Preconditioning discretizations of systems of partial differential equations

Kent-Andre Mardal[1]* and Ragnar Winther[2]

[1] *Center for Biomedical Computing, Simula Research Laboratory P.O. Box 134, No-134 Lysaker, Norway, e-mail:kent-and@simula.no, url:http://www.simula.no/people/kent-and*

[2] *Centre of Mathematics for Applications and Department of Informatics, University of Oslo, 0316 Oslo, Norway, e-mail:ragnar.winther@cma.uio.no, url:http://heim.ifi.uio.no/~rwinther/*

KEY WORDS: partial differential equations, finite element methods, discrete systems, preconditioning

SUMMARY

This survey paper is based on three talks given by the second author at the London Mathematical Society Durham Symposium on Computational Linear Algebra for Partial Differential Equations in the summer of 2008. The main focus will be on an abstract approach to the construction of preconditioners for symmetric linear systems in a Hilbert space setting. Typical examples which are covered by this theory are systems of partial differential equations which correspond to saddle point problems. We will argue that the mapping properties of the coefficient operators suggest that block diagonal preconditioners are natural choices for these systems. To illustrate our approach a number of examples will be considered. In particular, parameter–dependent systems arising in areas like incompressible flow, linear elasticity, and optimal control theory will be studied. The paper contains analysis of several models which have previously been discussed in the literature . However, here each example is discussed with reference to a more unified abstract approach.

KEY WORDS:

1. Introduction

It is an accepted fact that in order to properly design numerical methods for systems of partial differential equations one has to rely on the specific properties of the underlying differential systems themselves. In this paper we argue that the same is true for the construction of preconditioners, i.e., we argue that the structure of the preconditioners for the discrete systems are in some sense dictated by the properties of the corresponding continuous system. As a

---

Contract/grant sponsor: Both authors were supported by Center of Excellence grants from the Norwegian Research Council. The work of the first author was also supported by the Norwegian Research Council under grant 170650.

consequence, our study here will start by discussing so–called Krylov space methods in the setting of Hilbert spaces. The properties of these methods will then motivate the introduction of the concept of preconditioners for general continuous differential systems. This concept will be illustrated by various examples.

The discussion of preconditioned Krylov space methods for the continuous systems will be the starting point for a corresponding discrete theory. We will argue that if we have identified a proper preconditioner for the continuous problem and have a stable discretization of the problem, then we also have obtained the basic structure for a preconditioner of the corresponding discrete problems. In this respect the approach taken here corresponds to the analysis already given in [2, 3]. However, here we show how a number of different problems can be treated by the same abstract framework. Furthermore, in Section 5 we focus on the close link between the proper choice of preconditioner and the classical variational theory of Babuška [12, 13].

The strength of the approach taken in this paper is best illustrated by the construction of preconditioners for parameter–dependent problems. By a systematic approach we easily identify preconditioners which behave uniformly both with respect to these model parameters and the discretization parameter. The list of examples studied below includes stationary systems obtained from implicit time discretizations of the time dependent Stokes problem, where the time step is a critical parameter. Another example we will consider is the singular perturbation problem referred to as the Reissner–Mindlin plate model, where the thickness of the plate enters as a parameter. An example of an optimal control problem will also be discussed. It is well known that the solutions of such systems depend strongly on so–called regularization parameters, but below we show that for some problems it is possible to design preconditioners such that the conditioning of the preconditioned systems are bounded uniformly with respect to the regularization parameters.

For the construction of practical preconditioners for discrete systems, the computational cost of evaluating these operators and the memory requirements of these procedures are key factors. These issues will not be discussed in full detail in this paper. The aim of the discussions here is to identify what we refer to as *canonical preconditioners*, which are block diagonal operators suggested by the mapping properties of the coefficient operators of the systems, i.e., the blocks will typically correspond to exact inverses of discretizations of differential operators. These canonical operators identify the basic structure of the preconditioner. However, in order to use this approach to construct practical and efficient iterative schemes, the different blocks of the preconditioner have to be replaced by norm equivalent operators constructed, for example, by domain decomposition or multigrid techniques. A brief discussion of these issues will be included in Sections 6 and 7.

There are close similarities between the abstract approach taken in this paper and the discussion of "operator preconditioning" given in [48]. The common ingredients are the use of mapping properties of the underlying continuous operators and numerical stability to derive the basic structure of the preconditioners for the finite dimensional systems derived from the discretization procedure. A slightly alternative approach, taken for example in the texts [37, 93], is to rely more directly on the properties of the discrete problems, and then use a matrix–based framework to analyze the discrete preconditioned systems. Furthermore, the survey paper [92] studies the construction of preconditioners by exploiting a multiscale structure of the underlying function spaces.

An outline of the paper is as follows. In Section 2 we briefly discuss Krylov space methods and motivate the need of preconditioners for operators with unbounded spectrum, while in Section 3 we relate these concepts to examples of saddle point problems. Properties of some parameter–dependent problems are described in Section 4. In Section 5 we discuss how the structure of uniform preconditioners for stable finite element systems are related to the continuous systems, while Section 6 is mostly devoted to the representation of the finite element operators and the corresponding preconditioners. Finally, a number of fully discrete problems are studied in Section 7.

2. Krylov space methods and preconditioning

Let X be a separable, real Hilbert space, with inner product $\langle\cdot,\cdot\rangle$ and associated norm $\|\cdot\|$, and assume that $\mathcal{A}: X \to X$ is a symmetric isomorphism on X, i.e.,

$$\mathcal{A}, \mathcal{A}^{-1} \in \mathcal{L}(X, X).$$

Here $\mathcal{L}(X, X)$ is the set of bounded linear operators mapping X to itself. In general, for an operator $\mathcal{A} \in \mathcal{L}(X, Y)$ the corresponding operator norm is given by

$$\|\mathcal{A}\|_{\mathcal{L}(X,Y)} = \sup_{x \in X} \frac{\|\mathcal{A}x\|_Y}{\|x\|_X}.$$

The purpose of this section is to give an overview of so–called Krylov space methods for linear systems of the form

$$\mathcal{A}x = f, \tag{2.1}$$

where the right hand side $f \in X$ is given. Krylov space methods are among the most effective iterative methods for linear systems arising from discretizations of partial differential equations. As a consequence, such methods are usually discussed for systems defined on a finite–dimensional vector space. However, for the discussion given below, it is crucial to allow the space X to be infinite–dimensional.

Krylov space methods are composed of simple iterations that produce a sequence of approximate solutions $\{x_m\}$, which converges to the exact solution x as m increases. Of course, no exact inversion of the operator $\mathcal{A}$ is performed, but typically one evaluation of the operator $\mathcal{A}$ is required for each iteration. In some sense, the Krylov space methods can be seen as improvements of a simple fixed point iteration of the form

$$x_{m+1} = x_m - \alpha(\mathcal{A}x_m - f),$$

frequently referred to as the Richardson iteration in the numerical analysis literature. Here α is a real valued parameter which has to be properly chosen. However, the Krylov space methods are more robust and more efficient. If the operator $\mathcal{A}$ and the right hand side f are given, then the Krylov space of order m is given as

$$K_m = K_m(\mathcal{A}, f) = \mathrm{span}\{\, f, \mathcal{A}f, \ldots, \mathcal{A}^{m-1}f \,\}.$$

The most celebrated Krylov space method is the *conjugate gradient method* of Hestenes and Stiefel [47], see also [43, 45, 90], where it is assumed that the operator $\mathcal{A}$ is symmetric and positive definite. In this case the unique solution $x \in X$ of the system (2.1) can be characterized as

$$x = \arg\min_{y \in X} E(y),$$

where $E(y) = \langle \mathcal{A}y, y\rangle - 2\langle f, y\rangle$. The approximation $x_m \in K_m$ is defined as

$$x_m = \arg\min_{y \in K_m} E(y).$$

Alternatively, $x_m \in K_m$ solves the Galerkin system

$$\langle \mathcal{A}x_m, y\rangle = \langle f, y\rangle, \quad y \in K_m.$$

Furthermore, an effective computation of the sequence $\{x_m\}$ is based on a recurrence relation requiring a single evaluation of the operator $\mathcal{A}$ per iteration.

It is straightforward to check that x_m is the best approximation of x in K_m in the sense that

$$\|x - x_m\|_{\mathcal{A}} = \inf_{y \in K_m} \|x - y\|_{\mathcal{A}},$$

where $\|y\|_{\mathcal{A}}^2 = \langle \mathcal{A}y, y\rangle$. By combining this observation with spectral theory we obtain an error estimate of the form

$$\|x - x_m\|_{\mathcal{A}} \le \delta_m(\sigma(\mathcal{A}))\|x - x_0\|_{\mathcal{A}}, \tag{2.2}$$

where $\sigma(\mathcal{A})$ is the spectrum of $\mathcal{A}$, and for any set $J \subset \mathbb{R}$,

$$\delta_m(J) = \inf_{p \in \mathcal{P}_{m,1}} \sup_{\lambda \in J} |p(\lambda)|. \tag{2.3}$$

Here

$$\mathcal{P}_{m,1} = \{\, p \,|\, p \text{ is a polynomial of degree } m,\ p(0) = 1 \,\}.$$

By using this characterization it can be established that the conjugate gradient method converges in the energy norm $|| \cdot ||_{\mathcal{A}}$, with a rate which can be bounded by the condition number $\kappa(\mathcal{A}) \equiv \|\mathcal{A}\|_{\mathcal{L}(X,X)} \cdot \|\mathcal{A}^{-1}\|_{\mathcal{L}(X,X)}$. In fact, by using an exact characterization of $\delta_m(J)$ when J is an interval in $\mathbb{R}^+$, it can be shown that

$$\delta_m(\sigma(\mathcal{A})) \le 2(\frac{\sqrt{\kappa(\mathcal{A})} - 1}{\sqrt{\kappa(\mathcal{A})} + 1})^m, \tag{2.4}$$

cf. for example [90, Theorem 38.5]. This leads to the following well–known convergence result for the conjugate gradient method.

Theorem 2.1. *Assume that $\mathcal{A} : X \to X$ is a symmetric and positive definite isomorphism. If the sequence $\{x_m\}$ is generated by the conjugate gradient method then*

$$\|x - x_m\|_{\mathcal{A}} \le 2\alpha^m \|x - x_0\|_{\mathcal{A}},$$

where $\alpha = (\sqrt{\kappa(\mathcal{A})} - 1)/(\sqrt{\kappa(\mathcal{A})} + 1)$.

A key observation is that the constant α is independent of the dimension of the space X. Therefore, as long as the operator $\mathcal{A}$ can be evaluated, the conjugate gradient method can also be used in the infinite dimensional case.

Remark 2.1. If the spectrum of $\mathcal{A}$ is approximately uniformly distributed in the interval $[1/||\mathcal{A}^{-1}||_{\mathcal{L}(X,X)}, ||\mathcal{A}||_{\mathcal{L}(X,X)}]$ then the upper bound given in Theorem 2.1 is indeed sharp. However, if the operator $\mathcal{A}$ has a few eigenvalues far away from the rest of the spectrum, then the estimate is not sharp. In fact, a few "bad eigenvalues" will have almost no effect on the asymptotic convergence of the method, cf. [9, 10]. □

If the operator $\mathcal{A}$ is indefinite, but still a symmetric isomorphism mapping X to itself, then we can replace the conjugate gradient by the *minimum residual method* [14], [33], [74], [78] [83], [84]. In this case, $x_m \in K_m(\mathcal{A}, f)$ is characterized by

$$x_m = \arg\min_{y \in K_m} \|\mathcal{A}y - f\|^2,$$

i.e., x_m minimizes the residual. Again, the effective computation of $\{x_m\}$ is based on a recurrence relation requiring only evaluations of the operator $\mathcal{A}$, and x_m is the best approximation of x in K_m in the sense that

$$\|\mathcal{A}(x - x_m)\| = \inf_{v \in K_m} \|\mathcal{A}(x - v)\|.$$

Furthermore, an estimate similar to (2.2) still holds. More precisely,

$$\|\mathcal{A}(x - x_m)\| \le \delta_m(\sigma(\mathcal{A}))\|\mathcal{A}(x - x_0)\|, \tag{2.5}$$

but, of course, in the indefinite case the spectrum of $\mathcal{A}$, $\sigma(\mathcal{A})$, is not contained in the positive half–line. However, the bound (2.4) can still be used to obtain a convergence estimate which only depends on $\kappa(\mathcal{A})$, and not on the dimension of the space X. The following analog of Theorem 2.1 is obtained.

Theorem 2.2. *Assume that $\mathcal{A} : X \to X$ is a symmetric isomorphism, and that the sequence $\{x_m\}$ is generated by the minimum residual method. Then there is a constant $\alpha \in (0, 1)$, depending only on $\kappa(\mathcal{A})$, such that*

$$\|\mathcal{A}(x - x_m)\| \le 2\alpha^m \|\mathcal{A}(x - x_0)\|.$$

Remark 2.2. Note that, in contrast to Theorem 2.1, Theorem 2.2 above includes no estimate of the convergence rate α. In fact, in this more general setting, where the operator $\mathcal{A}$ has both positive and negative eigenvalues, it is in general not straightforward to obtain a sharp estimate for the quantity $\delta_m(\sigma(\mathcal{A}))$ and hence for the constant α. However, a crude upper bound for $\delta_m(\sigma(\mathcal{A}))$ can be obtained by restricting the admissible polynomials in (2.3) to even polynomials. By using (2.4) this approach leads to the estimate

$$\|\mathcal{A}(x - x_{2m})\| \le 2\alpha^m \|\mathcal{A}(x - x_0)\|, \tag{2.6}$$

where $\alpha = (\kappa(\mathcal{A}) - 1)/(\kappa(\mathcal{A}) + 1)$. Sharper bounds for the convergence of the minimum residual method can be obtained by taking into account the relative location of the positive and negative parts of the spectrum of the operator $\mathcal{A}$, cf. for example the text [37, Chapter 6] or [78, 83]. However, such sharper bounds will not be utilized in the discussions below. □

Remark 2.3. An alternative approach in the indefinite case is to consider the corresponding normal system. Since $\mathcal{A}$ is symmetric, this system is given by

$$\mathcal{A}^2 x = \mathcal{A}f,$$

and the corresponding coefficient operator, $\mathcal{A}^2$, is positive definite. This system can then be solved by the conjugate gradient method. In fact, in some sense this approach is related to the crude upper bound for the minimum residual method given by (2.6), since this bound is sharp for the iteration based on the normal equation if the spectrum of $\mathcal{A}^2$ is uniformly distributed, cf. Remark 2.1. Therefore, for most problems the iteration based on the normal equations will give slower convergence than the minimum residual method. This effect has also been observed experimentally, cf. for example [2, Section 8]. □

Remark 2.4. For symmetric operators the condition number, $\kappa(\mathcal{A}) = \|\mathcal{A}\|_{\mathcal{L}(X,X)} \cdot \|\mathcal{A}^{-1}\|_{\mathcal{L}(X,X)}$, may be characterized as

$$\kappa(\mathcal{A}) = \frac{\sup_{\lambda \in \sigma(\mathcal{A})} |\lambda|}{\inf_{\lambda \in \sigma(\mathcal{A})} |\lambda|}.$$

In particular, the condition number is independent of the choice of norm on the space X. □

Example 2.1. *Integral equation.*
Consider a Fredholm equation of the second kind of the form:

$$\mathcal{A}u(x) := u(x) + \int_\Omega k(x,y)u(y)\,dy = f(x),$$

where we assume that the kernel k is continuous and symmetric, and that the operator $\mathcal{A} : X \to X$ is one–to–one, where $X = L^2(\Omega)$. Since the integral operator is compact, it follows by the Fredholm alternative theorem that $\mathcal{A}^{-1} \in \mathcal{L}(X,X)$. Hence, the equation can be solved by the minimum residual method, and also by the conjugate gradient method if $\mathcal{A}$ is positive definite. In fact, since the operator $\mathcal{A}$ has the form "identity + compact" the convergence is superlinear [97]. □

Example 2.2. *The Laplace operator.*
Variants of the conjugate gradient method are frequently used to solve systems which are obtained from discretizations of second order elliptic operators. Therefore, a natural question is how such methods can be applied to the continuous problems themselves. So let Ω be a bounded domain in $\mathbb{R}^n$, and let $H_0^1(\Omega) \subset L^2(\Omega)$ be the Sobolev space consisting of L^2 functions with a weak gradient in L^2 and vanishing trace on the boundary. Furthermore, let $X^* = H^{-1}(\Omega) \supset L^2(\Omega)$ be the corresponding representation of the dual space, such that the duality pairing is an extension of the inner product in $L^2(\Omega)$. Hence, we have

$$X = H_0^1(\Omega) \subset L^2(\Omega) \subset H^{-1}(\Omega) = X^*,$$

and, in a standard manner, we define the negative Laplace operator $\mathcal{A} : X \to X^*$ by

$$\langle \mathcal{A}u, v \rangle = \int_\Omega \operatorname{grad} u \cdot \operatorname{grad} v \, dx, \quad u, v \in X,$$

cf. Figure 1.

The standard weak formulation for the corresponding Dirichlet problem is now

$$\mathcal{A}u = f,$$

where the right–hand side $f \in X^*$ is given and the unknown $u \in X$. A Krylov space method of the form discussed above is *not well defined,* since the operator $\mathcal{A}$ may map functions in X out of the space, cf. Figure 1. □

The key tool in order to be able to apply Krylov space methods to problems with *unbounded spectrum* like the one in Example 2.2 is to introduce *a preconditioner.* In general, consider the system

$$\mathcal{A}x = f, \tag{2.7}$$

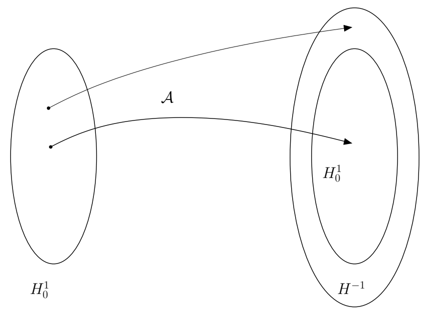

Figure 1. The mapping property of the operator $\mathcal{A}$.

where typically, $\mathcal{A}$ is an unbounded operator , or alternatively $\mathcal{A} \in \mathcal{L}(X, Y)$ with X strictly contained in the Hilbert space Y, i.e.,

$$X \xrightarrow{\mathcal{A}} Y \supset X$$

Remark 2.5. The assumption that $X \subset Y$ is typical, but not essential. The general approach outlined below applies to general isomorphism $\mathcal{A} \in \mathcal{L}(X, Y)$, where X and Y are separable Hilbert spaces. In particular, we do allow differential systems where the spectrum of the coefficient operator approaches both zero and infinity. $\Box$

In many problems arising from weak formulations of partial differential equations it is convenient to consider the space Y as a representation of the dual of X. Therefore, we will denote Y by X^*. Furthermore, in the rest of this section $\langle \cdot, \cdot \rangle$ is the duality pairing between X^* and X, while the notation $\langle \cdot, \cdot \rangle_X$ is used for the given inner product on X. We will assume that $\mathcal{A} \in \mathcal{L}(X, X^*)$ is symmetric in the sense that

$$\langle \mathcal{A}x, y \rangle = \langle \mathcal{A}y, x \rangle, \quad x, y \in X.$$

As we have argued above, when $X^* \neq X$ the Krylov space methods are in general not well defined. The preconditioner $\mathcal{B}$ that we will consider for the operator $\mathcal{A}$ is an isomorphism mapping X^* to X. Furthermore, we assume that the preconditioner $\mathcal{B}$ is symmetric and positive definite in the sense that $\langle \cdot, \mathcal{B} \cdot \rangle$ is an inner product on X^*. Hence, the preconditioner is a *Riesz operator* mapping X^* to X. As a consequence, $\langle \mathcal{B}^{-1} \cdot, \cdot \rangle$ is an inner product on X, with associated norm equivalent to $\| \cdot \|_X$. It is a direct consequence of these assumptions that the composition

$$\mathcal{B}\mathcal{A} : \quad X \xrightarrow{\mathcal{A}} X^* \xrightarrow{\mathcal{B}} X \tag{2.8}$$

is an isomorphism mapping X to itself. Furthermore, the operator $\mathcal{B}\mathcal{A} : X \to X$ is symmetric in the inner product $\langle \mathcal{B}^{-1} \cdot, \cdot \rangle$ on X. Therefore, the preconditioned system

$$\mathcal{B}\mathcal{A}x = \mathcal{B}f, \tag{2.9}$$

can be solved by a Krylov space method with a convergence rate bounded by $\kappa(\mathcal{B}\mathcal{A}) = ||\mathcal{B}\mathcal{A}||_{\mathcal{L}(X,X)} ||(\mathcal{B}\mathcal{A})^{-1}||_{\mathcal{L}(X,X)}$. Such an iteration is frequently referred to as a preconditioned Krylov space method for the system (2.7).

