FROM FUNCTIONAL ANALYSIS TO ITERATIVE METHODS

ROBERT C. KIRBY*

Abstract. We examine condition numbers, preconditioners, and iterative methods for finite element discretizations of coercive PDEs in the context of the fundamental solvability result, the Lax-Milgram Lemma. Working in this Hilbert space context is justified because finite element operators are restrictions of infinite-dimensional Hilbert space operators to finite-dimensional subspaces. Moreover, useful insight is gained as to the relationship between Hilbert space and matrix condition numbers, and translating Hilbert space fixed point iterations into matrix computations provides new ways of motivating and explaining some classic iteration schemes. In this framework, the "simplest" preconditioner for an operator from a Hilbert space into its dual is the Riesz isomorphism. Simple analysis gives spectral bounds and iteration counts bounded independent of the finite element subspaces chosen. Moreover, the abstraction allows us not only to consider Riesz map preconditioning for convection-diffusion equations in H^1, but also operators on other Hilbert spaces, such as planar elasticity in $\left(H^1\right)^2$.

Key words. preconditioner, functional analysis, iterative methods

AMS subject classifications. 65N30, 65N22

1. Introduction. The Lax-Milgram lemma [36] has played a critical role over the last half century by establishing existence and uniqueness of weak solutions of operator equations $Au = f$ where A is a continuous and coercive linear operator from a Hilbert space V into its topological dual V'. The result has had far-reaching impact on the analytic and numerical study of elliptic and parabolic partial differential equations; boundedness and coercivity of the weak operators are all that need be verified.

The Riesz map in functional analysis is used in establishing Lax-Milgram and will be central to this paper. It identifies linear functionals acting on a Hilbert space with members of the Hilbert space itself in a norm-preserving way. Members of Euclidean space are frequently denoted by column vectors, while linear functionals acting on these are denoted by row vectors. The Riesz map in this case is simply transposition. In function spaces, however, this identification is more complicated and may use differential operators. Incorporating this information into linear algebraic computations turns out to be useful in obtaining effective preconditioners.

While the 1954 paper of Lax and Milgram provided a nonconstructive proof of the result, constructive proofs appeared over the next several years. In 1960, Zarantanello [54] provided a generalization of Lax-Milgram to strongly monotone nonlinear operators via Banach fixed points. A similar fixed point proof in the linear case was given by Lions and Stampacchia in a 1967 paper on variational inequalities [37]. In both cases, the mapping shown to be contractive involves applying the Riesz map for V to a residual $Av - f \in V'$ for some $v \in V$. Also in 1967, Petryshyn gave a constructive proof of the result based on so-called upper and lower semi-orthogonal bases for subspaces of V. This result is more difficult to turn into iterative methods, and we shall follow the approach of Zarantanello/Lions/Stampacchia.

The discrete operators developed in conforming finite element methods are restrictions of weak differential operators to finite-dimensional subspaces of the under-

*Department of Mathematics and Statistics, Texas Tech University, Lubbock, TX 79409-1042. The author was supported by the US Department of Energy under grant DE-FG02-07ER25821. The author expresses his gratitude to Prof. Andy Wathen for his considerable input on the manuscript and pertinent literature.

lying Hilbert space. While this perspective led to early analysis of finite elements, such as is seen in the books of Strang and Fix [50] and Ciarlet [14], we here use it to study iterative methods. After developing notions of conditioning appropriate to Hilbert space and the linear algebraic representation of finite element operators, we study the Lax-Milgram fixed point iteration as a preconditioned stationary iteration. This starting point can actually give iterative methods with mesh-independent convergence rates, if the Riesz map can be computed efficiently. Moreover, we apply the ideas developed in stationary iterations to use Riesz maps as preconditioners in Krylov methods for convection-diffusion and elasticity equations. The theory also informs problems with mesh-dependent bilinear forms, although the bounds may have some mesh dependence,

Using Hilbert space to derive and analyze the implied matrix computations sits between two common approaches to iterative methods. Many approaches to iterative methods start from the linear algebraic system of equations, given in matrix form. We include the classic work on splitting and relaxation methods of Young [53], the original work on conjugate gradients by Hestenes and Stiefel [32], and the alternating direction implicit methods of Peaceman and Rachford [44] greatly generalized by Rachford, Douglas, and Gunn [19, 17, 18, 26, 25, 27]. It must be said that many of these techniques, at least implicitly, succeed by incorporating analytic information about elliptic equations. For example, the A-norm used in conjugate gradients for second order elliptic PDE is comparable to the H^1 norm. Still, in these and many other papers, the mathematical techniques derive primarily from linear algebra. Starting from Hilbert space makes the analytic structure explicit and the proofs of some known results much shorter. It also can be applied to coercive operators in other contexts besides H^1, and will tie the behavior of iterative methods to the infinite-dimensional continuity and coercivity constants.

Many authors have generalized iterative techniques to Hilbert space and other possibly infinite-dimensional settings. For example, conjugate gradients were considered in Hilbert space by Hayes [29] and Daniel [15] and more recently by Axelsson and Karátson [7, 8]. Kellogg analyzed ADI methods with "mesh spacing zero" [35]. Several papers [45, 41, 28] generalize the notion of splitting methods and overrelaxation to operator equations in Hilbert space. Nevanlinna [42] has written a book considering many iterative methods in the context of function spaces. In the multigrid literature, the early paper of Bank and Dupont [9] put known finite difference technology in an abstract framework suitable for finite elements. This has led to a large literature on the subject, including cascadic multigrid methods [11]. Our goal here is to see how such generalizations can feed back into solving finite element equations. Such an idea is not new; Axelsson and Karátson apply their nonsymmetric Hilbert space techniques to finite element convection-diffusion equations.

The work of Axelsson and Karátson cited above and the work of Manteuffel *et al.* on clusting singular values of differential operators preconditioned by differential operators [23] are very similar to the present paper, but there are two significant differences. First, these papers focus on either bounded operators from a Hilbert space to itself, or on densely defined unbounded operators. In Lax-Milgram, and hence in studying finite element methods, unbounded weak operators are naturally interpreted as *bounded* operators from a Hilbert space into its dual. This characterization will allow precise statements to be made in terms of the continuity and coercivity constants. Second, starting a discussion of iterative methods from the Lax-Milgram fixed point iteration provides a very different perspective that roots preconditioners and iterative

methods in the fundamental results of functional analysis.

The Lax-Milgram result relies heavily on the Riesz Representation Theorem. This theorem states that a Hilbert space V and its dual V' are isometrically isomorphic. This implies a mapping $\tau : V' \to V$ that we refer to as the *Riesz map*. We will state precisely what this map is in the next section. The fixed point iteration used in proving Lax-Milgram constructively requires the evaluation of this Riesz map at each step. Like most relaxation schemes, the iteration may be recast as a splitting method, but one based on *analytic* considerations (the function space setting) rather than *algebraic* considerations (the structure of a matrix). Rather than a diagonal or lower triangular part of a matrix, we will consider splitting a matrix representation of the Riesz map from the rest of the matrix. The rather abstract setting is useful, as it can lead to simple proofs of good convergence rate bounds for these relaxation schemes and also provides insight into methods with mesh-dependent variational forms. Relaxation methods are also referred to as preconditioned stationary iterations, in which some "simple" operator is inverted onto the residual at each step. As the Lax-Milgram fixed point iteration applies the Riesz map at each step, this suggests that Riesz maps can serve as preconditioners for other operators in the Hilbert space. Concretely, this motivates the technique of preconditioning some elliptic operator in H^1 with the inverse Laplace or Helmholtz operator and also simplifies the analysis of this. Moreover, it provides a framework to begin considering coercive operators in other Hilbert spaces, such as planar elasticity in $(H^1)^2$.

It is natural, then, to consider using Riesz maps, if they may be efficiently computed, as preconditioners for other iterative methods such as conjugate gradients. Linear algebra proofs for the effectiveness of preconditioners for CG can be fairly involved, as they require getting good bounds on the spectrum of the preconditioned matrix. However, working in a more abstract context, such as the Hilbert space formulation of Daniel [15] greatly simplifies analysis. The key is to focus on the *types* of the operators being inverted. In particular, Lax-Milgram provides unique weak solutions to operators from a Hilbert space V into its dual V'. However, CG requires mappings from V onto itself. In order to apply the conjugate gradient method formulated on V rather than on some matrix representation in Euclidean space, an operator $A : V \to V'$ must be composed with some other operator $B : V' \to V$ so that $BA : V \to V$ will "type check", to use the terminology of programming languages. Alternately, one could post-multiply A by some $B : V' \to V$ so that $AB : V' \to V'$ is a bounded operator on the dual space. Once the basic abstractions of functional analysis are absorbed, obtaining mesh-independent bounds on iteration counts for conforming methods from the basic theory is straightforward. Similar distinctions between Hilbert spaces and their duals in the context of preconditioning appear in various papers on mixed finite element methods such as [5], but this presentation makes more explicit connections to the Riesz Representation Theorem and the Lax-Milgram constants.

This paper is primarily expository in nature, intending to bridge between the functional analysis techniques common in finite elements and the numerical linear algebra community. Many of these results will be familiar to experts at the interface of finite elements and iterative methods, but it is difficult to find them in written form. It is hoped that this work will bring these ideas to a broader community.

2. Preliminaries. Throughout, let V denote a real Hilbert space equipped with inner product $(\cdot, \cdot)_V$ and associated norm $\|\cdot\|_V$. V' denotes the topological dual with $\langle \cdot, \cdot \rangle$ the $V' \times V$ duality pairing. For any $f \in V'$, we have the standard dual norm

$\|f\|_{V'} = \sup_{\substack{v\in V \\ \|v\|_V=1}} |\langle f, v\rangle|$.

For Hilbert spaces V_1 and V_2, we will also need the (Banach) space of bounded linear maps between V_1 and V_2, denoted $\mathcal{L}(V_1, V_2)$, with its associated norm $\|T\|_{\mathcal{L}(V_1,V_2)} = \sup_{\substack{v\in V_1 \\ \|v\|_{V_1}=1}} \|Tv\|_{V_2}$.

Let $L^2(\Omega)$ be the standard space of square-integrable functions over a domain $\Omega \subset \mathbb{R}^n$. The inner product on L^2 is given by

$$(u, v)_{L^2} = \int_\Omega uv \,\mathrm{dx}, \tag{2.1}$$

with associated norm $\|u\|^2_{L^2} = (u, u)$.

