BIT Numerical Mathematics
2003, Vol. 43, No. 1, pp. 001–018

0006-3835/03/4301-0001 $16.00

A black-box multigrid preconditioner for the biharmonic equation *

DAVID J. SILVESTER[1] and MILAN D. MIHAJLOVIĆ[2]

[1] *Department of Mathematics, University of Manchester Institute of Science and Technology, Manchester M60 1QD, UK. email: na.silvester@na-net.ornl.gov*

[2] *Department of Computer Science, University of Manchester, Oxford Road, Manchester M13 9PL, UK. email: milan@cs.man.ac.uk*

Abstract.

We examine the convergence characteristics of a preconditioned Krylov subspace solver applied to the linear systems arising from low-order mixed finite element approximation of the biharmonic problem. The key feature of our approach is that the preconditioning can be realized using any "black-box" multigrid solver designed for the discrete Dirichlet Laplacian operator. This leads to preconditioned systems having an eigenvalue distribution consisting of a tightly clustered set together with a small number of outliers. Numerical results show that the performance of the methodology is competitive with that of specialized fast iteration methods that have been developed in the context of biharmonic problems.

AMS subject classification (2000):

Key words: Biharmonic equation, mixed methods, finite elements, preconditioning, multigrid.

1 Introduction.

In this paper, we assess the performance of a new preconditioning methodology designed for discrete systems arising from the mixed approximation of the classical biharmonic problem: given a convex polygonal domain $\Omega \in \mathbb{R}^2$ with boundary $\partial\Omega$, and a sufficiently smooth load function (typically $f \in H^{-1}(\Omega)$), we seek the solution pair (ϕ, ω) satisfying

$$(1.1) \qquad \left.\begin{array}{rcl} -\Delta\phi & = & \omega \\ -\Delta\omega & = & f \end{array}\right\} \text{ in } \Omega, \qquad \phi = \frac{\partial\phi}{\partial n} = 0 \quad \text{on } \partial\Omega.$$

This problem arises in fluid mechanics, in which case ϕ and ω represent the streamfunction and vorticity, respectively. It is also a model for plate bending, in which case ϕ is the deflection and ω is the bending moment, see for example, Bjorstad & Tjostheim [6]. Our theoretical setting is that established in the review paper of Glowinski & Pironneau [8]. We outline the essential features here.

*Received August 2002. Communicated by .

The standard mixed approximation method for (1.1) is as follows: we seek $(\omega, \phi) \in H^1(\Omega) \times H_0^1(\Omega)$ satisfying

$$
\begin{array}{rcll}
(\omega, \upsilon) - (\nabla\upsilon, \nabla\phi) & = & 0 & \forall \upsilon \in H^1(\Omega) \\
-(\nabla\omega, \nabla\varphi) & = & -(f, \varphi) & \forall \varphi \in H_0^1(\Omega),
\end{array} \tag{1.2}
$$

where $(\cdot,\cdot)$ denotes the standard vector or scalar $L^2(\Omega)$ inner product. To generate a discrete system we take finite dimensional subspaces $X_h \subset H^1(\Omega)$ and $M_h = X_h \cap H_0^1(\Omega)$, where h is a representative mesh parameter, and enforce (1.2) over the subspaces X_h and M_h: that is compute $(\omega_h, \phi_h) \in X_h \times M_h$ satisfying

$$
\begin{array}{rcll}
(\omega_h, \upsilon_h) - (\nabla\upsilon_h, \nabla\phi_h) & = & 0 & \forall \upsilon_h \in X_h \\
-(\nabla\omega_h, \nabla\varphi_h) & = & -(f, \varphi_h) & \forall \varphi_h \in M_h.
\end{array} \tag{1.3}
$$

The fact that the discrete constraint space

$$Z_h = \{\omega_h \in X_h; (\nabla\omega_h, \nabla\varphi_h) = 0 \quad \forall \varphi_h \in M_h\}$$

is not a subspace of the continuous anologue means that the well-posedness of (1.3) does not automatically follow from the well-posedness of (1.2). Indeed, appropriate mesh dependent norms on $X_h \times M_h$, denoted $\|\omega\|_{0,h}$ and $\|\phi\|_{2,h}$ (see [2] for the definitions) need to be introduced if the natural approximation spaces of X_h consisting of C^0 piecewise polynomials of degree $k \geq 1$ and $M_h = X_h \cap H_0^1(\Omega)$, are to satisfy Brezzi's abstract stability conditions, see [5]. This intrinsic lack of stability is what ultimately limits the potential for developing spectrally equivalent preconditioners (and robust multigrid methods) for solving the discretised problem (1.3).

