J. KSIAM Vol.14, No.2, 125–140, 2010

ANALYSIS OF VELOCITY-FLUX FIRST-ORDER SYSTEM LEAST-SQUARES PRINCIPLES FOR THE OPTIMAL CONTROL PROBLEMS FOR THE NAVIER-STOKES EQUATIONS

YOUNGMI CHOI† AND HYUNG-CHUN LEE

DEPARTMENT OF MATHEMATICS, AJOU UNIVERSITY, SUWON 443-749, KOREA
E-mail address: ymchoi93@hanmail.net, hclee@ajou.ac.kr

ABSTRACT. This paper develops a least-squares approach to the solution of the optimal control problem for the Navier-Stokes equations. We recast the optimality system as a first-order system by introducing velocity-flux variables and associated curl and trace equations. We show that a least-squares principle based on L^2 norms applied to this system yields optimal discretization error estimates in the H^1 norm in each variable.

1. INTRODUCTION

In [8], Choi, Kim, Lee, and Shin developed first-order system least-squares functionals for formulation of the optimal control problem for the scaled Navier-Stokes equations. From the Lagrangian, one may derive an optimality system of equations for the solution of the optimal control problem for the Navier-Stokes equations. They recast the optimality system as a first-order system by introducing a velocity-flux variable. A least-squares principle based on L^2-norm applied to this first-order system and optimal discretization error estimates are obtained.

The goal of this paper is to extend this methodology to the optimal control problem for the original Navier-Stokes equations in two and three dimensions. We make this extension in the same way that the optimal control problem for the scaled Navier-Stokes equations were reformulated based on the velocity flux variables, but now we replace the data $\mathbf{u}_d$ by functions with known values. We first obtain a coupled optimality system related to two Navier-Stokes type equations associated with state variables and adjoint variables. The optimality system may be written as first-order system of partial differential equations by introducing velocity-flux variables. The Euler-Lagrange equations for the corresponding least-squares principle are then recast in the canonical form. This allows us to apply conventional abstract theory and our results to obtain optimal error estimates for least-squares finite element method.

Received by the editors May 7 2010; Accepted June 6 2010.
2000 *Mathematics Subject Classification.* 93B05.
Key words and phrases. optimal control, least-squares finite element methods, Navier-Stokes equations.
The first author's work was supported in part by the Korea Research Foundation Grant funded by the Korean Government(KRF-2008-359-C00007). The second author's work was supported in part by the Korea Research Foundation Grant 2009-0072201.
† Corresponding author.

The optimal control problem we consider is to minimize the functional

$$\mathcal{J}(\mathbf{u},p,\mathbf{f}) = \frac{1}{2}\,\|\mathbf{u}-\mathbf{u}_d\|^2 + \frac{\beta}{2}\|\mathbf{f}\|^2, \tag{1.1}$$

subject to the incompressible Navier-Stokes equations

$$-\nu\Delta\,\mathbf{u} + (\mathbf{u}\cdot\nabla)\mathbf{u} + \nabla p = \mathbf{f} \quad \text{in}\,\Omega, \tag{1.2}$$

$$\nabla\cdot\mathbf{u} = 0 \quad \text{in}\,\Omega, \tag{1.3}$$

$$\mathbf{u} = \mathbf{0} \quad \text{on}\,\partial\Omega \tag{1.4}$$

where $\mathbf{u}_d$ is a given desired function. Here, $\Omega\subset\mathbb{R}^n(n=2$ or $3)$ be an open, connected, and bounded domain with Lipschitz boundary $\partial\Omega=\Gamma$ and $\mathbf{u}$ a candidate velocity field, p the pressure, $\mathbf{f}$ a prescribed forcing term and ν the viscous constant. Assume that p satisfies the zero mean constraint, $\int_\Omega p\,dx=0$. The objective of this optimal control problem is to seek a state variables $\mathbf{u}$ and p, and the control $\mathbf{f}$ which minimize the L^2-norm distances between $\mathbf{u}$ and $\mathbf{u}_d$ and satisfy (1.2)–(1.4). The second term in (1.1) is added as a limiting the cost of control and the positive penalty parameter δ can be used to change the relative importance of the two terms appearing in the definition of the functional.

This paper consists of the following. In the next section, we give a precise statement of the optimization problem. Then we reformulate the optimality systems to the first-order system and define the L^2-norm least squares functional. In §3, we obtain the optimal error estimates for least-squares finite element method for the optimality system.

Throughout the paper, we use boldface lower case font to denote vectors and underline boldface upper case font to denote matrices.

2. The optimal control problem

2.1. **The optimization problem.** Let $\mathbf{u}\in H_0^1(\Omega)^n$ and $p\in L_0^2(\Omega)$ denote the state variables, and let $\mathbf{f}\in L^2(\Omega)^n$ denote the distributed control. The state and control variables are also constrained to satisfy the system (1.2)–(1.4), which recast into the weak form:

$$\nu a(\mathbf{u},\mathbf{w}) + c(\mathbf{u},\mathbf{u},\mathbf{w}) - b(\mathbf{w},p) = \langle\mathbf{f},\mathbf{w}\rangle \quad \forall\mathbf{w}\in H^1(\Omega)^n \tag{2.1}$$

$$b(\mathbf{u},r) = 0 \quad \forall r\in L^2(\Omega) \tag{2.2}$$

where

$$a(\mathbf{u},\mathbf{w}) = \int_\Omega \nabla\mathbf{u}:\nabla\mathbf{w}d\mathbf{x} = \frac{1}{2}\int_\Omega(\nabla\mathbf{u}+\nabla\mathbf{u}^T):(\nabla\mathbf{w}+\nabla\mathbf{w}^T)d\mathbf{x},$$

$$b(\mathbf{w},p) = \int_\Omega p\nabla\cdot\mathbf{w}d\mathbf{x},$$

$$c(\mathbf{u},\mathbf{v},\mathbf{w}) = \int_\Omega(\mathbf{u}\cdot\nabla)\mathbf{v}\cdot\mathbf{w}d\mathbf{x}$$

Then, since $\mathbf{f}\in L^2(\Omega)^n$, it is well known (see [10] or [14]) that $(\mathbf{u},p)\in\mathbf{H}^2(\Omega)^n\times H^1(\Omega)$ and $\|\mathbf{u}\|_2+\|p\|_1\le C\|\mathbf{f}\|$.

With $\mathcal{J}(\cdot)$ given by (1.1), the $admissibility\ set$ $\mathcal{U}_{ad}$ is defined by

$$\begin{aligned}\mathcal{U}_{ad} \;=\; \{&(\mathbf{u}, p, \mathbf{f}) \in H_0^1(\Omega)^n \times L_0^2(\Omega) \times L^2(\Omega)^n : \\ &\mathcal{J}(\mathbf{u}, p, \mathbf{f}) < \infty \;\text{ and } (\mathbf{u}, p, \mathbf{f}) \text{ satisfies (2.1) and (2.2)}\}\end{aligned} \tag{2.3}$$

Then $(\hat{\mathbf{u}}, \hat{p}, \hat{\mathbf{f}}) \in \mathcal{U}_{ad}$ is called an optimal solution if there exists $\epsilon > 0$ such that

$$\mathcal{J}(\hat{\mathbf{u}}, \hat{p}, \hat{\mathbf{f}}) \le \mathcal{J}(\mathbf{u}, p, \mathbf{f}) \qquad \forall (\mathbf{u}, p, \mathbf{f}) \in \mathcal{U}_{ad}$$

satisfying

$$\|\hat{\mathbf{u}} - \mathbf{u}\|_1 + \|\hat{p} - p\| + \|\hat{\mathbf{f}} - \mathbf{f}\| < \epsilon$$

The optimal control problem can now be formulated as a constrained minimization in a Hilbert space

$$\min_{(\mathbf{u},p,\mathbf{f}) \in \mathcal{U}_{ad}} \mathcal{J}(\mathbf{u}, p, \mathbf{f}) \tag{2.4}$$

Theorem 2.1. *Given $\mathbf{u}_d$, there exists a solution $(\tilde{\mathbf{u}}, \tilde{p}, \tilde{\mathbf{f}}) \in \mathcal{U}_{ad}$ such that (1.1) is minimized.*

Proof. It is similar to Thoerem 2.1 in [11]. □

2.2. **An optimality system.** From the Lagrangian

$$\mathcal{L}(\mathbf{u}, p, \mathbf{f}, \mathbf{v}, q : \mathbf{u}_d) = \mathcal{J}(\mathbf{u}, p, \mathbf{f}) - (\nu \Delta \mathbf{u} - (\mathbf{u} \cdot \nabla)\mathbf{u} - \nabla p + \mathbf{f}, \mathbf{v}) - (\nabla \cdot \mathbf{u}, q)$$

where $\mathcal{J}(\cdot, \cdot, \cdot)$ is defined by (1.1), one may derive an optimality system of equations for the solution of (2.4). The constrained problem (2.4) can now be recast as the unconstrained problem of finding stationary points of $\mathcal{L}(\cdot)$. We now apply the necessary conditions for the latter problem. Clearly, setting to zero the first variations with respect to $\mathbf{u}, p, \mathbf{f}, \mathbf{v}$ and q yields the optimality system

$$\left.\begin{aligned} -(\nu \Delta \mathbf{u} - (\mathbf{u} \cdot \nabla)\mathbf{u} - \nabla p + \mathbf{f}, \tilde{\mathbf{v}}) &= 0 \quad \forall \tilde{\mathbf{v}} \in H_0^1(\Omega)^n, \\ -(\nabla \cdot \mathbf{u}, \tilde{q}) &= 0 \quad \forall \tilde{q} \in L_0^2(\Omega), \\ \mathbf{u} &= \mathbf{0}, \quad \text{on}\, \Gamma, \\ (\mathbf{u} - \mathbf{u}_d, \tilde{\mathbf{u}}) - (\nu \Delta \tilde{\mathbf{u}} - (\tilde{\mathbf{u}} \cdot \nabla)\mathbf{u} - (\mathbf{u} \cdot \nabla)\tilde{\mathbf{u}}, \mathbf{v}) + (\nabla q, \tilde{\mathbf{u}}) &= 0 \quad \tilde{\mathbf{u}} \in H_0^1(\Omega)^n, \\ -(\nabla \cdot \mathbf{v}, \tilde{p}) &= 0 \quad \tilde{p} \in L_0^2(\Omega), \\ \mathbf{v} &= \mathbf{0} \quad \text{on}\, \Gamma, \\ (\beta \mathbf{f} - \mathbf{v}, \tilde{\mathbf{f}}) &= 0 \quad \tilde{\mathbf{f}} \in H^{-1}(\Omega)^n. \end{aligned}\right\} \tag{2.5}$$

Integrations by parts may be used to show that the system (2.5) constitutes a weak formulation of the problem

$$\left.\begin{aligned} -\nu\Delta\mathbf{u} + (\mathbf{u}\cdot\nabla)\mathbf{u} + \nabla p = \mathbf{f} &\quad \text{in}\,\Omega,\\ \nabla\cdot\mathbf{u} = 0 &\quad \text{in}\,\Omega,\\ \mathbf{u} = \mathbf{0} &\quad \text{on}\,\Gamma,\\ (\mathbf{u}-\mathbf{u}_d) - \nu\Delta\mathbf{v} + (\nabla\mathbf{u})^t\mathbf{v} - (\mathbf{u}\cdot\nabla)\mathbf{v} + \nabla q = 0 &\quad \text{in}\,\Omega,\\ \nabla\cdot\mathbf{v} = 0 &\quad \text{in}\,\Omega,\\ \mathbf{v} = \mathbf{0} &\quad \text{on}\,\Gamma,\\ \beta\mathbf{f} = \mathbf{v} &\quad \text{in}\,\Omega. \end{aligned}\right\} \tag{2.6}$$

Note that this system is coupled, i.e., the constraint equations for the state variables depend on the unknown controls, the adjoint equations for the Lagrange multipliers depend on the state, and optimality conditions for the controls depend on the Lagrange multipliers.

