Acta Numerica (2005), pp. 1–137
DOI: 10.1017/S0962492904000212
© Cambridge University Press, 2005
Printed in the United Kingdom

Numerical solution of saddle point problems

Michele Benzi*
Department of Mathematics and Computer Science,
Emory University, Atlanta, Georgia 30322, USA
E-mail: benzi@mathcs.emory.edu

Gene H. Golub†
Scientific Computing and Computational Mathematics Program,
Stanford University, Stanford,
California 94305-9025, USA
E-mail: golub@sccm.stanford.edu

Jörg Liesen‡
Institut für Mathematik, Technische Universität Berlin,
D-10623 Berlin, Germany
E-mail: liesen@math.tu-berlin.de

We dedicate this paper to Gil Strang on the occasion of his 70th birthday

Large linear systems of saddle point type arise in a wide variety of applications throughout computational science and engineering. Due to their indefiniteness and often poor spectral properties, such linear systems represent a significant challenge for solver developers. In recent years there has been a surge of interest in saddle point problems, and numerous solution techniques have been proposed for this type of system. The aim of this paper is to present and discuss a large selection of solution methods for linear systems in saddle point form, with an emphasis on iterative methods for large and sparse problems.

* Supported in part by the National Science Foundation grant DMS-0207599.
† Supported in part by the Department of Energy of the United States Government.
‡ Supported in part by the Emmy Noether Programm of the Deutsche Forschungsgemeinschaft.

https://doi.org/10.1017/S0962492904000212 Published online by Cambridge University Press

CONTENTS

1	Introduction	2
2	Applications leading to saddle point problems	5
3	Properties of saddle point matrices	14
4	Overview of solution algorithms	29
5	Schur complement reduction	30
6	Null space methods	32
7	Coupled direct solvers	40
8	Stationary iterations	43
9	Krylov subspace methods	49
10	Preconditioners	59
11	Multilevel methods	96
12	Available software	105
13	Concluding remarks	107
References		109

1. Introduction

In recent years, a large amount of work has been devoted to the problem of solving large linear systems in saddle point form. The reason for this interest is the fact that such problems arise in a wide variety of technical and scientific applications. For example, the ever-increasing popularity of mixed finite element methods in engineering fields such as fluid and solid mechanics has been a major source of saddle point systems (Brezzi and Fortin 1991, Elman, Silvester and Wathen 2005*c*). Another reason for this surge in interest is the extraordinary success of interior point algorithms in both linear and nonlinear optimization, which require at their heart the solution of a sequence of systems in saddle point form (Nocedal and Wright 1999, Wright 1992, Wright 1997).

Because of the ubiquitous nature of saddle point systems, methods and results on their numerical solution have appeared in a wide variety of books, journals and conference proceedings, justifying a comprehensive survey of the subject. The purpose of this article is to review many of the most promising solution methods, with an emphasis on iterative methods for large and sparse problems. Although many of these solvers have been developed with specific applications in mind (for example, Stokes-type problems in fluid dynamics), it is possible to discuss them in a fairly general setting using standard numerical linear algebra concepts, the most prominent being perhaps the Schur complement. Nevertheless, when choosing a preconditioner (or developing a new one), knowledge of the origin of the particular problem

https://doi.org/10.1017/S0962492904000212 Published online by Cambridge University Press

at hand is essential. Although no single 'best' method exists, very effective solvers have been developed for some important classes of problems. We therefore devote some space to a discussion of saddle point problems arising in a few selected applications.

It is hoped that the present survey will prove useful to practitioners who are looking for guidance in the choice of a solution method for their own application, to researchers in numerical linear algebra and scientific computing, and especially to graduate students as an introduction to this very rich and important subject.

Notation. We have used boldface to denote vectors in sections describing fluid dynamics applications, where this is traditional, but we have otherwise followed the standard practice of numerical linear algebra and employed no special notation.

1.1. Problem statement and classification

The subject of this paper is the solution of block 2×2 linear systems of the form

$$\begin{bmatrix} A & B_1^T \\ B_2 & -C \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} f \\ g \end{bmatrix}, \quad \text{or} \quad \mathcal{A}u = b, \tag{1.1}$$

$$A \in \mathbb{R}^{n \times n}, \quad B_1, B_2 \in \mathbb{R}^{m \times n}, \quad C \in \mathbb{R}^{m \times m} \quad \text{with} \quad n \geq m. \tag{1.2}$$

It is obvious that, under suitable partitioning, *any* linear system can be cast in the form (1.1)–(1.2). We explicitly exclude the case where A or one or both of B_1, B_2 are zero. When the linear system describes a (generalized) saddle point problem, the constituent blocks A, B_1, B_2 and C satisfy one or more of the following conditions:

C1 A is symmetric: $A = A^T$
C2 the *symmetric part* of A, $H \equiv \frac{1}{2}(A + A^T)$, is positive semidefinite
C3 $B_1 = B_2 = B$
C4 C is symmetric ($C = C^T$) and positive semidefinite
C5 $C = O$ (the zero matrix)

Note that C5 implies C4. The most basic case is obtained when *all* the above conditions are satisfied. In this case A is symmetric positive semidefinite and we have a *symmetric* linear system of the form

$$\begin{bmatrix} A & B^T \\ B & O \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} f \\ g \end{bmatrix}. \tag{1.3}$$

This system arises as the first-order optimality conditions for the following equality-constrained quadratic programming problem:

$$\min \ J(x) = \frac{1}{2} x^T A x - f^T x \tag{1.4}$$

$$\text{subject to} \ \ Bx = g. \tag{1.5}$$

https://doi.org/10.1017/S0962492904000212 Published online by Cambridge University Press

In this case the variable y represents the vector of Lagrange multipliers. Any solution (x_*, y_*) of (1.3) is a saddle point for the Lagrangian

$$\mathcal{L}(x, y) = \frac{1}{2} x^T A x - f^T x + (Bx - g)^T y,$$

hence the name 'saddle point problem' given to (1.3). Recall that a saddle point is a point $(x_*, y_*) \in \mathbb{R}^{n+m}$ that satisfies

$$\mathcal{L}(x_*, y) \le \mathcal{L}(x_*, y_*) \le \mathcal{L}(x, y_*) \quad \text{for any} \quad x \in \mathbb{R}^n \quad \text{and} \quad y \in \mathbb{R}^m,$$

or, equivalently,

$$\min_x \max_y \mathcal{L}(x, y) = \mathcal{L}(x_*, y_*) = \max_y \min_x \mathcal{L}(x, y).$$

Systems of the form (1.3) also arise in nonlinearly constrained optimization (sequential quadratic programming and interior point methods), in fluid dynamics (Stokes' problem), incompressible elasticity, circuit analysis, structural analysis, and so forth; see the next section for a discussion of applications leading to saddle point problems.

Another important special case is when conditions C1–C4 are satisfied, but not C5. In this case we have a block linear system of the form

$$\begin{bmatrix} A & B^T \\ B & -C \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} f \\ g \end{bmatrix}. \tag{1.6}$$

Problems of this kind frequently arise in the context of *stabilized* mixed finite element methods. Stabilization is used whenever the discrete variables x and y belong to finite element spaces that do not satisfy the Ladyzhenskaya–Babuška–Brezzi (LBB, or inf-sup) condition (Brezzi and Fortin 1991). Another situation leading to a nonzero C is the discretization of the equations describing slightly compressible fluids or solids (Braess 2001, Chapter 6.3). Systems of the form (1.6) also arise from regularized, weighted least-squares problems (Benzi and Ng 2004) and from certain interior point methods in optimization (Wright 1992, Wright 1997). Often the matrix C has small norm compared to the other blocks.

In the literature, the phrase *generalized saddle point problem* has been used primarily to allow for the possibility of a nonsymmetric matrix $\mathcal{A}$ in (1.1). In such problems either $A \neq A^T$ (with condition C2 usually satisfied), or $B_1 \neq B_2$, or both. The most important example is perhaps that of the linearized Navier–Stokes equations, where linearization has been obtained by Picard iteration or by some variant of Newton's method. See Ciarlet Jr., Huang and Zou (2003), Nicolaides (1982) and Szyld (1981) for additional examples. We note that our definition of generalized saddle point problem as a linear system of the form (1.1)–(1.2), where the blocks A, B_1, B_2 and C satisfy one or more of the conditions C1–C5, is the most general possible, and it contains previous definitions as special cases.

https://doi.org/10.1017/S0962492904000212 Published online by Cambridge University Press

In the vast majority of cases, linear systems of saddle point type have real coefficients, and in this paper we restrict ourselves to the real case. Complex coefficient matrices, however, do arise in some cases; see, *e.g.*, Bobrovnikova and Vavasis (2000), Mahawar and Sarin (2003) and Strang (1986, page 117). Most of the results and algorithms reviewed in this paper admit straightforward extensions to the complex case.

1.2. Sparsity, structure and size

Although saddle point systems come in all sizes and with widely different structural and sparsity properties, in this paper we are mainly interested in problems that are both large and sparse. This justifies our emphasis on iterative solvers. Direct solvers, however, are still the preferred method in optimization and other areas. Furthermore, direct methods are often used in the solution of subproblems, for example as part of a preconditioner solve. Some of the algorithms considered in this paper are also applicable if one or more of the blocks in $\mathcal{A}$ happen to be dense, as long as matrix-vector products with $\mathcal{A}$ can be performed efficiently, typically in $O(n+m)$ time. This means that if a dense block is present, it must have a special structure (*e.g.*, Toeplitz, as in Benzi and Ng (2004) and Jin (1996)) or it must be possible to approximate its action on a vector with (nearly) linear complexity, as in the fast multipole method (Mahawar and Sarin 2003).

Frequently, the matrices that arise in practice have quite a bit of structure. For instance, the A block is often block diagonal, with each diagonal block endowed with additional structure. Many of the algorithms discussed in this paper are able to exploit the structure of the problem to gain efficiency and save on storage. Sometimes the structure of the problem suggests solution algorithms that have a high degree of parallelism. This last aspect, however, is not emphasized in this paper. Finally we mention that in most applications n is larger than m, often much larger.

2. Applications leading to saddle point problems

As already mentioned, large-scale saddle point problems occur in many areas of computational science and engineering. The following is a list of some fields where saddle point problems naturally arise, together with some references:

- computational fluid dynamics (Glowinski 1984, Quarteroni and Valli 1994, Temam 1984, Turek 1999, Wesseling 2001)
- constrained and weighted least squares estimation (Björck 1996, Golub and Van Loan 1996)
- constrained optimization (Gill, Murray and Wright 1981, Wright 1992, Wright 1997)

https://doi.org/10.1017/S0962492904000212 Published online by Cambridge University Press

- economics (Arrow, Hurwicz and Uzawa 1958, Duchin and Szyld 1979, Leontief, Duchin and Szyld 1985, Szyld 1981)
- electrical circuits and networks (Bergen 1986, Chua, Desoer and Kuh 1987, Strang 1986, Tropper 1962)
- electromagnetism (Bossavit 1998, Perugia 1997, Perugia, Simoncini and Arioli 1999)
- finance (Markowitz 1959, Markowitz and Perold 1981)
- image reconstruction (Hall 1979)
- image registration (Haber and Modersitzki 2004, Modersitzki 2003)
- interpolation of scattered data (Lyche, Nilssen and Winther 2002, Sibson and Stone 1991)
- linear elasticity (Braess 2001, Ciarlet 1988)
- mesh generation for computer graphics (Liesen, de Sturler, Sheffer, Aydin and Siefert 2001)
- mixed finite element approximations of elliptic PDEs (Brezzi 1974, Brezzi and Fortin 1991, Quarteroni and Valli 1994)
- model order reduction for dynamical systems (Freund 2003, Heres and Schilders 2005, Stykel 2005)
- optimal control (Battermann and Heinkenschloss 1998, Battermann and Sachs 2001, Betts 2001, Biros and Ghattas 2000, Nguyen 2004)
- parameter identification problems (Burger and Mühlhuber 2002, Haber and Ascher 2001, Haber, Ascher and Oldenburg 2000).

Quite often, saddle point systems arise when a certain quantity (such as the energy of a physical system) has to be minimized, subject to a set of linear constraints. In this case the Lagrange multiplier y usually has a physical interpretation and its computation is also of interest. For example, in incompressible flow problems x is a vector of velocities and y a vector of pressures. In the complementary energy formulation of structural mechanics x is the vector of internal forces, y represents the nodal displacements of the structure. For resistive electrical networks y represents the nodal potentials, x being the vector of currents.

In some cases, such as fluid dynamics or linear elasticity, saddle point problems result from the discretization of systems of partial differential equations with constraints. Typically the constraints represent some basic conservation law, such as mass conservation in fluid dynamics. In other cases, such as resistive electrical networks or structural analysis, the equations are discrete to begin with. Now the constraints may correspond to the topology (connectivity) of the system being studied. Because saddle point equations can be derived as equilibrium conditions for a physical system, they are sometimes called *equilibrium equations*. See Strang (1986, 1988) for a very nice discussion of equilibrium equations throughout applied mathematics. Another popular name for saddle point systems, especially in the

https://doi.org/10.1017/S0962492904000212 Published online by Cambridge University Press

optimization literature, is 'KKT system', from the Karush–Kuhn–Tucker first-order optimality conditions; see Nocedal and Wright (1999, page 328) for precise definitions, and Golub and Greif (2003) and Kjeldsen (2000) for historical notes.

Systems of the form (1.1)–(1.2) also arise from non-overlapping domain decomposition when interface unknowns are numbered last, as well as from FETI-type schemes when Lagrange multipliers are used to ensure continuity at the interfaces; see for instance Chan and Mathew (1994), Farhat and Roux (1991), Hu, Shi and Yu (2004), Quarteroni and Valli (1999) and Toselli and Widlund (2004).

It is of course not possible for us to cover here all these different applications. We choose instead to give some details about three classes of problems leading to saddle point systems. The first comes from the field of computational fluid dynamics, the second from least squares estimation, and the third one from interior point methods in constrained optimization.

2.1. Incompressible flow problems

We begin with the (steady-state) Navier–Stokes equations governing the flow of a Newtonian, incompressible viscous fluid. Let $\Omega \subset \mathbb{R}^d$ $(d = 2, 3)$ be a bounded, connected domain with a piecewise smooth boundary Γ. Given a force field $\mathbf{f} : \Omega \to \mathbb{R}^d$ and boundary data $\mathbf{g} : \Gamma \to \mathbb{R}^d$, the problem is to find a velocity field $\mathbf{u} : \Omega \to \mathbb{R}^d$ and a pressure field $p : \Omega \to \mathbb{R}$ such that

$$-\nu\Delta\mathbf{u} + (\mathbf{u} \cdot \nabla)\,\mathbf{u} + \nabla p = \mathbf{f} \quad \text{in} \quad \Omega, \tag{2.1}$$

$$\nabla \cdot \mathbf{u} = 0 \quad \text{in} \quad \Omega, \tag{2.2}$$

$$\mathcal{B}\mathbf{u} = \mathbf{g} \quad \text{on} \quad \Gamma, \tag{2.3}$$

where $\nu > 0$ is the kinematic viscosity coefficient (inversely proportional to the Reynolds number Re), Δ is the Laplace operator in $\mathbb{R}^d$, ∇ denotes the gradient, $\nabla\cdot$ is the divergence, and $\mathcal{B}$ is some type of boundary operator (*e.g.*, a trace operator for Dirichlet boundary conditions). To determine p uniquely we may impose some additional condition, such as

$$\int_\Omega p \, \mathrm{d}\mathbf{x} = 0.$$

Equation (2.1) represents conservation of momentum, while equation (2.2) represents the incompressibility condition, or mass conservation. Owing to the presence of the convective term $(\mathbf{u} \cdot \nabla)\,\mathbf{u}$ in the momentum equations, the Navier–Stokes system is nonlinear. It can be linearized in various ways. An especially popular linearization process is the one based on Picard's iteration; see, *e.g.*, Elman *et al.* (2005*c*, Section 7.2.2). Starting with an initial guess $\mathbf{u}^{(0)}$ (with $\nabla \cdot \mathbf{u}^{(0)} = 0$) for the velocity field, Picard's iteration constructs a sequence of approximate solutions $(\mathbf{u}^{(k)}, p^{(k)})$ by solving the

https://doi.org/10.1017/S0962492904000212 Published online by Cambridge University Press

linear *Oseen problem*

$$-\nu\Delta\mathbf{u}^{(k)} + (\mathbf{u}^{(k-1)}\cdot\nabla)\,\mathbf{u}^{(k)} + \nabla p^{(k)} = \mathbf{f} \quad \text{in} \quad \Omega, \tag{2.4}$$

$$\nabla\cdot\mathbf{u}^{(k)} = 0 \quad \text{in} \quad \Omega, \tag{2.5}$$

$$\mathcal{B}\mathbf{u}^{(k)} = \mathbf{g} \quad \text{on} \quad \Gamma \tag{2.6}$$

($k = 1, 2, \ldots$). Note that no initial pressure needs to be specified. Under certain conditions on ν (which should not be too small) and $\mathbf{f}$ (which should not be too large in an appropriate norm), the steady Navier–Stokes equations (2.1)–(2.3) have a unique solution $(\mathbf{u}_*, p_*)$ and the iterates $(\mathbf{u}^{(k)}, p^{(k)})$ converge to it as $k \to \infty$ for any choice of the initial velocity $\mathbf{u}^{(0)}$. We refer to Girault and Raviart (1986) for existence and uniqueness results and to Karakashian (1982) for a proof of the global convergence of Picard's iteration.

Hence, at each Picard iteration one needs to solve an Oseen problem of the form

$$-\nu\Delta\mathbf{u} + (\mathbf{v}\cdot\nabla)\,\mathbf{u} + \nabla p = \mathbf{f} \quad \text{in} \quad \Omega, \tag{2.7}$$

$$\nabla\cdot\mathbf{u} = 0 \quad \text{in} \quad \Omega, \tag{2.8}$$

$$\mathcal{B}\mathbf{u} = \mathbf{g} \quad \text{on} \quad \Gamma \tag{2.9}$$

with a known, divergence-free coefficient $\mathbf{v}$. Discretization of (2.7)–(2.9) using, *e.g.*, finite differences (Peyret and Taylor 1983) or finite elements (Elman *et al.* 2005*c*, Quarteroni and Valli 1994) results in a generalized saddle point system of the form (1.6), in which x represents the discrete velocities and y the discrete pressure. Here $A = \mathrm{diag}(A_1, \ldots, A_d)$ is a block diagonal matrix, where each block corresponds to a discrete convection-diffusion operator with the appropriate boundary conditions. Note that A is nonsymmetric, but satisfies condition C2 when an appropriate (conservative) discretization is used. The rectangular matrix B^T represents the discrete gradient operator while B represents its adjoint, the (negative) divergence operator. A nonzero C may be present if stabilization is used.

The important special case $\mathbf{v} = \mathbf{0}$ corresponds to the (steady-state) Stokes equations:

$$-\Delta\mathbf{u} + \nabla p = \mathbf{f} \quad \text{in} \quad \Omega, \tag{2.10}$$

$$\nabla\cdot\mathbf{u} = 0 \quad \text{in} \quad \Omega, \tag{2.11}$$

$$\mathcal{B}\mathbf{u} = \mathbf{g} \quad \text{on} \quad \Gamma. \tag{2.12}$$

Note that without loss of generality we have set $\nu = 1$, since we can always divide the momentum equation by ν and rescale the pressure p and the forcing term $\mathbf{f}$ by ν. The Stokes equations can be interpreted as the Euler–Lagrange partial differential equations for the constrained variational

https://doi.org/10.1017/S0962492904000212 Published online by Cambridge University Press

problem

$$\min \ J(\mathbf{u}) = \frac{1}{2}\int_\Omega \|\nabla \mathbf{u}\|_2^2 \,\mathrm{d}\mathbf{x} - \int_\Omega \mathbf{f}\cdot\mathbf{u}\,\mathrm{d}\mathbf{x} \tag{2.13}$$

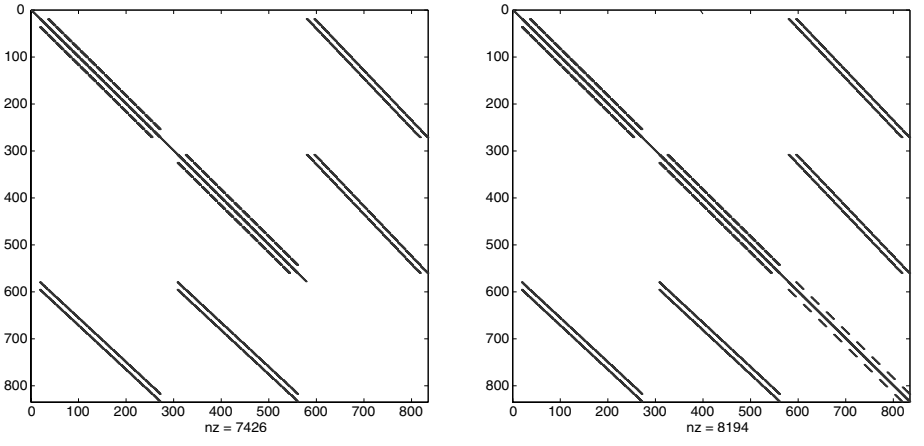
$$\text{subject to} \ \ \nabla\cdot\mathbf{u} = 0 \tag{2.14}$$

(see, *e.g.*, Gresho and Sani (1998, page 636)). Throughout this paper, $\|\mathbf{u}\|_2 = \sqrt{\mathbf{u}\cdot\mathbf{u}}$ denotes the Euclidean norm of the vector $\mathbf{u}$. Here the pressure p plays the role of the Lagrange multiplier. The Stokes equations describe the flow of a slow-moving, highly viscous fluid. They also arise as subproblems in the numerical solution of the Navier–Stokes equations by operator splitting methods (Glowinski 2003, Quarteroni and Valli 1994) and as the first step of Picard's iteration when the initial guess used is $\mathbf{u}^{(0)} = \mathbf{0}$.

Appropriate discretization of the Stokes system leads to a symmetric saddle point problem of the form (1.3) where A is now a block diagonal matrix, and each of its d diagonal blocks is a discretization of the Laplace operator $-\Delta$ with the appropriate boundary conditions. Thus, A is now symmetric and positive (semi-)definite. Again, a nonzero C may be present if stabilization is used. Typical sparsity patterns for $\mathcal{A}$ are displayed in Figure 2.1.

An alternative linearization of the Navier–Stokes equations can be derived on the basis of the identity

$$(\mathbf{u}\cdot\nabla)\mathbf{u} = \frac{1}{2}\nabla(\|\mathbf{u}\|_2^2) - \mathbf{u}\times(\nabla\times\mathbf{u}).$$



(a) without stabilization (b) with stabilization

Figure 2.1. Sparsity patterns for two-dimensional Stokes problem (leaky lid-driven cavity) using Q1-P0 discretization.

https://doi.org/10.1017/S0962492904000212 Published online by Cambridge University Press

See, for instance, Landau and Lifschitz (1959, page 5) or Chorin and Marsden (1990, page 47) as well as the discussion in Gresho (1991) and Gresho and Sani (1998). The corresponding linearized equations take the form

$$-\nu\Delta\mathbf{u} + \mathbf{w}\times\mathbf{u} + \nabla P = \mathbf{f} \quad \text{in} \quad \Omega, \tag{2.15}$$

$$\nabla\cdot\mathbf{u} = 0 \quad \text{in} \quad \Omega, \tag{2.16}$$

$$\mathcal{B}\mathbf{u} = \mathbf{g} \quad \text{on} \quad \Gamma, \tag{2.17}$$

where, for the two-dimensional case,

- $(\mathbf{w}\times) = \left(\begin{smallmatrix} 0 & w \\ -w & 0 \end{smallmatrix}\right)$
- $w = \nabla\times\mathbf{v} = -\frac{\partial v_1}{\partial x_2} + \frac{\partial v_2}{\partial x_1}$
- $P = p + \frac{1}{2}\|\mathbf{v}\|_2^2$ (the so-called *Bernoulli pressure*)

Here the divergence-free vector field $\mathbf{v}$ again denotes the approximate velocity from the previous Picard iteration. See Olshanskii (1999) for the three-dimensional case. Note that when the 'wind' function $\mathbf{v}$ is irrotational ($\nabla\times\mathbf{v} = 0$), equations (2.15)–(2.17) reduce to the Stokes problem. It is worth stressing that the linearizations (2.7)–(2.9) and (2.15)–(2.17), although both conservative (Olshanskii 1999, page 357), are not mathematically equivalent. The so-called *rotation form* (2.15) of the momentum equations, although popular in fluid mechanics, has not been widely known among numerical analysts until the recent work by Olshanskii and co-workers (Lube and Olshanskii 2002, Olshanskii 1999, Olshanskii and Reusken 2002), which showed its advantages over the standard (convective) form. We return to this in Section 10.3.

A related problem, also leading to large sparse linear systems in saddle point form upon discretization, is the potential fluid flow problem in porous media, often used to model groundwater contamination (Bear 1972, Maryška, Rozložník and Tůma 1995). This consists of a boundary value problem for a system of first-order partial differential equations representing, respectively, Darcy's Law for the velocity field $\mathbf{u}$ and the continuity equation:

$$K\mathbf{u} + \nabla p = 0 \quad \text{in} \quad \Omega, \tag{2.18}$$

$$\nabla\cdot\mathbf{u} = q \quad \text{in} \quad \Omega, \tag{2.19}$$

$$p = p_D \quad \text{on } \Gamma_D, \quad \mathbf{u}\cdot\mathbf{n} = u_N \quad \text{on} \quad \Gamma_N, \tag{2.20}$$

where p is a piezometric potential (fluid pressure), K is the symmetric and uniformly positive definite second-rank tensor of hydraulic permeability of the medium, and q represents density of potential sources (or sinks) in the medium. Here Γ_D and Γ_N are subsets of the boundary Γ of the bounded connected flow domain Ω, with $\Gamma = \bar{\Gamma}_D \cup \bar{\Gamma}_N$, $\Gamma_D \neq \emptyset$, and $\Gamma_D \cap \Gamma_N = \emptyset$; $\mathbf{n}$ is the outward normal vector defined (a.e.) on Γ. When discretized by mixed

https://doi.org/10.1017/S0962492904000212 Published online by Cambridge University Press

finite elements (Raviart–Thomas elements being a very popular choice for this problem), a linear system of the type (1.3) is obtained. The symmetric positive definite matrix A is now a discretization of the linear operator 'multiplication by K', a zeroth-order differential operator. The conditioning properties of A are independent of the discretization parameter h (for most discretizations), and depend only on properties of the hydraulic permeability tensor K. The matrix $-B$ represents, again, a discrete divergence operator and B^T a discrete gradient. We note that modelling the interaction between surface and subsurface flows leads to coupled (Navier–) Stokes and Darcy systems (Discacciati, Miglio and Quarteroni 2002, Discacciati and Quarteroni 2004). Problem (2.18)–(2.20) is just one example of a first-order system formulation of a second-order linear elliptic PDE (Brezzi and Fortin 1991). Saddle point systems also arise from mixed formulation of fourth-order (biharmonic) elliptic problems (Glowinski and Pironneau 1979).

In the course of this brief discussion we have restricted ourselves to stationary (steady-state) problems. The unsteady (time-dependent) case leads to sequences of saddle point systems when fully implicit time-stepping schemes are used, for example when the time derivative $\mathbf{u}_t$ is discretized using backward Euler or Crank–Nicolson schemes; see, *e.g.*, Turek (1999, Chapter 2). In the case of Stokes and Oseen, the resulting semi-discrete systems are often referred to as *generalized* Stokes and Oseen problems. The literature on numerical methods for incompressible flow problems is vast; see, *e.g.*, Elman *et al.* (2005*c*), Fortin (1993), Glowinski (2003), Gresho and Sani (1998), Gunzburger (1989), Quarteroni and Valli (1994), Temam (1984), Turek (1999) and Wesseling (2001).

2.2. Constrained and weighted least squares

Linear systems of saddle point type commonly arise when solving least squares problems. Consider the following least squares problem with linear equality constraints:

$$\min_x \ \|c - Gy\|_2 \tag{2.21}$$

$$\text{subject to} \ \ Ey = d, \tag{2.22}$$

where $c \in \mathbb{R}^p$, $G \in \mathbb{R}^{p\times m}$, $y \in \mathbb{R}^m$, $E \in \mathbb{R}^{q\times m}$, $d \in \mathbb{R}^q$ and $q < m$. Problems of this kind arise, for instance, in curve or surface fitting when the curve is required to interpolate certain data points; see Björck (1996, Chapter 5). The optimality conditions for problem (2.21)–(2.22) are

$$\begin{bmatrix} I_p & O & G \\ O & O & E \\ G^T & E^T & O \end{bmatrix} \begin{bmatrix} r \\ \lambda \\ y \end{bmatrix} = \begin{bmatrix} c \\ d \\ 0 \end{bmatrix}, \tag{2.23}$$

https://doi.org/10.1017/S0962492904000212 Published online by Cambridge University Press

where I_p is the $p \times p$ identity matrix and $\lambda \in \mathbb{R}^q$ is a vector of Lagrange multipliers. Clearly, (2.23) is a special case of the symmetric saddle point problem (1.3).

Next we consider the generalized linear least squares problem

$$\min_x (f - Gx)^T W^{-1} (f - Gx), \tag{2.24}$$

where $f \in \mathbb{R}^n$, $G \in \mathbb{R}^{n \times m}$ with $m < n$, and the (scaled) covariance matrix $W \in \mathbb{R}^{n \times n}$ is symmetric positive definite; see Björck (1996, Chapter 4). It is readily seen that (2.24) is equivalent to a standard saddle point system of the form (1.3) with $A = W$, $B = G^T$ and $g = 0$. This saddle point formulation of the generalized least squares problem is often referred to as the *augmented system* formulation. The matrix W is diagonal when the errors in the observations are uncorrelated. When $W = I$ (the $n \times n$ identity matrix) we have the usual linear least squares problem. An advantage of the augmented system formulation over the standard one is that the former allows for the case in which W is singular, which is important in some applications. In other words, W could be positive semidefinite rather than definite. Some applications even lead to an indefinite weight matrix W; see Bojanczyk, Higham and Patel (2003), Chandrasekaran, Gu and Sayed (1998), Hassibi, Sayed and Kailath (1996).

Finally, if the original problem is ill-posed and Tikhonov regularization is applied, one obtains a saddle point system of the form (1.6) with $C = \gamma L^T L$ where $\gamma > 0$ is the regularization parameter and L is either the $m \times m$ identity or some type of smoothing operator, such as a first-order finite difference operator. See Benzi and Ng (2004) and Gonzales and Woods (1992) for applications in image processing.

2.3. Saddle point systems from interior point methods

Here we show how saddle point systems arise when interior point methods are used to solve constrained optimization problems. Our presentation is based on the nice synopsis given in Bergamaschi, Gondzio and Zilli (2004). Consider a convex nonlinear programming problem,

$$\min \ f(x) \tag{2.25}$$

$$\text{subject to} \ \ c(x) \le 0, \tag{2.26}$$

where $f : \mathbb{R}^n \to \mathbb{R}$ and $c : \mathbb{R}^n \to \mathbb{R}^m$ are convex and twice differentiable. Introducing a nonnegative slack variable $z \in \mathbb{R}^m$, we can write the inequality constraint as the system of equalities $c(x) + z = 0$, and we can introduce the associated barrier problem:

$$\min \ f(x) - \mu \sum_{i=1}^{m} \ln z_i \tag{2.27}$$

$$\text{subject to} \ \ c(x) + z = 0. \tag{2.28}$$

https://doi.org/10.1017/S0962492904000212 Published online by Cambridge University Press

The corresponding Lagrangian is

$$\mathcal{L}(x, y, z; \mu) = f(x) + y^T(c(x) + z) - \mu \sum_{i=1}^{m} \ln z_i.$$

To find a stationary point of the Lagrangian we set

$$\nabla_x \mathcal{L}(x, y, z; \mu) = \nabla f(x) + \nabla c(x)^T y = 0, \tag{2.29}$$
$$\nabla_y \mathcal{L}(x, y, z; \mu) = c(x) + z = 0, \tag{2.30}$$
$$\nabla_z \mathcal{L}(x, y, z; \mu) = y - \mu Z^{-1} e = 0, \tag{2.31}$$

where $Z = \operatorname{diag}(z_1, z_2, \ldots, z_m)$ and $e = [1\,1\,\ldots\,1]^T$. Introducing the diagonal matrix $Y = \operatorname{diag}(y_1, y_2, \ldots, y_m)$, the first-order optimality conditions for the barrier problem become

$$\nabla f(x) + \nabla c(x)^T y = 0, \tag{2.32}$$
$$c(x) + z = 0, \tag{2.33}$$
$$YZe = \mu e, \tag{2.34}$$
$$y, z \geq 0. \tag{2.35}$$

This is a nonlinear system of equations with nonnegativity constraints and it can be solved by Newton's method. The barrier parameter μ is gradually reduced so as to ensure convergence of the iterates to the optimal solution of problem (2.25)–(2.26). At each Newton iteration, it is necessary to solve a linear system of the form

$$\begin{bmatrix} H(x,y) & B(x)^T & O \\ B(x) & O & I \\ O & Z & Y \end{bmatrix} \begin{bmatrix} \delta x \\ \delta y \\ \delta z \end{bmatrix} = \begin{bmatrix} -\nabla f(x) - B(x)^T y \\ -c(x) - z \\ \mu e - YZe \end{bmatrix}, \tag{2.36}$$

where

$$H(x, y) = \nabla^2 f(x) + \sum_{i=1}^{m} y_i \nabla^2 c_i(x) \in \mathbb{R}^{n \times n} \quad \text{and} \quad B(x) = \nabla c(x) \in \mathbb{R}^{m \times n}.$$

Here $\nabla^2 f(x)$ denotes the Hessian of f evaluated at x. The linear system (2.36) can be reduced to one of smaller dimensions by using the third equation to eliminate $\delta z = \mu Y^{-1} e - Ze - ZY^{-1}\delta y$ from the second equation. The resulting system is

$$\begin{bmatrix} -H(x,y) & B(x)^T \\ B(x) & ZY^{-1} \end{bmatrix} \begin{bmatrix} \delta x \\ -\delta y \end{bmatrix} = \begin{bmatrix} \nabla f(x) + B(x)^T y \\ -c(x) - \mu Y^{-1} e \end{bmatrix}. \tag{2.37}$$

Apart from the sign, (2.37) is a saddle point system of the form (1.6). If the objective function $f(x)$ and the constraints $c_i(x)$ are convex, the symmetric matrix $H(x, y)$ is positive semidefinite, and it is positive definite if $f(x)$ is strictly convex. The diagonal matrix ZY^{-1} is obviously posi-

https://doi.org/10.1017/S0962492904000212 Published online by Cambridge University Press

tive semidefinite. The coefficient matrix in (2.37) depends on the current approximation (x, y), and it changes at each Newton step.

Similar linear systems arise when interior point methods are used to solve linear and quadratic programming problems. Now the systems to be solved at each Newton iteration are of the form

$$\begin{bmatrix} -H - D & B^T \\ B & O \end{bmatrix} \begin{bmatrix} \delta x \\ \delta y \end{bmatrix} = \begin{bmatrix} \xi \\ \eta \end{bmatrix},$$

where the $n \times n$ matrix H is symmetric positive semidefinite if the problem is convex and D is a (positive) diagonal matrix. Now H and B remain constant ($H \equiv O$ in linear programming), while D changes at each Newton iteration.

There are many interesting linear algebra problems arising from the use of interior point methods; see in particular Bergamaschi *et al.* (2004), Czyzyk, Fourer and Mehrotra (1998), Forsgren, Gill and Shinnerl (1996), Fourer and Mehrotra (1993), Frangioni and Gentile (2004), Freund and Jarre (1996), Gill, Murray, Ponceleón and Saunders (1992), Nocedal and Wright (1999), Oliveira and Sorensen (2005), Wright (1992) and Wright (1997).

3. Properties of saddle point matrices

This section is devoted to establishing basic algebraic properties of the saddle point matrix $\mathcal{A}$ such as existence of various factorizations, invertibility, spectral properties, and conditioning. Knowledge of these properties is important in the development of solution algorithms.

3.1. Block factorizations and the Schur complement

If A is nonsingular, the saddle point matrix $\mathcal{A}$ admits the following block triangular factorization:

$$\mathcal{A} = \begin{bmatrix} A & B_1^T \\ B_2 & -C \end{bmatrix} = \begin{bmatrix} I & O \\ B_2 A^{-1} & I \end{bmatrix} \begin{bmatrix} A & O \\ O & S \end{bmatrix} \begin{bmatrix} I & A^{-1} B_1^T \\ O & I \end{bmatrix}, \tag{3.1}$$

where $S = -(C + B_2 A^{-1} B_1^T)$ is the *Schur complement* of A in $\mathcal{A}$. A number of important properties of the saddle point matrix $\mathcal{A}$ can be derived on the basis of (3.1): we do this in the next three subsections.

Also useful are the equivalent factorizations

$$\mathcal{A} = \begin{bmatrix} A & O \\ B_2 & S \end{bmatrix} \begin{bmatrix} I & A^{-1} B_1^T \\ O & I \end{bmatrix} \tag{3.2}$$

and

$$\mathcal{A} = \begin{bmatrix} I & O \\ B_2 A^{-1} & I \end{bmatrix} \begin{bmatrix} A & B_1^T \\ O & S \end{bmatrix}. \tag{3.3}$$

https://doi.org/10.1017/S0962492904000212 Published online by Cambridge University Press

The assumption that A is nonsingular may appear to be rather restrictive, since A is singular in many applications; see, *e.g.*, Haber and Ascher (2001). However, one can use augmented Lagrangian techniques (Fortin and Glowinski 1983, Glowinski and Le Tallec 1989, Golub and Greif 2003, Greif, Golub and Varah 2005) to replace the original saddle point system with an equivalent one having the same solution but in which the $(1,1)$ block A is now nonsingular. Hence, no great loss of generality is incurred. We shall return to augmented Lagrangian techniques in Section 3.5.

Besides being useful for deriving theoretical properties of saddle point matrices, the decompositions (3.1)–(3.3) are also the basis for many of the most popular solution algorithms for saddle point systems, as we shall see.

3.2. Solvability conditions

Assuming A is nonsingular, it readily follows from any of the block decompositions (3.1)–(3.3) that $\mathcal{A}$ is nonsingular if and only if S is. Unfortunately, very little can be said in general about the invertibility of the Schur complement $S = -(C + B_2 A^{-1} B_1^T)$. It is necessary to place some restrictions on the matrices A, B_1, B_2 and C.

Symmetric case

We begin with the standard saddle point system (1.3), where A is symmetric positive definite, $B_1 = B_2$ and $C = O$. In this case the Schur complement reduces to $S = -BA^{-1}B^T$, a symmetric negative semidefinite matrix. It is obvious that S, and thus $\mathcal{A}$, is invertible if and only if B^T has full column rank (hence, if and only if $\operatorname{rank}(B) = m$), since in this case S is symmetric negative definite. Then both problems (1.3) and (1.4)–(1.5) have a unique solution: if (x_*, y_*) is the solution of (1.3), x_* is the unique solution of (1.4)–(1.5). It can be shown that x_* is the A-orthogonal projection of the solution $\hat{x} = A^{-1}f$ of the unconstrained problem (1.4) onto the constraint set $\mathcal{C} = \{x \in \mathbb{R}^n \,|\, Bx = g\}$. Here A-orthogonal means orthogonal with respect to the inner product $\langle v, w \rangle_A \equiv w^T A v$. We will discuss this in more detail in Section 3.3.

Next we consider the case where A is symmetric positive definite, $B_1 = B_2 = B$, and $C \neq O$ is symmetric positive semidefinite. Then again $S = -(C + BA^{-1}B^T)$ is symmetric negative semidefinite, and it is negative definite (hence, invertible) if and only if $\ker(C) \cap \ker(B^T) = \{0\}$. Obvious sufficient conditions for invertibility are that C be positive definite or that B have full row rank. We can summarize our discussion so far in the following theorem.

Theorem 3.1. Assume A is symmetric positive definite, $B_1 = B_2 = B$, and C is symmetric positive semidefinite. If $\ker(C) \cap \ker(B^T) = \{0\}$, then

https://doi.org/10.1017/S0962492904000212 Published online by Cambridge University Press

the saddle point matrix $\mathcal{A}$ is nonsingular. In particular, $\mathcal{A}$ is invertible if B has full rank.

Now we relax the condition that A be positive definite. If A is indefinite, the following simple example shows that $\mathcal{A}$ may be singular, even if B has full rank:

$$\mathcal{A} = \left[\begin{array}{cc|c} 1 & 0 & -1 \\ 0 & -1 & 1 \\ \hline -1 & 1 & 0 \end{array}\right] = \begin{bmatrix} A & B^T \\ B & O \end{bmatrix}.$$

However, $\mathcal{A}$ will be invertible if A is positive definite on $\ker(B)$. When A is symmetric positive semidefinite, we have the following result (see, *e.g.*, the discussion of quadratic programming in Hadley (1964) or Luenberger (1984, page 424)). Although this is a well-known result, we include a proof to make our treatment more self-contained.

Theorem 3.2. Assume that A is symmetric positive semidefinite, $B_1 = B_2 = B$ has full rank, and $C = O$. Then a necessary and sufficient condition for the saddle point matrix $\mathcal{A}$ to be nonsingular is $\ker(A) \cap \ker(B) = \{0\}$.

Proof. Let $u = \begin{bmatrix} x \\ y \end{bmatrix}$ be such that $\mathcal{A}u = 0$. Hence, $Ax + B^T y = 0$ and $Bx = 0$. It follows that $x^T Ax = -x^T B^T y = -(Bx)^T y = 0$. Since A is symmetric positive semidefinite, $x^T Ax = 0$ implies $Ax = 0$ (see Horn and Johnson (1985, page 400)), and therefore $x \in \ker(A) \cap \ker(B)$, thus $x = 0$. Also, $y = 0$ since $B^T y = 0$ and B^T has full column rank. Therefore $u = 0$, and $\mathcal{A}$ is nonsingular. This proves the sufficiency of the condition.

Assume now that $\ker(A) \cap \ker(B) \neq \{0\}$. Taking $x \in \ker(A) \cap \ker(B)$, $x \neq 0$ and letting $u = \begin{bmatrix} x \\ 0 \end{bmatrix}$ we have $\mathcal{A}u = 0$, implying that $\mathcal{A}$ is singular. Hence, the condition is also necessary. □

Remark. It is clear from the proof of this theorem that the requirement that A be positive semidefinite can be somewhat relaxed: it suffices that A be definite on $\ker(B)$. In fact, all we need is that $x^T Ax \neq 0$ for $x \in \ker(B)$, $x \neq 0$. This implies that A is either positive definite or negative definite on $\ker(B)$. In any case, the rank of A must be at least $n - m$ for $\mathcal{A}$ to be nonsingular.

How restrictive is the assumption that B has full rank? A rank deficient B signifies that some of the constraints are redundant. It is generally easy to eliminate this redundancy. For instance, in the Stokes and Oseen case, where B^T represents a discrete gradient, a one-dimensional subspace (containing all the constant vectors) is often present. Hence $\mathcal{A}$ has one zero eigenvalue, corresponding to the so-called *hydrostatic pressure mode*, due to the fact that the pressure is defined up to a constant. A similar situation

https://doi.org/10.1017/S0962492904000212 Published online by Cambridge University Press

occurs with electric networks, where y, the vector of nodal potentials, is also defined up to an additive constant. The rank deficiency in B can easily be removed by 'grounding' one of the nodes, that is, by specifying the value of the potential (or of the pressure) at one point. One problem with this approach is that the resulting linear system may be rather ill-conditioned; see Bochev and Lehoucq (2005). Fortunately, since the system $\mathcal{A}u = b$ is consistent by construction, it may not be necessary to remove the singularity of $\mathcal{A}$. Iterative methods like GMRES (Saad and Schultz 1986) are largely unaffected by the presence of a single eigenvalue exactly equal to zero, at least when using a zero initial guess, $u_0 = 0$. The reader is referred to Elman *et al.* (2005*c*, Section 8.3.4) for a detailed discussion of this issue in the context of fluid flow problems; see further the remarks in Olshanskii and Reusken (2004, Section 4) and Zhang and Wei (2004, Section 4).

General case

When $C = O$, a *necessary* condition for invertibility is provided by the following theorem, a slight generalization of a similar result for the case $B_1 = B_2$: see Gansterer, Schneid and Ueberhuber (2003).

Theorem 3.3. If the matrix

$$\mathcal{A} = \begin{bmatrix} A & B_1^T \\ B_2 & O \end{bmatrix}$$

is nonsingular, then $\operatorname{rank}(B_1) = m$ and $\operatorname{rank}\binom{A}{B_2} = n$.

Proof. If $\operatorname{rank}(B_1) < m$ then there exists a nonzero vector $y \in \mathbb{R}^m$ with $B_1^T y = 0$. Therefore, letting $u = \begin{bmatrix} 0 \\ y \end{bmatrix}$, we get $\mathcal{A}u = 0$, a contradiction.

If $\operatorname{rank}\binom{A}{B_2} < n$ then there exists a nonzero vector $x \in \mathbb{R}^n$ such that $\binom{A}{B_2}x = 0$. Letting $u = \begin{bmatrix} x \\ 0 \end{bmatrix}$, we get $\mathcal{A}u = 0$, a contradiction. □

It is easy to show that these conditions are not sufficient to ensure the invertibility of $\mathcal{A}$. Some additional conditions are needed. Recall that for any matrix $A \in \mathbb{R}^{n\times n}$ we can write $A = H + K$ where $H = \frac{1}{2}(A + A^T)$ and $K = \frac{1}{2}(A - A^T)$ are the symmetric and skew-symmetric part of A, respectively. The following result provides a necessary and a sufficient condition for $\mathcal{A}$ to be invertible when $B_1 = B_2$.

Theorem 3.4. Assume that H, the symmetric part of A, is positive semidefinite, $B_1 = B_2 = B$ have full rank, and C is symmetric positive semidefinite (possibly zero). Then

(i) $\ker(H) \cap \ker(B) = \{0\} \Rightarrow \mathcal{A}$ invertible,
(ii) $\mathcal{A}$ invertible $\Rightarrow \ker(A) \cap \ker(B) = \{0\}$.

The converses of (i)–(ii) do not hold in general.

https://doi.org/10.1017/S0962492904000212 Published online by Cambridge University Press

Proof. The proof of part (i) is similar to the sufficiency proof in Theorem 3.1, and can be found in Benzi and Golub (2004, Lemma 1.1). The proof of part (ii) is exactly as the necessity proof in Theorem 3.1.

Now we show that the converses of (i) and (ii) are false in general. To see that the converse of (i) is not true in general, consider the matrix

$$\mathcal{A} = \left[\begin{array}{ccc|c} 1 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ \hline 0 & 0 & 1 & 0 \end{array}\right] = \begin{bmatrix} A & B^T \\ B & O \end{bmatrix}.$$

Here we have $\ker(H) \cap \ker(B) = \mathrm{span}([0\,1\,0]^T) \neq \{0\}$ and yet $\mathcal{A}$ is invertible.

To see that the converse of (ii) is not generally true, consider the matrix

$$\mathcal{A} = \left[\begin{array}{cc|c} 0 & -1 & 0 \\ 1 & 1 & 1 \\ \hline 0 & 1 & 0 \end{array}\right] = \begin{bmatrix} A & B^T \\ B & O \end{bmatrix}.$$

This matrix is manifestly singular, yet A is nonsingular and thus $\ker(A) \cap \ker(B) = \{0\}$. □

We note that when H is positive semidefinite, the inclusion $\ker(A) \subset \ker(H)$ holds. For if $x \in \ker(A)$, then $Ax = Hx + Kx = 0$ and thus $x^T Ax = x^T Hx = 0$. But since H is symmetric positive semidefinite the last equality implies $Hx = 0$, and therefore $x \in \ker(H)$. That this is a proper inclusion can be seen from the simple example

$$A = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}.$$

3.3. The inverse of a saddle point matrix

If A is nonsingular, then we know that $\mathcal{A}$ is invertible if and only if $S = -(C + B_2 A^{-1} B_1^T)$ is nonsingular, and we have the following explicit expression for the inverse:

$$\mathcal{A}^{-1} = \begin{bmatrix} A & B_1^T \\ B_2 & -C \end{bmatrix}^{-1} = \begin{bmatrix} A^{-1} + A^{-1} B_1^T S^{-1} B_2 A^{-1} & -A^{-1} B_1^T S^{-1} \\ -S^{-1} B_2 A^{-1} & S^{-1} \end{bmatrix}. \quad (3.4)$$

If A is singular but C is nonsingular, an analogous expression can be given if we assume that the matrix $A + B_1^T C^{-1} B_2$, the Schur complement of C in $\mathcal{A}$, is nonsingular. However, such an expression is of limited interest in the numerical solution of saddle point problems. See Lu and Shiou (2002) for additional expressions for the inverse of a block 2×2 matrix.

