SIAM J. Sci. Comput.
Vol. 21, No. 5, pp. 1728–1739

FIRST-ORDER SYSTEM LEAST SQUARES FOR THE STOKES AND LINEAR ELASTICITY EQUATIONS: FURTHER RESULTS*

Z. CAI†, C.-O. LEE‡, T. A. MANTEUFFEL§, AND S. F. MCCORMICK§

Abstract. First-order system least squares (FOSLS) was developed in [*SIAM J. Numer. Anal.*, 34 (1997), pp. 1727–1741; *SIAM J. Numer. Anal.*, 35 (1998), pp. 320–335] for Stokes and elasticity equations. Several new results for these methods are obtained here. First, the inverse-norm FOSLS scheme that was introduced but not analyzed in [*SIAM J. Numer. Anal.*, 34 (1997), pp. 1727–1741] is shown to be continuous and coercive in the L^2 norm. This result is shown to hold for pure displacement or pure traction boundary conditions in two or three dimensions, and for mixed boundary conditions in two dimensions. Next, the FOSLS schemes developed in [*SIAM J. Numer. Anal.*, 35 (1998), pp. 320–335] are applied to the pure displacement problem in planar and spatial linear elasticity by eliminating the pressure variable in the FOSLS formulations of [*SIAM J. Numer. Anal.*, 34 (1997), pp. 1727–1741]. The idea of two-dimensional variable rotation is then extended to three dimensions to make the intervariable coupling subdominant (uniformly so in the Poisson ratio for elasticity). This decoupling ensures optimal (uniform) performance of finite element discretization and multigrid solution methods. It also allows special treatment of the new trace variable, which corresponds to the divergence of velocity in the case of Stokes, so that conservation can be easily imposed, for example. Numerical results for various boundary value problems of planar linear elasticity are studied in a companion paper [*SIAM J. Sci. Comput.*, 21 (2000), pp. 1706–1727].

Key words. least squares, multigrid, Stokes, elasticity

AMS subject classifications. 65F10, 65F30

PII. S1064827598338652

1. Introduction.

Two *two-stage* first-order system least-squares (FOSLS) approaches were developed recently by Cai, Manteuffel, and McCormick (see [3, Remarks 3.2 and 3.3]) for the Stokes equations and the pure displacement problem in linear elasticity. Both approaches incorporate a first stage that solves for pressure and the gradient of the vector variable (velocity or displacement, immediately yielding other physically meaningful quantities such as vorticity or deformation and stress). A second stage can then be used to solve for the vector variable itself, if desired. One of these two-stage approaches uses L^2 norms to define the FOSLS functional. Under certain H^2 regularity assumptions, an immediate consequence of [3, Theorem 3.2] is that the homogeneous part of this functional (i.e., with zero source and boundary data) is continuous and coercive in an appropriately weighted H^1 product norm. This result assures optimal H^1-like performance for standard finite element discretization and multigrid solution methods for all variables—uniform in the Reynolds number or Poisson ratio. The second approach is based on H^{-1} norms. Equivalence of the homogeneous part of this FOSLS functional to an L^2 product norm is not a direct

*Received by the editors May 13, 1998; accepted for publication (in revised form) May 7, 1999; published electronically April 28, 2000.

http://www.siam.org/journals/sisc/21-5/33865.html

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

‡Department of Mathematics, Inha University, Inchon 402-751, Korea (colee@math.inha.ac.kr). This author was sponsored by BSRI-97-1436 and KOSEF 97-0701-01-01-3.

§Department of Applied Mathamatics, Campus Box 526, University of Colorado at Boulder, Boulder, CO 80309-0526 (tmanteuf@boulder.colorado.edu, stevem@boulder.colorado.edu). These authors were sponsored by the National Science Foundation under grant DMS-9706866 and the Department of Energy under grant DE-FG03-93ER25165.

consequence of the theorems in [3], so one aim here is to establish this property, which we do by means of a Helmholtz decomposition used, e.g., in [3, 4, 9]. This equivalence is shown to hold also for the case of two dimensions and mixed boundary conditions.

Two other two-stage FOSLS approaches for the pure traction problem in linear elasticity were also developed recently by Cai, Manteuffel, McCormick, and Parter in [4] for the planar case and by Kim, Manteuffel, and McCormick in [9] for the spatial case. They differ from those in [3] at the first stage primarily because they do not use the pressure as an independent variable. Many practical models of linear elasticity involve mixed boundary conditions. Thus, as a step toward this practical case, we reformulate the pure displacement approach of [3] so that it has the same form as the pure traction approach of [4, 9]. We will develop this reformulation by eliminating the pressure variable and incorporating some of the curl terms in the divergence equation. This has the effect of extending the approach in [4, 9] to the pure displacement problem for planar and spatial linear elasticity. Reversing the steps also has the effect of extending the approach of [3] to the pure traction problem of planar linear elasticity.

Since some of the components of the gradient of displacement for linear elasticity are strongly coupled in the trace term of our FOSLS functionals, a large Lamé constant λ leads to degrading performance of standard solvers. This difficulty is eliminated in the two-dimensional case by a simple rotation applied to the gradient of displacement (see [3, 4]). We extend this idea to three dimensions by a change of variables that uses an orthogonal mapping to define the trace as one of the new variables. This transformation can also be used for the Stokes equations to allow higher-order approximation of the trace variable (which corresponds to the divergence of velocity).

A companion paper [2] reports on numerical tests of these FOSLS methods applied to planar linear elasticity with pure traction, pure displacement, and mixed boundary conditions. The results confirm our theory by showing that standard finite elements and multigrid solvers achieve optimal performance uniformly in the Poisson ratio.

As in [3], we are content here with an abbreviated paper that focuses on establishing continuity and coercivity. We establish these results in section 3 for the first stage of the H^{-1} FOSLS functional applied to the Stokes equations, the pure displacement problem of linear elasticity, and the two-dimensional mixed problem of linear elasticity. In section 4, these results are extended to the pure displacement problem of linear elasticity. Section 2 introduces the generalized Stokes equations and some preliminaries. The change of variables and higher-order approximation for the trace variable are discussed in section 5.

