Angewandte Mathematik

Preconditioners for Indefinite Problems

Inaugural-Dissertation
zur Erlangung des Doktorgrades
der Naturwissenschaften im Fachbereich Mathematik
der Mathematisch-Naturwissenschaftlichen Fakultät
der Westfälischen Wilhelms–Universität Münster

vorgelegt von

Axel Klawonn

Westfälische
Wilhelms-Universität
Münster

November 1995

Dekan:	Prof. Dr. F.-K. Holtmeier
Erster Gutachter:	Prof. Dr. C. W. Cryer
Zweiter Gutachter:	Prof. Dr. F. Natterer
Tag der mündlichen Prüfungen:	01. , 06. , 13.02. 1996
Tag der Promotion:	13.02. 1996

Acknowledgements

My deepest gratitude goes to my adviser in New York, Professor Olof Widlund, for all the time and energy he has devoted to me and my work and I am very grateful to my adviser in Münster, Professor Colin Cryer, for his constant support and encouragement. I thank both of them for giving me the opportunity to spend a year at the Courant Institute of Mathematical Sciences in New York.

I also thank Dr. Barry Smith at Argonne National Laboratory for introducing me to the programming package PETSc.

In addition, I wish to thank all of the faculty, staff, and graduate students in Münster and New York for creating an interesting scientific atmosphere. I would like to acknowledge Mario Casarin and Dr. Marcus Sarkis for various discussions on domain decomposition.

Finally, I am very grateful for the financial support while visiting the Courant Institute. My stay during the academic year 1993/94 has been made possible through a scholarship of the Hochschulsonderprogramm II/AUFE des Deutschen Akademischen Austauschdienstes (DAAD). In March and April 1995, I was supported by the Department of Energy under contract DE-FG02-92ER25127.

Contents

Introduction

Many problems in the engineering sciences lead to indefinite systems of partial differential equations. Important examples are the Stokes equations of fluid dynamics, modeling the flow of an incompressible viscous fluid, and mixed formulations of problems from linear elasticity, e.g. for almost incompressible materials, beams and plates. These problems can be analyzed in the framework of saddlepoint problems with a penalty term. There are also important scalar indefinite partial differential equations, such as the Helmholtz equation which plays a fundamental role in many mathematical models of physical phenomena, e.g. the solution of Maxwell's equations; see Křižek and Neittaanmäki [59].

We now give an abstract formulation of the problem. Assume that $\mathcal{A} : X \longrightarrow Y^{'}$ is an isomorphism and that $\mathcal{F} \in Y^{'}$, where X, Y are Hilbert spaces and $X^{'}, Y^{'}$ their duals. Consider the abstract problem

$$\mathcal{A}x = \mathcal{F}. \tag{0.1}$$

Discretizing (0.1) by finite elements, we obtain a sparse linear system

$$\mathcal{A}_h x_h = \mathcal{F}_h,$$

which can be of very large dimension.

Ever since electronic computing became important, there has been a lot of research devoted to the development of (preconditioned) iterative methods for the solution of linear systems. It is not an exaggeration to say that the theory for symmetric, positive definite problems is very well advanced. Among the fastest algorithms are multigrid/-level and domain decomposition methods which often converge in $O(1)$ iterations when applied to this class of problems. Such a method is also often used as a preconditioner for a Krylov space method, which serves as an accelerator. The method of choice as an accelerator for symmetric, positive definite systems is the conjugate gradient algorithm. Turning to indefinite or nonsymmetric systems, no similar unified approach exists yet, although a lot of research has been carried out to extend the existing framework to these classes of problems.

It is the goal of this thesis to present new results on preconditioning indefinite linear systems arising from certain differential problems. As Krylov space methods, we use the conjugate residual algorithm for symmetric indefinite problems and GMRES and BI-CGSTAB for nonsymmetric systems.

We first consider saddlepoint problems with a penalty term. Then, we have

$$\mathcal{A} := \begin{pmatrix} A & B^t \\ B & -t^2 C \end{pmatrix}, \quad \mathcal{F} := \begin{pmatrix} f \\ g \end{pmatrix}, \quad t \in [0,1], \text{ and } X = Y = V \times M_c,$$

where V, M, M_c are certain Hilbert spaces with M_c dense in M and where A, C are linear operators fulfilling certain ellipticity conditions. The bilinear form associated with the operator B satisfies an inf-sup condition, i.e.

$$\inf_{p \in M_c} \sup_{u \in V} \frac{b(u,p)}{||p||_M \, ||u||_V} \geq \beta_0 > 0,$$

where $b(u,p) :=< Bu, p >$. From these assumptions, we know, from the famous theory of Babuška and Brezzi, that $\mathcal{A}$ defines an isomorphism. These results are also valid for properly chosen finite element spaces. We then require, additionally, that the constants are independent of the discretization parameter h. In the remainder of this indroduction, we restrict ourselves to the discretized versions of Problem (0.1), thus we can drop the index h.

Several techniques have been developed to solve saddlepoint problems, with or without a penalty term, iteratively. One of our goals is to develop and analyze methods with convergence rates independent of the penalty parameter. We have found little in the literature on preconditioned Krylov space methods concerning this issue.

The oldest algorithm for saddlepoint problems is Uzawa's algorithm; see Arrow, Hurwicz, and Uzawa [5]. One common variant of it is essentially a steepest descent algorithm applied to the Schur complement $t^2 C + BA^{-1}B^t$ of the indefinite linear system. When this algorithm is used, we have to solve a linear system of the form $A\delta = d$ quite accurately. This can require many iterations since this system normally is not well conditioned and it is often quite expensive to solve by a direct method. Therefore many authors have considered the effect of an inaccurate inner iteration for A^{-1}, see Bank, Welfert, and Yserentant [13], Bramble, Pasciak, and Vassilev [21], Elman and Golub [43], and Rønquist [68].

To avoid inner and outer iterations and to provide a much simpler approach, some authors have in recent years tried to precondition the whole indefinite system and using a conjugate residual algorithm as an accelerator; see e.g. Rusten and Winther [69], and Silvester and Wathen [85, 76]. This is closely related to our approach. In the first of these papers, a preconditioner for a saddle point problem without a penalty term is analyzed; in the others, problems arising from stabilized and unstabilized Stokes flow are considered. Only the stabilized case results in a saddle point problem with a penalty term. In contrast to certain problems arising in elasticity, where the penalty parameter normally arises from the material or the geometry, the penalty term in the stabilized Stokes case can and should be chosen to stabilize an otherwise unstable discretization and to ensure fast convergence of the iterative method. The main goal in Silvester and Wathen [85, 76] is to provide a good criterion for choosing this parameter in the context of preconditioning. We are not going to discuss this aspect here further. We refer the reader to the survey on stabilization techniques presented in Fortin [46] and to the references therein.

A third possibility is to transform the indefinite problem into a positive definite system by introducing a new inner product. Then the conjugate gradient method can be applied; see Bramble and Pasciak [20]. By using the results developed in Chapter 5, their transformation can be analyzed in the context of triangular preconditioners.

In our work, we present a blockdiagonal preconditioner,

$$\hat{\mathcal{B}}_D := \begin{pmatrix} \hat{A} & O \\ O & \hat{C} \end{pmatrix}, \tag{0.2}$$

where $\hat{A}, \hat{C}$ are $V-$elliptic and $M-$elliptic, respectively. Thus, $\hat{\mathcal{B}}_D$ is symmetric, positive definite. The preconditioner $\hat{\mathcal{B}}_D$ is assumed to be easily invertible and it ideally should define a norm equivalent to the standard norm in X; the preconditioning can be interpreted as a change of basis resulting in a well conditioned operator. We show that the conjugate residual method applied to the preconditioned system $\hat{\mathcal{B}}_D^{-1}\mathcal{A}$ can be made to converge independently of the discretization and the penalty parameters. We note that, except for our blockdiagonal preconditioner, all of the methods mentioned before explicitly assume that the subblock A of $\mathcal{A}$ is positive definite. However, the theory of Babuška and Brezzi indicates that this is not a necessary but only a sufficient condition for $\mathcal{A}$ being an isomorphism. Our new analysis of block-diagonal preconditioners can be applied, whenever $\mathcal{A}$ is an isomorphism, i.e. when $\mathcal{A}$ satisfies the assumptions of the Babuška-Brezzi theory.

Our second preconditioner has the form

$$\hat{\mathcal{B}}_U \ := \ \begin{pmatrix} \hat{A} & B^t \\ O & -\hat{C} \end{pmatrix},$$

where $\hat{A}$ and $\hat{C}$ again are $V-$ and $M-$elliptic, respectively. Thus, $\hat{\mathcal{B}}_U$ is nonsymmetric with an indefinite symmetric part. Using $\hat{\mathcal{B}}_U$ in combination with Krylov space methods like GMRES and BI-CGSTAB, we also obtain an iterative method that can be made to converge independently of the critical parameters. A new analysis of triangular preconditioners for symmetric saddlepoint problems is presented. We show that the preconditioned system is symmetrizable and use this property to derive GMRES convergence estimates dependent on $\sqrt{\kappa}$. Here κ is the condition number of the preconditioned system. Thus, we obtain a convergence estimate for a non-symmetric system matching, except for a constant leading factor, the well-known formula of the conjugate gradient method applied to symmetric, positive definite systems.

It turns out that our analysis of triangular preconditioners also provides a proof of a result due to Bramble and Pasciak [19].

As a second, somewhat different, example of indefinite problems, we consider scalar second order symmetric elliptic equations, such as the Helmholtz equation, i.e.

$$\mathcal{A} := A - C,$$

where $A, C > 0$ represent the part of second and zero order, respectively. The preconditioner is chosen as a good preconditioner for the principal part of the given differential operator, i.e. $\hat{\mathcal{B}} := \hat{A}$ is a good preconditioner for A. Then, our preconditioning strategy can again be interpreted as a change of basis. We show that the conjugate residual method applied to the preconditioned system can be made to converge independently of the discretization parameters.

The numerical experiments show that all of our preconditioners lead to robust iterative methods.

We note that our positive definite preconditioning techniques suggested for saddlepoint problems with a penalty term and for the Helmholtz equation, are generally valid for problems like (0.1), where $\mathcal{A}$ is a symmetric isomorphism. This result appears to be new.

The remainder of the thesis is organized as follows.

In the first chapter, we review Babuška's inf-sup and sup-sup conditions. They are necessary and sufficient for $\mathcal{A}$ being an isomorphism. Then, we give a short description of the Petrov-Galerkin method.

In the second chapter, we give a short general introduction to Krylov space methods and preconditioners, followed by a description of the conjugate residual method, GMRES and BI-CGSTAB. We also discuss some a priori convergence estimates which relate the convergence rates to the spectrum and the field of values of the preconditioned system. In the second part of this chapter, we review an abstract framework for the Schwarz methods which allows us to analyze a large family of preconditioners including multigrid and domain decomposition methods.

In the third chapter, we introduce the theory of saddlepoint problems with a penalty term. In the first part, we give conditions on the subblocks A, B, and C which guarantee that Babuška's inf-sup and sup-sup conditions for $\mathcal{A}$ are satisfied. In the second part, we discuss three examples of linear elasticity: the equations of linear elasticity for almost incompressible materials, the Timoshenko beam, and the Mindlin-Reissner plate in the Brezzi-Fortin formulation.

In the fourth chapter, we introduce our new analysis of blockdiagonal preconditioners for saddlepoint problems with a penalty term. This presentation is based on two reports [57, 58]. We give a condition number estimate and numerical results for a problem arising in linear elasticity.

In the fifth chapter, the triangular preconditioner is discussed. We show that the spectrum of the preconditioned system and its field of values, with respect to a certain energy metric, are bounded independently of the discretization and the penalty parameters. New convergence estimates for GMRES and numerical results using preconditioned GMRES and BI-CGSTAB are presented.

Finally, we show in the sixth chapter that the techniques developed in the fourth chapter also work well for the case of scalar symmetric second order elliptic equations. We give a condition number estimate and present numerical results for the Helmholtz equation.

Chapter 1

Background Results

1.1 The Abstract Problem

Let X, Y be Hilbert spaces, $X^{'}, Y^{'}$ be the duals of X, Y and let $< \cdot, \cdot >$ denote the dual pairing for $X^{'}, X$ or $Y^{'}, Y$. Let $\mathcal{A}(\cdot, \cdot) : X \times Y \longrightarrow \mathbf{R}$ be a continuous bilinear form. Associated with it is a linear operator $\mathcal{A} : X \longrightarrow Y^{'}$, defined by

$$< \mathcal{A}x, y >:= \mathcal{A}(x, y) \quad \forall y \in Y.$$

Consider the abstract problem:
Find $x \in X$, such that

$$\mathcal{A}x = \mathcal{F},$$

where $\mathcal{F} \in Y^{'}$. This problem is equivalent to the variational problem

$$\mathcal{A}(x, y) =< \mathcal{F}, y > \quad \forall y \in Y. \tag{1.1}$$

The next theorem on the uniqueness and existence of the solution of 1.1, is apparently due to Nečas [64], Theorem 3.1. Its importance for the finite element theory was pointed out by Babuška; see [12, 9, 11, 10].

Theorem 1.1 *Let X, Y be Hilbert spaces with norms $|||\cdot|||_X, |||\cdot|||_Y$. The variational problem (1.1) has a unique solution if the bilinear form $\mathcal{A}(\cdot, \cdot)$ satisfies the following three conditions*

(i) (sup-sup condition)
$\exists \gamma_1 > 0$*, such that*

$$\sup_{x \in X} \sup_{y \in Y} \frac{|\mathcal{A}(x, y)|}{|||x|||_X |||y|||_Y} \leq \gamma_1, \tag{1.2}$$

(ii) (inf-sup conditions)

$\exists \gamma_0 > 0$, *such that*

$$\inf_{x \in X} \sup_{y \in Y} \frac{\mathcal{A}(x,y)}{|||x|||_X |||y|||_Y} \geq \gamma_0, \tag{1.3}$$

$\exists \tilde{\gamma}_0 > 0$, *such that*

$$\inf_{y \in Y} \sup_{x \in X} \frac{\mathcal{A}(x,y)}{|||x|||_X |||y|||_Y} \geq \tilde{\gamma}_0. \tag{1.4}$$

For a proof of this theorem; see Nečas [64] or Babuška [9].
If the bilinear form $\mathcal{A}(\cdot,\cdot)$ is symmetric and $X = Y$, then we only have to prove one inf-sup condition. The well-known Lax-Milgram Lemma can easily be derived from the previous theorem.

Lemma 1.1 (Lax-Milgram) *Let X be a Hilbert space with the norm $|||\cdot|||$, $\mathcal{A}(\cdot,\cdot) : X \times X \longrightarrow \mathbf{R}$ a continuous $X-$elliptic bilinear form, i.e.*

$$\begin{aligned} \exists C_{\mathcal{A}} > 0, \textit{such that } \mathcal{A}(x,y) &\leq C_{\mathcal{A}} \, |||x|||\,|||y||| \quad \forall x, y \in X, \\ \exists \alpha > 0, \textit{such that } \mathcal{A}(x,x) &\geq \alpha \, |||x|||^2 \quad \forall x \in X. \end{aligned}$$

Let $\mathcal{F} : X \longrightarrow \mathbf{R}$ be a continuous linear functional, i.e. $\mathcal{F} \in X'$. Then, the variational problem

$$\mathcal{A}(x,y) = \mathcal{F}(y) \quad \forall y \in X,$$

has a unique solution.

Proof: The sup-sup condition of Theorem 1.1 follows directly from the continuity of the bilinear form $\mathcal{A}(\cdot,\cdot)$. We obtain the first inf-sup condition from

$$\sup_{y \in X} \frac{\mathcal{A}(x,y)}{|||y|||_Y} \geq \frac{\mathcal{A}(x,x)}{|||x|||_X} \geq \gamma_0 \, |||x|||_X.$$

The second inf-sup condition can be derived analogously since $X = Y$.

□

1.2 The Petrov-Galerkin Method

Let $h > 0$ denote a "small" parameter, which in the finite element method represents the meshsize, i.e. the diameter of a typical finite element. We assume that $X_h \subset X, Y_h \subset Y$ are finite dimensional subspaces with $\dim X_h =$

$\dim Y_h = n_h$ for all $h > 0$. The numerical approximation $x_h \in X_h$ of the solution of problem (1.1), is given by

$$\mathcal{A}(x_h, y_h) = <\mathcal{F}, y_h> \quad \forall y_h \in Y_h. \tag{1.5}$$

If the spaces X_h and Y_h are chosen, such that Theorem 1.1 is still valid, we obtain the following generalization of Cea's Lemma

Lemma 1.2 *Let $\mathcal{A}(\cdot,\cdot)$ and $\mathcal{F}$ satisfy the assumptions made in Theorem 1.1. Let X_h and Y_h be chosen, such that the sup-sup and inf-sup conditions are still valid uniformly. Then,*

$$|||x - x_h|||_X \leq (1 + \frac{\gamma_1}{\gamma_0}) \inf_{z_h \in Y_h} |||x - z_h|||_X.$$

For a proof of this lemma; see e.g. Braess [15], pp. 116-117. This application of Theorem 1.1 to the theory of finite elements is due to Babuška; see [12, 9, 11, 10].

The special case $X = Y$ (resp. $X_h = Y_h$) is also known as the Galerkin method.

Having derived a finite dimensional version of (1.1), we can also formulate it in a matrix/vector setting. Let $(\phi_j)_{j=1,\ldots,n_h}$ (resp. $(\psi_j)_{j=1,\ldots,n_h}$) be a basis of X_h (resp. Y_h), such that

$$x_h = \sum_{j=1}^{n_h} \xi_j \phi_j \text{ and } y_h = \sum_{j=1}^{n_h} \eta_j \psi_j.$$

From (1.5), we derive the matrix form

$$\mathcal{A}_h \xi = \mathcal{F}_h,$$

where $\mathcal{A}_h := (\mathcal{A}(\phi_i, \psi_j))_{i,j}, \mathcal{F}_h := (\mathcal{F}(\psi_j))_j$ and $\xi := (\xi_j)_j$, for $i, j = 1, \ldots, n_h$. The matrix $\mathcal{A}_h$ is called the *stiffness matrix* and the vector $\mathcal{F}_h$ called the *load vector*.

Chapter 2

Iterative Methods and Preconditioners

2.1 Iterative Methods for Indefinite Nonsymmetric Problems

We consider iterative methods for the solution of linear problems of the form

$$\mathcal{A}x = \mathcal{F},$$

where $\mathcal{A} : X \longrightarrow X'$, is a continuous symmetric operator on a Hilbert space X and $\mathcal{F} \in X'$. There exist many classical iterative methods to solve this problem and discretized versions thereof. Among them are the Richardson, Gauss-Seidel and Jacobi methods and the Chebyshev semi-iterative method. All these methods converge for positive definite symmetric problems, although the Richardson and the Jacobi method might require an adequate damping parameter. The Chebyshev and the Richardson algorithms converge even for some nonsymmetric problems but some spectral information is required to assure convergence, see e.g. Manteuffel [62]. More precisely, if we restrict ourselves to real operators, we require the spectrum to be contained in the right or left half plane. For the Chebyshev method, we then have to choose two parameters which depend on the knowledge of the convex hull of the spectrum of the nonsymmetric problem. The most common examples thereof are given by a positive real interval and by an ellipse that is contained in the right or left half plane and is symmetric with respect to the real axis. Additionally, it can easily be shown that Richardson's method fails to converge for all complex damping parameters if $\mathcal{A}$ has both, positive and negative real eigenvalues, as it is the case for symmetric indefinite problems.

Krylov space methods normally do not require any a priori spectral information and there are methods in this class especially designed for indefinite and nonsymmetric problems. We will therefore restrict ourselves to these methods.

2.1.1 Krylov Space Methods

The practical use of Krylov space methods for the solution of $\mathcal{A}x = \mathcal{F}$ are based on the assumption that the matrix-vector product $\mathcal{A}x$ is relatively inexpensive. The iterates of these methods are of the form

$$x_m := x_0 + p_{m-1}(\mathcal{A})r_0, \tag{2.1}$$

where $p_{m-1} \in \mathcal{P}_{m-1}$ is a polynomial of degree $m-1$ and $r_0 := \mathcal{F} - \mathcal{A}x_0$ is the initial residual associated with the initial guess x_0.Thus,

$$x_m \in x_0 + \mathcal{K}_m(r_0, \mathcal{A}),$$

where $\mathcal{K}_m(r_0, \mathcal{A}) := span\{r_0, \mathcal{A}r_0, \ldots, \mathcal{A}^{m-1}r_0\}$ is the Krylov space of dimension m. There are two common recipes to select the iterates. One is to minimize the residual $r_m := \mathcal{F} - \mathcal{A}x_m$ in a suitable norm, i.e.

$$\|r_m\| = \min_{x \in x_0 + \mathcal{K}_m(r_0, \mathcal{A})} \|\mathcal{F} - \mathcal{A}x\|.$$

The other is to require that the residuals satisfy a certain orthogonality condition,

$$(r_m, s) = 0 \quad \forall s \in \mathcal{S}_m,$$

where $\mathcal{S}_m$ is an appropriately chosen Krylov space of dimension m. The method of conjugate residuals and GMRES are examples for algorithms that minimize the residual while BI-CGSTAB is a method based on a combination of an orthogonality relation and a local minimization property.

At first glance, the construction of the iterates (2.1) seems to require long-term recurrences. It is clear that such an update in each iteration step could be quite expensive in terms of both, computing time and memory requirements with these costs increasing with m. When solving a problem, where $\mathcal{A}$ is symmetric, we are able to give a Krylov basis update that can be constructed from a three-term recurrence. More generally, Faber and

Manteuffel [45] show that $l-$term recurrences exist if $\mathcal{A}$ is $\mathcal{H}-$normal$(l-2)$, i.e. if

$$\mathcal{H}\mathcal{A} = (p(\mathcal{A}))^t\mathcal{H},$$

where $p \in \mathcal{P}_{l-2}$ is a polynomial of degree $l-2$ and $\mathcal{H}$ is a non-singular matrix. There is also another way of reducing the computing time and memory requirements of Krylov space methods using long-term recurrences: One fixes the dimension m of the Krylov space in advance. If the prescribed accuracy is not obtained after m iterations, then the method is restarted with the starting vector obtained from the previous cycle. Such a truncated version of GMRES is given in 2.1.4.

From the formula for the iterates (2.1), we can easily derive representations of the residuals $r_m := \mathcal{F} - \mathcal{A}x_m$ and of the errors $e_m := x - x_m$ with respect to the initial residual r_0 and the initial error e_0:

$$\begin{aligned} r_m &= (I - p_{m-1}(\mathcal{A})\mathcal{A})r_0 \\ &=: \tilde{p}_m(\mathcal{A})r_0, \\ e_m &= (I - p_{m-1}(\mathcal{A})\mathcal{A})e_0 \\ &=: \tilde{p}_m(\mathcal{A})e_0. \end{aligned}$$

For a more detailed discussion of Krylov space methods, we refer the reader to Bruaset [31].