An interesting special case arises when A is symmetric positive definite, $B_1 = B_2 = B$, $C = O$, $S = -BA^{-1}B^T$ is nonsingular, and $g = 0$. Then

https://doi.org/10.1017/S0962492904000212 Published online by Cambridge University Press

the explicit expression for $\mathcal{A}^{-1}$ shows that the solution (x_*, y_*) of (1.3) is given by

$$\begin{bmatrix} x_* \\ y_* \end{bmatrix} = \begin{bmatrix} (I + A^{-1}B^T S^{-1} B)\, A^{-1} f \\ S^{-1} B A^{-1} f \end{bmatrix}. \tag{3.5}$$

It is easy to see that the matrix

$$\Pi \equiv -A^{-1}B^T S^{-1} B = A^{-1}B^T (BA^{-1}B^T)^{-1} B$$

satisfies $\Pi^2 = \Pi$, *i.e.*, Π is a projector. Moreover, the relations

$$\Pi v \in \text{range}(A^{-1}B^T) \quad \text{and} \quad v - \Pi v \perp \text{range}(B^T), \quad \text{for all} \quad v \in \mathbb{R}^n,$$

show that Π represents an (oblique) projector onto range$(A^{-1}B^T)$ and orthogonal to range(B^T). The first component in (3.5) can then be written as

$$x_* = (I - \Pi)\, \hat{x},$$

where $\hat{x} \equiv A^{-1} f$ is the solution of the unconstrained problem (1.4). Hence x_* is orthogonal to range(B^T). Furthermore, $\hat{x} = \Pi \hat{x} + x_*$, which means that the solution of the unconstrained problem is decomposed into a part that is in range$(A^{-1}B^T)$ and a part that is orthogonal to range(B^T). By the nature of Π, this decomposition generally is oblique, and is orthogonal only when range$(A^{-1}B^T)$ = range(B^T) (the latter being true particularly for $A = I$). Next note that since $f - B^T y_* = A x_*$, and $B x_* = 0$,

$$0 = B x_* = (BA^{-1})(A x_*) = (A^{-1}B^T)^T\, (f - B^T y_*). \tag{3.6}$$

By assumption, A^{-1} is symmetric positive definite, so the function $\langle v, w \rangle_{A^{-1}}$ $\equiv w^T A^{-1} v$ is an inner product. Then (3.6) shows that the vector $f - B^T y_* \in$ $f +$ range(B^T) is orthogonal with respect to the A^{-1}-inner product (A^{-1}-orthogonal) to the space range(B^T). But this means that y_* is the solution of the (generalized) least squares problem $B^T u \approx f$ with respect to the A^{-1}-norm, $\|v\|_{A^{-1}} \equiv (\langle v, v \rangle_{A^{-1}})^{1/2}$, *i.e.*,

$$\|f - B^T y_*\|_{A^{-1}} = \min_u \|f - B^T u\|_{A^{-1}}. \tag{3.7}$$

The relation (3.7) is also derived in Benbow (1999, Section 2), where it is used to compute the solution of the standard saddle point problem (1.3) with $g = 0$ by first solving the generalized least squares problem with the matrix B^T and the right-hand side f (giving y_*), and then computing $x_* = A^{-1}(f - B^T y_*)$. To solve the generalized least squares problem (3.7) numerically, a generalized version of the LSQR method of Paige and Saunders (1982) is developed in Benbow (1999). The numerical experiments in Benbow (1999) show that this approach to computing y_* is often superior to applying the MINRES method (Paige and Saunders 1975) (see Section 9 below for details about this method) to the (symmetric) Schur complement system $S\, y = B A^{-1} f$, which is the second component in (3.5).

https://doi.org/10.1017/S0962492904000212 Published online by Cambridge University Press

Another important special case arises when A and C are both symmetric positive definite and $B_1 = B_2 = B$. Then the corresponding (symmetric) saddle point matrix $\mathcal{A}$ is said to be *quasidefinite*, regardless of what the rank of B may be (Vanderbei 1995). These matrices result from interior point methods in constrained optimization; see the discussion in Section 2.3. Properties of symmetric quasidefinite matrices have been studied, *e.g.*, in George, Ikramov and Kucherov (2000), Gill, Saunders and Shinnerl (1996) and Vanderbei (1995). One of the basic properties is that if $\mathcal{A}$ is quasidefinite, so is $\mathcal{A}^{-1}$ (under the same 2×2 block partitioning), as one can immediately see from (3.4). In George and Ikramov (2002), this property has been extended to the class of invertible *weakly quasidefinite* matrices, in which A and C are only assumed to be symmetric positive semidefinite.

An alternative expression of the inverse that does not require A to be invertible is the following. Assume that $B_1 = B_2 = B$ has full rank and $C = O$. Denote by $Z \in \mathbb{R}^{n\times(n-m)}$ any matrix whose columns form a basis for $\ker(B)$. If H, the symmetric part of A, is positive semidefinite, then it is easy to see that condition (i) in Theorem 3.4 implies that the $(n-m) \times (n-m)$ matrix $Z^T AZ$ is invertible; indeed, its symmetric part $Z^T HZ$ is positive definite. Letting $W = Z(Z^T AZ)^{-1}Z^T$, we have the following expression for the inverse of $\mathcal{A}$:

$$\begin{aligned} \mathcal{A}^{-1} &= \begin{bmatrix} A & B^T \\ B & O \end{bmatrix}^{-1} \\ &= \begin{bmatrix} W & (I - WA)B^T(BB^T)^{-1} \\ (BB^T)^{-1}B(I - AW) & -(BB^T)^{-1}B(A - AWA)B^T(BB^T)^{-1} \end{bmatrix}, \end{aligned} \tag{3.8}$$

which can be easily proved keeping in mind that $B^T(BB^T)^{-1}B = I - ZZ^T$; see Gansterer *et al.* (2003).

These explicit expressions for $\mathcal{A}^{-1}$ are of limited practical use, and their interest is primarily theoretical. See, however, Powell (2004) for a situation where the inverse of $\mathcal{A}$ is explicitly needed, and for a careful discussion of the problem of updating the inverse of $\mathcal{A}$ when a few of its rows and corresponding columns are modified.

It must be mentioned that in the finite element context the mere nonsingularity of $\mathcal{A}$ is not sufficient to ensure meaningful computed solutions. In order for the discrete problem to be well-posed it is essential that the saddle point matrix remain *uniformly invertible* as h, the mesh size parameter, goes to zero. This means that an appropriate (generalized) condition number of $\mathcal{A}$ remains bounded as $h \to 0$. Sufficient conditions for this to happen include the already-mentioned discrete LBB (or inf-sup) conditions; see Brezzi (2002), Brezzi and Fortin (1991), Ciarlet Jr. *et al.* (2003) and Nicolaides (1982). Some discussion of conditioning issues from a linear algebra standpoint can be found in Section 3.5.

https://doi.org/10.1017/S0962492904000212 Published online by Cambridge University Press

Finally, we point out that *singular* saddle point systems have been studied in Gansterer *et al.* (2003) and Wu, Silva and Yuan (2004); other relevant references are Wei (1992*a*) and Zhang and Wei (2004). Explicit forms for various generalized inverses of saddle point matrices arising in constrained optimization can be found in Campbell and Meyer Jr. (1979, Section 3.5); see also George and Ikramov (2002).

3.4. Spectral properties of saddle point matrices

In this section we collect a few facts on the spectral properties of saddle point matrices which are relevant when solving the equations by iterative methods. We also introduce an alternative formulation of the saddle point equations leading to a (nonsymmetric) positive definite coefficient matrix.

Eigenvalues: The symmetric case

Assume that A is symmetric positive definite, $B_1 = B_2 = B$ has full rank, and C is symmetric positive semidefinite (possibly zero). Then from (3.1) we obtain

$$\begin{bmatrix} I & O \\ -BA^{-1} & I \end{bmatrix} \begin{bmatrix} A & B^T \\ B & -C \end{bmatrix} \begin{bmatrix} I & -A^{-1}B^T \\ O & I \end{bmatrix} = \begin{bmatrix} A & O \\ O & S \end{bmatrix} \tag{3.9}$$

where $S = -(C + BA^{-1}B^T)$ is symmetric negative definite. Hence $\mathcal{A}$ is congruent to the block diagonal matrix $\left[\begin{smallmatrix} A & O \\ O & S \end{smallmatrix}\right]$. It follows from Sylvester's Law of Inertia (see Horn and Johnson (1985, page 224)) that $\mathcal{A}$ is indefinite, with n positive and m negative eigenvalues. The same is of course true if B is rank deficient, as long as S remains negative definite. Clearly, in case S is rank deficient, say $\operatorname{rank}(S) = m - r$, $\mathcal{A}$ has n positive, $m - r$ negative and r zero eigenvalues. A simple limiting argument shows that this result remains true if A is only assumed to be positive semidefinite, provided that the usual condition $\ker(A) \cap \ker(B) = \{0\}$ is satisfied. We refer the reader to Forsgren (2002), Forsgren and Murray (1993), Gill, Murray, Saunders and Wright (1991) and Gould (1985) for additional results on the inertia of symmetric saddle point matrices under various assumptions on A, B and C. Generally speaking, unless m is very small (which is seldom the case in practice), the matrix $\mathcal{A}$ is *highly indefinite*, in the sense that it has many eigenvalues of both signs.

The following result from Rusten and Winther (1992) establishes eigenvalue bounds for an important class of saddle point matrices.

Theorem 3.5. Assume A is symmetric positive definite, $B_1 = B_2 = B$ has full rank, and $C = O$. Let μ_1 and μ_n denote the largest and smallest eigenvalues of A, and let σ_1 and σ_m denote the largest and smallest singular

https://doi.org/10.1017/S0962492904000212 Published online by Cambridge University Press

values of B. Let $\sigma(\mathcal{A})$ denote the spectrum of $\mathcal{A}$. Then

$$\sigma(\mathcal{A}) \subset I^- \cup I^+,$$

where

$$I^- = \left[\frac{1}{2}\Big(\mu_n - \sqrt{\mu_n^2 + 4\sigma_1^2}\Big),\ \frac{1}{2}\Big(\mu_1 - \sqrt{\mu_1^2 + 4\sigma_m^2}\Big)\right]$$

and

$$I^+ = \left[\mu_n,\ \frac{1}{2}\Big(\mu_1 + \sqrt{\mu_1^2 + 4\sigma_1^2}\Big)\right].$$

An extension of this result to the case where $C \neq O$ (with C symmetric and positive semidefinite) can be found in Silvester and Wathen (1994), while the case where A is positive semidefinite with $\ker(A) \cap \ker(B) = \{0\}$ has been treated in Perugia and Simoncini (2000). These bounds can be used to obtain estimates for the condition number of $\mathcal{A}$ in specific cases. In turn, these estimates can be used to predict the rate of convergence of iterative methods like MINRES (Paige and Saunders 1975); see Fischer, Ramage, Silvester and Wathen (1998), Maryška, Rozložník and Tůma (1996), Wathen, Fischer and Silvester (1995) and Section 9 below. Eigenvalue bounds are also important when assessing the (inf-sup) stability of mixed finite element discretizations; see, *e.g.*, Bitar and Vincent (2000) and Malkus (1981).

Eigenvalues: The general case

In the general case, not much can be said about the eigenvalues of $\mathcal{A}$. However, in most cases of interest the convex hull of the eigenvalues of $\mathcal{A}$ contains the origin. If we consider for example the case where $A \neq A^T$, $B_1 = B_2 = B$, and $C = C^T$ (as in the Oseen problem, for example) then we have that the symmetric part of $\mathcal{A}$ is

$$\frac{1}{2}(\mathcal{A} + \mathcal{A}^T) = \begin{bmatrix} H & B^T \\ B & -C \end{bmatrix},$$

where H is the symmetric part of A. If H is positive definite and C positive semidefinite (as in the Oseen problem), then the symmetric part of $\mathcal{A}$ is indefinite and therefore $\mathcal{A}$ has eigenvalues on both sides of the imaginary axis. Figure 3.1(a) displays the eigenvalues of the discrete Oseen operator obtained from a Q1-P0 finite element approximation of problem (2.7)–(2.9) with $\nu = 0.01$ and $\Omega = [0,1] \times [0,1]$. The matrix was generated using the IFISS software package (Elman, Ramage, Silvester and Wathen 2005*b*).

Algorithms for solving both standard and generalized eigenvalue problems for saddle point matrices have been studied in the literature, particularly for investigating the stability of incompressible flows and in electromagnetism. See Arbenz and Geus (2005), Arbenz, Geus and Adam (2001), Cliffe, Garratt and Spence (1994), Graham, Spence and Vainikko (2003), Lehoucq and Salinger (2001) and Meerbergen and Spence (1997).

https://doi.org/10.1017/S0962492904000212 Published online by Cambridge University Press

An alternative formulation

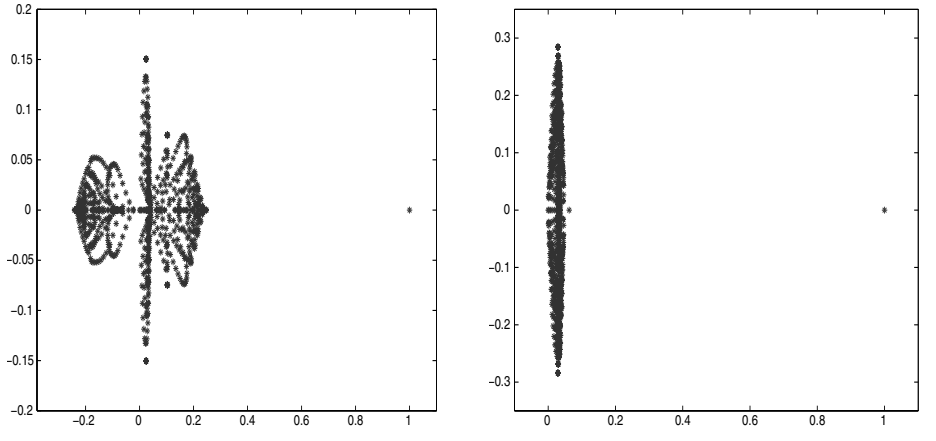
Eigenvalue distributions such as that shown in Figure 3.1(a) are generally considered unfavourable for solution by Krylov subspace methods and indeed, without preconditioning, Krylov subspace methods tend to converge poorly when applied to the corresponding linear system. It has been observed by several authors (*e.g.*, Benzi and Golub (2004), Fischer *et al.* (1998), Polyak (1970) and Sidi (2003), in addition to Glowinski (1984, page 20) and Quarteroni and Valli (1994, page 304)) that a simple transformation can be used to obtain an equivalent linear system with a coefficient matrix whose spectrum is entirely contained in the half-plane $\mathrm{Re}(z) > 0$. (Here we use $\mathrm{Re}(z)$ and $\mathrm{Im}(z)$ to denote the real and imaginary part of $z \in \mathbb{C}$.) Indeed, assuming that $B_1 = B_2 = B$, we can rewrite the saddle point system in the equivalent form

$$\begin{bmatrix} A & B^T \\ -B & C \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} f \\ -g \end{bmatrix}, \quad \text{or} \quad \hat{\mathcal{A}}u = \hat{b}. \tag{3.10}$$

Note that $\hat{\mathcal{A}} = \mathcal{J}\mathcal{A}$ where

$$\mathcal{J} = \begin{bmatrix} I_n & O \\ O & -I_m \end{bmatrix} \tag{3.11}$$

and therefore $\hat{\mathcal{A}}$ is nonsingular if $\mathcal{A}$ is. Moreover, we have the following result.



(a) indefinite form of $\mathcal{A}$ (b) positive definite form of $\mathcal{A}$

Figure 3.1. Eigenvalues for two-dimensional Oseen problem (leaky lid-driven cavity) using Q1-P0 discretization.

https://doi.org/10.1017/S0962492904000212 Published online by Cambridge University Press

Theorem 3.6. Let $\hat{\mathcal{A}}$ be the coefficient matrix in (3.10). Assume $H = \frac{1}{2}(A + A^T)$ is positive semidefinite, $B_1 = B_2 = B$ has full rank, $C = C^T$ is positive semidefinite, and $\ker(H) \cap \ker(B) = \{0\}$. Let $\sigma(\hat{\mathcal{A}})$ denote the spectrum of $\hat{\mathcal{A}}$. Then

(i) $\hat{\mathcal{A}}$ is positive semidefinite, in the sense that $v^T \hat{\mathcal{A}} v \geq 0$ for all $v \in \mathbb{R}^{n+m}$,

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

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

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

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

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

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

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

a real nonnegative quantity.

To prove (iii), assume (λ, v) is an eigenpair of $\hat{\mathcal{A}}$ with $v = \begin{bmatrix} x \\ y \end{bmatrix}$. Then

$$\begin{aligned} \mathrm{Re}(\lambda) &= x^* H x + y^* C y \\ &= \mathrm{Re}(x)^T H \mathrm{Re}(x) + \mathrm{Im}(x)^T H \mathrm{Im}(x) \\ &\quad + \mathrm{Re}(y)^T C \mathrm{Re}(y) + \mathrm{Im}(y)^T C \mathrm{Im}(y). \end{aligned}$$

This quantity is nonnegative, and it can be zero only if $x = 0$ (since H is assumed to be positive definite) and $Cy = 0$. But if $x = 0$ then from the first equation in the system $\hat{\mathcal{A}} v = \lambda v$ we get $B^T y = 0$, hence $y = 0$ since B^T has full column rank. Hence $v = 0$, a contradiction. □

Figure 3.1(b) displays the eigenvalues of the matrix $\hat{\mathcal{A}}$ corresponding to the same Oseen problem as before. As can be seen, all the eigenvalues lie in the right half-plane. However, such distribution of eigenvalues is not necessarily more favourable for Krylov subspace methods, and in fact Krylov subspace methods without preconditioning perform just as poorly as they do on the original problem. We revisit this topic repeatedly in the course of this survey.

https://doi.org/10.1017/S0962492904000212 Published online by Cambridge University Press

When $A = A^T$ and $C = C^T$ the matrix $\mathcal{A}$ is symmetric indefinite whereas $\hat{\mathcal{A}}$ is nonsymmetric positive (semi-)definite. Moreover, $\hat{\mathcal{A}}$ satisfies

$$\mathcal{J}\,\hat{\mathcal{A}} = \hat{\mathcal{A}}^T\,\mathcal{J}, \tag{3.12}$$

with $\mathcal{J}$ defined as in (3.11); that is, $\hat{\mathcal{A}}$ is $\mathcal{J}$*-symmetric* (or *pseudosymmetric*: see Mackey, Mackey and Tisseur (2003)). In other words, $\hat{\mathcal{A}}$ is symmetric with respect to the indefinite inner product defined on $\mathbb{R}^{n+m}$ by $[v, w] \equiv w^T \mathcal{J} v$. Conversely, any $\mathcal{J}$-symmetric matrix is of the form $\left[\begin{smallmatrix} A & B^T \\ -B & C \end{smallmatrix}\right]$ for some $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{m\times n}$, and $C \in \mathbb{R}^{m\times m}$ with $A = A^T$ and $C = C^T$. Note that the set of all $\mathcal{J}$-symmetric matrices

$$\mathbb{J} = \left\{ \begin{bmatrix} A & B^T \\ -B & C \end{bmatrix} \;\middle|\; A = A^T \in \mathbb{R}^{n\times n}, B \in \mathbb{R}^{m\times n}, C = C^T \in \mathbb{R}^{m\times m} \right\}$$

is closed under matrix addition and under the so-called *Jordan product*, defined as

$$\mathcal{F} \cdot \mathcal{G} \equiv \frac{1}{2}(\mathcal{F}\,\mathcal{G} + \mathcal{G}\,\mathcal{F}).$$

The triple $(\mathbb{J}, +, \cdot)$ is a non-associative, commutative algebra over the reals. It is known as the *Jordan algebra* associated with the real Lie group $O(n, m, \mathbb{R})$ of $\mathcal{J}$-orthogonal (or *pseudo-orthogonal*) matrices, *i.e.*, the group of all matrices $\mathcal{Q} \in \mathbb{R}^{n+m}$ that satisfy the condition $\mathcal{Q}^T\,\mathcal{J}\,\mathcal{Q} = \mathcal{J}$; see Ammar, Mehl and Mehrmann (1999) and Mackey *et al.* (2003). The spectral theory of these matrices has been investigated by several authors. A Schur-like decomposition for matrices in $\mathbb{J}$ has been given in Ammar *et al.* (1999, Theorem 8), and properties of invariant subspaces of $\mathcal{J}$-symmetric matrices have been studied in Gohberg, Lancaster and Rodman (1983).

Besides being mathematically appealing, these algebraic properties have implications from the point of view of iterative methods: see for example Freund, Golub and Nachtigal (1992, page 80), where it is shown how $\mathcal{J}$-symmetry can be exploited to develop transpose-free variants of basic Krylov methods using short recurrences. This is of course not enough to justify using the nonsymmetric form $\hat{\mathcal{A}}$ when $\mathcal{A}$ is symmetric, since in this case one may as well use a symmetric Krylov solver on the original (symmetric) formulation; see Fischer and Peherstorfer (2001), Fischer *et al.* (1998) and Section 9 below. Nevertheless, there are some advantages in using the transformed linear system $\hat{\mathcal{A}}v = \hat{b}$ instead of the original one, especially when certain preconditioners are used; see Benzi and Golub (2004), Sidi (2003) and Section 10.3 below. It can be shown that when A and C are symmetric, at most $2m$ of the $n + m$ eigenvalues of $\hat{\mathcal{A}}$ can have a nonzero imaginary part (Simoncini 2004*b*); furthermore, in some important special cases it turns out that the eigenvalues of $\hat{\mathcal{A}}$ are all real and positive. This implies the existence of a nonstandard inner product on $\mathbb{R}^{n+m}$ with respect

https://doi.org/10.1017/S0962492904000212 Published online by Cambridge University Press

to which $\hat{\mathcal{A}}$ is symmetric positive definite, a desirable property from the point of view of iterative methods. The following result gives a sufficient condition (easily checked in many cases) for the eigenvalues of $\hat{\mathcal{A}}$ to be real.

Theorem 3.7. Assume that A is symmetric positive definite, $B_1 = B_2 = B$ has full rank, and $C = O$. Let $S = BA^{-1}B^T$, and let μ_n denote the smallest eigenvalue of A. If $\mu_n \geq 4\|S\|_2$, then all the eigenvalues of the matrix $\hat{\mathcal{A}}$ in (3.10) are real and positive.

Proof. See Simoncini and Benzi (2005). □

We note that the conditions expressed in Theorem 3.7 are satisfied, for instance, for the stationary Stokes problem under a variety of finite differences and finite element discretization schemes.

A more detailed analysis is available in the special situation that $A = \eta I$, where $\eta > 0$ is a positive scaling parameter, $B_1 = B_2 = B$, and $C = O$. Denote the resulting usual saddle point matrix by $\mathcal{A}_\eta^+$, and the alternative formulation with negative (2,1) block by $\mathcal{A}_\eta^-$, *i.e.*,

$$\mathcal{A}_\eta^\pm = \begin{bmatrix} \eta I & B^T \\ \pm B & O \end{bmatrix}. \tag{3.13}$$

The following theorem characterizes the influence of the choice of $+$ or $-$ as well as η on the eigenvalues of the matrices $\mathcal{A}_\eta^\pm$.

Theorem 3.8. Suppose that the matrix B has rank $m - r$, and denote the nonzero singular values of B by $\sigma_1 \geq \cdots \geq \sigma_{m-r}$.

1 The $n + m$ eigenvalues of $\mathcal{A}_\eta^+$ in (3.13) are given by

(i) zero with multiplicity r,
(ii) η with multiplicity $n - m + r$,
(iii) $\frac{1}{2}\left(\eta \pm \sqrt{4\sigma_k^2 + \eta^2}\right)$ for $k = 1, \ldots, m - r$.

2 Furthermore, if $\sigma_1 \geq \cdots \geq \sigma_t > \frac{\eta}{2} \geq \sigma_{t+1} \geq \cdots \geq \sigma_{m-r}$, then the $n + m$ eigenvalues of $\mathcal{A}_\eta^-$ in (3.13) are given by

(i) zero with multiplicity r,
(ii) η with multiplicity $n - m + r$,
(iii) $\frac{1}{2}\left(\eta \pm \sqrt{\eta^2 - 4\sigma_k^2}\right)$ for $k = t + 1, \ldots, m - r$,
(iv) $\frac{1}{2}\left(\eta \pm i\sqrt{4\sigma_k^2 - \eta^2}\right)$ for $k = 1, \ldots, t$.

Proof. See Fischer *et al.* (1998, Section 2). □

This result shows that the eigenvalues of the symmetric indefinite matrix $\mathcal{A}_\eta^+$ (except for the multiple eigenvalues zero and η) always lie in

https://doi.org/10.1017/S0962492904000212 Published online by Cambridge University Press

two intervals symmetric about the point $\eta/2$. Changing η only leads to a scaling of these two intervals. On the other hand, the choice of η has a significant effect on the eigenvalues of the nonsymmetric matrix $\mathcal{A}_\eta^-$. For example, if $\eta > 2\sigma_1$, then all eigenvalues of $\mathcal{A}_\eta^-$ are real, while for $\eta < 2\sigma_{m-r}$ all eigenvalues (except zero and η) are purely imaginary. For intermediate values of η, the eigenvalues (except zero and η) form a cross in the complex plane with midpoint $\eta/2$. One is immediately tempted to determine what eigenvalue distribution is the most favourable for the solution by Krylov subspace methods; see Sidi (2003). We discuss this topic in Section 10.1, where the matrices $\mathcal{A}_\eta^\pm$ arise naturally as a result of block diagonal preconditioning of a symmetric saddle point matrix $\mathcal{A}$ with A positive definite and $C = O$.

3.5. Conditioning issues

Saddle point systems that arise in practice can be very poorly conditioned, and care must be taken when developing and applying solution algorithms. It turns out that in some cases the special structure of the saddle point matrix $\mathcal{A}$ can be exploited to avoid or mitigate the effect of ill-conditioning. Moreover, the structure of the right-hand side b in (1.1) also plays a role. Indeed, it is frequently the case that either f or g in (1.1) is zero. For instance, $f = 0$ in structural analysis (in the absence of dilation) and in mixed formulations of Poisson's equation, while $g = 0$ in incompressible flow problems and weighted least-squares. So if g (say) is zero, the $(1,2)$ and $(2,2)$ blocks in $\mathcal{A}^{-1}$ (see (3.4) and (3.8)) have no influence on the solution $u = \mathcal{A}^{-1}b$. In particular, any ill-conditioning that may be present in these blocks will not affect the solution, an important fact that should be taken into account in the development of robust solution algorithms; see Duff (1994), Gansterer *et al.* (2003) and Vavasis (1994).

Let us consider, for the sake of simplicity, a standard saddle point problem where $A = A^T$ is positive definite, $B_1 = B_2 = B$ has full rank, and $C = O$. In this case $\mathcal{A}$ is symmetric and its spectral condition number is given by

$$\kappa(\mathcal{A}) = \frac{\max |\lambda(\mathcal{A})|}{\min |\lambda(\mathcal{A})|}.$$

From Theorem 3.5 one can see that the condition number of $\mathcal{A}$ grows unboundedly as either $\mu_n = \lambda_{\min}(A)$ or $\sigma_m = \sigma_{\min}(B)$ goes to zero (assuming that $\lambda_{\max}(A)$ and $\sigma_{\max}(B)$ are kept constant). For mixed finite element formulations of elliptic PDEs, both μ_n and σ_m go to zero as h, the mesh size parameter, goes to zero, and the condition number of $\mathcal{A}$ grows like $O(h^{-p})$ for some positive value of p; see, *e.g.*, Maryška *et al.* (1996), Wathen *et al.* (1995). This growth of the condition number of $\mathcal{A}$ means that the rate of convergence of most iterative solvers (like Krylov subspace methods) deteriorates as the problem size increases. As discussed in Section 10,

https://doi.org/10.1017/S0962492904000212 Published online by Cambridge University Press

preconditioning may be used to reduce or even eliminate this dependency on h in many cases. Similar considerations apply to nonsymmetric saddle point problems.

A different type of ill-conditioning is encountered in saddle point systems from interior point methods. Consider, for instance, the case of linear programming, where the $(1,1)$ block A is diagonal. As the iterates generated by the interior point algorithm approach the solution, many of the entries of A tend to zero or infinity, and thus $\mathcal{A}$ becomes very ill-conditioned (the constraint matrix B remains constant throughout the iteration process). In particular, the norm of the inverse Schur complement $S^{-1} = -(BA^{-1}B^T)^{-1}$ goes to infinity. However, Stewart (1989) and Todd (1990) (see also Forsgren (1996)) have shown that the norm of the matrices $X = S^{-1}BA^{-1}$ (a *weighted pseudo-inverse* of B^T) and B^TX (the associated oblique projector onto the column space of B^T) are bounded by numbers that are independent of A. This important observation has been exploited in a series of papers by Vavasis and collaborators (Bobrovnikova and Vavasis 2001, Hough and Vavasis 1997, Vavasis 1994, Vavasis 1996) to develop stable algorithms for certain saddle point problems with a severely ill-conditioned $(1,1)$ block A.

When using direct methods based on triangular factorization, Björck (1996, Sections 2.5.3 and 4.4.2) has noted the importance of scaling the $(1,1)$ block A by a positive scalar quantity. Suitable tuning of this scaling factor can be interpreted as a form of preconditioning and has a dramatic impact on the accuracy attainable by sparse direct solvers (Arioli, Duff and De Rijk 1989, Duff 1994). On the other hand, such scaling seems to have little or no effect on the convergence behaviour of Krylov subspace methods (Fischer *et al.* 1998).

Another possible approach for dealing with an ill-conditioned or even singular $(1,1)$ block A is the augmented Lagrangian method; see Fortin and Glowinski (1983), Glowinski and Le Tallec (1989), Hestenes (1969), Powell (1969) and the more general treatment in Golub and Greif (2003) and Greif *et al.* (2005). Here we assume that $A = A^T$ (possibly singular), $B_1 = B_2 = B$ has full rank, and $C = O$. The idea is to replace the saddle point system (1.3) with the equivalent one

$$\begin{bmatrix} A + B^TWB & B^T \\ B & O \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} f + B^TWg \\ g \end{bmatrix}. \tag{3.14}$$

The $m \times m$ matrix W, to be suitably determined, is symmetric positive semidefinite. The simplest choice is to take $W = \gamma I$ ($\gamma > 0$). In this case the $(1,1)$ block in (3.14) is nonsingular, and indeed positive definite, provided that A is positive definite on $\ker(B)$. The goal is to choose W so that system (3.14) is easier to solve than the original one, particularly when using iterative methods. When $W = \gamma I$ is used, the choice $\gamma = \|A\|_2/\|B\|_2^2$ has been found to perform well in practice, in the sense that the condition

https://doi.org/10.1017/S0962492904000212 Published online by Cambridge University Press

number of both the $(1,1)$ block and of the whole coefficient matrix in (3.14) are approximately minimized. This choice also results in rapid convergence of classical iterative schemes like the method of multipliers; see Golub and Greif (2003) and Section 8.2 below.

The conditioning of equality-constrained and weighted least squares problems has been studied in depth by several authors; see Gulliksson, Jin and Wei (2002), Wei (1992*b*) and the references therein.

Conditioning properties of quasidefinite and saddle-point matrices arising from interior-point methods in linear programming have also been investigated in George and Ikramov (2000) and Korzak (1999). Finally, we mention that a numerical validation method for verifying the accuracy of approximate solutions of symmetric saddle point problems has been presented in Chen and Hashimoto (2003).

4. Overview of solution algorithms

Besides the usual (and somewhat simplistic) distinction between direct and iterative methods, solution algorithms for generalized saddle point problems can be subdivided into two broad categories, which we will call *segregated* and *coupled* (or 'all at once') methods. Segregated methods compute the two unknown vectors, x and y, separately; in some cases it is x to be computed first, in others it is y. This approach involves the solution of two linear systems of size smaller than $n+m$ (called *reduced systems*), one for each of x and y; in some cases a reduced system for an intermediate quantity is solved. Segregated methods can be either direct or iterative, or involve a combination of the two; for example, one of the reduced systems could be solved by a direct method and the other iteratively. The main representatives of the segregated approach are the Schur complement reduction method, which is based on a block LU factorization of $\mathcal{A}$, and the null space method, which relies on a basis for the null space for the constraints.

Coupled methods, on the other hand, deal with the system (1.1) as a whole, computing x and y (or approximations to them) simultaneously and without making explicit use of reduced systems. These methods include both direct solvers based on triangular factorizations of the global matrix $\mathcal{A}$, and iterative algorithms like Krylov subspace methods applied to the entire system (1.1), typically with some form of preconditioning. As we shall see, preconditioning tends to blur the distinction between direct and iterative solvers, and also that between segregated and coupled schemes. This is because direct solvers may be used to construct preconditioners, and also because preconditioners for coupled iterative schemes are frequently based on segregated methods.

In the next sections we review a number of solution methods, starting with direct solvers and continuing with stationary iterative methods, Krylov sub-

https://doi.org/10.1017/S0962492904000212 Published online by Cambridge University Press

space solvers, and preconditioners. We also include a brief discussion of multilevel methods, including multigrid and Schwarz-type algorithms. Within each group, we discuss segregated as well as coupled schemes and the interplay between them. It is simply not possible to cover every method that has been described in the literature; instead, we have striven to include, besides all of the 'classical' algorithms, those among the more recent methods that appear to be the most widely applicable and effective.

5. Schur complement reduction

Consider the saddle point system (1.1), or

$$Ax + B_1^T y = f, \quad B_2 x - Cy = g.$$

We assume that both A and $\mathcal{A}$ are nonsingular; by (3.1) this implies that $S = -(C + B_2 A^{-1} B_1^T)$ is also nonsingular. Pre-multiplying both sides of the first equation by $B_2 A^{-1}$, we obtain

$$B_2 x + B_2 A^{-1} B_1^T y = B_2 A^{-1} f.$$

Using $B_2 x = g + Cy$ and rearranging, we find

$$(B_2 A^{-1} B_1^T + C)\, y = B_2 A^{-1} f - g, \tag{5.1}$$

a reduced system of order m for y involving the (negative) Schur complement $-S = B_2 A^{-1} B_1^T + C$. Note that unless $f = 0$, forming the right-hand side of (5.1) requires solving a linear system of the form $Av = f$.

Once y_* has been computed from (5.1), x_* can be obtained by solving

$$Ax = f - B_1^T y_*, \tag{5.2}$$

a reduced system of order n for x involving the (1,1) block, A. Note that this is just block Gaussian elimination applied to (1.1). Indeed, using the block LU factorization (3.3) we get the transformed system

$$\begin{bmatrix} I & O \\ -B_2 A^{-1} & I \end{bmatrix} \begin{bmatrix} A & B_1^T \\ B_2 & -C \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} I & O \\ -B_2 A^{-1} & I \end{bmatrix} \begin{bmatrix} f \\ g \end{bmatrix},$$

that is,

$$\begin{bmatrix} A & B_1^T \\ O & S \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} f \\ g - B_2 A^{-1} f \end{bmatrix}.$$

Solving this block upper triangular system by block backsubstitution leads to the two reduced systems (5.1) and (5.2) for y and x. These systems can be solved either directly or iteratively. In the important special case where A and $-S$ are symmetric positive definite, highly reliable methods such as Cholesky factorization or the conjugate gradient (CG) method can be applied.

https://doi.org/10.1017/S0962492904000212 Published online by Cambridge University Press

The solution strategy outlined above is commonly used in the complementary energy formulation of structural mechanics, where it is known as the *displacement* method, since the vector of nodal displacements y is computed first; the reduction to the Schur complement system (5.1) is known as *static condensation*, and the Schur complement itself is called the *assembled stiffness matrix* (McGuire and Gallagher 1979). In electrical engineering it is known as the *nodal analysis* method, and in optimization as the *range-space* method (Vavasis 1994). In all these applications, A is symmetric positive (semi)definite, $B_1 = B_2$, and $C = O$.

This approach is attractive if the order m of the reduced system (5.1) is small and if linear systems with coefficient matrix A can be solved efficiently. The main disadvantages are the need for A to be nonsingular, and the fact that the Schur complement $S = -(BA^{-1}B^T + C)$ may be completely full and too expensive to compute or to factor. Numerical instabilities may also be a concern when forming S, especially when A is ill-conditioned (Vavasis 1994). Dense Schur complements occur in the case of Stokes and Oseen problems, where A corresponds to a (vector) differential operator. Other examples include problems from optimization when B contains one or more dense columns. Note, however, that when B contains no dense columns and A^{-1} is sparse (*e.g.*, A is diagonal or block diagonal with small blocks), then S is usually quite sparse. In this case efficient (graph-based) algorithms can be used to form S, and it is sometimes possible to apply the Schur complement reduction recursively and in a way that preserves sparsity through several levels, in the sense that the number of nonzeros to be stored remains nearly constant throughout the successive reduction steps; see Maryška, Rozložník and Tůma (2000) for an example arising from the solution of groundwater flow problems.

In cases where A is positive semidefinite and singular, Schur complement reduction methods may still be applied by making use of augmented Lagrangian techniques (3.14), which replace the original saddle point system with an equivalent one with a nonsingular (1,1) block. If S is too expensive to form or factor, Schur complement reduction can still be applied by solving (5.1) by iterative methods that do not need access to individual entries of S, but only need S in the form of matrix-vector products

$$p = -Sy = (B_2 A^{-1} B_1^T + C)y.$$

The action of S on y can be computed by means of matrix-vector products with B_1^T, B_2 and C and by solving a linear system with matrix A. If the latter can be performed efficiently and the iteration converges sufficiently fast, this is a viable option. The Schur complement system (5.1), however, may be rather ill-conditioned, in which case preconditioning will be required. Preconditioning the system (5.1) is nontrivial when S is not explicitly avail-

https://doi.org/10.1017/S0962492904000212 Published online by Cambridge University Press

able. Some options are discussed in Section 10.1 below, in the context of block preconditioners.

6. Null space methods

In this section we assume that $B_1 = B_2 = B$ has full rank and $C = O$. Furthermore, we assume that $\ker(H) \cap \ker(B) = \{0\}$, where H is the symmetric part of A. The saddle point system is then

$$Ax + B^T y = f, \quad Bx = g.$$

The null space method assumes that the following are available:

(1) a particular solution $\hat{x}$ of $B\,x = g$;
(2) a matrix $Z \in \mathbb{R}^{n\times(n-m)}$ such that $BZ = O$, that is, $\mathrm{range}(Z) = \ker(B)$ (the columns of Z span the null space of B).

Then the solution set of $Bx = g$ is described by $x = Zv + \hat{x}$ as v ranges in $\mathbb{R}^{n-m}$. Substituting $x = Zv + \hat{x}$ in $Ax + B^T y = f$, we obtain $A(Zv + \hat{x}) = f - B^T y$. Pre-multiplying by the full-rank matrix Z^T and using $Z^T B^T = O$, we get

$$Z^T AZ\, v = Z^T (f - A\hat{x}), \tag{6.1}$$

a reduced system of order $n-m$ for the auxiliary unknown v. This system is nonsingular under our assumptions. Once the solution v_* has been obtained, we set $x_* = Zv_* + \hat{x}$; finally, y_* can be obtained by solving

$$BB^T y = B(f - Ax_*), \tag{6.2}$$

a reduced system of order m with a symmetric positive definite coefficient matrix BB^T. Of course, (6.2) is just the normal equations for the overdetermined system $B^T y = f - Ax_*$, or

$$\min_y \|(f - Ax_*) - B^T y\|_2,$$

which could be solved, *e.g.*, by LSQR (Paige and Saunders 1982) or a sparse QR factorization (Matstoms 1994). Just as the Schur complement reduction method can be related to expression (3.4) for $\mathcal{A}^{-1}$, the null space method is related to the alternative expression (3.8). It is interesting to observe that when A is invertible, the null space method is just the Schur complement reduction method applied to the *dual* saddle point problem

$$\begin{bmatrix} A^{-1} & Z \\ Z^T & O \end{bmatrix} \begin{bmatrix} w \\ v \end{bmatrix} = \begin{bmatrix} -\hat{x} \\ -Z^T f \end{bmatrix}.$$

This strategy subsumes a whole family of *null space methods*, which differ primarily in the way the matrix Z (often called a *null basis*) is computed;

https://doi.org/10.1017/S0962492904000212 Published online by Cambridge University Press

see the discussion below. Null space methods are quite popular in optimization, where they are usually referred to as *reduced Hessian* methods; see Coleman (1984), Fletcher (1987), Gill *et al.* (1981), Nocedal and Wright (1999) and Wolfe (1962). In this setting the matrix A is the $(n \times n)$ Hessian of the function to be minimized subject to the constraint $Bx = g$, and $Z^T AZ$ is the reduced $((n-m) \times (n-m))$ Hessian, obtained by elimination of the constraints. When Z has orthonormal columns, the reduced system (6.1) can also be seen as a projection of the problem onto the constraint set. The null space approach has been extensively used in structural mechanics where it is known under the name of *force* method, because x, the vector of internal forces, is computed first; see, *e.g.*, Berry and Plemmons (1987), Heath, Plemmons and Ward (1984), Kaneko, Lawo and Thierauf (1982), Kaneko and Plemmons (1984), Plemmons and White (1990) and Robinson (1973). Other application areas where the null space approach is used include fluid mechanics (under the somewhat misleading name of *dual variable* method, see Amit, Hall and Porsching (1981), Arioli and Manzini (2002, 2003), Arioli, Maryška, Rozložník and Tůma (2001), Gustafson and Hartmann (1983), Hall (1985), Sarin and Sameh (1998)) and electrical engineering (under the name of *loop analysis*; see Chua *et al.* (1987), Strang (1986), Tropper (1962), Vavasis (1994)).

The null space method has the advantage of not requiring A^{-1}. In fact, the method is applicable even when A is singular, as long as the condition $\ker(H) \cap \ker(B) = \{0\}$ is satisfied. The null space method is often used in applications that require the solution of a sequence of saddle point systems of the type

$$\begin{bmatrix} A_k & B^T \\ B & O \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} f_k \\ g_k \end{bmatrix}, \quad k = 1, 2, \ldots,$$

where the A_k submatrix changes with k while B remains fixed. This situations arises, for instance, in the solution of unsteady fluid flow problems, and in the *reanalysis* of structures in computational mechanics; see, *e.g.*, Batt and Gellin (1985), Hall (1985) and Plemmons and White (1990). Another example is the analysis of resistive networks with a fixed connectivity and different values of the resistances. In all these cases the null basis matrix Z needs to be computed only once.

Null space methods are especially attractive when $n - m$ is small. If A is symmetric and positive semidefinite, then $Z^T AZ$ is symmetric positive definite and efficient solvers can be used to solve the reduced system (6.1). If Z is sparse then it may be possible to form and factor $Z^T AZ$ explicitly, otherwise iterative methods must be used, such as conjugate gradients or others.

The method is less attractive if $n - m$ is large, and cannot be applied if $C \neq O$. The main difficulty, however, is represented by the need for

https://doi.org/10.1017/S0962492904000212 Published online by Cambridge University Press

a null basis Z for B. We note that computing a particular solution for $Bx = g$ is usually not a difficult problem, and it can be obtained as a byproduct of the computations necessary to obtain Z. In the case where $g = 0$ (arising for instance from the divergence-free condition in incompressible flow problems), the trivial solution $\hat{x} = 0$ will do. Hence, the main issue is the computation of a null basis Z. There are a number of methods that one can use, at least in principle, to this end. In the large and sparse case, graph-based methods invariably play a major role.

Let P denote a permutation matrix chosen so that $BP = \begin{bmatrix} B_b & B_n \end{bmatrix}$, where B_b is $m \times m$ and nonsingular (this is always possible, since B is of rank m). Then it is straightforward to verify that the matrix

$$Z = P \begin{bmatrix} -B_b^{-1} B_n \\ I \end{bmatrix}, \tag{6.3}$$

where I denotes the identity matrix of order $n - m$, is a null basis for B. This approach goes back to Wolfe (1962); a basis of the form (6.3) is called a *fundamental basis*. Quite often, the matrix $B_b^{-1} B_n$ is not formed explicitly; rather, an LU factorization of B_b is computed and used to perform operations involving B_b^{-1}. For instance, if an iterative method like CG is used to solve (6.1), then matrix-vector products with $Z^T A Z$ can be performed by means of forward and backsubstitutions with the triangular factors of B_b, in addition to matrix-vector products with B_n, B_n^T, and A.

Since there are in principle many candidate submatrices B_b, (*i.e.*, many permutation matrices P) it is natural to ask whether one can find a matrix B_b with certain desirable properties. Ideally, one would like B_b to be easy to factor, well-conditioned, and to satisfy certain sparsity requirements (either in B_b, or in its factors, or in $B_b^{-1} B_n$). Another desirable property could be some kind of diagonal dominance. In the literature, this is known as the *nice basis* problem. This is a very difficult problem in general. Consider first the sparsity requirement. Unfortunately, not all sparse matrices admit a sparse null basis. To see this, consider the matrix $B = \begin{bmatrix} I & e \end{bmatrix}$, where e is the column vector all of whose components are equal to 1; clearly, there is no explicit sparse representation for its one-dimensional null space (Gilbert and Heath 1987). Moreover, even if a sparse null basis exists, the problem of computing a null basis Z (fundamental or not) with minimal number of nonzero entries has been shown to be NP-hard (Coleman and Pothen 1986, Pothen 1984). In spite of this, there are important situations where a sparse null basis exists and can be explicitly obtained. As we shall see, there may be no need to explicitly factor or invert any submatrix of B.

An example of this is Kirchhoff's classical method for finding the currents in a resistive electrical network (Kirchhoff 1847). Our discussion closely follows Strang (1986). For this problem, B is just the node–edge incidence

https://doi.org/10.1017/S0962492904000212 Published online by Cambridge University Press

matrix of the network, or directed graph, describing the connectivity of the network. More precisely, if the network consists of $m+1$ nodes and n edges, let B_0 be the $(m+1)\times n$ matrix with entries b_{ij} given by -1 if edge j starts at node i and by $+1$ if edge j ends at node i; of course, $b_{ij}=0$ if edge j does not meet node i. Hence, B_0 behaves as a discrete divergence operator on the network: each column contains precisely two nonzero entries, one equal to $+1$ and the other equal to -1. Matrix B_0 can be shown to be of rank m; note that $B_0^T e = 0$. A full-rank matrix B can be obtained by dropping the last row of B_0; that is, by 'grounding' the last node in the network: see Strang (1986, page 112).

A null space for B can be found using Kirchhoff's Voltage Law, which implies that the sum of the voltage drops around each closed loop in the network must be zero. In other words, for current flowing around a loop there is no buildup of charge. In matrix terms, each loop current is a solution to $By=0$. Since B has full rank, there are exactly $n-m$ independent loop currents, denoted by $z_1, z_2, \ldots z_{n-m}$. The loop currents can be determined by a procedure due to Kirchhoff, which consists of the following steps.

(1) Find a *spanning tree* for the network (graph); this is a connected subgraph consisting of the $m+1$ nodes and just m edges, so that between any two nodes there is precisely one path, and there are no loops. As shown by Kirchhoff, there are exactly $t = \det BB^T$ spanning trees in the network (which is assumed to be connected).

(2) Once a spanning tree has been picked, the remaining $n-m$ edges can be used to construct the $n-m$ loop currents by noticing that adding any of these edges to the spanning tree will create a loop. For each of these *fundamental loops* we construct the corresponding column z_i of Z by setting the jth entry equal to ± 1 if edge j belongs to the loop, and equal to 0 otherwise; the choice of sign specifies the orientation of the edge.

The resulting $Z = [z_1, z_2, \ldots, z_{n-m}]$ is then called an edge–loop matrix, and is a basis for the null space of B. As a simple example, consider the directed graph of Figure 6.1, with the spanning tree on the right. In this example, $m=4$ and $n=7$. The node–edge incidence matrix B_0 for this graph is

$$B_0 = \begin{bmatrix} -1 & -1 & 0 & -1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & -1 & 0 & -1 \\ 0 & 0 & -1 & 1 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & -1 & 0 \end{bmatrix}.$$

Note that $\operatorname{rank}(B_0) = 4$; the matrix B obtained from B_0 by grounding node 5 (*i.e.*, by dropping the last row of B_0) has full row rank, equal to 4.

https://doi.org/10.1017/S0962492904000212 Published online by Cambridge University Press

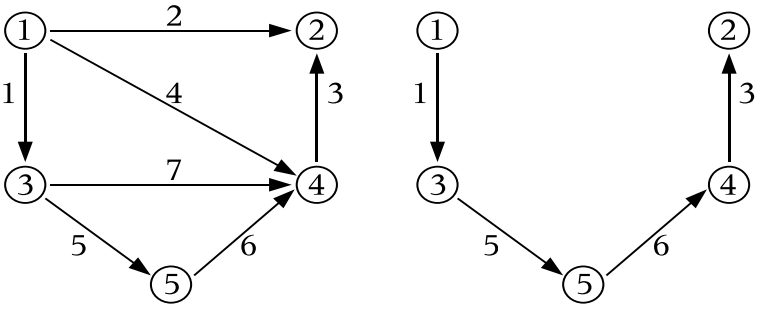


Figure 6.1. A directed graph with one of its spanning trees.

Consider now the spanning tree on the right of Figure 6.1. By adding the remaining edges (numbered 2, 4 and 7) to the tree we obtain, respectively, the following loops:

(1) $1 \to 3 \to 5 \to 4 \to 2 \to 1$, for the edge sequence (1, 5, 6, 3, -2);

(2) $1 \to 3 \to 5 \to 4 \to 1$, for the edge sequence (1, 5, 6, -4);

(3) $3 \to 5 \to 4 \to 3$, for the edge sequence (5, 6, -7).

Note that an edge was given the negative sign whenever its orientation required it. It follows that the edge–loop matrix for the network under consideration is

$$Z = \begin{bmatrix} 1 & 1 & 0 \\ -1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & -1 & 0 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \\ 0 & 0 & -1 \end{bmatrix}.$$

and it is straightforward to check that $BZ = O$.

It turns out that this elegant method is fairly general and can be applied to other problems besides the analysis of resistive networks. An important example is fluid dynamics, in particular the Darcy, Stokes and Oseen problems, where B represents a discrete divergence operator. In this case the null basis Z is called a *solenoidal basis*, since the columns of Z span the subspace of all discrete solenoidal (*i.e.*, divergence-free) functions. In other words, C can be regarded as a discrete curl operator (Chang, Giraldo and Perot 2002). In the case of finite differences on a regular grid, B is just the incidence matrix of a directed graph associated with the grid, and the cycles (loops) in this graph can be used to construct a sparse Z; see Amit *et al.* (1981), Burkardt, Hall and Porsching (1986), Chang *et al.* (2002), Hall

https://doi.org/10.1017/S0962492904000212 Published online by Cambridge University Press

(1985), Sameh and Sarin (2002) and Sarin and Sameh (1998). Note that in this context, solving the system (6.2) for y amounts to solving a Poisson equation. This methodology is not restricted to simple finite difference discretizations or to structured grids; see Alotto and Perugia (1999), Arioli and Manzini (2002, 2003), Arioli *et al.* (2001), Hall, Cavendish and Frey (1991), Sarin (1997) and Sarin and Sameh (2003) for applications to a variety of discretization methods on possibly unstructured grids.

