Negative Norm Least-Squares Methods for the Velocity-Vorticity-Pressure Navier-Stokes Equations

Pavel B. Bochev [1]
Department of Mathematics, Box 19408, University of Texas at Arlington, Arlington, TX 76019-0408 (bochev@utamat.uta.edu)

Received July, 1997

We develop and analyze a least-squares finite element method for the steady state, incompressible Navier-Stokes equations, written as a first-order system involving vorticity as new dependent variable. In contrast to standard L^2 least-squares methods for this system, our approach utilizes discrete negative norms in the least-squares functional. This allows to devise efficient preconditioners for the discrete equations, and to establish optimal error estimates under relaxed regularity assumptions. © 1998 John Wiley & Sons, Inc.

I. INTRODUCTION

The use of least-squares variational principles has been among the more recent developments in the numerical solution of elliptic boundary value problems (BVP). Such principles offer several important computational and analytic advantages that are not present in other discretization schemes. This is especially true for problems requiring approximation of different physical quantities, like velocity and pressure in the Stokes equations. Indeed, standard mixed Galerkin methods for such equations lead to saddle-point optimization problems. It is well-known that discretization of such problems is not stable, unless approximation spaces for the different unknowns satisfy the Ladyzhenskaya-Babuska-Brezzi (LBB) condition; see, e.g., [15]-[16]. Because algebraic problems associated with saddle-point optimization are indefinite, their numerical solution is more difficult. In addition, the LBB condition limits the choice of discretization spaces, which further complicates the algorithmic development of mixed finite element methods. For example, a mixed method for the Stokes problem is unstable if equal order spaces, defined with respect to the same triangulation, are used for both the velocity and pressure fields.

A least-squares method for a given BVP, on the other hand, is based on minimization of a problem-dependent, convex functional whose minimizer coincides with the solution of the PDE. As a result, a least-squares finite element method is stable without addi-

[1]This work was sponsored by the National Science Foundation under grant number DMS-9705793.

tional constraints on the finite element spaces, and the associated algebraic problems are symmetric and positive definite. In this paper we develop the least-squares methodology for the approximate solution of the steady-state, incompressible Navier-Stokes equations. In this context, the properties of least-squares principles have important impact on the algorithmic design of the finite element methods. For example, used in conjunction with Newton linearization, least-squares lead to symmetric and positive definite systems, at least in a neighborhood of the solution. Thus, in principle, one can devise a method for the approximate solution of the Navier-Stokes equations that will encounter only symmetric and positive definite algebraic systems in the solution process. These systems can be solved by robust and efficient iterative methods without assembling the discretization matrix, i.e., methods are well-suited for large-scale computations. The algorithmic development of the methods is further simplified by the possibility to use standard finite element spaces, including equal-order interpolation for all unknowns.

Before a necessarily brief review of the existing methods for the Stokes and Navier-Stokes equations let us introduce the notation that will be used throughout this paper. Ω and Γ will denote a bounded open region in $\mathbf{R}^n$, $n = 2, 3$, and its boundary, respectively. We use the standard notation $H^s(\Omega)$ for the Sobolev space of order $s \geq 0$, equipped with a norm $\| \cdot \|_s$, and an inner product $(\cdot, \cdot)_s$. When $s = 0$ we shall write $L^2(\Omega)$, instead of $H^0(\Omega)$, and omit s from the norm and inner product designations. As usual, $H^1_0(\Omega)$ will denote the subspace of all functions in $H^1(\Omega)$ which vanish on Γ, $L^2_0(\Omega)$ - the subspace of all zero mean functions in $L^2(\Omega)$, and $H^{-1}(\Omega)$ - the dual of $H^1_0(\Omega)$ equipped with the norm

$$\|\phi\|_{-1} = \sup_{\psi \in H^1_0(\Omega)} \frac{(\phi, \psi)}{\|\psi\|_1} \,. \tag{1.1}$$

Vector valued functions and corresponding Sobolev spaces will be denoted by bold face symbols, e.g., $\mathbf{H}^{-1}(\Omega)$, $\mathbf{H}^1(\Omega)$, $\mathbf{L}^2(\Omega)$, etc. Throughout this paper C will denote a generic positive constant whose value and meaning may change with context.

A. Least-squares methods for Stokes and Navier-Stokes equations

The inherent advantages of least-squares principles can be fully utilized only if development of the finite element methods is carried out in a setting which allows to combine practical implementation with optimal accuracy. In most of the existing literature the first task is accomplished by introducing new dependent variables effecting transformation of the Stokes, or the Navier-Stokes equations into an equivalent first-order system; see, e.g., [2]-[7], [9], [11], [12], [17]-[21]. Then, a least-squares functional for the first-order system can be defined by summing up L^2 norms of the equation residuals. Discretization of such *standard least-squares* functionals can be accomplished using simple, merely continuous finite element spaces. Most of the recent research has dealt with vorticity-based first order systems; see, e.g., [2], [4], [5], [6], [11], and [17]-[21]. Another possibilities include the use of the velocity gradient (see [3], [12]), or the stress tensor (see [7]) as new dependent variables. The second task, i.e., achieving optimal discretization errors, can be accomplished by using *norm equivalent* least-squares functionals. A norm-equivalent functional is a functional that generates a norm on the solution space for the particular BVP. As a result, the associated weak problems are coercive on the solution space, and error estimates can be derived using standard elliptic finite element arguments. Typically, norm-equivalent functionals are defined by using data norms from a valid a priori estimate for the given BVP. This approach works fine and leads to optimally accurate and practical finite element methods, if the BVP is well-posed in products of $L^2(\Omega)$ and $H^1(\Omega)$ spaces, for the data and the solution, respectively. For such first-order systems, known also as H^1-*coercive*, a standard least-squares functional is norm-equivalent, i.e., leads to optimally accurate finite element methods. Although it seems reasonable to expect that a first-order system will be H^1-coercive, this property is exhibited only by systems of Petrovsky type; see [22]. An example of a first-order system that is not H^1-coercive is furnished by the velocity-vorticity-pressure form of the Stokes equations. The combination of function spaces in which this system is well-posed depends, among other things, on the boundary condition and the space dimension. In particular, with the important velocity boundary condition, the velocity-vorticity-pressure system is not H^1-coercive, and data spaces include $H^1(\Omega)$; see [2] and [6]. As a result, a standard L^2 least-squares functional for this system cannot be optimally accurate; see [6].

There are several possibilities to deal with the lack of H^1-coercivity in the first-order system. The first one is to use a standard L^2 functional anyway. A recent analysis of such methods by Chang [13] indicates that, at least for the velocity approximation, one can still achieve optimal rates. However, rates for the vorticity and the pressure approximations are suboptimal. A second approach is to consider weighted least-squares functionals in which the stronger H^1 data norms, dictated by the a priori estimates, are replaced by mesh dependent L^2 norms; see [6]. A third approach, which is the subject of this paper, is to derive a norm equivalent functional starting from an a priori estimate which involves the norm of $H^{-1}(\Omega)$. Such functional does not immediately lend itself to a practical method because the norm $\|\cdot\|_{-1}$ is not computable. The main idea of the *negative norm* least-squares approach is to replace $\|\cdot\|_{-1}$ by a computable discrete equivalent, leading to a *discrete negative norm* functional. Although the new functional is not norm equivalent in the same sense as the original one, it retains norm equivalence for discrete functions, which allows to establish optimal error estimates. Least-squares methods based on such ideas were first proposed by Bramble et. al. in [9] for second order scalar linear elliptic BVP. In the context of the Stokes equations, negative norm methods were first developed and analyzed in [10] using both the velocity-vorticity-pressure and the primitive variable formulations of this problem. In [11] Cai et. al. considered negative norm least-squares methods for a perturbed form of the velocity-vorticity-pressure Stokes equations, which includes as a particular case the equations of linear elasticity. Compared with the Stokes problem, negative norm methods for the nonlinear Navier-Stokes equations are much less developed and studied. To the best of author's knowledge, the only existing such method was proposed and analyzed in [3]. This method is based on the velocity gradient form of the Navier-Stokes equations, and to this end no corresponding methods exist for the velocity-vorticity-pressure first-order system. At the same time, there are several reasons which make extension of the negative norm approach to the former first-order system desirable. Least-squares methods based on the velocity-vorticity-pressure Navier-Stokes equations have been extensively used in the engineering literature; see, e.g., [18]-[21]. However, existing methods are based predominantly on standard L^2, or weighted L^2 least-squares functionals. Although they have demonstrated robustness and efficiency, including in realistic applications, such methods also exhibit some shortcomings. First, as we have mentioned, standard functionals for the velocity-vorticity-pressure system with the velocity boundary condition are not optimal. Second, weighted functionals require more regularity in order to establish the optimal convergence rates. Lastly, in both cases it is not easy to devise efficient preconditioning techniques for the discrete equations. In this context the use of negative norms is potentially very attractive, because corresponding methods are optimally accurate under relaxed regularity assumptions, and the associated discrete problems can be easily preconditioned. Thus, the main goal of this paper is to develop the negative norm least-squares approach for the velocity-vorticity-pressure Navier-Stokes equations.

II. THE FIRST-ORDER SYSTEM AND A PRIORI ESTIMATES

In this section we state the velocity-vorticity-pressure form of the Stokes and Navier-Stokes equations and derive the a priori estimates relevant to the least-squares method. These estimates will be used in order to define a *norm-equivalent* functional. We consider the steady state, incompressible Navier-Stokes equations given by

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

$$\nabla\cdot\mathbf{u} = 0, \quad \text{in } \Omega \tag{2.2}$$

where $\mathbf{u}$, p and $\mathbf{f}$ denote velocity, pressure, and given body force, and ν is the inverse of the Reynolds number. System (2.1)-(2.2) will be considered along with the velocity boundary condition

$$\mathbf{u} = \mathbf{0} \quad \text{on } \Gamma \tag{2.3}$$

and the zero-mean constraint

$$\int_\Omega p\,dx = 0\,. \tag{2.4}$$

Following Jiang et. al. [17] we introduce the vorticity

$$\boldsymbol{\omega} = \nabla\times\mathbf{u} \tag{2.5}$$

as a new dependent variable. Note that in two-dimensions, $\boldsymbol{\omega}$ is a scalar function, whereas in three-dimensions $\boldsymbol{\omega}$ is a vector. Nevertheless, for simplicity we agree to use vector notation in both cases. In view of the vector identities

$$\nabla\times\nabla\times\mathbf{u} = -\triangle\mathbf{u} + \nabla(\nabla\cdot\mathbf{u}) \tag{2.6}$$

and

$$(\mathbf{u}\cdot\nabla)\mathbf{u} = \frac{1}{2}\nabla|\mathbf{u}|^2 - \mathbf{u}\times(\nabla\times\mathbf{u})\,, \tag{2.7}$$

equation (2.1) can be written in the form $\nu\nabla\times\boldsymbol{\omega}+\boldsymbol{\omega}\times\mathbf{u}+\nabla r = \mathbf{f}$ where $r = p+1/2|\mathbf{u}|^2$ is the total head (referred to as "pressure" in the sequel). For the subsequent developments it will be convenient to scale this equation by the Reynolds number $\lambda = 1/\nu$; for simplicity the scaled pressure and body force are denoted again by r and $\mathbf{f}$. As a result, we arrive at the following first-order velocity-vorticity-pressure form of (2.1)-(2.2).

$$\nabla\times\boldsymbol{\omega} + \lambda\boldsymbol{\omega}\times\mathbf{u} + \nabla r = \mathbf{f} \quad \text{in } \Omega \tag{2.8}$$

$$\nabla\cdot\mathbf{u} = 0 \quad \text{in } \Omega \tag{2.9}$$

$$\nabla\times\mathbf{u} - \boldsymbol{\omega} = 0 \quad \text{in } \Omega \tag{2.10}$$

along with (2.3) and (2.4). To obtain the Stokes problem associated with (2.8)-(2.10) we omit the convective term in (2.8):

$$\nabla\times\boldsymbol{\omega} + \nabla r = \mathbf{f} \quad \text{in } \Omega\,, \tag{2.11}$$

i.e., the velocity-vorticity-pressure Stokes equations are given by (2.11), (2.9) and (2.10).

