MULTIGRID PRECONDITIONING IN $\boldsymbol{H}(\mathrm{div})$ ON NON-CONVEX POLYGONS*

DOUGLAS N. ARNOLD†, RICHARD S. FALK‡, AND RAGNAR WINTHER§

Dedicated to Professor Jim Douglas, Jr. on the occasion of his seventieth birthday.

Abstract. In an earlier paper we constructed and analyzed a multigrid preconditioner for the system of linear algebraic equations arising from the finite element discretization of boundary value problems associated to the differential operator $\boldsymbol{I} - \mathbf{grad}\,\mathrm{div}$. In this paper we analyze the procedure without assuming that the underlying domain is convex and show that, also in this case, the preconditioner is spectrally equivalent to the inverse of the discrete operator.

Key words. preconditioner, finite element, multigrid, nonconvex domain

AMS(MOS) subject classifications (1991 revision). 65N55, 65N30

1. Introduction. In the earlier paper [1], we analyzed domain decomposition and multigrid precondtioners for the efficient solution of the equations which arise from the finite element discretization of boundary values problems for the operator $\boldsymbol{I} - \mathbf{grad}\,\mathrm{div}$. These results were then applied to construct efficient iterative methods for the solution of the equations which arise from the finite element discretization of scalar second order elliptic boundary value problems by mixed and least squares methods. In the case of the domain decomposition algorithm, the convergence results were obtained first for the case of a convex polygon, in which the solution of the scalar second order elliptic problem has H^2–regularity, and then extended to the case of a nonconvex polygon, where the solution has less regularity. In the case of the multigrid method, the analysis presented made essential use of H^2–regularity and hence did not apply to the case of a nonconvex polygon. The purpose of this paper is to present a different analysis for the multigrid method which applies also in the nonconvex case.

As in [1], we follow the outline of the modern theory of multigrid methods as, for example, presented in Bramble [2] or Xu [8]. However, the operator $\boldsymbol{I} - \mathbf{grad}\,\mathrm{div}$ lacks a number of properties possessed by standard elliptic operators. For example, when restricted to gradient fields this operator acts like a second order elliptic operator, while

*The first author was supported NSF grant DMS-9500672. The second author was supported by NSF grant DMS-9403552. The third author was supported by The Norwegian Research Council under grants 100331/431 and STP.29643.

†Department of Mathematics, Penn State, University Park, PA 16802. dna@math.psu.edu

‡Department of Mathematics, Rutgers University, New Brunswick, NJ 08903. falk@math.rutgers.edu

§Department of Informatics, University of Oslo, Oslo, Norway. ragnar@ifi.uio.no

when restricted to curl fields it coincides with the identity. The main tool for overcoming these problems is an appropriate discrete Helmholtz decomposition.

Our interest in preconditioning discrete approximations of the operator $\boldsymbol{I} - \mathbf{grad}\,\mathrm{div}$ is motivated by the fact that such operators can be used to build preconditioners for discretizations of many differential systems where the space $\boldsymbol{H}(\mathrm{div})$ appears naturally as a part of the solution space. Typical examples include the mixed method for second order elliptic problems, the least squares method of the form discussed by Lazarov, Manteuffel, and McCormick [5] and Pehlivanov, Carey, and Lazarov [7] and the sequential regularization method for the time dependent Navier–Stokes equation discussed in Lin [6]. For a more detailed discussion of applications and for numerical experiments we refer to [1].

In §2 we introduce notations and briefly recall the Raviart–Thomas finite element spaces and the associated Helmholtz decomposition. The V–cycle multigrid operator is analyzed in §3. Under suitable assumptions on the smoothing operators, we establish that the V–cycle operator leads to a uniform preconditioner for the discrete approximation of $\boldsymbol{I} - \mathbf{grad}\,\mathrm{div}$. In §4 we then verify these assumptions for appropriate smoothers.

2. Preliminaries. We suppose that the domain $\Omega \subset \mathbb{R}^2$ is polygonal, but not necessarily convex. The inner product and norm in $L^2 = L^2(\Omega)$ are denoted by $(\,\cdot\,,\,\cdot\,)$ and $\|\cdot\|$, respectively. The L^2–based Sobolev space of order m is denoted by H^m. We use boldface type for vectors in $\mathbb{R}^2$, vector-valued functions, spaces of such functions, and operators with range in such spaces. Thus, for example, $\boldsymbol{L}^2$ denotes the space of 2-vector-valued functions on Ω for which both components are square integrable. We also use the standard differential operators

$$\mathbf{grad} = \begin{pmatrix} \partial/\partial x \\ \partial/\partial y \end{pmatrix}, \qquad \mathbf{curl} = \begin{pmatrix} -\partial/\partial y \\ \partial/\partial x \end{pmatrix}, \qquad \mathrm{div} = (\,\partial/\partial x \quad \partial/\partial y\,).$$

The Hilbert space $\boldsymbol{H}(\mathrm{div})$ consists of square-integrable vectorfields on Ω with square integrable divergence: $\boldsymbol{H}(\mathrm{div}) = \{\boldsymbol{v} \in \boldsymbol{L}^2 : \mathrm{div}\,\boldsymbol{v} \in L^2\}$. The associated inner product is $\Lambda(\boldsymbol{u},\boldsymbol{v}) = (\boldsymbol{u},\boldsymbol{v}) + (\mathrm{div}\,\boldsymbol{u}, \mathrm{div}\,\boldsymbol{v})$. We note that this form is associated with the operator $\boldsymbol{I} - \mathbf{grad}\,\mathrm{div}$, in the sense that the problem of finding $\boldsymbol{u} \in \boldsymbol{H}(\mathrm{div})$ for which $\Lambda(\boldsymbol{u},\boldsymbol{v}) = (\boldsymbol{f},\boldsymbol{v})$ for all $\boldsymbol{v} \in \boldsymbol{H}(\mathrm{div})$, is the natural weak formulation of the problem $(\boldsymbol{I} - \mathbf{grad}\,\mathrm{div})\boldsymbol{u} = \boldsymbol{f}$ in Ω, $\mathrm{div}\,\boldsymbol{u} = 0$ on $\partial\Omega$.

