CONVERGENCE OF THE NEWTON'S METHOD FOR AN OPTIMAL CONTROL PROBLEMS FOR NAVIER-STOKES EQUATIONS

YOUNGMI CHOI, SANGDONG KIM, AND HYUNG-CHUN LEE

ABSTRACT. We consider the Newton's method for an direct solver of the optimal control problems of the Navier-Stokes equations. We show that the finite element solutions of the optimal control problem for Stokes equations may be chosen as the initial guess for the quadratic convergence of Newton's algorithm applied to the optimal control problem for the Navier-Stokes equations provided there are sufficiently small mesh size h and the moderate Reynold's number.

1. Introduction

The optimal control problem we consider is to minimize the functional

$$(1) \qquad \mathcal{J}(\mathbf{u}, p, \mathbf{f}) = \frac{1}{2} \int_\Omega |\mathbf{u} - \mathbf{u}_d|^2 \, d\Omega + \frac{\beta}{2} \int_\Omega |\mathbf{f}|^2 \, d\Omega,$$

subject to the incompressible Navier-Stokes equations

$$(2) \qquad \begin{aligned} -\nu \Delta \mathbf{u} + (\mathbf{u} \cdot \nabla)\mathbf{u} + \nabla p &= \mathbf{f} \quad \text{in } \Omega, \\ \nabla \cdot \mathbf{u} &= 0 \quad \text{in } \Omega, \\ \mathbf{u} &= \mathbf{0} \quad \text{on } \partial\Omega, \end{aligned}$$

where $\mathbf{u}_d$ is a given desired function. Here, $\Omega \subset \mathbb{R}^n (n = 2 \text{ or } 3)$ 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 (2). The second term in (1) is added as a limiting the cost of

control and the positive penalty parameter β can be used to change the relative importance of the two terms appearing in the definition of the functional.

It is well known that Newton and Newton-like methods combining with other methods, such as continuation method, are the most popular iteration methods in the course of solving nonlinear equations (see, [6, 7, 8]). If the initial guess is chosen sufficiently nearby the exact solution of the given nonlinear differential equations, the quadratic convergence is guaranteed (see [4]). In the continuation method for solving Navier-Stokes equations, the solution of Stokes equations is chosen as the initial guess. Based on the convergence analysis for the Navier-Stokes equations [5], we analyze the convergence of the Newton's method for an optimal control problem of Navier-Stokes equations.

The plan of the paper is as follows. In the next section, we give a precise statement of the optimal control problem. Then we derive an optimality system. In Section 3, we review the quadratic convergence theory of Newton's method and rewrite the Newton's method in an equivalent iterative method. In Section 4, for the implementing of Newton's method by the finite element method we choose the initial guess as the finite element solution of the optimal control problem for the Stokes equations. We show that the Newton's sequence converges to the unique solution of the optimal control problem for the Navier-Stokes equations quadratically and finally provide an error analysis combining the quadratic convergence with the finite element error of the initial guess and Reynold's number.

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 $\mathbf{f} \in H^{-1}(\Omega)^n$ denote the distributed control. The state and control variables are also constrained to satisfy the system (2), which recast into the weak form:

$$(3) \qquad \nu a(\mathbf{u}, \mathbf{w}) + c(\mathbf{u}, \mathbf{u}, \mathbf{w}) + b(\mathbf{w}, p) = \langle \mathbf{f}, \mathbf{w} \rangle \quad \forall \mathbf{w} \in H_0^1(\Omega)^n,$$

$$(4) \qquad b(\mathbf{u}, q) = 0 \quad \forall q \in L^2(\Omega),$$

where

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

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

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

With $\mathcal{J}(\cdot, \cdot, \cdot)$ given by (1), the *admissibility set* $\mathcal{U}_{ad}$ is defined by

$$\mathcal{U}_{ad} = \{ (\mathbf{u}, p, \mathbf{f}) \in H_0^1(\Omega)^n \times L_0^2(\Omega) \times H^{-1}(\Omega)^n :$$
$$\mathcal{J}(\mathbf{u}, p, \mathbf{f}) < \infty \text{ and } (\mathbf{u}, p, \mathbf{f}) \text{ satisfies (3) and (4)} \}.$$

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

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

satisfying

$$\|\hat{\mathbf{u}} - \mathbf{u}\|_{H^1(\Omega)^n} + \|\hat{p} - p\|_{L^2(\Omega)} + \|\hat{\mathbf{f}} - \mathbf{f}\|_{H^{-1}(\Omega)^n} < \epsilon.$$

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

$$(5) \qquad \min_{(\mathbf{u}, p, \mathbf{f}) \in \mathcal{U}_{ad}} \mathcal{J}(\mathbf{u}, p, \mathbf{f}).$$

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), one may derive an optimality system of equations for the solution of (5). The constrained problem (5) 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

$$\nu a(\mathbf{u}, \tilde{\mathbf{u}}) + c(\mathbf{u}, \mathbf{u}, \tilde{\mathbf{u}}) + b(\tilde{\mathbf{u}}, p) - (\mathbf{f}, \tilde{\mathbf{u}}) = 0, \quad \forall \tilde{\mathbf{u}} \in H_0^1(\Omega)^n,$$

$$b(\mathbf{u}, \tilde{p}) = 0, \quad \forall \tilde{p} \in L_0^2(\Omega)^n,$$

$$\nu a(\mathbf{v}, \tilde{\mathbf{v}}) + c(\tilde{\mathbf{v}}, \mathbf{u}, \mathbf{v}) + c(\mathbf{u}, \tilde{\mathbf{v}}, \mathbf{v}) + b(\tilde{\mathbf{v}}, q) + (\mathbf{u} - \mathbf{u}_d, \tilde{\mathbf{v}}) = 0, \quad \forall \tilde{\mathbf{v}} \in H_0^1(\Omega)^n,$$

$$b(\mathbf{v}, \tilde{q}) = 0, \quad \forall \tilde{q} \in L_0^2(\Omega)^n,$$

$$(\beta \mathbf{f} - \mathbf{v}, \tilde{\mathbf{f}}) = 0, \quad \forall \tilde{\mathbf{f}} \in H^{-1}(\Omega)^n.$$

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.

Using the optimality condition $\mathbf{f} = \frac{\mathbf{v}}{\beta}$,

(6)
$$\begin{cases} \nu a(\mathbf{u}, \tilde{\mathbf{u}}) + c(\mathbf{u}, \mathbf{u}, \tilde{\mathbf{u}}) + b(\tilde{\mathbf{u}}, p) - \left(\frac{\mathbf{v}}{\beta}, \tilde{\mathbf{u}} \right) = 0, \ \forall \tilde{\mathbf{u}} \in H_0^1(\Omega)^n, \\[2mm] b(\mathbf{u}, \tilde{p}) = 0, \ \forall \tilde{p} \in L_0^2(\Omega)^n, \\[2mm] \nu a(\mathbf{v}, \tilde{\mathbf{v}}) + c(\tilde{\mathbf{v}}, \mathbf{u}, \mathbf{v}) + c(\mathbf{u}, \tilde{\mathbf{v}}, \mathbf{v}) + b(\tilde{\mathbf{v}}, q) + (\mathbf{u} - \mathbf{u}_d, \tilde{\mathbf{v}}) = 0, \ \forall \tilde{\mathbf{v}} \in H_0^1(\Omega)^n, \\[2mm] b(\mathbf{v}, \tilde{q}) = 0, \ \forall \tilde{q} \in L_0^2(\Omega)^n. \end{cases}$$

