SIAM J. MATRIX ANAL. APPL.
Vol. 26, No. 1, pp. 20–41

A PRECONDITIONER FOR GENERALIZED SADDLE POINT PROBLEMS*

MICHELE BENZI† AND GENE H. GOLUB‡

Abstract. In this paper we consider the solution of linear systems of saddle point type by preconditioned Krylov subspace methods. A preconditioning strategy based on the symmetric/skew-symmetric splitting of the coefficient matrix is proposed, and some useful properties of the preconditioned matrix are established. The potential of this approach is illustrated by numerical experiments with matrices from various application areas.

Key words. saddle point problems, matrix splittings, iterative methods, preconditioning

AMS subject classifications. Primary, 65F10, 65N22, 65F50; Secondary, 15A06

DOI. 10.1137/S0895479802417106

1. Introduction.

We consider the solution of systems of linear equations with the following block 2×2 structure:

$$\begin{bmatrix} A & B^T \\ B & -C \end{bmatrix} \begin{bmatrix} u \\ p \end{bmatrix} = \begin{bmatrix} f \\ g \end{bmatrix}, \tag{1.1}$$

where $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{m\times n}$, $C \in \mathbb{R}^{m\times m}$, $f \in \mathbb{R}^n$, $g \in \mathbb{R}^m$, and $m \le n$. We further assume that matrices A, B, and C are large and sparse. Systems of the form (1.1) arise in a variety of scientific and engineering applications, including computational fluid dynamics [1, 20, 22, 24, 27, 44], mixed finite element approximation of elliptic PDEs [12, 48, 60], optimization [5, 25, 26, 32, 39, 43], optimal control [9, 35], weighted and equality constrained least squares estimation [10], structural analysis [56], electrical networks [56], inversion of geophysical data [34], computer graphics [42], and others.

An important special case of (1.1) is when A is symmetric positive semidefinite, $C = O$, $\mathrm{rank}(B) = m$, and $\mathcal{N}(A) \cap \mathcal{N}(B) = \{0\}$. In this case (1.1) corresponds to a saddle point problem, and it has a unique solution.

In this paper we consider *generalized* saddle point problems, i.e., systems of the form (1.1) satisfying all of the following assumptions:

- A has positive semidefinite symmetric part $H = \frac{1}{2}(A + A^T)$;
- $\mathrm{rank}(B) = m$;
- $\mathcal{N}(H) \cap \mathcal{N}(B) = \{0\}$;
- C is symmetric positive semidefinite.

As shown below (Lemma 1.1), these assumptions guarantee existence and uniqueness of the solution. Although very often A is symmetric positive definite, we are especially interested in cases where A is either symmetric and singular (i.e., only positive semidefinite), or nonsymmetric with positive definite symmetric part H (i.e., A is *positive real*). The latter situation arises when the steady-state Navier–Stokes

*Received by the editors October 30, 2002; accepted for publication (in revised form) by A. J. Wathen October 27, 2003; published electronically August 6, 2004.
http://www.siam.org/journals/simax/26-1/41710.html

†Department of Mathematics and Computer Science, Emory University, Atlanta, GA 30322 (benzi@mathcs.emory.edu). The work of this author was supported in part by National Science Foundation grant DMS-0207599.

‡Scientific Computing and Computational Mathematics Program, Stanford University, Stanford, CA 94305-9025 (golub@sccm.stanford.edu). The work of this author was supported in part by Department of Energy grant DOE-FC02-01ER4177.

equations are linearized by a Picard iteration, leading to the *Oseen equations*; see [20, 22]. In this case, the A block corresponds to an appropriate discretization of a convection-diffusion operator.

A number of solution methods have been proposed in the literature. Besides specialized sparse direct solvers [16, 17] we mention, among others, Uzawa-type schemes [11, 21, 24, 27, 62], block and approximate Schur complement preconditioners [4, 15, 20, 22, 41, 45, 46, 48, 51], splitting methods [18, 30, 31, 49, 57], indefinite preconditioning [23, 35, 39, 43, 48], iterative projection methods [5], iterative null space methods [1, 32, 54], and preconditioning methods based on approximate factorization of the coefficient matrix [25, 50]. Several of these algorithms are based on some form of reduction to a smaller system, for example, by projecting the problem onto the null space of B, while others work with the original (augmented) matrix in (1.1). The method studied in this paper falls in the second category.

When A is symmetric positive (semi-)definite, the coefficient matrix in (1.1) is symmetric indefinite, and indefinite solvers can be used to solve problem (1.1). Alternatively, one can solve instead of (1.1) the equivalent nonsymmetric system

$$\begin{bmatrix} A & B^T \\ -B & C \end{bmatrix} \begin{bmatrix} u \\ p \end{bmatrix} = \begin{bmatrix} f \\ -g \end{bmatrix}, \quad \text{or} \quad \mathcal{A}\mathbf{x} = \mathbf{b}, \tag{1.2}$$

where $\mathcal{A}$ is the coefficient matrix in (1.2), $\mathbf{x} = [u^T, p^T]^T$ and $\mathbf{b} = [f^T, -g^T]^T$. The nonsymmetric formulation is especially natural when A is nonsymmetric, but positive real. Whether A is symmetric or not, the nonsymmetric matrix $\mathcal{A}$ has certain desirable properties, which are summarized in the following result.

LEMMA 1.1. *Let $\mathcal{A} \in \mathbb{R}^{(n+m)\times(n+m)}$ be the coefficient matrix in* (1.2). *Assume $H = \frac{1}{2}(A + A^T)$ is positive semidefinite, B has full rank, $C = C^T$ is positive semidefinite, and $\mathcal{N}(H) \cap \mathcal{N}(B) = \{0\}$. Let $\sigma(\mathcal{A})$ denote the spectrum of $\mathcal{A}$. Then*

(i) *$\mathcal{A}$ is nonsingular.*

(ii) *$\mathcal{A}$ is* semipositive real*: $\langle \mathcal{A}\mathbf{v}, \mathbf{v}\rangle = \mathbf{v}^T\mathcal{A}\mathbf{v} \geq 0$ for all $\mathbf{v} \in \mathbb{R}^{n+m}$.*

(iii) *$\mathcal{A}$ is* positive semistable*, that is, the eigenvalues of $\mathcal{A}$ have nonnegative real part: $\Re(\lambda) \geq 0$ for all $\lambda \in \sigma(\mathcal{A})$.*

(iv) *If, in addition, $H = \frac{1}{2}(A + A^T)$ is positive definite, then $\mathcal{A}$ is* positive stable*: $\Re(\lambda) > 0$ for all $\lambda \in \sigma(\mathcal{A})$.*

Proof. To prove (i), let $\mathbf{x} = \left[\begin{smallmatrix} u \\ p \end{smallmatrix}\right]$ be such that $\mathcal{A}\mathbf{x} = \mathbf{0}$. Then

$$Au + B^T p = 0 \quad \text{and} \quad -Bu + Cp = 0. \tag{1.3}$$

Now, from $\mathcal{A}\mathbf{x} = \mathbf{0}$ we get $\mathbf{x}^T\mathcal{A}\mathbf{x} = u^TAu + p^TCp = 0$, and therefore it must be $u^TAu = 0$ and $p^TCp = 0$, since both of these quantities are nonnegative. But $u^TAu = u^THu = 0$, which implies $u \in \mathcal{N}(H)$ since H is symmetric positive semidefinite (see [36, p. 400]). Similarly, $p^TCp = 0$ with C symmetric positive semidefinite implies $Cp = 0$ and therefore (using the second of (1.3)) $Bu = 0$. Therefore $u = 0$ since $u \in \mathcal{N}(H) \cap \mathcal{N}(B) = \{0\}$. But if $u = 0$ then from the first of (1.3) we obtain $B^Tp = 0$ and therefore $p = 0$ since B has full column rank. Therefore the only solution to $\mathcal{A}\mathbf{x} = \mathbf{0}$ is the trivial solution, and $\mathcal{A}$ is nonsingular.

To prove (ii) we observe that for any $\mathbf{v} \in \mathbb{R}^{n+m}$ we have $\mathbf{v}^T\mathcal{A}\mathbf{v} = \mathbf{v}^T\mathcal{H}\mathbf{v}$, where

$$\mathcal{H} = \tfrac{1}{2}(\mathcal{A} + \mathcal{A}^T) = \begin{bmatrix} H & O \\ O & C \end{bmatrix}$$

is the symmetric part of $\mathcal{A}$. Clearly $\mathcal{H}$ is positive semidefinite, hence $\mathbf{v}^T\mathcal{A}\mathbf{v} \geq 0$.

To prove (iii), let $(\lambda, \mathbf{v})$ be an eigenpair of $\mathcal{A}$, with $||\mathbf{v}||_2 = 1$. Then $\mathbf{v}^*\mathcal{A}\mathbf{v} = \lambda$ and $(\mathbf{v}^*\mathcal{A}\mathbf{v})^* = \mathbf{v}^*\mathcal{A}^T\mathbf{v} = \bar{\lambda}$. Therefore $\frac{1}{2}\mathbf{v}^*(\mathcal{A} + \mathcal{A}^T)\mathbf{v} = \frac{\lambda+\bar{\lambda}}{2} = \Re(\lambda)$. To conclude the proof, observe that

$$\mathbf{v}^*(\mathcal{A} + \mathcal{A}^T)\mathbf{v} = \Re(\mathbf{v})^T(\mathcal{A} + \mathcal{A}^T)\Re(\mathbf{v}) + \Im(\mathbf{v})^T(\mathcal{A} + \mathcal{A}^T)\Im(\mathbf{v}),$$

a real nonnegative quantity.

To prove (iv), assume $(\lambda, \mathbf{v})$ is an eigenpair of $\mathcal{A}$ with $\mathbf{v} = \left[\begin{smallmatrix} u \\ p \end{smallmatrix}\right]$. Then

$$\Re(\lambda) = u^*Hu + p^*Cp = \Re(u)^TH\Re(u) + \Im(u)^TH\Im(u) + \Re(p)^TC\Re(p) + \Im(p)^TC\Im(p).$$

This quantity is nonnegative, and it can be zero only if $u = 0$ (since H is assumed to be positive definite) and $Cp = 0$. But if $u = 0$ then from the first of (1.3) we get $B^Tp = 0$, hence $p = 0$ since B has full column rank. Hence $\mathbf{v} = \mathbf{0}$, a contradiction. □

Thus, by changing the sign of the last m equations in (1.1) we may lose symmetry (when A is symmetric), but we gain positive (semi-)definiteness. This can be advantageous when using certain Krylov subspace methods, like restarted GMRES; see [19, 53].