Let $\{\mathcal{T}_h\}$ be a quasiuniform family of triangulations of Ω, where $h > 0$ is a parameter representative of the diameter of the elements of $\mathcal{T}_h$, and let r be a non-negative integer. The Raviart–Thomas space of index r is given by

$$\boldsymbol{V}_h = \{\boldsymbol{v} \in \boldsymbol{H}(\mathrm{div}) : \boldsymbol{v}|_T \in \boldsymbol{P}_r(T) + (x,y)P_r(T) \quad \text{for all } T \in \mathcal{T}_h\},$$

where $P_r(T)$ denotes the set of polynomial functions of degree at most r on T. The discrete system we wish to precondition is $\boldsymbol{\Lambda}_h \boldsymbol{u}_h = \boldsymbol{f}_h$, where $\boldsymbol{f}_h \in \boldsymbol{V}_h$ and $\boldsymbol{\Lambda}_h : \boldsymbol{V}_h \to \boldsymbol{V}_h$ is defined by

$$(\boldsymbol{\Lambda}_h \boldsymbol{u}, \boldsymbol{v}) = \Lambda(\boldsymbol{u},\boldsymbol{v}) \qquad \text{for all } \boldsymbol{u},\boldsymbol{v} \in \boldsymbol{V}_h. \tag{2.1}$$

We shall rely heavily on the discrete Helmholtz decomposition of $\boldsymbol{V}_h$. To recall it, we set

$$W_h = \{s \in H^1 : s|_T \in P_{r+1}(T)\}, \qquad S_h = \{q \in L^2 : q|_T \in P_r(T)\},$$

the usual spaces of continuous piecewise polynomials of degree $r+1$ and arbitrary piecewise polynomials of degree r, respectively. The discrete Helmholtz decomposition (cf. [3]) is then $\boldsymbol{V}_h = \mathbf{grad}_h S_h \oplus \mathbf{curl} W_h$, where $\mathbf{grad}_h : S_h \to \boldsymbol{V}_h$ is defined by the condition $(\mathbf{grad}_h q, \boldsymbol{v}) = -(q, \operatorname{div} \boldsymbol{v})$ for all $\boldsymbol{v} \in \boldsymbol{V}_h$. Note that the sum is orthogonal with respect to both the $\boldsymbol{L}^2$ and the $\boldsymbol{H}(\operatorname{div})$ inner products and the two summand spaces $\mathbf{grad}_h S_h$ and $\mathbf{curl} W_h$ are invariant under the action of $\boldsymbol{\Lambda}_h$. We also recall that the restriction of the divergence operator to $\boldsymbol{V}_h$ maps onto S_h and that its kernel is precisely $\mathbf{curl} W_h$.

3. Multigrid methods. In this section we analyze the V–cycle multigrid preconditioner $\boldsymbol{\Theta}_h$ for the operator $\boldsymbol{\Lambda}_h$. Under proper assumptions on the smoothing operators, we show that the operator $\boldsymbol{I} - \boldsymbol{\Theta}_h \boldsymbol{\Lambda}_h$ is a contraction uniformly with respect to h, and, *a fortiori*, that $\boldsymbol{\Theta}_h$ is spectrally equivalent to $\boldsymbol{\Lambda}_h^{-1}$. We usually suppress the subscript h. For example, we shall write $\mathcal{T}, \boldsymbol{V}, \boldsymbol{\Lambda}$, and $\boldsymbol{\Theta}$ instead of $\mathcal{T}_h, \boldsymbol{V}_h, \boldsymbol{\Lambda}_h$, and $\boldsymbol{\Theta}_h$.

In order to define a nested sequence of subspaces of $\boldsymbol{V}$, we assume that the triangulation $\mathcal{T}$ is constructed by a successive refinement process. More precisely, we assume that we have a nested sequence of quasi-uniform triangulations $\mathcal{T}_j$, $j = 1, 2 \ldots, J$, with characteristic mesh size h_j proportional to γ^{2j} for some positive constant $\gamma < 1$, and that $\mathcal{T} = \mathcal{T}_J$. Then $\boldsymbol{V}_1 \subset \boldsymbol{V}_2 \subset \cdots \subset \boldsymbol{V}_J = \boldsymbol{V}$, where $\boldsymbol{V}_j$ is the Raviart-Thomas space of index r relative to the triangulation $\mathcal{T}_j$. At each level j we have the discrete operator $\boldsymbol{\Lambda}_j : \boldsymbol{V}_j \to \boldsymbol{V}_j$, defined as in (2.1), i.e., $(\boldsymbol{\Lambda}_j \boldsymbol{v}, \boldsymbol{w}) = \Lambda(\boldsymbol{v}, \boldsymbol{w})$ for all $\boldsymbol{v}, \boldsymbol{w} \in \boldsymbol{V}_j$. Hence, $\boldsymbol{\Lambda}_J = \boldsymbol{\Lambda}$. The $\boldsymbol{L}^2$- and $\boldsymbol{H}(\operatorname{div})$-orthogonal projections onto $\boldsymbol{V}_j$, will be denoted by $\boldsymbol{Q}_j$ and $\boldsymbol{P}_j$, respectively, so the standard identity

$$\boldsymbol{Q}_{j-1} \boldsymbol{\Lambda}_j = \boldsymbol{\Lambda}_{j-1} \boldsymbol{P}_{j-1} \tag{3.1}$$

holds. In order to define a V–cycle operator from the nested sequence $\{\boldsymbol{V}_j\}$, we require appropriate smoothing operators. For each $j > 1$, we let $\boldsymbol{R}_j : \boldsymbol{V}_j \to \boldsymbol{V}_j$ be a linear operator which, as will be made more precise below, will be required to behave in some ways like an approximation to $\boldsymbol{\Lambda}_j^{-1}$.

The standard V–cycle multigrid algorithm with one smoothing recursively defines operators $\boldsymbol{\Theta}_j : \boldsymbol{V}_j \to \boldsymbol{V}_j$ beginning with $\boldsymbol{\Theta}_1 = \boldsymbol{\Lambda}_1^{-1}$. For $j > 1$ and $\boldsymbol{f} \in \boldsymbol{V}_j$, $\boldsymbol{\Theta}_j \boldsymbol{f} = \boldsymbol{x}_3$ where

$$\begin{aligned}
\boldsymbol{x}_1 &= \boldsymbol{R}_j \boldsymbol{f}, \\
\boldsymbol{x}_2 &= \boldsymbol{x}_1 + \boldsymbol{\Theta}_{j-1} \boldsymbol{Q}_{j-1} (\boldsymbol{f} - \boldsymbol{\Lambda}_j \boldsymbol{x}_1), \\
\boldsymbol{x}_3 &= \boldsymbol{x}_2 + \boldsymbol{R}_j (\boldsymbol{f} - \boldsymbol{\Lambda}_j \boldsymbol{x}_2).
\end{aligned}$$