2.1.2 Preconditioning

The performance of Krylov space methods can be often significantly improved by using appropriate preconditioners. Assume that we have two regular linear operators $\mathcal{B}_L$ and $\mathcal{B}_R$. Then, we understand by a preconditioned Krylov space method the original algorithm applied to

$$\mathcal{B}_L^{-1}\mathcal{A}\mathcal{B}_R^{-1}y = \mathcal{B}_L^{-1}\mathcal{F}, \quad \mathcal{B}_R^{-1}y = x.$$

The corresponding Krylov space is then $\mathcal{K}_m(\mathcal{B}_L^{-1}r_0, \mathcal{B}_L^{-1}\mathcal{A}\mathcal{B}_R^{-1})$. We make a distinction between left preconditioning ($\mathcal{B}_R = I$) and right preconditioning ($\mathcal{B}_L = I$). It often does not matter if left or right preconditioning is used as far as the cost of computation is regarded. However, for Krylov space methods that minimize the residual in the euclidean norm such as GMRES (cf. 2.1.4), the choice can make a difference. We write the preconditioned system in the form

$$\mathcal{A}^{(p)}x^{(p)} = \mathcal{F}^{(p)},$$

where

$$\mathcal{A}^{(p)} := \mathcal{B}_L^{-1}\mathcal{A}, \quad x^{(p)} := x, \quad \mathcal{F}^{(p)} := \mathcal{B}_L^{-1}\mathcal{F}$$

in the case of left preconditioning and

$$\mathcal{A}^{(p)} := \mathcal{A}\mathcal{B}_R^{-1}, \quad x^{(p)} := \mathcal{B}_R x, \quad \mathcal{F}^{(p)} := \mathcal{F}$$

for right preconditioning. Thus, the $k-$th residual $r_k^{(p)}$ is minimized in the euclidean norm over the Krylov space $\mathcal{K}(r_0^{(p)}, \mathcal{A}^{(p)})$.

In the case of right preconditioning, the quantity minimized is the euclidean norm of the residual of the unpreconditioned system, i.e.

$$\begin{aligned}
\|r_k^{(p)}\|_2 &= \|\mathcal{F}^{(p)} - \mathcal{A}^{(p)}x_k^{(p)}\|_2 \\
&= \|\mathcal{F} - \mathcal{A}x_k\|_2 \\
&= \|r_k\|_2.
\end{aligned}$$

In the case of left preconditioning, the norm of the residual $r_k^{(p)}$ might be a good approximation of the norm of the error of the unpreconditioned system, i.e. if $\mathcal{A}^{(p)} \approx I$, then

$$\begin{aligned}
\|r_k^{(p)}\|_2 &= \|\mathcal{F}^{(p)} - \mathcal{A}^{(p)}x_k^{(p)}\|_2 \\
&= \|\mathcal{A}^{(p)}(\mathcal{A}^{-1}\mathcal{F} - x_k)\|_2 \\
&\approx \|x - x_k\|_2.
\end{aligned}$$

The best choice of the preconditioners depends on the problem and on the Krylov space method applied to it. The construction of optimal preconditioners, i.e. of preconditioners that lead to iterative methods converging independently of parameters, such as the mesh size or the Lamé parameters, is a major concern of this thesis.

For some historical remarks on preconditioning, we refer to Golub and O'Leary [49].

2.1.3 The Conjugate Residual Method

The preconditioned conjugate gradient method (PCG) has gained great popularity for positive definite problems. A natural generalization for symmetric indefinite problems is the preconditioned conjugate residual method (PCR); see e.g. Ashby et al. [6], and Hackbusch [53]. We describe the PCR-method and give a convergence estimate that is determined by the condition number of the preconditioned linear system.

Algorithm 2.1

$Initialization:$

$$\begin{aligned}
r_0 &:= \mathcal{F} - \mathcal{A}x_0, \\
p_{-1} := s_{-1} := q_{-1} &:= 0, \\
p_0 &:= \hat{\mathcal{B}}^{-1} r_0, \\
q_0 &:= \mathcal{A}p_0, \\
b_{bot}^{old} &:= 1,
\end{aligned}$$

$Iteration:$

$$\begin{aligned}
\textit{while} \qquad \tfrac{\|r_m\|_2}{\|r_0\|_2} &> \epsilon \\
s_m &:= \hat{\mathcal{B}}^{-1} q_m, \\
b_{top} &:= (r_m, s_m)_2, \\
b_{bot} &:= (q_m, s_m)_2, \\
\lambda &:= b_{top}/b_{bot}, \\
x_{m+1} &:= x_m + \lambda p_m, \\
r_{m+1} &:= r_m - \lambda q_m, \\
t_m &:= \mathcal{A}s_m, \\
b_{top} &:= (t_m, s_m)_2, \\
\alpha_0 &:= b_{top}/b_{bot}, \\
b_{top} &:= (t_m, s_{m-1}), \\
\alpha_1 &:= b_{top}/b_{bot}^{old} \\
p_{m+1} &:= s_m - \alpha_0 p_m - \alpha_1 p_{m-1}, \\
/ * q_{m+1} &:= \mathcal{A}p_{m+1} * / \\
q_{m+1} &:= t_m - \alpha_0 q_m - \alpha_1 q_{m-1}, \\
p_{m+1} &:= p_{m+1}/\|p_{m+1}\|_2, \\
q_{m+1} &:= q_{m+1}/\|p_{m+1}\|_2, \\
s_{m-1} &:= s_m, \\
p_{m-1} &:= p_m, \\
p_m &:= p_{m+1}, \\
q_{m-1} &:= q_m, \\
q_m &:= q_{m+1}, \\
b_{bot}^{old} &:= b_{bot}, \\
m &:= m + 1,
\end{aligned}$$

end

The PCR-method is an algorithm to solve $\mathcal{A}x = \mathcal{F}$ with a symmetric indefinite operator $\mathcal{A}$ and a positive definite preconditioner $\hat{\mathcal{B}}$. We will give a stable version that is based on a three term recurrence; see Hackbusch [53], p. 270.

The implementation provided here, only requires one matrix-vector product per step with each of the operators $\mathcal{A}$ and $\hat{\mathcal{B}}^{-1}$. This is achieved at the expense of introducing three additional vectors $q_m := \mathcal{A}p_m$, $s_m := \hat{\mathcal{B}}^{-1}q_m$, and $t_m := \mathcal{A}s_m$. Here, the matrix-vector product $q_m := \mathcal{A}p_m$ can be computed from the three term recurrence for p_m.

Remark 2.1 *Attention should be paid to scaling: If we assume that $\hat{\mathcal{A}}$ is an optimal positive definite preconditioner but choose $\hat{\mathcal{B}} := c\hat{\mathcal{A}}$, $c \in \mathbf{R}^+$, then λ will grow in proportion to $\frac{1}{c^{m+1}}$. This can easily be seen by induction. This phenomenon seems to be well known but not discussed in the literature. The easiest way of fixing this problem is to normalize p_{m+1} in every iteration; we have done so in our implementation.*

We introduce the following notation

$$\kappa(\hat{\mathcal{B}}^{-1}\mathcal{A}) := \rho(\hat{\mathcal{B}}^{-1}\mathcal{A})\rho((\hat{\mathcal{B}}^{-1}\mathcal{A})^{-1}), \tag{2.2}$$

where $\rho(\hat{\mathcal{B}}^{-1}\mathcal{A})$ denotes the spectral radius of $\hat{\mathcal{B}}^{-1}\mathcal{A}$. The next theorem can be found in Hackbusch [53], p. 270. It gives an upper bound for the convergence rate of the PCR-method.

Theorem 2.1 *Let the regular matrix $\mathcal{A}$ be symmetric and $\hat{\mathcal{B}}$ be symmetric positive definite. Then the m-th iterate x_m of Algorithm 2.1 satisfies*

$$||\hat{\mathcal{B}}^{-1/2}(\mathcal{A}x^m - \mathcal{F})||_2 \leq \frac{2c^\mu}{1+c^{2\mu}}||\hat{\mathcal{B}}^{-1/2}(\mathcal{A}x^0 - \mathcal{F})||_2 \tag{2.3}$$

where $c := \frac{\kappa-1}{\kappa+1}$, $\kappa := \kappa(\hat{\mathcal{B}}^{-1}\mathcal{A})$ and $\frac{m}{2} - 1 < \mu \leq \frac{m}{2}$ $\forall \mu \in \mathbf{Z}$.

This estimate can be improved if $\mathcal{A}$ is only weakly indefinite, i.e. has only a few negative eigenvalues. In this case, it can be shown that the asymptotic convergence rate is determined by the square root of the condition number. The next theorem can be found in Hackbusch [53], p. 272. The bound depends on the number of negative eigenvalues k but for k fixed, we have the same asymptotic convergence rate as for the conjugate gradient method.

Theorem 2.2 *Let the regular matrix $\mathcal{A}$ be symmetric and $\hat{\mathcal{B}}$ be symmetric positive definite. Assume further that the number of negative eigenvalues k is bounded for $h \to 0$ and that there exist constants, $0 < \gamma_- \leq \Gamma_-$, and $0 < \gamma_+ \leq \Gamma_+$, such that all non-positive eigenvalues are contained in the interval $[-\Gamma_-, -\gamma_-]$ and all positive eigenvalues in $[\gamma_+, \Gamma_+]$ for all $h > 0$. Then the m-th iterate x_m of Algorithm 2.1 satisfies*

$$||\hat{\mathcal{B}}^{-1/2}(\mathcal{A}x^m - \mathcal{F})||_2 \leq 2 \left(\frac{1 + \Gamma_+/\gamma_-}{c}\right)^k \ c^m \ ||\hat{\mathcal{B}}^{-1/2}(\mathcal{A}x^0 - \mathcal{F})||_2 \qquad (2.4)$$

where $c := \frac{\sqrt{\tilde{\kappa}}-1}{\sqrt{\tilde{\kappa}}+1}$, $\tilde{\kappa} := \frac{\Gamma_+}{\gamma_+}$.

Remark 2.2 *Obviously, we can replace both, $\tilde{\kappa}$ and Γ_+/γ_-, by an upper bound of the condition number $\kappa(\hat{\mathcal{B}}^{-1}\mathcal{A}) := \rho(\hat{\mathcal{B}}^{-1}\mathcal{A})\rho((\hat{\mathcal{B}}^{-1}\mathcal{A})^{-1})$, where $\rho(\hat{\mathcal{B}}^{-1}\mathcal{A})$ is the spectral radius. Thus, in this special case when we only have a few negative eigenvalues, the convergence rate of the PCR-method is determined asymptotically by the square root of the condition number of the preconditioned system and by the number of negative eigenvalues.*

2.1.4 The Generalized Minimal Residual Method

We now consider a Krylov space method for general nonsymmetric matrices that is based on minimizing the residual. In contrast to the PCR method, the minimal residual algorithms for nonsymmetric problems typically use long term recurrences and work and storage requirements grow linearly with the number of iterations. Thus, it often becomes necessary to use restarted or truncated versions instead of running the full algorithm.

The best-known scheme of this type is GMRES, due to Saad and Schultz [71]. GMRES is a generalization of MINRES, see Paige and Saunders [66], to general nonsymmetric matrices, where the Arnoldi method replaces the symmetric Lanczos algorithm in the construction of the orthonormal basis for the Krylov space $\mathcal{K}_n(r_0, \mathcal{A})$; see Saad [70]. We will give a version of the algorithm using right preconditioning; see Bruaset [31], pp. 50-51. We have chosen right preconditioning for GMRES since then the quantity minimized is the euclidean norm of the residual of the original system; see 2.1.2.

For each iteration step, one matrix-vector product is required with the system matrix $\mathcal{A}$ as well as with the preconditioner $\hat{\mathcal{B}}^{-1}$. In addition, m

vectors have to be stored in the m-th iteration step and a small least squares problem has to be solved in each iteration.

Algorithm 2.2

$$
\begin{array}{l}
\textit{Initialization}:\\
\quad r_0 := \mathcal{F} - \mathcal{A}x_0,\\
\quad v_1 := r_0/\|r_0\|_2,\\
\textit{Iteration}:\\
\quad \textit{for } j = 1, 2, \ldots, m\\
\qquad q := \mathcal{A}\hat{\mathcal{B}}^{-1}v_j\\
\qquad \textit{for } i = 1, 2, \ldots, j\\
\qquad\quad h_{ij} := (q, v_i)_2\\
\qquad \textit{end}\\
\qquad w := q - \sum_{i=1}^{j} h_{ij}v_i\\
\qquad h_{j+1j} := \|w\|_2\\
\qquad v_{j+1} := w/h_{j+1j}\\
\quad \textit{end}\\
\quad \textit{Minimize } J(y_m) = \|\beta e_1 - \bar{H}_m y_m\|_2\\
\quad x_m := x_0 + \hat{\mathcal{B}}^{-1}V_m y_m,
\end{array}
$$

where e_i is the unit vector in the direction of the i-th coordinate direction, $V_m := (v_1, v_2, \ldots, v_m)$, $H_m := V_m^t \mathcal{A}\hat{\mathcal{B}}^{-1}V_m$ and

$$\bar{H}_m = \begin{pmatrix} H_m \\ h_{m+1m}e_m^t \end{pmatrix} \in \mathbf{R}^{m+1} \times \mathbf{R}^m.$$

If many iterations are needed to achieve the desired accuracy, GMRES can be restarted, i.e. the algorithm is repeated with $x_0 := x_m$. This procedure is repeated until the prescribed tolerance is obtained. The least squares problem that appears in the algorithm is normally solved by a QR factorization. For a more detailed analysis of GMRES, we refer the reader to the original work by Saad and Schultz [71].

We now give some convergence bounds for GMRES. We are mainly interested in a priori bounds based on properties of $\mathcal{A}\hat{\mathcal{B}}^{-1}$. Eigenvalue information of the preconditioned operator is generally not sufficient if the problem is nonsymmetric; see Nachtigal, Reddy, and Trefethen [63]. The best a priori bounds known today are based on the field of values of $\mathcal{A}\hat{\mathcal{B}}^{-1}$. For an arbitrary matrix $\mathcal{M} \in \mathbf{R}^{n\times n}$, the field of values is defined by

$$W(\mathcal{M}) := \{\frac{\bar{x}^t\mathcal{M}x}{\bar{x}^t x} \ : \ 0 \neq x \in \mathbf{C}^n\}.$$

The following theorem provides an upper bound for the convergence behavior of GMRES; see Starke [82].

Theorem 2.3 *The residuals computed by GMRES satisfy*

$$\frac{\|r_m\|_2}{\|r_0\|_2} \leq (1 - \tau\,\bar{\tau})^{m/2},$$

where

$$\begin{aligned} \tau &:= \min\{Re\, z \ : \ z \in W(\mathcal{A}\hat{\mathcal{B}}^{-1})\} \\ &= \min\{\frac{x^t \mathcal{A}\hat{\mathcal{B}}^{-1} x}{x^t x} \ : \ 0 \neq x \in \mathbf{R}^n\} \end{aligned}$$

and

$$\begin{aligned} \bar{\tau} &:= \min\{Re\, z \ : \ z \in W((\mathcal{A}\hat{\mathcal{B}}^{-1}))^{-1}\} \\ &= \min\{\frac{x^t (\mathcal{A}\hat{\mathcal{B}}^{-1})^{-1} x}{x^t x} \ : \ 0 \neq x \in \mathbf{R}^n\}. \end{aligned}$$

Let us point out that a well known bound given by Elman, cf. [41] Theorems 5.4 and 5.9, can be derived from Theorem 2.3; see Starke [82], p. 25. The bound due to Elman is

$$\frac{\|r_m\|_2}{\|r_0\|_2} \leq (1 - \frac{\lambda_{min}(\mathcal{M})^2}{\|\mathcal{A}\hat{\mathcal{B}}^{-1}\|_2^2})^{m/2},$$

where $\lambda_{min}(\mathcal{M})$ is the minimum eigenvalue of $\mathcal{M} := (\mathcal{A}\hat{\mathcal{B}}^{-1} + (\mathcal{A}\hat{\mathcal{B}}^{-1})^t)/2$. We observe that

$$\begin{aligned} \bar{\tau} &= \min\{\frac{x^t (\mathcal{A}\hat{\mathcal{B}}^{-1})^{-1} x}{x^t x} : 0 \neq x \in \mathbf{R}^n\} \\ &= \min\{\frac{y^t \mathcal{A}\hat{\mathcal{B}}^{-1} y}{y^t (\mathcal{A}\hat{\mathcal{B}}^{-1})^t \mathcal{A}\hat{\mathcal{B}}^{-1} y} : 0 \neq y \in \mathbf{R}^n\} \\ &\geq \frac{\tau}{\|\mathcal{A}\hat{\mathcal{B}}^{-1}\|_2^2}, \end{aligned}$$

and

$$\begin{aligned} \tau &= \min\{\frac{x^t \mathcal{A}\hat{\mathcal{B}}^{-1} x}{x^t x} : 0 \neq x \in \mathbf{R}^n\} \\ &= \min\{\frac{(x^t \mathcal{A}\hat{\mathcal{B}}^{-1} x)/2 + (x^t (\mathcal{A}\hat{\mathcal{B}}^{-1})^t x)/2}{x^t x} : 0 \neq x \in \mathbf{R}^n\} \\ &= \lambda_{min}(\mathcal{M}). \end{aligned}$$

Although, we have already mentioned that eigenvalue information, normally, is not sufficient to predict the convergence behavior of GMRES, there are some special cases where eigenvalue bounds give us some information.

When the preconditioned system is normal, we have the following

Theorem 2.4 *For GMRES applied to a normal matrix, we have*

$$\frac{||r_m||_2}{||r_0||_2} \leq \min_{p \in \mathcal{P}_m, p(0)=1} \max_{\lambda \in \sigma(\mathcal{A}\hat{\mathcal{B}}^{-1})} |p(\lambda)|,$$

where $\sigma(\mathcal{A}\hat{\mathcal{B}}^{-1})$ denotes the set of eigenvalues of $\mathcal{A}\hat{\mathcal{B}}^{-1}$.

For a proof, see Nachtigal, Reddy, and Trefethen [63], Theorem 2.

For diagonalizable systems, we have

Theorem 2.5 *If $\mathcal{A}\hat{\mathcal{B}}^{-1}$ is diagonalizable, i.e. there exists a regular matrix $\mathcal{Q}$, such that $\mathcal{A}\hat{\mathcal{B}}^{-1} = \mathcal{Q}\mathcal{D}\mathcal{Q}^{-1}$, $\mathcal{D} := diag\{\lambda_i\}$, where λ_i is an eigenvalue of $\mathcal{A}\hat{\mathcal{B}}^{-1}$, we have*

$$\frac{||r_m||_2}{||r_0||_2} \leq \kappa(\mathcal{Q}) \min_{p \in \mathcal{P}_m, p(0)=1} \max_{\lambda \in \sigma(\mathcal{A}\hat{\mathcal{B}}^{-1})} |p(\lambda)|,$$

where $\kappa(\mathcal{Q}) := ||\mathcal{Q}||_2 \, ||\mathcal{Q}^{-1}||_2$ is the condition number of $\mathcal{Q}$ and $\sigma(\mathcal{A}\hat{\mathcal{B}}^{-1})$ denotes the set of eigenvalues of $\mathcal{A}\hat{\mathcal{B}}^{-1}$.

For a reference to this theorem; see Saad and Schultz [71], Proposition 4.

Consequently, the convergence rate of GMRES can be bounded in terms of the eigenvalues of the preconditioned system when $\mathcal{A}\hat{\mathcal{B}}^{-1}$ is normal or close to normal. Unfortunately, nonsymmetric matrices are in practice rarely normal.

2.1.5 The Stabilized Bi-Conjugate Gradient Method

Another popular method for nonsymmetric indefinite problems is the stabilized bi-conjugate gradient method (BI-CGSTAB) due to van der Vorst [83]. An experimental comparison of BI-CGSTAB and GMRES (and other methods), is given in Schmid, Paffrath, and Hoppe [75]. Since there are no a priori bounds known to predict the convergence behavior of BI-CGSTAB, we only give the algorithm for BI-CGSTAB (with left preconditioning.)

BI-CGSTAB needs two matrix-vector products per iteration step with each of the linear operators $\mathcal{A}$ and $\hat{\mathcal{B}}^{-1}$. This is twice as many as for GMRES. However, unlike GMRES, it is based on a 3-term recurrence and thus requires less memory.

Algorithm 2.3

$$
\begin{aligned}
&\mathit{Initialization:}\\
&\quad \mathit{initial\ guess}\ x_0 \quad r_0 := \mathcal{F} - \mathcal{A}x_0,\\
&\quad \bar{r}_0\ \mathit{arbitrary\ such\ that}\\
&\quad (\bar{r}_0, r_0)_2 \neq 0,\ \ e.g.\ \bar{r}_0 = r_0\\
&\quad \rho_0 = \alpha = \omega_0 := 1\\
&\quad v_0 = p_0 := 0\\
&\mathit{Iteration:}\\
&\quad \mathit{while}\ \|r_i\|_2/\|r_0\|_2 > \epsilon\\
&\qquad \rho_i := (\bar{r}_0, r_{i-1})\\
&\qquad \beta := (\rho_i/\rho_{i-1})(\alpha/\omega_{i-1})\\
&\qquad p_i := r_{i-1} + \beta(p_{i-1} - \omega_{i-1}v_{i-1})\\
&\qquad y := \mathcal{A}p_i\\
&\qquad v_i := \hat{\mathcal{B}}^{-1}y\\
&\qquad \alpha := \rho_i/(\bar{r}_0, v_i)_2\\
&\qquad s := r_{i-1} - \alpha v_i\\
&\qquad z := \mathcal{A}s\\
&\qquad t := \hat{\mathcal{B}}^{-1}z\\
&\qquad \omega_i := (t,s)_2/(t,t)_2\\
&\qquad x_i := x_{i-1} + \alpha p_i + \omega_i s\\
&\qquad r_i := s - \omega_i t\\
&\qquad i := i+1\\
&\quad \mathit{end}
\end{aligned}
$$

2.2 Abstract Schwarz Methods

We now review an abstract framework for the construction of positive definite preconditioners. Such preconditioners provide essential tools in our design of preconditioners for symmetric indefinite problems. The framework includes domain decomposition and multilevel/multigrid techniques; it is named after H.A. Schwarz who, in 1869, published what is believed to be the first domain decomposition method.

For the time being, we consider the problem

$$\mathcal{A}(x,y) =< \mathcal{F}, y > \quad \forall y \in X,$$

where $\mathcal{A}(\cdot,\cdot)$ is a positive definite symmetric bilinear form and X is a Hilbert space. Let the exact solution of this variational problem be denoted by x^*.