2. The Stokes problem and preliminaries.

Let Ω be a bounded, open, connected domain in $\Re^n$ ($n = 2$ or 3) with Lipschitz boundary $\partial\Omega$. For simplicity, but without loss of generality, consider the pressure-perturbed form of the generalized stationary Stokes equations in dimensionless variables:

$$\left\{ \begin{array}{rcll} -\nabla\cdot(2\mu\boldsymbol{\epsilon}(\mathbf{u})) + \nabla p & = & \mathbf{f} & \text{in} \quad \Omega, \\ \nabla\cdot\mathbf{u} + \delta\, p & = & g & \text{in} \quad \Omega, \end{array} \right. \tag{2.1}$$

where $\boldsymbol{\epsilon}(\mathbf{u}) = (\epsilon_{ij}(\mathbf{u}))_{n\times n} = \frac{1}{2}(\partial_i u_j + \partial_j u_i)_{n\times n}$ is the *deformation tensor*, $\delta = 0$ for Stokes, and $\delta = \frac{1}{\lambda}$ for linear elasticity. (Stokes equations are usually posed with the inverse Reynolds number $\nu = \frac{1}{Re}$; here, for simplicity, we set $\nu = 1$, which is equivalent to absorbing Re into the pressure variable and source term, i.e., to replacing $Re\,p$ by p and $Re\,f$ by f.) Consider the (generalized) Stokes equations (2.1) together with homogeneous boundary conditions

$$\mathbf{u} = \mathbf{0} \quad \text{on } \Gamma_D \quad \text{and} \quad \sum_{j=1}^{n} \sigma_{ij}(\mathbf{u})\, n_j = 0 \quad \text{on } \Gamma_N \quad \text{for } i = 1, \ldots, n, \tag{2.2}$$

where $\Gamma_D \neq \emptyset$ and Γ_N partition the boundary of Ω, $\sigma_{ij}(\mathbf{u}) = -p\delta_{ij} + 2\mu\epsilon_{ij}(\mathbf{u})$ is the *stress tensor*, and δ_{ij} is the Kronecker delta symbol. (The case $\Gamma_D = \emptyset$ is treated in [4, 9] and, in effect, by the results of section 4 below.) Without loss of generality, we assume that $\mu = 1$ and that

$$\int_\Omega g\, dz = \int_\Omega p\, dz = 0 \quad \text{if } \Gamma_N = \emptyset. \tag{2.3}$$

We use the same notation as in [4] unless otherwise specified. In particular, $H_D^1(\Omega) := \{q \in H^1(\Omega) \,:\, q = 0 \text{ on } \Gamma_D\}$ and $H_N^1(\Omega) := \{q \in H^1(\Omega) \,:\, q = 0 \text{ on } \Gamma_N\}$. We use $H_D^{-1}(\Omega)$, $H_N^{-1}(\Omega)$, and $H^{-1}(\Omega)$ to denote the duals of $H_D^1(\Omega)$, $H_N^1(\Omega)$, and $H^1(\Omega)$ with norms defined by

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

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

respectively. Define the product spaces

$$H_D^{-1}(\Omega)^n = \prod_{i=1}^{n} H_D^{-1}(\Omega)\,, \quad H_N^{-1}(\Omega)^n = \prod_{i=1}^{n} H_N^{-1}(\Omega)\,, \quad H^{-1}(\Omega)^n = \prod_{i=1}^{n} H^{-1}(\Omega)$$

with standard product norms. Let

$$H(\mathrm{div};\, \Omega) := \{\mathbf{v} \in L^2(\Omega)^n \,:\, \nabla \cdot \mathbf{v} \in L^2(\Omega)\}$$

and

$$H(\mathbf{curl};\, \Omega) := \{\mathbf{v} \in L^2(\Omega)^n \,:\, \nabla\times \mathbf{v} \in L^2(\Omega)^{2n-3}\},$$

which are Hilbert spaces under the respective norms

$$\|\mathbf{v}\|_{H(\mathrm{div};\,\Omega)} := \left(\|\mathbf{v}\|^2 + \|\nabla \cdot \mathbf{v}\|^2\right)^{\frac{1}{2}} \quad \text{and} \quad \|\mathbf{v}\|_{H(\mathbf{curl};\,\Omega)} := \left(\|\mathbf{v}\|^2 + \|\nabla\times \mathbf{v}\|^2\right)^{\frac{1}{2}}.$$

Define the subspaces

$$H_0(\mathrm{div};\, \Omega) := \{\mathbf{v} \in H(\mathrm{div};\, \Omega) \,:\, \mathbf{n} \cdot \mathbf{v} = 0 \ \text{ on } \ \partial\Omega\}$$

and

$$H_0(\mathbf{curl};\, \Omega) := \{\mathbf{v} \in H(\mathbf{curl};\, \Omega) \,:\, \mathbf{n} \times \mathbf{v} = \mathbf{0} \ \text{ on } \ \partial\Omega\}.$$

3. First-order system least squares. We first review the two-stage FOSLS approaches developed in [3, Remarks 3.2 and 3.3]. For more detail on the two-stage algorithms, see [4]. The first stage solves for pressure and the gradient of the vector variable. One approach uses L^2 norms to define the FOSLS functional, and H^1

product norm equivalence of its homogeneous part is a direct consequence of the theorems in [3]. The other is based on H^{-1} norms, and L^2 product norm equivalence of its homogeneous part is established below. If the vector variables are needed, they can be recovered in a second stage that solves Poisson-like equations for each component (see Remark 3.1).

As in [3, 4], define the *velocity flux* variable $\underline{\mathbf{U}} = \nabla\mathbf{u} = ((\nabla u_1)^t, \ldots, (\nabla u_n)^t)^t = (\mathbf{U}_1^t, \ldots, \mathbf{U}_n^t)^t$, where $\mathbf{U}_i = (U_{i1}, \ldots, U_{in})^t$. This yields the equivalent extended first-order system

$$(3.1) \qquad \left\{ \begin{array}{rcll} -\nabla\cdot(2A_2\underline{\mathbf{U}}) + \nabla p & = & \mathbf{f} & \text{in} \quad \Omega, \\ \operatorname{tr}\underline{\mathbf{U}} + \delta\, p & = & g & \text{in} \quad \Omega, \\ \nabla\times\underline{\mathbf{U}} & = & \underline{\mathbf{0}} & \text{in} \quad \Omega, \\ \underline{\mathbf{U}} - \nabla\mathbf{u} & = & \underline{\mathbf{0}} & \text{in} \quad \Omega, \\ \nabla\cdot\mathbf{u} + \delta\, p & = & g & \text{in} \quad \Omega \end{array} \right.$$

with boundary conditions

$$\mathbf{u} = \mathbf{0} \quad \text{and} \quad \mathbf{n}\times\underline{\mathbf{U}} = \underline{\mathbf{0}} \quad \text{on } \Gamma_D \quad \text{and} \quad \mathbf{n}\cdot(2A_2\underline{\mathbf{U}}) - p\mathbf{n} = \mathbf{0} \quad \text{on } \Gamma_N.$$