To express the discrete problem (1.3) as a linear algebra problem we introduce operators $\mathcal{M} : X_h \mapsto X_h$ and $\mathcal{B} : X_h \mapsto M_h$ defined via

$$
\begin{array}{rcll}
(\mathcal{M}\omega_h, \upsilon_h) & = & (\omega_h, \upsilon_h) & \forall \omega_h, \upsilon_h \in X_h, \\
(\mathcal{B}\omega_h, \varphi_h) & = & (\omega_h, \mathcal{B}^T\varphi_h) = -(\nabla\omega_h, \nabla\varphi_h) & \forall \omega_h \in X_h, \forall \varphi_h \in M_h,
\end{array} \tag{1.4}
$$

so that $\mathcal{B}^T$ is the adjoint of $\mathcal{B}$. With these definitions (1.3) can be rewritten as a matrix system:

$$
\begin{pmatrix} \mathcal{M} & \mathcal{B}^T \\ \mathcal{B} & 0 \end{pmatrix} \begin{pmatrix} \omega_h \\ \phi_h \end{pmatrix} = \begin{pmatrix} 0 \\ f_h \end{pmatrix}, \tag{1.5}
$$

where f_h is the $L^2(\Omega)$ orthogonal projection of $-f$ into M_h. Furthermore, inf-sup stability and boundedness with respect to the mesh dependent norms $\|\cdot\|_{2,h}$ and $\|\cdot\|_{0,h}$, see [2, Thm. 3], implies that

$$
\gamma \|\phi_h\|_{2,h} \leq \sup_{\omega_h \in X_h} \frac{(\mathcal{B}\omega_h, \phi_h)}{(\mathcal{M}\omega_h, \omega_h)^{1/2}} \leq \Gamma \|\phi_h\|_{2,h} \qquad \forall \phi_h \in M_h, \tag{1.6}
$$

with constants $0 < \gamma < \Gamma$ independent of h. The denominator in (1.6) has a simple form since the norm $\|\omega_h\|_{0,h}$ is equivalent to the standard L^2 norm for functions in X_h.

The Schur complement operator $\mathcal{S} := \mathcal{B}\mathcal{M}^{-1}\mathcal{B}^t$ is the starting point for designing efficient multigrid preconditioners and solvers for the system (1.5), see [12], [4] and [9]. Note however that $\mathcal{S}$ has a condition number which deteriorates under mesh refinement like h^{-4}. To determine an optimal preconditioning strategy for $\mathcal{S}$ it is useful to express (1.5) and (1.6) in terms of the actual finite element matrices that arise in practice. To this end, we explicitly introduce the finite element basis set,

$$X_h = \mathrm{span}\{\psi_i\}_{i=1}^{n_I+n_B}, \qquad M_h = \mathrm{span}\{\psi_j\}_{j=1}^{n_I}; \tag{1.7}$$

so that we have n_I interior and n_B boundary degrees of freedom respectively. We then associate the functions ω_h, ϕ_h, f_h with the vectors $\underline{\omega} \in \mathbb{R}^{n_I+n_B}$, $\underline{\phi} \in \mathbb{R}^{n_I}$ and $\underline{f} \in \mathbb{R}^{n_I}$ of generalised coefficients, $\omega_h = \sum_{i=1}^{n_I+n_B} \underline{\omega}_i \psi_i$ etc. Defining the $(n_I + n_B) \times (n_I + n_B)$ "mass" matrix $M_{ij} = (\psi_i, \psi_j)$ and also the $n_I \times (n_I + n_B)$ constraint matrix $B_{ji} = -(\nabla\psi_j, \nabla\psi_i)$, gives the finite element version of (1.5)

$$\begin{pmatrix} M & B^t \\ B & 0 \end{pmatrix} \begin{pmatrix} \underline{\omega} \\ \underline{\phi} \end{pmatrix} = \begin{pmatrix} \underline{0} \\ \underline{f} \end{pmatrix}. \tag{1.8}$$

Furthermore, if the matrix operator Q is constructed (the realization of the mesh-dependent norm associated with the basis set defining M_h), so that $\|\phi_h\|_{2,h}^2 \equiv \underline{\phi}^t Q \underline{\phi}$, then rewriting the LHS of (1.6) leads to the following characterisation. For all $\underline{\phi} \in \mathbb{R}^{n_I}$, we have that

$$\gamma(\underline{\phi}^t Q \underline{\phi})^{1/2} \le \max_{\underline{\omega}} \frac{\underline{\phi}^t B \underline{\omega}}{(\underline{\omega}^t M \underline{\omega})^{1/2}} \tag{1.9}$$

$$= \max_{\underline{w}=M^{1/2}\underline{\omega}} \frac{\underline{\phi}^t B M^{-1/2} \underline{w}}{(\underline{w}^t \underline{w})^{1/2}} \tag{1.10}$$

$$= (\underline{\phi}^t B M^{-1} B^t \underline{\phi})^{1/2}, \tag{1.11}$$

since the maximum is attained when $\underline{w} = M^{-1/2} B^t \underline{\phi}$. Treating the RHS of (1.6) similarly implies the Rayleigh quotient bounds

$$\gamma^2 \le \frac{\underline{\phi}^t B M^{-1} B^t \underline{\phi}}{\underline{\phi}^t Q \underline{\phi}} \le \Gamma^2 \qquad \forall \underline{\phi} \in \mathbb{R}^{n_I} \tag{1.12}$$

with constants γ and Γ independent of h. The matrix Q thus represents an "ideal" preconditioner for the Schur complement $S = BM^{-1}B^t$ in (1.8). Whilst the construction of Q is certainly feasible, effecting the action of inverse of the operator Q in $O(n_I)$ flops is a challenging (and unresolved) problem. Our solution to this problem is to sacrifice optimality (in the sense of computational complexity) for practicality. Specifically, we explore using simple multigrid cycles, designed for Dirichlet Poisson problems, as an approximation to the "hard-to-handle" operator Q. Through a combination of analytical and experimental results we demonstrate that the resulting preconditioning strategy is extremely effective: in practice, iteration counts tend to grow very slowly with mesh refinement.