The desired preconditioner $\mathbf{\Theta} : \boldsymbol{V} \to \boldsymbol{V}$ is the final operator, i.e. $\mathbf{\Theta} = \mathbf{\Theta}_J$. By using the identity (3.1), we easily derive the relation

$$\boldsymbol{I} - \mathbf{\Theta}_j \mathbf{\Lambda}_j = (\boldsymbol{I} - \boldsymbol{R}_j \mathbf{\Lambda}_j)(\boldsymbol{I} - \mathbf{\Theta}_{j-1} \mathbf{\Lambda}_{j-1} \boldsymbol{P}_{j-1})(\boldsymbol{I} - \boldsymbol{R}_j \mathbf{\Lambda}_j) \qquad \text{for } j = 1, 2, \ldots, J, \tag{3.2}$$

from the V–cycle algorithm above. Here and below we conventionally define $\mathbf{\Theta}_0 = \mathbf{\Lambda}_0 = \boldsymbol{P}_0 = \boldsymbol{Q}_0 = 0$ and $\boldsymbol{R}_1 = \mathbf{\Lambda}_1^{-1}$.

Using the triangulation $\mathcal{T}_j$, we also define finite element spaces W_j and S_j and the discrete gradient operator $\mathbf{grad}_j : S_j \to \boldsymbol{V}_j$ as in §2. For each j the discrete Helmholtz decomposition, $\boldsymbol{V}_j = \mathbf{grad}_j\, S_j \oplus \mathbf{curl}\, W_j$, holds. We shall also use the L^2 projection $Q_j^W : L^2 \to W_j$ and $Q_j^S : L^2 \to S_j$.

We now state the main assumptions on the smoothing operators $\boldsymbol{R}_j : \boldsymbol{V}_j \mapsto \boldsymbol{V}_j$ needed for the theory below. These assumptions will be verified for proper additive Schwarz and multiplicative Schwarz smoothers in the next section. The smoothing operators $\boldsymbol{R}_j$ will be assumed to be L^2–symmetric and positive definite, whence the operators $\boldsymbol{R}_j \mathbf{\Lambda}_j : \boldsymbol{V}_j \mapsto \boldsymbol{V}_j$ are symmetric, positive definite with respect to the $\boldsymbol{H}(\mathrm{div})$ inner product. We make two further assumptions:

Assumption A1. *There is a constant ω, independent of h, such that the spectral radius of $\boldsymbol{R}_j \mathbf{\Lambda}_j$, $\rho(\boldsymbol{R}_j \mathbf{\Lambda}_j)$, satisfies*

$$\rho(\boldsymbol{R}_j \mathbf{\Lambda}_j) \le \omega < 2 \quad \text{for } j = 2, 3, \ldots, J. \tag{3.3}$$

Assumption A2. *There is a constant C, independent of h, such that*

$$(\boldsymbol{R}_j^{-1} \boldsymbol{v}, \boldsymbol{v}) \le C h_j^{-2} \|\boldsymbol{v}\|^2 \qquad \text{for all } \boldsymbol{v} \in \boldsymbol{V}_j,\ j = 2, 3, \ldots, J. \tag{3.4}$$

Furthermore, if $\boldsymbol{v} = \mathbf{curl}[(Q_j^W - Q_{j-1}^W)\phi]$ for $\phi \in W_j$ then

$$(\boldsymbol{R}_j^{-1} \boldsymbol{v}, \boldsymbol{v}) \le C \|\boldsymbol{v}\|^2 \quad \text{for } j = 2, 3, \ldots, J. \tag{3.5}$$

Note that both assumptions are met if $\boldsymbol{R}_j$ is replaced by $\mathbf{\Lambda}_j^{-1}$. With these assumptions we have the main result of this section.

Theorem 3.1. *The spectral condition number of $\mathbf{\Theta\Lambda} = \mathbf{\Theta}_J \mathbf{\Lambda}_J$ is bounded independently of h and J.*

Before proceeding to the proof, we need to establish some preliminary results. The following approximation result is a straightforward extension of the inequality (A.1) proved in Appendix A of [1] (the operator $I - Q_{j-1}^S$ doesn't appear in (A.1), but is present in the proof of that equation).

LEMMA 3.2. *Let $\boldsymbol{v}, \hat{\boldsymbol{v}} \in \mathbf{grad}_j S_j \subset \boldsymbol{V}_j$ be such that $\operatorname{div} \hat{\boldsymbol{v}} = Q_{j-1}^S \operatorname{div} \boldsymbol{v}$. Then there is a constant C, independent of h_j, such that*

$$\|\boldsymbol{v} - \hat{\boldsymbol{v}}\| \le C h_j \|(I - Q_{j-1}^S) \operatorname{div} \boldsymbol{v}\|.$$

The second result we shall need is a form of the strengthened Cauchy–Schwarz inequality (cf. §6.1 of [8]).

LEMMA 3.3. *Assume that $1 \le i < j \le J$. There is a constant C, independent of h and J, such that*

$$(\mathbf{curl}\,\mu, \boldsymbol{v}) \le C\gamma^{j-i} h_j^{-1} \|\mu\|_0 \|\boldsymbol{v}\|_0 \qquad \text{for all } \mu \in W_j,\ \boldsymbol{v} \in \boldsymbol{V}_i, \tag{3.6}$$

and

$$(\operatorname{div} \boldsymbol{u}, \operatorname{div} \boldsymbol{v}) \le C\gamma^{j-i} h_j^{-1} \|\boldsymbol{u}\|_0 \| \operatorname{div} \boldsymbol{v}\| \qquad \text{for all } \boldsymbol{u} \in \boldsymbol{V}_j,\ \boldsymbol{v} \in \boldsymbol{V}_i. \tag{3.7}$$

Proof. Let $T \in \mathcal{T}_i$. In order to show (3.6) it is enough to show that

$$\int_T (\mathbf{curl}\,\mu) \cdot \boldsymbol{v}\,dx \le C\gamma^{j-i} h_j^{-1} \|\mu\|_{0,T} \|\boldsymbol{v}\|_{0,T},$$

where the subscript T indicates that the L^2-norms are defined with respect to the domain T.