We introduce a subspace splitting of X into $N+1$ subspaces $X_i \subset X,\ i = 0, 1, \ldots, N$, i.e.

$$X = X_0 + X_1 + \ldots + X_N. \tag{2.5}$$

This subspace splitting does not necessarily have to be a direct sum; in fact, there are important splittings where the representation of an element of X can be constructed in different ways out of elements of the subspaces X_i. For each subspace X_i, we introduce a positive definite symmetric bilinear form $\mathcal{B}_i(\cdot,\cdot)\ :\ X_i \times X_i \longrightarrow \mathbf{R}$, and an operator $\mathcal{T}_i\ :\ X \longrightarrow X_i$, defined by

$$\mathcal{B}_i(\mathcal{T}_i x, y) = \mathcal{A}(x, y) \quad \forall y \in X_i.$$

The $\mathcal{B}_i(\cdot,\cdot)$ can be regarded as an approximation of $\mathcal{A}(\cdot,\cdot)$ on $X_i \times X_i$ and the $\mathcal{T}_i$ as an approximate projection; $\mathcal{T}_i$ is an $\mathcal{A}-$orthogonal projection onto X_i when we use $\mathcal{B}_i(\cdot,\cdot) = \mathcal{A}(\cdot,\cdot)$.

We can derive an operator representation of the $\mathcal{T}_i$ by using the operators associated with the bilinear forms. As already done for $\mathcal{A}$ in Section 1.1, we define $\mathcal{B}_i : X_i \longrightarrow X_i'$ by

$$< \mathcal{B}_i x, y >= \mathcal{B}_i(x, y) \quad \forall y \in X_i.$$

From the definition of $\mathcal{T}_i$, we obtain

$$\mathcal{T}_i = \mathcal{B}_i^{-1}\mathcal{A}.$$

The product $\mathcal{B}_i^{-1}\mathcal{A}$ is well defined since $X_i \subset X$ implies $X' \subset X_i'$.

There are basically two different paradigms: The multiplicative and the additive Schwarz method. Some hybrid methods have also been constructed mixing these two basic ideas; see Dryja, Smith, and Widlund [38] and the references therein.

Multiplicative Schwarz Method (MSM)

The MSM is an iterative method for solving the operator equation

$$\mathcal{T}_{ms} x = \mathcal{G}_{ms}, \tag{2.6}$$

where $\mathcal{T}_{ms}$ is defined by

$$\mathcal{T}_{ms} := I - \prod_{i=0}^{N} (I - \mathcal{T}_{N-i})$$

and $\mathcal{G}_{ms}$ by

$$\mathcal{G}_{ms} := \mathcal{T}_{ms} x^*.$$

We note that $\mathcal{G}_{ms}$ can be computed without knowing the exact solution since the polynomial $\mathcal{T}_{ms}$ does not contain any constant terms. Therefore, we can compute $\mathcal{G}_i$ by applying $\mathcal{B}_i^{-1}$ to a known vector

$$\begin{aligned}\mathcal{G}_i &= \mathcal{T}_i x^* \\ &= \mathcal{B}_i^{-1} \mathcal{A} x^* \\ &= \mathcal{B}_i^{-1} \mathcal{F}.\end{aligned}$$

The operator $\mathcal{T}_{ms}$ is normally nonsymmetric. Problem 2.6 can be solved by using any Krylov space method for nonsymmetric problems, e.g. GMRES or Bi-CGSTAB; see Section 2.1. It is also possible to symmetrize $\mathcal{T}_{ms}$; see Dryja, Smith and Widlund [38] for three alternatives:

- $\mathcal{T}_{sms1} := I - \left(\prod_{i=0}^{N}(I - \mathcal{T}_i)\right)\left(\prod_{i=0}^{N}(I - \mathcal{T}_{N-i})\right)$,
- $\mathcal{T}_{sms2} := I - (I - \mathcal{T}_0)\cdots(I - \mathcal{T}_N)\cdots(I - \mathcal{T}_0)$,
- $\mathcal{T}_{sms3} := \mathcal{T}_{ms} + \mathcal{T}_{ms}^t$.

If the operators $\mathcal{T}_{smsi}, i = 1, 2, 3$, are used, the right hand side $\mathcal{G}_{smsi}$ has to be modified appropriately.

Additive Schwarz Method (ASM)

The ASM is an iterative method for solving

$$\mathcal{T}_{as} x = \mathcal{G}_{as}, \tag{2.7}$$

where $\mathcal{T}_{as}$ is defined by

$$\mathcal{T}_{as} := \sum_{i=0}^{N} \mathcal{T}_i,$$

and $\mathcal{G}_{as}$ by

$$\mathcal{G}_{as} := \sum_{i=0}^{N} \mathcal{G}_i, \quad \mathcal{G}_i = \mathcal{T}_i x^*.$$

As in the multiplicative case, we can compute $\mathcal{G}_{as}$ without knowing the exact solution x.

Algorithms

The most prominent members of the family of Schwarz methods are domain decomposition and multigrid methods. Since this field consists of a large number of topics, we will not try to give a complete description of these methods but will only try to give a short overview of the two main types of methods. Normally, a domain decomposition method does not have to be a two-level algorithm. But without a second level which corresponds to a global communication operator, the performance of the methods often deteriorates when the number of subdomains is increased. We denote by Ω the domain, on which the differential problem is defined.

Two-Level Methods

A two-level method corresponds to a space decomposition into at least two spaces, i.e.

$$X = X_0 + X^{(1)}.$$

Here, the space X_0 is a coarse space that is used to provide a mechanism to transmit global information. It can for example be a finite element space on a coarse triangulation of the original domain Ω. The space $X^{(1)}$ could then correspond to a fine grid triangulation obtained from the coarse one by refinement; see Figure 2.1.

One possibility to construct a domain decomposition method, is to decompose the domain Ω into several overlapping subdomains $\Omega_i,\ i = 1, \ldots, N$, such that $\Omega = \bigcup_{i=1}^N \Omega_i$. For each subdomain Ω_i, we introduce a subspace X_i, such that $X^{(1)} = \sum_{i=1}^N X_i$.

To obtain a domain decomposition into overlapping subdomains, one can first divide Ω into non-overlapping subdomains Ω_i' and then extend them elementwise by the desired overlap; see Figure 2.1.

For overlapping Schwarz methods using nonnested spaces and unstructured grids, see Cai [32], Chan and Smith [36], and Chan, Smith and Zou [35]. There are also non-overlapping domain decomposition methods that can be analyzed within the abstract Schwarz framework. Since we are not using these methods in this thesis, we refer to Dryja, Smith, and Widlund [38] for further discussions and references.

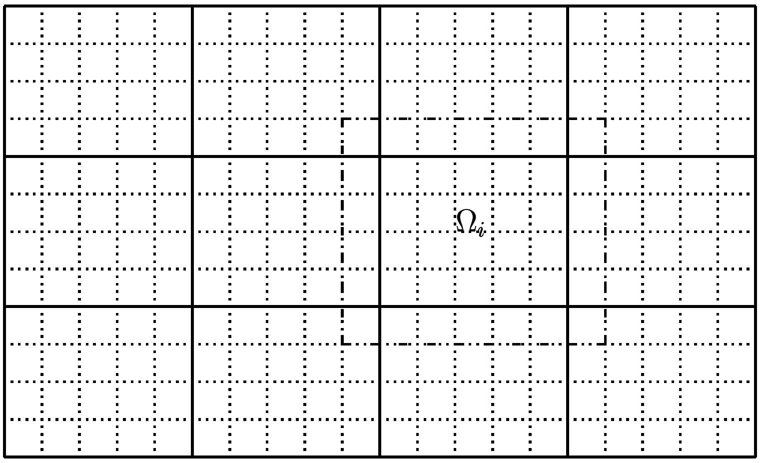

Figure 2.1: A two-level mesh and an overlapping subdomain.

Multilevel Methods

A natural extension of two-level methods can be obtained by introducing several coarse spaces or levels, such that we get the subspace splitting

$$X = X_0 + X^{(1)} + X^{(2)} + \ldots + X^{(l)},$$

where we assume that $X^{(l)}$ represents the finest level and X_0 the coarsest one. As in the two-level case, we could start from a coarse finite element triangulation and successively refine. For each triangulation obtained by refinement, we introduce a subspace $X^{(i)}$. By applying the multiplicative Schwarz method, the classical multigrid algorithm is obtained.

It is now possible to combine domain decomposition techniques with a multilevel splitting by decomposing the triangulation corresponding to the spaces $X^{(j)}$. Obviously, we obtain

$$X = X_0 + \sum_{j=1}^{l} \sum_{i=1}^{N_j} X_i^{(j)},$$

see e.g. Zhang [91, 90] or Smith, Bjørstad, and Gropp [81].

There are also hybrid methods, combining domain decomposition and multilevel/-grid techniques to improve performance, e.g. on parallel computers; see Griebel [50].

Convergence Theory

The convergence theory can be based on three parameters which measure the interactions between the subspaces X_i and the properties of the local operators $\mathcal{B}_i$. Before we discuss these parameters, let us present an important lemma.

Lemma 2.1 *The additive Schwarz operator is invertible and we have*

$$\mathcal{A}(\mathcal{T}_{as}^{-1}x, x) = \min_{\substack{x=\sum_{i=0}^{N} x_i \\ x_i \in X_i}} \sum_{i=0}^{N} \mathcal{B}_i(x_i, x_i).$$

For a proof, see Zhang [90], p.20, or [91].

We make the following

Assumptions

(i) Let C_0 be the smallest constant, such that for all $x \in X$ there exists a representation $x = \sum_{i=0}^{N} x_i, \;\; x_i \in X_i$, with

$$\sum_{i=0}^{N} \mathcal{B}_i(x_i, x_i) \leq C_0^2 \, \mathcal{A}(x, x).$$

(ii) Let the following strengthened Cauchy-Schwarz inequalities hold:

$$|\mathcal{A}(x_i, x_j)| \leq \epsilon_{ij} (\mathcal{A}(x_i, x_i))^{1/2} (\mathcal{A}(x_j, x_j))^{1/2} \quad \forall x_i \in X_i, x_j \in X_j,$$

for $i, j = 1, 2, \ldots, N$ and $0 \leq \epsilon_{ij} \leq 1$. The spectral radius $\rho(\mathcal{E})$ of the matrix $\mathcal{E} := (\epsilon_{ij})_{i,j}$ can be interpreted as a measure of the orthogonality of the subspaces X_i. Note that the coarse space X_0 is excluded.

(iii) Let $\omega > 0$ be the smallest constant, such that

$$\mathcal{A}(x, x) \leq \omega \mathcal{B}_i(x, x) \qquad \forall x \in X_i, \quad i = 0, 1, \ldots, N.$$

Obviously, we can always choose $\omega := \max_i \|\mathcal{B}_i\|_{\mathcal{A}}$. Note that it is always possible to scale the local operators, such that $1 \leq \omega < 2$, although this will also affect the value of C_0.

Subspace splittings satisfying these three assumption, are called stable; see Oswald [65]. The importance of these three parameters becomes clear in Theorem 2.6.

Theorem 2.6

$$\kappa(\mathcal{T}_{as}) \leq C_0^2 \, \omega \, (\rho(\mathcal{E}) + 1).$$

Proof: From the usual Rayleigh-quotient argument, we obtain

$$\begin{aligned}
\lambda_{min}^{-1}(\mathcal{T}_{as}) &= \lambda_{max}(\mathcal{T}_{as}^{-1}) \\
&= \max_{x \in X} \frac{\mathcal{A}(\mathcal{T}_{as}^{-1} x, x)}{\mathcal{A}(x, x)} \\
&= \max_{x \in X} \min_{\substack{x = \sum_{i=0}^{N} x_i \\ x_i \in X_i}} \frac{\sum_{i=0}^{N} \mathcal{B}_i(x_i, x_i)}{\mathcal{A}(x, x)}.
\end{aligned}$$

Analogously, we obtain

$$\lambda_{max}^{-1}(\mathcal{T}_{as}) = \min_{x \in X} \min_{\substack{x = \sum_{i=0}^{N} x_i \\ x_i \in X_i}} \frac{\sum_{i=0}^{N} \mathcal{B}_i(x_i, x_i)}{\mathcal{A}(x, x)}.$$

Here, we have used that $\mathcal{T}_{as}$ is symmetric with respect to $\mathcal{A}(\cdot, \cdot)$. Combining these two formulas with Assumptions (1)-(3), we get

$$\begin{aligned}
\lambda_{min}(\mathcal{T}_{as}) &\geq C_0^{-2}, \\
\lambda_{max}(\mathcal{T}_{as}) &\leq \omega(\rho(\mathcal{E}) + 1).
\end{aligned}$$

$\square$

The next convergence result on the unaccelerated multiplicative Schwarz method is due to Bramble, Pasciak, Wang, and Xu [22], see also Dryja and Widlund [39].

Theorem 2.7

$$\|\mathcal{T}_{ms}\|_{\mathcal{A}} \leq \left(1 - \frac{(2 - \hat{\omega})}{(1 + 2\hat{\omega}^2 \rho(\mathcal{E})^2 C_0^2}\right)^{1/2},$$

where $\hat{\omega} := \max\{1, \omega\}$.

The symmetric multiplicative Schwarz operator $\mathcal{T}_{sms1}$ can also be analyzed; cf. Dryja and Widlund [39].

Theorem 2.8
$$\kappa(\mathcal{T}_{sms1}) \leq \frac{(1 + 2\hat{\omega}^2 \rho(\mathcal{E})^2) C_0^2}{2 - \hat{\omega}},$$
where $\hat{\omega} := \max\{1, \omega\}$.

For a more detailed discussion on Schwarz methods, we refer to Smith, Bjørstad, and Gropp [81] and Oswald [65].

In our applications, we will assume that the parameters $C_0, \rho(\mathcal{E})$, and ω are independent of the mesh size h_j, the diameter of a typical finite element on the discretization level j, and of the number of levels. Such a result holds, for example, for the overlapping Schwarz method with generous overlap, i.e. an overlap of order H, the diameter of a typical subdomain. Another method that fulfills this optimality condition has been proposed by Smith [79]. This algorithm is based on the substructuring idea; the overlap is implemented on the edges and faces of the subdomains.

But there are also other interesting cases, where the convergence rate depends only weakly on the discretization parameters, e.g. (poly)logarithmically; see e.g. Dryja, Smith, and Widlund [38] and the references therein.

Chapter 3

Saddlepoint Problems with Penalty Term

3.1 Introduction

An important area, where saddlepoint problems arise, is fluid dynamics. Typical examples are the Navier-Stokes and the Stokes equations for incompressible flow. Saddlepoint or mixed formulations can also be of interest for scalar second order elliptic equations. A reformulation as a saddlepoint problem allows us for example to compute the gradient of the solution directly and not by numerical differentiation which could lead to a loss of accuracy. A third application area of saddlepoint problems is the field of mechanics. An important example is the Hellinger-Reissner principle of linear elasticity; cf. Braess [15], pp. 245-246. Saddlepoint reformulations also turn out to be a remedy for problems from solid mechanics that suffer from locking. Here, we will only give a short and informal definition of locking. For a detailed mathematical discussion of the locking phenomenon, we refer to Babuška and Suri [7, 8]. Engineers denote by locking "a condition of excessive stiffness of a particular deformation state", see MacNeal [60], p. 204. In other words, a finite element is said to lock when the results obtained by the finite element method are significantly smaller than the exact ones. Locking can occur when a parameter, e.g. the Poisson ratio of a material, approaches a critical limit and a (low order) conforming finite element approach is used.

In recent years, modern iterative methods, e.g. domain decomposition and multigrid methods, have been applied to parameter dependent problems arising in solid mechanics; see Braess [14], Braess and Blömer [16], Jung

[56], and Smith [78], p. 68 and Table 4.12. We note that one has to make a distinction between the convergence rate of the finite element model and the convergence rate of the iterative method. Both convergence rates can deteriorate severely when the limit of a certain parameter is approached, e.g. when the Poisson ratio tends to 1/2 in the problem of linear elasticity; see Braess [14] and Jung [56]. The loss of convergence can be explained as a problem of ill conditioning; see Braess [15], pp. 253-254.

There are different approaches to overcome the problem of locking in the finite element model; nonconforming finite element methods, reduced/selected integration and a reformulation in terms of a saddle point problem with a penalty term. Many of them can be analyzed as saddle point problems with a penalty term; see Braess [15], Brenner [24, 25], Brenner and Scott [26], Brezzi and Fortin [30], and Hughes [55]. For all of these approaches it can be proven, that the finite element solution converges uniformly with respect to the penalty parameter. There are still differences between these methods as far as the iterative solution of the resulting linear systems is concerned. Thus it has been observed in Braess and Blömer [16] that the mixed formulation is better suited for multigrid methods than reduced/selected integration; the convergence rate of the iterative method, considered by these authors, for the linear system arising from the latter model still deteriorates.

The remainder of this chapter is organized as follows. We first describe an abstract framework for saddle point problems with a penalty term. This framework is based on the famous theory of Babuška and Brezzi and it gives sufficient conditions to guarantee that Babuška's inf-sup and sup-sup conditions introduced in 1.1 are satisfied. Finally, we give some examples arising in solid mechanics.

3.2 The Abstract Framework

Let $(V, \|\cdot\|_V)$ and $(M, \|\cdot\|_M)$ be two Hilbert spaces, let M_c be a dense subspace of M, and let

$$a(\cdot,\cdot)\colon V\times V\to\mathbf{R},\ b(\cdot,\cdot)\colon V\times M_c\to\mathbf{R},\ c(\cdot,\cdot)\colon M_c\times M_c\to\mathbf{R}, \qquad (3.1)$$

be three continuous bilinear forms. Additionally, introduce V_0, a subspace of V, given by $V_0 := \{v\in V : b(v,q) = 0\ \forall q\in M\}$. We assume that $a(\cdot,\cdot)$ is V_0-elliptic and $V-$positive semi-definite and that $c(\cdot,\cdot)$ is symmetric M_c-

positive semi-definite. We consider the following problem:
Find $(u,p) \in V \times M_c$, such that

$$\begin{array}{rclcll} a(u,v) & + & b(v,p) & = & < f, v > & \forall v \in V \\ b(u,q) & - & t^2 c(p,q) & = & < g, q > & \forall q \in M_c \quad t \in [0,1]. \end{array} \tag{3.2}$$

We denote by X the product space $V \times M_c$, and by

$$\mathcal{A}(x,y) := a(u,v) + b(u,q) + b(v,p) - t^2 c(p,q),$$
$$x = (u,p) \in X, \; y = (v,q) \in X,$$

the bilinear form of problem (3.2) on X. Introducing

$$\mathcal{F}(y) :=< f, v > + < g, q >,$$

we obtain an equivalent formulation of problem (3.2):

$$\mathcal{A}(x,y) = \mathcal{F}(y) \quad \forall y \in X. \tag{3.3}$$

Assuming that we have an additional norm on M_c, $|||\cdot|||_M$, we introduce a new norm on X by

$$|||x||| := \|u\|_V + |||p|||_M \text{ for } x = (u,p) \in X.$$

If the bilinear form $c(\cdot,\cdot)$ is continuous on $M \times M$, let

$$|||p|||_M := \|p\|_M. \tag{3.4}$$

Otherwise, $|||p|||_M$ is defined by

$$|||p|||_M := \|p\|_M + t|p|_c, \tag{3.5}$$

where $|p|_c := \sqrt{c(p,p)}$ is a semi-norm on M_c.

We now apply the abstract theory given in Section 1.1 to our saddlepoint problem with a penalty term. Accordingly, we have to verify a sup-sup and an inf-sup condition to guarantee the well-posedness of the problem. Let us point out that for reasons of symmetry, we only have one inf-sup condition; see Section 1.1. From the assumptions, we can conclude that $\mathcal{A}(\cdot,\cdot)$ is a continuous bilinear form on X, i.e.

$$\sup_{y \in X} \sup_{x \in X} \frac{\mathcal{A}(x,y)}{|||x||| \, |||y|||} \leq \gamma_1, \tag{3.6}$$

where $\gamma_1 > 0$ is independent of $t \in [0,1]$. Additionally, $\mathcal{A}(\cdot,\cdot)$ has to fulfill an inf-sup condition,

$$\inf_{y\in X}\sup_{x\in X}\frac{\mathcal{A}(x,y)}{|||x|||\,|||y|||} \geq \gamma_0 > 0, \tag{3.7}$$

where γ_0 is independent of $t \in [0,1]$.

Theorem 3.1 *Let the following three assumptions be satisfied:*

(i) The continuous bilinear form $a(\cdot,\cdot)$ is V_0-elliptic, i.e.

$$\exists\alpha_0 > 0, \text{ such that } a(v,v) \geq \alpha_0\|v\|_V^2 \quad \forall v \in V_0$$

and $V-$positive semi-definite, i.e.

$$a(v,v) \geq 0, \quad \forall v \in V,$$

(ii) The continuous bilinear form $b(\cdot,\cdot)$ fulfills an inf-sup condition, i.e.

$$\exists\beta_0 > 0, \text{ such that } \inf_{q\in M_c}\sup_{v\in V}\frac{b(v,q)}{\|v\|_V\|q\|_M} \geq \beta_0$$

(iii) The continuous bilinear form $c(\cdot,\cdot)$ is symmetric and M_c-positive semi-definite, i.e.

$$c(q,q) \geq 0 \quad \forall q \in M_c.$$

Then, the inf-sup condition (3.7) holds if in addition one of the following conditions is satisfied:

1) The bilinear form $c(\cdot,\cdot)$ is continuous on $M \times M$.

2) The bilinear form $a(\cdot,\cdot)$ is $V-$elliptic.

Proof: Let us first assume that condition 1) holds. Then, we define the norm $|||\cdot|||_M$ by $|||p|||_M := \|p\|_M$; see (3.4). The proof that the inf-sup condition (3.7) holds, can be found in Braess and Blömer [16].

Now, let us assume that 2) is fulfilled. In this case, we define $|||p|||_M := \|p\|_M + t|p|_c$, see (3.5). For a special formulation of the Mindlin-Reissner plate a proof for (3.7) can be found in Huang [54], Lemma 3.1. The arguments given there immediately carry over to our more general setting.

$\square$

All these results are also valid for suitable finite element spaces; see Section 3.3. We then require, additionally, that the constants in Theorem 3.1 are independent of h. The continuity assumptions turn into uniform boundedness with respect to h; see e.g. Braess [15]. We note that there exist also useful methods for which β_0 goes to zero slowly with h, cf. Brezzi and Fortin [30].

3.3 Examples

We now discuss some problems from solid mechanics that can be treated within this abstract framework. We denote the finite element spaces approximating V and M (resp. M_c) by V^h and M^h (resp. M_c^h).

We introduce the following product for matrices

$$\sigma : \tau \quad := \quad \sum_{i=1}^{d}\sum_{j=1}^{d} \sigma_{ij}\tau_{ij},$$

where $\sigma, \tau \in \mathbf{R}^{d\times d}$.