2.3. **First-order system.** To formulate the least-squares method, system (2.6) will be transformed into an equivalent first-order system. Introduce the velocity-flux variable

$$\underline{\mathbf{U}} = \nabla\mathbf{u}^t \quad \text{and} \quad \underline{\mathbf{V}} = \nabla\mathbf{v}^t$$

which is a matrix with entries $U_{ij} = \partial u_j/\partial x_i$ and $V_{ij} = \partial v_j/\partial x_i$, $1 \le i, j \le n$. Then

$$(\nabla^t\underline{\mathbf{U}})^t = \Delta\mathbf{u} \quad \text{and} \quad (\nabla^t\underline{\mathbf{V}})^t = \Delta\mathbf{v}$$

and it is easy to see that the new variable satisfies the identities

$$\text{tr}\underline{\mathbf{U}} = 0, \quad \nabla\times\underline{\mathbf{U}} = \underline{\mathbf{0}}, \quad \text{tr}\underline{\mathbf{V}} = 0, \quad \nabla\times\underline{\mathbf{V}} = \underline{\mathbf{0}} \quad \text{in}\,\Omega$$

and

$$\mathbf{n}\times\underline{\mathbf{U}} = \underline{\mathbf{0}} \quad \mathbf{n}\times\underline{\mathbf{V}} = \underline{\mathbf{0}} \quad \text{on}\,\Gamma$$

where $\text{tr}\underline{\mathbf{U}} = \sum_{i=1}^n U_{ii}$ and $\mathbf{n}$ is the outward unit normal on Γ.

The optimality condition (the last equation in (2.6)) can be substituted into the state equations and thus, we have the first-order optimality system

$$\left.\begin{aligned} -\nu(\nabla^t\underline{\mathbf{U}})^t + \underline{\mathbf{U}}^t\mathbf{u} + \nabla p = \frac{\mathbf{v}}{\beta} &\quad \text{in}\ \Omega,\\ \nabla^t\mathbf{u} = 0 &\quad \text{in}\ \Omega,\\ \underline{\mathbf{U}} - \nabla\mathbf{u}^t = \underline{\mathbf{0}} &\quad \text{in}\ \Omega,\\ \nabla(\text{tr}\underline{\mathbf{U}}) = \mathbf{0} &\quad \text{in}\ \Omega,\\ \nabla\times\underline{\mathbf{U}} = \underline{\mathbf{0}} &\quad \text{in}\ \Omega,\\ \mathbf{u} = \mathbf{0} &\quad \text{on}\ \Gamma,\\ \int p\,dx = 0 &\quad \text{in}\ \Omega,\\ \mathbf{n}\times\underline{\mathbf{U}} = \underline{\mathbf{0}} &\quad \text{on}\ \Gamma, \end{aligned}\right\} \tag{2.7}$$

and

$$\left.\begin{aligned} (\mathbf{u}-\mathbf{u}_d) - \nu(\nabla^t\underline{\mathbf{V}})^t + \underline{\mathbf{U}}\mathbf{v} - \underline{\mathbf{V}}^t\mathbf{u} + \nabla q = \mathbf{0} \quad &\text{in } \Omega,\\ \nabla^t\mathbf{v} = 0 \quad &\text{in } \Omega,\\ \underline{\mathbf{V}} - \nabla\mathbf{v}^t = \underline{\mathbf{0}} \quad &\text{in } \Omega,\\ \nabla(\mathrm{tr}\underline{\mathbf{V}}) = \mathbf{0} \quad &\text{in } \Omega,\\ \nabla\times\underline{\mathbf{V}} = \underline{\mathbf{0}} \quad &\text{in } \Omega,\\ \mathbf{v} = \mathbf{0} \quad &\text{on } \Gamma,\\ \int q\,dx = 0 \quad &\text{in } \Omega,\\ \mathbf{n}\times\underline{\mathbf{V}} = \underline{\mathbf{0}} \quad &\text{on } \Gamma. \end{aligned}\right\} \tag{2.8}$$

The next step in the formulation of a first-order system is to scale the first equations in (2.7) and (2.8) by the Reynolds number and to replace the data $\mathbf{u}_d$ by functions with known values. The resulting form of the equations will provide insight into the overall approach and facilitate error analysis of the corresponding least-squares method. For this purpose, we assume that $\mathbf{u}_d \in L^2(\Omega)^n$ and consider the optimal solution $(\mathbf{u}_0, p_0, \mathbf{v}_0, q_0)$ of the scaled optimality system for the Stokes equation

$$\left.\begin{aligned} -\Delta\mathbf{u} + \nabla p = \frac{\mathbf{v}}{\nu\beta} \quad &\text{in}\,\Omega,\\ \nabla\cdot\mathbf{u} = 0 \quad &\text{in}\,\Omega,\\ \mathbf{u} = \mathbf{0} \quad &\text{on}\,\Gamma,\\ \frac{1}{\nu}(\mathbf{u}-\mathbf{u}_d) - \Delta\mathbf{v} + \nabla q = 0 \quad &\text{in}\,\Omega,\\ \nabla\cdot\mathbf{v} = 0 \quad &\text{in}\,\Omega,\\ \mathbf{v} = \mathbf{0} \quad &\text{on}\,\Gamma \end{aligned}\right\} \tag{2.9}$$

Letting $\underline{\mathbf{U}}_0 = \nabla\mathbf{u}_0^t$ and $\underline{\mathbf{V}}_0 = \nabla\mathbf{v}_0^t$, then the optimality system (2.7)–(2.8) is replaced by

$$\left.\begin{aligned} -(\nabla^t\underline{\mathbf{U}})^t + \frac{1}{\nu}(\underline{\mathbf{U}}+\underline{\mathbf{U}}_0)^t(\mathbf{u}+\mathbf{u}_0) + \nabla p = \frac{\mathbf{v}}{\nu\beta} \quad &\text{in } \Omega,\\ \nabla^t\mathbf{u} = 0 \quad &\text{in } \Omega,\\ \underline{\mathbf{U}} - \nabla\mathbf{u}^t = \underline{\mathbf{0}} \quad &\text{in } \Omega,\\ \nabla(\mathrm{tr}\underline{\mathbf{U}}) = \mathbf{0} \quad &\text{in } \Omega,\\ \nabla\times\underline{\mathbf{U}} = \underline{\mathbf{0}} \quad &\text{in } \Omega,\\ \mathbf{u} = \mathbf{0} \quad &\text{on } \Gamma,\\ \int p\,dx = 0 \quad &\text{in } \Omega,\\ \mathbf{n}\times\underline{\mathbf{U}} = \underline{\mathbf{0}} \quad &\text{on } \Gamma, \end{aligned}\right\} \tag{2.10}$$

and

$$\left.\begin{aligned} \frac{1}{\nu}\mathbf{u} - (\nabla^t\underline{\mathbf{V}})^t + \frac{1}{\nu}(\underline{\mathbf{U}}+\underline{\mathbf{U}}_0)(\mathbf{v}+\mathbf{v}_0) & \\ -\frac{1}{\nu}(\underline{\mathbf{V}}+\underline{\mathbf{V}}_0)^t(\mathbf{u}+\mathbf{u}_0) + \nabla q = \mathbf{0} &\quad \text{in } \Omega, \\ \nabla^t\mathbf{v} = 0 &\quad \text{in } \Omega, \\ \underline{\mathbf{V}} - \nabla\mathbf{v}^t = \underline{\mathbf{0}} &\quad \text{in } \Omega, \\ \nabla(\text{tr}\underline{\mathbf{V}}) = \mathbf{0} &\quad \text{in } \Omega, \\ \nabla\times\underline{\mathbf{V}} = \underline{\mathbf{0}} &\quad \text{in } \Omega, \\ \mathbf{v} = \mathbf{0} &\quad \text{on } \Gamma, \\ \int q\, dx = 0 &\quad \text{in } \Omega, \\ \mathbf{n}\times\underline{\mathbf{V}} = \underline{\mathbf{0}} &\quad \text{on } \Gamma. \end{aligned}\right\} \tag{2.11}$$

which is the principal system that relates the perturbation $(\underline{\mathbf{U}}, \mathbf{u}, \nu p, \underline{\mathbf{V}}, \mathbf{v}, \nu q)$ to the optimality system of the stokes equations $(\underline{\mathbf{U}}_0, \mathbf{u}_0, \nu p_0, \underline{\mathbf{V}}_0, \mathbf{v}_0, \nu q_0)$.

3. LEAST-SQUARES FINITE ELEMENT METHOD

3.1. **Least-Squares.** The L^2 least-squares functional for first-order system (2.10)–(2.11) is defined as follows:

$$\begin{aligned} &\mathcal{F}_1(\underline{\mathbf{U}}, \mathbf{u}, p, \underline{\mathbf{V}}, \mathbf{v}, q : \mathbf{u}_d) \\ &\quad = \| -(\nabla^t\underline{\mathbf{U}})^t + \frac{1}{\nu}(\underline{\mathbf{U}}+\underline{\mathbf{U}}_0)^t(\mathbf{u}+\mathbf{u}_0) + \nabla p - \frac{\mathbf{v}}{\nu\beta}\|^2 \\ &\quad\quad + \|\nabla^t\mathbf{u}\|^2 + \|\underline{\mathbf{U}} - \nabla\mathbf{u}^t\|^2 + \|\nabla(\text{tr}\underline{\mathbf{U}})\|^2 + \|\nabla\times\underline{\mathbf{U}}\|^2 \\ &\quad\quad + \| -(\nabla^t\underline{\mathbf{V}})^t + \frac{1}{\nu}(\underline{\mathbf{U}}+\underline{\mathbf{U}}_0)(\mathbf{v}+\mathbf{v}_0) - \frac{1}{\nu}(\underline{\mathbf{V}}+\underline{\mathbf{V}}_0)^t(\mathbf{u}+\mathbf{u}_0) + \nabla q + \frac{1}{\nu}\mathbf{u}\|^2 \\ &\quad\quad + \|\nabla^t\mathbf{v}\|^2 + \|\underline{\mathbf{V}} - \nabla\mathbf{v}^t\|^2 + \|\nabla(\text{tr}\underline{\mathbf{V}})\|^2 + \|\nabla\times\underline{\mathbf{V}}\|^2. \end{aligned} \tag{3.1}$$

To define the least-squares method, we need a suitable minimization problem.