To see how such a preconditioner might be defined, we follow [8] and [4], by decomposing ω_h into the sum of interior and boundary contributions:

$$\underbrace{\sum_{i=1}^{n_I+n_B} \underline{\omega}_i \psi_i}_{\omega_h \in X_h} = \underbrace{\sum_{j=1}^{n_I} \underline{v}_j \psi_j}_{v_h \in M_h} + \underbrace{\sum_{k=1}^{n_B} \underline{\lambda}_k \psi_{n_I+k}}_{\lambda_h \in T_h}, \tag{1.13}$$

where $T_h \approx L^2(\partial\Omega)$. This decomposition induces a partitioning of the matrices in (1.8): in particular, $B = (K_I \quad K_B)$, where the $n_I \times n_I$ matrix K_I is the standard finite element stiffness matrix obtained when discretising a Dirichlet problem using M_h. If the matrix M is lumped[1], e.g. using appropriate quadrature as discussed in [9], then M in (1.8) is replaced by the diagonal matrix $M_\ell = \mathrm{diag}\,(M_I, M_B)$ and (1.8) can be written in the form of a 3×3 block system

$$\begin{pmatrix} M_I & 0 & K_I \\ 0 & M_B & K_B^t \\ K_I & K_B & 0 \end{pmatrix} \begin{pmatrix} \underline{v} \\ \underline{\lambda} \\ \underline{\phi} \end{pmatrix} = \begin{pmatrix} \underline{0} \\ \underline{0} \\ \underline{f} \end{pmatrix}. \tag{1.14}$$

The Schur complement in (1.14) is of the form

$$S_\ell := K_I M_I^{-1} K_I + K_B M_B^{-1} K_B^t, \tag{1.15}$$

and may be readily approximated by the matrix

$$S_* := K_* M_I^{-1} K_*, \tag{1.16}$$

where K_* represents the action of multigrid applied to the model Poisson problem (e.g. the action of a single V-cycle with damped Jacobi as a smoother). This type of preconditioner $S_* \approx S_\ell \approx S$ is discussed in [4] but does not seem to have been fully developed because of its non-optimality: the matrix $S_*^{-1} S_\ell$ has n_* eigenvalues (relatively few, since $n_* < n_B$) increasing like h^{-1} under mesh refinement. We will demonstrate that preconditioners that rely on the approximation of (1.15) by (1.16) can be extremely effective nonetheless: indeed, iteration counts tend to grow very slowly with mesh refinement.

A summary of the paper is as follows. Two general ways of preconditioning (1.8) are described in sections 2 and 3. In section 2, motivated by our philosophy for solving discrete Stokes problems, a multigrid block diagonal preconditioning strategy is analysed. We go on to consider a novel approach in section 3 that is based on an exact constraint preconditioner. This gives a preconditioned system whose spectrum depends on $S_*^{-1} S$ with $K_* = K_I$. Inexact versions of this methodology have been suggested in the other contexts, specifically magnetostatics [14] and optimal control in ground-water flow modelling [3], but have not

[1]This means that the L^2 norm of ω_h is approximated so that $\underline{\omega}^t M \underline{\omega} = \|\omega_h\|^2 \approx \underline{v}^t M_I \underline{v} + \underline{\lambda}^t M_B \underline{\lambda}$, where M_I and M_B are diagonal matrices.

been considered in the context of biharmonic problems until now. In section 4 we present some numerical experiments illustrating the computational efficiency of our inexact constraint preconditioning approach. As an alternative approach, optimal preconditioning methods based on multigrid ideas have recently been derived by Arnold, Falk and Winther [1] in the context of stable finite element discretisation methods of the Reissner-Mindlin plate model. When taking the limit of an infinitely thin plate—so that the Reissner-Mindlin plate model reduces to the biharmonic problem—the results in section 4 are just as impressive as the numerical results in [1].

2 A conventional block preconditioner.

In this section we consider a block preconditioning approach that has proved to be extremely effective in the context of mixed approximation methods, see [15] and [16].

To provide insight we initially consider the mass lumped system (1.14), and the "exact" preconditioner (1.16) with $K_* = K_I$. The preconditioning matrix in this case is $\mathcal{P} := \operatorname{diag}(M_I, M_B, K_I M_I^{-1} K_I)$ and is symmetric and positive definite[2]. The eigenvalues $\{\mu_i\}$ of the preconditioned operator are real and satisfy

$$(2.1)\qquad \begin{pmatrix} M_I & 0 & K_I \\ 0 & M_B & K_B^t \\ K_I & K_B & 0 \end{pmatrix} \begin{pmatrix} \underline{v} \\ \underline{\lambda} \\ \underline{\phi} \end{pmatrix} = \mu \begin{pmatrix} M_I & 0 & 0 \\ 0 & M_B & 0 \\ 0 & 0 & K_I M_I^{-1} K_I \end{pmatrix} \begin{pmatrix} \underline{v} \\ \underline{\lambda} \\ \underline{\phi} \end{pmatrix}.$$

We note that any zero eigenvalue of (2.1) corresponds to the coefficient matrix being singular. For this to happen the Schur complement $S_\ell = K_I M_I^{-1} K_I + K_B M_B^{-1} K_B^t$ must be semi-definite, but this is impossible since $K_I M_I^{-1} K_I$ is positive definite. Note that this implies that the rank of $B = (K_I \quad K_B)$ is n_I.

