CONSTRUCTION OF PRECONDITIONERS BY MAPPING PROPERTIES FOR SYSTEMS OF PARTIAL DIFFERENTIAL EQUATIONS

KENT-ANDRE MARDAL[1] AND RAGNAR WINTHER[2]

Abstract. The purpose of this paper is to discuss a general approach to the construction of preconditioners for the linear systems of algebraic equations arising from discretizations of systems of partial differential equations. The discussion here is closely tied to our earlier paper [34], where we gave a comprehensive review of a mathematical theory for constructing preconditioners based on the mapping properties of the coefficient operator of the underlying differential systems. In the presentation given below we focus more on specific examples, while just an outline of the general theory is given.

1. Introduction

Discretization methods for partial differential equations are often designed to mimic key properties of the problems they are approximating. For example, discretizations of conservation laws are frequently constructed such that corresponding discrete quantities are conserved, while finite element spaces typically inherit continuity properties from the requirement that they should be subspaces of the Sobolev spaces they approximate. In the present paper we will demonstrate that also iterative solution algorithms for discretized differential systems should inherit key properties from the corresponding continuous systems. In particular, we will argue that the mapping properties of the governing differential operators suggest the basic structure of efficient preconditioners for the corresponding discrete systems.

Linear differential operators usually have unbounded spectrum, and as a consequence, standard iterative methods like the conjugate gradient and related Krylov space methods will not converge, or in many cases will not even be well defined, for such problems. These properties are reflected in the corresponding discrete problems. Even if these problems are of finite dimension, the spectrum of the coefficient operators will be unbounded as the mesh is refined, causing slower convergence of iterative methods for finer meshes. To overcome this effect, preconditioners are introduced. In fact, even the underlying differential equations can be transformed into problems which admit convergent iterations if the systems are properly preconditioned. As we will explain in Section 3 below, the natural preconditioner for such systems is an isomorphism mapping the space of right hand sides into the solution space. By applying the corresponding ideas to the discrete analogs we obtain so–called *canonical preconditioners* for the corresponding discrete systems.

Both authors were supported by Center of Excellence grants from the Norwegian Research Council.

Typically, these preconditioners lead to bounded condition numbers, and therefore to rates of convergence which are bounded uniformly with respect to the discretization. However, these preconditioners are usually not computationally efficient, since they typically are composed of inverses of discrete differential operators. Computationally efficient preconditioners can often be constructed by replacing these inverses by corresponding analogs generated by techniques like multigrid or domain decomposition methods.

This paper is closely tied to the review paper [34], where we present a comprehensive overview of an abstract theory to construct preconditioners for discrete systems from the mapping properties of the corresponding differential operators. This technique originates from the papers [2, 3] and has later been exploited in a variety of applications, see [34] and references given there. The strength of the approach to the construction of preconditioners for discretized differential systems presented here is most striking when we consider singular perturbation problems. For such problems we will construct preconditioners which behave uniformly well both with respect to the discretization and the perturbation parameter.

We remark that there are close similarities between the abstract approach to preconditioning taken here, and several other more abstract discussions of preconditioning. The relation between preconditioning of elliptic problems and the concept "equivalent operators" has been utilized by several authors, cf. for example [40, 4], while a more general approach to "operator preconditioning" is outlined in [21]. Furthermore, although the examples presented both in this paper and in [34] are mostly symmetric saddle point problems, the theory is not limited to such applications. Some discussion of nonsymmetric problems are for example given in [25, 30, 38]. Alternative block preconditioners for saddle point problems, on triangular or indefinite form, are described in [7, 8, 9, 10, 17, 26, 27, 28, 35].

The numerical examples presented in this paper are implemented in FEniCS [19] and are slight modifications of the examples described in [29]. In FEniCS we have used the linear algebra backend Trilinos [20] and its algebraic multigrid toolbox ML.

The outline of this paper is as follows. In Section 2 we present a series of numerical examples based on discretizations of a various differential systems, namely the Stokes problem, linear elasticity, the stabilized Stokes problem, and the time dependent Stokes problem. The general abstract framework given in [34] is briefly described in Section 3, before we revisit in Section 4 some of the examples studied previously.

2. Motivating Examples

In this section we will discuss a number of numerical experiments, where we present condition numbers and iteration counts for various discrete differential systems using different preconditioners. The purpose of these examples is to motivate the need for a better theoretical understanding on how to construct effective preconditioners. Such a theoretical overview will then be presented in the next section. Throughout the discussion below it is convenient to let the underlying system of partial differential equations be denoted $\mathcal{A}x = f$, where $\mathcal{A}$ is the governing differential operator, f is

the given data, while x is the solution. Here, x and f are elements of appropriate function spaces. Furthermore, we will use $\mathcal{B}$ to denote various preconditioners at the continuous level.

In all the examples below the domain Ω is taken to be the unit square in $\mathbb{R}^2$, and we will use $\partial\Omega$ to denote the boundary of Ω. Most commonly the computations will be done with respect to a uniform triangular mesh, obtained by dividing Ω into $h \times h$ squares, where $h = 1/(N-1)$, and then divided each square into two triangles. The corresponding discrete finite element systems are denoted $\mathcal{A}_h x_h = f_h$, and the corresponding preconditioners by $\mathcal{B}_h$. In the numerical examples below we will typically report estimates for the condition numbers of the operators $\mathcal{B}_h\mathcal{A}_h$. Furthermore, if $\mathcal{A}_h$ is positive definite we will give iteration counts for the conjugate gradient method (CG) applied to the preconditioned system $\mathcal{B}_h\mathcal{A}_h x_h = \mathcal{B}_h f_h$, while for indefinite problems we have used the conjugate gradient method on the normal equation (CGN). More precisely, we have used CG to solve

$$\mathcal{B}_h\mathcal{A}_h^T\mathcal{B}_h\mathcal{A}_h x = \mathcal{B}_h\mathcal{A}_h^T\mathcal{B}_h f_h.$$

This application of CGN requires that the preconditioner is symmetric and positive definite. Of course, since all our examples below are symmetric, we could also have used an iterative method which is more tailored, and usually more efficient, for such problems, cf. for example [36]. However, the purpose here is not efficiency, but the comparisons between different models and preconditioners. Therefore, we have preferred to use a robust method like CGN.

In some examples we compute the condition number by using a canonical preconditioner, i.e., the preconditioner is composed of inverses of discrete differential operators. However, since this approach requires the inversion of matrices, it is limited to coarse grids. On finer grids the canonical preconditioners are modified by introducing algebraic multigrid (AMG) operators as replacements for the exact inverses.