The first a priori estimate relevant to the least-squares method is established in the next theorem. The data for the momentum equation in this estimate is measured in the norm of $\mathbf{H}^{-1}(\Omega)$. There exist various techniques which can be used to establish such an estimate. For example, in [6] this has been accomplished using the elliptic regularity theory of Agmon, Douglis and Nirenberg [1]. For simplicity, here we give a direct proof which differs only in minor details from the proofs in [10] and [11]

Theorem 2.1. *Assume that Ω is a bounded domain in $\mathbf{R}^n$, $n = 2, 3$, with polygonal or polyhedral boundary Γ. Then, there exists a constant C such that for every $\mathcal{U} \equiv (\boldsymbol{\omega}, \mathbf{u}, r) \in \mathbf{L}^2(\Omega) \times \mathbf{H}^1_0(\Omega) \times L^2_0(\Omega)$*

$$\|\boldsymbol{\omega}\|_0 + \|\mathbf{u}\|_1 + \|r\|_0 \leq C\Big(\|\nabla\times\boldsymbol{\omega} + \nabla r\|_{-1} + \|\nabla\cdot\mathbf{u}\|_0 + \|\nabla\times\mathbf{u} - \boldsymbol{\omega}\|_0\Big)\,. \quad (2.12)$$

Proof. We prove (2.12) using a density argument. Let $(\boldsymbol{\omega}, \mathbf{u}, r) \in \mathbf{C}^\infty(\Omega)\times\mathbf{C}^\infty_0(\Omega)\times[C^\infty(\Omega)\cap L^2_0(\Omega)]$. We first estimate the seminorm $|\mathbf{u}|_1$ as follows. From (2.6) and the Cauchy-Schwarz inequality

$$\begin{aligned} |\mathbf{u}|_1^2 &= (\nabla\mathbf{u}, \nabla\mathbf{u}) \\ &= (\nabla\times\boldsymbol{\omega} + \nabla r, \mathbf{u}) + (\nabla\times\mathbf{u} - \boldsymbol{\omega}, \nabla\times\mathbf{u}) + \|\nabla\cdot\mathbf{u}\|_0^2 + (r, \nabla\cdot\mathbf{u}) \\ &\leq \frac{(\nabla\times\boldsymbol{\omega} + \nabla r, \mathbf{u})}{\|\mathbf{u}\|_1}\|\mathbf{u}\|_1 + \|\nabla\times\mathbf{u} - \boldsymbol{\omega}\|_0\|\mathbf{u}\|_1 + \|\nabla\cdot\mathbf{u}\|_0\|\mathbf{u}\|_1 + \|r\|_0\|\nabla\cdot\mathbf{u}\|_0 \\ &\leq \Big(\|\nabla\times\boldsymbol{\omega} + \nabla r\|_{-1} + \|\nabla\times\mathbf{u} - \boldsymbol{\omega}\|_0 + \|\nabla\cdot\mathbf{u}\|_0\Big)\|\mathbf{u}\|_1 + \|r\|_0\|\nabla\cdot\mathbf{u}\|\,. \end{aligned}$$

Using the norm equivalence of $|\cdot|_1$ in $\mathbf{H}^1_0(\Omega)$ and the ε-inequality we obtain the estimate

$$|\mathbf{u}|_1^2 \leq \Big(\|\nabla\times\boldsymbol{\omega} + \nabla r\|_{-1}^2 + \|\nabla\times\mathbf{u} - \boldsymbol{\omega}\|_0^2 + \|\nabla\cdot\mathbf{u}\|_0^2\Big) + 4/3\|r\|_0\|\nabla\cdot\mathbf{u}\|\,. \quad (2.13)$$

To bound $\|r\|_0$, recall that for any $r \in L^2_0(\Omega)$

$$\|r\|_0 \leq C\|\nabla r\|_{-1}\,,$$

see, e.g., Corollary 2.1, p.20 in [15]. To use this estimate, we rewrite $(\nabla r, \mathbf{v})$, where $\mathbf{v}$ is an arbitrary function in $\mathbf{C}^\infty_0(\Omega)$, as follows:

$$\begin{aligned} (\nabla r, \mathbf{v}) &= (\nabla\times\boldsymbol{\omega} + \nabla r, \mathbf{v}) - (\nabla\times\boldsymbol{\omega}, \mathbf{v}) \\ &= (\nabla\times\boldsymbol{\omega} + \nabla r, \mathbf{v}) - (\boldsymbol{\omega} - \nabla\times\mathbf{u}, \nabla\times\mathbf{v}) - (\nabla\times\mathbf{u}, \nabla\times\mathbf{v})\,. \end{aligned}$$

This easily yields the bound

$$\|r\|_0 \leq C\Big(\|\nabla\times\boldsymbol{\omega} + \nabla r\|_{-1} + \|\nabla\times\mathbf{u} - \boldsymbol{\omega}\|_0 + \|\mathbf{u}\|_1\Big)\,. \quad (2.14)$$

Now, (2.13) and (2.14) in combination with the ε-inequality result in the bound

$$|\mathbf{u}|_1^2 \leq C\Big(\|\nabla\times\boldsymbol{\omega} + \nabla r\|_{-1}^2 + \|\nabla\times\mathbf{u} - \boldsymbol{\omega}\|_0^2 + \|\nabla\cdot\mathbf{u}\|_0^2\Big)\,. \quad (2.15)$$

Then, (2.15) and (2.14) yield

$$\|r\|_0^2 \leq C\Big(\|\nabla\times\boldsymbol{\omega} + \nabla r\|_{-1}^2 + \|\nabla\times\mathbf{u} - \boldsymbol{\omega}\|_0^2 + \|\nabla\cdot\mathbf{u}\|_0^2\Big)\,. \quad (2.16)$$

Finally, to estimate $\|\boldsymbol{\omega}\|_0$ observe that

$$\|\boldsymbol{\omega}\|_0^2 \leq \|\nabla \times \mathbf{u} - \boldsymbol{\omega}\|_0 + \|\nabla \times \mathbf{u}\|_0 \,.$$

Using (2.15) we find

$$\|\boldsymbol{\omega}\|_0^2 \leq C\Big(\|\nabla \times \boldsymbol{\omega} + \nabla r\|_{-1}^2 + \|\nabla \times \mathbf{u} - \boldsymbol{\omega}\|_0^2 + \|\nabla \cdot \mathbf{u}\|_0^2\Big) \,. \qquad (2.17)$$

The bound (2.12) now follows from (2.15), (2.16) and (2.17) by density argument. ∎

To state the second a priori estimate, relevant to the least-squares method, we need to make some assumptions concerning regularity and uniqueness of the solutions to (2.8)-(2.10), (2.3). To this end, let

$$\mathbf{X}_m = \mathbf{H}^m(\Omega) \times [\mathbf{H}^{m+1}(\Omega) \cap \mathbf{H}_0^1(\Omega)] \times [H^m(\Omega) \cap L_0^2(\Omega)] \,. \qquad (2.18)$$

In what follows $\tilde{\mathcal{U}} \equiv (\tilde{\boldsymbol{\omega}}, \tilde{\mathbf{u}}, \tilde{r})$ will denote the solution of the velocity-vorticity-pressure Navier-Stokes equations (2.8)-(2.10), (2.3) that is being approximated. We shall assume that $\tilde{\mathcal{U}}$ is a *nonsingular* solution, and that it belongs to $\mathbf{X}_m$ for some integer $m \geq 1$. We recall that (see [15]) $\tilde{\mathcal{U}}$ is a nonsingular solution if and only if the linearized problem

$$\nabla \times \boldsymbol{\omega} + \lambda(\boldsymbol{\omega} \times \tilde{\mathbf{u}} + \tilde{\boldsymbol{\omega}} \times \mathbf{u}) + \nabla r = \mathbf{f} \quad \text{in } \Omega \qquad (2.19)$$

$$\nabla \cdot \mathbf{u} = 0 \quad \text{in } \Omega \qquad (2.20)$$

$$\nabla \times \mathbf{u} - \boldsymbol{\omega} = 0 \quad \text{in } \Omega \qquad (2.21)$$

along with boundary condition (2.3), has a unique solution $\mathcal{U} \equiv (\boldsymbol{\omega}, \mathbf{u}, r)$ for every $\mathbf{f} \in \mathbf{H}^{-1}(\Omega)$, and $\mathbf{f} \mapsto \mathcal{U}$ is continuous from $\mathbf{H}^{-1}(\Omega)$ into $\mathbf{X}_0$. As a result, for any $\mathcal{U} \in \mathbf{X}_0$ we have that

$$\|\boldsymbol{\omega}\|_0 + \|\mathbf{u}\|_1 + \|r\|_0 \leq C\Big(\|\nabla \times \boldsymbol{\omega} + \lambda(\boldsymbol{\omega} \times \tilde{\mathbf{u}} + \tilde{\boldsymbol{\omega}} \times \mathbf{u}) + \nabla r\|_{-1} + \|\nabla \cdot \mathbf{u}\|_0 + \|\nabla \times \mathbf{u} - \boldsymbol{\omega}\|_0\Big) \,. \qquad (2.22)$$

III. DISCRETE NEGATIVE NORM

In view of (2.12) and (2.22), a norm equivalent functional for the Navier-Stokes equations can be defined as

$$\mathcal{J}_{-1}(\mathcal{U}) = \frac{1}{2}\Big(\|\nabla \times \boldsymbol{\omega} + \lambda\boldsymbol{\omega} \times \mathbf{u} + \nabla r - \mathbf{f}\|_{-1}^2 + \|\nabla \cdot \mathbf{u}\|_0^2 + \|\nabla \times \mathbf{u} - \boldsymbol{\omega}\|_0^2\Big) \,. \quad (3.1)$$

However, this functional is not practical because the negative norm (1.1) is not computable. Following [9]-[10] we shall introduce a computable discrete negative analogue of (1.1). This will help us to define a practical analogue of the least-squares functional (3.1). Although the new functional will not be norm-equivalent in the same spaces as (3.1) it will, nevertheless, retain a similar property for discrete functions. To define the space of discrete functions that will be used in the sequel, let $\mathcal{T}_h$ denote a uniformly regular triangulation (see [14]) of Ω into finite elements where, as usual, h is some measure of the grid size. To approximate the solution $\tilde{\mathcal{U}}$ of the Navier-Stokes equations we introduce a discrete space $\mathbf{X}_h = \mathbf{W}_h \times \mathbf{U}_h \times \mathbf{R}_h$, where $\mathbf{W}_h$, $\mathbf{U}_h$ and $\mathbf{R}_h$ are finite element spaces used to approximate the vorticity, velocity and the pressure, respectively. Concerning the space $\mathbf{X}_h$ we make the following assumptions.

A.1 $\mathbf{X}_h$ is conforming in the sense that

$$\mathbf{X}_h \subset \mathbf{X}_0 \,, \tag{3.2}$$

and

$$\|\mathcal{U}_h\|_{\mathbf{X}_h} = \|\mathcal{U}_h\|_{\mathbf{X}_0} = \|\boldsymbol{\omega}_h\|_0 + \|\mathbf{u}_h\|_1 + \|r_h\|_0, \quad \forall \mathcal{U}_h \in \mathbf{X}_h \,.$$

A.2 There exists an integer $d \geq 1$ such that for every $\mathcal{U} \equiv (\boldsymbol{\omega}, \mathbf{u}, r) \in \mathbf{X}_d$ there exists $\mathcal{U}_h \equiv (\boldsymbol{\omega}_h, \mathbf{u}_h, r_h) \in \mathbf{X}_h$ such that

$$\begin{aligned} \|\boldsymbol{\omega} - \boldsymbol{\omega}_h\|_0 + h\|\boldsymbol{\omega} - \boldsymbol{\omega}_h\|_1 &\leq h^d C\|\boldsymbol{\omega}\|_d \,; && (3.3)\\ \|\mathbf{u} - \mathbf{u}_h\|_0 + h\|\mathbf{u} - \mathbf{u}_h\|_1 &\leq h^{d+1} C\|\mathbf{u}\|_{d+1} \,; && (3.4)\\ \|r - r_h\|_0 + h\|r - r_h\|_1 &\leq h^d C\|r\|_d \,. && (3.5) \end{aligned}$$

A.3 An inverse inequality of the form

$$\|u_h\|_1 \leq C\frac{1}{h}\|u_h\|_0 \tag{3.6}$$

holds for all components of $\mathbf{X}_h$.