In this paper we propose a new approach for preconditioning generalized saddle point problems based on an alternating symmetric/skew-symmetric splitting [2] applied to (1.2). This approach is very general in that it does not require the submatrix A to be nonsingular or symmetric; hence, it is applicable to a broad class of problems. The splitting method is described in section 2, and some of its convergence properties are studied in section 3. The use of the splitting as a preconditioner for Krylov subspace methods is considered in section 4. Numerical experiments are presented in section 5. Finally, in section 6 we draw our conclusions.

2. The alternating splitting iteration. In [2], the following stationary iterative methods for solving positive real linear systems $\mathcal{A}\mathbf{x} = \mathbf{b}$ was proposed. Write $\mathcal{A} = \mathcal{H} + \mathcal{S}$, where

$$\mathcal{H} = \tfrac{1}{2}(\mathcal{A} + \mathcal{A}^T), \quad \mathcal{S} = \tfrac{1}{2}(\mathcal{A} - \mathcal{A}^T)$$

are the symmetric and skew-symmetric part of $\mathcal{A}$, respectively. Let $\alpha > 0$ be a parameter. Similar in spirit to the classical alternating direction implicit (ADI) method [58], consider the following two splittings of $\mathcal{A}$:

$$\mathcal{A} = (\mathcal{H} + \alpha\mathcal{I}) - (\alpha\mathcal{I} - \mathcal{S})$$

and

$$\mathcal{A} = (\mathcal{S} + \alpha\mathcal{I}) - (\alpha\mathcal{I} - \mathcal{H}).$$

Here $\mathcal{I}$ denotes the identity matrix. The algorithm is obtained by alternating between these two splittings (see [7] for a general study of alternating iterations). Given an initial guess $\mathbf{x}^0$, the symmetric/skew-symmetric iteration computes a sequence $\{\mathbf{x}^k\}$ as follows:

$$\begin{cases} (\mathcal{H} + \alpha\mathcal{I})\mathbf{x}^{k+\frac{1}{2}} = (\alpha\mathcal{I} - \mathcal{S})\mathbf{x}^k + \mathbf{b}, \\ (\mathcal{S} + \alpha\mathcal{I})\mathbf{x}^{k+1} = (\alpha\mathcal{I} - \mathcal{H})\mathbf{x}^{k+\frac{1}{2}} + \mathbf{b}. \end{cases} \tag{2.1}$$

It is shown in [2] that if $\mathcal{H}$ is positive definite, the stationary iteration (2.1) converges for all $\alpha > 0$ to the solution of $\mathcal{A}\mathbf{x} = \mathbf{b}$.

Let us now consider the application of (2.1) to generalized saddle point problems in the form (1.2). In this case we have

$$\mathcal{H} = \begin{bmatrix} H & O \\ O & C \end{bmatrix} \quad \text{and} \quad \mathcal{S} = \begin{bmatrix} S & B^T \\ -B & O \end{bmatrix},$$

where $S = \frac{1}{2}(A - A^T)$ is the skew-symmetric part of A. Hence, $\mathcal{A}$ is positive real only when submatrices H and C are both symmetric positive definite (SPD), which is almost never the case in practice. Therefore, the convergence theory developed in [2] does not apply, and a more subtle analysis is required. We provide this analysis in the next section.

A few remarks are in order. At each iteration of (2.1), it is required to solve two sparse linear systems with coefficient matrices $\mathcal{H} + \alpha\mathcal{I}$ and $\mathcal{S} + \alpha\mathcal{I}$. Note that under our assumptions, both of these matrices are invertible for all $\alpha > 0$. Clearly, the choice of the solution methods used to perform the two half-steps in (2.1) is highly problem-dependent, and must be done on a case-by-case basis. The alternating algorithm (2.1) is just a general scheme that can incorporate whatever solvers are appropriate for a given problem.

Nevertheless, it is possible to make some general observations. The first half-step of algorithm (2.1) necessitates the solution of two (uncoupled) linear systems of the form

$$\begin{cases} (H + \alpha I_n)u^{k+\frac{1}{2}} = \alpha u^k - Su^k + f - B^T p^k, \\ (C + \alpha I_m)p^{k+\frac{1}{2}} = \alpha p^k - g + Bu^k. \end{cases} \tag{2.2}$$

Both systems in (2.2) are SPD, and any sparse solver for SPD systems can be applied. This could be a sparse Cholesky factorization, or a preconditioned conjugate gradient (PCG) scheme, or some specialized solver. Note that the addition of a positive term α to the main diagonal of H (and C) improves the condition number. This, in turn, tends to improve the rate of convergence of iterative methods applied to (2.2). More precisely, if H is normalized so that its largest eigenvalue is equal to 1, then for the spectral condition number of $H + \alpha I$ we have

$$\kappa(H + \alpha I) = \frac{1 + \alpha}{\lambda_{\min}(H) + \alpha} \leq 1 + \frac{1}{\alpha},$$

independent of the size of the problem. Note that even a fairly small value of α, such as $\alpha = 0.1$, yields a small condition number ($\kappa(H + \alpha I) \leq 11$). Unless α is very small, rapid convergence of the CG method applied to (2.2) can be expected, independent of the number n of unknowns.

The second half-step of algorithm (2.1) is less trivial. It requires the solution of two coupled linear systems of the form

$$\begin{cases} (\alpha I_n + S)u^{k+1} + B^T p^{k+1} = (\alpha I_n - H)u^{k+\frac{1}{2}} + f \equiv f^k, \\ -Bu^{k+1} + \alpha p^{k+1} = (\alpha I_m - C)p^{k+\frac{1}{2}} - g \equiv g^k. \end{cases} \tag{2.3}$$

This system can be solved in several ways. Of course, a sparse LU factorization could be used if the problem is not too large. An alternative approach is to eliminate u^{k+1} from the second equation using the first one (Schur complement reduction), leading to a smaller (order m) linear system of the form

$$[B(I_n + \alpha^{-1}S)^{-1}B^T + \alpha^2 I_m]p^{k+1} = B(I_n + \alpha^{-1}S)^{-1}f^k + \alpha g^k. \tag{2.4}$$

Once the solution p^{k+1} to (2.4) has been computed, the vector u^{k+1} is given by $u^{k+1} = (\alpha I_n + S)^{-1}(f^k - B^T p^{k+1})$. When $S = O$, system (2.4) simplifies to

$$(BB^T + \alpha^2 I_m)p^{k+1} = Bf^k + \alpha g^k, \tag{2.5}$$

and $u^{k+1} = \frac{1}{\alpha}(f^k - B^T p^{k+1})$. If BB^T is sufficiently sparse, system (2.5) could be formed explicitly and solved by a sparse Cholesky factorization. Otherwise, a PCG iteration with a simple preconditioner not requiring access to all the entries of the coefficient matrix $BB^T + \alpha^2 I_m$ could be used. However, when $S \neq O$ the coefficient matrix in (2.4) is generally dense. A nonsymmetric Krylov method could be used to solve (2.4), requiring matrix-vector products with the matrix $B(I_n + \alpha^{-1}S)^{-1}B^T + \alpha^2 I_m$. In turn, this requires solving a linear system of the form $(\alpha I_n + S)v = z$ at each step.

Also note that up to a scaling factor, the coefficient matrix of the coupled system in (2.3) is a normal matrix of the form "identity-plus-skew-symmetric." There are various Lanczos-type methods that can be applied to systems of this kind; see [14, 61] and, more generally, [38]. Other iterative methods for the solution of shifted skew-symmetric systems can be found, e.g., in [47] and [29].

Yet another possibility is to regard (2.3) as a general nonsymmetric system and to use preconditioned GMRES (say). Many of these schemes can benefit form the fact that for even moderate values of $\alpha > 0$, the condition number of $\mathcal{S} + \alpha\mathcal{I}$ is often rather small.

It is important to stress that the linear systems in (2.1) need not be solved exactly. The use of inexact solves was considered in [2] for the positive real case. The upshot is that inexact solves can be used to greatly reduce the cost of each iteration, at the expense of somewhat slower convergence. Typically, in practical implementations, inexact solves result in a much more competitive algorithm. Here we observe that when the alternating scheme is used as a preconditioner for a Krylov method, inexact solves are a natural choice, and there is no theoretical restriction on the accuracy of the inner solves. Inexact solutions are often obtained by iterative methods, leading to an inner-outer scheme; in this case, a flexible solver like FGMRES [52] should be used for the outer iteration. However, inexact solves may also be done by means of incomplete factorizations. In this case, standard GMRES can be used for the outer iteration.

Finally, we note that the scalar matrix $\alpha\mathcal{I}$ in (2.1) could be replaced by a matrix of the form $\alpha\mathcal{F}$, where $\mathcal{F}$ is SPD. This idea, in the context of ADI methods, goes back to Wachspress and Habetler [59]; see also [58, p. 242]. It is straightforward to see that this is equivalent to applying the alternating iteration (2.1) to the symmetrically preconditioned system

$$\hat{\mathcal{A}}\hat{\mathbf{x}} = \hat{\mathbf{b}}, \quad \hat{\mathcal{A}} := \mathcal{F}^{-1/2}\mathcal{A}\mathcal{F}^{-1/2}, \quad \hat{\mathbf{x}} = \mathcal{F}^{1/2}\mathbf{x}, \quad \hat{\mathbf{b}} = \mathcal{F}^{-1/2}\mathbf{b}. \tag{2.6}$$

In this paper we limit ourselves to the case where $\mathcal{F}$ is the $(n+m)\times(n+m)$ diagonal matrix having the ith diagonal entry equal to the ith diagonal entry of $\mathcal{A}$ if this is nonzero, and one otherwise. As we show in the section on numerical experiments, in many cases this simple diagonal preconditioning may considerably improve the rate of convergence.

In the next section we turn to the study of the convergence of the general scheme (2.1), assuming that the solves in (2.2) and (2.3) are performed exactly (rather than approximately, as in an inexact inner-outer setting).