Example 2.1. *The Stokes problem.*
The Stokes problem for an incompressible fluid is:

$$-\Delta u - \operatorname{grad} p = f, \text{ in } \Omega, \tag{1}$$

$$\operatorname{div} u = 0, \text{ in } \Omega, \tag{2}$$

$$u = 0, \text{ on } \partial\Omega, \tag{3}$$

where we refer to the unknown vector field u as the velocity and the unknown scalar field p as the pressure. We can write this system of equations formally as

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

The Stokes problem is a saddle-point problem, and for a stable finite element discretization the proper inf–sup condition should be fulfilled. In our experiment we have used the lowest order Taylor–Hood element, i.e., continuous piecewise quadratic velocity fields, and continuous piecewise linear pressures. As a preconditioner we have used the discrete analog of the

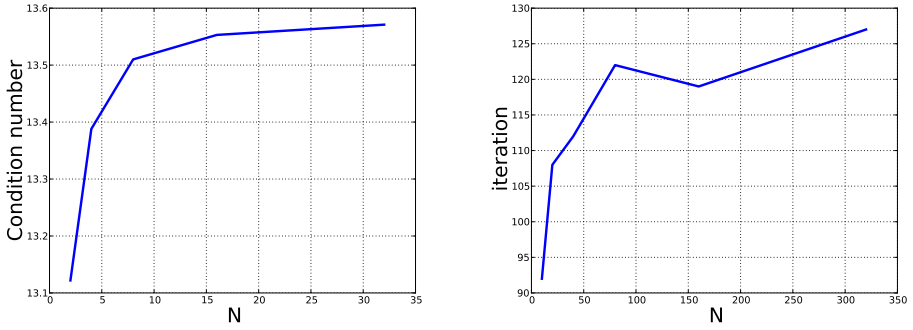

FIGURE 1. The left figure shows the condition number of the preconditioned Stokes operator operator $\mathcal{B}_h\mathcal{A}_h$ for different meshes using the canonical preconditititioners $\mathcal{B}_h$. The right figure shows the number of iterations required for convergence when using an AMG preconditioner combined with CGN for different meshes. The convergence criteria was a relative reduction of the preconditioned residual by a factor greater than 10^{10} in the discrete L_2 norm.

operator

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

on these spaces. In Figure 1, the condition number of the corresponding discrete analogs of the operator $\mathcal{BA}$ and iterations counts for the preconditioned CGN are given. Clearly, the condition number and the number of iterations remains bounded as the mesh is refined. We remark that the condition number of the operator $\mathcal{A}_h$ in this example is $1.1 \cdot 10^6$ for $N = 32$, and that it grows like h^{-2}. □

Example 2.2. *The linear elasticity problem.*
The elasticity problem for an isotropic material reads:

$$-(\lambda + \mu)\,\text{grad div}\, u - \mu\Delta u = f, \text{ in } \Omega, \tag{4}$$

$$u = 0, \text{ on } \partial\Omega, \tag{5}$$

where the positive constants λ and μ are the Lame's elasticity constants. For a nearly incompressible material $\lambda \gg \mu$, and for such problems the phenomena of *locking*, i.e., the case where the finite element computations produce significantly smaller deformation than it should, is well-known [13, 14]. Locking can be avoided by using reduced integration [18, 24] or special finite element spaces [31]. Another consequence of $\lambda \gg \mu$ is that the condition numbers of the corresponding finite element matrices increase, and hence the convergence of iterative methods deteriorates. We first investigate the linear system obtained by a finite element approximation of the system (4)–(5) using continuous piecewise quadratics to approximate the displacement u. The parameter $\mu = 1$, while λ and the mesh parameter varies. In Figure 2 we show the number of iterations needed for convergence of CG combined

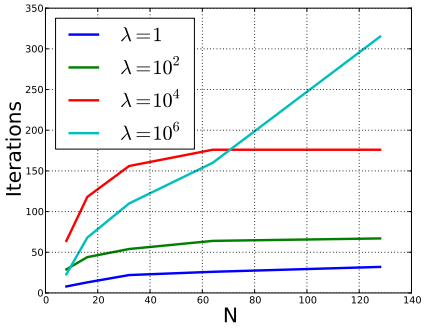

FIGURE 2. The figure shows the number of iterations required for convergence with respect to different mesh resolutions and values of λ. We used CG combined with a standard AMG preconditioner. The convergence criteria was a relative reduction of the preconditioned residual of a factor greater than 10^6 in the preconditioned discrete L_2 norm.

with an AMG preconditioner. Clearly, the number of iterations increases as $\lambda \to \infty$.

One common way of avoiding locking is to introduce a separate variable for the divergence, i.e., we let $p = (\lambda + \mu) \operatorname{div} u$. For $\mu = 1$ the problem can then be written as

$$\begin{aligned} -\Delta u - \operatorname{grad} p &= f, \text{ in } \Omega, \\ \operatorname{div} u - \epsilon^2 p &= 0, \text{ in } \Omega, \\ u &= 0, \text{ on } \partial\Omega, \end{aligned}$$

where $\epsilon^2 = 1/(1+\lambda)$. So as λ becomes large, ϵ tends to zero. This is a mixed system which formally can be written in the form

$$\mathcal{A}\begin{pmatrix} u \\ p \end{pmatrix} = \begin{pmatrix} -\Delta & -\operatorname{grad} \\ \operatorname{div} & -\epsilon^2 I \end{pmatrix}\begin{pmatrix} u \\ p \end{pmatrix} = \begin{pmatrix} f \\ 0 \end{pmatrix}.$$

Hence, when λ is large, ϵ tends to zero and the problem formally approaches the Stokes system. Therefore, as in the previous example, we discretize the problem by the lowest order Taylor–Hood element. In particular, as above the velocity field is approximated by piecewise quadratics. We consider the same preconditioner as we used in Example 2.1, i.e., the discrete analog of

$$\mathcal{B}_1 = \begin{pmatrix} -\Delta^{-1} & 0 \\ 0 & I \end{pmatrix}.$$

Motivated by the diagonal elements in $\mathcal{A}$ we also consider the alternative preconditioner

$$\mathcal{B}_2 = \begin{pmatrix} -\Delta^{-1} & 0 \\ 0 & \epsilon^{-2} I \end{pmatrix}.$$

These two preconditioners are compared with respect to different values of h and λ in Figure 3. We note that the two preconditioners coincide when $\lambda = 1$. The leftmost graph in Figure 3 clearly demonstrates that the condition numbers of the discrete analogs of $\mathcal{B}_1\mathcal{A}$ are bounded independently