Denote by H^1 be that subspace of L^2 whose members have square-integrable weak derivatives. The inner product is given by

$$(u, v)_{H^1} = \int_\Omega \nabla u \cdot \nabla v + uv \,\mathrm{dx}, \tag{2.2}$$

and associated norm $\|u\|^2_{H^1} = (u, u)_{H^1}$. The H^1 seminorm is given by

$$|u|^2_{H^1} = \int_\Omega \nabla u \cdot \nabla u \,\mathrm{dx}. \tag{2.3}$$

The Poincaré-Friedrichs inequality [12] states that if Γ is a subset of the boundary of Ω with positive measure and $V \subset H^1$ is a space of functions that vanish on Γ, then in fact the H^1 seminorm defines a norm on V. The result states that there exists some constant C_P such that

$$\|u\|_{L^2} \le C_P |u|_{H^1} \tag{2.4}$$

for all $u \in V$. The constant C_P depends on the size (appropriately defined) of the domain and the size of Γ relative to the whole boundary $\partial\Omega$. This result also implies that $(u, v) = \int_\Omega \nabla u \cdot \nabla v \,\mathrm{dx}$ defines an inner product on V that is equivalent to the standard H^1 inner product.

The Riesz Representation Theorem [47] establishes an isometric isomorphism between a Hilbert space V and its dual V', which we denote $\tau : V' \to V$. That is, for $f \in V'$, τf satisfies

$$(\tau f, v) = \langle f, v\rangle \tag{2.5}$$

for all $v \in V$. Moreover, $\|\tau f\|_V = \|f\|_{V'}$.

The space $H^{-1}(\Omega)$ is the dual space of $H^1_0(\Omega)$, which consists of those H^1 functions vanishing on the boundary of Ω. That is, it is the space of continuous linear functionals acting on $H^1_0(\Omega)$. In this case, the Riesz map $\tau : H^{-1}(\Omega) \to H^1_0(\Omega)$, is given by

$$\langle f, v\rangle = (\tau f, v) = \int_\Omega \nabla(\tau f) \cdot \nabla v + (\tau f)\, v \,\mathrm{dx} \tag{2.6}$$

for all $v \in H^1_0(\Omega)$. This is simply a weak form of the Helmholtz equation

$$-\Delta(\tau f) + (\tau f) = f. \tag{2.7}$$

In light of the Poincaré-Friedrichs inequality (2.4), if $V = \{v \in H^1(\Omega) : v|_\Gamma = 0\}$ for some nontrivial $\Gamma \subset \partial\Omega$, an equivalent Riesz map is

$$\langle f, v\rangle = (\tau f, v) = \int_\Omega \nabla(\tau f) \cdot \nabla v \, \mathrm{dx}, \tag{2.8}$$

which is a weak form of the Poisson equation

$$-\Delta(\tau f) = f. \tag{2.9}$$

Weak PDE and finite element methods are often written compactly in terms of bilinear forms. A bilinear form $a : V \times V \to \mathbb{R}$ is *continuous* if there exists $0 \le C < \infty$ with

$$C = \sup_{\substack{u,v \in V \\ \|u\|_V = \|v\|_V = 1}} |a(u, v)|, \tag{2.10}$$

and *coercive* if there exists $0 < \alpha < \infty$ such that

$$\alpha = \inf_{\substack{u \in V \\ \|u\|_V = 1}} a(u, u) \tag{2.11}$$

Note that we are defining C and α to be the "best" constants via the supremum and infimum. This slight increase in precision over the standard definitions of continuity and coercivity will be important in quantifying the behavior of iterative methods.

There is a natural equivalence between continuous bilinear forms on $V \times V$ and continuous linear operators from V to V'. If a is a continuous bilinear form, then a continuous operator $A : V \to V'$ can defined by

$$\langle Au, v\rangle = a(u, v), \quad v \in V. \tag{2.12}$$

Note that A must be continuous if a is; its norm in $\mathcal{L}(V, V')$ is C from (2.10). Similarly, $\langle Au, u\rangle \ge \alpha\|u\|_V^2$. The operator A is said to be coercive and continuous if the bilinear form a is. Of course, operators can also be used to define bilinear forms.

We state the famous result here.

THEOREM 2.1 (Lax-Milgram). *Let $A \in \mathcal{L}(V, V')$ be a continuous and coercive operator with constants C and α, respectively. Then, for any $f \in V'$, there exists a unique $u \in V$ such that*

$$\langle Au, v\rangle = \langle f, v\rangle \tag{2.13}$$

for all $v \in V$. Moreoever, the u depends continuously on f, in that

$$\|u\|_V \le \frac{1}{\alpha}\|f\|_{V'}. \tag{2.14}$$

Galerkin finite element methods succinctly define approximations by restricting the operator A on V to some finite-dimensional V_h. Such a subspace typically consists of continuous piecewise polynomial functions defined over a mesh, which we will assume comes from a quasiuniform family [13, p. 106]. Other methods, such as the streamline-diffusion method for convection-diffusion, do not come from simple restrictions, but our Hilbert space context also provides useful insight in this case as well.

If $\{\phi_i\}_{i=1}^{\dim V_h}$ is a basis for V_h, then one writes a function $u_h \in V_h$ as

$$u_h = \sum_{i=1}^{\dim V_h} \mathrm{u}_i \phi_i.$$

The vector $\mathrm{u} \in \mathbb{R}^{\dim V_h}$ of expansion coefficients is typeset with Roman type. Throughout, we will use the convention of Roman type to refer to vectors and matrices and italic type to refer to members of and operators on abstract Hilbert spaces.

The discrete problem is to find some $u_h \in V_h$ such that

$$\langle Au_h, v_h \rangle = a(u_h, v_h) = \langle f, v_h \rangle \tag{2.15}$$

for all $v_h \in V_h$.

The vector $\mathrm{u} \in \mathbb{R}^{\dim V_h}$ of coefficients of the solution of (2.15) satisfies the algebraic equations

$$\sum_{j=1}^{\dim V_h} \mathrm{A}_{i,j} \mathrm{u}_j = \mathrm{f}_i, \quad 1 \le i \le \dim V_h, \tag{2.16}$$

where the matrix

$$\mathrm{A}_{i,j} = a(\phi_j, \phi_i) \tag{2.17}$$

is typically called the *stiffness matrix* and the vector

$$\mathrm{f}_i = \langle f, \phi_i \rangle \tag{2.18}$$

is typically called the *load vector*.

By defining the finite element method as a restriction, the discrete operators inherit the analytic structure of the underlying PDE in Hilbert space. If we restrict a continuous and coercive operator $A : V \to V'$ to some subspace $V_h \subset V$, we obtain a continuous and coercive operator from $V_h \to (V_h)'$. To see this,

$$C_h = \sup_{\substack{u,v \in V_h \\ \|u\|=\|v\|=1}} \langle Au, v \rangle \le \sup_{\substack{u,v \in V \\ \|u\|=\|v\|=1}} \langle Au, v \rangle = C \tag{2.19}$$

$$\alpha_h = \inf_{\substack{u \in V_h \\ \|u\|=1}} \langle Au, u \rangle \ge \inf_{\substack{u \in V \\ \|u\|=1}} \langle Au, u \rangle = \alpha \tag{2.20}$$

In particular, for standard conforming methods, the continuity and coercivity constants for the finite-dimensional operators are bounded independently of the subspace chosen, which means they are bounded independently of the polynomial degree and mesh spacing. The continuity and coercivity constants on the discrete spaces are related to, but different than, standard matrix properties like norms and positive-definiteness. We shall describe these connections later.

Some finite element methods do not strictly follow the Galerkin formalism of restricting the weak operator to a finite-dimensional space. For nonconforming methods where $V_h \not\subset V$ or problems with mesh-dependent operators $A_h \neq A|_{V_h}$ such as streamline diffusion, the definitions of C_h and α_h are still valid, but their relationship to the infinite-dimensional constants C and α must be more carefully studied. However, these mesh-dependent constants still guide the behavior of iterative methods.

2.1. Examples. We continue with two example problems, including their finite element formulations. First, consider a linear convection-diffusion equation posed over the unit square

$$-\epsilon\Delta u + \beta \cdot \nabla u = f, \tag{2.21}$$

where β and f can vary spatially. At each point, the vector β is assumed divergence-free. While ϵ may also vary spatially, for simplicity we assume it constant. The boundary $\partial\Omega$ is partitioned into $\Gamma_D \cup \Gamma_N$, on which Dirichlet and Neumann boundary conditions will be imposed. Another important characterization of the boundary is based on the sign of $\beta \cdot n$, where n is the unit outward normal to the domain. $\partial\Omega$ may also be partitioned as $\Gamma_+ \cup \Gamma_- \cup \Gamma_0$, the portions of the boundary on which $\beta \cdot n$ is positive, negative, or zero. These are respectively referred to as the *outflow*, *inflow*, and *characteristic* boundaries. We pose boundary conditions

$$\begin{aligned} u &= 0, \quad \Gamma_D \\ \frac{\partial u}{\partial n} &= 0, \quad \Gamma_N \end{aligned} \tag{2.22}$$

Let $V^0 \subset H^1$ be the set of all functions vanishing on Γ_D. Typically, one has inhomogeneous Dirichlet boundary conditions and solves the problem in a set of functions matching the Dirichlet boundary conditions, denoted V^D. This set is closed under neither addition nor scalar multiplication. The problem is officially set in a Hilbert space by picking any known function $u^D \in V^D$, defining $V^0 \ni u^0 = u - u^D$ and solving the modified problem $a(u^0, v) = \langle f, v\rangle - a(u^D, v) \equiv \langle f^0, v\rangle$ for u^0. Without loss of generality, then, we will analyze the case of homogeneous Dirichlet data and not dwell further on this technicality.

The variational problem is to find $u \in V^0$ such that

$$a(u, v) = \langle f, v\rangle, \tag{2.23}$$

for all $v \in V^0$, where

$$a(u, v) = \int_\Omega \epsilon\nabla u \cdot \nabla v + (\beta \cdot \nabla u)v \,\mathrm{dx} \tag{2.24}$$

and

$$\langle f, v\rangle = \int_\Omega fv \,\mathrm{dx} - \int_{\Gamma_N} g\frac{\partial v}{\partial n} \,\mathrm{ds} \tag{2.25}$$

The continuity of this bilinear form is easy to establish using the Cauchy-Schwarz and Poincaré-Friedrichs inequalities, for

$$\begin{aligned} \left|\int_\Omega \epsilon\nabla u \cdot \nabla v + (\beta \cdot \nabla u)v \,\mathrm{dx}\right| &\le \epsilon\|\nabla u\|_{L^2}\|\nabla v\|_{L^2} + |\beta|\|\nabla u\|_{L^2}\|v\|_{L^2} \\ &\le \left(\epsilon^2 + C_P^2|\beta|^2\right)^{\frac{1}{2}} |u|_{H^1}|v|_{H^1}, \end{aligned} \tag{2.26}$$

where C_P is the constant in 2.4.