3. Analysis of the stationary iteration. To analyze the convergence of (2.1), we eliminate the intermediate vector $\mathbf{x}^{k+\frac{1}{2}}$ and write the iteration in fixed point form as

$$\mathbf{x}^{k+1} = \mathcal{T}_\alpha \mathbf{x}^k + \mathbf{c}, \tag{3.1}$$

where

$$\mathcal{T}_\alpha := (\mathcal{S} + \alpha\mathcal{I})^{-1}(\alpha\mathcal{I} - \mathcal{H})(\mathcal{H} + \alpha\mathcal{I})^{-1}(\alpha\mathcal{I} - \mathcal{S})$$

is the iteration matrix of the method, and $\mathbf{c} := (\mathcal{S}+\alpha\mathcal{I})^{-1}[\mathcal{I}+(\alpha\mathcal{I}-\mathcal{H})(\mathcal{H}+\alpha\mathcal{I})^{-1}]\mathbf{b}$. The fixed point iteration (3.1) converges for arbitrary initial guesses $\mathbf{x}^0$ and right-hand sides $\mathbf{b}$ to the solution $\mathbf{x} = \mathcal{A}^{-1}\mathbf{b}$ if and only if $\rho(\mathcal{T}_\alpha) < 1$, where $\rho(\mathcal{T}_\alpha)$ denotes the spectral radius of $\mathcal{T}_\alpha$.

It was shown in [2] that when $\mathcal{A}$ is positive real (i.e., $\mathcal{H}$ is SPD), the stationary iteration is unconditionally convergent: $\rho(\mathcal{T}_\alpha) < 1$ for all $\alpha > 0$. However, in the context of generalized saddle point problems (1.2), the matrix $\mathcal{H}$ is only positive semidefinite and generally singular. In this case, the analysis in [2] does not apply. Indeed, for matrices whose symmetric part is positive semidefinite and singular, the alternating iteration is not convergent in general.

However, as the following theorem shows, the alternating iteration converges for a broad class of generalized saddle point problems.

THEOREM 3.1. *Consider problem* (1.2) *and assume that A is positive real, C is symmetric positive semidefinite, and B has full rank. Then the iteration* (3.1) *is unconditionally convergent; that is, $\rho(\mathcal{T}_\alpha) < 1$ for all $\alpha > 0$.*

Proof. The iteration matrix $\mathcal{T}_\alpha$ is similar to

$$\hat{\mathcal{T}}_\alpha = (\alpha\mathcal{I} - \mathcal{H})(\alpha\mathcal{I} + \mathcal{H})^{-1}(\alpha\mathcal{I} - \mathcal{S})(\alpha\mathcal{I} + \mathcal{S})^{-1} = \mathcal{R}\,\mathcal{U},$$

where $\mathcal{R} := (\alpha\mathcal{I}-\mathcal{H})(\alpha\mathcal{I}+\mathcal{H})^{-1}$ is symmetric and $\mathcal{U} := (\alpha\mathcal{I}-\mathcal{S})(\alpha\mathcal{I}+\mathcal{S})^{-1}$ orthogonal (see [37, p. 440]). Now, $\mathcal{R}$ is orthogonally similar to the $(n+m)\times(n+m)$ diagonal matrix

$$\mathcal{D} = \begin{bmatrix} \frac{\alpha-\mu_1}{\alpha+\mu_1} & & & & & & & & \\ & \frac{\alpha-\mu_2}{\alpha+\mu_2} & & & & & & & \\ & & \ddots & & & & & & \\ & & & \frac{\alpha-\mu_n}{\alpha+\mu_n} & & & & & \\ & & & & \frac{\alpha-\nu_1}{\alpha+\nu_1} & & & & \\ & & & & & \frac{\alpha-\nu_2}{\alpha+\nu_2} & & & \\ & & & & & & \ddots & & \\ & & & & & & & \frac{\alpha-\nu_m}{\alpha+\nu_m} & \end{bmatrix},$$

where $\mu_1, \mu_2, \ldots, \mu_n$ are the (positive) eigenvalues of H and $\nu_1, \nu_2, \ldots, \nu_m$ are the (nonnegative) eigenvalues of C. That is, there is an orthogonal matrix $\mathcal{V}$ of order $n+m$ such that

$$\mathcal{V}^T\mathcal{R}\mathcal{V} = \mathcal{D} = \begin{bmatrix} D_1 & O \\ O & D_2 \end{bmatrix},$$

where D_1 and D_2 are diagonal matrices of order n and m, respectively.

Note that

$$\left|\frac{\alpha-\mu_i}{\alpha+\mu_i}\right| < 1 \text{ for } 1 \le i \le n \quad \text{and} \quad \left|\frac{\alpha-\nu_i}{\alpha+\nu_i}\right| \le 1 \text{ for } 1 \le i \le m.$$

It follows that $\mathcal{RU}$ is orthogonally similar to

$$\mathcal{V}^T\mathcal{RUV} = (\mathcal{V}^T\mathcal{RV})(\mathcal{V}^T\mathcal{UV}) = \mathcal{DQ},$$

where $\mathcal{Q} := \mathcal{V}^T\mathcal{UV}$, being a product of orthogonal matrices, is orthogonal. Hence, the iteration matrix $\mathcal{T}_\alpha$ is similar to $\mathcal{DQ}$, and therefore

$$\rho(\mathcal{T}_\alpha) = \rho(\mathcal{DQ}) = \rho(\mathcal{QD}).$$

We claim that $\rho(\mathcal{QD}) < 1$ for all $\alpha > 0$. To show this, partition $\mathcal{Q}$ conformally to $\mathcal{D}$:

$$\mathcal{Q} = \begin{bmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{bmatrix}.$$

Then

$$\mathcal{QD} = \begin{bmatrix} Q_{11}D_1 & Q_{12}D_2 \\ Q_{21}D_1 & Q_{22}D_2 \end{bmatrix}.$$

Now, let $\lambda \in \mathbb{C}$ be an eigenvalue of $\mathcal{QD}$ and let $\mathbf{x} \in \mathbb{C}^{n+m}$ be a corresponding eigenvector with $||\mathbf{x}||_2 = 1$. We assume $\lambda \neq 0$, or else there is nothing to prove. We want to show that $|\lambda| < 1$. Clearly, $\mathcal{QD}\mathbf{x} = \lambda\mathbf{x}$ implies $\mathcal{D}\mathbf{x} = \lambda\mathcal{Q}^T\mathbf{x}$ and taking norms:

$$||\mathcal{D}\mathbf{x}||_2 = |\lambda|\,||\mathcal{Q}^T\mathbf{x}||_2 = |\lambda|.$$

Therefore

$$|\lambda|^2 = ||\mathcal{D}\mathbf{x}||_2^2 = \sum_{i=1}^{n}\left(\frac{\alpha-\mu_i}{\alpha+\mu_i}\right)^2 x_i\bar{x}_i + \sum_{i=n+1}^{n+m}\left(\frac{\alpha-\nu_i}{\alpha+\nu_i}\right)^2 x_i\bar{x}_i \le ||\mathbf{x}||_2^2 = 1. \tag{3.2}$$

Hence, the spectral radius of $\mathcal{T}_\alpha$ cannot exceed unity.

To prove that $|\lambda| < 1$ (strictly), we show that there exists at least one i $(1 \le i \le n)$ such that $x_i \neq 0$. Using the assumption that B has full rank, we will show that $x_i = 0$ for all $1 \le i \le n$ implies $\mathbf{x} = \mathbf{0}$, a contradiction. Indeed, if the eigenvector $\mathbf{x}$ is of the form $\mathbf{x} = \left[\begin{smallmatrix} 0 \\ \hat{x} \end{smallmatrix}\right]$ (where $\hat{x} \in \mathbb{C}^m$), the identity $\mathcal{QD}\mathbf{x} = \lambda\mathbf{x}$ becomes

$$\mathcal{QD}\mathbf{x} = \begin{bmatrix} Q_{11}D_1 & Q_{12}D_2 \\ Q_{21}D_1 & Q_{22}D_2 \end{bmatrix}\begin{bmatrix} 0 \\ \hat{x} \end{bmatrix} = \begin{bmatrix} Q_{12}D_2\hat{x} \\ Q_{22}D_2\hat{x} \end{bmatrix} = \begin{bmatrix} 0 \\ \lambda\hat{x} \end{bmatrix} \tag{3.3}$$

so that, in particular, it must be $Q_{12}D_2\hat{x} = 0$. We will prove shortly that Q_{12} has full column rank; hence, it must be $D_2\hat{x} = 0$. But by (3.3) we have $\lambda\hat{x} = Q_{22}D_2\hat{x} = 0$, and since $\lambda \neq 0$ by assumption, it must be $\hat{x} = 0$ (a contradiction, since $\mathbf{x} \neq \mathbf{0}$).

To conclude the proof we need to show that $Q_{12} \in \mathbb{R}^{n\times m}$ has full column rank. Recall that $\mathcal{Q} = \mathcal{V}^T\mathcal{UV}$ with

$$\mathcal{V} = \begin{bmatrix} V_{11} & O \\ O & V_{22} \end{bmatrix},$$

where $V_{11} \in \mathbb{R}^n$ is the orthogonal matrix that diagonalizes $(\alpha I_n - H)(\alpha I_n + H)^{-1}$ and $V_{22} \in \mathbb{R}^m$ is the orthogonal matrix that diagonalizes $(\alpha I_m - C)(\alpha I_m + C)^{-1}$. Recall that the orthogonal matrix $\mathcal{U}$ is given by

$$(\alpha\mathcal{I} - \mathcal{S})(\alpha\mathcal{I} + \mathcal{S})^{-1} = \begin{bmatrix} \alpha I_n - S & -B^T \\ B & \alpha I_m \end{bmatrix} \begin{bmatrix} \alpha I_n + S & B^T \\ -B & \alpha I_m \end{bmatrix}^{-1} = \begin{bmatrix} U_{11} & U_{12} \\ U_{21} & U_{22} \end{bmatrix}.$$

An explicit calculation reveals that

$$U_{12} = -[(\alpha I_n - S)(\alpha I_n + S)^{-1} + I_n]B^T[\alpha I_m + B(\alpha I_n + S)^{-1}B^T]^{-1}.$$

Clearly, -1 cannot be an eigenvalue of the orthogonal matrix $(\alpha I_n - S)(\alpha I_n + S)^{-1}$, hence $(\alpha I_n - S)(\alpha I_n + S)^{-1} + I_n$ is nonsingular. The matrix $\alpha I_m + B(\alpha I_n + S)^{-1}B^T$ is also nonsingular, since $B(\alpha I_n + S)^{-1}B^T$ is positive real. Indeed $(\alpha I_n + S)^{-1}$, being the inverse of a positive real matrix, is itself positive real and since B has full column rank, so is $B(\alpha I_n + S)^{-1}B^T$.