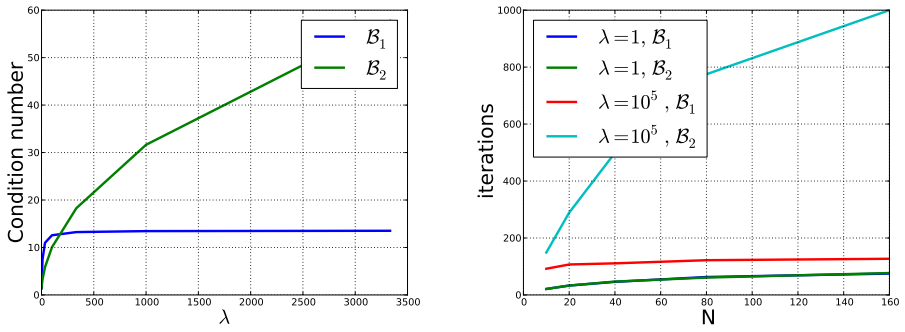

FIGURE 3. The left figure shows the condition number of the preconditioned discrete operators as a function of λ, using the two different canonical preconditititioners corresponding to $\mathcal{B}_1$ and $\mathcal{B}_2$. The right figure shows the number of iterations required for convergence with respect to the mesh resolutions, for two different values of λ, when using CGN with an AMG preconditioner. The convergence criteria was a relative reduction of the preconditioned residual of a factor greater than 10^{10} in the discrete L_2 norm.

of λ, while the condition numbers of the discrete analogs of $\mathcal{B}_2\mathcal{A}$ grows as λ increases. Similarly, in the rightmost graph the CGN combined with an AMG preconditioner of $\mathcal{B}_1$ gives convergence in a number of iterations that is bounded independently of both h and λ, while the preconditioner $\mathcal{B}_2$ behaves poorly. □

Example 2.3. *The stabilized Stokes problem.*
Consider the Stokes problem from Example 2.1 once more. By perturbing the system slightly one can obtain stability of the discretization by using the same polynomial order of the finite element spaces for both velocity and pressure. One popular stabilization method consists of perturbing the equation for mass conservation by adding a small diffusion term on the pressure, i.e., we consider the system

$$\mathcal{A}_\epsilon \begin{pmatrix} u \\ p \end{pmatrix} = \begin{pmatrix} -\Delta & -\operatorname{grad} \\ \operatorname{div} & \epsilon^2\Delta \end{pmatrix} \begin{pmatrix} u \\ p \end{pmatrix} = \begin{pmatrix} f \\ g \end{pmatrix}. \tag{6}$$

It has been shown in e.g. [22, 23] that the choice $\epsilon^2 = \beta h^2$, where $\beta > 0$ is fixed, gives a stable discretization with equal order elements in velocity and pressure. This method requires additional boundary conditions on the pressure which usually is assumed to be homogeneous Neumann conditions.

For this system we consider two preconditioners:

$$\mathcal{B}_1 = \begin{pmatrix} -\Delta^{-1} & 0 \\ 0 & I \end{pmatrix} \text{ and } \mathcal{B}_2 = \begin{pmatrix} -\Delta^{-1} & 0 \\ 0 & (-\epsilon^2\Delta)^{-1} \end{pmatrix}.$$

Here, the $\mathcal{B}_1$ is the good Stokes preconditioner from above. This preconditioner is independent of ϵ, while the preconditioner $\mathcal{B}_2$ includes the diffusion term in the pressure. In the computations we have used piecewise linears for both velocity and pressure.

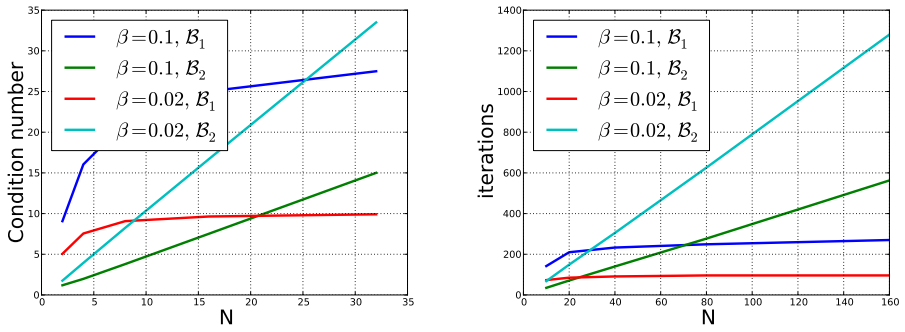

FIGURE 4. The left figure shows the condition number of the preconditioned system using the two different canonical preconditititioners $\mathcal{B}_1$ and $\mathcal{B}_2$ with respect to h, for two different values of β. The right figure shows the number of iterations required for convergence when using an AMG preconditioner combined with CGN for different meshes. The convergence criteria was a relative reduction of the preconditioned residual of a factor greater than 10^{10} in the discrete L_2 norm.

Figure 4 shows the behavior of the discrete versions of the preconditioners $\mathcal{B}_1$ and $\mathcal{B}_2$ in terms of h and β, where $\epsilon^2 = \beta h^2$. Clearly, the ϵ–independent preconditioner $\mathcal{B}_1$, combined with a Krylov solver like CGN, produces an efficient algorithm that obtains convergence in a number of iterations that is bounded independently of h, for a given β. On the other hand, $\mathcal{B}_2$ shows poor performance. □

Example 2.4. *The time dependent Stokes problem.*
The time dependent Stokes problem is an initial value problem of the form:

$$\begin{aligned} u_t - \Delta u - \operatorname{grad} p &= f, \text{ in } \Omega, \ t > 0, \\ \operatorname{div} u &= 0, \text{ in } \Omega, \ t > 0, \\ u &= 0, \text{ on } \partial\Omega, \ t > 0, \\ u &= u_0, \text{ in } \Omega, \ t = 0. \end{aligned}$$

Here, u is the unknown velocity vector, p is the uknown pressure, and u_0 is the initial condition. Various time stepping schemes applied to this equation result in systems of equations of the following form to be solved at each time step,

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

where g here includes both the right-hand side f and the solution at the previous time step. The small positive parameter ϵ is related to the time step.

Efficient preconditioners for this system have been described in [16, 39, 32, 33]. These are of the form:

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

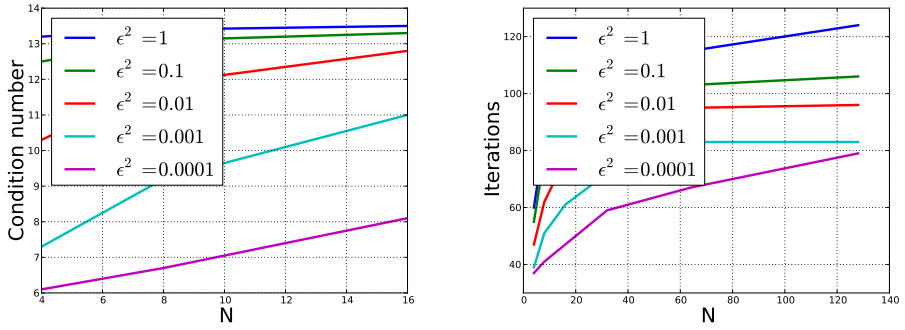

FIGURE 5. The left figure shows the condition number of the preconditioned system arising from discretizations of the time dependent Stokes problem, for different ϵ. The right figure shows the number of iterations required for convergence when using an AMG preconditioner combined with CGN for different mesh refinements and choices of ϵ. The convergence criteria was a relative reduction of the preconditioned residual of a factor greater than 10^{10} in the discrete L_2 norm.

