ELSEVIER Journal of Computational and Applied Mathematics 61 (1995) 323–343

JOURNAL OF COMPUTATIONAL AND APPLIED MATHEMATICS

Analysis and finite element approximation of optimal control problems for a Ladyzhenskaya model for stationary, incompressible, viscous flows[1]

Qiang Du[a], Max D. Gunzburger[b,*], L. Steven Hou[c]

[a] Department of Mathematics, Michigan State University, East Lansing, MI 48824, USA
[b] Interdisciplinary Center for Applied Mathematics, Department of Mathematics, Virginia Tech., Blacksburg, VA 24061, USA
[c] Department of Mathematics and Statistics, York University, North York, Ont., M3J 1P3, Canada

Received 6 December 1993; revised 16 May 1994

Abstract

We examine certain analytic and numerical aspects of optimal control problems for a Ladyzhenskaya model for stationary, incompressible, viscous flows. The control considered is of the distributed type; the functionals minimized are the L^2-distance of candidate flow to some desired flow and the viscous drag on bounding surfaces. We show the existence of optimal solutions and justify the use of Lagrange multiplier techniques to derive a system of partial differential equations from which optimal solutions may be deduced. We study the regularity of solutions of this system. Then, we consider approximations, by finite element methods, of solutions of the optimality system and examine their convergence properties.

Keywords: Optimal control; Ladyzhenskaya equations; Finite element methods

1. Introduction

The optimization problem we study is to seek a state pair $(\boldsymbol{u}, p)$ and a control $\boldsymbol{g}$ such that a functional of $\boldsymbol{u}$ and $\boldsymbol{g}$ is minimized subject to the constraint that the equations corresponding to a Ladyzhenskaya model of viscous, incompressible flow [9] are satisfied. Specifically, the state and

* Corresponding author.

[1] This work was supported in part by the US Department of Energy under grant number DE-FG02-931225172 for QD, the Air Force Office of Scientific Research under grants number AFOSR-93-1-0061 and AFOSR-F49620-93-1-0280 and the Office of Naval Research under grant number N-00014-91-J-1493 for MDG, and the National Science and Engineering Research Council of Canada under grant number OGP-0137436 for LSH.

SSDI 0377-0427(95)00072-7

control are required to satisfy

$$-\left(\nu_0 + \nu_1 \int_\Omega |\operatorname{grad} \boldsymbol{u}|^2 \, \mathrm{d}\Omega\right) \Delta \boldsymbol{u} + \boldsymbol{u} \cdot \operatorname{grad} \boldsymbol{u} + \operatorname{grad} p = \boldsymbol{f} + \boldsymbol{g} \quad \text{in } \Omega, \tag{1.1}$$

$$\operatorname{div} \boldsymbol{u} = 0 \quad \text{in } \Omega, \tag{1.2}$$

and

$$\boldsymbol{u} = 0 \quad \text{on } \Gamma, \tag{1.3}$$

where Ω denotes a $C^{1,1}$ or convex bounded domain in $\mathbb{R}^d$, $d = 2$ or 3, with a boundary Γ, and ν_0, ν_1 are positive constants. (When finite element approximations are considered, we will assume that Ω is a convex polyhedral domain.) If $\nu_1 = 0$, then (1.1)–(1.3) reduce to the well-known Navier–Stokes equations; in this case, if the variables in (1.1)–(1.3) are nondimensionalized, then ν_0 is simply the inverse of the Reynolds number Re. In (1.1)–(1.3), $\boldsymbol{u}$ and p denote the velocity and pressure fields, respectively, $\boldsymbol{f}$ a given body force, and $\boldsymbol{g}$ a distributed control. The constant density ρ has been absorbed into $p, \boldsymbol{f}$, and $\boldsymbol{g}$.

