SIAM J. NUMER. ANAL.
Vol. 35, No. 1, pp. 320–335, February 1998

© 1998 Society for Industrial and Applied Mathematics
016

FIRST-ORDER SYSTEM LEAST SQUARES (FOSLS) FOR PLANAR LINEAR ELASTICITY: PURE TRACTION PROBLEM*

ZHIQIANG CAI†, THOMAS A. MANTEUFFEL‡, STEPHEN F. MCCORMICK‡, AND SEYMOUR V. PARTER§

Abstract. This paper develops two first-order system least-squares (FOSLS) approaches for the solution of the pure traction problem in planar linear elasticity. Both are *two-stage* algorithms that first solve for the gradients of displacement (which immediately yield deformation and stress), then for the displacement itself (if desired). One approach, which uses L^2 norms to define the FOSLS functional, is shown under certain H^2 regularity assumptions to admit *optimal* H^1*-like performance* for standard finite element discretization and standard multigrid solution methods that is *uniform in the Poisson ratio for all variables.* The second approach, which is based on H^{-1} norms, is shown under general assumptions to admit optimal uniform performance for displacement flux in an L^2 norm and for displacement in an H^1 norm. These methods do not degrade as other methods generally do when the material properties approach the incompressible limit.

Key words. least-squares discretization, multigrid, linear elasticity, pure traction, Poisson ratio

AMS subject classifications. 65F10, 65F30

PII. S0036142995294930

1. Introduction.

The basic equations of elasticity are generally in self-adjoint form, so they lend themselves naturally to an energy minimization principle, cast in terms of the primitive *displacement* variables. Unfortunately, this direct approach seems to have many practical difficulties (e.g., degrading approximation properties of the discretization and convergence properties of the solution process) as the material tends to become incompressible (i.e., the *Lamé constant* λ tends to infinity for fixed *Lamé constant* μ, or, more precisely, the *Poisson ratio* ν tends to 0.5^-). There have been several attempts to develop alternate approaches (cf. [2], [3], [11], [13], and [21]) that are robust in the incompressible limit, but these alternatives are usually based on *mixed* formulations (see also [4] and [12]) that lead to discrete equations that are difficult to solve. Indeed, little attention seems to have been paid to the development of robust solution strategies for the matrix equations that arise in this context. Compounding these difficulties is the fact that what is often needed in practice are the deformations and stresses. These variables can be obtained by differentiating displacements, but this weakens the *order* (from h^2 to h) and *strength* (from H^1 to L^2) of the approximation.

---

*Received by the editors November 20, 1995; accepted for publication (in revised form) September 13, 1996.

http://www.siam.org/journals/sinum/35-1/29493.html

†Department of Mathematics, Purdue University, 1395 Mathematical Science Building, West Lafayette, IN 47907-1395 (zcai@math.purdue.edu). This work was sponsored by the National Science Foundation under grant DMS-9619792.

‡Program in Applied Mathematics, Campus Box 526, University of Colorado at Boulder, Boulder, CO 80309-0526 (tmanteuf@boulder.colorado.edu, stevem@newton.colorado.edu). The second author was sponsored by the National Science Foundation under grant DMS-8704169 and the Department of Energy under grant DE-FG03-93ER25217. The third author was sponsored by the Air Force Office of Scientific Research under grant F49620-92-J-0439 and the National Science Foundation under grant DMS-8704169.

§Departments of Computer Science and Mathematics, University of Wisconsin, Madison, WI 53705. This author was supported by the National Science Foundation under grant DMS-9501256 (parter@cs.wisc.edu).

In [7], *first-order system least squares* (FOSLS) was applied to the pure displacement problem of linear elasticity. This was accomplished by introducing *displacement flux* (i.e., gradients of displacement) and *pressure* and by recasting the problem as a perturbed Stokes equation. For the pure traction problem treated here, we instead take an approach that involves displacement flux as the only new variable and that is more in the spirit of earlier work on scalar elliptic equations (cf. [5] and [6]).

The aim of this paper is to develop two simple FOSLS formulations of the pure traction problem in planar linear elasticity. The first formulation is based on a least-squares functional involving L^2 norms and, under certain H^2 regularity assumptions, it yields uniform and optimal H^1 approximations of all variables, including deformations, stresses, and displacements. Of course, such H^1 approximations of the deformations and stresses require that the displacement be smooth (i.e., in $H^{2+\alpha}$). The second FOSLS formulation involves H^{-1} norms and, under general assumptions, yields uniform and optimal L^2 estimates of the deformations and stresses as well as H^1 approximations of the displacements. We know of no other problem formulation that is able to obtain such uniform H^1 and L^2 optimality.

Both methods are *two-stage*, in which one solves for the displacement flux variable first. (Deformations and stresses can then be readily obtained as simple algebraic combinations of the displacement fluxes.) The displacement components can then be obtained as solutions of two scalar Poisson equations (if desired). Both methods yield estimates that suggest effective numerical approaches for the actual computation of these approximations. We will discuss this aspect of our results in some detail.

For completeness, we develop a third FOSLS approach based on a *single-stage* method. In this case, the functional to be minimized depends on both the displacement and displacement flux. The estimates associated with this functional indicate that, while the discretization properties are quite good, it may be difficult to use standard multigrid methods to solve for the discrete approximations near the incompressible limit.

The paper is organized as follows. Section 2 introduces the pure traction problem in planar linear elasticity, some equivalent formulations, and notation. Section 3 discusses FOSLS functionals based on the extended system (displacement and displacement flux as primary variables) and on the reduced system (displacement flux as the primary variable), and establishes their ellipticity and continuity. We introduce two-stage algorithms based on the reduced system in sections 4 and 5, and section 6 provides the proofs of some basic regularity estimates.

2. The elasticity problem, its first-order system formulations, and other preliminaries.

Let Ω be a bounded, open, connected domain in $\Re^2$ with Lipschitz boundary $\partial\Omega$. Denote the Lamé constants by

$$\lambda = \frac{E\nu}{(1+\nu)(1-2\nu)} \quad \text{and} \quad \mu = \frac{E}{2(1+\nu)},$$

where $E > 0$ is the *modulus of elasticity*, $\mu > 0$, generally $\lambda > -\frac{3}{2\mu}$, and $\nu = \lambda/2(\lambda+\mu) \in (-1, 1/2)$ is the *Poisson ratio* of the elastic material. We write the system of equations of linear elasticity for the *displacement* $\mathbf{u} = (u_1,\, u_2)^t$, with *pure traction* boundary conditions, as follows (cf. [9]):

$$\left\{ \begin{array}{rcll} -\mu\Delta\,\mathbf{u} - (\lambda+\mu)\nabla\nabla\cdot\mathbf{u} &=& \mathbf{f}, & \text{in} \quad \Omega, \\ \sum_{j=1}^{2}\sigma_{ij}(\mathbf{u})\,n_j &=& 0, & \text{on} \quad \partial\Omega, \quad 1 \le i \le 2, \end{array} \right. \tag{2.1}$$

where the symbols Δ, ∇, and $\nabla\cdot$ stand for the Laplacian, gradient, and divergence operators, respectively ($\Delta\mathbf{u}$ is the vector of components $\Delta\, u_i$); $\mathbf{f}$ is a given vector function; $\sigma_{ij}(\mathbf{u}) = \lambda(\nabla\cdot\mathbf{u})\delta_{ij} + 2\mu\epsilon_{ij}(\mathbf{u})$ is the *stress*, $\epsilon_{ij}(\mathbf{u}) = \frac{1}{2}(\partial_j u_i + \partial_i u_j)$ is the *deformation*, and δ_{ij} is the Kronecker delta symbol; and $\mathbf{n} = (n_1,\, n_2)^t$ is the outward unit normal on the boundary.

We use standard notation and definitions for the Sobolev spaces $H^s(\Omega)^d$, associated inner products $(\cdot,\,\cdot)_s$, and respective norms $\|\cdot\|_s$, $s \geq 0$. (We suppress the designations d and Ω on the inner products and norms because dependence on dimension and region is clear by context.) $H^0(\Omega)^d$ coincides with $L^2(\Omega)^d$, in which case the norm and inner product will be denoted by $\|\cdot\|$ and $(\cdot,\,\cdot)$, respectively. As usual, $H_0^s(\Omega)$ will denote the closure with respect to the norm $\|\cdot\|_s$ of the space of infinitely differentiable functions with compact support in Ω, and $L_0^2(\Omega)$ will denote the space of $L^2(\Omega)$ functions p such that $\int_\Omega p dx = 0$.