In the context of saddlepoint problems, u, v (resp. p, q) will always denote vector valued (resp. scalar) functions. We use Sobolev spaces defined by:

$$\begin{aligned} H^1_\Gamma(\Omega) &:= \{v \in H^1(\Omega) : v_{|\Gamma_0} = 0\}, \\ H^1_0(\Omega) &:= H^1_\Gamma(\Omega) \text{ with } \Gamma_0 = \partial\Omega, \end{aligned}$$

where Γ_0 denotes the part of the boundary where a homogeneous Dirichlet condition is imposed. The finite element spaces are defined by:

$$\begin{aligned} \mathcal{M}^k(\mathcal{T}_h) &:= \{v \in L_2(\Omega) : v_{|T} \in \mathcal{P}^k \text{ for all } T \in \mathcal{T}_h\} \\ \mathcal{M}^k_0(\mathcal{T}_h) &:= \mathcal{M}^k(\mathcal{T}_h) \cap H^1(\Omega), \\ \mathcal{M}^k_{0,0}(\mathcal{T}_h) &:= \mathcal{M}^k(\mathcal{T}_h) \cap H^1_0(\Omega), \\ \mathcal{B}^3(\mathcal{T}_h) &:= \{v \in \mathcal{M}^3_0(\mathcal{T}_h) : v \text{ vanishes on the boundary of every element } \}, \\ \mathcal{N}^1_0(\mathcal{T}_h) &:= \mathcal{M}^1_0(\mathcal{T}_h) \oplus \mathcal{B}^3(\mathcal{T}_h). \end{aligned}$$

Here $\mathcal{T}_h$ is a triangulation of Ω and $\mathcal{P}^k$ the space of polynomials of total degree $\leq k$.

The differential operators are defined by:

$$div(v) \quad := \quad \sum_{i=1}^{d} \frac{\partial v_i}{\partial x_i},$$

$$\begin{aligned}
\nabla p &:= \left(\frac{\partial p}{\partial x_i}\right)_{i\in\{1,\ldots,d\}}, \\
\nabla v &:= (\nabla v_i)_{i=1,\ldots,d}, \\
\epsilon(v) &:= \frac{1}{2}\left(\nabla v + (\nabla v)^t\right).
\end{aligned}$$

In two dimensions, we set

$$\begin{aligned}
rot(v) &:= -\frac{\partial v_1}{\partial x_2} + \frac{\partial v_2}{\partial x_1}, \\
curl(q) &:= (-\frac{\partial q}{\partial x_2}, \frac{\partial q}{\partial x_1})^t.
\end{aligned}$$

All of these operators can be defined element by element on the space $\mathcal{M}^k(\mathcal{T}_h)$. The resulting discrete operators are marked by a subscript h.

We make use of the following inner products,

$$\begin{array}{ll}
(\varepsilon(v), \varepsilon(u))_0 := \int_\Omega \varepsilon(v) : \varepsilon(u) dx, & (v,u)_0 := \int_\Omega vu dx, \\
< f, v >:= \int_\Omega f v dx + \int_{\Gamma_1} g_1 v dx, & \Gamma_1 := \partial\Omega \setminus \Gamma_0.
\end{array}$$

3.3.1 The Equations of Linear Elasticity

An example of a saddlepoint problem with a penalty term arises from the displacement formulation of the equations of linear elasticity. The equations of linear elasticity model the displacement of an elastic material under the action of some external and internal forces. Denote the elastic body by $\Omega \subset \mathbf{R}^d, d = 2, 3$, and its boundary by Γ. Assume one part of the boundary, Γ_0, to be clamped, i.e. we impose Dirichlet boundary conditions, and let the other part of the boundary, $\Gamma_1 := \Gamma \backslash \Gamma_0$, be subject to a surface force of density g_1. We also introduce an internal volume force f, e.g. gravity; see Figure 3.1.

The variational formulation is given by

$$2\,\mu\,(\varepsilon(u), \varepsilon(v))_0 + \lambda\,(\mathrm{div}u, \mathrm{div}v)_0 =< f, v > \quad \forall v \in V := (H^1_\Gamma(\Omega))^d. \qquad (3.8)$$

The ellipticity of problem (3.8) follows from Korn's inequality.

Theorem 3.2 (Korn's inequality) *Let $\Omega \subset \mathbf{R}^d, d = 2, 3$, be an open and bounded domain with piecewise smooth boundary and assume that $\Gamma_0 \neq \emptyset$. Then there exists a positive constant $c = c(\Omega, \Gamma_0)$, such that*

$$(\varepsilon(v), \varepsilon(v))_0 \geq c\,\|v\|_1^2 \quad \forall v \in (H^1_\Gamma(\Omega))^d.$$

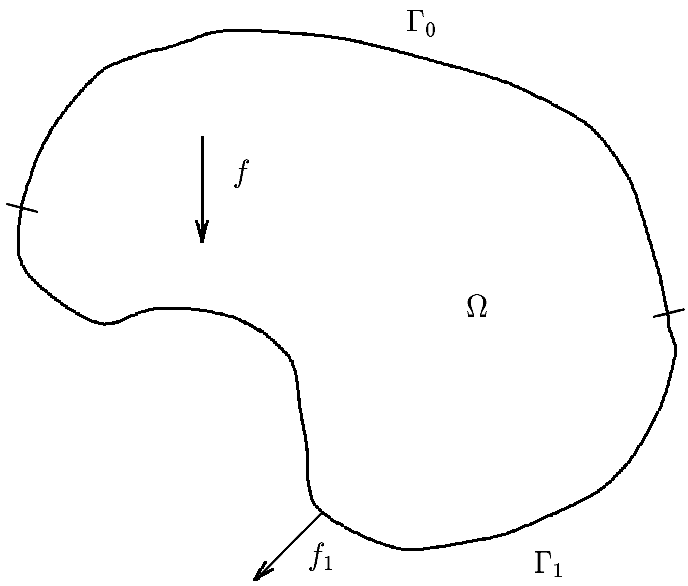

Figure 3.1: An elastic body.

For a proof, see e.g. Braess [15], p. 248.

Existence and uniqueness of problem (3.8) follow immediately from the Lax-Milgram Lemma; see Section 1.1.

Instead of using the Lamé constants λ and μ, we can also work with Young's elasticity modulus E and the Poisson ratio ν. These parameters are related to each other by the following equations

$$
\begin{array}{rclcrcl}
\lambda & = & \frac{E\nu}{(1+\nu)(1-2\nu)} & , & \nu & = & \frac{\lambda}{2(\lambda+\mu)}, \\
\mu & = & \frac{E}{2(1+\nu)} & , & E & = & \frac{\mu(3\lambda+2\mu)}{\lambda+\mu}.
\end{array}
\tag{3.9}
$$

Some materials, e.g. rubber, are nearly incompressible, i.e. small changes in the density of the material lead to a rapid growth of the energy. Almost incompressible materials are characterized by a Poisson ratio near $\frac{1}{2}$. In terms of the Lamé constants, this means that λ tends to infinity. This leads to the phenomenon of locking if a low order finite element model is used in a pure displacement setting. To obtain a finite element model that converges independently of locking as $h \to 0$, we reformulate the pure displacement model as a saddle point problem with a penalty term; see Brezzi and Fortin [30], and Braess [15]. We introduce a new variable $p := \lambda \operatorname{div} u$ and obtain

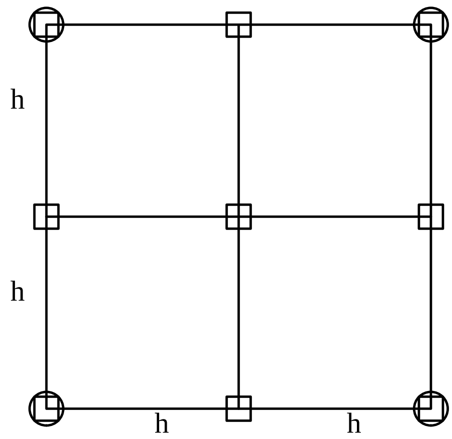

Figure 3.2: The Taylor-Hood Element: □ denotes the points of the displacement discretization, ◯ the ones of the Lagrange multiplier.

from (3.8)

$$\begin{array}{rcrcl} 2\,\mu\,(\varepsilon(u),\varepsilon(v))_0 & + & (\mathrm{div} v, p)_0 & = & <f,v> \qquad \forall v \in V := (H^1_\Gamma(\Omega))^d, \\ (\mathrm{div} u, q)_0 & - & \frac{1}{\lambda}\,(p,q)_0 & = & 0 \qquad \forall q \in M := L_2(\Omega). \end{array} \tag{3.10}$$

Remark 3.1 *If we have homogeneous Dirichlet conditions on the whole boundary it is possible to use an equivalent simpler formulation; see Brezzi and Fortin [30], p.201,*

$$\begin{array}{rcll} \mu\,(\nabla u, \nabla v)_0 + (\mathrm{div} v, p)_0 & = & <f,v> & \forall v \in V := (H^1_\Gamma(\Omega))^d, \\ (\mathrm{div} u, q)_0 - \dfrac{1}{\lambda+\mu}\,(p,q)_0 & = & 0 & \forall q \in M := L_2(\Omega). \end{array}$$

We note that in the case of pure Dirichlet boundary conditions, the pressure is only determined up to an additive constant. A remedy is, as known from the Stokes problem, to replace $M := L_2(\Omega)$ by $L_2(\Omega)/\mathbf{R}$.

We know from Korn's inequality that the bilinear form $(\varepsilon(u),\varepsilon(v))_0$ is V- elliptic. Thus (3.10) is very similar to the Stokes problem and we can use finite elements developed for that problem. For simplicity, we restrict ourselves to the Taylor-Hood element; see Fig. 3.2. For the displacements u, we use piecewise bilinear polynomials on quadrilaterals on a grid with the meshsize h and for the Lagrange parameter p piecewise bilinear polynomials on quadrilaterals with mesh size $2h$. The corresponding finite element spaces are

$$\begin{array}{rcl} V^h & := & \{v_h \in (\mathcal{C}(\Omega))^d \cap V : v_{h|T} \in \mathcal{Q}_1,\ T \in \tau_h\}, \\ M^h & := & \{q_h \in \mathcal{C}(\Omega) \cap M : q_{h|T} \in \mathcal{Q}_1,\ T \in \tau_{2h}\}. \end{array} \tag{3.11}$$

For a proof that the inf-sup condition holds for $b(\cdot,\cdot)$ in this case; see Verfürth [84], Girault and Raviart [48], or Brezzi and Fortin [30]. All the conditions for Theorem 3.1 are satisfied and the finite element method converges independently of the Lamé constants (resp. the Poisson ratio). There is of course no restriction to assume that $\frac{1}{\lambda} =: t^2$ is in $[0,1]$ since we are mainly interested in the nearly incompressible case.

3.3.2 The Timoshenko Beam Problem

Let $\Omega := I \subset \mathbf{R}$ be a finite interval and $f \in L_2(I)$. The mixed Ansatz for the Timoshenko beam is given by

$$(\theta', \psi')_0 + (v' - \psi, \gamma)_0 = (f, v)_0 \quad \forall (v,\psi) \in V, \tag{3.12}$$

$$(w' - \theta, \eta)_0 - t^2(\gamma, \eta)_0 = 0 \qquad \forall \eta \in M_c, \tag{3.13}$$

with $V := (H_0^1(I))^2$ and $M_c = M := L_2(I)$. Equations (3.12) and (3.13) represent a saddlepoint problem with a penalty term, in the sense of Theorem 3.1, with

$$\begin{array}{rcl} a((w,\theta),(v,\psi)) & := & (\theta', \psi')_0, \\ b((v,\psi),\eta) & := & (v' - \psi, \eta)_0 \\ c(\gamma,\eta) & := & (\gamma,\eta)_0. \end{array}$$

We obtain the finite element formulation of the Timoshenko beam by replacing $V \times M := (H_0^1(I))^2 \times L_2(I)$ with $V^h \times M^h := \left(\mathcal{M}_{0,0}^k(\mathcal{T}_h)\right)^2 \times \mathcal{M}^{k-1}(\mathcal{T}_h),\ k \geq 2$; see Arnold [1], Braess [15], p.278, and Braess and Blömer [16]. The proof that the assumptions of Theorem 3.1 hold, can also be found in these references.

3.3.3 The Mindlin-Reissner Plate Problem

Let $\Omega \subset \mathbf{R}^2$ be a polygonal domain and let $f \in L_2(\Omega)$. Brezzi and Fortin [29] consider a formulation of the Mindlin-Reissner plate problem which contains the following mixed Ansatz as a subproblem:

Find $(\psi, q) \in (H_0^1(\Omega))^2 \times H^1(\Omega)/\mathbf{R}$, such that

$$\begin{aligned} a(\theta,\psi) + (rot(\psi), p)_0 &= (f,\psi)_0 \quad \forall \psi \in \left(H_0^1(\Omega)\right)^2, \quad (3.14)\\ (rot(\theta), q)_0 - t^2(curl(p), curl(q))_0 &= 0 \qquad\qquad \forall q \in H^1(\Omega)/\mathbf{R}, \quad (3.15)\end{aligned}$$

with $a(\theta,\psi) := \int_\Omega \left\{2\mu\epsilon(\theta) : \epsilon(\psi) + \frac{\lambda}{2}\frac{2\mu}{\lambda+2\mu} div(\theta)\, div(\psi)\right\} dx$. The constants μ and λ denote the Lamé parameters and t represents the thickness of the plate.

By setting $x^\perp := (-x_2, x_1)$ for $x \in \mathbf{R}^2$, we have

$$rot(\psi) = div(\psi^\perp).$$

By using the definition of $curl(p)$, we obtain an equivalent Stokes problem with a penalty term from (3.14), (3.15):

$$\begin{aligned} a(\theta^\perp,\psi^\perp) + (div(\psi^\perp), p)_0 &= (f,\psi^\perp)_0 \quad \forall \psi^\perp \in \left(H_0^1(\Omega)\right)^2,\\ (div(\theta^\perp), q)_0 - t^2(p,q)_1 &= 0 \qquad\qquad \forall q \in H^1(\Omega)/\mathbf{R}.\end{aligned}$$

The finite element formulation is obtained by replacing $V \times M_c := (H_0^1(\Omega))^2 \times H^1(\Omega)/\mathbf{R}$ with $V^h \times M_c^h := (\mathcal{N}_{0,0}(\mathcal{T}_h))^2 \times \mathcal{M}_0^1(\mathcal{T}_h)/\mathbf{R}$, where $\mathcal{N}_{0,0}(\mathcal{T}) := \mathcal{M}_{0,0}^1(\mathcal{T}_h) \oplus \mathcal{B}^3(\mathcal{T}_h)$; see Arnold, Brezzi, and Fortin [3], or Arnold and Falk [4]. These finite element spaces correspond to the MINI-element of Arnold, Brezzi, and Fortin; see [3]. Since the problem considered is similar to the Stokes problem, it is also possible to use other elements that are common in fluid dynamics, e.g. the Taylor-Hood element; see Huang [54].

The assumptions of Theorem 3.1 hold with $V := (H_0^1(\Omega))^2, M := L_2(\Omega)/\mathbf{R}$, $M_c := H^1(\Omega)/\mathbf{R}$. A proof can be found in the references just provided.

It is also possible to use the MITCn-elements, introduced by Brezzi, Bathe, and Fortin in [28]; see also Peisker and Braess [67].

We note that the approach considered in this section is not the only one possible. There are direct approaches that do not use the Helmholtz decomposition; see Arnold and Brezzi [2].

Chapter 4

A Blockdiagonal Preconditioner

4.1 Introduction

In this chapter, we introduce a blockdiagonal preconditioner for saddlepoint problems with a penalty term; see Section 3.2. We show that the condition number of the preconditioned system is bounded independently of the discretization and the penalty parameters. Therefore the preconditioned conjugate residual method converges at a rate independent of the critical parameters; see Section 2.1.3.

The remainder of this chapter is organized as follows. In Section 4.2, we first describe the preconditioner and the assumptions made on it. In 4.2.1, we discuss a condition number estimate that is valid for the complete Babuška-Brezzi theory as described in Section 3.2, i.e. we do not assume A and C to be $V-$ and $M-$positive definite, respectively, whereas in 4.2.2, we restrict ourselves to this case and give a condition number estimate based on a different analysis. Finally, in Section 4.3, we present numerical results for a problem of linear elasticity.

The following analysis is using a matrix formulation of the discretized saddle point problem. Let us point out that it could have as well been presented in an abstract Hilbert space setting. The matrix representation obtained from discretization by finite elements provides bounds that are uniform. These bounds are natural counter parts of bounds for the solution of the continuous problem formulated in the proper Hilbert space.

The results presented in this chapter are based on Klawonn [57, 58].

4.2 The Preconditioner

To construct the preconditioner, we work with the matrix representation of the discretized saddle point problem,

$$\mathcal{A} := \begin{pmatrix} A & B^t \\ B & -t^2 C \end{pmatrix} \in \mathbf{R}^{n+m} \times \mathbf{R}^{n+m}, \tag{4.1}$$

and give the block diagonal preconditioner the form

$$\hat{\mathcal{B}} := \begin{pmatrix} \hat{A} & O \\ O & \hat{C} \end{pmatrix} \in \mathbf{R}^{n+m} \times \mathbf{R}^{n+m}. \tag{4.2}$$

Here $\hat{A}$ and $\hat{C}$ satisfy certain ellipticity conditions, i.e. there exist positive constants a_0, a_1 and c_0, c_1, such that

$$a_0^2 \, \|u\|_V^2 \leq u^t \hat{A} u \leq a_1^2 \, \|u\|_V^2,$$
$$c_0^2 \, |||p|||_M^2 \leq p^t \hat{C} p \leq c_1^2 \, |||p|||_M^2.$$

The next lemma shows that $\hat{\mathcal{B}}$ is positive definite and defines a norm on $\mathbf{R}^{n+m}$, which is equivalent to $||| \cdot |||$, cf. Section 3.2. Let us point out that the norms in this Chapter are defined on $\mathbf{R}^k, k = n, m, n+m$.

Lemma 4.1 *There exist positive constants b_0, b_1, such that*

$$b_0 \, |||x||| \leq \|\hat{\mathcal{B}}^{1/2} x\|_2 \leq b_1 \, |||x|||.$$

Proof: Using the ellipticity of $\hat{A}$ and $\hat{C}$, we obtain

$$\begin{aligned}
||\hat{\mathcal{B}}^{1/2} x||_2^2 &= x^t \hat{\mathcal{B}} x \\
&= u^t \hat{A} u + p^t \hat{C} p \\
&\leq a_1^2 \, \|u\|_V^2 + c_1^2 \, |||p|||_M^2 \\
&\leq \max\{a_1^2, c_1^2\}(\|u\|_V^2 + |||p|||_M^2) \\
&\leq \max\{a_1^2, c_1^2\}(\|u\|_V + |||p|||_M)^2 \\
&= \max\{a_1^2, c_1^2\}|||x|||^2.
\end{aligned}$$

Analogously, we get

$$\begin{aligned}
|||x|||^2 &= (\|u\|_V + |||p|||_M)^2 \\
&\leq 2 \, \|u\|_V^2 + 2|||p|||_M^2 \\
&\leq 2 \max\{\frac{1}{a_0^2}, \frac{1}{c_0^2}\} \, ||\hat{\mathcal{B}}^{1/2} x||_2^2 \\
&= 2 \left(\min\{a_0^2, c_0^2\}\right)^{-1} \, ||\hat{\mathcal{B}}^{1/2} x||_2^2.
\end{aligned}$$

□

Remark 4.1 *Examples of very fast and efficient methods, which fulfill the ellipticity requirements, are given by domain decomposition and multigrid methods or, more generally, by Schwarz methods; see Section 2.2. The ellipticity constants should preferably be independent of the discretization parameters but there are also interesting cases with a polylogarithmic dependence on H/h; see [37, 72, 73]. (The parameter H represents the diameter of a subdomain in a domain decomposition method.)*

4.2.1 A Condition Number Estimate

In view of Theorem 2.1, our goal is to give an estimate of the condition number $\kappa(\hat{\mathcal{B}}^{-1}\mathcal{A}) := \rho(\hat{\mathcal{B}}^{-1}\mathcal{A})\rho((\hat{\mathcal{B}}^{-1}\mathcal{A})^{-1})$. Since $\hat{\mathcal{B}}$ is positive definite, $\hat{\mathcal{B}}^{-1}\mathcal{A}$ and $\hat{\mathcal{B}}^{-1/2}\mathcal{A}\hat{\mathcal{B}}^{-1/2}$ have the same eigenvalues and since $\hat{\mathcal{B}}^{-1/2}\mathcal{A}\hat{\mathcal{B}}^{-1/2}$ is normal, we obtain $\rho(\hat{\mathcal{B}}^{-1}\mathcal{A}) = ||\hat{\mathcal{B}}^{-1/2}\mathcal{A}\hat{\mathcal{B}}^{-1/2}||_2$. The same argument applies to $\rho\left((\hat{\mathcal{B}}^{-1}\mathcal{A})^{-1}\right)$. Thus, we only have to provide upper bounds for $||\hat{\mathcal{B}}^{-1/2}\mathcal{A}\hat{\mathcal{B}}^{-1/2}||_2$ and $||\left(\hat{\mathcal{B}}^{-1/2}\mathcal{A}\hat{\mathcal{B}}^{-1/2}\right)^{-1}||_2$.

The next two Lemmata are well known and are given here for the sake of completeness only.

Lemma 4.2 *Let L be a $k \times k$ matrix. Then, the following three inequalities are equivalent:*

$$\inf_{y \neq 0} \sup_{x \neq 0} \frac{x^t L y}{||x||_2 ||y||_2} \geq \alpha,$$

$$||Ly||_2 \geq \alpha \, ||y||_2 \quad \forall y \in \mathbf{R}^k,$$

$$||L^{-1}||_2 \leq \frac{1}{\alpha}.$$

Lemma 4.3 *Let L be a $k \times k$ matrix. Then, the following three inequalities are equivalent:*

$$\sup_{y \neq 0} \sup_{x \neq 0} \frac{x^t L y}{||x||_2 ||y||_2} \leq C,$$

$$||Ly||_2 \leq C \, ||y||_2 \quad \forall y \in \mathbf{R}^k,$$

$$\|L\|_2 \le C.$$

In the following lemma, we prove lower and upper bounds for the inf-sup and sup-sup of $\hat{\mathcal{B}}^{-1/2}\mathcal{A}\hat{\mathcal{B}}^{-1/2}$, respectively.