A space $\mathbf{X}_h$ which satisfies **A.1 - A.3** can be defined as a direct product of standard piecewise polynomial finite element spaces defined with respect to the same triangulation $\mathcal{T}_h$. For an example, let $\mathcal{T}_h = \cup_{i=1}^M \mathcal{K}_i$ denote a uniformly regular triangulation of Ω into $n-$simplices $\mathcal{K}$, and consider the finite element space of continuous, piecewise polynomial functions

$$P_k = \{u_h \in C^0(\Omega) \mid u_{h|_\mathcal{K}} \in \mathcal{P}_\mathcal{K}^k\} \,, \tag{3.7}$$

where $\mathcal{P}_\mathcal{K}^k$ is the set of all polynomials of degree less than or equal k defined on $\mathcal{K}$. Then, $\mathbf{W}_h$ and $\mathbf{R}_h$ can be defined using the space P_{d-1}, and $\mathbf{U}_h$ can be defined using the space P_d. Since approximation spaces are not subject to any stability conditions, for the simplicity of implementation one can also use spaces of the same order for all unknowns.

Next, we introduce a discrete equivalent of (1.1). In this we follow with minor modifications the approach of [9]-[10]. For the convenience of the reader all essential details and proofs are summarized below. Let $\mathbf{S}^h$ denote solution operator for the weak Dirichlet problem

seek $\mathbf{u}_h \in \mathbf{U}_h$ *such that*

$$\int_\Omega \nabla \mathbf{u}_h : \nabla \mathbf{v}_h dx = \int_\Omega \mathbf{f} \cdot \mathbf{v}_h dx, \quad \forall \mathbf{v} \in \mathbf{U}_h \,. \tag{3.8}$$

It is not difficult to show that (see [9])

$$(\mathbf{S}^h \mathbf{f}, \mathbf{f}) \leq C\|\mathbf{f}\|_{-1}^2, \quad \forall \mathbf{f} \in \mathbf{H}^{-1}(\Omega) \,. \tag{3.9}$$

Let $\mathbf{B}^h : \mathbf{H}^{-1}(\Omega) \mapsto \mathbf{U}_h$ denote a preconditioner for (3.8), i.e., a symmetric and positive semidefinite operator that is spectrally equivalent to $\mathbf{S}^h$ in the sense that

$$C_1(\mathbf{S}^h \mathbf{v}, \mathbf{v}) \leq (\mathbf{B}^h \mathbf{v}, \mathbf{v}) \leq C_2(\mathbf{S}^h \mathbf{v}, \mathbf{v}), \quad \forall \mathbf{v} \in \mathbf{L}^2(\Omega) \,. \tag{3.10}$$

Although in most cases $\mathbf{B}^h$ can be identified with a matrix, we prefer to think of it as a "black-box" type algorithm for (3.8). The discrete equivalent of (1.1) is then defined as

$$\|\mathbf{v}\|_{-h} = ((h^2\mathbf{I} + \mathbf{B}^h)\mathbf{v}, \mathbf{v})^{1/2}, \quad \forall \mathbf{v} \in \mathbf{L}^2(\Omega) \,. \tag{3.11}$$

Evidently, the inner product associated with (3.11) is given by

$$(\mathbf{u},\mathbf{v})_{-h} = ((h^2\mathbf{I}+\mathbf{B}^h)\mathbf{u},\mathbf{v})^{1/2}, \quad \forall \mathbf{u},\mathbf{v}\in \mathbf{L}^2(\Omega)\,. \tag{3.12}$$

Next lemma establishes a fundamental equivalence property of (3.11).

Lemma 3.1. *For any* $\mathbf{u}\in\mathbf{L}^2(\Omega)$

$$C_1\|\mathbf{u}\|_{-1} \le \|\mathbf{u}\|_{-h} \le C_2(h\|\mathbf{u}\|_0 + \|\mathbf{u}\|_{-1})\,. \tag{3.13}$$

Proof. The upper bound follows from the definition (3.11), (3.10) and (3.9)

$$\begin{aligned}\|\mathbf{u}\|_{-h} &= ((h^2\mathbf{I}+\mathbf{B}^h)\mathbf{u},\mathbf{u})^{1/2}\\ &= (h^2\|\mathbf{u}\|_0^2 + (\mathbf{B}^h\mathbf{u},\mathbf{u}))^{1/2}\\ &\le C(h^2\|\mathbf{u}\|_0^2 + (\mathbf{S}^h\mathbf{u},\mathbf{u}))^{1/2}\\ &\le C(h^2\|\mathbf{u}\|_0^2 + \|\mathbf{u}\|_{-1}^2)^{1/2}\,.\end{aligned}$$

To show the lower bound, let $\mathbf{v}\in\mathbf{H}_0^1(\Omega)$ be an arbitrary function. From (3.4) it follows that there exists $\mathbf{v}_h\in\mathbf{U}_h$ such that

$$\|\mathbf{v}-\mathbf{v}_h\|_1 \le Ch\|\mathbf{v}\|_1 \quad \text{and} \quad \|\mathbf{v}_h\|_1 \le C\|\mathbf{v}\|_1\,. \tag{3.14}$$

Then,

$$\begin{aligned}(\mathbf{u},\mathbf{v}) &= (\mathbf{u},\mathbf{v}-\mathbf{v}_h) + (\mathbf{u},\mathbf{v}_h)\\ &\le \|\mathbf{u}\|_0\|\mathbf{v}-\mathbf{v}_h\|_0 + (\mathbf{u},\mathbf{v}_h)\\ &\le Ch\|\mathbf{u}\|_0\|\mathbf{v}\|_1 + (\mathbf{u},\mathbf{v}_h)\,.\end{aligned}$$

From the second inequality in (3.14) it now follows that

$$\frac{(\mathbf{u},\mathbf{v})}{\|\mathbf{v}\|_1} \le C\Big(h\|\mathbf{u}\|_0 + \frac{(\mathbf{u},\mathbf{v}_h)}{\|\mathbf{v}_h\|_1}\Big)\,,$$

which yields the upper bound in (3.13). ■

IV. NEGATIVE NORM LEAST-SQUARES METHOD

In this section we define the negative norm least-squares method using a discrete counterpart of (3.1). For this purpose, the norm $\|\cdot\|_{-1}$ in (3.1) is replaced by the discrete equivalent $\|\cdot\|_{-h}$, i.e., we consider a least-squares functional given by

$$\mathcal{J}_{-h}(\mathcal{U}) = \frac{1}{2}\Big(\|\nabla\times\boldsymbol{\omega} + \lambda\boldsymbol{\omega}\times\mathbf{u} + \nabla r - \mathbf{f}\|_{-h}^2 + \|\nabla\cdot\mathbf{u}\|_0^2 + \|\nabla\times\mathbf{u}-\boldsymbol{\omega}\|_0^2\Big)\,. \tag{4.1}$$

Then, approximations to the solution $\tilde{\mathcal{U}}$ of the Navier-Stokes equations (2.8)-(2.10), and (2.3) are obtained by finding the minimizer of (4.1) out of $\mathbf{X}_h$. This minimizer is subject to a necessary condition (Euler-Lagrange equation) given by

seek $\mathcal{U}_h\in\mathbf{X}_h$ *such that*

$$\mathcal{B}(\mathcal{U}_h,\mathcal{V}_h) \;=\; \Big(\nabla\times\boldsymbol{\omega}_h+\nabla r_h+\lambda\boldsymbol{\omega}_h\times\mathbf{u}_h-\mathbf{f},\qquad (4.2)$$
$$\nabla\times\boldsymbol{\xi}_h+\nabla q_h+\lambda(\boldsymbol{\omega}_h\times\mathbf{v}_h+\boldsymbol{\xi}_h\times\mathbf{u}_h)\Big)_{-h}$$
$$+\;(\nabla\cdot\mathbf{u}_h,\nabla\cdot\mathbf{v}_h)+(\nabla\times\mathbf{u}_h-\boldsymbol{\omega}_h,\nabla\times\mathbf{v}_h-\boldsymbol{\xi}_h)=0,\qquad \forall\mathcal{V}\in\mathbf{X}_h\,.$$

Problem (4.2) constitutes a nonlinear system of algebraic equations, which can be easily seen by, e.g., testing in (4.2) with the standard nodal basis of $\mathbf{X}_h$. In what follows, our attention will be focused on showing that (4.2) is a well-posed problem, i.e., that under certain assumptions it has a unique solution, and that this solution is close to the solution $\tilde{\mathcal{U}}$ of the Navier-Stokes equations. For this purpose let us write the nonlinear system (4.2) in the form

$$F(\mathcal{U}_h) \;=\; 0. \qquad (4.3)$$

To show that (4.3) is well-posed we employ a version of the implicit function theorem which can be found in [15]. Specialized to our needs this theorem is as follows.

Theorem 4.1. *Assume that there exists a function $\tilde{\mathcal{U}}_h\in\mathbf{X}_h$, such that $F'(\tilde{\mathcal{U}}_h)$ is an isomorphism of $\mathbf{X}_h$ onto itself. Let*

$$\varepsilon=\|F(\tilde{\mathcal{U}}_h)\|_{\mathbf{X}_h}\,; \qquad (4.4)$$

$$\gamma=\|F'(\tilde{\mathcal{U}}_h)^{-1}\|_{L(\mathbf{X}_h,\mathbf{X}_h)}; \qquad (4.5)$$

and

$$L(\alpha)=\sup_{\mathcal{V}_h\in B(\tilde{\mathcal{U}}_h,\alpha)}\|F'(\tilde{\mathcal{U}}_h)-F'(\mathcal{V}_h)\|_{L(\mathbf{X}_h,\mathbf{X}_h)}, \qquad (4.6)$$

where

$$B(\tilde{\mathcal{U}}_h,\alpha)=\{\mathcal{V}_h\in\mathbf{X}_h\,|\,\|\tilde{\mathcal{U}}_h-\mathcal{V}_h\|_{\mathbf{X}_h}\le\alpha\}.$$

Furthermore, assume that

$$2\gamma L(2\gamma\varepsilon)<1\,. \qquad (4.7)$$

Then, the problem $F(\mathcal{U}_h)=0$ has a solution $\mathcal{U}_h\in\mathbf{X}_h$ which belongs to the ball $B(\tilde{\mathcal{U}}_h,2\gamma\varepsilon)$, $F'(\mathcal{U}_h)$ is an isomorphism of $\mathbf{X}_h$ onto itself, and $\|F'(\mathcal{U}_h)^{-1}\|_{\mathbf{X}_h}\le 2\gamma$. Furthermore, $\mathcal{U}_h$ is the only solution in the ball $B(\tilde{\mathcal{U}}_h,\alpha)$ whose radius α satisfies $\gamma L(\alpha)<1$, and we have the error estimate

$$\|\mathcal{U}_h-\mathcal{V}_h\|_{\mathbf{X}_h}\le\frac{\gamma}{1-\gamma L(\alpha)}\|F(\mathcal{V}_h)\|_{\mathbf{X}_h},\quad \forall\mathcal{V}_h\in B(\tilde{\mathcal{U}}_h,\alpha)\,. \qquad (4.8)$$

For the proof of this theorem, which uses a fixed-point argument, we refer the reader to [15].