We should mention that even if the spaces X and X^* are fixed as sets, the preconditioner $\mathcal{B}$ is not unique, since the space X^* may have various inner products leading to the same topology. Different choices of preconditioners just correspond to different inner products on X^* of the form $\langle \cdot, \mathcal{B}\cdot \rangle$, which are equivalent in the sense of norms. Two such preconditioners $\mathcal{B}_1, \mathcal{B}_2 : X^* \to X$, which define norm equivalent inner products on X^*, are frequently referred to as *spectrally equivalent.* More precisely, $\mathcal{B}_1$ and $\mathcal{B}_2$ are spectrally equivalent if

$$c_0\langle f, \mathcal{B}_1 f\rangle \le \langle f, \mathcal{B}_2 f\rangle \le c_1\langle f, \mathcal{B}_1 f\rangle \quad f \in X^*,$$

for suitable positive constants $c_0, c_1 > 0$. In this case the operator $\mathcal{B}_2\mathcal{B}_1^{-1}$ is an isomorphism mapping X to itself, with $\sigma(\mathcal{B}_2\mathcal{B}_1^{-1}) \subset [c_0, c_1]$.

Remark 2.6. In the construction of preconditioners for discrete problems it is crucial that we are allowed to replace one preconditioner, or equivalently one inner product on X^*, by a spectrally equivalent operator. In other words, we utilize the fact that if $\kappa(\mathcal{B}_1\mathcal{A}) < \infty$, and $\mathcal{B}_1$ and B_2 are spectrally equivalent, then we can also conclude that $\kappa(\mathcal{B}_2\mathcal{A}) < \infty$. This is precisely the observation we can use to replace a computationally costly, and therefore impractical, preconditioner by an effective preconditioner. This will be explained more clearly in the discussions of discrete problems in Sections 6 and 7. □

Remark 2.7. Assume as above that $\mathcal{A} \in \mathcal{L}(X, X^*)$ is symmetric with respect to the duality pairing, and that $\mathcal{B} \in \mathcal{L}(X^*, X)$ is a corresponding preconditioner, such that $\langle \mathcal{B}^{-1}\cdot, \cdot\rangle$ defines an inner product on X. As observed above, the coefficient operator $\mathcal{B}\mathcal{A} \in \mathcal{L}(X, X)$ is symmetric in this inner product. However, note that this operator may not be symmetric in the original inner product $\langle \cdot, \cdot\rangle_X$ on X. Therefore, the Krylov space iteration should be defined with respect to the inner product $\langle \mathcal{B}^{-1}\cdot, \cdot\rangle$ on X. Furthermore, the error estimate derived from Theorem 2.2 is of the form

$$\langle \mathcal{B}\mathcal{A}(x - x_m), \mathcal{A}(x - x_m)\rangle^{1/2} \le 2\alpha^m \langle \mathcal{B}\mathcal{A}(x - x_0), \mathcal{A}(x - x_0)\rangle^{1/2} \qquad (2.10)$$

for this iteration. Here the constant $\alpha \in (0, 1)$ depends on $\kappa(\mathcal{B}\mathcal{A})$.

Assume in addition that the operator $\mathcal{A}$ is positive definite (or coercive) in the sense that there is a constant $\gamma > 0$ such that

$$\langle \mathcal{A}x, x\rangle \ge \gamma\|x\|_X^2, \quad x \in X.$$

Then $\mathcal{B}\mathcal{A}$ is a symmetric and positive definite operator with respect to the inner product $\langle \mathcal{B}^{-1}\cdot, \cdot\rangle$, and the corresponding convergence estimate for the preconditioned conjugate gradient method derived from Theorem 2.1 takes the form

$$\langle \mathcal{A}(x - x_m), (x - x_m)\rangle^{1/2} \le 2\alpha^m \langle \mathcal{A}(x - x_0), (x - x_0)\rangle^{1/2}, \qquad (2.11)$$

where $\alpha = (\sqrt{\kappa(\mathcal{B}\mathcal{A})} - 1)/(\sqrt{\kappa(\mathcal{B}\mathcal{A})} + 1)$. A nice property of this estimate is that the preconditioner $\mathcal{B}$ only enters through the convergence rate α. Furthermore, both the preconditioned minimum residual method and the preconditioned conjugate gradient method can be implemented such that there is no need to evaluate the operator $\mathcal{B}^{-1}$. Only evaluations of $\mathcal{B}$ are required.

Finally, we should note that when $\mathcal{A}$ is positive definite, the operator $\mathcal{A}$ itself defines an inner product on X. In other words, the bilinear form $\langle \mathcal{A}\cdot, \cdot\rangle$ is an inner product on X, and the coefficient operator $\mathcal{B}\mathcal{A}$ of the preconditioned system is symmetric and positive definite with respect to this inner product. Therefore, this inner product gives rise to an alternative variant of the preconditioned conjugate gradient method. The convergence estimate for this iteration, derived from Theorem 2.1, is exactly of the form (2.10), but with α given as in (2.11). □

Example 2.3. *The Laplace operator revisited.*
Above we saw that a Krylov space method is not well–defined for the Laplace operator in the continuous case. Let us therefore introduce a preconditioner $\mathcal{B} : H^{-1}(\Omega) \rightarrow H_0^1(\Omega)$. This preconditioner could for example be $(-\Delta)^{-1}$, cf. Figure 2.

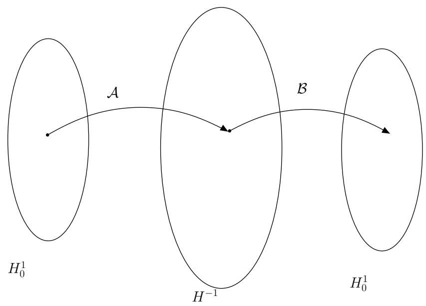

Figure 2. The mapping property of the composition of $\mathcal{A}$ and $\mathcal{B}$.

It is then obvious that $\mathcal{BA} = I$ is an isomorphism mapping $H_0^1(\Omega)$ onto itself. However, an immediate consequence is that all second order elliptic operators can be preconditioned this way. Consider for example the elliptic operator $\mathcal{A} : H_0^1(\Omega) \rightarrow H^{-1}(\Omega)$ defined by

$$\langle \mathcal{A}u, v\rangle = a(u, v) \quad u, v \in H_0^1(\Omega),$$

where the bilinear form $a : H_0^1(\Omega) \times H_0^1(\Omega) \rightarrow \mathbb{R}$ is of the form

$$a(u, v) = \int_\Omega (A\nabla u) \cdot \nabla v \, dx.$$

Here, we assume that $A = A(x) \in \mathbb{R}^{n\times n}$ is a symmetric and uniformly positive definite, i.e., there are positive constants c_0 and c_1 such that

$$c_0|\xi|^2 \le \xi^T A(x)\xi \le c_1|\xi|^2, \quad x \in \Omega,\ \xi \in \mathbb{R}^n.$$

If $\mathcal{B} = (-\Delta)^{-1} : H^{-1}(\Omega) \rightarrow H_0^1(\Omega)$ then the condition number of $\mathcal{BA} \in \mathcal{L}(H_0^1(\Omega), H_0^1(\Omega))$ is given by

$$\kappa(\mathcal{BA}) = ||\mathcal{BA}||_{\mathcal{L}(H_0^1,H_0^1)}||(\mathcal{BA})^{-1}||_{\mathcal{L}(H_0^1,H_0^1)} \le \frac{c_1}{c_0}.$$

Hence, the convergence rate of the corresponding Krylov space iteration can be bounded by the ratio c_1/c_0. We should also remark that the choice of preconditioner $\mathcal{B}$ for this example is not unique. Any symmetric and positive definite operator $\mathcal{B}$, with the proper mapping property, can be used. Different choices will only lead to different upper bounds for the condition numbers $\kappa(\mathcal{BA})$.

The fact that any second order elliptic operator can be preconditioned by a corresponding simpler operator has been utilized in a number of directions in the numerical analysis literature. For early pioneering papers, where this idea is applied to finite difference approximations, we refer for example to [35, 44]. □.

3. Saddle point problems

Our goal is to apply the theory for preconditioned Krylov space methods to examples of systems of partial differential equations. Most of the examples we will discuss can be characterized as saddle point systems. Therefore, we will first briefly recall parts of the Brezzi theory [27, 26] for abstract saddle point problems.

Consider the variational problem:

$$u = \arg\min_{v\in V} E(v) \quad \text{subject to } b(v,q) = G(q), \quad q \in Q, \tag{3.1}$$

where $E(v) = a(v,v) - 2F(v)$. Here we assume that

- V and Q are Hilbert spaces
- $F : V \to \mathbb{R}$ and $G : Q \to \mathbb{R}$ are bounded linear functionals
- $a : V \times V \to \mathbb{R}$ and $b : V \times Q \to \mathbb{R}$ are bilinear and bounded
- a is symmetric and positive semi definite

Associated the constrained minimization problem (3.1) we consider the saddle point system:

Find $(u,p) \in V \times Q$ such that

$$\begin{array}{lll} a(u,v) \quad +b(v,p) & = F(v), & v \in V, \\ b(u,q) & = G(q), & q \in Q. \end{array} \tag{3.2}$$

By introducing operators $A : V \to V^*$ and $B : V \to Q^*$ defined by

$$\langle Au, v\rangle = a(u,v) \quad \text{and } \langle Bu, q\rangle = b(u,q), \quad u,v \in V,\, q \in Q,$$

the system (3.2) can be rewritten in the form

$$\mathcal{A}\begin{pmatrix} u \\ p \end{pmatrix} = \begin{pmatrix} F \\ G \end{pmatrix}, \quad \text{with } \mathcal{A} = \begin{pmatrix} A & B^* \\ B & 0 \end{pmatrix}.$$

Here A and B are bounded operators, B^* is the adjoint of B, and V^* and Q^* are the dual spaces of V and Q, respectively. The two Brezzi conditions are necessary and sufficient conditions for ensuring that the coefficient operator $\mathcal{A} : V \times Q \to V^* \times Q^*$ is an isomorphism. These conditions can be stated as follows:

There are constants $\alpha, \beta > 0$ such that

$$a(v,v) \geq \alpha \|v\|_V^2, \quad v \in Z, \tag{3.3}$$

where $Z = \{v \in V \,|\, b(v,q) = 0, \quad q \in Q\,\}$, and

$$\inf_{q\in Q} \sup_{v\in V} \frac{b(v,q)}{\|v\|_V \|q\|_Q} \geq \beta. \tag{3.4}$$

If these conditions holds then we can conclude that the saddle point system (3.2) has a unique solution $(u,p) \in V \times Q$. Furthermore, under these conditions (3.2) and (3.1) are equivalent in the sense that if (u,p) solves (3.2), then u is the unique minimizer of (3.1). For more details we refer to [27, 26].

Next, we will consider preconditioners for saddle point problems. Recall from the general setup above, that if $\mathcal{A} : X \to X^*$ is an isomorphism, then the corresponding preconditioners $\mathcal{B}$

should be isomorphisms mapping X^* to X. Hence, in the setting of the abstract saddle point problem (3.2), where $X = V \times Q$ and $X^* = V^* \times Q^*$, the canonical choice is a Riesz operator mapping X^* to X of the form

$$\mathcal{B} = \begin{pmatrix} M & 0 \\ 0 & N \end{pmatrix},$$

where $M : V^* \to V$ and $N : Q^* \to Q$ are symmetric and positive definite isomorphisms. Therefore, *block diagonal preconditioners* are in some sense a natural choice for these problems, cf. [46, 78, 79, 83, 84].

Example 3.1. *Stokes problem.*
For a domain $\Omega \subset \mathbb{R}^n$ we consider the linear Stokes problem for incompressible flow given by

$$\begin{aligned} -\Delta u - \operatorname{grad} p &= f \quad \text{in } \Omega, \\ \operatorname{div} u &= g \quad \text{in } \Omega, \\ u &= 0 \quad \text{on } \partial\Omega. \end{aligned} \tag{3.5}$$

Furthermore, the corresponding weak formulation takes the form:
Find $(u, p) \in (H_0^1(\Omega))^n \times L_0^2(\Omega)$ such that

$$\begin{aligned} \langle \operatorname{grad} u, \operatorname{grad} v\rangle \quad + \langle p, \operatorname{div} v\rangle &= \langle f, v\rangle, & v &\in (H_0^1(\Omega))^n, \\ \langle \operatorname{div} u, q\rangle &= \langle g, q\rangle, & q &\in L_0^2(\Omega). \end{aligned}$$

Here L_0^2 is the space of L^2 functions with mean value zero. This is an example of a problem of the form (3.2). Furthermore, in this case the first condition (3.3) holds as a consequence of Poincaré's inequality. The second Brezzi condition, the inf–sup condition (3.4), can be expressed as

$$\inf_{q \in L_0^2} \sup_{v \in (H_0^1)^n} \frac{\langle q, \operatorname{div} v\rangle}{\|v\|_{H^1} \|q\|_{L^2}} \geq \beta > 0.$$

This estimate is for example established, under the condition that Ω is a Lipschitz domain in [69]. As a consequence of the theory outlined above, we therefore conclude that the coefficient operator

$$\mathcal{A} = \begin{pmatrix} -\Delta & -\operatorname{grad} \\ \operatorname{div} & 0 \end{pmatrix}$$

is an isomorphism mapping $(H_0^1(\Omega))^n \times L_0^2(\Omega)$ onto $(H^{-1}(\Omega))^n \times L_0^2(\Omega)$. The canonical choice of a preconditioner is therefore given as the block diagonal operator

$$\mathcal{B} = \begin{pmatrix} (-\Delta)^{-1} & 0 \\ 0 & I \end{pmatrix}$$

mapping the space $(H^{-1}(\Omega))^n \times L_0^2(\Omega)$ onto $(H_0^1(\Omega))^n \times L_0^2(\Omega)$. Here, the positive definite operator $(-\Delta)^{-1}$ can be replaced by any spectrally equivalent operator, corresponding to changing the inner product of the space $(H^{-1}(\Omega))^n$. □

Example 3.2. *Mixed formulation of the Poisson problem.*
Let Ω be a domain $\mathbb{R}^n$ and consider the system

$$\begin{aligned} u - \operatorname{grad} p &= f \quad \text{in } \Omega, \\ \operatorname{div} u &= g \quad \text{in } \Omega, \\ u \cdot n &= 0 \quad \text{on } \partial\Omega. \end{aligned} \tag{3.6}$$

By eliminating the vector variable u, we see that the scalar variable p must satisfy

$$-\Delta p = \operatorname{div} f - g \quad \text{in } \Omega, \quad \partial p/\partial n = -f \cdot n \quad \text{on } \partial\Omega. \tag{3.7}$$

The following weak formulation of the system (3.6) is again an example of a system of the form (3.2).

Find $(u, p) \in H_0(\operatorname{div}, \Omega) \times L_0^2(\Omega)$ such that

$$\begin{array}{llll} \langle u, v\rangle & +\langle p, \operatorname{div} v\rangle & = \langle f, v\rangle, & v \in H_0(\operatorname{div}, \Omega), \\ \langle \operatorname{div} u, q\rangle & & = \langle g, q\rangle, & q \in L_0^2(\Omega). \end{array}$$

Here, the space $H_0(\operatorname{div}, \Omega)$ consists of all vector fields in $(L^2(\Omega))^n$ with divergence in $L^2(\Omega)$. Furthermore, the zero subscript indicates that we restrict to vector fields with vanishing normal component on the boundary $\partial\Omega$. This is a Hilbert space with inner product given by

$$\langle u, v\rangle_{H_0(\operatorname{div},\Omega)} = \langle u, v\rangle + \langle \operatorname{div} u, \operatorname{div} v\rangle.$$

The two Brezzi conditions are easily verified in this case. In particular, the coercivity condition, (3.3), holds since $\|v\|_0 = \|v\|_{H_0(\operatorname{div})}$ for all divergence free vector fields, while the inf–sup condition (3.4) in the present case follows from the corresponding condition for the Stokes problem. As a consequence, we therefore can conclude that the coefficient operator

$$\mathcal{A} = \begin{pmatrix} I & -\operatorname{grad} \\ \operatorname{div} & 0 \end{pmatrix} : H_0(\operatorname{div}, \Omega) \times L_0^2(\Omega) \to H_0(\operatorname{div}, \Omega)^* \times L_0^2(\Omega),$$

is an isomorphism. Furthermore, for $n = 3$ the space $H_0(\operatorname{div}, \Omega)^*$ can be identified with

$$H^{-1}(\operatorname{curl}, \Omega) = \{ f \in H^{-1}(\Omega) \mid \operatorname{curl} f \in H^{-1}(\Omega) \}.$$

The preconditioner $\mathcal{B}$ should therefore be chosen as a block diagonal isomorphism mapping $H_0(\operatorname{div}, \Omega)^* \times L_0^2(\Omega)$ onto $H_0(\operatorname{div}, \Omega) \times L_0^2(\Omega)$, and the canonical choice is given by

$$\mathcal{B} = \mathcal{B}_1 = \begin{pmatrix} (I - \operatorname{grad}\operatorname{div})^{-1} & 0 \\ 0 & I \end{pmatrix}$$

Here the operator $I - \operatorname{grad}\operatorname{div}$ is the operator derived from the $H_0(\operatorname{div})$–inner product. We note that this operator is special in the sense that it acts as a second order elliptic operator on gradient–fields, but degenerates to the identity on curl–fields. This has the consequence that for the corresponding discrete operators, the construction of effective preconditioners for this operator is more delicate than for a standard second order elliptic operator. Of course, as before we are allowed to change inner products, i.e., the operator $(I - \operatorname{grad}\operatorname{div})^{-1}$ can be replaced by any spectrally equivalent operator. □

Example 3.3. *Alternative formulation of the Poisson equation.*
We consider again the system (3.6), but now with the alternative weak formulation:

Find $(u, p) \in (L^2(\Omega))^n \times H^1(\Omega) \cap L_0^2(\Omega)$ such that

$$\begin{array}{llll} \langle u, v\rangle & -\langle \operatorname{grad} p, v\rangle & = \langle f, v\rangle, & v \in (L^2(\Omega))^n, \\ -\langle u, \operatorname{grad} q\rangle & & = \langle g, q\rangle, & q \in H^1(\Omega) \cap L_0^2(\Omega). \end{array} \tag{3.8}$$

By eliminating the variable u we see that this formulation is indeed equivalent to the standard weak H^1 formulation of the boundary value problem (3.7), given by

$$\langle \operatorname{grad} p, \operatorname{grad} q\rangle = -\langle f, \operatorname{grad} q\rangle - \langle g, q\rangle, \quad q \in H^1(\Omega) \cap L_0^2(\Omega). \tag{3.9}$$

However, here we shall consider the system formulation (3.8). Hence, we formally consider the same system (3.6) as in the example above, and therefore with the same coefficient operator $\mathcal{A}$, but with a different weak interpretation. In the present formulation the operator $\mathcal{A}$ is an operator defined on the space $X = (L^2(\Omega))^n \times (H^1(\Omega) \cap L_0^2(\Omega))$. As above, the two Brezzi conditions are easily verified. In fact, in this case the coercivity condition (3.3) is obvious, while the inf–sup condition (3.4) is a consequence of Poincaré's inequality. The present choice of spaces leads to a canonical preconditioner $\mathcal{B} = \mathcal{B}_2 : X \to X^*$ of the form

$$\mathcal{B}_2 = \begin{pmatrix} I & 0 \\ 0 & (-\Delta)^{-1} \end{pmatrix}.$$

We therefore conclude that the operator

$$\mathcal{A} = \begin{pmatrix} I & -\operatorname{grad} \\ \operatorname{div} & 0 \end{pmatrix},$$

is well defined on two different spaces, either

$$X = H_0(\operatorname{div}, \Omega) \times L_0^2(\Omega) \quad \text{or} \quad X = (L^2(\Omega))^n \times (H^1(\Omega) \cap L_0^2(\Omega)),$$

and $\mathcal{A}$ maps these spaces isomorphically into their dual spaces, defined by extending the L^2 inner products into proper duality pairings. Furthermore, this leads to two corresponding preconditioners:

$$\mathcal{B}_1 = \begin{pmatrix} (I - \operatorname{grad}\operatorname{div})^{-1} & 0 \\ 0 & I \end{pmatrix} \quad \text{or} \quad \mathcal{B}_2 = \begin{pmatrix} I & 0 \\ 0 & (-\Delta)^{-1} \end{pmatrix},$$

where the operators $(I - \operatorname{grad}\operatorname{div})^{-1}$ and $(-\Delta)^{-1}$ can be replaced by spectrally equivalent operators. Note that the preconditioners $\mathcal{B}_1$ and $\mathcal{B}_2$ are not spectrally equivalent. In fact, they are quite different. □

Remark 3.1. Of course, the standard weak formulation of the Poisson problem, with the scalar variable p taken in H^1, is the formulation (3.9), where the vector field u has been eliminated. The system formulation (3.8) is just included above in order to illustrate that, in general, there is a possibility of non uniqueness in the choice of preconditioner. □

In order to better understand how one single operator can allow two completely different preconditioners, as in the example above, we will consider a more algebraic example.

