LEAST-SQUARES METHODS FOR LINEAR ELASTICITY BASED ON A DISCRETE MINUS ONE INNER PRODUCT

JAMES H. BRAMBLE, RAYTCHO D. LAZAROV,
AND JOSEPH E. PASCIAK

ABSTRACT. The purpose of this paper is to develop and analyze least-squares approximations for elasticity problems. The major advantage of the least-square formulation is that it does not require that the classical Ladyzhenskaya-Babŭska-Brezzi (LBB) condition be satisfied. By employing least-squares functionals which involve a discrete inner product which is related to the inner product in $H^{-1}(\Omega)$ (the Sobolev space of order minus one on Ω) we develop a finite element method which is unconditionally stable for problems with traction type of boundary conditions and for almost and incompressible elastic media. The use of such inner products (applied to second order problems) was proposed in an earlier paper by Bramble, Lazarov and Pasciak [7].

1. Introduction

There are many papers written on the subject of approximation schemes for Stokes equations and the equations of linear elasticity (see, [14], [16], [17], [18], [26], [39] and the included references). Mixed finite element methods involving a pair of approximation spaces are commonly used to handle the Stokes equations and avoid locking in linear elasticity problems. These spaces cannot be chosen independently of one another and, for stability, need to satisfy the so-called Ladyzhenskaya-Babuška-Brezzi (LBB) condition ([1], [15], [35]). To compute the resulting discrete approximation one must solve saddle point problems. Although much progress has been made in the development of efficient iterative procedures for solving such problems [9], [38], they still pose some difficulties.

To avoid restrictions on the pairs of approximation spaces used in the mixed formulations various stabilization techniques have been proposed and studied (see, e.g. [22], [25], [30], [31]). These stabilization techniques either add some new terms to the functional in order to make the corresponding finite element stiffness matrix uniformly stable (in the step-size h) stable, or introduce new variables (in general, these are the stresses) and again stabilize the corresponding bilinear forms. A common problem with these techniques is that the stabilization terms contain some parameters which have to be chosen in a proper way in order to have a stable scheme.

Another stabilization technique is based on the least-squares method. There are many papers dealing with the application of least-squares methods to Stokes equations and linear

Date: May 10, 1999.

elasticity (see, e.g. [3], [11], [17], [30], [31]). For a review of finite element methods of least-squares type, we refer to the recent review paper of Bochev and Gunzburger [4].

In this paper, we consider a least-squares method motivated by a regularity result (Theorem 1) for the equations of linear elasticity. This result is the most natural stability estimate for the system and is given in terms of the dual norm of the data, i.e., a negative norm. Any discrete method motivated by the stability estimate requires replacement of the negative norm in the computational algorithm. In fact, the convergence and stability properties of the resulting algorithm critically depend on proper replacement. In this paper, we develop such a replacement for a pressure/displacement formulation of the equations of linear elasticity that leads to a stable and convergent computational algorithm. Moreover, the discrete formulation is such that the approximation subspaces need only satisfy the usual essential boundary conditions for the mixed pressure/displacement formulation even when different types of boundary conditions are imposed on different parts of the boundary. Since the natural boundary conditions need not be imposed on the subspaces, the method is valid for problems with internal material interfaces such as those which result when different elastic materials are glued together. The approach is related to that given in [11]. Below we discuss some works related to the approach of our paper.

In [17], Cai, Manteffel, and McCormick discuss four different least-squares functionals for the Dirichlet boundary value problem for the Lamé equations of the linear elasticity, including the limiting case of an incompressible medium. Their formulation introduces the gradients of the velocity vector as new unknown functions and adds d^2 new unknowns, $d = 2, 3$ is the dimension of the space. One of the least-squares functionals introduced there is defined in terms of an H^{-1}-norm. The discrete replacement for the H^{-1}-norm alluded to in that paper imposes additional restrictions on the approximation spaces used for the new variables. A similar approach has been applied in [18] and [34] to the equations of the linear elasticity with pure traction boundary conditions. In contrast to the method of this paper, the above techniques do not extend to the case of mixed traction and Dirichlet boundary conditions or to the case of internal material interfaces.

The least-squares approach of this paper is based on a discrete negative norm, a technique developed in [8] and [11]. We note, that the first computable H^{-1}-norm was used by R. Falk in [24] to treat weakly the incompressibility condition $\nabla \cdot \mathbf{u} = 0$ for Stokes problems. We use the physical variables, the velocity/displacements and the pressure. By working with the original variables we have been able also to avoid the difficulties often arising in L^2-norm least-squares methods when imposing the boundary conditions. Our functional is properly scaled with respect to the parameter related to the compressibility of the medium and all estimates are independent of this parameter. Finally, the corresponding algebraic systems can be easily preconditioned by using preconditioners for standard second order problems, a task which is well understood (see, e.g., [2], [5], [10], [12], [13], [40]).

The outline of the remainder of the paper is as follows. In Section 2 we present the equations of linear elasticity and derive stability estimates which are used for the least-squares formulation. Section 3 describes and analyzes the least-squares method. By construction this method uses preconditioning and gives rise to an algebraic system for which assembly of the matrix is not feasible (a feature quite characteristic for all preconditioned systems). Nevertheless, we show that we can solve this system efficiently by an iterative method. In Section 4 we discuss an implementation of the proposed least-squares method and its computational complexity.

2. The equations of linear elasticity with mixed boundary conditions

In this section, we introduce the equations of the linear elasticity. Here, we define the necessary function spaces and provide an *a priori* inequality which is important for the stability and convergence of the least-squares methods studied in this paper.

Let Ω be a Lipschitz domain with a polygonal or polyhedral boundary in d dimensional Euclidean space for $d = 2$ or $d = 3$ with boundary $\Gamma = \Gamma_D \cup \Gamma_N$. The deformations of the elastic medium, Ω, due to given body forces $\boldsymbol{F}$ and external (boundary) forces $\mathbf{f}$ are described by the displacement vector $\boldsymbol{u} = \boldsymbol{u}(x) = (u_1, \dots, u_d)$. In the linear theory of elasticity the symmetric strain tensor is defined as

$$\epsilon_{ij} = \epsilon_{ij}(\boldsymbol{u}) \equiv \frac{1}{2}\left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i}\right), \quad i, j = 1, \dots, d.$$

Following [16], we introduce the Lagrange multiplier p by $\gamma p + \nabla \cdot \boldsymbol{u} = 0$. Then the relation between the strains, ϵ_{ij}, p, and the stresses, σ_{ij} is given by the linear Hooke's law

$$\sigma_{ij}(\boldsymbol{u}, p) = 2\mu\epsilon_{ij}(\boldsymbol{u}) - p\delta_{ij}, \quad i, j = 1, \cdots, d.$$

Here $\mu = \mu(x) > 0$ and $\lambda = \lambda(x) > 0$ are the Lamé coefficients, $\gamma = (1 - 2\sigma)/(2\mu\sigma) = 1/\lambda$, σ is Poisson's ratio and δ_{ij} is the Kronecker delta.

Remark 1. *Alternatively, we could define p to be the hydrostatic pressure. This results from using $\epsilon^D = \epsilon - d^{-1} tr(\epsilon) I$ and defining p so that*

$$\sigma_{ij} = 2\mu\epsilon^D_{ij} - p\delta_{ij}. \tag{2.1}$$

The hydrostatic pressure is thus given by

$$p = -(\lambda + 2\mu/d)\nabla \boldsymbol{u}. \tag{2.2}$$

The algorithms and analysis of this paper can be extended to this case and would be valid even when $\lambda < 0$ provided that $\lambda + 2\mu/d > 0$. However, unless λ is large, there is no particular reason to introduce the pressure since the standard finite element approximation is perfectly well behaved. We use the original definition of p for convenience.

The classical problem describing linear steady state elastic deformations of the medium Ω is: Find $\boldsymbol{u}$ and p satisfying

$$L(\boldsymbol{u}, p) = \boldsymbol{F} \quad \text{in } \Omega, \tag{2.3}$$

$$\gamma p + \nabla \cdot \boldsymbol{u} = 0 \quad \text{in } \Omega, \tag{2.4}$$

$$\boldsymbol{u} = 0 \quad \text{on } \Gamma_D, \tag{2.5}$$

$$\sigma_\nu = \mathbf{f} \quad \text{on } \Gamma_N. \tag{2.6}$$

Here

$$L(\boldsymbol{u}, p) = (L_1(\boldsymbol{u}, p), \dots, L_d(\boldsymbol{u}, p)) \text{ with } L_i(\boldsymbol{u}, p) = -\sum_{j=1}^{d} \frac{\partial \sigma_{ij}}{\partial x_j},$$

for $i = 1, \dots, d$, and

$$\sigma_\nu = \Big(\sum_{j=1}^{d} \sigma_{1j}\nu_j, \dots, \sum_{j=1}^{d} \sigma_{dj}\nu_j\Big),$$

where $\nu = (\nu_1, \dots, \nu_d)$ is the outward unit normal vector to Γ_N.

Remark 2. *The equations of elastic deformation can be written in various equivalent forms. The one given above is very convenient when the elastic medium is nonhomogeneous, i.e. $\sigma = \sigma(x)$ and $\mu = \mu(x)$. In particular, for materials with piecewise constant coefficients the following continuity conditions are satisfied on the surface Γ_0 of the coefficient discontinuity:*

$$[\boldsymbol{u}] = 0 \textit{ and } [\sigma_\nu] = 0 \textit{ on } \Gamma_0.$$

Here $[\cdot]$ denotes the difference between the limits from the two sides of Γ_0, σ_ν is as above, and ν is the normal to Γ_0.