In this example, it is harder to "guess" the form of an efficient preconditioner by simple means. However, as we will see in the discussion below, this preconditioner arise naturally from the abstract reasoning in following section. Again the system is discretized by the lowest order Taylor–Hood method. In Figure 5 we present the condition numbers of the discrete preconditioned system and corresponding iteration counts. The observed condition numbers are bounded by 14, independently of ϵ and mesh refinements, and the number of iterations is bounded by 130. Similar results can be found in e.g. [32, 34]. □

3. Variational problems and preconditioning

In this section we will first briefly review the abstract approach to preconditioning outlined in [34], and thereafter we will discuss how this theory relates to the examples presented above. The main motivation for designing preconditioners for linear systems of equations is related to the use of iterative solution algorithms. Consider a linear system of the form

$$\mathcal{A}x = f, \tag{7}$$

where $\mathcal{A}$ is a bounded and invertible linear operator mapping a real, separable Hilbert space X into itself. In other words, we assume that $\mathcal{A}, \mathcal{A}^{-1} \in \mathcal{L}(X, X)$, where in general $\mathcal{L}(X, Y)$ is the set of bounded linear operators from X to Y. If the operator $\mathcal{A}$ is also symmetric and positive definite then the equation (7) can be solved by the CG iteration in the sense that the approximate solutions $\{x_m\}$ satisfies

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

Here, the convergence parameter α satisfies $0 \leq \alpha < 1$, $x_0 \in X$ is an abritary start vector, and the energy norm, $\|\cdot\|_{\mathcal{A}}$, is given by $\|x\|_{\mathcal{A}}^2 = \langle \mathcal{A}x, x \rangle$,

where $\langle \cdot, \cdot \rangle$ denotes the inner product on X. More precisely, the convergence parameter α admits the bound

$$\alpha \leq \frac{\sqrt{\kappa(\mathcal{A})} - 1}{\sqrt{\kappa(\mathcal{A})} + 1},$$

where the condition number of the operator $\mathcal{A}$, $\kappa(\mathcal{A})$, is given as the product of the operator norms of $\mathcal{A}$ and $\mathcal{A}^{-1}$. In fact, x_m is the best approximation of the solution x in the associated Krylov space

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

in the sense that

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

A key to the efficiency of CG is that only one evaluation of the operator $\mathcal{A}$ is necessary to compute x_m from x_{m-1}.

If the operator $\mathcal{A}$ is a symmetric, but indefinite isomorphism on X, then we can instead apply CGN, i.e., CG to the normal equation $\mathcal{A}^2 x = \mathcal{A}b$. Also in this case the convergence rate depends on $\kappa(\mathcal{A})$, since $\kappa(\mathcal{A}^2) = \kappa(\mathcal{A})^2$ for symmetric operators. For more details on CG, CGN, and more general Krylov space methods we refer to [34, Section 2] and references given there.

3.1. **Preconditioning.** Krylov space methods like CG and CGN can in general not be applied directly to systems of partial differential equations, since the coefficient operators are not bounded. Consider for example the Stokes operator $\mathcal{A}$ studied in Example 2.1 above. This operator cannot be seen as a bounded operator of a Hilbert space into itself, since the eigenvalues accumulate at infinity. If the domain Ω is a bounded subset of $\mathbb{R}^n$, then the appropriate weak formulation of the operator $\mathcal{A}$ leads to a bounded operator from $X = (H_0^1(\Omega))^n \times L_0^2(\Omega)$ into $X^* = (H^{-1}(\Omega))^n \times L_0^2(\Omega)$. Here, $(H_0^1(\Omega))^n$ is the space of all L^2 vector fields with weak first order derivatives in L^2, and which are zero in the trace sense on the boundary $\partial\Omega$, while $L_0^2(\Omega)$ denote the set of L^2 scalar fields with mean value zero. Finally, the space $(H^{-1}(\Omega))^n \supset (L^2(\Omega))^n$ represents the dual of $(H_0^1(\Omega))^n$. In particular, $X \subsetneq X^*$. Since the operator $\mathcal{A}$ maps the solution space X out of itself, Krylov space methods like CG and CGN are in general not well–defined for such problems. However, if $\mathcal{B}$ is an operator such that $\mathcal{B} \in L(X^*, X)$ then the composition $\mathcal{B} \circ \mathcal{A} \in \mathcal{L}(X, X)$, cf. Figure 6. Hence, we can apply Krylov space methods to the corresponding preconditioned equation

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

Assume in general that $\mathcal{A} \in \mathcal{L}(X, X^*)$ is an isomorhism, i.e., $\mathcal{A}^{-1} \in \mathcal{L}(X^*, X)$. Here, X and X^* are assumed to be separable Hilbert spaces, where we should think of X^* as a representation of the dual of X. We also assume that $\mathcal{A}$ is symmetric in the sense that

$$\langle \mathcal{A}x, y \rangle = \langle x, \mathcal{A}y \rangle, \quad x, y \in X, \tag{8}$$

where $\langle \cdot, \cdot \rangle$ is the associated duality pairing between X and X^*.

We will 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. It is a direct consequence of these assumptions that the composition $\mathcal{BA}$ is an isomorphism mapping X to itself. Furthermore, the operator $\mathcal{BA} : X \to X$ is symmetric in the inner product $\langle \mathcal{B}^{-1}\cdot, \cdot\rangle$ on X. Therefore, the preconditioned system

$$\mathcal{BA}x = \mathcal{B}f,$$

can be solved by CGN with a convergence rate bounded by $\kappa(\mathcal{BA}) = ||\mathcal{BA}||_{\mathcal{L}(X,X)}||(\mathcal{BA})^{-1}||_{\mathcal{L}(X,X)}$.

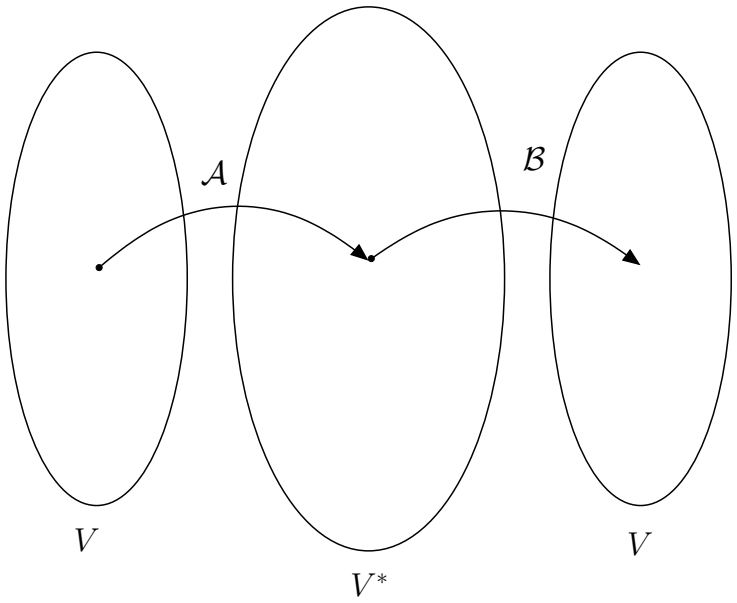

FIGURE 6. The mapping property of the composition of $\mathcal{A}$ and $\mathcal{B}$.