Let $\mathbf{X} := H^1(\Omega)^{n^2} \times H^1(\Omega)^n \times [H^1(\Omega)\cap L^2_0(\Omega)]$ and

$$\mathbf{V} := \{\ (\underline{\mathbf{U}}, \mathbf{u}, p, \underline{\mathbf{V}}, \mathbf{v}, q) \in \mathbf{X}\times\mathbf{X} \mid \mathbf{u} = \mathbf{0}, \mathbf{v} = \mathbf{0}, \mathbf{n}\times\underline{\mathbf{U}} = \underline{\mathbf{0}}, \mathbf{n}\times\underline{\mathbf{V}} = \underline{\mathbf{0}} \text{ on } \Gamma\ \}. \tag{3.2}$$

Then the least-squares principle is to find $(\underline{\mathbf{U}}, \mathbf{u}, p, \underline{\mathbf{V}}, \mathbf{v}, q) \in \mathbf{V}$ such that

$$\mathcal{F}_1(\underline{\mathbf{U}}, \mathbf{u}, p, \underline{\mathbf{V}}, \mathbf{v}, q : \mathbf{u}_d) = \inf_{(\tau,\mathbf{w},r,\psi,\mathbf{x},x)\in\mathbf{V}} \mathcal{F}_1(\tau, \mathbf{w}, r, \psi, \mathbf{x}, x : \mathbf{u}_d).$$

It is easy to see that the Euler-Lagrange equation for this minimization problem is given by the variational problem :

find $(\underline{\mathbf{U}}, \mathbf{u}, p, \underline{\mathbf{V}}, \mathbf{v}, q) \in \mathbf{V}$ such that

$$
\begin{aligned}
&B\Big((\underline{\mathbf{U}}, \mathbf{u}, p, \underline{\mathbf{V}}, \mathbf{v}, q), (\tilde{\underline{\mathbf{U}}}, \tilde{\mathbf{u}}, \tilde{p}, \tilde{\underline{\mathbf{V}}}, \tilde{\mathbf{v}}, \tilde{q})\Big) \qquad (3.3)\\
&= \Big(- (\nabla^t \underline{\mathbf{U}})^t + \frac{1}{\nu}(\underline{\mathbf{U}} + \underline{\mathbf{U}}_0)^t(\mathbf{u}+\mathbf{u}_0) + \nabla p - \frac{\mathbf{v}}{\nu\beta},\\
&\qquad - (\nabla^t \tilde{\underline{\mathbf{U}}})^t + \frac{1}{\nu}\tilde{\underline{\mathbf{U}}}^t(\mathbf{u}+\mathbf{u}_0) + \frac{1}{\nu}(\underline{\mathbf{U}} + \underline{\mathbf{U}}_0)^t\tilde{\mathbf{u}} + \nabla\tilde{p} - \frac{\tilde{\mathbf{v}}}{\nu\beta}\Big)\\
&\quad + (\nabla^t\mathbf{u}, \nabla^t\tilde{\mathbf{u}}) + (\underline{\mathbf{U}} - \nabla\mathbf{u}^t, \tilde{\underline{\mathbf{U}}} - \nabla\tilde{\mathbf{u}}^t) + (\nabla(\mathrm{tr}\underline{\mathbf{U}}), \nabla(\mathrm{tr}\tilde{\underline{\mathbf{U}}})) + (\nabla\times\underline{\mathbf{U}}, \nabla\times\tilde{\underline{\mathbf{U}}})\\
&\quad + \Big(- (\nabla^t \underline{\mathbf{V}})^t + \frac{1}{\nu}(\underline{\mathbf{U}} + \underline{\mathbf{U}}_0)(\mathbf{v}+\mathbf{v}_0) - \frac{1}{\nu}(\underline{\mathbf{V}} + \underline{\mathbf{V}}_0)^t(\mathbf{u}+\mathbf{u}_0) + \nabla q + \frac{1}{\nu}\mathbf{u},\\
&\qquad - (\nabla^t \tilde{\underline{\mathbf{V}}})^t + \frac{1}{\nu}\tilde{\underline{\mathbf{U}}}(\mathbf{v}+\mathbf{v}_0) + \frac{1}{\nu}(\underline{\mathbf{U}} + \underline{\mathbf{U}}_0)\tilde{\mathbf{v}}\\
&\qquad\qquad - \frac{1}{\nu}\tilde{\underline{\mathbf{V}}}^t(\mathbf{u}+\mathbf{u}_0) - \frac{1}{\nu}(\underline{\mathbf{V}} + \underline{\mathbf{V}}_0)^t\tilde{\mathbf{u}} + \nabla\tilde{q} + \frac{1}{\nu}\tilde{\mathbf{u}}\Big)\\
&\quad + (\nabla^t\mathbf{v}, \nabla^t\tilde{\mathbf{v}}) + (\underline{\mathbf{V}} - \nabla\mathbf{v}^t, \tilde{\underline{\mathbf{V}}} - \nabla\tilde{\mathbf{v}}^t) + (\nabla(\mathrm{tr}\underline{\mathbf{V}}), \nabla(\mathrm{tr}\tilde{\underline{\mathbf{V}}})) + (\nabla\times\underline{\mathbf{V}}, \nabla\times\tilde{\underline{\mathbf{V}}})\\
&= 0
\end{aligned}
$$

for all $(\tilde{\underline{\mathbf{U}}}, \tilde{\mathbf{u}}, \tilde{p}, \tilde{\underline{\mathbf{V}}}, \tilde{\mathbf{v}}, \tilde{q}) \in \mathbf{V}$.

Let $\mathbf{V}_h$ denote a finite-dimensional subspace of $\mathbf{V}$. Then the least-squares discretization method of the optimal control problem for the Navier-Stokes equations is defined by the following discrete variational problem:

find $(\underline{\mathbf{U}}^h, \mathbf{u}^h, p^h, \underline{\mathbf{V}}^h, \mathbf{v}^h, q^h) \in \mathbf{V}_h$ such that

$$
B((\underline{\mathbf{U}}^h, \mathbf{u}^h, p^h, \underline{\mathbf{V}}^h, \mathbf{v}^h, q^h), (\tilde{\underline{\mathbf{U}}}^h, \tilde{\mathbf{u}}^h, \tilde{p}^h, \tilde{\underline{\mathbf{V}}}^h, \tilde{\mathbf{v}}^h, \tilde{q}^h)) = 0
$$

$$
\text{for all } (\tilde{\underline{\mathbf{U}}}^h, \tilde{\mathbf{u}}^h, \tilde{p}^h, \tilde{\underline{\mathbf{V}}}^h, \tilde{\mathbf{v}}^h, \tilde{q}^h) \in \mathbf{V}_h \qquad (3.4)
$$

It is easy to see that the discrete variational problem (3.4) corresponds to the necessary condition for the following discrete least-squares principle for (3.1):

find $(\underline{\mathbf{U}}^h, \mathbf{u}^h, p^h, \underline{\mathbf{V}}^h, \mathbf{v}^h, q^h) \in \mathbf{V}_h$ such that

$$
\mathcal{F}_1(\underline{\mathbf{U}}^h, \mathbf{u}^h, p^h, \underline{\mathbf{V}}^h, \mathbf{v}^h, q^h : \mathbf{u}_d) \le \mathcal{F}_1(\tilde{\underline{\mathbf{U}}}^h, \tilde{\mathbf{u}}^h, \tilde{p}^h, \tilde{\underline{\mathbf{V}}}^h, \tilde{\mathbf{v}}^h, \tilde{q}^h : \mathbf{u}_d)
$$

$$
\text{for all } (\tilde{\underline{\mathbf{U}}}^h, \tilde{\mathbf{u}}^h, \tilde{p}^h, \tilde{\underline{\mathbf{V}}}^h, \tilde{\mathbf{v}}^h, \tilde{q}^h) \in \mathbf{V}_h \qquad (3.5)
$$

For space $\mathbf{V}_h$, we assume the following approximation property: there exists an integer $d \ge 1$ such that, for all $\underline{\mathbf{U}} \in H^{d+1}(\Omega)^{n^2}$, $\mathbf{u} \in H^{d+1}(\Omega)^n$, $p \in H^{d+1}(\Omega)$, $\underline{\mathbf{V}} \in H^{d+1}(\Omega)^{n^2}$, $\mathbf{v} \in H^{d+1}(\Omega)^n$ and $q \in H^{d+1}(\Omega)$, one can find $(\underline{\mathbf{U}}^h, \mathbf{u}^h, p^h, \underline{\mathbf{V}}^h, \mathbf{v}^h, q^h) \in \mathbf{V}_h$ such that

$$
\begin{aligned}
&\|\underline{\mathbf{U}} - \underline{\mathbf{U}}^h\|_\mu + \|\mathbf{u} - \mathbf{u}^h\|_\mu + \|p - p^h\|_\mu + \|\underline{\mathbf{V}} - \underline{\mathbf{V}}^h\|_\mu + \|\mathbf{v} - \mathbf{v}^h\|_\mu + \|q - q^h\|_\mu\\
&\qquad \le Ch^{d+1-\mu}\left(\|\underline{\mathbf{U}}\|_{d+1} + \|\mathbf{u}\|_{d+1} + \|p\|_{d+1} + \|\underline{\mathbf{V}}\|_{d+1} + \|\mathbf{v}\|_{d+1} + \|q\|_{d+1}\right), \qquad (3.6)
\end{aligned}
$$

$\mu = 0, 1$. Note, for example, that (3.6) can be satisfied with $d = 1$ by choosing continuous piecewise linears for all variables.

3.2. **Discretization error estimates.** The main goal of this section is to derive error estimates for least-squares method (3.4). For this purpose, we show how to cast nonlinear problems (3.3) and (3.4) in the respective canonical forms

$$F(\lambda,\mathcal{U}) \equiv \mathcal{U} + T \cdot G(\lambda,\mathcal{U}) = \mathbf{0} \tag{3.7}$$

and

$$F^h(\lambda,\mathcal{U}^h) \equiv \mathcal{U}^h + T_h \cdot G(\lambda,\mathcal{U}^h) = \mathbf{0}. \tag{3.8}$$

The following function spaces will be needed below (with m representing some nonnegative integer):

$$\mathbf{V}^m = [H^{m+1}(\Omega)^{n^2} \times H^{m+1}(\Omega)^n \times H^{m+1}(\Omega)]^2 \cap \mathbf{V}, \tag{3.9}$$

$$\mathbf{Y} = \mathbf{V}^*, \tag{3.10}$$

$$\mathbf{Z} = [L^{3/2}(\Omega)^{n^2} \times L^{3/2}(\Omega)^n \times L^{3/2}(\Omega)]^2, \tag{3.11}$$

where $\mathbf{V}^*$ denotes the dual of $\mathbf{V}$ with respect to the L^2 inner product.