Remark 3. *Poisson's ratio satisfies $0 < \sigma \le 1/2$. If σ is close to $1/2$ the elastic material is almost incompressible. When $\sigma = 1/2$ then (2.4) becomes the incompressibility condition and the equations coincide with the Stokes equations. In this paper we consider both incompressible and compressible elastic materials. This means that $0 < \sigma \le 1/2$ and $0 < \lambda \le \infty$. In all cases we assume that the coefficient $\mu(x)$ satisfies*

$$0 < \mu_0 \le \mu(x) \le C_0\mu_0 \textit{ in } \Omega$$

with positive constants C_0 and μ_0.

Remark 4. *If $\Gamma_D = \emptyset$ then a necessary condition for existence of a steady state deformation is equilibrium of the forces acting on the elastic body. This can be expressed in the following way: Introduce the set of all rigid body motions, for $d = 2$,*

$$\mathcal{R} = \{\boldsymbol{v} : \boldsymbol{v}(x) = A + b(-x_2, x_1) \quad \textit{for all} \quad A, x \in R^2 \quad b \in R^1\};$$

and, for $d = 3$,

$$\mathcal{R} = \{\boldsymbol{v} : \boldsymbol{v}(x) = a + b \times x \quad \textit{for all} \quad a, b, x \in R^3\}.$$

Then the steady-state elastic deformations are possible only if

$$\int_\Omega \boldsymbol{F}\boldsymbol{v}\, dx + \int_\Gamma \mathbf{f}\boldsymbol{v}\, ds = 0 \quad \textit{for all} \quad \boldsymbol{v} \in \mathcal{R}.$$

The solution $\boldsymbol{u}$ is unique provided that

$$\int_\Omega \boldsymbol{u}\boldsymbol{v}\, dx = 0 \quad \textit{for all} \quad \boldsymbol{v} \in \mathcal{R}.$$

Remark 5. *If $\Gamma_N = \emptyset$ and $\gamma = 0$ then p is determined only up to an additive constant. Thus, it is unique if we require that $\int_\Omega p\, dx = 0$.*

The existence, stability, and regularity properties of solutions of the above problem are most naturally described in terms of Sobolev spaces (see, e.g. [36], [37]). Let $(\cdot,\cdot)$ denote the $L^2(\Omega)$ inner product and $||\cdot||$ denote the corresponding norm. We will use the same inner product and norm notation for vector valued functions in the product space $(L^2(\Omega))^d$. For positive values of s, let $H^s(\Omega)$ denote the Sobolev space of order s and $||\cdot||_s$ denote the corresponding norm (cf. [29], [37]). We denote by $\mathbf{H}^1(\Omega)$ the space $(H^1(\Omega))^d$. Let $H^1_D(\Omega)$ be the set of functions in $H^1(\Omega)$ with vanishing trace on Γ_D. In the case that Γ_D is all of the boundary, we denote this space as $H^1_0(\Omega)$. Its dual will be called $H^{-1}(\Omega)$ with norm $\|\cdot\|_{-1}$. The Dirichlet form on Ω is defined by

$$D(v,w) \equiv \int_\Omega \nabla v \cdot \nabla w \; dx, \text{ for all } v,w \in H^1(\Omega).$$

For simplicity, we assume that Γ_D has positive measure. Since functions in $H^1_D(\Omega)$ vanish on Γ_D, the Poincaré inequality implies that

$$||v||_1 = D(v,v)^{1/2} \text{ for all } v \in H^1_D(\Omega),$$

is a norm equivalent to the usual Sobolev norm. Let $\mathbf{H}^1_D(\Omega)$ denote the product space $(H^1_D(\Omega))^d$. Its norm is induced by the form

$$\mathbf{D}(\boldsymbol{w},\boldsymbol{v}) \equiv \sum_{j=1}^{d} D(w_j,v_j). \tag{2.7}$$

Without ambiguity, we will use $||\cdot||_1$ to denote the norms in both $H^1_D(\Omega)$ and $\mathbf{H}^1_D(\Omega)$. We will also use Sobolev spaces with negative indices. In particular, the space $\mathbf{H}^{-1}_D(\Omega)$ is defined to be those linear functionals on $\mathbf{H}^1_D(\Omega)$ for which the norm

$$||\boldsymbol{v}||_{-1,D} = \sup_{\boldsymbol{w}\in\mathbf{H}^1_D(\Omega)} \frac{[\boldsymbol{v},\boldsymbol{w}]}{||\boldsymbol{w}||_1}$$

is finite. Here $[\boldsymbol{v},\boldsymbol{w}]$ denotes the value of the functional $\boldsymbol{v}$ at $\boldsymbol{w}$. For $\boldsymbol{v} \in (L^2(\Omega))^d$ the functional $[\boldsymbol{v},\cdot] = (\boldsymbol{v},\cdot)$ is identified with $\boldsymbol{v}$.

We will use $<\cdot,\cdot>_{\Gamma_N}$ to denote the inner product in $L^2(\Gamma_N)$. For $\gamma \neq 0$ we define the space Π to be $L^2(\Omega)$ except if $\gamma = 0$ and $\Gamma_N = \emptyset$ in which case Π is the set of functions in $L^2(\Omega)$ with zero mean value on Ω.

We introduce the bilinear form $A_0(\boldsymbol{u},\boldsymbol{v})$ by

$$A_0(\boldsymbol{u},\boldsymbol{v}) = 2\int_\Omega \mu \sum_{i,j=1}^{d} \epsilon_{ij}(\boldsymbol{u})\epsilon_{ij}(\boldsymbol{v})\, dx$$

and the bilinear form $A(\boldsymbol{u},p;\boldsymbol{v})$

$$A(\boldsymbol{u},p;\boldsymbol{v}) = \int_\Omega \sum_{i,j=1}^{d} \sigma_{ij}(\boldsymbol{u},p)\epsilon_{ij}(\boldsymbol{v})\, dx.$$

Clearly,

$$A(\boldsymbol{u}, p; \boldsymbol{v}) = A_0(\boldsymbol{u}, \boldsymbol{v}) - (p, \nabla \cdot \boldsymbol{v}).$$

The quadratic form $A(\boldsymbol{u}, p; \boldsymbol{u})$ represents the potential energy of elastic deformations.

The following lemma may be found in, for example, [26], Theorem 2.2. It is proved in [37]. We will show, for convenience of the reader, how this inequality implies Korn's inequality and another inequality important to our work in this paper.

Lemma 1. *Let Ω be a bounded domain with a Lipschitz boundary. There exists a constant $C > 0$, depending only on Ω, such that*

$$||p|| \leq C\left(||p||_{-1} + \sup_{\boldsymbol{w} \in \mathbf{H}_0^1(\Omega)} \frac{(p, \nabla \cdot \boldsymbol{w})}{||\boldsymbol{w}||_1} \right), \text{ for all } p \in L^2(\Omega). \tag{2.8}$$

¿From this lemma we may deduce the following version of Korn's inequality.

Proposition 1. *Assume that meas$(\Gamma_D) \neq 0$. Then there is a constant $C > 0$ such that*

$$C\mu_0 ||\boldsymbol{u}||_1^2 \leq A_0(\boldsymbol{u}, \boldsymbol{u}), \text{ for all } \boldsymbol{u} \in \mathbf{H}_D^1(\Omega). \tag{2.9}$$

Proof: Since $\mu(x) \geq \mu_0 > 0$ it is enough to show this inequality for $\mu \equiv 1$. To see this we first prove that

$$C||\boldsymbol{u}||_1^2 \leq ||\boldsymbol{u}||^2 + A_0(\boldsymbol{u}, \boldsymbol{u}), \text{ for all } \boldsymbol{u} \in \mathbf{H}^1(\Omega). \tag{2.10}$$

This follows by noting that, for i, j fixed and any k,

$$\frac{\partial^2 u_i}{\partial x_j \partial x_k} = \frac{\partial \epsilon_{ik}}{\partial x_j} + \frac{\partial \epsilon_{ij}}{\partial x_k} - \frac{\partial \epsilon_{jk}}{\partial x_i}$$

and applying the lemma to $p = \frac{\partial u_i}{\partial x_j}$. Inequality (2.9) follows from a standard contradiction argument, using the compact embedding of $\mathbf{H}^1(\Omega)$ in $(L^2(\Omega))^d$ and the fact that $A_0(\boldsymbol{u}, \boldsymbol{u}) = 0$, with $\boldsymbol{u} \in \mathbf{H}_D^1(\Omega)$ implies $\boldsymbol{u} = 0$. That is, if (2.9) is not true, there exists a sequence $\{\boldsymbol{u}_n\}$, with $\boldsymbol{u}_n \in \mathbf{H}_D^1(\Omega)$, such that $\|\boldsymbol{u}_n\|_1 = 1$ and $A_0(\boldsymbol{u}_n, \boldsymbol{u}_n) \to 0$. By compactness there is a subsequence, call it again $\{\boldsymbol{u}_n\}$, such that $\boldsymbol{u}_n \to \boldsymbol{u}$ in $(L^2(\Omega))^d$. Inequality (2.10) implies that $\boldsymbol{u}_n \to \boldsymbol{u}$ in $\mathbf{H}_D^1(\Omega)$. Thus $A_0(\boldsymbol{u}, \boldsymbol{u}) = 0$ which can hold only if $\boldsymbol{u} = 0$. This is a contradiction since the assumption that $\|\boldsymbol{u}_n\|_1 = 1$ implies that $\|\boldsymbol{u}\|_1 = 1$. Thus (2.9) is proved.