3.2. **Variational problems.** Many systems of partial differential equations can be formulated as variational problems. Throughout the discussion here, we will consider abstract variational problems of the form:
Find $x \in X$ such that for $f \in X^*$:

$$a(x, y) = \langle f, y\rangle, \quad y \in X, \tag{9}$$

where, as above, X is a Hilbert space, X^* its dual space, and $\langle \cdot, \cdot\rangle$ the associated duality pairing. Furthermore, we assume that $a : X \times X \to \mathbb{R}$ is a symmetric bilinear form. Following the classical variational theory of Babuška [5, 6], the following two conditions, referred to as a boundedness condition and an inf–sup condition, are sufficient to guarantee existence, uniqueness, and well-posedness of the problem:

$$|a(x, y)| \le c_1\|x\|_X\|y\|_X, \quad x, y \in X. \tag{10}$$

and

$$\inf_{x\in X}\sup_{y\in X}\frac{a(x, y)}{\|x\|_X\|y\|_X} \ge c_2 > 0, \tag{11}$$

where c_1 and c_2 are positive constants.

We may write the problem (9) as a linear system of the form

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

where $\mathcal{A} \in \mathcal{L}(X, X^*)$ is the linear operator defined by

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

The conditions (10) and (11) will imply that the operator $\mathcal{A}$ is an isomorphism mapping X to X^*. In particular,

$$\|\mathcal{A}\|_{\mathcal{L}(X,X^*)} \leq c_1 \text{ and } \|\mathcal{A}^{-1}\|_{\mathcal{L}(X^*,X)} \leq c_2^{-1}.$$

Furthermore, since the bilinear form a is symmetric, the operator $\mathcal{A}$ is symmetric in the sense of (8). Therefore, throughout the rest of the discussion here we assume that the bilinear form a is symmetric.

The canonical preconditioner in this case is the Riesz operator mapping X^* to X, i.e., $\mathcal{B} \in \mathcal{L}(X^*, X)$ is defined by

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

Here $\langle \cdot, \cdot \rangle_X$ denotes an inner product on the Hilbert space X. The operator $\mathcal{B}$ is symmetric and positive definite in the sense specified above. In fact, it is straightforward to verify that

$$\kappa(\mathcal{B}\mathcal{A}) = \|\mathcal{A}\|_{\mathcal{L}(X,X^*)} \cdot \|\mathcal{A}^{-1}\|_{\mathcal{L}(X^*,X)} \leq c_1/c_2, \tag{12}$$

and therefore the convergence rate of CGN, applied to the preconditioned system

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

can be bounded in terms of the constants c_1 and c_2.

Next, we consider the discrete variational problems approximating the system (9). Hence, let $\{X_h\}$ be a family of subspaces of the Hilbert space X, where $h > 0$ is refered to as the discretization parameter. As approximations of the system (9) we consider discrete problems of the form:
Find $x_h \in X_h$ such that:

$$a(x_h, y) = \langle f, y\rangle, \quad y \in X_h. \tag{13}$$

It is clear that the boundedness of the bilinear form a on X, given by (10), in particular implies boundedness on X_h. On the other hand, the inf–sup condition (11) will not imply that the corresponding condition holds on X_h. In general, the subspace X_h may even be constructed such that the system (13) is singular. However, the corresponding discrete inf–sup condition,

$$\inf_{x\in X_h} \sup_{y\in X_h} \frac{a(x, y)}{\|x\|_X \|y\|_X} \geq c_3 > 0, \tag{14}$$

where the constant c_3 is independent of h, will indeed ensure stability.

If the stability condition (14) holds then the system (13) can be written as a linear system of the form

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

where $\mathcal{A}_h : X_h \to X_h^*$ is defined by

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

Furthermore, in the discrete case we define the canonical preconditioner, $\mathcal{B}_h : X_h^* \to X_h$, by

$$\langle \mathcal{B}_h f, y \rangle_X = \langle f, y \rangle, \quad y \in X_h. \tag{15}$$

By arguing exactly as above we obtain that

$$\kappa(\mathcal{B}_h \mathcal{A}_h) \leq c_1/c_3,$$

and as a consequence CGN, applied to the preconditioned system

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

will converge with a convergence rate which can be bounded in terms of c_1 and c_3. We refer to [34, Section 5] for more details.

Remark 3.1. The canonical preconditioners defined from (15) will for most problems not lead to efficient preconditioners, since they typically will be composed of inverses of discrete differential operators that are expensive to evaluate. To obtain efficient preconditioners these operators have to be replaced by alternative *spectrally equivalent* operators, which can be evaluated cheaply. Here we refer to two symmetric and positive definite operators $\mathcal{B}_{1,h}, \mathcal{B}_{2,h} : X_h^* \to X_h$ as spectrally equivalent if there are constants α_0 and α_1, independent of the discretization parameter h, such that

$$\alpha_0 \langle f, B_{1,h} f \rangle \leq \langle f, B_{2,h} f \rangle \leq \alpha_1 \langle f, B_{1,h} f \rangle, \quad f \in X_h^*.$$

This is why we replaced the exact inverses by corresponding AMG operators in the numerical experiments above. We refer to [34, Section 6] for more details.

4. The numerical experiments – revisited

From the discussion above we can conclude that if discrete linear systems of the form (13) are preconditioned by a canonical preconditioner of the form (15), then the condition number of the preconditioned system can be bounded in terms of the boundedness constant c_1 and the inf–sup constant c_3. Below we will revisit the examples we presented in Section 2 above, and discuss how the results we observed there can be explained by the theory we just have outlined.