We use $H_0^{-1}(\Omega)$ and $H^{-1}(\Omega)$ to denote the dual spaces of $H_0^1(\Omega)$ and $H^1(\Omega)$ with norms defined by

$$\|\phi\|_{-1,\,0} = \sup_{0\neq\psi\in H_0^1(\Omega)} \frac{(\phi,\,\psi)}{\|\psi\|_1}$$

and

$$\|\phi\|_{-1} = \sup_{0\neq\psi\in H^1(\Omega)} \frac{(\phi,\,\psi)}{\|\psi\|_1},$$

respectively.

Let $\mathcal{N}$ denote the space of *infinitesimal rigid motions* (i.e., $\mathcal{N} = \{(a + cx_2, b - cx_1)^t\ :\ a, b, c \in \Re\}$), $\mathcal{N}^\perp$ its orthogonal complement in $L^2(\Omega)^2$, and $\mathcal{N}^c$ its complement in $H^1(\Omega)^2$ defined so that $\mathbf{u} \in \mathcal{N}^c$ if and only if

$$\int_\Omega u_1\, dx = \int_\Omega u_2\, dx = \int_\Omega (\partial_2 u_1 - \partial_1 u_2)\, dx = 0.$$

Then (the weak form of) boundary value problem (2.1) has a unique solution $\mathbf{u} \in \mathbf{X} \equiv H^1(\Omega)^2 \cap \mathcal{N}^c$ for any $\mathbf{f} \in H^{-1}(\Omega)^2 \cap \mathcal{N}^\perp$ (cf. [10]).

We will use standard curl notation for two dimensions by identifying $\Re^2$ with the $(x,\, y)$-plane in $\Re^3$. Thus, the curl of $\mathbf{u} = (u_1,\, u_2)^t$ means the scalar function

$$\nabla\times\mathbf{u} = \partial_1 u_2 - \partial_2 u_1,$$

and $\nabla^\perp$ denotes its formal adjoint:

$$\nabla^\perp q = \begin{pmatrix} \partial_2 q \\ -\partial_1 q \end{pmatrix}.$$

We will be introducing a new independent variable related to the 4-vector function of gradients of the u_i, $i = 1,\, 2$. It will be convenient to view the original 2-vector functions as column vectors and the new 4-vector functions as block column vectors. Thus, given

$$\mathbf{u} = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix},$$

an operator G defined on scalar functions is extended to 2-vectors componentwise:

$$G\mathbf{u} = \begin{pmatrix} Gu_1 \\ Gu_2 \end{pmatrix}.$$

For example, $\nabla \mathbf{u} = (\partial_1 u_1,\, \partial_2 u_1,\, \partial_1 u_2,\, \partial_2 u_2)^t$. If $\mathbf{U}_i \equiv G\, u_i$ is a vector function, then we write the block column vector

$$\mathbf{U} \equiv G\mathbf{u} = \begin{pmatrix} \mathbf{U}_1 \\ \mathbf{U}_2 \end{pmatrix}.$$

If D is an operator on vector functions, then its extension to block column vectors is defined by

$$D\mathbf{U} = \begin{pmatrix} D\mathbf{U}_1 \\ D\mathbf{U}_2 \end{pmatrix}.$$

For example, writing $\mathbf{U}_i = (U_{i1},\, U_{i2})^t$, then

$$\nabla \times \mathbf{U} = \begin{pmatrix} \partial_1 U_{12} - \partial_2 U_{11} \\ \partial_1 U_{22} - \partial_2 U_{21} \end{pmatrix}.$$

We also extend the respective normal and tangential operators $\mathbf{n}\cdot$ and $\mathbf{n}\times$ componentwise:

$$\mathbf{n} \cdot \mathbf{U} = \begin{pmatrix} \mathbf{n} \cdot \mathbf{U}_1 \\ \mathbf{n} \cdot \mathbf{U}_2 \end{pmatrix} \quad \text{and} \quad \mathbf{n} \times \mathbf{U} = \begin{pmatrix} \mathbf{n} \times \mathbf{U}_1 \\ \mathbf{n} \times \mathbf{U}_2 \end{pmatrix}.$$

Finally, inner products and norms on the block column vector functions are defined in the natural componentwise way:

$$\|\mathbf{U}\|^2 = \sum_{i=1}^{2} \|\mathbf{U}_i\|^2.$$

Let $A = \lambda A_1 + 2\mu A_2$ be the 4×4 matrix defined as follows:

$$A_1 = \mathbf{b}\,\mathbf{b}^t,$$

$$\mathbf{b} = (1,\, 0,\, 0,\, 1)^t,$$

and

$$A_2 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Then the elasticity equations in (2.1) may be rewritten in the compact form

$$\left\{ \begin{array}{rcll} -\nabla \cdot A\nabla \mathbf{u} & = & \mathbf{f}, & \text{in} \quad \Omega, \\ \mathbf{n} \cdot (A\nabla \mathbf{u}) & = & \mathbf{0}, & \text{on} \quad \partial\Omega. \end{array} \right. \tag{2.2}$$

We introduce the *displacement flux* variable $\mathbf{U} = \nabla \mathbf{u}$, that is,

$$\mathbf{U} = (U_1,\, U_2,\, U_3,\, U_4)^t = (\partial_1 u_1,\, \partial_2 u_1,\, \partial_1 u_2,\, \partial_2 u_2)^t. \tag{2.3}$$

Since the definition of $\mathbf{U}$ implies that

$$\nabla\times\mathbf{U}=\mathbf{0} \quad \text{in } \Omega, \tag{2.4}$$

then a system that is equivalent to (2.2) is

$$\begin{cases} \mathbf{U}-\nabla\mathbf{u} &= \mathbf{0}, & \text{in} & \Omega, \\ -\nabla\cdot A\mathbf{U} &= \mathbf{f}, & \text{in} & \Omega, \\ \nabla\times\mathbf{U} &= \mathbf{0}, & \text{in} & \Omega, \\ \mathbf{n}\cdot A\mathbf{U} &= \mathbf{0}, & \text{on} & \partial\Omega. \end{cases} \tag{2.5}$$

We will show in the next section that this extended system is well posed and suitable for treatment by FOSLS. However, what is probably more important in practice is the system that involves $\mathbf{U}$ only:

$$\begin{cases} -\nabla\cdot A\mathbf{U} &= \mathbf{f}, & \text{in} & \Omega, \\ \nabla\times\mathbf{U} &= \mathbf{0}, & \text{in} & \Omega, \\ \mathbf{n}\cdot A\mathbf{U} &= \mathbf{0}, & \text{on} & \partial\Omega. \end{cases} \tag{2.6}$$

We will show in the next section that this reduced system is also well posed and that it is perhaps better suited to FOSLS treatment, especially in the incompressible limit.

We define solution spaces for the primitive variables by

$$\mathcal{W}=\{\mathbf{u}\in H^1(\Omega)^2\cap\mathcal{N}^c : \nabla\cdot A\nabla\mathbf{u}\in L^2(\Omega)^2,\ \mathbf{n}\cdot A\nabla\mathbf{u}=\mathbf{0} \text{ on } \partial\Omega\}$$

and

$$\mathcal{Y}=H^1(\Omega)^2\cap\mathcal{N}^c.$$

Since we have posed (2.1) on the space for which $\nabla\mathbf{u}$ is orthogonal in $L^2(\Omega)^4$ to gradients of elements of $\mathcal{N}$ (we write $\nabla\mathbf{u}\in(\nabla\mathcal{N})^\perp$), we are at liberty to impose the condition that $\mathbf{U}\in(\nabla\mathcal{N})^\perp$ (i.e., $\int_\Omega(U_2-U_3)\,dx=0$). We thus define the solution space for the new variables by

$$\mathcal{U}=\{\mathbf{U}\in L^2(\Omega)^4\cap(\nabla\mathcal{N})^\perp\ :\ \nabla\cdot A\mathbf{U}\in L^2(\Omega)^2,\ \mathbf{n}\cdot A\mathbf{U}=\mathbf{0} \text{ on } \partial\Omega\}$$

for the case of general domain Ω and by

$$\mathcal{V}=\{\mathbf{U}\in H^1(\Omega)^4\cap(\nabla\mathcal{N})^\perp\ :\ \mathbf{n}\cdot A\mathbf{U}=\mathbf{0} \text{ on } \partial\Omega\}$$

for the case in which domain Ω is a convex polygon or has $C^{1,1}$ boundary. In the context of L^2 norms, $\mathcal{W}$ is a natural choice for (2.2) and $\mathcal{Y}$ is a natural choice for (2.5). Our theory will show that $\mathcal{U}$ is a natural choice for (2.6) in the context of H^{-1} norms, and that $\mathcal{V}$ is a natural choice for (2.5) or (2.6) in the context of L^2 norms.