We may also deduce from the Lemma 1 another inequality which is crucial to this paper.

Proposition 2. *Assume that meas$(\Gamma_D) \neq 0$. Then there is a constant $C > 0$ such that*

$$||p|| \leq C \sup_{\boldsymbol{w} \in \mathbf{H}_D^1(\Omega)} \frac{(p, \nabla \cdot \boldsymbol{w})}{||\boldsymbol{w}||_1}, \text{ for all } p \in \Pi. \tag{2.11}$$

Proof: To see this we note that by Lemma 1,

$$||p|| \leq C\left(||p||_{-1} + \sup_{\boldsymbol{w} \in \mathbf{H}_D^1(\Omega)} \frac{(p, \nabla \cdot \boldsymbol{w})}{||\boldsymbol{w}||_1} \right), \text{ for all } p \in L^2(\Omega).$$

Inequality (2.11) follows from a contradiction argument similar to that used in the previous proposition. We use the compact embedding of $L^2(\Omega)$ in $H^{-1}(\Omega)$ and the fact that if $p \in \Pi$ and

$$\sup_{\boldsymbol{w} \in \mathbf{H}^1_D(\Omega)} \frac{(p, \nabla \cdot \boldsymbol{w})}{||\boldsymbol{w}||_1} = 0$$

then $p = 0$.

The weak solution $(\boldsymbol{u}, p)$ in $\mathbf{H}^1_D(\Omega) \times \Pi$ of the problem (2.3)-(2.6) satisfies

$$\begin{aligned} A(\boldsymbol{u}, p; \boldsymbol{v}) + (\nabla \cdot \boldsymbol{u} + \gamma p, q) &= (\boldsymbol{F}, \boldsymbol{v}) + < \mathbf{f}, \boldsymbol{v} >_{\Gamma_N}, \\ \text{for all} \qquad & \boldsymbol{v} \in \mathbf{H}^1_D(\Omega), \text{ and } q \in \Pi. \end{aligned} \tag{2.12}$$

It follows from the above two propositions that problem (2.3)–(2.6) has unique solution in $\mathbf{H}^1_D(\Omega) \times \Pi$ for any $\boldsymbol{F} \in (L^2(\Omega))^d$ and $\mathbf{f} \in (L^2(\Gamma_N))^d$. Indeed, $(\boldsymbol{F}, \boldsymbol{v}) + < \mathbf{f}, \boldsymbol{v} >_{\Gamma_N}$, can be replaced by $[\widetilde{\boldsymbol{F}}, \boldsymbol{v}]$ where $\widetilde{\boldsymbol{F}}$ is any functional in $\mathbf{H}^{-1}_D(\Omega)$.

We define the operator $\mathcal{L}(\boldsymbol{v}, p) : \ \mathbf{H}^1_D(\Omega) \times \Pi \mapsto H^{-1}_D(\Omega)$ by

$$[\mathcal{L}(\boldsymbol{v}, p), \boldsymbol{w}] = A(\boldsymbol{v}, p; \boldsymbol{w}), \text{ for all } \boldsymbol{w} \in \mathbf{H}^1_D(\Omega).$$

Clearly

$$||\mathcal{L}(\boldsymbol{v}, p)||_{-1,D} = \sup_{\boldsymbol{w} \in \mathbf{H}^1_D(\Omega)} \frac{A(\boldsymbol{v}, p; \boldsymbol{w})}{||\boldsymbol{w}||_1} \le C(\mu_0 ||\boldsymbol{v}||_1 + ||p||).$$

The following theorem plays a leading role in motivating the least-squares method developed in the following section. This result follows directly from applying Proposition 2 and Theorem 1.2 of [16]. We include a proof since the technique is similar to that used in our subsequent analysis and gives some indication as to the behavior of the constants.

Theorem 1. *There exists a constant $C > 0$ independent of $\boldsymbol{v} \in \mathbf{H}^1_D(\Omega)$ and $p \in \Pi$ such that*

$$\begin{aligned} C(\mu_0 ||\boldsymbol{v}||_1 + ||p||) &\le ||\mathcal{L}(\boldsymbol{v}, p)||_{-1,D} + \mu_0 ||\nabla \cdot \boldsymbol{v} + \gamma p||, \\ & \textit{for all } \boldsymbol{v} \in \mathbf{H}^1_D(\Omega), \ p \in \Pi. \end{aligned} \tag{2.13}$$

Remark 6. *The above theorem holds for the hydrostatic pressure formulation, i.e., (2.1)-(2.2), provided that $\lambda + 2\mu/d > 0$. The proof is somewhat more complicated and involves two cases. Although the restriction $\lambda > 0$ in our theorem makes it less general, it is not important from a practical point of view. Indeed, the introduction of a pressure in the system is only of interest in the case of large λ. Otherwise, one should eliminate the pressure and approximate the resulting elliptic system by a standard Galerkin method.*

Proof: By Korn's inequality (2.9),

$$
\begin{aligned}
C\mu_0\|\boldsymbol{v}\|_1^2 &\le A_0(\boldsymbol{v},\boldsymbol{v}) = A(\boldsymbol{v},p;\boldsymbol{v}) + (p,\nabla\cdot\boldsymbol{v}) \\
&\le \|\boldsymbol{v}\|_1 \sup_{\boldsymbol{w}\in\mathbf{H}_D^1(\Omega)} \frac{A(\boldsymbol{v},p;\boldsymbol{w})}{\|\boldsymbol{w}\|_1} + (p,\nabla\cdot\boldsymbol{v}+\gamma p) - (\gamma p,p) \\
&\le \|\mathcal{L}(\boldsymbol{v},p)\|_{-1,D}\|\boldsymbol{v}\|_1 + \|p\|\,\|\nabla\cdot\boldsymbol{v}+\gamma p\|.
\end{aligned} \tag{2.14}
$$

By Proposition 2,

$$
\begin{aligned}
\|p\| &\le C \sup_{\boldsymbol{w}\in\mathbf{H}_D^1(\Omega)} \frac{(p,\nabla\cdot\boldsymbol{w})}{\|\boldsymbol{w}\|_1} \\
&\le C \sup_{\boldsymbol{w}\in\mathbf{H}_D^1(\Omega)} \frac{|(p,\nabla\cdot\boldsymbol{w}) - A_0(\boldsymbol{v},\boldsymbol{w})| + |A_0(\boldsymbol{v},\boldsymbol{w})|}{\|\boldsymbol{w}\|_1} \\
&\le C \sup_{\boldsymbol{w}\in\mathbf{H}_D^1(\Omega)} \frac{A(\boldsymbol{v},p;\boldsymbol{w})}{\|\boldsymbol{w}\|_1} + C \sup_{\boldsymbol{w}\in\mathbf{H}_D^1(\Omega)} \frac{A_0(\boldsymbol{v},\boldsymbol{w})}{\|\boldsymbol{w}\|_1} \\
&\le C(\|\mathcal{L}(\boldsymbol{v},p)\|_{-1,D} + \mu_0\|\boldsymbol{v}\|_1).
\end{aligned}
$$

Combining (2.14) and (2) we obtain

$$
\begin{aligned}
C\mu_0\|\boldsymbol{v}\|_1^2 &\le \|\mathcal{L}(\boldsymbol{v},p)\|_{-1,D}\|\boldsymbol{v}\|_1 \\
&+ C\|\nabla\cdot\boldsymbol{v}+\gamma p\|(\|\mathcal{L}(\boldsymbol{v},p)\|_{-1,D} + \mu_0\|\boldsymbol{v}\|_1),
\end{aligned} \tag{2.15}
$$

which easily leads to the required inequality (2.13).

We can now give a least-squares reformulation of (2.3)–(2.6) or (2.12) as follows: Find $\boldsymbol{u}\in\mathbf{H}_D^1(\Omega)$ and $p\in\Pi$ satisfying

$$
\begin{aligned}
&(\mathcal{L}(\boldsymbol{u},p),\mathcal{L}(\boldsymbol{v},q))_{-1} + (\nabla\cdot\boldsymbol{u}+\gamma p,\nabla\cdot\boldsymbol{v}+\gamma q) \\
&\quad = (\widetilde{\boldsymbol{F}},\mathcal{L}(\boldsymbol{v},q))_{-1} \text{ for all } \boldsymbol{v}\in\mathbf{H}_D^1(\Omega),\ q\in\Pi.
\end{aligned} \tag{2.16}
$$

Here $(\cdot,\cdot)_{-1}$ denotes the inner product in $\mathbf{H}_D^{-1}(\Omega)$ and $\widetilde{\boldsymbol{F}}$ is the functional given by

$$[\widetilde{\boldsymbol{F}},\boldsymbol{v}] = (\boldsymbol{F},\boldsymbol{v}) + <\mathbf{f},\boldsymbol{v}>_{\Gamma_N}.$$

Theorem 1 shows that the above bilinear form is coercive. It is straightforward to check that it is bounded and hence the solution pair exists, is unique, and satisfies

$$\mu_0\|\boldsymbol{u}\|_1 + \|p\| \le C\|\widetilde{\boldsymbol{F}}\|_{-1,D}.$$

3. The finite element spaces and their properties

To approximately solve (2.3)–(2.6), we introduce a pair of subspaces $\mathbf{V}_h\subset\mathbf{H}_D^1(\Omega)$ and $\Pi_h\subset\Pi$ indexed by h in the interval $0<h<1$. We do this by partitioning the region $\Omega=\cup_i\bar\tau_i$ into triangles or tetrahedra and denote by $\mathcal{T}=\{\tau\}$, the set of all finite elements. We further assume that Γ_D aligns with the mesh. This means that Γ_D consists of a union of edges of $\mathcal{T}$ in the two dimensional case and a union of faces of $\mathcal{T}$ in the three dimensional case. We let ϵ be an edge (face) of a $\tau \in \mathcal{T}$ and $\mathcal{E}$ be the set of all interior edges (faces). Let h_τ denote the diameter of the triangle τ. The mesh parameter h is defined to be

$$h = \max_{\tau \in \mathcal{T}} h_\tau.$$

As usual, the boundaries of two triangles or tetrahedra will intersect at either a vertex, an entire edge or an entire face. We assume that the triangulations are locally quasi-uniform. By this we mean that there is a constant $0 < c < 1$ such that each triangle contains a ball of radius ch_τ. With some abuse of semantics, we shall refer to τ as a triangle in both the two and three dimensional case. Spaces defined with respect to rectangular or parallelepiped partitioning of Ω pose no added difficulty.