By inspection, one solution of (2.1) is $\mu = 1$ and is of multiplicity n_B. The associated eigenvectors are of the form $(\underline{v} \quad \underline{\lambda} \quad \underline{0})^T$ where $B\underline{\omega} \equiv K_I \underline{v} + K_B \underline{\lambda} = \underline{0}$; i.e. $\underline{\omega}$ is in the nullspace of the constraint operator. We also note that the $n_I \times n_B$ matrix K_B is rank deficient. It maps a vector of boundary values onto those nodes that lie in elements on the boundary, and it has zero rows corresponding to all other interior nodes. In the case of a square domain using a cartesian product grid of bilinear elements, the dimension of the nullspace of K_B is 8. In general the rank deficiency depends on the shape of the domain Ω, and the order of approximation k.

We now consider $\mu \neq 1$. In this case eliminating $\underline{v}$ and $\underline{\lambda}$ from (2.1) leads to the eigenproblem

$$(2.2)\qquad (K_I M_I^{-1} K_I + K_B M_B^{-1} K_B^t)\underline{\phi} = \sigma K_I M_I^{-1} K_I \underline{\phi},$$

where $\sigma = \mu(\mu - 1)$. Moreover, writing (2.2) in terms of $\underline{v}$ and $\underline{\lambda}$ gives a natural characterisation of σ:

$$(2.3)\qquad \sigma = \frac{\underline{v}^t M_I \underline{v} + \underline{\lambda}^t M_B \underline{\lambda}}{\underline{v}^t M_I \underline{v}} = \frac{\|\omega_h\|_0^2}{\|v_h\|_0^2}.$$

[2]The matrices M_I, M_B and K_I are all symmetric and positive definite

Clearly $\sigma \geq 1$.

There are two distinct cases; either $\underline{\phi}$ is in the nullspace of K_B^t, or else $\underline{\phi}$ is in the range of K_B. In the first case, the component $\underline{\lambda}$ of the eigenvector is zero, and $\sigma = 1$. If K_B has rank deficiency n_r then the corresponding eigenvectors of (2.2) live in a subspace of dimension $n_I - n_B + n_r$. Furthermore, since every eigenvalue σ_j; $j = 1, \ldots, n_B$ generates two distinct eigenvalues $\mu_\pm$, it is clear that (2.1) has two different eigenvalues of multiplicity $n_I - n_B + n_r$, namely $\mu_-^* = (1-\sqrt{5})/2$ and $\mu_+^* = (1+\sqrt{5})/2$.

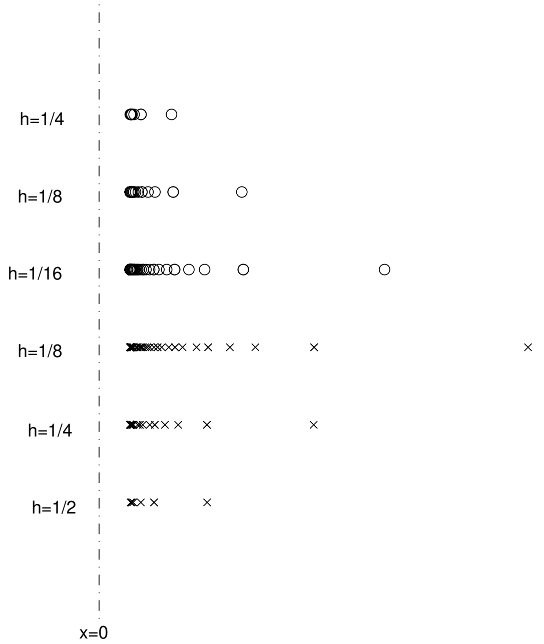

Figure 2.1: Computed eigenvalues σ using bilinear approximation (o): h=1/4, 1/8, 1/16 and biquadratic approximation (x): h=1/2, 1/4, 1/8.

The other possibility is that $\underline{\phi} \notin \mathrm{null}(K_B^t)$. In this case the eigenvectors of (2.2) form a subspace of dimension $n_B - n_r$ associated with $\underline{\lambda} \not\equiv \underline{0}$ in (2.3), and the associated eigenvalues σ_n lie in an interval $(1,\sigma_{max}]$. A bound for σ_{max} is given by Braess & Peisker [4, Prop. 3.3].

LEMMA 2.1. *Assuming a standard inverse estimate $\|\phi_h\|_1 \leq Ch^{-1}\|\phi_h\|_0$ for $\phi_h \in X_h$, we have that*

$$1 \leq \frac{\underline{\phi}^t(K_I M_I^{-1} K_I + K_B M_B^{-1} K_B^t)\underline{\phi}}{\underline{\phi}^t K_I M_I^{-1} K_I \underline{\phi}} \leq Ch^{-1}.$$

To see that this bound is tight, discrete eigenvalues $\{\sigma_n\}$ computed using bilinear and biquadratic approximation on a sequence of uniform square grids, are plotted in Fig. 2.1. The maximal eigenvalue is clearly growing like h^{-1}. In terms of the preconditioned operator (2.1) there is an eigenvalue cluster on each side of the origin; namely, $[-Ch^{-1/2}, \mu_-^*]$ and $[\mu_+^*, Ch^{-1/2}]$, both of which expand with increasing mesh refinement. Notice that both clusters contain at most n_B eigenvalues, which becomes a vanishingly small proportion of the total number of eigenvalues in the limit $h \to 0$.