Establishing the coercivity of a is a bit more delicate. As Brenner and Scott point out [13], with general boundary conditions this form need not be coercive for large enough β. On the other hand, the analysis of Elman, Silvester, and Wathen [21] shows

that if Neumann boundary conditions are only imposed on outflow or characteristic boundaries, the term $\int_\Omega (\beta \cdot \nabla u) v \, \mathrm{dx}$ is skew-symmetric with possibly a small positive-definite perturbation and hence does not degrade coercivity. In this case, one can then show

$$a(u,u) \geq \epsilon \|\nabla u\|_{L^2}^2 = \epsilon |u|_{H^1}. \tag{2.27}$$

Hence, in this case the operator is coercive, but the estimate degrades as the diffusivity $\epsilon \searrow 0$.

The standard Galerkin method is well-known to behave poorly in this case. A very fine mesh is required to resolve features, and the analysis of this is quite technical in the noncoercive case [13]. An alternate formulation of this problem with superior stability properties is the streamline-diffusion method of Hughes and Brooks [34]. This method solves a different, mesh-dependent variational problem that adds some artificial diffusion along the streamlines of the flow.

We consider only the piecewise-linear case, for which the variational problem is to find $u \in V_h \subset V^0$ such that $a_\delta^{SD}(u,v) = \langle f_\delta^{SD}, v \rangle$ for all $v \in V_h \subset V^0$, where

$$\begin{aligned} a_\delta^{SD}(u,v) &= \int_\Omega \epsilon \nabla u \cdot \nabla v + (\beta \cdot \nabla u) v \, \mathrm{dx} + \int_\Omega \delta (\beta \cdot \nabla u)(\beta \cdot \nabla v) \, \mathrm{dx} \\ \langle f_\delta^{SD}, v \rangle &= \int_\Omega f v \, \mathrm{dx} + \int_\Omega \delta f (\beta \cdot \nabla v) \, \mathrm{ds} \end{aligned} \tag{2.28}$$

Here, δ is a parameter controlling the additional term, many forms for which have been suggested in the literature. We follow Eriksson *et al.* [22] and choose the simple form

$$\delta = \frac{h}{2|\beta|}, \tag{2.29}$$

which is the optimal selection in the case of vanishing ϵ. Typically, the streamline-diffusion method operates in the regime where the cell Peclet number $P_h = \frac{h|\beta|}{2\epsilon}$ is greater than one, and production codes turn off δ when $P_h < 1$, so that the method locally reduces to standard Galerkin. We will perform our analysis assuming that $P_h > 1$.

While similar calculations as for the standard Galerkin method allow one to show H^1 continuity and coercivity constants, one may also introduce an alternate norm

$$\|v\|_{SD}^2 \equiv \epsilon \|u\|_{H^1}^2 + \delta \|\beta \cdot \nabla u\|_{L^2}^2. \tag{2.30}$$

This norm comes from the inner product

$$(u,v)_{SD} \equiv \epsilon \int_\Omega \nabla u \cdot \nabla v \, \mathrm{dx} + \delta \int_\Omega (\beta \cdot \nabla u)(\beta \cdot \nabla v) \, \mathrm{dx}, \tag{2.31}$$

which can be seen as the weak form of an elliptic operator with a certain tensor-valued diffusion coefficient [34]. Since the streamline-diffusion operator is continuous and coercive in this inner product, we may use it as a Riesz map. Using this inner product, we can rewrite the bilinear form compactly as

$$a_\delta^{SD}(u,v) = (u,v)_{SD} + \int_\Omega (\beta \cdot \nabla u) \, v \, \mathrm{dx}. \tag{2.32}$$

The coercivity estimate,

$$a_\delta^{SD}(u,u) \geq \|u\|_{SD}^2, \tag{2.33}$$

holds with $\alpha = 1$. For standard Galerkin, the continuity constant is independent of ϵ and the coercivity constant degrades. Because of the scaling in the streamline-diffusion norm, the coercivity constant is independent of ϵ while the continuity constant degrades. To see this, the Cauchy-Schwarz inequality gives

$$\begin{aligned} a_\delta^{SD}(u,v) &= (u,v)_{SD} + \int_\Omega (\beta \cdot \nabla u)\, v \,\mathrm{dx} \\ &\leq \|u\|_{SD}\|v\|_{SD} + \|\beta \cdot \nabla u\|_{L^2}\|v\|_{L^2}. \end{aligned} \tag{2.34}$$

Now, each of the two L^2 norms may be bounded in terms of the streamline-diffusion norm, for

$$\|\beta \cdot \nabla u\|_{L^2} = \frac{1}{\sqrt{\delta}}\sqrt{\delta}\|\beta \cdot \nabla u\|_{L^2} \leq \frac{1}{\sqrt{\delta}}\|u\|_{SD}, \tag{2.35}$$

and similarly,

$$\|v\|_{L^2} \leq C_P\|\nabla v\|_{L^2} = \frac{C_P}{\sqrt{\epsilon}}\sqrt{\epsilon}\|\nabla v\|_{L^2} \leq \frac{C_P}{\sqrt{\epsilon}}\|v\|_{SD}. \tag{2.36}$$

Combining these gives the bound

$$a_\delta^{SD}(u,v) \leq \left(1 + \frac{C_P}{\sqrt{\epsilon\delta}}\right)\|u\|_{SD}\|v\|_{SD}. \tag{2.37}$$

Comparing the estimates for coercivity and continuity constants for the two convection-diffusion discretizations turns out to shed light on the relative conditioning of the methods, which is important for the behavior of iterative methods. We shall return to these bounds in the following subsection.

For a second example problem, we consider conforming Galerkin discretizations of planar elasticity. The variational form is symmetric and gives rise to symmetric matrices, but requires a vector-valued space of functions. Our use of the Riesz map is applicable in this case, also. Simple models of planar elasticity give rise to coercive operators on $(H^1)^2$, the space of vector-valued functions whose components are in H^1. The weak form of elasticity is to find u in $(H^1)^2$ satisfying boundary conditions such that

$$2\mu\int_\Omega \epsilon(u) : \epsilon(v)\,\mathrm{dx} + \lambda\int_\Omega (\nabla \cdot u)(\nabla \cdot v)\,\mathrm{dx} = \int_\Omega f \cdot v\,\mathrm{dx} + \int_{\Gamma_N} (\sigma(u)n)v\,\mathrm{ds} \tag{2.38}$$

for all v.

The notation $\epsilon(u)$ refers to the symmetric part of the gradient tensor of u. The *stress* is $\sigma = 2\mu\epsilon(u) + \lambda \mathrm{tr}\epsilon(u)I$, where tr denotes the trace of a tensor and I is the identity. Here, the constants μ, λ are called the Lamé constants. These are related to the Young's modulus E and Poisson ratio ν by

$$E = \frac{4\mu\,(\mu+\lambda)}{2\mu+\lambda}, \qquad \nu = \frac{\lambda}{2\mu+\lambda}. \tag{2.39}$$

In the limiting case of $\nu \nearrow 0.5$, the material becomes nearly incompressible, leading to the phenomenon known as locking. Standard Galerkin discretizations perform poorly (this is actually due to a large Hilbert space condition number), and one must turn to nonconforming or mixed methods to handle the situation.

Coercivity of the weak form on $(H^1)^2$ is a result of Korn's inequality [12, 13]. Because of Poincaré-Friedrichs, we use the H^1 seminorm as a norm and the inverse vector Laplacian as the Riesz map.

2.2. Condition number. Following the standard definition for matrices, we define condition numbers of operators between Hilbert spaces.

DEFINITION 2.2. *Let $A \in \mathcal{L}(V_1, V_2)$ have a bounded inverse. The* condition number *of A is*

$$\kappa(A) = \|A\|_{\mathcal{L}(V_1,V_2)}\|A^{-1}\|_{\mathcal{L}(V_2,V_1)} \tag{2.40}$$

THEOREM 2.3. *Let $A \in \mathcal{L}(V, V')$ be a continuous and coercive operator with constants C and α. Then*

$$\kappa(A) \leq \frac{C}{\alpha} \tag{2.41}$$

Proof. Continuity gives $\|A\|_{\mathcal{L}(V,V')} = C$. The continuous dependence result (2.14) from Lax-Milgram gives $\|A^{-1}\|_{\mathcal{L}(V',V)} \leq \frac{1}{\alpha}$. □

COROLLARY 2.4. *Let a be a continuous and coercive bilinear form with constants C and α and let A be the associated operator from V to V'. For any $V_h \subset V$, let A_h be A restricted to V_h. Then, the condition number of $A_h \in \mathcal{L}(V_h, V_h')$ is bounded as*

$$\kappa(A_h) \leq \frac{C}{\alpha} \tag{2.42}$$

Proof. This is a simple consequence of the bounds (2.19) and (2.20) and the previous result. □

REMARK 2.1. *This result applies to the Galerkin methods for convection-diffusion and elasticity, but not the streamline-diffusion method. The variational operator for streamline-diffusion does not satisfy the property that $A_h = A|_{V_h}$. However, the condition number bound of $\frac{C_h}{\alpha_h}$ still makes sense in the streamline-diffusion norm on each mesh.*

Note that in the standard Galerkin case, the condition number bound in $\mathcal{L}(V_h, (V_h)')$ does not depend on the size of the mesh partition or kind of finite element basis used, unlike when the condition number of the stiffness matrix is measured in Euclidean norms.

Note also that in $\mathcal{L}(H^1, (H^1)')$, the condition number of the weak Helmholtz operator is one. To see this, the weak operator $a(u, v) = \int_\Omega \nabla u \cdot \nabla v + uv \, \mathrm{dx}$ is precisely the H^1 inner product. Trivially, $a(u, v) \leq \|u\|_{H^1}\|v\|_{H^1}$ because of the Cauchy-Schwarz inequality so that $C = 1$, and $a(u, u) = \|u\|^2$ so that $\alpha = 1$.