Example 3.4. *One operator, two different preconditioners.*
Let $H = \ell_2(\mathbb{R}^2)$ be the Hilbert space consisting of sequences of vectors in $\mathbb{R}^2$ with Euclidean length in ℓ_2. Hence, $x \in H$ if $x = \{x_j\}_{j=1}^{\infty}$ with each $x_j \in \mathbb{R}^2$ and $\sum_j |x_j|^2 < \infty$. Define an unbounded block diagonal operator $\mathcal{A}$ on H by

$$\mathcal{A} = \begin{pmatrix} A_1 & & \\ & A_2 & \\ & & \ddots \end{pmatrix}$$

where each block A_j is a 2×2 matrix of the form

$$A_j = \begin{pmatrix} 1 & a_j \\ a_j & 0 \end{pmatrix}$$

Here a_j are positive real numbers such that $1 = a_1 \leq a_2 \leq \ldots$ and $\lim_{j\to\infty} a_j = \infty$. The eigenvalues of $\mathcal{A}$ are

$$\lambda_j = \frac{1 \pm \sqrt{1+4a_j^2}}{2} \longrightarrow \pm\infty \quad \text{as } a_j \to \infty.$$

We will consider block diagonal preconditioners for the operator $\mathcal{A}$ of the form $\mathcal{B} = \mathrm{diag}(B_j)$ where each block is of the form:

$$B_j = \begin{pmatrix} \beta_j & 0 \\ 0 & \gamma_j \end{pmatrix}, \quad \beta_j, \gamma_j > 0.$$

Hence, for each j we need to study 2×2 matrices of the form

$$BA = \begin{pmatrix} \beta & 0 \\ 0 & \gamma \end{pmatrix} \begin{pmatrix} 1 & a \\ a & 0 \end{pmatrix} = \begin{pmatrix} \beta & \beta a \\ \gamma a & 0 \end{pmatrix}$$

for $a \in [1, \infty)$. The operator $\mathcal{B}$ will be a preconditioner if $\mathcal{B}\mathcal{A}$ is a bounded operator on X, with a bounded inverse. Hence, we need to study the eigenvalues of the blocks BA for different choices of β and γ. One possibility is to choose $\beta = \frac{1}{1+a^2}$ and $\gamma = 1$. This gives

$$B = \begin{pmatrix} \frac{1}{1+a^2} & 0 \\ 0 & 1 \end{pmatrix} \quad \text{and } BA = \begin{pmatrix} \frac{1}{1+a^2} & \frac{a}{1+a^2} \\ a & 0 \end{pmatrix} \quad \text{with eigenvalues } \lambda(a) \to \pm 1 \quad \text{as } a \to \infty.$$

We can therefore conclude that the corresponding operator $\mathcal{B} = \mathcal{B}_1$ is a preconditioner.

An alternative, non-spectrally equivalent preconditioner, is obtained by taking $\beta = 1$ and $\gamma = \frac{1}{1+a^2}$, which gives

$$B = \begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{1+a^2} \end{pmatrix} \quad \text{and } BA = \begin{pmatrix} 1 & a \\ \frac{a}{1+a^2} & 0 \end{pmatrix}.$$

Hence, the eigenvalues satisfy

$$\lambda(a) = \frac{1 \pm \sqrt{1 + \frac{4a^2}{1+a^2}}}{2} \longrightarrow \frac{1 \pm \sqrt{5}}{2} \quad \text{as } a \to \infty,$$

and therefore this choice defines another preconditioner $\mathcal{B} = \mathcal{B}_2$. As above, we note that the two preconditioners are not spectrally equivalent. In fact, the operator $\mathcal{B}_2\mathcal{B}_1^{-1}$ is a block diagonal operator where each 2×2 block is of the form

$$\begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{1+a^2} \end{pmatrix} \begin{pmatrix} 1+a^2 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1+a^2 & 0 \\ 0 & \frac{1}{1+a^2} \end{pmatrix}$$

and therefore $\mathcal{B}_2\mathcal{B}_1^{-1}$ has eigenvalues which approach both zero and infinity as $a \to \infty$. □

Remark 3.2. The example above should be seen a warning. As we have discussed above, cf. Remark 2.6, a key tool in constructing practical preconditioners, based on the identification of the canonical ones, is to utilize the fact that if $\kappa(\mathcal{B}_1\mathcal{A}) < \infty$, and $\mathcal{B}_1$ and B_2 are spectrally equivalent, then we can also conclude that $\kappa(\mathcal{B}_2\mathcal{A}) < \infty$. On the other hand, it is *not true* that if both $\kappa(\mathcal{B}_1\mathcal{A}) < \infty$ and $\kappa(\mathcal{B}_2\mathcal{A}) < \infty$ then $\mathcal{B}_1$ and B_2 have to be spectrally equivalent. This is exactly what Example 3.4 shows. □

4. Parameter–dependent problems

Many systems of partial differential equations depend crucially on small (or large) parameters. In order to develop computational procedures which are robust with respect to these parameters it is natural to consider preconditioners which are uniformly well–behaved. Below we consider three model problems which depend on small parameters, namely the so–called time dependent Stokes problem, the Reissner–Mindlin plate model, and an optimal control problem.

For each of these examples the goal is to produce preconditioners $\mathcal{B}$ for the coefficient operator $\mathcal{A}$ which results in iterations which converge uniformly with respect to the parameters, i.e., the condition number $\kappa(\mathcal{B}\mathcal{A})$ should be uniformly bounded. A key tool for achieving this is to introduce proper parameter–dependent spaces and norms, such that the corresponding operator norms of the coefficient operator is bounded uniformly with respect to the parameters. We start this discussion by considering an elementary example.

Example 4.1. *Reaction–diffusion equation.*
Consider the boundary value problem

$$-\epsilon^2 \Delta u + u = f \quad \text{in } \Omega, \quad u|_{\partial\Omega} = 0, \tag{4.1}$$

where $\epsilon > 0$ is a small parameter. This equation is often referred to as a reaction diffusion equation. From energy estimates we see that a natural norm for the solution u is

$$\|u\|_{L^2 \cap \epsilon H_0^1} \equiv (\|u\|_0^2 + \epsilon^2 \|\operatorname{grad} u\|_0^2)^{1/2},$$

where here, and below, we use $\|\cdot\|_s$ to denote the norm in $H^s(\Omega)$. What is the correct norm for f, $\|f\|_?$, such that we get a sharp bound of the form

$$\|u\|_{L^2 \cap \epsilon H_0^1} \le c\|f\|_?,$$

and where the constant c is independent of ϵ? Note that we formally have

$$u = (I - \epsilon^2 \Delta)^{-1} f \quad \text{and that} \quad \|u\|^2_{L^2 \cap \epsilon H_0^1} = \langle (I - \epsilon^2 \Delta)u, u\rangle,$$

where $\langle \cdot, \cdot \rangle$ is the duality pairing between $H^{-1}(\Omega)$ and $H_0^1(\Omega)$. Therefore,

$$\begin{aligned}
\|u\|^2_{L^2 \cap \epsilon H_0^1} &= \langle (I - \epsilon^2 \Delta)u, u\rangle = \langle (I - \epsilon^2 \Delta)^{-1} f, f\rangle \\
&= \langle (I - \epsilon^2 \Delta)^{-1} f, (I - \epsilon^2 \Delta)^{-1} f\rangle + \epsilon^2 \langle (-\Delta)(I - \epsilon^2 \Delta)^{-1} f, (I - \epsilon^2 \Delta)^{-1} f\rangle \\
&= \|f_0\|_0^2 + \epsilon^{-2} \langle (-\Delta)^{-1} f_1, f_1\rangle = \|f_0\|_0^2 + \epsilon^{-2} \|f_1\|_{-1}^2,
\end{aligned}$$

where

$$f_0 = (I - \epsilon^2 \Delta)^{-1} f \quad \text{and} \quad f_1 = -\epsilon^2 \Delta (I - \epsilon^2 \Delta)^{-1} f.$$

Note that $f_0 + f_1 = f$. In fact,

$$\langle (I - \epsilon^2 \Delta)^{-1} f, f\rangle = \inf_{\substack{f = f_0 + f_1 \\ f_0 \in L^2,\, f_1 \in H^{-1}}} \|f_0\|_0^2 + \epsilon^{-2} \|f_1\|_{-1}^2.$$

Furthermore, if we define an ϵ–dependent norm on f by

$$\|f\|^2_{*\,\epsilon} \equiv \langle (I - \epsilon^2 \Delta)^{-1} f, f\rangle,$$

then

$$\|u\|_{L^2 \cap \epsilon H_0^1} = \|f\|_{*\,\epsilon}.$$

Hence, in this sense the elliptic operator $(I - \epsilon^2 \Delta)$ is norm preserving. □

The example above motivates us to introduce the notions of intersections and sums of Hilbert spaces, cf. [19, Chapter 2]. If X and Y are Hilbert spaces, then $X \cap Y$ and $X+Y$ are themselves Hilbert spaces with the norms

$$\|z\|_{X\cap Y} = (\|z\|_X^2 + \|z\|_Y^2)^{1/2}$$

and

$$\|z\|_{X+Y} = \inf_{\substack{z=x+y\\ x\in X,\, y\in Y}} (\|x\|_X^2 + \|y\|_Y^2)^{1/2}.$$

Furthermore, if $X \cap Y$ is dense in both X and Y then $(X \cap Y)^* = X^* + Y^*$ and $(X + Y)^* = X^* \cap Y^*$. Finally, if T is a bounded linear operator mapping X_1 to Y_1 and X_2 to Y_2, respectively, then

$$T \in \mathcal{L}(X_1 \cap X_2, Y_1 \cap Y_2) \cap \mathcal{L}(X_1 + X_2, Y_1 + Y_2).$$

In particular, we have the bounds

$$\|T\|_{\mathcal{L}(X_1\cap X_2,Y_1\cap Y_2)}, \|T\|_{\mathcal{L}(X_1+X_2,Y_1+Y_2)} \le \max(\|T\|_{\mathcal{L}(X_1,Y_1)}, \|T\|_{\mathcal{L}(X_2,Y_2)}). \tag{4.2}$$

We refer to [19, Chapter 2] for these results.

Assume that X and Y are real, separable Hilbert spaces with $Y \subset X$. Hence, by scaling the norms properly, we can assume that $\|y\|_X \le \|y\|_Y$ for all $y \in Y$. For each $\epsilon > 0$ consider the spaces $X \cap \epsilon \cdot Y$ given by

$$\|z\|_{X\cap\epsilon\cdot Y}^2 = \|z\|_X^2 + \epsilon^2 \|z\|_Y^2.$$

The space $X \cap \epsilon \cdot Y$ is equal to Y as a set. However, the norm approaches $\|\cdot\|_X$ as ϵ tends to zero. On the other hand consider the space $Y + \epsilon^{-1} \cdot X$ with norm given by

$$\|z\|_{Y+\epsilon^{-1}X}^2 = \inf_{\substack{z=x+y\\ x\in X,\, y\in Y}} \|y\|_Y^2 + \epsilon^{-2} \|x\|_X^2.$$

This space is equal to X as a set, but the corresponding norm approaches $\|\cdot\|_Y$ as ϵ tends to zero.

Let X_ϵ be the space $X_\epsilon = L^2(\Omega) \cap \epsilon \cdot H_0^1(\Omega)$. If the duality pairing is an extension of the L^2 inner product, then $X_\epsilon^* = L^2(\Omega) + \epsilon^{-1} \cdot H^{-1}(\Omega)$. Hence, as ϵ tends to zero, the norms on both X_ϵ and X_ϵ^* approach the L^2 norm. Furthermore, $\|f\|_{X_\epsilon^*}^2$ is equivalent to $\langle (I - \epsilon^2 \Delta)^{-1} f, f\rangle$. Hence, the spaces X_ϵ and X_ϵ^*, and their norms, are exactly the proper tools to obtain ϵ–independent estimates for the reaction–diffusion equation (4.1).

Next, we study three examples of saddle point problems which all depend on a small parameter.

Example 4.2. *The time dependent Stokes problem.*
Consider the linear time dependent Stokes problem given by:

$$\begin{aligned} u_t - \Delta u - \operatorname{grad} p &= f \quad \text{in } \Omega \times \mathbb{R}^+,\\ \operatorname{div} u &= g \quad \text{in } \Omega \times \mathbb{R}^+,\\ u &= 0 \quad \text{on } \partial\Omega \times \mathbb{R}^+,\\ u &= u_0 \quad \text{on } \Omega \times \{t = 0\}, \end{aligned}$$

Prepared using nlaauth.cls

for $\Omega \subset \mathbb{R}^n$. Various implicit time stepping schemes lead to boundary value problems of the form

$$\begin{aligned} (I - \epsilon^2 \Delta)u - \operatorname{grad} p &= f \quad \text{in } \Omega, \\ \operatorname{div} u &= g \quad \text{in } \Omega, \\ u &= 0 \quad \text{on } \partial\Omega, \end{aligned} \tag{4.3}$$

where $\epsilon \in (0, 1]$ is the square root of the time step. This is a linear saddle point system with coefficient operator

$$\mathcal{A}_\epsilon = \begin{pmatrix} I - \epsilon^2 \Delta & -\operatorname{grad} \\ \operatorname{div} & 0 \end{pmatrix},$$

where $\mathcal{A}_\epsilon : H_0^1(\Omega) \times L^2(\Omega) \to H^{-1}(\Omega) \times L^2(\Omega)$ is defined by

$$\langle \mathcal{A}_\epsilon(u, p), (v, q)\rangle = \langle u, v\rangle + \epsilon^2\langle \operatorname{grad} u, \operatorname{grad} v\rangle + \langle p, \operatorname{div} v\rangle + \langle q, \operatorname{div} u\rangle, \quad (v, q) \in H_0^1(\Omega) \times L^2(\Omega).$$

Note that here the symbol $\langle \cdot, \cdot\rangle$ is used to denote both duality pairings and L^2 inner products. We would like to define $\mathcal{A}_\epsilon$ on a proper space X_ϵ such that we obtain bounds on the norms of $\mathcal{A}_\epsilon$ and $\mathcal{A}_\epsilon^{-1}$, independent of ϵ. Recall that for $\epsilon = 0$ the operator

$$\mathcal{A}_0 = \begin{pmatrix} I & -\operatorname{grad} \\ \operatorname{div} & 0 \end{pmatrix}$$

is bounded from $H_0(\operatorname{div}, \Omega) \times L_0^2(\Omega)$ into $H_0(\operatorname{div}, \Omega)^* \times L_0^2(\Omega)$. Let $X_\epsilon = (H_0(\operatorname{div}, \Omega) \cap \epsilon \cdot H_0^1(\Omega)) \times L_0^2(\Omega)$, and $X_\epsilon^* = (H_0(\operatorname{div}, \Omega)^* + \epsilon^{-1} \cdot H^{-1}(\Omega)) \times L_0^2(\Omega)$. Note that the space X_ϵ is equal to $H_0^1(\Omega) \times L^2(\Omega)$ as a set, and therefore $\mathcal{A}_\epsilon$ is well–defined on X_ϵ. It is also easy to see that $\|\mathcal{A}_\epsilon\|_{\mathcal{L}(X_\epsilon, X_\epsilon^*)}$ is bounded independently of ϵ. To check that the same is true for the inverse, we have to verify the two Brezzi conditions, (3.3) and (3.4), with appropriate constants independent of ϵ.

However, the coercivity condition (3.3) holds with constant $\alpha = 1$, while the inf–sup condition (3.4) follows since

$$\sup_{v \in H_0^1} \frac{\langle q, \operatorname{div} v\rangle}{\|v\|_{H(\operatorname{div}) \cap \epsilon \cdot H^1}} \geq \sup_{v \in H_0^1} \frac{\langle q, \operatorname{div} v\rangle}{2\|v\|_{H^1}}.$$

We conclude that a uniform preconditioner therefore should be a positive definite mapping $\mathcal{B}_\epsilon$, mapping X_ϵ^* isomorphically onto X_ϵ. Hence, the canonical choice is an operator of the form

$$\mathcal{B}_\epsilon = \mathcal{B}_{1,\epsilon} = \begin{pmatrix} (I - \operatorname{grad}\operatorname{div} - \epsilon^2 \Delta)^{-1} & 0 \\ 0 & I \end{pmatrix},$$

where the operator $(I - \operatorname{grad}\operatorname{div} - \epsilon^2 \Delta)^{-1} : H_0(\operatorname{div}, \Omega)^* + \epsilon^{-1} \cdot H^{-1}(\Omega) \to H_0(\operatorname{div}, \Omega) \cap \epsilon \cdot H_0^1(\Omega)$ can be replaced by any operator which is spectrally equivalent to this operator, uniformly in ϵ.