Lemma 4.4 *There exist positive constants C_0, C_1, such that*

$$C_0 \le \inf_{y\neq 0}\sup_{x\neq 0} \frac{x^t\hat{\mathcal{B}}^{-1/2}\mathcal{A}\hat{\mathcal{B}}^{-1/2}y}{\|x\|_2\|y\|_2},$$

$$\sup_{y\neq 0}\sup_{x\neq 0} \frac{x^t\hat{\mathcal{B}}^{-1/2}\mathcal{A}\hat{\mathcal{B}}^{-1/2}y}{\|x\|_2\|y\|_2} \le C_1.$$

Proof: The Lemma follows immediately from the discretized versions of (3.6), (3.7) by changing basis and by applying Lemma 4.1. From the theory of saddle point problems with penalty term, see (3.6), (3.7), we have

$$\begin{aligned}
\gamma_0 &\le \inf_{y\neq 0}\sup_{x\neq 0} \frac{x^t\mathcal{A}y}{\|x\|_X\|y\|_X} && (4.3)\\
&\le \sup_{y\neq 0}\sup_{x\neq 0} \frac{x^t\mathcal{A}y}{\|x\|_X\|y\|_X} && (4.4)\\
&\le \gamma_1. && (4.5)
\end{aligned}$$

Thus, we obtain by using Lemma 4.1

$$\begin{aligned}
\inf_{y\neq 0}\sup_{x\neq 0} \frac{x^t\hat{\mathcal{B}}^{-1/2}\mathcal{A}\hat{\mathcal{B}}^{-1/2}y}{\|x\|_2\|y\|_2} &= \inf_{\tilde{y}\neq 0}\sup_{\tilde{x}\neq 0} \frac{\tilde{x}^t\mathcal{A}\tilde{y}}{\|\hat{\mathcal{B}}^{1/2}\tilde{x}\|_2\|\hat{\mathcal{B}}^{1/2}\tilde{y}\|_2} && (4.6)\\
&\ge \frac{1}{b_1^2}\inf_{\tilde{y}\neq 0}\sup_{\tilde{x}\neq 0} \frac{\tilde{x}^t\mathcal{A}\tilde{y}}{\|\tilde{x}\|_X\|\tilde{y}\|_X} && (4.7)\\
&\ge \frac{\gamma_0}{b_1^2}. && (4.8)
\end{aligned}$$

Analogously, we get the upper bound

$$\sup_{y\neq 0}\sup_{x\neq 0} \frac{x^t\hat{\mathcal{B}}^{-1/2}\mathcal{A}\hat{\mathcal{B}}^{-1/2}y}{\|x\|_2\|y\|_2} \le \frac{\gamma_1}{b_0^2}. \tag{4.9}$$

From these estimates, we have $C_0 := \frac{\gamma_0}{b_1^2}$ and $C_1 := \frac{\gamma_1}{b_0^2}$.

□

The next theorem follows from the definition of the condition number and Lemmata 4.2, 4.3, and 4.4.

Theorem 4.1 *The condition number of $\hat{\mathcal{B}}^{-1}\mathcal{A}$ is bounded independently of the discretization and the penalty parameters, i.e.*

$$\kappa(\hat{\mathcal{B}}^{-1}\mathcal{A}) \leq \frac{C_1}{C_0}.$$

4.2.2 An Alternative Proof

In this section, we consider a special case of saddle point problems with penalty. We assume that A is $V-$ rather than V_0-elliptic and that C is $M-$elliptic. This is the case for the mixed formulation of the equations of linear elasticity and the Stokes problem of fluid dynamics. The restricted problem considered here, allows us to give a condition number estimate that is based on a generalized eigenvalue problem.

We assume that the preconditioner has the form

$$\hat{\mathcal{B}} := \begin{pmatrix} \hat{A} & O \\ O & \hat{C} \end{pmatrix}. \tag{4.10}$$

Here $\hat{A}$ is V-elliptic and $\hat{C}$ M-elliptic. Thus, $\hat{\mathcal{B}}$ is positive definite.

We denote the case of $\hat{A} = A$ and $\hat{C} = C$ by $\mathcal{B}$. As in the previous section, our goal is to give an estimate of the condition number $\kappa(\hat{\mathcal{B}}^{-1}\mathcal{A})$. A simple computation shows that

$$\begin{aligned} \hat{\mathcal{B}}^{-1/2}\mathcal{A}\hat{\mathcal{B}}^{-1/2} &= \begin{pmatrix} \hat{A}^{-1/2}A\hat{A}^{-1/2} & \hat{A}^{-1/2}B^t\hat{C}^{-1/2} \\ \hat{C}^{-1/2}B\hat{A}^{-1/2} & -t^2\,\hat{C}^{-1/2}C\hat{C}^{-1/2} \end{pmatrix} \\ &=: \begin{pmatrix} \tilde{A} & \tilde{B}^t \\ \tilde{B} & -t^2\,\tilde{C} \end{pmatrix}. \end{aligned} \tag{4.11}$$

We denote by μ an eigenvalue of $\tilde{B}\tilde{B}^t$.

We make the following assumptions on $\mathcal{A}$ and $\hat{\mathcal{B}}$:
The matrix $\hat{A}$ is a good preconditioner for A, i.e.

$$\exists a_0, a_1 > 0 \quad a_0^2 u^t\hat{A}u \leq u^tAu \leq a_1^2 u^t\hat{A}u \quad \forall u \in \mathbf{R}^n. \tag{4.12}$$

The constants a_0, a_1 should preferably be independent of the discretization parameters but there are also other interesting cases; see Remark 4.1.

We also require that $\hat{C}$ is a good preconditioner for the pressure mass matrix M_p, i.e.

$$\exists m_0, m_1 > 0 \quad m_0^2 p^t\hat{C}p \leq p^tM_pp \leq m_1^2 p^t\hat{C}p \quad \forall p \in \mathbf{R}^m \tag{4.13}$$

and we finally assume that C is spectrally equivalent to $\hat{C}$, i.e.

$$\exists c_0, c_1 > 0 \quad c_0^2 p^t \hat{C} p \leq p^t C p \leq c_1^2 p^t \hat{C} p \quad \forall p \in \mathbf{R}^m. \tag{4.14}$$

For lower order finite elements, a good choice for $\hat{C}$ is a one-level overlapping Schwarz method. This family of methods also includes the popular algorithms that use diagonal and block preconditioning. Although the (pressure) mass matrix is uniformly well conditioned, i.e. has a condition number independent of h, we can decrease the number of iterations at a small expense by replacing a diagonal preconditioner by a one-level overlapping Schwarz method using an overlap of one node and small subdomains; see Section 4.3.

We will next establish some bounds for $BA^{-1}B^t$. These bounds depend directly on those of the inf-sup condition and on the boundedness of B. We then consider the case of the special preconditioner $\mathcal{B}$ and use these results to obtain a condition number estimate for the more general choices of $\hat{\mathcal{B}}$.

From the inf-sup condition, we obtain

$$\beta_0^2 \, p^t M_p p \leq p^t B A^{-1} B^t p \quad \forall p \in \mathbf{R}^m; \tag{4.15}$$

see Babuška [10], Brezzi [27], Brezzi and Fortin [30], and Silvester and Wathen [85]. Since B is uniformly bounded, we have

$$\exists \beta_1 > 0, \text{ s.t. } u^t B^t p \leq \beta_1 \, (p^t M_p p)^{1/2} (u^t A u)^{1/2} \quad \forall u \in \mathbf{R}^n, \forall p \in \mathbf{R}^m,$$

where M_p is again the pressure mass matrix. Substituting $u = A^{-1}B^t p$ and cancelling a common factor, we get

$$p^t B A^{-1} B^t p \leq {\beta_1}^2 \, p^t M_p p \quad \forall p \in \mathbf{R}^m, \tag{4.16}$$

and the following inequality holds with positive constants β_0, β_1,

$$\beta_0^2 \, p^t M_p p \leq p^t B A^{-1} B^t p \leq \beta_1^2 \, p^t M_p p \quad \forall p \in \mathbf{R}^m. \tag{4.17}$$

These constants are independent of the penalty parameter t.

We next consider the case of the special preconditioner $\mathcal{B}$. By λ_{max}^{abs} and λ_{min}^{abs}, we denote the extrema of the absolute values of the eigenvalues.

Lemma 4.5

$$\begin{aligned}
\lambda_{max}^{abs}(\mathcal{B}^{-1}\mathcal{A}) &= \frac{1-t^2}{2} + \sqrt{\mu_{max} + \left(\frac{1+t^2}{2}\right)^2} \\
\lambda_{min}^{abs}(\mathcal{B}^{-1}\mathcal{A}) &= \frac{t^2-1}{2} + \sqrt{\mu_{min} + \left(\frac{1+t^2}{2}\right)^2}
\end{aligned}$$

Proof: We consider the preconditioned system $\mathcal{B}^{-1/2}\mathcal{A}\mathcal{B}^{-1/2}$ and the related eigenvalue problem

$$\begin{pmatrix} I & \tilde{B}^t \\ \tilde{B} & -t^2\, I \end{pmatrix} \begin{pmatrix} u \\ p \end{pmatrix} = \lambda \begin{pmatrix} u \\ p \end{pmatrix}$$

From this equation, it follows that

$$\tilde{B}\tilde{B}^t p = (\lambda - 1)(t^2 + \lambda)p.$$

Denoting the eigenvalues of $\tilde{B}\tilde{B}^t$ by μ, we get

$$(\lambda - 1)(t^2 + \lambda) = \mu.$$

Solving for λ, we obtain

$$\lambda = \frac{1-t^2}{2} \pm \sqrt{\mu + \left(\frac{1+t^2}{2}\right)^2}.$$

We get, by comparing the distances of the two values of λ to the origin,

$$\begin{aligned}
\lambda_{max}^{abs} &= \frac{1-t^2}{2} + \sqrt{\mu_{max} + \left(\frac{1+t^2}{2}\right)^2}, \\
\lambda_{min}^{abs} &= \left|\frac{1-t^2}{2} - \sqrt{\mu_{min} + \left(\frac{1+t^2}{2}\right)^2}\right| \\
&= \frac{t^2-1}{2} + \sqrt{\mu_{min} + \left(\frac{1+t^2}{2}\right)^2}.
\end{aligned}$$

□

The next lemma provides a condition number estimate for the case of the exact preconditioner.

Lemma 4.6

$$\kappa(\mathcal{B}^{-1}\mathcal{A}) \leq \frac{1/2 + \sqrt{\mu_{max} + 1/4}}{-1/2 + \sqrt{\mu_{min} + 1/4}}.$$

Proof: It is sufficient to provide an upper bound for

$$\frac{1-t^2}{2} + \sqrt{\mu_{max} + \left(\frac{1+t^2}{2}\right)^2}$$

and a lower bound for

$$\frac{t^2-1}{2} + \sqrt{\mu_{min} + \left(\frac{1+t^2}{2}\right)^2}.$$

From the assumptions that $0 \leq t \leq 1$, we easily, by computing the first derivative, find the upper and the lower bound.

$\square$

We are now able to prove

Theorem 4.2

$$\kappa(\mathcal{B}^{-1}\mathcal{A}) \leq \frac{1/2 + \sqrt{m_1^2\beta_1^2 + 1/4}}{-1/2 + \sqrt{m_0^2\beta_0^2 + 1/4}}.$$

Proof: Obviously, we only have to provide an upper bound for μ_{max} and a lower bound for μ_{min}, i.e. upper and lower bounds for the Rayleigh quotient $p^t\tilde{B}\tilde{B}^t p/p^t p$.

By construction, see (4.16), we have

$$\frac{p^t(BA^{-1}B^t)p}{p^t M_p p} \leq \beta_1^2$$

Using (4.13) and $C = \hat{C}$, we get

$$\frac{p^t(BA^{-1}B^t)p}{p^t\hat{C}p} = \frac{q^t\hat{C}^{-1/2}BA^{-1}B^t\hat{C}^{-1/2}q}{q^t q} \leq \beta_1^2 m_1^2.$$

Applying (4.12), results in

$$\frac{q^t(\hat{C}^{-1/2}B\hat{A}^{-1/2})(\hat{A}^{-1/2}B^t\hat{C}^{-1/2})q}{q^t q} \leq \beta_1^2 m_1^2 a_0^{-2}.$$

The definition of $\tilde{B}$ gives

$$\frac{q^t\tilde{B}\tilde{B}^t q}{q^t q} \leq \beta_1^2 m_1^2 a_0^{-2} = \beta_1^2 m_1^2.$$

Here, $a_0 = 1$, since we are using $\hat{A} = A$. Analogously, using (4.12),(4.13), and (4.15), we obtain

$$\frac{q^t\tilde{B}\tilde{B}^t q}{q^t q} \geq \beta_0^2 m_0^2 a_1^{-2} = \beta_0^2 m_0^2.$$

The theorem follows from Lemma 4.6.

$\square$

Remark 4.2 *In many applications, we have $C = M_p$. In this case the bound in Theorem 4.2 simplifies to*

$$\frac{1/2 + \sqrt{\beta_1^2 + 1/4}}{-1/2 + \sqrt{\beta_0^2 + 1/4}}.$$

We next give an upper bound of the condition number when a general preconditioner $\hat{\mathcal{B}}^{-1}$ is used.

Theorem 4.3

$$\kappa(\hat{\mathcal{B}}^{-1}\mathcal{A}) \leq \frac{\max\{a_1^2, c_1^2\}}{\min\{a_0^2, c_0^2\}}\kappa(\mathcal{B}^{-1}\mathcal{A}).$$

Proof: We consider

$$\begin{aligned}
\lambda_{max}^{abs}(\hat{\mathcal{B}}^{-1/2}\mathcal{A}\hat{\mathcal{B}}^{-1/2}) &= \rho(\hat{\mathcal{B}}^{-1/2}\mathcal{A}\hat{\mathcal{B}}^{-1/2}) \\
&= \|\hat{\mathcal{B}}^{-1/2}\mathcal{A}\hat{\mathcal{B}}^{-1/2}\|_2 \\
&= \sup_{x\neq 0}\left(\frac{x^t\hat{\mathcal{B}}^{-1/2}\mathcal{A}\hat{\mathcal{B}}^{-1/2}x}{x^tx}\right) \\
&\leq \sup_{x\neq 0}\left(\frac{x^t\mathcal{A}x}{x^t\mathcal{B}x}\right)\sup_{x\neq 0}\left(\frac{x^t\mathcal{B}x}{x^t\hat{\mathcal{B}}x}\right) \\
&\leq \lambda_{max}^{abs}(\mathcal{B}^{-1/2}\mathcal{A}\mathcal{B}^{-1/2})\ \max\{a_1^2, c_1^2\}.
\end{aligned}$$

Here, we have used that $M := \hat{\mathcal{B}}^{-1/2}\mathcal{A}\hat{\mathcal{B}}^{-1/2}$ is normal. Analogously, we obtain, by using that $\lambda_{min}^{abs}(M) = (\rho(M^{-1}))^{-1}$,

$$\lambda_{min}^{abs}(\hat{\mathcal{B}}^{-1/2}\mathcal{A}\hat{\mathcal{B}}^{-1/2}) \geq \lambda_{min}^{abs}(\mathcal{B}^{-1/2}\mathcal{A}\mathcal{B}^{-1/2})\min\{a_0^2, c_0^2\}.$$

Since $\hat{\mathcal{B}}^{-1}\mathcal{A}$ and $\hat{\mathcal{B}}^{-1/2}\mathcal{A}\hat{\mathcal{B}}^{-1/2}$ have the same spectrum, Theorem 4.3 follows.

$\square$

From Theorem 4.2 and 4.3 follows immediately

Corollary 4.1

$$\kappa(\hat{\mathcal{B}}^{-1}\mathcal{A}) \leq \frac{\max\{a_1^2, c_1^2\}}{\min\{a_0^2, c_0^2\}}\ \frac{1/2 + \sqrt{m_1^2\beta_1^2 + 1/4}}{-1/2 + \sqrt{m_0^2\beta_0^2 + 1/4}}.$$

Hence, we have now derived an estimate for the condition number which is independent of the discretization and the penalty parameter and for which we can guarantee that the convergence rate of the Krylov space method considered will not deteriorate when t and h decrease. Corollary 4.1, shows that the condition number estimate of $\hat{\mathcal{B}}^{-1}\mathcal{A}$ is completely determined by the quality of the preconditioners $\hat{A}, \hat{C}$, and the condition number of $BA^{-1}B^t$.

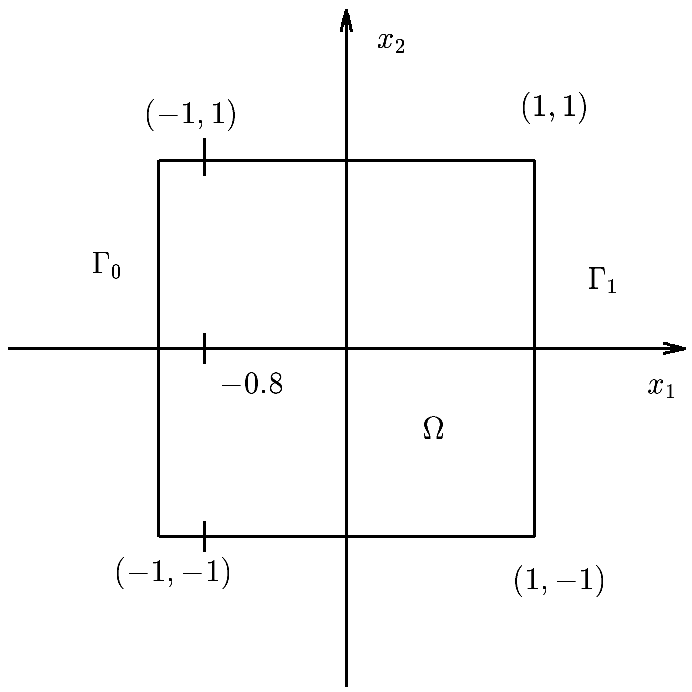

Figure 4.1: The elastic domain.

4.3 Numerical examples

We apply our preconditioner to the problem of planar, linear elasticity; see Section 3.3.1. For simplicity, we work with the formulation given in Remark 3.1. All the results shown are for mixed boundary conditions, $\Gamma_0 := \{x = (x_1, x_2) \in \partial\Omega : x_1 < -0.8\}$ and the region $[-1,1] \times [-1,1]$; see Figure 4.1.

Without loss of generality, we use $E = 1$ as the value for Young's modulus. We note that our model is mathematically equivalent to the full elasticity problem only in the case of homogeneous Dirichlet conditions. The numerical results confirm that the number of iterations is bounded independently of the critical parameters h and t.

All computations were carried out on a Sun workstation with 256 Mbyte memory using the numerical software package PETSc developed by William Gropp and Barry Smith at the Argonne National Laboratory; see Gropp and Smith [51] or Smith [80]. In the original version of PETSc, saddlepoint problems were not included. Thus, we have added this feature to the functionality of PETSc. The initial guess in our experiments is 0, and the stopping criterion is $\|r_k\|_2/\|r_0\|_2 < 10^{-5}$, where r_k is the k-th residual. The Krylov space method used is always the PCR-method.

Table 4.1: PCR-method with exact solvers as preconditioners for A and C, $\nu = 0.3$.

Grid	20×10	40×20	60×30	80×40	100×50	120×60	140×70
Iter	17	19	19	21	21	21	21

To see how the PCR-method behaves under the best of circumstances, we first conducted some experiments using exact solvers, i.e. $\hat{A} = A$ and $\hat{C} = C$. The method works as predicted for both limit cases, $h \to 0$ and $t \to 0$, see Tables 4.1 and 4.5.

In another series of experiments, we applied different preconditioners for A and C. We present results with $\hat{A}$ definied by a two-level multigrid preconditioner with a V-cycle including one pre- and one post-smoothing symmetric Gauss-Seidel step, and $\hat{C}$ defined by a one-level symmetric multiplicative overlapping Schwarz method with overlap of one node; see Tables 4.2, 4.3, 4.6, 4.7, 4.8. The use of a Schwarz method can be motivated as follows: We note, that although C is uniformly well conditioned, i.e. has a condition number independent of h, a simple diagonal preconditioning is not the best choice as far as the number of iterations is concerned; cf. Table 4.4, the iteration count can be reduced at a little extra expense. Interpreting diagonal preconditioning as a one-level additive non-overlapping Schwarz method with 1 node per subdomain, see e.g. Hackbusch [53], p. 343, it is natural to try to improve the convergence by introducing some overlap or by using a multiplicative scheme. Our experiments show that already a minimal overlap of one node, i.e. only the boundary nodes of adjacent substructures are common, combined with a one-level symmetric multiplicative Schwarz method yields results matching those obtained by the exact solver for C; see Tables 4.2 and 4.3 (resp. Tables 4.6 and 4.8). We also note that it would of course be more efficient to use more than two levels in the multigrid preconditioner but here, we primarily wish to analyze the parameter dependence of the preconditioned methods.

The experiments show that the preconditioned conjugate residual method represents an efficient and robust iterative solver for saddle point problems with a penalty term. For a comparison of the convergence rates and the efficiency of multigrid and Krylov subspace methods for saddle point problems; see Elman [42].

Table 4.2: PCR-method with a two-level multigrid preconditioner with a standard V-cycle defining $\hat{A}$, and $\hat{C} = C$, and $\nu = 0.3$.

Grid	20×10	40×20	60×30	80×40	100×50	120×60	140×70
Iter	20	23	24	26	26	26	26

Table 4.3: PCR-method with a two-level multigrid preconditioner with a standard V-cycle defining $\hat{A}$ and a one-level symmetric multiplicative overlapping Schwarz method with overlap of one node defining $\hat{C}$, and $\nu = 0.3$.

Grid	20×10	40×20	60×30	80×40	100×50	120×60	140×70
Iter	20	23	24	26	26	26	26

Table 4.4: PCR-method with a two-level multigrid preconditioner with a standard V-cycle defining $\hat{A}$, a diagonal preconditioner $\hat{C} = diag(C)$, and $\nu = 0.3$.

Grid	20×10	40×20	60×30	80×40	100×50	120×60	140×70
Iter	46	53	56	58	58	58	60

Table 4.5: PCR-method with $\hat{A} = A$ and $\hat{C} = C$ on a 80×40 grid.

ν	0.3	0.4	0.49	0.499	0.4999	0.49999	0.499999	0.5
Iter	21	21	23	25	25	25	25	25

Table 4.6: PCR-method with two-level multigrid with a standard V-cycle defining $\hat{A}$ and $\hat{C} = C$ on a 80×40 grid.

ν	0.3	0.4	0.49	0.499	0.4999	0.49999	0.499999	0.5
Iter	26	29	33	33	33	33	33	33

Table 4.7: PCR-method with an exact solver as $\hat{A}$ and a one-level symmetric multiplicative overlapping Schwarz method with overlap of one node as $\hat{C}$ on a 80×40 grid.

ν	0.3	0.4	0.49	0.499	0.4999	0.49999	0.499999	0.5
Iter	21	21	23	25	25	25	25	25

Table 4.8: PCR-method with two-level multigrid with a standard V-cycle as $\hat{A}$ and a one-level symmetric multiplicative overlapping Schwarz method with overlap of one node as $\hat{C}$ on a 80×40 grid.

ν	0.3	0.4	0.49	0.499	0.4999	0.49999	0.499999	0.5
Iter	26	29	33	33	33	33	33	33

Chapter 5

A Triangular Preconditioner

5.1 Introduction

In this chapter, we analyze a triangular preconditioner for saddlepoint problems with a penalty term. We only consider symmetric saddlepoint problems. For related work on the non-symmetric case, see Elman and Silvester [44], who analyze the Oseen operator which is obtained by applying a Picard iteration to the Navier-Stokes equations. It turns out that our analysis of triangular preconditioners also provides an alternative proof of a result due to Bramble and Pasciak [19].

As in the previous chapter, the following analysis is carried out using a matrix formulation of the discretized saddle point problem. Let us point out that it could have as well been presented in an abstract Hilbert space setting. The matrix representation obtained from discretization by finite elements provides bounds that are uniform. These bounds are natural counter parts of bounds for the solution of the continuous problem formulated in the proper Hilbert space.