For $\mu \in W_j$ we define $\mu_0 \in W_j$ by specifying its value at the usual nodal points of W_j, namely, μ_0 interpolates μ at the nodal points interior to T, but is zero at all other nodal points. In particular, μ_0 vanishes on ∂T, so

$$\begin{aligned}\int_T (\mathbf{curl}\,\mu_0) \cdot \boldsymbol{v}\,dx &= \int_T \mu_0 \operatorname{rot} \boldsymbol{v}\,dx \le \|\mu_0\|_{0,T} \| \operatorname{rot} \boldsymbol{v}\|_{0,T} \\ &\le C h_i^{-1} \|\mu\|_{0,T} \|\boldsymbol{v}\|_{0,T} \le C\gamma^{2(j-i)} h_j^{-1} \|\mu\|_{0,T} \|\boldsymbol{v}\|_{0,T}.\end{aligned}$$

Next, let $(\partial T)_j \subset T$ denote the union of all the triangles in $\mathcal{T}_j$ which are contained is T and meet ∂T. Considering the ratio of the areas of $(\partial T)_j$ and T, we see that $\|\boldsymbol{v}\|_{0,(\partial T)_j}^2 \le C\gamma^{2(j-i)} \|\boldsymbol{v}\|_{0,T}^2$, while

$$\|\mathbf{curl}(\mu - \mu_0)\|_{0,(\partial T)_j}^2 \le C \sum_x \left| \frac{\mu(x)}{h_j} \right|^2 h_j^2 \le C h_j^{-2} \|\mu\|_{0,(\partial T)_j}^2,$$

where the sum ranges of nodal points x belonging to ∂T. Thus

$$\int_T \mathbf{curl}(\mu - \mu_0) \cdot \boldsymbol{v}\,dx \le \|\mathbf{curl}(\mu - \mu_0)\|_{0,(\partial T)_j} \|\boldsymbol{v}\|_{0,(\partial T)_j} \le C\gamma^{j-i} h_j^{-1} \|\mu\|_{0,T} \|\boldsymbol{v}\|_{0,T},$$

which, combined with the preceding estimate, completes the proof of (3.6). The proof of (3.7) is completely similar. □

For $j = 1, 2, \ldots, J$ define operators $\boldsymbol{T}_j : \boldsymbol{V}_J \mapsto \boldsymbol{V}_j$ by $\boldsymbol{T}_j = \boldsymbol{R}_j \boldsymbol{Q}_j \boldsymbol{\Lambda} = \boldsymbol{R}_j \boldsymbol{\Lambda}_j \boldsymbol{P}_j$. Then $\boldsymbol{T}_j$ is symmetric and positive definite with respect to the bilinear form Λ and it follows from (3.3) that

$$\rho(\boldsymbol{T}_j) = \rho(\boldsymbol{R}_j \boldsymbol{\Lambda}_j) \le \omega < 2 \qquad \text{for } j = 1, 2, \ldots, J. \tag{3.8}$$

The operators $\boldsymbol{T}_j$ are useful in order to obtain a well-known algebraic characterization of the preconditioner $\boldsymbol{\Theta}$. The recurrence relation

$$\boldsymbol{I} - \boldsymbol{\Theta}_j \boldsymbol{\Lambda}_j \boldsymbol{P}_j = (\boldsymbol{I} - \boldsymbol{T}_j)(\boldsymbol{I} - \boldsymbol{\Theta}_{j-1} \boldsymbol{\Lambda}_{j-1} \boldsymbol{P}_{j-1})(\boldsymbol{I} - \boldsymbol{T}_j) \qquad \text{for } j = 1, 2, \ldots, J,$$

may be verified by considering separately its application to the elements of $\boldsymbol{V}_j$ and of its $\boldsymbol{H}(\mathrm{div})$–orthogonal complement, and invoking (3.2). In particular, this implies that $\boldsymbol{I} - \boldsymbol{\Theta} \boldsymbol{\Lambda} = \boldsymbol{E}_J \boldsymbol{E}_J^*$, where the operators $\boldsymbol{E}_j$ are defined by

$$\boldsymbol{E}_0 = \boldsymbol{I}, \quad \boldsymbol{E}_j = (\boldsymbol{I} - \boldsymbol{T}_j)(\boldsymbol{I} - \boldsymbol{T}_{j-1}) \ldots (\boldsymbol{I} - \boldsymbol{T}_1) \qquad \text{for } j = 1, 2, \ldots, J,$$

and $\boldsymbol{E}_J^*$ denotes the adjoint of $\boldsymbol{E}_J$ with respect to the $\boldsymbol{H}(\mathrm{div})$ inner product. Hence, to prove Theorem 3.1, it is suffices to show that

$$|||\boldsymbol{I} - \boldsymbol{\Theta} \boldsymbol{\Lambda}||| = |||\boldsymbol{E}_J|||^2 < \delta^2 < 1, \tag{3.9}$$

where δ is independent of h and J. Here $|||\cdot|||$ denotes the $\boldsymbol{H}(\mathrm{div})$ operator norm. It also follows from the spectral bound (3.8) and from the abstract multilevel theory (cf., for example, Lemma 4.3 of [8]) that

$$\Lambda(\boldsymbol{E}_J \boldsymbol{u}, \boldsymbol{E}_J \boldsymbol{u}) \le \Lambda(\boldsymbol{u}, \boldsymbol{u}) - (2 - \omega) \sum_{j=1}^{J} \Lambda(\boldsymbol{T}_j \boldsymbol{E}_{j-1} \boldsymbol{u}, \boldsymbol{E}_{j-1} \boldsymbol{u}). \tag{3.10}$$

If we can show that

$$\Lambda(\boldsymbol{u}, \boldsymbol{u}) \le c \sum_{j=1}^{J} \Lambda(\boldsymbol{T}_j \boldsymbol{E}_{j-1} \boldsymbol{u}, \boldsymbol{E}_{j-1} \boldsymbol{u}) \qquad \text{for all } \boldsymbol{u} \in \boldsymbol{V}_J, \tag{3.11}$$

with c independent of h and J, then we can combine with (3.10) to obtain

$$\Lambda(\boldsymbol{E}_J \boldsymbol{u}, \boldsymbol{E}_J \boldsymbol{u}) \le \Lambda(\boldsymbol{u}, \boldsymbol{u})[1 - (2 - \omega)/c],$$

from which (3.9) (and Theorem 3.1) follows.