When using the H^1 seminorm as a norm on some $V^0 \subset H^1$ where functions are required to vanish on a boundary segment, the Laplace operator with these boundary conditions has unit condition number in $\mathcal{L}(V^0, (V^0)')$. Condition numbers frequently are a proxy for the difficulty an operator poses for iterative methods. Because algorithms such as multigrid can efficiently solve Helmholtz and Laplace operators, it

makes sense to renormalize other operators by measuring their difficulty relative to these.

Now, we return to the analysis of the continuity and coercivity constants for the convection-diffusion methods. From (2.26) and (2.37), the H^1 condition number bound for standard Galerkin is

$$\kappa_{SG} \leq \frac{C}{\alpha} = \frac{\left(\epsilon^2 + C_P^2|\beta|^2\right)^{\frac{1}{2}}}{\epsilon}, \tag{2.43}$$

while for streamline diffusion, we have the bound in the modified norm of

$$\kappa_{SD} \leq \frac{C_h}{\alpha_h} = 1 + \frac{C_P}{\sqrt{\epsilon\delta}} \tag{2.44}$$

In (2.43), we see that the condition number bound for standard Galerkin scales like ϵ^{-1} for small ϵ. In (2.44), however, the streamline-diffusion condition number on a fixed mesh scales only like $\epsilon^{-\frac{1}{2}}$. Similarly, for fixed ϵ, the bound (2.43) scales linearly with the size of the convective velocity $|\beta|$. Using the simple δ given in (2.29) gives a streamline-diffusion bound that scales only with square root of $|\beta|$. On the other hand, the mesh size h appears in the streamline diffusion estimate through δ so that fixing ϵ, β, the condition number scales like $h^{-\frac{1}{2}}$. However, the point of streamline-diffusion is to stabilize relatively large h, and the stabilization is turned off for small cell Peclet numbers in practice.

2.3. Linear algebraic representations. Practical computation requires some simple representations of the functions and operators being considered, typically using vectors and matrices. In this section, we make explicit the connection between linear algebraic objects and members of finite element spaces.

The algebraic system (2.16) exactly encodes the discrete solution, but there are many possible interpretations of the equations. When an iterative method for solving the algebraic system is considered, an appropriate topology must be chosen to measure convergence. For example, one may work in the Euclidean topology, or else one may use an inner product topology that builds in knowledge of the underlying Hilbert spaces. After all, the Hilbert space settings emerge from physical considerations, and the natural norms typically have important application-specific meaning, such as the square root of the system's energy.

Consider the mapping $\mathcal{I}_h : \mathbb{R}^{\dim V_h} \to V_h$ defined by

$$\mathcal{I}_h \mathrm{u} = \sum_{i=1}^{\dim V_h} \mathrm{u}_i \phi_i \tag{2.45}$$

This mapping exactly represents what happens on a computer when vectors are used to represent finite element functions and also viewed simply as vectors. Note that $\mathcal{I}_h$ is trivially invertible, and $\mathcal{I}_h^{-1}$ interprets a function in V_h as its vector of coefficients. This operation is purely "psychological" in terms of a computer algorithm; it involves no actual computation and can be thought of as type casting in a programming language.

Similarly, one can view vectors as linear functionals by a mapping $\mathcal{I}_h'$, where

$$\langle \mathcal{I}_h' \mathrm{f}, u \rangle = \sum_{i=1}^{\dim V_h} \mathrm{f}_i \left(\mathcal{I}_h^{-1} u\right)_i = \mathrm{f}^t \left(\mathcal{I}_h^{-1} u\right) \tag{2.46}$$

$$\begin{array}{ccc} \mathbb{R}^{\dim V_h} & \xrightarrow{\mathcal{I}_h} & V_h \\ \downarrow \mathrm{A} & & \downarrow A_h \\ \mathbb{R}^{\dim V_h} & \xrightarrow{\mathcal{I}_h'} & V_h' \end{array}$$

FIG. 2.1. *Commuting diagram relating Euclidean space, the finite element space V_h, the Hilbert space operator A_h, and the stiffness matrix* A *via interpretation operators given in (2.45) and (2.46)*

That is, a vector is considered as a linear functional on V_h by computing its dot product with the vector of coefficients of the input function.

LEMMA 2.5. *The adjoint of $\mathcal{I}_h'$ is $(\mathcal{I}_h)^{-1}$.*

Proof. $\mathcal{I}_h \in \mathcal{L}(\mathbb{R}^{\dim V_h}, V_h)$, so $\mathcal{I}_h^* \in \mathcal{L}(V_h', \mathbb{R}^{\dim V_h})$ (we can identify Euclidean space with its dual isometrically in the trivial way). In light of the definition of $\mathcal{I}_h'$, the result is obvious. □

With these operators in hand, we may relate the stiffness matrix and its application to a vector to the Hilbert space setting. The linear algebraic representations are really just encodings of the Hilbert space members and operations.

PROPOSITION 2.6. *Let $u \in V_h$ with $u = \sum_{i=1}^{\dim V_h} \mathrm{u}_i \phi_i$, where $\mathrm{u} = \mathcal{I}_h^{-1} u$. Then $A_h u \in V_h'$ satisfies*

$$A_h u = \mathcal{I}_h' (\mathrm{Au}) \tag{2.47}$$

Proof. Let $v = \sum_{i=1}^{\dim V_h} \mathrm{v}_i \phi_i \in V_h$. Then

$$\begin{aligned} \langle Au, v \rangle = a(u,v) &= a\left(\sum_{i=1}^{\dim V_h} \mathrm{u}_i \phi_i, \sum_{j=1}^{\dim V_h} \mathrm{v}_j \phi_j \right) \\ &= \sum_{i,j=1}^{\dim V_h} \mathrm{u}_i \mathrm{v}_j a(\phi_i, \phi_j) = \sum_{i,j=1}^{\dim V_h} \mathrm{u}_i \mathrm{v}_j \mathrm{A}_{i,j} = (\mathrm{Au})^t \, \mathrm{v}, \end{aligned} \tag{2.48}$$

whence (2.47) follows by multiplying through by $\mathcal{I}_h'$. □

This implies a simple relationship between the stiffness matrix and the operator on V_h, namely

$$\mathrm{A} = (\mathcal{I}_h')^{-1} A_h \mathcal{I}_h. \tag{2.49}$$

Equivalently, one could write $\mathcal{I}_h' \mathrm{A} = A_h \mathcal{I}_h$, in which the interpretation operators for V_h and its dual, the matrix, and the Hilbert space operator can be placed in a commuting diagram, as seen in Figure 2.1.

This discussion of the duality between functions in a finite-dimensional Hilbert space and their vector representations holds in a fairly general setting. However, the following analysis of conditioning and iterative methods based on Riesz maps and the Lax-Milgram theory has elliptic equations particularly in mind.

The Riesz map plays a fundamental role in our iterative methods, so we provide a matrix representation of it here. On $V^0 \subset H^1$, it is just the inverse of the weak

Laplace operator. For a basis $\{\phi_i\}_{i=1}^{\dim V_h}$, the Riesz map is the inverse of the matrix defined by

$$\mathrm{H}_{i,j} = \int_\Omega \nabla\phi_i \cdot \nabla\phi_j \,\mathrm{dx}. \tag{2.50}$$

On a subdivision of the unit square into regular right triangles (a three-lines mesh), H is the well-known central difference matrix.

PROPOSITION 2.7. *Let $V_h \subset V^0$ and $f \in V_h'$ with $\mathrm{f} = (\mathcal{I}_h')^{-1} f$. Then,*

$$\mathcal{I}_h^{-1}(\tau f) = \mathrm{H}^{-1}\mathrm{f} \tag{2.51}$$

In a finite-dimensional subspace, converting a functional f to a function τf via the Riesz map is accomplished by multiplying the vector representation of f by the matrix representation of the inverse Laplace operator.

PROPOSITION 2.8. *Let $u \in V_h \subset V^0$ and $\mathrm{u} = \mathcal{I}_h^{-1} u$ be its vector of nodal coefficients. Then*

$$\|u\|_{V_h}^2 = \mathrm{u}^t\mathrm{H}\mathrm{u} \tag{2.52}$$

PROPOSITION 2.9. *Let $f \in (V_h)'$ and $\mathrm{f} = (\mathcal{I}_h')^{-1} f$ Then*

$$\|f\|_{V_h'}^2 = \mathrm{f}^t\mathrm{H}^{-1}\mathrm{f} \tag{2.53}$$

Proof. Using the Riesz representation theorem, the previous two propositions, and the symmetry of H,

$$\begin{aligned} \|f\|_{V_h'}^2 = \|\tau f\|_{V_h}^2 &= \left(\mathrm{H}^{-1}\mathrm{f}\right)^t \mathrm{H}\left(\mathrm{H}^{-1}\mathrm{f}\right) \\ &= \mathrm{f}^t\mathrm{H}^{-t}\mathrm{f} = \mathrm{f}^t\mathrm{H}^{-1}\mathrm{f} \end{aligned} \tag{2.54}$$

□

2.4. Matrix properties and analytic constants. Here, we explore the relationship between coercivity and continuity and the implied properties of the matrix and its spectrum. While coercive operators give rise to positive-definite matrices, coercivity is a stronger condition that applies uniformly to a family of matrices. To wit, if a is a coercive bilinear form on V, then we may say about the matrices that

$$\mathrm{u}^t\mathrm{A}\mathrm{u} = a(u,u) \geq \alpha\|u\|_V^2 = \alpha\mathrm{u}^t\mathrm{H}\mathrm{u} > 0. \tag{2.55}$$

For conforming Galerkin methods, this will hold for some α that is independent of the finite element subspace used to construct the matrices, and gives lower bounds much sharper than zero.

Similarly, continuity implies that there exists some C such that

$$\mathrm{u}^t\mathrm{A}\mathrm{v} = a(u,v) \leq C\|u\|_V\|v\|_V = C\sqrt{\mathrm{u}^t\mathrm{H}\mathrm{u}}\sqrt{\mathrm{v}^t\mathrm{H}\mathrm{v}} \tag{2.56}$$

As the matrices considered represent operators, we can say quite a bit about certain spectral properties based on the infinite-dimensional information. Suppose for a moment that the bilinear form a is symmetric so that $a(u,v) = a(v,u)$ (or equivalently, that the operator A is self-adjoint). Then, the finite element method

will give rise to symmetric matrices, for $\mathrm{A}_{i,j} = a(\phi_j, \phi_i) = a(\phi_i, \phi_j) = \mathrm{A}_{j,i}$. However, since A encodes an operator from V_h to V_h', a discussion of the eigenvalues of A implicitly involves identifying V_h with its dual.