V. ANALYSIS OF THE NEGATIVE NORM METHOD

In this section we show that all hypotheses of Theorem 4.1 are valid, provided the solution $\tilde{\mathcal{U}}$ of (2.8)-(2.10) is as described in §II., i.e., $\tilde{\mathcal{U}}$ is nonsingular and belongs to $\mathbf{X}_m$ for some integer $m \geq 1$. The asymptotic error estimates will then follow from (4.8). For the application of Theorem 4.1 it is of critical importance to assure existence of $\tilde{\mathcal{U}}_h \in \mathbf{X}_h$ such that $F'(\tilde{\mathcal{U}}_h)$ is invertible. It is reasonable to expect that a close approximation of $\tilde{\mathcal{U}}$ out of $\mathbf{X}_h$ will suit this purpose. According to (3.3)-(3.4), there exists $\tilde{\mathcal{U}}_h = (\tilde{\boldsymbol{\omega}}_h, \tilde{\mathbf{u}}_h, \tilde{r}_h)$ such that

$$\|\tilde{\boldsymbol{\omega}} - \tilde{\boldsymbol{\omega}}_h\|_0 + h\|\tilde{\boldsymbol{\omega}} - \tilde{\boldsymbol{\omega}}_h\|_1 \;\leq\; h^{\tilde{d}} C \|\tilde{\boldsymbol{\omega}}\|_{\tilde{d}}, \tag{5.1}$$

$$\|\tilde{\mathbf{u}} - \tilde{\mathbf{u}}_h\|_0 + h\|\tilde{\mathbf{u}} - \tilde{\mathbf{u}}_h\|_1 \;\leq\; h^{\tilde{d}+1} C \|\tilde{\mathbf{u}}\|_{\tilde{d}+1}, \tag{5.2}$$

$$\|\tilde{r} - \tilde{r}_h\|_0 + h\|\tilde{r} - \tilde{r}_h\|_1 \;\leq\; h^{\tilde{d}} C \|\tilde{r}\|_{\tilde{d}}, \tag{5.3}$$

where $\tilde{d} = \min\{d, m\}$. Note that (5.1)-(5.3) also imply

$$\|\tilde{\boldsymbol{\omega}}_h\|_0 \leq C\|\tilde{\boldsymbol{\omega}}\|_{\tilde{d}}, \quad \|\tilde{r}_h\|_0 \leq C\|\tilde{r}\|_{\tilde{d}}, \quad \text{and} \quad \|\tilde{\mathbf{u}}_h\|_1 \leq C\|\tilde{\mathbf{u}}\|_{\tilde{d}+1}. \tag{5.4}$$

Below we shall see that, for a sufficiently small h, this choice of $\tilde{\mathcal{U}}_h$ guarantees the existence of $F'(\tilde{\mathcal{U}}_h)^{-1}$.

We begin by separating form $\mathcal{B}(\cdot,\cdot)$ into a linear (Stokes) part

$$\begin{aligned}\mathcal{B}_S(\mathcal{U}, \mathcal{V}) = (\nabla \times \boldsymbol{\omega} + \nabla r, \nabla \times \boldsymbol{\xi} + \nabla q)_{-h} \\ + (\nabla \cdot \mathbf{u}, \nabla \cdot \mathbf{v}) + (\nabla \times \mathbf{u} - \boldsymbol{\omega}, \nabla \times \mathbf{v} - \boldsymbol{\xi}),\end{aligned} \tag{5.5}$$

and a nonlinear part

$$\begin{aligned}\mathcal{B}_G(\mathcal{U}, \mathcal{V}) = (\lambda \boldsymbol{\omega} \times \mathbf{u} - \mathbf{f}, \nabla \times \boldsymbol{\xi} + \nabla q + \lambda(\boldsymbol{\omega} \times \mathbf{v} + \boldsymbol{\xi} \times \mathbf{u}))_{-h} \\ + (\nabla \times \boldsymbol{\omega} + \nabla r, \lambda(\boldsymbol{\omega} \times \mathbf{v} + \boldsymbol{\xi} \times \mathbf{u}))_{-h}.\end{aligned} \tag{5.6}$$

It is not difficult to see that the Fréchet derivative of (5.6) is given by

$$\begin{aligned}\mathcal{B}'_G[\hat{\mathcal{U}}](\mathcal{U}, \mathcal{V}) \;=\; & (\nabla \times \hat{\boldsymbol{\omega}} + \nabla \hat{r} + \lambda \hat{\boldsymbol{\omega}} \times \hat{\mathbf{u}} - \mathbf{f}, \lambda(\boldsymbol{\omega} \times \mathbf{v} + \boldsymbol{\xi} \times \mathbf{u})_{-h} \\ + \; & \lambda(\nabla \times \boldsymbol{\omega} + \nabla r, \hat{\boldsymbol{\omega}} \times \mathbf{v} + \boldsymbol{\xi} \times \hat{\mathbf{u}})_{-h} \\ + \; & \lambda(\hat{\boldsymbol{\omega}} \times \mathbf{u} + \boldsymbol{\omega} \times \hat{\mathbf{u}}, \nabla \times \boldsymbol{\xi} + \nabla q + \lambda \hat{\boldsymbol{\omega}} \times \mathbf{v} + \boldsymbol{\xi} \times \hat{\mathbf{u}})_{-h}.\end{aligned} \tag{5.7}$$

As a result, form $\mathcal{B}(\cdot,\cdot)$ and its Fréchet derivative $\mathcal{B}'[\hat{\mathcal{U}}](\cdot,\cdot)$ can be written as

$$\mathcal{B}(\mathcal{U}, \mathcal{V}) = \mathcal{B}_S(\mathcal{U}, \mathcal{V}) + \mathcal{B}_G(\mathcal{U}, \mathcal{V})$$

and

$$\mathcal{B}'[\hat{\mathcal{U}}](\mathcal{U}, \mathcal{V}) = \mathcal{B}_S(\mathcal{U}, \mathcal{V}) + \mathcal{B}'_G[\hat{\mathcal{U}}](\mathcal{U}, \mathcal{V}), \tag{5.8}$$

respectively. Next we use these identities to effect a decomposition of the nonlinear function F in (4.3), and its Fréchet derivative into a linear and a nonlinear part. Let

$$\mathbf{Y} = \mathbf{L}^2(\Omega) \times \mathbf{H}^{-1}(\Omega) \times L_0^2(\Omega)$$

and consider the linear operator $T : \mathbf{Y} \mapsto \mathbf{X}_h$, defined by

$$\mathcal{U}_h = T\mathbf{g} \text{ if and only if } \mathcal{B}_S(\mathcal{U}_h, \mathcal{V}_h) = (\mathbf{g}, \mathcal{V}_h), \text{ for all } \mathcal{V}_h \in \mathbf{X}_h.$$

Using (5.6) we then introduce the nonlinear operator $G : \mathbf{X}_h \mapsto \mathbf{Y}$, defined by

$$\mathbf{g} = G(\mathcal{U}_h) \text{ if and only if } \mathcal{B}_G(\mathcal{U}_h, \mathcal{V}_h) = (\mathbf{g}, \mathcal{V}_h), \text{ for all } \mathcal{V}_h \in \mathbf{X}_h,$$

and its Fréchet derivative $G'(\hat{\mathcal{U}}) : \mathbf{X}_h \mapsto \mathbf{Y}$, defined by

$$\mathbf{g} = G'(\hat{\mathcal{U}}) \cdot \mathcal{U}_h \text{ if and only if } \mathcal{B}'_G[\hat{\mathcal{U}}](\mathcal{U}_h, \mathcal{V}_h) = (\mathbf{g}, \mathcal{V}_h), \text{ for all } \mathcal{V}_h \in \mathbf{X}_h.$$

With the above definitions, problem (4.3) can be written as

$$F(\mathcal{U}_h) \equiv U_h + T \cdot G(\mathcal{U}_h) = 0\,. \tag{5.9}$$

Similarly, the Fréchet derivative of F takes the form

$$F'(\hat{\mathcal{U}}) \cdot \mathcal{U}_h \equiv (\mathbf{I} + T \cdot G'(\hat{\mathcal{U}})) \cdot \mathcal{U}_h\,. \tag{5.10}$$

A. Coercivity estimates

In this subsection we show that the Stokes form $\mathcal{B}_S(\cdot,\cdot)$ and the Fréchet derivative form $\mathcal{B}'[\tilde{\mathcal{U}}_h](\cdot,\cdot)$ are coercive on $\mathbf{X}_h \times \mathbf{X}_h$. Essentially, this means that the discrete negative norm least-squares functional (4.1) is norm-equivalent on $\mathbf{X}_h$. In particular, coercivity of $\mathcal{B}_S(\cdot,\cdot)$ is a consequence of a priori estimate (2.12) and equivalence relation (3.13), whereas coercivity of $\mathcal{B}'[\tilde{\mathcal{U}}_h](\cdot,\cdot)$ follows from the nonsingularity assumption on $\tilde{\mathcal{U}}$, i.e., (3.13) and estimate (2.22). The subsection concludes with a Lipschitz-type estimate for the Frechet derivative F'. Our first lemma summarizes several technical results that will be frequently used in the sequel.

Lemma 5.1. *For $\boldsymbol{\omega} \in \mathbf{H}^1(\Omega)$, $\mathbf{u} \in \mathbf{H}^1(\Omega)$ and $r \in H^1(\Omega)$*

$$\|\nabla \times \boldsymbol{\omega}\|_{-h} \le C(h\|\nabla \times \boldsymbol{\omega}\|_0 + \|\boldsymbol{\omega}\|_0), \tag{5.11}$$

$$\|\nabla r\|_{-h} \le C(h\|\nabla r\|_0 + \|r\|_0), \tag{5.12}$$

$$\|\boldsymbol{\omega} \times \mathbf{u}\|_{-h} \le C(h\|\boldsymbol{\omega}\|_1 + \|\boldsymbol{\omega}\|_0)\|\mathbf{u}\|_1\,. \tag{5.13}$$

For discrete functions $\boldsymbol{\omega}_h \in \mathbf{W}_h$, $\mathbf{u}_h \in \mathbf{U}_h$ and $r_h \in \mathbf{R}_h$ (5.11)-(5.13) specialize further to

$$\|\nabla \times \boldsymbol{\omega}_h\|_{-h} \le C\|\boldsymbol{\omega}_h\|_0, \tag{5.14}$$

$$\|\nabla r_h\|_{-h} \le C\|r_h\|_0, \tag{5.15}$$

$$\|\boldsymbol{\omega}_h \times \mathbf{u}_h\|_{-h} \le C\|\boldsymbol{\omega}_h\|_0\|\mathbf{u}_h\|_1\,. \tag{5.16}$$

Proof. From (3.13)

$$\|\nabla \times \boldsymbol{\omega}\|_{-h} \le C(h\|\nabla \times \boldsymbol{\omega}\|_0 + \|\nabla \times \boldsymbol{\omega}\|_{-1})\,,$$

$$\|\nabla r\|_{-h} \le C(h\|\nabla r\|_0 + \|\nabla r\|_{-1}),$$

and

$$\|\boldsymbol{\omega} \times \mathbf{u}\|_{-h} \le C(h\|\boldsymbol{\omega} \times \mathbf{u}\|_0 + \|\boldsymbol{\omega} \times \mathbf{u}\|_{-1})\,.$$

Bounds (5.11), (5.12) and (5.13) now follow from the inequalities

$$\|\nabla \times \boldsymbol{\omega}\|_{-1} \le \|\boldsymbol{\omega}\|_0; \quad \|\nabla r\|_{-1} \le \|r\|_0\,,$$

and

$$\|\boldsymbol{\omega}\times\mathbf{u}\|_0 \le C\|\boldsymbol{\omega}\|_1\|\mathbf{u}\|_1; \quad \|\boldsymbol{\omega}\times\mathbf{u}\|_{-1} \le C\|\boldsymbol{\omega}\|_0\|\mathbf{u}\|_1 .$$

Remaining bounds (5.14)-(5.16) follow from (5.11)-(5.13) and the inverse inequality (3.6). ∎

Lemma 5.2. *The form $\mathcal{B}_S(\cdot,\cdot)$ is continuous and coercive on $\mathbf{X}_h\times\mathbf{X}_h$.*

Proof. Using (3.13) and (2.12)

$$\begin{aligned}\mathcal{B}_S(\mathcal{U}_h,\mathcal{U}_h) &= \|\nabla\times\boldsymbol{\omega}_h+\nabla r_h\|_{-h}^2+\|\nabla\cdot\mathbf{u}_h\|_0^2+\|\nabla\times\mathbf{u}_h-\boldsymbol{\omega}_h\|_0^2\\ &\ge C\Big(\|\nabla\times\boldsymbol{\omega}_h+\nabla r_h\|_{-1}^2+\|\nabla\cdot\mathbf{u}_h\|_0^2+\|\nabla\times\mathbf{u}_h-\boldsymbol{\omega}_h\|_0^2\Big)\\ &\ge C(\|\boldsymbol{\omega}_h\|_0^2+\|\mathbf{u}_h\|_1^2+\|r_h\|_0^2),\end{aligned}$$

which establishes coercivity of (5.5). Continuity follows directly using Cauchy inequality and (5.14)-(5.15). ∎

