quad ii)\ \delta_1 \delta_2^{-1} \neq \mu_i - \gamma. \tag{3.1}$$

Then EVP2 has eigenvalues σ_i, $i = 1, ..., n + m$ with

a) $\sigma_i = \mu_i - \gamma$, $i = 1, ..., n - m$

b) $\sigma_i = \delta_1 \delta_2^{-1}$, $i = n - m + 1, ..., n + m$.

Proof. Since M is positive definite it is straightforward to show, under (3.1.i), that the matrix on the right hand side of EVP2 is nonsingular and hence EVP2 has $n+m$ eigenvalues. Under the transformation introduced in the proof of Theorem 1, EVP2 is equivalent to

$$\begin{bmatrix} K_{11}-\gamma M_{11} & K_{12}-\gamma M_{12} & \delta_1 R_1 \\ K_{21}-\gamma M_{21} & K_{22}-\gamma M_{22} & 0 \\ \delta_1 R_1^T & 0 & 0 \end{bmatrix} v = \sigma \begin{bmatrix} M_{11} & M_{12} & \delta_2 R_1 \\ M_{21} & M_{22} & 0 \\ \delta_2 R_1^T & 0 & 0 \end{bmatrix} v$$

and premultiplication by the inverse of the matrix on the right hand side produces the eigenvalue problem

$$\begin{bmatrix} \delta_1\delta_2^{-1} & 0 & 0 \\ * & M_{22}^{-1}(K_{22}-\gamma M_{22}) & 0 \\ * & * & \delta_1\delta_2^{-1} \end{bmatrix} v = \sigma v$$

(where again $*$ denotes unique submatrices). Clearly this matrix has an eigenvalue at $\delta_1\delta_2^{-1}$ of multiplicity $2m$ and $n-m$ eigenvalues satisfying $(K_{22}-\gamma M_{22})z = \sigma M_{22} z$ which is precisely a simple shift of REVP1. Statements a) and b) now follow. □

Thus we see that the finite eigenvalues μ_i of EVP1 are shifted by γ, but the infinite eigenvalues are transformed to $\delta_1\delta_2^{-1}$. Thus any eigenvalue of EVP1 may be shifted to any location. Two examples of EVP2 which are the most likely to be useful are now introduced. Consider

$$\begin{bmatrix} K-\gamma M & 0 \\ 0 & 0 \end{bmatrix} v = \sigma \begin{bmatrix} M & C \\ C^T & 0 \end{bmatrix} v \qquad \text{(EVP3)}$$

Clearly this has $n-m$ eigenvalues $\sigma_i = \mu_i - \gamma, i = 1, ..., n-m$, and $2m$ eigenvalues at zero. Similarly the eigenvalue problem

$$\begin{bmatrix} K & \delta_1 C \\ \delta_1 C^T & 0 \end{bmatrix} v = \sigma \begin{bmatrix} M & C \\ C^T & 0 \end{bmatrix} \qquad \text{(EVP4)}$$

leaves the finite eigenvalues μ_i of EVP1 unchanged, but transforms the infinite eigenvalues to δ_1. Note however that not all eigenvectors are preserved by the shifts. Eigenvectors of EVP1 corresponding to the infinite eigenvalues have the form $(0, w_2)$ and these are unchanged. However the eigenvectors of the finite eigenvalues of EVP1 are changed, but only in the last m components. To be precise, we have the following result which may be deduced from the block form of the equations for EVP1 and EVP2.

LEMMA 3.2. *a) Let μ be a finite eigenvalue of EVP1. Assume (2.5) and (3.1). If $(w_1, w_2), w_1 \in \boldsymbol{R}^n, w_2 \in \boldsymbol{R}^m$ is an eigenvector of EVP1 associated with μ, then (v_1, v_2), where $v_1 = w_1, v_2 = (\delta_1 - (\mu-\gamma)\delta_2)^{-1} w_2$ is the corresponding eigenvector of EVP2.*

b) Let $(0, w_2), w_2 \in \boldsymbol{R}^m$ be an eigenvector corresponding to an infinite eigenvalue of EVP1. Then $(0, v_2)$ with $v_2 = w_2$ is an eigenvector of EVP2 corresponding to the eigenvalue $\delta_1\delta_2^{-1}$.