3. Newton's algorithm in a weak formulation

Let $L_0^2(\Omega)$ be the subspace of $L^2(\Omega)$ which consists of the functions of mean zero. The L^2 inner product and norm are denoted by $(\cdot, \cdot)$ and $\|\cdot\|$, respectively. Define the following function spaces:

$$\mathbf{X} := \mathbf{V} \times \mathcal{Q} \times \mathbf{V} \times \mathcal{Q}, \quad \text{where } \mathbf{V} = H_0^1(\Omega)^n, \ \mathcal{Q} = L_0^2(\Omega),$$

and

$$\mathbf{Y} := \mathbf{V}^* \times \mathbf{V}^*,$$

where $\mathbf{V}^*$ denotes the dual space of $\mathbf{V}$. With a little abuse of notation, let $\mathcal{U} := (\mathbf{u}, p, \mathbf{v}, q)$, $\mathcal{V} := (\tilde{\mathbf{u}}, \tilde{p}, \tilde{\mathbf{v}}, \tilde{q})$ and $\lambda = \frac{1}{\nu}$. Then the optimality system can be written as: find $\mathcal{U} \in \mathbf{X}$ such that, for all $\tilde{\mathbf{u}} \in \mathbf{V}$, $\tilde{p} \in \mathcal{Q}$, $\tilde{\mathbf{v}} \in \mathbf{V}$ and $\tilde{q} \in \mathcal{Q}$,

$$(7) \quad \begin{cases} a(\mathbf{u}, \tilde{\mathbf{u}}) + \lambda c(\mathbf{u}, \mathbf{u}, \tilde{\mathbf{u}}) + b(\tilde{\mathbf{u}}, p) - \lambda \left(\dfrac{\mathbf{v}}{\beta}, \tilde{\mathbf{u}} \right) &= 0, \\ b(\mathbf{u}, \tilde{p}) &= 0, \\ a(\mathbf{v}, \tilde{\mathbf{v}}) + \lambda c(\tilde{\mathbf{v}}, \mathbf{u}, \mathbf{v}) + \lambda c(\mathbf{u}, \tilde{\mathbf{v}}, \mathbf{v}) + b(\tilde{\mathbf{v}}, q) + \lambda(\mathbf{u} - \mathbf{u}_d, \tilde{\mathbf{v}}) &= 0, \\ b(\mathbf{v}, \tilde{q}) &= 0. \end{cases}$$

For the derivation of Newton's algorithm, we cast the optimality system (7) in the canonical form

$$(8) \quad F(\lambda, \mathcal{U}) \equiv \mathcal{U} + T \cdot G(\lambda, \mathcal{U}) = \mathbf{0},$$

in which a linear solution operator T for the Stokes equations is defined as $T : \mathbf{Y} \to \mathbf{X}$ by $\mathcal{U} = T\mathbf{g}$ for $\mathbf{g} = (g_1, g_2) \in \mathbf{Y}$ if and only if

$$(9) \quad \begin{aligned} B(\mathcal{U}, \mathcal{V}) &:= a(\mathbf{u}, \tilde{\mathbf{u}}) + b(\tilde{\mathbf{u}}, p) + b(\mathbf{u}, \tilde{p}) + a(\mathbf{v}, \tilde{\mathbf{v}}) + b(\tilde{\mathbf{v}}, q) + b(\mathbf{v}, \tilde{q}) \\ &= (g_1, \tilde{\mathbf{u}}) + (g_2, \tilde{\mathbf{v}}), \ \forall \mathcal{V} \in \mathbf{X}, \end{aligned}$$

and a C^∞ operator G for the optimality system (7) is defined as $G : \Lambda \times \mathbf{X} \to \mathbf{Y}$, by $\mathbf{g} = G(\lambda, \mathcal{U})$ for $(\lambda, \mathcal{U}) \in \Lambda \times \mathbf{X}$ if and only if

(10)
$$(g_1, \tilde{\mathbf{u}}) + (g_2, \tilde{\mathbf{v}})$$

$$= -\lambda c(\mathbf{u}, \mathbf{u}, \tilde{\mathbf{u}}) + \lambda \left(\frac{\mathbf{v}}{\beta}, \tilde{\mathbf{u}} \right) - \lambda c(\tilde{\mathbf{v}}, \mathbf{u}, \mathbf{v}) + \lambda c(\mathbf{u}, \mathbf{v}, \tilde{\mathbf{v}}) - \lambda(\mathbf{u} - \mathbf{u}_d, \tilde{\mathbf{v}}), \ \forall \mathcal{V} \in \mathbf{X},$$

where Λ is a compact interval in $(0, \infty)$. Now assume that $\mathcal{U}$ is the solution of problem (8) with T and G given by (9) and (10), respectively. Then $\mathcal{U} = T\mathbf{g}$ if and only if $B(\mathcal{U}, \mathcal{V}) = (\mathbf{g}, \mathcal{V}), \forall \mathcal{V} \in \mathbf{X}$, and $\mathbf{g} = G(\lambda, \mathcal{U})$ if and only if (10) holds. This argument shows that $\mathcal{U}$ is the solution of (7). Conversely, if $\mathcal{U}$ solves (7), let $\mathbf{g}$ be defined by (10). Then $B(\mathcal{U}, \mathcal{V}) = (\mathbf{g}, \mathcal{V}), \forall \mathcal{V} \in \mathbf{X}$. Thus (7) and (8) are equivalent.

Throughout this paper, we will use the following assumption.

(A) There is λ such that $\{(\lambda, \mathcal{U}(\lambda)) \, ; \, \lambda \in \Lambda\}$ is a branch of nonsingular solutions of the equation (8).

Under this assumption (A), the optimality system (6) have a unique solution $\mathcal{U} \in \mathbf{X}$ for a given $\mathbf{u}_d$ (see [4]).

Then the Newton's algorithm in a weak form for (6) reads as (see [4]): find $\mathcal{U}_{n+1} := (\mathbf{u}_{n+1}, p_{n+1}, \mathbf{v}_{n+1}, q_{n+1}) \in \mathbf{X}$ such that, for all $\tilde{\mathbf{u}} \in \mathbf{V}$, $\tilde{p} \in \mathcal{Q}$, $\tilde{\mathbf{v}} \in \mathbf{V}$ and $\tilde{q} \in \mathcal{Q}$,

(11)
$$
\begin{cases}
a(\mathbf{u}_{n+1}, \tilde{\mathbf{u}}) + \lambda c(\mathbf{u}_n, \mathbf{u}_{n+1}, \tilde{\mathbf{u}}) + \lambda c(\mathbf{u}_{n+1}, \mathbf{u}_n, \tilde{\mathbf{u}}) \\
\quad + b(\tilde{\mathbf{u}}, p_{n+1}) - \lambda \left(\dfrac{\mathbf{v}_{n+1}}{\beta}, \tilde{\mathbf{u}} \right) = \lambda c(\mathbf{u}_n, \mathbf{u}_n, \tilde{\mathbf{u}}), \\
b(\mathbf{u}_{n+1}, \tilde{p}) = 0, \\
a(\mathbf{v}_{n+1}, \tilde{\mathbf{v}}) + \lambda c(\tilde{\mathbf{v}}, \mathbf{u}_n, \mathbf{v}_{n+1}) + \lambda c(\tilde{\mathbf{v}}, \mathbf{u}_{n+1}, \mathbf{v}_n) + \lambda c(\mathbf{u}_n, \tilde{\mathbf{v}}, \mathbf{v}_{n+1}) \\
\quad + \lambda c(\mathbf{u}_{n+1}, \tilde{\mathbf{v}}, \mathbf{v}_n) + b(\tilde{\mathbf{v}}, q_{n+1}) + \lambda(\mathbf{u}_{n+1} - \mathbf{u}_d, \tilde{\mathbf{v}}) \\
\quad = \lambda c(\tilde{\mathbf{v}}, \mathbf{u}_n, \mathbf{v}_n) + \lambda c(\mathbf{u}_n, \tilde{\mathbf{v}}, \mathbf{v}_n), \\
b(\mathbf{v}_{n+1}, \tilde{q}) = 0.
\end{cases}
$$

Let $\mathcal{U}$ be the unique solution to (6) and let an ϵ-neighborhood of $\mathcal{U}$ be

(12) $S(\mathcal{U}; \epsilon) = \{ \mathcal{V} \in \mathbf{X} \, : \, \|\mathbf{u} - \tilde{\mathbf{u}}\|_1 + \|p - \tilde{p}\| + \|\mathbf{v} - \tilde{\mathbf{v}}\|_1 + \|q - \tilde{q}\| \leq \epsilon \}.$

The following convergence result for the sequence $\mathcal{U}_n$ by Newton's algorithm can be found in, for example, [4].