Here, tr is the trace operator defined by $\operatorname{tr}\underline{\mathbf{U}} = \sum_{i=1}^n U_{ii}$ and A_2 is the $n^2\times n^2$ matrix defined for $n = 2$ by

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

See [9] for the definition of A_2 when $n = 3$. Define the following first-stage FOSLS functionals:

$$(3.2) \quad G_0^{(1)}(\underline{\mathbf{U}}, p;\, \mathbf{f}, g) := \|\mathbf{f} + \nabla\cdot(2A_2\underline{\mathbf{U}}) - \nabla p\|^2 + \|\nabla\times\underline{\mathbf{U}}\|^2 + \|\nabla(\operatorname{tr}\underline{\mathbf{U}} + \delta\, p - g)\|^2$$

and

$$(3.3) \quad G_{-1}^{(1)}(\underline{\mathbf{U}}, p;\, \mathbf{f}, g) := \|\mathbf{f}+\nabla\cdot(2A_2\underline{\mathbf{U}})-\nabla p\|_{-1,D}^2 + \|\nabla\times\underline{\mathbf{U}}\|_{-1,N}^2 + \|\operatorname{tr}\underline{\mathbf{U}}+\delta\, p-g\|^2.$$

Let $L_D^2(\Omega)$ denote $L_0^2(\Omega) := \{q \in L^2(\Omega) : \int_\Omega q\, dx = 0\}$ if $\Gamma_N = \emptyset$ or $L^2(\Omega)$ otherwise. Define

$$\mathcal{V}_0 := \{(\underline{\mathbf{U}}, p) \in H^1(\Omega)^{n^2} \times \hat{H}^1(\Omega) : \mathbf{n}\times\underline{\mathbf{U}} = \underline{\mathbf{0}} \text{ on } \Gamma_D\,,\ \mathbf{n}\cdot(2A_2\underline{\mathbf{U}}) - p\mathbf{n} = \mathbf{0} \text{ on } \Gamma_N\},$$

where $\hat{H}^1(\Omega) = H^1(\Omega)/\mathcal{R}$ if $\Gamma_N = \emptyset$ or $H^1(\Omega)$ otherwise, and

$$\mathcal{V}_{-1} := \{(\underline{\mathbf{U}}, p) \in L^2(\Omega)^{n^2} \times L_D^2(\Omega) : \mathbf{n}\times\underline{\mathbf{U}} = \underline{\mathbf{0}} \text{ on } \Gamma_D\,,\ \mathbf{n}\cdot(2A_2\underline{\mathbf{U}}) - p\mathbf{n} = \mathbf{0} \text{ on } \Gamma_N\}.$$

(Boundary conditions in our set definitions are assumed to be defined in the appropriate sense.) The two-stage FOSLS algorithm proposed in [3] first involves solving for pressure and the gradient of the vector variable by minimizing either $G_0^{(1)}(\underline{\mathbf{U}}, p;\, \mathbf{f}, g)$

over $\mathcal{V}_0$ or $G^{(1)}_{-1}(\underline{\mathbf{U}}, p;\, \mathbf{f}, g)$ over $\mathcal{V}_{-1}$. The second stage then recovers $\mathbf{u}$ by fixing $(\underline{\mathbf{U}}, p)$ and minimizing the functional

$$G^{(2)}(\mathbf{u};\, \underline{\mathbf{U}}, p, g) := \|\underline{\mathbf{U}} - \nabla\,\mathbf{u}\|^2 + \|\nabla\cdot\mathbf{u} + \delta\, p - g\|^2 \tag{3.4}$$

over $H^1_0(\Omega)^n$ when $\Gamma_N = \emptyset$.

Remark 3.1. An alternative to the second stage is to minimize

$$\tilde{G}^{(2)}(\mathbf{u};\, \underline{\mathbf{U}}) := \|\underline{\mathbf{U}} - \nabla\,\mathbf{u}\|^2$$

over $H^1_0(\Omega)^n$. This minimization problem leads to n scalar Poisson equations. Another alternative is to apply the least-squares principle to the so-called div-curl system

$$\begin{cases} \nabla\times\mathbf{u} &= \quad \boldsymbol{\omega} & \text{in} \quad \Omega, \\ \nabla\cdot\mathbf{u} &= \quad g - \delta\, p & \text{in} \quad \Omega, \end{cases} \tag{3.5}$$

where $\boldsymbol{\omega} = U_{12} - U_{21}$ for $n = 2$ or $(U_{32} - U_{23}, U_{13} - U_{31}, U_{21} - U_{12})^t$ for $n = 3$. We may take (3.5) together with any one of the following boundary conditions:

$$\mathbf{n}\cdot\mathbf{u} = 0 \quad \text{or} \quad \mathbf{n}\times\mathbf{u} = \mathbf{0} \quad \text{or} \quad \mathbf{u} = \mathbf{0} \quad \text{on} \quad \partial\Omega.$$

Since the right-hand sides of (3.5) are always in L^2 in our applications, we consider the least-squares method only based on the L^2 norms. For the least-squares approach based on the H^{-1} norms, see [3, 4] as well as sections 3 and 4 below for similar but harder problems. We minimize the functional

$$\hat{G}^{(2)}(\mathbf{u};\, \boldsymbol{\omega}, p, g) := \|\nabla\times\mathbf{u} - \boldsymbol{\omega}\|^2 + \|\nabla\cdot\mathbf{u} + \delta\, p - g\|^2$$

over $H_0(\mathrm{div};\, \Omega)\cap H(\mathbf{curl};\, \Omega)$, $H(\mathrm{div};\, \Omega)\cap H_0(\mathbf{curl};\, \Omega)$, or $H_0(\mathrm{div};\, \Omega)\cap H_0(\mathbf{curl};\, \Omega)$. The theorems in [8] establish that the homogeneous part of the functional is equivalent to the H^1 norm on any of these three spaces.