3. FOSLS.

The primary objective of this section is to establish ellipticity of least-squares functionals based on (2.5) and (2.6) in appropriate Sobolev spaces. To this end, we assume that $\mathbf{f}\in L^2(\Omega)^2\cap\mathcal{N}^\perp$ and define the following:

$$G(\mathbf{U},\,\mathbf{u};\,\mathbf{f})=\|\mathbf{f}+\nabla\cdot A\mathbf{U}\|^2+\|\nabla\times\mathbf{U}\|^2+(A(\mathbf{U}-\nabla\mathbf{u}),\mathbf{U}-\nabla\mathbf{u}) \tag{3.1}$$

for $(\mathbf{U},\,\mathbf{u})\in\mathcal{V}\times\mathcal{Y}$;

$$G_{-1}(\mathbf{U};\,\mathbf{f})=\|\mathbf{f}+\nabla\cdot A\mathbf{U}\|_{-1}^2+\|\nabla\times\mathbf{U}\|_{-1,0}^2 \tag{3.2}$$

for $\mathbf{U} \in \mathcal{U}$; and

$$G_0(\mathbf{U};\, \mathbf{f}) = \|\mathbf{f} + \nabla \cdot A\mathbf{U}\|^2 + \|\nabla\times\mathbf{U}\|^2 \tag{3.3}$$

for $\mathbf{U} \in \mathcal{V}$.

In what follows, C, possibly with subscripts, will denote a generic constant that may vary in meaning with each occurrence and may depend on Ω, λ, and μ, but is independent of the Poisson ratio $\nu = \lambda/2(\lambda+\mu)$. We will frequently use the term *uniform* in reference to a relation to mean that it holds independent of ν.

We first establish uniform boundedness and ellipticity (i.e., equivalence) of the functionals $G_{-1}(\mathbf{U};\, \mathbf{0})$ and $G_0(\mathbf{U};\, \mathbf{0})$ in terms of the respective functionals $M_{-1}(\mathbf{U})$ and $M_0(\mathbf{U})$ defined on the respective spaces $\mathcal{U}$ and $\mathcal{V}$ by

$$M_{-1}(\mathbf{U}) = \|\mathbf{U}\|^2 + \lambda^2\|\mathrm{tr}\mathbf{U}\|^2$$

and

$$M_0(\mathbf{U}) = \|\mathbf{U}\|_1^2 + \lambda^2\|\nabla\mathrm{tr}\mathbf{U}\|^2,$$

where the *trace* operator tr is defined by $\mathrm{tr}\mathbf{U} = U_1 + U_4$. We then show that $G(\mathbf{U},\, \mathbf{u};\, \mathbf{0})$ is uniformly equivalent to the modified product H^1-type norm defined on $\mathcal{V} \times \mathcal{Y}$ by

$$M(\mathbf{U},\, \mathbf{u}) = \|\mathbf{U}\|_1^2 + \lambda^2\|\nabla\mathrm{tr}\mathbf{U}\|^2 + \lambda\,\|\mathrm{tr}\mathbf{U}\|^2 + \|A^{\frac{1}{2}}\nabla\mathbf{u}\|^2.$$

To this end, we appeal to standard H^1 regularity estimates (cf. [15, 16]) for elasticity equation (2.2):

$$\|\mathbf{v}\|_1 + \lambda\|\nabla\cdot\mathbf{v}\| \le C\,\|\nabla\cdot A\nabla\mathbf{v}\|_{-1} \tag{3.4}$$

for any $\mathbf{v} \in \mathcal{W}$. If the domain Ω is a convex polygon or its boundary is $C^{1,\,1}$, then we may appeal to standard H^2 regularity results (cf. [15, 16]):

$$\|\mathbf{v}\|_2 \le C\,\|\nabla\cdot A\nabla\mathbf{v}\| \tag{3.5}$$

for any $\mathbf{v} \in \mathcal{W} \cap H^2(\Omega)^2$. (If $\nabla \cdot A\nabla\mathbf{v} \in L^2(\Omega)^2$, then H^2 regularity implies that $\mathbf{v} \in H^2(\Omega)^2$.) Interestingly enough, we will also need similar estimates for the Stokes equations. In fact, for any constant $\rho > 0$, the usual Stokes H^1 and H^2 regularity results imply, respectively, that

$$\|\rho\mathbf{w}\|_1^2 + \|p\|^2 \le C\,\|-\rho\nabla\cdot\nabla\mathbf{w} + \nabla p\|_{-1,\,0}^2 \tag{3.6}$$

for any $p \in L_0^2(\Omega)$ and $\mathbf{w} \in H_0^1(\Omega)^2$ such that $\mathbf{w}$ is divergence free, and that (cf. [17] and [18])

$$\|\rho\mathbf{w}\|_2^2 + \|p\|_1^2 \le C\,\|-\rho\nabla\cdot\nabla\mathbf{w} + \nabla p\|^2 \tag{3.7}$$

for any $p \in H^1(\Omega)/\mathcal{R}$ and $\mathbf{w} \in H_0^1(\Omega)^2 \cap H^2(\Omega)^2$ such that $\mathbf{w}$ is divergence free. See section 6 for simple proofs of (3.4) and (3.6), which are provided for completeness.

THEOREM 3.1. *The functionals* $G_{-1}(\mathbf{U};\, \mathbf{0})$ *and* $M_{-1}(\mathbf{U})$ *satisfy the uniform equivalence relation*

$$\frac{1}{C}\,M_{-1}(\mathbf{U}) \le G_{-1}(\mathbf{U};\, \mathbf{0}) \le CM_{-1}(\mathbf{U}), \tag{3.8}$$

for all $\mathbf{U} \in \mathcal{U}$. *When* (3.5) *and* (3.7) *hold, the functionals* $G_0(\mathbf{U};\, \mathbf{0})$ *and* $M_0(\mathbf{U})$ *satisfy the uniform equivalence relation*

$$\frac{1}{C}\, M_0(\mathbf{U}) \le G_0(\mathbf{U};\, \mathbf{0}) \le C M_0(\mathbf{U}), \tag{3.9}$$

for all $\mathbf{U} \in \mathcal{V}$.