Lemma 5.3. *Assume that $\tilde{\mathcal{U}}_h$ satisfies (5.1)-(5.3). Then $\mathcal{B}(\tilde{\mathcal{U}}_h,\cdot)$ defines a continuous linear functional $\mathbf{X}_h\mapsto\mathbf{R}$, and*

$$\mathcal{B}(\tilde{\mathcal{U}}_h,\mathcal{V}_h)\le Ch^{\tilde d}\|\mathcal{V}_h\|_{\mathbf{X}_h}, \quad \forall\mathcal{V}_h\in\mathbf{X}_h . \tag{5.17}$$

Proof. From the definition of $\mathcal{B}(\cdot,\cdot)$ it is easy to see that, when the first argument is fixed, $\mathcal{B}(\tilde{\mathcal{U}}_h,\cdot)$ is linear with respect to the second argument. To establish (5.17), we first use the Cauchy inequality

$$\begin{aligned}&\mathcal{B}(\tilde{\mathcal{U}}_h,\mathcal{V}_h)\le\\ &\quad \|\nabla\times\tilde{\boldsymbol{\omega}}_h+\nabla\tilde r_h+\lambda\tilde{\boldsymbol{\omega}}_h\times\tilde{\mathbf{u}}_h-\mathbf{f}\|_{-h}\|\nabla\times\boldsymbol{\xi}_h+\nabla q_h+\lambda(\tilde{\boldsymbol{\omega}}_h\times\mathbf{v}_h+\boldsymbol{\xi}_h\times\tilde{\mathbf{u}}_h)\|_{-h}\\ &+\ \|\nabla\cdot\tilde{\mathbf{u}}_h\|_0\|\nabla\cdot\mathbf{v}_h\|_0+\|\nabla\times\tilde{\mathbf{u}}_h-\tilde{\boldsymbol{\omega}}_h\|_0\|\nabla\times\mathbf{v}_h-\boldsymbol{\xi}_h\|_0 .\end{aligned}$$

To estimate the last two terms we use (5.1), (5.2), and the fact that $\tilde{\mathcal{U}}$ solves (2.8)-(2.10):

$$\|\nabla\cdot\tilde{\mathbf{u}}_h\|_0\|\nabla\cdot\mathbf{v}_h\|_0\le Ch^{\tilde d}\|\tilde{\mathbf{u}}\|_{\tilde d+1}\|\mathcal{V}_h\|_{\mathbf{X}_h}, \tag{5.18}$$

$$\|\nabla\times\tilde{\mathbf{u}}_h-\tilde{\boldsymbol{\omega}}_h\|_0\|\nabla\times\mathbf{v}_h-\boldsymbol{\xi}_h\|_0\le Ch^{\tilde d}(\|\tilde{\mathbf{u}}\|_{\tilde d+1}+\|\tilde{\boldsymbol{\omega}}\|_{\tilde d})\|\mathcal{V}_h\|_{\mathbf{X}_h} . \tag{5.19}$$

Next we proceed with an estimate for the first term. Using again $\tilde{\mathcal{U}}$ and the triangle inequality,

$$\begin{aligned}\|\nabla\times\tilde{\boldsymbol{\omega}}_h+\nabla\tilde r_h+\lambda\tilde{\boldsymbol{\omega}}_h\times\tilde{\mathbf{u}}_h-\mathbf{f}\|_{-h}&\le\|\nabla\times(\tilde{\boldsymbol{\omega}}_h-\tilde{\boldsymbol{\omega}})\|_{-h}+\|\nabla(\tilde r_h-\tilde r)\|_{-h}\\ &+\|\tilde{\boldsymbol{\omega}}_h\times(\tilde{\mathbf{u}}_h-\tilde{\mathbf{u}})\|_{-h}+\|(\tilde{\boldsymbol{\omega}}_h-\tilde{\boldsymbol{\omega}})\times\tilde{\mathbf{u}}\|_{-h} .\end{aligned}$$

From (5.11) and (5.1) it follows that

$$\|\nabla\times(\tilde{\boldsymbol{\omega}}_h-\tilde{\boldsymbol{\omega}})\|_{-h}\le C(h\|\tilde{\boldsymbol{\omega}}_h-\tilde{\boldsymbol{\omega}}\|_1+\|\tilde{\boldsymbol{\omega}}_h-\tilde{\boldsymbol{\omega}}\|_0)\le Ch^{\tilde d}\|\tilde{\boldsymbol{\omega}}\|_{\tilde d} ,$$

from (5.12) and (5.3):

$$\|\nabla(\tilde r_h-\tilde r)\|_{-h}\le C(h\|\tilde r_h-\tilde r\|_1+\|\tilde r_h-\tilde r\|_0)\le Ch^{\tilde d}\|\tilde r\|_{\tilde d} ,$$

and from (5.13), (5.1), (5.2), and (5.4):

$$\|\tilde{\boldsymbol{\omega}}_h \times (\tilde{\mathbf{u}}_h - \tilde{\mathbf{u}})\|_{-h} + \|(\tilde{\boldsymbol{\omega}}_h - \tilde{\boldsymbol{\omega}}) \times \tilde{\mathbf{u}}\|_{-h} \le C h^{\tilde{d}} \|\tilde{\boldsymbol{\omega}}\|_{\tilde{d}} \|\tilde{\mathbf{u}}\|_{\tilde{d}+1} .$$

Using the triangle inequality for the remaining part of the first term,

$$\begin{aligned} &\|\nabla \times \boldsymbol{\xi}_h + \nabla q_h + \lambda(\tilde{\boldsymbol{\omega}}_h \times \mathbf{v}_h + \boldsymbol{\xi}_h \times \tilde{\mathbf{u}}_h)\|_{-h} \\ &\qquad \le \|\nabla \times \boldsymbol{\xi}_h\|_{-h} + \|\nabla q_h\|_{-h} + \lambda\Big(\|\tilde{\boldsymbol{\omega}}_h \times \mathbf{v}_h\|_{-h} + \|\boldsymbol{\xi}_h \times \tilde{\mathbf{u}}_h\|_{-h}\Big) . \end{aligned}$$

Then, terms in the right hand side are bounded using (5.14), (5.15), (5.16), and (5.4):

$$\|\nabla \times \boldsymbol{\xi}_h\|_{-h} \le C\|\boldsymbol{\xi}_h\|_0 ,$$

$$\|\nabla q_h\|_{-h} \le C\|q_h\|_0 ,$$

$$\|\tilde{\boldsymbol{\omega}}_h \times \mathbf{v}_h\|_{-h} \le C\|\tilde{\boldsymbol{\omega}}\|_{\tilde{d}}\|\mathbf{v}_h\|_1 ,$$

$$\|\boldsymbol{\xi}_h \times \tilde{\mathbf{u}}_h\|_{-h} \le C\|\boldsymbol{\xi}_h\|_0\|\tilde{\mathbf{u}}\|_{\tilde{d}+1} .$$

Combining the last four inequalities yields

$$\begin{aligned} &\|\nabla \times \tilde{\boldsymbol{\omega}}_h + \nabla \tilde{r}_h + \lambda\tilde{\boldsymbol{\omega}}_h \times \tilde{\mathbf{u}}_h - \mathbf{f}\|_{-h}\|\nabla \times \boldsymbol{\xi}_h + \nabla q_h + \lambda(\tilde{\boldsymbol{\omega}}_h \times \mathbf{v}_h + \boldsymbol{\xi}_h \times \tilde{\mathbf{u}}_h)\|_{-h} \\ &\qquad \le C h^{\tilde{d}} \|\mathcal{V}_h\|_{\mathbf{X}_h} . \end{aligned} \tag{5.20}$$

The lemma now follows from (5.18)-(5.20). ■

Lemma 5.4. *There exists a number $h_0 > 0$ such that, for all $h < h_0$, form $\mathcal{B}'[\tilde{\mathcal{U}}_h](\cdot,\cdot)$ is continuous and coercive on $\mathbf{X}_h \times \mathbf{X}_h$.*

Proof. From (5.8)

$$\begin{aligned} \mathcal{B}'[\tilde{\mathcal{U}}_h](\mathcal{U}_h,\mathcal{U}_h) &= \mathcal{B}_S(\mathcal{U}_h,\mathcal{U}_h) + \mathcal{B}'_G[\tilde{\mathcal{U}}_h](\mathcal{U}_h,\mathcal{U}_h) \\ &= \|\nabla \times \boldsymbol{\omega}_h + \nabla r_h + \lambda(\tilde{\boldsymbol{\omega}}_h \times \mathbf{u}_h + \boldsymbol{\omega}_h \times \tilde{\mathbf{u}}_h)\|_{-h}^2 \\ &\quad + \|\nabla \cdot \mathbf{u}_h\|_0^2 + \|\nabla \times \mathbf{u}_h - \boldsymbol{\omega}_h\|_0^2 \\ &\quad + 2\lambda\Big(\nabla \times \tilde{\boldsymbol{\omega}}_h + \nabla \tilde{r}_h + \lambda\tilde{\boldsymbol{\omega}}_h \times \tilde{\mathbf{u}}_h - \mathbf{f}, \boldsymbol{\omega}_h \times \mathbf{u}_h\Big) . \end{aligned} \tag{5.21}$$

To obtain a lower bound for the first term in (5.21) we use the solution $\tilde{\mathcal{U}}$ of the Navier-Stokes equations (2.8)-(2.10):

$$\begin{aligned} &\|\nabla \times \boldsymbol{\omega}_h + \nabla r_h + \lambda(\tilde{\boldsymbol{\omega}}_h \times \mathbf{u}_h + \boldsymbol{\omega}_h \times \tilde{\mathbf{u}}_h)\|_{-h} \\ &\ge \|\nabla \times \boldsymbol{\omega}_h + \nabla r_h + \lambda(\tilde{\boldsymbol{\omega}} \times \mathbf{u}_h + \boldsymbol{\omega}_h \times \tilde{\mathbf{u}})\|_{-h} \\ &\quad -\lambda\Big(\|(\tilde{\boldsymbol{\omega}}_h - \tilde{\boldsymbol{\omega}}) \times \mathbf{u}_h\|_{-h} + \|\boldsymbol{\omega}_h \times (\tilde{\mathbf{u}}_h - \tilde{\mathbf{u}})\|_{-h}\Big) . \end{aligned}$$

Then, (5.13), (5.1) and (5.2) combine to yield

$$\|(\tilde{\boldsymbol{\omega}}_h - \tilde{\boldsymbol{\omega}}) \times \mathbf{u}_h\|_{-h} \le C h^{\tilde{d}} \|\mathbf{u}_h\|_1 \|\tilde{\boldsymbol{\omega}}\|_{\tilde{d}},$$

and

$$\|\boldsymbol{\omega}_h \times (\tilde{\mathbf{u}}_h - \tilde{\mathbf{u}})\|_{-h} \le C h^{\tilde{d}} \|\boldsymbol{\omega}_h\|_0 \|\tilde{\mathbf{u}}\|_{\tilde{d}+1}.$$

As a result,

$$\begin{aligned}\|\nabla\times\boldsymbol{\omega}_h+\nabla r_h+\lambda(\tilde{\boldsymbol{\omega}}_h\times\mathbf{u}_h+\boldsymbol{\omega}_h\times\tilde{\mathbf{u}}_h)\|_{-h}^2 \\ \ge \|\nabla\times\boldsymbol{\omega}_h+\nabla r_h+\lambda(\tilde{\boldsymbol{\omega}}\times\mathbf{u}_h+\boldsymbol{\omega}_h\times\tilde{\mathbf{u}})\|_{-h}^2-Ch^{\tilde{d}}\|\mathcal{U}_h\|_{\mathbf{X}_h}^2 .\end{aligned} \tag{5.22}$$

The last term in (5.21) is estimated using (5.20) and (5.16)

$$2\lambda\Big(\nabla\times\tilde{\boldsymbol{\omega}}_h+\nabla\tilde{r}_h+\lambda\tilde{\boldsymbol{\omega}}_h\times\tilde{\mathbf{u}}_h-\mathbf{f},\boldsymbol{\omega}_h\times\mathbf{u}_h\Big)\le Ch^{\tilde{d}}\|\boldsymbol{\omega}_h\|_0\|\mathbf{u}_h\|_1 . \tag{5.23}$$

Combining (5.22), (5.23), and (2.22), i.e., the nonsingularity assumption on $\tilde{\mathcal{U}}$, yields

$$\begin{aligned}\mathcal{B}'[\tilde{\mathcal{U}}_h](\mathcal{U}_h,\mathcal{U}_h) &\ge \|\nabla\times\boldsymbol{\omega}_h+\nabla r_h+\lambda(\tilde{\boldsymbol{\omega}}\times\mathbf{u}_h+\boldsymbol{\omega}_h\times\tilde{\mathbf{u}})\|_{-h}^2 \\ &\quad+\|\nabla\cdot\mathbf{u}_h\|_0^2+\|\nabla\times\mathbf{u}_h-\boldsymbol{\omega}_h\|_0^2-Ch^{\tilde{d}}\|\mathcal{U}_h\|_{\mathbf{X}_h}^2 \\ &\ge C_1(\|\boldsymbol{\omega}_h\|_0^2+\|\mathbf{u}_h\|_1^2+\|r_h\|_0^2)-Ch^{\tilde{d}}\|\mathcal{U}_h\|_{\mathbf{X}_h}^2.\end{aligned}$$

Thus, one can find $h_0>0$ such that $\mathcal{B}'[\tilde{\mathcal{U}}_h](\cdot,\cdot)$ is coercive for all $h<h_0$. We omit the proof of continuity which easily follows from the Cauchy and triangle inequalities. ∎

The last lemma in this section establishes a Lipschitz bound for the Frechet derivative F'.