To establish (3.11), we begin by using the discrete Helmholtz decomposition with respect to the space $\boldsymbol{V}$ to write

$$\boldsymbol{u} = \mathbf{grad}_J\, p + \mathbf{curl}\,\mu, \quad \text{where } p \in S,\ \mu \in W.$$

Let $\boldsymbol{v} = \boldsymbol{v}^J = \mathbf{grad}_J\, p$ and decompose $\boldsymbol{v}$ as

$$\boldsymbol{v} = \sum_{j=1}^{J}(\boldsymbol{v}^j - \boldsymbol{v}^{j-1}) = \sum_{j=1}^{J}(\boldsymbol{v}^j - \hat{\boldsymbol{v}}^j) + \sum_{j=1}^{J}(\hat{\boldsymbol{v}}^j - \boldsymbol{v}^{j-1}),$$

where $\boldsymbol{v}^j$ and $\hat{\boldsymbol{v}}^j$ are defined by

$$\begin{aligned} \boldsymbol{v}^j &\in \mathbf{grad}_j\, S_j, \qquad \operatorname{div} \boldsymbol{v}^j = Q_j^S \operatorname{div} \boldsymbol{v}, \\ \hat{\boldsymbol{v}}^j &\in \mathbf{grad}_j\, S_j, \qquad \operatorname{div} \hat{\boldsymbol{v}}^j = Q_{j-1}^S \operatorname{div} \boldsymbol{v}, \end{aligned}$$

for $j = 1, 2, \ldots, J$, and $\boldsymbol{v}^0 = 0$. It follows from Lemma 3.2 that

$$\|\boldsymbol{v}^j - \hat{\boldsymbol{v}}^j\| \le C h_j \|(Q_j^S - Q_{j-1}^S) \operatorname{div} \boldsymbol{v}\|. \tag{3.12}$$

Furthermore, since $\operatorname{div}(\hat{\boldsymbol{v}}^j - \boldsymbol{v}^{j-1}) = 0$, we may write $\sum_{j=1}^{J}(\hat{\boldsymbol{v}}^j - \boldsymbol{v}^{j-1}) = \mathbf{curl}\,\rho$, for a suitable $\rho \in W$. Setting $\boldsymbol{w} = \sum_{j=1}^{J}(\boldsymbol{v}^j - \hat{\boldsymbol{v}}^j)$, we then have the decomposition

$$\boldsymbol{u} = \boldsymbol{w} + \mathbf{curl}(\mu + \rho) = \boldsymbol{w} + \mathbf{curl}\,\phi,$$

where $\phi = \mu + \rho \in W$. Hence, we have

$$\Lambda(\boldsymbol{u}, \boldsymbol{u}) = \Lambda(\boldsymbol{w}, \boldsymbol{u}) + \Lambda(\mathbf{curl}\,\phi, \boldsymbol{u}). \tag{3.13}$$

From the definition of the operators $\boldsymbol{E}_j$ we obtain the identity

$$\boldsymbol{I} = \boldsymbol{E}_j + \sum_{k=1}^{j} \boldsymbol{T}_k \boldsymbol{E}_{k-1} \tag{3.14}$$

and this implies

$$\begin{aligned} \Lambda(\boldsymbol{w}, \boldsymbol{u}) &= \sum_{j=1}^{J} \Lambda(\boldsymbol{v}^j - \hat{\boldsymbol{v}}^j, \boldsymbol{u}) \\ &= \sum_{j=1}^{J} \Lambda(\boldsymbol{v}^j - \hat{\boldsymbol{v}}^j, \boldsymbol{E}_{j-1}\boldsymbol{u}) + \sum_{j=2}^{J}\sum_{k=1}^{j-1} \Lambda(\boldsymbol{v}^j - \hat{\boldsymbol{v}}^j, \boldsymbol{T}_k \boldsymbol{E}_{k-1}\boldsymbol{u}). \end{aligned}$$

Now, from the definitions of the operators $\mathbf{\Lambda}$ and $\boldsymbol{T}_j$ we have

$$\begin{aligned}
\sum_{j=1}^{J} \Lambda(\boldsymbol{v}^j - \hat{\boldsymbol{v}}^j, \boldsymbol{E}_{j-1}\boldsymbol{u}) &= \sum_{j=1}^{J} (\boldsymbol{v}^j - \hat{\boldsymbol{v}}^j, \boldsymbol{Q}_j \mathbf{\Lambda} \boldsymbol{E}_{j-1}\boldsymbol{u}) \\
&= \sum_{j=1}^{J} (\boldsymbol{R}_j^{-1/2}[\boldsymbol{v}^j - \hat{\boldsymbol{v}}^j], \boldsymbol{R}_j^{1/2} \boldsymbol{Q}_j \mathbf{\Lambda} \boldsymbol{E}_{j-1}\boldsymbol{u}) \\
&\le \left\{ \sum_{j=1}^{J} (\boldsymbol{R}_j^{-1}[\boldsymbol{v}^j - \hat{\boldsymbol{v}}^j], [\boldsymbol{v}^j - \hat{\boldsymbol{v}}^j]) \right\}^{1/2} \left\{ \sum_{j=1}^{J} \Lambda(\boldsymbol{T}_j \boldsymbol{E}_{j-1}\boldsymbol{u}, \boldsymbol{E}_{j-1}\boldsymbol{u}) \right\}^{1/2}.
\end{aligned}$$

Using (3.4) and (3.12) we obtain

$$\begin{aligned}
\sum_{j=1}^{J} (\boldsymbol{R}_j^{-1}[\boldsymbol{v}^j - \hat{\boldsymbol{v}}^j], [\boldsymbol{v}^j - \hat{\boldsymbol{v}}^j]) &\le C \sum_{j=1}^{J} h_j^{-2} \|\boldsymbol{v}^j - \hat{\boldsymbol{v}}^j\|^2 \\
&\le C \sum_{j=1}^{J} \|(Q_j^S - Q_{j-1}^S) \operatorname{div} \boldsymbol{v}\|^2 = C \|\operatorname{div} \boldsymbol{v}\|^2.
\end{aligned}$$

Hence,

$$\sum_{j=1}^{J} \Lambda(\boldsymbol{v}^j - \hat{\boldsymbol{v}}^j, \boldsymbol{E}_{j-1}\boldsymbol{u}) \le C \|\operatorname{div} \boldsymbol{v}\| \left\{ \sum_{j=1}^{J} \Lambda(\boldsymbol{T}_j \boldsymbol{E}_{j-1}\boldsymbol{u}, \boldsymbol{E}_{j-1}\boldsymbol{u}) \right\}^{1/2}.$$