The use of a multigrid iteration in place of the Poisson operator K_I in the preconditioner in (2.1) is covered by the analysis in [16].

PROPOSITION 2.2. *Assume the conditions in Lemma 2.1 are satisfied, and that there exist constants θ, Θ, such that*

$$\theta^2 \le \frac{\underline{\omega}^t M \underline{\omega}}{\underline{\omega}^t M_\ell \underline{\omega}} \le \Theta^2 \qquad \forall \underline{\omega} \in \mathbb{R}^{n_I + n_B}.$$

If the matrix operator K_* is such that there exists a constant η satisfying

$$\|K_*^{-1} K_I - I\| \le \eta,$$

then the eigenvalues μ of the block preconditioned system

$$\begin{pmatrix} M & B^t \\ B & 0 \end{pmatrix} \begin{pmatrix} \underline{\omega} \\ \underline{\phi} \end{pmatrix} = \mu \begin{pmatrix} M_\ell & 0 \\ 0 & K_* M_I^{-1} K_* \end{pmatrix} \begin{pmatrix} \underline{\omega} \\ \underline{\phi} \end{pmatrix}$$

lie in the union of intervals

$$[-ah^{-1/2}, -b] \cup [c, dh^{-1/2}]$$

where a, b, c, d are independent of h.

PROOF. The Schur complement approximation can be expressed as a product

$$\frac{\underline{\omega}^t B M^{-1} B^t \underline{\omega}}{\underline{\omega}^t K_* M_I^{-1} K_* \underline{\omega}} = \underbrace{\frac{\underline{\omega}^t B M^{-1} B^t \underline{\omega}}{\underline{\omega}^t B M_\ell^{-1} B^t \underline{\omega}}}_{(i)} \times \underbrace{\frac{\underline{\omega}^t B M_\ell^{-1} B^t \underline{\omega}}{\underline{\omega}^t K_I M_I^{-1} K_I \underline{\omega}}}_{(ii)} \times \underbrace{\frac{\underline{\omega}^t K_I M_I^{-1} K_I \underline{\omega}}{\underline{\omega}^t K_* M_I^{-1} K_* \underline{\omega}}}_{(iii)},$$

of terms which can be bounded seperately. In particular, $(i) \in [1/\Theta^2, 1/\theta^2]$, from the bounds on $M_\ell^{-1} M$; $(ii) \in [1, Ch^{-1}]$, from Lemma 2.1; and using a duality argument as in [4], we have that $(iii) \in [(1-\eta)^2, (1+\eta)^2]$. The result then follows immediately from [16, Thm. 2.1]. □

REMARK 2.1. *As discussed in* [7] *there is a close relationship between the convergence of the Krylov subspace method* MINRES *applied to the preconditioned system in Proposition 2.2, and that of the conjugate gradient method applied to the Schur complement system $BM^{-1}B^t$, when it is preconditioned with the operator $S_* = K_* M_I^{-1} K_*$. Based on the bounds in Proposition 2.2, both the preconditioned* MINRES *and the preconditioned CG method will converge in at most $O(h^{-1/2})$ iterations (*MINRES *will typically take twice as many iterations to reach the same tolerance, see* [7, sect. 4]*).*

Our computational experience is that this $O(h^{-1/2})$ growth in iteration counts (when solving to a fixed tolerance) is observed in practice, i.e. when using one or more multigrid V-cycles as a preconditioning operator. In contrast, in the exact case, typical iteration counts grow much more slowly as h is reduced; see for example [13, Table 1]. We attribute this to the fact that there are a vanishing proportion of badly behaved eigenvalues in the exact case, as discussed above. In the next section we describe our method of choice for preconditioning the discrete biharmonic operator in (1.8).

3 A constraint preconditioner

In this section we consider a constraint preconditioning approach, see [10]. Although this approach has been used previously in the context of mixed approximations, our strategy is actually quite different—it is tailored to the special structure of the coefficient matrix in (1.8).

To provide insight, we consider the mass lumped system (1.14), and the following ideal version of the preconditioner:

$$\mathcal{P} = \begin{pmatrix} 0 & 0 & K_I \\ 0 & M_B & K_B^t \\ K_I & K_B & 0 \end{pmatrix} \equiv \begin{pmatrix} G & B^t \\ B & 0 \end{pmatrix}. \tag{3.1}$$

Note that $\mathcal{P}$ is symmetric but indefinite. Note also that the action of the inverse of $\mathcal{P}$ is simply achieved by back-substitution.