Proof. The upper bound in (3.8) for G_{-1} follows from the easily established bounds

$$\|\nabla \cdot A\mathbf{U}\|_{-1} \le \|A\mathbf{U}\| \quad \text{and} \quad \|\nabla \times \mathbf{U}\|_{-1,\,0} \le \|\mathbf{U}\|$$

from the triangle inequality and from noting that $A_1\mathbf{U} = (\mathrm{tr}\mathbf{U})\mathbf{b}$, where $\mathbf{b} = (1,\, 0,\, 0,\, 1)^t$. The upper bound in (3.9) for G_0 is a straightforward consequence of the triangle inequality and the fact that

$$\nabla \cdot A_1\mathbf{U} = \nabla \mathrm{tr}\mathbf{U}. \tag{3.10}$$

To show the validity of the lower bounds in (3.8) and (3.9) for the respective functional G_{-1} and G_0, we first show that

$$\|\mathbf{U}\|_1^2 \le C\, G_0(\mathbf{U};\, \mathbf{0}) \quad \forall\, \mathbf{U} \in \mathcal{V}. \tag{3.11}$$

To this end, write $\mathbf{U} \in \mathcal{V}$ as follows:

$$\mathbf{U} = \nabla \mathbf{v} + \mathbf{V} + \boldsymbol{\eta}, \tag{3.12}$$

where $\mathbf{v} \in \mathcal{W}$ solves (2.2) with $\mathbf{f} = -\nabla \cdot A\mathbf{U}$ and the decomposition $\mathbf{V} + \boldsymbol{\eta} = \mathbf{U} - \nabla \mathbf{v}$ is characterized by restricting $\mathbf{V}$ to the range of A_2 and $\boldsymbol{\eta}$ to the null space of A_2. Note that the second and third components of $\mathbf{V}$ must be equal and that $\boldsymbol{\eta} = (0,\, p,\, -p,\, 0)^t$ for some $p \in H^1(\Omega)/\mathcal{R}$ ($\int_\Omega p\, dx = 0$ because $\mathbf{U} \in (\nabla \mathcal{N})^\perp$). Hence, we have that

$$A_2\mathbf{V} = \mathbf{V} \tag{3.13}$$

and

$$\nabla \times \boldsymbol{\eta} = \nabla p. \tag{3.14}$$

Now, the definitions of A, A_1, A_2, and $\boldsymbol{\eta}$ imply that $A\boldsymbol{\eta} = A_1\boldsymbol{\eta} + A_2\boldsymbol{\eta} = 0$. Thus, since

$$\nabla \cdot A\mathbf{V} = \nabla \cdot A(\mathbf{V} + \boldsymbol{\eta}) = \nabla \cdot A\mathbf{U} - \nabla \cdot A\nabla \mathbf{v} = \mathbf{0}, \quad \text{in } \Omega,$$

and

$$\mathbf{n} \cdot\ A\mathbf{V} = \mathbf{0}, \quad \text{on } \partial\Omega,$$

it then follows (cf. [14]) that

$$A\mathbf{V} = \nabla^\perp \mathbf{w} \tag{3.15}$$

for some $\mathbf{w} \in H^1(\Omega)^2$. Comparing the second and third components of both sides of (3.15), we find that $\mathbf{w}$ is divergence free; i.e.,

$$\nabla \cdot \mathbf{w} = 0, \quad \text{in } \Omega.$$

Moreover,

$$\mathbf{n} \times \nabla \mathbf{w} = \mathbf{n} \cdot \nabla^{\perp} \mathbf{w} = \mathbf{n} \cdot A\mathbf{V} = \mathbf{0}, \quad \text{on } \partial\Omega,$$

and since Ω is simply connected, we may thus assume that $\mathbf{w}$ vanishes on the boundary $\partial\Omega$, i.e.,

$$\mathbf{w} = \mathbf{0}, \quad \text{on } \partial\Omega.$$

By taking the trace of both sides of (3.15) and using (3.13), we see that

$$\text{tr}\mathbf{V} = -\frac{1}{2(\lambda+\mu)} \nabla \times \mathbf{w}, \tag{3.16}$$

which implies that

$$2\mu\mathbf{V} = 2\mu A_2 \mathbf{V} = A\mathbf{V} - \lambda A_1 \mathbf{V} = \nabla^{\perp}\mathbf{w} - \lambda(\text{tr}\mathbf{V})\mathbf{b}$$

or

$$\mathbf{V} = \frac{1}{2\mu}\left(\nabla^{\perp}\mathbf{w} + \frac{\lambda}{2(\lambda+\mu)}(\nabla \times \mathbf{w})\mathbf{b}\right). \tag{3.17}$$

It now follows from $\nabla \times ((\nabla \times \mathbf{w})\,\mathbf{b}) = \nabla \cdot \nabla \mathbf{w}$ that

$$\nabla \times \mathbf{V} = -\rho \nabla \cdot \nabla \mathbf{w},$$

where

$$\rho = \frac{1}{2\mu}\left(1 - \frac{\lambda}{2(\lambda+\mu)}\right).$$

Combining this with (3.12) and (3.14) and using (3.5) and (3.7), we thus have that

$$\begin{aligned} G_0(\mathbf{U}; \mathbf{0}) &= \|\nabla \cdot A\mathbf{U}\|^2 + \|\nabla \times \mathbf{U}\|^2 \\ &= \|\nabla \cdot A\nabla\mathbf{v}\|^2 + \|\nabla \times (\mathbf{V} + \boldsymbol{\eta})\|^2 \\ &= \|\nabla \cdot A\nabla\mathbf{v}\|^2 + \| - \rho\nabla \cdot \nabla\mathbf{w} + \nabla p\|^2 \\ &\geq \frac{1}{C}\left(\|\mathbf{v}\|_2^2 + \|\rho\mathbf{w}\|_2^2 + \|p\|_1^2\right) \\ &\geq \frac{1}{C}\left(\|\nabla\mathbf{v}\|_1^2 + \|\mathbf{V}\|_1^2 + \|\boldsymbol{\eta}\|_1^2\right). \end{aligned}$$

Here we have used the bound

$$\|\mathbf{V}\|_1 \leq C \left\| \frac{1}{2\mu}\mathbf{w} \right\|_2,$$

which follows easily from (3.17). Thus, (3.11) follows from (3.12).

A similar argument based on (3.4) and (3.6) yields

$$G_{-1}(\mathbf{U}; \mathbf{0}) \geq \frac{1}{C}\|\mathbf{U}\|^2 \quad \forall\, \mathbf{U} \in \mathcal{U}. \tag{3.18}$$

We have thus established the lower bounds in (3.8) and (3.9) for the first terms in the respective M_{-1} and M_0. To establish these bounds for the second terms, first note that the definition of A_2 implies that

$$\|2\mu\nabla \cdot A_2\mathbf{U}\|^2 \leq C\,\|\mathbf{U}\|_1^2. \tag{3.19}$$

From (3.10), the definition of A, the triangle inequality, (3.19), and (3.11), we then have that

$$\begin{aligned} \|\lambda\nabla\mathrm{tr}\mathbf{U}\|^2 &\le C\left(\|\nabla\cdot A\mathbf{U}\|^2 + \|2\mu\nabla\cdot A_2\mathbf{U}\|^2\right) \\ &\le C\left(\|\nabla\cdot A\mathbf{U}\|^2 + \|\mathbf{U}\|_1^2\right) \\ &\le C\, G_0(\mathbf{U};\, \mathbf{0}). \end{aligned} \tag{3.20}$$

This proves the bound in (3.9) for the second term in M_0. To establish the bound in (3.8) for the second term in M_{-1}, using (3.12) and (3.16) (but now stemming from $\mathbf{U}\in\mathcal{U}$), we note that

$$\begin{aligned} \lambda A_1\mathbf{U} &= \lambda A_1\nabla\mathbf{v} + \lambda A_1\mathbf{V} \\ &= \left(\lambda\nabla\cdot\mathbf{v} - \frac{\lambda}{2(\lambda+\mu)}\nabla\times\mathbf{w}\right)\mathbf{b}. \end{aligned}$$

Hence,

$$\lambda\mathrm{tr}\mathbf{U} = \lambda\nabla\cdot\mathbf{v} - \frac{\lambda}{2(\lambda+\mu)}\nabla\times\mathbf{w}.$$

It then follows from (3.4) and (3.6) that

$$\begin{aligned} \lambda\|\mathrm{tr}\mathbf{U}\| &\le \|\lambda\nabla\cdot\mathbf{v}\| + \|\mathbf{w}\|_1 \\ &\le C\left(\|\nabla\cdot A\mathbf{U}\|_{-1} + \|-\rho\triangle\mathbf{w} + \nabla p\|_{-1,\,0}\right) \\ &= C\left(\|\nabla\cdot A\mathbf{U}\|_{-1} + \|\nabla\times\mathbf{U}\|_{-1,\,0}\right). \qquad \square \end{aligned}$$

Next, we use (3.9) to show uniform equivalence of the bilinear forms $G(\mathbf{U},\,\mathbf{u};\,\mathbf{0})$ and $M(\mathbf{U},\,\mathbf{u})$.