When discussing eigenvalues, it only makes sense to consider an operator from a space into itself. We restrict this discussion to standard conforming Galerkin-type methods. Since $A_h : V_h \to V_h'$ and the Riesz map $\tau_h = \tau|_{V_h} : V_h' \to V_h$, the composition $\tau_h A_h : V_h \to V_h$. The same is true at the infinite-dimensional level, where $\tau A : V \to V$. Now, what can be said about the spectrum of this operator?

THEOREM 2.10. *Let a be a symmetric, continuous, coercive bilinear form with constants C and α and $A \in \mathcal{L}(V, V')$ the operator associated with it. Then the spectrum of $\tau A \in \mathcal{L}(V, V)$ is contained in $[\alpha, C]$.*

Proof. τA is self-adjoint and bounded on V, then by the standard spectral theory for such operators (see, for example, Yosida [52]), the spectrum is contained in the interval $[\inf_{\|u\|=1}(Au, u), \sup_{\|u\|=1}(Au, u)]$. The lower bound of this interval is the constant α, and the upper bound is no larger than C. □

COROLLARY 2.11. *For a finite-dimensional subspace $V_h \subset V$, the eigenvalues of $\tau_h A_h$ are contained in $[\alpha_h, C_h] \subseteq [\alpha, C]$.*

The operator $\tau_h A_h : V_h \to V_h$ can be represented by the product of two matrices. When $V_h \subset H_0^1$, we can write that

$$\tau_h A_h = \mathcal{I}_h \mathrm{H}^{-1} \mathrm{A} \mathcal{I}_h^{-1}, \tag{2.57}$$

where H is the matrix representation of the Helmholtz operator defined above and A is the stiffness matrix. So, this $\mathrm{H}^{-1}\mathrm{A}$ is related to $\tau_h A_h$ by a kind of similarity transformation.

COROLLARY 2.12. *The spectrum of* $\mathrm{H}^{-1}\mathrm{A}$ *is contained in* $[\alpha_h, C_h] \subseteq [\alpha, C]$.

While it is well-known that an elliptic operator may be preconditioned with another elliptic operator, such as in Faber *et al* [23], estimates relying solely on linear algebra are more tedious. The approach taken here, though using more advanced techniques, gives much shorter proofs. Moreover, the quantification of the spectrum in terms of the same best constants appearing in the existence/uniqueness theory of Lax-Milgram also seems to be new.

Now, if the operator A is coercive but not symmetric, we may still bound the real parts of the eigenvalues in terms of the Lax-Milgram constants. Define the symmetric part of the bilinear form a by $a_s(u, v) = \frac{1}{2}(a(u, v) + a(v, u))$ and the symmetric part of the operator A_s to be the operator associated with a_s. Then, applying Theorem 2.10 and its corollary give bounds on the spectrum of A_s as well as its restriction to any subspace. One can also study the spectrum of the skew-Hermition portion and obtain bounds for A itself via Bendixson's result [10] or a generalization [51]. The symmetric bilinear form a_s so defined establishes a new inner product and serves as a generalization of symmetric-part preconditioning to the operator level. The work of Arioli *et al.* [3, 4] uses this kind of framework to study stopping criteria and orthogonalization for Krylov methods in appropriate Hilbert space contexts.

Besides the spectrum, we may also relate our discussion to field-of-values [33] analysis of the preconditioned operator, which can be useful in describing the early phases of convergence for certain iterative methods. For example, GMRES [48] is a parameter-free algorithm that approximates the solution to a linear system by minimizing the equation's residual over Krylov subspaces. Consider a preconditioned linear system of the form

$$\mathrm{C}^{-1}\mathrm{A}\mathrm{u} = \mathrm{C}^{-1}\mathrm{f}, \tag{2.58}$$

where $\mathrm{A} \in \mathbb{R}^{n\times n}$ comes from a standard finite element discretization of a coercive operator and C is a symmetric and positive-definite preconditioner. For a particular choice of C, Starke [49] derives linear convergence rates for GMRES in terms of the quantities

$$\beta = \inf_{0\neq \mathrm{w}\in\mathbb{R}^n} \frac{\mathrm{w}^t\mathrm{A}\mathrm{w}}{\mathrm{w}^t\mathrm{C}\mathrm{w}} \qquad \tilde{\beta} = \inf_{0\neq \mathrm{w}\in\mathbb{R}^n} \frac{\mathrm{w}^t\mathrm{A}^{-1}\mathrm{w}}{\mathrm{w}^t\mathrm{C}^{-1}\mathrm{w}} \tag{2.59}$$

While Starke interprets the quantities β and $\tilde{\beta}$ in terms of the field of values of A, when C=H we can relate these back to the infinite-dimensional operators. Using Propositions 2.6 through 2.9, it is clear that

$$\beta = \inf_{0\neq u\in V_h} \frac{a(u,u)}{\|u\|^2_{V_h}} \qquad \tilde{\beta} = \inf_{0\neq f\in (V_h)'} \frac{\langle f, A^{-1}f\rangle}{\|f\|^2_{(V_h)'}}. \tag{2.60}$$

That is, $\beta = \alpha_h$ is simply the coercivity constant of A on V_h, which is bounded below by the infinite-dimensional coercivity constant α. The constant $\tilde{\beta}$ is a coercivity constant for the inverse operator A^{-1}.

Even in the worst case, at each step GMRES reduces the error by at least as much as the best Lax-Milgram fixed point iteration. To see this, we will concentrate on $\tilde{\beta}$. Let $f \in (V_h)'$ and let $u = A^{-1}f$ so that $Au = f$. Then, we may use the coercivity and continuity of A to write

$$\frac{\langle f, A^{-1}f\rangle}{\|f\|^2_{(V_h)'}} = \frac{\langle Au, u\rangle}{\|Au\|^2_{V_h}} \geq \frac{\alpha\|u\|^2_{V_h}}{C^2\|u\|^2_{V_h}} = \frac{\alpha}{C^2} \tag{2.61}$$

If this estimate is used, then $\beta\tilde{\beta} \geq \frac{\alpha^2}{C^2}$. In this case, Starke's convergence rate bound degenerates to that for fixed point iterations studied in the next section. However, this estimate is wildly pessimistic, as the estimate on $\tilde{\beta}$ simultaneously minimizes $a(u,u)$ in the numerator while maximizing $a(u,v)$ in the denominator. The actual value of $\tilde{\beta}$ will typically be much larger, leading to GMRES converging more rapidly than the best fixed-point iteration, even before superlinear convergence is considered.

Now, we return to the origin of h^{-2} scaling for the matrix condition number of A on quasiunform meshes. This can be explained by an embedding of the finite element space into Euclidean space. The condition number is bounded via standard inverse inequalities. This result is not new; it appears at least as early as [6] for eigenvalue bounds and even earlier for norms by Descloux [16].

LEMMA 2.13. *The operators $\mathcal{I}_h$ and $\mathcal{I}_h'$ have condition numbers in their respective spaces of $\mathcal{L}(\mathbb{R}^{\dim V_h}, V_h)$ and $\mathcal{L}(\mathbb{R}^{\dim {V_h}'}, V_h')$ of $\mathcal{O}(h^{-1})$.*

Proof. Because of Lemma 2.5, it is sufficient to prove the result for $\mathcal{I}_h$. $\|\mathcal{I}_h(u)\|$ in Euclidean space defines an equivalent norm to the L^2 norm of u up to a factor of h^d, where d is the dimension of the domain. Since $\|u\|_{L^2} \leq \|u\|_{H^1_0} \leq Ch^{-1}\|u\|_{L^2}$ from an inverse inequality (see following remark), we are done. ∎

REMARK 2.2. *For readers less familiar with theoretical aspects of finite element analysis, inverse inequalities are tools that bound a polynomial in a some norm H^m in terms of a weaker norm H^ℓ, $\ell < m$ and negative powers of the mesh spacing (hence the term "inverse"). Under some standard assumptions such as quasiuniformity, one negative power of h is obtained for every order the space is decreased. Readers are referred to standard finite element texts such as [13] for further details.*

These results, together with Theorem 2.4 allow us to relate the matrix condition number to the operator condition number, recovering the standard $\mathcal{O}(h^{-2})$ estimate.

THEOREM 2.14. *The 2-norm matrix condition number for a finite element discretization of a coercive operator on V_h is bounded by*

$$\kappa_2(\mathrm{A}) = \frac{C}{\alpha}\mathcal{O}(h^{-2}) \tag{2.62}$$

Proof. Condition numbers satisfy a multiplicative inequality, as norms do, so

$$\begin{aligned}\kappa_2(\mathrm{A}) = \kappa_2((\mathcal{I}_h')^{-1} A_h \mathcal{I}_h) &\leq \kappa((\mathcal{I}_h')^{-1})\kappa(A_h)\kappa(\mathcal{I}_h) \\ &\leq c_1\mathcal{O}(h^{-1})\frac{C}{\alpha}c_2\mathcal{O}(h^{-1}) \leq \frac{C}{\alpha}\mathcal{O}(h^{-2})\end{aligned} \tag{2.63}$$

□

This theorem provides an alternate derivation of the h^{-2} condition number, but it also provides new insight into the structure of condition numbers by quantifying distinct contributions. The factor of $\frac{C}{\alpha}$ is "intrinsic" to the problem in Hilbert space; it describes the dissimilarity of the operator to the Riesz map. This is completely independent of the particular mesh or finite element space and depends on properties of the operator itself, such as the strength of nonsymmetry or the amount of heterogeneity or anisotropy in any coefficients. On the other hand, the $\mathcal{O}(h^{-2})$ results from embedding $V^0 \subset H^1$ nonisometrically into Euclidean space. Also, the constant multiplying the h^{-2} depends strongly on the assumption of quasiuniformity of the meshes. Though it is difficult to quantify precisely, the less regular the meshes, the larger the constant will be.

We also remark that the assumption of coercivity is not critical in bounding the matrix condition number. Provided merely that A is bounded with bounded inverse and that the discrete problems have stable, unique solutions, it is possible to get a bound of $\kappa(A_h)\mathcal{O}(h^{-2})$ on the matrix condition number.

