ANALYSIS OF LEAST-SQUARES FINITE ELEMENT METHODS FOR THE NAVIER-STOKES EQUATIONS *

PAVEL B. BOCHEV †

Abstract. In this paper we study finite element methods of least-squares type for the stationary, incompressible Navier-Stokes equations in 2 and 3 dimensions. We consider methods based on velocity-vorticity-pressure form of the Navier-Stokes equations augmented with several nonstandard boundary conditions. Least-squares minimization principles for these boundary value problems are developed with the aid of Agmon-Douglis-Nirenberg elliptic theory. Among the main results of this paper are optimal error estimates for conforming finite element approximations, and analysis of some nonstandard boundary conditions. Results of several computational experiments with least-squares methods which illustrate, among other things, the optimal convergence rates are also reported.

Key words. Navier-Stokes equations, least-squares principle, finite element methods, velocity-vorticity-pressure equations.

AMS subject classifications. 76D05, 76D07, 65F10, 65F30

1. Introduction. In the past few years finite element methods based on least-squares variational principles have drawn considerable attention from mathematicians and engineers. In particular, owing to a number of valuable theoretical and computational properties, there has been a substantial interest in the use of such principles in the context of the Stokes and Navier-Stokes equations. One approach has been to use terms of least-squares type for stabilization of standard mixed Galerkin methods; see, e.g., [3], and [20]-[21]. In fact, least-squares ideas can be found, either explicitly, or implicitly, in most of the known stabilization techniques; see, e.g., [26] for comprehensive summary of such techniques. A second approach, which is the subject of this paper, departs from mixed formulations and uses least-squares principles directly in the derivation of weak problems. This approach leads to *bona fide* least-squares finite element methods, see [5], [9], [29]-[33], [35] and [37], among others. A critical ingredient of this approach, which is largely responsible for the success of resulting finite element methods, is transformation of original boundary value problems into equivalent first-order systems, and formulation of least-squares variational principles in terms of these systems. Typically, a least-squares principle involves minimization of quadratic functional defined by summing up norms of residuals of the first-order system. Corresponding minimizers are sought out of a suitable functional space, and are subject to the Euler-Lagrange equations. The latter essentially represent an alternative weak formulation of the original boundary value problem. Because weak formulations are now associated with minimization problems, resulting least-squares finite element methods offer significant computational and theoretical advantages. Most notably, such methods circumvent the inf-sup (LBB) condition of Ladyzhenskaya-Babuska-Brezzi, see [24]. As a result, one has greater freedom in the choice of discretization spaces which lead to stable methods, including the possibility to use equal order interpolation for all unknowns. Likewise, application of least-squares principles in the context of the Navier-Stokes equations leads to discrete problems with symmetric and positive

*This work was supported by the Air Force Office of Scientific Research under grants number AFOSR-93-1-0061 and number AFOSR-93-1-0280

†Department of Mathematics, University of Texas at Arlington, Box 19408, Arlington, TX 76019-0408. E-mail: bochev@utamat.uta.edu. Part of this work has been performed while the author was with the Interdisciplinary Center for Applied Mathematics at Virginia Tech, Blacksburg, VA 24061-0531

definite Jacobian matrices, at least in a neighborhood of the solution. Influence of the Reynolds number on the positive definiteness of Jacobians is felt only through the size of this neighborhood. Thus, combined with properly implemented continuation with respect to the Reynolds number, one can devise algorithms for numerical solution of the Navier-Stokes equations which will encounter only symmetric and positive definite linear systems in the solution process.

Computational results reported by Jiang [29], Jiang et. al. [31], [33], [35], Lefebvre et. al. [37], Bochev [9], and Bochev and Gunzburger [6], among others, indicate that finite element methods for the Navier-Stokes equations, formulated along these lines, have great promise. At the same time, theoretical analysis of such methods has received very limited attention, especially when compared with the analyses available for the Stokes problem, (see [2], [7]-[9], [15], [30]); and for linear elliptic systems in general (see [10], [13]-[14], [16], [17], [19], [34], [39], [43]). Therefore, the aim of this paper is to develop theoretically least-squares approach for the stationary, incompressible Navier-Stokes equations, written as a first-order system involving velocity, vorticity and pressure as dependent variables. Our analysis draws upon several mathematical techniques among which central roles are played by the elliptic regularity theory of Agmon, Douglis and Nirenberg (ADN) [1], and the abstract approximation theory for branches of nonsingular solutions developed by Brezzi, Rappaz and Raviart (BRR) [12]. A distinctive feature of BRR theory is that it allows us to address existence, uniqueness and error estimates for least-squares finite element approximations of the Navier-Stokes equations, using results established in the context of the linear Stokes equations. Validity of such results depends largely on the existence of a priori estimates for boundary value problems involving a first-order *velocity-vorticity-pressure* Stokes operator. To establish the relevant a priori estimates here we shall use the *complementing condition* of [1] which is both necessary and sufficient for such estimates to hold. Compared with direct approaches (see [13]-[15], [32]), the use of complementing condition significantly simplifies analyses of corresponding boundary value problems, and allows us to study systematically a large number of nonstandard boundary operators. Besides being of theoretical interest, such boundary operators also come up in applications like electromagnetic field problems and decomposition of vector fields. Among the available mathematical literature on this subject is the work of Bendali et. al. [4], where nonstandard boundary conditions are considered in the context of vector potential formulations for the Stokes and Navier-Stokes equations, and the work of Girault [23] which analyzes methods for the Navier-Stokes equations with "no slip" type conditions. More recently, Bramble and Lee [11] have considered Stokes equations with "no penetration" type conditions arising from electromagnetic field applications, and in [32] Jiang et. al. studied several nonstandard boundary operators associated with the velocity-vorticity-pressure form of the Navier-Stokes equations. For a further discussion of nonstandard conditions we refer the reader to the comprehensive reviews by Gresho [25] and Gunzburger et. al. [27].

This paper is organized as follows. The velocity-vorticity-pressure Navier-Stokes equations in two and three space dimensions are introduced in Section 2. We show that corresponding first-order Stokes operators admit two different principal parts which, depending on the particular boundary operator, result in two different types of a priori estimates. Least-squares functionals, corresponding minimization problems and conforming finite element methods are defined in Section 3. Discretization error estimates for least-squares finite element approximations, which are the central result of this paper, are presented in Section 4. Section 5 is devoted to a discussion of implementation issues such as Newton's method and continuation techniques. Finally, in Section 6 we present computational study of least-squares methods. Our results illustrate, among other things, the optimal error estimates of Section 4.

1.1. Notation. We let Ω be an open bounded domain in $\mathbf{R}^3$ or $\mathbf{R}^2$ with a sufficiently smooth boundary Γ. The smoothness of Γ will be addressed in detail in Section 2. $\mathcal{D}(\Omega)$ will denote the space of smooth functions with compact support in Ω, and $\mathcal{D}(\bar{\Omega})$ will denote restriction of $\mathcal{D}(\mathbf{R}^n)$ on $\bar{\Omega}$. For $s \geq 0$ we use standard notation and definition for the Sobolev spaces $H^s(\Omega)$ and $H^s(\Gamma)$ with inner products and norms denoted by $(\cdot,\cdot)_{s,\Omega}$, $(\cdot,\cdot)_{s,\Gamma}$, $\|\cdot\|_{s,\Omega}$, and $\|\cdot\|_{s,\Gamma}$, respectively. When there is no ambiguity the measures Ω and Γ will be omitted from inner product and norm designation. As usual, $H^s_0(\Omega)$ will denote the closure of $\mathcal{D}(\Omega)$ with respect to the norm $\|\cdot\|_{s,\Omega}$ and $L^2_0(\Omega)$ will denote the subspace of square integrable functions with zero mean. We set $\tilde{\mathcal{D}}(\Omega) = \mathcal{D}(\Omega) \cap L^2_0(\Omega)$, $\tilde{\mathcal{D}}(\bar{\Omega}) = \mathcal{D}(\bar{\Omega}) \cap L^2_0(\Omega)$ and $\tilde{H}^s(\Omega) = H^s(\Omega) \cap L^2_0(\Omega)$. For negative values of s spaces $H^s(\Omega)$, $H^s_0(\Omega)$ and $\tilde{H}^s(\Omega)$ are defined as closures of $\mathcal{D}(\bar{\Omega})$, $\mathcal{D}(\Omega)$ and $\tilde{\mathcal{D}}(\bar{\Omega})$ with respect to the norm

$$\|\phi\|_s = \sup_{q \in D(\Omega)} \frac{\int_\Omega \phi q \, dx}{\|q\|_{-s}}\,; \tag{1.1}$$

where $D(\Omega) = \mathcal{D}(\bar{\Omega}), \mathcal{D}(\Omega)$ and $\tilde{\mathcal{D}}(\bar{\Omega})$ respectively. We identify $H^s(\Omega)$, $H^s_0(\Omega)$ and $\tilde{H}^s(\Omega)$ with the duals of $H^{-s}(\Omega)$, $H^{-s}_0(\Omega)$ and $\tilde{H}^{-s}(\Omega)$ respectively. By $(\cdot,\cdot)_{\mathbf{X}}$ and $\|\cdot\|_{\mathbf{X}}$ we denote inner products and norms, respectively, on product spaces $\mathbf{X} = H^{s_1}(\Omega) \times \cdots \times H^{s_n}(\Omega)$; when all s_i are equal we shall simply write $(\cdot,\cdot)_{s,\Omega}$ and $\|\cdot\|_{s,\Omega}$. Vectors will be denoted by bold face letters and C will denote a generic positive constant. We use $L(\mathbf{X},\mathbf{Y})$ for the set of all bounded linear operators $\mathbf{X} \mapsto \mathbf{Y}$.

We recall that in $\mathbf{R}^3$ the curl operator is defined by

$$\mathbf{curl}\, \mathbf{u} = \nabla \times \mathbf{u}\,, \tag{1.2}$$

and that in $\mathbf{R}^2$ there are two curl operators given by

$$\mathbf{curl}\, \phi = \begin{pmatrix} \phi_y \\ -\phi_x \end{pmatrix} \qquad \text{and} \qquad \mathrm{curl}\, \mathbf{u} = u_{2x} - u_{1y}\,,$$

respectively. Let us define two "vector" products $\phi \times \mathbf{u}$ and $\mathbf{v} \times \mathbf{u}$, where ϕ is scalar function, and $\mathbf{u}$, $\mathbf{v}$ are vectors in $\mathbf{R}^2$, by embedding ϕ, $\mathbf{u}$ and $\mathbf{v}$ into three-dimensional vectors $(0,0,\phi)$, $(u_1,u_2,0)$ and $(v_1,v_2,0)$, respectively. Then, $\mathbf{curl}\, \phi = \nabla \times \phi$ and $\mathrm{curl}\, \mathbf{u} = \nabla \times \mathbf{u}$. To avoid multiplicity of notation we agree to use **curl** in both cases and to denote the result as a vector. We recall that *vorticity* of a vector field $\mathbf{u}$ is defined by $\boldsymbol{\omega} = \mathbf{curl}\, \mathbf{u}$. Thus, depending on the space dimension, $\boldsymbol{\omega}$ is a scalar or a vector function. Finally, we recall the vector identities

$$\mathbf{curl}\, \mathbf{curl}\, \mathbf{u} = -\triangle \mathbf{u} + \mathrm{grad}\, \mathrm{div}\, \mathbf{u} \tag{1.3}$$

$$\mathbf{u} \cdot \mathrm{grad}\, \mathbf{u} = \frac{1}{2}\, \mathrm{grad}\, |\mathbf{u}|^2 - \mathbf{u} \times \mathbf{curl}\, \mathbf{u}\,, \tag{1.4}$$

which are valid for all sufficiently smooth vector functions $\mathbf{u}$ in $\mathbf{R}^2$ and $\mathbf{R}^3$.

2. The velocity-vorticity-pressure equations. The dimensionless equations governing the steady incompressible flow of a viscous fluid may be written in the form

$$-\nu\triangle\mathbf{u} + \mathbf{u}\cdot\operatorname{grad}\mathbf{u} + \operatorname{grad} p = \mathbf{f} \quad \text{in } \Omega \tag{2.1}$$

$$\operatorname{div}\mathbf{u} = 0 \quad \text{in } \Omega \tag{2.2}$$

where $\mathbf{u}$, p and $\mathbf{f} \in [L^2(\Omega)]^n$, $n = 2, 3$; denote velocity, pressure, and given body force, and ν is the inverse of the Reynolds number. Along with the system (2.1)-(2.2) we shall consider homogeneous boundary conditions of the form

$$\mathcal{R}(\mathbf{u}, p) = 0 \quad \text{on } \Gamma\,. \tag{2.3}$$

Following Jiang et. al. [30]-[35] we recast equations (2.1)-(2.2) into first-order system involving vorticity, velocity and pressure as dependent variables. Using identities (1.3)-(1.4) and in view of incompressibility constraint (2.2), momentum equation (2.1) can be written as $\nu\mathbf{curl}\,\boldsymbol{\omega} + \boldsymbol{\omega}\times\mathbf{u} + \operatorname{grad} r = \mathbf{f}$, where $r = p + 1/2|\mathbf{u}|^2$ denotes the total head (referred to as "pressure" in the sequel). Then we scale the new momentum equation by $Re = 1/\nu$ (this will be necessary for the subsequent application of some results in [12]). For simplicity, the scaled pressure (total head) and body force will be denoted again by r and $\mathbf{f}$, respectively. Thus, we consider the following first-order *velocity-vorticity-pressure* form of the steady state, incompressible Navier-Stokes equations

$$\mathbf{curl}\,\boldsymbol{\omega} + \frac{1}{\nu}\boldsymbol{\omega}\times\mathbf{u} + \operatorname{grad} r = \mathbf{f} \quad \text{in } \Omega \tag{2.4}$$

$$\mathbf{curl}\,\mathbf{u} - \boldsymbol{\omega} = 0 \quad \text{in } \Omega \tag{2.5}$$

$$\operatorname{div}\mathbf{u} = 0 \quad \text{in } \Omega\,. \tag{2.6}$$

It is important to note that the boundary operator (2.3) may not be suitable for the new system.

We recall that the abstract framework of [12] allows us to study methods for the nonlinear Navier-Stokes equations using results established in the context of the Stokes equations. As a result, for the purposes of our analysis it will be necessary to investigate well-posedness of boundary value problems involving the Stokes operator in velocity-vorticity-pressure form. Thus we continue with brief outline of ADN elliptic theory which will be used for this purpose.

2.1. Agmon-Douglis-Nirenberg elliptic theory. Below we review the elements of ADN elliptic theory [1] with a particular attention to conditions which guarantee existence of a priori estimates. Let us first establish the notation. For our purposes it suffices to consider only the constant coefficient case. Let $\mathcal{L} = \{\mathcal{L}_{ij}\}$, $i, j = 1, \dots, N$ denote a linear differential operator and let $\mathcal{R} = \{\mathcal{R}_{lj}\}$, $l = 1, \dots, m$, $j = 1, \dots, N$ denote a boundary operator. We consider boundary value problems of the form

$$\mathcal{L}(U) = F \quad \text{in } \Omega \tag{2.7}$$

$$\mathcal{R}(U) = G \quad \text{on } \Gamma\,. \tag{2.8}$$

We assign a system of integer indices $\{s_i\}$, $s_i \le 0$, for the equations and $\{t_j\}$, $t_j \ge 0$, for the unknown functions such that the order of $\mathcal{L}_{ij}$ is bounded by $s_i + t_j$. The principal part $\mathcal{L}^p$ of $\mathcal{L}$ is defined as all terms $\mathcal{L}_{ij}$ with orders exactly equal to $s_i + t_j$.