For some integer $r \geq 2$, let $V_h \subset H^1_D(\Omega)$ denote the functions which are piecewise polynomials of degree less than r with respect to the triangles, continuous on Ω and vanish on Γ_D. An obvious choice for $\mathbf{V}_h$ is $(V_h)^d$. Let Π_h denote a space of functions which are piecewise polynomial with respect to the triangles defining the mesh. We only assume that Π_h provides $r-1$'st order approximation. Note that the functions in Π_h can be discontinuous but need not be. There is a nodal basis associated with these spaces (see, e.g. [19]) and a corresponding nodal interpolation operator. In the case of $\Gamma_N = \emptyset$ and $\gamma = 0$, we set Π_h to be the subset of the functions defined above with zero mean value.

There exists a constant C_1 not depending on h such that for any $\boldsymbol{v} \in \mathbf{H}^1_D(\Omega)$, there exists $\boldsymbol{V} \in \mathbf{V}_h$ satisfying

$$\sum_{\tau \in \mathcal{T}} \{ h_\tau^{-2} \|\boldsymbol{V} - \boldsymbol{v}\|_\tau^2 + \|\boldsymbol{V} - \boldsymbol{v}\|_{1,\tau}^2 \} \leq C_1 \|\boldsymbol{v}\|_1^2. \tag{3.1}$$

To develop the least-squares method, we shall need some operators and additional norms and inner products on the discrete spaces just defined. We define a weighted L^2-inner product and corresponding norm,

$$(\boldsymbol{V}, \boldsymbol{W})_h = \sum_{\tau \in \mathcal{T}} h_\tau^2 \int_\tau \boldsymbol{V}(x) \cdot \boldsymbol{W}(x)\, dx, \quad ||\boldsymbol{V}||_h = (\boldsymbol{V}, \boldsymbol{V})_h^{1/2}.$$

We will often apply this norm to derivatives of piecewise smooth functions. In such cases, the derivative will be evaluated locally (not in the distributional sense).

We will also need edge norms and inner products. We introduce the bilinear form

$$< u, v >_{h,I} = \sum_{\epsilon \in \mathcal{E}} h_{\tau(\epsilon)} \int_\epsilon uv\, ds. \tag{3.2}$$

Here $\tau(\epsilon)$ denotes any triangle (tetrahedron) which has ϵ as an edge (face). Similarly,

$$< u, v >_{h,\Gamma_N} = \sum_{\epsilon \subset \Gamma_N} h_{\tau(\epsilon)} \int_\epsilon uv\, ds. \tag{3.3}$$

The corresponding seminorms are denoted by

$$\|v\|_{h,I} =< u, v >_{h,I}^{1/2} \quad \text{and} \quad \|v\|_{h,\Gamma_N} =< u, v >_{h,\Gamma_N}^{1/2} . \tag{3.4}$$

Finally, for $\boldsymbol{v} \in \mathbf{H}_D^{-1}(\Omega)$, we define the discrete semi-norm

$$||\boldsymbol{v}||_{-1,h} = \sup_{\boldsymbol{W} \in \mathbf{V}_h} \frac{[\boldsymbol{v}, \boldsymbol{W}]}{||\boldsymbol{W}||_1}. \tag{3.5}$$

This is a norm when restricted to the space $\boldsymbol{V}_h$.

We define the operator $\mathcal{L}_h(\boldsymbol{v}, q) : \ \mathbf{H}_D^1(\Omega) \times \Pi \mapsto \boldsymbol{V}_h$ by the identity:

$$(\mathcal{L}_h(\boldsymbol{v}, q), \boldsymbol{W}) = A(\boldsymbol{v}, q; \boldsymbol{W}) \text{ for all } \boldsymbol{W} \in \boldsymbol{V}_h. \tag{3.6}$$

Now we formulate the main *a priori* estimate for the least-squares finite element method.

Theorem 2. *There is a constant $C > 0$ independent of h such that*

$$\begin{aligned} C(\mu_0||\boldsymbol{U}||_1 + ||P||) \ &\leq \ ||\mathcal{L}_h(\boldsymbol{U}, P)||_{-1,h} + \mu_0||\nabla \cdot \boldsymbol{U} + \gamma P|| \\ &+ \ ||\sigma_\nu||_{h,\Gamma_N} + ||[\sigma_\nu]||_{h,I} + ||L(\boldsymbol{U}, P)||_h \end{aligned} \tag{3.7}$$

for all $P \in \Pi_h$ and $\boldsymbol{U} \in \boldsymbol{V}_h$. Here σ_ν is defined in terms of $\sigma_{ij} = \sigma_{ij}(\boldsymbol{U}, P)$ and $[\sigma_\nu]$ denotes the jump of σ_ν across the interelement boundaries.

Proof: Here and in the remainder of this paper, C with or without subscript will denote a generic positive constant independent of h, μ_0 and γ. These constants may represent different values in different occurrences.

We start by deriving an estimate for $||\boldsymbol{U}||_1$. First, by using the same argument as that for (2.14), we get

$$\mu_0||\boldsymbol{U}||_1^2 \leq C\Big(||\mathcal{L}_h(\boldsymbol{U}, P)||_{-1,h}||\boldsymbol{U}||_1 + ||P||\,||\nabla \cdot \boldsymbol{U} + \gamma P||\Big). \tag{3.8}$$

Next, we derive an estimate for $P \in \Pi_h$. By Proposition 2,

$$\begin{aligned} ||P|| &\leq C \sup_{\boldsymbol{v} \in \mathbf{H}_D^1(\Omega)} \frac{(P, \nabla \cdot \boldsymbol{v})}{||\boldsymbol{v}||_1} \\ &\leq C \sup_{\boldsymbol{v} \in \mathbf{H}_D^1(\Omega)} \frac{|(P, \nabla \cdot \boldsymbol{V})| + |(P, \nabla \cdot (\boldsymbol{v} - \boldsymbol{V}))|}{||\boldsymbol{v}||_1} \end{aligned} \tag{3.9}$$

where $\boldsymbol{V}$ satisfies (3.1). Now we estimate two terms in the right side of this inequality separately. For the first, we essentially repeat the proof of (2) and use the definition (3.6) of the operator $\mathcal{L}_h$ to obtain

$$|(P, \nabla \cdot \boldsymbol{V})| \leq C\Big(||\mathcal{L}_h(\boldsymbol{U}, P)||_{-1,h} + \mu_0||\boldsymbol{U}||_1\Big)\,||\boldsymbol{v}||_1.$$

The second term of (3.9) is handled in the following manner. Adding and subtracting $A_0(\boldsymbol{U}, \boldsymbol{v} - \boldsymbol{V})$ gives

$$|(P, \nabla \cdot (\boldsymbol{v} - \boldsymbol{V}))| \le |A(\boldsymbol{U}, P; \boldsymbol{v} - \boldsymbol{V})| + |A_0(\boldsymbol{U}, \boldsymbol{v} - \boldsymbol{V})|.$$

Next, by (3.1)

$$|A_0(\boldsymbol{U}, \boldsymbol{v} - \boldsymbol{V})| \le \mu_0 ||\boldsymbol{U}||_1 ||\boldsymbol{v} - \boldsymbol{V}||_1 \le C\mu_0 ||\boldsymbol{U}||_1 ||\boldsymbol{v}||_1.$$

The remaining term is split into integrals over all finite elements. Integrating by parts over each element yields

$$\begin{aligned} &|A(\boldsymbol{U}, P; \boldsymbol{v} - \boldsymbol{V})| \\ &\quad = |\sum_{\tau \in \mathcal{T}} \big(- \int_\tau L(\boldsymbol{U}, P) \cdot (\boldsymbol{v} - \boldsymbol{V}) dx + \int_{\partial\tau} \sigma_\nu \cdot (\boldsymbol{v} - \boldsymbol{V}) ds \big)| \\ &\quad \le \sum_{\tau \in \mathcal{T}} |\int_\tau L(\boldsymbol{U}, P) \cdot (\boldsymbol{v} - \boldsymbol{V}) dx| \\ &\qquad + \sum_{\epsilon \in \mathcal{E}} \int_\epsilon |[\sigma_\nu] \cdot (\boldsymbol{v} - \boldsymbol{V})| ds + |\int_{\Gamma_N} \sigma_\nu \cdot (\boldsymbol{v} - \boldsymbol{V}) ds|. \end{aligned}$$