Example 4.1. *The Stokes problem.*
The coefficient operator of the Stokes problem is of the form

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

and, as we stated above, this operator can be seen as an isomorphism mapping $X = (H_0^1(\Omega))^n \times L_0^2(\Omega)$ into $X^* = (H^{-1}(\Omega))^n \times L_0^2(\Omega)$, see e.g., [15]. The canonical preconditioner $\mathcal{B}$ should therefore be a Riesz mapping from X^* to X, and in this case this operator can be taken to be of the form

$$\mathcal{B} = \begin{pmatrix} -\Delta^{-1} & 0 \\ 0 & I \end{pmatrix}. \tag{16}$$

Furthermore, the corresponding discrete operator will be composed of the inverse of the discrete vector Laplacian, and a mass matrix (replacing the

identity). It is well known that multigrid methods leads to spectrally equivalent and computational efficient analogs of these operators. Therefore, both the condition number and the number of iterations required by CGN, shown in Figure 1, are bounded independently of the mesh parameter. □

Example 4.2. *The linear elasitiicy problem.*
In Example 2.2 we considered the mixed formulation of the linear elasticity problem. The coefficient operator of this problem is of the form,

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

where $\epsilon^2 = 1/(1+\lambda)$. It is easy to see that for $\epsilon \in [0,1]$ the operator $\mathcal{A}_\epsilon$ is an isomorphism mapping $X = (H_0^1(\Omega))^n \times L^2(\Omega)$ into $X^* = (H^{-1}(\Omega))^n \times L^2(\Omega)$, with operator norms bounded independently of ϵ. Therefore, the preconditioner used for Stokes problem is also the appropriate canonical preconditioner here, and the condition number of the preconditioned problem is bounded uniformly in ϵ. This explains the uniform results, both with respect to the elasticity constant λ and the mesh parameter h, observed for the discrete version of the preconditioner $\mathcal{B}_1$ in Figure 3. □

Example 4.3. *The stabilized Stokes problem.*
The coefficient operator of the stabilized Stokes problem studied in Example 2.3 is given by

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

Above we studied finite element approximations of this operator when the parameter ϵ was proportional to the mesh paramer h. The motivation for such studies is the desire to stabilize discretizations of the Stokes problem, where the velocity and the pressure are approximated by finite element spaces of the same polynomial order. However, the system (6) also appear in different settings, for example through time-stepping schemes for simplifications of the Biot equations describing a poroelastic problems c.f. e.g. [1, 12]. In these applications there are no relation between the positive parameter ϵ and the mesh parameter h. Therefore, we will discuss preconditioners for $\mathcal{A}_\epsilon$ in the more general situation, where no relation between ϵ and h is assumed. In particular, we will derive preconditioners for $\mathcal{A}_\epsilon$ in the continous case.

It is relative straightforward to check that for each positive ϵ the operator $\mathcal{A}_\epsilon$ is an isomorphism mapping $X = (H_0^1(\Omega))^n \times H^1(\Omega)$ onto $X^* = (H^{-1}(\Omega))^n \times H_0^{-1}(\Omega)$. Here, $H_0^{-1}(\Omega) \subset L_0^2(\Omega)$ represents the dual of $H^1 \cap L_0^2$. However, the bounds on the associated operator norms on $\mathcal{A}_\epsilon$ and $\mathcal{A}_\epsilon^{-1}$ will depend on ϵ. In order obtain corresponding ϵ–independent bounds on the operator norms, we need to introduce an ϵ–dependent norm on the solution space X, which degenerates to the L^2–norm as ϵ tends to zero. Let $X_\epsilon = (H_0^1(\Omega))^n \times (L_0^2 \cap \epsilon H^1)(\Omega)$, and X_ϵ^* the associated dual space. Here, we have used the notation of intersection of Hilbert spaces, which we will discuss in more details below. The space $(L_0^2 \cap \epsilon \cdot H^1)(\Omega)$ is equal to $H^1(\Omega)$ as a set for each $\epsilon > 0$, but its norm is given by

$$\|q\|^2_{L^2 \cap \epsilon H^1} = \|q\|^2_{L^2} + \epsilon^2 \|q\|^2_{H^1}.$$

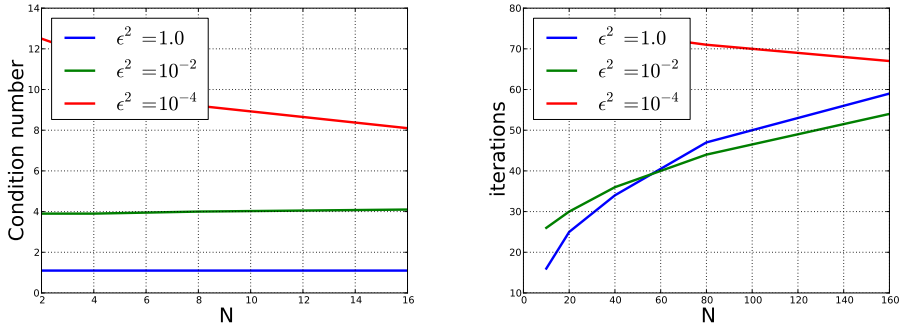

FIGURE 7. The left figure shows the condition number of the preconditioned system for the stabilized Stokes problem using Taylor–Hood elements, as a function of the mesh refinements and ϵ. The right figure shows the number of iterations required for CGN combined with an AMG preconditioner. The convergence criteria was a relative reduction of the preconditioned residual of a factor greater than 10^{10} in the discrete L_2 norm.

For $\epsilon = 0$ the space $(L_0^2 \cap \epsilon \cdot H^1)(\Omega)$ coincides with $L_0^2(\Omega)$, with identical norms, and hence the mapping property of $\mathcal{A}_\epsilon$ degenerates to the mapping property of the coefficient operator for Stokes problem as ϵ tends to zero.

The operator $\mathcal{A}_\epsilon$ is an isomorphism mapping X_ϵ to X_ϵ^*, and with associates operator norms bounded independently of ϵ. As a consequence, if we use the Riesz mapping $\mathcal{B}_\epsilon$ from X_ϵ^* to X_ϵ as a preconditioner, then the condition number $\kappa(\mathcal{B}_\epsilon \mathcal{A}_\epsilon)$ will be bounded independently of ϵ, cf. (12). The Riesz mapping from X_ϵ^* to X_ϵ is in the present case given by

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

In Figure 7 we provide numerical experiments for different values of ϵ and mesh size h. We use the Taylor–Hood element since we allow $\epsilon \to 0$ independently of h. Clearly, the condition numbers remain bounded independently of both h and ϵ. The number of iterations required for convergence seems to be increasing for large ϵ, but this is probably because the asymptotic limit is not obtained for large ϵ and h. In fact, the number of iterations required in this problem is lower than that for the Stokes problem, see Example 2.1.

It can also be seen that if ϵ^2 is bounded by βh^2, where the constant β is independent of h, then the operators $\mathcal{B}_{0,h}$ and $\mathcal{B}_{\epsilon,h}$ are spectrally equivalent. This explains the numerical results of Example 2.3, where we observes that the preconditioner $\mathcal{B}_{0,h}$ behaved uniformly well with respect to h for $\epsilon^2 = \beta h^2$. □

The two last examples we have encountered above both have a coefficient operator which depends on a parameter ϵ. For the linear elasticity problem we saw that we obtained a uniform conditioning of the preconditioned system by using an ϵ–independent preconditioner, while for the perturbed

Stokes system (6) we use ϵ–dependent preconditioner to obtain uniform conditioning in ϵ, although a standard Stokes preconditioner works fine for $\epsilon \approx h$ as seen in Example 2.3. The different behaviour of the two systems is reflected in the properties that the mixed linear elasticity problem is a regular perturbation problem, in the sense that the function spaces X and X^* remains unchanged as ϵ tends to zero, while the perturbed Stokes system (6) is a singular perturbation problem. For such problems we typically need to introduce parameter dependent function spaces in order to obtain parameter independent bounds on the coefficient operator and its inverse, and hence ϵ–independent conditioning of the preconditioned systems, cf. (12).