REMARK 3. We have not gone into the detailed structure of the infinite eigenvalue of EVP1 or the eigenvalue $\delta_1\delta_2^{-1}$ of EVP2 of algebraic multiplicity $2m$. However analysis using the Weierstrass-Kronecker canonical form reveals that $\delta_1\delta_2^{-1}$ has geometric

multiplicity m and that the eigenvalues occur in 2×2 pairs corresponding to Jordan blocks of the form $\begin{bmatrix} \delta_1\delta_2^{-1} & 1 \\ 0 & \delta_1\delta_2^{-1} \end{bmatrix}$. This is to be expected following the results in [14].

4. Computational Considerations. For K, C and M small and/or full it will probably be a reasonable strategy to form $K_{22} = Q_2^T K Q_2$ and $M_{22} = Q_2^T M Q_2$ and solve REVP1 directly. However, if K, C, and M are large and sparse the 'reduced' matrices K_{22} and M_{22} will be full, and transformation methods will probably not be feasible. Iterative methods, like simultaneous iteration or Arnoldi's method (probably applied to the EVP3 or EVP4) become attractive. It is therefore necessary to implement matrix $\times$ vector operations, which for the generalised eigenvalue problem reduce to solving systems of $(n+m)$ dimensional linear equations. These could be carried out directly on the $(n+m)$ dimensional systems [5], [7], [12], but an approach which involves only solving n dimensional systems is possible.

This is illustrated with reference to EVP3 with $\gamma = 0$. An iterative method will require the solution of a linear system of the form

$$\begin{bmatrix} M & C \\ C^T & 0 \end{bmatrix} \begin{bmatrix} u \\ p \end{bmatrix} = \begin{bmatrix} K & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} v \\ q \end{bmatrix}, \qquad v \text{ given}, \tag{4.1}$$

using an n dimensional system with coefficient matrix M.

It is readily shown (say by a block Gaussian elimination approach) that (4.1) can be solved by the following algorithm:

(i) solve $Mw = Kv$ for $w \in \boldsymbol{R}^n$
(ii) solve $[C^T M^{-1} C]p = C^T w$ for $p \in \boldsymbol{R}^m$
(iii) solve $Mx = Cp$ for $x \in \boldsymbol{R}^n$
(iv) set $u = w - x$.

This is the Uzawa algorithm [1] and in order to be efficient step (ii) would be carried out iteratively to avoid the direct computation of $C^T M^{-1} C$. For example, one could precondition $C^T M^{-1} C$ by $C^T M_l^{-1} C$, where M_l is the "lumped mass" matrix derived from M.

"Shift–invert" [16] or Cayley transform [7], [12] techniques are also possible, with iterative methods, like subspace iteration, requiring the solution of *nonsymmetric* systems of the form

$$\begin{bmatrix} K - \gamma M & C \\ C^T & 0 \end{bmatrix} \begin{bmatrix} u \\ p \end{bmatrix} = \begin{bmatrix} \nu \\ 0 \end{bmatrix} \tag{4.2}$$

for appropriate γ, and for 3–D pdes this system will invariably have to be solved iteratively. However, if good estimates of the wanted eigenvalues are available and if direct solution of (4.2) is possible, then these approaches are likely to prove very efficient.

A standard approach to estimate the accuracy of an approximate eigenpair $(\tilde{\mu}, \tilde{w})$ of $Aw = \mu w$ is to calculate the "residual" vector $r := A\tilde{w} - \tilde{\mu}\tilde{w}$. Standard backward error analysis (p171 of [23]) shows that typically a "small" residual indicates that $\tilde{\mu}$ is a "good" approximation to μ. A similar analysis holds for problems considered here. Let $(\tilde{\mu}, [\tilde{u}, \tilde{p}])$ denote an approximate eigenpair with $C^T \tilde{u} = 0$. Then the corresponding residual vector $r \in \mathbf{R}^{n+m}$ has the form

$$r = (r_1, 0), \quad r_1 := K\tilde{u} + C\tilde{p} - \tilde{\mu} M \tilde{u} \in \mathbf{R}^n. \tag{4.3}$$