The model (1.1)–(1.3) is one of a class of models having nonlinear constitutive relations that were introduced by Ladyzhenskaya [9] as possible alternatives to the Navier–Stokes model. These models have been recently attracting considerable attention; see, e.g., [3, 5, 8, 10]. Among the reasons for this interest is the realization that the Ladyzhenskaya models may be interpreted as algebraic turbulence models; see, e.g., [5].

The two functionals that we consider are given by

$$\mathscr{J}(\boldsymbol{u}, \boldsymbol{g}) = \frac{1}{2} \int_\Omega |\boldsymbol{u} - \boldsymbol{u}_0|^2 \, \mathrm{d}\Omega + \frac{1}{2} \int_\Omega |\boldsymbol{g}|^2 \, \mathrm{d}\Omega \tag{1.4}$$

and

$$\mathscr{K}(\boldsymbol{u}, \boldsymbol{g}) = \frac{1}{4}\left(2\nu_0 + \nu_1 \int_\Omega |\operatorname{grad} \boldsymbol{u}|^2 \, \mathrm{d}\Omega\right) \int_\Omega |\operatorname{grad} \boldsymbol{u}|^2 \, \mathrm{d}\Omega - \int_\Omega (\boldsymbol{f} + \boldsymbol{g}) \cdot \boldsymbol{u} \, \mathrm{d}\Omega + \frac{1}{2\delta} \int_\Omega |\boldsymbol{g}|^2 \, \mathrm{d}\Omega, \tag{1.5}$$

where $\delta > 0$ is a constant. The first of these effectively measures the difference between the velocity field $\boldsymbol{u}$ and a prescribed field $\boldsymbol{u}_0$, while the second measures the drag due to viscosity. For a discussion of the relation between (1.5) and the viscous drag, one may mimic the derivation given in [11] for the analogous expression in the case of the classical Navier–Stokes equations. The appearance of the control $\boldsymbol{g}$ in (1.4) and (1.5) is necessary since we will not impose any a priori constraints on the size of the control.

The plan of the paper is as follows. In the remainder of this section, we introduce the notation that will be used throughout the paper. Then, in Section 2, we give a precise statement of the optimization problem for the functional (1.4) and discuss the main results we have obtained concerning this problem. In Section 3, we define finite element algorithms for the approximation of solutions of the optimization problem; we also discuss the main results we have obtained concerning the existence and convergence of these approximations. In Section 4, we collect the proofs of the results given in Sections 2 and 3. Finally, in Section 5, we consider the drag functional (1.5).

Throughout, C will denote a positive constant whose meaning and value changes with context. $H^r(\mathscr{D})$, $r \in \mathbb{R}$, denotes the standard Sobolev space of order r with respect to the set $\mathscr{D}$, where $\mathscr{D}$ is either the flow domain Ω or its boundary Γ; note that $H^0(\mathscr{D}) = L^2(\mathscr{D})$. Norms of functions belonging to $H^r(\Omega)$ and $H^r(\Gamma)$ are denoted by $\|\cdot\|_r$ and $\|\cdot\|_{r,\Gamma}$, respectively. Corresponding Sobolev spaces of vector-valued functions will be denoted by $\boldsymbol{H}^r(\mathscr{D})$, e.g., $\boldsymbol{H}^1(\Omega) = [H^1(\Omega)]^d$. Norms for spaces of vector-valued functions will be denoted the by same notation as that used for their scalar counterparts. For example,

$$\|\boldsymbol{v}\|^r_{L^r(\Omega)} = \sum_{j=1}^{d} \|v_j\|^r_{L^r(\Omega)} \quad \text{and} \quad \|\boldsymbol{v}\|_1^2 = \sum_{j=1}^{d} \|v_j\|_1^2,$$

where v_j, $j = 1, \dots, d$, denotes the components of $\boldsymbol{v}$. We define, for (pq) and $(\boldsymbol{u}\cdot\boldsymbol{v}) \in L^1(\Omega)$,

$$(p,q) = \int_\Omega pq \,\mathrm{d}\Omega \quad \text{and} \quad (\boldsymbol{u},\boldsymbol{v}) = \int_\Omega \boldsymbol{u}\cdot\boldsymbol{v}\,\mathrm{d}\Omega$$