Now from (3.7) and (3.12), and the fact that $\rho(\boldsymbol{T}_k) \le \omega$, we obtain

$$\begin{aligned}
&\sum_{j=2}^{J} \sum_{k=1}^{j-1} \Lambda(\boldsymbol{v}^j - \hat{\boldsymbol{v}}^j, \boldsymbol{T}_k \boldsymbol{E}_{k-1}\boldsymbol{u}) \\
&= \sum_{j=2}^{J} \sum_{k=1}^{j-1} \left[(\boldsymbol{v}^j - \hat{\boldsymbol{v}}^j, \boldsymbol{T}_k \boldsymbol{E}_{k-1}\boldsymbol{u}) + (\operatorname{div}[\boldsymbol{v}^j - \hat{\boldsymbol{v}}^j], \operatorname{div} \boldsymbol{T}_k \boldsymbol{E}_{k-1}\boldsymbol{u}) \right] \\
&\le \sum_{j=2}^{J} \sum_{k=1}^{j-1} \|(Q_j^S - Q_{j-1}^S) \operatorname{div} \boldsymbol{v}\| \left[h_j \|\boldsymbol{T}_k \boldsymbol{E}_{k-1}\boldsymbol{u}\| + c\gamma^{j-k} \|\operatorname{div} \boldsymbol{T}_k \boldsymbol{E}_{k-1}\boldsymbol{u}\| \right] \\
&\le \sum_{j=2}^{J} \sum_{k=1}^{j-1} C\gamma^{j-k} \|(Q_j^S - Q_{j-1}^S) \operatorname{div} \boldsymbol{v}\| \|\boldsymbol{T}_k \boldsymbol{E}_{k-1}\boldsymbol{u}\|_{\boldsymbol{H}(\operatorname{div})}.
\end{aligned}$$

Using the elementary inequality

$$\sum_{j=1}^{\infty} \sum_{k=1}^{j} \gamma^{j-k} a_j b_k \le C_\gamma \left(\sum_{j=1}^{\infty} a_j^2 \right)^{1/2} \left(\sum_{k=1}^{\infty} b_k^2 \right)^{1/2},$$

we can bound the right hand side by

$$C\left\{\sum_{j=1}^{J}\|(Q_j^S-Q_{j-1}^S)\operatorname{div}\boldsymbol{v}\|^2\right\}^{1/2}\left\{\sum_{k=1}^{J}\Lambda(\boldsymbol{T}_k\boldsymbol{E}_{k-1}\boldsymbol{u},\boldsymbol{T}_k\boldsymbol{E}_{k-1}\boldsymbol{u})\right\}^{1/2}$$
$$\le C\omega^{1/2}\|\operatorname{div}\boldsymbol{v}\|\left\{\sum_{k=1}^{J}\Lambda(\boldsymbol{T}_k\boldsymbol{E}_{k-1}\boldsymbol{u},\boldsymbol{E}_{k-1}\boldsymbol{u})\right\}^{1/2}.$$

Since $\|\operatorname{div}\boldsymbol{v}\|^2=\|\operatorname{div}\boldsymbol{u}\|^2\le\Lambda(\boldsymbol{u},\boldsymbol{u})$, we can combine these results, to get

$$\Lambda(\boldsymbol{w},\boldsymbol{u})\le C\left\{\Lambda(\boldsymbol{u},\boldsymbol{u})\right\}^{1/2}\left\{\sum_{k=1}^{J}\Lambda(\boldsymbol{T}_k\boldsymbol{E}_{k-1}\boldsymbol{u},\boldsymbol{E}_{k-1}\boldsymbol{u})\right\}^{1/2}. \tag{3.15}$$

Next, we consider the second term $\Lambda(\mathbf{curl}\,\phi,\boldsymbol{u})$ on the right hand side of (3.13). Let $\phi_j=(Q_j^W-Q_{j-1}^W)\phi$. From the approximation property of the space W_j it follows that

$$\|\phi_j\|\le Ch_j\|\mathbf{curl}\,\phi_j\|. \tag{3.16}$$

We also recall the following result (cf., for example, Lemma 6.7 in [8]):

$$\sum_{j=1}^{J}\|\mathbf{curl}\,\phi_j\|^2\le C\|\mathbf{curl}\,\phi\|^2. \tag{3.17}$$

From the identity (3.14), we have

$$\Lambda(\mathbf{curl}\,\phi,\boldsymbol{u})=\sum_{j=1}^{J}\Lambda(\mathbf{curl}\,\phi_j,\boldsymbol{u})=\sum_{j=1}^{J}\Lambda(\mathbf{curl}\,\phi_j,\boldsymbol{E}_{j-1}\boldsymbol{u})+\sum_{j=2}^{J}\sum_{k=1}^{j-1}\Lambda(\mathbf{curl}\,\phi_j,\boldsymbol{T}_k\boldsymbol{E}_{k-1}\boldsymbol{u}).$$

Proceeding as before, we obtain

$$\sum_{j=1}^{J}\Lambda(\mathbf{curl}\,\phi_j,\boldsymbol{E}_{j-1}\boldsymbol{u})=\sum_{j=1}^{J}(\boldsymbol{R}_j^{-1/2}\,\mathbf{curl}\,\phi_j,\boldsymbol{R}_j^{1/2}\boldsymbol{Q}_j\boldsymbol{\Lambda}\boldsymbol{E}_{j-1}\boldsymbol{u})$$
$$\le\left\{\sum_{j=1}^{J}(\boldsymbol{R}_j^{-1}\,\mathbf{curl}\,\phi_j,\mathbf{curl}\,\phi_j)\right\}^{1/2}\left\{\sum_{j=1}^{J}\Lambda(\boldsymbol{T}_j\boldsymbol{E}_{j-1}\boldsymbol{u},\boldsymbol{E}_{j-1}\boldsymbol{u})\right\}^{1/2}.$$