Furthermore,

$$\mathcal{Q} = \mathcal{V}^T\mathcal{U}\mathcal{V} = \begin{bmatrix} V_{11}^T U_{11} V_{11} & V_{11}^T U_{12} V_{22} \\ V_{22}^T U_{21} V_{11} & V_{22}^T U_{22} V_{22} \end{bmatrix}$$

and therefore

$$Q_{12} = V_{11}^T U_{12} V_{22} = -V_{11}^T[(\alpha I_n - S)(\alpha I_n + S)^{-1} + I_n]B^T[\alpha I_m + B(\alpha I_n + S)^{-1}B^T]^{-1}V_{22},$$

showing that Q_{12} has full column rank since V_{11}^T and V_{22} are orthogonal and B^T has full column rank. This completes the proof. □

REMARK 3.1. *It is easy to see that there is a unique splitting $\mathcal{A} = \mathcal{M} - \mathcal{N}$ with $\mathcal{M}$ nonsingular such that the iteration matrix $\mathcal{T}_\alpha$ is the matrix induced by that splitting, i.e., $\mathcal{T}_\alpha = \mathcal{M}^{-1}\mathcal{N} = \mathcal{I} - \mathcal{M}^{-1}\mathcal{A}$. An easy calculation shows that*

$$\mathcal{M} \equiv \mathcal{M}_\alpha = \tfrac{1}{2\alpha}(\mathcal{H} + \alpha\mathcal{I})(\mathcal{S} + \alpha\mathcal{I}). \tag{3.4}$$

It is therefore possible to rewrite the iteration (2.1) *in* correction form:

$$\mathbf{x}^{k+1} = \mathbf{x}^k + \mathcal{M}_\alpha^{-1}\mathbf{r}^k, \quad \mathbf{r}^k = \mathbf{b} - \mathcal{A}\mathbf{x}^k.$$

This will be useful when we consider Krylov subspace acceleration.

The restriction in Theorem 3.1 that A be positive real is not essential. If A is only semipositive real (singular), the alternating iteration (2.1) is still well defined, but it may happen that $\rho(\mathcal{T}_\alpha) = 1$ for *all* values of $\alpha > 0$. A simple example with $n = 2$, $m = 1$ is given by

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 0 & 1 \end{bmatrix}, \quad C = [0].$$

Nevertheless, a simple modification of the basic algorithm yields a convergent iteration. To this end, recall that $\rho(\mathcal{T}_\alpha) \le 1$ for all $\alpha > 0$; see (3.2). Also, note that $1 \notin \sigma(\mathcal{T}_\alpha)$ since $\mathcal{A}$ is nonsingular. Let $\beta \in (0, 1)$ be a parameter; then the matrix $(1 - \beta)\mathcal{I} + \beta\mathcal{T}_\alpha$ has spectral radius less than 1 for all $\alpha > 0$. Indeed, the eigenvalues of $(1 - \beta)\mathcal{I} + \beta\mathcal{T}_\alpha$ are of the form $1 - \beta + \beta\lambda$, where $\lambda \in \sigma(\mathcal{T}_\alpha)$. It is easy to see that since $|\lambda| \le 1$ and $\lambda \ne 1$, all the quantities $1 - \beta + \beta\lambda$ have magnitude strictly less

than 1. This trick is routinely used in the solution of singular systems and Markov chains; see, e.g., [40].

Thus, for any choice of the initial guess $\hat{\mathbf{x}}^0 = \mathbf{x}^0$, the sequence $\{\hat{\mathbf{x}}^k\}$ defined by

$$\hat{\mathbf{x}}^{k+1} = (1-\beta)\hat{\mathbf{x}}^k + \beta\mathbf{x}^{k+1} = (1-\beta)\hat{\mathbf{x}}^k + \beta(\mathcal{T}_\alpha\hat{\mathbf{x}}^k + \mathbf{c})$$

($k = 0, 1, \ldots$) converges to the unique solution of problem (1.2) for all $\beta \in (0, 1)$ and all $\alpha > 0$. In this way, the alternating iteration is applicable to any generalized saddle point problem. The presence of the parameter β, unfortunately, adds another complication to the method. Numerical experiments suggest that a value of β slightly less than 1, like $\beta = 0.99$, should be used. When Krylov subspace acceleration is used, however, there is no need to use this technique (that is, one can use $\beta = 1$ even when H is singular).

Under the assumptions of Theorem 3.1, the asymptotic rate of convergence of the alternating iteration is governed by the spectral radius of $\mathcal{T}_\alpha$, so it makes sense to try to choose α so as to make $\rho(\mathcal{T}_\alpha)$ as small as possible. In general, finding such a value $\alpha = \alpha_{\text{opt}}$ is a difficult problem. Some results in this direction can be found in [2, 3, 6]. The results in [2] yield an expression of the optimal α for the case of $\mathcal{A}$ positive real, too restrictive in our setting where $\mathcal{H}$ is usually singular.

Of course, choosing α so as to minimize the spectral radius of the iteration matrix is not necessarily the best choice when the algorithm is used as a preconditioner for a Krylov subspace method. Remarkably, it can be shown that for certain problems the alternating iteration results in an h-independent preconditioner for GMRES when α is chosen sufficiently small, corresponding to a spectral radius very close to 1; see [6] and the numerical experiments in section 5.1 below.

Also, minimizing the spectral radius or even the number of GMRES iterations does not imply optimal performance in terms of CPU time. Indeed, the efficient implementation of the method almost invariably requires that the two linear systems (2.2) and (2.3) be solved inexactly. Clearly, the choice of α will influence the cost of performing the two solves. Indeed, "large" values of α will make the iterative solution of (2.2) and (2.3) easy; on the other hand, it is clear from (3.2) that the nonzero eigenvalues of $\mathcal{T}_\alpha$ approach 1 as $\alpha \to \infty$ (and also as $\alpha \to 0$), and convergence of the outer iteration slows down. Hence, there is a trade-off involved. If we define the "optimal" value of α as the one that minimizes the total amount of work needed to compute an approximate solution, this will not necessarily be the same as the α that minimizes the number of (outer) iterations. Overall, the analytic determination of such an optimal value for α appears to be daunting.

4. Krylov subspace acceleration. Even with the optimal choice of α, the convergence of the stationary iteration (2.1) is typically too slow for the method to be competitive. For this reason we propose using a nonsymmetric Krylov subspace method like GMRES, or its restarted version GMRES(m), to accelerate the convergence of the iteration.

It follows from Remark 3.1 that the linear system $\mathcal{A}\mathbf{x} = \mathbf{b}$ is equivalent to (i.e., has the same solution as) the linear system

$$(\mathcal{I} - \mathcal{T}_\alpha)\mathbf{x} = \mathcal{M}_\alpha^{-1}\mathcal{A}\mathbf{x} = \mathbf{c},$$

where $\mathbf{c} = \mathcal{M}_\alpha^{-1}\mathbf{b}$. This equivalent (left-preconditioned) system can be solved with GMRES. Hence, the matrix $\mathcal{M}_\alpha$ can be seen as a *preconditioner* for GMRES. Equivalently, we can say that GMRES is used to accelerate the convergence of the alternating iteration applied to $\mathcal{A}\mathbf{x} = \mathbf{b}$.

Note that as a preconditioner we can use

$$\mathcal{M}_\alpha = (\mathcal{H} + \alpha\mathcal{I})(\mathcal{S} + \alpha\mathcal{I})$$

instead of the expression given in (3.4), since the factor $\frac{1}{2\alpha}$ has no effect on the preconditioned system. Application of the alternating preconditioner within GMRES requires solving a linear system of the form $\mathcal{M}_\alpha \mathbf{z} = \mathbf{r}$ at each iteration. This is done by first solving

$$(\mathcal{H} + \alpha\mathcal{I})\mathbf{v} = \mathbf{r} \tag{4.1}$$

for $\mathbf{v}$, followed by

$$(\mathcal{S} + \alpha\mathcal{I})\mathbf{z} = \mathbf{v}. \tag{4.2}$$

The GMRES method can also be applied to the right-preconditioned system $\mathcal{A}\mathcal{M}_\alpha^{-1}\mathbf{y} = \mathbf{b}$ where $\mathbf{y} = \mathcal{M}_\alpha \mathbf{x}$. Note that $\mathcal{M}_\alpha^{-1}\mathcal{A}$ and $\mathcal{A}\mathcal{M}_\alpha^{-1}$ are similar and therefore have the same eigenvalues. In principle, the convergence behavior of GMRES can be different depending on whether left- or right-preconditioning is being used, but in our numerical experiments we noticed little or no difference.

Under the assumptions of Theorem 3.1, since $\mathcal{M}_\alpha^{-1}\mathcal{A} = \mathcal{I} - \mathcal{T}_\alpha$ it is readily seen that for all $\alpha > 0$ the eigenvalues of the preconditioned matrix $\mathcal{M}_\alpha^{-1}\mathcal{A}$ (or of $\mathcal{A}\mathcal{M}_\alpha^{-1}$) are entirely contained in the open disk of radius 1 centered at $(1, 0)$. In particular, the preconditioned matrix is positive stable. The smaller the spectral radius of $\mathcal{T}_\alpha$, the more clustered the eigenvalues of the preconditioned matrix (around 1); a clustered spectrum often translates in rapid convergence of GMRES.

If a matrix is positive real, then it is positive stable; the converse, however, is not true. A counterexample is given by a matrix of the form

$$A = \begin{bmatrix} 1 & 0 \\ a & 1 \end{bmatrix},$$

where a is any real number with $|a| \geq 2$. The question then arises whether $\mathcal{M}_\alpha^{-1}\mathcal{A}$ (or $\mathcal{A}\mathcal{M}_\alpha^{-1}$) is positive real, for in this case the convergence of GMRES(m) would be guaranteed for all restarts m; see [19] and [53, p. 866]. Unfortunately, this is not true in general. However, when A is SPD and $C = O$ we can prove that the preconditioned matrix is positive real provided that α is sufficiently large.