It is shown in [3] that, when $\Gamma_N = \emptyset$, then, under certain H^2 regularity assumptions, the homogeneous part $G^{(1)}_0(\underline{\mathbf{U}}, p;\, \mathbf{0}, 0)$ is continuous and coercive in $\mathcal{V}_0$ in the product H^1 norm for each variable: there exists a positive constant C independent of δ such that

$$\frac{1}{C}\left(\|\underline{\mathbf{U}}\|_1^2 + \|p\|_1^2\right) \le G^{(1)}_0(\underline{\mathbf{U}},\, p;\, \mathbf{0},\, 0) \le C\left(\|\underline{\mathbf{U}}\|_1^2 + \|p\|_1^2\right). \tag{3.6}$$

(We use C in this paper to denote a generic positive constant, possibly different at different occurrences, that is independent of δ but may depend on the domain Ω.) It is also shown in [3] that $G^{(2)}(\mathbf{u};\, \underline{\mathbf{0}},\, 0,\, 0)$ is continuous and coercive in $H^1_0(\Omega)^n$. Here, we establish continuity and coercivity of $G^{(1)}_{-1}(\underline{\mathbf{U}}, p;\, \mathbf{0},\, 0)$ by means of a Helmholtz decomposition used, e.g., in [3, 4, 9]. To this end, we need H^1 regularity results for Stokes problems (see Lemmas 3.1 and 3.2) and a certain three-dimensional curl regularity assumption (see Lemma 3.3, which imposes additional assumptions on the domain).

Lemma 3.1. *The following H^1 regularity result holds for the variational form of the generalized Stokes equation* (2.1)–(2.2):

$$\|\mathbf{u}\|_1 + \|p\| \le C\left(\|\mathbf{f}\|_{-1,D} + \|g\|\right), \tag{3.7}$$

where C is a positive constant independent of δ.

Proof. The variational form of (2.1)–(2.3) follows: find $(\mathbf{u},\, p) \in H^1_D(\Omega)^n \times L^2_D(\Omega)$ such that

$$\left\{ \begin{array}{rcll} (\boldsymbol{\epsilon}(\mathbf{u}),\, \boldsymbol{\epsilon}(\mathbf{v})) - (p,\, \nabla \cdot \mathbf{v}) & = & (\mathbf{f},\, \mathbf{v}) & \text{for all} \quad \mathbf{v} \in H^1_D(\Omega)^n, \\ (\nabla \cdot \mathbf{u},\, q) + \delta(p, q) & = & (g, q) & \text{for all} \quad q \in L^2_D(\Omega). \end{array} \right. \tag{3.8}$$

Now the following Korn inequality holds for all $\mathbf{u} \in H^1_D(\Omega)^n$ (see, e.g., [7]):

$$\|\mathbf{u}\|_1 \le C \|\boldsymbol{\epsilon}(\mathbf{u})\|. \tag{3.9}$$

Adding the two equations in (3.8) with $\mathbf{v} = \mathbf{u}$ and $q = p$ and using (3.9), the definition of the H^{-1}_D norm, and the Cauchy–Schwarz inequality, we have

$$\begin{aligned} \frac{1}{C} \|\mathbf{u}\|_1^2 + \delta \|p\|^2 &\le \|\boldsymbol{\epsilon}(\mathbf{u})\|^2 + \delta \|p\|^2 = (\mathbf{f}, \mathbf{u}) + (g, p) \\ &\le \|\mathbf{f}\|_{-1,D} \|\mathbf{u}\|_1 + \|g\|\, \|p\|. \end{aligned} \tag{3.10}$$

It follows from the first equation in (3.8) and the Cauchy–Schwarz inequality that

$$\|p\| \le C \sup_{\mathbf{v} \in H^1_D(\Omega)^n} \frac{(p, \nabla \cdot \mathbf{v})}{\|\mathbf{v}\|_1} = C \sup_{\mathbf{v} \in H^1_D(\Omega)^n} \frac{(\boldsymbol{\epsilon}(\mathbf{u}), \boldsymbol{\epsilon}(\mathbf{v})) - (\mathbf{f}, \mathbf{v})}{\|\mathbf{v}\|_1} \tag{3.11}$$

$$\le C \left(\|\mathbf{u}\|_1 + \|\mathbf{f}\|_{-1,D} \right) \tag{3.12}$$

for all $p \in L^2_D(\Omega)$. (See, e.g., [7] for the inequality in (3.11).) The lemma now is an immediate consequence of (3.12) and (3.10). □

Remark 3.2. Based on this estimate, it is possible to develop an inverse-norm FOSLS formulation applied directly to (2.1) by defining the least-squares functional

$$\|\mathbf{f} + \nabla \cdot (2\boldsymbol{\epsilon}(\mathbf{u})) - \nabla p\|^2_{-1,D} + \|g - \nabla \cdot \mathbf{u} - \delta p\|^2.$$

To be practical, the H^{-1}_D norm and divergence operator of the first term in the functional can be replaced by a computational feasible discrete H^{-1}_D norm and discrete divergence operator, respectively (see [5] for the treatment of the Reissner–Mindlin plate).

LEMMA 3.2. *Let $\hat{H}^1_N(\Omega)$ denote $H^1(\Omega)/\mathcal{R}$ if $\Gamma_N = \emptyset$ or $H^1_N(\Omega)$ otherwise. Assume that $(\mathbf{u},\, p) \in \hat{H}^1_N(\Omega)^n \times L^2(\Omega)$ is the solution of the Stokes equations*

$$\left\{ \begin{array}{rcll} \frac{1}{2}(\nabla \mathbf{u},\, \nabla \mathbf{v}) - (p,\, \nabla \cdot \mathbf{v}) & = & (\mathbf{f},\, \mathbf{v}) & \text{for all} \quad \mathbf{v} \in \hat{H}^1_N(\Omega)^n, \\ (\nabla \cdot \mathbf{u},\, q) & = & (g, q) & \text{for all} \quad q \in L^2(\Omega), \end{array} \right. \tag{3.13}$$

where $\mathbf{f}$ *satisfies the compatibility condition*

$$\int_\Omega \mathbf{f}\, dx = \mathbf{0} \quad \text{if } \Gamma_N = \emptyset.$$

Then there exists a positive constant C independent of δ such that

$$\|\mathbf{u}\|_1 + \|p\| \le C \left(\|\mathbf{f}\|_{-1,N} + \|g\| \right). \tag{3.14}$$