With $\tilde{u}^T\tilde{u} = 1$ it is readily shown that

$$\begin{bmatrix} K - r_1\tilde{u}^T & C \\ C^T & 0 \end{bmatrix} \begin{bmatrix} \tilde{u} \\ \tilde{p} \end{bmatrix} = \tilde{\mu} \begin{bmatrix} M & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \tilde{u} \\ \tilde{p} \end{bmatrix}$$

i.e. $(\tilde{\mu}, [\tilde{u}, \tilde{p}])$ is an exact eigenpair of a perturbed problem (with the same block structure as (1.1), (1.2)). In Section 5 we use $\|r_1\|_2 (= \|r\|_2)$ to test the accuracy of an approximate eigenpair (cf. p383 of [11]).

5. Applications. The eigenvalue problem arises in the determination of the the stability of steady solutions to the Navier-Stokes and related equations, and linearised stability theory [8], [19] shows that stability is determined by the eigenvalues of smallest real part of a linearised problem. The eigenvalue problem EVP1 arises if a mixed finite element method is used to discretize the linearised problem [5], [9], [15]. Of special interest is the case when the eigenvalues of smallest real part are complex, since algorithms for the detection of Hopf bifurcations in parameter dependent systems can be developed from knowledge of these eigenvalues. The matrices K, C and M are sparse and very large - in [12] a problem with over 2×10^5 degrees of freedom is studied.

Since EVP1 involves large sparse nonsymmetric matrices one has to fall back on iterative methods like Arnoldi's method or simultaneous iteration [2] to compute wanted eigenvalues. As mentioned in Section 4 shift–invert strategies are possible if good estimates of wanted eigenvalues are known. However one can also apply iterative methods directly to EVP3 or EVP4 with appropriate choices for γ or δ_1, δ_2 respectively. To be precise two approaches are:

a) Choose $\gamma \approx \{Re(\mu_1) + Re(\mu_{n-m})\}/2$ in EVP3. Thus with $\rho_i = \mu_i - \gamma$, we have $Re(\rho_1) + Re(\rho_{n-m}) \approx 0$. The zero eigenvalue of multiplicity $2m$ of EVP3 is "in the middle" of the spectrum.

b) Choose $\delta_1 \approx \{Re(\mu_1) + Re(\mu_{n-m})\}/2$ in EVP4. Thus the eigenvalue of multiplicity $2m$ of EVP2 is "in the middle" of the spectrum of EVP4, with $\mu_1 \ldots \mu_{n-m}$ being unchanged.

In both EVP3 and EVP4 the troublesome $2m$ multiple eigenvalue, which corresponds to the infinite eigenvalue of EVP1 (Lemma 1), will not be computed by the iterative algorithm. We note that the recent implicit polynomial filters algorithm of Sorensen [20] would appear to be an appropriate method to apply to reformulations like EVP2.

If the eigenvalues of EVP1 are known to be real then simple shift strategies based on EVP3 and EVP4 allied to iterative methods will provide the largest and smallest eigenvalues. However when the eigenvalues may be complex then one may have to further transform the matrix eigenvalue problem. One approach is to utilize the Chebyshev transformation ideas of Saad [17],[18]. For a standard eigenvalue problem $Aw = \mu w$, the idea is to carry out a shifted Chebyshev polynomial transformation of A, say to $p_s(A)$, where the eigenvalues μ_i of A lying inside a certain ellipse in the complex plane are mapped to eigenvalues $p_s(\mu_i)$ of $p_s(A)$ satisfying $|p_s(\mu_i)| < 1$. The aim is to choose the polynomial p_s (and hence the ellipse) so that only the desired eigenvalues of A lie outside the ellipse. These become dominant, well separated, eigenvalues of $p_s(A)$ and hence are computed by an iterative solver applied to $p_s(A)$. These techniques were applied successfully to EVP3 to find eigenvalues of smallest real part of two problems from fluid dynamics and the results were reported in [5] and

[7]. Transformation EVP4 was used to provide the numerical results in the following example.