Recall that the operator $\mathcal{A}_0$ is also bounded from $(L^2(\Omega))^n \times (H^1(\Omega) \cap L_0^2(\Omega))$ into its dual space. Hence, it seems that the operator $\mathcal{A}_\epsilon$ can also be defined on another space, where the velocity is allowed to be in $(L^2(\Omega) \cap \epsilon\, H_0^1(\Omega))^n$. Then the proper norm for the scalar variable should be

$$\sup_{v \in (H_0^1)^n} \frac{\langle q, \operatorname{div} v\rangle}{\|v\|_{L^2 \cap \epsilon\, H^1}} = \|\operatorname{grad} q\|_{L^2 + \epsilon^{-1} H^{-1}} \sim \|q\|_{H^1 + \epsilon^{-1} L^2},$$

where $\sim$ is used to indicate a possible equivalence of norms which needs to be checked. Note that for each fixed $\epsilon > 0$ the norm $\|\cdot\|_{H^1+\epsilon^{-1}L^2}$ is equivalent to the L^2 norm, but it approaches the H^1 norm as ϵ tends to zero. In agreement with this discussion, define the space

$$Y_\epsilon = (L^2(\Omega) \cap \epsilon \cdot H_0^1(\Omega))^n \times (H^1(\Omega) \cap L_0^2(\Omega)) + \epsilon^{-1} \cdot L_0^2(\Omega),$$

and let Y_ϵ^* be the corresponding representation of the dual space given by

$$Y_\epsilon^* = (L^2(\Omega) + \epsilon^{-1} \cdot H^{-1}(\Omega))^n \times (H^1(\Omega) \cap L_0^2(\Omega))^* \cap \epsilon \cdot L_0^2(\Omega).$$

The only real difficulty in establishing that $\mathcal{A}_\epsilon : Y_\epsilon \to Y_\epsilon^*$ is an isomorphism, with $\|\mathcal{A}_\epsilon\|_{\mathcal{L}(Y_\epsilon, Y_\epsilon^*)}$ and $\|\mathcal{A}_\epsilon^{-1}\|_{\mathcal{L}(Y_\epsilon^*, Y_\epsilon)}$ bounded uniformly in ϵ, is to establish the inf–sup condition

$$\sup_{v \in (H_0^1)^n} \frac{\langle q, \operatorname{div} v\rangle}{\|v\|_{L^2 \cap \epsilon H^1}} \geq \beta \|q\|_{H^1+\epsilon^{-1}L^2}, \quad q \in L_0^2,$$

for a suitable positive constant β, independent of ϵ. Alternatively, this can be written as a generalized Poincaré inequality of the form

$$\|q\|_{H^1+\epsilon^{-1}L_0^2} \leq \beta^{-1} \|\operatorname{grad} q\|_{L^2+\epsilon^{-1}H^{-1}}, \quad q \in L_0^2. \tag{4.4}$$

Consider the Stokes problem (3.5) with $g = 0$, and let $T \in \mathcal{L}((H^{-1}(\Omega))^n, L_0^2(\Omega))$ be the mapping given by $Tf = p$. Note that if grad : $(L^2(\Omega))^n \to (H^{-1}(\Omega))^n$ is the weakly defined gradient operator, then $T \circ \operatorname{grad} q = q$ for all $q \in L^2(\Omega)$.

If we assume that the Stokes problem admits H^2–regularity, which will hold if Ω is convex or has smooth boundary, cf. [34], then $T \in \mathcal{L}(L^2(\Omega), H^1(\Omega) \cap L_0^2(\Omega))$. Hence, using (4.2) we can conclude that

$$\|T\|_{\mathcal{L}(L^2+\epsilon^{-1}H^{-1}, (H^1 \cap L_0^2)+\epsilon^{-1}L_0^2)} \leq \beta^{-1}, \tag{4.5}$$

where $\beta^{-1} = \max(\|T\|_{\mathcal{L}(L^2, H^1 \cap L^2)}, \|T\|_{\mathcal{L}(H^{-1}, L_0^2)})$. Furthermore, we then have

$$\|q\|_{H^1+\epsilon^{-1}L^2} = \|T \circ \operatorname{grad} q\|_{H^1+\epsilon^{-1}L^2} \leq \beta^{-1} \|\operatorname{grad} q\|_{L^2+\epsilon^{-1}H^{-1}}$$

which is exactly the desired bound (4.4). For reasonable non-convex domains with non-smooth boundaries, the uniform inf–sup condition (4.4) still holds, but the proof has to be altered [67, 82].

From the mapping property of the operator $\mathcal{A}_\epsilon \in \mathcal{L}(Y_\epsilon, Y_\epsilon^*)$, we conclude that a uniform preconditioner for $\mathcal{A}_\epsilon$ is alternatively given by

$$\mathcal{B}_\epsilon = \mathcal{B}_{2,\epsilon} = \begin{pmatrix} (I - \epsilon^2 \Delta)^{-1} & 0 \\ 0 & (-\Delta)^{-1} + \epsilon^2 I \end{pmatrix}$$

Discrete preconditioners along these lines for the problem (4.3), and generalizations, have been suggested in [23, 29, 38, 36, 37, 58, 59, 66, 71, 91]. □

Example 4.3. *The Reissner-Mindlin plate model.*
This example is based on the presentation given in [2]. The study of elastic deformations of a clamped thin plate leads to the optimization problem

$$\min_{(\phi,u) \in (H_0^1)^2 \times H_0^1} E(\phi, u),$$

where the energy functional is given by

$$E(\phi, u) = \frac{1}{2}\int_\Omega \{(\mathcal{C}\mathcal{E}\phi) : (\mathcal{E}\phi) + t^{-2}|\phi - \operatorname{grad} u|^2\} dx - \int_\Omega gu\, dx.$$

Here the domain $\Omega \subset \mathbb{R}^2$, and $\mathcal{E}\phi$ denotes the symmetric part of $\operatorname{grad}\phi$. The parameter $t \in (0, 1]$ is the thickness of the plate, the function g represents a transverse load, while $\mathcal{C}$ is a uniformly bounded and positive definite operator mapping symmetric matrix fields into itself. The corresponding equilibrium system takes the form

$$\begin{aligned} -\operatorname{div}\mathcal{C}\mathcal{E}\phi + t^{-2}(\phi - \operatorname{grad} u) &= 0,\\ t^{-2}(-\Delta u + \operatorname{div}\phi) &= g,\\ \phi|_{\partial\Omega} = 0, \quad u|_{\partial\Omega} &= 0, \end{aligned}$$

or alternatively, with $\zeta = -t^{-2}(\phi - \operatorname{grad} u)$:

$$\begin{aligned} -\operatorname{div}\mathcal{C}\mathcal{E}\phi - \zeta &= 0,\\ -\operatorname{div}\zeta &= g,\\ -\phi + \operatorname{grad} u - t^2\zeta &= 0. \end{aligned}$$

This system can be given the following weak formulation:
Find $(\phi, u, \zeta) \in (H_0^1(\Omega))^2 \times H_0^1(\Omega) \times (L^2(\Omega))^2$ such that

$$\begin{aligned} \langle \mathcal{C}\mathcal{E}\phi, \mathcal{E}\psi\rangle - \langle \zeta, \psi - \operatorname{grad} v\rangle &= \langle g, v\rangle, \quad (\psi, v) \in (H_0^1(\Omega))^2 \times H_0^1(\Omega),\\ -\langle \phi - \operatorname{grad} u, \eta\rangle - t^2\langle \zeta, \eta\rangle &= 0, \quad \eta \in (L^2(\Omega))^2. \end{aligned}$$

This system can formally be written in the form

$$\mathcal{A}_t \begin{pmatrix} \phi \\ u \\ \zeta \end{pmatrix} = \begin{pmatrix} 0 \\ g \\ 0 \end{pmatrix},$$

where the coefficient operator, $\mathcal{A}_t$, is given by

$$\mathcal{A}_t = \begin{pmatrix} -\operatorname{div}\mathcal{C}\mathcal{E} & 0 & -I \\ 0 & 0 & -\operatorname{div} \\ -I & \operatorname{grad} & -t^2 I \end{pmatrix}.$$

The case $t = 0$ can be seen as an interpretation of the biharmonic problem, i.e., the limit case as the thickness of the plate tends to zero. Note that in this case the operator $\mathcal{A}_0$ is a saddle point operator of the form studied above, and hence the Brezzi theory can be applied. It is not hard to see, cf. [2], that

$$\sup_{(\psi,v)\in(H_0^1)^2\times H_0^1} \frac{\langle \eta, \psi - \operatorname{grad} v\rangle}{\|\psi\|_{H^1} + \|v\|_{H^1}} \sim \|\eta\|_{H^{-1}(\mathrm{div})}.$$

Here the symbol $\sim$ is used to indicate equivalence in norms, and the space $H^{-1}(\mathrm{div}, \Omega)$ is

$$H^{-1}(\mathrm{div}, \Omega) = \{\eta \in (H^{-1}(\Omega))^2 \mid \operatorname{div}\eta \in H^{-1}(\Omega)\},$$

with norm given by

$$\|\eta\|^2_{H^{-1}(\mathrm{div})} = \|\eta\|^2_{H^{-1}} + \|\operatorname{div}\eta\|^2_{H^{-1}}.$$

Therefore, the operator $\mathcal{A}_0$ can be seen to be a bounded operator from $X_0 = (H^1_0)^2 \times H^1_0(\Omega) \times H^{-1}(\mathrm{div},\Omega)$ into its L^2–dual given as $X_0^* = (H^{-1}(\Omega))^2 \times H^{-1}(\Omega) \times H_0(\mathrm{rot},\Omega)$, and with bounded inverse. Here rot denotes the differential operator mapping an $\mathbb{R}^2$ vector field η into the scalar field $\operatorname{rot}\eta$ given by $\operatorname{rot}\eta = \partial_y\eta_1 - \partial_x\eta_2$. Furthermore, the space $H_0(\mathrm{rot},\Omega)$ is given by

$$H_0(\mathrm{rot},\Omega) = \{\, \eta \in (L^2(\Omega))^2 \mid \operatorname{rot}\eta \in L^2(\Omega),\ P_T\eta = 0 \text{ on } \partial\Omega \,\},$$

where $P_T\eta$ denotes the tangential component of η on the boundary. Observe that in this case we do not have that $X_0 \subset X_0^*$, since

$$H_0(\mathrm{rot},\Omega) \subset (L^2(\Omega)^2 \subset H^{-1}(\mathrm{div},\Omega).$$

In fact, the coefficient operator $\mathcal{A}_0$ has a spectrum which approaches both zero and $\pm\infty$, cf. Remark 2.5.

The canonical operator mapping $H_0(\mathrm{rot},\Omega)$ onto $H^{-1}(\mathrm{div},\Omega)$ is *the differential operator*

$$D_0 = I + \operatorname{curl}\operatorname{rot},$$

where $\operatorname{curl} v = (-\partial_y v, \partial_x v)$. Therefore, the canonical preconditioner for this biharmonic system takes the form

$$\mathcal{B}_0 = \begin{pmatrix} (-\Delta)^{-1} & 0 & 0 \\ 0 & (-\Delta)^{-1} & 0 \\ 0 & 0 & D_0 \end{pmatrix}.$$

We note that the two first diagonal blocks are compact operators, corresponding to elliptic solution operators, while the third block is a differential operator.

For $t > 0$ the function spaces are modified such that

$$X_t = (H^1_0(\Omega))^2 \times H^1_0(\Omega) \times (H^{-1}(\mathrm{div},\Omega) \cap t \cdot (L^2(\Omega))^2)$$

and its L^2–dual

$$X_t^* = (H^{-1}(\Omega))^2 \times H^{-1}(\Omega) \times (H_0(\mathrm{rot},\Omega) + t^{-1} \cdot (L^2(\Omega))^2).$$

It can be checked that $\mathcal{A}_t \in \mathcal{L}(X_t, X_t^*)$ and $\mathcal{A}_t^{-1} \in \mathcal{L}(X_t^*, X_t)$ are both bounded uniformly in t for $t \in [0,1]$. Therefore, a corresponding uniform preconditioner takes the form

$$\mathcal{B}_t = \begin{pmatrix} (-\Delta)^{-1} & 0 & 0 \\ 0 & (-\Delta)^{-1} & 0 \\ 0 & 0 & D_t \end{pmatrix},$$

where the operator D_t is a suitable isomorphism mapping $H_0(\mathrm{rot},\Omega) + t^{-1} \cdot (L^2(\Omega))^2$ onto $H^{-1}(\mathrm{div},\Omega) \cap t \cdot (L^2(\Omega))^2$, and with proper operator norms of D_t and D_t^{-1} bounded independently of t. So for $t = 0$ the differential operator D_0 given above has the correct mapping property. For $t > 0$ the operator

$$D_t = I + (1 - t^2)\operatorname{curl}(I - t^2\Delta)^{-1}\operatorname{rot}$$

can be used. We note that for any $t > 0$, $D_t \in \mathcal{L}((L^2(\Omega))^2, (L^2(\Omega))^2)$, but with increasing operator norm as t approaches zero. However, the norms of D_t and D_t^{-1} are uniformly bounded in $\mathcal{L}(H_0(\mathrm{rot},\Omega) + t^{-1} \cdot (L^2(\Omega))^2, H^{-1}(\mathrm{div},\Omega) \cap t \cdot (L^2(\Omega))^2)$ and $\mathcal{L}(H^{-1}(\mathrm{div},\Omega) \cap t \cdot (L^2(\Omega))^2, H_0(\mathrm{rot},\Omega) + t^{-1} \cdot (L^2(\Omega))^2)$, respectively. We refer to [2] for more details. □

Example 4.4. *An optimal control problem.*
Let $\Omega \subset \mathbb{R}^n$ and for a given function $y_d \in L^2(\Omega)$ let

$$E_\epsilon(z,v) = \frac{1}{2}\int_\Omega (z-y_d)^2\,dx + \frac{\epsilon^2}{2}\int_\Omega v^2\,dx.$$

Consider the following optimal control problem:

$$(y,u) = \underset{(z,v)\in H^1\times L^2}{\arg\min}\; E_\epsilon(z,v) \tag{4.6}$$

subject to the elliptic constraint

$$\begin{aligned} z - \Delta z &= v, \\ \tfrac{\partial z}{\partial n}|_{\partial\Omega} &= 0. \end{aligned} \tag{4.7}$$

Here $\epsilon \in (0,1]$ is a regularization parameter. A weak formulation of the equilibrium system for this problem takes the form:
Find $(y,u,\lambda) \in H^1(\Omega)\times L^2(\Omega)\times H^1(\Omega)$ such that

$$\begin{aligned} \langle y,z\rangle + \langle \lambda,z\rangle + \langle \operatorname{grad}\lambda, \operatorname{grad} z\rangle &= \langle y_d, z\rangle, & z &\in H^1(\Omega), \\ \epsilon^2\langle u,v\rangle - \langle\lambda,v\rangle &= 0, & v &\in L^2(\Omega), \\ \langle y,\mu\rangle + \langle \operatorname{grad} y, \operatorname{grad}\mu\rangle - \langle u,\mu\rangle &= 0, & \mu &\in H^1(\Omega). \end{aligned} \tag{4.8}$$

Here the unknown $\lambda \in H^1(\Omega)$ corresponds to a Langrange multiplier.

Solution algorithms for variants of the system (4.8) have been studied by various authors, cf. for example [20, 68, 70, 80]. The discussion here follows closely the approach taken in [80]. The coefficient operator of the system (4.8) takes the form:

$$\mathcal{A}_\epsilon = \begin{pmatrix} I & 0 & I-\Delta \\ 0 & \epsilon^2 I & -I \\ I-\Delta & -I & 0 \end{pmatrix}. \tag{4.9}$$

Note that this operator is on saddle point form. Furthermore, for any fixed $\epsilon > 0$ it can easily be seen that $\mathcal{A}_\epsilon$ is a bounded operator from the space $H^1(\Omega)\times L^2(\Omega)\times H^1(\Omega)$ to the dual space $H^1(\Omega)^*\times L^2(\Omega)\times H^1(\Omega)^*$, where $H^1(\Omega)^*$ is the dual space of $H^1(\Omega)$ with respect to the L^2 inner product. However, the corresponding norm on $\mathcal{A}_\epsilon^{-1}$ will blow up as ϵ tends to zero. In order to avoid this blow up, we have to introduce ϵ–dependent spaces and norms. For each $\epsilon > 0$ we define the space X_ϵ by

$$X_\epsilon = (L^2(\Omega)\cap \epsilon^{1/2}\cdot H^1(\Omega)) \times \epsilon\cdot L^2(\Omega) \times (\epsilon^{-1}\cdot L^2(\Omega)\cap \epsilon^{-1/2}\cdot H^1(\Omega)).$$

It is easy to see that the bilinear forms defining the variational problem (4.8) are all uniformly bounded in these weighted norms. To show that the inverse, $\mathcal{A}_\epsilon^{-1}$ also admits the correct bounds, we need to show that the coercivity condition (3.3) and the inf–sup condition (3.4) holds uniformly in ϵ. To show condition (3.3) we need to establish that there is an $\alpha > 0$, independent of ϵ, such that

$$\|z\|^2 + \epsilon^2\,\|v\|^2 \ge \alpha(\|z\|^2 + \epsilon\,\|\operatorname{grad} z\|^2 + \epsilon^2\,\|v\|^2)$$

for all $(z,v) \in H^1(\Omega)\times L^2(\Omega)$ satisfying the constraint

$$\langle z,\mu\rangle + \langle \operatorname{grad} z, \operatorname{grad}\mu\rangle = \langle v,\mu\rangle \quad \mu \in H^1(\Omega).$$

Here and below $\|\cdot\|$ denotes L^2 norms. In fact, since the constraint implies that $\|\operatorname{grad} z\|^2 \leq \|v\|\|z\|$, we can easily conclude that the coercivity condition holds with $\alpha = 1/2$.

The inf–sup condition (3.4) in the present context reads

$$\sup_{(z,v)\in H^1\times L^2} \frac{\langle z,\mu\rangle + \langle \operatorname{grad} z, \operatorname{grad}\mu\rangle - \langle v,\mu\rangle}{(\|z\|^2 + \epsilon\,\|\operatorname{grad} z\|^2 + \epsilon^2\,\|v\|^2)^{1/2}} \geq \beta(\epsilon^{-2}\,\|\mu\|^2 + \epsilon^{-1}\,\|\operatorname{grad}\mu\|^2)^{1/2}$$

for all $\mu \in H^1(\Omega)$, where the positive constant β is required to be independent of ϵ. However, by choosing $v = -\,\epsilon^{-2}\,\mu$ and $z = \epsilon^{-1}\,\mu$ we obtain

$$\sup_{(z,v)\in H^1\times L^2} \frac{\langle z,\mu\rangle + \langle \operatorname{grad} z, \operatorname{grad}\mu\rangle - \langle v,\mu\rangle}{(\|z\|^2 + \epsilon\,\|\operatorname{grad} z\|^2 + \epsilon^2\,\|v\|^2)^{1/2}} \geq \frac{\epsilon^{-1}\,\|\mu\|^2 + \epsilon^{-1}\,\|\operatorname{grad}\mu\|^2 + \epsilon^{-2}\,\|\mu\|^2}{(\epsilon^{-2}\,\|\mu\|^2 + \epsilon^{-1}\,\|\operatorname{grad}\mu\|^2 + \epsilon^{-2}\,\|\mu\|^2)^{1/2}}$$

$$\geq (\frac{1}{\sqrt{2}})\frac{\epsilon^{-2}\,\|\mu\|^2 + \epsilon^{-1}\,\|\operatorname{grad}\mu\|^2}{(\epsilon^{-2}\,\|\mu\|^2 + \epsilon^{-1}\,\|\operatorname{grad}\mu\|^2)^{1/2}}.$$

So the uniform inf–sup condition holds with constant $\beta = 1/\sqrt{2}$.

Let X_ϵ^* be the associated representation of the dual space X_ϵ given by

$$X_\epsilon^* = (L^2(\Omega) + \epsilon^{-1/2}\cdot H^1(\Omega)^*) \times \epsilon^{-1}\cdot L^2(\Omega) \times \epsilon\cdot L^2(\Omega) + \epsilon^{1/2}\cdot H^1(\Omega)^*.$$

Then it follows from the bounds above that the operators $\mathcal{A}_\epsilon$ and $\mathcal{A}_\epsilon^{-1}$ are uniformly bounded in $\mathcal{L}(X_\epsilon, X_\epsilon^*)$ and $\mathcal{L}(X_\epsilon^*, X_\epsilon)$, respectively. Note that the canonical block diagonal operator mapping X_ϵ to X_ϵ^* is of the form

$$\begin{pmatrix} I-\epsilon\,\Delta & 0 & 0 \\ 0 & \epsilon^2\,I & 0 \\ 0 & 0 & \epsilon^{-2}(I-\epsilon\,\Delta) \end{pmatrix}.$$

Therefore, a uniform block diagonal preconditioner for the operator $\mathcal{A}_\epsilon$ takes the form

$$\mathcal{B}_\epsilon = \begin{pmatrix} (I-\epsilon\,\Delta)^{-1} & 0 & 0 \\ 0 & \epsilon^{-2}\,I & 0 \\ 0 & 0 & \epsilon^2(I-\epsilon\,\Delta)^{-1} \end{pmatrix}. \tag{4.10}$$

The analysis given in this example indicates that for the corresponding discrete problems proper uniform preconditioners should be composed from standard elliptic preconditioners and proper scalings. This hypothesis will be confirmed by our discussion below. □

Remark 4.1. In the above optimal control problem we managed to identify proper weighted Sobolev spaces such that the optimal control problem actually was an isomorphism. In general, it might not be easy to identify such spaces. Therefore, in [70], a different approach was taken. They demonstrated, by using weighted Sobolev spaces derived from abstract conditions, that the eigenvalues of the preconditioned system will be clustered in an interval close to zero and in two bounded intervals that do not approach zero. Moreover, the interval close to zero contains relatively few eigenvalues. Under such circumstances Krylov solvers such as the conjugate gradient method and the minimum residual method are known to be very efficient [9, 10, 70].