One way to introduce parameter dependent function spaces, which frequently occur in practice, is to consider weighted sums and intersections of Hilbert spaces. If X and Y are Hilbert spaces, then the intersection $X \cap Y$ is again a Hilbert space with norm

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

Furthermore, if $\epsilon > 0$ is a parameter then the corresponding weighted space, $X \cap \epsilon \cdot Y$, is equal to $X \cap Y$ as a set for $\epsilon > 0$ and to X for $\epsilon = 0$, while the corresponding norms are given by

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

Hence, formally the norms behave continuously as ϵ tends to zero. In Example 4.3 we encountered an example of a space of this form.

Another related parameter dependent space is the weighted sum of Hilbert spaces. In general the space $X+Y$ consists of all elements $z = x+y$, $x \in X$, $y \in Y$, with norm given by

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

while the corresponding weighted space, $X + \epsilon^{-1} \cdot Y$, has the norm

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

If $X \cap Y$ is dense in both X and Y, and $\epsilon > 0$, then

$$(X \cap \epsilon \cdot Y)^* = X^* + \epsilon^{-1} \cdot Y^*,$$

where the star indicates dual spaces. Furthermore, there exist a corresponding Riesz mapping $R_\epsilon : X^* + \epsilon^{-1} \cdot Y^* \to X \cap \epsilon \cdot Y$, such that the operator norms

$$\|R_\epsilon\|_{\mathcal{L}(X^*+\epsilon^{-1}\cdot Y^*, X\cap\epsilon\cdot Y)} \quad \text{and} \quad \|R_\epsilon^{-1}\|_{\mathcal{L}(X\cap\epsilon\cdot Y, X^*+\epsilon^{-1}\cdot Y^*)}$$

are bounded independently of ϵ. In fact, the operator R_ϵ is given by $z \mapsto x = R_\epsilon z$, where $x \in X \cap Y$ solves the problem:

$$\langle x, y\rangle_X + \epsilon^2\langle x, y\rangle_Y = \langle f, y\rangle + \langle g, y\rangle, \quad y \in X \cap Y.$$

Here we have assumed that $z = f + g$ and that $f \in X^*$ and $g \in Y^*$ is chosen such that $\|z\|^2_{X^*+\epsilon^{-1}\cdot Y^*} = \|f\|^2_{X^*} + \epsilon^{-2}\|g\|^2_{Y^*}$. As above, we have used $\langle\cdot,\cdot\rangle$ to denote the proper duality pairings. We refer to [11] for more details on sums and intersections of Hilbert spaces.

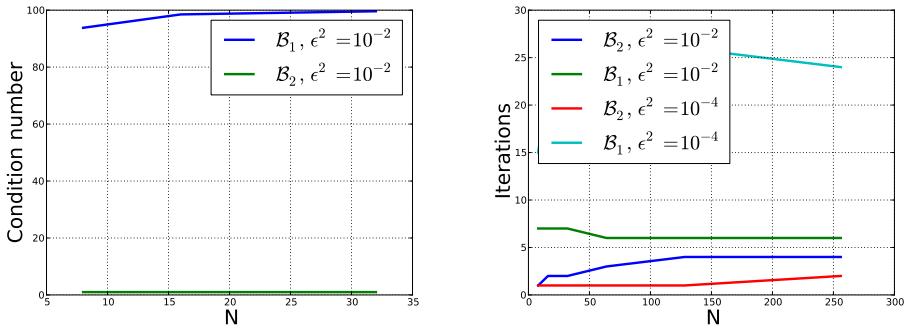

FIGURE 8. The rightmost figure shows the condition number of the preconditioned matrix for different mesh resolutions. The leftmost figure shows the number of iterations required for convergence using CG combined with AMG, where the convergence criterion was a relative reduction of a factor 10^6 of the residual in the discrete L_2 norm.

The significance of the weighted sums and intersections of Hilbert spaces for singular perturbation problem is illustrated by the following simple example.

Example 4.4. *Reaction–diffusion equation.*
Consider the reaction diffusion equation:

$$\begin{aligned} \mathcal{A}_\epsilon u = u - \epsilon^2 \Delta u &= f, \text{ in } \Omega, \\ u &= 0, \text{ on } \partial\Omega. \end{aligned}$$

For each fixed positive ϵ the operator $\mathcal{A}_\epsilon$ is a mapping from $H_0^1(\Omega)$ to $H^{-1}(\Omega)$, but it degenerates to the identity operator as ϵ tends to zero. Therefore, one possible choice for a preconditioner is,

$$\mathcal{B}_1 = -\Delta^{-1}.$$

However, this operator will not be an efficient preconditioner for ϵ close to zero. To obtain a preconditioner which is uniform with respect to ϵ we need to consider $\mathcal{A}_\epsilon$ as an operator from $(L^2 \cap \epsilon \cdot H_0^1)(\Omega)$ to its proper dual space $(L^2 + \epsilon H^{-1})(\Omega)$, with associated Riesz mapping given by

$$\mathcal{B}_2 = (I - \epsilon^2 \Delta)^{-1}.$$

Of course the preconditioned operator $\mathcal{B}_2 \mathcal{A}_\epsilon = I$ and using $\mathcal{B}_2$ involves solving the original reaction-diffusion problem. In practice however we replace $\mathcal{B}_2$ with equivalent and efficient operators, where it is crucial that the equivalence is independent of both ϵ and the characteristic mesh size. In Figure 8 we compute the condition numbers of the corresponding preconditioned discrete systems using continuous piecewise linears finite elements to approximate the solution u. The preconditioners are taken as discrete analogs of $\mathcal{B}_1$ and $\mathcal{B}_2$. We also show the number of iterations required by CG combined with AMG for a given convergence criteria. Clearly, as expected, we observe that the preconditioner $\mathcal{B}_2$ is superior to $\mathcal{B}_1$, in particular when ϵ is small. □

Example 4.5. *Linear elasitiсy in the primal variable.*
In Example 2.2 we saw that a standard multigrid method did not perform as a uniform preconditioner with respect to λ for the linear elasticity problem in primal variables. For $\mu = 1$ and $\epsilon^2 = 1/(1+\lambda)$, this equation can be written as

$$\begin{aligned} \mathcal{A}_\epsilon u = -\operatorname{grad}\operatorname{div} u - \epsilon^2 \Delta u &= \epsilon^2 f, \quad \text{in } \Omega, \\ u &= 0, \quad \text{on } \partial\Omega. \end{aligned}$$

Here the coefficient operator $\mathcal{A}_\epsilon$ can be seen as an isomorphism defined on $H_0(\operatorname{div},\Omega) \cap \epsilon \cdot H_0^1(\Omega)^n$ to its proper dual space. So it easy to identify the proper canonical preconditioner at the continuous level as the operator

$$\mathcal{B}_\epsilon = (-\operatorname{grad}\operatorname{div} -\epsilon^2\Delta)^{-1}.$$

As in Example 4.4 we must replace $\mathcal{B}_\epsilon$ with equivalent but efficient operators. However, standard multigrid algorithms do not produce preconditioners which behave uniformly with respect to ϵ for the most common finite element discretizations of this operator, and this is the reason for the poor performance in Example 2.2. Suitable multigrid methods are however described in [37]. □

Example 4.6. *Time dependent Stokes problem.*
In [32, 33] we showed that

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

was a bounded and continuously invertible linear operator mapping $X_\epsilon = (L^2 \cap \epsilon H^1)(\Omega) \times ((H^1 \cap L_0^2) + \epsilon L_0^2)(\Omega)$ onto its dual space. The preconditioner

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

corresponds to a canonical mapping from the dual space and back to X_ϵ. This explains the efficiency of the preconditioner studied in Example 2.4.