5.1. Example. We consider a matrix arising from a mixed finite element discretization of the equations modelling two dimensional double-diffusive convection in a box heated on the bottom boundary (see Chapter 8 of [22]). The governing equations are solved in the Boussinesq approximation and are given in [4]. We do not reproduce these equations here but note that the model has nondimensional parameters: Prandtl number Pr, Rayleigh number Ra, salinity Rayleigh number Rs and τ (see [4] p.254), and interest centres on the loss of stability as Ra increases (which corresponds to increasing the temperature difference between the top and bottom boundary). Our calculations were performed with $Pr = 10$ and $\tau = 10^{-2}$, which corresponds roughly to a salt solution and water, $Rs = 2000$ and $Ra = 2480$. The exact eigenvalues of the continuous problem, μ_i^e say, are known ((8.18) in [22]) and the three leftmost eigenvalues are $\mu_1^e, \mu_2^e = 0.047486 \pm 24.502i$, $\mu_3^e = 0.098696$. A mixed finite element approximation was obtained in the usual way [13], using nine-node quadrilateral elements with biquadratic interpolation for velocities, temperatures and salinities and discontinuous piecewise-linear interpolation for pressures [4]. The matrix was set up using ENTWIFE [24] with a 4×4 grid which leads to a matrix with the block structure of (1.1) with $n = 324$, $m = 48$ and hence $N(= n+m) = 372$. The three leftmost eigenvalues of the matrix problem to seven significant figures are $\mu_1, \mu_2 = 0.04932671 \pm 24.51725i$ $\mu_3 = 0.09874659$. (In fact numerical values are known with residuals (see (4.2)) less than 0.25×10^{-13}.)

In the following "Arnoldi (k, l)" means Arnoldi's method restarted l times with subspace of dimension k [18]. First Arnoldi (20,1) with a random starting vector was applied to $A_1 := \begin{bmatrix} M & C \\ C^T & 0 \end{bmatrix}^{-1} \begin{bmatrix} K & 0 \\ 0 & 0 \end{bmatrix}$ to obtain a very rough idea of the convex hull of the eigenvalues of EVP2 with $\gamma = 0$, $\delta_1 = 0$ and $\delta_2 = 1$, and hence a rough estimate of μ_{n-m} is obtained. (Of course, the matrix $\times$ vector operations with A_1 were performed by solving linear systems.) Next form EVP4 with $\delta_2 = 1$, $\delta_1 = Re(\mu_{n-m})/2$, in line with strategy b) above, and try to find the leftmost eigenvalues of $A_2 := \begin{bmatrix} M & C \\ C^T & 0 \end{bmatrix}^{-1} \begin{bmatrix} K & \delta_1 C \\ \delta_1 C^T & 0 \end{bmatrix}$. Arnoldi (20,50) failed to find any of the 3 leftmost eigenvalues, perhaps because of severe clustering of the eigenvalues ([18]). The hybrid Algorithm of Saad [18] (see above) utilizing the Chebyshev transformation to find the two leftmost eigenvalues (so that the ellipse passes through μ_3) is however successful. A two step procedure was used:

(i) Arnoldi (20,1) with a random starting vector was applied to $p_5(A_2)$ to obtain a "purified" starting vector,

(ii) Arnoldi (20,1) was applied to $p_{42}(A_2)$. This computed μ_1, μ_2 with residuals less than 5×10^{-12}, (see (4.3)).

Numerical experiments using EVP3 with γ chosen in a) above produce similar results. This is not surprising since the distribution of the extremal eigenvalues of EVP3 and EVP4 is the same.

Acknowledgements. The research of TJG was supported by an SERC CASE studentship at Bath University and AEA Technology, Harwell; that of KAC forms part of the Underlying Corporate Research Programme of AEA; and that of AS was

supported by SERC and the NSF under Cooperative Agreement No. CCR-8809615. Y. Saad kindly sent us a copy of his hybrid algorithm [18] which was used to carry out the computations in Section 5.

REFERENCES

[1] K. ARROW, L. HURWICZ AND H. UZAWA, *Studies in nonlinear programming*, Stanford University Press, Stanford (1958).