We make identifications $\mathcal{U} = (\underline{\mathbf{U}}, \mathbf{u}, p, \underline{\mathbf{V}}, \mathbf{v}, q)$, $\mathcal{U}^h = (\underline{\mathbf{U}}^h, \mathbf{u}^h, p^h, \underline{\mathbf{V}}^h, \mathbf{v}^h, q^h)$, $\mathcal{V} = (\tilde{\underline{\mathbf{U}}}, \tilde{\mathbf{u}}, \tilde{p}, \tilde{\underline{\mathbf{V}}}, \tilde{\mathbf{v}}, \tilde{q})$, $\mathcal{V}^h = (\tilde{\underline{\mathbf{U}}}^h, \tilde{\mathbf{u}}^h, \tilde{p}^h, \tilde{\underline{\mathbf{V}}}^h, \tilde{\mathbf{v}}^h, \tilde{q}^h)$ and $\lambda = \frac{1}{\nu}$, and we assume that $\lambda \in \Lambda$, where Λ is a compact subset of $\mathbb{R}^+$. We then introduce the following:

$T : \mathbf{Y} \mapsto \mathbf{V}$ defined by $\mathcal{U} = T\mathbf{g}$ for $\mathbf{g} \in \mathbf{Y}$ if and only if

$$\begin{aligned} B_S(\mathcal{U},\mathcal{V}) \equiv\ & (-(\nabla^t\underline{\mathbf{U}})^t + \nabla p, -(\nabla^t\tilde{\underline{\mathbf{U}}})^t + \nabla\tilde{p}) + (\nabla^t\mathbf{u}, \nabla^t\tilde{\mathbf{u}}) \\ & + (\underline{\mathbf{U}} - \nabla\mathbf{u}^t, \tilde{\underline{\mathbf{U}}} - \nabla\tilde{\mathbf{u}}^t) + (\nabla(\mathrm{tr}\underline{\mathbf{U}}), \nabla(\mathrm{tr}\tilde{\underline{\mathbf{U}}})) + (\nabla\times\underline{\mathbf{U}}, \nabla\times\tilde{\underline{\mathbf{U}}}) \\ & + (-(\nabla^t\underline{\mathbf{V}})^t + \nabla q, -(\nabla^t\tilde{\underline{\mathbf{V}}})^t + \nabla\tilde{q}) + (\nabla^t\mathbf{v}, \nabla^t\tilde{\mathbf{v}}) \\ & + (\underline{\mathbf{V}} - \nabla\mathbf{v}^t, \tilde{\underline{\mathbf{V}}} - \nabla\tilde{\mathbf{v}}^t) + (\nabla(\mathrm{tr}\underline{\mathbf{V}}), \nabla(\mathrm{tr}\tilde{\underline{\mathbf{V}}})) + (\nabla\times\underline{\mathbf{V}}, \nabla\times\tilde{\underline{\mathbf{V}}}) \\ =\ & (\mathbf{g}_1, \tilde{\underline{\mathbf{U}}}) + (\mathbf{g}_2, \tilde{\mathbf{u}}) + (\mathbf{g}_3, \tilde{p}) + (\mathbf{g}_4, \tilde{\underline{\mathbf{V}}}) + (\mathbf{g}_5, \tilde{\mathbf{v}}) + (\mathbf{g}_6, \tilde{q}) \end{aligned} \tag{3.12}$$

for all $(\tilde{\underline{\mathbf{U}}}, \tilde{\mathbf{u}}, \tilde{p}, \tilde{\underline{\mathbf{V}}}, \tilde{\mathbf{v}}, \tilde{q}) \in \mathbf{V}$,

$T_h : \mathbf{Y} \mapsto \mathbf{V}_h$ defined by $\mathcal{U}^h = T\mathbf{g}$ for $\mathbf{g} \in \mathbf{Y}$ if and only if

$$\begin{aligned} B_S(\mathcal{U}^h,\mathcal{V}^h) = (\mathbf{g}_1, \tilde{\underline{\mathbf{U}}}^h) + (\mathbf{g}_2, \tilde{\mathbf{u}}^h) + (\mathbf{g}_3, \tilde{p}^h) + (\mathbf{g}_4, \tilde{\underline{\mathbf{V}}}^h) + (\mathbf{g}_5, \tilde{\mathbf{v}}^h) + (\mathbf{g}_6, \tilde{q}^h) \\ \text{for all } (\tilde{\underline{\mathbf{U}}}^h, \tilde{\mathbf{u}}^h, \tilde{p}^h, \tilde{\underline{\mathbf{V}}}^h, \tilde{\mathbf{v}}^h, \tilde{q}^h) \in \mathbf{V}_h \end{aligned} \tag{3.13}$$

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

$$
\begin{aligned}
&(\mathbf{g}_1, \underline{\tilde{\mathbf{U}}}) + (\mathbf{g}_2, \tilde{\mathbf{u}}) + (\mathbf{g}_3, \tilde{p}) + (\mathbf{g}_4, \underline{\tilde{\mathbf{V}}}) + (\mathbf{g}_5, \tilde{\mathbf{v}}) + (\mathbf{g}_6, \tilde{q}) \qquad (3.14)\\
&= \Big(-(\nabla^t \underline{\mathbf{U}})^t + \nabla p, \frac{1}{\nu}\underline{\tilde{\mathbf{U}}}^t(\mathbf{u}+\mathbf{u}_0) + \frac{1}{\nu}(\underline{\mathbf{U}}+\underline{\mathbf{U}}_0)^t\tilde{\mathbf{u}} - \frac{\tilde{\mathbf{v}}}{\nu\beta}\Big)\\
&\quad + \Big(\frac{1}{\nu}(\underline{\mathbf{U}}+\underline{\mathbf{U}}_0)^t(\mathbf{u}+\mathbf{u}_0) - \frac{\mathbf{v}}{\nu\beta},\\
&\qquad\qquad -(\nabla^t\underline{\tilde{\mathbf{U}}})^t + \frac{1}{\nu}\underline{\tilde{\mathbf{U}}}^t(\mathbf{u}+\mathbf{u}_0) + \frac{1}{\nu}(\underline{\mathbf{U}}+\underline{\mathbf{U}}_0)^t\tilde{\mathbf{u}} + \nabla\tilde{p} - \frac{\tilde{\mathbf{v}}}{\nu\beta}\Big)\\
&\quad + \Big(-(\nabla^t\underline{\mathbf{V}})^t + \nabla q,\\
&\qquad\qquad \frac{1}{\nu}\underline{\tilde{\mathbf{U}}}(\mathbf{v}+\mathbf{v}_0) + \frac{1}{\nu}(\underline{\mathbf{U}}+\underline{\mathbf{U}}_0)\tilde{\mathbf{v}} - \frac{1}{\nu}\underline{\tilde{\mathbf{V}}}^t(\mathbf{u}+\mathbf{u}_0) - \frac{1}{\nu}(\underline{\mathbf{V}}+\underline{\mathbf{V}}_0)^t\tilde{\mathbf{u}} + \frac{1}{\nu}\tilde{\mathbf{u}}\Big)\\
&\quad + \Big(\frac{1}{\nu}(\underline{\mathbf{U}}+\underline{\mathbf{U}}_0)(\mathbf{v}+\mathbf{v}_0) - \frac{1}{\nu}(\underline{\mathbf{V}}+\underline{\mathbf{V}}_0)^t(\mathbf{u}+\mathbf{u}_0) + \frac{1}{\nu}\mathbf{u},\\
&\qquad\qquad -(\nabla^t\underline{\tilde{\mathbf{V}}})^t + \frac{1}{\nu}\underline{\tilde{\mathbf{U}}}(\mathbf{v}+\mathbf{v}_0) + \frac{1}{\nu}(\underline{\mathbf{U}}+\underline{\mathbf{U}}_0)\tilde{\mathbf{v}}\\
&\qquad\qquad\qquad - \frac{1}{\nu}\underline{\tilde{\mathbf{V}}}^t(\mathbf{u}+\mathbf{u}_0) - \frac{1}{\nu}(\underline{\mathbf{V}}+\underline{\mathbf{V}}_0)^t\tilde{\mathbf{u}} + \nabla\tilde{q} + \frac{1}{\nu}\tilde{\mathbf{u}}\Big)
\end{aligned}
$$

for all $(\underline{\tilde{\mathbf{U}}}, \tilde{\mathbf{u}}, \tilde{p}, \underline{\tilde{\mathbf{V}}}, \tilde{\mathbf{v}}, \tilde{q}) \in \mathbf{V}$.

Lemma 3.1. *Assume that T, T_h, and G are defined by* (3.12)*,* (3.13)*, and* (3.14)*, respectively. Then nonlinear problem* (3.3) *is equivalent to* (3.7) *and discrete nonlinear problem* (3.4) *is equivalent to* (3.8).

Proof. Assume that $\mathcal{U} = (\underline{\mathbf{U}}, \mathbf{u}, p, \underline{\mathbf{V}}, \mathbf{v}, q)$ solves problem (3.7) with T and G given by (3.12) and (3.14), respectively. Then $\mathcal{U} = -T\mathbf{g}$ if and only if

$$B_S(\mathcal{U}, \mathcal{V}) = (\mathbf{g}, \mathcal{V}) \quad \text{for all } \mathcal{V} \in \mathbf{V}$$

and $\mathbf{g} = G(\lambda, \mathcal{U})$ if and only if (3.14) holds. It follows that $\mathcal{U}$ also solves variational problem (3.3). Conversely, if $\mathcal{U}$ solves (3.3), let $\mathbf{g}$ be defined by (3.14). Then $B_S(\mathcal{U}, \mathcal{V}) = (\mathbf{g}, \mathcal{V})$ for all $\mathcal{V} \in \mathbf{V}$, i.e., $\mathcal{U} = -T\mathbf{g}$. Thus, (3.3) and (3.7) are equivalent. Proof of the equivlence of (3.4) and (3.8) is identical. □

Error estimates for least-squares method (3.4) will now be derived from the abstract approximation theory of [10]. Below we state the main result of this theory for general T and T_h, but otherwise specialized to our needs. Here we let $D_{\mathcal{U}}G(\lambda, \mathcal{U})$ and $D_{\mathcal{U}}F(\lambda, \mathcal{U})$ denote the Fréchet derivative of G and F with respect to $\mathcal{U}$. We refer to $\{(\lambda, \mathcal{U}(\lambda)) | \lambda \in \Lambda\}$ as a regular branch of solutions of (3.7) if $\mathcal{U} = \mathcal{U}(\lambda)$ is a weak solution of (3.7) for each $\lambda \in \Lambda$, $\lambda \mapsto \mathcal{U}(\lambda)$ is a continuous map $\Lambda \mapsto \mathbf{V}$, and $D_{\mathcal{U}}F(\lambda, \mathcal{U})$ is an isomorphism of $\mathbf{V}$.