We note that no floating-point arithmetic is needed to form Z. Furthermore, the sparsity pattern of the matrix Z^TAZ can be easily determined and the matrix Z^TAZ assembled rather cheaply, as long as it is sufficiently sparse. The sparsity will depend on the particular spanning tree used to form Z. Finding a tree that minimizes the number of nonzeros in Z is equivalent to finding the tree for which the sum of all the lengths of the fundamental loops is minimal, which is an NP-hard problem. Nevertheless, many efficient heuristics have been developed; see Tarjan (1983) for the fundamental concepts and algorithms. The relative size of n, m and $n-m$ depends on the discretization scheme used, and on whether the underlying fluid flow problem is posed in 2D or 3D. For lowest-order discretizations in 2D, $n-m$ and m are comparable, whereas $n-m$ is much larger than m in 3D or for certain mixed finite element discretizations. If sparse direct solvers are used to solve for the dual variable in (6.1), this makes the null space approach not viable in 3D. In this case iterative solvers must be used, and the spectral properties of Z^TAZ determine the convergence rate. When the matrix Z^TAZ is not formed explicitly, finding appropriate preconditioners for it requires some cleverness. Some work in this direction can be found in Coleman and Verma (2001) and in Nash and Sofer (1996) for constrained optimization problems. See also Saint-Georges, Notay and Warzée (1998) for closely related work in the context of constrained finite element analyses, and Barlow, Nichols and Plemmons (1988), James (1992) and James and Plemmons (1990) for earlier work on the use of preconditioned CG methods in the context of *implicit null space algorithms – i.e.*, null space algorithms in which the matrix Z is not formed explicitly.

For many mixed finite element formulations of second-order elliptic problems, A is symmetric positive definite and has condition number bounded independently of the discretization parameter h. In this case, fast CG convergence can be obtained by using incomplete factorization preconditioners based on Z^TZ: see Alotto and Perugia (1999). Point and block Jacobi preconditioners constructed without explicitly forming Z^TAZ have been tested in the finite element solution of the potential fluid flow problem (2.18)–(2.20) in Arioli and Manzini (2003).

Null space methods have been used for a long time in structural optimization. Some relevant references in this area include Cassell, Henderson and Kaveh (1974), Henderson and Maunder (1969), Kaveh (1979, 1992, 2004)

https://doi.org/10.1017/S0962492904000212 Published online by Cambridge University Press

and Pothen (1989). In this case B is an equilibrium matrix associated with Newton's Third Law (*i.e.*, action and reaction are equal and opposite). Now B is no longer an incidence matrix, but many of the above-described concepts can be extended to this more general situation. Algorithms to find a null basis for B developed by Coleman and Pothen (1987) consist of two phases, a combinatorial one and a numerical one. In the first phase, a maximum matching in the bipartite graph of B is used to locate the nonzero entries in the null basis. In the second phase (not needed when B is an incidence matrix), the numerical values of the nonzero entries in the basis are computed by solving certain systems of equations.

When additional structure, such as bandedness, is present in B, it is usually possible to exploit it so as to develop more efficient algorithms. Banded equilibrium matrices often arise in structural engineering. The so-called *turnback algorithm* (Topcu 1979) can be used to compute a banded Z; see also Kaneko *et al.* (1982), where an interpretation of the turnback algorithm in terms of matrix factorizations is given, and Gilbert and Heath (1987) for additional methods motivated by the turnback algorithm. We also mention Plemmons and White (1990) for approaches based on different graph-theoretic concepts with a focus on parallel implementation aspects, and Chow, Manteuffel, Tong and Wallin (2003) for another example of how structure in the constraints can be exploited to find a null basis resulting in a sparse reduced matrix $Z^T AZ$.

One problem that may occur in the null space method is that a computed null basis matrix Z may be very ill-conditioned, even numerically rank deficient. One way to avoid this problem is to compute a Z with orthonormal columns, which would be optimally conditioned. An orthonormal null basis for B can be computed by means of the QR factorization as follows. Let

$$B^T = Q \begin{bmatrix} R \\ O \end{bmatrix},$$

where Q is $n \times n$ orthogonal and R is $m \times m$, upper triangular and nonsingular. Then the first m columns of Q form an orthonormal basis for $\mathrm{range}(B^T)$ and the remaining $n - m$ columns form an orthonormal basis for $\mathrm{range}(B^T)^\perp = \ker(B)$. Therefore, if q_i denotes the ith column of Q, then

$$Z = \begin{bmatrix} q_{m+1} & q_{m+2} & \dots & q_n \end{bmatrix}$$

is the desired matrix. Of course, in the sparse case special ordering techniques must be utilized in order to maintain sparsity in Z (Amestoy, Duff and Puglisi 1996, Matstoms 1994). The fact that the columns of Z are orthonormal is advantageous not only from the point of view of conditioning, but also for other reasons. For example, in the computation of thin-plate splines

https://doi.org/10.1017/S0962492904000212 Published online by Cambridge University Press

it is often required to solve saddle point systems of the form

$$\begin{bmatrix} A+\rho I & B^T \\ B & O \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} f \\ 0 \end{bmatrix}, \tag{6.4}$$

where A is symmetric, $\rho > 0$ is a smoothing parameter, and B has full row rank; see Sibson and Stone (1991). Usually, problem (6.4) needs to be solved for several values of ρ. If Z is a null basis matrix with orthonormal columns, the coefficient matrix in the reduced system (6.1) is $Z^T(A+\rho I)Z = Z^TAZ + \rho I$. If $n-m$ is so small that a spectral decomposition $Z^TAZ = U\Lambda U^T$ of Z^TAZ can be computed, then for any ρ we have $Z^T(A+\rho I)Z = U(\Lambda+\rho I)U^T$ and the reduced linear systems can be solved efficiently.

One further advantage of having an orthonormal null basis is that the reduced system (6.1) is guaranteed to be well-conditioned if A is. For example, if A is symmetric positive definite and has condition number bounded independently of mesh size, the same is true of Z^TAZ and therefore the CG method applied to (6.1) converges in a number of iterations independent of mesh size, even without preconditioning; see Arioli and Manzini (2002) for an example from groundwater flow computations. This property may fail to hold if Z does not have orthonormal columns, generally speaking; see Arioli and Manzini (2003) and Perugia and Simoncini (2000).

Sparse orthogonal schemes have been developed by Berry, Heath, Kaneko, Lawo, Plemmons and Ward (1985), Gilbert and Heath (1987), Heath *et al.* (1984) and Kaneko and Plemmons (1984) in the context of structural optimization, and by Arioli (2000), Arioli and Manzini (2002) and Arioli *et al.* (2001) in the context of mixed-hybrid finite element formulations of the potential fluid flow problem (2.18)–(2.20). A parallel orthogonal null space scheme has been presented by Psiaki and Park (1995) for trajectory optimization problems in quadratic dynamic programming. One limitation of the QR factorization approach is that the null bases obtained by this method are often rather dense compared to those obtained by other sparse schemes; indeed, the sparsest orthogonal null basis may be considerably less sparse than an arbitrary null basis: see Coleman and Pothen (1987) and Gilbert and Heath (1987). Hence, there is a trade-off between good conditioning properties and sparsity.

Error analyses of various null space methods have been carried out by Cox and Higham (1999*a*) for dense problems, and by Arioli and Baldini (2001) for the sparse case. See further Barlow (1988), Barlow and Handy (1988), Björck and Paige (1994), Cox and Higham (1999*b*), Fletcher and Johnson (1997), Gulliksson (1994), Gulliksson and Wedin (1992), Hough and Vavasis (1997), Sun (1999) and Vavasis (1994) for stable implementations and other numerical stability aspects of algorithms for saddle point problems, in particular for equality-constrained and weighted least squares problems. Finally, appropriate stopping criteria for the CG method applied

https://doi.org/10.1017/S0962492904000212 Published online by Cambridge University Press

to the reduced system (6.1) in a finite element context have been given by Arioli and Manzini (2002).

7. Coupled direct solvers

In this section we give a brief overview of direct methods based on triangular factorizations of $\mathcal{A}$. Our discussion is limited to the symmetric case ($A = A^T$, $B_1 = B_2$ and $C = C^T$, possibly zero). As far as we know, no specialized direct solver exists for nonsymmetric saddle point problems. Although such problems are often *structurally* symmetric, in the sense that the nonzero pattern of $\mathcal{A}$ is symmetric, some form of numerical pivoting is almost certainly going to be needed for stability reasons; such pivoting would in turn destroy symmetry. See Duff, Erisman and Reid (1986) for a treatment of direct methods for general sparse matrices.

There are several ways to perform Gaussian elimination on a symmetric, possibly indefinite matrix in a way that exploits (and preserves) symmetry. A factorization of the form

$$\mathcal{A} = \mathcal{Q}^T \mathcal{L} \mathcal{D} \mathcal{L}^T \mathcal{Q}, \tag{7.1}$$

where $\mathcal{Q}$ is a permutation matrix, $\mathcal{L}$ is unit lower triangular, and $\mathcal{D}$ a block diagonal matrix with blocks of dimension 1 and 2 is usually referred to as an LDL^T factorization. The need for pivot blocks of size 2 is made clear by the following simple example:

$$\mathcal{A} = \left[\begin{array}{cc|c} 0 & 1 & 1 \\ 1 & 0 & 1 \\ \hline 1 & 1 & 0 \end{array}\right],$$

for which selecting pivots from the main diagonal is impossible. Diagonal pivoting may also fail on matrices with a zero-free diagonal due to instabilities. The use of 2×2 pivot blocks dates back to Lagrange (1759). In 1965, W. Kahan (in correspondence with R. de Meersman and L. Schotsmans) suggested that Lagrange's method could be used to devise stable factorizations for symmetric indefinite matrices. The idea was developed by Bunch and Parlett (1971), resulting in a stable algorithm for factoring symmetric indefinite matrices at a cost comparable to that of a Cholesky factorization for positive definite ones. The Bunch–Parlett pivoting strategy is akin to complete pivoting; in subsequent papers (Bunch 1974, Bunch and Kaufman 1977), alternative pivoting strategies requiring only $O(n^2)$ comparisons for a dense $n \times n$ matrix have been developed; see also Fletcher (1976*b*). The Bunch–Kaufman pivoting strategy (Bunch and Kaufman 1977) is widely accepted as the algorithm of choice for factoring dense symmetric indefinite matrices. In the sparse case, the pivoting strategy is usually relaxed in

https://doi.org/10.1017/S0962492904000212 Published online by Cambridge University Press

order to maintain sparsity in $\mathcal{L}$. Several sparse implementations are available; see Duff, Gould, Reid, Scott and Turner (1991), Duff and Reid (1983, 1995) and Liu (1987). In this case the permutation matrix $\mathcal{Q}$ is the result of symmetric row and column interchanges aimed at preserving sparsity in the factors as well as numerical stability. While the Bunch–Kaufman algorithm is normwise backward stable, the resulting factors can have unusual scaling, which may result in a degradation of the accuracy of computed solutions. As reported in Ashcraft, Grimes and Lewis (1998), such difficulties have been observed in the solution of saddle point systems arising in sparse nonlinear optimization codes. We refer the reader to Ashcraft *et al.* (1998) for a thorough discussion of such accuracy issues and ways to address these problems; see also Vavasis (1994).

We note that when A is positive definite and B has full rank, the saddle point matrix $\mathcal{A}$ admits an LDL^T factorization with $\mathcal{D}$ diagonal and $\mathcal{Q} = I$ (*i.e.*, no pivoting is needed). Indeed, since A is positive definite it can be decomposed as $A = L_A D_A L_A^T$ with L_A unit lower triangular and D_A diagonal (and positive definite); furthermore the Schur complement $S = -(C + BA^{-1}B^T)$ is negative definite and therefore it can be decomposed as $S = -L_S D_S L_S^T$. Hence, we can write

$$\mathcal{A} = \begin{bmatrix} A & B^T \\ B & -C \end{bmatrix} = \begin{bmatrix} L_A & O \\ L_B & L_S \end{bmatrix} \begin{bmatrix} D_A & O \\ O & -D_S \end{bmatrix} \begin{bmatrix} L_A^T & L_B^T \\ O & L_S^T \end{bmatrix} = \mathcal{L}\,\mathcal{D}\,\mathcal{L}^T, \quad (7.2)$$

where $L_B = BL_A^{-T}D_A^{-1}$; note that $BA^{-1}B^T = L_B D_A L_B^T$. In practice, however, the factors will be rather dense with the original ordering, and symmetric permutations have to be used in order to preserve sparsity. Note that L_S and L_B will be completely full if the Schur complement is. However, not all sparsity-preserving permutations are acceptable. It can be shown that there exist permutation matrices $\mathcal{Q}$ such that $\mathcal{Q}\mathcal{A}\mathcal{Q}^T$ does not have an LDL^T factorization with $\mathcal{D}$ diagonal. Furthermore, some permutations may lead to numerical instability problems.

For many symmetric indefinite codes the factorization consists of two phases, a symbolic and a numeric one. In the symbolic phase, an initial fill-reducing ordering is computed based on the structure of $\mathcal{A}$ only. This is often some variant of minimum degree or nested dissection (Duff *et al.* 1986). In the numeric phase, the actual factorization is computed. Frequently in the course of this phase, the pivot order from the symbolic phase may have to be altered for numerical stability reasons. There are, however, a few exceptions to this rule. An important one is the quasidefinite case discussed in Section 3.3, *i.e.*, when C (as well as A) is symmetric positive definite. In this case $\mathcal{Q}\mathcal{A}\mathcal{Q}^T$ always has an LDL^T factorization with $\mathcal{D}$ diagonal, regardless of the choice of $\mathcal{Q}$; see Vanderbei (1995). This is an important result: it suggests that the fill-reducing ordering computed in the symbolic phase of the factorization will not need to be altered in the course of the numeric

https://doi.org/10.1017/S0962492904000212 Published online by Cambridge University Press

phase because of stability concerns. Since no pivoting is used in the numeric phase, it is possible to exploit all the features of modern supernodal sparse Cholesky factorization codes (Ng and Peyton 1993). The resulting algorithm is more efficient than performing a Bunch–Parlett or Bunch–Kaufman factorization. Numerical stability considerations in Vanderbei (1995) suggest that the resulting factorization is usually sufficiently accurate. A stability analysis was given in Gill *et al.* (1996), where the close relationship between $\mathcal{A}$ and its nonsymmetric positive definite form (3.10) was used – together with results in Golub and Van Loan (1979) – to derive stability conditions.

A further exception (with $C = O$) has been identified by Tůma (2002). For a large class of saddle point matrices arising from mixed and hybrid finite element discretizations it is possible to prove the existence of static, fill-reducing pre-orderings $\mathcal{Q}$ such that the permuted matrix $\mathcal{Q}\mathcal{A}\mathcal{Q}^T$ has the LDL^T factorization with $\mathcal{D}$ diagonal. Such pre-orderings are characterized in terms of conditions on the resulting elimination tree. The factorization can be carried out in three phases: a first, symbolic phase in which an initial fill-reducing ordering is computed and the corresponding elimination tree is built; a second phase, also symbolic, where the initial ordering is modified so that the permuted matrix satisfies the conditions that guarantee the existence of the factorization; and a final, numeric phase where the LDL^T factorization itself is computed. The numerical experiments in Tůma (2002) show that this is an effective approach. As in the quasidefinite case, no numerical stability problems have appeared in practice; however, a formal error analysis has not yet been carried out. We mention that examples of saddle point systems that cause difficulties for symmetric indefinite factorization algorithms have been pointed out in Vavasis (1994).

For the general case, sophisticated strategies for computing sparse LDL^T factorizations with 1×1 and 2×2 pivot blocks have been developed over many years by Duff and Reid together with several collaborators; see Duff (1994), Duff *et al.* (1986), Duff *et al.* (1991), Duff and Pralet (2004), Duff and Reid (1983, 1995, 1996). This work has led to a series of widely used codes that are part of the HSL library; see Section 12 for information on how to access these codes. The first is MA27, developed in the early 1980s; the second is MA47, a code geared towards symmetric indefinite systems in saddle point form (with $C = O$); later came MA57 and, recently, the MA67 code. All these codes, except for MA67, are multifrontal codes. The need for codes specifically designed for saddle point systems (with $C = O$) is clear when one considers the presence of the zero block in position (2,2). Clearly, any form of symmetric pivoting must be restricted so that pivots are not chosen from the zero block. Failure to do so during the symbolic phase leads to a very large number of pivot order alterations during the numeric factorization phase, dramatically slowing down the computation. Furthermore, the structure of the matrix during the subsequent factorization

https://doi.org/10.1017/S0962492904000212 Published online by Cambridge University Press

steps must also be taken into account in order to avoid excessive fill-in. The code MA47 has been designed with both of these goals in mind; see Duff (1994) and Duff and Reid (1996) for details. The code MA67, also geared towards saddle point systems, is based on design concepts that are quite different from those of the previous codes; as already mentioned it is not a multifrontal code, and furthermore it does not have separate symbolic and numeric phases. Instead, the numerical values of the entries are taken into account during the selection of the pivots. A Markowitz-type strategy is used to balance sparsity and numerical stability needs. Unfortunately, the extensive comparison of HSL codes performed in Gould and Scott (2004) indicates that MA67 is generally inferior to its predecessors.

Other sparse direct solvers for symmetric indefinite systems, based on different design principles, exist; see for instance the recent reports (Meshar and Toledo 2005, Schenk and Gärtner 2004), and Section 12 below. While these codes have not been developed specifically for saddle point matrices, they may work quite well on such problems. For instance, Schenk and Gärtner (2004) report that their code factors a saddle point matrix from optimization of order approximately 2 million (with 6 million nonzeros) in less than a minute on a 2.4 GHz Intel 32-bit processor, producing a factor with about 1.4×10^8 nonzeros.

Although fairly reliable in practice, sparse LDL^T factorization methods are not entirely foolproof. Besides the examples given in Vavasis (1994), a few failures to compute acceptably accurate solutions have been reported in Gould and Scott (2004), even with the use of iterative refinement; see also Schenk and Gärtner (2004). Nevertheless, sparse LDL^T methods are the solvers of choice in various sparse optimization codes, where they are often preferred to methods based on Schur complement reduction ('normal equations methods') for both stability and sparsity reasons. Sparse direct solvers have been less popular in the numerical solution of PDE problems because of their intrinsic storage and computational limitations, although these solvers can be quite competitive for 2D problems; see, *e.g.*, Perugia *et al.* (1999). For saddle point systems arising from PDE problems on 3D meshes, it is necessary to turn to iterative methods.

8. Stationary iterations

We begin our discussion of iterative algorithms with stationary schemes. These methods have been popular for years as 'standalone' solvers, but nowadays they are most often used as preconditioners for Krylov subspace methods (equivalently, the convergence of these stationary iterations can be accelerated by Krylov subspace methods.) Another common use for stationary iterations is as smoothers for multigrid methods; we return to this in Section 11.

https://doi.org/10.1017/S0962492904000212 Published online by Cambridge University Press

8.1. The Arrow–Hurwicz and Uzawa methods

The first iterative schemes for the solution of saddle point problems of a rather general type were the ones developed by the mathematical economists Arrow, Hurwicz and Uzawa (see Arrow *et al.* (1958)). The original papers addressed the case of inequality constraints; see Polyak (1970) for an early study of these methods in the context of the equality-constrained problem (1.4)–(1.5).

The Arrow–Hurwicz and Uzawa methods are stationary schemes consisting of simultaneous iterations for both x and y, and can be expressed in terms of splittings of the matrix $\mathcal{A}$. By elimination of one of the unknown vectors, they can also be interpreted as iterations for the reduced (Schur complement) system. Hence, these algorithms may be regarded both as coupled and as segregated solvers.

We start with Uzawa's method (Uzawa 1958), which enjoys considerable popularity in fluid dynamics, especially for solving the (steady) Stokes problem (Fortin and Glowinski 1983, Glowinski 1984, Glowinski 2003, Temam 1984, Turek 1999). For simplicity, we assume A is invertible and we describe the algorithm in the case $B_1 = B_2 = B$ and $C = O$. Generalization to problems with $B_1 \neq B_2$ or $C \neq O$ is straightforward. Starting with initial guesses x_0 and y_0, Uzawa's method consists of the following coupled iteration:

$$\begin{cases} Ax_{k+1} = f - B^T y_k, \\ y_{k+1} = y_k + \omega(Bx_{k+1} - g), \end{cases} \tag{8.1}$$

where $\omega > 0$ is a relaxation parameter. As noted in Golub and Overton (1988, page 591) (see also Saad (2003, page 258)), this iteration can be written in terms of a matrix splitting $\mathcal{A} = \mathcal{P} - \mathcal{Q}$, *i.e.*, as the fixed-point iteration

$$\mathcal{P}u_{k+1} = \mathcal{Q}u_k + b,$$

where

$$\mathcal{P} = \begin{bmatrix} A & O \\ B & -\frac{1}{\omega}I \end{bmatrix}, \quad \mathcal{Q} = \begin{bmatrix} O & -B^T \\ O & -\frac{1}{\omega}I \end{bmatrix}, \quad \text{and} \quad u_k = \begin{bmatrix} x_k \\ y_k \end{bmatrix}. \tag{8.2}$$

Note that the iteration matrix is

$$\mathcal{T} = \mathcal{P}^{-1}\mathcal{Q} = \begin{bmatrix} O & -A^{-1}B^T \\ O & I - \omega BA^{-1}B^T \end{bmatrix},$$

and therefore the eigenvalues of $\mathcal{T}$ are all real (and at least n of them are exactly zero).

On the other hand, if we use the first equation in (8.1) to eliminate x_{k+1} from the second one we obtain

$$y_{k+1} = y_k + \omega\,(BA^{-1}f - g - BA^{-1}B^T y_k), \tag{8.3}$$

https://doi.org/10.1017/S0962492904000212 Published online by Cambridge University Press

showing that Uzawa's method is equivalent to a stationary Richardson iteration applied to the Schur complement system

$$BA^{-1}B^T y = BA^{-1}f - g. \tag{8.4}$$

If A is symmetric and positive definite, so is $BA^{-1}B^T$. Denoting the smallest and largest eigenvalues of $BA^{-1}B^T$ by $\lambda_{\min}$ and $\lambda_{\max}$, respectively, it is well known that Richardson's iteration (8.3) converges for all ω such that

$$0 < \omega < \frac{2}{\lambda_{\max}};$$

see, *e.g.*, Saad (2003, page 114). Furthermore, the spectral radius of the iteration matrix $I - \omega BA^{-1}B^T$ of (8.3) is minimized by taking

$$\omega_* = \frac{2}{\lambda_{\min} + \lambda_{\max}}.$$

In some special cases, the optimal value of ω can be estimated analytically (Langer and Queck 1986). An important example is that of (LBB-stable) discretizations of the steady-state Stokes system (2.10)–(2.12), for which the Schur complement is *spectrally equivalent* to the identity (Verfürth 1984*a*). This means that the eigenvalues of $BA^{-1}B^T$ are bounded below and above by positive constants, *i.e.*, by numbers that do not depend on the mesh size h. As a result, Uzawa's iteration converges at a rate independent of h. We note that this is not the case for the so-called *generalized Stokes* problem arising from the solution of the unsteady Stokes problems using implicit methods; see Cahouet and Chabard (1988). The convergence of Uzawa's method in this case is rather slow, as it is for most other problems, particularly for the Oseen problem at high Reynolds numbers (small ν) (Fortin and Fortin 1985). Improved convergence can be achieved by suitably preconditioning the Uzawa iteration: see Cahouet and Chabard (1988) and Elman and Golub (1994). Uzawa-type algorithms for the stabilized case ($C \neq O$) were first studied by Vincent and Boyer (1992). An Uzawa-type method with variable relaxation parameters was proposed by Hu and Zou (2001). Uzawa's method is still being actively developed by many researchers: recent papers discussing various extensions and improvements of Uzawa's classical algorithm include Bertrand and Tanguy (2002), Bramble, Pasciak and Vassilev (2000), Cao (2004*b*), Chen (1998), Cui (2004), Hu and Zou (2002), Liu and Xu (2001), Maday, Meiron, Patera and Ronquist (1993), Nochetto and Pyo (2004), Zsaki, Rixen and Paraschivoiu (2003). Not all applications of Uzawa's method are to fluid flow problems: see Ito and Kunisch (1999) for a recent application to image restoration.

The bulk of the computational effort in Uzawa's method is spent in the solution of linear systems involving A. These systems can be solved by direct methods or, more often, by an inner iterative scheme. For instance, in the

https://doi.org/10.1017/S0962492904000212 Published online by Cambridge University Press

case of the Stokes problem A is a direct sum of discrete Laplace operators, and multigrid methods are a natural choice. The case of inexact inner solves has been studied in detail in Bramble, Pasciak and Vassilev (1997), Chen (1998), Cheng (2000), Cheng and Zou (2003), Cui (2002), Elman and Golub (1994), Peters, Reichelt and Reusken (2004), Robichaud, Tanguy and Fortin (1990) and Zulehner (2002).

The Arrow–Hurwicz method (Arrow and Hurwicz 1958) may be regarded as an inexpensive alternative to Uzawa's method, useful when solves with A are too expensive. Here we follow the derivation given in Saad (2003, Section 8.4). By noting that iterate x_{k+1} given by the first of (8.1) is the minimizer of the objective function

$$\phi(x) = \frac{1}{2}x^T Ax - x^T(f - B^T y_k),$$

we can derive a less expensive method by taking one step in the direction of the (negative) gradient of $\phi(x)$, with fixed step length α. The resulting method is the Arrow–Hurwicz iteration:

$$\begin{cases} x_{k+1} = x_k + \alpha(f - Ax_k - B^T y_k), \\ y_{k+1} = y_k + \omega(Bx_{k+1} - g). \end{cases} \tag{8.5}$$

As in the case of Uzawa's method, the Arrow–Hurwicz method can be cast as a fixed-point iteration induced by the splitting

$$\mathcal{A} = \mathcal{P} - \mathcal{Q} \quad \text{where} \quad \mathcal{P} = \begin{bmatrix} \frac{1}{\alpha}I & O \\ B & -\frac{1}{\omega}I \end{bmatrix}, \quad \mathcal{Q} = \begin{bmatrix} \frac{1}{\alpha}I - A & -B^T \\ O & -\frac{1}{\omega}I \end{bmatrix}. \tag{8.6}$$

The convergence of this algorithm depends on the two relaxation parameters, α and ω. Convergence conditions and theoretical estimates for the optimal choice of parameters have been given in Fortin and Glowinski (1983), Polyak (1970), Queck (1989) and, more recently, in Astrakhantsev (2001) and Bychenkov (2002). Because the convergence of the Arrow–Hurwicz method is usually rather slow, various improvements have been proposed, including preconditioned variants of the form

$$\begin{cases} x_{k+1} = x_k + \alpha\, Q_A^{-1}(f - Ax_k - B^T y_k), \\ y_{k+1} = y_k + \omega\, Q_B^{-1}(Bx_{k+1} - g), \end{cases} \tag{8.7}$$

where Q_A and Q_B are appropriately chosen 'preconditioning matrices': see Astrakhantsev (2001), Queck (1989) and Robichaud *et al.* (1990). Obviously, the line between preconditioned versions of the Arrow–Hurwicz algorithm and inexact/preconditioned variants of Uzawa's method is blurred. Additional hybrids can be obtained by combining the Arrow–Hurwicz and CG algorithms: see, *e.g.*, Aboulaich and Fortin (1989) and Stoyan (2001).

https://doi.org/10.1017/S0962492904000212 Published online by Cambridge University Press

8.2. Penalty and multiplier methods

Here we assume that $A = A^T$ is positive semidefinite, $B_1 = B_2 = B$ is of full rank, and $C = O$. We further assume that $\ker(A) \cap \ker(B) = \{0\}$, so that the saddle point system has a unique solution. As we have noted in Section 1, the saddle point problem is then equivalent to the constrained minimization problem

$$\min \ J(x) = \frac{1}{2} x^T A x - f^T x, \tag{8.8}$$

$$\text{subject to} \ \ Bx = g. \tag{8.9}$$

A very old method for finding the solution x_* of (8.8)–(8.9) is based on the observation that

$$x_* = \lim_{\gamma \to \infty} x(\gamma),$$

where $x(\gamma)$ is the unique solution of the unconstrained minimization problem

$$\min \ \hat{J}(x) \equiv J(x) + \frac{\gamma}{2} \|Bx - g\|_2^2.$$

In mechanical terms, rigid constraints can be thought of as limiting cases of very large restoring forces, *i.e.*, in this case, forces with potential energy of the form $U(x) = \frac{\gamma}{2}(Bx - g)^T(Bx - g)$; see Courant (1943). The minimizer $x(\gamma)$ can be found by setting the gradient of $J(x) + \frac{\gamma}{2}(Bx - g)^T(Bx - g)$ to zero, leading to the linear system

$$(A + \gamma B^T B)\, x = f + \gamma B^T g. \tag{8.10}$$

If we let $y(\gamma) = \gamma(Bx(\gamma) - g)$, where $x = x(\gamma)$ is the solution of (8.10), then it is possible to prove that

$$\|x_* - x(\gamma)\|_2 = O(\gamma^{-1}) \quad \text{and} \quad \|y_* - y(\gamma)\|_2 = O(\gamma^{-1}) \quad \text{for} \ \ \gamma \to \infty;$$

see Glowinski (1984, pages 21–22). Therefore, provided that γ is taken large enough, $x(\gamma)$ and $y(\gamma)$ are approximate solutions of the original saddle point problem. The penalty method can be thought of as an *approximate* direct method. Since a monotonically increasing sequence $\gamma_1, \gamma_2, \ldots$ of values of γ may be used to compute better and better approximations to x_*, it may also be regarded as a stationary iterative method. In some cases, the choice of γ may be made on the basis of physical considerations: see, *e.g.*, Destuynder and Nevers (1990), Nour-Omid and Wriggers (1986).

Since the matrix $A + \gamma B^T B$ in (8.10) is symmetric and positive definite for $\gamma > 0$, one can in principle use a Cholesky factorization or the CG method to compute $x(\gamma)$. Unfortunately, such an approach cannot be recommended, since the condition number of $A + \gamma B^T B$ grows like a (possibly large) multiple of γ; see Glowinski (1984, pages 22–23) and Van Loan (1985) for an analysis of the penalty method applied to equality-constrained least squares

https://doi.org/10.1017/S0962492904000212 Published online by Cambridge University Press

problems. In practice, for large values of γ the coefficient matrix $A+\gamma B^T B$ is dominated by the (highly singular) term $\gamma B^T B$, and accurate solutions of (8.10) are difficult to obtain.

This drawback of the penalty method can be overcome in two ways. One way is to observe that $(x(\gamma), y(\gamma))$ is the unique solution of the *regularized* saddle point system

$$\begin{bmatrix} A & B^T \\ B & -\varepsilon I \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} f \\ g \end{bmatrix}, \quad \varepsilon = \gamma^{-1}. \tag{8.11}$$

Stable solution of this linear system is now possible, for instance using a sparse direct solver; this approach is popular in optimization. However, using a direct solver is not always feasible.

The other option is to modify the penalty method so as to avoid ill-conditioning. This leads to the *method of multipliers*, developed independently by Arrow and Solow (1958, page 172), Hestenes (1969) and Powell (1969). A further advantage of this method, which combines the use of penalty with Lagrange multipliers, is that it produces the *exact* solution (x_*, y_*) rather than an approximate one. The method of multipliers can be described as follows. Select $\gamma > 0$ and consider the *augmented Lagrangian*

$$\mathcal{L}(x, y) = J(x) + (Bx - g)^T y + \frac{\gamma}{2} \|Bx - g\|_2^2. \tag{8.12}$$

Given an approximation y_k for the Lagrange multiplier vector y, we compute the minimum x_{k+1} of the function $\psi(x) \equiv \mathcal{L}(x, y_k)$. This requires solving the linear system

$$(A + \gamma B^T B)x = f - B^T y_k + \gamma B^T g. \tag{8.13}$$

Now we use the computed solution x_{k+1} to obtain the new Lagrange multiplier approximation y_{k+1} according to

$$y_{k+1} = y_k + \gamma\,(Bx_{k+1} - g),$$

and so on. Clearly, the method of multipliers is precisely Uzawa's iteration applied to the saddle point system

$$\begin{bmatrix} A + \gamma B^T B & B^T \\ B & O \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} f + \gamma B^T g \\ g \end{bmatrix}, \tag{8.14}$$

which has exactly the same solution (x_*, y_*) as the original one. Note that the parameter γ does double duty here, in that it appears both in the definition of the augmented Lagrangian and as the relaxation parameter for the Uzawa iteration. As we know from our discussion of Uzawa's method, the iteration converges for $\gamma \in (0, 2/\rho)$ where ρ denotes the largest eigenvalue of the Schur complement $B(A+\gamma B^T B)^{-1} B^T$. This interval becomes unbounded, and the rate of convergence arbitrarily large, as $\gamma \to \infty$. Again, taking too large a value of γ results in extreme ill-conditioning of the coefficient

https://doi.org/10.1017/S0962492904000212 Published online by Cambridge University Press

matrix in (8.13). It is necessary to strike a balance between the rate of convergence of the method and the conditioning properties of (8.13). The choice of γ and many other aspects of the multiplier method have been discussed by a number of authors, including Bertsekas (1982), Chen and Zou (1999), Fortin and Glowinski (1983), Greif *et al.* (2005), Hestenes (1975), Luenberger (1984), Zienkiewicz, Vilotte, Toyoshima and Nakazawa (1985). See further Awanou and Lai (2005) for a study of the nonsymmetric case. Another possibility is to combine the augmented Lagrangian method with a (preconditioned) Arrow–Hurwicz scheme; see Fortin and Glowinski (1983, page 26) and Kouhia and Menken (1995) for an application of this idea to problems in structural mechanics.

8.3. Other stationary iterations

In addition to the foregoing algorithms, a number of other stationary iterations based on matrix splittings $\mathcal{A} = \mathcal{P} - \mathcal{Q}$ can be found in the literature. In particular, SOR- and block-SOR-type schemes have been proposed in Strikwerda (1984) for the Stokes problem, in Barlow *et al.* (1988), Benzi (1993) and Plemmons (1986) for structural analysis computations, and in Chen (1998), Golub, Wu and Yuan (2001) and Li, Li, Evans and Zhang (2003) for general saddle point systems. Some of these schemes can be interpreted as preconditioned or inexact variants of the classical Uzawa algorithm. Alternating-direction iterative methods for saddle point problems have been studied in Brown (1982), Douglas Jr., Durán and Pietra (1986, 1987). Other stationary iterative methods for saddle point problems have been studied in Bank, Welfert and Yserentant (1990), Benzi and Golub (2004), Dyn and Ferguson (1983), Golub and Wathen (1998) and Tong and Sameh (1998); since these methods are most often used as preconditioners for Krylov subspace methods, we defer their description to Section 10.

9. Krylov subspace methods

In this section we discuss Krylov subspace methods for solving (preconditioned) saddle point problems. Our goal is not to survey all existing methods and implementations (more complete surveys can be found, *e.g.*, in the monographs by Greenbaum (1997), Saad (2003) and van der Vorst (2003) or in the papers by Eiermann and Ernst (2001) and Freund *et al.* (1992)), but to describe the main properties of the most commonly used methods. We discuss the general theory, the main convergence results, and implementation details. For simplicity, we describe the basics of Krylov subspace methods for the unpreconditioned and nonsingular system (1.1)–(1.2). The later sections will describe the general ideas of preconditioning (Section 10), and different preconditioning techniques specifically constructed for (generalized) saddle point systems (Sections 10.1–10.4).

https://doi.org/10.1017/S0962492904000212 Published online by Cambridge University Press

9.1. General theory

Suppose that u_0 is an initial guess for the solution u of (1.1)–(1.2), and define the initial residual $r_0 = b - \mathcal{A}u_0$. Krylov subspace methods are iterative methods whose kth iterate u_k satisfies

$$u_k \in u_0 + \mathcal{K}_k(\mathcal{A}, r_0), \quad k = 1, 2, \ldots, \tag{9.1}$$

where

$$\mathcal{K}_k(\mathcal{A}, r_0) \equiv \text{span}\,\{r_0, \mathcal{A}r_0, \ldots, \mathcal{A}^{k-1}r_0\} \tag{9.2}$$

denotes the kth Krylov subspace generated by $\mathcal{A}$ and r_0. It is well known that the Krylov subspaces form a nested sequence that ends with dimension $d \equiv \dim \mathcal{K}_{n+m}(\mathcal{A}, r_0) \leq n + m$, *i.e.*,

$$\mathcal{K}_1(\mathcal{A}, r_0) \subset \cdots \subset \mathcal{K}_d(\mathcal{A}, r_0) = \cdots = \mathcal{K}_{n+m}(\mathcal{A}, r_0).$$

In particular, for each $k \leq d$, the Krylov subspace $\mathcal{K}_k(\mathcal{A}, r_0)$ has dimension k. Because of the k degrees of freedom in the choice of the iterate u_k, k constraints are required to make u_k unique. In Krylov subspace methods this is achieved by requiring that the kth residual $r_k = b - \mathcal{A}u_k$ is orthogonal to a k-dimensional space $\mathcal{C}_k$, called the constraints space:

$$r_k = b - \mathcal{A}u_k \in r_0 + \mathcal{A}\mathcal{K}_k(\mathcal{A}, r_0), \qquad r_k \perp \mathcal{C}_k. \tag{9.3}$$

Orthogonality here is meant with respect to the Euclidean inner product. The relations (9.1)–(9.3) show that Krylov subspace methods are based on a general type of *projection process* that can be found in many areas of mathematics. For example, in the language of the finite element method, we may consider $\mathcal{K}_k(\mathcal{A}, r_0)$ the test and $\mathcal{C}_k$ the trial space for constructing the approximate solution u_k. In this sense the projection process (9.1)–(9.3) corresponds to the Petrov–Galerkin framework; see, *e.g.*, Quarteroni and Valli (1994, Chapter 5). The interpretation of Krylov subspace methods as projection processes was popularized by Saad in a series of papers in the early 1980s (Saad 1981, 1982). A survey of his approach can be found in his book, Saad (2003). For additional analyses of Krylov subspace methods in terms of projections see, *e.g.*, Barth and Manteuffel (1994) and Eiermann and Ernst (2001).

Knowing the properties of the system matrix $\mathcal{A}$ it is possible to determine constraint spaces $\mathcal{C}_k$ that lead to uniquely defined iterates u_k, $k = 1, 2, \ldots$, in (9.1)–(9.3). Examples for such spaces are given in the following theorem.

Theorem 9.1. Suppose that the Krylov subspace $\mathcal{K}_k(\mathcal{A}, r_0)$ has dimension k. If

(C) $\mathcal{A}$ is symmetric positive definite and $\mathcal{C}_k = \mathcal{K}_k(\mathcal{A}, r_0)$, or
(M) $\mathcal{A}$ is nonsingular and $\mathcal{C}_k = \mathcal{A}\mathcal{K}_k(\mathcal{A}, r_0)$,

https://doi.org/10.1017/S0962492904000212 Published online by Cambridge University Press

then there exists a uniquely defined iterate u_k of the form (9.1) for which the residual $r_k = b - \mathcal{A}u_k$ satisfies (9.3).

Proof. See Saad (2003, Proposition 5.1). □

Items (C) and (M) in Theorem 9.1 represent mathematical characterizations of the projection properties of well-known Krylov subspace methods. Item (C) characterizes the conjugate gradient (CG) method of Hestenes and Stiefel for symmetric positive definite matrices (Hestenes and Stiefel 1952). Note that if $\mathcal{A}$ is *not* symmetric positive definite, an approximate solution u_k satisfying both (9.1)–(9.2) and (9.3) with $\mathcal{C}_k = \mathcal{K}_k(\mathcal{A}, r_0)$ may not exist; *cf.*, *e.g.*, Brown (1991) and Cullum and Greenbaum (1996). Nevertheless, there are several implementations of this projection process, in particular the full orthogonalization method (FOM) of Saad (1981). Implementations of the projection process characterized by item (M) are the minimal residual (MINRES) method of Paige and Saunders (1975) for nonsingular symmetric (possibly indefinite) matrices, and the generalized minimal residual (GMRES) method of Saad and Schultz (1986) for general nonsingular matrices. Further mathematically equivalent implementations are discussed in Liesen, Rozložník and Strakoš (2002).

Numerous other choices of constraint spaces for constructing Krylov subspace methods exist. For example, in case of a nonsymmetric matrix $\mathcal{A}$ one may choose $\mathcal{C}_k = \mathcal{K}_k(\mathcal{A}^T, r_0)$, which represents a generalization of the projection process characterized in item (C). Specific implementations based on this choice include the method of Lanczos (1950) and the biconjugate gradient (BiCG) method of Fletcher (1976*a*). However, for a general nonsymmetric matrix $\mathcal{A}$ the process based on $\mathcal{C}_k = \mathcal{K}_k(\mathcal{A}^T, r_0)$ is not well defined, because it may happen that no iterate u_k satisfying both (9.1)–(9.2) and (9.3) exists. In an actual implementation such as BiCG this will lead to a breakdown. In practice such breakdowns are unlikely to occur, but near-breakdowns may cause irregular convergence and serious build-up of rounding errors, see Greenbaum (1997, Chapter 5) for further discussion. Such instabilities are often overcome by the stabilized BiCG (BiCGStab) method of van der Vorst (1992), which combines the BiCG projection principle with an additional minimization step in order to 'stabilize' the convergence behaviour. It is important to note that although BiCGStab is based on BiCG, it avoids using the transposed matrix $\mathcal{A}^T$. Closely related is the transpose-free quasi-minimal residual method (TFQMR) for general nonsymmetric matrices developed by Freund (1993). While both BiCGStab and TFQMR typically produce smoother convergence curves than BiCG, none of these methods is guaranteed to be free of breakdowns. A related BiCG-like method, which overcomes some of the numerical instabilities of BiCG, is the quasi-minimal residual (QMR) method for general nonsymmetric matrices of Freund and Nachtigal (1991). We point out that, despite the

https://doi.org/10.1017/S0962492904000212 Published online by Cambridge University Press

naming similarities, QMR and TFQMR are not mathematically equivalent, and hence may produce completely different approximate solutions even in exact arithmetic. In case of a symmetric (possibly indefinite) matrix $\mathcal{A}$ the Lanczos process underlying the QMR method can be simplified, leading to a mathematically equivalent but numerically more efficient implementation. This is exploited in the symmetric QMR (SQMR) method of Freund and Nachtigal (1994).

Additional Krylov subspace methods can be constructed by using different search spaces in (9.1). For example, the projection process

$$u_k \in u_0 + \mathcal{A}\mathcal{K}_k(\mathcal{A}, r_0), \qquad r_k \in r_0 + \mathcal{A}\mathcal{K}_k(\mathcal{A}, \mathcal{A}r_0) \perp \mathcal{K}_k(\mathcal{A}, r_0), \tag{9.4}$$

yields a uniquely defined iterate u_k whenever $\mathcal{A}$ is nonsingular and symmetric. An implementation of this mathematical principle is the SYMMLQ method of Paige and Saunders (1975) for nonsingular symmetric (possibly indefinite) matrices.

9.2. Convergence analysis

In exact arithmetic, the methods that are mathematically described by items (C) and (M) in Theorem 9.1 as well as by (9.4) for symmetric nonsingular $\mathcal{A}$ terminate with the exact solution in step $d \equiv \dim \mathcal{K}_{n+m}(\mathcal{A}, r_0)$, *i.e.*, they yield $u_d = u$. This feature is called the *finite termination property*. In practice, however, one is typically not interested in computing the exact solution. For example, when the linear system represents a discretized partial differential equation, then an *approximate solution* u_k of the linear system with an error norm on the level of the discretization error is often sufficient. Once this error level is reached, the iterative method can be stopped. In this way iterative methods such as Krylov subspace methods may significantly outperform any direct method which cannot be stopped prematurely, because their intermediate results do not represent approximate solutions. The question is how fast can a given Krylov subspace method reach a given accuracy level.

To analyse this question we use the fact that the geometric *orthogonality condition* expressed in (9.3) is often equivalent to an algebraic *optimality condition* for a certain norm of either the error $e_k = u - u_k$ or the residual $r_k = b - \mathcal{A}u_k$. This optimality condition can be derived using the following well-known theorem for best approximations in Hilbert spaces; see, *e.g.*, Conway (1990, Chapter 1).

Theorem 9.2. Suppose that $\mathcal{H}$ is a Hilbert space with inner product $\langle \cdot, \cdot \rangle$ and associated norm $\| \cdot \|$. If $\mathcal{M} \subset \mathcal{H}$ is a closed linear subspace, then for each fixed $h \in \mathcal{H}$ there exists a unique element $m_0 \in \mathcal{M}$ with $\|h - m_0\| = \inf_{m \in \mathcal{M}} \|h - m\|$. Moreover, $h - m_0$ is orthogonal to $\mathcal{M}$ with respect to

https://doi.org/10.1017/S0962492904000212 Published online by Cambridge University Press

$\langle \cdot, \cdot \rangle$. Conversely, if $m_0 \in \mathcal{M}$ is such that $h - m_0$ is orthogonal to $\mathcal{M}$ with respect to $\langle \cdot, \cdot \rangle$, then $\|h - m_0\| = \inf_{m \in \mathcal{M}} \|h - m\|$.

To apply this result in our context, consider a symmetric positive definite matrix $\mathcal{A}$ and the projection process (9.1)–(9.3) with $\mathcal{C}_k = \mathcal{K}_k(\mathcal{A}, r_0)$. According to item (C) in Theorem 9.1, this process leads to a uniquely defined iterate $u_k \in u_0 + \mathcal{K}_k(\mathcal{A}, r_0)$. The residual $r_k = \mathcal{A}e_k$ is orthogonal to $\mathcal{K}_k(\mathcal{A}, r_0)$ with respect to the Euclidean inner product, or, equivalently, the error $e_k = u - u_k = (u - u_0) - (u_k - u_0)$ is orthogonal to $\mathcal{K}_k(\mathcal{A}, r_0)$ with respect to the $\mathcal{A}$-inner product defined by $\langle v, w \rangle_{\mathcal{A}} \equiv w^T \mathcal{A} v$. With $\mathbb{R}^{n+m}$, $\mathcal{K}_k(\mathcal{A}, r_0)$, $u - u_0 \in \mathbb{R}^{n+m}$, and $u_k - u_0 \in \mathcal{K}_k(\mathcal{A}, r_0)$ taking the roles of $\mathcal{H}$, $\mathcal{M}$, h, and m_0 in Theorem 9.2, respectively, we see that the orthogonality condition (9.3) is equivalent to

$$\|e_k\|_{\mathcal{A}} = \min_{z \in u_0 + \mathcal{K}_k(\mathcal{A}, r_0)} \|u - z\|_{\mathcal{A}},$$

where $\| \cdot \|_{\mathcal{A}}$ denotes the $\mathcal{A}$-norm (sometimes called energy norm) associated with the $\mathcal{A}$-inner product. As mentioned above, item (C) represents a mathematical characterization of the CG method; what we have just derived is its well-known optimality property (error minimization in the energy norm). Analogously, we may show that the methods characterized by item (M) in Theorem 9.1 minimize the Euclidean norm of the residual $r_k = b - \mathcal{A}u_k$ over the affine subspace $u_0 + \mathcal{K}_k(\mathcal{A}, r_0)$. These results are summarized in the following theorem.

Theorem 9.3. Suppose that the Krylov subspace $\mathcal{K}_k(\mathcal{A}, r_0)$ in the projection process (9.1)–(9.3) has dimension k. Then the iterate u_k satisfies the following optimality properties.

(C) If $\mathcal{A}$ is symmetric positive definite and $\mathcal{C}_k = \mathcal{K}_k(\mathcal{A}, r_0)$, then

$$\|u - u_k\|_{\mathcal{A}} = \|e_k\|_{\mathcal{A}} = \min_{z \in u_0 + \mathcal{K}_k(\mathcal{A}, r_0)} \|u - z\|_{\mathcal{A}} = \min_{p \in \Pi_k} \|p(\mathcal{A}) e_0\|_{\mathcal{A}}. \quad (9.5)$$

(M) If $\mathcal{A}$ is nonsingular and $\mathcal{C}_k = \mathcal{A}\mathcal{K}_k(\mathcal{A}, r_0)$, then

$$\|b - \mathcal{A}u_k\|_2 = \|r_k\|_2 = \min_{z \in u_0 + \mathcal{K}_k(\mathcal{A}, r_0)} \|b - \mathcal{A}z\|_2 = \min_{p \in \Pi_k} \|p(\mathcal{A}) r_0\|_2. \quad (9.6)$$

Here Π_k denotes the set of polynomials of degree at most k with value 1 at the origin.

For reasons apparent from (9.6) we refer to methods characterized by item (M) as *minimal residual methods*. Note that the Euclidean residual norm in these methods gives a lower bound for the residual norm of all Krylov subspace methods based on (9.1)–(9.3). Therefore significant research efforts have been made to understand the convergence behaviour of

https://doi.org/10.1017/S0962492904000212 Published online by Cambridge University Press

these methods. The interpretation of the kth error and residual in (9.5) and (9.6) in terms of the initial error and residual multiplied by a certain polynomial in the matrix $\mathcal{A}$, respectively, is the typical starting point for the convergence analysis of the Krylov subspace methods characterized by items (C) and (M).