Proof. The proof of (3.14) is the same as that of Lemma 3.1. □

LEMMA 3.3. *Assume that the domain Ω is a convex polyhedron or has $C^{1,1}$ boundary. If $\mathbf{v} \in H_0(\mathbf{curl};\,\Omega)$ is divergence free, then there exists a positive constant C such that*

$$\|\mathbf{v}\| \le C\|\nabla\times\,\mathbf{v}\|_{-1}. \tag{3.15}$$

Proof. For any $\mathbf{v} \in H_0(\mathbf{curl};\,\Omega)$ that is divergence free, there exists $\boldsymbol{\phi} \in H^1(\Omega)^3$ such that (cf. [8])

$$\mathbf{v} = \nabla\times\,\boldsymbol{\phi} \quad \text{in } \Omega, \quad \nabla\cdot\boldsymbol{\phi} = 0 \quad \text{in } \Omega, \quad \boldsymbol{\phi}\cdot\mathbf{n} = 0 \quad \text{on } \partial\Omega, \quad \|\boldsymbol{\phi}\|_1 \le C\|\mathbf{v}\|.$$

It then follows from integration by parts that

$$\|\mathbf{v}\|^2 = (\mathbf{v}, \nabla\times\,\boldsymbol{\phi}) = (\nabla\times\mathbf{v}, \boldsymbol{\phi}) \le C\|\nabla\times\,\mathbf{v}\|_{-1}\|\mathbf{v}\|,$$

which implies (3.15). $\square$

THEOREM 3.4. *For $n = 2$, there exists a positive constant C independent of δ such that*

$$\frac{1}{C}\left(\|\underline{\mathbf{U}}\|^2 + \|p\|^2\right) \le G_{-1}^{(1)}(\underline{\mathbf{U}},\, p;\, \mathbf{0},\, 0) \le C\left(\|\underline{\mathbf{U}}\|^2 + \|p\|^2\right) \tag{3.16}$$

for any $(\underline{\mathbf{U}},\, p) \in \mathcal{V}_{-1}$. For $n = 3$, this equivalence is also valid when $\Gamma_N = \emptyset$ and the domain Ω is a convex polyhedron or has $C^{1,1}$ boundary.

Proof. The upper bound in (3.16) follows from the easily established bounds

$$\|\nabla\cdot(2A_2\underline{\mathbf{U}}) - \nabla p\|_{-1,D} \le \|\underline{\mathbf{U}}\| + \|p\| \quad \text{and} \quad \|\nabla\times\underline{\mathbf{U}}\|_{-1,N} \le \|\underline{\mathbf{U}}\|$$

and from the triangle inequality. We prove the lower bound in (3.16) first for $n = 2$. As in [4], we have the following decomposition:

$$\underline{\mathbf{U}} = \nabla\mathbf{v} + \underline{\mathbf{V}} + \boldsymbol{\eta}, \tag{3.17}$$

where $\mathbf{v} \in H_D^1(\Omega)^2$ is the unique solution of

$$\begin{cases} \nabla\cdot(2A_2\nabla\mathbf{v}) & = & \nabla\cdot(2A_2\underline{\mathbf{U}}) & \text{in} & \Omega, \\ \mathbf{v} & = & \mathbf{0} & \text{on} & \Gamma_D, \\ \mathbf{n}\cdot(2A_2\nabla\mathbf{v}) & = & \mathbf{0} & \text{on} & \Gamma_N; \end{cases}$$

$\underline{\mathbf{V}} = \frac{1}{2}\nabla^{\perp}\mathbf{w}$, with $\mathbf{w} \in H_N^1(\Omega)^2$ divergence free in Ω and satisfying

$$\frac{1}{2}\mathbf{n}\cdot\nabla\mathbf{w} - q\mathbf{n} = \mathbf{0} \quad \text{on } \Gamma_D;$$

and $\boldsymbol{\eta} = (0,\, q,\, -q,\, 0)^t$ for some $q \in L^2(\Omega)$. Notice that boundary condition for $(\mathbf{w},\, q)$ on Γ_D follows from the fact that

$$\mathbf{0} = \mathbf{n}\times\underline{\mathbf{U}} = \mathbf{n}\times\underline{\mathbf{V}} + \mathbf{n}\times\boldsymbol{\eta}.$$

The following relations are easily verified:

$$\mathrm{tr}\nabla\mathbf{v} = \nabla\cdot\mathbf{v}, \quad \mathrm{tr}\nabla^{\perp}\mathbf{w} = \nabla\times\mathbf{w}, \quad \mathrm{tr}\boldsymbol{\eta} = 0, \quad \nabla\times\boldsymbol{\eta} = \nabla q. \tag{3.18}$$

By substituting decomposition (3.17) into the functional and using relations (3.18), bound (3.14), the triangle inequality, and bound (3.7), we have that

$$\begin{aligned}
&G_{-1}^{(1)}(\underline{\mathbf{U}},\, p;\, \mathbf{0},\, 0)\\
&= \|\nabla\cdot(2A_2\nabla\mathbf{v}) - \nabla p\|_{-1,D}^2 + \left\|-\frac{1}{2}\Delta\mathbf{w} + \nabla q\right\|_{-1,N}^2 + \left\|\nabla\cdot\mathbf{v} + \frac{1}{2}\nabla\times\mathbf{w} + \delta p\right\|^2\\
&\geq \|\nabla\cdot(2A_2\nabla\mathbf{v}) - \nabla p\|_{-1,D}^2 + \frac{1}{C}\left(\|\mathbf{w}\|_1^2 + \|q\|^2\right) + \left\|\nabla\cdot\mathbf{v} + \frac{1}{2}\nabla\times\mathbf{w} + \delta p\right\|^2\\
&\geq \|\nabla\cdot(2A_2\nabla\mathbf{v}) - \nabla p\|_{-1,D}^2 + \|\nabla\cdot\mathbf{v} + \delta p\|^2 + \frac{1}{C}\left(\|\mathbf{w}\|_1^2 + \|q\|^2\right)\\
&\geq \frac{1}{C}\left(\|\mathbf{v}\|_1^2 + \|p\|^2 + \|\mathbf{w}\|_1^2 + \|q\|^2\right)\\
&\geq \frac{1}{C}\left(\|\underline{\mathbf{U}}\|^2 + \|p\|^2\right),
\end{aligned}$$

which is the lower bound in (3.16) for $n = 2$.