The outline of the remainder of this chapter is as follows. In Section 5.2, we discuss the preconditioning strategy. In the case of exact solvers for A and C, we show that the spectrum of the preconditioned system is bounded independently of the critical parameters by solving a generalized eigenvalue problem. This technique cannot be applied to inexact preconditioners. Neither can a perturbation argument be used since the preconditioner is indefinite. We find that the preconditioned system $\mathcal{A}\hat{\mathcal{B}}^{-1}$ is symmetric positive definite in a certain metric which is defined by a symmetric positive definite matrix $\mathcal{H}^{-1}$ and that the generalized eigenvalues of $\mathcal{H}^{-1}\mathcal{A}\hat{\mathcal{B}}^{-1} = \lambda\mathcal{H}^{-1}$ are bounded independently of the critical parameters. From this, we obtain the bounds of the eigenvalues of $\mathcal{A}\hat{\mathcal{B}}^{-1}$. We also use this fact to show that GMRES, in a certain metric defined by $\hat{\mathcal{H}}^{-1}$, converges independently of these parameters. Here $\hat{\mathcal{H}}$ is an arbitrary symmetric positive definite matrix with $\bar{C}_0^2\mathcal{H}^{-1} \leq \hat{\mathcal{H}}^{-1} \leq \bar{C}_1^2\mathcal{H}^{-1}$, where $\bar{C}_0, \bar{C}_1 > 0$ are constants independent of the discretization and the penalty parameters. To prove this convergence estimate, we introduce an apparently new technique. We exploit that both, GMRES and the method of conjugate residuals (CR), minimize the residual in the norm used and that $\mathcal{H}^{-1}$ and $\hat{\mathcal{H}}^{-1}$ define equivalent norms. Then, we use that the convergence of CR in the $\mathcal{H}^{-1}-$metric can be estimated in terms of the generalized eigenvalues of $\mathcal{H}^{-1}\mathcal{A}\hat{\mathcal{B}}^{-1} = \lambda\mathcal{H}^{-1}$. Finally, we show that $\mathcal{A}\hat{\mathcal{B}}^{-1}$ is diagonalizable and use this property to give a convergence estimate for GMRES in the Euclidean metric.

In Section 5.3, we discuss numerical results for a problem of linear elasticity obtained by using GMRES and BI-CGSTAB.

5.2 Preconditioning Techniques

In this section, we consider a triangular preconditioner for problem (3.2). As in Section 4.2.2, we restrict ourselves to the case of A being $V-$elliptic and C being $M-$elliptic. The preconditioned system is either of the form $\mathcal{A}\hat{\mathcal{B}}^{-1}$ or $\hat{\mathcal{B}}^{-1}\mathcal{A}$ where $\hat{\mathcal{B}}$ is the triangular preconditioner.

We use the following notation

$$\hat{\mathcal{B}}_U \;:=\; \begin{pmatrix} \hat{A} & B^t \\ O & -\hat{C} \end{pmatrix}, \qquad \hat{\mathcal{B}}_L \;:=\; \begin{pmatrix} \hat{A} & O \\ B & -\hat{C} \end{pmatrix},$$

Here $\hat{A}$ and $\hat{C}$ are positive definite. We make the following assumptions on $\mathcal{A}$ and $\hat{\mathcal{B}}$:
The matrix $\hat{A}$ is a good preconditioner for A, i.e.

$$\exists a_0, a_1 > 0 \quad a_0^2 u^t\hat{A}u \leq u^tAu \leq a_1^2 u^t\hat{A}u \quad \forall u \in \mathbf{R}^n. \tag{5.1}$$

The constants a_0, a_1 should preferably be close to each other and be independent of the discretization parameters but there are also other interesting cases; see Remark 4.1. Multigrid and domain decomposition methods are examples of preconditioners that meet these requirements; see Section 2.2.

We also require that $\hat{C}$ is a good preconditioner for the pressure mass matrix M_p, i.e.

$$\exists m_0, m_1 > 0 \quad m_0^2\, p^t\hat{C}p \le p^t M_p p \le m_1^2\, p^t \hat{C}p \quad \forall p \in \mathbf{R}^m \tag{5.2}$$

and we finally assume that C is spectrally equivalent to $\hat{C}$, i.e.

$$\exists c_0, c_1 > 0 \quad c_0^2\, p^t\hat{C}p \le p^t C p \le c_1^2\, p^t \hat{C}p \quad \forall p \in \mathbf{R}^m. \tag{5.3}$$

A good choice for $\hat{C}$ is a one-level overlapping Schwarz method; see Klawonn [57] and Section 4.3, Tables 4.3,4.7,4.8.

From the inf-sup condition for B and the uniform boundedness of B, we obtain the following inequality

$$\beta_0^2\, p^t M_p p \le p^t B A^{-1} B^t p \le \beta_1^2\, p^t M_p p \quad \forall p \in \mathbf{R}^m,$$

with positive constants β_0, β_1; see Brezzi and Fortin [30]. These constants are independent of the discretization and the penalty parameters. We denote by μ an eigenvalue of the generalized eigenvalue problem

$$BA^{-1}B^t p = \mu\, Cp$$

and by μ_{min}, μ_{max} its extreme eigenvalues.

We first restrict our analysis to $\hat{\mathcal{B}}_U$ and drop the subscript U. In the case of $\hat{A} = A$ and $\hat{C} = C$, we use $\mathcal{B}$ rather than $\hat{\mathcal{B}}$.

To get bounds for the spectrum of $\mathcal{A}\mathcal{B}^{-1}$, we consider the generalized eigenvalue problem

$$\mathcal{A}x = \lambda \mathcal{B}x.$$

Now, we can proceed as in the case of the block-diagonal preconditioner that is discussed in Section 4.2.2.

Lemma 5.1

$$\sigma(\mathcal{A}\mathcal{B}^{-1}) \subset [\mu_{min}, \mu_{max} + 1] \cup \{1\}$$

Proof: The proof follows by considering the generalized eigenvalue problem

$$\mathcal{A}x = \lambda\, \mathcal{B}x.$$

We obtain

$$\begin{pmatrix} A & B^t \\ B & -t^2 C \end{pmatrix} \begin{pmatrix} u \\ p \end{pmatrix} = \lambda \begin{pmatrix} A & B^t \\ O & -C \end{pmatrix} \begin{pmatrix} u \\ p \end{pmatrix},$$

which is equivalent to

$$\begin{aligned} Au + B^t p &= \lambda\,(Au + B^t p) \\ Bu - t^2\,Cp &= -\lambda\,Cp. \end{aligned}$$

We first show that $\lambda = 1$ is an eigenvalue. This can be easily seen by considering

$$Bu = (t^2 - 1)\,Cp$$

and looking at the cases $t^2 = 1$ and $t^2 < 1$, separately. In the first case, we can choose $p \neq 0$ arbitrarily and $u \in ker\,B$ and in the second one, we can choose $u \neq 0$ arbitrarily and set $p := (t^2 - 1)^{-1} C^{-1} Bu$. In both cases, we obtain eigenvectors for the eigenvalue $\lambda = 1$.

We now assume that $\lambda \neq 1$. Then, we get

$$BA^{-1}B^t p = (\lambda - t^2)\,Cp.$$

Thus, we obtain the relation

$$\mu = \lambda - t^2.$$

By using that $t \in [0, 1]$, we immediately have

$$\begin{aligned} \lambda_{max} &\leq 1 + \mu_{max}, \\ \lambda_{min} &\geq \mu_{min}. \end{aligned}$$

$\square$

From this lemma, we obtain

Theorem 5.1

$$\sigma(\mathcal{A}\mathcal{B}^{-1}) \subset [\beta_0^2 m_0^2, \beta_1^2 m_1^2 + 1] \cup \{1\}$$

Proof: Use the bounds for μ_{min} and μ_{max} given in Theorem 4.2.

$\square$

To provide bounds for $\mathcal{A}\hat{\mathcal{B}}^{-1}$, we cannot solve the generalized eigenvalue problem easily anymore. Nor is it possible to apply the techniques of Section 4.2.2, since $\hat{\mathcal{B}}^{-1}$ is not positive definite.

By making the assumption that

$$1 < a_0 \leq a_1, \tag{5.4}$$

which can always be achieved by an appropriate scaling, we can show that the spectrum of $\mathcal{A}\hat{\mathcal{B}}^{-1}$ stays bounded independently of the discretization and the penalty parameters. Consider

$$\begin{aligned} \mathcal{A}\hat{\mathcal{B}}^{-1} &= \begin{pmatrix} A & B^t \\ B & -t^2C \end{pmatrix} \begin{pmatrix} \hat{A}^{-1} & \hat{A}^{-1}B^t\hat{C}^{-1} \\ O & -\hat{C}^{-1} \end{pmatrix} \\ &= \begin{pmatrix} A\hat{A}^{-1} & (A-\hat{A})\hat{A}^{-1}B^t\hat{C}^{-1} \\ B\hat{A}^{-1} & (t^2C + B\hat{A}^{-1}B^t)\hat{C}^{-1} \end{pmatrix}. \end{aligned}$$

Introduce the notation,

$$\mathcal{H} := \begin{pmatrix} A-\hat{A} & O \\ O & \hat{C} \end{pmatrix}.$$

Multiplying $\mathcal{A}\hat{\mathcal{B}}^{-1}$ by $\mathcal{H}$ from the right, we obtain

$$\mathcal{A}\hat{\mathcal{B}}^{-1}\mathcal{H} = \begin{pmatrix} A\hat{A}^{-1}(A-\hat{A}) & (A-\hat{A})\hat{A}^{-1}B^t \\ B\hat{A}^{-1}(A-\hat{A}) & t^2C + B\hat{A}^{-1}B^t \end{pmatrix},$$

which is a symmetric positive definite matrix; cf. Lemma 5.2. Thus, the preconditioned system $\mathcal{A}\hat{\mathcal{B}}^{-1}$ is $\mathcal{H}-$normal(1); see Section 2.1.1. Introduce the notation

$$\mathcal{T} := \begin{pmatrix} I & 0 \\ BA^{-1} & I \end{pmatrix} \quad \text{and} \quad \tilde{\mathcal{H}} := \begin{pmatrix} A & 0 \\ 0 & S \end{pmatrix},$$

where $S := t^2C + BA^{-1}B^t$ and $-S$ denotes the Schur complement of $\mathcal{A}$ which is obtained by block Gaussian elimination.

Lemma 5.2 *There exist positive constants $\tilde{C}_0, \tilde{C}_1$, independent of t, h, such that*

$$\tilde{C}_0\,\tilde{\mathcal{H}} \leq \mathcal{A}\hat{\mathcal{B}}^{-1}\mathcal{H} \leq \tilde{C}_1\,\tilde{\mathcal{H}},$$

with $\tilde{C}_0 = \min\{(a_0-1), 1\}/3$ and $\tilde{C}_1 = 3\max\{(a_1-1), 1\}$.[1]

Proof: The proof is based on a block Cholesky factorization of $\mathcal{A}\hat{\mathcal{B}}^{-1}\mathcal{H}$. A direct computation shows:

$$\mathcal{A}\hat{\mathcal{B}}^{-1}\mathcal{H} = \begin{pmatrix} I & 0 \\ BA^{-1} & I \end{pmatrix} \begin{pmatrix} A\hat{A}^{-1}A - A & 0 \\ 0 & S \end{pmatrix} \begin{pmatrix} I & A^{-1}B^t \\ 0 & I \end{pmatrix}. \quad (5.5)$$

[1]As usual, $A \leq B$ means $B - A$ is symmetric positive semi-definite.

From the assumptions (5.1), (5.4), we obtain

$$\hat{C}_0\tilde{\mathcal{H}} \le \begin{pmatrix} A\hat{A}^{-1}A - A & 0 \\ 0 & S \end{pmatrix} \le \hat{C}_1\tilde{\mathcal{H}}, \tag{5.6}$$

where $\hat{C}_0 := \min\{(a_0 - 1), 1\}$ and $\hat{C}_1 := \max\{(a_1 - 1), 1\}$ are positive constants.

We now show that the eigenvalues of $\mathcal{T}\tilde{\mathcal{H}}\mathcal{T}^t$ are bounded by the extreme eigenvalues of $\tilde{\mathcal{H}}$,

$$\frac{1}{3}\tilde{\mathcal{H}} \le \mathcal{T}\tilde{\mathcal{H}}\mathcal{T}^t \le 3\tilde{\mathcal{H}}. \tag{5.7}$$

We have

$$\begin{aligned} \mathcal{T}\tilde{\mathcal{H}}\mathcal{T}^t &= \begin{pmatrix} I & 0 \\ BA^{-1} & I \end{pmatrix} \begin{pmatrix} A & 0 \\ 0 & S \end{pmatrix} \begin{pmatrix} I & A^{-1}B^t \\ 0 & I \end{pmatrix} \\ &= \begin{pmatrix} A & B^t \\ B & S + BA^{-1}B^t \end{pmatrix}. \end{aligned}$$

From this, we get

$$\begin{aligned} x^t\mathcal{T}\tilde{\mathcal{H}}\mathcal{T}^t x &= \begin{pmatrix} u \\ p \end{pmatrix}^t \begin{pmatrix} A & B^t \\ B & S + BA^{-1}B^t \end{pmatrix} \begin{pmatrix} u \\ p \end{pmatrix} \\ &= u^tAu + 2p^tBu + p^tSp + p^tBA^{-1}B^tp \\ &\le u^tAu + 2|p^tBu| + 2p^tSp. \end{aligned} \tag{5.8}$$

Applying the Cauchy-Schwarz inequality and $ab \le a^2/2 + b^2/2$, we obtain

$$\begin{aligned} |p^tBu| &= |p^tBA^{-1/2}A^{1/2}u| \\ &\le (p^tBA^{-1}B^tp)^{1/2}(u^tAu)^{1/2} \\ &\le \frac{1}{2}p^tBA^{-1}B^tp + \frac{1}{2}u^tAu \\ &\le \frac{1}{2}(p^tSp + u^tAu). \end{aligned}$$

Hence, we get from (5.8)

$$\begin{aligned} x^t\mathcal{T}\tilde{\mathcal{H}}\mathcal{T}^t x &\le 2u^tAu + 3p^tSp \\ &\le 3x^t\tilde{\mathcal{H}}x. \end{aligned}$$

To obtain a lower bound, consider the generalized Rayleigh-quotient

$$\frac{x^t\mathcal{T}\tilde{\mathcal{H}}\mathcal{T}^t x}{x^t\tilde{\mathcal{H}}x} = \frac{y^t\tilde{\mathcal{H}}y}{y^t\mathcal{T}^{-1}\tilde{\mathcal{H}}\mathcal{T}^{-t}y},$$

where we have used the substitution $y := \mathcal{T}^t x$. As in the case of the upper bound, we obtain

$$
\begin{aligned}
y^t \mathcal{T}^{-1} \tilde{\mathcal{H}} \mathcal{T}^{-t} y &= \begin{pmatrix} v \\ q \end{pmatrix}^t \begin{pmatrix} I & 0 \\ -BA^{-1} & I \end{pmatrix} \begin{pmatrix} A & 0 \\ 0 & S \end{pmatrix} \begin{pmatrix} I & -A^{-1}B^t \\ 0 & I \end{pmatrix} \begin{pmatrix} v \\ q \end{pmatrix} \\
&= \begin{pmatrix} v \\ q \end{pmatrix}^t \begin{pmatrix} A & -B^t \\ -B & S + BA^{-1}B^t \end{pmatrix} \begin{pmatrix} v \\ q \end{pmatrix} \\
&= v^t A v - 2q^t B v + q^t S q + q^t B A^{-1} B^t q \\
&\leq v^t A v + 2|q^t B v| + 2 q^t S q \\
&\leq 3 y^t \tilde{\mathcal{H}} y.
\end{aligned}
$$

From (5.5), (5.6) and (5.7), we obtain

$$\tilde{C}_0 \tilde{\mathcal{H}} \leq \mathcal{A} \hat{\mathcal{B}}^{-1} \mathcal{H} \leq \tilde{C}_1 \tilde{\mathcal{H}}.$$

□

From this lemma, we derive

Lemma 5.3 *There exist positive constants C_0, C_1, independent of t, h, such that*

$$C_0 \mathcal{H} \leq \mathcal{A} \hat{\mathcal{B}}^{-1} \mathcal{H} \leq C_1 \mathcal{H}.$$

Proof: The lemma follows immediately from Lemma 5.2 since $\hat{C}$ is spectrally equivalent to $t^2 C + BA^{-1}B^t$ and $A - \hat{A}$ to A.

□

Remark 5.1 *This lemma could also be used to prove a result due to Bramble and Pasciak [20], cf. Theorem 1.*

Since we have made the assumption that $1 < a_0 \leq a_1$, $A - \hat{A}$ and $\hat{C}$ are positive definite, and $\mathcal{H}$ defines a new inner product on $\mathbf{R}^{n+m}$. We are now able to give bounds for the spectrum of $\mathcal{A}\hat{\mathcal{B}}^{-1}$.

Theorem 5.2 *There exist positive constants C_0, C_1, independent of t, h, such that*

$$\sigma(\mathcal{A}\hat{\mathcal{B}}^{-1}) \subset [C_0, C_1].$$

Proof: The constants C_0, C_1 in Lemma 5.3 provide lower and upper bounds for the eigenvalues of the generalized eigenvalue problem

$$\mathcal{A}\hat{\mathcal{B}}^{-1}\mathcal{H}x = \lambda\mathcal{H}x.$$

Since $\mathcal{H}$ is non-singular, this problem has the same eigenvalues as

$$\mathcal{A}\hat{\mathcal{B}}^{-1}y = \lambda y.$$

□

Corollary 5.1 *$\mathcal{H}^{-1}$ also defines an inner product on $\mathbf{R}^{n+m}$ and $\mathcal{A}\hat{\mathcal{B}}^{-1}$ is symmetric positive definite in this inner product, i.e. $\mathcal{H}^{-1}\mathcal{A}\hat{\mathcal{B}}^{-1}$ is symmetric and there exist positive constants C_0, C_1, such that $C_0\ \mathcal{H}^{-1} \leq \mathcal{H}^{-1}\mathcal{A}\hat{\mathcal{B}}^{-1} \leq C_1\ \mathcal{H}^{-1}$.*

Proof: The symmetry of $\mathcal{H}^{-1}\mathcal{A}\hat{\mathcal{B}}^{-1}$ follows immediately from a direct computation:

$$\begin{aligned}
\mathcal{H}^{-1}\mathcal{A}\hat{\mathcal{B}}^{-1} &= \begin{pmatrix} (A-\hat{A})^{-1}A\hat{A}^{-1} & \hat{A}^{-1}B^t\hat{C}^{-1} \\ \hat{C}^{-1}B\hat{A}^{-1} & \hat{C}^{-1}(t^2C + B\hat{A}^{-1}B^t)\hat{C}^{-1} \end{pmatrix} \\
&= \begin{pmatrix} (\hat{A}-\hat{A}A^{-1}\hat{A})^{-1} & \hat{A}^{-1}B^t\hat{C}^{-1} \\ \hat{C}^{-1}B\hat{A}^{-1} & \hat{C}^{-1}(t^2C + B\hat{A}^{-1}B^t)\hat{C}^{-1} \end{pmatrix}.
\end{aligned}$$

The generalized eigenvalue problem

$$\mathcal{H}^{-1}\mathcal{A}\hat{\mathcal{B}}^{-1}x = \lambda\, \mathcal{H}^{-1}x$$

has the same eigenvalues as

$$\mathcal{A}\hat{\mathcal{B}}^{-1}x = \lambda\, x.$$

From Theorem 5.2, we know that $\mathcal{A}\hat{\mathcal{B}}^{-1}$ has only positive real eigenvalues.

□

Remark 5.2 *Since it can be shown that*

$$\mathcal{A}\hat{\mathcal{B}}_U^{-1}\mathcal{H} = \mathcal{H}\hat{\mathcal{B}}_L^{-1}\mathcal{A},$$

the previous results obtained for $\mathcal{A}\hat{\mathcal{B}}_U^{-1}$ also apply to the spectrum of $\hat{\mathcal{B}}_L^{-1}\mathcal{A}$.

We now use the bounds of Corollary 5.1 to provide bounds of the convergence rate of the GMRES method used with a norm equivalent to the $\mathcal{H}^{-1}$–norm. Here, we exploit the fact that both, CR and GMRES, minimize the residual in the norm used; see e.g. Bruaset [31], Freund, Golub, and Nachtigal [47] or Saad and Schultz [71].

Theorem 5.3 *Let $\hat{\mathcal{H}}$ be a positive definite matrix, such that $\bar{C}_0^2 \ \ \mathcal{H}^{-1} \leq \hat{\mathcal{H}}^{-1} \leq \bar{C}_1^2 \ \ \mathcal{H}^{-1}$, where $\bar{C}_0, \bar{C}_1$ are positive constants independent of the discretization and penalty parameters. Then,*

$$\|r_n\|_{\hat{\mathcal{H}}^{-1}} \leq \frac{\bar{C}_1}{\bar{C}_0} \, 2 \, \left(\frac{\sqrt{\kappa}-1}{\sqrt{\kappa}+1}\right)^n \ \ \|r_0\|_{\hat{\mathcal{H}}^{-1}},$$

where r_n is the n-th residual, $r_0 = b - \mathcal{A}\hat{\mathcal{B}}^{-1}x_0$ and $\kappa := \kappa(\mathcal{A}\hat{\mathcal{B}}^{-1}) \leq \frac{C_1}{C_0}$ is the condition number of $\mathcal{A}\hat{\mathcal{B}}^{-1}$ in the $\mathcal{H}^{-1}$–inner product.