Theorem 3.1. *Let $F(\lambda, \mathcal{U}) = \mathbf{0}$ denote abstract form* (3.7) *and assume that $\{(\lambda, \mathcal{U}(\lambda)) | \lambda \in \Lambda\}$ is a branch of regular solutions of* (3.7)*. Furthermore, assume that $T \in L(\mathbf{Y}, \mathbf{V})$, that G is a C^2 map $\Lambda \times \mathbf{V} \mapsto \mathbf{Y}$ such that all second derivatives of G are bounded on bounded* *subsets of* $\Lambda \times \mathbf{V}$*, and that there exists a space* $\mathbf{Z} \subset \mathbf{Y}$*, with continuous imbedding, such that* $D_{\mathcal{U}}G(\lambda,\mathcal{U}) \in L(\mathbf{V},\mathbf{Z})$ *for all* $\lambda \in \Lambda$ *and* $\mathcal{U} \in \mathbf{V}$*. If approximate problem* (3.8) *is such that*

$$\lim_{h\to 0} \|(T - T_h)\mathbf{g}\|_{\mathbf{V}} = 0 \tag{3.15}$$

for all $\mathbf{g} \in \mathbf{Y}$ *and*

$$\lim_{h\to 0} \|(T - T_h)\|_{L(\mathbf{Z},\mathbf{V})} = 0. \tag{3.16}$$

Then:

1. there exists a neighborhood $\mathcal{O}$ *of the origin in* $\mathbf{V}$ *and, for* h *sufficiently small, a unique* C^2 *function* $\lambda \mapsto \mathcal{U}^h(\lambda) \in \mathbf{V}_h$ *such that* $\{(\lambda,\mathcal{U}^h(\lambda))|\lambda \in \Lambda\}$ *is a branch of regular solutions of discrete problem* (3.8) *and* $\mathcal{U}(\lambda) - \mathcal{U}^h(\lambda) \in \mathcal{O}$ *for all* $\lambda \in \Lambda$ *;*

2. for all $\lambda \in \Lambda$ *we have*

$$\|\mathcal{U}^h(\lambda) - \mathcal{U}(\lambda)\|_{\mathbf{V}} \le C\|(T - T_h)G(\lambda,\mathcal{U}(\lambda))\|_{\mathbf{V}}; \tag{3.17}$$

3. if the regular branch is such that $\mathcal{U}(\lambda) \in \mathbf{V}^m$ *for some integer* $m \ge 1$ *and* $\tilde{d} \equiv \min\{d,m\}$*, where d is the largest integer satisfying* (3.6)*, then*

$$\begin{aligned} &\|\underline{\mathbf{U}}(\lambda) - \underline{\mathbf{U}}^h(\lambda)\|_1 + \|\mathbf{u}(\lambda) - \mathbf{u}^h(\lambda)\|_1 + \|p(\lambda) - p^h(\lambda)\|_1 \\ &\qquad + \|\underline{\mathbf{V}}(\lambda) - \underline{\mathbf{V}}^h(\lambda)\|_1 + \|\mathbf{v}(\lambda) - \mathbf{v}^h(\lambda)\|_1 + \|q(\lambda) - q^h(\lambda)\|_1 \\ &\le Ch^{\tilde{d}}\left(\|\underline{\mathbf{U}}(\lambda)\|_{\tilde{d}+1} + \|\mathbf{u}(\lambda)\|_{\tilde{d}+1} + \|p(\lambda)\|_{\tilde{d}+1} + \|\underline{\mathbf{V}}(\lambda)\|_{\tilde{d}+1} + \|\mathbf{v}(\lambda)\|_{\tilde{d}+1} + \|q(\lambda)\|_{\tilde{d}+1}\right) \end{aligned} \tag{3.18}$$

In the next few lemmas, we verify the hypotheses of Theorem3.1 for our least-squares formulation. We begin by establishing essential properties of operators T and T_h, which we assume, for this and the next section, are defined by (3.12) and (3.13), respectively.

Lemma 3.2. $T \in L(\mathbf{Y},\mathbf{V})$ *and* $T_h \in L(\mathbf{Y},\mathbf{V}_h)$*.*

Proof. From $B_S(\cdot,\cdot)$ is continuous and coercive on $\mathbf{V} \times \mathbf{V}$ (see [6]) and, by virtue of the inclusion $\mathbf{V}_h \subset \mathbf{V}$, it is also continuous and coercive on $\mathbf{V}_h \times \mathbf{V}_h$. Furthermore, for each $\mathbf{g} \in \mathbf{Y}$, $(\mathbf{g},\mathcal{V})$ defines a continuous functional on $\mathbf{V}$. Thus, the Lax-Milgram theorem implies that, for all $\mathbf{g} \in \mathbf{Y}$, variational problems (3.12) and (3.13) have unique respective solutions $\mathcal{U} \in \mathbf{V}$ and $\mathcal{U}_h \in \mathbf{V}_h$, i.e., $T : \mathbf{Y} \mapsto \mathbf{V}$ and $T_h : \mathbf{Y} \mapsto \mathbf{V}_h$ are well-defined linear operators. From

$$C\|\mathcal{U}\|^2_{\mathbf{V}} \le B_S(\mathcal{U},\mathcal{U}) = (\mathbf{g},\mathcal{U}) \le \|\mathbf{g}\|_{\mathbf{Y}}\|\mathcal{U}\|_{\mathbf{V}},$$

it follows that

$$\|T\mathbf{g}\|_{\mathbf{V}} = \|\mathcal{U}\|_{\mathbf{V}} \le C\|\mathbf{g}\|_{\mathbf{Y}};$$

i.e., T is in $L(\mathbf{Y},\mathbf{V})$. The proof that $T_h \in L(\mathbf{Y},\mathbf{V}_h)$ is similar. □

Before continuing with the approximation properties of T_h, consider the choice of $\mathbf{Y}$ and $\mathbf{Z}$ in (3.10) and (3.11). When $\mathbf{Z} \subset \mathbf{Y}$ with compact imbedding, the proof of (3.16) in Theorem3.1 can be simplified. Since $L^{3/2}(\Omega)$ is compactly imbedded the duals of $H^1_0(\Omega)$, $\mathbf{H}^1_t(\Omega) = \{\mathbf{v} \in H^1(\Omega)^n | \mathbf{n} \times \mathbf{v} = 0 \text{ on } \Gamma\}$, and $H^1(\Omega)$, the imbedding $\mathbf{Z} \subset \mathbf{Y}$ is compact. (see[3])

Lemma 3.3. *Convergence properties (3.15) and (3.16) hold. If, in addition,* $\mathbf{g} \in \mathbf{Y}$ *is such that* $T\mathbf{g} \in \mathbf{V}^m$ *for some* $m \geq 1$ *and* $\tilde{d} = \min(d, m)$*, where* d *is the largest integer satisfying (3.6), then*

$$\|(T - T_h)\mathbf{g}\|_{\mathbf{V}} \leq Ch^{\tilde{d}}\|T\mathbf{g}\|_{\mathbf{V}^{\tilde{d}+1}}. \tag{3.19}$$

Proof. It is similar to Lemma 3 in [3]. □

The only hypotheses of Theorem3.1 that remain to be verified are the assumptions concerning the nonlinear operator G. For this purpose, we need the weak and strong forms of the first Fréchet derivative $D_{\mathcal{U}}G(\lambda,\mathcal{U})$ and second Fréchet derivative $D^2_{\mathcal{U}}G(\lambda,\mathcal{U})$. To determine the weak form of $D_{\mathcal{U}}G(\lambda,\mathcal{U})$, let $\hat{\mathcal{U}} \in \mathbf{V}$, substitute $\mathcal{U} + \hat{\mathcal{U}}$ into (3.14), and expand about $\mathcal{U}$. This yields the following weak representation of $D_{\mathcal{U}}G(\lambda,\mathcal{U})$:

$D_{\mathcal{U}}G(\lambda,\mathcal{U}) : \Lambda \times \mathbf{V} \to \mathbf{Y}$ defined by $\mathbf{g} = D_{\mathcal{U}}G(\lambda,\mathcal{U})\hat{\mathcal{U}}$ for $\mathcal{U} \in \mathbf{V}$ if and only if