COROLLARY 3.1. *Assume that* (3.5) *and* (3.7) *hold. Then the functionals* $G(\mathbf{U},\,\mathbf{u};\,\mathbf{0})$ *and* $M(\mathbf{U},\,\mathbf{u})$ *satisfy the uniform equivalence relation*

$$\frac{1}{C}\,M(\mathbf{U},\,\mathbf{u}) \le G(\mathbf{U},\,\mathbf{u};\,\mathbf{0}) \le CM(\mathbf{U},\,\mathbf{u}) \tag{3.21}$$

for all $(\mathbf{U},\,\mathbf{u})\in\mathcal{V}\times\mathcal{Y}$.

Proof. The upper bound in (3.21) for G follows easily from the triangle inequality by noting that $\|A^{1/2}\mathbf{U}\|^2 = \lambda\|\mathrm{tr}\mathbf{U}\|^2 + 2\mu\|A_2^{1/2}\mathbf{U}\|^2$ and using (3.9). To prove the lower bound, from (3.9) it suffices to prove that

$$\frac{1}{C}\left(\|A^{\frac12}\mathbf{U}\|^2 + \|A^{\frac12}\nabla\mathbf{u}\|^2\right) \le \|\nabla\cdot A\mathbf{U}\|^2 + \|A^{\frac12}(\mathbf{U}-\nabla\mathbf{u})\|^2 \tag{3.22}$$

for any $(\mathbf{U},\,\mathbf{u})\in\mathcal{V}\times\mathcal{Y}$. To this end, note that our restriction on $\mathcal{Y}$ that $\int_\Omega u_1dx = \int_\Omega u_2dx = 0$ yields the Poincaré–Friedrichs inequality

$$\|\mathbf{u}\| \le C_1\,|\mathbf{u}|_1. \tag{3.23}$$

By Korn's second inequality (cf. [20]), we thus have that

$$\|\mathbf{u}\| \le C_2\,\|A_2^{\frac12}\nabla\mathbf{u}\|. \tag{3.24}$$

Since $A = \lambda A_1 + 2\mu A_2$ and both A_1 and A_2 are nonnegative definite symmetric matrices, then (3.24) implies that

$$\|\mathbf{u}\| \le C_2\,\|A^{\frac12}\nabla\mathbf{u}\|. \tag{3.25}$$

But the boundary conditions imposed on $\mathcal{V}$ then give that

$$\begin{aligned}\|A^{\frac{1}{2}}\nabla\mathbf{u}\|^2 &= (A^{\frac{1}{2}}(\nabla\mathbf{u}-\mathbf{U}),\, A^{\frac{1}{2}}\nabla\mathbf{u}) - (\nabla\cdot A\mathbf{U},\, \mathbf{u})\\ &\le \|A^{\frac{1}{2}}(\nabla\mathbf{u}-\mathbf{U})\|\,\|A^{\frac{1}{2}}\nabla\mathbf{u}\| + \|\nabla\cdot A\mathbf{U}\|\,\|\mathbf{u}\|,\end{aligned}$$

which, together with (3.25), establishes the upper bound for the first term in (3.22). The bound for the second term then follows from the triangle inequality. □

4. A two-stage algorithm based on G_0.

The basic aim of FOSLS is to develop a functional whose homogeneous form is equivalent to a product norm composed of individual scalar L^2- or H^1-like norms. The essential purpose of this construction is to reduce the original problem to a system of easily solved scalar equations whose coupling is weak enough to enable relatively easy solution of the full system. While the theory of the previous section achieves this basic goal, the quality of the relevant coupling degrades as the material properties tend to the incompressible limit (i.e., as $\lambda \to \infty$ for fixed μ, or as $\nu \to 0.5^-$). The sources of this trouble are the terms involving λ in the definitions of the functionals M_{-1}, M_0, and M. First, since the expression $\text{tr}\mathbf{U} = U_1 + U_4$ represents an intimate coupling between U_1 and U_4, large λ implies that the coupling between these two variables must tend to become dominant in the functionals. This difficulty, which causes degrading performance of standard solvers, will be eliminated here by a simple rotation applied to $\mathbf{U}$. Second, the expression $A^{1/2}\nabla\mathbf{u}$ present in M also implies dominant coupling for large λ, this time between the variables u_1 and u_2, which again implies serious difficulties in the design of fast solvers. To overcome this more imposing difficulty, we will make use of the new variables and the uniform equivalence of $G_0(\mathbf{U};\,\mathbf{0})$ and $M_0(\mathbf{U})$ to develop a two-stage algorithm for computing $\mathbf{U}$ and recovering $\mathbf{u}$ from the definition $\mathbf{U} = \nabla\mathbf{u}$.

The rotation we consider is defined by the matrix

$$Q = \begin{pmatrix} \frac{1}{\sqrt{2}} & 0 & 0 & \frac{1}{\sqrt{2}}\\ 0 & 1 & 0 & 0\\ 0 & 0 & 1 & 0\\ \frac{1}{\sqrt{2}} & 0 & 0 & -\frac{1}{\sqrt{2}} \end{pmatrix} \tag{4.1}$$

and the space $\tilde{\mathcal{V}} \equiv Q\mathcal{V} = \{\mathbf{V} = Q\mathbf{U} : \mathbf{U} \in \mathcal{V}\}$. Note that $\mathcal{V} = Q\tilde{\mathcal{V}}$ and that each vector $\mathbf{U} \in \mathcal{V}$ is of the form

$$\mathbf{U} = Q\mathbf{V}, \quad \mathbf{V} \in \tilde{\mathcal{V}}.$$

Note also that spaces $\mathcal{V}$ and $\tilde{\mathcal{V}}$ are the same up to boundary conditions.

The solution $(\mathbf{U},\, \mathbf{u})$ of the extended system (2.5) can be obtained as the solution of the following two-stage algorithm.

Stage 1: Let $\mathbf{V} \in \tilde{\mathcal{V}}$ be the unique solution of

$$G_0(Q\mathbf{V};\,\mathbf{f}) = \min\{G_0(Q\mathbf{W};\,\mathbf{f}) :\ \mathbf{W} \in \tilde{\mathcal{V}}\} \tag{4.2}$$

and set $\mathbf{U} = Q\mathbf{V}$.

Stage 2: Define

$$\mathcal{Z} = \left\{\mathbf{u} \in H^1(\Omega)^2 : \int_\Omega u_1 dx = \int_\Omega u_2 dx = 0\right\} \tag{4.3}$$

and let $\mathbf{u} \in \mathcal{Z}$ be the unique solution of

$$\|\nabla\mathbf{u} - \mathbf{U}\| = \min\{\|\nabla\mathbf{v} - \mathbf{U}\| :\ \mathbf{v} \in \mathcal{Z}\}. \tag{4.4}$$

Corollary 4.1. *Assume that* (3.5) *and* (3.7) *hold. Then*

$$\frac{1}{C}\left(\|\mathbf{V}\|_1^2+\lambda^2\|\nabla V_1\|^2\right)\le G_0(Q\mathbf{V};\,\mathbf{0})\le C\left(\|\mathbf{V}\|_1^2+\lambda^2\|\nabla V_1\|^2\right) \tag{4.5}$$

for all $\mathbf{V}\in\tilde{\mathcal{V}}$.

Remark 4.1. H^1 equivalence (4.5) immediately implies (cf. [6]) that standard finite elements and standard multigrid for minimizing $G_0(Q\mathbf{V};\,\mathbf{f})$ will achieve uniform and optimal H^1 approximations to $\mathbf{V}$ and, hence, to the deformations and stresses.

It is clear that the uniqueness of the solution of the two minimization problems guarantees that $(Q\mathbf{V},\,\mathbf{u})$ is the unique solution of the pure traction problem as expressed in equation (2.5). Nevertheless, a few comments on some interesting features of this representation are in order.