and, for (pq) and $(\boldsymbol{u}\cdot\boldsymbol{v}) \in L^1(\Gamma)$,

$$(p,q)_\Gamma = \int_\Gamma pq\,\mathrm{d}\Gamma \quad \text{and} \quad (\boldsymbol{u},\boldsymbol{v})_\Gamma = \int_\Gamma \boldsymbol{u}\cdot\boldsymbol{v}\,\mathrm{d}\Gamma.$$

Thus, the inner products in $L^2(\Omega)$ and $\boldsymbol{L}^2(\Omega)$ are both be denoted by $(\cdot,\cdot)$ and those in $L^2(\Gamma)$ and $\boldsymbol{L}^2(\Gamma)$ by $(\cdot,\cdot)_\Gamma$. If X denotes a Banach space, X^* will denote its dual. Also, since in our context $\boldsymbol{L}^2(\Omega)$ or $\boldsymbol{L}^2(\Gamma)$ will play the role of a pivot space between X and X^*, $(\cdot,\cdot)$ or $(\cdot,\cdot)_\Gamma$ (as the case may be) also denotes the duality pairing of X and X^*. For details concerning these matters, see [1, 2 or 6].

We will use the forms

$$a(\boldsymbol{u},\boldsymbol{v}) = \int_\Omega (\operatorname{grad}\boldsymbol{u}):(\operatorname{grad}\boldsymbol{v})\,\mathrm{d}\Omega \quad \forall\, \boldsymbol{u},\boldsymbol{v} \in \boldsymbol{H}^1(\Omega),$$

$$b(\boldsymbol{v},q) = -\int_\Omega q \operatorname{div}\boldsymbol{v}\,\mathrm{d}\Omega \quad \forall\, \boldsymbol{v} \in \boldsymbol{H}^1(\Omega) \quad \text{and} \quad \forall\, q \in L^2(\Omega),$$

and

$$c(\boldsymbol{u},\boldsymbol{v},\boldsymbol{w}) = \int_\Omega \boldsymbol{u}\cdot\operatorname{grad}\boldsymbol{v}\cdot\boldsymbol{w}\,\mathrm{d}\Omega \quad \forall\, \boldsymbol{u},\boldsymbol{v},\boldsymbol{w} \in \boldsymbol{H}^1(\Omega),$$

where, for $\boldsymbol{u} = (u_1, u_2, \dots, u_d)$ and $\boldsymbol{v} = (v_1, v_2, \dots, v_d)$, we have

$$(\operatorname{grad}\boldsymbol{u}):(\operatorname{grad}\boldsymbol{v}) = \sum_{i,j=1}^{d} \frac{\partial u_j}{\partial x_i}\frac{\partial v_j}{\partial x_i}.$$

These forms are continuous in the sense that there exist constants $C_1 > 0$, $C_2 > 0$, and $C_3 > 0$ such that

$$|a(\boldsymbol{u},\boldsymbol{v})| \leqslant C_1 \|\boldsymbol{u}\|_1 \|\boldsymbol{v}\|_1 \quad \forall\, \boldsymbol{u},\boldsymbol{v} \in \boldsymbol{H}^1(\Omega), \tag{1.6}$$

$$|b(\boldsymbol{v},q)| \leqslant C_2 \|\boldsymbol{v}\|_1 \|q\|_0 \quad \forall\, \boldsymbol{v} \in \boldsymbol{H}^1(\Omega) \quad \text{and} \quad q \in L^2(\Omega), \tag{1.7}$$

and

$$|c(\boldsymbol{u},\boldsymbol{v},\boldsymbol{w})| \leqslant C_3 \,\|\boldsymbol{u}\|_1 \,\|\boldsymbol{v}\|_1 \,\|\boldsymbol{w}\|_1 \quad \forall\, \boldsymbol{u},\boldsymbol{v},\boldsymbol{w} \in \boldsymbol{H}^1(\Omega). \tag{1.8}$$