$$\begin{aligned}
&(\mathbf{g}_1, \tilde{\underline{\mathbf{U}}}) + (\mathbf{g}_2, \tilde{\mathbf{u}}) + (\mathbf{g}_3, \tilde{p}) + (\mathbf{g}_4, \tilde{\underline{\mathbf{V}}}) + (\mathbf{g}_5, \tilde{\mathbf{v}}) + (\mathbf{g}_6, \tilde{q})\\
&= \Big(-(\nabla^t \underline{\mathbf{U}})^t + \nabla p, \frac{1}{\nu}\tilde{\underline{\mathbf{U}}}^t\hat{\mathbf{u}} + \frac{1}{\nu}\hat{\underline{\mathbf{U}}}^t\tilde{\mathbf{u}}\Big)\\
&\quad + \Big(-(\nabla^t \hat{\underline{\mathbf{U}}})^t + \nabla \hat{p}, \frac{1}{\nu}\tilde{\underline{\mathbf{U}}}^t(\mathbf{u}+\mathbf{u}_0) + \frac{1}{\nu}(\underline{\mathbf{U}} + \underline{\mathbf{U}}_0)^t\tilde{\mathbf{u}} - \frac{\tilde{\mathbf{v}}}{\nu\beta}\Big)\\
&\quad + \Big(\frac{1}{\nu}(\underline{\mathbf{U}} + \underline{\mathbf{U}}_0)^t(\mathbf{u}+\mathbf{u}_0) - \frac{\mathbf{v}}{\nu\beta}, \frac{1}{\nu}\tilde{\underline{\mathbf{U}}}^t\hat{\mathbf{u}} + \frac{1}{\nu}\hat{\underline{\mathbf{U}}}^t\tilde{\mathbf{u}}\Big)\\
&\quad + \Big(\frac{1}{\nu}(\underline{\mathbf{U}} + \underline{\mathbf{U}}_0)^t\hat{\mathbf{u}} + \frac{1}{\nu}\hat{\underline{\mathbf{U}}}^t(\mathbf{u}+\mathbf{u}_0) - \frac{\hat{\mathbf{v}}}{\nu\beta},\\
&\qquad\qquad -(\nabla^t \tilde{\underline{\mathbf{U}}})^t + \frac{1}{\nu}\tilde{\underline{\mathbf{U}}}^t(\mathbf{u}+\mathbf{u}_0) + \frac{1}{\nu}(\underline{\mathbf{U}} + \underline{\mathbf{U}}_0)^t\tilde{\mathbf{u}} + \nabla\tilde{p} - \frac{\tilde{\mathbf{v}}}{\nu\beta}\Big)\\
&\quad + \Big(-(\nabla^t \underline{\mathbf{V}})^t + \nabla q, \frac{1}{\nu}\tilde{\underline{\mathbf{U}}}\hat{\mathbf{v}} + \frac{1}{\nu}\hat{\underline{\mathbf{U}}}\tilde{\mathbf{v}} - \frac{1}{\nu}\tilde{\underline{\mathbf{V}}}^t\hat{\mathbf{u}} - \frac{1}{\nu}\hat{\underline{\mathbf{V}}}^t\tilde{\mathbf{u}}\Big)\\
&\quad + \Big(-(\nabla^t \hat{\underline{\mathbf{V}}})^t + \nabla\hat{q},\\
&\qquad\qquad \frac{1}{\nu}\tilde{\underline{\mathbf{U}}}(\mathbf{v}+\mathbf{v}_0) + \frac{1}{\nu}(\underline{\mathbf{U}} + \underline{\mathbf{U}}_0)\tilde{\mathbf{v}} - \frac{1}{\nu}\tilde{\underline{\mathbf{V}}}^t(\mathbf{u}+\mathbf{u}_0) - \frac{1}{\nu}(\underline{\mathbf{V}} + \underline{\mathbf{V}}_0)^t\tilde{\mathbf{u}} + \frac{\tilde{\mathbf{u}}}{\nu}\Big)\\
&\quad + \Big(\frac{1}{\nu}(\underline{\mathbf{U}} + \underline{\mathbf{U}}_0)(\mathbf{v}+\mathbf{v}_0) - \frac{1}{\nu}(\underline{\mathbf{V}} + \underline{\mathbf{V}}_0)^t(\mathbf{u}+\mathbf{u}_0) + \frac{\mathbf{u}}{\nu},\\
&\qquad\qquad \frac{1}{\nu}\tilde{\underline{\mathbf{U}}}\hat{\mathbf{v}} + \frac{1}{\nu}\hat{\underline{\mathbf{U}}}\tilde{\mathbf{v}} - \frac{1}{\nu}\tilde{\underline{\mathbf{V}}}^t\hat{\mathbf{u}} - \frac{1}{\nu}\hat{\underline{\mathbf{V}}}^t\tilde{\mathbf{u}}\Big)\\
&\quad + \Big(\frac{1}{\nu}(\underline{\mathbf{U}} + \underline{\mathbf{U}}_0)\hat{\mathbf{v}} + \frac{1}{\nu}\hat{\underline{\mathbf{U}}}(\mathbf{v}+\mathbf{v}_0) - \frac{1}{\nu}(\underline{\mathbf{V}} + \underline{\mathbf{V}}_0)^t\hat{\mathbf{u}} - \frac{1}{\nu}\hat{\underline{\mathbf{V}}}^t(\mathbf{u}+\mathbf{u}_0) + \frac{\hat{\mathbf{u}}}{\nu},\\
&\quad -(\nabla^t \tilde{\underline{\mathbf{V}}})^t + \frac{1}{\nu}\tilde{\underline{\mathbf{U}}}(\mathbf{v}+\mathbf{v}_0) + \frac{1}{\nu}(\underline{\mathbf{U}} + \underline{\mathbf{U}}_0)\tilde{\mathbf{v}} - \frac{1}{\nu}\tilde{\underline{\mathbf{V}}}^t(\mathbf{u}+\mathbf{u}_0) - \frac{1}{\nu}(\underline{\mathbf{V}} + \underline{\mathbf{V}}_0)^t\tilde{\mathbf{u}} + \nabla\tilde{q} + \frac{\tilde{\mathbf{u}}}{\nu}\Big)
\end{aligned} \tag{3.20}$$

for all $(\tilde{\underline{\mathbf{U}}}, \tilde{\mathbf{u}}, \tilde{p}, \tilde{\underline{\mathbf{V}}}, \tilde{\mathbf{v}}, \tilde{q}) \in \mathbf{V}$.

The strong form of $D_{\mathcal{U}}G(\lambda,\mathcal{U})\hat{\mathcal{U}}$ can be found from (3.20) using standard integration by parts:

$$\begin{aligned}\mathbf{g}_1 = &\frac{1}{\nu}\hat{\mathbf{u}}\Big(-(\nabla^t\underline{\mathbf{U}})^t + \nabla p + \frac{1}{\nu}(\underline{\mathbf{U}}+\underline{\mathbf{U}}_0)^t(\mathbf{u}+\mathbf{u}_0) - \frac{\mathbf{v}}{\nu\beta}\Big)^t \\ &+ \frac{1}{\nu}(\mathbf{u}+\mathbf{u}_0)\Big(-(\nabla^t\hat{\underline{\mathbf{U}}})^t + \nabla\hat{p} + \frac{1}{\nu}(\underline{\mathbf{U}}+\underline{\mathbf{U}}_0)^t\hat{\mathbf{u}} + \frac{1}{\nu}\hat{\underline{\mathbf{U}}}^t(\mathbf{u}+\mathbf{u}_0) - \frac{\hat{\mathbf{v}}}{\nu\beta}\Big)^t \\ &+ \nabla\Big(\frac{1}{\nu}(\underline{\mathbf{U}}+\underline{\mathbf{U}}_0)^t\hat{\mathbf{u}} + \frac{1}{\nu}\hat{\underline{\mathbf{U}}}^t(\mathbf{u}+\mathbf{u}_0) - \frac{\hat{\mathbf{v}}}{\nu\beta}\Big)^t \\ &+ \frac{1}{\nu}\Big(-(\nabla^t\underline{\mathbf{V}})^t + \nabla q + \frac{1}{\nu}(\underline{\mathbf{U}}+\underline{\mathbf{U}}_0)(\mathbf{v}+\mathbf{v}_0) - \frac{1}{\nu}(\underline{\mathbf{V}}+\underline{\mathbf{V}}_0)^t(\mathbf{u}+\mathbf{u}_0) + \frac{\mathbf{u}}{\nu}\Big)\hat{\mathbf{v}}^t \\ &+ \frac{1}{\nu}\Big(-(\nabla^t\hat{\underline{\mathbf{V}}})^t + \nabla\hat{q} + \frac{1}{\nu}(\underline{\mathbf{U}}+\underline{\mathbf{U}}_0)\hat{\mathbf{v}} + \frac{1}{\nu}\hat{\underline{\mathbf{U}}}(\mathbf{v}+\mathbf{v}_0) \\ &\qquad - \frac{1}{\nu}(\underline{\mathbf{V}}+\underline{\mathbf{V}}_0)^t\hat{\mathbf{u}} - \frac{1}{\nu}\hat{\underline{\mathbf{V}}}^t(\mathbf{u}+\mathbf{u}_0) + \frac{\hat{\mathbf{u}}}{\nu}\Big)(\mathbf{v}+\mathbf{v}_0)^t\end{aligned} \tag{3.21}$$

$$\begin{aligned}\mathbf{g}_2 = &\frac{1}{\nu}\hat{\underline{\mathbf{U}}}\Big(-(\nabla^t\underline{\mathbf{U}})^t + \nabla p + \frac{1}{\nu}(\underline{\mathbf{U}}+\underline{\mathbf{U}}_0)^t(\mathbf{u}+\mathbf{u}_0) - \frac{\mathbf{v}}{\nu\beta}\Big) \\ &+ \frac{1}{\nu}(\underline{\mathbf{U}}+\underline{\mathbf{U}}_0)\Big(-(\nabla^t\hat{\underline{\mathbf{U}}})^t + \nabla\hat{p} + \frac{1}{\nu}(\underline{\mathbf{U}}+\underline{\mathbf{U}}_0)^t\hat{\mathbf{u}} + \frac{1}{\nu}\hat{\underline{\mathbf{U}}}^t(\mathbf{u}+\mathbf{u}_0) - \frac{\hat{\mathbf{v}}}{\nu\beta}\Big) \\ &- \frac{1}{\nu}\hat{\underline{\mathbf{V}}}\Big(-(\nabla^t\underline{\mathbf{V}})^t + \nabla q + \frac{1}{\nu}(\underline{\mathbf{U}}+\underline{\mathbf{U}}_0)(\mathbf{v}+\mathbf{v}_0) - \frac{1}{\nu}(\underline{\mathbf{V}}+\underline{\mathbf{V}}_0)^t(\mathbf{u}+\mathbf{u}_0) + \frac{\mathbf{u}}{\nu}\Big) \\ &- \frac{1}{\nu}(\underline{\mathbf{V}}+\underline{\mathbf{V}}_0)\Big(-(\nabla^t\hat{\underline{\mathbf{V}}})^t + \nabla\hat{q} + \frac{1}{\nu}(\underline{\mathbf{U}}+\underline{\mathbf{U}}_0)\hat{\mathbf{v}} + \frac{1}{\nu}\hat{\underline{\mathbf{U}}}(\mathbf{v}+\mathbf{v}_0) \\ &\qquad - \frac{1}{\nu}(\underline{\mathbf{V}}+\underline{\mathbf{V}}_0)^t\hat{\mathbf{u}} - \frac{1}{\nu}\hat{\underline{\mathbf{V}}}^t(\mathbf{u}+\mathbf{u}_0) + \frac{\hat{\mathbf{u}}}{\nu}\Big) \\ &+ \frac{1}{\nu}\Big(-(\nabla^t\hat{\underline{\mathbf{V}}})^t + \nabla\hat{q} + \frac{1}{\nu}(\underline{\mathbf{U}}+\underline{\mathbf{U}}_0)\hat{\mathbf{v}} \\ &\qquad + \frac{1}{\nu}\hat{\underline{\mathbf{U}}}(\mathbf{v}+\mathbf{v}_0) - \frac{1}{\nu}(\underline{\mathbf{V}}+\underline{\mathbf{V}}_0)^t\hat{\mathbf{u}} - \frac{1}{\nu}\hat{\underline{\mathbf{V}}}^t(\mathbf{u}+\mathbf{u}_0) + \frac{\hat{\mathbf{u}}}{\nu}\Big)\end{aligned} \tag{3.22}$$

$$\mathbf{g}_3 = -\nabla^t\Big(\frac{1}{\nu}(\underline{\mathbf{U}}+\underline{\mathbf{U}}_0)^t\hat{\mathbf{u}} + \frac{1}{\nu}\hat{\underline{\mathbf{U}}}^t(\mathbf{u}+\mathbf{u}_0) - \frac{\hat{\mathbf{v}}}{\nu\beta}\Big) \tag{3.23}$$