The function $\mathbf{u}$ that solves the minimization problem (4.4) of Stage 2 satisfies the Euler–Lagrange equation

$$(\nabla\mathbf{u}-\mathbf{U},\,\nabla\mathbf{v})=0\quad\forall\,\mathbf{v}\in\mathcal{Z}, \tag{4.6}$$

where $\mathbf{U}=Q\mathbf{V}$. Hence, $\mathbf{u}$ is the weak solution of the Poisson problem

$$\left\{\begin{array}{rclcc}\triangle\mathbf{u} & = & \nabla\cdot\mathbf{U}, & \text{in} & \Omega,\\ \mathbf{n}\cdot\nabla\mathbf{u} & = & \mathbf{n}\cdot\mathbf{U}, & \text{on} & \partial\Omega.\end{array}\right. \tag{4.7}$$

It may not be immediately clear that the boundary condition in (4.7) implies that $\mathbf{u}$ satisfies the boundary condition of equation (2.2), namely,

$$\mathbf{n}\cdot(A\nabla\mathbf{u})=\mathbf{0}. \tag{4.8}$$

However, this must follow from existence and uniqueness of the solutions of (2.2) and the two minimization problems.

Similarly, the conditions on $\mathcal{Z}$ do not imply directly that $\mathbf{u}$ satisfies the third compatibility condition

$$\int_\Omega(\partial_2u_1-\partial_1u_2)\,dx=0, \tag{4.9}$$

but this follows from the observation that $\nabla\mathbf{u}=\mathbf{U}\in\mathcal{V}\subset(\nabla\mathcal{N})^\perp$.

We now turn to a numerical method for the approximation of the solution $(\mathbf{U},\,\mathbf{u})$ based on this two-stage algorithm. Let $\tilde{\mathcal{V}}_h$ be a finite-dimensional subspace of $\tilde{\mathcal{V}}$ and $\mathcal{Z}_h$ a finite-dimensional subspace of $\mathcal{Z}$. Assume that they satisfy the following approximation property: there exists a constant C and two integers s and q such that for all $(\mathbf{V},\,\mathbf{v})\in\left(\tilde{\mathcal{V}}\cap H^k(\Omega)^4\right)\times\left(\mathcal{Z}\cap H^l(\Omega)^2\right)$, $1\le k\le s$, $1\le l\le q$, there exists a pair $(\mathbf{V}^h,\,\mathbf{v}^h)\in\tilde{\mathcal{V}}_h\times\mathcal{Z}_h$ such that

$$\|V_j-V_j^h\|+h\,\|V_j-V_j^h\|_1\le C\,h^k\|V_j\|_k,\quad j=1,\,\ldots,\,4, \tag{4.10}$$

and

$$\|v_j-v_j^h\|+h\,\|v_j-v_j^h\|_1\le C\,h^l\|v_j\|_l,\quad j=1,\,2. \tag{4.11}$$

Note that the boundary conditions on $\tilde{\mathcal{V}}_h$ can be implemented for polygonal domains by imposing simple algebraic relations on the boundary nodes. This can also be done for $\mathcal{Z}_h$, but the relations involve nodes on the elements that intersect the boundary. Curved boundaries require the usual special care.

Let $\mathbf{u} \in \mathcal{W} \cap H^l(\Omega)^2$ be the solution of the pure traction problem (2.1), where l is an integer, $1 \le l \le q$, and assume that $\mathbf{u}$ also satisfies (4.9). Let

$$p = \lambda \nabla \cdot \mathbf{u} \in H^{l-1}(\Omega) \tag{4.12}$$

be the *pressure* and let

$$\mathbf{U} = \nabla \mathbf{u} \in \mathcal{V} \cap H^{l-1}(\Omega)^4 \tag{4.13}$$

be the displacement flux. Assume that

$$l - 1 \le s. \tag{4.14}$$

Let $\mathbf{U}^h = Q\mathbf{V}^h$ and $\mathbf{u}^h$ be the solution of the two-stage algorithm restricted to $\tilde{\mathcal{V}}_h \times \mathcal{Z}_h$, that is, $\mathbf{V}^h \in \tilde{\mathcal{V}}_h$ is the unique solution of

$$G_0(Q\mathbf{V}^h;\, \mathbf{f}) = \min\{G_0(Q\mathbf{W}^h;\, \mathbf{f}) :\ \mathbf{W}^h \in \tilde{\mathcal{V}}_h\}, \tag{4.15}$$

and $\mathbf{u}^h \in \mathcal{Z}_h$ is the unique solution of

$$\|\nabla \mathbf{u}^h - \mathbf{U}^h\| = \min\{\|\nabla \mathbf{w}^h - \mathbf{U}^h\| :\ \mathbf{w}^h \in \mathcal{Z}_h\}. \tag{4.16}$$

THEOREM 4.1. *We have the error estimates*

$$\|\mathbf{U} - \mathbf{U}^h\|_1^2 + \lambda^2 \|\nabla(\mathrm{tr}\mathbf{U} - \mathrm{tr}\mathbf{U}^h)\|^2 \le C\, G_0(\mathbf{U}^h;\, \mathbf{f}), \tag{4.17}$$

$$\begin{aligned} \|\mathbf{U} - \mathbf{U}^h\|_1^2 + \lambda^2 \|\nabla(\mathrm{tr}\mathbf{U} - \mathrm{tr}\mathbf{U}^h)\|^2 &\le \frac{C_1}{C_0} C^2 h^{2l-4} \left(\|\mathbf{U}\|_{l-1}^2 + \|p\|_{l-1}^2\right) \\ &\le \frac{C_1}{C_0} C^2 h^{2l-4} \left(\|\mathbf{u}\|_l^2 + \|p\|_{l-1}^2\right), \end{aligned} \tag{4.18}$$

$$\|\nabla \mathbf{u} - \nabla \mathbf{u}^h\| \le \|\mathbf{U} - \mathbf{U}^h\| + \|\mathbf{U}^h - \nabla \mathbf{u}^h\|, \tag{4.19}$$

$$\|\nabla \mathbf{u} - \nabla \mathbf{u}^h\| \le \|\mathbf{U} - \mathbf{U}^h\| + C\, h^{l-1} \|\mathbf{u}\|_l. \tag{4.20}$$

Proof. Estimate (4.17) comes from the standard observation that

$$G_0(Q(\mathbf{V} - \mathbf{V}^h);\, \mathbf{0}) = G_0(Q\mathbf{V}^h;\, \mathbf{f}).$$

Estimate (4.18) is just the usual FOSLS *a priori* bound based on Theorem 3.1 and the approximation estimates (4.10) and (4.11). In order to obtain (4.19), we observe that

$$\|\nabla \mathbf{u} - \nabla \mathbf{u}^h\| \le \|\nabla \mathbf{u} - \mathbf{U}\| + \|\mathbf{U} - \mathbf{U}^h\| + \|\mathbf{U}^h - \nabla \mathbf{u}^h\|,$$

where the first term on the right-hand side is zero. Now consider the fact that the minimization conditions yield

$$(\nabla \mathbf{u},\, \nabla \mathbf{v}^h) = (\mathbf{U},\, \nabla \mathbf{v}^h) \quad \forall\ \mathbf{v}^h \in \mathcal{Z}_h$$

and

$$(\nabla \mathbf{u}^h,\, \nabla \mathbf{v}^h) = (\mathbf{U},\, \nabla \mathbf{v}^h) + (\mathbf{U}^h - \mathbf{U},\, \nabla \mathbf{v}^h) \quad \forall\ \mathbf{v}^h \in \mathcal{Z}_h.$$

Let $\tilde{\mathbf{u}}^h \in \mathcal{Z}_h$ be the solution of

$$(\nabla \tilde{\mathbf{u}}^h,\, \nabla \mathbf{v}^h) = (\mathbf{U},\, \nabla \mathbf{v}^h) \quad \forall\ \mathbf{v}^h \in \mathcal{Z}_h.$$

Then

$$\|\mathbf{u} - \tilde{\mathbf{u}}^h\|_1 \le C h^{l-1} \|\mathbf{u}\|_l$$

and

$$(\nabla(\mathbf{u}^h - \tilde{\mathbf{u}}^h),\, \nabla \mathbf{v}^h) = (\mathbf{U}^h - \mathbf{U},\, \nabla \mathbf{v}^h) \quad \forall\ \mathbf{v}^h \in \mathcal{Z}_h,$$

which yields

$$\|\mathbf{u}^h - \tilde{\mathbf{u}}^h\|_1 \le \|\mathbf{U}^h - \mathbf{U}\|.$$

Estimate (4.20) now follows from the triangle inequality. □

Remark 4.2. Bound (4.17) is an *a posteriori* error estimate, while bound (4.18) is an *a priori* estimate. Bound (4.19) is a "hybrid" estimate involving the term $\|\mathbf{U} - \mathbf{U}^h\|$, which we have not yet bounded, and the term $\|\mathbf{U}^h - \nabla \mathbf{u}^h\|$, which we can bound *a posteriori*, as in error estimate (4.20). The term $\|\mathbf{U} - \mathbf{U}^h\|$ can certainly be bounded by $\|\mathbf{U} - \mathbf{U}^h\|_1$, so that (4.19) combined with (4.18) then yields the usual H^1 optimal *a priori* estimate for displacement. However, it may be possible to establish bounds directly on $\|\mathbf{U} - \mathbf{U}^h\|$ to yield even higher-order *a priori* estimates.