Here $[\sigma_\nu]$ is the jump of σ_ν across the inter-element boundary. Note, that σ_ν is computed using $(\boldsymbol{U}, P)$, i.e from $\sigma_{ij} = 2\mu\epsilon_{ij}(\boldsymbol{U}) - P\delta_{ij}$. Using the well known inequality

$$\int_{\partial\tau} |\theta|^2 \, ds \le C(h_\tau^{-1} ||\theta||^2_{L^2(\tau)} + h_\tau ||\theta||^2_{H^1(\tau)}), \tag{3.10}$$

it follows from (3.1) that

$$\begin{aligned} \frac{|(P, \nabla \cdot \boldsymbol{v})|}{||\boldsymbol{v}||_1} \quad &\le \quad C\Big(||\mathcal{L}_h(\boldsymbol{U}, P)||_{-1,h} + \mu_0 ||\boldsymbol{U}||_1 \\ &+ \quad ||\sigma_\nu||_{h,\Gamma_N} + ||[\sigma_\nu]||_{h,I} + ||L(\boldsymbol{U}, P)||_h \Big), \end{aligned} \tag{3.11}$$

that is,

$$\begin{aligned} C||P|| \le ||\mathcal{L}_h(\boldsymbol{U}, P)||_{-1,h} + \mu_0 ||\boldsymbol{U}||_1 + ||\sigma_\nu||_{h,\Gamma_N} \\ + ||[\sigma_\nu]||_{h,I} + ||L(\boldsymbol{U}, P)||_h. \end{aligned} \tag{3.12}$$

Combining (3.12) with (3.8) yields (3.7). This completes the proof of the theorem.

Remark 7. *It is easy to show that for $(\boldsymbol{U}, P) \in \boldsymbol{V}_h \times \Pi_h$, the right-hand side of the inequality (3.7) is bounded from above by $C(\mu_0 ||\boldsymbol{U}||_1 + ||P||)$. Therefore, the right-hand side of (3.7) gives an equivalent norm on $\boldsymbol{V}_h \times \Pi_h$.*

4. The least-squares method

In this section, we introduce a least-squares finite element method for the equations of linear elasticity which involves direct approximations of the original variables $\boldsymbol{u}$ and q.

Let us denote by $(\cdot,\cdot)_{-1,h}$ the inner product corresponding to the norm $||\cdot||_{-1,h}$. It is not difficult to show that

$$(\boldsymbol{v},\boldsymbol{w})_{-1,h} = [\boldsymbol{w}, \mathbf{T}_h \boldsymbol{v}] \tag{4.1}$$

where $\mathbf{T}_h : \mathbf{H}^{-1}_D(\Omega) \mapsto \mathbf{V}_h$ is the solution operator defined by

$$\mathbf{D}(\mathbf{T}_h \boldsymbol{v}, \boldsymbol{X}) = [\boldsymbol{v}, \boldsymbol{X}] \text{ for all } \boldsymbol{X} \in \mathbf{V}_h. \tag{4.2}$$

For $\boldsymbol{V} \in \mathbf{V}_h$, since $\boldsymbol{V} \in L^2(\Omega)$, $(\mathbf{T}_h \boldsymbol{V}, \boldsymbol{V}) = [\boldsymbol{V}, \mathbf{T}_h \boldsymbol{V}] = (\boldsymbol{V},\boldsymbol{V})_{-1,h}$. Now let $\mathbf{B}_h$: $\mathbf{H}^{-1}_D(\Omega) \mapsto \mathbf{V}_h$ be an operator which is symmetric on $L^2(\Omega)$ and positive semidefinite and is spectrally equivalent to $\mathbf{T}_h$ on $\mathbf{V}_h$. This means that there exist constants C_0 and C_1 independent of h such that

$$C_0(\mathbf{T}_h \boldsymbol{V}, \boldsymbol{V}) \le (\mathbf{B}_h \boldsymbol{V}, \boldsymbol{V}) \le C_1(\mathbf{T}_h \boldsymbol{V}, \boldsymbol{V}), \text{ for all } \boldsymbol{V} \in \mathbf{V}_h. \tag{4.3}$$

Thus, on $\mathbf{V}_h$, $(\mathbf{B}_h\cdot,\cdot)^{1/2}$ is a norm equivalent to $||\cdot||_{-1,h}$.

There is a vast literature concerning techniques for developing preconditioners for symmetric positive definite problems, especially for discretizations of elliptic boundary value problems (see, e.g., [5], [20], [21], [27]). The best preconditioners satisfy (4.3) with constants C_0 and C_1 independent of the mesh parameter. In addition, a good preconditioner is economical to evaluate. This means that the cost of computing the action of $\mathbf{B}_h$ applied to an arbitrary vector should be much less than that of applying $\mathbf{T}_h$. For our application, low cost preconditioners are known for which (4.3) holds with C_0 and C_1 independent of the mesh size and hence the number of unknowns (see, e.g., [2], [5], [10], [12], [13], [40]).

The least-squares method which we shall consider is based on the form

$$\begin{aligned}
\langle\!\langle (\boldsymbol{u},p),(\boldsymbol{v},q)\rangle\!\rangle_1 &\equiv (\mathcal{L}_h(\boldsymbol{u},p), \mathbf{B}_h \mathcal{L}_h(\boldsymbol{v},q)) \\
&\quad + (L(\boldsymbol{u},p), L(\boldsymbol{v},q))_h + < \sigma_\nu(\boldsymbol{u},p), \sigma_\nu(\boldsymbol{v},q) >_{h,\Gamma_N} \\
&\quad + < [\sigma_\nu(\boldsymbol{u},p)], [\sigma_\nu(\boldsymbol{v},q)] >_{h,I} + (\nabla\cdot\boldsymbol{u} + \gamma p, \nabla\cdot\boldsymbol{v} + \gamma q).
\end{aligned} \tag{4.4}$$

The least-squares solution is the pair $(\boldsymbol{U},P) \in \mathbf{V}_h \times \Pi_h$ satisfying

$$\begin{aligned}
\langle\!\langle (\boldsymbol{U},P),(\boldsymbol{V},Q)\rangle\!\rangle_1 &= (\boldsymbol{F}, \mathbf{B}_h \mathcal{L}_h(\boldsymbol{V},Q)) + < \mathbf{f}, \mathbf{B}_h \mathcal{L}_h(\boldsymbol{V},Q) >_{\Gamma_N} \\
&\quad + < \mathbf{f}, \sigma_\nu(\boldsymbol{V},Q) >_{h,\Gamma_N} + (\boldsymbol{F}, L(\boldsymbol{V},Q))_h
\end{aligned} \tag{4.5}$$

for all $(\boldsymbol{V},Q)$ in $\mathbf{V}_h \times \Pi_h$. It is a direct consequence of Theorem 2 and (4.3) that for $\boldsymbol{F} \in (L^2(\Omega))^d$ and $\mathbf{f} \in (L^2(\Gamma_N))^d$, the solution $(\boldsymbol{U},P)$ of (4.5) exists and is unique. The following theorem shows that the solution $(\boldsymbol{U},P)$ of the approximate problem (4.5) is close to the solution $(\boldsymbol{u},p)$ of (2.3)–(2.6). For convenience, we give a proof of the theorem in the case of a globally quasi-uniform mesh. It is based on the following well know approximation properties for the subspaces:

(1) For $\boldsymbol{v} \in (H^r(\Omega) \cap H^1_D(\Omega))^d$,

$$\inf_{\boldsymbol{V} \in \mathbf{V}_h} ||\boldsymbol{v} - \boldsymbol{V}||_1 \leq C h^{r-1} ||\boldsymbol{v}||_r; \tag{4.6}$$

(2) For $q \in H^{r-1}(\Omega) \cap \Pi$,

$$\inf_{Q \in \Pi_h} ||q - Q|| \leq C h^{r-1} ||q||_{r-1}. \tag{4.7}$$

The constant C appearing above is independent of the approximation parameter h.

Theorem 3. *Let $(\boldsymbol{U}, P)$ solve (4.5) and $(\boldsymbol{u}, p)$ solve (2.3)–(2.6). Assume that the triangulation is globally quasi-uniform and let $\mathbf{V}_h$ and Π_h be as described in the previous section and satisfy the approximation conditions (4.6), (4.7) with $r \geq 2$. Assume that $\boldsymbol{F} \in (L^2(\Omega))^d$, $\mathbf{f} \in (L^2(\Gamma_N))^d$, and that the solution $(\boldsymbol{u}, p)$ is in $(H^r(\Omega) \cap H^1_D(\Omega))^d \times H^{r-1}(\Omega)$. Then*

$$||\boldsymbol{U} - \boldsymbol{u}||_1 + ||P - p|| \leq C h^{r-1} \big(||\boldsymbol{u}||_r + ||p||_{r-1} \big).$$

Proof: By (2.12) and (3.6), since $\boldsymbol{V}_h \subset H^1_D(\Omega)$,

$$(\mathcal{L}_h(\boldsymbol{u}, p), \boldsymbol{V}) = (\boldsymbol{F}, \boldsymbol{V}) + < \mathbf{f}, \boldsymbol{V} >_{\Gamma_N} \quad \text{for all } \boldsymbol{V} \in \boldsymbol{V}_h. \tag{4.8}$$