We remark that an alternative preconditioner at the continuous level is

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

This follows since $\mathcal{A}_\epsilon$ can be seen as an isomorphism mapping $H(div,\Omega) \cap \epsilon H^1(\Omega) \times L^2(\Omega)$ onto its dual space, and with appropriate operator norms bounded independently of ϵ. For discussions of discrete analogs of this preconditioner we refer to [37, 9, 10]. □

5. Conclusions

In this paper we have described how preconditioners for systems of partial differential equations can be constructed based on the mapping properties of the differential operator in properly chosen Sobolev spaces. In particular, the canonical preconditioners can be seen as Riesz isomorphisms mapping the space of the right hand side into the solution space. These Riesz isomorphisms are then replaced by equivalent and efficient preconditititioners constructed by using for example multigrid or domain decomposition techniques.

References

1. G. Aguilar, F. Gaspar, F. Lisbona, and C. Rodrigo, *Numerical stabilization of Biots consolidation model by a perturbation on the flow equation*, Int. J. Numer. Meth. Engng **75** (2008), 1282–1300.
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. O. Axelsson and J. Karatson, *Equivalent operator preconditioning for elliptic problems*, Numer. Algor. **50** (2009), 297–380.
5. I. Babuška, *Error bounds for finite element methods*, Numerische Mathematik **16** (1971), 322–333.
6. 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.
7. M. Benzi, G. H. Golub, and J. Liesen, *Numerical solution of saddle point problems*, Acta Numerica **14** (2005), 1–137.
8. 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.
9. M. Benzi and M A. Olshanskii, *An augumented Lagrangian-based approach to the Oseen problem*, SIAM Journal on Scientific Computing **28** (2006), 2095–2113.
10. 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.
11. J. Bergh and J. Löfström, *Interpolation spaces*, Springer-Verlag, 1976.
12. M. A. Biot, *General theory of three-dimensional consolidation*, Journal of Applied Physics **12** (1941), 155–164.
13. D. Braess, *Finite elements; theory, fast solvers, and applications in solid mechanics*, Cambridge University Press, 1997.
14. S. C. Brenner and L. R. Scott, *The mathematical theory of finite element methods*, Text in Applied Mathematics, Springer-Verlag, 1994.
15. F. Brezzi and M. Fortin, *Mixed and hybrid finite element methods*, Springer-Verlag, 1991.
16. J. Cahouet and J.-P. Chabard, *Some fast* 3*D finite element solvers for the generalized Stokes problem*, International Journal for Numerical Methods in Fluids **8** (1988), no. 8, 869–895.
17. H. C. Elman, D. J. Silvester, and A. J. Wathen, *Finite elements and fast iterative solvers: with applications in incompressible fluid dynamics*, Numerical Mathematics and Scientific Computation, Oxford University Press, New York, 2005.
18. M. S. Engelman, R. I. Sani, P. M. Gresho, and M. Bercovier, *Consistent vs. reduced integration penalty methods for incompressible media using several old and new elements*, Int. J. Num. Meth. in Fluids **2** (1982), 25–42.
19. *FEniCS – finite element software package*, URL: *http://www.fenics.org*, 2010.
20. M. A. Heroux, R. A. Bartlett, V. E. Howle, R. J. Hoekstra, J. J. Hu, T. G. Kolda, R. B. Lehoucq, K. R. Long, R. P. Pawlowski, E. T. Phipps, A. G. Salinger, H. K. Thornquist, R. S. Tuminaro, J. M. Willenbring, A. Williams, and K. S. Stanley, *An overview of the trilinos project*, ACM Trans. Math. Softw. **31** (2005), no. 3, 397–423.
21. R. Hiptmair, *Operator preconditioning*, Computers & Mathematics with Applications **52** (2006), no. 5, 699–706.
22. 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.

23. 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.
24. T. J. R. Hughes, W. K. Liu, and A. Brooks, *Finite element analysis of incompressible viscous flows by the penalty function formulation*, J. Comp. Phys. **30** (1979), 1–60.
25. R. C. Kirby, *From functional analysis to iterative methods*, Accepted in SIAM Review (2010).
26. 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).
27. A. Klawonn and G. Starke, *Block triangular preconditioners for nonsymmetric saddle point problems: field-of-values analysis*, Numerische Mathematik **81** (1999), 577594.
28. D. Loghin and A. J. Wathen, *Analysis of preconditioners for saddle-point problems*, SIAM Journal on Scientific Computing **25** (2004), 2029–2049.
29. K.-A. Mardal, *Block preconditioning of systems of PDEs*, Accepted to Automated Scientific Computing (A. Logg, K.-A. Mardal, and G. Wells, eds.), Springer, 2010.
30. 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.
31. 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.
32. K.-A. Mardal and R. Winther, *Uniform preconditioners for the time dependent Stokes problem*, Numerische Mathematik **98** (2004), no. 2, 305–327.
33. _____, *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.
34. _____, *Preconditioning discretizations of systems of partial differential equations*, Accepted to Numerical Linear Algebra with Applications (2010).
35. M. A. Olshanskii and Y. Vassilevski, *Pressure Schur complement preconditioners for the discrete Oseen problem*, SIAM Journal on Scientific Computing **29** (2007), 2686–2704.
36. 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).
37. J. Schöberl, *Multigrid methods for a parameter dependent problem in primal variables*, Numerische Mathematik **84** (1999), 97–119.
38. 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.
39. S. Turek, *Efficient solvers for incompressible flow problems*, Springer, 1999.
40. T. Manteffel V. Faber and S.V. Parter, *On the theory of equivalent operators and application to the numerical solution of uniformly elliptic partial differential equations*, Adv. Appl. Math. **11** (1990), 109–163.

[1] Center for Biomedical Computing, Simula Research Laboratory, P.O. Box 134, 1325 Lysaker, Norway. kent-and@simula.no,, [2] Centre of Mathematics for Applications and Department of Informatics, University of Oslo, 0316 Oslo, Norway, ragnar.winther@cma.uio.no