2.5. Left and right preconditioning. We have typically been considering applying Riesz maps as left preconditioners. In right preconditioning, instead of the linear system

$$\mathrm{Ax} = \mathrm{b}, \tag{2.64}$$

one solves

$$\left(\mathrm{AH}^{-1}\right)\mathrm{y} = \mathrm{b}, \tag{2.65}$$

followed by the computation

$$\mathrm{x} = \mathrm{H}^{-1}\mathrm{y}. \tag{2.66}$$

Analysis of left and right preconditioning is very similar. As we commented before, the interpretation of left and right preconditioning is somewhat different. With left preconditioning, one takes an operator $A : V \to V'$ and constructs a bounded linear map $\tilde{A} = \tau A : V \to V$, giving an operator equation on V. On the other hand, right Riesz preconditioning leads one to the operator equation

$$(A\tau)\, y = f \tag{2.67}$$

for some $y \in V'$ such that $x = \tau y$. This inverts the operator $A\tau : V' \to V'$ on the dual space rather than the space itself.

3. Stationary iterative methods.

3.1. Basic theory. The proof of Lax-Milgram given by Lions and Stampacchia [37] and followed in finite element texts such as [13, 14] solves the operator equation $Au = f$ in V' by finding a fixed point of the mapping

$$Tv = v + \rho\tau\left(f - Av\right). \tag{3.1}$$

Here, ρ is some positive number, τ is the Riesz map from V' onto V. Under the assumption of coercivity and continuity of A with constants α, C, the operator T in (3.1) is contractive and hence has a unique Banach fixed point provided $0 < \rho < \frac{2\alpha}{C^2}$, in which case the simple iteration

$$v^{n+1} = Tv^n = v^n + \rho\tau\left(f - Av^n\right) \tag{3.2}$$

converges to the unique fixed point of T for any initial guess $v^0 \in V$. Obviously, the fixed point u satisfies $f = Au$ in V'.

It can be shown(e.g. [13, Section 2.7]) that the norm of T in terms of ρ, C, α is bounded by

$$\|T\|^2_{\mathcal{L}(V,V)} \leq 1 - 2\rho\alpha + \rho^2 C^2 \tag{3.3}$$

In particular the bound on this quantity is minimized if

$$\rho = \frac{\alpha}{C^2}, \tag{3.4}$$

when the bound becomes

$$\|T\|_{\mathcal{L}(V,V)} \leq \sqrt{1 - \frac{\alpha^2}{C^2}} = \sqrt{1 - \kappa(A)^{-2}}. \tag{3.5}$$

It is easy to show that a fixed error reduction per step will lead to a relative error of no greater than ϵ in

$$N > \frac{\log \epsilon}{\log \sqrt{1 - \kappa(A)^{-2}}} \tag{3.6}$$

iterations.

This analysis does not guarantee sharpness, only bounds on convergence rates. An obvious question is whether the exactly optimal choice of ρ has been made. We have so far used only estimates for the coercivity and continuity constants and have not proven our bounds sharp. Moreover, it is not clear even that the estimates leading to (3.3) are sharp. It could also be possible to "overrelax" the method, much as with SOR. In the literature, exact answers are known for particular matrices, and Petryshyn [45] considers this question for general splitting methods in Hilbert space for a limited class of operators. We may use the earlier characterization of Hilbert space operators as matrices to describe (3.2) in terms of matrices.

3.2. Matrix representation. The Lax-Milgram iteration (3.2) has a simple interpretation in terms of matrix operations as a Richardson iteration. With A the stiffness matrix from (2.17) and H the matrix obtained by discretizing the Riesz map (e.g. the Laplace operator) with finite elements, the computation is

$$\mathrm{v}^{n+1} = \mathrm{v}^n + \rho\mathrm{H}^{-1}\left(\mathrm{f} - \mathrm{A}\mathrm{v}^n\right). \tag{3.7}$$

This is just a stationary iteration using a Laplace solver as the preconditioner. Algebraically, this is equivalent to a splitting method in which ρ times the Riesz matrix is split from the rest of the matrix

$$\begin{aligned} \mathrm{v}^{n+1} &= \mathrm{v}^n + \rho \mathrm{H}^{-1} \left(\mathrm{f} - \mathrm{A}\mathrm{v}^n \right) \\ &= \frac{1}{\rho} \mathrm{H}^{-1} \left(\mathrm{f} - (\mathrm{A} - \rho \mathrm{H}) \, \mathrm{v}^n \right) \end{aligned} \tag{3.8}$$

Suppose that the matrix A were considered simply as a mapping on Euclidean space and the Lax-Milgram iteration were applied in the Euclidean topology. The Riesz map is trivially the identity matrix, and the fixed point iteration becomes

$$\begin{aligned} \mathrm{v}^{n+1} &= \mathrm{v}^n + \rho \mathrm{I} \left(\mathrm{f} - \mathrm{A}\mathrm{v}^n \right) \\ &= \rho \frac{1}{\rho} \mathrm{I}\mathrm{v}^n + \rho \left(\mathrm{f} - \mathrm{A}\mathrm{v}^n \right) \\ &= \rho \mathrm{I} \left(\mathrm{f} - \left(\mathrm{A} - \frac{1}{\rho} \mathrm{I} \right) \mathrm{v}^n \right). \end{aligned} \tag{3.9}$$

If the diagonal entries of the matrix are all some constant γ (such as happens in standard finite differences for constant coefficient operators on regular meshes), then $\mathrm{D} = \gamma \mathrm{I}$. We can select $\rho = \frac{1}{\gamma}$ to write the iteration as

$$\mathrm{v}^{n+1} = \mathrm{D}^{-1} \left(\mathrm{f} - (\mathrm{A} - \mathrm{D}) \, \mathrm{v}^n \right), \tag{3.10}$$

which is nothing more than the Jacobi iteration with well-known $\mathcal{O}(h^{-2})$ complexity. Other choices of ρ would correspond to under- or over-relaxed Jacobi iterations.

3.3. Generalizations and comments. Starting from the Lax-Milgram iteration to derive relaxation methods leads to a different perspective on splitting techniques. Classically, methods such as Jacobi, Gauss-Seidel, and SOR [53] were motivated by splitting off an easily-invertible part of the matrix. At the time of Young's analysis, typically only matrices with particular algebraic structure, such as diagonal or triangular matrices, could be inverted quickly. However, methods such as ADI and multigrid extend the class of matrices that can be inverted quickly. In light of these developments, it makes sense to consider splitting off portions of the matrix that look like a Laplace or Helmholtz operator. Doing so means working with splittings based on *analytic* rather than algebraic properties of the operator.

In the context of splitting, we need not restrict to the Riesz map or even symmetric operators. Suppose that $A : V \to V'$ may be written as

$$A = A_1 + A_2, \tag{3.11}$$

where A_1 is coercive and continuous with constants α_1 and C_1 and A_2 is continuous with constant C_2. We do not assume coercivity or even invertibility of A_2. The operator equation $\langle Au, v \rangle = \langle f, v \rangle$ may be written as

$$\langle A_1 u, v \rangle = \langle f - A_2 u, v \rangle, \tag{3.12}$$

which motivates a fixed point iteration. Let $u^0 \in V$ be some initial starting point, and define u^{n+1} iteratively as the solution of

$$\langle A_1 u^{n+1}, v \rangle = \langle f - A_2 u^n, v \rangle, \quad v \in V \tag{3.13}$$

Methods such as these were famously analyzed for matrices by Young [53]. This work, including overrelaxation, was extended to abstract Hilbert space by several authors [28, 41, 45] through the 1960's and early 1970's. Our approach here differs from these papers in a few ways. First, we work in terms of the Lax-Milgram constants C, α rather than the spectra of the operators being considered. Working in terms of the spectrum potentially gives sharper results [20], but the present discussion is concerned primarily with the relation between baseline convergence rates and analytic constants. Second, we suggest particular splittings of coercive operators based on later developments in iterative methods. The idea of splitting something "easy" to invert is a moving target; for example, multigrid was largely unknown when these splitting techniques were developed, but now is widely used both as a solver and a preconditioner.

Before proceeding, we note that we do not even require linearity of A; A_2 may be some Lipschitz function from V into V'. Before stating the convergence result, we require a lemma.

LEMMA 3.1. *Let $A : V \to V'$ be continuous and coercive with constants C_A and α_A, and let $F : V \to V'$ be Lipschitz with $\|F(u)\|_{V'} \leq C_F \|u\|_V$. Then the equation*

$$\langle Au, v\rangle = \langle F(u), v\rangle, \quad v \in V \tag{3.14}$$

is solvable if $C_F < \alpha_A$.

Proof. The proof is by a Banach fixed point argument and Lax-Milgram. Let $v_1, v_2 \in V$ and consider the mapping $w_i = Tv_i$ defined by solving

$$a(w_i, x) = \langle F(v_i), x\rangle \tag{3.15}$$

for all $x \in V$ and for $i = 1, 2$. These linear problems have unique solutions with continuous dependence by Lax-Milgram, so

$$\begin{aligned} \|w_1 - w_2\|_V^2 &\leq \frac{1}{\alpha_A} a(w_1 - w_2, w_1 - w_2) \\ &= \frac{1}{\alpha_A} \langle F(v_1) - F(v_2), w_1 - w_2\rangle \\ &\leq \frac{1}{\alpha_A} \|F(v_1) - F(v_2)\|_{V'} \|w_1 - w_2\|_V \\ &\leq \frac{C_F}{\alpha_A} \|v_1 - v_2\|_V \|w_1 - w_2\|_V, \end{aligned} \tag{3.16}$$

so that

$$\|w_1 - w_2\|_V \leq \frac{C_F}{\alpha_A} \|v_1 - v_2\|_V. \tag{3.17}$$

Hence, the mapping is contractive provided that $C_F < \alpha_A$. ∎

THEOREM 3.2. *Iteration (3.13) converges to the unique solution independent of initial guess if*

$$C_2 < \alpha_1 \tag{3.18}$$

Proof. Our splitting iteration is an instance of Lemma 3.1, with A_1 playing the role of A and $F(u) = f - A_2(u)$. Then for $u_1, u_2 \in V$,

$$\|F(u_1) - F(u_2)\|_{V'} = \|(f - A_2(u_1)) - (f - A_2(u_2))\|_{V'} \leq C_2 \|u_1 - u_2\|_{V'}, \tag{3.19}$$

so that F is Lipschitz with constant C_2. So, the mapping $v \mapsto A_1^{-1}(f - A_2 v)$ must have a fixed point by the previous lemma, and this fixed point solves the equation. □

In practice, even efficient solvers are typically more expensive to apply than preconditioners. It is possible to replace the Riesz map in the Lax-Milgram iteration with an "approximate" Riesz map. This approximation could either be a few steps of an iterative method (inexact iteration) or else an incomplete factorization technique. Provided that a uniform preconditioner (one which does not degrade with mesh refinement) is used, mesh independent bounds can still be obtained.