First consider item (C) (as implemented by the CG method). Since $\mathcal{A}$ is assumed to be symmetric positive definite, it is orthogonally diagonalizable, $\mathcal{A} = \mathcal{V}\mathcal{D}\mathcal{V}^T$, with $\mathcal{V}^T\mathcal{V} = I$ and $\mathcal{D} = \mathrm{diag}(\lambda_j)$. The $\mathcal{A}$-norm of the error at step k satisfies

$$\begin{aligned}\|e_k\|_{\mathcal{A}} &= \min_{p\in\Pi_k} \|p(\mathcal{A})e_0\|_{\mathcal{A}} = \min_{p\in\Pi_k} \|\mathcal{A}^{1/2}p(\mathcal{A})e_0\|_2 = \min_{p\in\Pi_k} \|p(\mathcal{A})\mathcal{A}^{1/2}e_0\|_2 \\ &\le \min_{p\in\Pi_k} \|p(\mathcal{A})\|_2 \, \|\mathcal{A}^{1/2}e_0\|_2 \\ &= \|e_0\|_{\mathcal{A}} \min_{p\in\Pi_k} \|p(\mathcal{D})\|_2 \\ &= \|e_0\|_{\mathcal{A}} \min_{p\in\Pi_k} \max_{\lambda_j} |p(\lambda_j)|. \end{aligned} \tag{9.7}$$

Hence $\|e_k\|_{\mathcal{A}}/\|e_0\|_{\mathcal{A}}$, the kth relative $\mathcal{A}$-norm of the error, is bounded by the value of a polynomial approximation problem on the eigenvalues of $\mathcal{A}$. This bound is sharp in the sense that for each (symmetric positive definite) matrix $\mathcal{A}$ and each iteration step k there exists an initial error $e_0^{(k)}$ for which equality holds; see Greenbaum (1979). Consequently, the *worst-case* behaviour of the CG method is completely determined by the eigenvalues of the matrix. An immediate question is how small can a kth degree polynomial with value 1 at the origin become on a given set of matrix eigenvalues. While the polynomial that solves the min-max problem is explicitly known (see Greenbaum (1979) and Liesen and Tichý (2004*b*) for different derivations), no simple expression for the min-max value itself exists. The sharp bound (9.7), however, provides some intuition of how the eigenvalue distribution influences the worst-case convergence behaviour. For example, if all eigenvalues are tightly clustered around a *single* point that is far away from the origin, one may expect fast convergence. Widely spread eigenvalues, on the other hand, will potentially lead to slow convergence. The standard approach for estimating the right-hand side of (9.7) is to replace the min-max problem on the discrete set of eigenvalues by a min-max approximation problem on its convex hull (*i.e.*, on an interval from the smallest eigenvalue $\lambda_{\min}$ to the largest eigenvalue $\lambda_{\max}$ of $\mathcal{A}$). The latter is solved by scaled and shifted Chebyshev polynomials of the first kind, giving the well-known bound

$$\min_{p\in\Pi_k} \max_{\lambda_j} |p(\lambda_j)| \le 2\left(\frac{\sqrt{\kappa(\mathcal{A})}-1}{\sqrt{\kappa(\mathcal{A})}+1}\right)^k, \quad \text{where} \quad \kappa(\mathcal{A}) = \frac{\lambda_{\max}}{\lambda_{\min}}; \tag{9.8}$$

see, *e.g.*, Greenbaum (1997, Theorem 3.1.1). The bounds (9.7)–(9.8) show

https://doi.org/10.1017/S0962492904000212 Published online by Cambridge University Press

that a small condition number of $\mathcal{A}$ is *sufficient* (but not necessary) for a fast decrease of the relative $\mathcal{A}$-norm of the error. This fact motivates the classical goal of *preconditioning*, which is to modify the given linear system in order to reduce the condition number of the system matrix. However, while (9.7) is sharp, the right-hand side of (9.8) often overestimates the left-hand side and thus the worst-case relative $\mathcal{A}$-norm of the error. Moreover, the actual behaviour for a specific right-hand side vector b depends not only on the eigenvalue distribution but also on the coefficients of b in the eigenvectors of $\mathcal{A}$. Several case studies on model problems have recently been performed; see, *e.g.*, Beckermann and Kuijlaars (2001, 2002), Liesen and Tichý (2004*a*) and Naiman, Babuška and Elman (1997) for more details.

Next consider the minimal residual methods. The resulting projection process (9.1)–(9.3) is well defined for each nonsingular matrix $\mathcal{A}$. For simplicity, suppose that $\mathcal{A}$ is diagonalizable, $\mathcal{A} = \mathcal{X}\mathcal{D}\mathcal{X}^{-1}$, with $\mathcal{D} = \mathrm{diag}(\lambda_j)$. Then the kth Euclidean norm of the residual satisfies

$$\begin{aligned} \|r_k\|_2 &= \min_{p\in\Pi_k} \|p(\mathcal{A})r_0\|_2 \\ &\le \min_{p\in\Pi_k} \|\mathcal{X}p(\mathcal{D})\mathcal{X}^{-1}\|_2 \, \|r_0\|_2 \\ &\le \|\mathcal{X}\|_2 \min_{p\in\Pi_k} \|p(\mathcal{D})\|_2 \, \|\mathcal{X}^{-1}\|_2 \, \|r_0\|_2 \\ &= \|r_0\|_2 \, \kappa(\mathcal{X}) \min_{p\in\Pi_k} \max_{\lambda_j} |p(\lambda_j)|. \end{aligned} \tag{9.9}$$

Clearly,

$$\|r_k\|_2/\|r_0\|_2,$$

the kth relative Euclidean residual norm, is bounded by the value of the same type of polynomial approximation problem as in (9.7), multiplied by the condition number of the eigenvector matrix $\mathcal{X}$ of $\mathcal{A}$.

If $\mathcal{A}$ is *normal*, then $\kappa(\mathcal{X}) = 1$, and it can be shown that the bound (9.9) is sharp in the same sense as the bound (9.7); see Greenbaum and Gurvits (1994) and Joubert (1994). In this case the same intuition as described above for the worst-case behaviour of CG also applies to the worst-case behaviour of minimal residual methods. In particular, a single eigenvalue cluster far away from the origin implies fast convergence (here measured by the relative Euclidean residual norm). Additionally, (9.8) can be used if $\mathcal{A}$ is symmetric positive definite, which shows that in this case a small condition number of $\mathcal{A}$ is *sufficient* (but not necessary) for fast convergence. As in the case of CG, reducing the condition number of a symmetric positive definite system matrix $\mathcal{A}$ also represents a reasonable goal of preconditioning for minimal residual methods.

In the case of a nonsingular symmetric indefinite matrix $\mathcal{A}$, the min-max approximation problem on the matrix eigenvalues in (9.9) cannot be replaced by the min-max problem on their convex hull, as eigenvalues lie on

https://doi.org/10.1017/S0962492904000212 Published online by Cambridge University Press

both sides of the origin. Here one may replace the discrete set of eigenvalues by the union of two intervals containing all of them and excluding the origin, say $I^- \cup I^+ \equiv [\lambda_{\min}, \lambda_s] \cup [\lambda_{s+1}, \lambda_{\max}]$ with $\lambda_{\min} \leq \lambda_s < 0 < \lambda_{s+1} \leq \lambda_{\max}$.

When both intervals are of the same length, *i.e.*, $\lambda_{\max} - \lambda_{s+1} = \lambda_s - \lambda_{\min}$, the solution of the corresponding min-max approximation problem

$$\min_{p \in \Pi_k} \max_{\lambda \in I^- \cup I^+} |p(\lambda)|, \tag{9.10}$$

is characterized by a result of de Boor and Rice (1982). This leads to the bound

$$\min_{p \in \Pi_k} \max_{\lambda_j} |p(\lambda_j)| \leq 2 \left(\frac{\sqrt{|\lambda_{\min}\lambda_{\max}|} - \sqrt{|\lambda_s \lambda_{s+1}|}}{\sqrt{|\lambda_{\min}\lambda_{\max}|} + \sqrt{|\lambda_s \lambda_{s+1}|}} \right)^{[k/2]}, \tag{9.11}$$

where $[k/2]$ denotes the integer part of $k/2$; see Greenbaum (1997, Chap. 3). For an illustration of this bound suppose that $|\lambda_{\min}| = \lambda_{\max} = 1$ and $|\lambda_s| = \lambda_{s+1}$. Then $\kappa(\mathcal{A}) = \lambda_{s+1}^{-1}$, and the right-hand side of (9.11) reduces to

$$2 \left(\frac{1/\lambda_{s+1} - 1}{1/\lambda_{s+1} + 1} \right)^{[k/2]}. \tag{9.12}$$

Note that (9.12) corresponds to the value of the right-hand side of (9.8) at step $[k/2]$ for a symmetric positive definite matrix having all its eigenvalues in the interval $[\lambda_{s+1}^2, 1]$, and thus a condition number of λ_{s+1}^{-2}. Hence the convergence bound for an indefinite matrix with condition number κ needs twice as many steps to decrease to the value of the bound for a definite matrix with condition number κ^2. Although neither of the two bounds is sharp, this clearly indicates that solving indefinite problems represents a significant challenge.

In the general case when the two intervals I^- and I^+ are not of the same length, the explicit solution of (9.10) becomes quite complicated (see, *e.g.*, Fischer (1996, Chapter 3)), and no simple and explicit bound on the min-max value is known. One may of course extend the smaller interval to match the length of the larger one, and still apply (9.11). But this usually results in a significantly weaker convergence bound, which fails to give relevant information about the actual convergence behaviour. An alternative is to consider the asymptotic behaviour of the min-max value (9.10), and in particular the *asymptotic convergence factor*

$$\rho(I^- \cup I^+) \equiv \lim_{k \to \infty} \left(\min_{p \in \Pi_k} \max_{\lambda \in I^- \cup I^+} |p(\lambda)| \right)^{1/k}. \tag{9.13}$$

Obviously, $\rho(I^- \cup I^+)$ may be estimated even if the value (9.10) for each step k is unknown. Asymptotic convergence results are common in the theory of semi-iterative methods (Eiermann and Niethammer 1983, Eiermann, Niethammer and Varga 1985) and of classical iterative methods such as

https://doi.org/10.1017/S0962492904000212 Published online by Cambridge University Press

SOR (Varga 1962). Because of the finite termination property, asymptotic convergence results for Krylov subspace methods have to be put into proper perspective. They certainly can be useful in the convergence analysis of minimal residual methods for sequences of linear systems of growing dimension, *e.g.*, when studying the dependence of the convergence behaviour on the mesh size in a discretized differential equation. An example is discussed in Section 10.1 below.

If $\mathcal{A}$ is *nonnormal*, then $\kappa(\mathcal{X})$ may be very large, and (9.9) may be a very large overestimate even of the worst-case relative residual norm of minimal residual methods. In particular, the matrix eigenvalues may in the nonnormal case give misleading information about the convergence behaviour. In fact, it has been shown that any nonincreasing convergence curve of relative residual norms can be obtained by a (nonnormal) matrix $\mathcal{A}$ having any prescribed set of eigenvalues (Greenbaum, Pták and Strakoš 1996). While the examples constructed by this theory may be artificial, misleading eigenvalue information was also demonstrated and analysed in the more practical context of discretized convection-diffusion problems; see Ernst (2000) and Liesen and Strakoš (2005) for further details. Several other sets associated with the matrix $\mathcal{A}$ have been used in the convergence analysis for nonnormal problems, among them the field of values (Ernst 2000, Starke 1997), pseudospectra (Nachtigal, Reddy and Trefethen 1992), and the polynomial numerical hull (Greenbaum 2002). In the context of saddle point problems, the field of values seems to be quite useful. In particular, rate of convergence estimates obtained from this set for preconditioned saddle point systems arising from mixed finite element discretizations of PDEs are sometimes optimal in the sense that they are independent of the mesh size parameter; see, *e.g.*, Klawonn and Starke (1999) and Loghin and Wathen (2004) for details. However, the first inequality in (9.9) has been shown to be strict for some nonnormal matrices $\mathcal{A}$; see Faber, Joubert, Knill and Manteuffel (1996) and Toh (1997). Hence no convergence analysis based solely on the matrix $\mathcal{A}$ can in the nonnormal case give a sharp bound on the worst-case residual norm, so that the initial residual should be included in the convergence analysis whenever possible. This is sometimes only possible in the context of the specific application; see Liesen and Strakoš (2005) for an example. An important and often overlooked fact is that the convergence of minimal residual methods is *not* slower for nonnormal than for normal matrices. In particular, it has been shown that for each nonnormal matrix there exists a normal matrix for which the same convergence behaviour can be observed (for the same initial residual); see Greenbaum and Strakoš (1994). Yet, as described above, the convergence behaviour in the nonnormal case is significantly more difficult to analyse than in the normal case. Sharp convergence results can usually only be obtained by considering the specific properties of $\mathcal{A}$ (*e.g.*, its eigenvalue-eigenvector structure) in relation to the given initial residual.

https://doi.org/10.1017/S0962492904000212 Published online by Cambridge University Press

9.3. Implementation details

When implementing Krylov subspace methods, one needs to generate (at least implicitly) a basis of the Krylov subspace $\mathcal{K}_k(\mathcal{A}, r_0)$. For reasons of numerical stability this basis should preferably be *orthogonal*, and for reasons of computational efficiency, the basis should be generated by a *short-term recurrence*.

In the case of a *symmetric* matrix $\mathcal{A}$, these two goals are achieved simultaneously by the symmetric Lanczos method (Lanczos 1950), which generates an orthogonal basis $v_1, \ldots, v_k$ of $\mathcal{K}_k(\mathcal{A}, r_0)$ using a three-term recurrence, meaning that in step k only the vectors v_{k-1} and v_k are required to compute v_{k+1}. This highly efficient method is the basic ingredient of many Krylov subspace methods for symmetric matrices, among them CG, MINRES and SYMMLQ. Symmetry of the system matrix (even positive definiteness in case of CG) turns out to be a strong restriction in the case of generalized saddle point systems (1.1)–(1.2). First of all, $\mathcal{A}$ needs to be symmetric, which requires $A = A^T$, $C = C^T$, and $B_1 = B_2$. These conditions are satisfied in many applications: see Section 2. However, if the system is preconditioned, the preconditioned system matrix needs to be symmetric (positive definite) as well. This in general requires that the preconditioner $\mathcal{P}$ be symmetric positive definite. For example, consider left preconditioning by $\mathcal{P}$, *i.e.*, the preconditioned system

$$\mathcal{P}^{-1}\,\mathcal{A}\,u = \mathcal{P}^{-1}\,b.$$

If $\mathcal{P}$ is symmetric positive definite, its Cholesky decomposition exists, $\mathcal{P} = \mathcal{L}\,\mathcal{L}^T$, and hence the preconditioned system is equivalent to

$$(\mathcal{L}^{-1}\,\mathcal{A}\,\mathcal{L}^{-T})\,(\mathcal{L}^T u) = \mathcal{L}^{-1}\,b.$$

The matrix $\mathcal{L}^{-1}\mathcal{A}\mathcal{L}^{-T}$ is again symmetric (positive definite), and one can apply the same method to solve the unpreconditioned as well as the preconditioned system. If $\mathcal{P}$ is symmetric but *not* positive definite, then the preconditioned system matrix $\mathcal{P}^{-1}\mathcal{A}$ is in general nonsymmetric, regardless of $\mathcal{A}$ being symmetric or not (unless $\mathcal{P}$ and $\mathcal{A}$ commute, *e.g.*, when $\mathcal{P}$ is a polynomial in $\mathcal{A}$). If no good symmetric positive definite preconditioner is known, or if a very good nonsymmetric preconditioner is available, the possible advantage of $\mathcal{A}$ being symmetric is lost, and one usually must use a solution method for nonsymmetric matrices. See Simoncini (2004*a*) and Section 10 below.

For a general, *nonsymmetric* matrix $\mathcal{A}$, the two goals mentioned above cannot be achieved at once (with few exceptions possessing a particular structure; see Faber and Manteuffel (1984) and Liesen and Saylor (2005)). Here one has to choose between a full recurrence and an orthogonal Krylov subspace basis, or a short-term recurrence and a nonorthogonal basis. The former approach, implemented by the Arnoldi method (Arnoldi 1951), is

https://doi.org/10.1017/S0962492904000212 Published online by Cambridge University Press

Table 9.1. Summary of Krylov subspace methods discussed in Section 9.

Method	Required $\mathcal{A}$	Type	Recurrence	Required $\mathcal{P}$
CG	symm. def.	optimal	three-term	symm. def.
MINRES, SYMMLQ	symm.	optimal	three-term	symm. def.
SQMR	symm.	non-optimal	three-term	symm.
GMRES	general	optimal	full	general
QMR, BiCGStab, TFQMR	general	non-optimal	three-term	general

used in the GMRES algorithm. The latter approach is implemented in the nonsymmetric (or two-sided) Lanczos method (Lanczos 1950), which forms a main ingredient of methods like BiCG, BiCGStab, and QMR. Methods based on orthogonal bases are sometimes called optimal methods, while the other class is referred to as non-optimal methods. The non-optimal methods also include truncated and restarted versions of the optimal methods. Recent research shows that when an optimal method such as GMRES converges quickly, the related non-optimal methods based on (9.1)–(9.2) often also converge quickly (Simoncini and Szyld 2005). The nonsymmetric Lanczos method is also used in SQMR, which represents a non-optimal method for symmetric indefinite matrices. However, unlike the optimal methods for this class of problems discussed above (in particular MINRES and SYMMLQ), the SQMR method does not require a symmetric positive definite preconditioner, since its underlying Lanczos method works for nonsymmetric matrices as well. A summary of the methods discussed in this section is given in Table 9.1.

10. Preconditioners

The use of preconditioning has been already referred to several times in previous sections. As we saw, preconditioning may be used in the iterative solution of reduced systems arising from Schur complement reduction or from the application of various null space methods. In this section we discuss preconditioners in greater detail, and provide a description of some of the most widely used or promising techniques. In view of the fact that preconditioning has been and remains a most active area of research, accounting for the vast majority of papers on the numerical solution of saddle point problems (more generally, on linear solvers) in the last several years, a completely exhaustive survey is impossible. In some cases we limit ourselves to very brief notes and to pointers to the literature.

As is well known, the term *preconditioning* refers to transforming the

https://doi.org/10.1017/S0962492904000212 Published online by Cambridge University Press

linear system $\mathcal{A}u = b$ into another system with more favourable properties for iterative solution. A *preconditioner* is a matrix $\mathcal{P}$ (or $\mathcal{P}^{-1}$) that effects such a transformation. Generally speaking, preconditioning attempts to improve the spectral properties of the system matrix. For symmetric problems, the (worst-case) rate of convergence of Krylov subspace methods like CG or MINRES depends on the distribution of the eigenvalues of $\mathcal{A}$. Ideally, the preconditioned matrix $\mathcal{M} = \mathcal{P}^{-1}\mathcal{A}$ (or $\mathcal{M} = \mathcal{A}\mathcal{P}^{-1}$) will have a smaller spectral condition number, and/or eigenvalues clustered around 1. Another favourable situation is when the preconditioned matrix has a minimum polynomial of small degree.

For nonsymmetric (nonnormal) problems the situation is more complicated, and the eigenvalues may not describe the convergence of nonsymmetric matrix iterations like GMRES; see the discussion in Section 9. Nevertheless, a clustered spectrum (away from 0) often results in rapid convergence, especially if the departure from normality of the preconditioned matrix is not too high.

Generally speaking, there are two approaches to constructing preconditioners. One is based on purely algebraic techniques, like incomplete factorizations, sparse approximate inverses, and algebraic multilevel methods. These preconditioners require little knowledge of the problem at hand besides the entries of $\mathcal{A}$, and can be applied – at least in principle – in a more or less black-box fashion. This type of preconditioning has proved quite effective in the solution of linear systems arising from the discretization of *scalar* partial differential equations of elliptic type, and is widely used in many areas of computational science and engineering; see Benzi (2002), Meurant (1999), Saad (2003), van der Vorst (2003) for recent treatments. When applied to saddle point systems, on the other hand, standard algebraic preconditioners are often found to perform poorly. Because of the indefiniteness and lack of diagonal dominance, these preconditioners are often unstable. Even when the computation of the preconditioner does not suffer from some type of breakdown (*e.g.*, zero pivots in an incomplete factorization), the quality of the resulting preconditioner is often not very satisfactory, and slow convergence is observed. Also, because of the absence of decay in $\mathcal{A}^{-1}$, it is difficult to construct good sparse approximate inverse preconditioners for saddle point matrices.

The second approach develops preconditioners that are tailored to the particular application at hand. This approach requires knowledge of the origin of the problem, including (for PDEs) details about the discretization used, the underlying geometry, properties of the coefficients, and so forth. Of course, the more information one can use, the better the quality of the resulting preconditioner. The drawback of this approach is that the range of problems that can be treated with a particular preconditioner will necessarily be narrow, but this may not be a problem from the user's viewpoint.

https://doi.org/10.1017/S0962492904000212 Published online by Cambridge University Press

For saddle point problems, the construction of high-quality preconditioners necessitates exploiting the block structure of the problem, together with detailed knowledge about the origin and structure of the various blocks. Because the latter varies greatly from application to application, there is no such thing as the 'best' preconditioner for saddle point problems. The choice of a preconditioner is strongly problem-dependent. For instance, techniques that give excellent results for the time-dependent Stokes problem may be completely inadequate for the steady-state case, or with the Oseen equations. Preconditioners that have been successfully used in optimization may be useless in fluid dynamics, and conversely. The good news is that powerful preconditioning techniques have been developed for many problems of practical interest.

We review several such techniques below. We begin with block preconditioners, especially popular in fluid dynamics, and continue with constraint preconditioners, especially popular in optimization. As it turns out, the first class of preconditioners has close ties with Schur complement reduction, while the second is related to the null space approach. Next we describe a more recent, promising class of methods based on the Hermitian and skew-Hermitian splitting of $\hat{\mathcal{A}}$ in (3.10). This approach is rather general and has already been applied to a fairly wide range of problems, from fluid dynamics to weighted least squares. We conclude this section with a brief discussion of recent attempts to develop reliable incomplete factorization techniques for symmetric indefinite systems. It is also worth noting that both the stationary iterative methods described in Section 8 and the multilevel methods described in Section 11 can be accelerated by (used as preconditioners for) Krylov subspace methods.

The survey of preconditioners in this section is by no means exhaustive; see Axelsson and Neytcheva (2003) and Zulehner (2002) for other overviews of preconditioning techniques for saddle point problems.

10.1. Block preconditioners

In this section we consider block diagonal and block triangular preconditioners for Krylov subspace methods applied to the coupled system $\mathcal{A}\,u = b$. As we shall see, the performance of such preconditioners depends on whether fast, approximate solvers for linear systems involving A and the Schur complement S are available. Therefore, the preconditioners considered in this section are related to the 'segregated' solvers considered in Section 5.

10.1.1. Block diagonal preconditioners

Here we consider block diagonal preconditioning for the case of an invertible but possibly nonsymmetric saddle point matrix $\mathcal{A}$ with $C = O$. Then the

https://doi.org/10.1017/S0962492904000212 Published online by Cambridge University Press

basic block diagonal preconditioner is given by

$$\mathcal{P}_d = \begin{bmatrix} A & O \\ O & -S \end{bmatrix}, \tag{10.1}$$

where $S = -B_2 A^{-1} B_1^T$ is the Schur complement. Left preconditioning of $\mathcal{A}$ with $\mathcal{P}_d$ results in the matrix

$$\mathcal{M} = \mathcal{P}_d^{-1}\mathcal{A} = \begin{bmatrix} I & A^{-1}B_1^T \\ -S^{-1}B_2 & O \end{bmatrix}. \tag{10.2}$$

The matrix $\mathcal{M}$ is nonsingular by assumption and, as pointed out in Kuznetsov (1995) and Murphy, Golub and Wathen (2000), it satisfies

$$(\mathcal{M} - I)\left(\mathcal{M} - \frac{1}{2}(1+\sqrt{5})I\right)\left(\mathcal{M} - \frac{1}{2}(1-\sqrt{5})I\right) = O.$$

Hence $\mathcal{M}$ is diagonalizable and has only three distinct eigenvalues, namely $1, \frac{1}{2}(1+\sqrt{5})$, and $\frac{1}{2}(1-\sqrt{5})$ (see de Sturler and Liesen (2005) for the complete form of the eigendecomposition of $\mathcal{M}$). Hence, for each initial residual r_0, $\dim \mathcal{K}_{n+m}(\mathcal{M}, r_0) \leq 3$, which means that GMRES applied to the preconditioned system with matrix $\mathcal{M}$ will terminate after at most 3 steps. The same can be shown for right preconditioning with $\mathcal{P}_d$, or any centred preconditioning of the form $\mathcal{P}_1^{-1}\mathcal{A}\mathcal{P}_2^{-1}$ with $\mathcal{P}_1\mathcal{P}_2 = \mathcal{P}_d$. Furthermore, these results generalize to the case of a nonzero (2,2) block C of the matrix $\mathcal{A}$ (Ipsen 2001).

At first sight this looks promising. However, a simple calculation using the formula (3.4) for the inverse of $\mathcal{A}$ shows that

$$\mathcal{A}^{-1} = \left(\mathcal{M} + \begin{bmatrix} A^{-1}B_1^T S^{-1} B_2 & O \\ O & -I \end{bmatrix}\right)\mathcal{P}_d^{-1}.$$

We see that forming the preconditioned system $\mathcal{M}u = \mathcal{P}_d^{-1}b$ out of the given saddle point system $\mathcal{A}u = b$ using the block diagonal preconditioner (10.1) is essentially as expensive as computing the inverse of $\mathcal{A}$ directly using (3.4). In practice, the exact preconditioner (10.1) needs to be replaced by an approximation,

$$\hat{\mathcal{P}}_d = \begin{bmatrix} \hat{A} & O \\ O & -\hat{S} \end{bmatrix}, \tag{10.3}$$

where both $\hat{A}$ and $\hat{S}$ are approximations of A and S, respectively.

Several different approximations have been considered in the literature. Examples from a few specific applications are given in Section 10.1.3. Here we describe a fairly general framework developed in de Sturler and Liesen (2005). There the preconditioner (10.3) is obtained by considering a splitting of A into

$$A = D - E,$$

https://doi.org/10.1017/S0962492904000212 Published online by Cambridge University Press

where D is invertible. Then $\hat{A} = D$ and $\hat{S} = -B_2 D^{-1} B_1^T$ are chosen, and bounds for the eigenvalues of the resulting preconditioned matrix

$$\hat{\mathcal{M}} = \hat{\mathcal{P}}_d^{-1}\mathcal{A} = \begin{bmatrix} I - D^{-1}E & D^{-1}B_1^T \\ -\hat{S}^{-1}B_2 & O \end{bmatrix} = \begin{bmatrix} I & D^{-1}B_1^T \\ -\hat{S}^{-1}B_2 & O \end{bmatrix} - \begin{bmatrix} D^{-1}E & O \\ O & O \end{bmatrix} \tag{10.4}$$

are given in terms of the eigenvalues of $\mathcal{M} = \mathcal{P}_d^{-1}\mathcal{A}$. These bounds show that the distance of the eigenvalues of $\hat{\mathcal{M}}$ to the three distinct eigenvalues of $\mathcal{M}$ depends on several factors, including the norm of the matrix $D^{-1}E$. In particular, when A is diagonally dominant, and $D = \mathrm{diag}(A)$, then a good clustering of the eigenvalues of $\hat{\mathcal{M}}$ around 1, $\frac{1}{2}(1+\sqrt{5})$, and $\frac{1}{2}(1-\sqrt{5})$ can be expected. However, the preconditioned matrix $\hat{\mathcal{M}}$ still has (now potentially $n+m$ distinct) eigenvalues that may be on both sides of the imaginary axis.

To overcome this drawback, de Sturler and Liesen (2005) propose combining the approximate block diagonal preconditioner (10.3) with an Uzawa-type fixed point iteration (see Section 8.1 above) that is based on a splitting of $\hat{\mathcal{M}}$ as in the rightmost expression in (10.4) into $\hat{\mathcal{M}} = \mathcal{P}_1 - \mathcal{P}_2$. The inverse of $\mathcal{P}_1$ is known explicitly, and if D is simple enough (*e.g.*, diagonal), it can be efficiently computed. Then the preconditioned system $\hat{\mathcal{M}}u = \hat{\mathcal{P}}_d^{-1}b = \hat{b}$ can be written as

$$(\mathcal{P}_1 - \mathcal{P}_2)u = \hat{b} \quad \Leftrightarrow \quad u = \mathcal{P}_1^{-1}\mathcal{P}_2 u + \mathcal{P}_1^{-1}\hat{b},$$

which yields the fixed point iteration

$$u_{k+1} = \mathcal{P}_1^{-1}\mathcal{P}_2 u_k + \mathcal{P}_1^{-1}\hat{b}, \tag{10.5}$$

$$\text{where} \quad \mathcal{P}_1^{-1}\mathcal{P}_2 = \begin{bmatrix} (I + D^{-1}B_1^T\hat{S}^{-1}B_2)D^{-1}E & O \\ -\hat{S}^{-1}B_2 D^{-1}E & O \end{bmatrix}.$$

Because of the form of the iteration matrix $\mathcal{P}_1^{-1}\mathcal{P}_2$, it is easy to see that this fixed point iteration depends only on the first n components of u, *i.e.*, it is sufficient to consider the iteration

$$x_{k+1} = (I + D^{-1}B_1^T\hat{S}^{-1}B_2)D^{-1}E\,x_k + [\mathcal{P}_1^{-1}\hat{b}]_{1:n},$$

where $[\mathcal{P}_1^{-1}\hat{b}]_{1:n}$ denotes the first n components of the vector $\mathcal{P}_1^{-1}\hat{b}$. Hence this approach can be considered a segregated solution method, in which (an approximation to) x_* is computed first, and is then used to compute (an approximation to) y_*. The *reduced system* of order n here is given by

$$\big(I - (I + D^{-1}B_1^T\hat{S}^{-1}B_2)D^{-1}E\big)\,x = [\mathcal{P}_1^{-1}\hat{b}]_{1:n}. \tag{10.6}$$

Obviously, this solution technique reduces the number of unknowns compared to the preconditioned system with the matrix $\hat{\mathcal{M}}$ from $n+m$ to n. This might result in a significant saving of computational resources, especially when m is relatively large. Moreover, and possibly even more importantly,

https://doi.org/10.1017/S0962492904000212 Published online by Cambridge University Press

the eigenvalues of the system matrix in (10.6) tend to be clustered around 1, with tighter clustering corresponding to a smaller norm of the matrix $D^{-1}E$. In particular, the spectrum has only one instead of the three clusters of $\hat{\mathcal{M}}$; see de Sturler and Liesen (2005). Because of this, the performance of a Krylov subspace method such as GMRES applied to (10.6) is typically better than the performance for the block diagonally preconditioned system with the matrix $\hat{\mathcal{M}}$. It is important to note that the approach through (10.6) is closely related to (left) *constraint preconditioning* described below in Section 10.2. For simplicity, consider the case with $B_1 = B_2 = B$, and the constraint preconditioner $\mathcal{P}_c$ in (10.15) with $G = D$ resulting from the splitting $A = D - E$. A straightforward calculation shows that the system matrix in (10.6) is exactly the same as the (1,1) block of the preconditioned matrix $\mathcal{M}$ in (10.17). Also note that the block diagonal preconditioner $\hat{\mathcal{P}}_d$ in this case is equal to the block diagonal matrix in the factorization of the constraint preconditioner $\mathcal{P}_c$ in (10.15). Further details about the relation between these two approaches can be found in de Sturler and Liesen (2005).

Next, consider the case of a symmetric saddle point matrix $\mathcal{A}$ with A positive definite and $C = O$. Then Fischer *et al.* (1998) consider block diagonal preconditioners of the form

$$\hat{\mathcal{P}}_d^{\pm} = \begin{bmatrix} \eta^{-1}A & O \\ O & \pm\hat{S} \end{bmatrix}, \tag{10.7}$$

where $\eta > 0$ is a scaling parameter, and $\hat{S}$ is a symmetric positive definite approximation of the (negative) Schur complement. Note that $\hat{\mathcal{P}}_d^{+}$ is positive definite, while $\hat{\mathcal{P}}_d^{-}$ is indefinite. Since both A and $\hat{S}$ are symmetric positive definite, their Cholesky decompositions $A = LL^T$ and $\hat{S} = MM^T$ can be used to transform the system $\mathcal{A}u = b$ into the equivalent system

$$\begin{bmatrix} \eta I & (M^{-1}BL^{-T})^T \\ \pm M^{-1}BL^{-T} & O \end{bmatrix} \begin{bmatrix} \eta^{-1}L^T x \\ M^T y \end{bmatrix} = \begin{bmatrix} L^{-1}f \\ (\eta M)^{-1}g \end{bmatrix}. \tag{10.8}$$

Clearly, the system matrix in (10.8) is of the same form as $\mathcal{A}_\eta^{\pm}$ in (3.13), and hence its eigenvalues depending on the choice of $+$ or $-$ and of η are characterized in Theorem 3.8. There it is shown that the eigenvalue distributions vary greatly from each other. In addition, the positive definite preconditioner $\hat{\mathcal{P}}_d^{+}$ yields a symmetric system matrix $\mathcal{A}_\eta^{+}$ in (10.8), while the indefinite preconditioner $\hat{\mathcal{P}}_d^{-}$ leads to the nonsymmetric matrix $\mathcal{A}_\eta^{-}$. For a fixed η it is therefore interesting to understand which choice yields the better performance of a Krylov subspace method applied to the preconditioned system (10.8). Curiously, there may be no difference at all. In fact, assuming that $g = 0$ in the saddle point system, Fischer *et al.* (1998) show that the preconditioned matrices $\mathcal{A}_\eta^{\pm}$ in (10.8) both generate the same Krylov subspace iterates u_k when special initial residuals are used. In this case the

https://doi.org/10.1017/S0962492904000212 Published online by Cambridge University Press

residual norms of MINRES applied to the symmetric system coincide with the residual norms of GMRES applied to the nonsymmetric system in every step of the iteration. This exact equivalence is lost when random initial residuals are used, but as is shown numerically in Fischer *et al.* (1998), the convergence curves are very close in such cases as well. Since the MINRES method is based on three-term recurrences due to symmetry of the system matrix (*cf.* Section 9), the positive definite preconditioner clearly represents the superior strategy.

Block diagonal preconditioners for saddle point problems arising from (stabilized) discretized Stokes problems (2.10)–(2.12) have been studied, for example, in Silvester and Wathen (1994) and Wathen and Silvester (1993). The system matrix of the discrete problem is a generalized saddle point matrix of the form

$$\mathcal{A} = \begin{bmatrix} A & B^T \\ B & -\beta C \end{bmatrix}, \tag{10.9}$$

where A is block diagonal and symmetric positive definite, C is symmetric positive semidefinite, and $\beta > 0$ is a stabilization parameter. Silvester and Wathen (Silvester and Wathen 1994, Wathen and Silvester 1993) provide several results on the eigenvalue distribution of the preconditioned matrix $\hat{\mathcal{P}}_d^{-1}\mathcal{A}$ for different symmetric positive definite preconditioners $\hat{\mathcal{P}}_d$ of the form (10.3).

For example, consider the case $\hat{A} = \mathrm{diag}(A)$, and $-\hat{S} = \beta\,\mathrm{diag}(C)$ if $C \neq O$ or $-\hat{S} = \beta h^d I$ if $C = O$ (here d is the spatial dimension of the Stokes problem, and h is the mesh size parameter), *i.e.*, a positive definite diagonal preconditioner $\hat{\mathcal{P}}_d$. Then it can be shown that the eigenvalues of the symmetric indefinite preconditioned matrix $\hat{\mathcal{P}}_d^{-1}\mathcal{A}$ are contained in the union of two real intervals of the form

$$(-a, -bh) \cup (ch^2, d), \tag{10.10}$$

where a, b, c, d are positive constants that are independent of h (Wathen and Silvester 1993, Theorems 1 and 2). As discussed in Section 9, the (worst-case) convergence of MINRES applied to the preconditioned system depends on the eigenvalue distribution of $\hat{\mathcal{P}}_d^{-1}\mathcal{A}$, and an upper convergence bound can be found by using (the smallest) two intervals containing the eigenvalues. If these intervals are of the same length, then such a convergence bound is given by (9.9)–(9.11). But in the case of (10.10), the two intervals are in fact not of the same length, since the negative eigenvalues spread out less rapidly under mesh refinement than the positive eigenvalues. Extending the two intervals in (10.10) to make their lengths equal, say by replacing (10.10) with $(-d, -ch^2) \cup (ch^2, d)$ (where we assume, for simplicity, $d > a$ and $b > c$), would predict an asymptotic convergence rate of $O(1 - h^2 c/d)$ for MINRES. However, with some careful analysis it can be

https://doi.org/10.1017/S0962492904000212 Published online by Cambridge University Press

shown that the asymptotic convergence factor of (10.10) is actually of order $O(1 - h^{3/2}\sqrt{bc/ad})$; see Wathen *et al.* (1995) for details. More sophisticated positive definite block diagonal preconditioners in this context leading to even better asymptotic convergence rates are studied in Silvester and Wathen (1994); see also Section 10.1.3.

Apart from the work in Silvester and Wathen (1994) and Wathen and Silvester (1993), block diagonal preconditioners have been studied in the context of numerous additional applications. The upshot is that for suitably chosen approximations to A and S, block diagonal preconditioners resulting in mesh-independent rates of convergence exist for mixed finite element formulations of many problems. For example, an analysis of the eigenvalue distribution of block diagonally preconditioned saddle point systems arising from mixed finite element discretizations of magnetostatics problems is given in Perugia and Simoncini (2000, Section 4). Other analyses of block diagonal preconditioners in specific applications can be found in Battermann and Heinkenschloss (1998), Chizhonkov (2001, 2002), Klawonn (1998*a*), Krzyżanowski (2001), Kuznetsov (1995, 2004), Lukšan and Vlček (1998), Lyche *et al.* (2002), Mardal and Winther (2004), Pavarino (1997, 1998), Peters *et al.* (2004), Powell and Silvester (2004), Toh, Phoon and Chan (2004) and Vassilevski and Lazarov (1996).

10.1.2. Block triangular preconditioners

Block triangular preconditioners of the form

$$\mathcal{P}_t = \begin{bmatrix} A & B^T \\ O & S \end{bmatrix}$$

were first considered in Bramble and Pasciak (1988) and extensively developed in the last few years. Note that Uzawa's method may also be regarded as a block (lower) triangular preconditioner; *cf.* (8.2). This class of preconditioners includes some of the most effective solvers for saddle point problems, both symmetric and nonsymmetric. Note that the first block row of $\mathcal{P}_t$ coincides with the first block row of $\mathcal{A}$; hence, for the initial guess $u_0 = 0$, the solution of the initial preconditioning equation $\mathcal{P}_t z_0 = r_0 = b$ must in particular satisfy the first of the two equations in the saddle point system (1.6). In Section 10.2 we will examine preconditioning strategies that require satisfying the *second* equation in the saddle point system (constraint preconditioning). In the setting of fluid flow problems, the first approach requires the preconditioner to respect the conservation of momentum, whereas the second one imposes a mass balance condition.

Recalling the block factorization (3.3), we immediately see that the spectrum of $\mathcal{P}_t^{-1}\mathcal{A}$ is $\sigma(\mathcal{P}_t^{-1}\mathcal{A}) = \{1\}$ and moreover, the preconditioned matrix has minimum polynomial of degree 2, so that a method like GMRES would converge in at most two steps. In practice, of course, approximations $\hat{A}$ and

https://doi.org/10.1017/S0962492904000212 Published online by Cambridge University Press

$\hat{S}$ to A and S have to be used. The computation of $z_k = \mathcal{P}_t^{-1} r_k$ at each step of a nonsymmetric Krylov subspace method can be implemented on the basis of the factorization

$$\mathcal{P}_t^{-1} = \begin{bmatrix} \hat{A}^{-1} & O \\ O & I \end{bmatrix} \begin{bmatrix} I & B^T \\ O & -I \end{bmatrix} \begin{bmatrix} I & O \\ O & -\hat{S}^{-1} \end{bmatrix}. \tag{10.11}$$

Note that the cost of applying the preconditioner is only slightly higher than in the block diagonal case: besides the solves with $\hat{A}$ and $\hat{S}$, there is an additional multiplication by B^T. Once again, the choice of the approximations $\hat{A}$ and $\hat{S}$ is problem-dependent; see Section 10.1.3. Generally speaking, the better the approximations, the faster the convergence. However, since the preconditioned matrix $\mathcal{P}_t^{-1}\mathcal{A}$ (or $\mathcal{A}\mathcal{P}_t^{-1}$) is nonnormal, convergence estimates are not easy to obtain. Field of values analysis can be used in some instances to obtain eigenvalue bounds and convergence estimates for GMRES. These analyses and numerical experiments indicate that h-independent convergence rates can be achieved for some important problems, including the Oseen equations; see Klawonn and Starke (1999), Loghin and Wathen (2004) and the discussion in the next section. Other analyses of block triangular preconditioners for specific applications can be found in Battermann and Heinkenschloss (1998), Cao (2004*a*), Elman, Silvester and Wathen (2002*a*), Kanschat (2003), Klawonn (1998*b*), Krzyżanowski (2001) and Pavarino (1998); see also Bramble and Pasciak (1988), Simoncini (2004*a*) and Zulehner (2002) for analyses of the inexact case. An apparent disadvantage in the symmetric case is that block triangular preconditioning destroys symmetry. A symmetrized version of block triangular preconditioning has been proposed in Bramble and Pasciak (1988); see also Axelsson and Neytcheva (2003), D'Yakonov (1987) and Zulehner (2002). However, symmetrization is seldom necessary in practice: if good approximations to A and S are available, a method like GMRES with block triangular preconditioning will converge quickly and the overhead incurred from the use of a nonsymmetric solver will be negligible.

We also mention preconditioners based on incomplete block triangular factorizations of the form

$$\mathcal{P} = \begin{bmatrix} \hat{A} & O \\ B & \hat{S} \end{bmatrix} \begin{bmatrix} I & \hat{A}^{-1}B^T \\ O & I \end{bmatrix} \approx \mathcal{A},$$

where $\hat{A}$ and $\hat{S}$ are easily invertible approximations of A and the Schur complement S. Note that application of this preconditioner requires two solves with $\hat{A}$ rather than one, in addition to the solve with $\hat{S}$. Multiplying out the factors, we obtain

$$\mathcal{P} = \begin{bmatrix} \hat{A} & B^T \\ B & -\hat{C} \end{bmatrix}, \qquad \hat{C} = -(B\hat{A}^{-1}B^T + \hat{S}).$$

https://doi.org/10.1017/S0962492904000212 Published online by Cambridge University Press

Hence, generally speaking, $\mathcal{P}$ is an indefinite preconditioner of the type considered in Section 10.2. Note that $\mathcal{P}$ is symmetric if $\hat{A}$ and $\hat{S}$ are. Again, the key issue here is the choice of the approximations $\hat{A} \approx A$ and $\hat{S} \approx S$. Such incomplete block LU factorization preconditioners correspond to classical solution methods known in computational fluid dynamics as SIMPLE schemes (for 'Semi-Implicit Method for Pressure-Linked Equations'): see Patankar (1980) and Patankar and Spalding (1972). In the original SIMPLE scheme $\hat{A} = D_A$ and $\hat{S} = -BD_A^{-1}B^T$, respectively, where D_A denotes the main diagonal of A. Different choices of the approximations involved lead to different preconditioners. Variants of SIMPLE are also often used as smoothers for multigrid; see the discussion in Section 11. A spectral analysis of the SIMPLE preconditioner applied to the Oseen problem can be found in Li and Vuik (2004). See also Benzi and Liu (2005) for numerical experiments and comparisons with other preconditioners. These experiments confirm the intuition that SIMPLE preconditioning is effective when A is strongly diagonally dominant, as is the case for generalized Stokes and Oseen problems with sufficiently small time steps. On the other hand, SIMPLE is not competitive when the time steps are large or 'infinite' (steady case), for in this case the approximations to A and especially to S are poor. It should be mentioned that for unsteady problems classical (Chorin-style (Chorin 1968)) pressure-correction schemes are also very efficient; see Wesseling (2001, page 303). Further references on incomplete block factorizations for the treatment of time-dependent Navier–Stokes equations include Gauthier, Saleri and Quarteroni (2004), Perot (1993), Quarteroni, Saleri and Veneziani (2000), Saleri and Veneziani (2005) and Veneziani (2003).

10.1.3. Approximating A and S

As we have seen, many algorithms for solving saddle point systems depend on the availability of good approximations for the (1,1) block A and for the Schur complement S. Such approximations are needed in segregated approaches and in preconditioned Uzawa schemes where the Schur complement system (5.1) is solved by a preconditioned iteration, and in the construction of block diagonal and block triangular preconditioners for coupled Krylov iterations applied to $\mathcal{A}u = b$.

The construction of such approximations is a strongly problem-dependent process, and a large body of literature exists on this topic. The problem has been especially well-studied for mixed finite element formulations of elliptic PDEs, for the Stokes and Oseen problems, and for linear elasticity problems. Furthermore, some work has been done towards the construction of preconditioners for the Schur complement ('normal equations') systems arising from interior point methods in optimization.

https://doi.org/10.1017/S0962492904000212 Published online by Cambridge University Press

In the case of mixed finite element approximations of elliptic boundary value problems and of Stokes- and Oseen-type equations, the submatrix A usually corresponds either to a zeroth-order differential operator (multiplication by a function or tensor K, as in (2.18)), or to a direct sum of second-order differential operators of diffusion or convection-diffusion type. When A represents multiplication by a function (a mass matrix), it can be well approximated by either a (scaled) identity matrix or by diagonal lumping of A (that is, the ith diagonal entry of $\hat{A}$ is equal to the sum of the entries in the ith row of A). If the coefficients in K are sufficiently smooth, the resulting approximation $\hat{A}$ is then spectrally equivalent to A, in the sense that the eigenvalues of $\hat{A}^{-1}A$ are positive and contained in an interval whose endpoints do not depend on the mesh size; see Rusten and Winther (1992, 1993) and Wathen (1987). When A represents a discrete (vector) diffusion or convection-diffusion operator, a spectrally equivalent approximation $\hat{A}$ can be obtained in many cases by performing one or more multigrid sweeps on the linear system $Ax = v$. Thus, the approximation $\hat{A}$ of A is not constructed explicitly, but is defined implicitly by the action $\hat{A}^{-1}v$ on a given vector v. We refer the reader to Elman *et al.* (2005*c*) for details. Here we note that for the generalized (unsteady) Stokes and Oseen problems the matrix A is usually well-conditioned because of the presence of an additional term inversely proportional to the time step Δt. This makes the (1,1) block strongly diagonally dominant, and a single multigrid V-cycle with an appropriate Gauss–Seidel smoother is often sufficient: see Mardal and Winther (2004) and Turek (1999).

When A does not arise from the discretization of a differential operator, approximate solves, if necessary, may be obtained by either incomplete factorizations or by a few iterations of a Krylov subspace method, possibly with an appropriate preconditioner.

The approximation $\hat{S}$ of the Schur complement matrix S is usually more delicate. In the context of saddle point systems arising from the discretization of PDEs, good approximations to the Schur complement are critical to the performance of block preconditioners. In many cases, excellent approximations can be constructed by thinking in terms of the underlying differential operators. We recommend Loghin and Wathen (2003) for a highly readable discussion of this problem in terms of pseudodifferential operators.

For mixed finite element discretizations of elliptic problems of the type (2.18)–(2.20), the Schur complement $S = -BA^{-1}B^T$ can be interpreted as a discretization of the second-order diffusion operator $\nabla K^{-1} \cdot \nabla$ acting on the unknown function corresponding to the y variable (*e.g.*, on the pressure for incompressible flow problems). In this case, a number of approximations $\hat{S}$ are possible. Provided that K is sufficiently well-behaved, the action of S^{-1} can be approximated by an iteration of multigrid. For higher-order methods, a low-order discretization of the same diffusion operator may also

https://doi.org/10.1017/S0962492904000212 Published online by Cambridge University Press

be expected to give a very good approximation. The literature in this area is vast; see, *e.g.*, Kanschat (2003), Kuznetsov (1995, 2004), Loghin and Wathen (2003), Martikainen (2003), Moulton, Morel and Ascher (1998), Perugia and Simoncini (2000), Powell and Silvester (2004), Rusten, Vassilevski and Winther (1996), Rusten and Winther (1992, 1993) and Warsa, Benzi, Wareing and Morel (2004).

For LBB-stable discretizations of the linear elasticity and steady-state Stokes problem, the Schur complement is spectrally equivalent to a mass matrix (Verfürth 1984*a*). This is not surprising if one thinks of S as a discrete counterpart of a pseudodifferential operator of the form $\operatorname{div}\Delta^{-1}\operatorname{grad}$ and keeping in mind the identity $\Delta = \operatorname{div}\operatorname{grad}$. Hence, under appropriate discretization the Schur complement has condition number bounded independently of mesh size and can be approximated effectively and cheaply in a number of ways; see the previous discussion on approximating zeroth-order differential operators, as well as the references Elman (1996), Elman *et al.* (2005*c*), Klawonn (1998*a*, 1998*b*), Loghin and Wathen (2003, 2004), Pavarino (1997), Silvester and Wathen (1994), Wathen and Silvester (1993) and Yang (2002).

For mixed finite element discretizations of the generalized Stokes problem that arises from the implicit treatment of the time-dependent Stokes problem, on the other hand, the Schur complement is of the form $S = -B(A + \beta I)^{-1}B^T$, where $\beta > 0$ is inversely proportional to the time step. In the 2D case, A is the block diagonal matrix

$$A = \begin{bmatrix} \nu H & O \\ O & \nu H \end{bmatrix},$$

where H is a discrete Laplace operator and $\nu > 0$ denotes the viscosity. In the 3D case, there will be three blocks along the diagonal instead of two. Clearly, for very large time steps (β small) the matrix $S = -B(A + \beta I)^{-1}B^T$ is close to the Schur complement of the steady Stokes problem and is well-conditioned independent of mesh size. In this case a scaled pressure mass matrix will be a good approximation. On the other hand, for small step sizes (β large) this operator has a condition number that grows as $h \to 0$. Note that in the limit as $\beta \to \infty$, the Schur complement behaves essentially as a discrete diffusion operator. The same is true for the case of small viscosity ν (high Reynolds numbers Re), regardless of the size of the time step, for in this case the matrix A, which contains terms proportional to $\nu = \mathrm{Re}^{-1}$, will be negligible. In these cases the SIMPLE approximation for the Schur complement can be quite good for sufficiently small β. Robust preconditioners that work well for small as well as large time steps and over a wide range of Reynolds numbers were first proposed and analysed in Cahouet and Chabard (1988); see also Bramble and Pasciak (1997) and Kobelkov and Olshanskii (2000). The idea is to

https://doi.org/10.1017/S0962492904000212 Published online by Cambridge University Press

build preconditioners that are discrete counterparts of the pseudodifferential operator L implicitly defined by

$$L^{-1} = \nu I - \beta \Delta^{-1},$$

where I denotes the identity operator and Δ is a pressure Laplacian. Here Δ^{-1} stands for the data-to-solution mapping for a pressure-Poisson equation with suitable boundary conditions. Thus, the action of the approximate Schur complement $\hat{S}$ on a vector v consists of

$$\hat{S}^{-1} v = \nu M_p v - \beta (B M_u^{-1} B^T)^{-1} v,$$

where M_p is a pressure mass matrix and M_u a velocity mass matrix. Note that $BM_u^{-1}B^T$ is simply the so-called *compatible discretization* of the Laplacian (Cahouet and Chabard 1988). Again, inexact solves can be used to approximate $(BM_u^{-1}B^T)^{-1}v$. For further details see Turek (1999) and the recent paper by Mardal and Winther (2004).