Proposition 3.1. *The matrix $\mathcal{P}$ in* (3.1) *is non-singular.*

Proof. Since B is of full rank, a sufficient condition for $\mathcal{P}$ to be invertible is that G is positive definite on the constraint space Z_h i.e. for all non-trivial vectors $\underline{\omega} \in \mathbb{R}^{n_I}$ such that $B\underline{\omega} = \underline{0}$. By contradiction, assume that there exists $\underline{\omega}_* = [\underline{v} \quad \underline{\lambda}]^t$ such that $B\underline{\omega}_* = \underline{0}$ and $\underline{\omega}_*^t G \underline{\omega}_* = 0$. Noting that $\underline{\omega}_*^t G \underline{\omega}_* = \underline{\lambda}^t M_B \underline{\lambda}$ we deduce that $\underline{\lambda} = \underline{0}$. Imposing the constraint $B\underline{\omega}_* = \underline{0}$ then means that $K_I \underline{v} = \underline{0}$, so that $\underline{v} = \underline{0}$, hence $\underline{\omega}_* = \underline{0}$! ◻

We now consider the eigenvalues of the preconditioned operator:

$$\begin{pmatrix} M_I & 0 & K_I \\ 0 & M_B & K_B^t \\ K_I & K_B & 0 \end{pmatrix} \begin{pmatrix} \underline{v} \\ \underline{\lambda} \\ \underline{\phi} \end{pmatrix} = \mu \begin{pmatrix} 0 & 0 & K_I \\ 0 & M_B & K_B^t \\ K_I & K_B & 0 \end{pmatrix} \begin{pmatrix} \underline{v} \\ \underline{\lambda} \\ \underline{\phi} \end{pmatrix}. \tag{3.2}$$

We first note that although $\mathcal{P}$ is indefinite, the eigenvalues $\{\mu_i\}$ are real and nonzero, see below. The structure of the eigenvectors corresponding to the multiple eigenvalue $\mu = 1$ is discussed in [10, Thm. 2.1].

Lemma 3.2. *Given that the preconditioner* (3.1) *is invertible, the preconditioned system* (3.2) *has an eigenvalue $\mu = 1$ with multiplicity $\geq 2n_I$.*

We now investigate the case $\mu \neq 1$. By inspection, eliminating $\underline{v}$ and $\underline{\lambda}$ from (3.2) leads to the eigenproblem

$$(\mu - 1) K_I M_I^{-1} K_I \underline{\phi} = K_B M_B^{-1} K_B^t \underline{\phi}. \tag{3.3}$$

There are again two possibilities to consider. If $\underline{\phi}$ is in the nullspace of K_B^t, we again have that $\mu = 1$ (if $\underline{\phi} = \underline{0}$ then $\underline{\lambda} = \underline{v} = \underline{0}$). The second possibility corresponds to the eigenproblem

$$(K_I M_I^{-1} K_I + K_B M_B^{-1} K_B^t)\underline{\phi} = \mu K_I M_I^{-1} K_I \underline{\phi}, \tag{3.4}$$

i.e. identical to (2.2) except that $\underline{\phi}$ is restricted to the subspace $\underline{\phi} \notin \text{null}(K_B^t)$, which is of dimension $n_B - n_r$. The upshot here is that the preconditioned matrix operator corresponding to (3.2) has a multiple eigenvalue at unity, together with a small cluster of $n_B - n_r$ eigenvalues that live in the interval $(1, Ch^{-1}]$, as illustrated in Fig. 2.1. This means that we can anticipate extremely rapid convergence of a Krylov subspace method applied to the preconditioned system. Indeed, given such a distribution it can be shown[3] that the Krylov subspace method GMRES will terminate in at most n_B+2 iterations using exact arithmetic.

Since $\mathcal{P}$ is not definite, there is no guarantee that the standard MINRES method is going to work. Whilst we have not made an exhaustive study of alternative methods, our experience is that BICGSTAB(2) is an effective choice in the sense of giving comparable iteration counts to that of GMRES at minimal cost/iteration. Specialist Krylov subspace methods that exploit the symmetry of $\mathcal{A}$ and $\mathcal{P}$ could possibly be even more cost effective. For further discussion, see [14].

Using the ideal preconditioner (3.1) in the context of the system (1.8):

$$\begin{pmatrix} M & B^t \\ B & 0 \end{pmatrix} \begin{pmatrix} \underline{\omega} \\ \underline{\phi} \end{pmatrix} = \mu \begin{pmatrix} G & B^t \\ B & 0 \end{pmatrix} \begin{pmatrix} \underline{\omega} \\ \underline{\phi} \end{pmatrix}. \tag{3.5}$$

also generates an eigenvalue $\mu = 1$ with multiplicity $2n_I + n_r$. Moreover, using the bound (i) in the proof of Proposition 2.2, it can be shown that all the eigenvalues μ lie in the interval $[\theta^2, Ch^{-1}]$. By way of illustration, extremal eigenvalues computed using a set of uniformly refined triangular meshes are given in Tables 3.1–3.3. Iteration counts using BICGSTAB(2) to solve the associated preconditioned systems to a fixed tolerance are given in the next section. These show that the condition number bounds give a pessimistic picture of the convergence rate that is achieved in practice. Using the exact constraint preconditioner, (or if sufficiently many multigrid V-cycles are used to define K_* in place of K_I in (3.1)), iteration counts are essentially independent of h.