The principal part $\mathcal{R}^p$ is defined in a similar way by assigning nonpositive weights r_l to each row in $\mathcal{R}$ such that the order of $\mathcal{R}$ is bounded by $r_l + t_j$. We shall say that $\mathcal{L}$ is *elliptic* of total order $2m$ if there exists a set of indices t_j and s_i, and a positive integer m, such that $\deg(\det \mathcal{L}^p(\boldsymbol{\xi})) = 2m$ and $\det \mathcal{L}^p(\boldsymbol{\xi}) \neq 0$ for all real $\boldsymbol{\xi} \neq 0$. We shall say that $\mathcal{L}$ is *uniformly elliptic* if there exists a constant C_e, such that

$$C_e^{-1}|\boldsymbol{\xi}|^{2m} \leq |\det \mathcal{L}^p(\boldsymbol{\xi})| \leq C_e|\boldsymbol{\xi}|^{2m} \tag{2.9}$$

Let us now state conditions on $\mathcal{L}$ and $\mathcal{R}$ that are necessary for well-posedness of the boundary value problem (2.7)-(2.8). We assume that $\mathcal{L}$ is elliptic of total order $2m$. The first condition is to require that the number of rows in $\mathcal{R}$ equals m. Second, we require that the following *supplementary condition* is satisfied [1].

Supplementary Condition on $\mathcal{L}$. *(1.)* $\det \mathcal{L}^p(\boldsymbol{\xi})$ *is of even degree* $2m$ (*with respect to* $\boldsymbol{\xi}$). *(2.) For every pair of linearly independent real vectors* $\boldsymbol{\xi}$, $\boldsymbol{\xi}'$, *the polynomial* $\det \mathcal{L}^p(\boldsymbol{\xi} + \tau\boldsymbol{\xi}')$ *in the complex variable* τ *has exactly* m *roots with positive imaginary part.*

For any elliptic system in three or more dimensions, the supplementary condition holds [1], however in two-dimensions it must be verified for any given $\mathcal{L}^p$. The final, third condition is the *complementing condition* which is both necessary and sufficient for coercivity type estimates to hold. It is a local, algebraic condition on the principal parts $\mathcal{L}^p$ and $\mathcal{R}^p$ which guarantees that $\mathcal{R}$ is compatible with $\mathcal{L}$. Let $\tau_k^+(\boldsymbol{\xi})$ denote the m roots of $\det \mathcal{L}^p(\boldsymbol{\xi} + \tau\boldsymbol{\xi}')$ having positive imaginary part. Let

$$M^+(\boldsymbol{\xi}, \tau) = \prod_{k=1}^{m} \left(\tau - \tau_k^+(\boldsymbol{\xi})\right),$$

and let $\mathcal{L}'$ denote the adjoint matrix to $\mathcal{L}^p$. Then we have the following definition [1].

Complementing Condition. *For any point* $P \in \Gamma$ *let* $\mathbf{n}$ *denote the unit outward normal vector to the boundary* Γ *at the point* P. *Then, for any real, non-zero vector* $\boldsymbol{\xi}$ *tangent to* Γ *at* P, *regard* $M^+(\boldsymbol{\xi}, \tau)$ *and the elements of the matrix*

$$\sum_{j=1}^{N} \mathcal{R}_{lj}^p(\boldsymbol{\xi} + \tau\mathbf{n})\mathcal{L}'_{jk}(\boldsymbol{\xi} + \tau\mathbf{n})$$

as polynomials in τ. *The operators* $\mathcal{L}$ *and* $\mathcal{R}$ *satisfy the complementing condition if the rows of the latter matrix are linearly independent modulo* $M^+(\boldsymbol{\xi}, \tau)$, *that is,*

$$\sum_{l=1}^{m} C_l \sum_{j=1}^{N} \mathcal{R}_{lj}^p \mathcal{L}'_{jk} \equiv 0 \pmod{M^+} \tag{2.10}$$

if and only if the constants C_l *all vanish.*

In [1], the following result is proved.

Theorem 2.1. *Let* $\mathcal{L}$ *be uniformly elliptic operator of order* $2m$ *which in 2D satisfies the Supplementary Condition, and let* $\mathcal{R}$ *be a boundary operator which satisfies the complementing condition on* Γ. *Assume that for some* $q \geq 0$, *the boundary* Γ *of the domain* Ω *is of class* C^{r+t}, *where* $t = \max_j\{t_j\}$ *and* $r = \max\{\max_l\{r_l + 1\}, q\}$. *Furthermore, assume that* $U \in \prod_{j=1}^{N} H^{q+t_j}(\Omega)$, $F \in \prod_{i=1}^{N} H^{q-s_i}(\Omega)$, *and*

$G \in \prod_{l=1}^{m} H^{q-r_l-1/2}(\Gamma)$. *Then, there exists a constant* $C > 0$ *such that*

$$(2.11)\quad \sum_{j=1}^{N} \|u_j\|_{q+t_j,\Omega} \le C \left(\sum_{i=1}^{N} \|F_i\|_{q-s_i,\Omega} + \sum_{l=1}^{m} \|G_l\|_{q-r_l-1/2,\Gamma} + \sum_{j=1}^{N} \|u_j\|_{0,\Omega} \right) .$$

Moreover, if the problem $\mathcal{L}(U) = F$, $\mathcal{R}(U) = G$ *has a unique solution in the indicated spaces, then the* L^2*-norm on the right-hand side of* (2.11) *can be omitted.*

A notable feature of ADN theory is that the indices s_i and t_j for which problem (2.7)-(2.8) is well-posed are not necessarily unique. As a result, a given differential operator $\mathcal{L}$ may possess several uniformly elliptic principal parts, moreover, the number of these principal parts may be different for different forms of the operator. For example, the principal part of the Stokes operator in both primitive variables and velocity-pressure-stress forms is unique, see [8]. At the same time, the velocity-vorticity-pressure form of this operator admits *two different principal parts.* Moreover, for some boundary operators considered in the sequel, the critical complementing condition holds with only one of these principal parts.

2.2. Boundary conditions for the velocity-vorticity-pressure equations. In this section we use the complementing condition to examine admissibility of candidate boundary operators for the velocity-vorticity-pressure Stokes equations

$$(2.12)\qquad \mathbf{curl}\,\boldsymbol{\omega} + \operatorname{grad} r = \mathbf{f} \quad \text{in } \Omega$$

$$(2.13)\qquad \mathbf{curl}\,\mathbf{u} - \boldsymbol{\omega} = 0 \quad \text{in } \Omega$$

$$(2.14)\qquad \operatorname{div} \mathbf{u} = 0 \quad \text{in } \Omega\,.$$

First we shall discuss how the space dimension affects the ellipticity and the total order of this linear system. We assume that unknowns are ordered as $U = (\boldsymbol{\omega}, r, \mathbf{u})$.

It is not difficult to see that in two-dimensions the choices

$$(2.15)\qquad t_1 = \ldots = t_4 = 1; \quad s_1 = \ldots = s_4 = 0$$

$$(2.16)\qquad t_1 = t_2 = 1; \;\; t_3 = t_4 = 2; \quad s_1 = s_2 = 0; \;\; s_3 = s_4 = -1\,,$$

result in uniformly elliptic principal parts for (2.12)-(2.14). Indeed, the principal parts of (2.12)-(2.14) for the indices (2.15) and (2.16) are given by

$$(2.17)\qquad \mathcal{L}_1^p = \begin{pmatrix} \mathbf{curl}\,\boldsymbol{\omega} & + & \operatorname{grad} r \\ & & \mathbf{curl}\,\mathbf{u} \\ & & \operatorname{div} \mathbf{u} \end{pmatrix}$$

and

$$(2.18)\qquad \mathcal{L}_2^p = \begin{pmatrix} \mathbf{curl}\,\boldsymbol{\omega} & + & \operatorname{grad} r \\ -\boldsymbol{\omega} & + & \mathbf{curl}\,\mathbf{u} \\ & & \operatorname{div} \mathbf{u} \end{pmatrix} .$$

respectively. In both cases $\det \mathcal{L}_i^p(\boldsymbol{\xi}) = -(\xi_1^2 + \xi_2^2)^2 = -|\boldsymbol{\xi}|^4$; $i = 1, 2$; that is, the uniform ellipticity condition (2.9) holds with $m = 2$ and $C_e = 1$. The supplementary condition also holds, see [7]. As a result, $\mathcal{R}$ should provide $l = 2$ conditions on Γ (the total order of (2.12)-(2.14) is four). Since the Stokes operator in primitive variables has the same total order, an operator of the form (2.3) formally satisfies the first compatibility condition of Section 2.1.

Remark. We recall that for first-order elliptic systems in the plane Lopatinski condition [43] plays similar role to that of the complementing condition. The latter however is more general in the following sense. If $\mathcal{R}$ satisfies Lopatinski condition for (2.12)-(2.14) then the complementing condition holds for this boundary operator with the indices (2.15). In this case we shall say that $\mathcal{R}$ satisfies *equal differentiability assumption*. However, a boundary operator which satisfies the complementing condition with the indices (2.16), but not with (2.15), will fail Lopatinski condition. In such a case we shall say that $\mathcal{R}$ satisfies *different differentiability assumption*. A typical example of such operator is velocity boundary condition (see [7] for the details).

It is easy to see that in three-dimensions system (2.12)-(2.14) has seven equations and unknowns, and is not elliptic in the sense of ADN. Following Chang [16]-[17] we augment equations (2.12)-(2.14) with the seemingly redundant (in view of (2.5)) equation

$$\operatorname{div} \boldsymbol{\omega} = 0 \quad \text{in } \Omega \tag{2.19}$$

and introduce a new "slack" variable ϕ in (2.13):

$$\mathbf{curl}\,\mathbf{u} - \boldsymbol{\omega} + \operatorname{grad}\phi = 0 \quad \text{in } \Omega\,.$$

We assume that the eight unknowns are ordered as $U = (\boldsymbol{\omega}, r, \phi, \mathbf{u})$ and that the eight equations are ordered as (2.12), (2.19), (2.13), (2.14). Then the two sets of indices analogous to (2.15) and (2.16) are

$$t_1 = \ldots = t_8 = 1; \quad s_1 = \ldots = s_8 = 0 \tag{2.20}$$

and

$$t_1 = \ldots = t_4 = 1;\ t_5 = \ldots = t_8 = 2; \quad s_1 = \ldots = s_4 = 0;\ s_5 = \ldots = s_8 = -1\,, \tag{2.21}$$

respectively. The principal parts are now given by

$$\mathcal{L}_1^p = \left(\begin{array}{ccc} \mathbf{curl}\,\boldsymbol{\omega} & + & \operatorname{grad} r \\ & & \operatorname{div}\boldsymbol{\omega} \\ \mathbf{curl}\,\mathbf{u} & + & \operatorname{grad}\phi \\ & & \operatorname{div}\mathbf{u} \end{array} \right) \tag{2.22}$$

and

$$\mathcal{L}_2^p = \left(\begin{array}{ccc} \mathbf{curl}\,\boldsymbol{\omega} & + & \operatorname{grad} r \\ & & \operatorname{div}\boldsymbol{\omega} \\ -\boldsymbol{\omega} + \mathbf{curl}\,\mathbf{u} & + & \operatorname{grad}\phi \\ & & \operatorname{div}\mathbf{u} \end{array} \right), \tag{2.23}$$

respectively. A short computation shows that $\det \mathcal{L}_i^p(\boldsymbol{\xi}) = -|\boldsymbol{\xi}|^8;\ i = 1,2$. Thus, the augmented system (2.12), (2.19), (2.13), (2.14) is uniformly elliptic of total order eight. In contrast, total order of the three-dimensional Stokes operator in primitive variables is only six. As a result, in three-dimensions a boundary operator for (2.12), (2.19), (2.13) and (2.14) should specify $l = 4$ conditions on Γ, that is (2.3) fails the first compatibility condition of Section 2.1. Sometimes this problem can be resolved

TABLE 2.1
Classification of boundary conditions for the Navier-Stokes equations

Boundary conditions		$\mathbf{R}^3$	$\mathbf{R}^2$	Type
BC1	Velocity	$\mathbf{u}$	$\mathbf{u}$	2
	Slack variable	ϕ	-	
BC1A	Velocity	$\mathbf{u}$	$\mathbf{u}$	2
	Normal vorticity	$\boldsymbol{\omega}\cdot\mathbf{n}$	-	
BC2	Normal velocity	$\mathbf{u}\cdot\mathbf{n}$	$\mathbf{u}\cdot\mathbf{n}$	
	Normal vorticity	$\boldsymbol{\omega}\cdot\mathbf{n}$	-	1
	Pressure	r	r	
	Slack variable	ϕ	-	
BC2A	Normal velocity	$\mathbf{u}\cdot\mathbf{n}$	$\mathbf{u}\cdot\mathbf{n}$	
	Tangential vorticity	$\mathbf{n}\times\boldsymbol{\omega}\times\mathbf{n}$	$\boldsymbol{\omega}$	1
	Slack variable	ϕ	-	
BC2B	Normal velocity	$\mathbf{u}\cdot\mathbf{n}$	$\mathbf{u}\cdot\mathbf{n}$	not well-posed
	Tangential vorticity	$\mathbf{n}\times\boldsymbol{\omega}\times\mathbf{n}$	$\boldsymbol{\omega}$	(r is redundant in $\mathbf{R}^2$
	Pressure	r	r	
BC2C	Normal velocity	$\mathbf{u}\cdot\mathbf{n}$	$\mathbf{u}\cdot\mathbf{n}$	not well-posed in $\mathbf{R}^3$
	Vorticity	$\boldsymbol{\omega}$	$\boldsymbol{\omega}$	1 in $\mathbf{R}^2$
BC3	Tangential velocity	$\mathbf{n}\times\mathbf{u}\times\mathbf{n}$	$\mathbf{u}\cdot\mathbf{t}$	2 in $\mathbf{R}^3$
	Pressure	r	r	1 in $\mathbf{R}^2$
	Slack variable	ϕ	-	
BC3A	Tangential velocity	$\mathbf{n}\times\mathbf{u}\times\mathbf{n}$	$\mathbf{u}\cdot\mathbf{t}$	
	Normal vorticity	$\boldsymbol{\omega}\cdot\mathbf{n}$	-	1
	Pressure	r	r	
BC3B	Tangential velocity	$\mathbf{n}\times\mathbf{u}\times\mathbf{n}$	$\mathbf{u}\cdot\mathbf{t}$	not
	Normal vorticity	$\boldsymbol{\omega}\cdot\mathbf{n}$	-	well-posed
	Slack variable	ϕ	-	
BC3C	Tangential velocity	$\mathbf{n}\times\mathbf{u}\times\mathbf{n}$	$\mathbf{u}\cdot\mathbf{t}$	1
	Tangential vorticity	$\mathbf{n}\times\boldsymbol{\omega}\times\mathbf{n}$	$\boldsymbol{\omega}$	
BC4	Vorticity	$\boldsymbol{\omega}$	$\boldsymbol{\omega}$	not
	Pressure	r	r	well-posed
BC4A	Vorticity	$\boldsymbol{\omega}$	$\boldsymbol{\omega}$	not
	Slack variable	ϕ	-	well-posed
BC5	Tangential vorticity	$\mathbf{n}\times\boldsymbol{\omega}\times\mathbf{n}$	$\boldsymbol{\omega}$	not
	Pressure	r	r	well-posed
	Slack variable	ϕ	-	

by augmenting (2.3) with a fourth condition derived from the already specified data. For example, if $\mathbf{u} = \mathbf{U}$ is given on Γ then $\boldsymbol{\omega}\cdot\mathbf{n} = \mathbf{n}\cdot\mathbf{curl}\,\mathbf{U}$ may be added to (2.3) in view of the fact that $\mathbf{n}\cdot\mathbf{curl}\,\mathbf{U}$ involves only tangential derivatives of $\mathbf{U}$. However, a larger class of boundary conditions will result if, instead, we consider boundary operators $\mathcal{R}$ of the form

$$\mathcal{R}(\boldsymbol{\omega}, r, \phi, \mathbf{u}) = 0 \quad \text{on } \Gamma\,. \tag{2.24}$$