Lemma 5.5. *For* $\hat{\mathcal{U}}_h^1,\hat{\mathcal{U}}_h^2\in\mathbf{X}_h$

$$\|F'(\hat{\mathcal{U}}_h^1)-F'(\hat{\mathcal{U}}_h^2)\|_{L(\mathbf{X}_h,\mathbf{X}_h)}\le C\|\hat{\mathcal{U}}_h^1-\hat{\mathcal{U}}_h^2\|_{\mathbf{X}_h}\Big(1+\|\hat{\mathcal{U}}_h^1\|_{\mathbf{X}_h}+\|\hat{\mathcal{U}}_h^2\|_{\mathbf{X}_h}\Big) \tag{5.24}$$

Proof. From (5.10) and definitions of T and G' it follows that for $\mathcal{U}_h\in\mathbf{X}_h$

$$U_h^i=F'(\hat{\mathcal{U}}_h^i)\cdot\mathcal{U}_h,\quad i=1,2;$$

if and only if $\mathcal{U}_h^i$ solve

$$\textit{seek } \mathcal{U}_h^i\in\mathbf{X}_h \textit{ such that } \quad \mathcal{B}_S(\mathcal{U}_h^i,\mathcal{V}_h)=\mathcal{B}'[\hat{\mathcal{U}}_h^i](\mathcal{U}_h,\mathcal{V}_h), \quad \textit{for all } \mathcal{V}_h\in\mathbf{X}_h. \tag{5.25}$$

Therefore,

$$\mathcal{B}_S(\mathcal{U}_h^1-\mathcal{U}_h^2,\mathcal{V}_h)=\mathcal{B}'[\hat{\mathcal{U}}_h^1](\mathcal{U}_h,\mathcal{V}_h)-\mathcal{B}'[\hat{\mathcal{U}}_h^2](\mathcal{U}_h,\mathcal{V}_h) .$$

Next we proceed with an estimate of the right-hand side above. From (5.8)

$$\begin{aligned}&\mathcal{B}'[\hat{\mathcal{U}}_h^1](\mathcal{U}_h,\mathcal{V}_h)-\mathcal{B}'[\hat{\mathcal{U}}_h^2](\mathcal{U}_h,\mathcal{V}_h) \\ &= \Big(\nabla\times(\hat{\boldsymbol{\omega}}_h^1-\hat{\boldsymbol{\omega}}_h^2)+\nabla(\hat{r}_h^1-\hat{r}_h^2)+\lambda(\hat{\boldsymbol{\omega}}_h^1\times\hat{\mathbf{u}}_h^1-\hat{\boldsymbol{\omega}}_h^2\times\hat{\mathbf{u}}_h^2),\lambda(\boldsymbol{\omega}_h\times\mathbf{v}_h+\boldsymbol{\xi}_h\times\mathbf{u}_h)\Big)_{-h} \\ &\quad+\lambda\Big(\nabla\times\boldsymbol{\omega}_h+\nabla r_h,(\hat{\boldsymbol{\omega}}_h^1-\hat{\boldsymbol{\omega}}_h^2)\times\mathbf{v}_h+\boldsymbol{\xi}_h\times(\hat{\mathbf{u}}_h^1-\hat{\mathbf{u}}_h^2)\Big)_{-h} \\ &\quad+\lambda\Big((\hat{\boldsymbol{\omega}}_h^1-\hat{\boldsymbol{\omega}}_h^2)\times\mathbf{u}_h+\boldsymbol{\omega}_h\times(\hat{\mathbf{u}}_h^1-\hat{\mathbf{u}}_h^2),\nabla\times\boldsymbol{\xi}_h+\nabla r_h\Big)_{-h} \\ &\quad+\lambda^2\Big((\hat{\boldsymbol{\omega}}_h^1\times\mathbf{u}_h+\boldsymbol{\omega}_h\times\hat{\mathbf{u}}_h^1,\hat{\boldsymbol{\omega}}_h^1\times\mathbf{v}_h+\boldsymbol{\xi}_h\times\hat{\mathbf{u}}_h^1)_{-h} \\ &\qquad-(\hat{\boldsymbol{\omega}}_h^2\times\mathbf{u}_h+\boldsymbol{\omega}_h\times\hat{\mathbf{u}}_h^2,\hat{\boldsymbol{\omega}}_h^2\times\mathbf{v}_h+\boldsymbol{\xi}_h\times\hat{\mathbf{u}}_h^2)_{-h}\Big) .\end{aligned} \tag{5.26}$$

For the first term in (5.26) we use the Cauchy and triangle inequalities, and (5.14)-(5.16):

$$
\begin{aligned}
&\Big(\nabla\times(\hat{\boldsymbol\omega}_h^1-\hat{\boldsymbol\omega}_h^2)+\nabla(\hat r_h^1-\hat r_h^2)+\lambda(\hat{\boldsymbol\omega}_h^1\times\hat{\mathbf u}_h^1-\hat{\boldsymbol\omega}_h^2\times\hat{\mathbf u}_h^2),\lambda(\boldsymbol\omega_h\times\mathbf v_h+\boldsymbol\xi_h\times\mathbf u_h)\Big)_{-h}\\
&\quad\le\ \Big\{\|\nabla\times(\hat{\boldsymbol\omega}_h^1-\hat{\boldsymbol\omega}_h^2)\|_{-h}+\|\nabla(\hat r_h^1-\hat r_h^2)\|_{-h}+\|\lambda(\hat{\boldsymbol\omega}_h^1\times\hat{\mathbf u}_h^1-\hat{\boldsymbol\omega}_h^2\times\hat{\mathbf u}_h^2)\|_{-h}\Big\}\\
&\qquad\lambda\Big\{\|\boldsymbol\omega_h\times\mathbf v_h\|_{-h}+\|\boldsymbol\xi_h\times\mathbf u_h\|_{-h}\Big\}\\
&\quad\le\ \Big\{\|\hat{\boldsymbol\omega}_h^1-\hat{\boldsymbol\omega}_h^2\|_0+\|\hat r_h^1-\hat r_h^2\|_0+(\|\hat{\boldsymbol\omega}_h^1-\hat{\boldsymbol\omega}_h^2\|_0^2+\|\hat{\mathbf u}_h^1-\hat{\mathbf u}_1^2\|_1^2)^{1/2}(\|\hat{\boldsymbol\omega}_h^1\|_0^2+\|\hat{\mathbf u}_h^2\|_1^2)^{1/2}\Big\}\\
&\qquad(\|\boldsymbol\omega_h\|_0^2+\|\mathbf u_h\|_1^2)^{1/2}(\|\boldsymbol\xi_h\|_0^2+\|\mathbf v_h\|_1^2)^{1/2}\\
&\quad\le\ C\|\hat{\mathcal U}_h^1-\hat{\mathcal U}_h^2\|_{\mathbf X_h}\|\mathcal U_h\|_{\mathbf X_h}\|\mathcal V_h\|_{\mathbf X_h}\Big(2+\lambda(\|\hat{\mathcal U}_h^1\|_{\mathbf X_h}+\|\hat{\mathcal U}_h^2\|_{\mathbf X_h})\Big)\,.
\end{aligned}
$$

Similarly, for the second and the third terms in (5.26) we obtain

$$
\begin{aligned}
\lambda\Big(\nabla\times\boldsymbol\omega_h+\nabla r_h,(\hat{\boldsymbol\omega}_h^1-\hat{\boldsymbol\omega}_h^2)\times\mathbf v_h+\boldsymbol\xi_h\times(\hat{\mathbf u}_h^1-\hat{\mathbf u}_h^2)\Big)_{-h}&\\
\le\ &\lambda\Big\{\|\nabla\times\boldsymbol\omega_h\|_{-h}+\|\nabla r_h\|_{-h}\Big\}\\
&\Big\{\|(\hat{\boldsymbol\omega}_h^1-\hat{\boldsymbol\omega}_h^2)\times\mathbf v_h\|_{-h}+\|\boldsymbol\xi_h\times(\hat{\mathbf u}_h^1-\hat{\mathbf u}_h^2)\|_{-h}\Big\}\\
\le\ &C\|\hat{\mathcal U}_h^1-\hat{\mathcal U}_h^2\|_{\mathbf X_h}\|\mathcal U_h\|_{\mathbf X_h}\|\mathcal V_h\|_{\mathbf X_h}\,,
\end{aligned}
$$

and

$$
\begin{aligned}
\lambda\Big((\hat{\boldsymbol\omega}_h^1-\hat{\boldsymbol\omega}_h^2)\times\mathbf u_h+\boldsymbol\omega_h\times(\hat{\mathbf u}_h^1-\hat{\mathbf u}_h^2),\nabla\times\boldsymbol\xi_h+\nabla r_h\Big)_{-h}&\\
\le\ &\lambda\Big\{\|(\hat{\boldsymbol\omega}_h^1-\hat{\boldsymbol\omega}_h^2)\times\mathbf u_h\|_{-h}+\|\boldsymbol\omega_h\times(\hat{\mathbf u}_h^1-\hat{\mathbf u}_h^2)\|_{-h}\Big\}\\
&\Big\{\|\nabla\times\boldsymbol\xi_h\|_{-h}+\|\nabla r_h\ \|_{-h}\Big\}\\
\le\ &C\|\hat{\mathcal U}_h^1-\hat{\mathcal U}_h^2\|_{\mathbf X_h}\|\mathcal U_h\|_{\mathbf X_h}\|\mathcal V_h\|_{\mathbf X_h}\,.
\end{aligned}
$$