THEOREM 4.1. *Assume A is SPD, $C = O$, and B has full rank. Then there exists $\alpha^* > 0$ such that $\mathcal{M}_\alpha^{-1}\mathcal{A}$ is positive real for all $\alpha > \alpha^*$. An analogous result holds for the right-preconditioned matrix, $\mathcal{A}\mathcal{M}_\alpha^{-1}$.*

Proof. For brevity, we prove the theorem only for the left-preconditioned matrix; the proof for the right-preconditioned one is similar. Up to a positive scalar multiple, the symmetric part of the preconditioned matrix $\mathcal{M}_\alpha^{-1}\mathcal{A}$ is given by

$$\mathcal{B} = (\mathcal{S} + \alpha\mathcal{I})^{-1}(\mathcal{H} + \alpha\mathcal{I})^{-1}\mathcal{A} + \mathcal{A}^T(\mathcal{H} + \alpha\mathcal{I})^{-1}(\alpha\mathcal{I} - \mathcal{S})^{-1}$$

(where we have used the fact that $\mathcal{S}^T = -\mathcal{S}$). This matrix is congruent to

$$(\mathcal{S} + \alpha\mathcal{I})\mathcal{B}(\mathcal{S} + \alpha\mathcal{I})^T = (\mathcal{H} + \alpha\mathcal{I})^{-1}\mathcal{A}(\alpha\mathcal{I} - \mathcal{S}) + (\mathcal{S} + \alpha\mathcal{I})\mathcal{A}^T(\mathcal{H} + \alpha\mathcal{I})^{-1},$$

which, in turn, is congruent to the inverse-free matrix

$$\mathcal{Z} = \mathcal{A}(\alpha\mathcal{I} - \mathcal{S})(\mathcal{H} + \alpha\mathcal{I}) + (\mathcal{H} + \alpha\mathcal{I})(\mathcal{S} + \alpha\mathcal{I})\mathcal{A}^T.$$

A direct calculation shows that

$$\mathcal{Z} = \begin{bmatrix} Z_\alpha & -2\alpha AB^T \\ -2\alpha BA & 2\alpha BB^T \end{bmatrix},$$

where

$$Z_\alpha := 2\alpha A^2 + 2\alpha B^T B + 2\alpha^2 A + B^T BA + AB^T B.$$

We want to show that $\mathcal{Z}$ is SPD for sufficiently large α. To this end, we observe that $\mathcal{Z}$ can be split as

$$\mathcal{Z} = 2\begin{bmatrix} \alpha A^2 & -\alpha AB^T \\ -\alpha BA & \alpha BB^T \end{bmatrix} + \begin{bmatrix} M_\alpha & O \\ O & O \end{bmatrix}, \tag{4.3}$$

where

$$M_\alpha := 2\alpha^2 A + 2\alpha B^T B + B^T BA + AB^T B.$$

The first matrix on the right-hand side of (4.3) is symmetric positive semidefinite, since

$$\begin{bmatrix} \alpha A^2 & -\alpha AB^T \\ -\alpha BA & \alpha BB^T \end{bmatrix} = \begin{bmatrix} \alpha A & O \\ -\alpha B & I_m \end{bmatrix} \begin{bmatrix} \alpha^{-1} I_n & O \\ O & O \end{bmatrix} \begin{bmatrix} \alpha A & -\alpha B^T \\ O & I_m \end{bmatrix}.$$

Next, we observe that

$$M_\alpha = \alpha(2B^T B + 2\alpha A) + (B^T BA + AB^T B)$$

is similar to a matrix of the form $\alpha I_n + W$, where $W = W^T$ is generally indefinite. This matrix can be made SPD by taking α sufficiently large. Specifically, M_α is SPD for all $\alpha > \alpha^*$, where

$$\alpha^* = -\lambda_{\min}(B^T BA + AB^T B)$$

(note that $B^T BA + AB^T B$ is generally indefinite). Hence, for $\alpha > \alpha^*$ the matrix $\mathcal{Z}$ is the sum of two symmetric positive semidefinite matrices; therefore, it is itself symmetric positive semidefinite. Finally, it must be nonsingular for all $\alpha > \alpha^*$ (and therefore positive definite). Indeed, it is clear from (4.3) that when M_α is positive definite, any null vector of $\mathcal{Z}$ must be of the form

$$\mathbf{x} = \begin{bmatrix} 0 \\ \hat{x} \end{bmatrix}, \quad \text{where } \hat{x} \in \mathbb{R}^m.$$

But then

$$\mathcal{Z}\mathbf{x} = 2\begin{bmatrix} \alpha A^2 & -\alpha AB^T \\ -\alpha BA & \alpha BB^T \end{bmatrix} \begin{bmatrix} 0 \\ \hat{x} \end{bmatrix} = \begin{bmatrix} -2\alpha AB^T \hat{x} \\ 2\alpha BB^T \hat{x} \end{bmatrix},$$

which cannot be zero unless $\hat{x} = 0$, since B^T has full column rank and A is nonsingular. Hence $\mathcal{Z}$ has no nontrivial null vectors for $\alpha > \alpha^*$. This shows that the symmetric part of the preconditioned matrix is SPD for all $\alpha > \alpha^*$, since it is congruent to a matrix which is SPD for all such values of α. □

It is worth mentioning that in all cases that we were able to check numerically, we found the symmetric part of the preconditioned operator to be positive definite already for rather small values of α.

More refined bounds and clustering results for the eigenvalues of $\mathcal{M}_\alpha^{-1}\mathcal{A}$ can be found in [55].

5. Numerical experiments. In this section we present a sample of numerical experiments conducted in order to assess the effectiveness of the alternating algorithm (2.1) both as a stationary iterative scheme and as a preconditioner for GMRES. All experiments were performed in Matlab. Our codes have not been optimized for highest efficiency and therefore we do not report timings, but we do provide cost estimates for some of the test problems. We think that the results of the experiments presented here provide evidence of the fact that our approach is worth further consideration.

We target matrices from different application areas, but mostly from PDE problems. In all our runs we used a zero initial guess and stopped the iteration when the relative residual had been reduced by at least six orders of magnitude (i.e., when $||\mathbf{b} - \mathcal{A}\mathbf{x}^k||_2 \leq 10^{-6}||\mathbf{b}||_2$).

5.1. Second order equations in first order system form. Let $\Omega \subset \mathbb{R}^d$ $(d = 2, 3)$ be a bounded open set. Here we consider the numerical solution of boundary value problems for the following second order elliptic PDE:

$$-\nabla \cdot (K \nabla p) = g \quad \text{in } \Omega, \tag{5.1}$$

where $K = K(\mathbf{r})$ is a strictly positive function or tensor for $\mathbf{r} \in \bar{\Omega}$ and $g(\mathbf{r})$ is a given forcing term. Equation (5.1) is complemented by appropriate boundary conditions.

The PDE (5.1) is equivalent to the following system of two first order PDEs:

$$\begin{cases} K^{-1}\,\mathbf{u} - \nabla p = \mathbf{0}, \\ -\nabla \cdot \mathbf{u} = g. \end{cases} \tag{5.2}$$

Discretization of these equations leads to large sparse linear systems in saddle point form (1.2).

We begin with the simplest possible case, namely, Poisson's equation on the unit square:

$$-\Delta p = -\nabla \cdot (\nabla p) = g \quad \text{in } \Omega = [0, 1] \times [0, 1].$$

This corresponds to taking $K \equiv 1$ in (5.1). We discretize form (5.2) of the problem using finite differences with a forward difference for the gradient and a backward difference for the divergence. Using an $N \times N$ uniform grid with mesh size $h = \frac{1}{N+1}$ results in a linear system of type (1.2) with $n = 2N^2$ and $m = N^2$, for a total system size of $3N^2$ equations in as many unknowns.

As shown in [6], for this model problem Fourier analysis at the continuous (differential operator) level can be used to completely analyze the spectrum of the iteration operator $\mathcal{T}_\alpha$. This allows us to find the optimal value α_{opt} of the parameter as a function of h, showing that the spectral radius for the stationary iteration (2.1) behaves as $1 - c\sqrt{h}$ as $h \to 0$. The optimal value α_{opt} itself behaves as $h^{-\frac{1}{2}}$ as $h \to 0$. More interestingly, the spectral analysis in [6] indicates that when GMRES acceleration is used, a better choice is to use a small value of α, for it can be shown that for $\alpha \in (0, 1)$ the eigenvalues of the preconditioned matrix lie in two intervals which depend on α, but do not depend on h, resulting in h-independent convergence. In particular, α can always be chosen so as to have convergence within 2–3 iterations, uniformly in h.

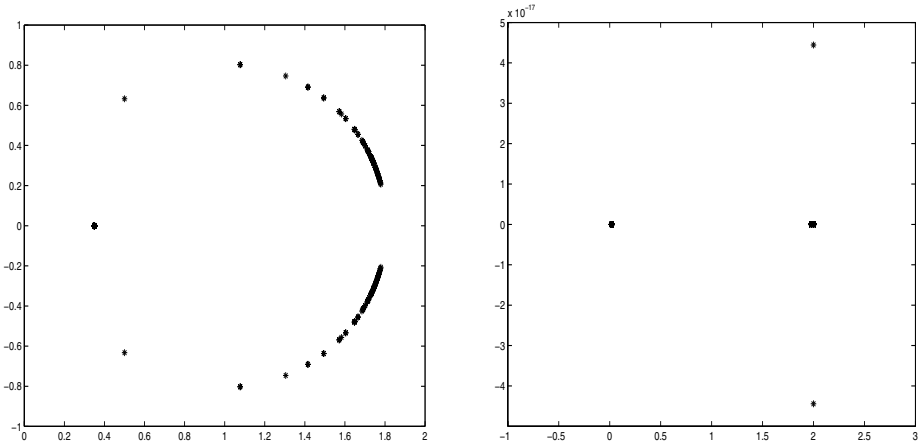
This behavior is illustrated in Table 5.1. We take the forcing term to be the function $g(x, y) = \sin \pi x \sin \pi y$ and we impose Neumann boundary conditions for $x = 0$, $x = 1$, and homogeneous Dirichlet boundary conditions for $y = 0$, $y = 1$. The numerical results are in agreement with the theoretical analysis. In particular, note

TABLE 5.1
Two-dimensional Poisson's equation. Comparison of iterative scheme optimized as an iterative solver, full GMRES without preconditioner, GMRES with the optimized iterative scheme as a preconditioner and iterative scheme optimized for GMRES.

		GMRES		
h	Iterative	No Prec.	Preconditioned	Optimized
1/10	66	54	14	2
1/25	103	140	19	2
1/50	146	286	25	2
1/100	207	574	34	2

that convergence is attained in two steps (independent of h) when the iteration is optimized for GMRES acceleration. Here we used $\alpha = 0.001$, but the behavior of the preconditioned iteration is not very sensitive to the choice of $\alpha \in (0, 1)$.

In Figure 5.1 we display the eigenvalues of the preconditioned matrix $\mathcal{M}_\alpha^{-1}\mathcal{A}$ in the case of $h = \frac{1}{10}$ for two values of α. On the left we used the value $\alpha = \alpha_{\rm opt}$ that minimizes the spectral radius, which is given by $\rho(\mathcal{T}_{\alpha_{\rm opt}}) = 0.8062$. On the right we used $\alpha = 0.01$, showing the clustering near 0 and 2 predicted by the theory developed in [6]. Now the spectral radius of the iteration matrix is very close to 1. The cluster near 0 contains $m = 81$ eigenvalues, the one near 2 the remaining $n = 162$. It should be noted that the (tiny) imaginary part in Figure 5.1(b) is due to round-off error, since the eigenvalues are real for small α; see [6, 55].