For $n = 3$ and $\Gamma_N = \emptyset$, the following decomposition is admitted (see [3]):

$$\underline{\mathbf{U}} = \nabla\mathbf{v} + \nabla\times\underline{\Phi},$$

where $\mathbf{v} \in H_0^1(\Omega)^2$ satisfies $\Delta\mathbf{v} = \nabla\cdot\underline{\mathbf{U}}$ in Ω and $\underline{\Phi}$ is columnwise divergence free with $\mathbf{n}\times(\nabla\times\underline{\Phi}) = \underline{\mathbf{0}}$ on $\partial\Omega$. The triangle inequality and bound (3.15) imply that

$$\|\mathrm{tr}\nabla\times\underline{\Phi}\| \leq 3\|\nabla\times\underline{\Phi}\| \leq C\|\nabla\times\nabla\times\underline{\Phi}\|_{-1} = C\|\Delta\underline{\Phi}\|_{-1}. \tag{3.19}$$

By relation $\mathrm{tr}\nabla\mathbf{v} = \nabla\cdot\mathbf{v}$, bound (3.19), the triangle inequality, and bound (3.7), the proof now parallels that for $n = 2$ above. This completes the proof of the theorem. ∎

4. Linear elasticity. Here and in the next section, we restrict ourselves to linear elasticity ($\delta > 0$) with pure displacement boundary conditions ($\Gamma_N = \emptyset$), unless otherwise stated. For this case, the first-stage FOSLS functionals, $G_k^{(1)}(\underline{\mathbf{U}},\, p;\, \mathbf{f},\, g)$ ($k = -1, 0$), make use of both the displacement flux variable $\underline{\mathbf{U}}$ and the pressure variable p. But the functionals developed in [4, 9] only use the variable $\underline{\mathbf{U}}$ for the pure traction problem in planar linear elasticity. Mixed boundary conditions are common in practice, so it is important to unify these two slightly different approaches. In this section, we take the first step in this direction by eliminating the pressure variable in the divergence equation in (3.1). This is done by way of an algebraic relation between $\mathrm{tr}\underline{\mathbf{U}}$ and p that comes from the second equation in (2.1). After we analyze norm equivalence for the corresponding reduced-variable FOSLS functionals, we then take the second step using a linear transformation of equations (3.1). This transformation yields equations that match those for the pure traction formulations in [4, 9]. Since this linear transformation is invertible, norm equivalence of the corresponding functionals is immediate. This has the effect of proving that the functionals developed for the pure traction problem apply equally well to the pure displacement case. Reversing these two steps also establishes norm equivalence of the functionals that correspond to the generalized Stokes equations with pure traction boundary conditions. This suggests, of course, that both types of functionals can be effectively applied to the mixed boundary value problem, which we showed in two dimensions for the inverse-norm functional theoretically in the previous section and numerically in the companion paper [2].

To take the first step, note that solving the second equation in (3.1) for p gives

$$p = \frac{1}{\delta}(g - \mathrm{tr}\underline{\mathbf{U}}).$$

Substitution into the first equation in (3.1) and rearrangement yields the following first-order system:

$$\left\{ \begin{array}{rclcl} -\nabla \cdot \underline{\mathbf{U}} - \frac{1}{\delta}\nabla \mathrm{tr}\underline{\mathbf{U}} & = & \mathbf{f} - \frac{1}{\delta}\nabla g & \text{in} & \Omega, \\ \nabla\times \underline{\mathbf{U}} & = & \underline{\mathbf{0}} & \text{in} & \Omega \end{array} \right. \tag{4.1}$$

with the boundary condition

$$\mathbf{n} \times \underline{\mathbf{U}} = \underline{\mathbf{0}} \quad \text{on } \partial\Omega.$$

Consider the following FOSLS functionals:

$$\tilde{G}_0^{(1)}(\underline{\mathbf{U}};\, \mathbf{f},\, g) := \left\| \mathbf{f} - \frac{1}{\delta}\nabla g + \nabla \cdot \underline{\mathbf{U}} + \frac{1}{\delta}\nabla \mathrm{tr}\underline{\mathbf{U}} \right\|^2 + \|\nabla\times\underline{\mathbf{U}}\|^2 \tag{4.2}$$

for $\underline{\mathbf{U}} \in \mathcal{U}_0 := \{\underline{\mathbf{V}} \in (H(\mathrm{div};\, \Omega) \cap H_0(\mathbf{curl};\, \Omega))^n \, : \, \int_\Omega \mathrm{tr}\underline{\mathbf{V}}\, dx = 0\}$ and

$$\tilde{G}_{-1}^{(1)}(\underline{\mathbf{U}};\, \mathbf{f},\, g) := \left\| \mathbf{f} - \frac{1}{\delta}\nabla g + (\nabla \cdot \underline{\mathbf{U}})^t + \frac{1}{\delta}\nabla \mathrm{tr}\underline{\mathbf{U}} \right\|_{-1,0}^2 + \|\nabla\times\underline{\mathbf{U}}\|_{-1}^2 \tag{4.3}$$

for $\underline{\mathbf{U}} \in \mathcal{U}_{-1} := \{\underline{\mathbf{V}} \in L^2(\Omega)^{n^2} \, : \, \mathbf{n} \times \underline{\mathbf{U}} = \underline{\mathbf{0}} \text{ on } \partial\Omega\,, \int_\Omega \mathrm{tr}\underline{\mathbf{V}}\, dx = 0\}$. (Here we have $\Gamma_N = \emptyset$, so $\|\cdot\|_{-1,N}$ becomes $\|\cdot\|_{-1}$ and we write $\|\cdot\|_{-1,0}$ in place of $\|\cdot\|_{-1,D}$.)

The change of variables that we introduce in the next section produces the trace as one of the new variables. This makes it relatively easy to impose the integral condition, $\int_\Omega \mathrm{tr}\underline{\mathbf{U}}\, dx = 0$, in the definition of $\mathcal{U}_k$ $(k = -1, 0)$ on the finite element space.

THEOREM 4.1. *Assume that the domain Ω is a bounded convex polyhedron or has $C^{1,1}$ boundary. Then there exists a positive constant C independent of δ such that*

$$\frac{1}{C}\left(\|\underline{\mathbf{U}}\|_1^2 + \frac{1}{\delta^2}\|\nabla \mathrm{tr}\underline{\mathbf{U}}\|^2 \right) \le \tilde{G}_0^{(1)}(\underline{\mathbf{U}};\, \mathbf{0},\, 0) \le C\left(\|\underline{\mathbf{U}}\|_1^2 + \frac{1}{\delta^2}\|\nabla \mathrm{tr}\underline{\mathbf{U}}\|^2 \right) \tag{4.4}$$

for any $\underline{\mathbf{U}} \in \mathcal{U}_0$.