5. A general approach to preconditioning finite element systems

Up to this point we have only discussed how to precondition continuous systems of partial differential equations. In particular, we have focused on how to precondition various parameter–dependent problems uniformly with respect to the parameters. Most of the examples we have discussed are saddle point problems. Therefore, our main tool has been the Brezzi theory for saddle point problems outlined in Section 3, and the introduction of proper parameter–dependent spaces and norms discussed in the beginning of Section 4. Of course, our *real goal* is to study preconditioners for discrete versions of systems of partial differential equations. In this section we argue that:

- if we have identified the correct preconditioner for the continuous problem, and
- if we have applied a stable finite element discretization to this system,

$\Rightarrow$ then we know the basic structure of a preconditioner for the discrete problem.

This approach will lead to preconditioners such that the condition numbers of the preconditioned systems are bounded uniformly with respect to the discretization parameter. Furthermore, for parameter–dependent problems of the form studied in the previous section we obtain condition numbers for the preconditioned system which are bounded independently of both the discretization parameter and the model parameters. In order to achieve our goal we have found it convenient to utilize the general variational theory developed by Babuška [12, 13]. This theory is not restricted to saddle point problems.

As we have indicated above, the general procedure just outlined, based on the mapping properties of the coefficient operator, will only identify a canonical block diagonal preconditioner, i.e., a preconditioner which might be rather costly to evaluate and therefore not efficient. As a consequence, to construct a practical preconditioner we usually need to replace the different blocks of the canonical preconditioner by more cost efficient operators with equivalent mapping properties. Typically these operators will be constructed by domain decomposition methods, multigrid methods or similar techniques. These issues will be further discussed in Sections 6 and 7.

Let X be a real, separable Hilbert space and $a : X \times X \to \mathbb{R}$ is a bounded and symmetric, but not necessarily coercive, bilinear form satisfying an inf–sup condition of the form

$$\inf_{x\in X} \sup_{y\in X} \frac{a(x,y)}{\|x\|_X \|y\|_X} \geq \gamma > 0. \tag{5.1}$$

We let $C_a > 0$ be the bound on the bilinear form a, i.e.,

$$|a(x,y)| \leq C_a \|x\|_X \|y\|_X \quad x, y \in X. \tag{5.2}$$

For a given $f \in X^*$ consider the variational problem: Find $x \in X$ such that

$$a(x,y) = \langle f, y\rangle \quad y \in X, \tag{5.3}$$

or equivalently

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

where $\langle \cdot, \cdot \rangle$ is the duality pairing between X^* and X, and $\mathcal{A} : X \to X^*$ is given by

$$\langle \mathcal{A}x, y\rangle = a(x,y) \quad x, y \in X.$$

The following theorem is established in [12, 13].

Theorem 5.1. *The variational problem* (5.3) *has a unique solution. Furthermore, the coefficient operator* $\mathcal{A}$ *satisfies the bounds*

$$\|\mathcal{A}\|_{\mathcal{L}(X,X^*)} \le C_a, \quad \textit{and} \quad \|\mathcal{A}^{-1}\|_{\mathcal{L}(X^*,X)} \le \gamma^{-1}.$$

Remark 5.1. The theory developed in [12, 13] is more general than indicated above, since no symmetry of the bilinear form a is required. In fact, even different test and trial spaces are allowed. However, in the present paper we will only use this result in the symmetric setting. □

Example 5.1. *The mixed Poisson problem.*
Consider the mixed weak formulation of the Poisson problem:
Find $(u,p) \in H_0(\mathrm{div},\Omega) \times L_0^2(\Omega)$ such that

$$\begin{aligned} \langle u, v\rangle + \langle p, \operatorname{div} v\rangle &= \langle f, v\rangle, \qquad v \in H_0(\mathrm{div},\Omega), \\ \langle \operatorname{div} u, q\rangle &= \langle g, q\rangle, \qquad q \in L_0^2(\Omega). \end{aligned} \tag{5.4}$$

We have seen in Section 3 that existence and uniqueness of solutions of this problem follow from the Brezzi theory. However, we could also have used Theorem 5.1 above. To see this, let $X = H_0(\mathrm{div},\Omega) \times L_0^2(\Omega)$. This is a Hilbert space with inner product

$$\langle x, y\rangle_X = \langle u, v\rangle + \langle \operatorname{div} u, \operatorname{div} v\rangle + \langle p, q\rangle$$

for $x = (u,p)$ and $y = (v,q)$. Define a bilinear form on X by

$$a(x,y) = \langle u, v\rangle + \langle p, \operatorname{div} v\rangle + \langle \operatorname{div} u, q\rangle.$$

Then the problem (5.4) corresponds to the variational problem (5.3). Furthermore, it is straightforward to check that the bilinear form a is bounded on X. Therefore, it only remains to check that the general inf–sup condition (5.1) holds for the bilinear form a. However, since we have already seen that the Brezzi conditions (3.3) and (3.4) hold, this will be a consequence of Theorem 5.2 below. □

In fact, all saddle point problems studied in Section 3 can be seen as examples of problems of the form (5.3) with an associated bilinear form a satisfying the conditions (5.1) and (5.2). To see this, consider an abstract saddle point problem of the form studied in Section 3:
Find $(u,p) \in V \times Q$ such that

$$\begin{aligned} a_0(u,v) \quad + b(v,p) &= F(v), \qquad v \in V, \\ b(u,q) \qquad\qquad &= G(q), \qquad q \in Q, \end{aligned} \tag{5.5}$$

where we have changed notation slightly by replacing the bilinear form a in (3.2) by a_0. Let us assume that the bilinear forms a_0 and b are both bounded and satisfy (3.3) and (3.4). Let X be the Hilbert space $X = V \times Q$ with inner product

$$\langle x, y\rangle_X = \langle u, v\rangle_V + \langle p, q\rangle_Q,$$

and define the bilinear form $a : X \times X \to \mathbb{R}$ by

$$a(x,y) = a_0(u,v) + b(v,p) + b(u,q). \tag{5.6}$$

Here $x = (u,p)$ and $y = (v,q)$. We then have the following theorem.

Theorem 5.2. *Assume that the bilinear forms* a_0 *and* b *are both bounded and satisfy* (3.3) *and* (3.4). *Then the corresponding bilinear form* a*, defined by* (5.6) *on* $X = V \times Q$*, is bounded and satisfies* (5.1).

This result, which basically follows from the fundamental theory of [12, 13, 27], is well–known, see for example the discussion in the introduction of [39]. However, as a service to the reader, we include a sketch of a proof in the present setting.

Sketch of a proof of Theorem 5.2. It is straightforward to check that the boundedness of a_0 and b, implies that also the bilinear form a is bounded on X. Hence, it only remains to establish the inf–sup condition (5.1). As a first step in this direction we establish that $(a_0(v,v) + \|Bv\|_{Q^*}^2)^{1/2}$ is a norm, equivalent to $\|\cdot\|_V$, on V. Here, the operator $B \in \mathcal{L}(V, Q^*)$ is defined by $\langle Bv, q\rangle = b(v,q)$ for $v \in V$ and $q \in Q$. Therefore, (5.1) will follow if we can show that there are constants $c_1, c_2 > 0$ such that for each $x = (u,p) \in X$, we can find $y = (v,q) \in X$ such that

$$\|y\|_X \le c_1\|x\|_X \quad \text{and} \quad a(x,y) \ge c_2(a_0(u,u) + \|Bu\|_{Q^*}^2 + \|p\|_Q^2). \tag{5.7}$$

For $x = (u,p) \in X$ given and for any $t > 0$, define $y = (v,q) \in X$ by

$$v = u + tw, \quad q = -p + tR_QBu,$$

where $w \in Z^\perp$ satisfies $R_QBw = p$. Here $R_Q : Q^* \to Q$ is the corresponding Riesz operator. It can now be checked that the positive parameter t can be chosen such that (5.7) holds, and this will complete the proof. □

Let us return to the variational problem (5.3). Define $\mathcal{B} : X^* \to X$ as the Riesz operator

$$\langle \mathcal{B}f, y\rangle_X = \langle f, y\rangle \quad y \in X.$$

Note that

$$\langle \mathcal{B}\mathcal{A}x, y\rangle_X = \langle \mathcal{A}x, y\rangle = a(x,y) = \langle x, \mathcal{B}\mathcal{A}y\rangle_X,$$

so $\mathcal{B}\mathcal{A}$ is a symmetric operator mapping X to itself as a consequence of the symmetry of the bilinear form a. Furthermore, from (5.2) we obtain

$$\|\mathcal{B}\mathcal{A}\|_{\mathcal{L}(X,X)} = \sup_{x\in X} \frac{\langle \mathcal{B}\mathcal{A}x, y\rangle_X}{\|x\|_X\|y\|_X} = \sup_{x\in X} \frac{|a(x,y)|}{\|x\|_X\|y\|_X} \le C_a$$

and from (5.1)

$$\|(\mathcal{B}\mathcal{A})^{-1}\|_{\mathcal{L}(X,X)}^{-1} = \inf_{x\in X} \frac{\|\mathcal{B}\mathcal{A}x\|_X}{\|x\|_X} = \inf_{x\in X}\sup_{y\in X} \frac{\langle \mathcal{B}\mathcal{A}x, y\rangle_X}{\|x\|_X\|y\|_X} = \inf_{x\in X}\sup_{y\in X} \frac{a(x,y)}{\|x\|_X\|y\|_X} \ge \gamma > 0.$$

In other words, $\mathcal{B}$ is a preconditioner for $\mathcal{A}$ and the condition number of the preconditioned operator $\mathcal{B}\mathcal{A}$ satisfies $\kappa(\mathcal{B}\mathcal{A}) \le C_a/\gamma$.

Next, we consider the discrete versions of (5.3) obtained from finite element discretizations. So let $\{X_h\}$, with each $X_h \subset X$, be a family of finite element spaces indexed by the discretization parameter h, and consider the corresponding discrete variational problems: Find $x_h \in X_h$ such that

$$a(x_h, y) = \langle f, y\rangle \quad y \in X_h \quad \text{or equivalently} \quad \mathcal{A}_hx_h = f_h, \tag{5.8}$$

where $\langle\cdot,\cdot\rangle$ denotes the duality pairing between X_h^* and X_h, and $\mathcal{A}_h : X_h \to X_h^*$ is given by

$$\langle \mathcal{A}_hx, y\rangle = a(x,y) \quad x, y \in X_h.$$

Since, a is not positive definite it is not clear that this discretization is stable, in fact the system can even be singular. The stable discretizations are characterized by a discrete inf–sup condition corresponding to (5.1). Hence, we assume that there is a constant γ_1, independent of h, such that

$$\inf_{x \in X_h} \sup_{y \in X_h} \frac{a(x,y)}{\|x\|_X \|y\|_X} \geq \gamma_1 > 0. \tag{5.9}$$

This condition does not follow from the corresponding condition for the continuous case. However, for saddle problems, where the bilinear form is given of the form (5.6), then it follows from Theorem 5.2 that the bound (5.9) will follow from the standard discrete versions of the Brezzi conditions (3.3) and (3.4).

As in the continuous case we define the preconditioner $\mathcal{B}_h : X_h^* \to X_h$ by

$$\langle \mathcal{B}_h f, y \rangle_X = \langle f, y \rangle \quad y \in X_h.$$

The same arguments as in the continuous case show that $\mathcal{B}_h \mathcal{A}_h : X_h \to X_h$ is symmetric (with respect to $\langle \cdot, \cdot \rangle_X$) and that

$$\|\mathcal{B}_h \mathcal{A}_h\|_{\mathcal{L}(X_h, X_h)} \leq C_a, \quad \text{and} \quad \|(\mathcal{B}_h \mathcal{A}_h)^{-1}\|_{\mathcal{L}(X_h, X_h)} \leq \gamma_1^{-1},$$

such that $\kappa(\mathcal{B}_h \mathcal{A}_h) \leq C_a / \gamma_1$. So we have confirmed the claim that for stable discretizations of the variational problem (5.3), *the structure of the preconditioner for the discrete problems* follows from the structure of the preconditioner in the continuous case. Furthermore, the inner product, $\langle \cdot, \cdot \rangle_X$ on X_h is only determined up to equivalence of norms, uniformly in h. As we shall see below, this observation can be utilized to introduce various preconditioners for the discrete finite element systems.

Remark 5.2. The bound on the condition number, $\kappa(\mathcal{B}_h \mathcal{A}_h) \leq C_a / \gamma_1$, obtained above, only depends on the upper bound (5.2) restricted to the subspace X_h, and the discrete inf–sup condition (5.9). In fact, this observation also applies to *mesh dependent bilinear forms* and *mesh dependent norms.* Assume that we consider the discrete problem (5.8), where the bilinear form $a = a_h$ depends on h, and that the bilinear form a_h satisfies an upper bound of the form

$$|a_h(x,y)| \leq C_a \|x\|_{X_h} \|y\|_{X_h}$$

for a suitable mesh dependent norm $\| \cdot \|_{X_h}$ on X_h, derived from an inner product $\langle \cdot, \cdot \rangle_{X_h}$. Furthermore, assume that the corresponding discrete inf–sup condition

$$\inf_{x \in X_h} \sup_{y \in X_h} \frac{a_h(x,y)}{\|x\|_{X_h} \|y\|_{X_h}} \geq \gamma_1 > 0$$

holds. If we define a preconditioner $\mathcal{B}_h : X_h^* \to X_h$ by

$$\langle \mathcal{B}_h f, y \rangle_{X_h} = \langle f, y \rangle, \quad y \in X_h,$$

then the same arguments as above shows that $\kappa(\mathcal{B}_h \mathcal{A}_h) \leq C_a / \gamma_1$. Therefore, if the upper bound C_a, and the inf–sup constant γ_1 can be taken to be independent of h, then we can conclude that $\kappa(\mathcal{B}_h \mathcal{A}_h)$ is bounded independently of h. Note, that since we do allow mesh dependent bilinear forms and mesh dependent norms, this approach applies, in particular, to nonconforming finite element methods, i.e., to the cases when X_h is not a subspace of X. □

6. The building blocks of the discrete preconditioners

We have seen above that the individual blocks of the canonical preconditioners for linear systems arising from discretizations of partial differential equations are naturally associated with the proper function spaces for a well–posed weak formulation of the problem. The function spaces $H^1(\Omega)$, $H(\mathrm{curl}, \Omega)$, and $H(\mathrm{div}, \Omega)$ appears frequently in weak formulations of various systems of partial differential equations. As a consequence, the canonical preconditioners for these systems will be block diagonal operators, where the corresponding blocks are exact inverses of the operators $I - \Delta$, $I + \mathrm{curl}\,\mathrm{curl}$, and $I - \mathrm{grad}\,\mathrm{div}$, derived from the inner products

$$\langle u, v\rangle + \langle \mathrm{grad}\, u, \mathrm{grad}\, v\rangle, \quad \langle u, v\rangle + \langle \mathrm{curl}\, u, \mathrm{curl}\, v\rangle, \quad \langle u, v\rangle + \langle \mathrm{div}\, u, \mathrm{div}\, v\rangle.$$

More precisely, for proper stable finite element discretization of these systems we need *efficient* preconditioners for the discrete operators corresponding to $I - \Delta$, $I + \mathrm{curl}\,\mathrm{curl}$, and $I - \mathrm{grad}\,\mathrm{div}$, since such preconditioners will be the building blocks for the preconditioners of the complete systems. Furthermore, in many parameter–dependent problems, weighted inner products of the form

$$\langle u, v\rangle + \epsilon^2\langle \mathrm{grad}\, u, \mathrm{grad}\, v\rangle, \quad \langle u, v\rangle + \epsilon^2\langle \mathrm{curl}\, u, \mathrm{curl}\, v\rangle, \quad \langle u, v\rangle + \epsilon^2\langle \mathrm{div}\, u, \mathrm{div}\, v\rangle.$$

may occur, where $\epsilon \geq 0$ is a small parameter. Therefore, we need preconditioners for the discretizations of the operators $I - \epsilon^2\,\Delta$, $I + \epsilon^2\,\mathrm{curl}\,\mathrm{curl}$, and $I - \epsilon^2\,\mathrm{grad}\,\mathrm{div}$ which are well behaved uniformly in both the discretization parameter and the parameter ϵ.

Of course, there is a large collection of literature on how to construct efficient preconditioners for discrete versions of these operators, and we only mention a few here. Multigrid and domain decomposition preconditioners for elliptic problems, like $I - \Delta$, have been studied by many authors, consider for example [24, 28, 45, 85] and references given there. Many of these algorithms, like the multigrid V–cycle algorithm, will in fact adapt naturally to reaction–diffusion operators of the form $I - \epsilon^2\,\Delta$, while special attention to the parameter dependency in these constructions is given in [21, 30, 31, 72]. In contrast to this, the operators $I + \mathrm{curl}\,\mathrm{curl}$, and $I - \mathrm{grad}\,\mathrm{div}$ are not elliptic operators. As a consequence, some extra care has to be taken to construct preconditioners for these operators. For example, the most straightforward generalizations of the elliptic multigrid algorithms may not work, cf. [32]. However, one remedy to compensate for the lack of ellipticity in these algorithms is to replace the simplest diagonal smoothers by suitable block diagonal smoothers, cf. for example [3, 4, 5]. Another approach is to use a proper Helmholtz decomposition to construct appropriate smoothers, cf. [50, 51, 94, 95], and also [6, Section 10], or to use the Helmholtz decomposition more indirectly as in [49]. In particular, the preconditioners proposed in [5] are shown to be robust with respect to ϵ when applied to the operators $I + \epsilon^2\,\mathrm{curl}\,\mathrm{curl}$, and $I - \epsilon^2\,\mathrm{grad}\,\mathrm{div}$.

We will not carry out a more detailed discussion of the preconditioners associated with the spaces $H^1(\Omega)$, $H(\mathrm{curl}, \Omega)$, and $H(\mathrm{div}, \Omega)$ here. Instead, we briefly discuss the representation of the discrete differential operators, and their preconditioners, as matrices. In order to make this discussion as simple as possible we restrict ourselves to the finite element representation of the operator $I - \Delta$ and its preconditioners, but our conclusions also hold for other operators like for example $I + \mathrm{curl}\,\mathrm{curl}$, and $I - \mathrm{grad}\,\mathrm{div}$.

So let $a = a(u, v)$ be the bilinear form

$$a(u, v) = \langle u, v\rangle + \langle \mathrm{grad}\, u, \mathrm{grad}\, v\rangle = \int_\Omega uv + \mathrm{grad}\, u \cdot \mathrm{grad}\, v\, dx$$

defined on the space $H^1(\Omega) \times H^1(\Omega)$. We consider a finite element discretization of the problem

$$-\Delta u + u = f, \quad \text{in } \Omega, \quad \partial_n u = 0, \quad \text{on } \partial\Omega,$$

where ∂_n denotes the normal derivative on the boundary $\partial\Omega$. Hence, for a given finite element space $V_h \subset H^1(\Omega)$, the finite element approximation $u_h \in V_h$ solves the discrete system

$$a(u_h, v) = \langle f, v \rangle, \quad v \in V_h.$$

Frequently, in the finite element literature, this problem is written as $\mathcal{A}_h u_h = f_h$, where the operator $\mathcal{A}_h : V_h \to V_h^*$ is defined by

$$\langle \mathcal{A}_h u, v \rangle = a(u, v), \quad u, v \in V_h.$$

Here, as above, $\langle \cdot, \cdot \rangle$ denotes the duality pairing between V_h^* and V_h.