5. A two-stage algorithm based on G_{-1}. Here we consider an analogous two-stage algorithm based on the functional G_{-1}, which for computational purposes we represent as follows:

$$G_{-1}(\mathbf{U},\, \mathbf{f}) = (B(\mathbf{f} + \nabla \cdot A\mathbf{U}),\, \mathbf{f} + \nabla \cdot A\mathbf{U}) + (B_0 \nabla \times \mathbf{U},\, \nabla \times \mathbf{U}), \tag{5.1}$$

where $B : H^{-1}(\Omega)^2 \longrightarrow H^1(\Omega)^2$ and $B_0 : H_0^{-1}(\Omega)^2 \longrightarrow H_0^1(\Omega)^2$ denote the respective solution operators ($\mathbf{u} = B\mathbf{f}$ and $\mathbf{u}_0 = B_0\mathbf{f}$) for the elliptic boundary value problems

$$\left\{ \begin{array}{rcll} -\triangle \mathbf{u} + \mathbf{u} & = & \mathbf{f}, & \text{in} \quad \Omega, \\ \frac{\partial}{\partial \mathbf{n}} \mathbf{u} & = & 0, & \text{on} \quad \partial\Omega \end{array} \right. \tag{5.2}$$

and

$$\left\{ \begin{array}{rcll} -\triangle \mathbf{u}_0 & = & \mathbf{f}, & \text{in} \quad \Omega, \\ \mathbf{u}_0 & = & 0, & \text{on} \quad \partial\Omega. \end{array} \right. \tag{5.3}$$

The solution $(\mathbf{U},\, \mathbf{u})$ of extended system (2.5) can be obtained as the solution of the following two-stage algorithm, which uses the notation $\tilde{\mathcal{U}} \equiv Q\mathcal{U}$.

Stage 1: Let $\mathbf{V} \in \tilde{\mathcal{U}}$ be the unique function that minimizes $G_{-1}(Q\mathbf{V};\, \mathbf{f})$ over $\tilde{\mathcal{U}}$ and set $\mathbf{U} = Q\mathbf{V}$.

Stage 2: Let $\mathbf{u} \in \mathcal{Z}$ be the unique function that minimizes $\|\nabla \mathbf{u} - \mathbf{U}\|$ over $\mathcal{Z}$.

The optimal uniform solvability of the minimization problem in Stage 1 is guaranteed by the following immediate consequence of (3.8).

COROLLARY 5.1. *We have that*

$$\frac{1}{C} \left(\|\mathbf{V}\|^2 + \lambda^2 \|V_1\|^2 \right) \le G_{-1}(Q\mathbf{V};\, \mathbf{0}) \le C \left(\|\mathbf{V}\|^2 + \lambda^2 \|V_1\|^2 \right) \tag{5.4}$$

for all $\mathbf{V} \in \tilde{\mathcal{U}}$.

Remark 5.1. Corollary 5.1 implies that standard discrete H^{-1} norms (for more details, see [1] and [8]) can be used to develop a discretization and solution process that achieves uniform and optimal L^2 approximations to the deformations and stresses, and that displacements can be recovered as in section 4 with uniform and optimal H^1 performance.

Let $\tilde{\mathcal{U}}_h$ and $\mathcal{Z}_h$ be finite-dimensional subspaces of $\tilde{\mathcal{U}}$ and $\mathcal{Z}$, respectively, that (as in section 4) satisfy estimates (4.10) and (4.11). Let $\tilde{\mathcal{Z}}_h$ and $\tilde{\mathcal{Z}}_{0,\,h}$ be finite-dimensional subspaces of $H^1(\Omega)^2$ and $H_0^1(\Omega)^2$, respectively, that (as in section 4) satisfy estimate (4.11) and the inverse inequality

$$\|\mathbf{w}^h\|_1 \le C\,h^{-1}\|\mathbf{w}^h\| \quad \forall\ \mathbf{w}^h \in \tilde{\mathcal{Z}}_h \text{ or } \tilde{\mathcal{Z}}_{0,\,h}.$$

Let B_h and $B_{0,\,h}$ be discrete solution operators associated with boundary value problems (5.2) and (5.3) posed on $\tilde{\mathcal{Z}}_h$ and $\tilde{\mathcal{Z}}_{0,\,h}$, respectively. These inverse operators are easily approximated by standard symmetric multigrid operators. Assume that we have preconditioners $\bar{B}_h : L^2(\Omega)^2 \longrightarrow \tilde{\mathcal{Z}}_h$ and $\bar{B}_{0,\,h} : L^2(\Omega)^2 \longrightarrow \tilde{\mathcal{Z}}_{0,\,h}$ that are symmetric with respect to the L^2 inner product and that are spectrally equivalent to B_h and $B_{0,\,h}$, respectively. Define

$$\tilde{B}_h = h^2 I + \bar{B}_h \quad \text{and} \quad \tilde{B}_{0,\,h} = h^2 I + \bar{B}_{0,\,h},$$

where I denotes the identity operator and

$$\tilde{G}_{-1,\,h}(\mathbf{U};\,\mathbf{f}) \equiv (\tilde{B}_h(\mathbf{f} + \nabla\cdot A\mathbf{U}),\, \mathbf{f} + \nabla\cdot A\mathbf{U}) + (\tilde{B}_{0,\,h}\nabla\times\mathbf{U},\, \nabla\times\mathbf{U}). \tag{5.5}$$

THEOREM 5.1. *Let* $\mathbf{V}^h$ *be the unique function in* $\tilde{\mathcal{U}}_h$ *that satisfies*

$$\tilde{G}_{-1,\,h}(Q\mathbf{V}^h;\,\mathbf{f}) = \min\{\tilde{G}_{-1,\,h}(Q\mathbf{W}^h;\,\mathbf{f}) :\ \mathbf{W}^h \in \tilde{\mathcal{U}}_h\}, \tag{5.6}$$

and let $\mathbf{u}^h \in \mathcal{Z}_h$ *be the unique function that satisfies*

$$\|\nabla\mathbf{u}^h - Q\mathbf{V}^h\| = \min\{\|\nabla\mathbf{w}^h - Q\mathbf{V}^h\| :\ \mathbf{w}^h \in \mathcal{Z}_h\}. \tag{5.7}$$

Then

$$\|\mathbf{U} - \mathbf{U}^h\|^2 + \lambda^2\|\mathrm{tr}\mathbf{U} - \mathrm{tr}\mathbf{U}^h\|^2 \le C\,\tilde{G}_{-1,\,h}(\mathbf{U}^h;\,\mathbf{f}), \tag{5.8}$$

$$\|\mathbf{U} - \mathbf{U}^h\|^2 + \lambda^2\|\mathrm{tr}\mathbf{U} - \mathrm{tr}\mathbf{U}^h\|^2 \le \frac{C_1}{C_0}C\,h^{2l-2}\left(\|\mathbf{u}\|_l^2 + \|p\|_{l-1}^2\right), \tag{5.9}$$

$$\|\nabla\mathbf{u} - \nabla\mathbf{u}^h\| \le \|\mathbf{U} - \mathbf{U}^h\| + \|\mathbf{U}^h - \nabla\mathbf{u}^h\|, \tag{5.10}$$

$$\|\nabla\mathbf{u} - \nabla\mathbf{u}^h\|^2 \le C\,\tilde{G}_{-1,\,h}(\mathbf{U}^h;\,\mathbf{f}) + \|\mathbf{U}^h - \nabla\mathbf{u}^h\|^2, \tag{5.11}$$

$$\|\nabla\mathbf{u} - \nabla\mathbf{u}^h\| \le C\,h^{l-1}\left(\|\mathbf{u}\|_l + \|p\|_{l-1}\right). \tag{5.12}$$