Theorem 3.1. *Assume that the first Fréchet derivative $D_\mathcal{U} F(\lambda, \mathcal{V})$ of $F(\lambda, \mathcal{V})$ is Lipschitz continuous with respect to $\mathcal{V}$ in the ball $S(\mathcal{U}, \epsilon)$. Then there exists ϵ' with $0 < \epsilon' \leq \epsilon$ such that for each initial guess $\mathcal{U}^0$ in $S(\mathcal{U}, \epsilon')$ in Newton's algorithm (11) determines a unique sequence $\{\mathcal{U}_{n+1}\} \subset S(\mathcal{U}; \epsilon')$ that converges to the unique solution $\mathcal{U}$ of (6). Furthermore the convergence is quadratic*

$$(\|\mathbf{u}_{n+1} - \mathbf{u}\|_1 + \|p_{n+1} - p\| + \|\mathbf{v}_{n+1} - \mathbf{v}\|_1 + \|q_{n+1} - q\|)$$
$$\leq C(\|\mathbf{u}_n - \mathbf{u}\|_1 + \|p_n - p\| + \|\mathbf{v}_n - \mathbf{v}\|_1 + \|q_n - q\|)^2.$$

We will choose the initial guess as the weak solution of (13) to rewrite an iterative scheme (14) and show that (14) is in fact Newton's algorithm.

First, solve for all $\tilde{\mathbf{u}} \in \mathbf{V}$, $\tilde{p} \in \mathcal{Q}$, $\tilde{\mathbf{v}} \in \mathbf{V}$ and $\tilde{q} \in \mathcal{Q}$,

(13)
$$
\begin{cases}
a(\mathbf{u}^{(0)}, \tilde{\mathbf{u}}) + b(\tilde{\mathbf{u}}, p^{(0)}) - \lambda(\dfrac{\mathbf{v}^{(0)}}{\beta}, \tilde{\mathbf{u}}) &= 0, \\
b(\mathbf{u}^{(0)}, \tilde{p}) &= 0, \\
a(\mathbf{v}^{(0)}, \tilde{\mathbf{v}}) + b(\tilde{\mathbf{v}}, q^{(0)}) + \lambda(\mathbf{u}^{(0)} - \mathbf{u}_d, \tilde{\mathbf{v}}) &= 0, \\
b(\mathbf{v}^{(0)}, \tilde{q}) &= 0.
\end{cases}
$$

With the solution $\mathcal{U}^{(0)} = (\mathbf{u}^{(0)}, p^{(0)}, \mathbf{v}^{(0)}, q^{(0)})$ of the optimality system of the optimal control problem for the Stokes equation (13) as the initial guess, proceed with the following iterative scheme: find $\mathcal{U}^{(j)} = (\mathbf{u}^{(j)}, p^{(j)}, \mathbf{v}^{(j)}, q^{(j)}) \in$

X where $j = 1, 2, \ldots$ such that for all $\tilde{\mathbf{u}} \in \mathbf{V}$, $\tilde{p} \in \mathcal{Q}$, $\tilde{\mathbf{v}} \in \mathbf{V}$ and $\tilde{q} \in \mathcal{Q}$,

(14)
$$
\begin{cases}
a(\mathbf{u}^{(j)}, \tilde{\mathbf{u}}) + \lambda c(\mathbf{w}^{(j-1)}, \mathbf{u}^{(j)}, \tilde{\mathbf{u}}) + \lambda c(\mathbf{u}^{(j)}, \mathbf{w}^{(j-1)}, \tilde{\mathbf{u}}) \\
\quad + b(\tilde{\mathbf{u}}, p^{(j)}) - \lambda(\dfrac{\mathbf{v}^{(j)}}{\beta}, \tilde{\mathbf{u}}) = -\lambda c(\mathbf{u}^{(j-1)}, \mathbf{u}^{(j-1)}, \tilde{\mathbf{u}}), \\
b(\mathbf{u}^{(j)}, \tilde{p}) = 0, \\
a(\mathbf{v}^{(j)}, \tilde{\mathbf{v}}) + \lambda c(\tilde{\mathbf{v}}, \mathbf{w}^{(j-1)}, \mathbf{v}^{(j)}) + \lambda c(\tilde{\mathbf{v}}, \mathbf{u}^{(j)}, \mathbf{z}^{(j-1)}) + \lambda c(\mathbf{w}^{(j-1)}, \tilde{\mathbf{v}}, \mathbf{v}^{(j)}) \\
\quad + \lambda c(\mathbf{u}^{(j)}, \tilde{\mathbf{v}}, \mathbf{z}^{(j-1)}) + b(\tilde{\mathbf{v}}, q^{(j)}) + \lambda(\mathbf{u}^{(j)} - \mathbf{u}_d, \tilde{\mathbf{v}}) \\
\quad = -\lambda c(\tilde{\mathbf{v}}, \mathbf{u}^{(j-1)}, \mathbf{v}^{(j-1)}) - \lambda c(\mathbf{u}^{(j-1)}, \tilde{\mathbf{v}}, \mathbf{v}^{(j-1)}), \\
b(\mathbf{v}^{(j)}, \tilde{q}) = 0,
\end{cases}
$$

where, with a nonnegative integer k,

$$
\mathbf{w}^{(k)} = \sum_{i=0}^{k} \mathbf{u}^{(i)} \quad \text{and} \quad \mathbf{z}^{(k)} = \sum_{i=0}^{k} \mathbf{v}^{(i)}.
$$

Now let

$$
\mathcal{U}_n = (\mathbf{u}_n, p_n, \mathbf{v}_n, q_n) := \mathcal{U}^{(0)} + \mathcal{U}^{(1)} + \cdots + \mathcal{U}^{(n)},
$$

where

$$
\mathbf{u}_n = \sum_{j=0}^{n} \mathbf{u}^{(j)}, \; p_n = \sum_{j=0}^{n} p^{(j)}, \; \mathbf{v}_n = \sum_{j=0}^{n} \mathbf{v}^{(j)} \quad \text{and} \quad q_n = \sum_{j=0}^{n} q^{(j)},
$$

which will be shown as the solution to (6) by Newton's method in the rest section.

Theorem 3.2. *The solution $\mathcal{U}_n$ by the iterative scheme* (14) *is the solution generated by Newton's algorithm* (11).

Proof. First summing (14) from $j = 1$ to $n+1$, and then adding (13), we have

$$
a(\mathbf{u}_{n+1}, \tilde{\mathbf{u}}) + \lambda \sum_{j=1}^{n+1} c(\mathbf{u}_{j-1}, \mathbf{u}^{(j)}, \tilde{\mathbf{u}}) + \lambda \sum_{j=1}^{n+1} c(\mathbf{u}^{(j)}, \mathbf{u}_{j-1}, \tilde{\mathbf{u}}) +
$$

$$
b(\tilde{\mathbf{u}}, p_{n+1}) - \lambda(\frac{\mathbf{v}_{n+1}}{\beta}, \tilde{\mathbf{u}}) = -\lambda \sum_{j=1}^{n+1} c(\mathbf{u}^{(j-1)}, \mathbf{u}^{(j-1)}, \tilde{\mathbf{u}}),
$$

$$
b(\mathbf{u}_{n+1}, \tilde{p}) = 0,
$$

$$
a(\mathbf{v}_{n+1}, \tilde{\mathbf{v}}) + \lambda \sum_{j=1}^{n+1} c(\tilde{\mathbf{v}}, \mathbf{u}_{j-1}, \mathbf{v}^{(j)}) + \lambda \sum_{j=1}^{n+1} c(\tilde{\mathbf{v}}, \mathbf{u}^{(j)}, \mathbf{v}_{j-1})
$$

$$
+ \lambda \sum_{j=1}^{n+1} c(\mathbf{u}_{j-1}, \tilde{\mathbf{v}}, \mathbf{v}^{(j)}) + \lambda \sum_{j=1}^{n+1} c(\mathbf{u}^{(j)}, \tilde{\mathbf{v}}, \mathbf{v}_{j-1}) + b(\tilde{\mathbf{v}}, q_{n+1}) + \lambda(\mathbf{u}_{n+1} - \mathbf{u}_d, \tilde{\mathbf{v}})
$$