Using (2.3) - (2.6) gives

$$\begin{aligned} \langle\!\langle (\boldsymbol{u}, p), (\boldsymbol{V}, Q) \rangle\!\rangle_1 = (\boldsymbol{F}, \mathbf{B}_h \mathcal{L}_h(\boldsymbol{V}, Q)) &+ < \mathbf{f}, \mathbf{B}_h \mathcal{L}_h(\boldsymbol{V}, Q) >_{\Gamma_N} \\ &+ < \mathbf{f}, \sigma_\nu(\boldsymbol{V}, Q) >_{h,\Gamma_N} + (\boldsymbol{F}, L(\boldsymbol{V}, Q))_h \end{aligned} \tag{4.9}$$

for all $(\boldsymbol{V}, Q)$ in $\mathbf{V}_h \times \Pi_h$. By (4.6) and (4.7), there exists $\widetilde{P} \in \Pi_h$ and $\widetilde{\boldsymbol{U}} \in \boldsymbol{V}_h$ satisfying

$$||p - \widetilde{P}|| \leq C h^{r-1} ||p||_{r-1} \tag{4.10}$$

and

$$||\boldsymbol{u} - \widetilde{\boldsymbol{U}}|| + h ||\boldsymbol{u} - \widetilde{\boldsymbol{U}}||_1 \leq C h^r ||\boldsymbol{u}||_r. \tag{4.11}$$

Setting $(\widetilde{\boldsymbol{E}}, \tilde{e}) = (\widetilde{\boldsymbol{U}} - \boldsymbol{U}, \widetilde{P} - P)$, Theorem 1, (4.3) and (4.9) give that

$$||\widetilde{\boldsymbol{E}}||_1^2 + ||\tilde{e}||^2 \leq C \langle\!\langle (\widetilde{\boldsymbol{E}}, \tilde{e}), (\widetilde{\boldsymbol{E}}, \tilde{e}) \rangle\!\rangle_1 = C \langle\!\langle (\widetilde{\boldsymbol{U}} - \boldsymbol{u}, \widetilde{P} - p), (\widetilde{\boldsymbol{E}}, \tilde{e}) \rangle\!\rangle_1.$$

It immediately follows that

$$\begin{aligned} ||\widetilde{\boldsymbol{E}}||_1^2 + ||\tilde{e}||^2 &\leq C \langle\!\langle (\widetilde{\boldsymbol{U}} - \boldsymbol{u}, \widetilde{P} - p), (\widetilde{\boldsymbol{U}} - \boldsymbol{u}, \widetilde{P} - p) \rangle\!\rangle_1 \\ &\leq C \big(||\mathcal{L}_h(\widetilde{\boldsymbol{U}} - \boldsymbol{u}, \widetilde{P} - p)||_{-1,h}^2 + ||L(\widetilde{\boldsymbol{U}} - \boldsymbol{u}, \widetilde{P} - p)||_h^2 \\ &\quad + ||[\sigma_\nu(\widetilde{\boldsymbol{U}} - \boldsymbol{u}, \widetilde{P} - p)]||_{h,I}^2 + ||\sigma_\nu(\widetilde{\boldsymbol{U}} - \boldsymbol{u}, \widetilde{P} - p)||_{h,\Gamma_N}^2 \\ &\quad + ||\nabla \cdot (\widetilde{\boldsymbol{U}} - \boldsymbol{u}) + \gamma(\widetilde{P} - p)||^2 \big). \end{aligned} \tag{4.12}$$

The last inequality above follows from (4.3) and (4.1).

We now bound the terms on the right-hand side of (4.12). It easily follows from (4.10) and (4.11) that

$$
\begin{aligned}
\|\mathcal{L}_h(\widetilde{\boldsymbol{U}}-\boldsymbol{u},\widetilde{P}-p)\|_{-1,h} &\le C(\|\widetilde{\boldsymbol{U}}-\boldsymbol{u}\|_1+\|\widetilde{P}-p\|) \\
&\le Ch^{r-1}\big(\|\boldsymbol{u}\|_r+\|p\|_{r-1}\big).
\end{aligned} \tag{4.13}
$$

Let $\bar{\Pi}_h$ denote the set of discontinuous piecewise polynomial functions of degree less than $r-1$ with respect to the triangulation defining Π_h and let $\bar{P}$ be the $L^2(\Omega)$ projection of p into $\bar{\Pi}_h$. Then,

$$\|\nabla(\widetilde{P}-p)\|_h \le C(\|\nabla(\widetilde{P}-\bar{P})\|_h+\|\nabla(\bar{P}-p)\|_h).$$

Since the mesh is quasi-uniform, we may apply the inverse inequality for the term $\nabla(\widetilde{P}-\bar{P})$. Then using the approximation property for both $\widetilde{P}$ and the local projection $\bar{P}$, it follows that

$$\|\nabla(\widetilde{P}-p)\|_h \le C\|\widetilde{P}-\bar{P}\|+Ch^{r-1}\|p\|_{r-1} \le Ch^{r-1}\|p\|_{r-1}. \tag{4.14}$$

Now we estimate $\|L(\widetilde{\boldsymbol{U}}-\boldsymbol{u},\widetilde{P}-p)\|_h$. Note, that the i-component of the operator L is given by

$$
\begin{aligned}
L_i(\widetilde{\boldsymbol{U}}-\boldsymbol{u},\widetilde{P}-p) &= -\sum_{j=1}^{d}\frac{\partial\sigma_{ij}(\widetilde{\boldsymbol{U}}-\boldsymbol{u},\widetilde{P}-p)}{\partial x_j} \\
&= -\sum_{j=1}^{d}\frac{\partial(\mu\epsilon_{ij}(\widetilde{\boldsymbol{U}}-\boldsymbol{u})-(\widetilde{P}-p)\delta_{ij})}{\partial x_j}.
\end{aligned}
$$

The derivatives of $(\widetilde{P}-p)$ have been already estimated. Here we need to estimate the norm of the derivatives of the strains computed for $\widetilde{\boldsymbol{U}}-\boldsymbol{u}$ on an element by element basis. We use a similar argument as that given above for the pressure (which again involves local L^2-projections of $\boldsymbol{u}$) and the assumption that the coefficient μ is piecewise smooth to get

$$\Big\|\sum_{i,j=1}^{d}\frac{\partial(\mu\epsilon_{ij}(\widetilde{\boldsymbol{U}}-\boldsymbol{u}))}{\partial x_j}\Big\|_h \le Ch^{r-1}\|\boldsymbol{u}\|_r. \tag{4.15}$$

The third and fourth terms on the right-hand side of (4.12) are dealt with in the same manner. We first note that the stress tensor σ_{ij} contains two terms, the strains and the pressure. For the pressure we again use the projection $\bar{P}$ as defined above and get

$$|||[(\widetilde{P}-p)]|||_{h,I} \le C(|||[\widetilde{P}-\bar{P}]|||_{h,I}+|||[\bar{P}-p]|||_{h,I}).$$

Since $\bar{P}$ is defined as a local $L^2(\Omega)$-projection on each triangle or tetrahedron $\tau\in\mathcal{T}$,

$$
\begin{aligned}
h\int_{\partial\tau}(\bar{P}(s)-p(s))^2\,ds &\le C\bigg(\int_\tau(\bar{P}(x)-p(x))^2\,dx \\
&\qquad + h^2\int_\tau|\nabla(\bar{P}(x)-p(x))|^2\,dx\bigg) \\
&\le Ch^{2r-2}\|p\|^2_{H^{r-1}(\tau)}.
\end{aligned}
$$

Summing the above inequalities over all edges (faces) in $\mathcal{E}$ gives

$$|||[\bar{P} - p]|||_{h,I} \leq Ch^{r-1}||p||_{r-1}.$$

In addition, since $\widetilde{P} - \bar{P}$ is a polynomial on τ, standard reference element mapping arguments imply that

$$h \int_{\partial\tau} (\widetilde{P}(s) - \bar{P}(s))^2 \, ds \leq C \int_{\tau} (\widetilde{P}(x) - \bar{P}(x))^2 \, dx$$

and hence

$$|||[\widetilde{P} - \bar{P}]|||_{h,I} \leq Ch^{r-1}||p||_{r-1}.$$

Combining the above inequalities shows that

$$|||[\widetilde{P} - p]|||_{h,I} \leq Ch^{r-1}||p||_{r-1} \text{ and } ||\widetilde{P} - p||_{h,\Gamma_N} \leq Ch^{r-1}||p||_{r-1}. \tag{4.16}$$

Similar arguments can also be applied to estimate the part of the error related to the strains ϵ_{ij}, which are essentially the jumps in the derivatives of $\widetilde{\boldsymbol{U}} - \boldsymbol{u}$ across the inter-element boundaries.

For the last term on the right-hand side of (4.12) we apply the estimates (4.10) and (4.11) to get

$$||\nabla \cdot (\widetilde{\boldsymbol{U}} - \boldsymbol{u}) + \gamma(\widetilde{P} - p)|| \leq Ch^{r-1}\big(||\boldsymbol{u}||_r + ||p||_{r-1}\big). \tag{4.17}$$

Combining (4.12)–(4.17) gives

$$||\widetilde{\boldsymbol{E}}||_1 + ||\tilde{e}|| \leq Ch^{r-1}\big(||\boldsymbol{u}||_r + ||p||_{r-1}\big).$$

The theorem follows from (4.10), (4.11) and the triangle inequality.