Let us now discuss various choices for $\mathcal{R}$ in (2.24). We shall say that the boundary operator $\mathcal{R}$ is of *type one* if $\mathcal{R}$ satisfies equal differentiability assumption, that is, when complementing condition holds with the indices (2.15) or (2.20). We shall say that $\mathcal{R}$ is of *type two* when $\mathcal{R}$ satisfies different differentiability assumption. This classification is used in Table 2.2 where a list of boundary conditions for the velocity-vorticity-pressure equations is presented. We state boundary operators in both three and two-dimensions. Three-dimensional operators are specialized to $\mathbf{R}^2$ using two-dimensional "vector" products defined in Section 1.1. Thus, in $\mathbf{R}^2$ the term $\boldsymbol{\omega}\cdot\mathbf{n}$ does not give condition and $\mathbf{n}\times\boldsymbol{\omega}\times\mathbf{n}$ is replaced by $\boldsymbol{\omega}$. Admissibility and type of each boundary condition are established as follows. For a given boundary operator we first check whether equal differentiability assumption holds, that is, we try to verify complementing condition with the principal parts (2.17) or (2.22). If complementing condition fails then we try to verify it with the principal parts (2.18) or (2.23), that is, we check whether different differentiability assumption is valid for $\mathcal{R}$. If complementing condition fails again, then $\mathcal{R}$ is deemed inadmissible. Verification of complementing condition requires elementary but tedious and lengthy algebraic manipulations and for this reason the details are not presented here. For more detailed examples of this verification process the reader can consult [1], [7]-[9] and [40].

One conclusion which can be drawn from Table 1 is that dimensionality can change the type of the boundary operator. For example, the type of BC3 changes from one in $\mathbf{R}^2$ to two in $\mathbf{R}^3$. Another observation is that a two-dimensional boundary condition can have two different three-dimensional counterparts. Such pairs are BC3-BC3A, and BC2A-BC2C.

The type of the boundary operator is also important for the smoothness of Γ required in Theorem 2.1. Recall that Γ must be of class C^{t+r} where t and r were defined in Theorem 2.1. It is not difficult to see that for type one operators $t = 1$, for type two operators $t = 2$, and that for all operators in Table 1 $\max_l\{r_l + 1\} = 1 - \min_j\{t_j\} = 0$. As a result, Γ must be of class C^{1+q} for type one, and of class C^{2+q} for type two operators, respectively.

Let us conclude this section with a lemma which will allow us to ignore the slack variable in all further developments. The proof of this lemma is standard, and is omitted.

LEMMA 2.2. *Assume that:*

1. *The operator (2.24) is one of BC1 or BC2A and* $r \in L^2_0(\Omega)$*;*
2. *The operator (2.24) is one of BC3 or BC3A and* $\phi \in L^2_0(\Omega)$ *in three-dimensions;*
3. *The operator (2.24) is one of BC1A or BC3C and* $r \in L^2_0(\Omega)$*, and* $\phi \in L^2_0(\Omega)$ *in three dimensions;*
4. *The operator (2.24) is BC2.*

Then the boundary value problem (2.12)-(2.14), (and (2.19) in $\mathbf{R}^3$*), (2.24) has at most one solution.*

If $\mathbf{u}$ and r solve the Stokes problem in primitive variables then $(\boldsymbol{\omega} = \mathbf{curl}\,\mathbf{u}, r, \mathbf{u})$ and $(\boldsymbol{\omega} = \mathbf{curl}\,\mathbf{u}, r, \phi = 0, \mathbf{u})$ solve (2.12)-(2.14) in two and three-dimensions, respectively. By virtue of Lemma 2.2 it follows that the slack variable ϕ is identically zero and, therefore, it can be completely ignored. In fact, we can carry out the analyses including the slack variable and then specialize all results for $\phi = 0$. Thus, in what follows we will not make use of this variable. However, we stress upon the fact that in three-dimensions, the "redundant" equation (2.19) is essential for the ellipticity of the Stokes problem and cannot be ignored.

2.3. A priori estimates. In this section we specialize results of Theorem 2.1 for the *generalized* Stokes problem

$$
\begin{aligned}
\mathbf{curl}\,\boldsymbol{\omega} + \operatorname{grad} r &= \mathbf{f}_1 \quad \text{in } \Omega \\
\mathbf{curl}\,\mathbf{u} - \boldsymbol{\omega} &= \mathbf{f}_2 \quad \text{in } \Omega \\
\operatorname{div} \mathbf{u} &= f_3 \quad \text{in } \Omega\,,
\end{aligned}
$$

augmented with $\operatorname{div} \boldsymbol{\omega} = f_4$ in three dimensions. The functions f_3 and f_4 are subject to the solvability conditions $\int_\Omega f_3\, d\Omega = \int_\Gamma \mathbf{u}\cdot\mathbf{n}\, d\Gamma$ and $\int_\Omega f_4\, d\Omega = \int_\Gamma \boldsymbol{\omega}\cdot\mathbf{n}\, d\Gamma$, whenever the boundary operator $\mathcal{R}$ prescribes $\mathbf{u}\cdot\mathbf{n}$ or $\boldsymbol{\omega}\cdot\mathbf{n}$.

In the following σ will denote parameter which depends on the type of the boundary condition:

$$\sigma = \begin{cases} 0 & \text{for type 1 boundary operator} \\ 1 & \text{for type 2 boundary operator} \end{cases} .$$

We will also assume that the boundary Γ of the domain Ω is of class $C^{1+q+\sigma}$. In addition, $\bar{H}^{q+1}(\Omega)$ will denote the space $H^{q+1}(\Omega)$ whenever the operator (2.24) prescribes the pressure r on Γ, and $\tilde{H}^{q+1}(\Omega)$ otherwise. Let $q \geq 0$ and let

$$\mathbf{X}_{\sigma,q} = \{(\boldsymbol{\omega}, r, \mathbf{u}) \in H^{1+q}(\Omega) \times \bar{H}^{1+q}(\Omega) \times [H^{1+q+\sigma}(\Omega)]^2 \,|\, \mathcal{R}(\boldsymbol{\omega}, r, \mathbf{u}) = 0\} \tag{2.25}$$

in two-dimensions and

$$\mathbf{X}_{\sigma,q} = \{(\boldsymbol{\omega}, r, \mathbf{u}) \in [H^{1+q}(\Omega)]^3 \times \bar{H}^{1+q}(\Omega) \times [H^{1+q+\sigma}(\Omega)]^3 \,|\, \mathcal{R}(\boldsymbol{\omega}, r, \mathbf{u}) = 0\} \tag{2.26}$$

in three-dimensions.

Owing to Lemma 2.2 the L^2-norm in (2.11) can be omitted, and since we consider only homogeneous boundary conditions all boundary terms in (2.11) will vanish. Then, letting f_i to correspond to differential equations (2.12)-(2.14) evaluated at $\boldsymbol{\omega}$, r and $\mathbf{u}$, the estimate (2.11) specializes to

$$\begin{aligned} &\|\mathbf{u}\|_{1+q+\sigma} + \|\boldsymbol{\omega}\|_{1+q} + \|r\|_{1+q} \\ &\leq C\,\left(\|\mathbf{curl}\,\boldsymbol{\omega} + \mathrm{grad}\, r\|_q + \|\mathbf{curl}\,\mathbf{u} - \omega\|_{q+\sigma} + \|\mathrm{div}\,\mathbf{u}\|_{q+\sigma}\right) \end{aligned} \tag{2.27}$$

in two-dimensions and to

$$\begin{aligned} &\|\mathbf{u}\|_{1+q+\sigma} + \|\boldsymbol{\omega}\|_{1+q} + \|r\|_{1+q} \\ &\leq C\,\left(\|\mathbf{curl}\,\boldsymbol{\omega} + \mathrm{grad}\, r\|_q + \|\mathrm{div}\,\boldsymbol{\omega}\|_q + \|\mathbf{curl}\,\mathbf{u} - \boldsymbol{\omega}\|_{q+\sigma} + \|\mathrm{div}\,\mathbf{u}\|_{q+\sigma}\right) \end{aligned} \tag{2.28}$$

in three-dimensions.

Because the complementing condition is *necessary and sufficient*, estimates (2.27) and (2.28) with $\sigma = 0$ are not valid for boundary operators of type 2. For an example, consider first the cube $\Omega = (-1,1)^3 \subset \mathbf{R}^3$ and let $\mathcal{R}$ correspond to the velocity boundary condition BC1A in Table 1. Let $q = 0$ and let: $\mathbf{u} = \mathbf{0}$; $r = \sin nx \cdot \mathrm{e}^{ny}(z^2 - 1)$, and $\boldsymbol{\omega} = (0, 0, -(\cos nx \cdot \mathrm{e}^{ny})(z^2 - 1))^T$. Then, $\mathbf{u} = \mathbf{0}$ and $\boldsymbol{\omega} \cdot \mathbf{n} = 0$ on Γ. Since $\|\boldsymbol{\omega}\|_1 \sim O(n\mathrm{e}^n)$ and $\|r\|_1 \sim O(n\mathrm{e}^n)$ it follows that

$$\|\mathbf{u}\|_1 + \|\boldsymbol{\omega}\|_1 + \|r\|_1 \sim O(n\mathrm{e}^n)\,.$$

On the other hand, we have that

$$\mathbf{curl}\,\boldsymbol{\omega} + \mathrm{grad}\, r = (0, 0, 2z \sin nx \cdot \mathrm{e}^{ny}))^T\,,$$

$$\mathrm{div}\,\boldsymbol{\omega} = -2z \cos nx \cdot \mathrm{e}^{ny}\,,$$

and as a result,

$$\|\mathbf{curl}\,\boldsymbol{\omega} + \mathrm{grad}\, r\|_0 + \|\mathbf{curl}\,\mathbf{u} - \boldsymbol{\omega}\|_0 + \|\mathrm{div}\,\mathbf{u}\|_0 + \|\mathrm{div}\,\boldsymbol{\omega}\|_0 \sim O(\mathrm{e}^n)\,.$$

Hence, (2.28) cannot hold for BC1A with $\sigma = 0$. This counterexample can be easily extended to smooth boundaries [38]. For a two-dimensional counterexample we refer the reader to [7].

3. Least-squares finite element methods. In this section we define conforming least-squares finite element methods for the velocity-vorticity-pressure Navier-Stokes equations augmented with homogeneous boundary conditions of the form (2.24). We consider the three-dimensional case with the tacit understanding that the slack variable is ignored from the equations and the boundary conditions. Analysis of the methods will use (2.28) with $q = 0$. Thus, in what follows we assume that the boundary Γ of the domain Ω is of class $C^{1+\sigma}$. All our results can be easily specialized in two-dimensions by dropping equation (2.19) and using appropriate curl and vector product definitions from Section 1.1.

3.1. The least-squares principle. Let $(\boldsymbol{\omega}_0, r_0, \mathbf{u}_0) \in \mathbf{X}_{\sigma,0}$ denote the unique solution of the Stokes problem (2.12)-(2.14), (2.19), and (2.24). We first replace (2.4) by the equation

$$\mathbf{curl}\,\boldsymbol{\omega} + \operatorname{grad} r + \frac{1}{\nu}(\boldsymbol{\omega} + \boldsymbol{\omega}_0) \times (\mathbf{u} + \mathbf{u}_0) = 0 \quad \text{in } \Omega\,, \tag{3.1}$$

in which $\mathbf{f}$ has been eliminated. Evidently, if ω_1, r_1 and $\mathbf{u}_1$ solve the problem (3.1), (2.5), (2.6), (2.19), and (2.24), then $\omega_1 + \omega_0$, $r_1 + r_0$ and $\mathbf{u}_1 + \mathbf{u}_0$ solve the original problem (2.4)-(2.6), (2.19), and (2.24). The reason to prefer (3.1) over (2.4) is purely for technical convenience in the analyses. Then we consider the following least-squares functional

$$\begin{aligned} J_\sigma(\boldsymbol{\omega}, r, \mathbf{u}) = \frac{1}{2}(&\|\mathbf{curl}\,\boldsymbol{\omega} + \operatorname{grad} r + \frac{1}{\nu}(\boldsymbol{\omega} + \boldsymbol{\omega}_0) \times (\mathbf{u} + \mathbf{u}_0)\|_0^2 \\ &+\|\operatorname{div}\boldsymbol{\omega}\|_0^2 + \|\mathbf{curl}\,\mathbf{u} - \boldsymbol{\omega}\|_\sigma^2 + \|\operatorname{div}\mathbf{u}\|_\sigma^2\ )\ . \end{aligned} \tag{3.2}$$

Residual of each equation in (3.2) appears in the norm of the Sobolev space $H^{-s_i}(\Omega)$, where s_i are the equation indices given by (2.20) or (2.21). We also note that scaling of equation (2.4) by Re, (see Section 2) can be viewed as weighting of its residual in (3.2) by Re^2. Similar weighting, although with a different motivation, appears in [15] for the Stokes problem.