$$= -\lambda \sum_{j=1}^{n+1} c(\tilde{\mathbf{v}}, \mathbf{u}^{(j-1)}, \mathbf{v}^{(j-1)}) - \lambda \sum_{j=1}^{n+1} c(\mathbf{u}^{(j-1)}, \tilde{\mathbf{v}}, \mathbf{v}^{(j-1)}),$$

$$b(\mathbf{v}_{n+1}, \tilde{q}) = 0,$$

which can be written as the Newton's iteration form

$$a(\mathbf{u}_{n+1}, \tilde{\mathbf{u}}) + \lambda c(\mathbf{u}_n, \mathbf{u}_{n+1}, \tilde{\mathbf{u}}) + \lambda c(\mathbf{u}_{n+1}, \mathbf{u}_n, \tilde{\mathbf{u}}) + b(\tilde{\mathbf{u}}, p_{n+1}) - \lambda(\frac{\mathbf{v}_{n+1}}{\beta}, \tilde{\mathbf{u}})$$

$$= \lambda c(\mathbf{u}_n, \mathbf{u}_n, \tilde{\mathbf{u}}) + \lambda(\epsilon_1(n), \tilde{\mathbf{u}}),$$

$$b(\mathbf{u}_{n+1}, \tilde{p}) = 0,$$

$$a(\mathbf{v}_{n+1}, \tilde{\mathbf{v}}) + \lambda c(\tilde{\mathbf{v}}, \mathbf{u}_n, \mathbf{v}_{n+1}) + \lambda c(\tilde{\mathbf{v}}, \mathbf{u}_{n+1}, \mathbf{v}_n) + \lambda c(\mathbf{u}_n, \tilde{\mathbf{v}}, \mathbf{v}_{n+1})$$

$$+ \lambda c(\mathbf{u}_{n+1}, \tilde{\mathbf{v}}, \mathbf{v}_n) + b(\tilde{\mathbf{v}}, q_{n+1}) + \lambda(\mathbf{u}_{n+1} - \mathbf{u}_d, \tilde{\mathbf{v}})$$

$$= \lambda c(\tilde{\mathbf{v}}, \mathbf{u}_n, \mathbf{v}_n) + \lambda c(\mathbf{u}_n, \tilde{\mathbf{v}}, \mathbf{v}_n) + \lambda(\epsilon_2(n), \tilde{\mathbf{v}}),$$

$$b(\mathbf{v}_{n+1}, \tilde{q}) = 0,$$

where

$$(\epsilon_1(n), \tilde{\mathbf{u}}) = c(\mathbf{u}_n, \mathbf{u}_{n+1}, \tilde{\mathbf{u}}) + c(\mathbf{u}_{n+1}, \mathbf{u}_n, \tilde{\mathbf{u}}) - \sum_{j=1}^{n+1} c(\mathbf{u}_{j-1}, \mathbf{u}^{(j)}, \tilde{\mathbf{u}})$$

$$- \sum_{j=1}^{n+1} c(\mathbf{u}^{(j)}, \mathbf{u}_{j-1}, \tilde{\mathbf{u}}) - \sum_{j=1}^{n+1} c(\mathbf{u}^{(j-1)}, \mathbf{u}^{(j-1)}, \tilde{\mathbf{u}}) - c(\mathbf{u}_n, \mathbf{u}_n, \tilde{\mathbf{u}}),$$

and

$$(\epsilon_2(n), \tilde{\mathbf{v}}) = c(\tilde{\mathbf{v}}, \mathbf{u}_n, \mathbf{v}_{n+1}) + c(\tilde{\mathbf{v}}, \mathbf{u}_{n+1}, \mathbf{v}_n) + c(\mathbf{u}_n, \tilde{\mathbf{v}}, \mathbf{v}_{n+1})$$

$$+ c(\mathbf{u}_{n+1}, \tilde{\mathbf{v}}, \mathbf{v}_n) - \sum_{j=1}^{n+1} c(\tilde{\mathbf{v}}, \mathbf{u}_{j-1}, \mathbf{v}^{(j)}) - \sum_{j=1}^{n+1} c(\tilde{\mathbf{v}}, \mathbf{u}^{(j)}, \mathbf{v}_{j-1})$$

$$- \sum_{j=1}^{n+1} c(\mathbf{u}_{j-1}, \tilde{\mathbf{v}}, \mathbf{v}^{(j)}) - \sum_{j=1}^{n+1} c(\mathbf{u}^{(j)}, \tilde{\mathbf{v}}, \mathbf{v}_{j-1}) - \sum_{j=1}^{n+1} c(\tilde{\mathbf{v}}, \mathbf{u}^{(j-1)}, \mathbf{v}^{(j-1)})$$

$$- \sum_{j=1}^{n+1} c(\mathbf{u}^{(j-1)}, \tilde{\mathbf{v}}, \mathbf{v}^{(j-1)}) - c(\tilde{\mathbf{v}}, \mathbf{u}_n, \mathbf{v}_n) - c(\mathbf{u}_n, \tilde{\mathbf{v}}, \mathbf{v}_n).$$

Hence, it is enough to show that $(\epsilon_i(n), \tilde{\mathbf{v}}) = 0$ for $i = 1, 2$, which can be shown by using the following identity.

$$\left(\sum_{i=0}^{n} a_i \right) \left(\sum_{i=0}^{n+1} b_i \right) = \sum_{i=1}^{n+1} \left(\sum_{j=0}^{i-1} a_j \right) b_i + \sum_{i=1}^{n+1} a_i \left(\sum_{j=0}^{i-1} b_j \right) + \sum_{i=1}^{n+1} a_{i-1} b_{i-1}.$$

This completes the assertion of theorem.   □

4. A convergence analysis

Let $\mathcal{T}_h$ be a family of triangulations of Ω by a standard finite-element subdivisions of Ω into quasi-uniform triangles with $h = \max\{\text{diam}(K) : K \in \mathcal{T}_h\}$. Assume that the finite-element subspaces $\mathbf{V}_h \subset \mathbf{V}$ and $\mathcal{Q}_h \subset \mathcal{Q}$ satisfy the following approximation properties:

$$\inf_{\mathbf{w}_h \in \mathbf{V}_h} \|\mathbf{w} - \mathbf{w}_h\|_1 \le Ch^k\|\mathbf{w}\|_{k+1}, \quad \forall \mathbf{w} \in \mathbf{V} \cap H^{k+1}(\Omega)^2, \, 1 \le k \le l,$$

$$\inf_{r_h \in \mathcal{Q}_h} \|r - r_h\|_0 \le Ch^m\|r\|_m, \quad \forall r \in \mathcal{Q} \cap H^m(\Omega), \, 1 \le m \le l,$$

where $\mathbf{w}$ is $\mathbf{u}$ or $\mathbf{v}$ and r is p or q. Further assume that the following uniform inf-sup condition holds on $\mathbf{V}_h \times \mathcal{Q}_h$: for each $r_h \in \mathcal{Q}_h$, there exists a $\mathbf{w}_h \in \mathbf{V}_h$ such that

$$(r_h, \nabla \cdot \mathbf{w}_h) = \|r_h\|^2, \qquad |\mathbf{w}_h| \le C\|r_h\|,$$

with a positive constant C independent of h.