(a) Spectrum for $\alpha = \alpha_{\rm opt}$ (b) Spectrum for $\alpha = 0.01$

FIG. 5.1. *Eigenvalues of preconditioned matrices for the Poisson problem on a* 10×10 *grid.*

Next we consider a somewhat harder problem, namely, the anisotropic equation

$$-100\, p_{xx} - p_{yy} = g \quad \text{in } \Omega = [0, 1] \times [0, 1].$$

Since this problem has constant coefficients, the technique used in [6] for Poisson's equation can be used to optimize the method. The results in Table 5.2 show that the anisotropy in the coefficients drastically decreases the rate of convergence. However, in this case there is an easy fix: as the results reported in Table 5.3 show, it is enough to apply the scaling (2.6) to restore the effectiveness of the solver. We note that a similar scaling has been used in [13] in a somewhat different context.

TABLE 5.2
Results for two-dimensional problem with anisotropic coefficients.

		GMRES		
h	Iterative	No Prec.	Preconditioned	Optimized
1/10	709	186	34	29
1/25	> 1000	651	44	31
1/50	> 1000	> 1000	52	31
1/100	> 1000	> 1000	59	31

TABLE 5.3
Results for two-dimensional problem with anisotropic coefficients, diagonally scaled.

		GMRES		
h	Iterative	No Prec.	Preconditioned	Optimized
1/10	138	100	15	2
1/25	210	344	17	2
1/50	292	> 500	22	2
1/100	400	> 500	29	2

Finally, we consider a more difficult problem with large jumps in the coefficients K. The system is discretized using a discontinuous Galerkin finite element scheme. This radiation diffusion problem arises in a nuclear engineering application and was supplied to us by James Warsa of Los Alamos National Laboratory. For more details, see [60] and the references therein. For this problem $n = 2592$, $m = 864$, $n+m = 3456$, and $\mathcal{A}$ contains 93,612 nonzero entries. Here $C \neq O$ (and indeed it is SPD).

The results for this problem are presented in Table 5.4, where the entries in the first row correspond to GMRES with diagonal preconditioning (2.6). We give results for full GMRES and for restarted GMRES with restart every 20 steps. Here we cannot apply Fourier analysis to optimize the choice of α as we did in the constant coefficient cases. Therefore, we experiment with different values of α. While the fastest convergence rate for the stationary iterative methods correspond to $\alpha = 0.25$, a somewhat bigger α works best if the method is used as a preconditioner for GMRES. In any case the method is not overly sensitive to the choice of α when GMRES acceleration is used. We stress here again the importance of the diagonal scaling (2.6), which results in a reduction by a factor of two in the number of iterations for this problem.

5.2. Stokes and Oseen problems. In this section we present a few results for discretizations of Stokes and Oseen problems. Recall that the Stokes system is

$$\begin{cases} -\Delta \mathbf{u} + \nabla p = \mathbf{f}, \\ \nabla \cdot \mathbf{u} = 0 \end{cases} \tag{5.3}$$

in $\Omega \subset \mathbb{R}^d$, together with suitable boundary conditions. Here $\mathbf{u}$ denotes the velocity vector field and p the pressure scalar field. Discretization of (5.3) using stabilized finite elements leads to saddle point problems of the type (1.2) with a symmetric positive definite A and a symmetric positive semidefinite C.

The Oseen equations are obtained when the steady-state Navier–Stokes equations are linearized by Picard iteration:

$$\begin{cases} -\nu \Delta \mathbf{u} + (\mathbf{v} \cdot \nabla)\, \mathbf{u} + \nabla p = \mathbf{f}, \\ \nabla \cdot \mathbf{u} = 0\,. \end{cases} \tag{5.4}$$

Here the vector field $\mathbf{v}$ is the approximation of $\mathbf{u}$ from the previous Picard iteration. The parameter $\nu > 0$ represents viscosity. Various approximation schemes can be used

TABLE 5.4
Results for discontinuous radiation diffusion equations.

α	Iterative	GMRES	GMRES(20)
–		697	> 1000
0.1	750	155	239
0.2	375	100	113
0.25	257	97	101
0.3	304	89	100
0.4	404	85	95
0.5	504	85	100
0.6	604	81	98
0.7	704	80	108
0.8	805	80	115
0.9	905	80	120
1.0	1005	82	135

to discretize the Oseen problem (5.4) leading to a generalized saddle point system of type (1.2). Now the A block corresponds to a discretization of the convection-diffusion operator $L[\mathbf{u}] := -\nu\Delta\mathbf{u} + (\mathbf{v}\cdot\nabla)\,\mathbf{u}$. It is nonsymmetric, but for conservative discretizations, the symmetric part is positive definite.

We generated several test problems using the IFISS software package written by Howard Elman, Alison Ramage, and David Silvester. We used this package to generate discretizations of leaky lid driven cavity problems for both the Stokes and Oseen equations. The discretization used is stabilized Q1-P0 finite elements. In all cases the default value of the stabilization parameter ($\beta = 0.25$) was used. It should be mentioned that the matrices generated by this package are actually singular, since B has rank $m - 2$. This does not cause any difficulty to the iterative solvers considered here. In particular, even if $\lambda = 1$ is now an eigenvalue of the iteration matrix $\mathcal{T}_\alpha = \mathcal{I} - \mathcal{M}_\alpha^{-1}\mathcal{A}$, the stationary iteration is still convergent, with a rate of convergence governed by $\gamma(\mathcal{T}_\alpha) := \max\{\,|\lambda|;\, \lambda \in \sigma(\mathcal{T}_\alpha),\, \lambda \neq 1\}$.

For the Stokes problem we used a 16×16 grid. For the Oseen problem we used two grids, 16×16 and 32×32. The first grid corresponds to $n = 578$ and $m = 256$, for a total of 834 unknowns. For the second grid $n = 2178$ and $m = 1024$, for a total of 3202 unknowns. Two values of the viscosity parameter were used for the Oseen problems, $\nu = 0.01$ and $\nu = 0.001$. We experiment with both full GMRES and GMRES(20). Diagonal scaling (2.6) greatly improves the rate of convergence in all cases, and it is used throughout.

Table 5.5 contains results for the Stokes problem with both exact and inexact solves. Although there is no value of α that yields convergence in two steps, the alternating iteration is able to significantly improve the convergence of GMRES. Note that the behavior of the preconditioned iteration is not overly sensitive to the choice of α; in contrast, the rate of convergence of the stationary iteration without GMRES acceleration depends strongly on α. Since the (average) cost of a preconditioned GMRES(20) iteration is approximately three times the cost of an unpreconditioned iteration, the preconditioner allows for a saving of about a factor of two over unpreconditioned GMRES(20), when using the "best" values of α. Better results are obtained with inexact solves corresponding to incomplete factorizations. We used drop tolerance-based incomplete Cholesky for the first system in (2.1) and ILU for the second one. In both cases the drop tolerance was set to $tol = 0.05$. For $\alpha \geq 0.1$ the incomplete Cholesky factor of $\mathcal{H} + \alpha\mathcal{I}$ is very sparse, with about 25% of the nonzeros in the coefficient matrix itself. The ILU factors of $\mathcal{S} + \alpha\mathcal{I}$, for this particular example,

TABLE 5.5
Results for Stokes problem.

	Exact solves			Inexact solves	
α	Iterative	GMRES	GMRES(20)	GMRES	GMRES(20)
–		103	194		
0.01	> 1000	101	205	210	> 500
0.1	801	53	60	58	62
0.2	135	33	36	35	36
0.3	78	29	30	32	30
0.4	107	29	34	33	35
0.5	134	30	39	37	41
0.6	137	32	47	40	48
0.7	165	35	51	42	54
0.8	222	37	58	44	59
0.9	250	39	63	45	64
1.0	277	42	67	46	68

TABLE 5.6
Results for Oseen problem on 16×16 *grid,* $\nu = 0.01$.

	Exact solves			Inexact solves	
α	Iterative	GMRES	GMRES(20)	GMRES	GMRES(20)
–		353	> 500		
0.01	> 1000	105	268	179	> 500
0.1	> 1000	58	66	163	> 500
0.2	> 1000	38	39	107	> 500
0.3	> 1000	29	29	82	166
0.4	842	25	25	70	114
0.5	474	23	23	61	89
0.6	301	22	22	50	62
0.7	203	23	23	43	52
0.8	149	23	23	40	50
0.9	157	24	25	39	49
1.0	170	26	27	40	55
1.1	183	27	30	37	46
1.2	197	29	33	38	51

have around 35% of the nonzeros in the complete factors. As a result, the cost of applying the preconditioner is reduced by about a factor of four (from 62.6×10^3 to 15.4×10^3 operations, per iteration). It can be seen that the rate of convergence deteriorates only slightly. This deterioration is more than compensated by the lower cost per iteration. Moreover, the set-up cost goes down from 292×10^3 operations for the complete factorizations to 81×10^3 for the incomplete ones. Compared to the exact case, the overall reduction in the total number of operations is more than a factor of two for α between 0.4 and 1, while total storage for the preconditioner is reduced by almost a factor of three. Also note that with inexact solves, the (average) cost of a preconditioned GMRES(20) iteration is approximately one and a half times the cost of an unpreconditioned iteration. Hence, for the best values of α, the preconditioner results in a reduction of the cost of GMRES(20) by more than a factor of four.

Table 5.6 contains experimental results for the Oseen problem on the small (16×16) grid with viscosity $\nu = 0.01$. The results for GMRES with diagonal scaling (2.6), reported in the first row, indicate that the Oseen problem is harder than the Stokes problem. Here we see a surprising result: while the stationary iteration tends to converge more slowly than for the Stokes problem, the preconditioned GMRES iteration now tends to converge faster. We think this could be due to the fact

TABLE 5.7
Results for Oseen problem on 16×16 *grid,* $\nu = 0.001$.