Let $U = (\boldsymbol{\omega}, r, \mathbf{u})$ and $V = (\boldsymbol{\xi}, q, \mathbf{v})$. We consider minimization of $J_\sigma(U)$ over the space $\mathbf{X}_{\sigma,0}$. Thus, the least-squares principle for (3.2) is given by

$$\textit{seek } U \in \mathbf{X}_{\sigma,0} \textit{ such that} \quad J_\sigma(U) \le J_\sigma(V), \quad \textit{for all} \quad V \in \mathbf{X}_{\sigma,0}\,. \tag{3.3}$$

Minimizers of (3.2) are subject to the necessary condition (Euler-Lagrange equation)

$$\lim_{\varepsilon\to 0} \frac{d}{d\varepsilon} J_\sigma(U + \varepsilon V) = 0 \quad \text{for all} \quad V \in \mathbf{X}_{\sigma,0}\,, \tag{3.4}$$

which has the following variational form: *seek* $U \in \mathbf{X}_{\sigma,0}$ *such that*

$$\begin{aligned} &\mathcal{B}(U, V) = \\ &= (\mathbf{curl}\,\boldsymbol{\omega} + \operatorname{grad} r, \mathbf{curl}\,\boldsymbol{\xi} + \operatorname{grad} q)_0 + (\operatorname{div}\boldsymbol{\omega}, \operatorname{div}\boldsymbol{\xi})_0 \\ &+ (\mathbf{curl}\,\mathbf{u} - \boldsymbol{\omega}, \mathbf{curl}\,\mathbf{v} - \boldsymbol{\xi})_\sigma + (\operatorname{div}\mathbf{u}, \operatorname{div}\mathbf{v})_\sigma \\ &+ \Bigg(\frac{1}{\nu}(\boldsymbol{\omega} + \boldsymbol{\omega}_0) \times (\mathbf{u} + \mathbf{u}_0), \\ &\qquad\qquad \mathbf{curl}\,\boldsymbol{\xi} + \operatorname{grad} q + \frac{1}{\nu}(\boldsymbol{\xi} \times (\mathbf{u} + \mathbf{u}_0) + (\boldsymbol{\omega} + \boldsymbol{\omega}_0) \times \mathbf{v})\Bigg)_0 \\ &+ \frac{1}{\nu}\,(\mathbf{curl}\,\boldsymbol{\omega} + \operatorname{grad} r, \boldsymbol{\xi} \times (\mathbf{u} + \mathbf{u}_0) + (\boldsymbol{\omega} + \boldsymbol{\omega}_0) \times \mathbf{v})_0 = 0 \end{aligned} \tag{3.5}$$

for all $V \in \mathbf{X}_{\sigma,0}$. Let $\mathcal{L}$ denote the differential operator of velocity-vorticity-pressure Stokes problem, and let

$$\mathcal{M}(U,V) = \frac{1}{\nu}(\boldsymbol{\xi} \times (\mathbf{u}+\mathbf{u}_0) + (\boldsymbol{\omega}+\boldsymbol{\omega}_0) \times \mathbf{v})\,,$$

$$\mathcal{N}(U) = \mathbf{curl}\,\boldsymbol{\omega} + \operatorname{grad} r\,,$$

$$\mathcal{H}(U) = \frac{1}{\nu}(\boldsymbol{\omega}+\boldsymbol{\omega}_0) \times (\mathbf{u}+\mathbf{u}_0)\,.$$

Then (3.5) can be put into the more compact form

$$(\mathcal{L}U, \mathcal{L}V)_{\mathbf{X}_{\sigma,-1}} + (\mathcal{H}(U), \mathcal{N}(V))_0 + (\mathcal{H}(U)+\mathcal{N}(U), \mathcal{M}(U,V))_0 = 0 \quad \text{for all} \quad V \in \mathbf{X}_{\sigma,0}\,.$$

Note that in (3.5) linear terms depend on the type of boundary conditions but not on the parameter ν, whereas nonlinear terms depend only on ν.

3.2. Finite element methods. Starting with the weak formulation (3.5) a conforming finite element method can be defined in a completely standard manner. We choose a finite element space $\mathbf{X}^h_\sigma$, parameterized by h, and such that

$$\mathbf{X}^h_\sigma \subset \mathbf{X}_{\sigma,0}\,.$$

Furthermore, we assume that $\mathbf{X}^h_\sigma$ approximates optimally with respect to $\mathbf{X}_{\sigma,d}$ for some fixed $d \geq 1$ in the following sense: for every $U \in \mathbf{X}_{\sigma,d}$ there exists $U^h \in \mathbf{X}^h_\sigma$ such that for $r = -1, 0$

$$\|U - U^h\|_{\mathbf{X}_{\sigma,r}} \leq C\, h^{d-r}\, \|U\|_{\mathbf{X}_{\sigma,d}}\,. \tag{3.6}$$

Then, discrete analogue of problem (3.5) is given by: *seek* $U^h \in \mathbf{X}^h_\sigma$ *such that*

$$\begin{aligned}
\mathcal{B}(U^h, V^h) &= \\
&= (\mathbf{curl}\,\boldsymbol{\omega}^h + \operatorname{grad} r^h, \mathbf{curl}\,\boldsymbol{\xi}^h + \operatorname{grad} q^h)_0 + (\operatorname{div}\boldsymbol{\omega}^h, \operatorname{div}\boldsymbol{\xi}^h)_0 \\
&+ (\mathbf{curl}\,\mathbf{u}^h - \boldsymbol{\omega}^h, \mathbf{curl}\,\mathbf{v}^h - \boldsymbol{\xi}^h)_\sigma + (\operatorname{div}\mathbf{u}^h, \operatorname{div}\mathbf{v}^h)_\sigma \\
&+ \Bigg(\frac{1}{\nu}(\boldsymbol{\omega}^h + \boldsymbol{\omega}_0) \times (\mathbf{u}^h + \mathbf{u}_0), \\
&\qquad \mathbf{curl}\,\boldsymbol{\xi}^h + \operatorname{grad} r^h + \frac{1}{\nu}\left(\boldsymbol{\xi}^h \times (\mathbf{u}^h + \mathbf{u}_0) + (\boldsymbol{\omega}^h + \boldsymbol{\omega}_0) \times \mathbf{v}^h\right)\Bigg)_0 \\
&+ \left(\mathbf{curl}\,\boldsymbol{\omega}^h + \operatorname{grad} r^h, \frac{1}{\nu}\left(\boldsymbol{\xi}^h \times (\mathbf{u}^h + \mathbf{u}_0) + (\boldsymbol{\omega}^h + \boldsymbol{\omega}_0) \times \mathbf{v}^h\right)\right)_0 = 0\,.
\end{aligned} \tag{3.7}$$

for all $V^h \in \mathbf{X}^h_\sigma$. Problem (3.7) can be derived directly as necessary condition for the finite dimensional least-squares principle:

$$\textit{seek } U^h \in \mathbf{X}^h_\sigma \textit{ such that} \quad J_\sigma(U^h) \leq J_\sigma(V^h), \quad \textit{for all} \quad V^h \in \mathbf{X}^h_\sigma\,.$$

The finite element space $\mathbf{X}^h_\sigma$ can be constructed in the following manner. Let $\mathcal{T}_h$ denote triangulation of Ω into finite elements. Triangulation $\mathcal{T}_h$ is not necessarily uniform, however we assume that it is uniformly regular (see [18] or [24]). Next, we choose finite element spaces $S^h_1 \subset H^1(\Omega)$ and $S^h_2 \subset H^{1+\sigma}(\Omega)$, defined with respect to

$\mathcal{T}_h$, such that for every $u \in H^{1+d}(\Omega)$, and every $v \in H^{1+d+\sigma}(\Omega)$ there exist elements $u^h \in S_1^h$ and $v^h \in S_2^h$ with

$$\|u - u^h\|_r \le C\, h^{1+d-r}\, \|u\|_{1+d}\,, \quad r = 0, 1\,;$$

$$\|v - v^h\|_r \le C\, h^{1+d+\sigma-r}\, \|v\|_{1+d+\sigma}\,, \quad r = 0, .., 1+\sigma\,.$$

If the boundary operator is of type 1 then one can choose $S_2^h = S_1^h$. For particular examples of such finite element spaces we refer the reader to monographs [18], [24], and [26]. Lastly, let $\bar{S}_1^h = S_1^h$ if the boundary operator $\mathcal{R}$ prescribes r on Γ, and let $\bar{S}_1^h = S_1^h \cap L_0^2(\Omega)$ otherwise. Then we set

$$\mathbf{X}_\sigma^h = \{U^h \in [S_1^h]^3 \times \bar{S}_1^h \times [S_2^h]^3 \,|\, \mathcal{R}(U^h) = 0 \text{ on } \Gamma\}\,. \tag{3.8}$$

4. Error estimates. The goal of this section is derivation of error estimates. Our main result will be to establish that least-squares finite element approximations defined by (3.7) converge to all sufficiently smooth solutions of the Navier-Stokes equations at the best possible rate. As noted earlier, for this purpose here we shall use the abstract approximation theory of Brezzi, Rappaz and Raviart [12]. For the sake of completeness below we quote the relevant results of [12] specialized to our needs. In this we follow version of the abstract theory given in [24].

Let $\mathbf{X}$ and $\mathbf{Y}$ be Banach spaces and let $\Lambda \subset \mathbf{R}$ be a compact interval. We consider nonlinear problems of the form

$$F(\lambda, \phi) \equiv \phi + T \cdot G(\lambda, \phi) = 0\,, \tag{4.1}$$

where $T \in L(\mathbf{Y}, \mathbf{X})$ and G is a C^2 map $\Lambda \times \mathbf{X} \mapsto \mathbf{Y}$. The set $\{(\lambda, \phi(\lambda)) \,|\, \lambda \in \Lambda\}$ is called branch of solutions of (4.1) if $F(\lambda, \phi(\lambda)) = 0$ for $\lambda \in \Lambda$, and the map $\lambda \to \phi(\lambda)$ is continuous function from Λ into $\mathbf{X}$. If, in addition, Frèchet derivative $D_\phi F(\lambda, \phi(\lambda))$ of F with respect to ϕ is an isomorphism of $\mathbf{X}$ for all $\lambda \in \Lambda$, then the branch $\{(\lambda, \phi(\lambda)) \,|\, \lambda \in \Lambda\}$ is called regular.

Approximations for (4.1) are defined in the following manner. We introduce a finite dimensional subspace $\mathbf{X}^h \subset \mathbf{X}$, a linear operator $T^h \in L(\mathbf{Y}, \mathbf{X}^h)$, which presumably approximates T, and consider the discrete problem

$$F^h(\lambda, \phi^h) \equiv \phi^h + T^h \cdot G(\lambda, \phi^h) = 0\,. \tag{4.2}$$

Both T and T^h are assumed to be independent from λ. The error estimates for nonlinear approximations ϕ^h are derived under the following hypotheses. First, we assume that there exists a Banach space $\mathbf{Z}$ continuously imbedded in $\mathbf{Y}$, such that

$$D_\phi G(\lambda, \phi) \in L(\mathbf{X}, \mathbf{Z}) \quad \text{for all} \quad \lambda \in \Lambda \;\; \text{and} \;\; \phi \in \mathbf{X}\,. \tag{4.3}$$

Second, we assume that

$$\lim_{h \to 0} \|(T^h - T)g\|_{\mathbf{X}} = 0 \quad \text{for all} \quad g \in \mathbf{Y}\,, \tag{4.4}$$

and that

$$\lim_{h \to 0} \|T^h - T\|_{L(\mathbf{Z}, \mathbf{X})} = 0\,. \tag{4.5}$$

With these assumptions we may state the results that will be needed in the sequel.

THEOREM 4.1. *Let $\mathbf{X}$ and $\mathbf{Y}$ be Banach spaces and let Λ be a compact subset of $\mathbf{R}$. Assume that G is a C^2 mapping from $\Lambda \times \mathbf{X}$ into $\mathbf{Y}$ and that all second Frèchet derivatives of G are bounded on all bounded sets of $\Lambda \times \mathbf{X}$. Assume that (4.3)-(4.5) hold and that $\{(\lambda, \phi(\lambda)) \mid \lambda \in \Lambda\}$ is branch of regular solutions of (4.1). Then there exists a neighborhood $\mathcal{O}$ of the origin in $\mathbf{X}$, and for h sufficiently small, a unique C^2 function $\lambda \to \phi^h \in \mathbf{X}^h$, such that $\{(\lambda, \phi^h(\lambda)) \mid \lambda \in \Lambda\}$ is branch of regular solutions of (4.2) and $\phi^h(\lambda) - \phi(\lambda) \in \mathcal{O}$ for all $\lambda \in \Lambda$. Moreover, there exists a constant $C > 0$, independent of h and λ, such that*

$$\|\phi(\lambda) - \phi^h(\lambda)\|_{\mathbf{X}} \leq C\,\|(T - T^h) \cdot G(\lambda, \phi(\lambda))\|_{\mathbf{X}} \quad \forall \lambda \in \Lambda . \tag{4.6}$$

4.1. Abstract form of the least-squares method. Let us show that variational problems (3.5) and (3.7) can be cast into canonical forms (4.1) and (4.2), respectively. To this end in addition to the spaces $\mathbf{X}_{\sigma,q}$ we shall need the dual spaces

$$\mathbf{Y}_{\sigma,q} = \mathbf{X}^*_{\sigma,q} \tag{4.7}$$

and the space

$$\mathbf{Z}_\sigma = [L^{(3+\sigma)/2}(\Omega)]^3 \times L^{(3+\sigma)/2}(\Omega) \times [L^{3/2}(\Omega)]^3 . \tag{4.8}$$

LEMMA 4.2. *We have that $\mathbf{Z}_\sigma \subset \mathbf{Y}_{\sigma,0}$ with compact imbedding.*

Proof. For the proof of this lemma we use an argument communicated to us by M. Renardy. First we note that by virtue of Sobolev's imbedding theorem, each component of $\mathbf{X}_{\sigma,0}$ imbedds compactly into $L^q(\Omega)$; $2 \leq q \leq 6$. Then, since the adjoint of a compact operator is also compact, it follows that each component of $\mathbf{Z}_\sigma$ is compactly imbedded into the respective component of the dual space $\mathbf{Y}_{\sigma,0}$. $\square$

We make the association

$$\mathbf{X} = \mathbf{X}_{\sigma,0}\,,\ \mathbf{X}^h = \mathbf{X}^h_\sigma\,,\ \mathbf{Y} = \mathbf{Y}_{\sigma,0}\,, \mathbf{Z} = \mathbf{Z}_\sigma \quad \text{and} \quad \lambda = \frac{1}{\nu}\,.$$

Next, operator T is defined as follows:

$T : \mathbf{Y}_{\sigma,0} \mapsto \mathbf{X}_{\sigma,0}$ with $U = T \cdot \mathbf{g}$ for $\mathbf{g} \in \mathbf{Y}_{\sigma,0}$ if and only if

$$(\mathcal{L}U, \mathcal{L}V)_{\mathbf{X}_{\sigma,-1}} = (\mathbf{g}, V)_0 \quad \text{for all} \quad V \in \mathbf{X}_{\sigma,0}\,. \tag{4.9}$$

For the definition of T^h we consider conforming discretization of T:

$T^h : \mathbf{Y}_{\sigma,0} \mapsto \mathbf{X}^h_\sigma$ with $U^h = T^h \cdot \mathbf{g}$ for $\mathbf{g} \in \mathbf{Y}_{\sigma,0}$ if and only if

$$(\mathcal{L}U^h, \mathcal{L}V^h)_{\mathbf{X}_{\sigma,-1}} = (\mathbf{g}, V^h)_0 \quad \text{for all} \quad V^h \in \mathbf{X}^h_\sigma\,. \tag{4.10}$$

Thanks to the scaling of momentum equation (2.4) operators T and T^h are independent from the parameter λ. Finally, nonlinear operator G is defined as follows:

$G : \Lambda \times \mathbf{X} \to \mathbf{Y}$ with $\mathbf{g} = G(\lambda, U)$ for $U \in \mathbf{X}_{\sigma,0}$ if and only if

$$(\mathcal{H}(U), \mathcal{N}(V))_0 + (\mathcal{H}(U) + \mathcal{N}(U), \mathcal{M}(U,V))_0 = (\mathbf{g}, V)_0 \quad \text{for all} \quad V \in \mathbf{X}_{\sigma,0}\,. \tag{4.11}$$

Next Lemma verifies that (3.5) and (3.7) can be cast into the canonical forms (4.1) and (4.2).