The Newton's algorithm in the finite-element approach (11) can be stated as: find $(\mathbf{u}_{n+1,h}, p_{n+1,h}, \mathbf{v}_{n+1,h}, q_{n+1,h}) \in \mathbf{V}_h \times \mathcal{Q}_h \times \mathbf{V}_h \times \mathcal{Q}_h$ such that, for all $\tilde{\mathbf{u}} \in \mathbf{V}_h$, $\tilde{p} \in \mathcal{Q}_h$, $\tilde{\mathbf{v}} \in \mathbf{V}_h$ and $\tilde{q} \in \mathcal{Q}_h$,

(15)
$$\begin{cases} a(\mathbf{u}_{n+1,h}, \tilde{\mathbf{u}}) + \lambda c(\mathbf{u}_{n,h}, \mathbf{u}_{n+1,h}, \tilde{\mathbf{u}}) + \lambda c(\mathbf{u}_{n+1,h}, \mathbf{u}_{n,h}, \tilde{\mathbf{u}}) + b(\tilde{\mathbf{u}}, p_{n+1,h}) \\ \quad - \lambda \left(\dfrac{\mathbf{v}_{n+1,h}}{\beta}, \tilde{\mathbf{u}} \right) = \lambda c(\mathbf{u}_{n,h}, \mathbf{u}_{n,h}, \tilde{\mathbf{u}}), \\ b(\mathbf{u}_{n+1,h}, \tilde{p}) = 0, \\ a(\mathbf{v}_{n+1,h}, \tilde{\mathbf{v}}) + \lambda c(\tilde{\mathbf{v}}, \mathbf{u}_{n,h}, \mathbf{v}_{n+1,h}) + \lambda c(\tilde{\mathbf{v}}, \mathbf{u}_{n+1,h}, \mathbf{v}_{n,h}) + \lambda c(\mathbf{u}_{n,h}, \tilde{\mathbf{v}}, \mathbf{v}_{n+1,h}) \\ \quad + \lambda c(\mathbf{u}_{n+1,h}, \tilde{\mathbf{v}}, \mathbf{v}_{n,h}) + b(\tilde{\mathbf{v}}, q_{n+1,h}) + \lambda(\mathbf{u}_{n+1,h} - \mathbf{u}_d, \tilde{\mathbf{v}}) \\ \quad = \lambda c(\tilde{\mathbf{v}}, \mathbf{u}_{n,h}, \mathbf{v}_{n,h}) + \lambda c(\mathbf{u}_{n,h}, \tilde{\mathbf{v}}, \mathbf{v}_{n,h}), \\ b(\mathbf{v}_{n+1,h}, \tilde{q}) = 0, \end{cases}$$

where $n = 0, 1, 2, \ldots$, with the initial guess $\mathcal{U}_h^{(0)} := (\mathbf{u}_{0,h}, p_{0,h}, \mathbf{v}_{0,h}, q_{0,h})$ which is the finite-element approximation solution of Stokes problem

(16)
$$\begin{aligned} a(\mathbf{u}_{0,h}, \tilde{\mathbf{u}}) + b(\tilde{\mathbf{u}}, p_{0,h}) - \lambda(\frac{\mathbf{v}_{0,h}}{\beta}, \tilde{\mathbf{u}}) &= 0, \\ b(\mathbf{u}_{0,h}, \tilde{p}) &= 0, \\ a(\mathbf{v}_{0,h}, \tilde{\mathbf{v}}) + b(\tilde{\mathbf{v}}, q_{0,h}) + \lambda(\mathbf{u}_{0,h} - \mathbf{u}_d, \tilde{\mathbf{v}}) &= 0, \\ b(\mathbf{v}_{0,h}, \tilde{q}) &= 0. \end{aligned}$$

Equivalently, the discrete Newton's algorithm can be written as: find $\mathcal{U}_h^{(j)} = (\mathbf{u}_h^{(j)}, p_h^{(j)}, \mathbf{v}_h^{(j)}, q_h^{(j)}) \in \mathbf{X}_h$ such that

$$
\begin{cases}
a(\mathbf{u}_h^{(j)}, \tilde{\mathbf{u}}) + \lambda c(\mathbf{w}_h^{(j-1)}, \mathbf{u}_h^{(j)}, \tilde{\mathbf{u}}) + \lambda c(\mathbf{u}_h^{(j)}, \mathbf{w}_h^{(j-1)}, \tilde{\mathbf{u}}) + b(\tilde{\mathbf{u}}, p_h^{(j)}) - \lambda(\dfrac{\mathbf{v}_h^{(j)}}{\beta}, \tilde{\mathbf{u}}) \\[2mm]
\quad = -\lambda c(\mathbf{u}_h^{(j-1)}, \mathbf{u}_h^{(j-1)}, \tilde{\mathbf{u}}), \\[2mm]
b(\mathbf{u}_h^{(j)}, \tilde{p}) = 0, \\[2mm]
a(\mathbf{v}_h^{(j)}, \tilde{\mathbf{v}}) + \lambda c(\tilde{\mathbf{v}}, \mathbf{w}_h^{(j-1)}, \mathbf{v}_h^{(j)}) + \lambda c(\tilde{\mathbf{v}}, \mathbf{u}_h^{(j)}, \mathbf{z}_h^{(j-1)}) + \lambda c(\mathbf{w}_h^{(j-1)}, \tilde{\mathbf{v}}, \mathbf{v}_h^{(j)}) \\[2mm]
\quad + \lambda c(\mathbf{u}_h^{(j)}, \tilde{\mathbf{v}}, \mathbf{z}_h^{(j-1)}) + b(\tilde{\mathbf{v}}, q_h^{(j)}) + \lambda(\mathbf{u}_h^{(j)} - \mathbf{u}_d, \tilde{\mathbf{v}}) \\[2mm]
\quad = -\lambda c(\tilde{\mathbf{v}}, \mathbf{u}_h^{(j-1)}, \mathbf{v}_h^{(j-1)}) - \lambda c(\mathbf{u}_h^{(j-1)}, \tilde{\mathbf{v}}, \mathbf{v}_h^{(j-1)}), \\[2mm]
b(\mathbf{v}_h^{(j)}, \tilde{q}) = 0,
\end{cases}
$$

where, with a nonnegative integer j,

$$
\mathbf{w}_h^{(j-1)} = \sum_{i=0}^{j-1} \mathbf{u}_h^{(i)} \quad \text{and} \quad \mathbf{z}_h^{(j-1)} = \sum_{i=0}^{j-1} \mathbf{v}_h^{(i)}.
$$

Then, according to Theorem 3.2, the discrete solution to (15) can be written as

$$
\mathbf{u}_{n,h} = \sum_{j=0}^{n} \mathbf{u}_h^{(j)}, \quad p_{n,h} = \sum_{j=0}^{n} p_h^{(j)}, \quad \mathbf{v}_{n,h} = \sum_{j=0}^{n} \mathbf{v}_h^{(j)} \quad \text{and} \quad q_{n,h} = \sum_{j=0}^{n} q_h^{(i)}.
$$

The following technical known results are necessary for a future use.

Lemma 4.1. (i) *For every $p \in \mathcal{Q}$, there is a $\mathbf{u} \in \mathbf{V}$ such that*

$$
\nabla \cdot \mathbf{u} = p \quad \text{and} \quad \|\mathbf{u}\|_1 \leq C_1 \|p\|,
$$

where C_1 is a positive constant.

(ii) *The following Poincaŕe inequality holds;*

$$
C_p \|\mathbf{u}\|_1^2 \leq \|\nabla \mathbf{u}\|^2, \quad \forall \mathbf{u} \in \mathbf{V},
$$

where C_p is a positive constant depending only on Ω.

(iii) *For all $\mathbf{u}$, $\mathbf{v}$ and $\mathbf{w} \in H^1(\Omega)$*

$$
(17) \qquad \left| \int_\Omega D_i \mathbf{u} \mathbf{v} \mathbf{w} \, d\Omega \right| \leq C_2 \|\mathbf{u}\|_1 \|\mathbf{v}\|_1 \|\mathbf{w}\|_1, \quad i = 1, 2.
$$

Proof. See [5]. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

Let us define the bilinear forms $a_1(\cdot, \mathbf{z}, \cdot : \cdot)$ and $a_2(\cdot, \mathbf{z}, \mathbf{w}, \cdot : \cdot)$ as follows:

$$
a_1(\mathbf{u}, \mathbf{z}, \mathbf{v} : \tilde{\mathbf{u}}) = (\nabla \mathbf{u}, \nabla \tilde{\mathbf{u}}) + \lambda((\mathbf{z} \cdot \nabla)\mathbf{u} + (\mathbf{u} \cdot \nabla)\mathbf{z}, \tilde{\mathbf{u}}) - \frac{\lambda}{\beta}(\mathbf{v}, \tilde{\mathbf{u}})
$$

and

$$
a_2(\mathbf{v}, \mathbf{z}, \mathbf{w}, \mathbf{u} : \tilde{\mathbf{v}}) = (\nabla \mathbf{v}, \nabla \tilde{\mathbf{v}}) + \lambda((\tilde{\mathbf{v}} \cdot \nabla)\mathbf{z} + (\mathbf{z} \cdot \nabla)\tilde{\mathbf{v}}, \mathbf{v})
$$

$$+ \lambda((\mathbf{u} \cdot \nabla)\tilde{\mathbf{v}} + (\tilde{\mathbf{v}} \cdot \nabla)\mathbf{u}, \mathbf{w}) + \lambda(\mathbf{u}, \tilde{\mathbf{v}}).$$