Remark 6.1. It is probably more common in the finite element literature to identify V_h and its dual space. Hence, $\mathcal{A}_h$ is an operator mapping V_h into itself, and $\langle \cdot, \cdot \rangle$ should be interpreted as the inner product on $L^2(\Omega)$. However, for the discussion below it seems favorable to distinguish between V_h and its dual space V_h^*. □

The operator $\mathcal{A}_h \in \mathcal{L}(V_h, V_h^*)$ introduced above depends on the finite element space V_h, but is independent of any basis in this space. Another representation of the discrete differential operator, which is needed for computations, is the corresponding stiffness matrix,

$$\mathbb{A}_h : \mathbb{R}^n \to \mathbb{R}^n, \quad \text{given by} \quad (\mathbb{A}_h)_{i,j} = a(\phi_j, \phi_i),$$

where $\{\phi_j\}_{j=1}^n$ is a basis for the finite element space V_h. This matrix obviously depends on the choice of basis for the space V_h. Furthermore, since the basis functions ϕ_j usually have local support, the stiffness matrix will be sparse.

It is instructive to clarify the relation between the stiffness matrix $\mathbb{A}_h$ and the operator $\mathcal{A}_h$. We define two "representation operators," $\pi_h : V_h \to \mathbb{R}^n$ and $\mu_h : V_h^* \to \mathbb{R}^n$, by

$$v = \sum_j (\pi_h v)_j \phi_j, \quad (\mu_h f)_j = \langle f, \phi_j \rangle, \quad v \in V_h, f \in V_h^*.$$

We refer to the vectors $\pi_h v$ and $\mu_h v$ as the *primal* and *dual representations.* Note that

$$(\pi_h u) \cdot (\mu_h v) = \sum_j (\pi_h u)_j \langle v, \phi_j \rangle = \langle u, v \rangle$$

so π_h^{-1} is the adjoint of μ_h and μ_h^{-1} is the adjoint of π_h, i.e., $\pi_h^{-1} = \mu_h^*$ and $\mu_h^{-1} = \pi_h^*$. Furthermore, for any $v \in V_h$ we have

$$(\mu_h(\mathcal{A}_h v))_i = \langle \mathcal{A}_h v, \phi_i \rangle = a(v, \phi_i) = \sum_j (\pi_h v)_j a(\phi_j, \phi_i) = (\mathbb{A}_h \pi_h v)_i.$$

Hence, $\mu_h \mathcal{A}_h = \mathbb{A}_h \pi_h$, i.e., the diagram

$$\begin{array}{ccc} V_h & \xrightarrow{\mathcal{A}_h} & V_h^* \\ \Big\downarrow \pi_h & & \Big\downarrow \mu_h \\ \mathbb{R}^n & \xrightarrow{\mathbb{A}_h} & \mathbb{R}^n \end{array}$$

commutes. Alternatively, the stiffness matrix is given by

$$\mathbb{A}_h = \mu_h \mathcal{A}_h \pi_h^{-1} = \mu_h \mathcal{A}_h \mu_h^*.$$

We should observe here that if $I_h : V_h \to V_h^*$ is the Riesz operator given by

$$\langle I_h v, w \rangle = \langle v, w \rangle_{L^2}, \quad v, w \in V_h,$$

then the corresponding stiffness matrix, usually referred to as *the mass matrix*, is

$$\mu_h I_h \pi_h^{-1} \neq \text{the identity matrix.}$$

In fact, $(\mu_h I_h \pi_h^{-1})_{i,j} = \langle \phi_j, \phi_i \rangle_{L^2}$. In particular, its inverse $\pi_h \mu_h^{-1}$, representing the operator $I_h^{-1} : V_h^* \to V_h$, is in general not sparse.

We remark that the square of the stiffness matrix is given by

$$\mathbb{A}_h^2 = \mu_h \mathcal{A}_h \pi_h^{-1} \mu_h \mathcal{A}_h \pi_h^{-1} \neq \mu_h \mathcal{A}_h^2 \pi_h^{-1}.$$

Hence, $\mathbb{A}_h^2$ is not a sparse representation of $\mathcal{A}_h^2$. Instead the matrix

$$\mathbb{A}_h \pi_h \mu_h^{-1} \mathbb{A}_h = \mu_h \mathcal{A}_h \pi_h^{-1} (\pi_h \mu_h^{-1}) \mu_h \mathcal{A}_h \pi_h^{-1} = \mu_h \mathcal{A}_h^2 \pi_h^{-1}$$

represents $\mathcal{A}_h^2$. This observation is relevant for Krylov space methods, since it shows that if the stiffness matrix is used to represent the operator $\mathcal{A}_h$ then, independent of the conditioning of the operators, such iterative schemes will always "require" a preconditioner, which transforms the dual representation of the residual to a corresponding primal representation. In other words, we need an operator which maps the residual as an element of V_h^* into V_h.

Consider now a general finite element system of the form

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

where the coefficient operator $\mathcal{A}_h$ maps the finite element space X_h into X_h^*, and let $\mathcal{B}_h : X_h^* \to X_h$ be a preconditioner as discussed in the previous section. Assume further that we want to approximate the solution x_h by applying a Krylov space method to the preconditioned system

$$\mathcal{B}_h \mathcal{A}_h x_h = \mathcal{B}_h f_h.$$

In agreement with the discussion above we assume that the operator $\mathcal{A}_h$ is represented by a sparse square stiffness matrix $\mathbb{A}_h = \mu_h \mathcal{A}_h \pi_h^{-1}$ mapping the primal representation of a function $x \in X_h$ into the dual representation of $\mathcal{A}_h x \in X_h^*$. Therefore, in order for the preconditioned coefficient operator, $\mathcal{B}_h \mathcal{A}_h$ to be cheaply computed, the preconditioner $\mathcal{B}_h$ should have the property that the primal representation of $\mathcal{B}_h f$ can be cheaply computed from the dual representation of $f \in X_h^*$, i.e., the operator $\pi_h \mathcal{B}_h \mu_h^{-1} : \mathbb{R}^n \to \mathbb{R}^n$ can be cheaply evaluated. The operator $\pi_h \mathcal{B}_h \mathcal{A}_h \pi_h^{-1} : \mathbb{R}^n \to \mathbb{R}^n$ is then computed by

$$\pi_h \mathcal{B}_h \mathcal{A}_h \pi_h^{-1} : \quad \text{primal repr.} \xrightarrow{\mathbb{A}_h} \text{dual repr.} \xrightarrow{\pi_h \mathcal{B}_h \mu_h^{-1}} \text{primal repr.}$$

In fact, this diagram should be seen as a discrete, and computational, version of the diagram (2.8) which expresses the mapping properties of the corresponding continuous operators.

7. Preconditioning discrete saddle point problems

In this final section of the paper we consider a few examples of discrete linear systems arising from stable discretizations of systems of partial differential equations of the form discussed above. In each example our goal is to construct preconditioners which give rise to condition numbers for the preconditioned systems which are bounded independently of the discretization parameter, and to present some numerical experiments which show the effect of the constructions. Of course, by now our strategy for the construction of such preconditioners is clear. As was done in Sections 3 and 4, we first identify the canonical block diagonal preconditioners for the corresponding continuous systems. Then, according to the conclusions of Section 5, the discrete preconditioner should simply be a corresponding discrete analog of this operator. However, for computational efficiency the exact inverses appearing in the canonical preconditioners should be replaced by proper cost effective, and norm equivalent, operators of the form discussed in Section 6.

The general approach we have taken in this paper is that we utilize the close connection between the discrete systems and the corresponding continuous counterparts to construct preconditioners. Before we consider some examples of discrete systems below, we should mention that many papers on preconditioning differ from the presentation given here in the sense that the discrete problems are analyzed more directly in a matrix framework, cf. for example [15, 36, 37, 55, 78, 83, 84, 93]. Furthermore, in this paper we have restricted the discussion to positive definite preconditioners, since such operators appear naturally as a consequence of the mapping properties of the coefficient operator of the corresponding continuous problem. However, many alternative techniques exist, where the preconditioners are more general operators.

A technique which may be quite effective for certain saddle point problems is so–called constraint preconditioning, where the preconditioner is also indefinite. In particular, this approach has been applied successfully to saddle point systems where the $(1, 1)$ block corresponds to a zero order differential operator. Consider a variational problem of the form (3.1). The idea is to build a preconditioner for the corresponding saddle point system from a simplified variational problem with the same constraint as the original problem. Therefore, the practical application of this approach is limited to problems where it is possible to find such simplified systems which can be inverted by a fast algorithm. As a consequence of the construction, the coefficient operator for the preconditioned system will be invariant on the constrained set. Furthermore, since the saddle point system is positive definite on the subspace associated the constraint, it is possible to use the standard conjugate gradient method as an iterative solver for the preconditioned system, and the convergence properties of this iteration will be determined by the behavior of the preconditioned operator restricted to this subspace. Indefinite preconditioners of this type are for example proposed in [11, 76], and a general discussion of constraint preconditioning is given in [15, Section 10].

Block triangular preconditioners for saddle point operators has also been discussed by many authors. A special example of such a preconditioner is the so–called Bramble–Pasciak transformation introduced in [22], see also [93, Chapter 9]. Here, the saddle point system is transformed into a positive definite system by a triangular transformation, and, as a consequence, the system can be solved by the standard conjugate gradient method. For certain problems this approach is very effective, but a possible difficulty is that the method requires a proper choice of a critical scaling parameter to obtain a positive definite operator. Another motivation for studying block triangular preconditioners is the desire to generalize, and therefore to improve, block diagonal preconditioners. Of course, a negative effect of this approach for symmetric problems of the form studied in this paper is that the symmetry, in general, will be lost. However, for generalized saddle point problems, which are nonsymmetric, these preconditioners have proved to be rather effective. We refer for example to [15, 16, 18, 17, 37, 57, 56, 60, 73] for various studies of such problems.

The list of examples of systems of partial differential equations discussed below is by no means complete. The construction of preconditioners and effective iterative methods are discussed for many other problems in the literature. One example, which formally seems very close to the mixed formulation of the Poisson problem, studied in Example 3.2 above and Example 7.3 below, is the Hellinger–Reissner mixed formulation of linear elasticity. However, for this problem the construction of a stable pair of finite element spaces has proved to be surprisingly hard. An analysis of preconditioners in the two dimensional case is done in [75, 96], based on the stable finite element spaces proposed in [8]. The main challenge in this study, successfully handled in [75, 96], is to construct preconditioners for a proper analog of the operator $I - \operatorname{grad}\operatorname{div}$ applied to piecewise polynomial spaces of symmetric matrix fields of at least cubic order. For the corresponding problem in three space dimensions it seems that similar finite element discretizations, based on a strong symmetry constraint, are too complex, cf. [1]. Therefore other approaches, like the weak symmetry approach, cf. [7] and references given there, should be preferred. For such methods the construction of preconditioners should be more straightforward, since all that will be required are several copies of standard $H(\operatorname{div})$ preconditioners. However, to the authors' knowledge, no such numerical studies have been reported so far.

An important class of linear systems are the systems that arise during Newton, or inexact Newton, iterations of nonlinear systems of partial differential equations. For such problems, preconditioning the family of linear systems arising during the iteration is often crucial. The text [37] is devoted to fast iterative solvers and preconditioners for discrete Navier–Stokes problems. Another study which fit into this category is [54], where numerical schemes for the nonlinear harmonic map problem are investigated.

Finally, we mention that preconditioners for bidomain systems are discussed in [62, 88, 87], while preconditioning and scattered data interpolation are investigated in [61]. Preconditioners for the Babuška Langrange multiplier method are analyzed in [46], and some linear systems arising from Runge-Kutta discretizations of parabolic equations are considered in [63, 86].

We start our study of examples of discrete systems by returning to Stokes problem.

Example 7.1. *Stokes problem.*
We recall the Stokes problem in Example 3.1, where the coefficient operator

$$\mathcal{A} = \begin{pmatrix} -\Delta & -\operatorname{grad} \\ \operatorname{div} & 0 \end{pmatrix}$$

is a mapping from $(H_0^1(\Omega))^n \times L_0^2(\Omega)$ onto $(H^{-1}(\Omega))^n \times L_0^2(\Omega)$. Stable finite element discretizations of this problem are for example discussed in the texts [26, 42]. If V_h and Q_h are proper finite element spaces such that $V_h \times Q_h \subset (H_0^1(\Omega))^n \times L_0^2(\Omega)$, then stability will follow from an inf–sup condition of the form

$$\inf_{q \in Q_h} \sup_{v \in V_h} \frac{\langle q, \operatorname{div} v\rangle}{\|v\|_{H^1}\|q\|_{L^2}} \geq \beta > 0, \tag{7.1}$$

h	2^{-1}	2^{-2}	2^{-3}	2^{-4}
$\kappa(\mathcal{B}_h\mathcal{A}_h)$	13.3	13.5	13.6	13.7

Table I. Condition numbers for the preconditioned Stokes system.

where the constant β is independent of h. Various preconditioners for the corresponding discrete problems are analyzed in [37] and references given there. Here we just present some simple numerical results obtained by using the Taylor–Hood element on a unit square domain Ω in two space dimensions. Hence, the space V_h consists of continuous piecewise quadratic vector fields and Q_h of continuous piecewise linear scalar fields with respect to a triangulation $\mathcal{T}_h$ of Ω. Here, and below, h indicates the mesh size associated with the triangulation $\mathcal{T}_h$. From Example 3.1 we know that the discrete preconditioner should be of the form

$$\mathcal{B}_h = \begin{pmatrix} (-\Delta_h)^{-1} & 0 \\ 0 & (I_h)^{-1} \end{pmatrix},$$

In Table I the condition number of the preconditioned system is shown. Corresponding numerical experiments by using a multigrid preconditioner combined with the minimum residual method typically result in reduction of the norm of the preconditioned residual $(\mathcal{B}r_k, r_k)$ by a factor 10^4 in less than 20 iterations. Similar experiments can be found in [37, 64, 78].

Example 7.2. *Stabilized Stokes problem.*
Due to the inf–sup condition (7.1) a number of simple pairs of finite element spaces will not be stable for the Stokes problem. For example, if both spaces V_h and Q_h are taken to be spaces of piecewise linear functions with respect to the same triangulation then this pair is in general not stable. However, in order to overcome this problem it has been proposed to perturb the problem from an incompressible flow into a "moderately compressible flow" by modifying the divergence constraint, $\operatorname{div} u = 0$, into an equation of the form $\operatorname{div} u + \epsilon^2 \Delta p = 0$. Here $\epsilon > 0$ is a small parameter which will be related to the discretization parameter h. We refer for example to [25, 41, 53, 52] for a discussion of stabilization techniques. The coefficient operator of the perturbed system is of the form

$$\mathcal{A}_\epsilon = \begin{pmatrix} -\Delta & -\operatorname{grad} \\ \operatorname{div} & \epsilon^2 \Delta \end{pmatrix}.$$

It is straightforward to check that this operator is an isomorphism mapping $(H_0^1(\Omega))^n \times (L_0^2(\Omega) \cap \epsilon H^1(\Omega))$ into the corresponding dual space $(H^{-1}(\Omega))^n \times (L_0^2(\Omega) + \epsilon^{-1} H^1(\Omega)^*)$, with corresponding operator norms bounded independently of ϵ. Therefore, at the continuous level, a uniform preconditioner is given of the form:

$$\mathcal{B}_\epsilon = \begin{pmatrix} (-\Delta)^{-1} & 0 \\ 0 & (I - \epsilon^2 \Delta)^{-1} \end{pmatrix}.$$

Consider a discretization of this problem with finite element spaces V_h and Q_h such that $V_h \times Q_h \subset (H_0^1(\Omega))^n \times (H^1(\Omega) \cap L_0^2(\Omega))$. Assume that both the velocity space V_h and the pressure space Q_h are spanned by continuous piecewise polynomials, and that the space of piecewise linear vector fields are contained in V_h. Then it follows from the inf–sup condition

	$\alpha\backslash h$	2^{-1}	2^{-2}	2^{-3}	2^{-4}
$\kappa(\mathcal{B}_{0,h}\mathcal{A}_{\alpha,h})$	0.02	4.0	5.6	7.0	7.7
	0.04	5.5	7.7	9.8	11.2
	0.08	7.4	12.4	16.7	20.0
$\kappa(\mathcal{B}_{\alpha,h}\mathcal{A}_{\alpha,h})$	0.02	3.9	5.7	7.1	7.8
	0.04	3.1	4.8	6.5	7.5
	0.08	2.45	3.9	5.8	7.1

Table II. Condition numbers for the preconditioned stabilized Stokes operator with different stabilization parameter $\epsilon^2 = \alpha h^2$.

in the continuous case that the spaces V_h and Q_h satisfy a modified inf–sup condition of the form

$$\sup_{v \in V_h} \frac{\langle q, \operatorname{div} v\rangle}{\|v\|_{H^1}} \geq \beta_1 \|q\|_{L^2} - \beta_2 h \|\operatorname{grad} q\|_{L^2}, \quad q \in Q_h,$$

where the constants β_1 and β_2 are independent of h. By using this condition we can obtain stability of the method in the norm of $(H_0^1(\Omega))^n \times (L_0^2(\Omega) \cap \epsilon\, H^1(\Omega))$, as long as the parameter ϵ is proportional to h. More precisely, if we assume that $\epsilon^2 = \alpha h^2$, where the positive constant α is independent of h, then the stability constant will be independent of h, but of course dependent of α. As a consequence, for such choices of ϵ discrete preconditioners of the form

$$\mathcal{B}_{\epsilon,h} = \begin{pmatrix} (-\Delta_h)^{-1} & 0 \\ 0 & (I_h - \epsilon^2\, \Delta_h)^{-1} \end{pmatrix}$$

will in fact be uniform preconditioners. Furthermore, if a standard inverse property of the form

$$\|\operatorname{grad} q\|_{L^2} \leq c h^{-1} \|q\|_{L^2}$$

holds, then the preconditioners $\mathcal{B}_{\epsilon,h}$ and $\mathcal{B}_{0,h}$ are spectrally equivalent, uniformly in h, but depending on α.

In Table II we give the condition numbers of the preconditioned operators $\mathcal{B}_{0,h}\mathcal{A}_{\epsilon,h}$ and $\mathcal{B}_{\epsilon,h}\mathcal{A}_{\epsilon,h}$ for different values of the proportionality constant α. The stabilized Stokes problem is discretized on different refinements of the unit square and using piecewise linear continuous finite elements for both velocity and pressure. Homogeneous Dirichlet conditions are assigned for the velocity. The condition numbers in Table II show a slight increase as the mesh is refined. Still there is a clear sign that the increase is about to level off, even for these rather coarse meshes.

Example 7.3. *The mixed Poisson problem.*
Recall the mixed formulation of the Poisson problem discussed in Example 3.2. The coefficient operator

$$\mathcal{A} = \begin{pmatrix} I & -\operatorname{grad} \\ \operatorname{div} & 0 \end{pmatrix} \tag{7.2}$$

is an isomorphism mapping $H_0(\operatorname{div}) \times L_0^2$ into its dual space, such that in the continuous case the preconditioner takes the form

$$\mathcal{B} = \begin{pmatrix} (I - \operatorname{grad}\operatorname{div})^{-1} & 0 \\ 0 & I \end{pmatrix}. \tag{7.3}$$

h	1	2^{-1}	2^{-2}	2^{-3}	2^{-4}
$\kappa(\mathcal{A}_h)$	8.25	15.0	29.7	59.6	119
$\kappa(\mathcal{B}_h\mathcal{A}_h)$	1.04	1.32	1.68	2.18	2.34

Table III. Condition numbers for the coefficient operator and the corresponding preconditioned operator for the mixed Poisson problem using the Raviart-Thomas element.