Proof. An analysis similar to that in [1] (see also [8]) shows that (5.4) is valid for $\tilde{G}_{-1,\,h}(Q\mathbf{V};\,\mathbf{f})$ and for any $\mathbf{V} \in \tilde{\mathcal{U}}_h$. Thus, these estimates follow directly from the usual FOSLS *a posteriori* error bounds and approximation estimates (4.10) and (4.11). ☐

Remark 5.2. The theory here assures that the discrete functional $\tilde{G}_{-1,\,h}(Q\mathbf{V}^h;\,\mathbf{0})$ is uniformly equivalent to the simple functional $D_h(\mathbf{V}^h) \equiv \|\mathbf{V}^h\|^2 + \lambda^2\|V_1^h\|^2$ on $\tilde{\mathcal{U}}_h$. Thus, rescaling the first component of vector $\mathbf{V}$ by $(1+\lambda^2)^{1/2}$ makes $\tilde{G}_{-1,\,h}$ uniformly equivalent to the L^2 norm; that is, the matrix that arises from the Euler–Lagrange equation for (5.6) using standard nodal basis functions is uniformly well conditioned. This means that basic iterative methods like Gauss–Seidel or conjugate gradients could be used to solve these equations with optimal efficiency.

6. Regularity estimates. Although estimates (3.4) and (3.6) are standard, we provide proofs here for completeness.

Proof of (3.4). For any $\mathbf{v} \in \mathcal{W}$, it follows from Korn's inequality and integration by parts that

$$\begin{aligned} C\,\|\mathbf{v}\|_1^2 &\le \lambda\|\nabla\cdot\mathbf{v}\|^2 + 2\mu\|\epsilon(\mathbf{v})\|^2 \\ &= -(\nabla\cdot A\nabla\mathbf{v},\,\mathbf{v}) \le \|\nabla\cdot A\nabla\mathbf{v}\|_{-1}\|\mathbf{v}\|_1. \end{aligned}$$

This establishes the upper bound for the first term in (3.4). To bound the second term, choose $\mathbf{w} \in \mathcal{N}^c$ such that (cf. [2])

$$\nabla\cdot\mathbf{w} = \nabla\cdot\mathbf{v} \quad \text{and} \quad \|\mathbf{w}\|_1 \le C\,\|\nabla\cdot\mathbf{v}\|.$$

We then have that

$$\begin{aligned} \lambda\|\nabla\cdot\mathbf{v}\|^2 &= \lambda(\nabla\cdot\mathbf{v},\,\nabla\cdot\mathbf{w}) \\ &= (A\nabla\mathbf{v},\,\nabla\mathbf{w}) - 2\mu(\epsilon(\mathbf{v}),\,\epsilon(\mathbf{w})) \\ &= -(\nabla\cdot A\nabla\mathbf{v},\,\mathbf{w}) - 2\mu(\epsilon(\mathbf{v}),\,\epsilon(\mathbf{w})) \\ &\le C\,\|\mathbf{w}\|_1\,(\|\nabla\cdot A\nabla\mathbf{v}\|_{-1} + \|\mathbf{v}\|_1) \\ &\le C\,\|\nabla\cdot\mathbf{v}\|\,\|\nabla\cdot A\nabla\mathbf{v}\|_{-1}. \end{aligned}$$

Canceling $\|\nabla\cdot\mathbf{v}\|$, we obtain

$$\lambda\|\nabla\cdot\mathbf{v}\| \le C\,\|\nabla\cdot A\nabla\mathbf{v}\|_{-1},$$

which completes the proof for (3.4). ☐

Proof of (3.6). Since $\mathbf{w} \in H_0^1(\Omega)^2$ is divergence free, we have that

$$\begin{aligned} C\,\|\rho\mathbf{w}\|_1^2 &\le |\rho\mathbf{w}|_1^2 \\ &= (-\rho\triangle\mathbf{w} + \nabla p,\,\rho\mathbf{w}) \\ &\le \|-\rho\triangle\mathbf{w} + \nabla p\|_{-1,\,0}\|\rho\mathbf{w}\|_1, \end{aligned}$$

which proves the bound for the first term in (3.6). The bound for the second term in (3.6) follows directly from the well-known inequality (cf. [19]) $\|p\| \le C\,\|\nabla p\|_{-1,\,0}$ for $p \in L_0^2(\Omega)$, the triangle inequality, and the fact that $\|\rho\triangle\mathbf{w}\|_{-1,\,0} \le \|\rho\mathbf{w}\|_1$. ☐

REFERENCES

[1] J. H. Bramble, R. D. Lazarov, and J. E. Pasciak, *A least-squares approach based on a discrete minus one inner product for first order system*, manuscript.

[2] S. C. Brenner, *A nonconforming mixed multigrid method for the pure traction problem in planar linear elasticity*, SIAM J. Numer. Anal., 30 (1993), pp. 116–135.

[3] F. Brezzi and J. Douglas, *Stablized mixed methods for Stokes problem*, Numer. Math., 53 (1988), pp. 225–236.

[4] F. Brezzi and M. Fortin, *Mixed and Hybrid Finite Element Methods*, Springer-Verlag, New York, 1991.

[5] Z. Cai, R. D. 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] Z. Cai, T. Manteuffel, and S. McCormick, *First-order system least squares for second-order partial differential equations: Part* II, SIAM J. Numer. Anal., 34 (1997), pp. 425–454.

[7] Z. Cai, T. Manteuffel, and S. McCormick, *First-order system least squares for the Stokes equations, with application to linear elasticity*, SIAM J. Numer. Anal., 34 (1997), pp. 1727–1741.

[8] Z. Cai, T. Manteuffel, and S. McCormick, *First-order system least squares for velocity-vorticity-pressure form of the Stokes equations, with application to linear elasticity*, Electron. Trans. Numer. Anal., 3 (1995), pp. 150–159.

[9] P. G. Ciarlet, *The Finite Element Method for Elliptic Problems*, North-Holland, New York, 1978.

[10] P. G. Ciarlet, *Mathematical Elasticity: Volume* I*: Three-Dimensional Elasticity*, North-Holland, New York, 1988.

[11] J. Douglas and J. Wang, *An absolutely stabilized finite element method for the Stokes problem*, Math. Comp., 52 (1989), pp. 495–508.

[12] M. Fortin and R. Glowinski, *Augmented Lagrangian Methods*, North-Holland, Amsterdam, 1983.

[13] L. P. Franca and R. Stenberg, *Error analysis of some Garlerkin least squares methods for the elasticity equations*, SIAM J. Numer. Anal., 28 (1991), pp. 1680–1697.

[14] V. Girault and P. A. Raviart, *Finite Element Methods for Navier-Stokes Equations: Theory and Algorithms*, Springer-Verlag, New York, 1986.

[15] P. Grisvard, *Problémes aux limits dans les polygones. Mode d'emploi*, EDF Bull. de la Direction des Etudes et Recherches Ser. C Mathematiques, Informatique, 1 (1986), pp. 21–59.

[16] P. Grisvard, *Singularités en elasticité*, Arch. Rational Mech. Anal., 107 (1989), pp. 157–180.

[17] R. B. Kellogg and J. E. Osborn, *A regularity result for the Stokes problem in a convex polygon*, J. Functional Anal., 21 (1976), pp. 397–431.

[18] O. A. Ladyhenskaya, *The Mathematical Theory of Viscous Incompressible Flows*, Gordon and Breach, New York, 1963.

[19] J. Nečas, *Equations aux dérivées partielles*, Presses de l'Université de Montréal, Canada, 1965.

[20] J. Nitche, *On Korn's second inequality*, RAIRO Numer. Anal., 15 (1981), pp. 237–248.

[21] R. Stenberg, *A family of mixed finite elements for the elasticity problem*, Numer. Math., 53 (1988), pp. 513–538.