Remark 8. *The theorem still holds in the case of locally quasi-uniform meshes. Its proof is similar to that given but requires replacing the approximation inequalities (4.10) and (4.11) by inequalities which are valid locally. However, the error estimate is no better.*

Remark 9. *The regularity of the solution of the 2-D Stokes equations with Dirichlet boundary conditions has been studied by Kellogg and Osborn in* [33]. *In particular, for convex domains with Lipschitz boundaries it is proven there that the solution $\boldsymbol{u} \in (H^2(\Omega))^d \cap \mathbf{H}_0^1(\Omega)$ when $\boldsymbol{F} \in (L^2(\Omega))^d$ and $\mathbf{f} = 0$ and therefore the least-squares method converges with a rate of at least $O(h)$.*

5. Implementation and the iterative solution of the least-squares system

In this section we consider the implementation aspects of the least-squares methods described in the preceding two sections. As already noted, the matrices corresponding to the algebraic systems resulting from the least-squares forms described in the previous section are full. Nevertheless, we shall see that effective preconditioned iterative schemes can be developed which converge rapidly and avoid assembly of the full matrix.

Let Π_h and $\mathbf{V}_h$ consist, respectively, of discontinuous piecewise constant functions and continuous piecewise linear functions. The implementation of higher order spaces is completely analogous.

There are three major aspects involved in setting up the algebraic system and its subsequent solution by a preconditioned iteration. All of these operations are performed with respect to a computational basis. Let n_1 and n_2 respectively denote the dimension of $\mathbf{V}_h$ and Π_h and set $n = n_1 + n_2$. Let $\{\Theta_i\}$ and $\{\theta_i\}$ denote the local nodal bases for $\mathbf{V}_h$ and Π_h, respectively. These can be combined into a basis $\{\Psi_j\} = \{(\Theta_i, 0)\} \cup \{(0, \theta_i)\} = \{(\Phi_j, \phi_j)\}$ for $\mathbf{V}_h \times \Pi_h$.

Let $(\boldsymbol{U}, P)$ be the solution of (4.5) and $\tilde{c}$ be the corresponding vector of nodal values, i.e., $(\boldsymbol{U}, P) = \sum_{i=1}^n \tilde{c}_i \Psi_j$. Then,

$$\widetilde{M}\tilde{c} = \tilde{d} \qquad \text{where} \qquad \widetilde{M}_{ij} = \langle\!\langle \Psi_i, \Psi_j \rangle\!\rangle_1. \tag{5.1}$$

The right-hand side of (5.1) is given by

$$\begin{aligned} \tilde{d}_i = {} & (\mathbf{B}_h \widetilde{\boldsymbol{F}}, \mathcal{L}_h(\Phi_i, \phi_i)) + (\boldsymbol{F}, L(\Phi_i, \phi_i))_h \\ & + < \mathbf{f}, \sigma_\nu(\Phi_i, \phi_i) >_{h,\Gamma_N}, \end{aligned} \tag{5.2}$$

for $i = 1, \ldots, n$.

In previous sections of this paper, we defined $\mathbf{B}_h$ as a symmetric positive definite operator on $\mathbf{V}_h$. In terms of the implementation, the preconditioner can be more naturally thought of in terms of a $n_1 \times n_1$ matrix N. The operator $\mathbf{B}_h$ is defined in terms of this matrix as follows. Fix $\boldsymbol{V} \in \mathbf{V}_h$ and expand

$$\mathbf{B}_h \boldsymbol{V} = \sum_i G_i \Theta_i.$$

Then,

$$NG = \widetilde{G} \qquad \text{where} \qquad \widetilde{G}_i = (\boldsymbol{V}, \Theta_i). \tag{5.3}$$

The operator $\mathbf{B}_h$ is a good preconditioner for $\mathbf{T}_h$ provided that the matrix $N^{-1}\widetilde{N}$ has small condition number. Here $\widetilde{N}$ is the stiffness matrix for the form $\mathbf{D}(\cdot, \cdot)$, i.e.,

$$\widetilde{N}_{ij} = \mathbf{D}(\Theta_i, \Theta_j).$$

The matrix N need not explicitly appear in the computation of the action of the preconditioner. Instead, one often has a process or algorithm which acts on the vector $\widetilde{G}$ and produces the vector G, i.e., computes $N^{-1}\widetilde{G}$. Thus, the practical application of the preconditioner on a function in $\boldsymbol{V}$ reduces to a predefined algorithm for computing the action of N^{-1} and the evaluation of the vector $\widetilde{G}$ defined by (5.3).

The first step in computing the coefficient vector $\tilde{c}$ solving (5.1) is to compute the right-hand side vector $\tilde{d}$. We shall assume that some method for computing integrals of the form

$$\int_\tau \boldsymbol{F} \cdot \eta \, dx \qquad \text{and} \quad \int_\epsilon \mathbf{f} \cdot \eta \, dx \tag{5.4}$$

is available when η is a vector valued polynomial. Here τ is a finite element in the mesh and ϵ is an edge (face) of a finite element. Thus, we can compute the data $[\widetilde{\boldsymbol{F}}, \Theta_j] = (\boldsymbol{F}, \Theta_j) + < \mathbf{f}, \Theta_j >_{\Gamma_N}$, for $j = 1, \ldots, n_1$ from which $\mathbf{B}_h \widetilde{\boldsymbol{F}}$ can be computed as discussed above. With $\mathbf{B}_h \widetilde{\boldsymbol{F}}$ known, the first term on the right-hand side of (5.2) reduces to $A(\Phi_i, \phi_i; \mathbf{B}_h \widetilde{\boldsymbol{F}})$. This involves the integration of polynomials over the triangles $\tau \in \mathcal{T}$. The remaining two terms in (5.2) reduce to more integrals of the form of (5.4) and the integration of polynomials over the triangles $\tau \in \mathcal{T}$ and their edges. These actions are local in the sense that the result for each Φ_j, Θ_j and ϕ_j only involves the triangles containing the support of respective function. The number of operations (work) involved is of the order of n.

The next action required for the implementation of the preconditioned iteration is the application of $\widetilde{M}$ to arbitrary vectors $c \in R^n$. The vector c represents the coefficients of a function pair $(\boldsymbol{V}, \delta)$; i.e.,

$$(\boldsymbol{V}, \delta) = \sum_{i=1}^{n} c_i(\Phi_i, \phi_i).$$

We are required to evaluate

$$\begin{aligned}(\widetilde{M}c)_j &= \langle\!\langle(\boldsymbol{V}, \delta), (\Phi_j, \phi_j)\rangle\!\rangle_1 \\ &= (\mathbf{B}_h\mathcal{L}_h(\boldsymbol{V}, \delta), \mathcal{L}_h(\Phi_j, \phi_j)) + (L(\boldsymbol{V}, \delta), L(\Phi_j, \phi_j))_h \\ &\quad + < \sigma_\nu(\boldsymbol{V}, \delta), \sigma_\nu(\Phi_j, \phi_j) >_{h,\Gamma_N} + < [\sigma_\nu(\boldsymbol{V}, \delta)], [\sigma_\nu(\Phi_j, \phi_j)] >_{h,I} \\ &\quad + (\nabla \boldsymbol{V} + \gamma\delta, \nabla\Phi_j + \gamma\phi_j),\end{aligned} \tag{5.5}$$

for $j = 1, \ldots, n$. The data for the preconditioner evaluation is

$$(\mathcal{L}_h(\boldsymbol{V}, \delta), \Theta_i) = A_0(\boldsymbol{V}, \Theta_i) - (\delta, \nabla\Theta_i)$$

and reduces to integrals of polynomials over $\tau \in \mathcal{T}$. After application of the preconditioner, the coefficients for the function $\mathbf{B}_h\mathcal{L}_h(\boldsymbol{V}, \delta)$ are known. All quantities appearing on the right-hand side of (5.5) can then be computed by integrals of polynomials over the triangles and their edges. The work required for computing $(MG)_j$, $j = 1, \ldots, n$ is of the order of n plus the work involved in applying the preconditioning process.

The final step required for a preconditioned iteration is the action of an appropriate preconditioning matrix M. By Remark 7, there exist positive constants C_0 and C_1, not depending on h, satisfying

$$C_0(||\boldsymbol{V}||_1 + ||P||) \le \langle\!\langle(\boldsymbol{V}, Q), (\boldsymbol{V}, Q)\rangle\!\rangle_1 \le C_1(||\boldsymbol{V}||_1 + ||P||). \tag{5.6}$$

The above inequalities hold for all $(\boldsymbol{V}, Q)$ in the product space $\mathbf{V}_h \times \Pi_h$. Consequently the task of defining a preconditioner for $\widetilde{M}$ is the same as finding a preconditioner for the block diagonal system

$$\begin{pmatrix} \widetilde{N} & 0 \\ 0 & \widetilde{N}_0 \end{pmatrix}$$

where $\widetilde{N}_0$ is the Gram matrix

$$(\widetilde{N}_0)_{ij} = (\phi_i, \phi_j) \text{ for } i, j = 1, \ldots, n_2.$$

Define

$$M = \begin{pmatrix} N & 0 \\ 0 & D \end{pmatrix}$$

where D is the diagonal matrix with entries $D_{ii} = (\phi_i, \phi_i)$. It follows from (5.6) that the condition number of $M^{-1}\widetilde{M}$ is independent of h. Thus, the reduction rate per step in, for example, the preconditioned conjugate gradient iteration can be bounded independently of h. The application of M^{-1} involves multiplying the Π_h data by D^{-1} and the application of the preconditioning process to the $\mathbf{V}_h$ data. The work involved in one step of the conjugate gradient iteration is on the order of n plus twice the cost of the application of the preconditioning process N^{-1}.