Proof. Since $\int_\Omega \mathrm{tr}\underline{\mathbf{U}}\, dx = 0$, then we may choose

$$p = -\frac{1}{\delta}\mathrm{tr}\underline{\mathbf{U}} \in H^1(\Omega)/\mathcal{R}$$

in $G_0^{(1)}(\underline{\mathbf{U}},\, p;\, \mathbf{0},\, 0)$. Together with (3.6), this implies the validity of (4.4) and, hence, the theorem. $\square$

THEOREM 4.2. *For $n = 2$, there exists a positive constant C independent of δ such that*

$$\frac{1}{C}\left(\|\underline{\mathbf{U}}\|^2 + \frac{1}{\delta^2}\|\mathrm{tr}\underline{\mathbf{U}}\|^2 \right) \le \tilde{G}_{-1}^{(1)}(\underline{\mathbf{U}};\, \mathbf{0},\, 0) \le C\left(\|\underline{\mathbf{U}}\|^2 + \frac{1}{\delta^2}\|\mathrm{tr}\underline{\mathbf{U}}\|^2 \right) \tag{4.5}$$

for any $\underline{\mathbf{U}} \in \mathcal{U}_{-1}$. For $n = 3$, this equivalence is also valid when the domain Ω is a convex polyhedron or has $C^{1,1}$ boundary.

Proof. Since $\int_\Omega \mathrm{tr}\underline{\mathbf{U}}\,dx = 0$, then we may choose

$$p = -\frac{1}{\delta}\mathrm{tr}\underline{\mathbf{U}} \in L_0^2(\Omega)$$

in $G_{-1}^{(1)}(\underline{\mathbf{U}},\, p;\, \mathbf{0},\, 0)$ which, together with Theorem 3.4, implies the validity of (4.5) and, hence, the theorem. □

To complete the transformation of the Stokes-type elasticity functionals, we take the second step by considering the equations on which the functionals in [4, 9] are based:

$$\left\{ \begin{array}{rcll} -\nabla\cdot(2A_2\underline{\mathbf{U}}) - (\frac{1}{\delta}-1)\nabla\mathrm{tr}\underline{\mathbf{U}} & = & \mathbf{f} - \frac{1}{\delta}\nabla g & \text{in} \quad \Omega, \\ \nabla\times\underline{\mathbf{U}} & = & \underline{\mathbf{0}} & \text{in} \quad \Omega. \end{array} \right. \tag{4.6}$$

It is easy to see that this system can be obtained from the Stokes-type system in (4.1) by adding some of the curl equations to the div equations:

$$-\nabla\cdot(2A_2\underline{\mathbf{U}}) - \left(\frac{1}{\delta}-1\right)\nabla\mathrm{tr}\underline{\mathbf{U}} = -\nabla\cdot\underline{\mathbf{U}} - \frac{1}{\delta}\nabla\mathrm{tr}\underline{\mathbf{U}} + \mathbf{r}(\underline{\mathbf{U}}), \tag{4.7}$$

where $\mathbf{r}(\underline{\mathbf{U}}) = (-\partial_1 U_{22} + \partial_2 U_{12}\,,\ \partial_1 U_{21} - \partial_2 U_{11})^t$ for $n=2$. (The formula for $\mathbf{r}(\underline{\mathbf{U}})$ is analogous but more complicated for $n=3$.) Note that $\mathbf{r}(\underline{\mathbf{U}}) = \mathbf{0}$ at the solution of (3.1). Denote the FOSLS functionals based on (4.6) by

$$\hat{G}_0^{(1)}(\underline{\mathbf{U}};\, \mathbf{f},\, g) = \left\| \mathbf{f} - \frac{1}{\delta}\nabla g + \nabla\cdot(2A_2\underline{\mathbf{U}}) + \left(\frac{1}{\delta}-1\right)\nabla\mathrm{tr}\underline{\mathbf{U}} \right\|^2 + \|\nabla\times\underline{\mathbf{U}}\|^2$$

and

$$\hat{G}_{-1}^{(1)}(\underline{\mathbf{U}};\, \mathbf{f},\, g) = \left\| \mathbf{f} - \frac{1}{\delta}\nabla g + \nabla\cdot(2A_2\underline{\mathbf{U}}) + \left(\frac{1}{\delta}-1\right)\nabla\mathrm{tr}\underline{\mathbf{U}} \right\|^2_{-1,0} + \|\nabla\times\underline{\mathbf{U}}\|^2_{-1}.$$

THEOREM 4.3. *The homogeneous parts of FOSLS functionals $\tilde{G}_k^{(1)}$ and $\hat{G}_k^{(1)}$ $(k=-1,0)$ are equivalent for $n=2$ or 3:*

$$\frac{1}{n+1}\tilde{G}_k^{(1)}(\underline{\mathbf{U}};\, \mathbf{0},\, 0) \le \hat{G}_k^{(1)}(\underline{\mathbf{U}};\, \mathbf{0},\, 0) \le (n+1)\tilde{G}_k^{(1)}(\underline{\mathbf{U}};\, \mathbf{0},\, 0)\,. \tag{4.8}$$

Proof. (4.8) is an immediate consequence of the relation (4.7), the triangle inequality, and the dominance of the H^{-1} norm over the H_0^{-1} norm. □

5. Practical issues. Since the expression $\mathrm{tr}\underline{\mathbf{U}} = \sum_{i=1}^n U_{ii}$ represents an intimate coupling between the U_{ii} $(i=1,2,\ldots,n)$ in the $\tilde{G}_k^{(1)}$ $(k=-1,0)$, then small δ (i.e., large λ) implies that this coupling must tend to dominate. This causes degrading performance of standard solvers, but it is eliminated in the two-dimensional case by a simple rotation applied to $\underline{\mathbf{U}}$ (see [3, 4]). We extend this idea here to three dimensions and discuss its application to the Stokes equations (see Remark 5.3).