4 Numerical results

In this section we assess the performance of exact and inexact versions of the indefinite constraint preconditioner (3.1) when solving the discrete version of the biharmonic problem (1.8). Note that we are not lumping the mass matrix here. We illlustrate the performance of our methodology on the systems obtained by discretization of (1.1) on the unit square domain $\Omega = [0, 1]^2$. The discretization is performed on uniform meshes (congruent right-angled triangles) using the simplest posible Lagrangian element (piecewise linear approximation). The right hand side vector is chosen randomly (although the elements of $\underline{f}$ in (1.8) behave

[3]The preconditioned matrix is not diagonalisable (it only has $n_I + 2n_B - n_r$ linearly independent eigenvectors, see [10, Thm. 2.3]). This means that the standard convergence theory of Krylov subspace methods needs to be modified.

Table 3.1: Linear approximation ($k = 1$).

mesh	h	#dof	$\mu_{\min}$	$\mu_{\max}$	$\kappa(\mathcal{P}^{-1}\mathcal{A})$
12×12	0.0833	290	0.6851	14.85	21.68
24×24	0.0417	1154	0.6930	28.67	41.37
48×48	0.0208	4610	0.6956	56.29	80.93

Table 3.2: Quadratic approximation ($k = 2$).

mesh	h	#dof	$\mu_{\min}$	$\mu_{\max}$	$\kappa(\mathcal{P}^{-1}\mathcal{A})$
6×6	0.1667	290	0.7991	13.10	16.40
12×12	0.0833	1154	0.8121	25.54	31.46
24×24	0.0417	4610	0.8130	50.47	62.07

Table 3.3: Cubic approximation ($k = 3$).

mesh	h	#dof	$\mu_{\min}$	$\mu_{\max}$	$\kappa(\mathcal{P}^{-1}\mathcal{A})$
4×4	0.2500	290	0.8075	13.37	16.56
8×8	0.1250	1154	0.8169	25.65	31.40
16×16	0.0625	4610	0.8169	50.22	61.48

like $O(h^2)$, so as to simulate the realistic situation of a randomly distributed load).

Following our philosophy of designing a black-box implementation, we use publicly available code, namely the BICGSTAB(ℓ) algorithm supplied as a part of the routine F11BEF from the NAG library Mark 19 (see [11]). The stopping criterion for BICGSTAB(2) is given by

$$\|\underline{r}_i\|_\infty \leq \varepsilon\,(\|\underline{b}\|_\infty + \|A\|_\infty \|\underline{x}_i\|_\infty), \tag{4.1}$$

as provided by the routine F11BEF. In (4.1), $\underline{r}_i$ denotes the residual vector, $\underline{b}$ is the RHS vector, A is the coefficient matrix, and $\underline{x}_i$ is the current approximation of the solution. We give two different values for ε: 10^{-6} and 10^{-9}. In the first case $\|\underline{r}_i\|_\infty$ is typically 10^{-4} and in the latter case it is around 10^{-7}. By comparing the resulting solution with that obtained using a direct solver we anticipate at least three correct decimal digits using the coarser tolerance, with six correct decimal digits for the tighter tolerance.

In our experiments we cover exact and inexact approaches for solving the systems that involve the discrete Laplacian operator K_I (which occur as a result of computing the action of the preconditioner $\mathcal{P}$). In the inexact approach the discrete operator K_I is replaced by a multigrid approximation K_*, where a fixed number of V-cycles are typically performed. In our implementation we use the NAG routine D03EDF—which uses Wesseling's MGD1 algorithm [17] to solve

Table 4.1: Performance of the preconditioned BICGSTAB(2) method.

exact factorization for K_I

		$\varepsilon = 10^{-6}$		$\varepsilon = 10^{-9}$	
mesh	#`dof`	N_{it}	$\|\|\underline{r}_k\|\|_\infty$	N_{it}	$\|\|\underline{r}_k\|\|_\infty$
30×30	1802	5	1.1E-4	13	9.2E-8
42×42	3530	5	1.3E-4	17	3.8E-8
66×66	8714	5	1.7E-4	15	3.5E-7
114×114	25994	5	2.3E-4	23	1.2E-7
162×162	52490	5	2.3E-4	23	8.0E-7
258×258	133130	5	2.1E-4	29	7.2E-7

multigrid preconditioning: $V_n = 1$

		$\varepsilon = 10^{-6}$			$\varepsilon = 10^{-9}$		
mesh	#`dof`	N_{it}	$\|\|\underline{r}_k\|\|_\infty$	WU	N_{it}	$\|\|\underline{r}_k\|\|_\infty$	WU
30×30	1802	10	1.6E-5	0.21	16	2.1E-8	0.32
42×42	3530	8	2.0E-4	0.34	18	1.8E-8	0.71
66×66	8714	12	2.9E-4	1.23	26	9.3E-8	2.59
114×114	25994	16	6.1E-5	5.83	28	4.3E-7	9.72
162×162	52490	18	2.8E-4	13.95	34	8.5E-7	25.98
258×258	133130	26	1.3E-3	61.65	46	7.4E-7	107.31

multigrid preconditioning: $V_n = 3$

		$\varepsilon = 10^{-6}$			$\varepsilon = 10^{-9}$		
mesh	#`dof`	N_{it}	$\|\|\underline{r}_k\|\|_\infty$	WU	N_{it}	$\|\|\underline{r}_k\|\|_\infty$	WU
30×30	1802	6	1.1E-4	0.16	14	1.7E-8	0.36
42×42	3530	8	6.1E-5	0.42	16	6.2E-8	0.81
66×66	8714	6	1.7E-4	0.84	18	2.5E-7	2.42
114×114	25994	6	4.3E-4	3.42	18	6.1E-7	8.06
162×162	52490	4	9.1E-4	4.37	24	1.4E-7	24.26
258×258	133130	4	6.2E-4	12.92	20	1.2E-6	60.34