Now using (3.5) and (3.17), we get $\sum_{j=1}^{J}(\boldsymbol{R}_j^{-1}\,\mathbf{curl}\,\phi_j,\mathbf{curl}\,\phi_j)\le C\|\mathbf{curl}\,\phi\|^2$. Furthermore, using the strengthened Cauchy–Schwarz inequality (3.6) and (3.16), we obtain

$$\begin{aligned}\sum_{j=2}^{J}\sum_{k=1}^{j-1}\Lambda(\mathbf{curl}\,\phi_j,\boldsymbol{T}_k\boldsymbol{E}_{k-1}\boldsymbol{u}) &= \sum_{j=2}^{J}\sum_{k=1}^{j-1}(\mathbf{curl}\,\phi_j,\boldsymbol{T}_k\boldsymbol{E}_{k-1}\boldsymbol{u})\\ &\le C\sum_{j=2}^{J}\sum_{k=1}^{j-1}\gamma^{j-k}h_j^{-1}\|\phi_j\|\|\boldsymbol{T}_k\boldsymbol{E}_{k-1}\boldsymbol{u}\|\\ &\le C\sum_{j=2}^{J}\sum_{k=1}^{j-1}\gamma^{j-k}\|\,\mathbf{curl}\,\phi_j\|\|\boldsymbol{T}_k\boldsymbol{E}_{k-1}\boldsymbol{u}\|\\ &\le C\left\{\sum_{j=1}^{J}\|\,\mathbf{curl}\,\phi_j\|^2\right\}^{1/2}\left\{\sum_{k=1}^{J}\Lambda(\boldsymbol{T}_k\boldsymbol{E}_{k-1}\boldsymbol{u},\boldsymbol{T}_k\boldsymbol{E}_{k-1}\boldsymbol{u})\right\}^{1/2}\\ &\le C\|\,\mathbf{curl}\,\phi\|\omega^{1/2}\left\{\sum_{k=1}^{J}\Lambda(\boldsymbol{T}_k\boldsymbol{E}_{k-1}\boldsymbol{u},\boldsymbol{E}_{k-1}\boldsymbol{u})\right\}^{1/2}.\end{aligned}$$

Combining these results, we obtain

$$\Lambda(\mathbf{curl}\,\phi,\boldsymbol{u})\le C\|\,\mathbf{curl}\,\phi\|\left\{\sum_{k=1}^{J}\Lambda(\boldsymbol{T}_k\boldsymbol{E}_{k-1}\boldsymbol{u},\boldsymbol{E}_{k-1}\boldsymbol{u})\right\}^{1/2}. \tag{3.18}$$

If we can show that $\|\,\mathbf{curl}\,\phi\|\le C\left\{\Lambda(\boldsymbol{u},\boldsymbol{u})\right\}^{1/2}$, then (3.13), (3.15), and (3.18) will imply the desired estimate (3.11). Observe that

$$\|\,\mathbf{curl}\,\phi\|^2=\|\boldsymbol{u}-\boldsymbol{w}\|^2\le 2(\|\boldsymbol{u}\|^2+\|\boldsymbol{w}\|^2).$$

Therefore, to complete the proof it remains to show that $\|\boldsymbol{w}\|^2\le C\Lambda(\boldsymbol{u},\boldsymbol{u})$. However, by (3.12) and the fact that $\sum h_j^2\le\sum\gamma^{4j}\le C_\gamma$, we obtain

$$\begin{aligned}\|\boldsymbol{w}\|^2 &\le\left(\sum_{j=1}^{J}\|\boldsymbol{v}^j-\hat{\boldsymbol{v}}^j\|\right)^2\le C\left[\sum_{j=1}^{J}h_j\|(Q_j^S-Q_{j-1}^S)\operatorname{div}\boldsymbol{v}\|\right]^2\\ &\le C\left(\sum_{j=1}^{J}h_j^2\right)\left[\sum_{j=1}^{J}\|(Q_j^S-Q_{j-1}^S)\operatorname{div}\boldsymbol{v}\|^2\right]\le C\|\operatorname{div}\boldsymbol{v}\|^2\le C\|\operatorname{div}\boldsymbol{u}\|^2.\end{aligned}$$

Hence the theorem is established. □

4. Smoothing operators. To complete the description of the multigrid algorithm, we must construct proper L^2–symmetric, positive definite smoothing operators $\boldsymbol{R}_j : \boldsymbol{V}_j \to \boldsymbol{V}_j$ satisfying the Assumptions A1 and A2. Let us first observe that a simple Richardson smoother of the form $\boldsymbol{R}_j = \alpha_j \boldsymbol{I}$, where α_j is a positive constant, will not have this property. This is because (3.3) will imply a bound of the form $\alpha_j \le c h_j^2$, where c is independent of h and j. On the other hand, (3.5) will only be satisfied if α_j is uniformly bounded away from zero. In fact, this observation is consistent with numerical observations done by Cai, Goldstein, and Pasciak [4]. They observed that, in general, standard multigrid methods may not lead to a uniform preconditioner for the operator $\boldsymbol{\Lambda}$.

We now show how to define additive Schwarz and multiplicative Schwarz smoothers that do satisfy A1 and A2. These are the same smoothers as were used in [1] and analyzed in the convex case, and the construction is described in more detail there.

To define the additive Schwarz smoother, let $\mathcal{N}_j$ be the set of vertices in the triangulation $\mathcal{T}_j$, and for each $\nu \in \mathcal{N}_j$ let $\mathcal{T}_{j,\nu}$ be the set of triangles in $\mathcal{T}_j$ meeting at the vertex ν. These form a triangulation of a small subdomain, $\Omega_{j,\nu}$, and the family of domains $\{\Omega_{j,\nu}\}_{\nu\in\mathcal{N}_j}$ form an overlapping covering of Ω. Let $\boldsymbol{V}_{j,\nu}$, $W_{j,\nu}$, and $S_{j,\nu}$ be the subsets of functions in $\boldsymbol{V}_j$, W_j, and S_j, respectively, which are supported in $\bar{\Omega}_{j,\nu}$. The discrete operator $\boldsymbol{\Lambda}_{j,\nu} : \boldsymbol{V}_{j,\nu} \to \boldsymbol{V}_{j,\nu}$ and the $\boldsymbol{L}^2$ and $\boldsymbol{H}(\mathrm{div})$ projections, $\boldsymbol{Q}_{j,\nu}$ and $\boldsymbol{P}_{j,\nu}$, onto $\boldsymbol{V}_{j,\nu}$ are defined in the obvious way. The additive Schwarz operator with respect to the decomposition $\{\boldsymbol{V}_{j,\nu}\}$ of $\boldsymbol{V}_j$ is the operator $\boldsymbol{B}_j : \boldsymbol{V}_j \to \boldsymbol{V}_j$ given by

$$\boldsymbol{B}_j = \sum_{\nu\in\mathcal{N}_j} \boldsymbol{P}_{j,\nu}\Lambda_j^{-1} \equiv \sum_{\nu\in\mathcal{N}_j} \Lambda_{j,\nu}^{-1} Q_{j,\nu},$$

and the additive smoother we shall use is $\boldsymbol{R}_j = \eta \boldsymbol{B}_j$, where $\eta > 0$ is a scaling factor. As discussed in [1], $\boldsymbol{R}_j$ can be easily and efficiently evaluated.