Lastly, to estimate the fourth term in (5.26) we add and subtract

$$(\hat{\boldsymbol\omega}_h^2\times\mathbf u_h+\boldsymbol\omega_h\times\hat{\mathbf u}_h^2,\hat{\boldsymbol\omega}_h^1\times\mathbf v_h+\boldsymbol\xi_h\times\hat{\mathbf u}_h^1)_{-h},$$

and then apply repeatedly the Cauchy and triangle inequalities along with (5.16):

$$
\begin{aligned}
\lambda^2\Big(&(\hat{\boldsymbol\omega}_h^1\times\mathbf u_h+\boldsymbol\omega_h\times\hat{\mathbf u}_h^1,\hat{\boldsymbol\omega}_h^1\times\mathbf v_h+\boldsymbol\xi_h\times\hat{\mathbf u}_h^1)_{-h}\\
&-(\hat{\boldsymbol\omega}_h^2\times\mathbf u_h+\boldsymbol\omega_h\times\hat{\mathbf u}_h^2,\hat{\boldsymbol\omega}_h^2\times\mathbf v_h+\boldsymbol\xi_h\times\hat{\mathbf u}_h^2)_{-h}\Big)\\
=\ &\lambda^2\Big(((\hat{\boldsymbol\omega}_h^1-\hat{\boldsymbol\omega}_h^2)\times\mathbf u_h+\boldsymbol\omega_h\times(\hat{\mathbf u}_h^1-\hat{\mathbf u}_h^2),\hat{\boldsymbol\omega}_h^1\times\mathbf v_h+\boldsymbol\xi_h\times\hat{\mathbf u}_h^1)_{-h}\\
&+(\hat{\boldsymbol\omega}_h^2\times\mathbf u_h+\boldsymbol\omega_h\times\hat{\mathbf u}_h^2,(\hat{\boldsymbol\omega}_h^1-\hat{\boldsymbol\omega}_h^2)\times\mathbf v_h+\boldsymbol\xi_h\times(\hat{\mathbf u}_h^1-\hat{\mathbf u}_h^2))_{-h}\Big)\\
\le\ &C\Big\{\|\hat{\boldsymbol\omega}_h^1-\hat{\boldsymbol\omega}_h^2\|_0\|\mathbf u_h\|_1+\|\hat{\mathbf u}_h^1-\hat{\mathbf u}_h^2\|_1\|\boldsymbol\omega_h\|_0\Big\}\Big\{\|\hat{\boldsymbol\omega}_h^1\|_0\|\mathbf v_h\|_1+\|\boldsymbol\xi_h\|_0\|\hat{\mathbf u}^1\|_1\Big\}\\
+\ &C\Big\{\|\hat{\boldsymbol\omega}_h^1-\hat{\boldsymbol\omega}_h^2\|_0\|\mathbf v_h\|_1+\|\hat{\mathbf u}_h^1-\hat{\mathbf u}_h^2\|_1\|\boldsymbol\xi_h\|_0\Big\}\Big\{\|\hat{\boldsymbol\omega}_h^2\|_0\|\mathbf u_h\|_1+\|\boldsymbol\omega_h\|_0\|\hat{\mathbf u}^2\|_1\Big\}\\
\le\ &C\|\hat{\mathcal U}_h^1-\hat{\mathcal U}_h^2\|_{\mathbf X_h}\|\mathcal U_h\|_{\mathbf X_h}\|\mathcal V_h\|_{\mathbf X_h}\Big(\|\hat{\mathcal U}_h^1\|_{\mathbf X_h}+\|\hat{\mathcal U}_h^2\|_{\mathbf X_h}\Big).
\end{aligned}
$$

As a result,

$$|\mathcal{B}'[\hat{\mathcal{U}}_h^1](\mathcal{U}_h,\mathcal{V}_h) - \mathcal{B}'[\hat{\mathcal{U}}_h^2](\mathcal{U}_h,\mathcal{V}_h)| \\ \leq C\|\hat{\mathcal{U}}_h^1 - \hat{\mathcal{U}}_h^2\|_{\mathbf{X}_h}\|\mathcal{U}_h\|_{\mathbf{X}_h}\|\mathcal{V}_h\|_{\mathbf{X}_h}\left(1+\|\hat{\mathcal{U}}_h^1\|_{\mathbf{X}_h}+\|\hat{\mathcal{U}}_h^2\|_{\mathbf{X}_h}\right). \tag{5.27}$$

The lemma now follows by observing that coercivity of $\mathcal{B}_S(\cdot,\cdot)$ and (5.27) yield

$$\|(F'(\hat{\mathcal{U}}_h^1) - F'(\hat{\mathcal{U}}_h^2))\cdot\mathcal{U}_h\|_{\mathbf{X}_h} \\ = \|\mathcal{U}_h^1 - \mathcal{U}_h^2\|_{\mathbf{X}_h} \leq C\|\hat{\mathcal{U}}_h^1 - \hat{\mathcal{U}}_h^2\|_{\mathbf{X}_h}\|\mathcal{U}_h\|_{\mathbf{X}_h}\left(1+\|\hat{\mathcal{U}}_h^1\|_{\mathbf{X}_h}+\|\hat{\mathcal{U}}_h^2\|_{\mathbf{X}_h}\right).$$

■

B. Error estimates

In this section, the technical results of §A. are used to show that all assumptions of Theorem 4.1 are valid for the nonlinear least-squares problem (4.2). In addition to establishing the error estimates, this result will also assert the well-posedness of the discrete equations.

Lemma 5.6. *Assume that $\tilde{\mathcal{U}}_h$ satisfies (5.1)-(5.3). Then*

$$\|F(\tilde{\mathcal{U}}_h)\|_{\mathbf{X}_h} \leq Ch^{\tilde{d}}. \tag{5.28}$$

Proof. From (5.9), and the definitions of T and G, it follows that

$$\mathcal{U}_h = F(\tilde{\mathcal{U}}_h) \equiv \tilde{\mathcal{U}}_h + T\cdot G(\tilde{\mathcal{U}}_h),$$

if and only if $\mathcal{U}_h$ solves the variational problem

$$\textit{seek } \mathcal{U}_h\in\mathbf{X}_h \textit{ such that } \quad \mathcal{B}_S(\mathcal{U}_h,\mathcal{V}_h) = \mathcal{B}(\tilde{\mathcal{U}}_h,\mathcal{V}_h), \quad \textit{for all } \mathcal{V}_h\in\mathbf{X}_h. \tag{5.29}$$

Since $\mathcal{B}_S(\cdot,\cdot)$ is coercive (Lemma 5.2) and $\mathcal{B}(\tilde{\mathcal{U}}_h,\cdot)$ defines a continuous linear functional (Lemma 5.3), Lax-Milgramm Theorem implies that (5.29) has a unique solution $\mathcal{U}_h$. The norm of $\mathcal{U}_h$, which equals $\|F(\tilde{\mathcal{U}}_h)\|_{\mathbf{X}_h}$, can be bounded from above using Lemma 5.2 and (5.17):

$$C\|\mathcal{U}_h\|^2_{\mathbf{X}_h} \leq \mathcal{B}_S(\mathcal{U}_h,\mathcal{U}_h) = \mathcal{B}(\tilde{\mathcal{U}}_h,\mathcal{U}_h) \leq Ch^{\tilde{d}}\|\mathcal{U}_h\|_{\mathbf{X}_h}.$$

■

Lemma 5.7. *Assume that $\tilde{\mathcal{U}}_h$ satisfies (5.1)-(5.3). Then*

1. *there exists $h_0 > 0$ such that, for all $h < h_0$, $F'(\tilde{\mathcal{U}}_h)$ is an isomorphism of $\mathbf{X}_h$ onto itself;*
2. *the norm $\|F'(\tilde{\mathcal{U}}_h)^{-1}\|_{L(\mathbf{X}_h,\mathbf{X}_h)}$ is bounded from above independently of h.*

Proof. To prove 1. we must show that for every $\hat{\mathcal{U}}_h\in\mathbf{X}_h$, problem

$$F'(\tilde{\mathcal{U}}_h)\cdot\mathcal{U}_h = \hat{\mathcal{U}}_h$$

has a unique solution $\mathcal{U}_h\in\mathbf{X}_h$. From (5.10), and the definitions of T and G', we now have

$$(\mathbf{I} + T\cdot G'(\tilde{\mathcal{U}}_h))\cdot\hat{\mathcal{U}} = \hat{\mathcal{U}}_h$$

if and only if $\hat{\mathcal{U}}$ solves the problem

$$\textit{seek } \mathcal{U}_h \in \mathbf{X}_h \textit{ such that } \quad \mathcal{B}'[\tilde{\mathcal{U}}_h](\mathcal{U}_h, \mathcal{V}_h) = \mathcal{B}_S(\hat{\mathcal{U}}_h, \mathcal{V}_h), \quad \textit{for all } \mathcal{V}_h \in \mathbf{X}_h. \tag{5.30}$$

Let h_0 be the positive real number from Lemma 5.4. Then, for $\hat{\mathcal{U}}_h$ fixed, $\mathcal{B}_S(\hat{\mathcal{U}}_h, \cdot)$ defines a continuous linear functional on $\mathbf{X}_h$, and for $h < h_0$, $\mathcal{B}'[\tilde{\mathcal{U}}_h](\mathcal{U}_h, \mathcal{V}_h)$ is continuous and coercive. As a result, Lax-Milgramm Theorem implies existence and uniqueness of solutions to (5.30), i.e., $F'(\tilde{\mathcal{U}}_h)$ is an isomorphism.

To show 2. we use again problem (5.30). Recall that

$$\gamma \equiv \|F'(\tilde{\mathcal{U}}_h)^{-1}\|_{L(\mathbf{X}_h,\mathbf{X}_h)} = \sup_{0\neq\hat{\mathcal{U}}_h\in\mathbf{X}_h} \frac{\|F'(\tilde{\mathcal{U}}_h)^{-1}\cdot\hat{\mathcal{U}}_h\|_{\mathbf{X}_h}}{\|\hat{\mathcal{U}}_h\|_{\mathbf{X}_h}}.$$

For $\hat{\mathcal{U}}_h \in \mathbf{X}_h$, arbitrary, let $\mathcal{U}_h = F'(\tilde{\mathcal{U}}_h)^{-1}\cdot\hat{\mathcal{U}}_h$. Then,

$$\begin{aligned} C_1\|\mathcal{U}_h\|^2_{\mathbf{X}_h} &\le \mathcal{B}'[\tilde{\mathcal{U}}_h](\mathcal{U}_h,\mathcal{U}_h) = \mathcal{B}_S(\hat{\mathcal{U}}_h,\mathcal{U}_h) \\ &\le C_2\|\hat{\mathcal{U}}_h\|_{\mathbf{X}_h}\|\mathcal{U}_h\|_{\mathbf{X}_h}, \end{aligned}$$

that is,

$$\|F'(\tilde{\mathcal{U}}_h)^{-1}\cdot\hat{\mathcal{U}}_h\|_{\mathbf{X}_h} \le C\|\hat{\mathcal{U}}_h\|_{\mathbf{X}_h}.$$

Thus, $\gamma \le C$, independently from h. ∎

We are now ready to state and prove the main result concerning the negative norm least-squares method of §IV.

Theorem 5.8. *Assume that $\tilde{\mathcal{U}}$ is a nonsingular solution of the Navier-Stokes equations (2.8)-(2.10), and (2.3), such that $\tilde{\mathcal{U}} \in \mathbf{X}_m$ for some integer $m \ge 1$. Then, for h sufficiently small, there exists $\alpha > 0$ such that the least-squares problem (4.2) has a unique solution $\mathcal{U}_h \equiv (\boldsymbol{\omega}_h, \mathbf{u}_h, r_h) \in \mathbf{X}_h$ in the ball $B(\tilde{\mathcal{U}}, \alpha)$. Moreover, there exists a constant $C >$, independent of h, such that*

$$\|\boldsymbol{\omega}_h - \tilde{\boldsymbol{\omega}}\|_0 + \|\mathbf{u}_h - \tilde{\mathbf{u}}\|_1 + \|r_h - \tilde{r}\|_0 \le Ch^{\tilde{d}} \tag{5.31}$$

Proof. We apply Theorem 4.1 with $\tilde{\mathcal{U}}_h$ chosen to be a discrete function satisfying (5.1)-(5.3). First we note that (5.24) in Lemma 5.5 and the definition of $\tilde{\mathcal{U}}_h$ imply

$$\begin{aligned} L(\alpha) &= \sup_{\mathcal{V}_h\in B(\tilde{\mathcal{U}}_h,\alpha)} \|F'(\tilde{\mathcal{U}}_h) - F'(\mathcal{V}_h)\|_{L(\mathbf{X}_h,\mathbf{X}_h)} \\ &\le \sup_{\mathcal{V}_h\in B(\tilde{\mathcal{U}}_h,\alpha)} \|\tilde{\mathcal{U}}_h - \mathcal{V}_h\|\Big(1 + \|\tilde{\mathcal{U}}_h\|_{\mathbf{X}_h} + \|\mathcal{V}_h\|_{\mathbf{X}_h}\Big) \le C\alpha, \end{aligned}$$

that is, $L(\alpha)$ is a decreasing function of α. From (5.28) and Lemma 5.7 it now follows that there exists a positive number h_1 such that for $h < h_1$,

$$2\gamma L(2\gamma\varepsilon) < 1.$$

According to Lemma 5.7 there is another positive number h_0, such that for $h < h_0$, $F'(\tilde{\mathcal{U}}_h)$ is an isomorphism of $\mathbf{X}_h$ onto itself. As a result, all hypotheses of Theorem 4.1 are satisfied for $h < \min\{h_0, h_1\}$ and we can conclude that the least-squares problem has a unique solution in $B(\tilde{\mathcal{U}}_h, \alpha)$.