In this case, the fixed-point iteration on (3.1) is modified to

$$v^{n+1} = v^n + \rho\tilde{\tau}_h (f - Av^n). \tag{3.20}$$

However, we can write $\tilde{\tau_h} = \tau\tau^{-1}\tilde{\tau}_h$ and see that this iteration is equivalent to applying the exact Riesz map to the modified system

$$\tau^{-1}\tilde{\tau}Au = \tau^{-1}\tilde{\tau}f. \tag{3.21}$$

These ideas are well-known in the literature. We will not pursue this further in this work, but simply remark that such inexact Riesz maps may also be interpreted in the Lax-Milgram framework.

4. Applications. In this section, we consider the two model problems of convection-diffusion and planar elasticity, applying Lax-Milgram iterations and Krylov methods to each. For convection-diffusion, we consider GMRES. For planar elasticity, which is symmetric, we consider conjugate gradients. All of our calculations are performed using the Sundance library [38, 39] to generate matrices from high-level descriptions of variational forms, and the actual linear algebra is performed using the solver packages of Trilinos [30, 31].

4.1. Convection-diffusion. We consider an example problem, similar to Example 3.1.2 in [21] on the unit square with $\beta = (0, 1 + \frac{(1+x)^2}{4})$. Homogeneous Neumann conditions are imposed along $y = 1$. The Dirichlet condition $u = 1$ is set along $y = 0$. The left and right edges have the Dirichlet conditions $u(0, y) = 1 - y^2$, and $u(1, y) = 1 - y^3$. The source term is 0. We apply the Lax-Milgram relaxation and GMRES preconditioned with the Riesz map to both standard Galerkin and streamline-diffusion methods. The meshes are constructed by dividing the unit square into an $N \times N$ grid of squares, then dividing each square into right triangles.

For the Lax-Milgram iterations, we iterated until the residual $A\tilde{u} - f \in H^{-1}$ reached an H^{-1} norm of 10^{-4}. We used the GMRES implementation in the AztecOO package within Trilinos, which performs orthogonalization with respect to the Euclidean inner product. We iterated until the Eucldean norm of the relative residual reached 10^{-8}. The results for the Lax-Milgram iteration are shown in Tables 4.1 and 4.3 and for GMRES in Tables 4.2 and 4.4.

Table 4.1 shows that the Lax-Milgram iterations behave as expected. For each ϵ, the number of Lax-Milgram iterations seems bounded independently of the mesh spacing. The number of iterations varies strongly as a function of ϵ, indicating an increase in the Hilbert space condition number. While the case of $\epsilon = 0.01$ shows some variability with respect to the mesh, note that the iteration counts seem to be levelling. This is constistent with the fact that the discrete condition number $\frac{C_h}{\alpha_h}$ is bounded above by the actual condition number, and the bound is only asymptotic.

The story for GMRES for standard Galerkin is similar. Table 4.2 shows the iteration counts as a function of ϵ and N. Note that the number of iterations is

TABLE 4.1
Number of Lax-Milgram iterations for standard Galerkin required to reach a residual norm of 10^{-4} in the dual norm for a range of ϵ and N. The second column, ρ shows a manually selected relaxation parameter that gave optimal results.

ϵ	ρ	$N = 8$	$N = 16$	$N = 32$	$N = 64$	$N = 128$
1.0	1.0	6	6	7	7	7
0.1	0.3	35	39	41	42	42
0.01	0.018	625	660	877	978	1006

TABLE 4.2
Number of GMRES iterations for standard Galerkin required to reach a Euclidean residual norm of 10^{-8} for a range of ϵ and N.

ϵ	$N = 8$	$N = 16$	$N = 32$	$N = 64$	$N = 128$
1	7	7	6	6	5
0.1	16	17	17	16	14
0.01	45	67	73	72	68
0.001	60	174	291	326	325

considerably less as a function of ϵ than for the Lax-Milgram relaxation, as expected. The only surprise here is that the iteration counts actually start to decrease on the finest meshes. In fact, this trend continues if the mesh is refined further to $N = 256$ and $N = 512$. This phenonmenon is not fully explained by the present analysis. The bound of the condition number by $\frac{C}{\alpha}$ is just a bound, and moreover the linear convergence analysis of Starke does not fully determine the final iteration count.

For streamline-diffusion, we also studied the Lax-Milgram iteration and GMRES preconditioned with the Riesz map defined by the streamline diffusion inner product. One suspects the mesh-dependence to affect the convergence properties of the iterations. This is in fact the case, as seen in Table 4.3, as the number of Lax-Milgram iterations increases more sharply for mesh refinement than in standard Galerkin. Moreover, the best relaxation parameter seems to depend on the mesh, approaching that for standard Galerkin as the mesh is refined.

The GMRES results are shown in Table 4.4. For each fixed mesh, the number of iterations required as $\epsilon \searrow 0$ scales better than for standard Galerkin, as suggested in our earlier condition number analysis. Also, for fixed ϵ, the iterations required in streamline-diffusion vary more under mesh refinement, again as suggested by the analysis.

4.2. Planar elasticity. We consider the Galerkin discretization of a simple model elasticity problem. Let the reference configuration Ω be the unit square. The top boundary is assumed fixed (homogeneous boundary condition). Homogeneous Neumann conditions are posed on the left and right boundaries (stress-free), and a traction boundary condition is applied on the bottom, wherein the vertical stress has unit value. We consider the Lax-Milgram and conjugate gradient iterations preconditioned with the vector Laplace operator for Young's modulus $E = 1.0$ and Poisson ratio $\nu = 0.3, 0.4$. The conjugate gradient implementation was again from AztecOO, using the relative residual in the Euclidean norm as a stopping criterion. This is a standard conjugate gradient implementation. The number of iterations as a function of N for these two Poisson ratios appears in Tables 4.5 and 4.6. In both cases, the number of iterations required to reach a particular tolerance remained nearly flat with

TABLE 4.3
Number of Lax-Milgram iterations to generate a residual of 10^{-4} *in the dual streamline-diffusion norm. The optimal relaxation parameter is in parentheses beside each iteration count.*

ϵ	$N = 8$	$N = 16$	$N = 32$	$N = 64$	$N = 128$
1.0	6 (1.0)	6(1.0)	6(1.0)	6(1.0)	7(1.0)
0.1	26 (0.6)	32(0.4)	36 (0.35)	39 (0.325)	40 (0.3)
0.01	83 (0.5)	194 (0.15)	337 (0.05)	422 (0.03)	513 (0.025)

TABLE 4.4
Number of GMRES iterations to generate a Euclidean residual of 10^{-8} *for the streamline diffusion discretization.*

ϵ	$N = 8$	$N = 16$	$N = 32$	$N = 64$	$N = 128$
1	7	7	6	6	5
0.1	16	16	16	15	13
0.01	35	57	65	62	57
0.001	47	121	250	307	236

the mesh size. With all other parameters fixed, the condition number increases with $\nu \nearrow 0.5$. The case of $\nu = 0.4$ has a worse Hilbert space condition number than $\nu = 0.3$, which is reflected in the increased number of iterations required. However, the dependence of conjugate gradient on this condition number is much weaker than the fixed point iteration. Also, all of the experiments for elasticity were repeated with piecewise quadratic elements rather than linears, giving nearly identical iteration counts. This is in accordance with our statements about having spectral bounds independent of the finite element subspace. Finally, if ν is increased toward the incompressible limit of 0.5, these methods behave badly; the phenomenon of locking leads to ill-conditioned systems as well as inaccurate answers and must be addressed by alternate discretizations such as mixed or nonconforming methods.

4.3. Some remarks. Methods tailored to particular problems, such as multigrid methods specialized to particular problems such as convection-diffusion or elasticity typically can provide better performance than Riesz maps. For example, see the geometric and algebraic multigrid algorithms for elasticity in [1, 2, 24] or the convection-diffusion multigrid in [43, 46]. Without those specialized implementations, Riesz maps can still give preconditioners that scale well with respect to mesh size but less so with respect to system parameters such as the Peclet number or Poisson ratio. As solver technology typically lags behind when new discretizations or new application areas are developed, Riesz maps can provide an intermediate, scalable technique until special-purpose solvers become available to the community.

5. Conclusions. The analytic structure of PDEs inherited by finite element methods leads to a natural framework to discuss conditioning and iterative methods. By seeing finite element operators in Hilbert space, we are able to succinctly show the separate dependence of matrix condition numbers on the operator itself and the particular features of the discretization. Moreover, obtaining a baseline preconditioner that will provide mesh-independence becomes much simpler. While more specialized preconditioners will obviously outperform Riesz maps, they may not always be immediately available in software. Moreover, when new kinds of applications are considered, discretizations must be developed before scalable algorithms for those discretizations. Given high-level software tools for finite elements such as Sundance,

TABLE 4.5
Lax-Milgram iteration counts for elasticity. The Young's modulus $E = 1.0$. The iteration is run to a tolerance of 10^{-4} in the dual norm.

ν	ρ	$N = 8$	$N = 16$	$N = 32$	$N = 64$	$N = 128$
0.3	0.9	22	22	22	22	22
0.4	0.45	44	46	48	48	48

TABLE 4.6
Conjugate gradient iteration for elasticity. The Young's modulus $E = 1.0$. The iteration is run until the relative Euclidean norm of the residual is 10^{-4}.

ν	$N = 8$	$N = 16$	$N = 32$	$N = 64$	$N = 128$
0.3	15	18	17	17	17
0.4	22	24	27	21	23

Riesz maps are available at the same time as the discretization. This gives simple scalability for large meshes and provides a reference point for solver development to target.

It is interesting to ask how portable this paradigm is to other kinds of finite element discretizations, especially in the context of coupled systems and saddle point problems. A recent manuscript of Mardal and Winther [40] extends many of these ideas to the context of Babuška inf-sup theory. Also, Riesz preconditioning may prove useful for least squares methods, which are typically have H^1 coercivity.

REFERENCES

[1] MARK ADAMS, *Parallel multigrid algorithms for unstructured 3D large deformation elasticity and plasticity finite element problems*, Tech. Report CSD-99-1036, 25, 1999.