The situation is more complicated for the nonsymmetric saddle point problems arising from discretizations of the (steady and unsteady) Oseen equations. For steady problems with large viscosity (small Reynolds numbers), using a pressure mass matrix to approximate the Schur complement is usually sufficient and results in rates of convergence independent of mesh size when block diagonal or block triangular preconditioners are used. However, the rate of convergence tends to deteriorate quickly as the viscosity gets smaller. The problem is that as $\nu \to 0$, the exact Schur complement becomes increasingly nonsymmetric and cannot be well approximated by a (necessarily symmetric) mass matrix; see, *e.g.*, Cao (2004*a*), Elman and Silvester (1996) and Krzyżanowski (2001). Much progress has been made in the last few years towards developing robust preconditioners of the block triangular type for the Oseen equations; see Elman (1999, 2002), Elman, Howle, Shadid, Shuttleworth and Tuminaro (2005*a*), Elman, Howle, Shadid and Tuminaro (2003*a*), Elman, Silvester and Wathen (2002*a*, 2002*b*, 2005*c*), Kay, Loghin and Wathen (2002), Loghin and Wathen (2002) and Silvester, Elman, Kay and Wathen (2001). Out of this effort, two especially interesting techniques have emerged. The first idea is to approximate the inverse of the Schur complement matrix $S = BA^{-1}B^T$ with

$$\hat{S}^{-1} = (BB^T)^{-1} BAB^T (BB^T)^{-1} \tag{10.12}$$

(here and in the remainder of this section we dispense with the negative sign for ease of notation). This approximation, originally proposed by Elman (1999) (see also Saleri and Veneziani (2005)), has been given various justifications. A simple one is to observe that if B were square and invertible, then the inverse of $BA^{-1}B^T$ would be $B^{-T}AB^{-1}$. However, B is rectangular. Therefore it makes sense to replace the inverse of B with the Moore–Penrose pseudo-inverse $B^\dagger$; see Campbell and Meyer Jr. (1979).

https://doi.org/10.1017/S0962492904000212 Published online by Cambridge University Press

If we assume B to have full row rank, the pseudo-inverse of B is given by $B^\dagger = B^T(BB^T)^{-1}$, thus justifying the approximation (10.12). Now the approximate Schur complement is (strongly) nonsymmetric if A is. Note that evaluating the action of $\hat{S}^{-1}$ on a given vector v requires the solution of two pressure-Poisson-type problems, plus matrix-vector multiplies involving B^T, A and B. It follows from (10.11) that the block triangular preconditioner can be applied by performing matrix-vector products involving B^T, A and B, and a few iterations of multigrid. For 3D problems, multigrid is needed to approximately solve three convection-diffusion problems for the three components of the velocity field (approximate inversion of A), and to solve the two Poisson-type problems approximately on the pressure space. Numerical experiments for simple model problems indicate that the performance of this preconditioner depends only mildly on the mesh size and viscosity coefficient; see Elman (1999). However, it was found to perform poorly (occasionally failing to converge) on some difficult problems; see Vainikko and Graham (2004). As noted in Elman, Howle, Shadid, Shuttleworth and Tuminaro (2005*a*), for finite element problems the performance of this preconditioner can be greatly improved by the use of appropriate scalings. Denoting by M_1 and M_2 suitable diagonal matrices, Elman *et al.* (2005*a*) derive the more accurate approximation

$$\hat{S}^{-1} = (BM_2^{-2}B^T)^{-1}BM_2^{-2}AM_1^{-1}B^T(BM_1^{-2}B^T)^{-1}. \tag{10.13}$$

The matrix M_1 is taken to be the diagonal of the velocity mass matrix M_u; taking $M_2 = M_1^{1/2}$ makes the two variable-coefficient discrete diffusion operators in (10.13) identical. Although the cost of the two variants (10.12) and (10.13) is almost the same, the improvement in performance obtained by the use of scaling can be quite dramatic; see the numerical experiments in Elman *et al.* (2005*a*).

Another effective approximation of the Schur complement for the Oseen problem was introduced in Kay, Loghin and Wathen (2002) and analysed in Elman, Silvester and Wathen (2002*b*). Again, several justifications have been given for this approach, the original one being based on the Green's tensor of the Oseen operator on the continuous level. Another elegant derivation in terms of pseudodifferential operators can be found in Loghin and Wathen (2003). A simpler, heuristic argument is to first consider the case where A and B^T are both square and commute: $AB^T = B^TA$. If A is nonsingular, its inverse must also commute with B^T, hence $A^{-1}B^T = B^TA^{-1}$. If we think of A as representing a second-order differential operator and B a first-order one then, apart from possible problems near the boundary, assuming commutativity is reasonable. It follows from these assumptions that $S^{-1} = (BA^{-1}B^T)^{-1} = A(BB^T)^{-1}$. In practice, however, A and B^T have different dimensions and represent

https://doi.org/10.1017/S0962492904000212 Published online by Cambridge University Press

operators acting on different function spaces. Note that BB^T is, again, a (scalar) pressure-Poisson-type operator; on the other hand, A is a (vector) convection-diffusion operator acting on the velocities. This suggests introducing a discrete convection-diffusion operator A_p acting on the pressures. Including the necessary mass matrix scalings, the resulting approximation to the inverse of the Schur complement is thus

$$\hat{S}^{-1} = M_p^{-1} A_p (B M_u^{-1} B^T)^{-1}. \tag{10.14}$$

The block triangular preconditioner based on this approximate Schur complement has been extensively tested on a variety of steady and unsteady model problems, in both 2D and 3D; see, *e.g.*, Kay, Loghin and Wathen (2002), Elman, Howle, Shadid and Tuminaro (2003*a*), Elman, Silvester and Wathen (2002*a*, 2002*b*), Hemmingsson-Frändén and Wathen (2001) and Loghin and Wathen (2002). This method, which is referred to as the *pressure convection-diffusion* preconditioner in Elman *et al.* (2005*c*), appears to be very robust with respect to mesh refinement, and fairly robust with respect to the viscosity, exhibiting only a mild deterioration as $\nu \to 0$ on a wide range of problems and discretizations. Note that this approximation is less expensive than the one based on (10.13), since computing the action of $\hat{S}^{-1}$ only requires solving a single pressure-Poisson problem and a matrix-vector product with A_p.

One drawback of this approach is the need to construct an additional operator, A_p, which is not required by the formulation of the Oseen problem but is needed only to form the preconditioner. This discrete convection-diffusion operator is not usually available in standard fluid dynamics codes. This has motivated new, fully algebraic approaches to compute approximations to A_p from the existing blocks A, B and B^T based on the notion of *sparse approximate commutators* (SPAC); see Elman *et al.* (2005*a*). The approximate A_p is the matrix that minimizes the cost functional

$$F(X) = \|AB^T - B^T X\|_F$$

(where $\|\cdot\|_F$ denotes the Frobenius norm) over a space of matrices with prescribed sparsity pattern. Note that the solution of the *unconstrained* minimization problem is $A_p = (BB^T)^{-1} BAB^T$, corresponding to Elman's approximation $\hat{S}^{-1} = (BB^T)^{-1} BAB^T (BB^T)^{-1}$.

The block triangular preconditioner based on the pressure convection-diffusion operator was also used by Syamsudhuha and Silvester (2003) with good results as a smoother for multigrid applied to the Oseen problem. It was also used by Elman, Loghin and Wathen (2003*b*) as a preconditioner for the linear systems arising from Newton's method applied to the Navier–Stokes equations. In this case the method was again found to result in h-independent convergence, but some growth in the iteration number was observed for small ν.

https://doi.org/10.1017/S0962492904000212 Published online by Cambridge University Press

The pressure convection-diffusion operator, on the other hand, does not seem to be applicable to the rotation form (2.15)–(2.17) of the Oseen equations. Different approximations need to be used in this case; see Olshanskii (1999) and Section 10.3 below.

Approximations for the Schur complement matrices arising in interior point methods for constrained optimization have been studied by several authors. Here we give a brief discussion based on Frangioni and Gentile (2004). As already pointed out in Section 3.5, such linear systems can be extremely ill-conditioned. An interesting technique is the *tree preconditioner* described in Resende and Veiga (1993), based on ideas first put forth in Vaidya (1990). The approximation to $S = BA^{-1}B^T$ is of the type $\hat{S} = \hat{B}M\hat{B}^T$ where $\hat{B}$ and M are now both square matrices of order m. The actual construction of these matrices depends on the application. In network flow problems (the motivating application in Resende and Veiga (1993)), the constraint matrix B is the node–edge incidence matrix of a directed graph G, and $\hat{B}$ is defined as the node–edge incidence matrix of a spanning tree of G. Furthermore, A is diagonal and M is obtained by restricting A^{-1} to the arcs that comprise the tree. (Note that when $A = I$, the matrix $S = BB^T$ is closely related to the Laplacian matrix of the undirected version of G.) More precisely, $\hat{B}$ is chosen as the node–edge incidence matrix of an (approximate) maximum weight spanning tree, where the weight of each edge (i, j) is given by the corresponding entry of M. Such spanning trees can be constructed in $O(m)$ work using a variant of the classical Kruskal algorithm. The resulting linear systems involving $\hat{S}$ can be solved in $O(n)$ work by solving linear systems with matrices $\hat{B}^T$, M and $\hat{B}$. These linear systems can be solved by visiting the spanning tree, without fill-in. This preconditioner can be expected to be especially effective in the final stages of the interior point iteration, since in this case the weights m_{ij} tend to zero on all arcs, except on those corresponding to the basic optimal solution that form a spanning tree. The preconditioner is less effective in the initial stages of the interior point algorithm. This observation has motivated the use of hybrid strategies where a diagonal preconditioner is used initially and the tree preconditioner in later stages of the algorithm: see Resende and Veiga (1993) for details. Improvements of the tree preconditioner were proposed in Mehrotra and Wang (1996). A different strategy, based on an incomplete QR factorization, was proposed in Júdice, Patricio, Portugal, Resende and Veiga (2003), and found to be superior to the tree preconditioner for the special case of transportation problems.

New preconditioners based on combinatorial principles have recently been introduced by Frangioni and Gentile (2004). Motivated by the tree preconditioner idea, these authors construct approximations of S of the form $\hat{S} = \hat{B}M\hat{B}^T$, where now $\hat{B}$ corresponds to a *subgraph* of the graph G of B; this subgraph, which may contain (strictly) a spanning tree of G, should be

https://doi.org/10.1017/S0962492904000212 Published online by Cambridge University Press

such that linear systems associated with the matrix $\hat{S}$ can be easily solved. For example, if the chosen subgraph happens to be *chordal*, then there is an ordering of $\hat{S}$ such that no fill-in occurs when factoring $\hat{S}$; see Rose (1970). Fast algorithms exist to compute approximate maximum weight chordal subgraphs. We refer the reader to Frangioni and Gentile (2004) for further details and for numerical experiments showing that in many cases the new subgraph-based preconditioners improve on existing techniques in terms both of iteration counts and CPU time. An interesting open question is to what extent these ideas can be exploited to construct approximate Schur complements for saddle point problems arising from mixed finite element discretizations of PDE problems, where the matrix B is a discrete divergence operator.

We conclude this section with a brief discussion of purely algebraic methods that have been used to construct approximations $\hat{S}$ or $\hat{S}^{-1}$ to the Schur complement matrix or to its inverse. See also the general treatment in Axelsson (1994, Chapter 9) and the discussion in Siefert and de Sturler (2004, Section 4). For simplicity, we confine ourselves to the case where $B_1 = B_2$ and $C = O$. Extension to cases with $B_1 \neq B_2$ is usually straightforward; the case $C \neq O$, on the other hand, may be less obvious, especially when seeking explicit approximations to S^{-1}.

The simplest approximation is perhaps to set $\hat{S} = BD_A^{-1}B^T$ where D_A is either diagonal or possibly block diagonal with blocks of small size. Note that in this case $\hat{S}$ will be sparse if B is, except for special situations (*e.g.*, when B contains one or more dense columns). The diagonal approximation $D_A = \mathrm{diag}(A)$ is used, for instance, in the already-mentioned SIMPLE-type schemes popular in computational fluid dynamics. Better approximations can be obtained by means of incomplete factorizations of the form $A \approx LU \equiv \hat{A}$. Then $\hat{S} = XY$ where $X = BU^{-1}$ and $Y = L^{-1}B^T$. Matrices X and Y may be computed by solving two lower triangular linear systems involving U^T and L and with multiple right-hand sides given by the columns of B^T. *A posteriori* dropping of small entries may be necessary to keep X and Y sparse. Alternatively, one can use (explicit) sparse approximate inverse techniques directly applied to A. For example, the approximate inverse M_A may be computed as the minimizer of the matrix functional

$$F(X) = \|I - AX\|_F$$

over a prescribed set of sparse matrices. Many other approximate inverse techniques have been developed; see Benzi (2002) for an overview. Whatever the technique employed, the resulting sparse approximate inverse $M_A \approx A^{-1}$ can then be used to form an approximation $\hat{S} = BM_AB^T$. See Chow and Saad (1997*a*), Little and Saad (2003) and Little, Saad and Smoch (2003) for implementation details and for experimental comparisons of various strategies. These techniques may be expected to work well when the entries

https://doi.org/10.1017/S0962492904000212 Published online by Cambridge University Press

in A^{-1} decay rapidly away from the main diagonal. This can be expected to hold, for instance, for implicit time discretizations of time-dependent Stokes and Oseen problems. Provided that sufficiently small time steps are taken, the diagonal entries of A will be large compared to the off-diagonals and will result in fast decay in A^{-1}; see Demko, Moss and Smith (1984).

Other approximations that have been used in the literature for specific applications can be used, at least in principle, in more general settings. Juvigny (1997) proposed the simple approximation $(BA^{-1}B^T)^{-1} \approx BAB^T$ as an approximate inverse of the Schur complement for saddle point matrices arising from the use of Lagrange multipliers in domain decomposition schemes. In this case B is often a very simple matrix representing a restriction operator, and the approximation is not unreasonable. Although this approximation is not very accurate, it has the advantage of having zero construction cost; furthermore, computing the action of $\hat{S}^{-1}$ on a vector only requires matrix vector products with sparse matrices. See also Lyche *et al.* (2002) and Tyrtyshnikov and Vassilevski (2003) for more sophisticated approximations to the Schur complements arising in similar contexts.

Another possible approximation is the already-mentioned Elman preconditioner of the form

$$\hat{S}^{-1} = (BB^T)^{-1}BAB^T(BB^T)^{-1}.$$

Although it was developed specifically for fluid flow problems, this approximation should, at least in principle, be feasible whenever the solution of linear systems with coefficient matrix BB^T (or the equivalent least squares problems) can be computed efficiently. Note that accommodating a nonzero C in this and the previous case is not easy.

Yet another possibility is to approximate the action of S^{-1} on a vector v by performing a few steps of an iterative method on the Schur complement system $Sz = v$. Krylov subspace methods only require S in the form of matrix-vector products. These in turn necessitate multiplications with B and its transpose and solves with A. The latter operation may in turn be performed by an (inner) iterative process. If a variable number of inner iterations is performed, then a generalized conjugate gradient method (Axelsson and Vassilevski 1991) or a flexible solver like FGMRES (Saad 1993) should be used for the outer iteration. These inner solves need not be performed to high accuracy. Even if an accurate approximation $z \approx S^{-1}v$ is sought, the accuracy of the inner solves with A can be relaxed in the course of the outer iteration; see Simoncini and Szyld (2003). We further mention that matrix-vector products involving S and selected vectors are used in the *probing* technique to compute banded approximations to S; see, *e.g.*, Axelsson (1994, Chapter 8.5.1). Finally, Bomhof and van der Vorst (2000) have proposed a parallel direct-iterative approximate Schur complement algorithm for linear systems arising in circuit simulation.

https://doi.org/10.1017/S0962492904000212 Published online by Cambridge University Press

10.2. Constraint and indefinite preconditioning

Constraint preconditioning is based on the principle that the preconditioning matrix should have the same 2×2 block structure as the original saddle point matrix. In other words, the saddle point matrix is preconditioned by another, 'easier-to-invert' saddle point matrix. Constraint preconditioning has been extensively used in the solution of saddle point systems arising from mixed finite element formulations of elliptic partial differential equations (Axelsson and Neytcheva 2003, Bank *et al.* 1990, Ewing, Lazarov, Lu and Vassilevski 1990, Mihajlović and Silvester 2004, Perugia and Simoncini 2000, Rozložník and Simoncini 2002, Tong and Sameh 1998). The method is also popular in the solution of saddle point ('KKT') systems in optimization (Bergamaschi *et al.* 2004, Dyn and Ferguson 1983, Freund and Nachtigal 1995, Gould, Hribar and Nocedal 2001, Keller, Gould and Wathen 2000, Lukšan and Vlček 1998, Shi 1995). Komzsik, Poschmann and Sharapov (1997) applied the preconditioner to problems arising in the finite element analysis of structures.

We describe constraint preconditioning in the case where $B_1 = B_2 = B$ has full row rank and $C = O$. The preconditioning matrix is then

$$\mathcal{P}_c = \begin{bmatrix} G & B^T \\ B & O \end{bmatrix}, \tag{10.15}$$

where G is some approximation of A, $G \neq A$. Of course, G should be chosen so that $\mathcal{P}_c$ is invertible and solving linear systems involving $\mathcal{P}_c$ is significantly easier than solving the original linear system $\mathcal{A}\,u = b$. The name *constraint preconditioning* comes from the fact that $\mathcal{P}_c$ is the coefficient matrix for a saddle point problem with the same constraints as the original problem. Moreover, the constraint preconditioner projects the problem onto the null space of the constraint matrix B, as noted in Lukšan and Vlček (1998) and Perugia *et al.* (1999). With an appropriate choice of the initial guess u_0, the component x_k of every iterate u_k generated by a Krylov method preconditioned with $\mathcal{P}_c$ satisfies the constraint $Bx_k = g$; see, *e.g.*, Rozložník and Simoncini (2002).

When A has positive diagonal entries, G is often taken to be the diagonal part of A, and if $A = (a_{ij})$ has been scaled so that $a_{ii} = 1$ for $1 \leq i \leq n$, then $G = I$. In this case the preconditioning matrix is just

$$\mathcal{P}_c = \begin{bmatrix} I & B^T \\ B & O \end{bmatrix},$$

i.e., the augmented matrix for a least squares problem of the form

$$\min_z \|h - B^T z\|_2,$$

and application of the preconditioner within a Krylov subspace method involves repeated solution of such problems, either explicitly or implicitly.

https://doi.org/10.1017/S0962492904000212 Published online by Cambridge University Press

More generally, solution of linear systems with coefficient matrix $\mathcal{P}_c$ given by (10.15) can be accomplished based on the identity

$$\begin{bmatrix} G & B^T \\ B & O \end{bmatrix}^{-1} = \begin{bmatrix} I & -G^{-1}B^T \\ O & I \end{bmatrix} \begin{bmatrix} G^{-1} & O \\ O & -(BG^{-1}B^T)^{-1} \end{bmatrix} \begin{bmatrix} I & O \\ -BG^{-1} & I \end{bmatrix}. \tag{10.16}$$

If G is sufficiently simple (*e.g.*, diagonal) and linear systems with coefficient matrix $BG^{-1}B^T$ can be solved efficiently, then the action of the preconditioner on a vector can be computed at a reasonable cost. In many cases, however, the 'exact' solution of linear systems involving $BG^{-1}B^T$ is too expensive, and approximations must be used. For instance, incomplete factorizations or sparse approximate inverse techniques can be used to approximately form and invert $BG^{-1}B^T$; see Chow and Saad (1997*a*), Haws (2002), Perugia and Simoncini (2000) and Toh *et al.* (2004). These strategies are often more efficient than methods based on a direct LDL^T factorization of $\mathcal{P}_c$. However, a sparse factorization of $\mathcal{P}_c$ would be advantageous when B contains a dense column, as sometimes happens in optimization problems (Bergamaschi *et al.* 2004), since in this case $BG^{-1}B^T$ is completely dense. Recently, an interesting new factorization (alternative to (10.16)) has been introduced by Schilders; see Dollar and Wathen (2004) and Section 10.4 below. This can also be used to compute the action of $\mathcal{P}_c^{-1}$ efficiently.

For a symmetric positive definite G, the (left) preconditioned matrix is given by

$$\mathcal{M} = \begin{bmatrix} G & B^T \\ B & O \end{bmatrix}^{-1} \begin{bmatrix} A & B^T \\ B & O \end{bmatrix} = \begin{bmatrix} (I-\Pi)\,G^{-1}A + \Pi & O \\ X & I \end{bmatrix} \tag{10.17}$$

where

$$\Pi = G^{-1}B^T(BG^{-1}B^T)^{-1}B$$

is the G-orthogonal projector onto range$(G^{-1}B^T)$ and

$$X = (BG^{-1}B^T)^{-1}B(G^{-1}A - I).$$

Note that in the special case $m = n$, we have $\Pi = I$ and all the eigenvalues of the preconditioned matrix are equal to 1; furthermore,

$$\mathcal{M} - I = \begin{bmatrix} O & O \\ X & O \end{bmatrix} \Rightarrow (\mathcal{M} - I)^2 = O.$$

Therefore the minimal polynomial of the preconditioned matrix has degree 2, and GMRES applied to the preconditioned system delivers the solution in at most two steps, independently of A and G. Of course, if $m = n$ and B is nonsingular then the solution of the saddle point system is simply given by $Bx_* = g$ and $B^T y_* = f - Ax_*$, and the augmented system formulation is neither necessary nor recommended.

https://doi.org/10.1017/S0962492904000212 Published online by Cambridge University Press

In the following, we deal with the more interesting case $m < n$. The eigenvalues and eigenvectors of the preconditioned matrix $\mathcal{M}$ in this case have been studied in several publications, *e.g.*, Durazzi and Ruggiero (2003*a*), Lukšan and Vlček (1998), Perugia and Simoncini (2000) and Rozložník and Simoncini (2002). The following result was given in Keller *et al.* (2000, Theorems 2.1 and 2.3).

Theorem 10.1. Let $A \in \mathbb{R}^{n\times n}$ be symmetric, let $B \in \mathbb{R}^{m\times n}$ have full rank ($= m < n$), and suppose $\mathcal{A}$ is nonsingular. Let $Z \in \mathbb{R}^{n\times(n-m)}$ be a matrix whose columns form a basis for $\ker(B)$. Furthermore, let $G \in \mathbb{R}^{n\times n}$, with $G = G^T \neq A$, be such that $\mathcal{P}_c$ in (10.15) is nonsingular. Then the preconditioned matrix $\mathcal{M}$ in (10.17) has:

(1) the eigenvalue 1 with multiplicity $2m$;
(2) $n - m$ eigenvalues that are defined by the generalized eigenproblem $Z^T A Z x_z = \lambda Z^T G Z x_z$.

Assume, in addition, that $Z^T G Z$ is positive definite. Then $\mathcal{M}$ has the following $m + i + j$ linearly independent eigenvectors:

(1) m eigenvectors of the form $[0^T, y^T]^T$ corresponding to the eigenvalue 1 of $\mathcal{M}$;
(2) i ($0 \le i \le n$) eigenvectors of the form $[w^T, y^T]^T$ corresponding to the eigenvalue 1 of $\mathcal{M}$, where the components w arise from the generalized eigenproblem $Aw = Gw$;
(3) j ($0 \le j \le n-m$) eigenvectors of the form $[x_z^T, 0^T, y^T]^T$ corresponding to the eigenvalues of $\mathcal{M}$ not equal to 1, where the components x_z arise from the generalized eigenproblem $Z^T A Z x_z = \lambda Z^T G Z x_z$ with $\lambda \neq 1$.

Proof. See Keller *et al.* (2000). □

As pointed out in Keller *et al.* (2000), if either $Z^T A Z$ or $Z^T G Z$ are positive definite, then the indefinite preconditioner $\mathcal{P}_c$ (10.15) applied to the indefinite saddle point matrix $\mathcal{A}$ yields a preconditioned matrix $\mathcal{M}$ (10.17) that has real eigenvalues. Clearly, the eigenvalues will be more clustered near 1 the better G approximates A. Spectral properties of the preconditioned matrices in the presence of inexact solves with G and $BG^{-1}B^T$ have been studied by Perugia and Simoncini (2000) and Toh *et al.* (2004). We further mention the approach proposed by Coleman and Verma (2001), where the blocks B and B^T in the constraint preconditioner $\mathcal{P}_c$ are replaced by approximations $\bar{B}$, $\bar{B}^T$ in order to simplify solves involving the preconditioning matrix. In Coleman and Verma (2001), $\bar{B}$ is obtained from B by sparsification; in this case, care must be taken to avoid a (nearly) rank deficient approximation. See also Arbenz *et al.* (2001), Biros and Ghattas (2005*a*, 2005*b*) and Haber and Ascher (2001) for similar approaches in different application contexts.

https://doi.org/10.1017/S0962492904000212 Published online by Cambridge University Press

If, in the notation of Theorem 10.1, $Z^T G Z$ is positive definite, then for each initial residual r_0, $\dim \mathcal{K}_{n+m}(\mathcal{M}, r_0) \leq n-m+2$: see, *e.g.*, Lukšan and Vlček (1998, Theorem 3.4) and Keller *et al.* (2000, Theorem 3.5). Hence GMRES applied to the preconditioned system with matrix $\mathcal{M}$ will terminate after at most $n-m+2$ steps. This is a very attractive feature of constraint preconditioning when $n-m$ is small.

In many applications $n-m+2$ is quite large, so that a total number of $n-m+2$ iterations may be prohibitive. In addition, one is typically interested only in an approximate rather than the exact solution, which means that the iteration can be stopped after a prescribed tolerance is achieved. Then a convergence analysis of the type described in Section 9 is required. Unfortunately, the matrix $\mathcal{M}$ is not diagonalizable, so that standard convergence estimates like (9.9) are not applicable. In some cases, however, it is still possible to derive convergence bounds similar to (9.9) for GMRES applied to the preconditioned system. For example, assume that $A = A^T$ is positive definite, B has full rank, and $G = I$. Then, for a special initial residual of the form $r_0 = [s_0^T, 0^T]^T$, it can be shown that the convergence of GMRES for the *right* constraint preconditioned system only depends on the (1,1) block of the preconditioned matrix $\mathcal{A}\mathcal{P}_c^{-1}$. In our notation this (1,1) block is given by the (diagonalizable) matrix $A(I-\Pi)+\Pi$; see Rozložník and Simoncini (2002, Section 6) for details.

It should be pointed out that GMRES might not be the best method for solving the constraint preconditioned linear system. The fact that constraint preconditioning incorporates a projection onto $\ker(B)$ suggests that this preconditioning technique is closely related to the null space method discussed in Section 7. The two methods are, in fact, mathematically equivalent (Gould *et al.* 2001, Lukšan and Vlček 1998, Perugia *et al.* 1999). Exploitation of this equivalence allows the construction of a CG method for a constraint preconditioned positive definite reduced system (Gould *et al.* 2001). Unlike GMRES (when applied to a system with the nonsymmetric matrix $\mathcal{M}$), this conjugate gradient method is based on efficient three-term recurrences. In the numerical experiments in Keller *et al.* (2000), this preconditioned conjugate gradient method represents a viable alternative to a direct solver (here MA27 from the Harwell subroutine library (Duff and Reid 1983)), even for systems of relatively small order, and it typically outperforms GMRES applied to $\mathcal{M}$.

All numerical experiments in Keller *et al.* (2000) are performed using test problems arising in linear and nonlinear optimization taken from the *constrained and unconstrained testing environment* (CUTE) (Bongartz, Conn, Gould and Toint 1995). Another situation where constraint preconditioning performs well is when the saddle point system arises from a mixed finite element formulation of a second-order, elliptic PDE. In this case it is often possible to find an easily invertible (*e.g.*, diagonal) G that is spectrally equi-

https://doi.org/10.1017/S0962492904000212 Published online by Cambridge University Press

valent to A. In many cases of practical interest, $G = I$ will do. The resulting constraint preconditioner (10.15) leads to asymptotic rates of convergence for the *right* preconditioned system that are independent of h, the mesh discretization parameter, when the initial guess $u_0 = [0^T, g^T]^T$ (meaning that r_0 is of the form $r_0 = [s_0^T, 0^T]^T$) is used. As in the analysis in Rozložník and Simoncini (2002) mentioned above, the convergence of GMRES then only depends on the (diagonalizable) matrix $A(I-\Pi)+\Pi$. The eigenvalues of this matrix are either 1, or contained in an interval $[\alpha_0, \alpha_1]$, $0 < \alpha_0 \leq \alpha_1 < 1$, where both α_0 and α_1 are independent of h (Perugia and Simoncini 2000, Theorem 1) (*cf.* also Ewing *et al.* (1990)). Furthermore, the condition number of the eigenvector matrix of $A(I-\Pi)+\Pi$ is bounded independently of h (Rozložník and Simoncini 2002, Proposition 6.3). The asymptotic optimality of these methods, however, does not always translate in computational efficiency. As shown in Perugia and Simoncini (2000), it may be much more efficient to replace the exact (but costly) preconditioner solves with inexact ones, even if this introduces a dependency on h in the number of preconditioned iterations. This is especially true of saddle point systems that arise from the solution of 3D PDE problems. See also Mihajlović and Silvester (2004) for a discussion of inexact constraint preconditioning for mixed formulations of biharmonic problems in 2D.

Some of these results have been extended to the case where both A and G are allowed to be nonsymmetric; see Cao (2002). The case where $B_1 \neq B_2$ was investigated in Ikramov (2003) and Wei and Zhang (2004). Dyn and Ferguson (1983) studied the case where A is symmetric but G may be nonsymmetric, and $B_1 = B_2 = B$. Assuming that A is symmetric positive semidefinite with $\ker(A) \cap \ker(B) = \{0\}$ and that G is nonsingular and such that $G + G^T - A$ is positive definite, they showed that the matrix $\mathcal{P}_c$ in (10.15) is nonsingular, and that the stationary iteration

$$\begin{bmatrix} G & B^T \\ B & O \end{bmatrix} \begin{bmatrix} x_{k+1} \\ y_{k+1} \end{bmatrix} = \begin{bmatrix} G-A & O \\ O & O \end{bmatrix} \begin{bmatrix} x_k \\ y_k \end{bmatrix} + \begin{bmatrix} f \\ g \end{bmatrix}, \quad k = 0, 1, \ldots \tag{10.18}$$

is convergent for any choice of the initial guess. Therefore, the spectrum of $\mathcal{P}_c^{-1}\mathcal{A}$ is contained in the disk $\{z \in \mathbb{C}\,;\, |z-1| < 1\}$. A matrix splitting of the form $A = G - (G - A)$ such that $G + G^T - A$ is positive definite is called a *P-regular splitting*; see Ortega (1972, page 122). Examples of such splittings when A is symmetric positive semidefinite with a nonzero diagonal include the damped Jacobi, SOR, and SSOR splittings (Dyn and Ferguson 1983). We mention that variants of the coupled iteration (10.18) have been used as smoothers for multigrid methods; see Braess and Sarazin (1997), Leem, Oliveira and Stewart (2004), Wabro (2004), Zulehner (2000) and Section 11 below.

The important case where A is nonsymmetric but has positive definite symmetric part has been studied in Golub and Wathen (1998), motivated

https://doi.org/10.1017/S0962492904000212 Published online by Cambridge University Press

by the Oseen (linearized Navier–Stokes) equations, and more recently in Baggag and Sameh (2004), motivated by particulate flow problems. In both of these papers, the idea is to use $G = H \equiv \frac{1}{2}(A + A^T)$, the symmetric part of A. For the nonsymmetric linear system arising from the discretization of the Oseen equations (2.7)–(2.9), the preconditioner becomes the (symmetric) matrix of the corresponding Stokes system (2.10)–(2.12). When the kinematic viscosity coefficient ν in (2.7) is not too small, A is not too far from being symmetric and this preconditioning strategy is very effective, yielding convergence rates that are independent of the mesh size. In practice the exact solution of the preconditioning equation

$$\mathcal{P}_c z = r, \quad \mathcal{P}_c = \begin{bmatrix} H & B^T \\ B & O \end{bmatrix} \tag{10.19}$$

may be too expensive, and inexact solves may be required. For example, (10.19) may be solved approximately by an iterative method, leading to a nested (or inner-outer) iterative scheme. As shown in Baggag and Sameh (2004), it is often possible in this way to retain the excellent convergence properties of the outer scheme, while at the same time reducing the work per iteration significantly. On the other hand, this approach does not work well when A is far from symmetric (large Reynolds number, or small ν in (2.7)); see the numerical experiments reported in Benzi and Liu (2005) and Botchev and Golub (2004). In this case, better results can be obtained with preconditioners of the form

$$\mathcal{P}_c = \begin{bmatrix} T & B^T \\ B & O \end{bmatrix},$$

where the submatrix T is nonsymmetric positive definite with a 'large' skew-symmetric part that incorporates information from K, the skew-symmetric part of A; see Botchev and Golub (2004).

The case where $A = A^T$, $B_1 = B_2 = B$ and $C = C^T \neq O$ has been considered in Axelsson and Neytcheva (2003), Bergamaschi *et al.* (2004), Dollar (2005), Perugia and Simoncini (2000), Toh *et al.* (2004) and Zulehner (2002). Here A and C are positive semidefinite and G is chosen so that the preconditioning matrix

$$\mathcal{P}_c = \begin{bmatrix} G & B^T \\ B & -C \end{bmatrix} \tag{10.20}$$

is nonsingular. Furthermore, linear systems involving G and $C + BG^{-1}B^T$ should be easy to (approximately) solve. The spectral properties of this preconditioner have been studied in Axelsson and Neytcheva (2003), Perugia and Simoncini (2000) and Toh *et al.* (2004). We mention the following results from Axelsson and Neytcheva (2003).

https://doi.org/10.1017/S0962492904000212 Published online by Cambridge University Press

Lemma 10.2. Let B, C, E be real matrices of order $m \times n$, $m \times m$ and $n \times n$, respectively, where B has full rank, $C = C^T$ is positive semidefinite, and $E = E^T$. Then the eigenvalues of the generalized eigenproblem

$$\gamma \begin{bmatrix} I & B^T \\ B & -C \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} E & O \\ O & O \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}, \quad \|x\| + \|y\| \neq 0, \tag{10.21}$$

where $x \in \mathbb{C}^n$ and $y \in \mathbb{C}^m$, satisfy

(i) $\gamma = \frac{x^H E x}{x^H (I + B^T C^{-1} B) x} \neq 0$ if $Ex \neq 0$ and C is positive definite,

(ii) $\gamma = 0$ if and only if $Ex = 0$, $y \neq 0$ and the dimension of the eigenspace corresponding to the zero eigenvalue is $m + q$, where $q = \dim \ker(E)$,

(iii) the nonzero eigenvalues are contained in the interval $\min\{0, \lambda_{\min}(E)\} \leq \gamma \leq \lambda_{\max}(E)$.

Consider now a symmetric saddle point matrix $\mathcal{A}$ and the preconditioner $\mathcal{P}_c$ given by (10.20), with G symmetric and positive definite. The eigenvalues of the preconditioned matrix $\mathcal{P}_c^{-1}\mathcal{A}$ are those of the generalized eigenproblem $\mathcal{A}u = \lambda \mathcal{P}_c u$; but these are of the form $\lambda = \gamma + 1$, where γ are the eigenvalues of the generalized eigenproblem

$$\gamma \begin{bmatrix} G & B^T \\ B & -C \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} A - G & O \\ O & O \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}.$$

The latter is equivalent to

$$\gamma \begin{bmatrix} I & \hat{B}^T \\ \hat{B} & -C \end{bmatrix} \begin{bmatrix} \hat{x} \\ y \end{bmatrix} = \begin{bmatrix} \hat{E} & O \\ O & O \end{bmatrix} \begin{bmatrix} \hat{x} \\ y \end{bmatrix}, \tag{10.22}$$

where $\hat{B} = BG^{-1/2}$, $\hat{E} = G^{-1/2} A G^{-1/2} - I$ and $\hat{x} = G^{1/2} x$. This generalized eigenproblem is of the form (10.21) and therefore Lemma 10.2 can be applied. In particular, the eigenvalues of $\mathcal{P}_c^{-1}\mathcal{A}$ are real. Clearly, the better G approximates A, the more clustered around 1 are the eigenvalues of $\mathcal{P}_c^{-1}\mathcal{A}$.

Another form of indefinite preconditioning is based on the observation that 'stabilized' saddle point systems (*i.e.*, systems of the type (1.6) with $C \neq O$) are generally easier to solve than standard saddle point systems (*i.e.*, with $C = O$), in the sense that iterative methods tend to converge faster when $C \neq O$. It is also easier to construct effective block preconditioners when $C \neq O$, since the (negative) Schur complement $C + BA^{-1}B^T$ is typically better conditioned than $BA^{-1}B^T$. This observation justifies the use of preconditioners of the form (10.20), with a suitably chosen C, even if the $(2,2)$ block is zero in the original problem. This approach, referred to as *regularized preconditioning*, was first considered in Axelsson (1979) in the special case $G = A$ and $C = \varepsilon I$, where $\varepsilon > 0$. The preconditioning

https://doi.org/10.1017/S0962492904000212 Published online by Cambridge University Press

equation itself, $\mathcal{P}_c z = r$, is solved by an iterative method, leading again to an inner-outer process.

The case where $G = A$ is symmetric positive definite, $B_1 = B_2 = B$ has full rank, and C is symmetric positive semidefinite is easily analysed. In this case we have

$$\mathcal{A} - \mathcal{P}_c = \begin{bmatrix} O & O \\ O & C \end{bmatrix} \;\Rightarrow\; I - \mathcal{A}^{-1}\mathcal{P}_c = \mathcal{A}^{-1}(\mathcal{A} - \mathcal{P}_c) = \begin{bmatrix} O & Y \\ O & S^{-1}C \end{bmatrix},$$

where $S = -BA^{-1}B^T$ and Y is a certain nonzero matrix. It follows that

$$\mathcal{P}_c^{-1}\mathcal{A} = \begin{bmatrix} I & \hat{Y} \\ O & (I - S^{-1}C)^{-1} \end{bmatrix},$$

where $\hat{Y}$ is again a nonzero matrix. Since S is negative definite and C is positive semidefinite, $I - S^{-1}C$ has positive real eigenvalues. Therefore the spectrum of the preconditioned matrix $\mathcal{P}_c^{-1}\mathcal{A}$ is real and contained in the interval $(\beta, 1]$, where

$$\beta = \frac{1}{1 + \lambda_{\max}(-S^{-1}C)} \geq \frac{1}{1 + \|S^{-1}\|_2 \, \|C\|_2}.$$

Also, $\mathcal{P}_c^{-1}\mathcal{A}$ has eigenvalue 1 with multiplicity $n + p$, where $p = \dim \ker(C)$. This approach is especially efficient for the Stokes problem, for which S is spectrally equivalent to the identity. Then $\|S^{-1}\|_2 \approx 1$ and as long as $\|C\|_2$ is bounded above independently of h, the spectrum of $\mathcal{P}_c^{-1}\mathcal{A}$ also remains bounded as $h \to 0$ and fast convergence of the preconditioned iteration can be expected, independent of mesh size. In practice there is no need to solve with $\mathcal{P}_c$ exactly: an approximate solution will suffice and will be more efficient. Spectral bounds for the inexact case can be found in Gorelova and Chizhonkov (2004) for $C = O$ and in Dohrmann and Lehoucq (2004) for $C = C^T \neq O$. It should be mentioned that regularization of indefinite (constraint) preconditioners has also been used with good results in optimization problems; see Bergamaschi *et al.* (2004) as well as Durazzi and Ruggiero (2003*a*, 2003*b*); the latter papers also include results for a parallel implementation.

Concerning the choice of a Krylov subspace method to use with constrained preconditioning, we distinguish between the symmetric case ($A = A^T$, $B_1 = B_2$, $C = C^T$) and the nonsymmetric case. In the nonsymmetric case it is generally necessary to use a nonsymmetric Krylov subspace method, such as GMRES or BiCGStab, whether the constraint preconditioner used is symmetric or not. In the case of a symmetric $\mathcal{A}$ with a symmetric constraint preconditioner $\mathcal{P}_c$, on the other hand, there are alternatives to the use of nonsymmetric Krylov subspace methods. Although the preconditioned matrix $\mathcal{M}$ itself is nonsymmetric, the SQMR method can be used; see the discussion at the end of Section 9, in particular Table 9.1.

https://doi.org/10.1017/S0962492904000212 Published online by Cambridge University Press

Furthermore, as described above, in some cases a combination of the null space method and the CG method may be very effective. As noted in Bramble and Pasciak (1988) and Lukšan and Vlček (1998), in the symmetric case it is sometimes even possible to use the CG method for the preconditioned system with $\mathcal{M}$. However, the indefiniteness of both $\mathcal{P}_c$ and $\mathcal{A}$ does not in general lead to a robust algorithm, and safeguard strategies have to be applied to overcome potential breakdowns. For the case of right constraint preconditioning and assuming $g = 0$ in (1.3), such strategies are discussed in detail in Lukšan and Vlček (1998) and Rozložník and Simoncini (2002). When such strategies are applied, the preconditioned conjugate gradient method can be competitive with direct solution methods, particularly for problems from nonlinear equality-constrained optimization (Lukšan and Vlček 1998, 2001). Finally, use of the CG method with an inexact form of regularized constraint preconditioning has been rigorously justified in Dohrmann and Lehoucq (2004).

10.3. Hermitian/skew-Hermitian preconditioning

The Hermitian and skew-Hermitian splitting (HSS) was introduced as a stationary iterative method in Bai, Golub and Ng (2003), where it was shown to converge for non-Hermitian positive definite systems, *i.e.*, linear systems $Ax = b$ with $A + A^*$ positive definite. In the real case (which is the only one we consider here), such systems are said to be *positive real*. Problems of this type arise, for instance, in the numerical solution of convection-diffusion equations. The use of HSS as a preconditioner for rather general saddle point problems has been studied in Benzi, Gander and Golub (2003), Benzi and Golub (2004) and Simoncini and Benzi (2004).

The HSS preconditioner is based on the nonsymmetric formulation (3.10). Here we are under the assumptions of Theorem 3.6. In particular, $B_1 = B_2 = B$ and $C = C^T$. Letting $H \equiv \frac{1}{2}(A + A^T)$ and $K \equiv \frac{1}{2}(A - A^T)$ we have the following splitting of $\hat{\mathcal{A}}$ into its symmetric and skew-symmetric part:

$$\hat{\mathcal{A}} = \begin{bmatrix} A & B^T \\ -B & C \end{bmatrix} = \begin{bmatrix} H & O \\ O & C \end{bmatrix} + \begin{bmatrix} K & B^T \\ -B & O \end{bmatrix} = \mathcal{H} + \mathcal{K}. \tag{10.23}$$

Note that $\mathcal{H}$, the symmetric part of $\hat{\mathcal{A}}$, is symmetric positive semidefinite since both H and C are. Let $\alpha > 0$ be a parameter. In the same spirit as the classical ADI (Alternating-Direction Implicit) method (Varga 1962), we consider the following two splittings of $\hat{\mathcal{A}}$:

$$\hat{\mathcal{A}} = (\mathcal{H} + \alpha\mathcal{I}) - (\alpha\mathcal{I} - \mathcal{K}) \quad \text{and} \quad \hat{\mathcal{A}} = (\mathcal{K} + \alpha\mathcal{I}) - (\alpha\mathcal{I} - \mathcal{H}).$$

Here $\mathcal{I}$ denotes the identity matrix of order $n + m$. The stationary HSS iteration is obtained by alternating between these two splittings. Given an

https://doi.org/10.1017/S0962492904000212 Published online by Cambridge University Press

initial guess u_0, the HSS iteration computes a sequence $\{u_k\}$ as follows:

$$\begin{cases} (\mathcal{H} + \alpha\mathcal{I})\, u_{k+\frac{1}{2}} = (\alpha\mathcal{I} - \mathcal{K})\, u_k + \hat{b}, \\ (\mathcal{K} + \alpha\mathcal{I})\, u_{k+1} = (\alpha\mathcal{I} - \mathcal{H})\, u_{k+\frac{1}{2}} + \hat{b}. \end{cases} \tag{10.24}$$

Note that both $\mathcal{H} + \alpha\mathcal{I}$ and $\mathcal{K} + \alpha\mathcal{I}$ are nonsingular. The first matrix is symmetric positive definite while the second one is a shifted skew-symmetric matrix with eigenvalues of the form $\alpha + i\nu_j$, where $\nu_j \in \mathbb{R}$ for all $j = 1, \ldots, n+m$.

The two-step process (10.24) can be written as a fixed point iteration by eliminating the intermediate vector $u_{k+\frac{1}{2}}$, yielding

$$u_{k+1} = \mathcal{T}_\alpha u_k + c. \tag{10.25}$$

Here

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

is the iteration matrix of the method, and

$$c := (\mathcal{K} + \alpha\mathcal{I})^{-1}[\mathcal{I} + (\alpha\mathcal{I} - \mathcal{H})(\mathcal{H} + \alpha\mathcal{I})^{-1}]\hat{b}.$$

The fixed point iteration (10.25) converges for arbitrary initial guesses u_0 and right-hand sides $\hat{b}$ to the solution $u = \hat{\mathcal{A}}^{-1}\hat{b}$ if and only if $\rho(\mathcal{T}_\alpha) < 1$, where $\rho(\mathcal{T}_\alpha)$ denotes the spectral radius of $\mathcal{T}_\alpha$. It follows from the results in Bai *et al.* (2003) that when $\mathcal{H}$ is positive definite, the stationary iteration (10.24) converges for all $\alpha > 0$ to the solution of $\hat{\mathcal{A}}\, u = \hat{b}$. For saddle point problems, $\mathcal{H}$ is positive definite if and only if H and C are. These conditions are rather restrictive in practice, especially the one on C. However, it was shown in Benzi and Golub (2004) that the HSS iteration converges for all $\alpha > 0$ if H is positive definite, B is of full rank and C is positive semidefinite (possibly zero). The method can be made to converge even when H is positive semidefinite and singular, provided that $\ker(H) \cap \ker(B) = \{0\}$; see Benzi and Golub (2004).

It was shown in Bai *et al.* (2003) that when $\mathcal{H}$ is positive definite, the choice

$$\alpha = \sqrt{\lambda_{\min}(\mathcal{H})\lambda_{\max}(\mathcal{H})}$$

minimizes an upper bound on the spectral radius of the iteration matrix $\mathcal{T}_\alpha$. Unfortunately, in most saddle point problems $\lambda_{\min}(\mathcal{H}) = 0$, so the result does not apply. Furthermore, the rate of convergence of the HSS iteration is rather slow, even with the 'optimal' choice of α. For these reasons Benzi and Golub (2004) proposed that GMRES or other Krylov subspace methods should be used to accelerate the convergence of the HSS method. In other words, the HSS method is best used as a preconditioner for (say) GMRES rather than as a stationary iterative method. As observed in Benzi and Golub (2004), there is a unique splitting $\hat{\mathcal{A}} = \mathcal{P} - \mathcal{Q}$ with $\mathcal{P}$ nonsingular

https://doi.org/10.1017/S0962492904000212 Published online by Cambridge University Press

such that the iteration matrix $\mathcal{T}_\alpha$ is the matrix induced by that splitting, *i.e.*, $\mathcal{T}_\alpha = \mathcal{P}^{-1}\mathcal{Q} = \mathcal{I} - \mathcal{P}^{-1}\mathcal{A}$. An easy calculation shows that the HSS iteration (10.24) can be written in *correction form* as

$$u_{k+1} = u_k + \mathcal{P}_\alpha^{-1} r_k, \qquad r_k = \hat{b} - \hat{\mathcal{A}} u_k,$$

where the preconditioner $\mathcal{P}$ is given by

$$\mathcal{P} \equiv \mathcal{P}_\alpha = \tfrac{1}{2\alpha}(\mathcal{H} + \alpha\mathcal{I})(\mathcal{K} + \alpha\mathcal{I}). \tag{10.26}$$

Note that as a preconditioner we can use $\mathcal{P}_\alpha = (\mathcal{H} + \alpha\mathcal{I})(\mathcal{K} + \alpha\mathcal{I})$ instead of the expression given in (10.26), since the factor $\frac{1}{2\alpha}$ has no effect on the preconditioned system. It is just a normalization factor that allows us to conclude that the eigenvalues of the preconditioned matrix $\mathcal{P}_\alpha^{-1}\hat{\mathcal{A}}$ (or $\hat{\mathcal{A}}\,\mathcal{P}_\alpha^{-1}$, which has the same spectrum) are all contained in the disk $\{z \in \mathbb{C}\,;\, |z-1| < 1\}$. In particular, the spectrum of the preconditioned matrix, like that of $\hat{\mathcal{A}}$, lies entirely in the right half-plane: the preconditioned matrix is positive stable.

The rate of convergence of nonsymmetric Krylov iterations (like GMRES) preconditioned by $\mathcal{P}_\alpha$ depends on the particular choice of α. Finding the value of α that optimizes the rate of convergence is a very difficult problem in general. Note that the value of α that minimizes the number of GMRES iterations may be quite different from the one that minimizes the spectral radius of $\mathcal{T}_\alpha$; see Benzi *et al.* (2003). Numerical experiments show that with an appropriate scaling of the system (such that the nonzero diagonal entries of $\hat{\mathcal{A}}$ are equal to 1), there is a unique value α_* of α for which the number of preconditioned iterations is minimized, and this α_* is usually a small number, between 0 and 1. In some cases, but not always, the optimal α can be determined by trial and error on a small example (*e.g.*, a coarse discretization of the continuous problem to be solved) and then used with good results on larger problems corresponding to finer discretizations; see Benzi and Ng (2004), where HSS preconditioning was used to solve weighted least squares problems arising in image processing.

Spectral properties of the preconditioned matrix as a function of α have been studied, under different sets of assumptions, in Benzi *et al.* (2003), Benzi and Ng (2004) and Simoncini and Benzi (2004). A Fourier analysis of HSS preconditioning for saddle point formulations of Poisson's equation (including the anisotropic case) was given in Benzi *et al.* (2003). The analysis showed that using a sufficiently small value of α results in h-independent convergence. Furthermore, as $\alpha \to 0$ the eigenvalues of the preconditioned matrix are all real and fall within two small intervals $(0, \varepsilon_1)$ and $(2-\varepsilon_2, 2)$, with $\varepsilon_1, \varepsilon_2 > 0$ and $\varepsilon_1, \varepsilon_2 \to 0$ as $\alpha \to 0$. This clustering result was generalized in Simoncini and Benzi (2004) to general saddle point systems with $A = A^T$ positive definite and $C = O$, using purely algebraic arguments. However, h-independent convergence is not always guaranteed: for example,

https://doi.org/10.1017/S0962492904000212 Published online by Cambridge University Press

it does not occur for the Stokes problem. Nevertheless, good results have been obtained in the solution of the Oseen problem, where HSS preconditioning appears to be competitive with other preconditioners for large Reynolds number (small ν). Numerical experiments indicate that the optimal value of α for steady 2D problems is largely independent of ν and is approximately given by $h^{1/2}$, where h denotes the mesh size. Furthermore, rapid convergence independently of h and ν is observed for unsteady problems; see the numerical experiments in Benzi (2004) and Benzi and Liu (2005).