Consider the following problem: to find $(\mathbf{u}, p, \mathbf{v}, q) \in \mathbf{X}$ such that

(18)
$$\begin{cases} a_1(\mathbf{u}, \mathbf{z}, \mathbf{v} : \tilde{\mathbf{u}}) + (\nabla p, \tilde{\mathbf{u}}) = (\mathbf{h}, \tilde{\mathbf{u}}) \\ \qquad\qquad\qquad (\nabla \cdot \mathbf{u}, \tilde{p}) = 0 \\ a_2(\mathbf{v}, \mathbf{z}, \mathbf{w}, \mathbf{u} : \tilde{\mathbf{v}}) + (\nabla q, \tilde{\mathbf{v}}) = (\mathbf{g}, \tilde{\mathbf{v}}) \\ \qquad\qquad\qquad (\nabla \cdot \mathbf{v}, \tilde{q}) = 0 \end{cases}$$

for $\mathbf{z}, \mathbf{w} \in \mathbf{V}_\lambda$ where

$$\mathbf{V}_\lambda := \Big\{ \mathbf{z} \in \mathbf{V} | \; \|\mathbf{z}\|_1 \leq \frac{C_p}{kC_2} \frac{1}{\lambda}, C_p \text{ is the Poincar\'e constant,}$$

$$k > 0 \text{ is an arbitrary constant} \Big\}.$$

Lemma 4.2. *For any* $\mathbf{u}, \mathbf{v}, \tilde{\mathbf{u}}, \tilde{\mathbf{v}} \in \mathbf{V}$ *and* $\mathbf{z}, \mathbf{w} \in \mathbf{V}_\lambda$, *the following holds:*
$$|((\mathbf{z} \cdot \nabla)\mathbf{u}, \tilde{\mathbf{u}})| \leq \frac{C_p}{k\lambda}\|\mathbf{u}\|_1\|\tilde{\mathbf{u}}\|_1, \quad |((\mathbf{u} \cdot \nabla)\mathbf{z}, \tilde{\mathbf{u}})| \leq \frac{C_p}{k\lambda}\|\mathbf{u}\|_1\|\tilde{\mathbf{u}}\|_1,$$
$$|((\tilde{\mathbf{v}} \cdot \nabla)\mathbf{z}, \mathbf{v})| \leq \frac{C_p}{k\lambda}\|\mathbf{v}\|_1\|\tilde{\mathbf{v}}\|_1, \quad |((\mathbf{z} \cdot \nabla)\tilde{\mathbf{v}}, \mathbf{v})| \leq \frac{C_p}{k\lambda}\|\mathbf{v}\|_1\|\tilde{\mathbf{v}}\|_1,$$
$$|((\mathbf{u} \cdot \nabla)\tilde{\mathbf{v}}, \mathbf{w})| \leq \frac{C_p}{k\lambda}\|\mathbf{u}\|_1\|\tilde{\mathbf{v}}\|_1, \text{ and } |((\tilde{\mathbf{v}} \cdot \nabla)\mathbf{u}, \mathbf{w})| \leq \frac{C_p}{k\lambda}\|\mathbf{u}\|_1\|\tilde{\mathbf{v}}\|_1.$$

Proof. A simple application of (17) yields the conclusion. $\qquad\square$

Lemma 4.3. *Let* $\mathbf{z}, \mathbf{w} \in \mathbf{V}_\lambda$. *Then we have*
$$|a_1(\mathbf{u}, \mathbf{z}, \mathbf{v} : \tilde{\mathbf{u}})| \leq C_3(\|\mathbf{u}\|_1 + \|\mathbf{v}\|_1)\|\tilde{\mathbf{u}}\|_1,$$
and
$$|a_2(\mathbf{v}, \mathbf{z}, \mathbf{w}, \mathbf{u} : \tilde{\mathbf{v}})| \leq C_4(\|\mathbf{u}\|_1 + \|\mathbf{v}\|_1)\|\tilde{\mathbf{v}}\|_1,$$
$\forall \mathbf{u}, \mathbf{v}, \tilde{\mathbf{u}}, \tilde{\mathbf{v}} \in \mathbf{V}$.

Proof. The proof comes from Cauchy-Schwarz inequality and Lemma 4.2. That is,

$$|a_1(\mathbf{u}, \mathbf{z}, \mathbf{v} : \tilde{\mathbf{u}})| \leq |(\nabla \mathbf{u}, \nabla \tilde{\mathbf{u}})| + \lambda|((\mathbf{z} \cdot \nabla)\mathbf{u} + (\mathbf{u} \cdot \nabla)\mathbf{z}, \tilde{\mathbf{u}})| + \frac{\lambda}{\beta}|(\mathbf{v}, \tilde{\mathbf{u}})|$$

$$\leq \left(1 + \frac{2C_p}{k} + \frac{\lambda}{\beta}\right)(\|\mathbf{u}\|_1 + \|\mathbf{v}\|_1)\|\tilde{\mathbf{u}}\|_1,$$

$$|a_2(\mathbf{v}, \mathbf{z}, \mathbf{w}, \mathbf{u} : \tilde{\mathbf{v}})| \leq |(\nabla \mathbf{v}, \nabla \tilde{\mathbf{v}})| + \lambda|((\tilde{\mathbf{v}} \cdot \nabla)\mathbf{z} + (\mathbf{z} \cdot \nabla)\tilde{\mathbf{v}}, \mathbf{v})|$$
$$\qquad\qquad + \lambda|((\mathbf{u} \cdot \nabla)\tilde{\mathbf{v}} + (\tilde{\mathbf{v}} \cdot \nabla)\mathbf{u}, \mathbf{w})| + \lambda|(\mathbf{u}, \tilde{\mathbf{v}})|$$

$$\leq \left(1 + \frac{2C_p}{k} + \lambda\right)(\|\mathbf{u}\|_1 + \|\mathbf{v}\|_1)\|\tilde{\mathbf{v}}\|_1,$$

where C_3 and C_4 depend on C_p and λ. $\qquad\square$

Lemma 4.4. *Let* $\mathbf{z}, \mathbf{w} \in \mathbf{V}_\lambda$. *Then we have*

$$a_1(\mathbf{u}, \mathbf{z}, \mathbf{v} : \mathbf{u}) \geq \frac{k-1}{k} C_p \|\mathbf{u}\|_1^2 - \frac{\lambda}{\beta} \|\mathbf{u}\|_1 \|\mathbf{v}\|_1$$

and

$$a_2(\mathbf{v}, \mathbf{z}, \mathbf{w}, \mathbf{u} : \mathbf{v}) \geq \frac{k-1}{k} C_p \|\mathbf{v}\|_1^2 - \left(\frac{2C_p}{k} + \lambda\right) \|\mathbf{u}\|_1 \|\mathbf{v}\|_1,$$

$\forall \mathbf{u}, \mathbf{v}, \tilde{\mathbf{u}}, \tilde{\mathbf{v}} \in \mathbf{V}$.