$$\begin{aligned}
\mathbf{g}_4 = &-\frac{1}{\nu}\hat{\mathbf{u}}\Big(-(\nabla^t\underline{\mathbf{V}})^t + \nabla q + \frac{1}{\nu}(\underline{\mathbf{U}}+\underline{\mathbf{U}}_0)(\mathbf{v}+\mathbf{v}_0) - \frac{1}{\nu}(\underline{\mathbf{V}}+\underline{\mathbf{V}}_0)^t(\mathbf{u}+\mathbf{u}_0) + \frac{\mathbf{u}}{\nu}\Big)^t \\
&-\frac{1}{\nu}(\mathbf{u}+\mathbf{u}_0)\Big(-(\nabla^t\underline{\hat{\mathbf{V}}})^t + \nabla\hat{q} + \frac{1}{\nu}(\underline{\mathbf{U}}+\underline{\mathbf{U}}_0)\hat{\mathbf{v}} \\
&\qquad + \frac{1}{\nu}\underline{\hat{\mathbf{U}}}(\mathbf{v}+\mathbf{v}_0) - \frac{1}{\nu}(\underline{\mathbf{V}}+\underline{\mathbf{V}}_0)^t\hat{\mathbf{u}} - \frac{1}{\nu}\underline{\hat{\mathbf{V}}}^t(\mathbf{u}+\mathbf{u}_0) + \frac{\hat{\mathbf{u}}}{\nu}\Big)^t \\
&+\nabla\Big(\frac{1}{\nu}(\underline{\mathbf{U}}+\underline{\mathbf{U}}_0)\hat{\mathbf{v}} + \frac{1}{\nu}\underline{\hat{\mathbf{U}}}(\mathbf{v}+\mathbf{v}_0) - \frac{1}{\nu}(\underline{\mathbf{V}}+\underline{\mathbf{V}}_0)^t\hat{\mathbf{u}} - \frac{1}{\nu}\underline{\hat{\mathbf{V}}}^t(\mathbf{u}+\mathbf{u}_0) + \frac{\hat{\mathbf{u}}}{\nu}\Big)^t
\end{aligned} \tag{3.24}$$

$$\begin{aligned}
\mathbf{g}_5 = &-\frac{1}{\nu\beta}\Big(-(\nabla^t\underline{\hat{\mathbf{U}}})^t + \nabla\hat{p} + \frac{1}{\nu}(\underline{\mathbf{U}}+\underline{\mathbf{U}}_0)^t\hat{\mathbf{u}} + \frac{1}{\nu}\underline{\hat{\mathbf{U}}}^t(\mathbf{u}+\mathbf{u}_0) - \frac{\hat{\mathbf{v}}}{\nu\beta}\Big) \\
&+\frac{1}{\nu}\underline{\hat{\mathbf{U}}}^t\Big(-(\nabla^t\underline{\mathbf{V}})^t + \nabla q + \frac{1}{\nu}(\underline{\mathbf{U}}+\underline{\mathbf{U}}_0)(\mathbf{v}+\mathbf{v}_0) - \frac{1}{\nu}(\underline{\mathbf{V}}+\underline{\mathbf{V}}_0)^t(\mathbf{u}+\mathbf{u}_0) + \frac{\mathbf{u}}{\nu}\Big) \\
&+\frac{1}{\nu}(\underline{\mathbf{U}}+\underline{\mathbf{U}}_0)^t\Big(-(\nabla^t\underline{\hat{\mathbf{V}}})^t + \nabla\hat{q} + \frac{1}{\nu}(\underline{\mathbf{U}}+\underline{\mathbf{U}}_0)\hat{\mathbf{v}} \\
&\qquad + \frac{1}{\nu}\underline{\hat{\mathbf{U}}}(\mathbf{v}+\mathbf{v}_0) - \frac{1}{\nu}(\underline{\mathbf{V}}+\underline{\mathbf{V}}_0)^t\hat{\mathbf{u}} - \frac{1}{\nu}\underline{\hat{\mathbf{V}}}^t(\mathbf{u}+\mathbf{u}_0) + \frac{\hat{\mathbf{u}}}{\nu}\Big)
\end{aligned} \tag{3.25}$$

$$\mathbf{g}_6 = -\nabla^t\Big(\frac{1}{\nu}(\underline{\mathbf{U}}+\underline{\mathbf{U}}_0)\hat{\mathbf{v}} + \frac{1}{\nu}\underline{\hat{\mathbf{U}}}(\mathbf{v}+\mathbf{v}_0) - \frac{1}{\nu}(\underline{\mathbf{V}}+\underline{\mathbf{V}}_0)^t\hat{\mathbf{u}} - \frac{1}{\nu}\underline{\hat{\mathbf{V}}}^t(\mathbf{u}+\mathbf{u}_0) + \frac{\hat{\mathbf{u}}}{\nu}\Big) \tag{3.26}$$

for all $(\underline{\tilde{\mathbf{U}}}, \tilde{\mathbf{u}}, \tilde{p}, \underline{\tilde{\mathbf{V}}}, \tilde{\mathbf{v}}, \tilde{q}) \in \mathbf{V}$.

Finally, the weak form of the second Fréchet derivative is $D^2_{\mathcal{U}}G(\lambda,\mathcal{U}) : \Lambda \times [\mathbf{V}\times\mathbf{V}] \to \mathbf{Y}$ defined by $\mathbf{g} = D^2_{\mathcal{U}}G(\lambda,\mathcal{U})[\hat{\mathcal{U}},\hat{\hat{\mathcal{U}}}]$ for $\mathcal{U} \in \mathbf{V}$ if and only if

$$\begin{aligned}
&(\mathbf{g}_1, \underline{\tilde{\mathbf{U}}}) + (\mathbf{g}_2, \tilde{\mathbf{u}}) + (\mathbf{g}_3, \tilde{p}) + (\mathbf{g}_4, \underline{\tilde{\mathbf{V}}}) + (\mathbf{g}_5, \tilde{\mathbf{v}}) + (\mathbf{g}_6, \tilde{q}) \\
&= \Big(-(\nabla^t\underline{\hat{\hat{\mathbf{U}}}})^t + \nabla\hat{\hat{p}} + \frac{1}{\nu}(\underline{\mathbf{U}}+\underline{\mathbf{U}}_0)^t\hat{\hat{\mathbf{u}}} + \frac{1}{\nu}\underline{\hat{\hat{\mathbf{U}}}}^t(\mathbf{u}+\mathbf{u}_0) - \frac{\hat{\hat{\mathbf{v}}}}{\nu\beta}, \frac{1}{\nu}\underline{\tilde{\mathbf{U}}}^t\hat{\mathbf{u}} + \frac{1}{\nu}\underline{\hat{\mathbf{U}}}^t\tilde{\mathbf{u}}\Big) \\
&\quad + \Big(-(\nabla^t\underline{\hat{\mathbf{U}}})^t + \nabla\hat{p} + \frac{1}{\nu}(\underline{\mathbf{U}}+\underline{\mathbf{U}}_0)^t\hat{\mathbf{u}} + \frac{1}{\nu}\underline{\hat{\mathbf{U}}}^t(\mathbf{u}+\mathbf{u}_0) - \frac{\hat{\mathbf{v}}}{\nu\beta}, \frac{1}{\nu}\underline{\tilde{\mathbf{U}}}^t\hat{\hat{\mathbf{u}}} + \frac{1}{\nu}\underline{\hat{\hat{\mathbf{U}}}}^t\tilde{\mathbf{u}}\Big) \\
&\quad + \Big(\frac{1}{\nu}\underline{\hat{\hat{\mathbf{U}}}}^t\hat{\mathbf{u}} + \frac{1}{\nu}\underline{\hat{\mathbf{U}}}^t\hat{\hat{\mathbf{u}}}, -(\nabla^t\underline{\tilde{\mathbf{U}}})^t + \frac{1}{\nu}\underline{\tilde{\mathbf{U}}}^t(\mathbf{u}+\mathbf{u}_0) + \frac{1}{\nu}(\underline{\mathbf{U}}+\underline{\mathbf{U}}_0)^t\tilde{\mathbf{u}} + \nabla\tilde{p} - \frac{\tilde{\mathbf{v}}}{\nu\beta}\Big) \\
&\quad + \Big(-(\nabla^t\underline{\hat{\hat{\mathbf{V}}}})^t + \nabla\hat{\hat{q}} + \frac{1}{\nu}(\underline{\mathbf{U}}+\underline{\mathbf{U}}_0)\hat{\hat{\mathbf{v}}} + \frac{1}{\nu}\underline{\hat{\hat{\mathbf{U}}}}(\mathbf{v}+\mathbf{v}_0) - \frac{1}{\nu}(\underline{\mathbf{V}}+\underline{\mathbf{V}}_0)^t\hat{\hat{\mathbf{u}}} \\
&\qquad - \frac{1}{\nu}\underline{\hat{\hat{\mathbf{V}}}}^t(\mathbf{u}+\mathbf{u}_0) + \frac{\hat{\hat{\mathbf{u}}}}{\nu}, \frac{1}{\nu}\underline{\tilde{\mathbf{U}}}\hat{\mathbf{v}} + \frac{1}{\nu}\underline{\hat{\mathbf{U}}}\tilde{\mathbf{v}} - \frac{1}{\nu}\underline{\tilde{\mathbf{V}}}^t\hat{\mathbf{u}} - \frac{1}{\nu}\underline{\hat{\mathbf{V}}}^t\tilde{\mathbf{u}}\Big)
\end{aligned}$$

$$+\Big(-(\nabla^t\hat{\underline{\mathbf{V}}})^t+\nabla\hat{q}+\frac{1}{\nu}(\underline{\mathbf{U}}+\underline{\mathbf{U}}_0)\hat{\mathbf{v}}+\frac{1}{\nu}\hat{\underline{\mathbf{U}}}(\mathbf{v}+\mathbf{v}_0)-\frac{1}{\nu}(\underline{\mathbf{V}}+\underline{\mathbf{V}}_0)^t\hat{\mathbf{u}}$$
$$-\frac{1}{\nu}\hat{\underline{\mathbf{V}}}^t(\mathbf{u}+\mathbf{u}_0)+\frac{\hat{\mathbf{u}}}{\nu},\quad \frac{1}{\nu}\tilde{\underline{\mathbf{U}}}\hat{\hat{\mathbf{v}}}+\frac{1}{\nu}\hat{\hat{\underline{\mathbf{U}}}}\tilde{\mathbf{v}}-\frac{1}{\nu}\tilde{\underline{\mathbf{V}}}^t\hat{\hat{\mathbf{u}}}-\frac{1}{\nu}\hat{\hat{\underline{\mathbf{V}}}}^t\tilde{\mathbf{u}}\Big)$$
$$+\Big(\frac{1}{\nu}\hat{\hat{\underline{\mathbf{U}}}}\hat{\mathbf{v}}+\frac{1}{\nu}\hat{\underline{\mathbf{U}}}\hat{\hat{\mathbf{v}}}-\frac{1}{\nu}\hat{\hat{\underline{\mathbf{V}}}}^t\hat{\mathbf{u}}-\frac{1}{\nu}\hat{\underline{\mathbf{V}}}^t\hat{\hat{\mathbf{u}}},-(\nabla^t\tilde{\underline{\mathbf{V}}})^t+\frac{1}{\nu}\tilde{\underline{\mathbf{U}}}(\mathbf{v}+\mathbf{v}_0)$$
$$+\frac{1}{\nu}(\underline{\mathbf{U}}+\underline{\mathbf{U}}_0)\tilde{\mathbf{v}}-\frac{1}{\nu}\tilde{\underline{\mathbf{V}}}^t(\mathbf{u}+\mathbf{u}_0)-\frac{1}{\nu}(\underline{\mathbf{V}}+\underline{\mathbf{V}}_0)^t\tilde{\mathbf{u}}+\nabla\tilde{q}+\frac{\tilde{\mathbf{u}}}{\nu}\Big) \tag{3.27}$$

for all $(\tilde{\underline{\mathbf{U}}},\tilde{\mathbf{u}},\tilde{p},\tilde{\underline{\mathbf{V}}},\tilde{\mathbf{v}},\tilde{q})\in\mathbf{V}$.