[2] F. CHATELIN, *Spectral approximation of linear operators*, Academic Press, New York (1983).

[3] K.A. CLIFFE, *Numerical calculations of two-cell and single-cell Taylor flows*, J. Fluid Mech., 135 (1983), pp. 219-233.

[4] K.A. CLIFFE AND K.H. WINTERS, *Convergence properties of the finite-element method for Bénard convection in an infinite layer*, J. Comp. Physics, 60, (1985), pp. 346-351.

[5] K.A. CLIFFE, T.J. GARRATT AND A. SPENCE, *Calculation of eigenvalues of the discretised Navier-Stokes and related equations*, in The mathematics of finite elements and applications VII MAFELAP, J.R. Whiteman, ed., Academic Press, 1990, pp. 470-486.

[6] A.J. CONLEY, *Spectral methods for 3-D Navier-Stokes equations*, PhD Thesis, Caltech, 1992.

[7] T.J. GARRATT, *The numerical detection of Hopf bifurcations in large systems arising in fluid mechanics*, PhD Thesis, School of Mathematical Sciences, University of Bath, 1991.

[8] A. GEORGESCU, *Hydrodynamic stability theory*, Martinus Nijhoff, Dordrecht, Netherlands, 1985.

[9] V. GIRAULT AND P.A. RAVIART, *Finite element approximation of the Navier-Stokes equations*, Lecture Notes in Mathematics, Springer Verlag, Berlin, 1979.

[10] G.H. GOLUB, *Some modified matrix eigenvalue problems*, SIAM Rev., 15, (1973), pp. 318-334.

[11] G.H. GOLUB AND C.F. VAN LOAN, *Matrix Computations*, The Johns Hopkins University Press, Baltimore, Maryland, 1983.

[12] P.M. GRESHO, D.K. GARTLING, J.R. TORCZYNSKI, T.J. GARRATT, K.A. CLIFFE, A. SPENCE, K.H. WINTERS AND J.W. GOODRICH, *Is the steady viscous incompressible 2D flow over a backward-facing step at Re = 800 stable?*, Int. J. Num. Meth. Fluids, (1993), to appear.

[13] C. JOHNSON, *Numerical solutions of partial differential equations by the finite element method*, Cambridge University Press, Cambridge, 1987.

[14] D.S. MALKUS, *Eigenproblems associated with the discrete LBB condition for incompressible finite elements*, Int. J. for Eng. Sci, 19, (1981), pp. 1299-1310.

[15] B. MERCIER, J. OSBORN, J. RAPPAZ AND P.A. RAVIART, *Eigenvalue approximation by mixed and hybrid methods*, Mathematics of Computation, 36, number 154, (1981), pp. 427-453.

[16] B.N. PARLET AND Y. SAAD, *Complex shift and invert strategies for real matrices*, Lin. Alg. and Applic, 88/89, (1987), pp. 575-595.

[17] Y. SAAD, *Practical use of some Krylov subspace methods for solving indefinite and unsymmetric linear equations*, SIAM J. Sci. Stat. Comput., 5, (1984), pp. 203-228.

[18] Y. SAAD, *Chebyshev acceleration techniques for solving nonsymmetric eigenvalue problems*, Math. Comp., 42, (1984), pp. 567-588.

[19] D.H. SATTINGER, *Transformation groups and bifurcation at multiple eigenvalues*, Bull. Amer. Math. Soc., 79, (1973), pp. 709-711.

[20] D.G. SORENSEN, *Implicit application of polynomial filters in a k-step Arnoldi method*, SIAM J. Matrix Anal. Appl., 13, (1992), pp. 357-385.

[21] G.W. STEWART, *Perturbation Theory for the Generalised Eigenvalue Problem*, in Advances in Numerical Analysis, G.H. Golub and C. de Boor, eds., Academic Press, 1978.

[22] J.S. TURNER, *Buoyancy effects in fluids*, CUP, Cambridge, 1973.

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

[24] K.H. WINTERS, *ENTWIFE user manual (release 1)*, Harwell report AERE-R 11577, 1985.