Moreover, there exist constants $\alpha > 0$ and $\beta > 0$ such that

$$a(\boldsymbol{v},\boldsymbol{v}) \geqslant \alpha \,\|\boldsymbol{v}\|_1^2 \quad \forall\, \boldsymbol{v} \in \boldsymbol{H}_0^1(\Omega) \tag{1.9}$$

and

$$\sup_{0 \neq \boldsymbol{v} \in \boldsymbol{H}_0^1(\Omega)} \frac{b(\boldsymbol{v},q)}{\|\boldsymbol{v}\|_1} \geqslant \beta \,\|q\|_0 \quad \forall\, q \in L_0^2(\Omega), \tag{1.10}$$

where $\boldsymbol{H}_0^1(\Omega) = \{\boldsymbol{v} \in \boldsymbol{H}^1(\Omega)\colon \boldsymbol{v} = \boldsymbol{0} \text{ on } \Gamma\}$ and $L_0^2(\Omega) = \{q \in L^2(\Omega)\colon \int_\Omega q \,\mathrm{d}\Omega = 0\}$. For details concerning these forms and inequalities (1.6)–(1.10), one may consult [6, 7 or 9].

2. The optimization problem and the optimality system

We begin by giving a precise statement of the first optimization problem we consider. Let $\boldsymbol{g} \in \boldsymbol{L}^2(\Omega)$ denote the distributed control and let $\boldsymbol{u} \in V := \{\boldsymbol{v} \in \boldsymbol{H}_0^1(\Omega)\colon \operatorname{div} \boldsymbol{v} = 0\}$ denote the state, i.e., the velocity field. The state and control variables are constrained to satisfy the Ladyzhenskaya equations in the weak form (see, e.g., [5, 9])

$$[\nu_0 + \nu_1 a(\boldsymbol{u},\boldsymbol{u})]\, a(\boldsymbol{u},\boldsymbol{v}) + c(\boldsymbol{u},\boldsymbol{u},\boldsymbol{v}) = (\boldsymbol{f} + \boldsymbol{g}, \boldsymbol{v}) \quad \forall\, \boldsymbol{v} \in V, \tag{2.1}$$

where $\boldsymbol{f} \in \boldsymbol{L}^2(\Omega)$ is a given function.

The functional (1.4), using the notation introduced in Section 1, is given by

$$\mathscr{J}(\boldsymbol{u},\boldsymbol{g}) = \tfrac{1}{2} \,\|\boldsymbol{u} - \boldsymbol{u}_0\|_0^2 + \tfrac{1}{2} \,\|\boldsymbol{g}\|_0^2, \tag{2.2}$$

where $\boldsymbol{u}_0 \in \boldsymbol{L}^2(\Omega)$ is a given function. The *admissibility set* $\mathscr{U}_{\mathrm{ad}}$ is defined by

$$\mathscr{U}_{\mathrm{ad}} = \{(\boldsymbol{u},\boldsymbol{g}) \in V \times \boldsymbol{L}^2(\Omega)\colon (2.1) \text{ is satisfied}\}. \tag{2.3}$$

Then, $(\hat{\boldsymbol{u}},\hat{\boldsymbol{g}}) \in \mathscr{U}_{\mathrm{ad}}$ is called an *optimal solution* if there exists $\varepsilon > 0$ such that

$$\mathscr{J}(\hat{\boldsymbol{u}},\hat{\boldsymbol{g}}) \leqslant \mathscr{J}(\boldsymbol{u},\boldsymbol{g}) \quad \forall\, (\boldsymbol{u},\boldsymbol{g}) \in \mathscr{U}_{\mathrm{ad}} \text{ satisfying } \|\boldsymbol{u} - \hat{\boldsymbol{u}}|_1 + \|\boldsymbol{g} - \hat{\boldsymbol{g}}\|_0 \leqslant \varepsilon. \tag{2.4}$$

Hence, optimal solutions are defined as local minima. The first main result that we obtain is (see Theorem 4.6):

there exists a $(\hat{\boldsymbol{u}},\hat{\boldsymbol{g}}) \in \mathscr{U}_{\mathrm{ad}}$ *such that* (2.2) *is minimized in the sense of* (2.4).