LEMMA 4.3. *Assume that T, T^h and G are defined by (4.9), (4.10) and (4.11). Then (3.5) and (3.7) are equivalent to (4.1) and (4.2), respectively.*

Proof. Assume that $U \in \mathbf{X}_{\sigma,0}$ solves (4.1). Then $-U = T \cdot \mathbf{g}$ if and only if

$$-(\mathbf{g}, V)_0 = (\mathcal{L}U, \mathcal{L}V)_{\mathbf{X}_{\sigma,-1}} \quad \text{for all} \quad V \in \mathbf{X}_{\sigma,0}\,,$$

and $\mathbf{g} = G(\lambda, U)$ if and only if

$$(\mathbf{g}, V)_0 = (\mathcal{H}(U), \mathcal{N}(V))_0 + ((\mathcal{H}(U) + \mathcal{N}(U), \mathcal{M}(U,V))_0 \quad \text{for all} \quad V \in \mathbf{X}_{\sigma,0}\,.$$

Thus, $U + T \cdot G(\lambda, U) = 0$ if and only if $\mathcal{B}(U,V) = 0$ for all $V \in \mathbf{X}_{\sigma,0}$. The proof for (3.7) and (4.2) is identical. ∎

4.1.1. Properties of T and T^h. First we note that variational problem (4.9) is associated with a least-squares principle for the Stokes problem (2.12)-(2.14), (2.24). Indeed, if $(\mathbf{g}, V)_0 = (\mathbf{f}, \mathbf{curl}\,\xi + \operatorname{grad} q)_0$ for all $V \in \mathbf{X}_{\sigma,0}$, then problem (4.9) is exactly the Euler-Lagrange equation for the Stokes least-squares functional

$$\tilde{J}_\sigma(\boldsymbol{\omega}, r, \mathbf{u}) = \frac{1}{2}\left(\|\mathbf{curl}\,\boldsymbol{\omega} + \operatorname{grad} r - \mathbf{f}\|_0^2 + \|\operatorname{div}\boldsymbol{\omega}\|_0^2 + \|\mathbf{curl}\,\mathbf{u} - \boldsymbol{\omega}\|_\sigma^2 + \|\operatorname{div}\mathbf{u}\|_\sigma^2\right)\,.$$

Similarly, discrete problem (4.10) can be associated with a least-squares finite element method for the Stokes problem. In the next three lemmas we generalize some results of [7] concerning T and T^h, and establish several other facts that are needed for our analysis.

LEMMA 4.4. *The operators T and T^h given by (4.9) and (4.10) are well-defined linear operators in $L(\mathbf{Y}_{\sigma,0}, \mathbf{X}_{\sigma,0})$ and $L(\mathbf{Y}_{\sigma,0}, \mathbf{X}_\sigma^h)$, respectively.*

Proof. It is not difficult to see that the bilinear form $(\mathcal{L}U, \mathcal{L}V)_{\mathbf{X}_{\sigma,-1}}$ is continuous in $\mathbf{X}_{\sigma,0} \times \mathbf{X}_{\sigma,0}$:

$$(\mathcal{L}U, \mathcal{L}V)_{\mathbf{X}_{\sigma,-1}} \le C_1\left(\|\boldsymbol{\omega}\|_1 + \|r\|_1 + \|\mathbf{u}\|_{1+\sigma}\right)\left(\|\boldsymbol{\xi}\|_1 + \|q\|_1 + \|\mathbf{v}\|_{1+\sigma}\right)$$

Using (2.28) with $q = 0$ it also follows that this form is coercive on $\mathbf{X}_{\sigma,0} \times \mathbf{X}_{\sigma,0}$:

$$\begin{aligned} &C_2\left(\|\boldsymbol{\omega}\|_1^2 + \|r\|_1^2 + \|\mathbf{u}\|_{1+\sigma}^2\right) \\ &\le \|\mathbf{curl}\,\boldsymbol{\omega} + \operatorname{grad} r\|_0^2 + \|\operatorname{div}\boldsymbol{\omega}\|_0^2 + \|\mathbf{curl}\,\mathbf{u} - \boldsymbol{\omega}\|_\sigma^2 + \|\operatorname{div}\mathbf{u}\|_\sigma^2 \\ &= (\mathcal{L}U, \mathcal{L}U)_{\mathbf{X}_{\sigma,-1}}\,. \end{aligned}$$

Finally, for all $\mathbf{g} \in \mathbf{Y}_{\sigma,0}$ the product $(\mathbf{g}, V)_0$ defines continuous linear functional on $\mathbf{X}_{\sigma,0}$. Then, by virtue of Lax-Milgram Lemma problem (4.9) has unique solution $U \in \mathbf{X}_{\sigma,0}$, and

$$\|U\|_{\mathbf{X}_{\sigma,0}} = \|\omega\|_1 + \|r\|_1 + \|\mathbf{u}\|_{1+\sigma} \le C\left(\|g_1\|_{-1} + \|g_2\|_{-1} + \|g_3\|_{-(1+\sigma)}\right) = C\|\mathbf{g}\|_{\mathbf{Y}_{\sigma,0}}\,.$$

As a result, it follows that $T \in L(\mathbf{Y}_{\sigma,0}, \mathbf{X}_{\sigma,0})$. Thanks to the inclusion $\mathbf{X}_\sigma^h \subset \mathbf{X}_{\sigma,0}$, it also follows that $T^h \in L(\mathbf{Y}_{\sigma,0}, \mathbf{X}_\sigma^h)$. ∎

It remains to verify that T^h approximates T in the sense of (4.4) and (4.5).

LEMMA 4.5. *Assume that (3.6) holds for the finite element space $\mathbf{X}_{\sigma,0}$ with some $d \ge 1$. Then, for any $\mathbf{g} \in \mathbf{Y}_{\sigma,0}$*

$$\lim_{h\to 0}\|(T^h - T)\mathbf{g}\|_{\mathbf{X}_{\sigma,0}} = 0\,. \tag{4.12}$$

If $\mathbf{g}$ is such that $U = T \cdot \mathbf{g}$ belongs to $\mathbf{X}_{\sigma,q}$ for some $q \ge 1$ then

$$\|(T^h - T)\cdot\mathbf{g}\|_{\mathbf{X}_{\sigma,0}} \le Ch^{\tilde{d}}\left(\|\boldsymbol{\omega}\|_{1+\tilde{d}} + \|r\|_{1+\tilde{d}} + \|\mathbf{u}\|_{1+\tilde{d}+\sigma}\right)\,, \tag{4.13}$$

where $\tilde{d} = \min\{d, q\}$.

Proof. Let $\mathbf{g} \in \mathbf{Y}_{\sigma,0}$ and let $U^h = T^h \cdot \mathbf{g}$, $U = T \cdot \mathbf{g}$. We must show that

$$\lim_{h\to 0} \left(\|\boldsymbol{\omega}^h - \boldsymbol{\omega}\|_1 + \|r^h - r\|_1 + \|\mathbf{u}^h - \mathbf{u}\|_{1+\sigma} \right) = 0\,.$$

By virtue of Cèa's lemma [18]

$$\begin{aligned} &\|\boldsymbol{\omega}^h - \boldsymbol{\omega}\|_1 + \|r^h - r\|_1 + \|\mathbf{u}^h - \mathbf{u}\|_{1+\sigma} \\ &\leq \inf_{V^h\in \mathbf{X}^h_\sigma} C \left(\|\boldsymbol{\xi}^h - \boldsymbol{\omega}\|_1 + \|q^h - r\|_1 + \|\mathbf{v}^h - \mathbf{u}\|_{1+\sigma} \right) . \end{aligned} \tag{4.14}$$

Consider first the term $\|\boldsymbol{\omega}^h - \boldsymbol{\omega}\|_1$. Since $\mathcal{D}(\bar{\Omega})$ is dense in $H^1(\Omega)$, for any fixed $\varepsilon > 0$ there exists $\boldsymbol{\omega}_\varepsilon \in \mathcal{D}(\bar{\Omega})$ such that $\|\boldsymbol{\omega} - \boldsymbol{\omega}_\varepsilon\|_1 < \varepsilon$. Owing to (3.6) there exists function $\boldsymbol{\xi}^h \in S^h_1$ such that

$$\|\boldsymbol{\omega}_\varepsilon - \boldsymbol{\xi}^h\|_1 \leq C\, h^d\, \|\boldsymbol{\omega}_\varepsilon\|_{d+1}$$

with C independent from h. Therefore $\lim_{h\to 0} \|\boldsymbol{\omega} - \boldsymbol{\xi}^h\|_1 \leq \varepsilon$. Since ε was arbitrary number, it follows that $\lim_{h\to 0} \|\boldsymbol{\omega}^h - \boldsymbol{\omega}\|_1 = 0$. A similar argument is valid for the remaining terms, and (4.12) follows. If $(\boldsymbol{\omega}, r, \mathbf{u})$ has higher regularity determined by an index $q \geq 1$, then (4.14) in conjunction with (3.6) yields (4.13). $\square$

Our last result is to establish that T^h converges to T in $L(\mathbf{Z}_\sigma, \mathbf{X}_{\sigma,0})$.

LEMMA 4.6. *The following holds true for* T *and* T^h

$$\lim_{h\to 0} \|T - T^h\|_{L(\mathbf{Z}_\sigma, \mathbf{X}_{\sigma,0})} = 0\,. \tag{4.15}$$

Proof. Denote $S^h = T - T^h$. Then (4.12) and the Uniform Boundedness Theorem imply that S^h is uniformly bounded, that is, there exists $C > 0$, independent from h such that

$$\|S^h\|_{L(\mathbf{Y}_{\sigma,0}, \mathbf{X}_{\sigma,0})} \leq C\,.$$

Suppose that (4.15) does not hold. Then there exists $\varepsilon > 0$ such that for every $h = 1/n, n = 1, 2, 3, \ldots$ one can find $\mathbf{g}_n \in \mathbf{Z}_\sigma$ with $\|\mathbf{g}_n\|_{\mathbf{Z}_\sigma} = 1$ and $\|S^h \mathbf{g}_n\|_{\mathbf{X}_{\sigma,0}} \geq \varepsilon$. Since $\{\mathbf{g}_n\}$ is bounded in $\mathbf{Z}_\sigma$ and the latter is compactly imbedded in $\mathbf{Y}_{\sigma,0}$ it follows that $\mathbf{g}_n \to \mathbf{g}$ for some $\mathbf{g} \in \mathbf{Y}_{\sigma,0}$. Then, as $n \to \infty$

$$\|S^h \mathbf{g}_n\|_{\mathbf{X}_{\sigma,0}} \leq C \left(\|\mathbf{g}_n - \mathbf{g}\|_{\mathbf{Y}_{\sigma,0}} + \|S^h \mathbf{g}\|_{\mathbf{X}_{\sigma,0}} \right) \to 0\,,$$

a contradiction. $\square$

4.1.2. Properties of G. In this section we verify assumptions of Theorem 4.1 concerning the nonlinear operator (4.11). For this purpose we shall need some well-known inequalities and embedding results, which are quoted below for the convenience of the reader. We recall (see e.g. [42]) the inequalities

$$\left| \int_\Omega uvwz\, d\Omega \right| \leq C_1 \|u\|_{1,\Omega} \|v\|_{1,\Omega} \|w\|_{1,\Omega} \|z\|_{1,\Omega} \tag{4.16}$$

and

$$\left| \int_\Omega u \frac{\partial v}{\partial x_i} w\, d\Omega \right| \leq C_2 \|u\|_{1,\Omega} \|v\|_{1,\Omega} \|w\|_{1,\Omega}\,. \tag{4.17}$$

which hold for all functions u, v, w and z in $H^1(\Omega)$. We also recall that (see [24], Corollary 1.1)

$$\|uv\|_{0,3/2} \le C\|u\|_{0,2}\|v\|_{1,2} \quad \text{for all } u \in L^2(\Omega) \text{ and } v \in H^1(\Omega)\,; \tag{4.18}$$

$$\|uv\|_{0,2} \le C\|u\|_{0,2}\|v\|_{2,2} \quad \text{for all } u \in L^2(\Omega) \text{ and } v \in H^2(\Omega)\,; \tag{4.19}$$

$$\|uv\|_{0,2} \le C\|u\|_{1,2}\|v\|_{1,2} \quad \text{for all } u, v \in H^1(\Omega)\,; \tag{4.20}$$

$$\|uvw\|_{0,3/2} \le C\|u\|_{1,2}\|v\|_{1,2}\|w\|_{1,2} \quad \text{for all } u, v, w \in H^1(\Omega)\,, \tag{4.21}$$

that is, $(u,v) \mapsto uv$ is continuous bilinear mapping $L^2(\Omega)\times H^1(\Omega) \mapsto L^{3/2}(\Omega)$; $L^2(\Omega)\times H^2(\Omega) \mapsto L^2(\Omega)$ and $H^1(\Omega)\times H^1(\Omega) \mapsto L^2(\Omega)$, and that $(u,v,w) \mapsto uvw$ is continuous trilinear mapping $H^1(\Omega) \times H^1(\Omega) \times H^1(\Omega) \mapsto L^{3/2}(\Omega)$.

LEMMA 4.7. *Let Λ be a compact subset of $\mathbf{R}^+$, $\lambda = 1/\nu \in \Lambda$ and let G be defined by (4.11). Then*

1. *G is a C^2 mapping from $\Lambda \times \mathbf{X}_{\sigma,0}$ into $\mathbf{Y}_{\sigma,0}$;*
2. *$D_U G(\lambda, U) \in L(\mathbf{X}_{\sigma,0}, \mathbf{Z}_\sigma)$ for all $U \in \mathbf{X}_{\sigma,0}$;*
3. *all second derivatives of G are bounded on bounded subsets of $\Lambda \times \mathbf{X}$.*

Proof.

1. $G(\lambda, U)$ is polynomial map in λ and the components of U. Thus it can be shown that G is in fact a C^∞ mapping $\Lambda \times \mathbf{X}_{\sigma,0} \mapsto \mathbf{Y}_{\sigma,0}$.