a general elliptic PDE on a rectangular domain discretised by a seven point difference operator—to effect this approximation. We emphasize that instead of the routine D03EDF one could use any other efficient Poisson solver based on multigrid (on a parallel architecture, domain decomposition solvers would be an appropriate alternative). The specific choice of meshes that we used in our experiments is governed by the fact that the routine D03EDF essentially requires that the number of grid points in each coordinate direction can be expressed in the form $n_x = n_y = m2^{l-1} + 1$.

The results are summarized in Table 4.1. We use the following notation; `mesh` denotes the number of grid cells in the x and y direction, and #`dof` is the total dimension of a discrete system (#`dof` $= 2n_I + n_B$). N_{it} denotes the number of

BICGSTAB(2) iterations, and $\|\underline{r}_k\|_\infty$ the final residual. For the inexact version we run one and three V-cycles respectively of the MGD1 multigrid method encoded in D03EDF, and we report a measure WU of the computational effort required to solve the discrete problem to the specified tolerance.

Surprisingly, the convergence rate seems to be independent of the subdivision in the case of the exact factorization and for the coarser tolerance. For the tighter tolerance, the iteration count slowly increases as h is reduced. Consistent with theoretical expectations, N_{it} seems to grow like $O(h^{-1/2})$ using the inexact preconditioner with a single multigrid cycle. On the other hand, taking three V-cycles gives the convergence rate that is seen in the case of the exact preconditioner. (Although the number of iterations is smaller than in the exact case on the finer grids, the accuracy of the solution is not as good.) In the latter case, WU is essentially proportional to the number of degrees of freedom, so that optimal complexity is achieved.

REFERENCES

1. D. N. Arnold, R. S. Falk and R. Winther, *Preconditioning discrete approximations of the Reissner-Mindlin plate model*, Math. Modelling Numer. Anal., 31 (1997), pp. 517–557.
2. I. Babuška, J. Osborn and J. Pitkäranta, *Analysis of mixed methods using mesh dependent norms*, Math. Comput., 35 (1980), pp. 1039–1062.
3. A. Battermann and E.W. Sachs, *Block Preconditioners for KKT systems in PDE-Governed Optimal Control Problems*, in Proceedings of Workshop on Fast Solutions of Discretised Optimisation Problems, R. H. W. Hoppe, K.-H. Hoffmann, V. Schulz, eds., Birkhäuser Verlag, Berlin, 2001, pp. 1–18.
4. D. Braess and P. Peisker, *On the numerical solution of the biharmonic equation and the role of squaring matrices for preconditioning*, IMA J. Numer. Anal., 6 (1986), pp. 393–404.
5. F. Brezzi and M. Fortin, *Mixed and Hybrid Finite Element Methods*, Springer-Verlag, New York, 1991.
6. P.E. Bjørstad, B.P. Tjøstheim, *Efficient algorithms for solving a fourth order equation with the spectral Galerkin method*, SIAM J. Sci. Comp., 18 (1997), pp. 621–632.
7. B. Fischer, A. Ramage, D. Silvester, and A. J. Wathen, *Minimum residual methods for augmented systems*, BIT, 38 (1998), pp. 527–543.
8. R. Glowinski and O. Pironneau, *Numerical methods for the first biharmonic equation and for the two-dimensional Stokes problem,* SIAM Rev., 21 (1979), pp. 167–212.
9. M. R. Hanisch, *Multigrid preconditioning for the biharmonic Dirichlet problem*, SIAM J. Numer. Anal., 30 (1993), pp. 184–214.
10. C. Keller, N. I. M. Gould and A. J. Wathen, *Constraint preconditioning for indefinite linear systems*, SIAM J. Matrix. Anal. Appl., 21 (2000), 1300–1317.
11. Numerical Algorithms Group, *NAG manual, Fortran Library Mark 19*, 1999.
12. P. Peisker, *A multilevel algorithm for the biharmonic problem*, Numer. Math., 46 (1985), pp. 623–634.
13. P. Peisker and D. Braess, *A conjugate gradient method and a multigrid algorithm for Morley's finite element approximation of the biharmonic equation*, Numer. Math., 50 (1987), pp. 567–586.

14. I. Perugia and V. Simoncini, *Block-diagonal and indefinite symmetric preconditioners for mixed finite element formulations*, Numer. Lin. Alg. Appl., 7 (2000), pp. 585–616.
15. T. Rusten and R. Winther, *A preconditioned iterative method for saddle point problems*, SIAM J. Matrix Anal. Appl., 13 (1992), pp. 887–904.
16. D. Silvester and A. Wathen, *Fast iterative solution of stabilized Stokes systems. Part II: Using block preconditioners*, SIAM J. Numer. Anal., 31 (1994), pp. 1352–1367.
17. P. Wesseling, *MGD1—A robust and efficient multigrid method*, in Lecture Notes in Mathematics 960, Springer-Verlag, 1982, pp. 614–630.