Due to the definition (2.3) for $\mathscr{U}_{\mathrm{ad}}$, we see that the problem of finding $(\hat{\boldsymbol{u}},\hat{\boldsymbol{g}}) \in \mathscr{U}_{\mathrm{ad}}$ satisfying (2.4) is a constrained optimization problem. We wish to use the Lagrange multiplier rule to turn this constrained optimization problem into an unconstrained one. Proceeding formally, we introduce the Lagrange multipliers $\boldsymbol{\xi} \in \boldsymbol{H}_0^1(\Omega)$ and $\sigma \in L_0^2(\Omega)$ and define the product space

$$\chi = \boldsymbol{H}_0^1(\Omega) \times L_0^2(\Omega) \times \boldsymbol{L}^2(\Omega) \times \boldsymbol{H}_0^1(\Omega) \times L_0^2(\Omega)$$

and the Lagrangian

$$\mathscr{M}(\boldsymbol{u},p,\boldsymbol{g},\boldsymbol{\xi},\sigma) = \mathscr{J}(\boldsymbol{u},\boldsymbol{g}) - \{[\nu_0 + \nu_1 a(\boldsymbol{u},\boldsymbol{u})]a(\boldsymbol{u},\boldsymbol{\xi}) + c(\boldsymbol{u},\boldsymbol{u},\boldsymbol{\xi}) + b(\boldsymbol{\xi},p) + b(\boldsymbol{u},\sigma) - (\boldsymbol{f}+\boldsymbol{g},\boldsymbol{\xi})\} \quad \forall (\boldsymbol{u},p,\boldsymbol{g},\boldsymbol{\xi},\sigma) \in \chi. \tag{2.5}$$

We now seek stationary points of $\mathscr{M}(\boldsymbol{u},p,\boldsymbol{g},\boldsymbol{\xi},\sigma)$ over χ.

Again, proceeding in a formal manner, using standard techniques of the calculus of variations, one may derive the Euler–Lagrange equations that correspond to the minimization of (2.5). This process yields the *optimality system*

$$[\nu_0 + \nu_1 a(\hat{\boldsymbol{u}},\hat{\boldsymbol{u}})]a(\hat{\boldsymbol{u}},\boldsymbol{v}) + c(\hat{\boldsymbol{u}},\hat{\boldsymbol{u}},\boldsymbol{v}) + b(\boldsymbol{v},\hat{p}) = (\boldsymbol{f}+\hat{\boldsymbol{g}},\boldsymbol{v}) \quad \forall \boldsymbol{v} \in \boldsymbol{H}_0^1(\Omega), \tag{2.6}$$

$$b(\hat{\boldsymbol{u}},q) = 0 \quad \forall q \in L_0^2(\Omega), \tag{2.7}$$

$$[\nu_0 + \nu_1 a(\hat{\boldsymbol{u}},\hat{\boldsymbol{u}})]a(\boldsymbol{\xi},\boldsymbol{\omega}) + 2\nu_1 a(\boldsymbol{\xi},\hat{\boldsymbol{u}})a(\hat{\boldsymbol{u}},\boldsymbol{\omega}) + c(\boldsymbol{\omega},\hat{\boldsymbol{u}},\boldsymbol{\xi}) + c(\hat{\boldsymbol{u}},\boldsymbol{\omega},\boldsymbol{\xi}) + b(\boldsymbol{\omega},\sigma) = (\hat{\boldsymbol{u}} - \boldsymbol{u}_0,\boldsymbol{\omega}) \quad \forall \boldsymbol{\omega} \in \boldsymbol{H}_0^1(\Omega) \tag{2.8}$$

$$b(\boldsymbol{\xi},\psi) = 0 \quad \forall \psi \in L_0^2(\Omega), \tag{2.9}$$

and

$$(\hat{\boldsymbol{g}},\boldsymbol{s}) + (\boldsymbol{s},\boldsymbol{\xi}) = 0 \quad \forall \boldsymbol{s} \in \boldsymbol{L}^2(\Omega). \tag{2.10}$$