2. From definition (4.11) of G it is not difficult to see that for given U and $\hat{U}$ in $\mathbf{X}_{\sigma,0}$ we have $D_U G(\lambda, U)[\hat{U}] = \mathbf{g}$, if and only if

$$\begin{aligned}
(\mathbf{g}, V)_0 &= (g_1, \boldsymbol{\xi})_0 + (g_2, q)_0 + (g_3, \mathbf{v})_0 \\
&= \lambda \left(\mathbf{curl}\,\boldsymbol{\omega} + \operatorname{grad} r + \lambda(\boldsymbol{\omega} + \boldsymbol{\omega}_0) \times (\mathbf{u} + \mathbf{u}_0), \boldsymbol{\xi} \times \hat{\mathbf{u}} + \hat{\boldsymbol{\omega}} \times \mathbf{v}\right)_0 \\
&+ \lambda \big(\mathbf{curl}\,\hat{\boldsymbol{\omega}} + \operatorname{grad} \hat{r} + \lambda(\hat{\boldsymbol{\omega}} \times (\mathbf{u} + \mathbf{u}_0) + (\boldsymbol{\omega} + \boldsymbol{\omega}_0) \times \hat{\mathbf{u}}), \\
&\qquad\quad \boldsymbol{\xi} \times (\mathbf{u} + \mathbf{u}_0) + (\boldsymbol{\omega} + \boldsymbol{\omega}_0) \times \mathbf{v}\big)_0 \\
&+ \lambda \left(\hat{\boldsymbol{\omega}} \times (\mathbf{u} + \mathbf{u}_0) + (\boldsymbol{\omega} + \boldsymbol{\omega}_0) \times \hat{\mathbf{u}}, \mathbf{curl}\,\boldsymbol{\xi} + \operatorname{grad} q\right)_0
\end{aligned} \tag{4.22}$$

for all $V \in \mathbf{X}_{\sigma,0}$. Through an examination of (4.22) we further see that

$$\begin{aligned}
g_1 = &-\lambda \left(\mathbf{curl}\,\boldsymbol{\omega} + \operatorname{grad} r + \lambda(\boldsymbol{\omega} + \boldsymbol{\omega}_0) \times (\mathbf{u} + \mathbf{u}_0)\right) \times \hat{\mathbf{u}} \\
&+ \lambda \mathbf{curl}\, \left(\hat{\boldsymbol{\omega}} \times (\mathbf{u} + \mathbf{u}_0) + (\boldsymbol{\omega} + \boldsymbol{\omega}_0) \times \hat{\mathbf{u}}\right) \\
&- \lambda \left(\mathbf{curl}\,\hat{\boldsymbol{\omega}} + \operatorname{grad} \hat{r} + \lambda(\hat{\boldsymbol{\omega}} \times (\mathbf{u} + \mathbf{u}_0) + (\boldsymbol{\omega} + \boldsymbol{\omega}_0) \times \hat{\mathbf{u}})\right) \\
&\qquad \times (\mathbf{u} + \mathbf{u}_0)
\end{aligned} \tag{4.23}$$

$$g_2 = -\lambda \operatorname{div}\, \left(\hat{\boldsymbol{\omega}} \times (\mathbf{u} + \mathbf{u}_0) + (\boldsymbol{\omega} + \boldsymbol{\omega}_0) \times \hat{\mathbf{u}}\right) \tag{4.24}$$

$$\begin{aligned}
g_3 = &\;\lambda \left(\mathbf{curl}\,\boldsymbol{\omega} + \operatorname{grad} r + \lambda(\boldsymbol{\omega} + \boldsymbol{\omega}_0) \times (\mathbf{u} + \mathbf{u}_0)\right) \times \hat{\boldsymbol{\omega}} \\
&+ \lambda \left(\mathbf{curl}\,\hat{\boldsymbol{\omega}} + \operatorname{grad} \hat{r} + \lambda(\hat{\boldsymbol{\omega}} \times (\mathbf{u} + \mathbf{u}_0) + (\boldsymbol{\omega} + \boldsymbol{\omega}_0) \times \hat{\mathbf{u}})\right) \\
&\qquad \times (\boldsymbol{\omega} + \boldsymbol{\omega}_0)\;,
\end{aligned} \tag{4.25}$$

independently from the type of boundary conditions. Next we note that all terms in equations (4.23)-(4.25) are of the form uvw or $u\frac{\partial v}{\partial x_i}$, where u, w, and v denote various components of U, U_0, and $\hat{U}$. Recall that for type two boundary conditions

$$\boldsymbol{\omega}, \hat{\boldsymbol{\omega}}, \boldsymbol{\omega}_0 \in [H^1(\Omega)]^3\,; \quad r, \hat{r}, r_0 \in H^1(\Omega)\,; \quad \mathbf{u}, \hat{\mathbf{u}}, \mathbf{u}_0 \in [H^2(\Omega)]^3\,.$$

Using (4.19)-(4.21) we can see that $g_1 \in [L^2(\Omega)]^3$ and $g_2 \in L^2(\Omega)$. Then, using (4.18) and (4.21) for the last equation it follows that $g_3 \in [L^{3/2}(\Omega)]^3$.
For type one boundary conditions

$$\boldsymbol{\omega}, \hat{\boldsymbol{\omega}}, \boldsymbol{\omega}_0 \in [H^1(\Omega)]^3\,; \quad r, \hat{r}, r_0 \in H^1(\Omega)\,; \quad \mathbf{u}, \hat{\mathbf{u}}, \mathbf{u}_0 \in [H^1(\Omega)]^3\,.$$

Using (4.18) and (4.21) it follows that for such boundary conditions $g_1 \in [L^{3/2}(\Omega)]^3$, $g_2 \in L^{3/2}(\Omega)$, and $g_3 \in [L^{3/2}(\Omega)]^3$.

3. Similarly to (4.22), one can show that for U' and U'' in $\mathbf{X}_{\sigma,0}$, we have $\mathbf{g} = D^2_U G(\lambda, U)[U', U'']$, if and only if

$$\begin{aligned}
(4.26)\quad (g_1, \boldsymbol{\xi})_0 = {} & \lambda(\mathbf{curl}\,\boldsymbol{\omega}'' + \operatorname{grad} r'' + \lambda(\boldsymbol{\omega}'' \times (\mathbf{u}+\mathbf{u}_0) + (\boldsymbol{\omega}+\boldsymbol{\omega}_0) \times \mathbf{u}''), \\
& \quad \boldsymbol{\xi} \times \mathbf{u}')_0 \\
& + \lambda(\mathbf{curl}\,\boldsymbol{\omega}' + \operatorname{grad} r' + \lambda(\boldsymbol{\omega}' \times (\mathbf{u}+\mathbf{u}_0) + (\boldsymbol{\omega}+\boldsymbol{\omega}_0) \times \mathbf{u}'), \\
& \quad \boldsymbol{\xi} \times \mathbf{u}'')_0 \\
& + \lambda(\boldsymbol{\omega}'' \times \mathbf{u}' + \boldsymbol{\omega}' \times \mathbf{u}'', \mathbf{curl}\,\boldsymbol{\xi} + \lambda\boldsymbol{\xi} \times (\mathbf{u}_0 + \mathbf{u}))_0 \\
(4.27)\quad (g_2, q)_0 = {} & \lambda\left(\boldsymbol{\omega}'' \times \mathbf{u}' + \boldsymbol{\omega}' \times \mathbf{u}'', \operatorname{grad} q\right)_0 \\
(4.28)\quad (g_3, \mathbf{v})_0 = {} & \lambda(\mathbf{curl}\,\boldsymbol{\omega}'' + \operatorname{grad} r'' + \lambda(\boldsymbol{\omega}'' \times (\mathbf{u}+\mathbf{u}_0) + (\boldsymbol{\omega}+\boldsymbol{\omega}_0) \times \mathbf{u}''), \\
& \quad \boldsymbol{\omega}' \times \mathbf{v})_0 \\
& + \lambda(\mathbf{curl}\,\boldsymbol{\omega}' + \operatorname{grad} r' + \lambda(\boldsymbol{\omega}' \times (\mathbf{u}+\mathbf{u}_0) + (\boldsymbol{\omega}+\boldsymbol{\omega}_0) \times \mathbf{u}'), \\
& \quad \boldsymbol{\omega}'' \times \mathbf{v})_0 \\
& + \lambda^2(\boldsymbol{\omega}'' \times \mathbf{u}' + \boldsymbol{\omega}' \times \mathbf{u}'', (\boldsymbol{\omega}_0 + \boldsymbol{\omega}) \times \mathbf{v})_0
\end{aligned}$$

for all $V \in \mathbf{X}_{\sigma,0}$. All terms in (4.26)-(4.28) are of the form $uvwz$ or $u\frac{\partial v}{\partial x_i}w$ where u, v, w, and z denote again various components of U, U_0, U', U'' and V. For both types of boundary operators these functions are at least in $H^1(\Omega)$. As a result, using (4.16) and (4.17) to estimate all terms in (4.26)-(4.28) yields

$$|(\mathbf{g}, V)_0| \le C(\lambda, \|U_0\|_{\mathbf{X}_{\sigma,0}}, \|U\|_{\mathbf{X}_{\sigma,0}}) \|U'\|_{\mathbf{X}_{\sigma,0}} \|U''\|_{\mathbf{X}_{\sigma,0}} \|V\|_{\mathbf{X}_{\sigma,0}}\,,$$

where $C(\lambda, \|U_0\|_{\mathbf{X}_{\sigma,0}}, \|U\|_{\mathbf{X}_{\sigma,0}})$ is polynomial function of λ, $\|U_0\|_{\mathbf{X}_{\sigma,0}}$, and $\|U\|_{\mathbf{X}_{\sigma,0}}$. As a result,

$$\|D^2_U G(\lambda, U)\|_{L^2(\mathbf{X},\mathbf{Y})} \le C(\lambda, \|U_0\|_{\mathbf{X}_{\sigma,0}}, \|U\|_{\mathbf{X}_{\sigma,0}})\,. \qquad \square$$

4.2. Approximation result. We are now ready to apply Theorem 4.1 and derive error estimates for the least-squares finite element method (3.7).

THEOREM 4.8. *Assume that Λ is a compact interval of $\mathbf{R}^+$ and that $\{(\lambda, U(\lambda)) \mid \lambda \in \Lambda\}$ is a branch of regular solutions of the problem (3.5). Assume that the finite element space $\mathbf{X}^h_\sigma$ satisfies (3.6) for some integer $d \ge 1$. Then, there exists a neighborhood $\mathcal{O}$ of the origin in $\mathbf{X}_{\sigma,0}$ and, for h sufficiently small, a unique branch $\{(\lambda, U^h(\lambda)) \mid \lambda \in \Lambda\}$ of solutions of the discrete problem (3.7) such that $U(\lambda) - U^h(\lambda) \in \mathcal{O}$ for all $\lambda \in \Lambda$. Moreover,*

$$(4.29)\qquad \|\boldsymbol{\omega}(\lambda) - \boldsymbol{\omega}^h(\lambda)\|_1 + \|r(\lambda) - r^h(\lambda)\|_1 + \|\mathbf{u}(\lambda) - \mathbf{u}^h(\lambda)\|_{1+\sigma} \to 0$$

as $h \to 0$, uniformly in λ. If the solution $U(\lambda)$ of (3.5) belongs to $\mathbf{X}_{\sigma,q}$ for some $q \ge 1$ then there exists a constant C, independent of h, such that

$$\begin{aligned}
(4.30)\qquad & \|\boldsymbol{\omega}(\lambda) - \boldsymbol{\omega}^h(\lambda)\|_1 + \|r(\lambda) - r^h(\lambda)\|_1 + \|\mathbf{u}(\lambda) - \mathbf{u}^h(\lambda)\|_{1+\sigma} \\
& \quad \le Ch^{\tilde{d}}\left(\|\omega(\lambda)\|_{\tilde{d}+1} + \|r(\lambda)\|_{\tilde{d}+1} + \|\mathbf{u}(\lambda)\|_{\tilde{d}+1+\sigma}\right),
\end{aligned}$$

where $\tilde{d} = \min\{d, q\}$.

Proof. Lemma 4.7 establishes (4.3). Assumptions (4.4) and (4.5) have been verified in Lemmas 4.4 - 4.6. Thus, all hypotheses of Theorem 4.1 are valid for problems (3.5) and (3.7), and (4.29) follows from (4.6) and (4.12).

To establish (4.30) assume that $U(\lambda) \in \mathbf{X}_{\sigma,q}$ for some $q \geq 1$, and let $\mathbf{g} = G(\lambda, U(\lambda))$. Then $T \cdot \mathbf{g} = -U(\lambda) \in \mathbf{X}_{\sigma,q}$, and estimate (4.30) follows from (4.6) and (4.13). $\square$

5. Implementation of least-squares methods. We begin this section with remarks concerning theoretical settings for the results in Section 4. Then we continue with a brief discussion of Newton's method for solution of (3.7), and an outline of continuation techniques which can be used in conjunction with Newton's method.

Error estimates of Section 4 are valid for all bona fide conforming finite element methods. In practice, however, one is often forced to commit "variational crimes" in order to implement the method efficiently. For example, ADN a priori estimates are valid under the assumption that the boundary Γ is of class $C^{1+\sigma}$. This formally rules out from consideration domains such as polygons and polytopes. At the same time, boundary conditions which involve normal or tangential components of vector fields are very difficult to satisfy, unless Γ consists of straight line segments. Least-squares approach offers an elegant solution of this problem which is to include residuals of boundary conditions into least-squares quadratic functionals. Such are methods for linear elliptic boundary value problems proposed by Aziz et. al. [2] and Wendland [43]. Typical for these methods is the use of weighted L^2-norms of the boundary terms in order to avoid computation of inner products in fractional order Sobolev spaces. Although method (3.7) can be easily modified along these lines, the error analysis of such method will be significantly complicated by the presence of mesh dependent boundary norms, and is beyond the scope of this paper.

We stress upon the fact that smoothness of Γ enters our analysis only through the use of ADN theory for derivation of a priori estimates. Consequently, if these estimates can be established under less stringent conditions on Γ all results of Section 4 concerning the method (3.7) will remain valid. For some recent developments in this direction the reader may consult [14] and [32], where a priori estimates for the Stokes problem are established by means of Friedrichs-Poincare type inequalities, and various vector field decomposition results. At first it might appear that this somewhat reduces the value of ADN theory in the context of least-squares methods. However, because complementing condition is necessary and sufficient for the a priori estimates, ADN approach remains the most effective analytical tool for systematic identification of both the admissible boundary conditions and the appropriate function analytic settings for elliptic boundary value problems. Moreover, features like multiplicity of principal parts and the associated a priori estimates, and their dependence on dimensionality and boundary operators are difficult to determine by other means. Once the proper functional framework for a given boundary value problem has been established with the help of ADN theory, one may try to relax the smoothness of Γ, needed at this stage, using alternative techniques.

5.1. The Newton's method. Discrete problem (3.7) constitutes nonlinear system of algebraic equations that must be solved in an iterative manner. There are many methods that one might use for such a purpose; here we only consider Newton's method. We write the discrete problem (3.7) formally as

$$U^h + T^h \cdot G(\lambda, U^h) = \mathbf{0}.$$

Then, for a given initial guess $U_0^h = (\boldsymbol{\omega}_0^h, r_0^h, \mathbf{u}_0^h)$ the sequence of Newton iterates

$$U_k^h = (\boldsymbol{\omega}_k^h, r_k^h, \mathbf{u}_k^h)_{k>0} = (\boldsymbol{\omega}_{k-1}^h + \Delta\boldsymbol{\omega}_k^h, r_{k-1}^h + \Delta r_k^h, \mathbf{u}_{k-1}^h + \Delta\mathbf{u}_k^h) = U_k^h + \Delta U_k^h$$

is generated recursively by solving, for $k = 1, 2, \ldots$, the linear system

$$(I + T^h \cdot D_U G(\lambda, U_{k-1}^h)) \cdot \Delta U_k^h = -(U_{k-1}^h + T^h \cdot G(\lambda, U_{k-1}^h)). \tag{5.1}$$

The explicit form of the system of algebraic equations (5.1) is rather formidable. However it also has some very good features. First, it is easy to see that this system is symmetric. Indeed in a neighborhood of a solution of (3.7), the Hessian matrix for the functional (3.2) is necessarily positive definite; but this Hessian matrix is exactly the coefficient matrix $(I + T^h \cdot D_U G(\lambda, U_{k-1}^h))$ of (5.1). As a result, in a neighborhood of a solution of (3.7) the system (5.1) is *symmetric* and *positive definite* independently from the value of the Reynolds number. This valuable property of the least-squares method can be used to devise an algorithm for numerical solution of the Navier-Stokes equations which will encounter only symmetric and positive definite algebraic systems in the solution process. Solution of these systems can be accomplished by efficient iterative solvers, such as conjugate gradients method. As a result, method (3.7) can be implemented without assembling the matrix in (5.1). Along with the guaranteed local and quadratic convergence of Newton's method this makes the least-squares algorithm very attractive for large scale computations, see [29], [31] and [37].

5.2. Continuation methods. In this section we briefly describe some continuation techniques that can be used in conjunction with the least-squares methods for the Navier-Stokes equations. For more detailed discussion of this subject we will refer the reader to [6].