Well known stable finite element spaces $V_h \times Q_h \subset H_0(\mathrm{div}) \times L_0^2$ are for example the Raviart-Thomas spaces or Brezzi–Douglas–Marini spaces in two space dimensions and the spaces of the two Nedelec families in three dimensions. We refer to [26] for details. Hence, to construct a discrete preconditioner $\mathcal{B}_h$, the main challenge is to construct a preconditioner for the discrete analog of the operator $I - \mathrm{grad}\,\mathrm{div}$, corresponding to the bilinear form

$$\langle u, v\rangle + \langle \mathrm{div}\, u, \mathrm{div}\, v\rangle.$$

Here $\langle\cdot,\cdot\rangle$ denotes the corresponding L^2 inner product. The construction of such preconditioners was discussed in [3, 4, 5, 50, 49, 94]. In Table III we show the condition number of the coefficient matrix (7.2) when the problem is discretized with the lowest order Raviart–Thomas element and the condition number of the preconditioned coefficient matrix when a multigrid method is used to precondition the corresponding discrete operator $I - \mathrm{grad}_h\,\mathrm{div}$. These results are taken from [3], where more details are given. □

Remark 7.1. We recall that in the continuous case the mixed Poisson problem also allowed a second preconditioner $\mathcal{B} = \mathcal{B}_2$ of the form

$$\mathcal{B} = \mathcal{B}_2 = \begin{pmatrix} I & 0 \\ 0 & (-\Delta)^{-1} \end{pmatrix}.$$

This is based on the fact that the continuous system is well posed in $(L^2(\Omega)^n \times H^1(\Omega)$. However, for most standard stable finite element spaces for this problem, the pressure space Q_h is not a subspace of H^1. Therefore, since the discrete Laplace operator is not defined in an obvious way, it does not seem appropriate to consider a discrete analog of the preconditioner $\mathcal{B}_2$ for these spaces. However, it is possible to define a nonconforming discrete Laplace operator on Q_h and use a preconditioner for this operator to construct a discrete variant of $\mathcal{B}_2$. In fact, this is exactly what is done in [77]. Furthermore, in the construction of preconditioners for discrete versions of the time–dependent Stokes problem, studied below, such operators will indeed enter. □

Example 7.4. *A Maxwell problem.*
Let $\Omega \subset \mathbb{R}^3$ with boundary $\partial\Omega$. In connection with models based on Maxwell's equations one may encounter boundary value problems of the form

$$\begin{aligned} \mathrm{curl}\,\mathrm{curl}\, u &= f \quad \text{in } \Omega, \\ u \times n &= 0 \quad \text{on } \partial\Omega. \end{aligned} \tag{7.4}$$

Here n is the outwards-pointing unit normal on $\partial\Omega$, such that the boundary condition simply means that the tangential components of the vector field u must vanish on the boundary. In fact, there is no uniqueness of solution of the system (7.4), since any vector field of the form

$u = \operatorname{grad} q$, for $q \in H_0^1(\Omega)$, will be a solution of the corresponding homogeneous problem. Furthermore, the problem (7.4) cannot be solved for a general right–hand side f, since, if u is a smooth solution then

$$\langle f, \operatorname{grad} q\rangle = \langle \operatorname{curl}\operatorname{curl} u, \operatorname{grad} q\rangle = 0, \quad q \in H_0^1(\Omega), \tag{7.5}$$

where here, and below, $\langle\cdot,\cdot\rangle$ denote L^2 inner products. The system (7.4) will have at most one solution if we add the extra requirement that $\langle u, \operatorname{grad} q\rangle = 0$ for all $q \in H_0^1(\Omega)$, or equivalently, we require that $\operatorname{div} u = 0$. Furthermore, if u is such a smooth solution, where the right–hand side f satisfies (7.5), then u will solve the system

$$\begin{aligned} \operatorname{curl}\operatorname{curl} u - \operatorname{grad}\operatorname{div} u &= f \quad \text{in } \Omega, \\ \operatorname{div} u &= 0 \quad \text{in } \Omega, \\ u \times n &= 0 \quad \text{on } \partial\Omega. \end{aligned} \tag{7.6}$$

The new coefficient operator, $\operatorname{curl}\operatorname{curl} - \operatorname{grad}\operatorname{div}$, is simply the vector Laplace operator. By introducing a new variable, $p = \operatorname{div} u$, the system (7.6) can be given the following weak formulation:

Find $(u, p) \in H_0(\operatorname{curl}, \Omega) \times H_0^1(\Omega)$ such that

$$\begin{aligned} \langle \operatorname{curl} u, \operatorname{curl} v\rangle - \langle \operatorname{grad} p, v\rangle &= \langle f, v\rangle, \quad v \in H_0(\operatorname{curl}, \Omega), \\ -\langle u, \operatorname{grad} q\rangle - \langle p, q\rangle &= 0, \quad q \in H_0^1(\Omega). \end{aligned} \tag{7.7}$$

If we assume that the domain is contractible, then this system will have a unique solution. In fact, this system is just a special case of the Hodge–Laplace problem discussed in [6, Section 7]. The coefficient operator of this system, formally given as

$$\mathcal{A} = \begin{pmatrix} \operatorname{curl}\operatorname{curl} & -\operatorname{grad} \\ \operatorname{div} & -I \end{pmatrix},$$

is an isomorphism of $H_0(\operatorname{curl}, \Omega) \times H_0^1(\Omega)$ onto the corresponding representation of the dual space, defined by an extension of the L^2 inner product. Therefore, the canonical preconditioner is the operator of the form

$$\mathcal{B} = \begin{pmatrix} (I + \operatorname{curl}\operatorname{curl})^{-1} & 0 \\ 0 & (-\Delta)^{-1} \end{pmatrix}.$$

In order to construct a stable pair of finite elements for the system (7.7) some care has to be taken, since there is a delicate balance between choice of finite element subspaces V_h of $H_0(\operatorname{curl}, \Omega)$ and Q_h of $H_0^1(\Omega)$. However, as explained in [6], there are a number of stable finite elements. In Table IV we show the condition numbers of the preconditioned matrix on various refinements of the cube $\Omega = [0, \pi]^3$ using the first family of Nedelec elements of order zero for the unknown vector field, combined with continuous piecewise linears functions for the scalar field. Here, we taken the preconditioner $\mathcal{B}_h$ to be the discrete analog of $\mathcal{B}$, i.e., exact inverses of discrete differential operators are used. Therefore, the fact that the observed condition numbers $\kappa(\mathcal{B}_h\mathcal{A}_h)$ are nearly independent of h reflects that they are all close to $\kappa(\mathcal{B}\mathcal{A})$. However, to check the robustness of this preconditioner we have also tested the same preconditioner for a corresponding problem with a mildly varying coefficent M. More precisely, we consider the system with coefficient operator of the form:

$$\mathcal{A}_M = \begin{pmatrix} \operatorname{curl}(M \operatorname{curl}) & -\operatorname{grad} \\ \operatorname{div} & -I \end{pmatrix}, \text{ where } M(x, y, z) = \begin{pmatrix} x + 5 & z & \frac{1}{10}y \\ z & 10 + cos(z) & \frac{1}{10}x \\ \frac{1}{10}y & \frac{1}{10}x & 10 + sin(y) \end{pmatrix}.$$

h	1/2	1/4	1/6	1/8
$\kappa(\mathcal{B}_h\mathcal{A}_h)$	1.71	1.73	1.75	1.76
$\kappa(\mathcal{B}_h\mathcal{A}_{M,h})$	10.9	12.5	13.0	13.2

Table IV. Condition numbers for the preconditioned Maxwell system.

The coefficient M is uniformly positive definite on Ω, with $\kappa(M) < 4$. Again, the condition numbers $\kappa(\mathcal{B}_h\mathcal{A}_{M,h})$ of the preconditioned system appear to be bounded independently of h, in complete agreement with the theoretical predictions. □

Finally, we return to the three parameter–dependent problems studied in Section 4, i.e., the time dependent Stokes problem, the Reissner–Mindlin plate model, and the optimal control problem (4.6)–(4.7). For all three examples we identify stable discretizations, and therefore derive preconditioners which are uniform both with respect to the model parameters and the discretization parameter. Furthermore, these theoretical results will also be verified by numerical experiments.

Example 7.5. *The time dependent Stokes problem.*
Recall that we considered the time dependent Stokes problem in Example 4.2, with the following coefficient operator

$$\mathcal{A}_\epsilon = \begin{pmatrix} I - \epsilon^2\,\Delta & -\operatorname{grad} \\ \operatorname{div} & 0 \end{pmatrix}. \tag{7.8}$$

As we discussed above, the operator $\mathcal{A}_\epsilon$ is an isomorphism mapping $Y_\epsilon = (L^2 \cap \epsilon\, H^1_0)^n \times ((H^1 \cap L^2_0) + \epsilon^{-1}\, L^2_0)$ into its L^2–dual Y^*_ϵ, with bounds on $\mathcal{A}_\epsilon$ in $\mathcal{L}(Y_\epsilon, Y^*_\epsilon)$ and $\mathcal{A}^{-1}_\epsilon$ in $\mathcal{L}(Y^*_\epsilon, Y_\epsilon)$ independent of ϵ. We briefly review some of the results in [66]. As above we use the Taylor–Hood space, $V_h \times Q_h$, for the discretization. This leads to a uniformly stable discretization in the norm of Y_ϵ. Therefore, the canonical discrete preconditioner is of the form

$$\mathcal{B}_{\epsilon,h} = \begin{pmatrix} (I - \epsilon^2\,\Delta)^{-1}_h & 0 \\ 0 & \epsilon^2\, I^{-1}_h + (-\Delta_h)^{-1} \end{pmatrix}.$$

In our computations we replace "the inverse mass matrix" $I^{-1}_h : Q^*_h \to Q_h$ by an operator J_h obtained from a simple symmetric Gauss–Seidel iteration, and the operator $(-\Delta_h)^{-1} : Q_h \to Q_h$ by a multigrid operator N_h, computed with one symmetric Gauss-Seidel sweep as smoother. The operator $(I - \epsilon^2\,\Delta)^{-1}_h : V_h \to V_h$ is replaced by the corresponding multigrid operator $M_{\epsilon,h}$ on V_h.

In Tables V and VI we report the proper condition numbers for the building blocks of the full preconditioner, while the condition numbers $\kappa(\mathcal{B}_{\epsilon,h}\mathcal{A}_{\epsilon,h})$ are given for different values of ϵ and h in Table VII. In agreement with the theory above, these condition numbers seem to be uniformly bounded with respect to both ϵ and h.

In [66] similar results were also reported for various other finite element discretizations, like the Mini-element, the $P_2 - P_0$ element, and Crouzeix-Raviart elements. For the two latter elements, the auxiliary space method [98] was used to define preconditioners for the suitable nonconforming Laplacian on the space of discontinuous piecewise constants, cf. Remark 7.1. For other studies of preconditioners for problems motivated from time dependent Stokes problems we refer to [23, 29, 38, 36, 37, 58, 59, 66, 71, 91]. □

h	2^{-2}	2^{-3}	2^{-4}	2^{-5}	2^{-6}	2^{-7}
$\kappa(N_h(-\Delta_h))$	1.71	1.50	1.47	1.47	1.47	1.47
$\kappa(J_h I_h)$	1.66	1.62	1.61	1.60	1.60	1.60

Table V. Condition numbers for the operators $N_h(-\Delta_h)$ and $J_h I_h$

$h\backslash\,\epsilon$	0	0.001	0.01	0.1	0.5	1.0
2^{-3}	1.11	1.11	1.03	1.14	1.22	1.22
2^{-5}	1.11	1.09	1.03	1.23	1.24	1.24
2^{-7}	1.11	1.02	1.20	1.24	1.24	1.24

Table VI. Condition numbers for the operators $M_{\epsilon,h}(I-\epsilon^2\,\Delta)_h$

$h\backslash\,\epsilon$	0	0.001	0.01	0.1	0.5	1.0
2^{-3}	6.03	6.05	6.92	13.42	15.25	15.32
2^{-5}	6.07	6.23	10.62	15.14	15.59	15.61
2^{-7}	6.08	7.81	14.18	15.55	15.64	15.65

Table VII. $\kappa(\mathcal{B}_{\epsilon,h}\mathcal{A}_{\epsilon,h})$ for the Taylor–Hood element

Remark 7.2. Following the continuous theory in Example 4.2 we should also be able to find a uniform preconditioner for the discrete versions of the operator $\mathcal{A}_\epsilon$, given by 7.8, of the form

$$\mathcal{B}_{\epsilon,h} = \begin{pmatrix} (I - \operatorname{grad}\operatorname{div} - \epsilon^2\,\Delta)_h^{-1} & 0 \\ 0 & I_h \end{pmatrix}.$$

In fact, from the discussion given in Section 5 we know that such preconditioners can be found, if the corresponding finite element discretization is uniformly stable in the norm of $X_\epsilon = (H_0(\operatorname{div}) \cap \epsilon\, H_0^1) \times L_0^2$. However, the Taylor-Hood element, like most other Stokes elements, is not uniformly stable in X_ϵ, cf. [65]. Nonconforming finite elements, which are uniformly stable, have been constructed in [65] in the two dimensional case, and in [89] in three dimensions. Uniform preconditioners for discrete versions of the operator $(I - \operatorname{grad}\operatorname{div} - \epsilon^2\,\Delta)^{-1}$ have been developed in [81]

Example 7.6. *The Reissner–Mindlin plate model.*
We recall from Example 4.3 that the continuous model is a linear system of partial differential equations, on a domain $\Omega \subset \mathbb{R}^2$, with a coefficient operator given by

$$\mathcal{A}_t = \begin{pmatrix} -\operatorname{div}\mathcal{C}\mathcal{E} & 0 & -I \\ 0 & 0 & -\operatorname{div} \\ -I & \operatorname{grad} & -t^2 I \end{pmatrix},$$

where $t \geq 0$ is the thickness parameter. Furthermore, the operator $\mathcal{A}_t$ is an isomorphism mapping X_t to X_t^*, where

$$X_t = (H_0^1)^2 \times H_0^1 \times (H^{-1}(\operatorname{div}) \cap t \cdot (L^2)^2) \quad \text{and } X_t^* = (H^{-1})^2 \times H^{-1} \times (H_0(\operatorname{rot}) + t^{-1} \cdot (L^2)^2),$$

$t\backslash h$	2^{-3}	2^{-4}	2^{-5}	2^{-6}	2^{-7}
0	8.17	10.7	11.1	10.6	9.62
0.01	8.15	10.7	11.2	11.1	9.68
0.1	8.64	10.8	11.1	11.2	11.1
1	17.5	18.4	19.0	19.0	18.9

Table VIII. Condition numbers for the preconditioned Reissner–Mindlin operators $\mathcal{B}_{t,h}\mathcal{A}_{t,h}$
.

with t–independent bounds on $\mathcal{A}_t$ in $\mathcal{L}(X_t, X_t^*)$ and $\mathcal{A}_t^{-1}$ in $\mathcal{L}(X_t^*, X_t)$. As a consequence, the canonical uniform preconditioner in the continuous case takes the form

$$\mathcal{B}_t = \begin{pmatrix} (-\Delta)^{-1} & 0 & 0 \\ 0 & (-\Delta)^{-1} & 0 \\ 0 & 0 & D_t \end{pmatrix},$$

where

$$D_t = I + (1 - t^2)\,\mathrm{curl}(I - t^2\Delta)^{-1}\,\mathrm{rot}\,.$$

A number of uniformly stable finite element discretizations have been proposed for the Reissner–Mindlin plate model, cf. for example the survey paper [40] and references given there. One of the simplest methods in this class is the Arnold–Falk method. Here the finite element approximation of the space X_t is of the form $X_{t,h} = V_h \times W_h \times \Gamma_h$, where V_h consists of the span of all continuous piecewise linear vector fields and cubic bubble functions on each triangle of $\mathcal{T}_h$. Furthermore, the scalar space W_h is the nonconforming Crouzeix–Raviart space, i.e., piecewise linear functions with mean value continuity on each edge, and the space Γ_h is the space of piecewise constant vector fields. Note that $X_{t,h} \not\subseteq X_t$ since $W_h \not\subseteq H^1$. Hence, this method is a nonconforming finite element approximation of the Reissner–Mindlin plate model, cf Remark 5.2. To approximate the operator $D_t = I + (1-t^2)\,\mathrm{curl}(I - t^2\Delta)^{-1}\,\mathrm{rot}$ on the space Γ_h^*, we use duality to introduce an operator rot_h mapping Γ_h^* into the dual space of continuous piecewise linears. We refer to [2] for details.

The discrete preconditioner $\mathcal{B}_{t,h}$ is of the form

$$\mathcal{B}_{t,h} = \begin{pmatrix} L_h & 0 & 0 \\ 0 & M_h & 0 \\ 0 & 0 & D_{t,h} \end{pmatrix}, \tag{7.9}$$

Here L_h and M_h are standard V–cycle multigrid preconditioners for the discrete Laplacian on the spaces V_h and W_h. The operator $D_{t,h} : \Gamma_h^* \to \Gamma_h$ takes the form

$$D_{t,h} = I_h + (1 - t^2)\,\mathrm{curl}\,N_{t,h}\,\mathrm{rot}_h,$$

where $N_{t,h}$ is a standard V–cycle multigrid preconditioner on the space of continuous piecewise linears. The condition numbers $\kappa(\mathcal{B}_{t,h}\mathcal{A}_{t,h})$, for different values of t and h, are given in Table VIII. These results are taken from [2]. We observe that $\kappa(\mathcal{B}_{t,h}\mathcal{A}_{t,h})$ varies with t and h, but in agreement with the theory, the results appear to be uniformly bounded with respect to both parameters. □

Remark 7.3. For several of the suggested finite element discretizations of the Reissner–Mindlin plate model discussed above, the discrete spaces are of the form $X_{t,h} = V_h \times W_h \times \Gamma_h$, where the space Γ_h is a subspace of $H_0(\text{rot})$. This may appear attractive, since in contrast to the situation for the Arnold–Falk element discussed above, the operator rot is now a well–defined operator mapping Γ_h onto the corresponding space of piecewise constants. For simplicity, consider the case $t = 0$. To define the preconditioner (7.9) we need to evaluate a discrete version of the differential operator $D_t = I + \text{curl}\,\text{rot}$. This operator has a natural discretization $D_{0,h} : \Gamma_h \to \Gamma_h^*$, defined from the $H(\text{rot})$ inner product, given by

$$\langle \zeta, \eta \rangle_{H(\text{rot})} = \langle \zeta, \eta \rangle + \langle \text{rot}\,\zeta, \text{rot}\,\eta \rangle,$$

where the inner products on the right–hand side are in L^2. Furthermore, the corresponding matrix, $\mu_h D_{0,h} \pi_h^{-1}$, is a sparse matrix, reflecting the fact that the differential operator D_0 is a local operator. Here π_h and μ_h are the representation operators for the space Γ_h, corresponding to the primal and dual representation, cf. Section 6 above. Hence, it appears that this block of the preconditioner is a simple sparse matrix, which can be cheaply evaluated. However, if the operator $D_{0,h}$ is to be used as a building block for the preconditioner, then we need to evaluate the matrix $\pi_h D_{0,h} \mu_h^{-1}$, and not $\mu_h D_{0,h} \pi_h^{-1}$, and the matrix $\pi_h D_{0,h} \mu_h^{-1}$ is not sparse, due to the continuity requirements on the space Γ_h. Discussions on how to overcome this problem is given in [2, Section 7]. □

Example 7.7. *An optimal control problem.*
Finally, we recall the optimal control problem in Example 4.4, where the coefficient operator is given by

$$\mathcal{A}_{\epsilon,h} = \begin{pmatrix} I & 0 & I - \Delta \\ 0 & \epsilon^2 I & -I \\ I - \Delta & -I & 0 \end{pmatrix}, \tag{7.10}$$

mapping X_ϵ to X_ϵ^*, and with corresponding operator norms bounded uniformly with respect to ϵ. Here,

$$X_\epsilon = (L^2(\Omega) \cap \epsilon^{1/2} \cdot H^1(\Omega)) \times \epsilon \cdot L^2(\Omega) \times (\epsilon^{-1} \cdot L^2(\Omega) \cap \epsilon^{-1/2} \cdot H^1(\Omega))$$

and X_ϵ^* is the corresponding dual space defined by extending the L^2 inner product. Hence, the canonical preconditioner in the continuous case is given by

$$\mathcal{B}_\epsilon = \begin{pmatrix} (I - \epsilon\,\Delta)^{-1} & 0 & 0 \\ 0 & \epsilon^{-2} I & 0 \\ 0 & 0 & \epsilon^2 (I - \epsilon\,\Delta)^{-1} \end{pmatrix}. \tag{7.11}$$

In Table IX we show condition numbers $\kappa(\mathcal{B}_{\epsilon,h}\mathcal{A}_{\epsilon,h})$ for various values of ϵ and on different refinements of the unit square in two dimensions using continuous piecewise linear elements for all variables. Similar experiments, using the minimum residual method combined with an algebraic multigrid method show that the norm of the preconditioned residual $(\mathcal{B}r_k, r_k)$ is reduced by a factor 10^6 in less than 30 iterations independently of ϵ and h. Similar results were obtained in [80]. Algorithms for other optimal control problems with elliptic PDE constraints are discussed in, e.g., [20, 68, 70, 80]. □

$h\backslash\epsilon$	1	0.1	0.01	0.001
2^{-1}	2.39	2.67	3.04	2.82
2^{-2}	2.41	2.73	3.04	3.05
2^{-3}	2.42	2.74	3.04	3.06
2^{-4}	2.42	2.74	3.04	3.06
2^{-5}	2.42	2.75	3.04	3.06

Table IX. The condition number $\kappa(\mathcal{B}_{\epsilon,h}\mathcal{A}_{\epsilon,h})$ for the optimal control problem.