[2] M. F. ADAMS, *Parallel multigrid solvers for 3d unstructured finite element problems in large deformation elasticity and plasticity*, International Journal for Numerical Methods in Engineering, 48 (2000), pp. 1241–1262.

[3] M. ARIOLI, *A stopping criterion for the conjugate gradient algorithms in a finite element method framework*, Numer. Math., 97 (2004), pp. 1–24.

[4] M. ARIOLI, D. LOGHIN, AND A. J. WATHEN, *Stopping criteria for iterations in finite element methods*, Numer. Math., 99 (2005), pp. 381–410.

[5] DOUGLAS N. ARNOLD, RICHARD S. FALK, AND RAGNAR WINTHER, *Preconditioning discrete approximations of the Reissner-Mindlin plate model*, RAIRO Modél. Math. Anal. Numér., 31 (1997), pp. 517–557.

[6] O. AXELSSON AND V. A. BARKER, *Finite element solution of boundary value problems*, vol. 35 of Classics in Applied Mathematics, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2001. Theory and computation, Reprint of the 1984 original.

[7] O. AXELSSON AND J. KARÁTSON, *On the rate of convergence of the conjugate gradient method for linear operators in Hilbert space*, Numer. Funct. Anal. Optim., 23 (2002), pp. 285–302.

[8] ———, *Symmetric part preconditioning for the conjugate gradient method in Hilbert space*, Numer. Funct. Anal. Optim., 24 (2003), pp. 455–474.

[9] RANDOLPH E. BANK AND TODD DUPONT, *An optimal order process for solving finite element equations*, Math. Comp., 36 (1981), pp. 35–51.

[10] IVAR BENDIXSON, *Sur les racines d'une équation fondamentale*, Acta Math., 25 (1902), pp. 359–365.

[11] FOLKMAR A. BORNEMANN AND PETER DEUFLHARD, *The cascadic multigrid method for elliptic problems*, Numer. Math., 75 (1996), pp. 135–152.

[12] DIETRICH BRAESS, *Finite elements*, Cambridge University Press, Cambridge, third ed., 2007. Theory, fast solvers, and applications in elasticity theory, Translated from the German by Larry L. Schumaker.

[13] SUSANNE C. BRENNER AND L. RIDGWAY SCOTT, *The mathematical theory of finite element*

methods, vol. 15 of Texts in Applied Mathematics, Springer, New York, third ed., 2008.
[14] P. G. Ciarlet, *The Finite Element Method for Elliptic Problems*, North-Holland, Amsterdam, New York, Oxford, 1978.
[15] James W. Daniel, *The conjugate gradient method for linear and nonlinear operator equations*, SIAM J. Numer. Anal., 4 (1967), pp. 10–26.
[16] Jean Descloux, *On finite element matrices*, SIAM J. Numer. Anal., 9 (1972), pp. 260–265.
[17] Jim Douglas, Jr. and James E. Gunn, *Alternating direction methods for parabolic systems in m space variables*, J. Assoc. Comput. Mach., 9 (1962), pp. 450–456.
[18] ———, *A general formulation of alternating direction methods. I. Parabolic and hyperbolic problems*, Numer. Math., 6 (1964), pp. 428–453.
[19] Jim Douglas, Jr. and H. H. Rachford, Jr., *On the numerical solution of heat conduction problems in two and three space variables*, Trans. Amer. Math. Soc., 82 (1956), pp. 421–439.
[20] Tobin A. Driscoll, Kim-Chuan Toh, and Lloyd N. Trefethen, *From potential theory to matrix iterations in six steps*, SIAM Rev., 40 (1998), pp. 547–578 (electronic).
[21] Howard C. Elman, David J. Silvester, and Andrew 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.
[22] K. Eriksson, D. Estep, P. Hansbo, and C. Johnson, *Computational differential equations*, Cambridge University Press, Cambridge, 1996.
[23] V. Faber, Thomas A. Manteuffel, and Seymour V. Parter, *On the theory of equivalent operators and application to the numerical solution of uniformly elliptic partial differential equations*, Adv. in Appl. Math., 11 (1990), pp. 109–163.
[24] Michael Griebel, Daniel Oeltz, and Marc Alexander Schweitzer, *An algebraic multigrid method for linear elasticity*, SIAM J. Sci. Comput., 25 (2003), pp. 385–407 (electronic).
[25] James E. Gunn, *The numerical solution of* $\nabla \cdot a \nabla u = f$ *by a semi-explicit alternating-direction iterative technique*, Numer. Math., 6 (1964), pp. 181–184.
[26] ———, *On the two-stage iterative method of Douglas for mildly nonlinear elliptic difference equations*, Numer. Math., 6 (1964), pp. 243–249.
[27] ———, *The solution of elliptic difference equations by semi-explicit iterative techniques*, J. Soc. Indust. Appl. Math. Ser. B Numer. Anal., 2 (1965), pp. 24–45.
[28] Michael P. Hanna, *Generalized Overrelaxation and Gauss-Seidel Convergence on Hilbert space*, Proceedings of the American Mathematical Society, 35 (1972), pp. 524–530.
[29] R. M. Hayes, *Iterative methods of solving linear problems on Hilbert space*, in Contributions to the solution of systems of linear equations and the determination of eigenvalues, National Bureau of Standards Applied Mathematics Series No. 39, U. S. Government Printing Office, Washington, D. C., 1954, pp. 71–103.
[30] Michael Heroux, Roscoe Bartlett, Vicki Howle Robert Hoekstra, Jonathan Hu, Tamara Kolda, Richard Lehoucq, Kevin Long, Roger Pawlowski, Eric Phipps, Andrew Salinger, Heidi Thornquist, Ray Tuminaro, James Willenbring, and Alan Williams, *An Overview of Trilinos*, Tech. Report SAND2003-2927, Sandia National Laboratories, 2003.
[31] Michael A. Heroux, Roscoe A. Bartlett, Vicki E. Howle, Robert J. Hoekstra, Jonathan J. Hu, Tamara G. Kolda, Richard B. Lehoucq, Kevin R. Long, Roger P. Pawlowski, Eric T. Phipps, Andrew G. Salinger, Heidi K. Thornquist, Ray S. Tuminaro, James M. Willenbring, Alan Williams, and Kendall S. Stanley, *An overview of the trilinos project*, ACM Trans. Math. Softw., 31 (2005), pp. 397–423.
[32] Magnus R. Hestenes and Eduard Stiefel, *Methods of conjugate gradients for solving linear systems*, J. Research Nat. Bur. Standards, 49 (1952), pp. 409–436 (1953).
[33] Roger A. Horn and Charles R. Johnson, *Topics in matrix analysis*, Cambridge University Press, Cambridge, 1991.
[34] T. J. R. Hughes and A. Brooks, *A multidimensional upwind scheme with no crosswind diffusion*, in Finite element methods for convection dominated flows (Papers, Winter Ann. Meeting Amer. Soc. Mech. Engrs., New York, 1979), vol. 34 of AMD, Amer. Soc. Mech. Engrs. (ASME), New York, 1979, pp. 19–35.
[35] R. B. Kellogg, *An alternating direction method for operator equations*, J. Soc. Indust. Appl. Math., 12 (1964), pp. 848–854.
[36] P. D. Lax and A. N. Milgram, *Parabolic equations*, in Contributions to the theory of partial differential equations, Annals of Mathematics Studies, no. 33, Princeton University Press, Princeton, N. J., 1954, pp. 167–190.
[37] J.-L. Lions and G. Stampacchia, *Variational inequalities*, Comm. Pure Appl. Math., 20 (1967), pp. 493–519.
[38] Kevin Long, *Sundance, a rapid prototyping tool for parallel PDE-constrained optimization*, in

Large-Scale PDE-Constrained Optimization, Lecture notes in computational science and engineering, Springer-Verlag, 2003.
[39] ———, *Sundance 2.0 tutorial*, Tech. Report TR–2004–09, Sandia National Laboratories, 2004.
[40] Kent-Andre Mardal and Ragnar Winther, *Preconditioning discretizations of systems of partial differential equations*. submitted to *Numerical Linear Algebra*.
[41] Ivo Marek, *On the SOR method for solving linear equations in Banach spaces*, Wiss. Z. Techn. Hochsch. Karl-Marx-Stadt, 11 (1969), pp. 335–341.
[42] Olavi Nevanlinna, *Convergence of iterations for linear equations*, Lectures in Mathematics ETH Zürich, Birkhäuser Verlag, Basel, 1993.
[43] Maxim A. Olshanskii and Arnold Reusken, *Convergence analysis of a multigrid method for a convection-dominated model problem*, SIAM J. Numer. Anal., 42 (2004), pp. 1261–1291 (electronic).
[44] D. W. Peaceman and H. H. Rachford, Jr., *The numerical solution of parabolic and elliptic differential equations*, J. Soc. Indust. Appl. Math., 3 (1955), pp. 28–41.
[45] W. V. Petryshyn, *Remarks on the generalized overrelaxation and the extrapolated Jacobi methods for operator equations in Hilbert space*, J. Math. Anal. Appl., 29 (1970), pp. 558–568.
[46] Arnold Reusken, *Convergence analysis of a multigrid method for convection-diffusion equations*, Numer. Math., 91 (2002), pp. 323–349.
[47] H. L. Royden, *Real analysis*, Macmillan Publishing Company, New York, third ed., 1988.
[48] Yousef Saad, *Iterative methods for sparse linear systems*, Society for Industrial and Applied Mathematics, Philadelphia, PA, second ed., 2003.
[49] Gerhard Starke, *Field-of-values analysis of preconditioned iterative methods for nonsymmetric elliptic problems*, Numer. Math., 78 (1997), pp. 103–117.
[50] G. Strang and G. J. Fix, *An Analysis of the Finite Element Method*, Prentice-Hall, Englewood Cliffs, 1973.
[51] Helmut Wielandt, *On eigenvalues of sums of normal matrices*, Pacific J. Math., 5 (1955), pp. 633–638.
[52] Kōsaku Yosida, *Functional analysis*, Classics in Mathematics, Springer-Verlag, Berlin, 1995. Reprint of the sixth (1980) edition.
[53] David Young, *Iterative methods for solving partial difference equations of elliptic type*, Trans. Amer. Math. Soc., 76 (1954), pp. 92–111.
[54] Eberhard Zeidler, *Nonlinear functional analysis and its applications. I*, Springer-Verlag, New York, 1986. Fixed-point theorems, Translated from the German by Peter R. Wadsack.