The need to incorporate continuation strategies into the method stems from the following observations. First, as the Reynolds number increases, the attraction ball for Newton's method will decrease. Second, positive definiteness of the Hessian matrix is guaranteed only in a neighborhood of the minimizer. As a result, for an arbitrary initial guess we may have that Newton's method does not converge and/or that the coefficient matrix in (5.1) is not positive definite. In order to guarantee that the initial guess is within the attraction ball of Newton's method and that the coefficient matrix in (5.1) is positive definite, one can use continuation or homotopy methods, among others. A simple continuation method can be defined as follows (see e.g. [36], [41]). Let us symbolically express system (3.7) in the form

$$F(U^h; Re) = 0$$

where $Re = 1/\nu$ is the target Reynolds number. Consider a sequence of increasing Reynolds numbers $\{Re_m\}_{m=1}^M$ with $Re_M = Re$. Let U_m denote solution of system (3.7) for Re_m. This solution, for any m, is obtained by solving the sequence of linear systems (5.1) for $k = 1, 2, \ldots$ and a given initial approximation U_m^0. To start the continuation procedure we can choose Re_1 to be sufficiently small, e.g. $Re_1 = 1$ so that iteration (5.1) converges if U_1^0 is defined to be the solution of the linear Stokes problem (2.12)-(2.14), (2.24). The remaining initial guesses U_m^0 can be determined by "continuing along the tangent", that is, by solving the linear system

$$D_U F(U_{m-1}; Re_{m-1}) \cdot \left(U_m^0 - U_{m-1}\right) = -(Re_m - Re_{m-1}) D_{Re} F(U_{m-1}; Re_{m-1})$$

or even by a simpler "continuation along a constant" method

$$U_m^0 = U_{m-1}\,.$$

The combined Newton-continuation method is now completely defined. Since Newton's method is guaranteed to be locally convergent, and since the neighborhood of a minimizer where the Hessian matrix is positive definite is also nontrivial, one can guarantee, by choosing $Re_m - Re_{m-1}$ sufficiently small, that the Newton-continuation method should only encounter symmetric and positive definite matrices.

6. Numerical examples. In this section we present computational study of the finite element least-squares method (3.7) in $\mathbb{R}^2$. The main objectives of our numerical experiments are:

1. to illustrate error estimates (4.30) with the boundary condition BC2A;
2. to asses importance of enforcing the zero mean constraint for pressure approximations;
3. to study the impact of the ill-posedness of BC4 on computations.

For numerical experiments with the velocity condition we refer to [7], [31], [37], and [35]. Numerical examples with normal velocity-pressure boundary condition can be found in [6] and [7]. In all experiments Ω is taken to be the unit square. We consider four examples of artificial planar flows, that is, we begin with a known velocity and pressure fields and then compute the data by evaluating Navier-Stokes equations (2.4)-(2.6) at the exact solution. The four examples are as follows.

Example 1.
$$\begin{aligned} \mathbf{u} &= (\sin(\pi x)\cos(\pi y), -\cos(\pi x)\sin(\pi y))^T \\ r &= y(1-y)\sin(\pi x)\,. \end{aligned}$$

Example 2.
$$\begin{aligned} \mathbf{u} &= \left(x(1-x)\cos(\pi y), -\frac{(1-2x)}{\pi}\sin(\pi y)\right)^T \\ r &= y(1-y)\sin(\pi x) \end{aligned}$$

Example 3.
$$\begin{aligned} \mathbf{u} &= \big(\exp(x)\cos(y)+\sin(y), -\exp(x)\sin(y)+1-x^3\big)^T \\ r &= \sin(y)\cos(x)+xy^2-\frac{1}{6}-\sin(1)(1-\cos(1))\,. \end{aligned}$$

Example 4.
$$\begin{aligned} \mathbf{u} &= \left(\frac{y}{(x^2+y^2+0.05)}, -\frac{x}{(x^2+y^2+0.05)}\right)^T \\ r &= \sin(y)\cos(x)+xy^2-\frac{1}{6}-\sin(1)(1-\cos(1))\,. \end{aligned}$$

These examples provide various combinations of homogeneous and inhomogeneous boundary data on Γ. For Examples 1 and 2 we have that $\mathbf{u}\cdot\mathbf{n}=0$, $\boldsymbol{\omega}=0$, $r=0$ on Γ, and that $\mathbf{u}\cdot\mathbf{n}=0$, $r=0$, respectively. For the last two examples we have that $\mathbf{u}\cdot\mathbf{n}\neq 0$, $\boldsymbol{\omega}\neq 0$, $r\neq 0$ on Γ, $r\in L^2_0(\Omega)$.

Since BC2A is of type one the method (3.7) is implemented with $\sigma=0$. The finite element space $\mathbf{X}^h_\sigma$ used in computations is defined as follows. For a given region $S\subset\mathbb{R}^2$, let $Q_2(S)$ denote the set of all functions which are polynomials of degree less than or equal to 2 in each of the coordinate directions. We consider uniform triangulation of Ω into rectangles and let $S^h_1\equiv S^h_2$ be the space of biquadratic finite elements

$$S^h_1 = \{u^h\in C^0(\Omega)\,|\,u^h|_\Box\in Q_2(\Box),\quad \Box\in\mathcal{T}_h\}\,. \tag{6.1}$$

The resulting finite element space $\mathbf{X}^h_\sigma$ uses equal order interpolation for all unknowns and can be shown to satisfy (3.6) with $d=2$, see for instance, [18] or [26]. Thus, we

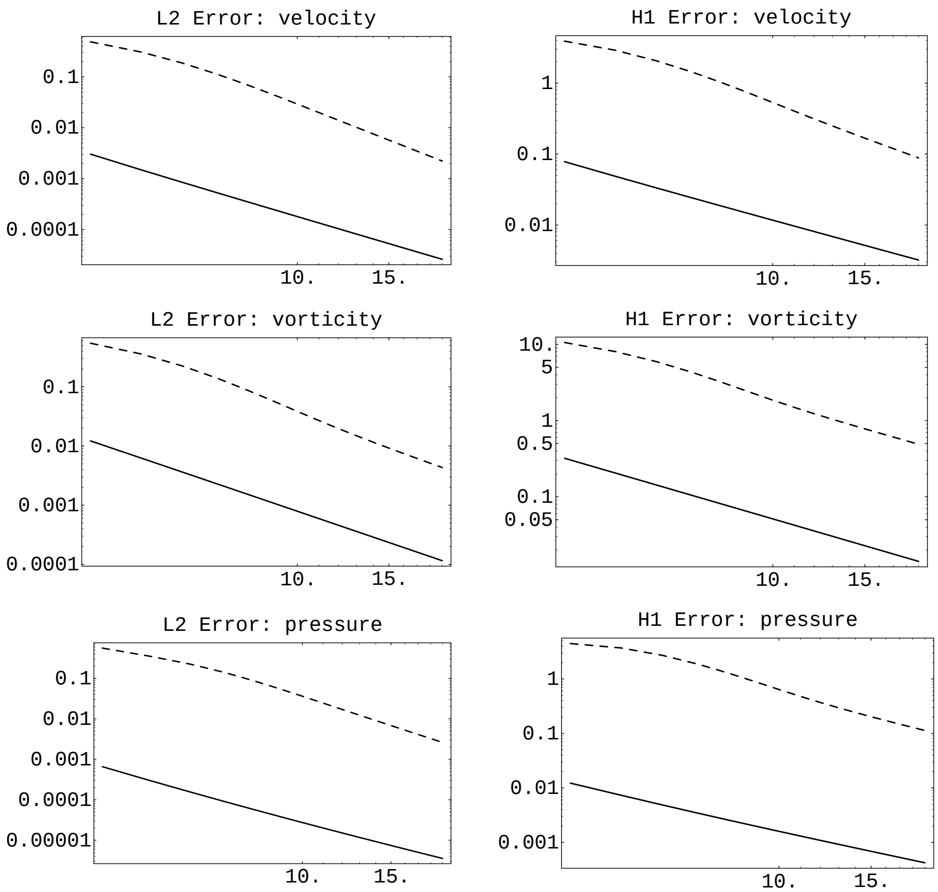

FIG. 6.1. *L^2 and H^1 errors for Example 1 (solid line) vs. Example 4 (dashed line) with BC2A.*

expect that for $\mathbf{u} \in [H^3(\Omega)]^2$, $\boldsymbol{\omega} \in H^3(\Omega)$ and $r \in H^3(\Omega)$,

$$\|\mathcal{E}\|_{\mathbf{X}_{\sigma,0}} \equiv \|\mathbf{u} - \mathbf{u}^h\|_1 + \|\boldsymbol{\omega} - \boldsymbol{\omega}^h\|_1 + \|r - r^h\|_1 = O(h^2)\,. \tag{6.2}$$

For all examples errors are computed using Gauss-Legendre quadrature formula of degree seven. Convergence rates are estimated by computing approximate solutions on sequence of uniform grids and finding the slope of the best least-squares straight line fit for the log-log coordinates of the errors vs. grid size. Since the main goal of our experiments is to illustrate numerically convergence rates (6.2), computations are restricted to grid sizes which allow one to enter into the asymptotic range of the error estimates. We found that triangulations of up to 20 by 20 biquadratic elements (39 by 39 grid points in each direction) are sufficient for this purpose. For Examples 1 and 4 this observation is illustrated by logarithmic plots of L^2 and H^1 errors vs. the number of grid intervals in each direction presented in Fig. 1. Similar results were obtained for Examples 2 and 3. In addition to H^1 rates we also report results for L^2 rates. Although our analysis does not include such estimates we expect to observe rates of order $O(h^3)$. Finally, to attain better understanding of the errors, for each experiment we present both the rates in $\mathcal{E}$, and in the individual components of the

Table 6.1
Rates of convergence for Examples 1, 2, 3 and 4 with BC2A and $r^h \in L^2_0(\Omega)$.

Error	Example 1	Example 2	Example 3	Example 4
	L^2 error rates			
$\mathbf{u}$	3.01	3.00	3.00	3.63
$\boldsymbol{\omega}$	3.00	3.00	3.02	3.29
r	3.12	3.00	3.04	3.63
$\mathcal{E}$	**3.00**	**3.00**	**3.00**	**3.29**
	H^1 error rates			
$\mathbf{u}$	2.01	2.00	2.01	2.58
$\boldsymbol{\omega}$	2.00	2.00	2.01	2.09
r	2.07	2.07	2.03	2.59
$\mathcal{E}$	**2.00**	**2.00**	**2.01**	**2.09**

solution.

All computations for Examples 1-4 were carried out with $\nu = 1$. Initial approximations for the Newton's method were computed using a least-squares solver for the Stokes problem. At each Newton step the linearized system has been solved by conjugate gradients method with Jacobi preconditioning. Inhomogeneous boundary conditions in Examples 2-4 were treated by using boundary interpolants of the data in order to define boundary conditions which can be satisfied by the finite element functions.

Remark. In actual computations the zero mean constraint for the pressure approximations can be imposed a posteriori [28] or by fixing the value of the pressure at some point [6].

In the first experiment we have computed convergence rates for Examples 1, 2, 3 and 4. These results are summarized in Table 2. We observe that computed H^1 rates for all unknowns, including the vorticity, are in very good agreement with the theoretical estimate (6.2). The L^2 rates also appear to be optimal, at least computationally. In contrast, recall that other methods involving vorticity frequently need artificial boundary conditions for this variable and often yield poor vorticity approximations; see, e.g., Gunzburger et. al. [27]. We note that the sides of our computational domain are parallel to the coordinate axes, and as a result, the boundary condition $\mathbf{u} \cdot \mathbf{n} = 0$ for Example 1 can be implemented exactly. Similarly, approximation of inhomogeneous boundary conditions for Examples 2-4 only involves interpolation of boundary data. Since convergence rates for these examples are not affected by the presence of inhomogeneities, one can conclude that use of boundary interpolants in this context is acceptable.

Convergence rates for Examples 1, 2, 3 and 4, without enforcing the zero mean constraint in computations, are reported in Table 3. The H^1 rates in Table 3 are identical with the H^1 rates in Table 2, that is computations with pressure approximations $r^h \notin L^2_0(\Omega)$ does not affect seriously H^1 convergence. At the same time the L^2 rates for $\mathcal{E}$ are much worse in Examples 1 and 2. However, a closer look at the individual rates reveals that the loss of convergence is entirely due to the poor pressure approximation. It also appears that when $r \in L^2_0(\Omega)$, as it is for Examples 3 and 4, the L^2 convergence is not completely destroyed.

For the third experiment computations were performed with the theoretically ill-posed boundary condition BC4. Corresponding rates, except for Example 4 where Newton's method diverged after one iteration, are summarized in Table 4.

TABLE 6.2
Rates of convergence for Examples 1, 2, 3 and 4 with BC2A and r^h not in $L^2_0(\Omega)$.

Error	Example 1	Example 2	Example 3	Example 4
	L^2 error rates			
$\mathbf{u}$	3.01	3.00	3.00	3.63
$\boldsymbol{\omega}$	3.00	3.00	3.02	3.29
r	0.00	0.00	2.04	3.57
$\mathcal{E}$	**0.00**	**0.00**	**2.04**	**3.29**
	H^1 error rates			
$\mathbf{u}$	2.01	2.00	2.01	2.58
$\boldsymbol{\omega}$	2.00	2.00	2.01	2.09
r	2.07	2.07	2.03	2.59
$\mathcal{E}$	**2.00**	**2.00**	**2.01**	**2.09**

TABLE 6.3
Rates of convergence for Examples 1, 2, 3 and 4 with BC4.