Acknowledgement

The first author is grateful to Bjørn Fredrik Nielsen for many discussions on optimal control problems. The second author is grateful to Douglas N. Arnold and Richard S. Falk for their cooperation over many years. In fact, most of his knowledge on the topic of this paper is developed through joint research projects with them. He would also thank Joachim Schöberl for many useful discussions on problems related to this paper.

REFERENCES

1. D. N. Arnold, G. Awanou, and R. Winther, *Finite elements for symmetric tensors in three dimensions*, Mathematics of Computation **77** (2008), no. 263, 1229–1251.
2. D. N. Arnold, R. S. Falk, and R. Winther, *Preconditioning discrete approximations of the Reissner-Mindlin plate model*, RAIRO Modélisation Mathématique et Analyse Numérique **31** (1997), no. 4, 517–557.
3. ______, *Preconditioning in H(div) and applications*, Mathematics of Computation **66** (1997), 957–984.
4. ______, *Multigrig preconditioning in H(div) on non-convex polygons*, Journal of Computational and Applied Mathematics **17** (1998), 307–319.
5. ______, *Multigrid in H(div) and H(curl)*, Numerische Mathematik **85** (2000), 197–217.
6. ______, *Finite element exterior calculus, homological techniques, and applications*, Acta Numerica **15** (2006), 1–155.
7. ______, *Mixed finite element methods for linear elasticity with weakly imposed symmetry*, Mathematics of Computation **76** (2007), no. 260, 1699–1723.
8. D. N. Arnold and R. Winther, *Mixed finite elements for elasticity*, Numerische Mathematik **92** (2002), no. 3, 401–419.
9. O. Axelsson and G. Lindskog, *On the eigenvalue distribution of a class of preconditioning methods*, Numerische Mathematik (1986), 479–498.
10. ______, *On the rate of convergence of the preconditioned conjugate gradient method*, Numerische Mathematik (1986), 499–523.
11. O. Axelsson and M. Neytcheva, *Preconditioning methods for linear systems arising in constrained optimization problems*, Numerical Linear Algebra with Applications **10** (2003), 3–31.
12. I. Babuška and A. K. Aziz, *Survey lectures on the mathematical foundation of the finite element method*, The Mathematical Foundation of the Finite Element Method with Applications to Partial Differential Equations (A. K. Aziz, ed.), Academic Press, New York - London, 1972, pp. 3–345.
13. I. Babuška, *Error bounds for finite element methods*, Numerische Mathematik **16** (1971), 322–333.
14. A. Backem, M. Grötschel, and B. Korte (eds.), *Solution of large linear systems of equtions by the conjugate gradient method*, Springer-Verlag, 1983.
15. M. Benzi, G. H. Golub, and J. Liesen, *Numerical solution of saddle point problems*, Acta Numerica **14** (2005), 1–137.
16. M. Benzi and J. Liu, *Block preconditioning for saddle point systems with indefinite* $(1,1)$ *block*, International Journal of Computer Mathematics **84** (2007), no. 8, 1117–1129.
17. M. Benzi, M. A. Olshanskii, and Z. Wang, *Augmented Lagrangian preconditioners for the incompressible Navier-Stokes equations*, url:http://www.mathcs.emory.edu/ molshan/OLSH/publ.html.

18. M. Benzi and M A. Olshanskii, *An augumented Lagrangian-based approach to the Oseen problem*, SIAM Journal on Scientific Computing **28** (2006), 2095–2113.
19. J. Bergh and J. Löfström, *Interpolation spaces*, Springer-Verlag, 1976.
20. A. Borzì, K. Kunisch, and Do Y. Kwak, *Accuracy and convergence properties of the finite difference multigrid solution of an optimal control optimality system*, SIAM Journal on Control and Optimization **41** (2003), no. 5, 1477–1497.
21. J. H. Bramble, J. E. Pasciak, and P. Vassilevski, *Computational scales of sobolev norms with application to preconditioning*, Mathematics of Computation **69** (2000), 463–480.
22. J. H. Bramble and J. E. Pasciak, *A preconditioning technique for indefinite systems resulting from mixed approximations of elliptic problems*, Mathematics of Computation **50** (1988), 117.
23. ———, *Iterative techniques for time dependent Stokes problems*, Computers & Mathematics with Applications. An International Journal **33** (1997), no. 1-2, 13–30, Approximation theory and applications.
24. J. H. Bramble, *Multigrid methods*, Pitman Research Notes in Mathematical Sciences, vol. 294, Longman Scientific & Technical, Essex, England, 1993.
25. F. Brezzi and J. Douglas, *Stabilized mixed methods for the Stokes problem*, Numerische Mathematik **53** (1988), no. 1-2, 225–235.
26. F. Brezzi and M. Fortin, *Mixed and hybrid finite element methods*, Springer-Verlag, 1991.
27. F. Brezzi, *On the existence, uniqueness and approximation of saddle–point problems arising from Lagrangian multipliers*, RAIRO. Analyse Numérique **8** (1974), 129–151.
28. W. L. Briggs, V. E. Henson, and S. F. McCormick, *A multigrid tutorial*, 2nd ed., SIAM Books, 2000.
29. J. Cahouet and J.-P. Chabard, *Some fast 3D finite element solvers for the generalized Stokes problem*, International Journal for Numerical Methods in Fluids **8** (1988), no. 8, 869–895.
30. X.-C. Cai, *Additive Schwarz algorithms for parabolic convection-diffusion equations*, Numerische Mathematik **60** (1991), 41–62.
31. ———, *Multiplicative Schwarz methods for parabolic problems*, SIAM Journal on Scientific Computing **15(3)** (1994), 587–603.
32. Z. Cai, C.I. Goldstein, and J.E. Pasciak, *Multilevel iterationm for mixed finite element system with penalty*, SIAM Journal on Scientific Computing **14** (1993), 1072–1088.
33. R. Chandra, *Conjugate Gradient methods for partial differential equations*, Report 129, Computer Science Department, Yale University, New Haven, CT, 1978.
34. M. Dauge, *Stationary Stokes and Navier-Stokes systems on two- or three-dimensional domains with corners. I. Linearized equations*, SIAM Journal on Mathematical Analysis **20** (1989), no. 1, 74–97.
35. E.G. D'Yakonov, *On an iterative method for the solution of finite difference equations*, Doklady Akademii Nauk SSSR **138** (1961), 522–525.
36. H. C. Elman, D. J. Silvester, and A. J. Wathen, *Performance and analysis of saddle point preconditioners for the discrete steady-state Navier-Stokes equations*, Numerische Mathematik **90** (2002), no. 4, 665–688.
37. ———, *Finite elements and fast iterative solvers: with applications in incompressible fluid dynamics*, Numerical Mathematics and Scientific Computation, Oxford University Press, New York, 2005.
38. H. C. Elman, *Preconditioners for saddle point problems arising in computational fluid dynamics*, Applied Numerical Mathematics. An IMACS Journal **43** (2002), no. 1-2, 75–89, 19th Dundee Biennial Conference on Numerical Analysis (2001).
39. R. S. Falk and J. E. Osborn, *Error estimates for mixed methods*, RAIRO Analyse Numérique **14** (1980), 249–277.
40. R. S. Falk, *Finite elements for the Reissner-Mindlin plate*, Mixed Finite Elements, Compatibility Conditions, and Applications, Lecture Notes in Mathematics, 1939, Springer-Verlag, 2008, pp. 195–232.
41. L. P. Franca and R. Stenberg, *Error analysis of some Galerkin least squares methods for the elasticity equations*, SIAM Journal on Numerical Analysis **28** (1991), no. 6, 1680–1697.
42. V. Girault and P.-A. Raviart, *Finite element methods for Navier-Stokes equations*, Springer Series in Computational Mathematics, vol. 5, Springer-Verlag, Berlin, 1986, Theory and algorithms.
43. G.H. Golub and C.F. van Loan, *Matrix computations*, Johns Hopkins University Press, 1989.
44. J.E. Gunn, *The solution of difference equations by semi-explicit iterative techniques*, SIAM Journal on Numerical Analysis **2** (1965), 24–45.
45. W. Hackbusch, *Iterative solution of large sparse systems of equations*, Springer-Verlag, 1994.
46. E. Haug and R. Winther, *A domain embedding preconditioner for the Lagrange multiplier system*, Mathematics of Computation **69** (1999), 65–82.
47. M.R. Hestenes and E. Stiefel, *Methods of conjungate gradients for solving linear systems*, Journal of Research of the National Bureau of Standards of J. Res. Nat. Bur. Standards **49** (1952), 409–436.
48. Ralf Hiptmair, *Operator preconditioning*, Computers & Mathematics with Applications **52** (2006), no. 5, 699–706.
49. R. Hiptmair and J. Xu, *Nodal auxiliary space preconditioning in* H(curl) *and* H(div) *spaces*, SIAM Journal on Numerical Analysis **45** (2007), no. 6, 2483–2509.

50. R. Hiptmair, *Multigrid method for* **H**(div) *in three dimensions*, Electronic Transactions on Numerical Analysis **6** (1997), no. Dec., 133–152, Special issue on multilevel methods (Copper Mountain, CO, 1997).
51. ———, *Multigrid method for Maxwell's equations*, SIAM Journal on Numerical Analysis **36** (1999), no. 1, 204–225.
52. T. J. R. Hughes, L. P. Franca, and M. Balestra, *A new finite element formulation for computational fluid dynamics: V. Circumventing the Babuska-Brezzi condition: a stable Petrov-Galerkin formulation of the Stokes problem accommodating equal-order interpolations*, Computer Methods in Applied Mechanics and Engineering **59** (1986), 85–99.
53. T. J. R. Hughes and L. P. Franca, *A new finite element formulation for computational fluid dynamics: VII. The Stokes problem with various well-posed boundary conditions: symmetric formulations that converge for all velocity/pressure spaces*, Computer Methods in Applied Mechanics and Engineering **65** (1987), 85–96.
54. Q. Hu, X.-C Tai, and R. Winther, *A saddle point approach to the computation of harmonic maps*, SIAM Journal on Numerical Analysis **47** (2009), no. 2, 1500–1523.
55. D. Kay, D. Loghin, and A. Wathen, *A preconditioner for the steady-state Navier-Stokes equations*, SIAM Journal on Scientific Computing **24** (2002), no. 1, 237–256.
56. A. Klawonn and G. Starke, *Block triangular preconditioners for nonsymmetric saddle point problems: field-of-values analysis*, Numerische Mathematik **81** (1999), 577594.
57. A. Klawonn, *Block-triangular preconditioners for saddle point problems with a penalty term*, SIAM Journal on Scientific Computing **19** (1998), no. 1, 172–184, Special issue on iterative methods (Copper Mountain, CO, 1996).
58. G. M. Kobelkov and M. A. Olshanskii, *Effective preconditioning of Uzawa type schemes for generalized Stokes problem*, Numerische Mathematik **86** (2000), 443–470.
59. H. P. Langtangen, K.-A. Mardal, and R. Winther, *Numerical methods for incompressible viscous flow*, Advances in Water Resources **25** (2002), no. 8, 1125–1146.
60. D. Loghin and A. J. Wathen, *Analysis of preconditioners for saddle-point problems*, SIAM Journal on Scientific Computing **25** (2004), 2029–2049.
61. T. Lyche, T. K. Nilssen, and R. Winther, *Preconditioned iterative methods for scattered data interpolation*, Advances in Computational Mathematics **17** (2002), no. 3, 237–256.
62. K.-A. Mardal, B. F. Nielsen, X. Cai, and A. Tveito, *An order optimal solver for the discretized Bidomain equations*, Numerical Linear Algebra with Applications **14** (2007), no. 2, 83–98.
63. K.-A. Mardal, T. K. Nilssen, and G. A. Staff, *Order optimal preconditioners for implicit Runge-Kutta schemes applied to parabolic PDEs*, SIAM Journal of Scientific Computing **29** (2007), no. 1, 361–375.
64. K.-A. Mardal, J. Sundnes, and H. P. Langtangen, *Systems of PDEs and block preconditionering*, Computational Partial Differential Equations using Diffpack - Advanced Topics (H. P. Langtangen and A. Tveito, eds.), Springer, 2003.
65. K.-A. Mardal, X.-C. Tai, and R. Winther, *A robust finite element method for Darcy–Stokes flow*, SIAM Journal on Numerical Analysis **40** (2002), 1605–1631.
66. K.-A. Mardal and R. Winther, *Uniform preconditioners for the time dependent Stokes problem*, Numerische Mathematik **98** (2004), no. 2, 305–327.
67. ———, *Erratum: "Uniform preconditioners for the time dependent Stokes problem" [Numer. Math.* **98** *(2004), no. 2, 305–327]*, Numerische Mathematik **103** (2006), no. 1, 171–172.
68. T. P. Mathew, M. Sarkis, and C. E. Schaerer, *Analysis of block matrix preconditioners for elliptic optimal control problems*, Numerical Linear Algebra with Applications **14** (2007), no. 4, 257–279.
69. J. Nečas, *Equations aux dérivée partielles*, Presses de l'Université de Montréal, 1965.
70. B. F. Nielsen and K.-A. Mardal, *Efficient preconditioners for optimality systems arising in connection with inverse problems.*
71. M. A. Olshanskii, J. Peters, and A. Reusken, *Uniform preconditioners for a parameter dependent saddle point problem with application to generalized Stokes interface equations*, Numerische Mathematik **105** (2006), 159–191.
72. M. A. Olshanskii and A. Reusken, *On the convergence of a multigrid method for linear reaction-diffusion problems*, Computing **65** (2000), no. 3, 193–202.
73. M. A. Olshanskii and Y. Vassilevski, *Pressure Schur complement preconditioners for the discrete Oseen problem*, SIAM Journal on Scientific Computing **29** (2007), 2686–2704.
74. C. C. Paige and M. A. Saunders, *Solution of sparse indefinite systems of linear equations*, SIAM Journal on Numerical Analysis **12** (1975), 617–629.
75. J. E. Pasciak and Y. Wang, *A multigrid preconditioner for the mixed formulation of linear plane elasticity*, SIAM Journal on Numerical Analysis **44** (2006), no. 2, 478–493.
76. I. Perugia and V. Simoncini, *Block-diagonal and indefinite symmetric preconditioners for mixed finite element formulations*, Numerical Linear Algebra with Applications **7** (2000), 585–616.
77. T. Rusten, P. S. Vassilevski, and R. Winther, *Interior penalty preconditioners for mixed finite element*

approximations of elliptic problems, Mathematics of Computation **65** (1996), no. 214, 447–466.
78. T. Rusten and R. Winther, *A preconditioned iterative method for saddlepoint problems*, SIAM Journal on Matrix Analysis and Applications **13** (1992), no. 3, 887–904, Iterative methods in numerical linear algebra (Copper Mountain, CO, 1990).
79. ———, *Substructure preconditioners for elliptic saddle point problems*, Mathematics of Computation **60** (1993), no. 201, 23–48.
80. J. Schöberl and W. Zulehner, *Symmetric indefinite preconditioners for saddle point problems with applications to PDE-constrained optimization problems*, SIAM Journal on Matrix Analysis and Applications **29** (2007), no. 3, 752–773.
81. J. Schöberl, *Multigrid methods for a parameter dependent problem in primal variables*, Numerische Mathematik **84** (1999), 97–119.
82. ———, Private communication.
83. D. Silvester and A. Wathen, *Fast iterative solution of stabilised Stokes systems. Part I: Using simple diagonal preconditioners*, SIAM Journal on Numerical Analysis **30** (1993), no. 3, 630–649.
84. ———, *Fast iterative solution of stabilized Stokes systems. Part II: Using block diagonal preconditioners*, SIAM Journal on Numerical Analysis **31** (1994), no. 5, 1352–1367.
85. B. Smith, P. Bjørstad, and W. Gropp, *Domain decomposition*, Cambridge University Press, 1996.
86. G. A. Staff, K. A. Mardal, and T. K. Nilssen, *Preconditioning of fully implicit Runge-Kutta schemes for parabolic PDEs*, Modeling, Identification and Control **27** (2006), no. 2, 109–123.
87. J. Sundnes, G. T. Lines, X. Cai, B. F. Nielsen, K.-A. Mardal, and A. Tveito, *Computing the electrical activity in the heart*, Springer-Verlag, 2006.
88. J. Sundnes, G. Lines, K.-A Mardal, and A. Tveito, *Multigrid block preconditioning for a coupled system of partial differential equations modeling the electrical activity in the heart*, Computer Methods in Biomechanics and Biomedical Engineering **5** (2002), no. 6, 397–409.
89. X.-C. Tai and R. Winther, *A discrete de Rham complex with enhanced smoothness*, Calcolo **43** (2006), 287–306.
90. L. N. Trefethen and D. Bau, *Numerical linear algebra*, SIAM, 1997.
91. S. Turek, *Efficient solvers for incompressible flow problems*, Springer, 1999.
92. Panayot S. Vassilevski and Junping Wang, *Wavelet-like methods in the design of efficient multilevel preconditioners for elliptic PDEs*, Multiscale wavelet methods for partial differential equations, Wavelet Anal. Appl., vol. 6, Academic Press, San Diego, CA, 1997, pp. 59–105. MR MR1474997 (98i:65040)
93. Panayot S. Vassilevski, *Multilevel block factorization preconditioners*, Springer, New York, 2008, Matrix-based analysis and algorithms for solving finite element equations. MR MR2427040
94. P. S. Vassilevski and J. P. Wang, *Multilevel iterative methods for mixed finite element discretizations of elliptic problems*, Numerische Mathematik **63** (1992), no. 4, 503–520.
95. P. S. Vassilevski and J. Wang, *An application of the abstract multilevel theory to nonconforming finite element methods*, SIAM Journal on Numerical Analysis **32** (1995), no. 1, 235–248.
96. Y. Wang, *Overlapping Schwarz preconditioner for the mixed formulation of plane elasticity*, Applied Numerical Mathematics. An IMACS Journal **54** (2005), no. 2, 292–309.
97. R. Winther, *Some superlinear convergence results for the conjugate gradient method*, SIAM Journal on Numerical Analysis **17** (1980), no. 1, 14–17.
98. J. Xu, *The auxiliary space method and optimal multigrid preconditioning techniques for unstructured grids*, Computing **56** (1996), 215–235.