Application of the HSS preconditioner within GMRES requires solving a linear system of the form $\mathcal{P}_\alpha z_k = r_k$ at each iteration. This is done by first solving the system

$$(\mathcal{H} + \alpha\mathcal{I})\, w_k = r_k \tag{10.27}$$

for w_k, followed by

$$(\mathcal{K} + \alpha\mathcal{I})\, z_k = w_k. \tag{10.28}$$

Recalling the form of $\mathcal{H}$, see (10.23), equation (10.27) consists of two decoupled systems with coefficient matrices $H + \alpha I$ and $C + \alpha I$, respectively. Both matrices are symmetric positive definite, and a number of efficient methods can be applied, including Cholesky factorization, preconditioned CG (PCG) schemes, or multigrid, either geometric or algebraic. Multigrid methods are the solvers of choice for a number of problems arising from the discretization of partial differential equations, particularly in the case of the Oseen problem (2.10)–(2.11) in $\Omega \subset \mathbb{R}^3$, where H is a direct sum of three discrete Laplace operators. The solution of the system with matrix $C + \alpha I$ is often much easier and reduces to a scaling by α when $C = O$. When solving regularized weighted least squares problems, both H and C are diagonal and the cost of solving (10.27) is negligible; see Benzi and Ng (2004).

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 (10.27). 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

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

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 ($\kappa\,(H + \alpha I) \le 11$). For most problems, both multigrid and the CG method applied to (10.27) can be expected to converge rapidly, independent of the number n of unknowns.

Solving (10.28) is usually more involved, especially when $K \neq O$. It

https://doi.org/10.1017/S0962492904000212 Published online by Cambridge University Press

requires the solution of a linear system of the form

$$\begin{cases} (\alpha I + K)\, x_{k+1} + B^T y_{k+1} = f_k, \\ -B x_{k+1} + \alpha y_{k+1} = g_k. \end{cases} \tag{10.29}$$

This system can be solved in several ways. One approach is to eliminate x_{k+1} from the second equation using the first one (Schur complement reduction), leading to a smaller (order m) linear system of the form

$$[B(I + \alpha^{-1}K)^{-1}B^T + \alpha^2 I]y_{k+1} = B(I + \alpha^{-1}K)^{-1}f_k + \alpha g_k. \tag{10.30}$$

Once the solution y_{k+1} to (10.30) has been computed, the vector x_{k+1} is given by

$$(\alpha I + K)x_{k+1} = f_k - B^T y_{k+1}.$$

When $K = O$, system (10.30) simplifies to

$$(BB^T + \alpha^2 I)\, y_{k+1} = Bf_k + \alpha g_k, \tag{10.31}$$

and $x_{k+1} = \alpha^{-1}(f_k - B^T y_{k+1})$. If BB^T is sufficiently sparse, system (10.31) could be formed explicitly and solved by a sparse Cholesky factorization. Otherwise, an iterative method like LSQR with a simple preconditioner could be used. If B represents a discrete divergence operator, then $BB^T + \alpha^2 I$ is just a shifted discrete Laplace operator, and many fast solvers for system (10.31) are available. When $K \neq O$ the coefficient matrix in (10.30) is generally dense. An important exception is the Oseen problem in rotation form (2.15)–(2.17). In the 2D case the coefficient matrix has the form

$$\hat{\mathcal{A}} = \begin{bmatrix} A_1 & D & B_1^T \\ -D & A_2 & B_2^T \\ -B_1 & -B_2 & C \end{bmatrix} = \begin{bmatrix} A_1 & O & O \\ O & A_2 & O \\ O & O & C \end{bmatrix} + \begin{bmatrix} O & D & B_1^T \\ -D & O & B_2^T \\ -B_1 & -B_2 & O \end{bmatrix} = \mathcal{H} + \mathcal{K}. \tag{10.32}$$

Here A_1, A_2 are discrete Laplace operators with appropriate boundary conditions, $B = \begin{bmatrix} B_1 & B_2 \end{bmatrix}$, $C = C^T$ is a stabilization term that can be assumed to be zero if the discretization used is already stable, and $D = D^T$ is a matrix that corresponds to multiplication by $w = \nabla \times \mathbf{v}$; see Section 2.1 for details. For finite difference schemes (for instance, MAC (Harlow and Welch 1965)) D is a diagonal matrix; for finite elements, it will be a scaled mass matrix. When D is diagonal, the Schur complement $B(I + \alpha^{-1}K)^{-1}B^T + \alpha^2 I$ is a sparse matrix and can be formed explicitly. This follows from the fact that

$$(I + \alpha^{-1}K)^{-1} = \begin{bmatrix} I & \alpha^{-1}D \\ -\alpha^{-1}D & I \end{bmatrix}^{-1} = \begin{bmatrix} E_1 & -E_2 \\ E_2 & E_3 \end{bmatrix},$$

https://doi.org/10.1017/S0962492904000212 Published online by Cambridge University Press

where E_1, E_2 and E_3 are diagonal matrices given by

$$E_1 = I - \alpha^{-2} D (I + \alpha^{-2} D^2)^{-1} D,$$
$$E_2 = \alpha^{-1} D (I + \alpha^{-2} D^2)^{-1},$$
$$E_3 = (I + \alpha^{-2} D^2)^{-1}.$$

When D is not diagonal, we can replace it with a spectrally equivalent diagonal approximation and still have a sparse Schur complement; since we are constructing a preconditioner, the action of $(\mathcal{K} + \alpha\mathcal{I})^{-1}$ need not be computed exactly. Hence, the Schur complement $B(I + \alpha^{-1}K)^{-1}B^T + \alpha^2 I$ (or the approximation of it obtained by replacing D with a diagonal matrix) is sparse, and system (10.28) can be efficiently solved via (10.30) using sparse matrix techniques. It is also possible to use multigrid, since the Schur complement can be interpreted as a discretization of a second-order, elliptic operator with variable coefficients. Alternatively, an ILU-preconditioned GMRES can be used. While we have focused here on the 2D case, the 3D case can be treated along the same lines; see Benzi and Liu (2005).

Besides the Schur complement reduction, there are other approaches that can be used to solve linear systems with matrix $\mathcal{K} + \alpha\mathcal{I}$. Note that this is a normal matrix of the form 'identity-plus-skew-symmetric'. Various Lanczos-type methods can be applied to systems of this kind; see Concus and Golub (1976), Widlund (1978) and, more generally, Huhtanen (2002). Other approaches to the solution of shifted skew-symmetric systems are studied in Golub and Vanderstraeten (2000). Yet another possibility is to regard (10.29) as a general nonsymmetric system and to use preconditioned GMRES (say). Many of these schemes can benefit from the fact that for even moderate values of $\alpha > 0$, the condition number of $\mathcal{K} + \alpha\mathcal{I}$ is often rather small.

It is important to stress that the linear systems in (10.24) need not be solved exactly. The use of inexact solves was considered in Bai *et al.* (2003) for the positive real case. The upshot is that inexact solves can 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 method like FGMRES (Saad 1993) should be used for the outer iteration. Inexact solves may also be performed by means of matrix splittings or incomplete factorizations; see Benzi and Golub (2004). In this case, standard GMRES can be used for the outer iteration.

https://doi.org/10.1017/S0962492904000212 Published online by Cambridge University Press

Other iterative solvers for the Navier–Stokes equations in rotation form have been introduced and studied in Olshanskii (1999) and Olshanskii and Reusken (2002). Although the rotation form does not appear to be widely used in practice, it has some advantages over the (standard) convective form (2.7)–(2.9). As mentioned in Olshanskii and Reusken (2002, page 1685), the numerical solution of the Navier–Stokes equations in rotation form is a topic that deserves further study.

10.4. Approximate and incomplete factorization methods

We conclude this section on preconditioners with a brief discussion of recent attempts to develop effective and robust approximate and incomplete factorization methods for saddle point matrices.

Approximate factorizations for symmetric saddle point systems have been used in Gill *et al.* (1992) in the context of interior point methods for constrained optimization. Here we assume that $A = A^T$ is positive definite, $B_1 = B_2 = B$ is of full rank, and $C = C^T$ is positive semidefinite (possibly zero). The approach taken in Gill *et al.* (1992) is to define the preconditioner in terms of an *exact* LDL^T factorization of an approximation $\mathcal{P} \approx \mathcal{A}$. More precisely, the idea is to compute a sparse Bunch–Parlett factorization

$$\mathcal{P} = \mathcal{Q}^T \mathcal{L} \mathcal{D} \mathcal{L}^T \mathcal{Q},$$

where $\mathcal{P}$ is symmetric indefinite, $\mathcal{Q}$ is a permutation matrix, $\mathcal{L}$ is unit lower triangular, and $\mathcal{D}$ a block diagonal matrix with blocks of dimension 1 and 2. The resulting factorization can be used as a preconditioner with SQMR or with a nonsymmetric Krylov subspace solver. Gill *et al.* (1992) modify $\mathcal{P}$ to guarantee it is positive definite; the resulting preconditioner can then be used with a symmetric Krylov subspace solver, such as SYMMLQ or MINRES. In order to do so, it is sufficient to compute the eigendecomposition of $\mathcal{D}$ (which can be done very easily), and then to change the sign of the negative eigenvalues of $\mathcal{D}$. If $\bar{\mathcal{D}}$ denotes the resulting block diagonal matrix, the symmetric positive definite preconditioner is

$$\bar{\mathcal{P}} = \mathcal{Q}^T \mathcal{L} \bar{\mathcal{D}} \mathcal{L}^T \mathcal{Q}.$$

The main issue is the choice of the approximation $\mathcal{P} \approx \mathcal{A}$. As usual, a trade-off is involved. On one hand, $\mathcal{P}$ must be sufficiently close to $\mathcal{A}$ so that convergence of the preconditioned iteration will be rapid; on the other, $\mathcal{P}$ must be such that an LDL^T factorization can be computed rapidly and without too much fill-in. The choice of such $\mathcal{P}$ is clearly problem-dependent. One possibility is to introduce a re-partitioning of $\mathcal{A}$ in the form

$$\mathcal{A} = \begin{bmatrix} A_{11} & A_{12}^T \\ A_{12} & A_{22} \end{bmatrix}, \tag{10.33}$$

https://doi.org/10.1017/S0962492904000212 Published online by Cambridge University Press

where the order of A_{11} does not exceed $n-m$. This condition is necessary in order to have a nonsingular block A_{22}. If A_{11} is strongly diagonally dominant, then a reasonable approximation could be

$$\mathcal{P} = \begin{bmatrix} D_{11} & O \\ O & A_{22} \end{bmatrix},$$

where D_{11} denotes the main diagonal of A_{11}. For linear systems arising in interior point methods, it is often possible to find a permutation of the (1,1) block A of $\mathcal{A}$ so that $\mathcal{A}$ can be cast in the form (10.33) with A_{11} having very large entries on the main diagonal. Indeed, when a variable x_i is approaching the boundary of the feasible set, the ith diagonal entry of the Hessian A becomes very large; by numbering these variables first, the saddle point matrix $\mathcal{A}$ can be given the desired form. The block A_{22} is now a smaller saddle point-type matrix. If $m \ll n$, the submatrix A_{22} may be much smaller than $\mathcal{A}$, and a sparse Bunch–Parlett factorization of A_{22} can be computed efficiently. In the context of interior point methods, this preconditioner can be expected to be especially effective in later stages of Newton's iteration, when many of the diagonal entries in A_{11} are large. See Gill *et al.* (1992) for additional discussion and alternative approaches.

A different approximate factorization method has been presented by Dollar and Wathen (2004) in the context of constraint preconditioning. Consider a constraint preconditioner (10.15) partitioned as

$$\mathcal{P}_c = \begin{bmatrix} G_{11} & G_{12} & B_1^T \\ G_{21} & G_{22} & B_2^T \\ B_1 & B_2 & O \end{bmatrix},$$

where the block B_1 is nonsingular. Assume we choose matrices $L_1 \in \mathbb{R}^{m\times m}$ and $L_2 \in \mathbb{R}^{(n-m)\times(n-m)}$, with L_2 nonsingular; for instance, $L_2 = I$. Then $\mathcal{P}_c$ can be factorized as follows:

$$\mathcal{P}_c = \begin{bmatrix} B_1^T & O & L_1 \\ B_2^T & L_2 & E \\ O & O & I \end{bmatrix} \begin{bmatrix} D_1 & O & I \\ O & D_2 & O \\ I & O & O \end{bmatrix} \begin{bmatrix} B_1 & B_2 & O \\ O & L_2^T & O \\ L_1^T & E^T & I \end{bmatrix}, \tag{10.34}$$

where

$$\begin{aligned} D_1 &= B_1^{-T} G_{11} B_1^{-1} - L_1^T B_1^{-1} - B_1^{-T} L_1, \\ D_2 &= L_2^{-1}(G_{22} - B_2^T D_1 B_2 - E B_2 - B_2^T E^T) L_2^{-T}, \\ E &= G_{21} B_1^{-1} - B_2^T D_1 - B_2^T L_1^T B_1^{-1}. \end{aligned}$$

A decomposition of the form (10.34) is known as a *Schilders factorization*. Dollar and Wathen (2004) show how such a factorization can be used to

https://doi.org/10.1017/S0962492904000212 Published online by Cambridge University Press

solve linear systems of the form $\mathcal{P}_c z_k = r_k$ efficiently at each iteration of an iterative method. Since (10.34) is an exact factorization of an approximation to $\mathcal{A}$, we can think of it as a rather special approximate factorization of the original coefficient matrix $\mathcal{A}$. An alternative factorization (for symmetric G), related to the null space method, is the following (Saunders 2005):

$$\mathcal{P}_c = \begin{bmatrix} I & O & O \\ B_2^T B_1^{-T} & I & O \\ O & O & I \end{bmatrix} \begin{bmatrix} G_{11} & X^T & B_1^T \\ X & Z^T G Z & O \\ B_1 & O & O \end{bmatrix} \begin{bmatrix} I & B_1^{-1} B_2 & O \\ O & I & O \\ O & O & I \end{bmatrix}, \quad (10.35)$$

where $Z = \begin{bmatrix} -B_1^{-1} B_2 \\ I \end{bmatrix}$ and $X^T = G_{12} - G_{11} B_1^{-1} B_2$. Approximate variants of (10.35) may be used to construct factorized preconditioners.

A more standard approach is to compute an *incomplete* factorization of the *exact* matrix $\mathcal{A}$:

$$\mathcal{A} = \mathcal{Q}^T \mathcal{L} \mathcal{D} \mathcal{L}^T \mathcal{Q} + \mathcal{R} \equiv \mathcal{P} + \mathcal{R},$$

where $\mathcal{R}$ represents a remainder matrix that contains terms that have been discarded in the course of the incomplete factorization. The size of $\mathcal{R}$ depends on the permutation matrix $\mathcal{Q}$ and on the dropping strategy used: although levels of fill can be used, threshold-based dropping is likely to be far more effective for indefinite systems. Unfortunately, the development of reliable and effective incomplete LDL^T factorizations for highly indefinite systems has turned out to be a very difficult problem, and until recently not much progress has been made in this area. This is in contrast with the symmetric positive definite case and even with the general, nonsymmetric case, for which many successful techniques exist (Axelsson 1994, Benzi 2002, Meurant 1999, Saad 2003, van der Vorst 2003).

A possible explanation for the difficulties encountered in developing reliable incomplete factorizations for indefinite systems has been offered by van der Vorst (Dongarra, Duff, Sorensen and van der Vorst 1998, pages 198–199). When the matrix $\mathcal{A}$ is highly indefinite, it has many eigenvalues on both sides of the imaginary axis. The eigenvalues of the preconditioned matrix $\mathcal{P}^{-1}\mathcal{A}$ depend continuously on the entries of $\mathcal{P}$. For $\mathcal{P} = I$, they coincide with the eigenvalues of $\mathcal{A}$; for $\mathcal{P} = \mathcal{A}$, they are all equal to 1. As the preconditioner $\mathcal{P}$ approaches $\mathcal{A}$, for instance by allowing more and more fill-in in the incomplete factors, the eigenvalues of $\mathcal{P}^{-1}\mathcal{A}$ approach the value 1; in particular, the negative eigenvalues, *en route* to 1, have to cross the imaginary axis. There is a high risk that some (perhaps many) of these eigenvalues will come arbitrarily close to the origin. Hence, 'improving' the preconditioner (by allowing additional fill-in) may actually cause the preconditioned matrix to become very close to singular, which in turn may cause the preconditioned iteration to converge more slowly or even fail.

https://doi.org/10.1017/S0962492904000212 Published online by Cambridge University Press

This non-monotonic behaviour of incomplete factorization preconditioners with respect to fill-in has been often observed in numerical experiments with indefinite matrices. Moreover, serious difficulties are often met in the course of computing an incomplete factorization, owing to various types of instabilities originating from the highly indefinite and non-diagonally dominant nature of saddle point matrices (Chow and Saad 1997*b*).

It is plausible that using the alternative, nonsymmetric positive definite form (3.10) of the saddle point system may alleviate this problem, particularly when $C \neq O$; however, we are not aware of any experimental studies in this direction. Another possibility would be to ignore the structure or symmetry of $\mathcal{A}$ altogether and apply one of the numerous, time-tested incomplete LU (ILU) factorization algorithms that have been developed for general sparse matrices, combined with some form of pivoting to promote stability. However, since fill-in tends to be very heavy with the original ordering of $\mathcal{A}$, large numbers of fill-ins have to be discarded, often resulting in preconditioners of low quality. Band-reducing and sparsity-preserving symmetric orderings (such as reverse Cuthill–McKee, minimum degree or nested dissection; see Duff *et al.* (1986)) are of limited use here and often produce unstable pivot sequences.

Some degree of success has been achieved through the use of nonsymmetric permutations and scalings aimed at increasing the diagonal dominance of the coefficient matrix; see Benzi, Haws and Tůma (2000), Haws (2002) and Haws and Meyer (2003), where the HSL MC64 preprocessing subroutines (Duff and Koster 1999, 2001) have been used in combination with various ILU and sparse approximate inverse techniques. When applied to saddle point matrices with $C = O$, these permutations produce a matrix with a zero-free diagonal; moreover, the zero (or small) diagonal entries are replaced by (relatively) large nonzeros. The net effect of this preprocessing is that stable ILU-type factorizations can now be computed in most cases. Unfortunately, this preprocessing destroys the symmetry and other structural properties of $\mathcal{A}$, and may lead to unnecessarily high storage demands in some cases.

Very recently, some new approaches have been introduced for constructing sparse incomplete LDL^T factorizations of symmetric indefinite matrices (Fish and Qu 2000, Freund 1997, Hagemann and Schenk 2004, Li and Saad 2004, Qu 2000, Qu and Fish 2000, Qu and Fish 2002). While not necessarily targeted to indefinite matrices in saddle point form, these techniques may be quite effective when applied to such systems; see the results reported in Hagemann and Schenk (2004) and Li and Saad (2004). In particular, maximum weighted matching techniques have been used in Hagemann and Schenk (2004) to minimize the need for dynamic (Bunch-type (Bunch 1974)) pivoting in the course of the factorization (Duff and Pralet 2004). Such matching techniques can be regarded as symmetric analogues of the non-

https://doi.org/10.1017/S0962492904000212 Published online by Cambridge University Press

symmetric permutations used in Duff and Koster (1999, 2001) to permute large entries to the diagonal. Here the purpose is to preprocess the matrix in a symmetry-preserving fashion with the goal of creating 'large' 2×2 diagonal blocks, thus enabling a more stable (incomplete) LDL^T factorization. Li and Saad (2004) instead promote stability by a different kind of preprocessing, referred to as *Bunch equilibration*: see Bunch (1971). These new preconditioners appear to be reasonably efficient and robust. In particular, one of the approaches presented in Hagemann and Schenk (2004) appears to be especially well suited to saddle point matrices. Since no comparison between these general-purpose methods and more specialized preconditioners for saddle point problems has been carried out yet, it is not yet clear how competitive this class of preconditioners really is. Nevertheless, in the engineering community there is certainly strong interest in general-purpose techniques (and software) that can be applied with little or no modification to a wide range of problems, even though such preconditioners may not be the best for any particular problem.

Some work has been done on developing incomplete factorization preconditioners specifically for saddle point matrices. Unlike general-purpose methods, these approaches take into account the block structure of $\mathcal{A}$. One approach is based on the following observation (Zhao 1998, Ren and Zhao 1999). If $A = A^T$ is positive definite, $B_1 = B_2 = B$ is of full rank and $C = C^T$ is positive semidefinite (possibly zero), then

$$\mathcal{A} = \begin{bmatrix} A & B^T \\ B & -C \end{bmatrix} = \begin{bmatrix} L_{11} & O \\ L_{21} & L_{22} \end{bmatrix} \begin{bmatrix} L_{11}^T & L_{21}^T \\ O & -L_{22}^T \end{bmatrix} = \mathcal{L}\mathcal{U}, \tag{10.36}$$

where $A = L_{11}L_{11}^T$ is the Cholesky factorization of A, $L_{21} = BL_{11}^{-T}$, and $-S = C + L_{21}L_{21}^T = L_{22}L_{22}^T$ is the Cholesky factorization of the (negative) Schur complement. Note that (10.36) is a triangular, Cholesky-like factorization of $\mathcal{A}$; such a factorization always exists, without the need for pivoting (*cf.* Section 7). If we could use $\mathcal{P} = \mathcal{L}\,\mathcal{L}^T$ as a (split) preconditioner, it is easily verified that

$$\mathcal{L}^{-1}\mathcal{A}\mathcal{L}^{-T} = \begin{bmatrix} I & O \\ O & -I \end{bmatrix},$$

and a symmetric Krylov subspace algorithm like MINRES or SYMMLQ would deliver the solution in at most two iterations. In practice, the exact factor $\mathcal{L}$ is replaced by an incomplete one, as follows. First, an incomplete Cholesky factorization $A \approx \bar{L}_{11}\bar{L}_{11}^T$ is computed; several stable and efficient algorithms exist for this task. Next, we compute a sparse approximation $\bar{L}_{21}$ to L_{21} by solving a matrix equation of the form $\bar{L}_{11}X = B^T$ using back-substitutions; some form of dropping may be needed to preserve sparsity in $\bar{L}_{21}$. Finally, an incomplete Cholesky factorization $C + \bar{L}_{21}\bar{L}_{21}^T \approx \bar{L}_2\bar{L}_2^T$ is

https://doi.org/10.1017/S0962492904000212 Published online by Cambridge University Press

computed. The resulting incomplete factor

$$\bar{\mathcal{L}} = \begin{bmatrix} \bar{L}_{11} & O \\ \bar{L}_{21} & \bar{L}_{22} \end{bmatrix}$$

and its transpose can be used to define a positive definite, factorized preconditioner $\mathcal{P} = \bar{\mathcal{L}}\,\bar{\mathcal{L}}^T$ for a method like SYMMLQ or MINRES. Numerical experiments in Ren and Zhao (1999) indicate that this can be an effective preconditioning strategy. Moreover, the resulting incomplete Cholesky factorization of the Schur complement ($-S \approx \bar{L}_{22}\bar{L}_{22}^T$) can be used as a preconditioner for Uzawa's method, resulting in a significant acceleration. Clearly, this approach will only work well when the Schur complement S is sparse or can be well approximated by a sparse matrix; see the discussion in Section 10.1.3. This approach may be extended to nonsymmetric saddle point systems; in this case, however, the existence of the corresponding incomplete ILU factorizations may not be guaranteed. Using an ordering different from the standard one may help; see Vasconcelos and D'Almeida (1998) for a study of incomplete LU preconditioning of the discretized Navier–Stokes equations with a nodal ordering of the unknowns.

Finally, we mention that incomplete factorizations for special classes of nonsymmetric saddle point problems have been developed by Wille and collaborators for solving the linear systems arising from Newton's method applied to mixed finite element discretizations of the steady Navier–Stokes equations; see Dahl and Wille (1992), Wille and Loula (2002) and Wille, Staff and Loula (2003). Here A is nonsymmetric (but positive definite), B_1 may or may not be equal to B_2, and $C = O$. In this work, fill-in is allowed only in predetermined locations within the factors, corresponding to the pressure block – *i.e.*, in the (2,2) block of the factors. More precisely, the guiding principle is that fill-in is accepted at locations in the global matrix where the nodes belong to the same finite element. The resulting incomplete factorizations appear to be numerically stable, and satisfactory convergence rates are observed for problems in both 2D and 3D. In Wille *et al.* (2003), parallel implementation aspects are also discussed.

11. Multilevel methods

A survey on the numerical solution of saddle point systems would not be complete without some discussion of multilevel solvers. In this section we give a very brief overview of multigrid and domain decomposition methods for saddle point systems arising from PDE problems. The literature on this topic is vast and highly technical, the construction of multilevel solvers usually being tied to a particular discretization and to analytical properties of the coefficients at the continuous (infinite-dimensional) level. Since our main emphasis is on the linear algebra aspects of finite-dimensional saddle point

https://doi.org/10.1017/S0962492904000212 Published online by Cambridge University Press

problems, we refrain from going into details and mostly restrict ourselves to providing pointers to the literature.

11.1. Multigrid methods

As already mentioned repeatedly in the course of this survey, multigrid and, more generally, multilevel methods can be used in segregated approaches to solve reduced systems, as in Uzawa-type or block preconditioning schemes. In the incompressible flow setting these subproblems correspond to discretizations of scalar elliptic PDEs for the velocities and/or the pressure field, such as diffusion- or convection-diffusion-type problems. Multigrid methods are ideally suited for such problems, often achieving optimal computational complexity, in the sense that the number of operations scales linearly with the number of unknowns. In the context of preconditioning inexact solves are usually sufficient, and it may be enough to perform just one iteration (V-, W- or F-cycle). The choice of multigrid components (smoothers, coarse grid operators and intergrid transfer operators) is well understood for this type of PDEs, and a plethora of algorithmic and software tools is available to deal with a wide range of discretizations, properties of coefficients, and problem geometries. Standard treatments of multigrid methods can be found in Hackbusch (1985), Trottenberg, Oosterlee and Schüller (2001) and Wesseling (1992). For detailed discussions of the use of multigrid components in the solution of saddle point problems from finite element discretizations, the reader is referred to Elman *et al.* (2005*c*) and Turek (1999); see also the recent survey by Wesseling and Oosterlee (2001).

More challenging is the construction of *coupled* multilevel methods, *i.e.*, multilevel methods that are applied to the entire system $\mathcal{A}u = b$. This topic has seen extensive development (especially in the last 15 years) and there are now a number of coupled multilevel methods for incompressible flow problems (Braess and Dahmen 1999, Braess and Sarazin 1997, Brandt and Dinar 1979, Elman 1996, Pernice 2000, Vanka 1986, Verfürth 1984*b*, Wesseling 1992, Wesseling 2001, Wittum 1989, Wittum 1990), mixed formulations of second- and fourth-order elliptic equations and of Maxwell's equations (Arnold, Falk and Winther 1997, Braess, Dahmen and Wieners 1999, Hiptmair 1996, Hiptmair 1997, Hiptmair and Hoppe 1999, Hiptmair, Shiekofer and Wohlmuth 1996, Trottenberg *et al.* 2001, Vassilevski and Wang 1992), optimization (Dreyer, Maar and Schulz 2000), and parameter estimation problems (Ascher and Haber 2003). These multigrid schemes may be used alone or as preconditioners for Krylov subspace methods; see Trottenberg *et al.* (2001, pages 287–288) for a discussion of the *solver versus preconditioner* question for multigrid.

For simplicity we consider here the two-grid case only. Our discussion, although fairly general, is geared towards Stokes- and Oseen-type problems

https://doi.org/10.1017/S0962492904000212 Published online by Cambridge University Press

and is based on Deng, Piquet, Queutey and Visonneau (1996, Section 4) and Wesseling (2001, Chapter 7.6). Introducing subscripts h and H for the fine and the coarse grid discretization parameters (*e.g.*, $H = 2h$), we rewrite $\mathcal{A}\,u = b$ as $\mathcal{A}_h\,u_h = b_h$, and we use the H subscripts for the coarse grid problem. The restriction and prolongation operators are denoted $\mathcal{I}_H^h$ and $\mathcal{I}_h^H$, respectively. The so-called 'full approximation storage' (FAS) two-grid method consists of the following steps:

(i) perform ν_1 pre-smoothing steps on $\mathcal{A}_h\,u_h = b_h$;
(ii) compute the fine grid residual $r_h = b_h - \mathcal{A}_h\,u_h$;
(iii) apply the restriction to the fine grid residual, $r_H = \mathcal{I}_H^h r_h$;
(iv) solve the coarse grid problem $\mathcal{A}_H\,u_H = r_H$;
(v) add the correction, $u_h = u_h + \alpha\,\mathcal{I}_h^H u_H$;
(vi) perform ν_2 post-smoothing steps.

At step (v), the correction control parameter α is either held fixed (possibly $\alpha = 1$) or chosen adaptively so as to minimize an appropriate measure of the error. A truly multilevel algorithm (V-cycle) is obtained by recursively applying the above procedure in step (iv). There are many variants of the basic scheme just described.

The choice of the restriction and prolongation operators has been discussed at length in the above-mentioned monographs (*e.g.*, in Turek (1999, Chapter 3.4.2)) and will not be pursued here. More critical are the construction of the coarse grid operator $\mathcal{A}_H$ and the choice of the coarse grid solver in step (iv), and the choice of the smoother in steps (i) and (vi). The matrix $\mathcal{A}_H$ can be constructed in at least two ways. One possibility is to simply rediscretize the problem on the coarse grid: this strategy, which is referred to as DCGA (discretization coarse grid approximation), may also be applied to nonlinear problems and is frequently used for the Navier–Stokes equations. An alternative is to compute $\mathcal{A}_H$ as $\mathcal{A}_H = \mathcal{I}_H^h \mathcal{A}_h \mathcal{I}_h^H$. This choice, which is referred to as GCGA (Galerkin coarse grid approximation) is restricted to linear problems, but has the advantage that it can be computed without knowledge of the underlying discretization. Hence, it is the preferred choice for algebraic multigrid (AMG) algorithms, which only exploit information contained in the coefficient matrix (Stüben 2001, Trottenberg *et al.* 2001). If the prolongation operator $\mathcal{I}_h^H$ is taken to be the transpose of the restriction operator $\mathcal{I}_H^h$, then the coarse grid operator $\mathcal{A}_H$ is symmetric if the fine grid one is. Preserving symmetry may be important if the multigrid iteration is to be accelerated by a symmetric Krylov subspace method. In this case the pre- and post-smoothing steps must also be performed in such a way as to preserve symmetry. The restriction operator is usually of the form

$$\mathcal{I}_H^h = \begin{bmatrix} I_H^h & O \\ O & \hat{I}_H^h \end{bmatrix},$$

https://doi.org/10.1017/S0962492904000212 Published online by Cambridge University Press

where I_H^h and $\hat{I}_H^h$ are restriction operators for the x and y unknowns, respectively. If we assume for simplicity that $B_1 = B_2 = B_h$ and $C_h = O$ and take $\mathcal{I}_h^H = (\mathcal{I}_H^h)^T$, the coarse grid matrix has the same form as the fine one,

$$\mathcal{A}_H = \begin{bmatrix} I_H^h & O \\ O & \hat{I}_H^h \end{bmatrix} \begin{bmatrix} A_h & B_h^T \\ B_h & O \end{bmatrix} \begin{bmatrix} I_h^H & O \\ O & \hat{I}_h^H \end{bmatrix} = \begin{bmatrix} A_H & B_H^T \\ B_H & O \end{bmatrix},$$

where $A_H = I_H^h A_h I_h^H$ and $B_H = \hat{I}_H^h B_h I_h^H$. Note that A_H is symmetric if A_h is.

The choice of the solver for the coarse grid problem is not as simple as it may appear at first. Only seldom can the grid be refined to the point that a direct method can be used to solve the coarse problem. This is because grid geometries can be highly irregular, thus allowing only a modest amount of coarsening. It is therefore necessary to resort to iterative methods and inexact coarse grid solves. In this case the multigrid method should be used as a preconditioner for an (outer) flexible Krylov subspace iteration in order to guarantee convergence. Alternatively, a hybrid nonlinear multigrid-inexact Newton method like the one proposed in Pernice (2000) can be used.

We now come to the choice of the smoother. To appreciate the importance of this component, it should be kept in mind that the smoothing steps is where most of the computational effort is usually spent, at least for geometric multigrid. Moreover, using the wrong smoother will destroy the efficiency of the entire multigrid algorithm. The definition of appropriate smoothing operations for saddle point problems is highly problem-dependent, and far from obvious. The smoothers used for scalar elliptic PDEs, such as diffusion- or convection-diffusion-type problems, are typically Gauss–Seidel or damped Jacobi relaxation, possibly combined with an appropriate renumbering of the grid points. These smoothers are not appropriate for saddle point problems, and they are not even defined when $C = O$. One approach is to use Richardson or Kaczmarz-type iterations, which is equivalent to using Jacobi or Gauss–Seidel smoothing on the normal equations (or squared system for symmetric problems): see Hackbusch (1994) and Verfürth (1984*b*). A better approach is to use some of the stationary iterations (or preconditioners) for saddle point problems discussed in Sections 8–10, but even this is not entirely straightforward.

The so-called *distributive relaxation schemes* (or *transforming iterations*) provide a unified framework for the description of smoothers for saddle point problems: see Brandt and Dinar (1979) and Wittum (1989, 1990). The idea is to transform the original problem $\mathcal{A}u = b$ by right preconditioning (often referred to as *postconditioning* in the multigrid literature) so that the transformed system

$$\mathcal{A}\mathcal{B}v = b, \qquad u = \mathcal{B}v$$

https://doi.org/10.1017/S0962492904000212 Published online by Cambridge University Press

is more easily solved. A splitting

$$\mathcal{A}\mathcal{B} = \mathcal{S} - \mathcal{T}$$

of the transformed matrix induces the following splitting of the original matrix $\mathcal{A}$:

$$\mathcal{A} = \mathcal{S}\,\mathcal{B}^{-1} - \mathcal{T}\,\mathcal{B}^{-1}.$$

This splitting defines a stationary iterative method for the original system,

$$\mathcal{S}\,\mathcal{B}^{-1}u_{k+1} = \mathcal{T}\,\mathcal{B}^{-1}u_k + b$$

or, in correction form,

$$u_{k+1} = u_k + \mathcal{B}\,\mathcal{S}^{-1}(b - \mathcal{A}\,u_k). \tag{11.1}$$

As long as $\mathcal{B}$ and $\mathcal{S}$ are nonsingular, the iteration (11.1) is consistent with $\mathcal{A}\,u = b$, in the sense that if it converges, it converges to the solution $u_* = \mathcal{A}^{-1}b$. The scheme (11.1) is called a *distributive relaxation* in the multigrid literature. The reason is that the correction $\mathcal{S}^{-1}(b - \mathcal{A}\,u_k)$ corresponding to non-distributive ($\mathcal{B} = I$) iterations is distributed over the entries of u_{k+1}; see Brandt and Dinar (1979) and Wesseling (2001, page 295).

A number of different distributive iterations can be obtained by special choices of $\mathcal{B}$ and $\mathcal{S}$. One possibility is to pick $\mathcal{B}$ such that $\mathcal{A}\mathcal{B}$ is block lower triangular. For instance, the block factorization (3.2) suggests the choice

$$\mathcal{B} = \begin{bmatrix} I & -A^{-1}B^T \\ O & I \end{bmatrix} \quad \Rightarrow \quad \mathcal{A}\mathcal{B} = \begin{bmatrix} A & O \\ B & S \end{bmatrix}.$$

This leads to a decoupling (segregation) of the x and y variables, and various choices of $\mathcal{S}$ lead essentially to preconditioned versions of Uzawa's method. A slightly more general form of $\mathcal{B}$ is

$$\mathcal{B} = \begin{bmatrix} I & -A^{-1}B^TQ \\ O & Q \end{bmatrix},$$

which leads to the transformed matrix

$$\mathcal{A}\mathcal{B} = \begin{bmatrix} A & O \\ B & M \end{bmatrix}, \qquad M = -BA^{-1}B^TQ.$$

A variety of methods result from the choice of Q and $\mathcal{S}$. For example, taking $Q = I$ and

$$\mathcal{S} = \begin{bmatrix} \hat{A} & O \\ B & \hat{S} \end{bmatrix},$$

where $\hat{A}$ and $\hat{S}$ are easily invertible approximations of A and of the Schur complement S, results in the SIMPLE scheme (Patankar and Spalding 1972, Patankar 1980) already described in the context of block preconditioning in Section 10.1.2. In the original SIMPLE scheme, $\hat{A} = D_A$ and $\hat{S} =$

https://doi.org/10.1017/S0962492904000212 Published online by Cambridge University Press

$-BD_A^{-1}B^T$, where D_A denotes the main diagonal of A. The matrix $\mathcal{B}$ used for the distribution step is further approximated by

$$\hat{\mathcal{B}} = \begin{bmatrix} I & -\hat{A}^{-1}B^T \\ O & I \end{bmatrix}.$$

Different choices of the approximations involved lead to variations on the original SIMPLE scheme (Patankar 1980, 1981, Shaw and Sivaloganathan 1988). Other options include the distributive Gauss–Seidel iteration and distributive ILU methods; see Brandt and Dinar (1979), Wittum (1989, 1990) and the nice description in Wesseling (2001, Chapter 7.6). For the Stokes problem, good results have been reported in Braess and Sarazin (1997) using as a smoother an iteration of the form (10.18), *i.e.*, a constraint-type preconditioner. This paper also discusses several variants of the SIMPLE scheme and shows that not all variants result in good smoothers. See also Zulehner (2000) for a generalization to the case $C \neq O$. Smoothers of the form (10.18) have also been used in connection with AMG methods for the Oseen problem in Wabro (2004), and in Leem *et al.* (2004) in the context of meshless discretizations. Recently, block triangular preconditioners of the kind discussed in Section 10.1 (so-called pressure convection-diffusion preconditioners) have been used with good results as smoothers for multigrid methods applied to the Oseen problem in Syamsudhuha and Silvester (2003). See Brandt (1998) for a discussion of the barriers that remain to be overcome before *textbook multigrid efficiency* (*i.e.*, solution of the governing equations in a number of operations not exceeding a small multiple of the nonzeros in the discretized equations) can be achieved in realistic computational fluid dynamics simulations. See also Thomas, Diskin and Brandt (2001) for recent progress in this direction.

All the smoothers discussed so far can be implemented in a segregated fashion, with each relaxation involving solution of decoupled linear systems for x and y. A strongly coupled smoothing procedure, not of the distributive type, has been introduced by Vanka (1986). The main principle of Vanka smoothing as applied to fluid flow problems is to visit each cell in some order and to apply relaxation simultaneously to each of the variables associated with that cell. This is equivalent to a reordering of the unknowns so that all the velocities and the pressures associated with each cell are numbered consecutively. Both multiplicative (Gauss–Seidel-type) and additive (Jacobi-type) variants have been developed. Vanka-type smoothing requires solving, at each step, a number of tiny saddle point problems, one for each cell. Various approximations and suitable damping (under-relaxation) are used to improve efficiency and to ensure the smoothing property. We refer to the original paper (Vanka 1986) as well as to Wesseling (2001, pages 300–301) for additional details. We note that this type of smoothing can also be

https://doi.org/10.1017/S0962492904000212 Published online by Cambridge University Press

interpreted as a form of domain decomposition, each subdomain reducing to a single grid cell.

Problems involving unstructured meshes on complicated geometries have motivated the development of AMG methods. AMG algorithms for scalar problems have been around for almost 20 years and have been used for some time in the solution of subproblems from segregated approaches; see, *e.g.*, Griebel, Neunhoeffer and Regler (1998), Stüben (2001) and Wagner, Kinzelbach and Wittum (1997). Coupled AMG methods for saddle point systems are barely 10 years old and are currently under development. We mention Raw (1995), Wabro (2004) and Webster (1994) for applications to incompressible flow problems, Adams (2004) for the development of AMG solvers in solid mechanics, and Leem *et al.* (2004) for AMG methods for saddle point systems arising from mesh-free discretizations. The reported results on difficult problems are promising. However, many open questions concerning coarsening strategies, the choice of appropriate smoothers, and parallel implementation issues remain to be addressed before coupled AMG solvers for saddle point problems can be considered fully mature.

Multilevel preconditioners for discontinuous Galerkin mixed finite element discretizations of radiation-diffusion problems on unstructured meshes have been developed in Warsa *et al.* (2004). Here the main idea is to use a continuous finite element approximation of the same problem to precondition the discontinuous one. Although only one mesh needs to be generated, suitable restriction and prolongation operators are needed to transfer information between the two discretizations. The continuous approximation can be solved efficiently by standard algorithms like AMG or preconditioned CG. This approach, which results in nearly optimal rates of convergence, is akin to an approximate Schur complement method. It is worth noting that in both Leem *et al.* (2004) and Warsa *et al.* (2004) the nonsymmetric formulation (3.10) was found to be advantageous, even though the original problems were symmetric.

Finally we mention some recent work on wavelet-based, *multiscale* methods. Originally developed for positive definite elliptic operators, these techniques have recently been extended to symmetric indefinite saddle point problems; see Dahlke, Hochmuth and Urban (2000), Hochmuth (1998) and Kunoth (2002). These are discretization schemes that produce saddle point systems that are well-conditioned, and can be efficiently solved by either Uzawa-type schemes or by diagonally preconditioned Krylov methods, resulting in convergence rates independent of discretization parameters.

11.2. Domain decomposition methods

Just as in the case of multigrid, domain decomposition methods can be applied straightforwardly to the (elliptic, scalar) subproblems arising from

https://doi.org/10.1017/S0962492904000212 Published online by Cambridge University Press

segregated approaches; see, *e.g.*, Fischer and Lottes (2004) for a recent example. Coupled algorithms have recently been developed by several authors, in particular for fluid flow and linear elasticity problems. Such methods are seldom (if ever) used as solvers: rather, they are used either as preconditioners for Krylov subspace methods, or as smoothers for multigrid, often motivated by parallel processing. We refer the reader to Chan and Mathew (1994) and Quarteroni and Valli (1999) for general treatments of domain decomposition methods, and to Toselli and Widlund (2004, Chapter 9) for an excellent discussion of methods designed specifically for saddle point problems. Because an up-to-date overview is already available, and also because of space limitations, we do not go into details and we largely limit ourselves to providing pointers to the literature.

Substructuring and additive Schwarz-type preconditioners for the symmetric saddle point problems arising from mixed finite element discretizations of second-order elliptic PDEs have been proposed and analysed in Rusten *et al.* (1996) and Rusten and Winther (1993). Different boundary conditions (Dirichlet or Neumann) on the interior interfaces are considered. The preconditioned systems are symmetrizable (so that MINRES can be used), and have condition numbers that are independent of mesh size.

An algebraic additive Schwarz domain decomposition preconditioner for solving saddle point problems arising from mixed finite element simulations of stochastic PDEs modelling flow in heterogeneous media has been described in Cliffe, Graham, Scheichl and Stals (2000). In this paper the preconditioner is not applied to to the coupled saddle point problem, but rather to a reduced (SPD) system of the type (6.1), where Z is a suitably constructed solenoidal basis. The resulting solver exhibits good robustness with respect to problem parameters and almost optimal levels of parallel efficiency. See also Oswald (1998) for a multilevel preconditioner for systems of the form (6.1) arising in the context of Stokes problems.

Several substructuring and domain decomposition preconditioners exist for linear elasticity and Stokes problems. Overlapping Schwarz preconditioners for saddle point systems containing a nonzero (2,2) block that depends on a penalty parameter (as in (10.9)) were proposed by Klawonn and Pavarino (1998). These preconditioners require the solution of local saddle point problems on overlapping subdomains, plus the solution of a coarse-level saddle point problem. Numerical experiments show that these preconditioners result in convergence rates that are independent of the mesh size, the number of subdomains, and the penalty parameter. The additive preconditioners are scalable also in the sense of parallel efficiency. In Klawonn and Pavarino (2000), the additive Schwarz preconditioner proposed in Klawonn and Pavarino (1998) is compared experimentally with a block diagonal and with a block triangular preconditioner (both with inexact solves). The results indicate that the additive Schwarz preconditioner (used with

https://doi.org/10.1017/S0962492904000212 Published online by Cambridge University Press

GMRES) is superior to the other methods in terms of iteration counts and very often in terms of operation counts, with the block triangular preconditioner being occasionally slightly better in terms of operation count. A convergence analysis for the additive Schwarz preconditioner, however, is still lacking. The preconditioned matrices have complex eigenvalues and an indefinite symmetric part. On the other hand, they appear to be positive stable with eigenvalues contained in a rectangular region of the right half-plane bounded away from zero, which explains in part the good convergence properties of GMRES. In Pavarino (2000) the overlapping Schwarz preconditioner is applied to mixed spectral or finite element discretizations of time-dependent Stokes problems, again with excellent results.

A different approach, based on iterative substructuring, was introduced and analysed by Pavarino and Widlund (2000) in the context of spectral element mixed approximations. A related balancing Neumann–Neumann approach for the Stokes problem is proposed in Pavarino and Widlund (2002), and extended to heterogeneous problems from linear elasticity in Goldfeld, Pavarino and Widlund (2003). This is a non-overlapping hybrid domain decomposition method in which the local problems are treated additively while the coarse problem is treated multiplicatively. Theoretical and numerical evidence show that these methods have excellent convergence and scalability properties; see in particular Goldfeld (2000) for numerical experiments with problems with up to 100 million unknowns using parallel machines with up to 2000 processors. We further mention recent work by Li on FETI methods for incompressible flow problems; see Li (2001, 2002*a*, 2002*b*).

Another domain decomposition method for linear elasticity problems was introduced by Klawonn and Widlund (2000). This preconditioner uses inexact subdomain solves and Lagrange multipliers and is based on a reformulation of the popular FETI method as a saddle point problem with both primal and dual variables as unknowns. It is shown in Klawonn and Widlund (2000) that the condition number of the preconditioned matrices is bounded independently of the number of subdomains and grows only polylogarithmically with the number of unknowns in each subdomain. See also Klawonn and Widlund (2001), and, for recent developments, Klawonn and Widlund (2004). We also mention the overlapping additive and multiplicative two-level Schwarz methods proposed in Wang (2005) for plane elasticity problems, which exhibit convergence rates independent of mesh size.

In Ainsworth and Sherwin (1999), p and h-p finite element discretizations of the Stokes problem are considered. Both segregated (block diagonal) and coupled (indefinite) preconditioners of the additive Schwarz type are studied. For each type of scheme, eigenvalue bounds for the preconditioned matrices are derived in terms of p (the polynomial degree), the fine and coarse mesh sizes, and the inf-sup constant for the method. These estimates

https://doi.org/10.1017/S0962492904000212 Published online by Cambridge University Press

and the results of actual computations on representative problems show that the rate of convergence for both methods does not deteriorate significantly as the mesh is refined and the polynomial degree is increased.

We note that while there is a rich literature on domain decomposition methods for symmetric saddle point problems and in particular for mixed discretizations of the Stokes equations, there seems to be very few papers concerned with coupled domain decomposition methods for nonsymmetric saddle point systems and in particular for the Oseen equations. One such paper is Vainikko and Graham (2004), where domain decomposition methods are implemented and experimentally compared with (and found to be superior to) block triangular preconditioners for the Oseen problem. Theoretical understanding, however, is still largely lacking.

Substructuring preconditioners for saddle point problems arising from edge element discretizations of Maxwell's equation in 3D have been introduced and analysed by Hu and Zou (2003). These preconditioners are shown to result in nearly optimal convergence rates, with the condition number of the preconditioned matrices growing as the logarithm of the ratio between the subdomain diameter and the finite element mesh size.

Domain decomposition methods for saddle point systems arising from PDE-constrained optimal control problems have been studied in Heinkenschloss and Nguyen (2004) and Nguyen (2004). Numerical experiments indicate that the performance of these (overlapping and non-overlapping) preconditioners with respect to mesh size and number of subdomains is close to that of the corresponding domain decomposition preconditioners for scalar PDEs. Furthermore, the preconditioners proposed in Nguyen (2004) appear to be rather insensitive to control regularization parameters.

Finally, the application of additive Schwarz-type iterations as smoothers in coupled multigrid methods for saddle point problems has been studied in Schöberl and Zulehner (2003). These are additive variants of the already-mentioned (multiplicative) Vanka smoother (Vanka 1986). The additive smoothers are shown to be related to an inexact version of Uzawa's method. An analysis of the smoothing property in the symmetric case is also given. Numerical experiments indicate that the rates of convergence for the multiplicative Vanka-type smoothers are significantly better than for the additive smoothers, as one would expect; however, a theoretical analysis of the convergence and smoothing properties in the multiplicative case is still missing.

12. Available software

In spite of vigorous algorithmic and theoretical developments, the production of high-quality, widely accessible software for solving linear systems in saddle point form has been somewhat lagging. This may be in part

https://doi.org/10.1017/S0962492904000212 Published online by Cambridge University Press

a reflection of the fact that many saddle point systems arising in the applications require a customized approach – for example, a particular Schur complement approximation. Another explanation is that when segregated approaches are used, standard sparse linear equations packages are often sufficient to solve the reduced linear systems. For example, many of the existing algorithms for saddle point problems from incompressible flow calculations require the availability of efficient solvers for diffusion- or convection-diffusion-type problems, and can utilize a number of 'standard' multigrid codes. In other situations, sparse direct solvers may be used to solve the subproblems arising from segregated approaches. Because many general-purpose linear solver packages exist and are readily accessible, there has been no great demand for software specifically designed to solve linear systems in saddle point form. An exception is represented by the field of optimization, where the interior point revolution has sparked a strong need for reliable and efficient solvers for augmented systems.