Error	Example 1	Example 2	Example 3	Example 4
	L^2 error rates			
$\mathbf{u}$	3.01	2.09	2.58	-
$\boldsymbol{\omega}$	3.00	3.00	3.05	-
r	3.63	3.10	3.05	-
$\mathcal{E}$	**3.00**	**2.09**	**2.58**	-
	H^1 error rates			
$\mathbf{u}$	2.44	1.43	1.96	-
$\boldsymbol{\omega}$	1.99	2.00	2.01	-
r	2.18	2.01	2.03	-
$\mathcal{E}$	**1.99**	**1.43**	**1.96**	-

Interestingly, the ill-posedness of BC4 does not seem to have disastrous effect on convergence rates, in particular, for Example 1. Still, we observe reduced L^2 rates for Examples 2 and 3, and reduced H^1 rates for Example 2. At the same time, the ill-posedness of BC4 does have serious impact on conditioning of the linearized algebraic problems. In the context of the Stokes problem one can show that the spectral condition number of least-squares discretization matrices for boundary operators of type one (such as BC2A) is of order h^{-2}, see [9]. It is not difficult to extend these results to linearized algebraic problems corresponding to the Navier-Stokes equations with boundary conditions of the same type. Thus, we shall compare conditioning of linearized problems with BC2A and BC4. Assuming that condition number is of order $O(h^{-\alpha})$, we have that for uniform triangulations with $n \times n$ finite elements the asymptotic number $m(n)$ of conjugate gradient iterations necessary for convergence to a prescribed tolerance is $m(n) \approx Cn^{\alpha/2}$. Thus, we expect that the number of iterations $m(n)$ for BC2A will behave like a linear function in n. This observation is confirmed by the plots of $m(n)$ vs. number of grid intervals given in Fig. 2. The plots of $m(n)$ for BC4, on the other hand, behave like powers of n, i.e., we can conclude that condition numbers in this case are likely to be higher than $O(h^{-2})$. Conclusions drawn from Figure 2 can be quantized by estimating α according to the formula

$$\alpha \approx 2 \frac{\log\left(\frac{m(i)}{m(j)}\right)}{\log\left(\frac{i}{j}\right)} .$$

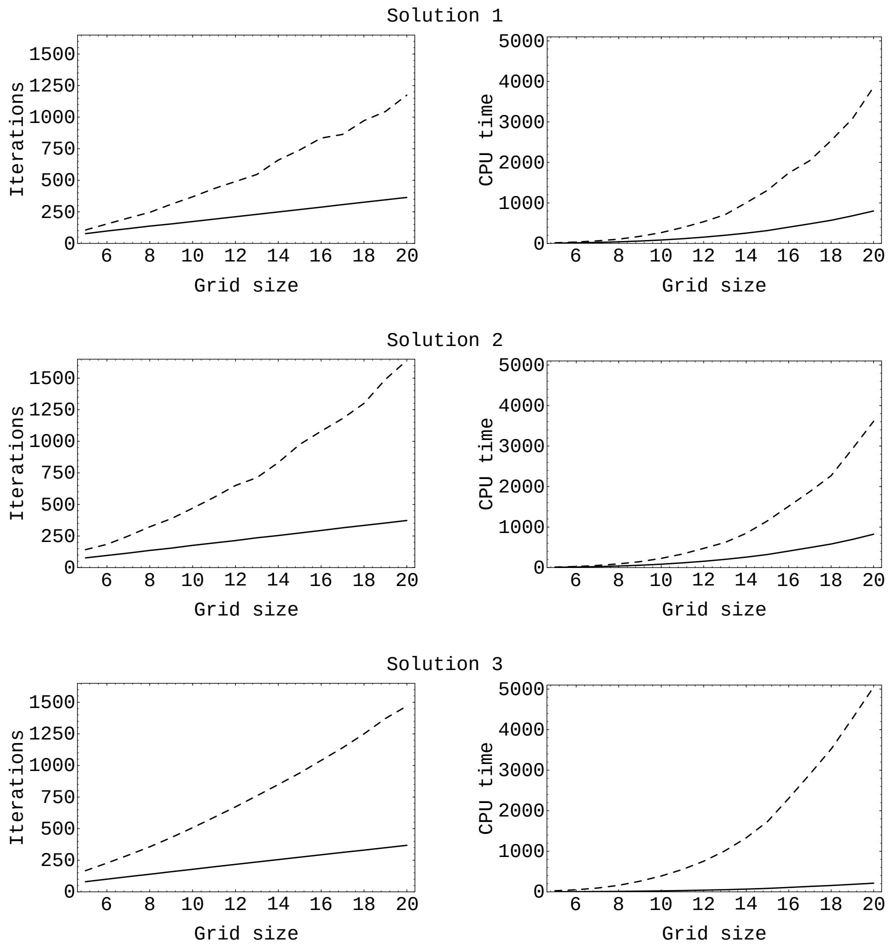

FIG. 6.2. *Average number of conjugate gradients iterations per Newton step and total CPU times for Examples 1, 2 and 3 with BC2A (solid lines) vs. BC4 (dashed lines).*

TABLE 6.4
Condition number estimates for BC2A and BC4.

Boundary Condition	Example 1	Example 2	Example 3	Example 4
BC2A	2.11	2.12	2.04	2.09
BC4	3.57	3.90	3.02	-

The results are given in Table 5, where α has been computed by averaging values obtained with $i = 20$ and $j = 19, ..., 15$. We see that the estimates for α with BC2A in Table 5 are, indeed, very close to the expected theoretical value of 2, whereas conditioning of the problems with BC4 approaches $O(h^{-4})$. The higher condition numbers for BC4 are also reflected by substantially higher CPU times needed for solution of corresponding discrete equations, see Fig. 2. In fact, using a similar argument, one can show that CPU times with BC4 approach $O(h^{-4})$.

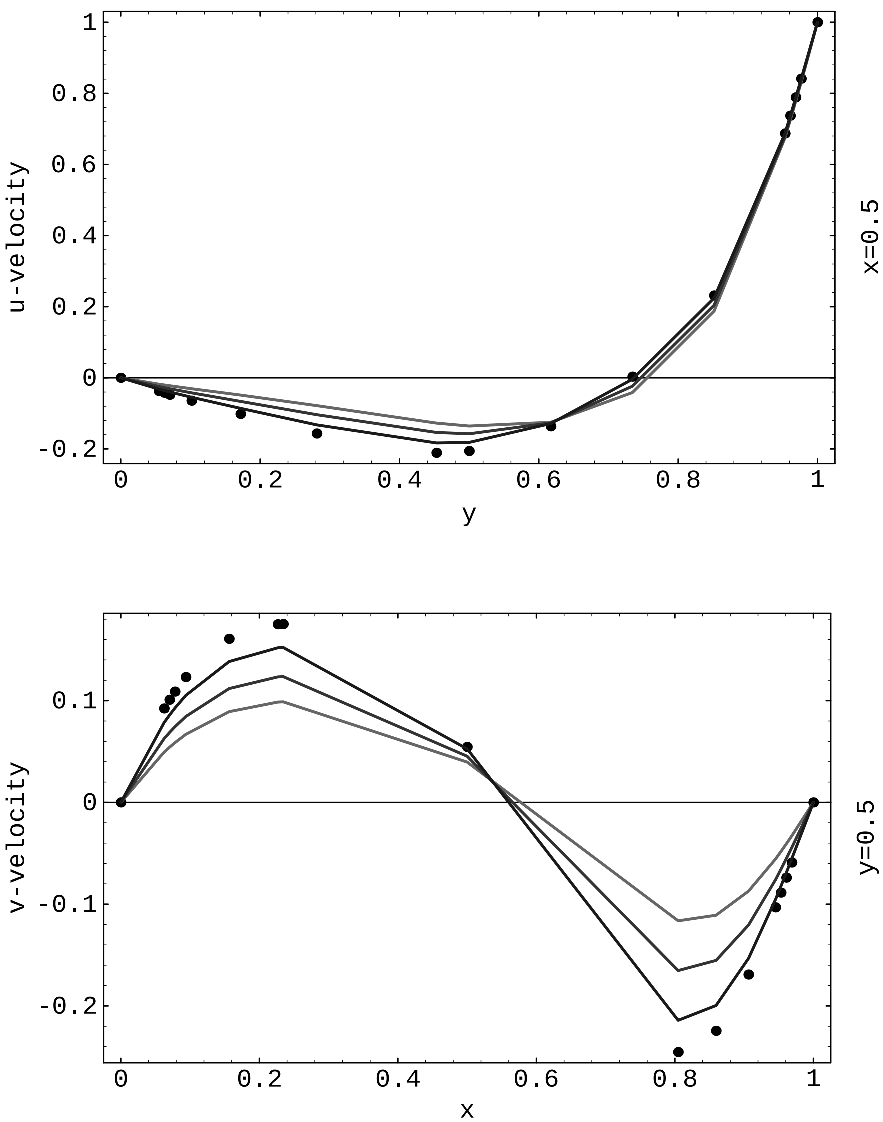

FIG. 6.3. *Profiles of velocity components for 18, 25 and 33 uniform grid lines vs. benchmark results (solid dots) [22]; $Re = 100$.*

We conclude with one final numerical experiment in which the fictitious, driven cavity flow is approximated using the method (3.7) with $\sigma = 0$. For this flow $\mathbf{f} = \mathbf{0}$ and $\mathbf{u} = \mathbf{0}$ on all sides of the cavity, except for the top one where $\mathbf{u} = (1, 0)$. A well-documented benchmark results for the driven cavity are given by Ghia et. al. in [22]. We have computed three approximate solutions for $Re = 100$ using uniformly spaced grids with 18, 23 and 33 grid lines in each direction. These results are compared with the benchmark data of [22] on Fig. 3. Our experiment shows that when the exact solution lacks sufficient regularity, theoretically suboptimal method (for the velocity boundary condition (3.7) results in an optimal method when $\sigma = 1$) can still provide reasonable results. This conclusion is also supported by similar results reported in [29], [31] and [37].

Acknowledgments. - During the preparation of this paper I had many stimulating and fruitful discussions with Prof. M. D. Gunzburger who introduced me to least-squares methods. His constant encouragement was essential for the completion of this work.

I also would like to thank the anonymous referees for the thoughtful and thorough remarks and suggestions which helped to improve the content and the style of this paper, and Prof. J. Bramble for pointing out references [4], [11] and [23].

REFERENCES

[1] S. Agmon, A. Douglis, and L. Nirenberg, *Estimates near the boundary for solutions of elliptic partial differential equations satisfying general boundary conditions II*, Comm. Pure Appl. Math., 17 (1964), pp. 35–92.

[2] A. Aziz, R. Kellogg, and A. Stephens, *Least-squares methods for elliptic systems*, Math. Comp., 44 (1985), pp. 53–70.

[3] M. A. Behr, L. P. Franca, and T. E. Tezduyar, *Stabilized finite element methods for the velocity-pressure-stress formulation of incompressible flows*, Computer Methods in Appl. Mech. Engrg., 104 (1993), pp. 31–48.

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

[5] P. Bochev and M. Gunzburger, *A least-squares finite element method for the Navier-Stokes equations*, Appl. Math. Lett. 6/2 (1993), pp. 27–30.

[6] ———, *Accuracy of least-squares methods for the Navier-Stokes equations*, Comput. & Fluids, 22 (1993), pp. 549–563.

[7] ———, *Analysis of least-squares finite element methods for the Stokes equations*, Math. Comp., 63 (1994), pp. 479–506.

[8] ———, *Least-squares methods for the velocity-pressure-stress formulation of the Stokes equations*, Comp. Meth. Appl. Mech. Engrg., 126 (1995), pp. 267–287.

[9] P. Bochev, *Least-squares methods for the Stokes and Navier-Stokes equations*, PhD thesis, Department of Mathematics, Virginia Polytechnic Institute and State University, Blacksburg, VA, 1994.

[10] 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*, Mathematical Science Institute, Cornell University, Technical report 94-32 (1994), 23 pages.

[11] J. H. Bramble and P. Lee, *On variational formulations for the Stokes equations with nonstandard boundary conditions*, Math. Model. and Numer. Analysis, 28 (1994), pp. 903–919.

[12] F. Brezzi, J. Rappaz, and P.-A. Raviart, *Finite-dimensional approximation of nonlinear problems, Part I: Branches of nonsingular solutions*, Numer. Math., 36 (1980), pp. 1–25.

[13] Z. Cai, R. D Lazarov, T. A. Manteuffel, and S. F. McCormick, *First-order system least-squares for partial differential equations: discretization*, SIAM J. Numer. Anal., 31 (1994), pp. 1785–1799.

[14] Z.Cai, T. A. Manteuffel, and S. F. McCormick, *First-order system least-squares for partial differential equations: Part II*, to appear in SIAM J. Numer. Anal., 34 (1997).

[15] ———, *First-Order System Least Squares for Velocity-Vorticity-Pressure Form of the Stokes Equations, with Application to Linear Elasticity*, Electronic Transactions in Numerical Analysis, 3 (1995), pp. 150–159.

[16] C.-L. Chang, *Finite element approximation for grad-div type systems in the plane*, SIAM J. Num. Anal., 29 (1992), pp. 452–461.

[17] C.-L. Chang and M. Gunzburger, *A finite element method for first-order systems in three dimensions*, Appl. Math. Comput., 23 (1987), pp. 171–184.

[18] P. G. Ciarlet, *The finite element method for elliptic problems*, North-Holland, Amsterdam, 1978.

[19] G. J. Fix, M. Gunzburger, and R. Nicolaides, *On finite element methods of the least-squares type*, Comput. Math. Appl., 5 (1979), pp. 87–98.

[20] L. P. Franca, S. L. Frey, and T. J. R. Hughes, *Stabilized finite element methods: I. Application to the advective-diffusive model*, Comput. Methods Appl. Mech. Engrg., 95 (1992), pp. 253–276.

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

[22] U. Ghia, K. N. Ghia, and C. T. Shin, *High-Re solutions for incompressible flow using the Navier-Stokes equations and a multigrid method*, J. Comp. Phys., 48 (1982), pp. 387–411.

[23] V. Girault, *Incompressible finite element methods for Navier-Stokes equations with nonstandard boundary conditions in* $\mathbf{R}^3$, Math. Comp., 51 (1988), pp. 55–74.

[24] V. Girault and P.-A. Raviart, *Finite Element Methods for Navier-Stokes Equations*, Springer Verlag, Berlin, 1986.

[25] P. M. Gresho, *Incompressible fluid dynamics: some fundamental formulation issues*, Ann. Rev. Fluid Mech., 23 (1991), pp. 413–453.

[26] M. Gunzburger, *Finite Element Methods for Viscous Incompressible Flows*, Academic Press, Boston, 1989.

[27] M. Gunzburger, M. Mundt, and J. Peterson, *Experiences with computational methods for the velocity-vorticity formulation of incompressible viscous flows*, Computational Methods in Viscous Aerodynamics, Elsevier, Amsterdam, pp. 231–271, 1990.

[28] M. Gunzburger and J. Peterson, *On conforming finite element methods for the inhomogeneous stationary Navier-Stokes equations*, Numer. Math., 42 (1983), pp. 173–194.

[29] B.-N. Jiang, *A least-squares finite element method for incompressible Navier-Stokes problems*, Inter. J. Numer. Meth. Fluids, 14 (1992), pp. 843–859.

[30] B.-N. Jiang and C. L. Chang, *Least-squares finite elements for the Stokes problem*, Comput. Methods Appl. Mech. Engrg., 78 (1990), pp. 297–311.

[31] B.-N. Jiang, T. Lin, and L. Povinelli, *Large-scale computation of incompressible viscous flow by least-squares finite element method*, Comput. Methods Appl. Mech. Engrg., 114 (1994), pp. 213–231.

[32] B.-N. Jiang, C. Y. Loh, and L. Povinelli, *Theoretical study of the incompressible Navier-Stokes equations by the least-squares method*, NASA Technical Memorandum 106535, ICOMP-94-04, NASA, Cleveland, 1994.

[33] B.-N. Jiang and L. Povinelli, *Least-squares finite element method for fluid dynamics*, Comput. Meth. Appl. Mech. Engrg., 81 (1990), pp. 13–37.

[34] ———, *Optimal Least-Squares finite element methods for elliptic problems*, Comp. Meth. Appl. Mech. Engrg., 102 (1993), pp. 199–212.

[35] B.-N. Jiang and V. Sonnad, *Least-squares solution of incompressible Navier-Stokes equations with the p-version of finite elements*, NASA Technical Memorandum 105203, ICOMP-91-14, NASA, Cleveland, 1991.

[36] H. Keller, *Numerical methods in Bifurcation Problems*, Springer Verlag, Berlin, 1987.

[37] D. Lefebvre, J. Peraire, and K. Morgan, *Least-squares finite element solution of compressible and incompressible flows*, Int. J. Num. Meth. Heat Fluid Flow, 2 (1992), pp. 99–113.

[38] T. Manteuffel, private communication.

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

[40] M. Renardy and R. Rogers, *Introduction to Partial Differential Equations*, Springer Verlag, Berlin, 1993.

[41] W. Rheinboldt, *Solution fields of nonlinear equations and continuation methods*, SIAM J. Numer. Anal., 17 (1980), pp. 221–246.

[42] R. Témam, *Nonlinear Functional Analysis and Navier-Stokes Equations*, SIAM, Philadelphia, 1983.

[43] W. Wendland, *Elliptic Systems in the Plane*, Pitman, London, 1979.