Proof:

$$\begin{aligned}
\|r_n\|_{\hat{\mathcal{H}}^{-1}} &= \min_{\Phi_n \in \mathcal{P}_n, \Phi_n(0)=1} \|\Phi_n(\mathcal{A}\hat{\mathcal{B}}^{-1})r_0\|_{\hat{\mathcal{H}}^{-1}} \\
&\leq \bar{C}_1 \min_{\Phi_n \in \mathcal{P}_n, \Phi_n(0)=1} \|\Phi_n(\mathcal{A}\hat{\mathcal{B}}^{-1})r_0\|_{\mathcal{H}^{-1}} \\
&\leq \bar{C}_1 \, 2 \, \left(\frac{\sqrt{\kappa}-1}{\sqrt{\kappa}+1}\right)^n \ \ \|r_0\|_{\mathcal{H}^{-1}} \\
&\leq \frac{\bar{C}_1}{\bar{C}_0} \, 2 \, \left(\frac{\sqrt{\kappa}-1}{\sqrt{\kappa}+1}\right)^n \ \ \|r_0\|_{\hat{\mathcal{H}}^{-1}}.
\end{aligned}$$

□

Remark 5.3 *Our convergence estimate only depends on the square root of the condition number of the preconditioned problem. Note that this estimate matches, except for a leading factor, the standard estimate for the conjugate gradient method applied to positive definite symmetric problems.*

Finally, we give a well-known convergence estimate for GMRES in the l_2–norm; see e.g. Saad and Schultz [71]. This estimate is based on the eigenvalues of the preconditioned system. We make use of the fact that $\mathcal{A}\hat{\mathcal{B}}^{-1}$ is diagonalizable. This can easily be seen from the following arguments:
$\mathcal{A}\hat{\mathcal{B}}^{-1}\mathcal{H}$ is symmetric positive definite; so is

$$\mathcal{H}^{-1/2}(\mathcal{A}\hat{\mathcal{B}}^{-1}\mathcal{H})\mathcal{H}^{-1/2} = \mathcal{H}^{-1/2}\mathcal{A}\hat{\mathcal{B}}^{-1}\mathcal{H}^{1/2}.$$

Thus, $\mathcal{H}^{-1/2}\mathcal{A}\hat{\mathcal{B}}^{-1}\mathcal{H}^{1/2}$ is diagonalizable, i.e. there exists an unitary matrix $\tilde{\mathcal{Q}}$, such that

$$\mathcal{H}^{-1/2}\mathcal{A}\hat{\mathcal{B}}^{-1}\mathcal{H}^{1/2} = \tilde{\mathcal{Q}}\mathcal{D}\tilde{\mathcal{Q}}^{-1},$$

where $\mathcal{D} := diag\{\lambda_i\}$ with λ_i the eigenvalues of $\mathcal{A}\hat{\mathcal{B}}^{-1}$. Note that $\mathcal{A}\hat{\mathcal{B}}^{-1}$ and $\mathcal{H}^{-1/2}\mathcal{A}\hat{\mathcal{B}}^{-1}\mathcal{H}^{1/2}$ have the same eigenvalues. We also obtain

$$\begin{aligned} \mathcal{A}\hat{\mathcal{B}}^{-1} &= \left(\mathcal{H}^{1/2}\tilde{\mathcal{Q}}\right)\mathcal{D}(\mathcal{H}^{1/2}\tilde{\mathcal{Q}})^{-1} \\ &=: \mathcal{Q}\mathcal{D}\mathcal{Q}^{-1}, \end{aligned}$$

i.e. $\mathcal{A}\hat{\mathcal{B}}^{-1}$ is diagonalizable with the diagonalization matrix $\mathcal{Q} := \mathcal{H}^{1/2}\tilde{\mathcal{Q}}$. Unfortunately, our estimate depends on the condition number of the matrix $\mathcal{H}^{1/2}$.

Theorem 5.4 *We have*

$$\|r_n\|_2 \le \|\mathcal{H}^{1/2}\|_2 \;\; \|\mathcal{H}^{-1/2}\|_2 \;\; 2 \left(\frac{\sqrt{\kappa}-1}{\sqrt{\kappa}+1}\right)^n \;\; \|r_0\|_2,$$

where r_n is the n-th residual and $\kappa := \kappa(\mathcal{A}\hat{\mathcal{B}}^{-1}) := \lambda_{max}/\lambda_{min}$ is the spectral condition number of the preconditioned system.

Proof: Under the assumptions made, the following estimate is satisfied,

$$\|r_n\|_2 \le \|\mathcal{Q}\|_2 \;\; \|\mathcal{Q}^{-1}\|_2 \min_{\Phi_n \in \mathcal{P}_n, \Phi_n(0)=1} \max_{\lambda \in \sigma(\mathcal{A}\hat{\mathcal{B}}^{-1})} |\Phi_n(\lambda)| \, \|r_0\|_2,$$

see Saad and Schultz [71], Proposition 4. Since the spectrum of $\mathcal{A}\hat{\mathcal{B}}^{-1}$ is real and positive, we can use the Chebyshev polynomials to construct an upper bound. The given estimate follows from the invariance of the Euclidean norm under unitary transformations, i.e.

$$\begin{aligned} \|\mathcal{Q}\|_2 &= \|\mathcal{H}^{1/2}\tilde{\mathcal{Q}}\|_2 = \|\mathcal{H}^{1/2}\|_2, \\ \|\mathcal{Q}^{-1}\|_2 &= \|\tilde{\mathcal{Q}}^{-1}\mathcal{H}^{-1/2}\|_2 = \|\mathcal{H}^{-1/2}\|_2. \end{aligned}$$

□

Note that the convergence estimate in Theorem 5.4 is independent of the penalty parameter but unfortunately it normally still depends on h. To give an estimate of $\kappa(\mathcal{H}^{1/2})$ for a particular discretization, we now restrict ourselves to the mixed formulation of the equations of linear elasticity discretized by the Taylor-Hood element; see Section 3.3.1. In this case, we have

$\kappa(\mathcal{H}^{1/2}) \in O(1/h)$. Since $\mathcal{H}^{1/2}$ is symmetric positive definite, this can be immediately seen from the extreme eigenvalues:

$$\begin{aligned} \lambda_{max}(\mathcal{H}^{1/2}) &\leq const\,(1/h), \\ \lambda_{min}(\mathcal{H}^{1/2}) &\geq const\,(1+h). \end{aligned}$$

A factor of $1/h$ appears in the bound given by Theorem 5.4. However, our numerical experiments do not reflect the presence of such a factor; see Section 5.3.

5.3 Numerical Examples

In this section, we apply the triangular preconditioner to the problem of planar, linear elasticity; see Section 3.3.1. For simplicity, we work with the formulation given in Remark 3.1. All results shown are for mixed boundary conditions and the region $[-1,1] \times [-1,1]$; see Figure 4.1. We note that our model is mathematically equivalent to the full elasticity problem only in the case of homogeneous Dirichlet conditions. The relation between the penalty parameter t and the Poisson ratio ν is given by $t := (1+\nu)(1-2\nu)/(E\nu)$, where E is Young's modulus. Without loss of generality, we use $E = 1$. We discretize by a Taylor-Hood element, see Section 3.3.1.

All computations were carried out on a SUN SPARC 10 workstation using the numerical software package PETSc developed by William Gropp and Barry Smith at the Argonne National Laboratory; see Gropp and Smith [51] or Smith [80]. The initial guess is 0, and the stopping criterion is $\|r_k\|_2/\|r_0\|_2 < 10^{-5}$, where r_k is the k-th residual.

We give numerical results for two Krylov space methods, GMRES and BI-CGSTAB; see e.g. [77, 71, 40]. We use a version of GMRES without restarts but we also ran a version with a restart every 10 iterations. The number of iterations for this version was always just 1-2 iterations larger than for the GMRES method without a restart. We use right-oriented preconditioning with $\hat{\mathcal{B}}_U^{-1}$ for GMRES and left-oriented preconditioning with $\hat{\mathcal{B}}_L^{-1}$ for BI-CGSTAB and we only use the L_2- rather than the $\hat{\mathcal{H}}^{-1}-$metric. The numerical results suggest that the number of iterations is bounded independently of the critical parameters h and t. We point out that BI-CGSTAB always converges smoothly in our experiments. We test each combination of preconditioners $\hat{A}, \hat{C}$ for both limit cases, $h \to 0$ and $t \to 0$.

Table 5.1: Iteration counts for exact solvers as preconditioners for A and C, and $\nu = 0.3$.

Grid	GMRES	Bi-CGSTAB
20×10	10	5
40×20	11	5
60×30	11	6
80×40	12	6
100×50	12	6
120×60	12	6
140×70	12	6

To see how the Krylov space methods behave under the best of circumstances, we first conducted some experiments using exact solvers, i.e. $\hat{A} = A$ and $\hat{C} = C$; see Tables 5.1 and 5.4.

In other series of experiments, we use different preconditioners for A and C. We present results with a two-level multigrid preconditioner with a V-cycle including one pre- and one post-smoothing symmetric Gauss-Seidel step defining $\hat{A}$, and a one-level symmetric multiplicative overlapping Schwarz method with the minimal overlap of one node as $\hat{C}$; see Tables 5.2, 5.3, 5.5, 5.6, 5.7.

Finally, we show that the assumption $a_0 > 1$ does not appear to be necessary for the convergence of the Krylov space methods. We conducted some experiments with $\hat{A} := \epsilon A$, where ϵ is a parameter at our disposal. The numerical results show that the methods still converge, even when $\epsilon > 1$; see Table 5.8.

In all experiments, BI-CGSTAB converges twice as fast as GMRES. On the other hand, BI-CGSTAB requires twice as many matrix/vector products with the system matrix $\mathcal{A}$ and the preconditioner $\hat{\mathcal{B}}^{-1}$; see Section 2.1.4 and 2.1.5. Hence, by just comparing matrix/vector products, the two methods appear to be equally efficient. But one also has to take into account that GMRES needs more memory than BI-CGSTAB. An implementation on a parallel computer might also be faster if BI-CGSTAB is used since it is based on a 3-term recurrence.

Table 5.2: Iteration counts for a two-level multigrid preconditioner with a standard V-cycle defining $\hat{A}$, $\hat{C} = C$, and $\nu = 0.3$.

Grid	GMRES	Bi-CGSTAB
20×10	13	7
40×20	14	7
60×30	14	7
80×40	15	7
100×50	15	7
120×60	15	7
140×70	15	7

Table 5.3: Iteration counts for a two-level multigrid preconditioner with a standard V-cycle defining $\hat{A}$ and a one-level symmetric multiplicative overlapping Schwarz method with the minimal overlap of one node defining $\hat{C}$, and $\nu = 0.3$.

Grid	GMRES	Bi-CGSTAB
20×10	13	7
40×20	14	7
60×30	14	7
80×40	15	7
100×50	15	7
120×60	15	7
140×70	15	7

Table 5.4: Iteration counts for exact solvers as preconditioners for A and C on a 80×40 grid.

ν	GMRES	Bi-CGSTAB
0.3	12	6
0.4	13	7
0.49	14	7
0.499	14	7
0.4999	14	7
0.49999	14	7
0.499999	14	7
0.5	14	7

Table 5.5: Iteration counts for a two-level multigrid method with a standard V-cycle defining $\hat{A}$ and $\hat{C} = C$ on a 80×40 grid.

ν	GMRES	Bi-CGSTAB
0.3	15	7
0.4	16	7
0.49	17	7
0.499	17	7
0.4999	17	7
0.49999	17	7
0.499999	17	7
0.5	17	7

Table 5.6: Iteration counts for an exact solver as $\hat{A}$ and a one-level symmetric multiplicative overlapping Schwarz method with the minimal overlap of one node as $\hat{C}$ on a 80×40 grid.

ν	GMRES	Bi-CGSTAB
0.3	12	6
0.4	13	7
0.49	14	7
0.499	14	7
0.4999	14	7
0.49999	14	7
0.499999	14	7
0.5	14	7

Table 5.7: Iteration counts for a two-level multigrid method with a standard V-cycle as $\hat{A}$ and a one-level symmetric multiplicative overlapping Schwarz method with the minimal overlap of one node as $\hat{C}$ on a 80×40 grid.

ν	GMRES	Bi-CGSTAB
0.3	15	7
0.4	16	7
0.49	17	7
0.499	17	7
0.4999	17	7
0.49999	17	7
0.499999	17	7
0.5	17	7

Table 5.8: Iteration counts for $\hat{A} = \epsilon A$ and $\hat{C} = C$ on a 80×40 Grid with $\nu = 0.3$.

ϵ	GMRES	Bi-CGSTAB
0.8	13	6
0.9	13	6
0.99	13	6
1.01	13	6
1.1	14	6
1.2	14	6

Chapter 6

Preconditioning Second Order Symmetric Problems

6.1 Introduction

In the last decade, several methods for indefinite elliptic problems have been proposed; see Bramble, Kwak and Pasciak [17], Bramble, Leyk and Pasciak [18], Bramble, Pasciak and Xu [23], Cai and Widlund [33, 34], Mandel [61], Xu and Cai [87], Xu [86], and Yserentant [89]. All of these methods use indefinite preconditioners or a reduced (positive definite) system.

Here, we show, that under certain assumptions, it is sufficient to precondition indefinite elliptic problems with a positive definite preconditioner. By applying techniques developed in Section 4.2 to second order elliptic problems, we have rediscovered results established, by different means, by Yserentant [88]. In our opinion, the proof given in this chapter is simpler and shorter, and gives new insight into the method.

We consider indefinite linear systems arising from second order elliptic operators, such as the indefinite Helmholtz equation. The preconditioner is chosen to be a good preconditioner for the principal part of the given differential operator. The problem is discretized by conforming finite elements and we assume that the resulting linear system satisfies Babuška's inf-sup and sup-sup conditions. Our preconditioning strategy can be interpreted as a change of basis. The proof of the bound of the condition number uses the same arguments as in Section 4.2.1.

The remainder of this chapter is organized as follows. In Section 6.2, we describe the elliptic second-order problem and its discretization by finite elements. In Section 6.3, we analyze our preconditioner and give an estimate of the conditon number of the preconditioned system. In Section 6.4, we present numerical results for the indefinite Helmholtz equation.

6.2 Second Order Elliptic Problems

Let $\Omega \subset \mathbf{R}^d, d = 2, 3$, be an open, bounded polygon (resp. polyhedron) with boundary $\partial\Omega$. We consider the homogeneous Dirichlet boundary value problem

$$\begin{aligned} Lu &= f \quad in \quad \Omega \\ u &= 0 \quad on \quad \partial\Omega. \end{aligned}$$

The elliptic operator L has the form

$$Lu(x) = -\sum_{i,j=1}^{d} \frac{\partial}{\partial x_i}\left(a_{ij}(x)\frac{\partial u(x)}{\partial x_j}\right) - c(x)u(x), \tag{6.1}$$

where the matrix $(a_{ij})_{i,j}$ is symmetric and uniformly positive definite for all $x \in \Omega$. Since we only consider indefinite problems, we restrict ourselves to the case of $c(x) > 0$ for all $x \in \Omega$. All the coefficients are, by assumption, sufficiently smooth. We note that the assumption of $c(x)$ being positive is not necessary. As pointed out below, we only need that the bilinear form associated with the operator L satisfies Gårding's inequality.

Let $(\cdot,\cdot)_0$ denote the standard L_2-inner product and $(\cdot,\cdot)_1$ the usual H^1-inner product. As usual, the corresponding norms are defined by $||\cdot||_0 := \sqrt{(\cdot,\cdot)_0}$ and $||\cdot||_1 := \sqrt{(\cdot,\cdot)_1 + (\cdot,\cdot)_0}$. From the Friedrichs-Poincaré inequality, we know that $|\cdot|_1 := \sqrt{(\cdot,\cdot)_1}$ defines a norm on $X := H_0^1(\Omega)$ equivalent to $||\cdot||_1$.

The weak formulation of (6.1) is

$$\begin{aligned} \mathcal{A}(u,v) &:= \sum_{i,j=1}^{d}\int_\Omega a_{ij}(x)\frac{\partial u(x)}{\partial x_i}\frac{\partial v(x)}{\partial x_j} - \int_\Omega c(x)u(x)v(x) \\ &=: a(u,v) - c(u,v), \end{aligned}$$

for all $u, v \in X := H_0^1(\Omega)$. From the assumptions made on $(a_{ij})_{i,j}$, we see that $a(\cdot,\cdot)$ is a symmetric and positive definite bilinear form defining a norm $||\cdot||_X := \sqrt{a(\cdot,\cdot)}$ on X. Since $c(x)$ is positive on Ω, $c(\cdot,\cdot)$ is a positive definite bilinear form. Thus, the bilinear form $\mathcal{A}(\cdot,\cdot)$ satisfies Gårding's inequality, i.e. there exist constants $C_a > 0, C_c \in \mathbf{R}$, such that

$$\mathcal{A}(u,u) \geq C_a\|u\|_X^2 - C_c\|u\|_0^2 \quad \forall u \in X.$$

Furthermore, we can easily prove that $\mathcal{A}(\cdot,\cdot)$ is continuous, i.e. there exists a positive constant $C > 0$, such that

$$|\mathcal{A}(u,v)| \leq C\|u\|_X\|v\|_X \quad \forall u,v \in X. \tag{6.2}$$

Finally, we assume that

$$\mathcal{A}(u,v) = (f,v)_0 \quad \forall v \in X \tag{6.3}$$

is solvable for all $f \in X'$.

To discretize problem (6.3), we use conforming finite elements. For simplicity, we restrict ourselves to bilinear quadratic (resp. trilinear cubic) elements. Let $\mathcal{T}$ denote a triangulation of Ω. The corresponding finite element spaces are denoted by $X^h := \{v_h \in \mathcal{C}(\Omega) \cap H_0^1(\Omega) : v_h \in \mathcal{Q}_1(T),\ T \in \mathcal{T}\}$, where h represents the diameter of a typical element and $\mathcal{Q}_1(T)$ denotes the set of bilinear (resp. trilinear) functions on T. The discretized version of (6.3) is

$$\mathcal{A}(u_h,v_h) = (f,v_h)_0 \quad \forall v_h \in X^h. \tag{6.4}$$

The following lemma shows that (6.4) is solvable if h is small enough; see e.g. Hackbusch [52], p. 155. To the best of our knowledge, the earliest reference to this result is an article by Schatz; see [74].

Lemma 6.1 *Let $\mathcal{A}(\cdot,\cdot)$ satisfy Gårding's inequality and let $X^h \subset X$, such that $\lim_{h\to 0} dist(u,X^h) = 0 \ \ \forall u \in X$.*

If h is small enough, the inf-sup condition

$$\inf_{u_h\in X^h}\sup_{v_h\in X^h}\frac{\mathcal{A}(u_h,v_h)}{\|u_h\|_X\|v_h\|_X} \geq \alpha > 0 \tag{6.5}$$

holds, with α independent of h.

From the continuity assumption (6.2), we obtain

$$\sup_{u_h\in X^h}\sup_{v_h\in X^h}\frac{\mathcal{A}(u_h,v_h)}{\|u_h\|_X\|v_h\|_X} \leq C. \tag{6.6}$$

According to the well-known theory of Babuška, see Section 1.1, we now know that (6.4) is uniquely solvable. Note that the second inf-sup condition in Theorem 1.1 follows from (6.5) by symmetry.

6.3 The Preconditioning Strategy

In this section, we construct a positive definite preconditioner for the indefinite problem (6.4). We prove that already a good preconditioner for $a(\cdot,\cdot)$ can guarantee an $h-$independent convergence rate of the PCR-method.

We denote the matrix representation of the bilinear forms $\mathcal{A}(\cdot,\cdot), a(\cdot,\cdot)$ and $c(\cdot,\cdot)$ by $\mathcal{A}, A$ and C. Hence, we have $\mathcal{A} = A - C$.

We assume that $\hat{\mathcal{B}} := \hat{A}$ is symmetric and positive definite and defines a norm on $\mathbf{R}^n$ that is equivalent to $\|\cdot\|_X$, i.e. there exist positive constants a_0, a_1, independent of h, such that

$$a_0\|u\|_X \leq \|\hat{A}^{1/2}u\|_2 \leq a_1\|u\|_X, \tag{6.7}$$

where $\|\cdot\|_2$ denotes the l_2-norm.

Examples of methods that meet these requirements are given by domain decomposition and multigrid methods or, more generally, by Schwarz methods; see Section 2.2.

In view of Theorem 2.2 and Remark 2.2, our goal is to give an estimate of the condition number $\kappa(\hat{\mathcal{B}}^{-1}\mathcal{A})$. Since $\hat{\mathcal{B}} := \hat{A}$ is positive definite, $\hat{\mathcal{B}}^{-1}\mathcal{A}$ and $\hat{\mathcal{B}}^{-1/2}\mathcal{A}\hat{\mathcal{B}}^{-1/2}$ have the same eigenvalues and $\rho(\hat{\mathcal{B}}^{-1}\mathcal{A}) = \|\hat{\mathcal{B}}^{-1/2}\mathcal{A}\hat{\mathcal{B}}^{-1/2}\|_2$. Here, we have used that $\hat{\mathcal{B}}^{-1/2}\mathcal{A}\hat{\mathcal{B}}^{-1/2}$ is normal. Thus, we only have to provide upper bounds for $\|\hat{\mathcal{B}}^{-1/2}\mathcal{A}\hat{\mathcal{B}}^{-1/2}\|_2$ and $\|(\hat{\mathcal{B}}^{-1/2}\mathcal{A}\hat{\mathcal{B}}^{-1/2})^{-1}\|_2$.

We obtain the next lemma from (6.5) and (6.6) by performing a change of basis; see also Lemma 4.4.

Lemma 6.2 *There exist positive constants C_0, C_1, independent of h, such that*

$$C_0 \leq \inf_{v_h\neq 0}\sup_{u_h\neq 0}\frac{(\hat{\mathcal{B}}^{-1/2}\mathcal{A}\hat{\mathcal{B}}^{-1/2}u_h, v_h)_2}{\|u_h\|_2\|v_h\|_2},$$
$$\sup_{v_h\neq 0}\sup_{u_h\neq 0}\frac{(\hat{\mathcal{B}}^{-1/2}\mathcal{A}\hat{\mathcal{B}}^{-1/2}u_h, v_h)_2}{\|u_h\|_2\|v_h\|_2} \leq C_1.$$

Now, we are able to give our condition number estimate. It follows immediately from the definition of the condition number in combination with Lemmata 4.2, 4.3, and 6.2.

Theorem 6.1 *The condition number of $\hat{\mathcal{B}}^{-1}\mathcal{A}$ is bounded independently of the discretization parameter $h < h_0$, i.e.*

$$\kappa(\hat{\mathcal{B}}^{-1}\mathcal{A}) \leq \frac{C_1}{C_0},$$

where h_0 is small enough to guarantee that the inf-sup condition (6.5) is satisfied.

Remark 6.1 *There exist other parameters that might influence the convergence rate of the iterative method, such as the number of levels in a multigrid method or the number of subdomains in a domain decomposition method. The PCR-method also converges independently of these parameters if the condition number of $\hat{A}^{-1}A$ is independent of them. This condition is satisfied for many domain decomposition and multilevel methods.*

Remark 6.2 *Assuming $h < h_0$, we exclude the case of $\mathcal{A}$ being (nearly) singular. The numerical experiments show that the proposed method still behaves well in this case; see Section 6.4.*

6.4 Numerical Results

In this section, we apply our preconditioner to the indefinite Helmholtz equation with homogeneous Dirichlet boundary conditions on a square, i.e.

$$\begin{aligned} -\Delta u - qu &= f \quad in \quad \Omega := [0,1]^2, \\ u &= 0 \quad on \quad \partial\Omega, \end{aligned}$$

where the right hand side $f(x, y)$ is always chosen so that the exact solution is $u(x, y) = x\, e^{xy} \sin(\pi\, x) \sin(\pi\, y)$; see also Figure 6.1.

All the computations were carried out on a SPARC 10 workstation using Matlab 4.2a. The initial guess is 0 and the stopping criterion $\|r_k\|_2/\|r_0\|_2 < 10^{-5}$, where r_k is the k-th residual.

To see how the PCR-method behaves under the best of circumstances, we conducted some experiments using an exact solver as a preconditioner, i.e. $\hat{A} = A$; see Table 6.1. The method works as predicted for the limit case $h \to 0$.