We apply an orthogonal matrix Q to the column vector $\underline{\mathbf{U}}$ designed in part so that the first component of $Q\underline{\mathbf{U}}$ is $\frac{1}{\sqrt{3}}\mathrm{tr}\underline{\mathbf{U}}$:

$$Q = \begin{pmatrix} \frac{1}{\sqrt{3}} & 0 & 0 & 0 & \frac{1}{\sqrt{3}} & 0 & 0 & 0 & \frac{1}{\sqrt{3}} \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ \frac{1}{\sqrt{6}} & 0 & 0 & 0 & \frac{1}{\sqrt{6}} & 0 & 0 & 0 & -\frac{2}{\sqrt{6}} \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ \frac{1}{\sqrt{2}} & 0 & 0 & 0 & -\frac{1}{\sqrt{2}} & 0 & 0 & 0 & 0 \end{pmatrix}.$$

It is easy to see that the mapping Q changes the U_{ii} elements only:

$$\underline{\mathbf{V}} := Q\underline{\mathbf{U}} = (V_{11}, U_{21}, U_{31}, U_{12}, V_{22}, U_{32}, U_{13}, U_{23}, V_{33})^t,$$

where $V_{11} = \frac{1}{\sqrt{3}}\mathrm{tr}\underline{\mathbf{U}}$, $V_{22} = \frac{1}{\sqrt{6}}(U_{11} + U_{22} - 2U_{33})$, and $V_{33} = \frac{1}{\sqrt{2}}(U_{11} - U_{22})$.

As in [4], define $\tilde{\mathcal{U}}_k \equiv Q\mathcal{U}_k = \{\underline{\mathbf{V}} = Q\underline{\mathbf{U}} : \underline{\mathbf{U}} \in \mathcal{U}_k\}$ for $k = -1, 0$. Note that $\mathcal{U}_k = Q^t\tilde{\mathcal{U}}_k$ and that each vector $\underline{\mathbf{U}} \in \mathcal{U}_k$ is of the form

$$\underline{\mathbf{U}} = Q^t\underline{\mathbf{V}}, \quad \underline{\mathbf{V}} \in \tilde{\mathcal{U}}_k.$$

Note also that spaces $\mathcal{U}_k$ and $\tilde{\mathcal{U}}_k$ are the same up to boundary conditions. The solution $\underline{\mathbf{U}}$ of the first-order system (4.1) can be obtained by minimizing the quadratic functional $\tilde{G}_k^{(1)}(Q^t\underline{\mathbf{V}}; \mathbf{f}, g)$ over $\tilde{\mathcal{U}}_k$:

$$\underline{\mathbf{V}} = \mathrm{argmin}\left\{\tilde{G}_k^{(1)}(Q^t\underline{\mathbf{W}}; \mathbf{f}, g) : \underline{\mathbf{W}} \in \tilde{\mathcal{U}}_k\right\}. \tag{5.1}$$

COROLLARY 5.1. *Under the assumptions in Theorem* 4.1, *we have that*

$$\frac{1}{C}\left(\|\underline{\mathbf{V}}\|_1^2 + \frac{1}{\delta^2}\|\nabla V_{11}\|^2\right) \le \tilde{G}_0^{(1)}(Q^t\underline{\mathbf{V}}; \mathbf{0}, 0) \le C\left(\|\underline{\mathbf{V}}\|_1^2 + \frac{1}{\delta^2}\|\nabla V_{11}\|^2\right) \tag{5.2}$$

for all $\underline{\mathbf{V}} \in \tilde{\mathcal{U}}_0$. *Under the assumptions in Theorem* 4.2, *we have that*

$$\frac{1}{C}\left(\|\underline{\mathbf{V}}\|^2 + \frac{1}{\delta^2}\|V_{11}\|^2\right) \le \tilde{G}_{-1}^{(1)}(Q^t\underline{\mathbf{V}}; \mathbf{0}, 0) \le C\left(\|\underline{\mathbf{V}}\|^2 + \frac{1}{\delta^2}\|V_{11}\|^2\right) \tag{5.3}$$

for all $\underline{\mathbf{V}} \in \tilde{\mathcal{U}}_{-1}$.

Proof. Since the orthogonal mapping Q preserves the H^1 and L^2 norms, then (5.2) and (5.3) are direct consequences of Theorems 4.1 and 4.2, respectively. ∎

Remark 5.1. Note that the orthogonality of Q is not essential: it can be any invertible matrix provided one of the new variables is proportional to the trace of the old variables and its condition number is close to one.

Remark 5.2. The norm equivalence confirmed by Corollary 5.1 ensures good behavior of the numerical solution process even as the limiting incompressible case is approached. In fact, these strong statements of equivalence for the respective L^2 and H^{-1} norm functionals mean that standard finite element discretization and multigrid solution methods can be used to obtain accurate H^1 and L^2 approximation of the new variables and, hence, of the deformation and stress; cf. [4] for more detail. It may seem troublesome that the parameter δ remains in these bounds, but it is only there in a diagonal sense: as a multiplier of one of the variables, not in any coupling term. This strong δ-uniform sense of equivalence can perhaps be more easily understood by just absorbing δ into V_{11} (that is, by rescaling V_{11} by $1 + \frac{1}{\delta}$) so that δ disappears from the equivalence statement.

Remark 5.3. The change of variables determined here by Q can also be used for the Stokes equations ($\delta = 0$). This would allow either elimination of V_{11} (by replacing it by g) or approximation of V_{11} by finite elements that are more accurate than those used for the other variables. An important practical consequence of this elimination or high-order approximation is that the velocity flux variables would either satisfy the continuity equation exactly or very accurately, respectively.

REFERENCES

[1] A. Bendali and J. M. Dominguez, *A variational approach for the vector potential formulation of the Stokes and Navier-Stokes problems in three dimensional domains*, J. Math. Anal., 107 (1985), pp. 537–560.

[2] Z. Cai, C.-O. Lee, T. Manteuffel, and S. McCormick, *First-order system least squares for linear elasticity: Numerical results*, SIAM J. Sci. Comp., 21 (2000), pp. 1706–1727.

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

[4] Z. Cai, T. A. Manteuffel, S. F. McCormick, and S. V. Parter, *First-order system least squares (FOSLS) for planar linear elasticity: Pure traction problem*, SIAM J. Numer. Anal, 35 (1998), pp. 320–335.

[5] Z. Cai, X. Ye, and H. Zhang, *Least-squares finite element approximations for the the Reissner-Mindlin plate*, Numer. Linear Algebra Appl., 6 (1999), pp. 479–496.

[6] C. L. Chang, *A mixed finite element method for the Stokes problem: An acceleration-pressure formulation*, Appl. Math. Comp., 36 (1990), pp. 135–146.

[7] G. Duvaut and J. L. Lions, *Inequalities in Mechanics and Physics*, Springer-Verlag, New York, 1976.

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

[9] S.-D. Kim, T. Manteuffel, and S. McCormick, *First-order system least squares (FOSLS) for spatial linear elasticity: Pure traction*, SIAM J. Numer. Anal., submitted.