As usual, the software situation tends to be better for direct solvers than for iterative ones. Professional implementations of sparse direct solvers for symmetric indefinite systems include the MA27/MA47/MA57 suite of codes from the Harwell Subroutine Library (HSL) and the PARDISO package from the University of Basel (Schenk and Gärtner 2004). Of the HSL codes, the one that is best suited for saddle point matrices is perhaps MA47, although MA57 appears to be the most efficient overall of the symmetric solvers among the HSL codes (Gould and Scott 2004). For further information on the HSL symmetric indefinite solvers, including licensing information, the reader is referred to the web page

`www.cse.clrc.ac.uk/nag/hsl/contents.shtml`

For information about the PARDISO package, see

`www.computational.unibas.ch/cs/scicomp/software/pardiso/`

Although not specifically designed with saddle point systems in mind, the out-of-core sparse symmetric indefinite code developed as part of the TAUCS package at Tel Aviv University is also of interest; see

`www.tau.ac.il/~stoledo/taucs/`

and the accompanying paper by Meshar and Toledo (2005).

Software tools for incorporating block preconditioners geared towards incompressible flow problems into Krylov subspace methods are now available in the Meros package, which is part of the Trilinos project from Sandia National Laboratories (Heroux, Bartlett, Howle, Hoekstra, Hu, Kolda, Lehoucq, Long, Pawlowski, Phipps, Salinger, Thornquist, Tuminaro, Willenbring, Williams and Stanley 2005); see

`http://software.sandia.gov/trilinos/`

https://doi.org/10.1017/S0962492904000212 Published online by Cambridge University Press

Many of the solvers within Trilinos may also be used as part of segregated approaches; for instance, multilevel solvers are part of the ML package, which is part of Trilinos.

Algebraic multigrid codes for incompressible flow problems, including Markus Wabro's AMuSE (Algebraic Multigrid for Stokes-type Equations) package, can be accessed from the Institute of Computational Mathematics at the Johannes Kepler University (Linz, Austria) web site

`www.numa.uni-linz.ac.at/Research/Projects/P14953.html`

A number of other freely available codes implementing a host of multigrid and domain decomposition methods can be downloaded from the MGNET web site; see

`www.mgnet.org/mgnet-codes.html`

We conclude this brief – and by no means complete – overview of available software with the IFISS ('Incompressible Flow Iterative Solvers Software') package developed by Howard Elman, Alison Ramage, David Silvester and Andy Wathen: see

`www.cs.umd.edu/~elman/ifiss/`

See also the book by Elman *et al.* (2005*c*). This is a finite element package that can be used to generate a variety of incompressible flow problems on both structured and unstructured 2D meshes. The code, written in MATLAB, allows the user to experiment with several preconditioner and solver options.

13. Concluding remarks

In this paper we have surveyed a large number of solution methods for solving linear systems in saddle point form, with a focus on iterative methods suitable for large and sparse problems. We have discussed classical algorithms based on Schur complement reduction, null space methods, triangular factorization and stationary iterations like the Arrow–Hurwicz and Uzawa schemes and their variants. We have further discussed the use of Krylov subspace methods and described a number of preconditioning techniques including block diagonal, block triangular, constraint and incomplete factorization preconditioners. We have reviewed a promising new approach based on the Hermitian and skew-Hermitian splitting which appears to be especially well suited to the Navier–Stokes equations in rotation form. We saw that many preconditioning techniques rely (either explicitly or implicitly) on the availability of good and inexpensive approximations of a Schur complement matrix; similar comments apply to other solvers, such as multigrid or domain decomposition methods. We have seen that very effective

https://doi.org/10.1017/S0962492904000212 Published online by Cambridge University Press

solvers exist for some classes of problems, including mixed finite element formulations of elliptic PDEs and Stokes problems. Furthermore, great strides have been made in recent years towards the development of effective preconditioners for strongly nonsymmetric problems, such as the Oseen equations with low viscosity.

In spite of this, many challenges remain. Effective preconditioners are yet to be developed for large classes of linear systems arising from interior point methods in constrained optimization. Some degree of success has been achieved in this area with various types of approximate Schur complement and constraint preconditioners, but there is plenty of room for improvement. For many saddle point systems arising from optimal control problems, the (1,1) block A is often symmetric indefinite. Although these systems may sometimes be transformed to reduced saddle point systems with a definite (1,1) block, this reduction is not always easy to accomplish and it may be necessary to deal with the unreduced system. Hence, there is a need for effective solvers that are able to handle saddle point matrices with an indefinite (1,1) block. While some of the techniques described in this paper may be adapted to this situation, there is a need for new approaches. Also needed is a thorough study of stopping criteria for iterative methods applied to saddle point problems. In particular, criteria that take into account the different nature of the unknowns x and y should be developed.

Space and time limitations did not allow us to give very detailed treatments of many interesting techniques. In particular, important subjects such as direct solvers and multilevel methods have been touched upon only very briefly. Sparse direct solvers have been around for a long time and are widely used in some areas (such as optimization), but there are still some difficulties to overcome before these methods can be considered fully mature for solving saddle point systems: see for instance Gould and Scott (2004), where it is mentioned that the symmetric HSL sparse direct solvers could not solve four out of sixty-one of the test problems considered there. The same holds true, of course (even more so!), for iterative solvers. The ever-increasing complexity and size of the linear systems to be solved is already making the use of iterative methods absolutely mandatory for many applications, thus requiring continually improving solvers and preconditioners, with scalability more than ever a central concern. In line with current trends, it is quite likely that the future will see many new and improved multilevel algorithms, both geometric and algebraic, for solving large-scale saddle point problems from PDEs and PDE-constrained optimization.

Acknowledgements

During the preparation of this paper we have benefited from the help and advice of numerous colleagues and friends. Many thanks to Peter Arbenz,

https://doi.org/10.1017/S0962492904000212 Published online by Cambridge University Press

Mario Arioli, Howard Elman, Chen Greif, Eldad Haber, Amnon Meir, Luca Pavarino, Miro Rozložník, Wil Schilders, Daniel Szyld and Andy Wathen for informative discussions and pointers to the literature. We are especially thankful to Jacek Gondzio, Jim Nagy, Maxim Olshanskii, Michael Saunders and Valeria Simoncini for their careful reading of the manuscript and for their helpful comments.

REFERENCES

R. Aboulaich and M. Fortin (1989), 'Iterative methods for the solution of Stokes equations', *Comput. Methods Appl. Mech. Engrg.* **75**, 317–324.

M. F. Adams (2004), 'Algebraic multigrid methods for constrained linear systems with applications to contact problems in solid mechanics', *Numer. Linear Algebra Appl.* **11**, 141–153.

M. Ainsworth and S. Sherwin (1999), 'Domain decomposition preconditioners for p and hp finite element approximation of Stokes equations', *Comput. Methods Appl. Mech. Engrg.* **175**, 243–266.

P. Alotto and I. Perugia (1999), 'Mixed finite element methods and tree–cotree implicit condensation', *Calcolo* **36**, 233–248.

P. Amestoy, I. S. Duff and C. Puglisi (1996), 'Multifrontal QR factorization in a multiprocessor environment', *Numer. Linear Algebra Appl.* **3**, 275–300.

R. Amit, C. A. Hall and T. A. Porsching (1981), 'An application of network theory to the solution of implicit Navier–Stokes difference equations', *J. Comput. Phys.* **40**, 183–201.

G. Ammar, C. Mehl and V. Mehrmann (1999), 'Schur-like forms for matrix Lie groups, Lie algebras and Jordan algebras', *Linear Algebra Appl.* **287**, 11–39.

P. Arbenz and R. Geus (2005), 'Multilevel preconditioned iterative eigensolvers for Maxwell eigenvalue problems', *Appl. Numer. Math.*, to appear.

P. Arbenz, R. Geus and S. Adam (2001), 'Solving Maxwell eigenvalue problems for accelerating cavities', *Phys. Rev. ST Accel. Beams* **4**, # 022001.

M. Arioli (2000), 'The use of QR factorization in sparse quadratic programming', *SIAM J. Matrix Anal. Appl.* **21**, 825–839.

M. Arioli and L. Baldini (2001), 'A backward error analysis of a null space algorithm in sparse quadratic programming', *SIAM J. Matrix Anal. Appl.* **23**, 425–442.

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

M. Arioli and G. Manzini (2003), 'Null space algorithm and spanning trees in solving Darcy's equation', *BIT* **43**, 839–848.

M. Arioli, I. S. Duff and P. P. M. De Rijk (1989), 'On the augmented system approach to sparse least-squares problems', *Numer. Math.* **55**, 667–684.

M. Arioli, J. Maryška, M. Rozložník and Tůma (2001), Dual variable methods for mixed-hybrid finite element approximation of the potential fluid flow problem in porous media, Technical Report RAL-TR-2001-023, Rutherford Appleton Laboratory, RAL-CLRC, Oxfordshire, UK.

https://doi.org/10.1017/S0962492904000212 Published online by Cambridge University Press

D. N. Arnold, R. S. Falk and R. Winther (1997), 'Preconditioning in H(div) and applications', *Math. Comp.* **66**, 957–984.

W. E. Arnoldi (1951), 'The principle of minimized iteration in the solution of the matrix eigenvalue problem', *Quart. Appl. Math.* **9**, 17–29.

K. J. Arrow and L. Hurwicz (1958), Gradient method for concave programming I: Local results, in *Studies in Linear and Nonlinear Programming* (K. J. Arrow, L. Hurwicz and H. Uzawa, eds), Stanford University Press, Stanford, CA, pp. 117–126.

K. J. Arrow and R. M. Solow (1958), Gradient methods for constrained maxima, with weakened assumptions, in *Studies in Linear and Nonlinear Programming* (K. J. Arrow, L. Hurwicz and H. Uzawa, eds), Stanford University Press, Stanford, CA, pp. 166–176.

K. J. Arrow, L. Hurwicz and H. Uzawa (1958), *Studies in Linear and Nonlinear Programming*, Stanford University Press, Stanford, CA.

U. M. Ascher and E. Haber (2003), 'A multigrid method for distributed parameter estimation problems', *Electron. Trans. Numer. Anal.* **15**, 1–17.

C. Ashcraft, R. Grimes and J. G. Lewis (1998), 'Accurate symmetric indefinite linear equation solvers', *SIAM J. Matrix Anal. Appl.* **20**, 513–561.

G. P. Astrakhantsev (2001), 'Analysis of algorithms of the Arrow–Hurwicz type', *Comput. Math. Math. Phys.* **41**, 15–26.

G. M. Awanou and M. J. Lai (2005), 'On convergence rate of the augmented Lagrangian algorithm for nonsymmetric saddle point problems', *Appl. Numer. Math.*, to appear.

O. Axelsson (1979), 'Preconditioning of indefinite problems by regularization', *SIAM J. Numer. Anal.* **16**, 58–69.

O. Axelsson (1994), *Iterative Solution Methods*, Cambridge University Press, Cambridge.

O. Axelsson and M. Neytcheva (2003), 'Preconditioning methods for linear systems arising in constrained optimization problems', *Numer. Linear Algebra Appl.* **10**, 3–31.

O. Axelsson and P. S. Vassilevski (1991), 'A black-box generalized conjugate gradient solver with inner iterations and variable-step preconditioning', *SIAM J. Matrix Anal. Appl.* **12**, 625–644.

A. Baggag and A. Sameh (2004), 'A nested iterative scheme for indefinite linear systems in particulate flows', *Comput. Methods Appl. Mech. Engrg.* **193**, 1923–1957.

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

R. E. Bank, B. D. Welfert and H. Yserentant (1990), 'A class of iterative methods for solving saddle point problems', *Numer. Math.* **56**, 645–666.

J. L. Barlow (1988), 'Error analysis and implementation aspects of deferred correction for equality-constrained least squares problems', *SIAM J. Numer. Anal.* **25**, 1340–1358.

J. L. Barlow and S. L. Handy (1988), 'The direct solution of weighted and equality constrained least squares problems', *SIAM J. Sci. Statist. Comput.* **9**, 704–716.

https://doi.org/10.1017/S0962492904000212 Published online by Cambridge University Press

J. L. Barlow, N. K. Nichols and R. J. Plemmons (1988), 'Iterative methods for equality-constrained least squares problems', *SIAM J. Sci. Statist. Comput.* **9**, 892–906.

T. Barth and T. Manteuffel (1994), Variable metric conjugate gradient methods, in *Advances in Numerical Methods for Large Sparse Sets of Linear Equations: Proc. 10th International Symposium on Matrix Analysis and Parallel Computing, PCG 94* (M. Natori and T. Nodera, eds), pp. 165–188.

J. R. Batt and S. Gellin (1985), 'Rapid reanalysis by the force method', *Comput. Methods Appl. Mech. Engrg.* **53**, 105–117.

A. Battermann and M. Heinkenschloss (1998), Preconditioners for Karush–Kuhn–Tucker systems arising in the optimal control of distributed systems, in *Optimal Control of Partial Differential Equations* (W. Desch, F. Kappel and K. Kunisch, eds), Birkhäuser, pp. 15–32.

A. Battermann and E. W. Sachs (2001), Block preconditioners for KKT systems arising in PDE-governed optimal control problems, in *Fast Solution of Discretized Optimization Problems, Berlin 2000* (K.-H. Hoffmann, R. H. W. Hoppe and W. Schulz, eds), Birkhäuser, pp. 1–18.

J. Bear (1972), *Dynamics of Fluids in Porous Media*, American Elsevier, New York.

B. Beckermann and A. B. J. Kuijlaars (2001), 'Superlinear convergence of conjugate gradients', *SIAM J. Numer. Anal.* **39**, 300–329.

B. Beckermann and A. B. J. Kuijlaars (2002), 'Superlinear CG convergence for special right-hand sides', *Electron. Trans. Numer. Anal.* **14**, 1–19.

S. J. Benbow (1999), 'Solving generalized least-squares problems with LSQR', *SIAM J. Matrix Anal. Appl.* **21**, 166–177.

M. Benzi (1993), 'Solution of equality-constrained quadratic programming problems by a projection iterative method', *Rend. Mat. Appl.* (7) **13**, 275–296.

M. Benzi (2002), 'Preconditioning techniques for large linear systems: a survey', *J. Comput. Phys* **182**, 418–477.

M. Benzi (2004), HSS preconditioning for the Oseen problem, in *Proc. European Congress on Computational Methods in Applied Sciences and Engineering: ECCOMAS 2004, Jyväskylä, Finland* (P. Neittaanmäki, T. Rossi, S. Korotov, E. Oñate, J. Périaux and D. Knörzer, eds), CD-ROM.

M. Benzi and G. H. Golub (2004), 'A preconditioner for generalized saddle point problems', *SIAM J. Matrix Anal. Appl.* **26**, 20–41.

M. Benzi and J. Liu (2005), 'An efficient solver for the incompressible Navier–Stokes equations in rotation form', *J. Comput. Phys.*, submitted.

M. Benzi and M. K. Ng (2004), 'Preconditioned iterative methods for weighted Toeplitz least squares', *SIAM J. Matrix Anal. Appl.*, submitted.

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

M. Benzi, J. C. Haws and M. Tůma (2000), 'Preconditioning highly indefinite and nonsymmetric matrices', *SIAM J. Sci. Comput.* **22**, 1333–1353.

L. Bergamaschi, J. Gondzio and G. Zilli (2004), 'Preconditioning indefinite systems in interior point methods for optimization', *Comput. Optim. Appl.* **28**, 149–171.

https://doi.org/10.1017/S0962492904000212 Published online by Cambridge University Press

A. R. Bergen (1986), *Power Systems Analysis*, Prentice-Hall, Englewood Cliffs, NJ.
M. W. Berry and R. J. Plemmons (1987), 'Algorithms and experiments for structural mechanics on high-performance architectures', *Comput. Methods Appl. Mech. Engrg.* **64**, 487–507.
M. W. Berry, M. T. Heath, I. Kaneko, M. Lawo, R. J. Plemmons and R. C. Ward (1985), 'An algorithm for computing a sparse basis of the null space', *Numer. Math.* **47**, 483–504.
F. Bertrand and P. A. Tanguy (2002), 'Krylov-based Uzawa algorithms for the solution of the Stokes equations using discontinuous-pressure tetrahedral finite elements', *J. Comput. Phys.* **181**, 617–638.
D. P. Bertsekas (1982), *Constrained Optimization and Lagrange Multiplier Methods*, Academic Press, New York.
J. T. Betts (2001), *Practical Methods for Optimal Control Using Nonlinear Programming*, Advances in Design and Control, SIAM, Philadelphia, PA.
G. Biros and O. Ghattas (2000), Parallel preconditioners for KKT systems arising in optimal control of viscous incompressible flows, in *Proceedings of Parallel CFD '99, Williamsburg, VA, May 23–26, 1999* (D. Keyes, A. Ecer, J. Periaux and N. Satofuka, eds), North-Holland, pp. 1–8.
G. Biros and O. Ghattas (2005*a*), 'Parallel Lagrange–Newton–Krylov methods for PDE-constrained optimization, Part I: The Krylov-Schur solver', *SIAM J. Sci. Comput.*, to appear.
G. Biros and O. Ghattas (2005*b*), 'Parallel Lagrange–Newton–Krylov methods for PDE-constrained optimization, Part II: The Lagrange–Newton solver and its application to optimal control of steady viscous flows', *SIAM J. Sci. Comput.*, to appear.
L. Bitar and C. Vincent (2000), 'Eigenvalue upper bounds for the discretized Stokes operator', *Commun. Numer. Meth. Engng.* **16**, 449–457.
Å. Björck (1996), *Numerical Methods for Least Squares Problems*, SIAM, Philadelphia, PA.
Å. Björck and C. C. Paige (1994), 'Solution of augmented linear systems using orthogonal factorizations', *BIT* **34**, 1–24.
E. Y. Bobrovnikova and S. A. Vavasis (2000), 'A norm bound for projections with complex weights', *Linear Algebra Appl.* **307**, 69–75.
E. Y. Bobrovnikova and S. A. Vavasis (2001), 'Accurate solution of weighted least squares by iterative methods', *SIAM J. Matrix Anal. Appl.* **22**, 1153–1174.
P. Bochev and R. B. Lehoucq (2005), 'On finite element solution of the pure Neumann problem', *SIAM Rev.* **47**, 50–66.
A. Bojanczyk, N. J. Higham and H. Patel (2003), 'Solving the indefinite least squares problem by hyperbolic QR factorization', *SIAM J. Matrix Anal. Appl.* **24**, 914–931.
C. W. Bomhof and H. A. van der Vorst (2000), 'A parallel linear system solver for circuit-simulation problems', *Numer. Linear Algebra Appl.* **7**, 649–665.
I. Bongartz, A. R. Conn, N. I. M. Gould and P. L. Toint (1995), 'CUTE: Constrained and unconstrained testing environment', *ACM Trans. Math. Software* **21**, 123–160.
A. Bossavit (1998), ' "Mixed" systems of algebraic equations in computational electromagnetics', *COMPEL* **17**, 59–63.

https://doi.org/10.1017/S0962492904000212 Published online by Cambridge University Press

M. A. Botchev and G. H. Golub (2004), A class of nonsymmetric preconditioners for saddle point problems, Technical Report SCCM-04-14, Scientific Computing and Computational Mathematics, Stanford University, Stanford, California.

D. Braess (2001), *Finite Elements. Theory, Fast Solvers, and Applications in Solid Mechanics*, second edn, Cambridge University Press, Cambridge.

D. Braess and W. Dahmen (1999), 'A cascadic multigrid algorithm for the Stokes equations', *Numer. Math.* **82**, 179–191.

D. Braess and R. Sarazin (1997), 'An efficient smoother for the Stokes problem', *Appl. Numer. Math.* **23**, 3–19.

D. Braess, W. Dahmen and C. Wieners (1999), 'A multigrid algorithm for the mortar finite element method', *SIAM J. Numer. Anal.* **37**, 48–69.

J. H. Bramble and J. E. Pasciak (1988), 'A preconditioning technique for indefinite systems resulting from mixed approximations of elliptic problems', *Math. Comp.* **50**, 1–17.

J. H. Bramble and J. E. Pasciak (1997), 'Iterative techniques for time-dependent Stokes problems', *Comput. Math. Appl.* **33**, 13–30.

J. H. Bramble, J. E. Pasciak and A. T. Vassilev (1997), 'Analysis of the inexact Uzawa algorithm for saddle point problems', *SIAM J. Numer. Anal.* **34**, 1072–1092.

J. H. Bramble, J. E. Pasciak and A. T. Vassilev (2000), 'Uzawa type algorithms for nonsymmetric saddle point problems', *Math. Comp.* **69**, 667–689.

A. Brandt (1998), Barriers to achieving textbook multigrid efficiency (TME) in CFD, Gauss Center Technical Report WI-GC 10, Weizmann Institute of Science, Rehovot, Israel.

A. Brandt and D. Dinar (1979), Multigrid solutions to elliptic flow problems, in *Numerical Methods for Partial Differential Equations* (S. Parter, ed.), Academic Press, New York, pp. 53–147.

F. Brezzi (1974), 'On the existence, uniqueness, and approximation of saddle-point problems arising from Lagrangian multipliers', *RAIRO Anal. Numér.* **8**, 129–151.

F. Brezzi (2002), Stability of saddle-points in finite dimensions, in *Frontiers in Numerical Analysis* (J. F. Blowey, A. W. Craig and T. Shardlow, eds), Springer, Berlin, pp. 17–61.

F. Brezzi and M. Fortin (1991), *Mixed and Hybrid Finite Element Methods*, Vol. 15 of *Springer Series in Computational Mathematics*, Springer, New York.

D. C. Brown (1982), Alternating-direction iterative schemes for mixed finite element methods for second order elliptic problems, PhD thesis, Department of Mathematics, University of Chicago, Chicago, IL.

P. N. Brown (1991), 'A theoretical comparison of the Arnoldi and GMRES algorithms', *SIAM J. Sci. Statist. Comput.* **12**, 58–78.

J. R. Bunch (1971), 'Equilibration of symmetric matrices in the Max-norm', *J. Assoc. Comput. Mach.* **18**, 566–572.

J. R. Bunch (1974), 'Partial pivoting strategies for symmetric matrices', *SIAM J. Numer. Anal.* **11**, 521–528.

J. R. Bunch and L. Kaufman (1977), 'Some stable methods for calculating inertia and solving symmetric linear systems', *Math. Comp.* **31**, 163–179.

https://doi.org/10.1017/S0962492904000212 Published online by Cambridge University Press

J. R. Bunch and B. N. Parlett (1971), 'Direct methods for solving symmetric indefinite systems of linear equations', *SIAM J. Numer. Anal.* **8**, 639–655.

M. Burger and W. Mühlhuber (2002), 'Iterative regularization of parameter identification problems by sequential quadratic programming methods', *Inverse Problems* **18**, 943–969.

J. Burkardt, C. Hall and T. Porsching (1986), 'The dual variable method for the solution of compressible fluid flow problems', *SIAM J. Alg. Disc. Meth.* **7**, 476–483.

Y. V. Bychenkov (2002), 'Optimization of one class of nonsymmetrizable algorithms for saddle point problems', *Russian J. Numer. Anal. Math. Modelling* **17**, 521–546.

J. Cahouet and J. P. Chabard (1988), 'Some fast 3-D finite element solvers for the generalized Stokes problem', *Internat. J. Numer. Meth. Fluids* **8**, 869–895.

S. L. Campbell and C. D. Meyer Jr. (1979), *Generalized Inverses of Linear Transformations*, Pitman, London. Reprinted by Dover, New York (1991).

Z.-H. Cao (2002), 'A note on constraint preconditioning for nonsymmetric indefinite matrices', *SIAM J. Matrix Anal. Appl.* **24**, 121–125.

Z.-H. Cao (2004*a*), 'Fast iterative solution of stabilized Navier–Stokes systems', *Appl. Math. Comput.* **157**, 219–241.

Z.-H. Cao (2004*b*), 'Fast Uzawa algorithms for solving nonsymmetric stabilized saddle point problems', *Numer. Linear Algebra Appl.* **11**, 1–24.

A. Cassell, J. C. Henderson and A. Kaveh (1974), 'Cycle bases for the flexibility analysis of structures', *Internat. J. Numer. Methods Engng.* **8**, 521–528.

T. Chan and T. P. Mathew (1994), 'Domain decomposition algorithms', *Acta Numerica*, Vol. 3, Cambridge University Press, pp. 61–143.

S. Chandrasekaran, M. Gu and A. H. Sayed (1998), 'A stable and efficient algorithm for the indefinite linear least-squares problem', *SIAM J. Matrix Anal. Appl.* **20**, 354–362.

W. Chang, F. Giraldo and J. B. Perot (2002), 'Analysis of an exact fractional step method', *J. Comput. Phys.* **180**, 183–199.

X. Chen (1998), 'On preconditioned Uzawa methods and SOR methods for saddle-point problems', *J. Comput. Appl. Math.* **100**, 207–224.

X. Chen and K. Hashimoto (2003), 'Numerical validation of solutions of saddle point matrix equations', *Numer. Linear Algebra Appl.* **10**, 661–672.

Z. M. Chen and J. Zou (1999), 'An augmented Lagrangian method for identifying discontinuous parameters in elliptic systems', *SIAM J. Contr. Optim.* **37**, 892–910.

X. L. Cheng (2000), 'On the nonlinear inexact Uzawa algorithm for saddle-point problems', *SIAM J. Numer. Anal.* **37**, 1930–1934.

X. L. Cheng and J. Zou (2003), 'An inexact Uzawa-type iterative method for solving saddle point problems', *Internat. J. Computer Math.* **80**, 55–64.

E. V. Chizhonkov (2001), 'On solving an algebraic system of Stokes type under block diagonal preconditioning', *Comput. Math. Math. Phys.* **41**, 514–521.

E. V. Chizhonkov (2002), 'Improving the convergence of the Lanczos method in solving algebraic saddle point problems', *Comput. Math. Math. Phys.* **42**, 483–491.

https://doi.org/10.1017/S0962492904000212 Published online by Cambridge University Press

A. J. Chorin (1968), 'Numerical solution of the Navier–Stokes equations', *Math. Comp.* **22**, 745–762.

A. J. Chorin and J. E. Marsden (1990), *A Mathematical Introduction to Fluid Mechanics*, Vol. 4 of *Texts in Applied Mathematics*, Springer, New York.

E. Chow and Y. Saad (1997*a*), 'Approximate inverse techniques for block-partitioned matrices', *SIAM J. Sci. Comput.* **18**, 1657–1675.

E. Chow and Y. Saad (1997*b*), 'Experimental study of ILU preconditioners for indefinite matrices', *J. Comput. Appl. Math.* **86**, 387–414.

E. Chow, T. Manteuffel, C. Tong and B. K. Wallin (2003), 'Algebraic elimination of slide surface constraints in implicit structural analysis', *Internat. J. Numer. Methods Engng.* **57**, 1129–1144.

L. O. Chua, C. A. Desoer and E. S. Kuh (1987), *Linear and Nonlinear Circuits*, McGraw-Hill, New York.

P. Ciarlet Jr., J. Huang and J. Zou (2003), 'Some observations on generalized saddle-point problems', *SIAM J. Matrix Anal. Appl.* **25**, 224–236.

P. Ciarlet (1988), *Mathematical Elasticity, Vol. I: Three Dimensional Elasticity*, Vol. 20 of *Studies in Mathematics and its Applications*, North-Holland, Amsterdam.

K. A. Cliffe, T. J. Garratt and A. Spence (1994), 'Eigenvalues of block matrices arising from problems in fluid mechanics', *SIAM J. Matrix Anal. Appl.* **15**, 1310–1318.

K. A. Cliffe, I. G. Graham, R. Scheichl and L. Stals (2000), 'Parallel computation of flow in heterogeneous media modelled by mixed finite elements', *J. Comput. Phys.* **164**, 258–282.

T. F. Coleman (1984), *Large Sparse Numerical Optimization*, Springer, New York.

T. F. Coleman and A. Pothen (1986), 'The null space problem I: complexity', *SIAM J. Alg. Disc. Meth.* **7**, 527–537.

T. F. Coleman and A. Pothen (1987), 'The null space problem II: algorithms', *SIAM J. Alg. Disc. Meth.* **8**, 544–563.

T. F. Coleman and A. Verma (2001), 'A preconditioned conjugate gradient approach to linear equality constrained minimization', *Comput. Optim. Appl.* **20**, 61–72.

P. Concus and G. H. Golub (1976), A generalized conjugate gradient method for nonsymmetric systems of linear equations, in *Computing Methods in Applied Sciences and Engineering* (R. Glowinski and J. L. Lions, eds), Springer, pp. 56–65.

J. B. Conway (1990), *A Course in Functional Analysis*, Vol. 96 of *Graduate Texts in Mathematics*, second edn, Springer, New York.

R. Courant (1943), 'Variational methods for the solution of problems of equilibrium and vibration', *Bull. Amer. Math. Soc.* **69**, 1–23.

A. J. Cox and N. J. Higham (1999*a*), 'Accuracy and stability of the null space method for solving the equality constrained least squares problem', *BIT* **39**, 34–50.

A. J. Cox and N. J. Higham (1999*b*), 'Row-wise backward stable elimination methods for the equality constrained least squares problem', *SIAM J. Matrix Anal. Appl.* **21**, 313–326.

https://doi.org/10.1017/S0962492904000212 Published online by Cambridge University Press

M.-R. Cui (2002), 'A sufficient condition for the convergence of the inexact Uzawa algorithm for saddle point problems', *J. Comput. Appl. Math.* **139**, 189–196.

M.-R. Cui (2004), 'Analysis of iterative algorithms of Uzawa type for saddle point problems', *Appl. Numer. Math.* **50**, 133–146.

J. Cullum and A. Greenbaum (1996), 'Relations between Galerkin and norm-minimizing iterative methods for solving linear systems', *SIAM J. Matrix Anal. Appl.* **17**, 223–247.

J. Czyzyk, R. Fourer and S. Mehrotra (1998), 'Using a massively parallel processor to solve large sparse linear programs by an interior-point method', *SIAM J. Sci. Comput.* **19**, 553–565.

O. Dahl and S. Ø. Wille (1992), 'An ILU preconditioner with coupled node fill-in for iterative solution of the mixed finite element formulation of the 2D and 3D Navier–Stokes equations', *Internat. J. Numer. Meth. Fluids* **15**, 525–544.

S. Dahlke, R. Hochmuth and K. Urban (2000), 'Adaptive wavelet methods for saddle point problems', *Math. Model. Numer. Anal.* **34**, 1003–1022.

C. de Boor and J. R. Rice (1982), 'Extremal polynomials with application to Richardson iteration for indefinite linear systems', *SIAM J. Sci. Statist. Comput.* **3**, 47–57.

E. de Sturler and J. Liesen (2005), 'Block-diagonal and constraint preconditioners for nonsymmetric indefinite linear systems, Part I: Theory', *SIAM J. Sci. Comput.*, to appear.

S. Demko, W. F. Moss and P. W. Smith (1984), 'Decay rates for inverses of band matrices', *Math. Comp.* **44**, 491–499.

G. B. Deng, J. Piquet, P. Queutey and M. Visonneau (1996), Navier–Stokes equations for incompressible flows: finite-difference and finite volume methods, in *Handbook of Computational Fluid Mechanics* (R. Peyret, ed.), Academic Press, London, pp. 25–97.

P. Destuynder and T. Nevers (1990), 'Some numerical aspects of mixed finite elements for bending plates', *Comput. Methods Appl. Mech. Engrg.* **78**, 73–87.

M. Discacciati and A. Quarteroni (2004), 'Convergence analysis of a subdomain iterative method for the finite element approximation of the coupling of Stokes and Darcy equations', *Comput. Visual. Sci.* **6**, 93–103.

M. Discacciati, E. Miglio and A. Quarteroni (2002), 'Mathematical and numerical models for coupling surface and groundwater flows', *Appl. Numer. Math.* **43**, 57–74.

C. Dohrmann and R. B. Lehoucq (2004), A primal based penalty preconditioner for elliptic saddle point systems, Research Report SAND-2004-59644, Sandia National Laboratories, Albuquerque, New Mexico.

H. S. Dollar (2005), Extending constraint preconditioners for saddle point problems, Research Report NA-05/02, Numerical Analysis Group, Oxford University.

H. S. Dollar and A. J. Wathen (2004), Incomplete factorization constraint preconditioners for saddle-point matrices, Research Report NA-04/01, Numerical Analysis Group, Oxford University.

J. J. Dongarra, I. S. Duff, D. C. Sorensen and H. A. van der Vorst (1998), *Numerical Linear Algebra for High-Performance Computers*, Vol. 11 of *Software, Environments and Tools*, SIAM, Philadelphia, PA.

https://doi.org/10.1017/S0962492904000212 Published online by Cambridge University Press

J. Douglas Jr., R. Durán and P. Pietra (1986), Alternating-direction iteration for mixed finite element methods, in *Computing Methods in Applied Sciences and Engineering VII: Versailles, France, 1985*, Springer, New York, pp. 181–196.

J. Douglas Jr., R. Durán and P. Pietra (1987), Formulation of alternating-direction iteration for mixed methods in three-space, in *Numerical Approximation of Partial Differential Equations*, Vol. 133 of *Studies in Mathematics*, North-Holland, Amsterdam, pp. 21–30.

T. Dreyer, B. Maar and V. Schulz (2000), 'Multigrid optimization in applications', *J. Comput. Appl. Math.* **120**, 67–84.

F. Duchin and D. B. Szyld (1979), 'Application of sparse matrix techniques to inter-regional input-output analysis', *Economics of Planning* **15**, 142–167.

I. S. Duff (1994), The solution of augmented systems, in *Numerical Analysis 1993: Proc. 15th Dundee Conf., Dundee University, June–July 1993*, Vol. 303 of *Pitman Research Notes in Mathematics*, Longman Scientific & Technical, Harlow, UK, pp. 40–55.

I. S. Duff and J. Koster (1999), 'The design and use of algorithms for permuting large entries to the diagonal of sparse matrices', *SIAM J. Matrix Anal. Appl.* **20**, 889–901.

I. S. Duff and J. Koster (2001), 'An algorithm for permuting large entries to the diagonal of a sparse matrix', *SIAM J. Matrix Anal. Appl.* **22**, 973–996.

I. S. Duff and S. Pralet (2004), Strategies for scaling and pivoting for sparse symmetric indefinite problems, Research Report TR/PA/04/59, CERFACS, Toulouse, France.

I. S. Duff and J. K. Reid (1983), 'The multifrontal solution of indefinite sparse symmetric linear systems', *ACM Trans. Math. Software* **9**, 302–325.

I. S. Duff and J. K. Reid (1995), MA47, a FORTRAN code for direct solution of indefinite sparse symmetric linear systems, Research Report RAL-95-001, Rutherford Appleton Laboratory.

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

I. S. Duff, A. M. Erisman and J. K. Reid (1986), *Direct Methods for Sparse Matrices*, Monographs on Numerical Analysis, Oxford University Press, Oxford.

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

C. Durazzi and V. Ruggiero (2003*a*), 'Indefinitely preconditioned conjugate gradient method for large sparse equality and inequality constrained quadratic problems', *Numer. Linear Algebra Appl.* **10**, 673–688.

C. Durazzi and V. Ruggiero (2003*b*), 'Numerical solution of special linear and quadratic programs via a parallel interior-point method', *Parallel Comput.* **29**, 485–503.

E. D'Yakonov (1987), 'On iterative methods with saddle operators', *Soviet Math. Dokl.* **35**, 166–170.

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

https://doi.org/10.1017/S0962492904000212 Published online by Cambridge University Press

M. Eiermann and O. G. Ernst (2001), 'Geometric aspects of the theory of Krylov subspace methods', *Acta Numerica*, Vol. 10, Cambridge University Press, pp. 251–312.

M. Eiermann and W. Niethammer (1983), 'On the construction of semi-iterative methods', *SIAM J. Numer. Anal.* **20**, 1153–1160.

M. Eiermann, W. Niethammer and R. S. Varga (1985), 'A study of semi-iterative methods for nonsymmetric systems of linear equations', *Numer. Math.* **47**, 505–533.

H. C. Elman (1996), 'Multigrid and Krylov subspace methods for the discrete Stokes equations', *Internat. J. Numer. Meth. Fluids* **22**, 755–770.

H. C. Elman (1999), 'Preconditioning for the steady-state Navier–Stokes equations with low viscosity', *SIAM J. Sci. Comput.* **20**, 1299–1316.

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

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

H. C. Elman and D. J. Silvester (1996), 'Fast nonsymmetric iterations and preconditioning for Navier–Stokes equations', *SIAM J. Sci. Comput.* **17**, 33–46.

H. C. Elman, V. E. Howle, J. Shadid and R. Tuminaro (2003*a*), 'A parallel block multi-level preconditioner for the 3D incompressible Navier–Stokes equations', *J. Comput. Phys.* **187**, 504–523.

H. C. Elman, V. E. Howle, J. Shadid, R. Shuttleworth and R. Tuminaro (2005*a*), 'Block preconditioners based on approximate commutators', *SIAM J. Sci. Comput.*, to appear.

H. C. Elman, D. Loghin and A. J. Wathen (2003*b*), 'Preconditioning techniques for Newton's method for the incompressible Navier–Stokes equations', *BIT* **43**, 961–974.

H. C. Elman, A. R. Ramage, D. J. Silvester and A. J. Wathen (2005*b*). `http://www.cs.umd.edu/~elman/ifiss/`.

H. C. Elman, D. J. Silvester and A. J. Wathen (2002*a*), 'Block preconditioners for the discrete incompressible Navier–Stokes equations', *Internat. J. Numer. Methods Fluids* **40**, 333–344.

H. C. Elman, D. J. Silvester and A. J. Wathen (2002*b*), 'Performance and analysis of saddle point preconditioners for the discrete steady-state Navier–Stokes equations', *Numer. Math.* **90**, 665–688.

H. C. Elman, D. J. Silvester and A. J. Wathen (2005*c*), *Finite Elements and Fast Iterative Solvers*, Numerical Mathematics and Scientific Computation, Oxford University Press, Oxford.

O. G. Ernst (2000), 'Residual-minimizing Krylov subspace methods for stabilized discretizations of convection-diffusion equations', *SIAM J. Matrix Anal. Appl.* **21**, 1079–1101.

R. E. Ewing, R. D. Lazarov, P. Lu and P. S. Vassilevski (1990), Preconditioning indefinite systems arising from mixed finite element discretization of second-order elliptic problems, in *Preconditioned Conjugate Gradient Methods: Nijmegen 1989*, Vol. 1457 of *Lecture Notes in Mathematics*, Springer, Berlin, pp. 28–43.

https://doi.org/10.1017/S0962492904000212 Published online by Cambridge University Press

V. Faber and T. Manteuffel (1984), 'Necessary and sufficient conditions for the existence of a conjugate gradient method', *SIAM J. Numer. Anal.* **21**, 352–362.

V. Faber, W. Joubert, E. Knill and T. Manteuffel (1996), 'Minimal residual method stronger than polynomial preconditioning', *SIAM J. Matrix Anal. Appl.* **17**, 707–729.

C. Farhat and F.-X. Roux (1991), 'A method of finite element tearing and interconnecting and its parallel solution algorithm', *Internat. J. Numer. Methods Engng.* **32**, 1205–1227.

B. Fischer (1996), *Polynomial Based Iteration Methods for Symmetric Linear Systems*, Wiley–Teubner Series Advances in Numerical Mathematics, Wiley, Chichester, UK.

B. Fischer and F. Peherstorfer (2001), 'Chebyshev approximation via polynomial mappings and the convergence behaviour of Krylov subspace methods', *Electr. Trans. Numer. Anal.* **12**, 205–215.

B. Fischer, A. Ramage, D. J. Silvester and A. J. Wathen (1998), 'Minimum residual methods for augmented systems', *BIT* **38**, 527–543.

P. F. Fischer and J. W. Lottes (2004), Hybrid Schwarz-multigrid methods for the spectral element method: Extensions to Navier–Stokes, in *Domain Decomposition Methods in Science and Engineering* (R. Kornhuber, R. H. W. Hoppe, J. Périaux, O. Pironneau, O. B. Widlund and J. Xu, eds), Vol. 40 of *Lecture Notes in Computational Science and Engineering*, Springer, pp. 35–49.

J. Fish and Y. Qu (2000), 'Global basis two-level method for indefinite systems. Part 1: Convergence studies', *Internat. J. Numer. Methods Engng.* **49**, 439–460.

R. Fletcher (1976*a*), Conjugate gradient methods for indefinite systems, in *Numerical Analysis 1975: Proc. 6th Biennial Dundee Conf., Univ. Dundee, Dundee, 1975*, Vol. 506 of *Lecture Notes in Mathematics*, Springer, Berlin, pp. 73–89.

R. Fletcher (1976*b*), 'Factorizing symmetric indefinite matrices', *Linear Algebra Appl.* **14**, 257–272.

R. Fletcher (1987), *Practical Methods of Optimization*, 2nd edn, Wiley, Chichester, UK.

R. Fletcher and T. Johnson (1997), 'On the stability of null-space methods for KKT systems', *SIAM J. Matrix Anal. Appl.* **18**, 938–958.

A. Forsgren (1996), 'On linear least-squares problems with diagonally dominant weight matrices', *SIAM J. Matrix Anal. Appl.* **17**, 763–788.

A. Forsgren (2002), 'Inertia-controlling factorizations for optimization algorithms', *Appl. Numer. Math.* **43**, 91–107.

A. Forsgren and W. Murray (1993), 'Newton methods for large-scale linear equality-constrained minimization', *SIAM J. Matrix Anal. Appl.* **14**, 560–587.

A. Forsgren, P. E. Gill and J. R. Shinnerl (1996), 'Stability of symmetric ill-conditioned systems arising in interior methods for constrained optimization', *SIAM J. Matrix Anal. Appl.* **17**, 187–211.

M. Fortin (1993), 'Finite element solution of the Navier–Stokes equations', *Acta Numerica*, Vol. 2, Cambridge University Press, pp. 239–284.

https://doi.org/10.1017/S0962492904000212 Published online by Cambridge University Press

M. Fortin and A. Fortin (1985), 'A generalization of Uzawa's algorithm for the solution of the Navier–Stokes equations', *Comm. Appl. Numer. Meth.* **1**, 205–208.

M. Fortin and R. Glowinski (1983), *Augmented Lagrangian Methods: Applications to the Numerical Solution of Boundary Value Problems*, Vol. 15 of *Studies in Mathematics and its Applications*, North-Holland, Amsterdam. Translated from the French by B. Hunt and D. C. Spicer.

R. Fourer and S. Mehrotra (1993), 'Solving symmetric indefinite systems in an interior-point method for linear programming', *Math. Program.* **62**, 15–39.

A. Frangioni and C. Gentile (2004), 'New preconditioners for KKT systems of network flow problems', *SIAM J. Optim.* **14**, 894–913.

R. W. Freund (1993), 'A transpose-free quasi-minimal residual algorithm for non-Hermitian linear systems', *SIAM J. Sci. Comput.* **14**, 470–482.

R. W. Freund (1997), Preconditioning of symmetric, but highly indefinite linear systems, in *Proc. 15th IMACS World Congress on Scientific Computation, Modeling and Applied Mathematics* (A. Sydow, ed.), Wissenschaft und Technik Verlag, pp. 551–556.

R. W. Freund (2003), 'Model reduction methods based on Krylov subspaces', *Acta Numerica*, Vol. 12, Cambridge University Press, pp. 267–319.

R. W. Freund and F. Jarre (1996), 'A QMR-based interior-point algorithm for solving linear programs', *Math. Program.* **76**, 183–210.

R. W. Freund and N. M. Nachtigal (1991), 'QMR: a quasi-minimal residual method for non-Hermitian linear systems', *Numer. Math.* **60**, 315–339.

R. W. Freund and N. M. Nachtigal (1994), A new Krylov-subspace method for symmetric indefinite linear systems, in *Proc. 14th IMACS World Congress on Computational and Applied Mathematics* (W. F. Ames, ed.), IMACS, pp. 1253–1256.

R. W. Freund and N. M. Nachtigal (1995), 'Software for simplified Lanczos and QMR algorithms', *Appl. Numer. Math.* **19**, 319–341.

R. W. Freund, G. H. Golub and N. M. Nachtigal (1992), 'Iterative solution of linear systems', *Acta Numerica*, Vol. 1, Cambridge University Press, pp. 57–100.

W. N. Gansterer, J. Schneid and C. W. Ueberhuber (2003), Mathematical properties of equilibrium systems, Technical Report AURORA TR2003-13, University of Vienna and Vienna University of Technology.

A. Gauthier, F. Saleri and A. Quarteroni (2004), 'A fast preconditioner for the incompressible Navier–Stokes equations', *Comput. Visual. Sci.* **6**, 105–112.

A. George and K. Ikramov (2000), 'On the condition of symmetric quasi-definite matrices', *SIAM J. Matrix Anal. Appl.* **21**, 970–977.

A. George and K. Ikramov (2002), 'The closedness of certain classes of matrices with respect to pseudoinversion', *Comput. Math. Math. Phys.* **42**, 1242–1246.

A. George, K. Ikramov and B. Kucherov (2000), 'Some properties of symmetric quasi-definite matrices', *SIAM J. Matrix Anal. Appl.* **21**, 1318–1323.

J. Gilbert and M. T. Heath (1987), 'Computing a sparse basis for the null space', *SIAM J. Alg. Disc. Meth.* **8**, 446–459.

P. E. Gill, W. Murray and M. H. Wright (1981), *Practical Optimization*, Academic Press Inc. [Harcourt Brace Jovanovich Publishers], London.

https://doi.org/10.1017/S0962492904000212 Published online by Cambridge University Press

P. E. Gill, W. Murray, D. B. Ponceleón and M. A. Saunders (1992), 'Preconditioners for indefinite systems arising in optimization', *SIAM J. Matrix Anal. Appl.* **13**, 292–311.

P. E. Gill, W. Murray, M. A. Saunders and M. H. Wright (1991), 'Inertia-controlling methods for general quadratic programming', *SIAM Rev.* **33**, 1–36.

P. E. Gill, M. A. Saunders and J. R. Shinnerl (1996), 'On the stability of Cholesky factorization for symmetric quasidefinite systems', *SIAM J. Matrix Anal. Appl.* **17**, 35–46.

V. Girault and P. A. Raviart (1986), *Finite Element Approximation of the Navier–Stokes Equations*, Vol. 749 of *Lecture Notes in Mathematics*, Springer, New York.

R. Glowinski (1984), *Numerical Methods for Nonlinear Variational Problems*, Springer Series in Computational Physics, Springer, New York.

R. Glowinski (2003), *Finite Element Methods for Incompressible Viscous Flow*, Vol. IX of *Handbook of Numerical Analysis, Part 3: Numerical Methods for Fluids*, North-Holland, Amsterdam.

R. Glowinski and P. Le Tallec (1989), *Augmented Lagrangian and Operator Splitting Methods in Nonlinear Mechanics*, SIAM, Philadelphia, PA.

R. Glowinski and O. Pironneau (1979), 'Numerical methods for the first biharmonic equation and for the two-dimensional Stokes problem', *SIAM Rev.* **21**, 167–212.

I. Gohberg, P. Lancaster and L. Rodman (1983), *Matrices and Indefinite Scalar Products*, Vol. 8 of *Operator Theory: Advances and Applications*, Birkhäuser, Basel.

P. Goldfeld (2000), Balancing Neumann–Neumann preconditioners for mixed formulation of almost-incompressible linear elasticity, PhD thesis, Courant Institute of Mathematical Sciences, New York University, New York.

P. Goldfeld, L. F. Pavarino and O. B. Widlund (2003), 'Balancing Neumann–Neumann preconditioners for mixed approximations of heterogeneous problems in linear elasticity', *Numer. Math.* **95**, 283–324.

G. H. Golub and C. Greif (2003), 'On solving block-structured indefinite linear systems', *SIAM J. Sci. Comput.* **24**, 2076–2092.

G. H. Golub and M. L. Overton (1988), 'The convergence of inexact Chebyshev and Richardson iterative methods for solving linear systems', *Numer. Math.* **53**, 571–593.

G. H. Golub and C. F. Van Loan (1979), 'Unsymmetric positive definite linear systems', *Linear Algebra Appl.* **28**, 85–97.

G. H. Golub and C. F. Van Loan (1996), *Matrix Computations*, Johns Hopkins Studies in the Mathematical Sciences, third edn, Johns Hopkins University Press, Baltimore, MD.

G. H. Golub and D. Vanderstraeten (2000), 'On the preconditioning of matrices with skew-symmetric splittings', *Numer. Algorithms* **25**, 223–239.

G. H. Golub and A. J. Wathen (1998), 'An iteration for indefinite systems and its application to the Navier–Stokes equations', *SIAM J. Sci. Comput.* **19**, 530–539.

G. H. Golub, X. Wu and J.-Y. Yuan (2001), 'SOR-like methods for augmented systems', *BIT* **41**, 71–85.

https://doi.org/10.1017/S0962492904000212 Published online by Cambridge University Press

R. Gonzales and R. Woods (1992), *Digital Image Processing*, Addison-Wesley, New York.

M. V. Gorelova and E. V. Chizhonkov (2004), 'Preconditioning saddle point problems with the help of saddle point operators', *Comput. Math. Math. Phys.* **44**, 1445–1455.

N. I. M. Gould (1985), 'On practical conditions for the existence and uniqueness of solutions to the general equality quadratic programming problem', *Math. Program.* **32**, 90–99.

N. I. M. Gould and J. A. Scott (2004), 'A numerical evaluation of HSL packages for the direct solution of large sparse, symmetric linear systems of equations', *ACM Trans. Math. Software* **30**, 300–325.

N. I. M. Gould, M. E. Hribar and J. Nocedal (2001), 'On the solution of equality constrained quadratic programming problems arising in optimization', *SIAM J. Sci. Comput.* **23**, 1376–1395.

I. G. Graham, A. Spence and E. Vainikko (2003), 'Parallel iterative methods for Navier–Stokes equations and application to eigenvalue calculation', *Concurrency and Computation: Practice and Experience* **15**, 1151–1168.

A. Greenbaum (1979), 'Comparison of splittings used with the conjugate gradient algorithm', *Numer. Math.* **33**, 181–194.

A. Greenbaum (1997), *Iterative Methods for Solving Linear Systems*, Vol. 17 of *Frontiers in Applied Mathematics*, SIAM, Philadelphia, PA.

A. Greenbaum (2002), 'Generalizations of the field of values useful in the study of polynomial functions of a matrix', *Linear Algebra Appl.* **347**, 233–249.

A. Greenbaum and L. Gurvits (1994), 'Max-min properties of matrix factor norms', *SIAM J. Sci. Comput.* **15**, 348–358.

A. Greenbaum and Z. Strakoš (1994), Matrices that generate the same Krylov residual spaces, in *Recent Advances in Iterative Methods*, Vol. 60 of *IMA Vol. Math. Appl.*, Springer, New York, pp. 95–118.