THEOREM 4.1. *If the scaling factor $\eta \in (0, 2/3)$, then the additive Schwarz smoother $\boldsymbol{R}_j$ is L^2-symmetric, positive definite and satisfies Assumptions A1 and A2.*

Proof. It is clear that $\boldsymbol{R}_j$ is L^2-symmetric and positive definite. Furthermore, by a standard argument from domain decomposition theory, the spectral radius $\rho(\boldsymbol{B}_j\boldsymbol{\Lambda}_j)$ is bounded by the maximum number of overlaps in the covering $\{\Omega_{j,\nu}\}_{\nu\in\mathcal{N}_J}$, namely 3 (cf. the proof of Theorem 4.1 in [1]). Hence, $\rho(\boldsymbol{R}_j\boldsymbol{\Lambda}_j) \le 3\eta$ for $j = 2, 3, \dots, J$, and Assumption A1 is satisfied.

In order to establish Assumption A2 for the additive smoother, we will use the fact that

$$(\boldsymbol{B}_j^{-1}\boldsymbol{v}, \boldsymbol{v}) = \inf_{\substack{\boldsymbol{v}_\nu \in \boldsymbol{V}_{j,\nu} \\ \sum_\nu \boldsymbol{v}_\nu = \boldsymbol{v}}} \sum_{\nu\in\mathcal{N}_j} \Lambda(\boldsymbol{v}_\nu, \boldsymbol{v}_\nu). \tag{4.1}$$

This property is frequently used in the analysis of domain decomposition methods, and a proof can be found in Appendix B of [1].

Given $\boldsymbol{v} \in \boldsymbol{V}_j$, we can construct by local interpolation, elements $\boldsymbol{v}_\nu \in \boldsymbol{V}_{j,\nu}$ such that $\boldsymbol{v} = \sum_\nu \boldsymbol{v}_\nu$ and

$$\sum_{\nu \in \mathcal{N}_j} \|\boldsymbol{v}_\nu\|^2 \le C\|\boldsymbol{v}\|^2.$$

Using an inverse inequality we have

$$\sum_{\nu \in \mathcal{N}_j} \Lambda(\boldsymbol{v}_\nu, \boldsymbol{v}_\nu) \le C \sum_{\nu \in \mathcal{N}_j} h_j^{-2}\|\boldsymbol{v}_\nu\|^2 \le C h_j^{-2}\|\boldsymbol{v}\|^2.$$

In view of (4.1) this establishes (3.4).

Next, let $\boldsymbol{v} = \mathbf{curl}\,\phi_j$, with $\phi_j = (Q_j^W - Q_{j-1}^W)\phi$ for some $\phi \in W_j$. Then (3.16) holds and, again by local interpolation, we can find $\phi_{j,\nu} \in W_{j,\nu}$ such that

$$\phi_j = \sum_\nu \phi_{j,\nu}, \qquad \sum_{\nu \in \mathcal{N}_j} \|\phi_{j,\nu}\|^2 \le C\|\phi_j\|^2.$$

Let $\boldsymbol{v}_{j,\nu} = \mathbf{curl}\,\phi_{j,\nu}$. Then

$$\sum_{\nu \in \mathcal{N}_j} \Lambda(\boldsymbol{v}_{j,\nu}, \boldsymbol{v}_{j,\nu}) = \sum_{\nu \in \mathcal{N}_j} \|\,\mathbf{curl}\,\phi_{j,\nu}\|^2$$
$$\le C h_j^{-2} \sum_{\nu \in \mathcal{N}_j} \|\phi_{j,\nu}\|^2 \le C h_j^{-2}\|\phi_j\|^2 \le C\|\,\mathbf{curl}\,\phi_j\|^2 = C\|\boldsymbol{v}\|^2. \ \square$$

Finally we mention that the assumptions are satisfied as well by the symmetric multiplicative Schwarz smoother $\widetilde{\boldsymbol{R}}_j : \boldsymbol{V}_j \to \boldsymbol{V}_j$ with respect to the decomposition $\{\boldsymbol{V}_{j,\nu}\}$. For the precise definition of this operator see [1], where we also show (in Corollary 4.3 and the last equation of §5) that

$$\Lambda(\widetilde{\boldsymbol{R}}_j \boldsymbol{\Lambda}_j \boldsymbol{v}, \boldsymbol{v}) \le \Lambda(\boldsymbol{v}, \boldsymbol{v}), \qquad (\widetilde{\boldsymbol{R}}_j^{-1} \boldsymbol{v}, \boldsymbol{v}) \le C\Lambda(\boldsymbol{v}, \boldsymbol{v}).$$

Assumption A1 follow easily from the first of these equations, and Assumption A2 from the second.

Acknowledgements. The authors are grateful to Professors J. Bramble, J. Pasciak, and J. Xu for many useful discussions.

REFERENCES

[1] D. N. Arnold, R. S. Falk, and R. Winther, *Preconditioning in* $\boldsymbol{H}(\mathrm{div})$ *and applications*, Math. Comp., 66 (1997 (to appear)).

[2] J. H. Bramble, *Multigrid methods*, Pitman Research Notes in Mathematics Series, 294 (1993).

[3] F. Brezzi, M. Fortin, and R. Stenberg, *Error analysis of mixed-interpolated elements for Reissner-Mindlin plates*, Math. Models and Methods in Applied Sciences, 1 (1991), pp. 125-151.

[4] Z. Cai, C. I. Goldstein and J. E. Pasciak, *Multilevel iteration for mixed finite element systems with penalty*, SIAM J. Sci. Comput., 14 (1993), pp. 1072–1088.

[5] Z. Cai, R. Lazarov, T. Manteuffel, and S. McCormick, *First-order system least squares for second-order partial differential equations: Part I*, SIAM J. Numer. Anal., 31 (1994), pp. 1785-1799.

[6] P. Lin, *A sequential regularization method for time-dependent incompressible Navier–Stokes equations*, to appear in SIAM J. Numer. Anal. (1997).

[7] A. I. Pehlivanov, G. F. Carey, and R. D. Lazarov, *Least-squares mixed finite elements for second-order elliptic problems*, SIAM J. Numer. Anal., 31 (1994), pp. 1368-1377.

[8] J. Xu, *Iterative methods by space decomposition and subspace correction*, SIAM Rev., 34 (1992), pp. 581-613.