	Exact solves			Inexact solves	
α	Iterative	GMRES	GMRES(20)	GMRES	GMRES(20)
–		616	> 1000		
0.01	> 1000	69	177	162	> 1000
0.1	> 1000	42	55	149	> 1000
0.2	> 1000	32	37	131	> 1000
0.3	> 1000	22	28	112	> 1000
0.4	> 1000	22	22	101	> 1000
0.5	> 1000	22	22	90	> 1000
0.6	> 1000	21	21	86	> 1000
0.7	965	21	21	70	664
0.8	713	21	21	76	> 1000
0.9	552	21	21	71	275
1.0	444	22	22	60	166
1.1	364	22	23	53	228
1.2	302	23	24	54	108
1.5	239	25	29	53	137
2.0	286	31	42	56	151

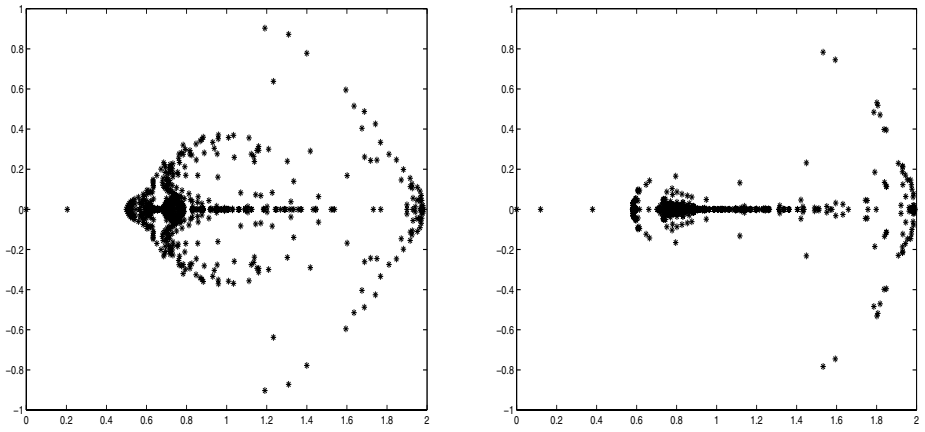
that the coefficient matrix has a more substantial skew-symmetric part in this case, and preconditioning with the (shifted) skew-symmetric part becomes more effective. Now GMRES(20) does not converge within 500 iterations without preconditioning. Full GMRES requires about 5.7 times more flops than the stationary iteration with $\alpha = 0.8$, and about 17 times more than the preconditioned iteration. Note that this estimate includes the set-up time for the preconditioner. The results obtained with inexact solves (by incomplete factorization) show some deterioration (about a factor of two for the "best" α) in convergence rates. This deterioration is more than compensated by the reduced cost of each preconditioned iteration.

In Table 5.7 we report results for the Oseen problem on the 16×16 grid and a viscosity parameter $\nu = 0.001$. Generally speaking, the Oseen problem becomes harder to solve as the viscosity gets smaller; see the results for diagonally scaled GMRES, and for the stationary iteration. However, the combination of the iteration and GMRES acceleration results in even faster convergence than in the previous case of $\nu = 0.01$. In Figure 5.2 we display the eigenvalues of the preconditioned matrix corresponding to the Oseen problem on the 16×16 grid. The plot on the left corresponds to a viscosity $\nu = 0.01$ and the one on the right to $\nu = 0.001$; we used the values of α that resulted in the smallest number of preconditioned GMRES iterations ($\alpha = 0.6$ and $\alpha = 0.8$, respectively). Note the stronger clustering of the spectrum for the case with $\nu = 0.001$.

Unfortunately, this apparent robustness with respect to ν is lost as soon as the exact solves in (2.1) are replaced by inexact solves by incomplete factorization, especially with restarted GMRES. The same value of the drop tolerance $tol = 0.05$ was used in all cases. Whether it is possible to solve the inner problems inexactly and still preserve robustness with respect to ν remains an open question.

Finally, in Table 5.8 we present results for the Oseen problem with $\nu = 0.001$ on the finer grid. The preconditioned GMRES iteration appears to be fairly robust with respect to the mesh size h and the viscosity parameter ν when exact solves are used.

5.3. A problem with singular A. Finally, we consider a saddle point problem arising in geophysics and supplied to us by Eldad Haber of Emory University; see [28, 33, 34]. In this application the submatrix A is symmetric positive semidefinite



(a) Spectrum for $\nu = 0.01$ (b) Spectrum for $\nu = 0.001$

FIG. 5.2. *Eigenvalues of preconditioned matrices for the Oseen problem on a* 16×16 *grid.*

TABLE 5.8
Results for Oseen problem with exact solves on 32×32 *grid,* $\nu = 0.001$.

α	Iterative	GMRES	GMRES(20)
–		> 1000	> 1000
0.1	> 1000	52	58
0.2	> 1000	36	38
0.3	> 1000	32	32
0.4	> 1000	29	30
0.5	> 1000	28	29
0.6	> 1000	28	34
0.7	> 1000	28	40
0.8	> 1000	39	42
0.9	757	31	44
1.0	574	33	47
1.5	684	43	55
2.0	864	53	121

and singular. In the example at hand $n = 1241$, $m = 729$, $n + m = 1970$, and $\mathcal{A}$ contains 25,243 nonzeros. The A block has $\text{rank}(A) = 876$. In this problem, $C = O$.

We present results for this problem in Table 5.9. Diagonal scaling (2.6) drastically improves the convergence of the preconditioned iterations. However, the convergence of the stationary iteration (2.1) without GMRES acceleration remains extremely slow. Likewise for GMRES with no preconditioning or diagonal preconditioning alone. The results with inexact solves in Table 5.9 were obtained by replacing the exact solve with no-fill incomplete factorizations, IC(0) and ILU(0). Again we see a deterioration in convergence rates, but each iteration is now far cheaper than in the case of exact solves, resulting in huge savings. When $\alpha = 0.3$ (but similar results hold for all the other values of α in the table), the Cholesky factorization of $\mathcal{H} + \alpha\mathcal{I}$ requires 1.7×10^6 operations using a minimum degree ordering, resulting in a triangular factor with 31.5×10^3 nonzeros. The complete factorization of $\mathcal{S} + \alpha\mathcal{I}$ (using minimum degree as the initial ordering) costs a staggering 207×10^6 operations with a total number of nonzeros in the factors exceeding 837×10^3. In contrast, the IC(0) factorization of $\mathcal{H} + \alpha\mathcal{I}$ only required 17.6×10^3 operations and resulted in an incomplete Cholesky factor with just 5.2×10^3 nonzeros; the ILU(0) factorization of $\mathcal{S} + \alpha\mathcal{I}$ took

TABLE 5.9
Results for geophysics problem with singular A.

	Exact solves			Inexact solves	
α	Iterative	GMRES	GMRES(30)	GMRES	GMRES(30)
–		> 500	> 500		
0.1	$> 1{,}000$	49	60	85	177
0.2	$> 1{,}000$	42	42	68	110
0.3	$> 1{,}000$	44	62	60	112
0.4	$> 1{,}000$	48	92	58	93
0.5	$> 1{,}000$	51	98	62	104

76.1×10^3 operations and resulted in a total of 21.4×10^3 nonzeros in the incomplete factors.

6. Conclusions and future work. In this paper we have studied the extension of the alternating method of [2] to generalized saddle point problems. Because these linear systems have coefficient matrices with singular symmetric part, they are not positive real. Thus, the convergence analysis carried out in [2] for the positive real case does not apply, and convergence has to be established using different arguments from those used in [2]. Other approaches to studying convergence have been proposed recently in [3] and [6]; see also [8] and [55].

Rather than used as a stand-alone solver, the stationary iteration is best used as a preconditioner for a nonsymmetric Krylov subspace method, such as GMRES. Here we have established theoretical properties of the preconditioned matrices that were relevant for restarted GMRES, at least from a qualitative point of view.

Our numerical experiments with test matrices from several different applications suggest that the combination of GMRES and the alternating iteration is fairly robust, and not overly sensitive to the choice of the parameter α. As demonstrated already in [6] for some model problems, there are important examples of systems of PDEs where the combination of iteration (2.1) with an appropriate choice of the optimization parameter α and GMRES acceleration results in an h-independent solver, or with a weak dependence on h.

Our numerical experiments show that diagonal scaling (2.6) greatly improves the convergence of the outer iteration. We have also performed some experiments with inexact solves. For several of our test problems, the rate of convergence suffered relatively little deterioration, leading to a reduction in overall costs in many cases. However, we also found problems where inexactness in the inner solves resulted in slow convergence, at least when incomplete factorizations were used.

Future work should focus on developing efficient implementations of the algorithm, with particular attention to the problem of striking a balance between the rate of convergence of the outer (preconditioned) iteration, and the amount of work spent performing the inner (inexact) solves. Here we have presented a few results using incomplete factorizations, but iterative methods may be a better (more flexible) option. For the Oseen equations with small viscosity parameter ν, it may be difficult to find inexact inner solves that do not lead to a serious deterioration of the rate of convergence of the outer iteration. The shifted symmetric part (4.1) has condition numbers often of the order of 10 or less, and is typically very easy to solve, at least in PDE problems. The solution of the shifted skew-symmetric part (4.2), on the other hand, is somewhat more problematic and warrants further research. Preliminary results show that when α is not too small, fairly accurate approximate solutions to the linear system (4.2) can be obtained in just 3–4 iterations of GMRES preconditioned

with an incomplete factorization. This inner-outer scheme, which requires using a flexible Krylov method (like FGMRES) as the outer iteration, is currently being investigated.

Acknowledgments. We gratefully acknowledge the many excellent suggestions from the anonymous referees, as well as helpful discussions with Howard Elman and Valeria Simoncini.

REFERENCES

[1] M. Arioli and G. Manzini, *A null space algorithm for mixed finite-element approximations of Darcy's equation*, Comm. Numer. Meth. Engrg., 18 (2002), pp. 645–657.

[2] Z. Z. Bai, G. H. Golub, and M. K. Ng, *Hermitian and skew-Hermitian splitting methods for non-Hermitian positive definite linear systems*, SIAM J. Matrix Anal. Appl., 24 (2003), pp. 603–626.

[3] Z. Z. Bai, G. H. Golub, and J. Y. Pan, *Preconditioned Hermitian and Skew-Hermitian Splitting Methods for Non-Hermitian Positive Semidefinite Linear Systems*, Technical Report SCCM-02-12, Scientific Computing and Computational Mathematics Program, Department of Computer Science, Stanford University, Stanford, CA, 2002.

[4] 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.

[5] M. Benzi, *Solution of equality-constrained quadratic programming problems by a projection iterative method*, Rend. Mat. Appl., 13 (1993), pp. 275–296.

[6] M. Benzi, M. J. Gander, and G. H. Golub, *Optimization of the Hermitian and skew-Hermitian splitting iteration for saddle-point problems*, BIT, 43 (2003), pp. 881–900.

[7] M. Benzi and D. B. Szyld, *Existence and uniqueness of splittings for stationary iterative methods with applications to alternating methods*, Numer. Math., 76 (1997), pp. 309–321.