A. Greenbaum, V. Pták and Z. Strakoš (1996), 'Any nonincreasing convergence curve is possible for GMRES', *SIAM J. Matrix Anal. Appl.* **17**, 465–469.

C. Greif, G. H. Golub and J. M. Varah (2005), 'Augmented Lagrangian techniques for solving saddle point linear systems', *SIAM J. Matrix Anal. Appl.*, to appear.

P. M. Gresho (1991), 'Incompressible fluid dynamics: some fundamental formulation issues', *Annual Rev. Fluid Mech.* **23**, 413–453.

P. M. Gresho and R. L. Sani (1998), *Incompressible Flow and the Finite Element Method. Advection-Diffusion and Isothermal Laminar Flow*, Wiley, Chichester, UK.

M. Griebel, T. Neunhoeffer and H. Regler (1998), 'Algebraic multigrid methods for the solution of the Navier–Stokes equations in complicated geometries', *Internat. J. Numer. Meth. Fluids* **26**, 281–301.

M. Gulliksson (1994), 'Iterative refinement for constrained and weighted linear least squares', *BIT* **34**, 239–253.

M. Gulliksson and P.-Å. Wedin (1992), 'Modifying the QR-decomposition to constrained and weighted linear least squares', *SIAM J. Matrix Anal. Appl.* **13**, 1298–1313.

https://doi.org/10.1017/S0962492904000212 Published online by Cambridge University Press

M. Gulliksson, X.-Q. Jin and Y. Wei (2002), 'Perturbation bounds for constrained and weighted least squares problems', *Linear Algebra Appl.* **349**, 221–232.
M. Gunzburger (1989), *Finite Element Methods for Viscous Incompressible Flows*, Academic Press, San Diego.
K. Gustafson and R. Hartmann (1983), 'Divergence-free bases for finite element schemes in hydrodynamics', *SIAM J. Numer. Anal.* **20**, 697–721.
E. Haber and U. M. Ascher (2001), 'Preconditioned all-at-once methods for large, sparse parameter estimation problems', *Inverse Problems* **17**, 1847–1864.
E. Haber and J. Modersitzki (2004), 'Numerical methods for volume-preserving image registration', *Inverse Problems* **20**, 1621–1638.
E. Haber, U. M. Ascher and D. Oldenburg (2000), 'On optimization techniques for solving nonlinear inverse problems', *Inverse Problems* **16**, 1263–1280.
W. Hackbusch (1985), *Multi-grid Methods and Applications*, Springer, Berlin.
W. Hackbusch (1994), *Iterative Solution of Large Sparse Systems of Equations*, Springer, Berlin.
G. Hadley (1964), *Nonlinear and Dynamic Programming*, Addison-Wesley, Reading, MA.
M. Hagemann and O. Schenk (2004), Weighted matchings for the preconditioning of symmetric indefinite linear systems, Technical Report CS-2004-005, Department of Computer Science, University of Basel.
C. A. Hall (1985), 'Numerical solution of Navier–Stokes problems by the dual variable method', *SIAM J. Alg. Disc. Meth.* **6**, 220–236.
C. A. Hall, J. C. Cavendish and W. H. Frey (1991), 'The dual variable method for solving fluid flow difference equations on Delaunay triangulations', *Comput. & Fluids* **20**, 145–164.
E. L. Hall (1979), *Computer Image Processing and Recognition*, Academic Press, New York.
F. H. Harlow and J. E. Welch (1965), 'Numerical calculation of time-dependent viscous incompressible flow of fluid with free surface', *Phys. Fluids* **8**, 2182–2189.
B. Hassibi, A. H. Sayed and T. Kailath (1996), 'Linear estimation in Krein spaces, Part I: Theory', *IEEE Trans. Automat. Control* **41**, 18–33.
J. C. Haws (2002), Preconditioning KKT systems, PhD thesis, Department of Mathematics, North Carolina State University, Raleigh, NC.
J. C. Haws and C. D. Meyer (2003), Preconditioning KKT systems, Technical Report M&CT-Tech-01-021, The Boeing Company.
M. T. Heath, R. J. Plemmons and R. C. Ward (1984), 'Sparse orthogonal schemes for structural optimization using the force method', *SIAM J. Sci. Statist. Comput.* **5**, 514–532.
M. Heinkenschloss and H. Nguyen (2004), Domain decomposition preconditioners for linear quadratic elliptic optimal control problems, Technical Report TR04–20, Department of Computational and Applied Mathematics, Rice University.
L. Hemmingsson-Frändén and A. Wathen (2001), 'A nearly optimal preconditioner for the Navier–Stokes equations', *Numer. Linear Algebra Appl.* **8**, 229–243.
J. C. Henderson and E. A. W. Maunder (1969), 'A problem in applied topology: on the selection of cycles for the flexibility analysis of skeletal structures', *J. Inst. Math. Appl.* **5**, 254–269.

https://doi.org/10.1017/S0962492904000212 Published online by Cambridge University Press

P. Heres and W. Schilders (2005), Practical issues of model order reduction with Krylov-subspace methods, in *Progress in Electromagnetic Research Symposium: PIERS 2004*. To appear.

M. A. Heroux, R. A. Bartlett, V. E. Howle, R. J. Hoekstra, J. J. Hu, T. G. Kolda, R. B. Lehoucq, K. R. Long, R. P. Pawlowski, E. T. Phipps, A. G. Salinger, H. K. Thornquist, R. S. Tuminaro, J. S. Willenbring, A. Williams and K. S. Stanley (2005), 'An overview of the Trilinos project', *ACM Trans. Math. Software*, to appear.

M. R. Hestenes (1969), 'Multiplier and gradient methods', *J. Optim. Theory Appl.* **4**, 303–320.

M. R. Hestenes (1975), *Optimization Theory: The Finite Dimensional Case*, Wiley, New York.

M. R. Hestenes and E. Stiefel (1952), 'Methods of conjugate gradients for solving linear systems', *J. Research Nat. Bur. Standards* **49**, 409–436.

R. Hiptmair (1996), *Multilevel Preconditioning for Mixed Problems in Three Dimensions*, Vol. 8 of *Augsburger Mathematisch-Naturwissenschaftliche Schriften*, Wissner, Augsburg, Germany.

R. Hiptmair (1997), 'Multigrid method for H(div) in three dimensions', *Electron. Trans. Numer. Anal.* **6**, 133–152.

R. Hiptmair and R. H. W. Hoppe (1999), 'Multilevel methods for mixed finite elements in three dimensions', *Numer. Math.* **82**, 253–279.

R. Hiptmair, T. Shiekofer and B. Wohlmuth (1996), 'Multilevel preconditioned augmented Lagrangian techniques for 2nd order mixed problems', *Comput.* **57**, 25–48.

R. Hochmuth (1998), 'Stable multiscale discretization for saddle point problems and preconditioning', *Numer. Funct. Anal. Optim.* **19**, 789–806.

R. A. Horn and C. R. Johnson (1985), *Matrix Analysis*, Cambridge University Press, Cambridge.

P. D. Hough and S. D. Vavasis (1997), 'Complete orthogonal decomposition for weighted least squares', *SIAM J. Matrix Anal. Appl.* **18**, 369–392.

Q. Hu and J. Zou (2001), 'An iterative method with variable relaxation parameters for saddle-point problems', *SIAM J. Matrix Anal. Appl.* **23**, 317–338.

Q. Hu and J. Zou (2002), 'Two new variants of nonlinear inexact Uzawa algorithms for saddle-point problems', *Numer. Math.* **93**, 333–359.

Q. Hu and J. Zou (2003), 'Substructuring preconditioners for saddle-point problems arising from Maxwell's equations in three dimensions', *Math. Comp.* **73**, 35–61.

Q. Hu, Z. Shi and D. Yu (2004), 'Efficient solvers for saddle-point problems arising from domain decompositions with Lagrange multipliers', *SIAM J. Numer. Anal.* **42**, 905–933.

M. Huhtanen (2002), 'A Hermitian Lanczos method for normal matrices', *SIAM J. Matrix Anal. Appl.* **23**, 1092–1108.

K. D. Ikramov (2003), 'On a certain type of preconditioning for KKT-matrices', *Moscow Univ. Comput. Math. Cybernet.* **2**, 1–3.

I. C. F. Ipsen (2001), 'A note on preconditioning nonsymmetric matrices', *SIAM J. Sci. Comput.* **23**, 1050–1051.

https://doi.org/10.1017/S0962492904000212 Published online by Cambridge University Press

K. Ito and K. Kunisch (1999), 'An active set strategy based on the augmented Lagrangian formulation for image restoration', *Math. Model. Numer. Anal.* **33**, 1–21.

D. James (1992), 'Implicit null space iterative methods for constrained least squares problems', *SIAM J. Matrix Anal. Appl.* **13**, 962–978.

D. James and R. J. Plemmons (1990), 'An iterative substructuring algorithm for equilibrium equations', *Numer. Math.* **57**, 625–633.

X.-Q. Jin (1996), 'A preconditioner for constrained and weighted least squares problems with Toeplitz structure', *BIT* **36**, 101–109.

W. Joubert (1994), 'A robust GMRES-based adaptive polynomial preconditioning algorithm for nonsymmetric linear systems', *SIAM J. Sci. Comput.* **15**, 427–439.

J. Júdice, J. Patricio, L. Portugal, M. Resende and G. Veiga (2003), 'A study of preconditioners for network interior point methods', *Comput. Optim. Appl.* **24**, 5–35.

X. Juvigny (1997), Solution of large linear systems on massively parallel machines, PhD thesis, Applied Mathematics, University of Paris VI, Paris, France. In French.

I. Kaneko and R. J. Plemmons (1984), 'Minimum norm solutions to linear elastic analysis problems', *Internat. J. Numer. Methods Engng.* **20**, 983–998.

I. Kaneko, M. Lawo and G. Thierauf (1982), 'On computational procedures for the force method', *Internat. J. Numer. Methods Engng.* **18**, 1469–1495.

G. Kanschat (2003), 'Preconditioning methods for local discontinuous Galerkin discretizations', *SIAM J. Sci. Comput.* **25**, 815–831.

O. A. Karakashian (1982), 'On a Galerkin–Lagrange multiplier method for the stationary Navier–Stokes equations', *SIAM J. Numer. Anal.* **19**, 909–923.

A. Kaveh (1974), Application of topology and matroid theory to the analysis of structures, PhD thesis, Imperial College of Science and Technology, London.

A. Kaveh (1979), 'A combinatorial optimization problem: optimal generalized cycle bases', *Comput. Methods Appl. Mech. Engrg.* **20**, 39–51.

A. Kaveh (1992), 'Recent developments in the force method of structural analysis', *Appl. Mech. Rev.* **45**, 401–418.

A. Kaveh (2004), *Structural Mechanics: Graph and Matrix Methods*, third edn, Research Studies Press, UK.

D. Kay, D. Loghin and A. J. Wathen (2002), 'A preconditioner for the steady-state Navier–Stokes equations', *SIAM J. Sci. Comput.* **24**, 237–256.

C. Keller, N. I. M. Gould and A. J. Wathen (2000), 'Constraint preconditioning for indefinite linear systems', *SIAM J. Matrix Anal. Appl.* **21**, 1300–1317.

G. R. Kirchhoff (1847), 'Über die Auflösung der Gleichungen, auf welche man bei der Untersuchung der linearen Verteilung galvanischer Ströme geführt wird', *Ann. Phys. Chem.* **72**, 497–508.

T. H. Kjeldsen (2000), 'A contextualized historical analysis of the Kuhn–Tucker theorem in nonlinear programming: the impact of World War II', *Historia Mathematica* **27**, 331–361.

A. Klawonn (1998*a*), 'Block-triangular preconditioners for saddle point problems with a penalty term', *SIAM J. Sci. Comput.* **19**, 172–184.

https://doi.org/10.1017/S0962492904000212 Published online by Cambridge University Press

A. Klawonn (1998*b*), 'An optimal preconditioner for a class of saddle point problems with a penalty term', *SIAM J. Sci. Comput.* **19**, 540–552.

A. Klawonn and L. F. Pavarino (1998), 'Overlapping Schwarz methods for mixed linear elasticity and Stokes problems', *Comput. Methods Appl. Mech. Engrg.* **165**, 233–245.

A. Klawonn and L. F. Pavarino (2000), 'A comparison of overlapping Schwarz methods and block preconditioners for saddle point problems', *Numer. Linear Algebra Appl.* **7**, 1–25.

A. Klawonn and G. Starke (1999), 'Block triangular preconditioners for nonsymmetric saddle point problems: field-of-values analysis', *Numer. Math.* **81**, 577–594.

A. Klawonn and O. B. Widlund (2000), 'A domain decomposition method with Lagrange multipliers and inexact solves for linear elasticity', *SIAM J. Sci. Comput.* **22**, 1199–1219.

A. Klawonn and O. B. Widlund (2001), 'FETI and Neumann–Neumann iterative substructuring methods: Connections and new results', *Comm. Pure Appl. Math.* **54**, 57–90.

A. Klawonn and O. B. Widlund (2004), Dual–primal FETI methods for linear elasticity, Technical Report TR2004-855, Courant Institute of Mathematical Sciences, New York University.

G. M. Kobelkov and M. A. Olshanskii (2000), 'Effective preconditioning of Uzawa type schemes for a generalized Stokes problem', *Numer. Math.* **86**, 443–470.

L. Komzsik, P. Poschmann and I. Sharapov (1997), 'A preconditioning technique for indefinite linear systems', *Finite Elements Anal. Des.* **26**, 253–258.

J. Korzak (1999), 'Eigenvalue relations and conditions of matrices arising in linear programming', *Comput.* **62**, 45–54.

R. Kouhia and C. M. Menken (1995), 'On the solution of second-order post-buckling fields', *Commun. Numer. Meth. Engng.* **11**, 443–453.

P. Krzyżanowski (2001), 'On block preconditioners for nonsymmetric saddle point problems', *SIAM J. Sci. Comput.* **23**, 157–169.

A. Kunoth (2002), 'Fast iterative solution of saddle point problems in optimal control based on wavelets', *Comput. Optim. Appl.* **22**, 225–259.

Y. A. Kuznetsov (1995), 'Efficient iterative solvers for elliptic finite element problems on nonmatching grids', *Russian J. Numer. Anal. Math. Modelling* **10**, 187–211.

Y. A. Kuznetsov (2004), 'Efficient preconditioner for mixed finite element methods on nonmatching meshes', *Russian J. Numer. Anal. Math. Modelling* **19**, 163–172.

C. Lanczos (1950), 'An iteration method for the solution of the eigenvalue problem of linear differential and integral operators', *J. Research Nat. Bur. Standards* **45**, 255–282.

L. D. Landau and E. M. Lifschitz (1959), *Fluid Mechanics*, Pergamon Press, London.

U. Langer and W. Queck (1986), 'On the convergence factor of Uzawa's algorithm', *J. Comput. Appl. Math.* **15**, 191–202.

https://doi.org/10.1017/S0962492904000212 Published online by Cambridge University Press

K. H. Leem, S. Oliveira and D. E. Stewart (2004), 'Algebraic multigrid (AMG) for saddle point systems from meshfree discretizations', *Numer. Linear Algebra Appl.* **11**, 293–308.

R. B. Lehoucq and A. G. Salinger (2001), 'Large-scale eigenvalue calculations for stability analysis of steady flows on massively parallel computers', *Internat. J. Numer. Meth. Fluids* **36**, 309–327.

W. Leontief, F. Duchin and D. B. Szyld (1985), 'New approaches in economic analysis', *Science* **228**, 419–422.

C. Li and C. Vuik (2004), 'Eigenvalue analysis of the SIMPLE preconditioning for incompressible flow', *Numer. Linear Algebra Appl.* **11**, 511–523.

C. Li, Z. Li, D. J. Evans and T. Zhang (2003), 'A note on an SOR-like method for augmented systems', *IMA J. Numer. Anal.* **23**, 581–592.

J. Li (2001), A dual–primal FETI method for incompressible Stokes equations, Technical Report TR2001-816, Courant Institute of Mathematical Sciences, New York University.

J. Li (2002*a*), A dual–primal FETI method for incompressible Stokes and linearized Navier–Stokes equations, Technical Report TR2002-828, Courant Institute of Mathematical Sciences, New York University.

J. Li (2002*b*), Dual–primal FETI methods for stationary Stokes and Navier–Stokes equations, PhD thesis, Courant Institute of Mathematical Sciences, New York. TR-830, Department of Computer Science, New York University.

N. Li and Y. Saad (2004), Crout versions of ILU factorization with pivoting for sparse symmetric matrices, Technical Report UMSI-2004-044, Minnesota Supercomputer Institute.

J. Liesen and P. E. Saylor (2005), 'Orthogonal Hessenberg reduction and orthogonal Krylov subspace bases', *SIAM J. Numer. Anal.* **42**, 2148–2158.

J. Liesen and Z. Strakoš (2005), 'GMRES convergence analysis for a convection-diffusion model problem', *SIAM J. Sci. Comput.*, to appear.

J. Liesen and P. Tichý (2004*a*), Behavior of CG and MINRES for symmetric tridiagonal Toeplitz matrices, Technical Report 34-2004, Technical University of Berlin, Department of Mathematics.

J. Liesen and P. Tichý (2004*b*), 'The worst-case GMRES for normal matrices', *BIT* **44**, 79–98.

J. Liesen, E. de Sturler, A. Sheffer, Y. Aydin and C. Siefert (2001), Preconditioners for indefinite linear systems arising in surface parameterization, in *Proc. 10th International Meshing Round Table*, pp. 71–81.

J. Liesen, M. Rozložník and Z. Strakoš (2002), 'Least squares residuals and minimal residual methods', *SIAM J. Sci. Comput.* **23**, 1503–1525.

L. Little and Y. Saad (2003), 'Block preconditioners for saddle point problems', *Numer. Algorithms* **33**, 367–379.

L. Little, Y. Saad and L. Smoch (2003), 'Block LU preconditioners for symmetric and nonsymmetric saddle point problems', *SIAM J. Sci. Comput.* **25**, 729–748.

J. W.-H. Liu (1987), 'A partial pivoting strategy for sparse symmetric matrix decomposition', *ACM Trans. Math. Software* **13**, 173–182.

W. M. Liu and S. Xu (2001), 'A new improved Uzawa method for finite element solution of Stokes' problem', *Comput. Mech.* **27**, 305–310.

https://doi.org/10.1017/S0962492904000212 Published online by Cambridge University Press

D. Loghin and A. J. Wathen (2002), 'Schur complement preconditioners for the Navier–Stokes equations', *Internat. J. Numer. Methods Fluids* **40**, 403–412.

D. Loghin and A. J. Wathen (2003), 'Schur complement preconditioning for elliptic systems of partial differential equations', *Numer. Linear Algebra Appl.* **10**, 423–443.

D. Loghin and A. J. Wathen (2004), 'Analysis of preconditioners for saddle-point problems', *SIAM J. Sci. Comput.* **25**, 2029–2049.

T.-T. Lu and S.-H. Shiou (2002), 'Inverses of 2×2 block matrices', *Comput. Math. Appl.* **43**, 119–129.

G. Lube and M. A. Olshanskii (2002), 'Stable finite element calculation of incompressible flows using the rotation form of convection', *IMA J. Numer. Anal.* **22**, 437–461.

G. Luenberger (1984), *Linear and Nonlinear Programming*, Addison-Wesley, Reading, MA.

L. Lukšan and J. Vlček (1998), 'Indefinitely preconditioned inexact Newton method for large sparse equality constrained non-linear programming problems', *Numer. Linear Algebra Appl.* **5**, 219–247.

L. Lukšan and J. Vlček (2001), 'Numerical experience with iterative methods for equality constrained non-linear programming problems', *Optim. Methods Software* **16**, 257–287.

T. Lyche, T. K. Nilssen and R. Winther (2002), 'Preconditioned iterative methods for scattered data interpolation', *Adv. Comput. Math.* **17**, 237–256.

D. S. Mackey, N. Mackey and F. Tisseur (2003), 'Structured tools for structured matrices', *Electr. J. Linear Algebra* **10**, 106–145.

Y. Maday, D. Meiron, A. T. Patera and E. M. Ronquist (1993), 'Analysis of iterative methods for the steady and unsteady Stokes problem: Application to spectral element discretizations', *SIAM J. Sci. Comput.* **14**, 310–337.

H. Mahawar and V. Sarin (2003), 'Parallel iterative methods for dense linear systems in inductance extraction', *Parallel Comput.* **29**, 1219–1235.

D. S. Malkus (1981), 'Eigenproblems associated with the discrete LBB condition for incompressible finite elements', *Internat. J. Engng. Sci.* **19**, 1299–1310.

K.-A. Mardal and R. Winther (2004), 'Uniform preconditioners for the time dependent Stokes problem', *Numer. Math.* **98**, 305–327.

H. M. Markowitz (1959), *Portfolio Selection: Efficient Diversification of Investments*, Wiley, New York.

H. M. Markowitz and A. F. Perold (1981), Sparsity and piecewise linearity in large portfolio optimization problems, in *Sparse Matrices and Their Uses* (I. S. Duff, ed.), Academic Press, London, pp. 89–108.

J. Martikainen (2003), 'Numerical study of two sparse AMG methods', *Math. Model. Numer. Anal.* **37**, 133–142.

J. Maryška, M. Rozložník and M. Tůma (1995), 'Mixed-hybrid finite element approximation of the potential fluid flow problem', *J. Comput. Appl. Math.* **63**, 383–392.

J. Maryška, M. Rozložník and M. Tůma (1996), 'The potential fluid flow problem and the convergence rate of the minimal residual method', *Numer. Linear Algebra Appl.* **3**, 525–542.

https://doi.org/10.1017/S0962492904000212 Published online by Cambridge University Press

J. Maryška, M. Rozložník and M. Tůma (2000), 'Schur complement systems in the mixed-hybrid finite element approximation of the potential fluid flow problem', *SIAM J. Sci. Comput.* **22**, 704–723.
P. Matstoms (1994), Sparse QR factorization with applications to linear least squares problems, PhD thesis, Linköping University, Linköping, Sweden.
W. McGuire and R. H. Gallagher (1979), *Matrix Structural Analysis*, Wiley, New York.
K. Meerbergen and A. Spence (1997), 'Implicitly restarted Arnoldi with purification for the shift-invert transformation', *Math. Comp.* **66**, 667–689.
S. Mehrotra and J. Wang (1996), Conjugate gradient based implementation of interior point methods for network flow problems, in *Linear and Nonlinear Conjugate Gradient-Related Methods* (L. Adams and L. Nazareth, eds), SIAM, Philadelphia, PA, pp. 124–142.
O. Meshar and S. Toledo (2005), 'An out-of-core sparse symmetric indefinite factorization method', *ACM Trans. Math. Software*, to appear.
G. Meurant (1999), *Computer Solution of Large Linear Systems*, Vol. 28 of *Studies in Mathematics and its Applications*, North-Holland, Amsterdam.
M. D. Mihajlović and D. J. Silvester (2004), 'Efficient parallel solvers for the biharmonic equation', *Parallel Comput.* **30**, 35–55.
J. Modersitzki (2003), *Numerical Methods for Image Registration*, Numerical Mathematics and Scientific Computation, Oxford University Press, Oxford.
J. D. Moulton, J. E. Morel and U. M. Ascher (1998), 'Approximate Schur complement preconditioning of the lowest-order nodal discretization', *SIAM J. Sci. Comput.* **19**, 185–205.
M. F. Murphy, G. H. Golub and A. J. Wathen (2000), 'A note on preconditioning for indefinite linear systems', *SIAM J. Sci. Comput.* **21**, 1969–1972.
N. M. Nachtigal, S. C. Reddy and L. N. Trefethen (1992), 'How fast are nonsymmetric matrix iterations?', *SIAM J. Matrix Anal. Appl.* **13**, 778–795.
A. E. Naiman, I. M. Babuška and H. C. Elman (1997), 'A note on conjugate gradient convergence', *Numer. Math.* **76**, 209–230.
S. G. Nash and A. Sofer (1996), 'Preconditioning reduced matrices', *SIAM J. Matrix Anal. Appl.* **17**, 47–68.
E. G. Ng and B. W. Peyton (1993), 'Block sparse Cholesky algorithms on advanced uniprocessor computers', *SIAM J. Sci. Comput.* **14**, 1034–1056.
H. Q. Nguyen (2004), Domain decomposition methods for linear-quadratic elliptic optimal control problems, PhD thesis, Department of Computational and Applied Mathematics, Rice University, Houston, TX.
R. A. Nicolaides (1982), 'Existence, uniqueness and approximation for generalized saddle point problems', *SIAM J. Numer. Anal.* **19**, 349–357.
J. Nocedal and S. J. Wright (1999), *Numerical Optimization*, Springer Series in Operations Research, Springer, Berlin.
R. H. Nochetto and J.-H. Pyo (2004), 'Optimal relaxation parameter for the Uzawa method', *Numer. Math.* **98**, 695–702.
B. Nour-Omid and P. Wriggers (1986), 'A two-level iteration method for solution of contact problems', *Comput. Methods Appl. Mech. Engrg.* **54**, 131–144.
A. R. L. Oliveira and D. C. Sorensen (2005), 'A new class of preconditioners for large-scale linear systems for interior point methods for linear programming', *Linear Algebra Appl.* **394**, 1–24.

https://doi.org/10.1017/S0962492904000212 Published online by Cambridge University Press

M. A. Olshanskii (1999), 'An iterative solver for the Oseen problem and numerical solution of incompressible Navier–Stokes equations', *Numer. Linear Algebra Appl.* **6**, 353–378.

M. A. Olshanskii and A. Reusken (2002), 'Navier–Stokes equations in rotation form: a robust multigrid solver for the velocity problem', *SIAM J. Sci. Comput.* **23**, 1683–1706.

M. A. Olshanskii and A. Reusken (2004), Analysis of a Stokes interface problem, Technical Report 237, RWTH Aachen, Aachen, Germany.

J. M. Ortega (1972), *Numerical Analysis: A Second Course*, Computer Science and Applied Mathematics, Academic Press, New York and London.

P. Oswald (1998), 'An optimal multilevel preconditioner for solenoidal approximations of the two-dimensional Stokes problem', *IMA J. Numer. Anal.* **18**, 207–228.

C. C. Paige and M. A. Saunders (1975), 'Solution of sparse indefinite systems of linear equations', *SIAM J. Numer. Anal.* **12**, 617–629.

C. C. Paige and M. A. Saunders (1982), 'LSQR: An algorithm for sparse linear equations and sparse least squares', *ACM Trans. Math. Software* **8**, 43–71.

S. V. Patankar (1980), *Numerical Heat Transfer and Fluid Flow*, McGraw-Hill, New York.

S. V. Patankar (1981), 'A calculation procedure for two-dimensional elliptic situations', *Numer. Heat Transfer* **4**, 409–425.

S. V. Patankar and D. B. Spalding (1972), 'A calculation procedure for heat, mass and momentum transfer in three-dimensional parabolic flows', *Internat. J. Heat Mass Transfer* **15**, 1787–1806.

L. F. Pavarino (1997), 'Preconditioned conjugate residual methods for mixed spectral discretizations of elasticity and Stokes problems', *Comput. Methods Appl. Mech. Engrg.* **146**, 19–30.

L. F. Pavarino (1998), 'Preconditioned mixed spectral element methods for elasticity and Stokes problems', *SIAM J. Sci. Comput.* **19**, 1941–1957.

L. F. Pavarino (2000), 'Indefinite overlapping Schwarz methods for time-dependent Stokes problems', *Comput. Methods Appl. Mech. Engrg.* **187**, 35–51.

L. F. Pavarino and O. B. Widlund (2000), 'Iterative substructuring methods for spectral element discretizations of elliptic systems II: Mixed methods for linear elasticity and Stokes flow', *SIAM J. Numer. Anal.* **37**, 375–402.

L. F. Pavarino and O. B. Widlund (2002), 'Balancing Neumann–Neumann methods for incompressible Stokes equations', *Comm. Pure Appl. Math.* **55**, 302–335.

M. Pernice (2000), 'A hybrid multigrid method for the steady-state incompressible Navier–Stokes equations', *Electr. Trans. Numer. Anal.* **10**, 74–91.

J. B. Perot (1993), 'An analysis of the fractional step method', *J. Comput. Phys.* **108**, 51–58.

I. Perugia (1997), 'A field-based mixed formulation for the two-dimensional magnetostatic problem', *SIAM J. Numer. Anal.* **34**, 2382–2391.

I. Perugia and V. Simoncini (2000), 'Block-diagonal and indefinite symmetric preconditioners for mixed finite element formulations', *Numer. Linear Algebra Appl.* **7**, 585–616.

https://doi.org/10.1017/S0962492904000212 Published online by Cambridge University Press

I. Perugia, V. Simoncini and M. Arioli (1999), 'Linear algebra methods in a mixed approximation of magnetostatic problems', *SIAM J. Sci. Comput.* **21**, 1085–1101.

J. Peters, V. Reichelt and A. Reusken (2004), Fast iterative solvers for discrete Stokes equations, Technical Report 241, RWTH Aachen, Aachen, Germany.

R. Peyret and T. D. Taylor (1983), *Computational Methods for Fluid Flow*, Springer Series in Computational Physics, Springer, Berlin.

R. J. Plemmons (1986), 'A parallel block iterative scheme applied to computations in structural analysis', *SIAM J. Algebraic Discrete Methods* **7**, 337–347.

R. J. Plemmons and R. E. White (1990), 'Substructuring methods for computing the nullspace of equilibrium matrices', *SIAM J. Matrix Anal. Appl.* **11**, 1–22.

B. T. Polyak (1970), 'Iterative methods using Lagrange multipliers for solving extremal problems with constraints of the equation type', *USSR Comput. Math. Math. Phys.* **10**, 42–52.

A. Pothen (1984), Sparse null bases and marriage theorems, PhD thesis, Department of Computer Science, Cornell University, Ithaca, NY.

A. Pothen (1989), 'Sparse null basis computations in structural optimization', *Numer. Math.* **55**, 501–519.

C. E. Powell and D. Silvester (2004), 'Optimal preconditioning for Raviart-Thomas mixed formulation of second-order elliptic problems', *SIAM J. Matrix Anal. Appl.* **25**, 718–738.

M. J. D. Powell (1969), A method for nonlinear constraints in minimization problems, in *Optimization* (R. Fletcher, ed.), Academic Press, London, pp. 283–298.

M. J. D. Powell (2004), On updating the inverse of a KKT matrix, Technical Report DAMTP2004/NA01, Department of Applied Mathematics and Theoretical Physics, Cambridge University.

M. L. Psiaki and K. Park (1995), 'Parallel orthogonal factorization null-space method for dynamic-quadratic programming', *J. Optim. Theory Appl.* **85**, 409–434.

Y. Qu (2000), Multilevel methods for indefinite systems, PhD thesis, Department of Civil Engineering, Renssealer Polytechnic Institute, Troy, NY.

Y. Qu and J. Fish (2000), 'Global basis two-level method for indefinite systems, Part 2: Computational issues', *Internat. J. Numer. Methods Engng.* **49**, 461–478.

Y. Qu and J. Fish (2002), 'Multifrontal incomplete factorization for indefinite and complex symmetric systems', *Internat. J. Numer. Methods Engng.* **53**, 1433–1459.

A. Quarteroni and A. Valli (1994), *Numerical Approximation of Partial Differential Equations*, Vol. 23 of *Springer Series in Computational Mathematics*, Springer, Berlin.

A. Quarteroni and A. Valli (1999), *Domain Decomposition Methods for Partial Differential Equations*, Numerical Mathematics and Scientific Computation, Oxford University Press, Oxford.

A. Quarteroni, F. Saleri and A. Veneziani (2000), 'Factorization methods for the numerical approximation of Navier–Stokes equations', *Comput. Methods Appl. Mech. Engrg.* **188**, 505–526.

https://doi.org/10.1017/S0962492904000212 Published online by Cambridge University Press

W. Queck (1989), 'The convergence factor of preconditioned algorithms of the Arrow–Hurwicz type', *SIAM J. Numer. Anal.* **26**, 1016–1030.

M. J. Raw (1995), A coupled algebraic multigrid method for the 3D Navier–Stokes equations, in *Fast Solvers for Flow Problems, Proc. 10th GAMM-Seminar* (W. Hackbusch and G. Wittum, eds), Notes on Numerical Fluid Mechanics, Vieweg, Wiesbaden, Germany, pp. 204–215.

W.-Q. Ren and J.-X. Zhao (1999), 'Iterative methods with preconditioners for indefinite systems', *J. Comput. Math.* **17**, 89–96.

M. Resende and G. Veiga (1993), 'An implementation of the dual affine scaling algorithm for minimum-cost flow on bipartite uncapacitated networks', *SIAM J. Optim.* **3**, 516–537.

M. P. Robichaud, P. A. Tanguy and M. Fortin (1990), 'An iterative implementation of the Uzawa algorithm for 3-D fluid flow problems', *Internat. J. Numer. Meth. Fluids* **10**, 429–442.

J. Robinson (1973), *Integrated Theory of Finite Element Methods*, Wiley, New York.

D. Rose (1970), 'Triangulated graphs and the elimination process', *J. Math. Anal. Appl.* **32**, 597–609.

M. Rozložník and V. Simoncini (2002), 'Krylov subspace methods for saddle point problems with indefinite preconditioning', *SIAM J. Matrix Anal. Appl.* **24**, 368–391.

T. Rusten and R. Winther (1992), 'A preconditioned iterative method for saddle-point problems', *SIAM J. Matrix Anal. Appl.* **13**, 887–904.

T. Rusten and R. Winther (1993), 'Substructure preconditioners for elliptic saddle-point problems', *Math. Comp.* **60**, 23–48.

T. Rusten, P. S. Vassilevski and R. Winther (1996), 'Interior penalty preconditioners for mixed finite element approximations of elliptic problems', *Math. Comp.* **65**, 447–466.

Y. Saad (1981), 'Krylov subspace methods for solving large unsymmetric linear systems', *Math. Comp.* **37**, 105–126.

Y. Saad (1982), 'The Lanczos biorthogonalization algorithm and other oblique projection methods for solving large unsymmetric systems', *SIAM J. Numer. Anal.* **19**, 485–506.

Y. Saad (1993), 'A flexible inner-outer preconditioned GMRES algorithm', *SIAM J. Sci. Comput.* **14**, 461–469.

Y. Saad (2003), *Iterative Methods for Sparse Linear Systems*, second edn, SIAM, Philadelphia, PA.

Y. Saad and M. H. Schultz (1986), 'GMRES: a generalized minimal residual algorithm for solving nonsymmetric linear systems', *SIAM J. Sci. Statist. Comput.* **7**, 856–869.

P. Saint-Georges, Y. Notay and G. Warzée (1998), 'Efficient iterative solution of constrained finite element analyses', *Comput. Methods Appl. Mech. Engrg.* **160**, 101–114.

F. Saleri and A. Veneziani (2005), 'Pressure-correction algebraic splitting methods for the incompressible Navier–Stokes equations', *SIAM J. Numer. Anal.*, to appear.

https://doi.org/10.1017/S0962492904000212 Published online by Cambridge University Press

A. Sameh and V. Sarin (2002), 'Parallel algorithms for indefinite linear systems', *Parallel Comput.* **28**, 285–299.
V. Sarin (1997), Efficient iterative methods for saddle point problems, PhD thesis, Department of Computer Science, University of Illinois, Urbana-Champaign, IL.
V. Sarin and A. Sameh (1998), 'An efficient iterative method for the generalized Stokes problem', *SIAM J. Sci. Comput.* **19**, 206–226.
V. Sarin and A. Sameh (2003), 'Hierarchical divergence-free bases and their application to particulate flows', *Trans. ASME* **70**, 44–49.
M. A. Saunders (2005), private communication.
O. Schenk and K. Gärtner (2004), On fast factorization pivoting methods for sparse symmetric indefinite systems, Technical Report CS-2004-004, Department of Computer Science, University of Basel.
J. Schöberl and W. Zulehner (2003), 'On Schwarz-type smoothers for saddle point problems', *Numer. Math.* **95**, 377–399.
G. J. Shaw and S. Sivaloganathan (1988), 'On the smoothing properties of the SIMPLE pressure correction algorithm', *Internat. J. Numer. Meth. Fluids* **8**, 441–461.
Y. Shi (1995), 'Solving linear systems involved in constrained optimization', *Linear Algebra Appl.* **229**, 175–189.
R. Sibson and G. Stone (1991), 'Computation of thin-plate splines', *SIAM J. Sci. Stat. Comput.* **12**, 1304–1313.
A. Sidi (2003), 'A zero-cost preconditioning for a class of indefinite linear systems', *WSEAS Trans. Math.* **2**, 142–150.
C. Siefert and E. de Sturler (2004), Preconditioners for generalized saddle-point problems, Technical Report UIUCDCS-R-2004-2448, Department of Computer Science, University of Illinois.
D. J. Silvester and A. J. Wathen (1994), 'Fast iterative solution of stabilised Stokes systems II: Using general block preconditioners', *SIAM J. Numer. Anal.* **31**, 1352–1367.
D. J. Silvester, H. C. Elman, D. Kay and A. J. Wathen (2001), 'Efficient preconditioning of the linearized Navier–Stokes equations for incompressible flow', *J. Comput. Appl. Math.* **128**, 261–279.
V. Simoncini (2004*a*), 'Block triangular preconditioners for symmetric saddle point problems', *Appl. Numer. Math.* **49**, 63–80.
V. Simoncini (2004*b*), private communication.
V. Simoncini and M. Benzi (2004), 'Spectral properties of the Hermitian and skew-Hermitian splitting preconditioner for saddle point problems', *SIAM J. Matrix Anal. Appl.* **26**, 377–389.
V. Simoncini and M. Benzi (2005), 'On the eigenvalues of a class of saddle point matrices'. In preparation.
V. Simoncini and D. B. Szyld (2003), 'Theory of inexact Krylov subspace methods and applications to scientific computing', *SIAM J. Sci. Comput.* **25**, 454–477.
V. Simoncini and D. B. Szyld (2005), 'The effect of non-optimal bases on the convergence of Krylov subspace methods', *Numer. Math.*, to appear.
G. Starke (1997), 'Field-of-values analysis of preconditioned iterative methods for nonsymmetric elliptic problems', *Numer. Math.* **78**, 103–117.

https://doi.org/10.1017/S0962492904000212 Published online by Cambridge University Press

G. W. Stewart (1989), 'On scaled projections and pseudoinverses', *Linear Algebra Appl.* **112**, 189–193.
G. Stoyan (2001), 'Iterative Stokes solvers in the harmonic Velte subspace', *Comput.* **67**, 13–33.
G. Strang (1986), *Introduction to Applied Mathematics*, Wellesley–Cambridge Press, Wellesley, MA.
G. Strang (1988), 'A framework for equilibrium equations', *SIAM Rev.* **30**, 283–297.
J. C. Strikwerda (1984), 'An iterative method for solving finite difference approximations to the Stokes equations', *SIAM J. Numer. Anal.* **21**, 447–458.
K. Stüben (2001), 'A review of algebraic multigrid', *J. Comput. Appl. Math.* **128**, 281–309.
T. Stykel (2005), 'Balanced truncation model reduction for semidiscretized Stokes equations', *Linear Algebra Appl.*, to appear.
J. G. Sun (1999), 'Structured backward errors for KKT systems', *Linear Algebra Appl.* **288**, 75–88.
Syamsudhuha and D. J. Silvester (2003), 'Efficient solution of the steady-state Navier–Stokes equations using a multigrid preconditioned Newton–Krylov method', *Internat. J. Numer. Meth. Fluids* **43**, 1407–1427.
D. B. Szyld (1981), Using sparse matrix techniques to solve a model of the world economy, in *Sparse Matrices and Their Uses* (I. S. Duff, ed.), Academic Press, pp. 357–365.
R. E. Tarjan (1983), *Data Structures and Network Algorithms*, SIAM, Philadelphia, PA.
R. Temam (1984), *Navier–Stokes Equations: Theory and Numerical Analysis*, Vol. 2 of *Studies in Mathematics and its Applications*, third edn, North-Holland, Amsterdam.
J. L. Thomas, B. Diskin and A. Brandt (2001), 'Textbook multigrid efficiency for the incompressible Navier–Stokes equations: high Reynolds number wakes and boundary layers', *Comput. & Fluids* **30**, 853–874.
M. J. Todd (1990), 'A Dantzig–Wolfe-like variant of Karmarkar's interior point linear programming algorithm', *Oper. Res.* **38**, 1006–1018.
K.-C. Toh (1997), 'GMRES vs. ideal GMRES', *SIAM J. Matrix Anal. Appl.* **18**, 30–36.
K.-C. Toh, K.-K. Phoon and S.-H. Chan (2004), 'Block preconditioners for symmetric indefinite linear systems', *Internat. J. Numer. Methods Engng.* **60**, 1361–1381.
Z. Tong and A. Sameh (1998), 'On an iterative method for saddle point problems', *Numer. Math.* **79**, 643–646.
A. Topcu (1979), A contribution to the systematic analysis of finite element structures using the force method, PhD thesis, University of Essen, Germany. In German.
A. Toselli and O. B. Widlund (2004), *Domain Decomposition Methods: Algorithms and Theory*, Vol. 34 of *Springer Series in Computational Mathematics*, Springer.
A. M. Tropper (1962), *Matrix Theory for Electrical Engineers*, Vol. 1 of *International Series in the Engineering Sciences*, Addison-Wesley, New York.

https://doi.org/10.1017/S0962492904000212 Published online by Cambridge University Press

U. Trottenberg, C. W. Oosterlee and A. Schüller (2001), *Multigrid*, Academic Press, London.

M. Tůma (2002), 'A note on the LDL^T decomposition of matrices from saddle-point problems', *SIAM J. Matrix Anal. Appl.* **23**, 903–915.

S. Turek (1999), *Efficient Solvers for Incompressible Flow Problems: An Algorithmic and Computational Approach*, Vol. 6 of *Lecture Notes in Computational Science and Engineering*, Springer, Berlin.

E. Tyrtyshnikov and Y. Vassilevski (2003), A mosaic preconditioner for a dual Schur complement, in *Numerical Mathematics and Advanced Applications* (F. Brezzi, A. Buffa, S. Corsaro and A. Murli, eds), Springer, pp. 867–877.

H. Uzawa (1958), Iterative methods for concave programming, in *Studies in Linear and Nonlinear Programming* (K. J. Arrow, L. Hurwicz and H. Uzawa, eds), Stanford University Press, Stanford, CA, pp. 154–165.

P. Vaidya (1990), Solving linear equations with diagonally dominant matrices by constructing good preconditioners, Technical report, Department of Computer Science, University of Illinois.

E. Vainikko and I. G. Graham (2004), 'A parallel solver for PDE systems and application to the incompressible Navier–Stokes equations', *Appl. Numer. Math.* **49**, 97–116.

H. A. van der Vorst (1992), 'Bi-CGSTAB: a fast and smoothly converging variant of Bi-CG for the solution of nonsymmetric linear systems', *SIAM J. Sci. Statist. Comput.* **13**, 631–644.

H. A. van der Vorst (2003), *Iterative Krylov Methods for Large Linear Systems*, Vol. 13 of *Cambridge Monographs on Applied and Computational Mathematics*, Cambridge University Press, Cambridge.

C. F. Van Loan (1985), 'On the method of weighting for equality constrained least squares problems', *SIAM J. Numer. Anal.* **22**, 851–864.

R. J. Vanderbei (1995), 'Symmetric quasidefinite matrices', *SIAM J. Optim.* **5**, 100–113.

S. P. Vanka (1986), 'Block-implicit multigrid solution of Navier–Stokes equations in primitive variables', *J. Comput. Phys.* **65**, 138–158.

R. S. Varga (1962), *Matrix Iterative Analysis*, Prentice-Hall, Englewood Cliffs, NJ.

P. B. Vasconcelos and F. D. D'Almeida (1998), 'Preconditioned iterative methods for coupled discretizations of fluid flow problems', *IMA J. Numer. Anal.* **18**, 385–397.

P. S. Vassilevski and R. D. Lazarov (1996), 'Preconditioning mixed finite element saddle-point elliptic problems', *Numer. Linear Algebra Appl.* **3**, 1–20.

P. S. Vassilevski and J. Wang (1992), 'Multilevel iterative methods for mixed finite element discretizations of elliptic probems', *Numer. Math.* **63**, 503–520.

S. A. Vavasis (1994), 'Stable numerical algorithms for equilibrium systems', *SIAM J. Matrix Anal. Appl.* **15**, 1108–1131.

S. A. Vavasis (1996), 'Stable finite elements for problems with wild coefficients', *SIAM J. Numer. Anal.* **33**, 890–916.

A. Veneziani (2003), 'Block factorized preconditioners for high-order accurate in time approximation of the Navier–Stokes equations', *Numer. Methods Partial Differential Equations* **19**, 487–510.

https://doi.org/10.1017/S0962492904000212 Published online by Cambridge University Press

R. Verfürth (1984*a*), 'A combined conjugate gradient-multigrid algorithm for the numerical solution of the Stokes problem', *IMA J. Numer. Anal.* **4**, 441–455.
R. Verfürth (1984*b*), 'A multilevel algorithm for mixed problems', *SIAM J. Numer. Anal.* **21**, 264–271.
C. Vincent and R. Boyer (1992), 'A preconditioned conjugate gradient Uzawa-type method for the solution of the Stokes problem by mixed Q_1-P_0 stabilized finite elements', *Internat. J. Numer. Meth. Fluids* **14**, 289–298.
M. Wabro (2004), 'Coupled algebraic multigrid methods for the Oseen problem', *Comput. Visual. Sci.* **7**, 141–151.
C. Wagner, W. Kinzelbach and G. Wittum (1997), 'Schur-complement multigrid: a robust method for groundwater flow and transport problems', *Numer. Math.* **75**, 523–545.
Y. Wang (2005), 'Overlapping Schwarz preconditioner for the mixed formulation of plane elasticity', *Appl. Numer. Math.*, to appear.
J. S. Warsa, M. Benzi, T. A. Wareing and J. E. Morel (2004), 'Preconditioning a mixed discontinuous finite element method for radiation diffusion', *Numer. Linear Algebra Appl.* **11**, 795–811.
A. J. Wathen (1987), 'Realistic eigenvalue bounds for the Galerkin mass matrix', *IMA J. Numer. Anal.* **7**, 449–457.
A. J. Wathen and D. J. Silvester (1993), 'Fast iterative solution of stabilised Stokes systems I: Using simple diagonal preconditioners', *SIAM J. Numer. Anal.* **30**, 630–649.
A. J. Wathen, B. Fischer and D. J. Silvester (1995), 'The convergence rate of the minimal residual method for the Stokes problem', *Numer. Math.* **71**, 121–134.
R. Webster (1994), 'An algebraic multigrid solver for Navier–Stokes problems', *Internat. J. Numer. Meth. Fluids* **18**, 761–780.
M. Wei (1992*a*), 'Algebraic properties of the rank-deficient equality constrained and weighted least squares problems', *Linear Algebra Appl.* **161**, 27–43.
M. Wei (1992*b*), 'Perturbation theory for the rank-deficient equality constrained least squares problem', *SIAM J. Numer. Anal.* **29**, 1462–1481.
Y. Wei and N. Zhang (2004), 'Further note on constraint preconditioning for non-symmetric indefinite matrices', *Appl. Math. Comput.* **152**, 43–46.
P. Wesseling (1992), *An Introduction to Multigrid Methods*, Wiley, Chichester, UK.
P. Wesseling (2001), *Principles of Computational Fluid Dynamics*, Vol. 29 of *Springer Series in Computational Mathematics*, Springer, Berlin.
P. Wesseling and C. W. Oosterlee (2001), 'Geometric multigrid with applications to fluid dynamics', *J. Comput. Appl. Math.* **128**, 311–334.
O. B. Widlund (1978), 'A Lanczos method for a class of nonsymmetric systems of linear equations', *SIAM J. Numer. Anal.* **15**, 801–812.
S. Ø. Wille and A. F. D. Loula (2002), '*A priori* pivoting in solving the Navier–Stokes equations', *Commun. Numer. Meth. Engng.* **18**, 691–698.
S. Ø. Wille, O. Staff and A. F. D. Loula (2003), 'Parallel ILU preconditioning, *a priori* pivoting and segregation of variables for iterative solution of the mixed finite element formulation of the Navier–Stokes equations', *Internat. J. Numer. Meth. Fluids* **41**, 977–996.
G. Wittum (1989), 'Multi-grid methods for Stokes and Navier–Stokes equations with transforming smoothers: Algorithms and numerical results', *Numer. Math.* **54**, 543–563.

https://doi.org/10.1017/S0962492904000212 Published online by Cambridge University Press

G. Wittum (1990), 'On the convergence of multi-grid methods with transforming smoothers', *Numer. Math.* **57**, 15–38.
P. Wolfe (1962), The reduced gradient method, Technical report, The RAND Corporation, Santa Monica, CA. Unpublished.
M. H. Wright (1992), 'Interior methods for constrained optimization', *Acta Numerica*, Vol. 1, Cambridge University Press, pp. 341–407.
S. J. Wright (1997), *Primal–Dual Interior Point Methods*, SIAM, Philadelphia, PA.
X. Wu, B. P. B. Silva and J. Y. Yuan (2004), 'Conjugate gradient method for rank deficient saddle point problems', *Numer. Algorithms* **35**, 139–154.
D. Yang (2002), 'Iterative schemes for mixed finite element methods with applications to elasticity and compressible flow problems', *Numer. Math.* **93**, 177–200.
N. Zhang and Y. Wei (2004), 'Solving EP singular linear systems', *Internat. J. Computer Math.* **81**, 1395–1405.
J.-X. Zhao (1998), 'The generalized Cholesky factorization method for saddle point problems', *Appl. Math. Comput.* **92**, 49–58.
O. C. Zienkiewicz, J. P. Vilotte, S. Toyoshima and S. Nakazawa (1985), 'Iterative method for constrained and mixed approximation: An inexpensive improvement of FEM performance', *Comput. Methods Appl. Mech. Engrg.* **51**, 3–29.
A. Zsaki, D. Rixen and M. Paraschivoiu (2003), 'A substructure-based iterative inner solver coupled with Uzawa's algorithm for the Stokes problem', *Internat. J. Numer. Meth. Fluids* **43**, 215–230.
W. Zulehner (2000), 'A class of smoothers for saddle point problems', *Comput.* **65**, 227–246.
W. Zulehner (2002), 'Analysis of iterative methods for saddle point problems: a unified approach', *Math. Comp.* **71**, 479–505.

https://doi.org/10.1017/S0962492904000212 Published online by Cambridge University Press