Another critical parameter is the constant q since it determines k, the number of negative eigenvalues. According to Theorem 2.2 we should expect that the number of iterations will grow as q is increased. This is confirmed by the experiments; see Table 6.1.

In the theory developed in the previous sections, we have excluded the (nearly) singular case. In our model problem, $\mathcal{A}$ is singular when q is an eigenvalue of $-\Delta$. It is well-known that the eigenvalues of $-\Delta u = \lambda u$ on

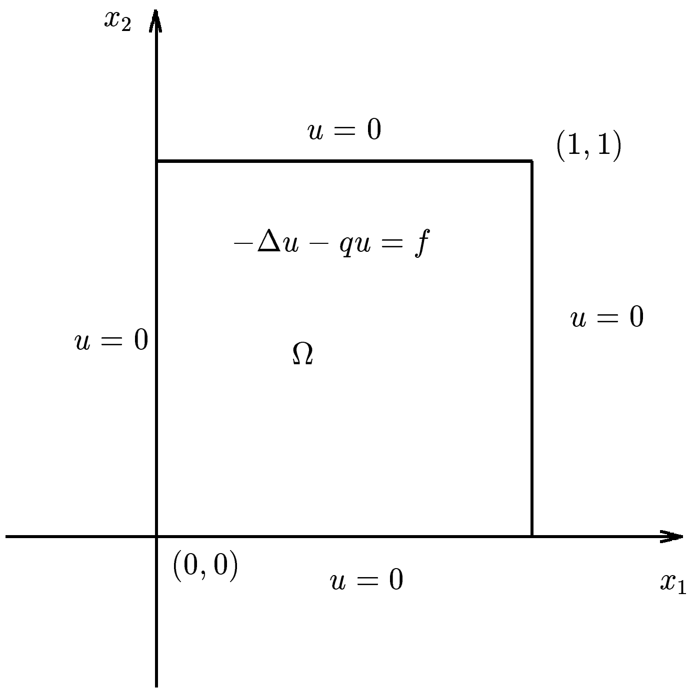

Figure 6.1: The domain.

the unit square are given by $(i^2 + j^2)\,\pi^2, \;\; i, j = 1, 2, 3, \ldots$. To demonstrate the robustness of the proposed algorithm, we consider two cases for q, one when it is an exact eigenvalue and the other when it is an approximation of an eigenvalue. This is done for the two different eigenvalues $2\,\pi^2$ and $8\,\pi^2$, i.e. $q = 2\,\pi^2 \approx 19.7392$, $q = 19.72$ and $q = 8\,\pi^2 \approx 78.9568$, $q = 78.94$. The experiments show that the algorithm does not deteriorate, see Tables 6.2,6.3,6.4,6.5, although the finite element solution might be quite inaccurate. The same accuracy is obtained by solving the system by Gaussian elimination. This shows that the loss of precision cannot be traced back to the stopping criterion but is inherent to the finite element discretization. We also ran the same experiments with different random right hand sides, to eliminate the possibility that our results are due to a special right hand side. We could not detect any significant difference in the number of iterations. As we did not know the exact solution, we cannot say anything about the error of the finite element solution.

Table 6.1: PCR-method with an exact solver as a preconditioner, i.e. $\hat{A} = A$.

q	$h = \frac{1}{10}$	$h = \frac{1}{20}$	$h = \frac{1}{30}$	$h = \frac{1}{40}$	$h = \frac{1}{50}$	$h = \frac{1}{60}$	$h = \frac{1}{70}$
10	4	5	5	5	5	5	5
20	8	7	7	7	7	7	7
30	7	7	7	7	7	7	7
40	9	9	9	9	9	9	9
50	10	11	11	11	11	11	11
60	10	11	11	11	11	11	11
70	10	11	12	12	12	12	12
80	12	15	15	15	16	15	15
90	12	13	13	14	16	16	16
100	12	17	17	17	17	17	17
160	14	18	19	19	19	19	19
300	24	43	33	33	33	33	33

Table 6.2: PCR-method with an exact solver as a preconditioner, i.e. $\hat{A} = A$, and $q = 19.72$.

q	$h = \frac{1}{10}$	$h = \frac{1}{20}$	$h = \frac{1}{30}$	$h = \frac{1}{40}$	$h = \frac{1}{50}$	$h = \frac{1}{60}$	$h = \frac{1}{70}$
Iterations	7	9	9	9	9	9	9
$\|u - u_h\|_\infty$	0.9376	0.7049	0.5020	0.3581	0.2616	0.1969	0.1523

Table 6.3: PCR-method with an exact solver as a preconditioner, i.e. $\hat{A} = A$, and $q = 2\,\pi^2 \approx 19.7392$.

q	$h = \frac{1}{10}$	$h = \frac{1}{20}$	$h = \frac{1}{30}$	$h = \frac{1}{40}$	$h = \frac{1}{50}$	$h = \frac{1}{60}$	$h = \frac{1}{70}$
Iterations	7	9	9	9	10	10	10
$\|u - u_h\|_\infty$	1.0508	1.0408	1.0389	1.0377	1.0380	1.0378	1.0377

Table 6.4: PCR-method with an exact solver as a preconditioner, i.e. $\hat{A} = A$, and $q = 78.94$.

q	$h = \frac{1}{10}$	$h = \frac{1}{20}$	$h = \frac{1}{30}$	$h = \frac{1}{40}$	$h = \frac{1}{50}$	$h = \frac{1}{60}$	$h = \frac{1}{70}$
Iterations	12	13	16	15	11	11	11
$\|u - u_h\|_\infty$	0.1410	0.1260	0.1119	0.0994	0.0723	0.0685	0.0660

Table 6.5: PCR-method with an exact solver as a preconditioner, i.e. $\hat{A} = A$, and $q = 8\,\pi^2 \approx 78.9568$.

q	$h = \frac{1}{10}$	$h = \frac{1}{20}$	$h = \frac{1}{30}$	$h = \frac{1}{40}$	$h = \frac{1}{50}$	$h = \frac{1}{60}$	$h = \frac{1}{70}$
Iterations	12	13	16	15	11	11	11
$\|u - u_h\|_\infty$	0.1414	0.1279	0.1162	0.1070	0.0725	0.0687	0.0661

Bibliography

[1] Douglas N. Arnold. Discretization by finite elements of a model parameter dependent problem. *Numer. Math.*, 37:405–421, 1981.

[2] Douglas N. Arnold and Franco Brezzi. Some new elements for the Reissner-Mindlin plate model. In *Boundary Value Problems for Partial Differential Equations and Applications*, pages 287–292. Masson, 1993.

[3] Douglas N. Arnold, Franco Brezzi, and Michel Fortin. A stable finite element for the Stokes equations. *Calcolo*, 21:337–344, 1984.

[4] Douglas N. Arnold and Richard S. Falk. A uniformly accurate finite element method for the Reissner- Mindlin plate. *SIAM J. Numer. Anal.*, 26(6):1276–1290, 1989.

[5] Kenneth J. Arrow, Leonid Hurwicz, and Hirofumi Uzawa. *Studies in Linear and Nonlinear Programming.* Stanford University Press, 1958.

[6] Steven F. Ashby, Thomas A. Manteuffel, and Paul E. Saylor. A taxonomy for conjugate gradient methods. *SIAM J.Numer.Anal.*, 27:1542–1568, 1990.

[7] Ivo Babuška and Manil Suri. Locking effects in the finite element approximation of elasticity problems. *Numer. Math.*, 62(4):439–463, 1992.

[8] Ivo Babuška and Manil Suri. On locking and robustness in the finite element method. *SIAM J. Numer. Anal.*, 29(5):1261–1293, 1992.

[9] Ivo Babuška. Error bounds for finite element method. *Numer. Math.*, 16:322–333, 1971.

[10] Ivo Babuška. The finite element method with Lagrange multipliers. *Numer. Math.*, 20:179–192, 1973.

[11] Ivo Babuška. The finite element method with penalty. *Math. Comp.*, 27:221–228, 1973.

[12] Ivo Babuška and A.K. Aziz. Survey lectures on the mathematical foundations of the finite element method. In A.K. Aziz, editor, *The Mathematical Foundations of the Finite Element Method with Applications to Partial Differential Equations*, pages 3–359. Academic Press, 1972.

[13] Randolph E. Bank, Bruno D. Welfert, and Harry Yserentant. A class of iterative methods for solving saddle point problems. *Numer. Math.*, 56:645–666, 1990.

[14] Dietrich Braess. A multigrid method for the membrane problem. *Comp. Mech.*, 3:321–329, 1988.

[15] Dietrich Braess. *Finite Elemente.* Springer-Verlag, 1992.

[16] Dietrich Braess and C. Blömer. A multigrid method for a parameter dependent problem in solid mechanics. *Numer. Math.*, 57:747–761, 1990.

[17] James H. Bramble, Do Y. Kwak, and Joseph E. Pasciak. Uniform convergence of multigrid V-cycle iterations for indefinite and nonsymmetric problems. *SIAM J. Numer. Anal.*, 31(6):1746–1763, 1994.

[18] James H. Bramble, Zbigniew Leyk, and Joseph E. Pasciak. Iterative schemes for non-symmetric and indefinite elliptic boundary value problems. *Math. Comp.*, 60:1–22, 1993.

[19] James H. Bramble and Joseph E. Pasciak. A preconditioning technique for indefinite systems resulting from mixed approximations of elliptic problems. Technical report, Cornell University, 1987.

[20] James H. Bramble and Joseph E. Pasciak. A preconditioning technique for indefinite systems resulting from mixed approximations of elliptic problems. *Math. Comp.*, 50:1–17, 1988.

[21] James H. Bramble, Joseph E. Pasciak, and Apostol E. Vasilev. Analysis of the inexact Uzawa algorithm for saddle point problems. Technical report, Cornell University, 1994.

[22] James H. Bramble, Joseph E. Pasciak, Junping Wang, and Jinchao Xu. Convergence estimates for product iterative methods with applications to domain decomposition. *Math. Comp.*, 57(195):1–21, 1991.

[23] James H. Bramble, Joseph E. Pasciak, and Jinchao Xu. The analysis of multigrid algorithms for nonsymmetric and indefinite elliptic problems. *Math. Comp.*, 51:389–414, 1988.

[24] Susanne C. Brenner. A nonconforming mixed multigrid method for the pure displacement problem in planar linear elasticity. *SIAM J. Numer. Anal.*, 30(1):116–135, 1993.

[25] Susanne C. Brenner. A nonconforming mixed multigrid method for the pure traction problem in planar linear elasticity. *Math. Comp.*, 63(208):435–460, October 1994.

[26] Susanne C. Brenner and L. Ridgway Scott. *The Mathematical Theory of Finite Element Methods.* Springer-Verlag, 1994.

[27] Franco Brezzi. On the existence, uniqueness, and approximation of saddle-point problems arising from Lagrange multipliers. *R.A.I.R.O. Anal. Numer.*, 8:129–151, 1974.

[28] Franco Brezzi, Klaus-Jürgen Bathe, and Michel Fortin. Mixed-interpolated elements for Reissner-Mindlin plates. *Int. J. Num. Meth. Eng. 28*, 28:1787–1801, 1989.

[29] Franco Brezzi and Michel Fortin. Numerical approximation of Mindlin-Reissner plates. *Math. Comp.*, 47:151–158, 1986.

[30] Franco Brezzi and Michel Fortin. *Mixed and Hybrid Finite Element Methods.* Springer-Verlag, 1991.

[31] Are Magnus Bruaset. *A survey of preconditioned iterative methods.* Pitman Research Notes in Mathematics Series. Longman Scientific and Technical, 1995.

[32] Xiao-Chuan Cai. The Use of Pointwise Interpolation in Domain Decomposition Methods with Nonnested Meshes. *SIAM J. Sci. Comput.*, 16(1):250–256, Jan. 1995.

[33] Xiao-Chuan Cai and Olof Widlund. Domain decomposition algorithms for indefinite elliptic problems. *SIAM J. Sci. Statist. Comput.*, 13(1):243–258, January 1992.

[34] Xiao-Chuan Cai and Olof Widlund. Multiplicative Schwarz algorithms for some nonsymmetric and indefinite problems. *SIAM J. Numer. Anal.*, 30(4):936–952, August 1993.

[35] T. F. Chan, B. F. Smith, and J. Zou. Overlapping Schwarz Methods on Unstructured Meshes using Non-matching Coarse Grids. *Numer. Math.* To appear. Also available as Technical Report, UCLA CAM Report 94-04.

[36] T.F. Chan and B.F. Smith. Multigrid and Domain Decomposition on Unstructured Grids. *Electr. Trans. Numer. Anal.*, 2:171–182, Dec. 1994.

[37] Maksymilian Dryja, Marcus Sarkis, and Olof B. Widlund. Multilevel Schwarz methods for elliptic problems with discontinuous coefficients in three dimensions. Technical Report 662, Department of Computer Science, Courant Institute, March 1994. To appear in Numer. Math.

[38] Maksymilian Dryja, Barry F. Smith, and Olof B. Widlund. Schwarz analysis of iterative substructuring algorithms for elliptic problems in three dimensions. *SIAM J. Numer. Anal.*, 31(6):1662–1694, December 1994.

[39] Maksymilian Dryja and Olof B. Widlund. Schwarz methods of Neumann-Neumann type for three-dimensional elliptic finite element problems. *Comm. Pure Appl. Math.*, 48(2), February 1995.

[40] Stanley C. Eisenstat, Howard C. Elman, and Martin H. Schultz. Variational iterative methods for nonsymmetric systems of linear equations. *SIAM J. Numer. Anal.*, 20 (2):345–357, 1983.

[41] Howard Elman. *Iterative Methods for Large, Sparse, Nonsymmetric Systems of Linear Equations.* PhD thesis, Yale University, 1982.

[42] Howard Elman. Multigrid and Krylov subspace methods for the discrete Stokes equations. Technical Report UMIA CS-TR-94-76, University of Maryland, June 1994.

[43] Howard Elman and Gene Golub. Inexact and preconditioned Uzawa algorithms for saddle point problems. *SIAM J. Numer. Anal.*, 31(6), December 1994.

[44] Howard Elman and David Silvester. Fast Nonsymmetric Iterations and Preconditioning for Navier-Sokes Equations. Technical Report UMIAC CS-TR-94-66, University of Maryland, June 1994.

[45] V. Faber and T. A. Manteuffel. Orthogonal error methods. *SIAM J. Numer. Anal.*, 24:170–187, 1987.

[46] Michel Fortin. *Finite Element Solutions of the Navier-Stokes equations*, pages 239–284. Acta Numerica 1993. Cambridge University Press, 1993.

[47] R. Freund, G. H. Golub, and N. Nachtigal. *Iterative Solution of Linear Systems*, pages 57–100. Acta Numerica 1992. Cambridge University Press, 1991.

[48] Vivette Girault and Pierre-Arnaud Raviart. *Finite Element Methods for Navier-Stokes Equations*. Springer-Verlag, 1986.

[49] Gene H. Golub and Diane P. O'Leary. Some history of the conjugate gradient and Lanczos algorithms: 1948-1976. *SIAM Review*, 31:50–102, 1989.

[50] Michael Griebel. *Multilevelmethoden als Iterationsverfahren über Erzeugendensystemen*. Teubner Skripten zur Numerik, Teubner-Verlag, Stuttgart, 1994.

[51] William D. Gropp and Barry F. Smith. Scalable, extensible, and portable numerical libraries. In *Proceedings of Scalable Parallel Libraries Conference*, pages 87–93. IEEE, 1994.

[52] Wolfgang Hackbusch. *Theorie und Numerik elliptischer Differentialgleichungen*. Teubner-Verlag, Stuttgart, 1986. Engl. Transl.: Elliptic Differential Equations. Theory and Numerical Treatment. Springer, New York, 1992.

[53] Wolfgang Hackbusch. *Iterative Lösung großer schwachbesetzter Gleichungssysteme*. Teubner-Verlag, Stuttgart, 1991. Engl. Transl.: Iterative Solution of Large Sparse Systems of Equations. Springer, New York, 1994.

[54] Ziping Huang. A Multi-Grid Algorithm for Mixed Problems with Penalty. *Numer. Math.*, 57:227–247, 1990.

[55] Thomas J.R. Hughes. *The Finite Element Method.* Prentice-Hall, 1987.

[56] Michael Jung. Konvergenzfaktoren von Mehrgitterverfahren für Probleme der ebenen, linearen Elastizitätstheorie. *ZAMM*, 67:165–173, 1987.

[57] Axel Klawonn. An optimal preconditioner for a class of saddle point problems with a penalty term. Technical Report 676, Courant Institute of Mathematical Sciences, New York University, December 1994.

[58] Axel Klawonn. An optimal preconditioner for a class of saddle point problems with a penalty term, Part II: General theory. Technical Report 14/95, Westfälische Wilhelms-Universität Münster, Germany, April 1995. Also available as Technical Report 683 at the Courant Institute of Mathematical Sciences, New York University.

[59] M. Křížek and P. Neittaanmäki. *Finite Element Approximation of Variational Problems and Applications.* Pitman Monographs and Surveys in Pure and Applied Mathematics. Longman Scientific and Technical, 1990.

[60] Richard H. MacNeal. *Finite Elements: Their Design and Performance.* Marcel Dekker, Inc., New York-Basel-Hong Kong, 1994.

[61] Jan Mandel. Multigrid convergence for nonsymmetric, indefinite variational problems and one smoothing step. *Appl. Math. Comput.*, 19:201–216, 1986.

[62] T. A. Manteuffel. The Tschebychev iteration for nonsymmetric linear systems. *Numer. Math.*, 28:307–327, 1977.

[63] N.M. Nachtigal, S.C. Reddy, and L.N. Trefethen. How fast are nonsymmetric matrix iterations? *SIAM J. Matrix Anal.*, 13:778–795, 1992.

[64] J. Nečas. Sur une méthode pour resoudre les equations aux derivées partielles du type elliptique, voisine de la variationelle. *Ann. Scuola Norm. Sup. Pisa*, 16:305–326, 1962.

[65] Peter Oswald. *Multilevel Finite Element Approximation. Theory and Applications.* Teubner Skripten zur Numerik, Teubner Verlag, Stuttgart, 1994.

[66] Christopher C. Paige and Michael A. Saunders. Solutions of sparse linear systems of linear equations. *SIAM J. Numer. Anal.*, 12:617–629, 1975.

[67] Petra Peisker and Dietrich Braess. Uniform convergence of mixed interpolated elements for Reissner-Mindlin plates. M^2AN, 26:557–574, 1992.

[68] Einar M. Rønquist. A domain decomposition solver for the incompressible Navier-Stokes equations, 1994. Workshop on Spectral Element Methods, North Carolina State University.

[69] Torgeir Rusten and Ragnar Winther. A preconditioned iterative method for saddle point problems. *SIAM J. Matrix Anal. Appl.*, 13:887–904, 1992.

[70] Y. Saad. Variations of Arnoldi's method for computing eigenelements of large unsymmetric matrices. *Lin. Alg. Appl.*, 34:269–295, 1980.

[71] Y. Saad and M. H. Schultz. GMRES: A generalized minimal residual algorithm for solving nonsymmetric linear systems. *SIAM J. Sci. Stat. Comp.*, 7:856–869, 1986.

[72] Marcus Sarkis. Two-level Schwarz methods for nonconforming finite elements and discontinuous coefficients. In N. Duane Melson, Thomas A. Manteuffel, and Steve F. McCormick, editors, *Proceedings of the Sixth Copper Mountain Conference on Multigrid Methods, Volume 2*, number 3224, pages 543–566, Hampton VA, 1993. NASA.

[73] Marcus Sarkis. *Schwarz preconditioners for elliptic problems with discontinuous coefficients using conforming and non-conforming elements.* PhD thesis, Courant Institute of the Mathematical Sciences, New York University, 1994.

[74] Alfred H. Schatz. An observation concerning Ritz-Galerkin methods with indefinite bilinear forms. *Math.Comp.*, 28:959–962, 1974.

[75] Werner Schmid, Meinhard Paffrath, and Ronald H.-W. Hoppe. Application of iterative methods for solving nonsymmetric linear systems in the simulation of semiconductor processing. *Surv. Math. Ind.*, 5(1):1–26, 1995.

[76] David Silvester and Andrew Wathen. Fast iterative solutions of stabilised Stokes systems Part II: Using general block preconditioners. *SIAM J. Numer. Anal.*, 31(5):1352–1367, 1994.

[77] Gerard L.G. Sleijpen, Henk A. Van der Vorst, and Diederik R. Fokkema. BiCGstab(l) and other hybrid Bi-CG methods. Technical report, University of Utrecht, February 1994.

[78] Barry F. Smith. *Domain decomposition algorithms for the partial differential equations of linear elasticity.* PhD thesis, Courant Institute of the Mathematical Sciences, New York University, 1990.

[79] Barry F. Smith. An optimal domain decomposition preconditioner for the finite element solution of linear elasticity problems. *SIAM J. Sci. Stat. Comput.*, 13(1):364–378, January 1992.

[80] Barry F. Smith. Extensible PDE solvers package users manual. Technical Report ANL 94/40, Argonne National Laboratory, September 1994.

[81] Barry F. Smith, Petter Bjørstad, and William Gropp. *Domain Decomposition: Parallel Multilevel Methods for Elliptic Partial Differential Equations.* Cambridge University Press, New York, 1995. To appear.

[82] Gerhard Starke. Iterative Methods and Decomposition-Based Preconditioners for Nonsymmetric Elliptic Boundary Value Problems, July 1994. Habilitationsschrift.

[83] Henk A. van der Vorst. BI-CGSTAB: A fast and smoothly converging variant of Bi-CG for the solution of nonsymmetric linear systems. *SIAM J. Sci. Comput.*, 13(2):631–644, Mar. 1992.

[84] Rüdiger Verfürth. Error estimates for a mixed finite element approximation of the Stokes equations. *R.A.I.R.O. Num. Anal.*, 18(2):175–182, 1984.

[85] Andrew Wathen and David Silvester. Fast iterative solution of stabilised Stokes systems Part I: Using simple diagonal preconditioners. *SIAM J. Numer. Anal.*, 30(3):630–649, 1993.

[86] Jinchao Xu. A new class of iterative methods for nonselfadjoint or indefinite problems. *SIAM J. Numer. Anal.*, 29:303–319, 1992.

[87] Jinchao Xu and Xiao-Chuan Cai. A preconditioned GMRES method for nonsymmetric or indefinite problems. *Math. Comp.*, 59:311–319, 1992.

[88] H. Yserentant. Preconditioning indefinite discretization matrices. *Numer. Math.*, 54:719–734, 1988.

[89] Harry Yserentant. On the multi-level splitting of finite element spaces for indefinite elliptic boundary value problems. *SIAM J. Numer. Anal.*, 23:581–595, 1986.

[90] Xuejun Zhang. *Studies in Domain Decomposition: Multilevel Methods and the Biharmonic Dirichlet Problem.* PhD thesis, Courant Institute, New York University, September 1991.

[91] Xuejun Zhang. Multilevel Schwarz methods. *Numer. Math.*, 63(4):521–539, 1992.