The next lemma summarizes technical results that we use below.

Lemma 3.4. *Let D_i denote the derivative with respect to the ith coordinate variable in $\mathbb{R}^n$, $1\le i\le n$, and assume that u,v,w, and z are in $H^1(\Omega)$. Then*

$$\left|\int_\Omega D_i u\, v\, w\, d\Omega\right|\le C\|u\|_1\|v\|_1\|w\|_1, \tag{3.28}$$

$1\le i\le n$, and

$$\left|\int_\Omega u\, v\, w\, z\, d\Omega\right|\le C\|u\|_1\|v\|_1\|w\|_1\|z\|_1. \tag{3.29}$$

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

$$\|uv\|_{0,3/2}\le C\|u\|_{0,2}\|v\|_{1,2}\quad \text{for all } u\in L^2(\Omega) \text{ and } v\in H^1(\Omega), \tag{3.30}$$

$$\|uvw\|_{0,3/2}\le C\|u\|_{1,2}\|v\|_{1,2}\|w\|_{1,2}\quad \forall u,v,w\in H^1(\Omega). \tag{3.31}$$

Proof. It is similar to Lemma 4 in [3]. The first part of the lemma follows easily from the imbedding $H^1(\Omega)\subset L^4(\Omega)$ in two and three dimensions and the Hölder inequality. The second part follows directly from a result in [10](see Corollary 1.1, p. 5). □

In the next lemma, we establish properties of G that are required for the validity of the approximation result in Theorem3.1.

Lemma 3.5. *Assume that mapping G is defined by* (3.14). *For $\mathbf{V},\mathbf{Y}$, and $\mathbf{Z}$ given by* (3.2), (3.10) *and* (3.11), *respectively, the following are true.*

1. *For all $\mathcal{U}\in\mathbf{V}$, $D_\mathcal{U}G(\lambda,\mathcal{U})\in L(\mathbf{V},\mathbf{Z})$.*
2. *The second Fréchet derivative $D^2_\mathcal{U}G(\lambda,\mathcal{U})$ is bounded on bounded subsets of $\Lambda\times\mathbf{V}$.*

Proof. To prove 1, consider strong form (3.21)–(3.26) of $D_\mathcal{U}G(\lambda,\mathcal{U})$. By assumption, $\mathcal{U}\in\mathbf{V}$; i.e., $\underline{\mathbf{U}}\in H^1(\Omega)^{n^2}$, $\mathbf{u}\in H^1(\Omega)^n$, $p\in H^1(\Omega)$, $\underline{\mathbf{V}}\in H^1(\Omega)^{n^2}$, $\mathbf{v}\in H^1(\Omega)^n$, and $q\in H^1(\Omega)$. Now each equation (3.21)–(3.25) and (3.26) consists of terms of the form D_iuv

and uvw, where u, v, and w belong to $H^1(\Omega)$, so the second part of Lemma3.4 implies that $(\mathbf{g}_1, \mathbf{g}_2, \mathbf{g}_3, \mathbf{g}_4, \mathbf{g}_5, \mathbf{g}_6) \in \mathbf{Z}$. Using (3.30) and (3.31), it also follows that

$$\|D_{\mathcal{U}}G(\lambda, \mathcal{U})\hat{\mathcal{U}}\|_{\mathbf{Z}} \leq C\|\hat{\mathcal{U}}\|_{\mathbf{V}}, \tag{3.32}$$

i.e., that $D_{\mathcal{U}}G(\lambda, \mathcal{U}) \in L(\mathbf{V}, \mathbf{Z})$.

To prove 2, consider weak form (3.27) of the second Fréchet derivative. Assume that $(\lambda, \mathcal{U})$ belongs to a bounded subset of $\Lambda \times \mathbf{V}$ and let $\hat{\mathcal{U}}$, $\hat{\hat{\mathcal{U}}} \in \mathbf{V}$ be arbitrary. Then it is not difficult to see that weak form (3.27) involves only terms of the form $D_i uvw$ and $uvwz$, where u, v, w, and z belong to $H^1(\Omega)$. Thus, each term can be estimated using (3.28) or (3.29):

$$\begin{aligned}
|(\mathbf{g}_1, \underline{\tilde{\mathbf{U}}})| &\leq C_1(\lambda, \mathcal{U}, \mathcal{U}_0)(\|\hat{\mathcal{U}}\|_{\mathbf{V}} + \|\hat{\hat{\mathcal{U}}}\|_{\mathbf{V}})\|\underline{\tilde{\mathbf{U}}}\|_1, \\
|(\mathbf{g}_2, \tilde{\mathbf{u}})| &\leq C_2(\lambda, \mathcal{U}, \mathcal{U}_0)(\|\hat{\mathcal{U}}\|_{\mathbf{V}} + \|\hat{\hat{\mathcal{U}}}\|_{\mathbf{V}})\|\tilde{\mathbf{u}}\|_1, \\
|(\mathbf{g}_3, \tilde{p})| &\leq C_3(\lambda, \mathcal{U}, \mathcal{U}_0)(\|\hat{\mathcal{U}}\|_{\mathbf{V}} + \|\hat{\hat{\mathcal{U}}}\|_{\mathbf{V}})\|\tilde{p}\|_1, \\
|(\mathbf{g}_4, \underline{\tilde{\mathbf{V}}})| &\leq C_4(\lambda, \mathcal{U}, \mathcal{U}_0)(\|\hat{\mathcal{U}}\|_{\mathbf{V}} + \|\hat{\hat{\mathcal{U}}}\|_{\mathbf{V}})\|\underline{\tilde{\mathbf{V}}}\|_1, \\
|(\mathbf{g}_5, \tilde{\mathbf{v}})| &\leq C_5(\lambda, \mathcal{U}, \mathcal{U}_0)(\|\hat{\mathcal{U}}\|_{\mathbf{V}} + \|\hat{\hat{\mathcal{U}}}\|_{\mathbf{V}})\|\tilde{\mathbf{v}}\|_1, \\
|(\mathbf{g}_6, \tilde{q})| &\leq C_6(\lambda, \mathcal{U}, \mathcal{U}_0)(\|\hat{\mathcal{U}}\|_{\mathbf{V}} + \|\hat{\hat{\mathcal{U}}}\|_{\mathbf{V}})\|\tilde{q}\|_1
\end{aligned}$$

where C_i is polynomial function of λ, $\|\mathcal{U}\|_{\mathbf{V}}$, and $\|\mathcal{U}_0\|_{\mathbf{V}}$. In combination with the fact that λ and $\|\mathcal{U}\|_{\mathbf{V}}$ are in bounded subsets of $\Lambda \times \mathbf{V}$, and that $\|\mathcal{U}_0\|_{\mathbf{V}}$ is fixed, it follows that $D^2_{\mathcal{U}}G(\lambda, \mathcal{U})$ is bounded in the norm of L(**V**, L(**V**, **Y**)). □

This completes verification of all assumptions of Theorem3.1. As a result, we can conclude that error estimates (3.17) and (3.18) hold for the least-squares finite element approximation as long as problem (3.3) has a regular branch of solutions with sufficient regularity.

REFERENCES

[1] P. B. BOCHEV, *Analysis of least-squares finite element methods for the Navier-Stokes equations*, Siam J. Numer. Anal., 34 (1997), pp. 1817–1844.

[2] P. B. BOCHEV AND M. D. GUNZBURGER, *Analysis of least-squares finite element methods for the Stokes equations*, Math. Comp., 63 (1994), pp. 479–506.

[3] P. B. BOCHEV, Z. CAI, T. A. MANTEUFFEL AND S. F. MCCORMICK, *Analysis of velocity-flux least-squares principles for the Navier-Stokes equations. Part I.*, SIAM J. Numer. Anal., 35 (1998), pp. 990–1009.

[4] P. B. BOCHEV, T. A. MANTEUFFEL AND S. F. MCCORMICK, *Analysis of velocity-flux least-squares principles for the Navier-Stokes equations. Part II.*, SIAM J. Numer. Anal., 36 (1999), pp. 1125–1144.

[5] J. H. BRAMBLE AND J. E. PASCIAK, *Least-squares methods for the Stokes equations on a discrete minus one inner product*, Jour. comput. appl. math, 74 (1996), pp. 155–173.

[6] Z. CAI, T. MANTEUFFEL, AND S. MCCORMICK, *First-order system least squares for the Stokes equations, with application to linear elasticity*, SIAM J. Numer. Anal., 34 (1997) pp. 1727–1741.

[7] P. CIARLET, *Finite element method for elliptic problems*, North-Holland, Amsterdam, 1978.

[8] Y. CHOI, S. D. KIM, H.-C. LEE AND B.-C. SHIN, *Analysis of first-order system least-squares for the optimal control problems for the Navier-Stokes equations*, J. KSIAM, vol.11 No.4, 55-68, 2007

[9] Y. CHOI, H.-C. LEE AND S. D. KIM, *Analysis and computations of least-squares method for optimal control problems for the Stokes equations*, J. Korean Math. Soc. 46 (2009), no. 5, 1007–1025

[10] V. GIRAULT AND P.-A. RAVIART, Finite element methods for Navier-Stokes equations, Springer, Berlin, 1986.

[11] M. D. GUNZBURGER, L. HOU, AND T. P. SVOBODNY , *Analysis and finite element approximation of optimal control problems for the stationary Navier-Stokes equations with distributed and Neumann controls*, Math. Comp., 57 (1991), pp. 123–151.

[12] H.-C. LEE AND Y. CHOI, *A least-squares method for optimal control problems for a second-order elliptic systems in two dimensions*, Jour. Math. Anal. Appl., 242 (2000), pp. 105-128.

[13] SOOROK RYU, HYUNG-CHUN LEE AND SANG DONG KIM, *First- order system least-squares methods for an optimal control problem by the Stokes flow*, SIAM J. Numer. Anal. 47 (2009), no. 2, 1524–1545.

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