[8] D. Bertaccini, G. H. Golub, S. Serra Capizzano, and C. Tablino Possio, *Preconditioned HSS Method for the Solution of Non-Hermitian Positive Definite Linear Systems*, Technical Report SCCM-02-11, Scientific Computing and Computational Mathematics Program, Department of Computer Science, Stanford University, Stanford, CA, 2002.

[9] J. T. Betts, *Practical Methods for Optimal Control Using Nonlinear Programming*, SIAM, Philadelphia, 2001.

[10] Å. Björck, *Numerical Methods for Least Squares Problems*, SIAM, Philadelphia, 1996.

[11] J. H. Bramble, J. E. Pasciak, and A. T. Vassilev, *Analysis of the inexact Uzawa algorithm for saddle point problems*, SIAM J. Numer. Anal., 34 (1997), pp. 1072–1092.

[12] F. Brezzi and M. Fortin, *Mixed and Hybrid Finite Element Methods*, Springer-Verlag, New York, 1991.

[13] P. Concus and G. H. Golub, *Use of fast direct methods for the efficient numerical solution of nonseparable elliptic equations*, SIAM J. Numer. Anal., 10 (1973), pp. 1103–1120.

[14] P. Concus and G. H. Golub, *A generalized conjugate gradient method for nonsymmetric systems of linear equations*, in Computing Methods in Applied Sciences and Engineering, Lecture Notes in Econom. and Math. Systems 134, R. Glowinski and J. L. Lions, eds., Springer-Verlag, Berlin, 1976, pp. 56–65.

[15] E. de Sturler and J. Liesen, *Block-Diagonal Preconditioners for Indefinite Linear Algebraic Systems*, Report UIUCDCS-R-2002-2279, University of Illinois, Champaign, IL, 2002.

[16] I. S. Duff, N. I. M. Gould, J. K. Reid, J. A. Scott, and K. Turner, *The factorization of sparse symmetric indefinite matrices*, IMA J. Numer. Anal., 11 (1991), pp. 181–204.

[17] I. S. Duff and J. K. Reid, *Exploiting zeros on the diagonal in the direct solution of indefinite sparse symmetric linear systems*, ACM Trans. Math. Software, 22 (1996), pp. 227–257.

[18] N. Dyn and W. E. Ferguson, Jr., *The numerical solution of equality constrained quadratic programming problems*, Math. Comp., 41 (1983), pp. 165–170.

[19] S. C. Eisenstat, H. C. Elman, and M. H. Schultz, *Variational iterative methods for nonsymmetric systems of linear equations*, SIAM J. Numer. Anal., 20 (1983), pp. 345–357.

[20] H. C. Elman, *Preconditioners for saddle point problems arising in computational fluid dynamics*, Appl. Numer. Math., 43 (2002), pp. 75–89.

[21] H. C. Elman and G. H. Golub, *Inexact and preconditioned Uzawa algorithms for saddle point problems*, SIAM J. Numer. Anal., 31 (1994), pp. 1645–1661.

[22] H. C. Elman, D. J. Silvester, and A. J. Wathen, *Performance and analysis of saddle point preconditioners for the discrete steady-state Navier–Stokes equations*, Numer. Math., 90

(2002), pp. 665–688.
[23] R. E. Ewing, R. D. Lazarov, P. Lu, and P. S. Vassilevski, *Preconditioning indefinite systems arising from mixed finite element discretization of second-order elliptic problems*, in Preconditioned Conjugate Gradient Methods, Lecture Notes in Math. 1457, Springer-Verlag, Berlin, 1990, pp. 28–43.
[24] M. Fortin and R. Glowinski, *Augmented Lagrangian Methods: Application to the Numerical Solution of Boundary-Value Problems*, Stud. Math. Appl. 15, North-Holland, Amsterdam, 1983.
[25] P. E. Gill, W. Murray, D. B. Ponceleón, and M. A. Saunders, *Preconditioners for indefinite systems arising in optimization*, SIAM J. Matrix Anal. Appl., 13 (1992), pp. 292–311.
[26] P. E. Gill, W. Murray, and M. H. Wright, *Practical Optimization*, Academic Press, New York, 1981.
[27] R. Glowinski, *Numerical Methods for Nonlinear Variational Problems*, Springer-Verlag, New York, 1984.
[28] G. H. Golub and C. Greif, *On solving block-structured indefinite linear systems*, SIAM J. Sci. Comput., 24 (2003), pp. 2076–2092.
[29] G. H. Golub and D. Vanderstraeten, *On the preconditioning of matrices with skew-symmetric splittings*, Numer. Algorithms, 25 (2000), pp. 223–239.
[30] G. H. Golub and A. J. Wathen, *An iteration for indefinite systems and its application to the Navier–Stokes equations*, SIAM J. Sci. Comput., 19 (1998), pp. 530–539.
[31] G. H. Golub, X. Wu, and J.-Y. Yuan, *SOR-like methods for augmented systems*, BIT, 41 (2001), pp. 71–85.
[32] N. I. M. Gould, M. E. Hribar, and J. Nocedal, *On the solution of equality constrained quadratic programming problems arising in optimization*, SIAM J. Sci. Comput., 23 (2001), pp. 1376–1395.
[33] E. Haber and U. M. Ascher, *Preconditioned all-at-once methods for large, sparse parameter estimation problems*, Inverse Problems, 17 (2001), pp. 1847–1864.
[34] E. Haber, U. M. Ascher, and D. Oldenburg, *On optimization techniques for solving nonlinear inverse problems*, Inverse Problems, 16 (2000), pp. 1263–1280.
[35] J. C. Haws, *Preconditioning KKT Systems*, Ph.D. thesis, Department of Mathematics, North Carolina State University, Raleigh, NC, 2002.
[36] R. A. Horn and C. R. Johnson, *Matrix Analysis*, Cambridge University Press, Cambridge, 1985.
[37] R. A. Horn and C. R. Johnson, *Topics in Matrix Analysis*, Cambridge University Press, Cambridge, 1991.
[38] M. Huhtanen, *A Hermitian Lanczos method for normal matrices*, SIAM J. Matrix Anal. Appl., 23 (2002), pp. 1092–1108.
[39] C. Keller, N. I. M. Gould, and A. J. Wathen, *Constraint preconditioning for indefinite linear systems*, SIAM J. Matrix Anal. Appl., 21 (2000), pp. 1300–1317.
[40] J. G. Kemeny and J. L. Snell, *Finite Markov Chains*, D. Van Nostrand Co., New York, 1960.
[41] A. Klawonn, *Block-triangular preconditioners for saddle point problems with a penalty term*, SIAM J. Sci. Comput., 19 (1998), pp. 172–184.
[42] J. Liesen, E. de Sturler, A. Sheffer, Y. Aydin, and C. Siefert, *Preconditioners for indefinite linear systems arising in surface parameterization*, in Proceedings of the 10th International Meshing Round Table, Sandia National Laboratories, eds., 2001, pp. 71–81.
[43] L. Lukšan and J. Vlček, *Indefinitely preconditioned inexact Newton method for large sparse equality constrained non-linear programming problems*, Numer. Linear Algebra Appl., 5 (1998), pp. 219–247.
[44] J. Maryška, M. Rozložník, and M. Tůma, *Mixed-hybrid finite element approximation of the potential fluid flow problem*, J. Comput. Appl. Math., 63 (1995), pp. 383–392.
[45] J. Maryška, M. Rozložník, and M. Tůma, *Schur complement systems in the mixed-hybrid finite element approximation of the potential fluid flow problem*, SIAM J. Sci. Comput., 22 (2000), pp. 704–723.
[46] M. F. Murphy, G. H. Golub, and A. J. Wathen, *A note on preconditioning for indefinite linear systems*, SIAM J. Sci. Comput., 21 (2000), pp. 1969–1972.
[47] W. Niethammer and R. S. Varga, *Relaxation methods for non-Hermitian linear systems*, Results Math., 16 (1989), pp. 308–320.
[48] I. Perugia and V. Simoncini, *Block-diagonal and indefinite symmetric preconditioners for mixed finite element formulations*, Numer. Linear Algebra Appl., 7 (2000), pp. 585–616.
[49] R. J. Plemmons, *A parallel block iterative scheme applied to computations in structural analysis*, SIAM J. Algebraic Discrete Methods, 7 (1986), pp. 337–347.

[50] W. Ren and J. Zhao, *Iterative methods with preconditioners for indefinite systems*, J. Comput. Math., 17 (1999), pp. 89–96.
[51] T. Rusten and R. Winther, *A preconditioned iterative method for saddlepoint problems*, SIAM J. Matrix Anal. Appl., 13 (1992), pp. 887–904.
[52] Y. Saad, *A flexible inner-outer preconditioned GMRES algorithm*, SIAM J. Sci. Comput., 14 (1993), pp. 461–469.
[53] Y. Saad and M. H. Schultz, *GMRES: A generalized minimal residual algorithm for solving nonsymmetric linear systems*, SIAM J. Sci. Statist. Comput., 7 (1986), pp. 856–869.
[54] V. Sarin and A. Sameh, *An efficient iterative method for the generalized Stokes problem*, SIAM J. Sci. Comput., 19 (1998), pp. 206–226.
[55] V. Simoncini and M. Benzi, *Spectral properties of the Hermitian and skew-Hermitian splitting preconditioner for saddle point problems*, SIAM J. Matrix Anal. Appl., to appear.
[56] G. Strang, *Introduction to Applied Mathematics*, Wellesley-Cambridge Press, Wellesley, MA, 1986.
[57] J. C. Strikwerda, *An iterative method for solving finite difference approximations to the Stokes equations*, SIAM J. Numer. Anal., 21 (1984), pp. 447–458.
[58] R. S. Varga, *Matrix Iterative Analysis*, Prentice-Hall, Englewood Cliffs, NJ, 1962.
[59] E. L. Wachspress and G. J. Habetler, *An alternating-direction-implicit iteration technique*, J. Soc. Indust. Appl. Math., 8 (1960), pp. 403–424.
[60] J. S. Warsa, M. Benzi, T. A. Wareing, and J. E. Morel, *Preconditioning a Mixed Discontinuous Finite Element Method for Radiation Diffusion*, Technical Report LA-UR-01-4754, Los Alamos National Laboratory, 2001; Numer. Linear Algebra Appl., to appear.
[61] O. Widlund, *A Lanczos method for a class of nonsymmeric systems of linear equations*, SIAM J. Numer. Anal., 15 (1978), pp. 801–812.
[62] W. Zulehner, *Analysis of iterative methods for saddle point problems: A unified approach*, Math. Comp., 71 (2002), pp. 479–505.