The error estimate (5.31) now easily follows from (5.28) by choosing $\mathcal{V}_h = \tilde{\mathcal{U}}_h$ in (4.8). ∎

VI. IMPLEMENTATION AND NUMERICAL RESULTS

In this section we briefly discuss several implementation issues, including methods for solving the nonlinear system (4.2), the choice of preconditioners for the discrete equations, and the choice of $\mathbf{B}_h$, needed for the computation of the discrete negative norm (3.11). Then we report some preliminary numerical results obtained with the least-squares method (4.2). A more detailed account of the algorithmic development of this method, as well as of its numerical performance will appear in a forthcoming paper.

A. Implementation

We discuss implementation of the negative norm method in two space dimensions. Recall that the choice of the finite element spaces $\mathbf{W}_h$, $\mathbf{U}_h$ and $\mathbf{R}_h$ is subject only to the approximation conditions (3.3)-(3.5). Thus, for simplicity of implementation, all variables can be approximated by the same finite element space. Here for this purpose we choose the *biquadratic* finite element space Q_2 (see [14]), that is

$$\mathbf{X}_h = [Q_2] \times [Q_2 \cap H^1_0(\Omega)]^2 \times [Q_2 \cap L^2_0(\Omega)] \,.$$

Once a basis for $\mathbf{X}_h$ is chosen, problem (4.2) becomes a nonlinear system of algebraic equations of the form (4.3), that must be solved in an iterative manner. This system has some very attractive computational properties which can be exploited in the algorithmic design. First, F' is exactly the Hessian matrix of the functional (3.1) and thus, it is symmetric. Second, if $\tilde{\mathcal{U}}$ is a nonsingular solution of the Navier-Stokes equations (2.8)-(2.10), Theorem 5.8 assures that F' is invertible in a neighborhood of $\tilde{\mathcal{U}}$. As a result, we are guaranteed the existence of a nontrivial neighborhood of $\tilde{\mathcal{U}}$ such that F' is symmetric and positive definite. These features suggest the Newton's method as a good candidate for the solution of (4.2). Indeed, symmetry and positive definiteness of F' imply that the linearized equations can be solved using robust iterative methods. This fact has critical importance for the efficient implementation of the least-squares method, because the use of negative norms leads to dense discretization matrices. Consequently, solution of the linearized systems must be accomplished without forming the discretization matrix. For this purpose we consider a preconditioned conjugate gradients method, implemented without matrix assembly. To define the preconditioner for the conjugate gradient method, let us consider the linearized system that must be solved at the $k+1$st Newton step. If $\mathcal{U}_h^k$ denotes the kth Newton iterate, this system is given by

$$\textit{seek } \delta\mathcal{U}_h^{k+1} \in \mathbf{X}_h \textit{ such that } \quad \mathcal{B}'[\mathcal{U}_h^k](\delta\mathcal{U}_h^{k+1}, \mathcal{V}_h) = -\mathcal{B}(\mathcal{U}_h^k, \mathcal{V}_h), \text{for all } \mathcal{V}_h \in \mathbf{X}_h. \quad (6.1)$$

If $\mathcal{U}_h^k$ is sufficiently close to $\tilde{\mathcal{U}}$, the form $\mathcal{B}'[\mathcal{U}_h^k](\cdot,\cdot)$ is coercive (Lemma 5.4), i.e.,

$$C_1\|\mathcal{V}_h\|^2_{\mathbf{X}_h} \leq \mathcal{B}'[\mathcal{U}_h^k](\mathcal{V}_h, \mathcal{V}_h) \leq C_2\|\mathcal{V}_h\|^2_{\mathbf{X}_h}, \quad \forall \mathcal{V}_h \in \mathbf{X}_h \,. \quad (6.2)$$

Since $\|\mathcal{V}_h\|_{\mathbf{X}_h} = \|\boldsymbol{\omega}_h\|_0 + \|\mathbf{u}_h\|_1 + \|r_h\|_0$, (6.2) implies that the matrix in (6.1) is spectrally equivalent to a block diagonal matrix of the form

$$A = \begin{pmatrix} G_{\mathbf{W}} & 0 & 0 \\ 0 & D_{\mathbf{U}} & 0 \\ 0 & 0 & G_{\mathbf{R}} \end{pmatrix},$$

where $G_{\mathbf{W}}$ and $G_{\mathbf{R}}$ are the Gramm matrices of the bases of $\mathbf{W}_h$ and $\mathbf{R}_h$, and $D_{\mathbf{U}}$ is the Dirichlet matrix for the basis of $\mathbf{U}_h$. In view of (3.10), it follows that the preconditioner for (6.1) can be defined as

$$\tilde{A} = \begin{pmatrix} h^2\mathbf{I} & 0 & 0 \\ 0 & \mathbf{B}_h^{-1} & 0 \\ 0 & 0 & h^2\mathbf{I} \end{pmatrix} . \tag{6.3}$$

In conclusion, let us discuss the operator $\mathbf{B}_h$. Although the only condition that must be satisfied by this operator is the spectral equivalence relation (3.10), computational cost is an important consideration for the choice of $\mathbf{B}_h$. This cost must be significantly lower than the cost of solving (3.8). Thus, it is typical to consider operators $\mathbf{B}_h$, defined in terms of several multigrid V-cycles; see, e.g., [9]-[10], and [11]. Other choices of $\mathbf{B}_h$ include various preconditioners for (3.8), or even few conjugate gradient iterations applied to (3.8).

B. Numerical examples

The principal objective of the numerical experiments presented in this section is to demonstrate convergence rates of Theorem 5.8, and performance of the preconditioner (6.3). For all experiments Ω is taken to be the unit square in $\mathbf{R}^2$. We consider an exact solution given by

$$\begin{aligned} \tilde{\mathbf{u}} &= \left(\exp(x)\cos(y) + \sin(y), -\exp(x)\sin(x) + (1-x^3)\right)^t \\ \tilde{\boldsymbol{\omega}} &= \nabla \times \tilde{\mathbf{u}} \\ \tilde{r} &= \sin(y)\cos(x) + xy^2 - \frac{1}{6} - \sin(1)(1-\cos(1)) . \end{aligned} \tag{6.4}$$

The data for the least-squares method is computed by substituting (6.4) in the Navier-Stokes equations (2.8)-(2.10). In view of the choice of $\mathbf{X}_h$ we expect that (5.31) holds with $\tilde{d} = 2$, i.e.,

$$\|\boldsymbol{\omega}_h - \tilde{\boldsymbol{\omega}}\|_0 + \|\mathbf{u}_h - \tilde{\mathbf{u}}\|_1 + \|r_h - \tilde{r}\|_0 \le Ch^2$$

To estimate convergence rates numerically, computations were carried out using triangulations with 8x8 and 16x16 uniformly spaced grid lines in each coordinate direction (15x15 and 31x31 uniform grid points). L^2 and H^1 errors were computed using a six point Gauss-Legendre quadrature rule. Convergence rates estimates are summarized in Table I. The underlined rates in this table correspond to the errors included in (5.31). Although our analysis does not assert optimal L^2 rates for the velocity, and optimal H^1

TABLE I. Convergence rates of the negative norm least-squares method.

Variable	L^2 rates	H^1 rates
$\mathbf{u}$	3.082	<u>2.023</u>
$\boldsymbol{\omega}$	<u>2.996</u>	1.989
r	<u>2.430</u>	1.998

rates for the vorticity and the pressure, these rates also appear optimal. Most likely, this can be attributed to the fact that all components of the exact solution (6.4) are C^∞ functions.

Next we consider how well the problem (4.2) is preconditioned by (6.3). Table II reports the number of conjugate gradient iterations, with and without the preconditioner, for each step of the Newton's method for uniform grids with 3x3, 4x4, 8x8, and 16x16 grid lines in each coordinate direction. Initial approximation was computed by solving a Stokes problem with a right hand side given by (6.4) evaluated at the Navier-Stokes equations (2.8)-(2.10). In all cases, iterations were carried out until the relative error in the residual vector became less than $0.5/10^{-5}$. The same tolerance was used for the conjugate gradients solver. With the exception of the 3x3 case, Newton's method converged in two steps. From the data in Table II, we see that the number of conjugate gradient iterations with the preconditioner (6.3) grows very slowly. This suggests that the condition number of the discretization matrix, preconditioned by (6.3), is bounded, or also grows very slowly. Similar results are observed when (6.3) is applied to the Stokes equations. In that case, the preconditioner does even better job in keeping the condition numbers of the discretization matrix bounded; see Table III.

TABLE II. Effects of preconditioning: Navier-Stokes equations

Grid	3x3		4x4		8x8		16x16	
Newton step	PCG	CG	PCG	CG	PCG	CG	PCG	CG
step 1	77	102	118	257	161	657	175	2059
step 2	76	93	109	243	151	610	152	1784
step 3	69	89	-	-	-	-	-	-

TABLE III. Effects of preconditioning: Stokes equations

Grid	3x3	4x4	8x8	16x16
PCG	59	77	88	96
CG	73	173	455	999

REFERENCES

1. Agmon, S., Douglis, A., and Nirenberg, L., "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. Bochev, P., "Analysis of least-squares finite element methods for the Navier-Stokes equations," SIAM J. Num. Anal., 34, 1997, pp. 1817–1844.
3. Bochev, P., Cai, Z., Manteuffel, T. A., and McCormick, S. F., "Analysis of velocity-flux least-squares principles for the Navier-Stokes equations: Part I," SIAM J. Num. Anal., 35-3, 1998, pp. 990–1009.
4. Bochev, P., and Gunzburger, M., "A least-squares finite element method for the Navier-Stokes equations," Appl. Math. Lett. 6/2, 1993, pp. 27–30.
5. Bochev, P., and Gunzburger, M., "Accuracy of least-squares methods for the Navier-Stokes equations," Comput. & Fluids, 22, 1993, pp. 549–563.
6. Bochev, P., and Gunzburger, M., "Analysis of least-squares finite element methods for the Stokes equations," Math. Comp., 63, 1994, pp. 479–506.
7. Bochev, P., and Gunzburger, M., "Least-squares methods for the velocity-pressure-stress formulation of the Stokes equations," Comp. Meth. Appl. Mech. Engrg., 126, 1995, pp. 267–287.

8. Bochev, P., and Gunzburger, M., "Finite element methods of least-squares type", to appear in SIAM Review.
9. Bramble, J. H., Lazarov, R. D., and Pasciak, J. E., "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.
10. Bramble, J. H., and Pasciak, J. E., "Least-squares methods for Stokes equations based on a discrete minus one inner product," manuscript, 1994.
11. Cai, Z., Manteuffel, T. A., and McCormick, S. F., "First-order system least-squares for velocity-vorticity-pressure form of the Stokes equations, with application to linear elasticity," ETNA, 3, 1995, pp. 150–159.
12. Cai, Z., Manteuffel, T. A., and McCormick, S. F., "First-order system least squares for the Stokes equations, with application to linear elasticity," SIAM J. Num. Anal., to appear.
13. CHANG, C. L., "Piecewise linear approximations to the Stokes problem with velocity boundary conditions," submitted.
14. Ciarlet, P. G. "The finite element method for elliptic problems," North-Holland, Amsterdam, 1978.
15. Girault, V. and Raviart, P.-A., "Finite Element Methods for Navier-Stokes Equations," Springer, Berlin, 1986.
16. Gunzburger, M. D. "Finite Element Methods for Viscous Incompressible Flows," Academic, Boston, 1989.
17. Jiang, B.-N. and Chang, C. L., "Least-squares finite elements for the Stokes problem," Comput. Methods Appl. Mech. Engrg., 78, 1990, pp. 297–311.
18. Jiang, B.-N., Lin, T. and Povinelli, L., "A least-squares finite element method for 3D incompressible Navier-Stokes equations," AIAA Report 93-0338, AIAA, New York, 1993.
19. Jiang, B.-N., Lin, T. and Povinelli, L., "Large-scale computation of incompressible viscous flow by least-squares finite element method," Comput. Methods Appl. Mech. Engrg., 114, 1994, pp. 213–231.
20. Jiang, B.-N., Loh, C. Y. and Povinelli, L., "Theoretical study of the incompressible Navier-Stokes equations by the least-squares method," NASA Technical Memorandum 106535, ICOMP-94-04, NASA, Cleveland, 1994.
21. Lefebvre, D., Peraire, J. and Morgan, K., "Least-squares finite element solution of compressible and incompressible flows," Int. J. Num. Meth. Heat Fluid Flow, 2, 1992, pp. 99–113.
22. Wendland, W., "Elliptic Systems in the Plane," Pitman, London, 1979.