The above discussion is summarized in the following algorithm for computing the solution of (4.5).

Algorithm: The solution of (4.5) involves the following two steps.

- **(1):** The computation of the right-hand side vector $\tilde{d}$ of (5.1).
  - **(a):** Compute $\{(\boldsymbol{F}, \Theta_j) + < \mathbf{f}, \Theta_j >_{\Gamma_N}\}$ by assembling the quantities given in (5.4) (Work $\sim \mathrm{O}(n_1)$).
  - **(b):** Solve the preconditioning problem (5.3) with data computed from $\widetilde{G} = \{(\boldsymbol{F}, \Theta_j) + < \mathbf{f}, \Theta_j >_{\Gamma_N}\}$. This gives the coefficients for $\mathbf{B}_h \widetilde{\boldsymbol{F}}$.
  - **(c):** Compute $\tilde{d}$. This involves additional integrals of the form (5.4) and integrals of polynomials on the triangles and edges (Work $\sim \mathrm{O}(n_1)$).
- **(2):** Compute $\tilde{c}$ solving (5.1) by preconditioned conjugate gradient iteration. The entries of $\tilde{c}$ are the coefficients for the solution of (4.5). Each iterative step requires the evaluation of the matrix operator and preconditioner.
  - **(a):** Evaluation of the matrix operator on a given vector $\{c_i\}$ corresponding to the function pair $(\boldsymbol{V}, \delta)$ involves the following steps.
    - **(i):** Compute the data $\widetilde{G} = \{A_0(\boldsymbol{V}, \Theta_i) - (\delta, \nabla \Theta_i)\}$ for the $\mathbf{B}_h$ evaluation (Work $\sim \mathrm{O}(n)$).
    - **(ii):** Apply the preconditioning process to obtain the coefficients for $\mathbf{B}_h \mathcal{L}_h(\boldsymbol{V}, \delta)$.
    - **(iii):** Compute the quantities $(Mc)_j$ $j = 1, \ldots, n$ given in (5.5) (Work $\sim \mathrm{O}(n)$).
  - **(b):** Evaluation of the block preconditioner on a given vector $\{\bar{c}_j,\ j = 1, \ldots, n\}$. This involves multiplying the last n_2 coefficients by D^{-1} (Work $\sim \mathrm{O}(n_2)$) and evaluating the action of the preconditioning process.

Remark 10. *As already noted, the appearance of $\mathbf{B}_h$ gives rise to a full upper left hand block in the stiffness matrix for the least-squares operator. Consequently, it is not feasible to assemble the matrix. However, some efficiency may be gained by assembling parts of the matrix. For example, it would be feasible to assemble a sparse matrix for all terms in (5.5) excluding the one involving $\mathbf{B}_h$. Additionally to compute more efficiently the first term of (5.5) one could assemble the matrices $\{A_0(\Theta_j, \Theta_i)\}$ and $\{(\nabla_h \phi_j, \Theta_i)\}$.*

References

[1] I. Babuška, On the Schwarz algorithm in the theory of differential equations of mathematical physics, Tchecosl. Math. J., **8** (1958), 328-342 (Russian).

[2] P.E. Bjorstad and O.B. Widlund, Solving elliptic problems on regions partitioned into substructures, in: G. Birkhoff and A. Schoenstadt, Eds., *Elliptic Problem Solvers II*, Acad. Press, New York, 1984, 245-256.

[3] P.B. Bochev and M.D. Gunzburger, Analysis of least-squares finite element methods for Stokes equations, Math. Comp., **63** (1994), 479-506.

[4] P.B. Bochev and M.D. Gunzburger, Finite element methods of least-squares type, SIAM Review, **40** (1998), 789–837.

[5] J.H. Bramble, *Multigrid Methods*, Pitman research Notes in Mathematics Series, (Longman Scientific & Technical, London, Copublished with Wiley, New York, 1993).

[6] J.H. Bramble, Interpolation between Sobolev spaces in Lipschitz domains with an application to multigrid theory, Math. Comp., **64** (1995), 1359-1366.

[7] 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 systems, Math. Comp., **66** (1997), 935-955.

[8] J.H. Bramble, R.D. Lazarov and J.E. Pasciak, Least-squares for second order elliptic problems, Comput. Meth. Appl. Mech. Eng., **152** (1998) 195-210.

[9] J.H. Bramble and J.E. Pasciak, A preconditioning technique for indefinite systems resulting from mixed approximations of elliptic problems, Math. Comp., **50** (1988) 1-17.

[10] J.H. Bramble and J.E. Pasciak, New convergence estimates for multigrid algorithms, Math. Comp., **49** (1987) 311-329.

[11] J.H. Bramble and J.E. Pasciak, Least-squares methods for Stokes equations based on a discrete minus one inner product, J. Comput. Appl. Math., **74** (1996), 155-173.

[12] J.H. Bramble and J.E. Pasciak, New estimates for multigrid algorithms including V-cycle, Math. Comp., **60** (1993) 447-471.

[13] J.H. Bramble, J.E. Pasciak, J. Wang and J. Xu, Convergence estimates for product iterative methods with applications to domain decomposition, Math. Comp., **57** (1991) 1-21.

[14] S.C. Brenner and L.Y. Sung, Linear finite elements methods for planar linear elasticity, Math. Comp., **59** (1992) 321-338.

[15] F. Brezzi, On the existence, uniqueness, and approximation of saddle-point problems arising from Lagrange multipliers, R.A.I.R.O., **8** (1974), 479-506.

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

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

[18] 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 Numer. Anal., **35** (1998) 320-335.

[19] P. Ciarlet, Basic error estimates for elliptic problems, in *Finite Element Methods: Handbook of Numer. Analysis*, P. Ciarlet and J. Lions, Eds, v. II, North Holland, New York, 1991, pp. 18-352.

[20] T.F. Chan, R. Glowinski, J. Periaux, O.B. Widlund, Eds., *Third Int. Symposium on Domain Decomposition Methods for Partial Differential Equations*, SIAM, Philadelphia, PA, 1990.

[21] T.F. Chan, R. Glowinski, J. Periaux, O.B. Widlund, Eds., *Domain Decomposition Methods*, SIAM, Philadelphia, PA, 1989.

[22] J. Douglas and J. Wang, An absolutely stabilized finite element methods for the Stokes problem, Math. Comp., **52** (1989), 495-508.

[23] G. Duvaut and J.L. Lions, *Inequalities in Mechanics and Physics*, Series of Comprehensive Studies in Math., 219, Springer, 1976.

[24] R. Falk, A finite element method for the stationary Stokes equations using trial functions which do not satisfy $div v = 0$, Math. Comp, **30** (1976) 698-702.

[25] L.P. Franca and R. Stenberg, Error analysis of some Galerkin least-squares methods for the elasticity equations, SIAM J. Numer. Anal., **28** (1991), 1680-1697.
[26] V. Girault and P.-A. Raviart, *Finite Element Methods for Navier-Stokes equations*, Springer Series in Computational Mathematics, 5, Springer, 1986.
[27] R. Glowinski, Yu.A. Kuznetsov, G.A. Meurant and J. Periaux, *Fourth Int. Symposium on Domain Decomposition Methods for Partial Differential Equations*, SIAM, Philadelphia, PA, 1991.
[28] J. Gobert, Une inégalité fondamentale de la théorie de l'élasticité, Bull. Soc. Royale Science Liège, **31** année, No 3-4 (1962) 182-191.
[29] P. Grisvard, *Elliptic Problems in Nonsmooth Domains*, Pitman, Boston, 1985.
[30] T.J.R. Hughes and L.P. Franca, A new finite element formulation for computational fluid dynamics. VII. The Stokes problems with various well posed boundary conditions: symmetric formulation that converges for all velocity pressure spaces, Comput. Meth. Appl. Mech. Eng., **65** (1987) 85-96.
[31] T.J.R. Hughes and L.P. Franca, A new finite element formulation for computational fluid dynamics. V. Circumventing the Babuška-Brezzi condition: a stable Petrov-Galerkin formulation of the Stokes problem accommodating equal-order interpolations, Comput. Meth. Appl. Mech. Eng., **59** (1986) 85-99.
[32] B.N. Jiang and C.L. Chang, Least-squares finite elements for Stokes equations, Comput. Meth. Appl. Mech. Engrg., **78** (1990), 297-311.
[33] R.B. Kellog and J.E. Osborn, A regularity result for the Stokes problem in a convex polygon, J. Funct. Anal., **21** (1976) 397-431.
[34] S.-D. Kim, T. Manteuffel, and S. McCormick, First order system least-squares (FOSLS) for spatial linear elasticity, Preprint, 1999.
[35] O.A. Ladyzhenskaya, *The Mathematical Theory of Viscous Incompressible Flows*, Gordon and Breach, London, 1969.
[36] J.L. Lions and E. Magenes, *Non-homogeneous boundary value problems and applications*, Springer, Berlin-Heidelberg-New York, v. 181 and 182 (1972) and v. 183 (1973).
[37] J. Nečas, *Les Méthodes Directes en Théorie des Équations Elliptiques*, Masson, 1967.
[38] T. Rusten and R. Winter, A preconditioned iterative methods for saddle point problems, SIAM J. Matrix Anal. Appl., **13** (1992), 489-512.
[39] R. Temam, *Navier-Stokes Equations*, North-Holland, New York, 1977.
[40] P.S. Vassilevski, Hybrid V-cycle algebraic multilevel method for second order elliptic problems, Preprint, 1987.