Proof.

$$a_1(\mathbf{u}, \mathbf{z}, \mathbf{v} : \mathbf{u}) = \|\nabla \mathbf{u}\|^2 + \lambda \left(((\mathbf{z} \cdot \nabla)\mathbf{u}, \mathbf{u}) + ((\mathbf{u} \cdot \nabla)\mathbf{z}, \mathbf{u})\right) - \frac{\lambda}{\beta}(\mathbf{v}, \mathbf{u})$$

$$\geq \frac{k-1}{k} C_p \|\mathbf{u}\|_1^2 - \frac{\lambda}{\beta} \|\mathbf{u}\|_1 \|\mathbf{v}\|_1,$$

$$a_2(\mathbf{v}, \mathbf{z}, \mathbf{w}, \mathbf{u} : \mathbf{v}) = \|\nabla \mathbf{v}\|^2 + \lambda \left(((\mathbf{v} \cdot \nabla)\mathbf{z}, \mathbf{v}) + ((\mathbf{z} \cdot \nabla)\mathbf{v}, \mathbf{v})\right)$$

$$+ \lambda \left(((\mathbf{u} \cdot \nabla)\mathbf{v}, \mathbf{w}) + ((\mathbf{v} \cdot \nabla)\mathbf{u}, \mathbf{w})\right) + \lambda(\mathbf{u}, \mathbf{v})$$

$$\geq \frac{k-1}{k} C_p \|\mathbf{v}\|_1^2 - \left(\frac{2C_p}{k} + \lambda\right) \|\mathbf{u}\|_1 \|\mathbf{v}\|_1. \qquad \square$$

Lemma 4.5. *Let* $\mathbf{h}, \mathbf{g} \in \mathbf{V}^*$. *For* $\mathbf{z}, \mathbf{w} \in \mathbf{V}_\lambda$, *the solution* $(\mathbf{u}, p, \mathbf{v}, q) \in \mathbf{X}$ *of the weak form* (18) *satisfies a priori estimate*

$$\|\mathbf{u}\|_1 + \|\mathbf{v}\|_1 \leq C_5(\|\mathbf{h}\|_{-1} + \|\mathbf{g}\|_{-1}),$$

$$\|p\| \leq C_6(\|\mathbf{h}\|_{-1} + \|\mathbf{g}\|_{-1})$$

and

$$\|q\| \leq C_7(\|\mathbf{h}\|_{-1} + \|\mathbf{g}\|_{-1}).$$

Proof. Since $(\mathbf{u}, p, \mathbf{v}, q)$ is the weak solution of (18), we have

$$a_1(\mathbf{u}, \mathbf{z}, \mathbf{v} : \mathbf{u}) = (\mathbf{h}, \mathbf{u}), \ a_2(\mathbf{v}, \mathbf{z}, \mathbf{w}, \mathbf{u} : \mathbf{v}) = (\mathbf{g}, \mathbf{v}).$$

Then using Lemma 4.4, it follows that

$$\frac{k-1}{k} C_p \|\mathbf{u}\|_1^2 - \frac{\lambda}{\beta} \|\mathbf{u}\|_1 \|\mathbf{v}\|_1 \leq a_1(\mathbf{u}, \mathbf{z}, \mathbf{v} : \mathbf{u}) = (\mathbf{h}, \mathbf{u}) \leq \|\mathbf{h}\|_{-1} \|\mathbf{u}\|_1$$

which implies

(19) $$\frac{k-1}{k} C_p \|\mathbf{u}\|_1 - \frac{\lambda}{\beta} \|\mathbf{v}\|_1 \leq \|\mathbf{h}\|_{-1}.$$

Also,

$$\frac{k-1}{k} C_p \|\mathbf{v}\|_1^2 - \left(\frac{2C_p}{k} + \lambda\right) \|\mathbf{u}\|_1 \|\mathbf{v}\|_1 \leq a_2(\mathbf{v}, \mathbf{z}, \mathbf{w}, \mathbf{u} : \mathbf{v}) = (\mathbf{g}, \mathbf{v})$$

$$\leq \|\mathbf{g}\|_{-1} \|\mathbf{v}\|_1$$

implies

$$(20) \qquad \frac{k-1}{k}C_p\|\mathbf{v}\|_1 - \left(\frac{2C_p}{k} + \lambda\right)\|\mathbf{u}\|_1 \leq \|\mathbf{g}\|_{-1}.$$

Using (19) and (20),

$$\left(\frac{k-1}{k}C_p - \frac{2C_p}{k} - \lambda\right)\|\mathbf{u}\|_1 + \left(\frac{k-1}{k}C_p - \frac{\lambda}{\beta}\right)\|\mathbf{v}\|_1 \leq \|\mathbf{h}\|_{-1} + \|\mathbf{g}\|_{-1}.$$

Thus we have

$$(21) \qquad \|\mathbf{u}\|_1 + \|\mathbf{v}\|_1 \leq C_5(\|\mathbf{h}\|_{-1} + \|\mathbf{g}\|_{-1}),$$

where $C_5 = 1/\min\left\{\frac{k-3}{k}C_p - \lambda, \frac{k-1}{k}C_p - \frac{\lambda}{\beta}\right\}$, $\frac{k-3}{k}C_p - \lambda > 0$ and $\frac{k-1}{k}C_p - \frac{\lambda}{\beta} > 0$ for suitable k.

Since $p \in \mathcal{Q}$, there is a $\mathbf{u} \in \mathbf{V}$ such that

$$(22) \qquad \nabla \cdot \mathbf{u} = p, \qquad \|\mathbf{u}\|_1 \leq C\|p\|.$$

By taking such $\mathbf{u}$ into (18), Lemma 4.3 yields

$$\|p\|^2 = (p, \nabla \cdot \mathbf{u}) = a_1(\mathbf{u}, \mathbf{z}, \mathbf{v} : \mathbf{u}) - (\mathbf{h}, \mathbf{u}) \leq C_8(\|\mathbf{u}\|_1 + \|\mathbf{v}\|_1 + \|\mathbf{h}\|_{-1})\|\mathbf{u}\|_1,$$

where $C_8 = \max\{1, C_3\}$, which implies, using (21) and (22),

$$\|p\| \leq C_9(\|\mathbf{u}\|_1 + \|\mathbf{v}\|_1 + \|\mathbf{h}\|_{-1}) \leq C_6(\|\mathbf{h}\|_{-1} + \|\mathbf{g}\|_{-1}).$$

Similarly,

$$\|q\| \leq C_7(\|\mathbf{h}\|_{-1} + \|\mathbf{g}\|_{-1}). \qquad \qquad \square$$

Corollary 4.1. $\|\mathbf{u}\|_1 + \|\mathbf{v}\|_1 + \|p\| + \|q\| \leq C(\|\mathbf{h}\|_{-1} + \|\mathbf{g}\|_{-1}).$

By triangle inequality, it follows that

$$\|\mathbf{u} - \mathbf{u}_{0,h}\|_1 + \|p - p_{0,h}\| + \|\mathbf{v} - \mathbf{v}_{0,h}\|_1 + \|q - q_{0,h}\|$$
$$\leq \|\mathbf{u} - \mathbf{u}^{(0)}\|_1 + \|p - p^{(0)}\| + \|\mathbf{v} - \mathbf{v}^{(0)}\|_1 + \|q - q^{(0)}\|$$
$$+ \|\mathbf{u}^{(0)} - \mathbf{u}_{0,h}\|_1 + \|p^{(0)} - p_{0,h}\| + \|\mathbf{v}^{(0)} - \mathbf{v}_{0,h}\|_1 + \|q^{(0)} - q_{0,h}\|.$$

Now subtracting (13) from (7), we have, $\forall \tilde{\mathbf{u}} \in \mathbf{V}$, $\forall \tilde{p} \in \mathcal{Q}$, $\forall \tilde{\mathbf{v}} \in \mathbf{V}$,

$$a(\mathbf{u} - \mathbf{u}^{(0)}, \tilde{\mathbf{u}}) + b(\tilde{\mathbf{u}}, p - p^{(0)}) - \lambda(\frac{\mathbf{v} - \mathbf{v}^{(0)}}{\beta}, \tilde{\mathbf{u}}) = -\lambda c(\mathbf{u}, \mathbf{u}, \tilde{\mathbf{u}}),$$
$$b(\mathbf{u} - \mathbf{u}^{(0)}, \tilde{p}) = 0,$$
$$a(\mathbf{v} - \mathbf{v}^{(0)}, \tilde{\mathbf{v}}) + b(\tilde{\mathbf{v}}, q - q^{(0)}) + \lambda(\mathbf{u} - \mathbf{u}^{(0)}, \tilde{\mathbf{v}}) = -\lambda c(\tilde{\mathbf{v}}, \mathbf{u}, \mathbf{v}) - \lambda c(\mathbf{u}, \tilde{\mathbf{v}}, \mathbf{v})$$
$$b(\mathbf{v} - \mathbf{v}^{(0)}, \tilde{q}) = 0 \quad \forall \tilde{q} \in \mathcal{Q}.$$

By $(\mathbf{h}, \tilde{\mathbf{u}}) = -\lambda c(\mathbf{u}, \mathbf{u}, \tilde{\mathbf{u}})$ and $(\mathbf{g}, \tilde{\mathbf{v}}) = -\lambda c(\tilde{\mathbf{v}}, \mathbf{u}, \mathbf{v}) - \lambda c(\mathbf{u}, \tilde{\mathbf{v}}, \mathbf{v})$ in (18), it can be shown from Lemma 4.5 that

(23)
$$\|\mathbf{u} - \mathbf{u}^{(0)}\|_1 + \|p - p^{(0)}\| + \|\mathbf{v} - \mathbf{v}^{(0)}\|_1 + \|q - q^{(0)}\| \leq C\lambda(\|\mathbf{u}\|_1^2 + \|\mathbf{u}\|_1\|\mathbf{v}\|_1)$$

and it can be shown that the finite-error estimate for optimal control problem for Stokes equations

$$
(24) \quad \begin{aligned}
&\|\mathbf{u}^{(0)} - \mathbf{u}_{0,h}\|_1 + \|p^{(0)} - p_{0,h}\| + \|\mathbf{v}^{(0)} - \mathbf{v}_{0,h}\|_1 + \|q^{(0)} - q_{0,h}\| \\
&\leq Ch^m(\|\mathbf{u}\|_{m+1} + \|p\|_m + \|\mathbf{v}\|_{m+1} + \|q\|_m).
\end{aligned}
$$

Combining (23) and (24), we have the following.

Proposition 4.1. *Let* $(\mathbf{u}, p, \mathbf{v}, q) \in (\mathbf{V} \times \mathcal{Q} \times \mathbf{V} \times \mathcal{Q}) \cap (H^{k+1}(\Omega) \times H^m(\Omega) \times H^{k+1}(\Omega) \times H^m(\Omega))$ *be the solution of* (7) *and* $(\mathbf{u}_{0,h}, p_{0,h}, \mathbf{v}_{0,h}, q_{0,h})$ *be the solution of* (16). *Then*

$$
(25) \quad \|\mathbf{u} - \mathbf{u}_{0,h}\|_1 + \|p - p_{0,h}\| + \|\mathbf{v} - \mathbf{v}_{0,h}\|_1 + \|q - q_{0,h}\|
$$
$$
\leq C(h^m\|p\|_m + h^k\|\mathbf{u}\|_{k+1} + \lambda\|\mathbf{u}\|_1^2 + h^m\|q\|_m + h^k\|\mathbf{v}\|_{k+1} + \lambda\|\mathbf{u}\|_1\|\mathbf{v}\|_1).
$$

The above Proposition 4.1 tells us that the finite element approximate solutions $(\mathbf{u}_{0,h}, p_{0,h}, \mathbf{v}_{0,h}, q_{0,h})$ can be chosen as an initial guess for the Newton's method (15) and (16) to solve the optimality system for the Navier-Stokes equations if for an $\epsilon' < 1$, there are a mesh size h and λ such that

$$
(26) \quad C(h^m\|p\|_m + h^k\|\mathbf{u}\|_{k+1} + \lambda\|\mathbf{u}\|_1^2 + h^m\|q\|_m + h^k\|\mathbf{v}\|_{k+1} + \lambda\|\mathbf{u}\|_1\|\mathbf{v}\|_1) \leq \epsilon'.
$$

In order to show the quadratic convergence of the sequence $(\mathbf{u}_{n,h}, p_{n,h}, \mathbf{v}_{n,h}, q_{n,h})$ generated by Newton's method for the chosen initial guess, we will prove the validity of the assumptions of Theorem 2.1 found in [4].

Theorem 4.1. *Let* ϵ *be in* (12). *Assume that* (26) *holds for* $\epsilon' \leq \epsilon$, *i.e., the solution* $(\mathbf{u}_{0,h}, p_{0,h}, \mathbf{v}_{0,h}, q_{0,h})$ *of the optimal control problem for the Stokes equations* (16) *belongs to* $S(\mathcal{U}; \epsilon')$ *where* $\mathcal{U} = (\mathbf{u}, p, \mathbf{v}, q)$ *is the solution of* (7). *Then Newton's algorithm* (15) *determines a unique sequence* $\{(\mathbf{u}_{n+1,h}, p_{n+1,h}, \mathbf{v}_{n+1,h}, q_{n+1,h})\} \subset S(\mathcal{U}; \epsilon') \cap (\mathbf{V}_h \times \mathcal{Q}_h \times \mathbf{V}_h \times \mathcal{Q}_h)$ *that converges to the unique solution* $\mathcal{U} = (\mathbf{u}, p, \mathbf{v}, q) \in H^{k+1}(\Omega) \times H^m(\Omega) \times H^{k+1}(\Omega) \times H^m(\Omega)$ *of* (7). *Furthermore the convergence is quadratic with an initial guess* $(\mathbf{u}_{0,h}, p_{0,h}, \mathbf{v}_{0,h}, q_{0,h})$

$$
(27) \quad \begin{aligned}
&(\|\mathbf{u}_{n+1,h} - \mathbf{u}\|_1 + \|p_{n+1,h} - p\| + \|\mathbf{v}_{n+1,h} - \mathbf{v}\|_1 + \|q_{n+1,h} - q\|) \\
&\leq C(\|\mathbf{u}_{n,h} - \mathbf{u}\|_1 + \|p_{n,h} - p\| + \|\mathbf{v}_{n,h} - \mathbf{v}\|_1 + \|q_{n,h} - q\|)^2 \\
&\leq C_n(h^m\|p\|_m + h^k\|\mathbf{u}\|_{k+1} + \lambda\|\mathbf{u}\|_1^2 \\
&\quad + h^m\|q\|_m + h^k\|\mathbf{v}\|_{k+1} + \lambda\|\mathbf{u}\|_1\|\mathbf{v}\|_1)^{2^{n+1}},
\end{aligned}
$$

where C_n *is a positive constant dependent on the iteration number* n.

Proof. Since the hypothesis holds the assumptions of Theorem 2.1, the Newton's sequence converges to the unique solution. Furthermore, the last inequality in (27) comes from (25). These arguments may complete the proof. □

References

[1] P. B. Bochev, Z. Cai, T. A. Manteuffel, and S. F. McCormick, *Analysis of velocity-flux first-order system least-squares principles for the Navier-Stokes equations. I*, SIAM J. Numer. Anal. **35** (1998), no. 3, 990–1009.

[2] F. Brezzi, J. Rappaz, and P. A. Raviart, *Finite-dimensional approximation of nonlinear problems. I. Branches of nonsingular solutions*, Numer. Math. **36** (1980/81), no. 1, 1–25.

[3] P. Ciarlet, *Finite Element Method for Elliptic Problems*, North-Holland, Amsterdam, 1978.

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

[5] S. D. Kim, Y. H. Lee, and B. C. Shin, *Newton's method for the Navier-Stokes equations with finite-element initial guess of Stokes equations*, Comput. Math. Appl. **51** (2006), no. 5, 805–816.

[6] D. A. Knoll and V. A. Mousseau, *On Newton-Krylov multigrid methods for the incompressible Navier-Stokes equations*, J. Comput. Phys. **163** (2000), 262–267.

[7] D. A. Knoll and W. Rider, *A multigrid preconditioned Newton-Krylov method*, SIAM J. Sci. Comput. **21** (1999), no. 2, 691–710.

[8] M. Pernice and M. D. Tocci, *A multigrid-preconditioned Newton-Krylov Method for the incompressible Navier-Stokes equations*, SIAM J. Sci. Comput. **23** (2001), no. 2, 398–418.

Youngmi Choi
Department of Mathematics
Ajou University
Suwon 443-749, Korea
E-mail address: ymchoi93@hanmail.net

Sangdong Kim
Department of Mathematics
Kyungpook National University
Daegu 702-701, Korea
E-mail address: skim@knu.ac.kr

Hyung-Chun Lee
Department of Mathematics
Ajou University
Suwon 443-749, Korea
E-mail address: hclee@ajou.ac.kr