Variations in the Lagrange multipliers $\boldsymbol{\xi}$ and σ recover the constraints (2.6) and (2.7). Variations in the state variables $\boldsymbol{u}$ and p yield the co-state equations (2.8) and (2.9) and variations in the control $\boldsymbol{g}$ yield (2.10). Thus, the optimal solution necessarily satisfies the optimality system (2.6)–(2.10).

Our second main result (see Theorem 4.8) is to make the above formal process of obtaining the optimality system through the use of the Lagrange multiplier rule a rigorous one:

let $(\hat{\boldsymbol{u}},\hat{p},\hat{\boldsymbol{g}}) \in \boldsymbol{H}_0^1(\Omega) \times L_0^2(\Omega) \times \boldsymbol{L}_0^2(\Omega)$ *denote an optimal solution in the sense of* (2.4); *then, there exists a nonzero multiplier* $(\boldsymbol{\xi},\sigma) \in \boldsymbol{H}_0^1(\Omega) \times L_0^2(\Omega)$ *satisfying the optimality system* (2.6)–(2.10).

Note that (2.10) enables us to eliminate the optimal control $\hat{\boldsymbol{g}}$ from (2.6), resulting in

$$[\nu_0 + \nu_1 a(\hat{\boldsymbol{u}},\hat{\boldsymbol{u}})]a(\hat{\boldsymbol{u}},\boldsymbol{v}) + c(\hat{\boldsymbol{u}},\hat{\boldsymbol{u}},\boldsymbol{v}) + b(\boldsymbol{v},\hat{p}) = (\boldsymbol{f} - \boldsymbol{\xi},\boldsymbol{v}) \quad \forall \boldsymbol{v} \in \boldsymbol{H}_0^1(\Omega). \tag{2.11}$$

Then, the optimality system in terms of the optimal state $(\hat{\boldsymbol{u}},\hat{p})$ and co-state $(\boldsymbol{\xi},\sigma)$ is given by (2.7)–(2.9) and (2.11). Once the state variables $\hat{\boldsymbol{u}}$ and $\hat{p}$ and the Lagrange multipliers $\boldsymbol{\xi}$ and σ are determined, the optimal control may be easily deduced from the optimality condition (2.10), i.e., we essentially have that $\hat{\boldsymbol{g}} = -\boldsymbol{\xi}$.

Remark. A *strong form* of the optimality system may be obtained by the usual application of integration by parts. Indeed, one sees that (2.6)–(2.10) constitutes a weak formulation of the problem

$$-[\nu_0 + \nu_1 a(\hat{\boldsymbol{u}},\hat{\boldsymbol{u}})]\Delta\hat{\boldsymbol{u}} + \hat{\boldsymbol{u}}\cdot\operatorname{grad}\hat{\boldsymbol{u}} + \operatorname{grad}\hat{p} = \boldsymbol{f} + \hat{\boldsymbol{g}} \quad \text{in } \Omega, \tag{2.12}$$

$$\operatorname{div}\hat{\boldsymbol{u}} = 0 \quad \text{in } \Omega, \tag{2.13}$$

$$\hat{\boldsymbol{u}} = \boldsymbol{0} \quad \text{on } \Gamma, \tag{2.14}$$

$$-[\nu_0 + \nu_1 a(\hat{\boldsymbol{u}},\hat{\boldsymbol{u}})]\Delta\boldsymbol{\xi} - 2\nu_1 a(\hat{\boldsymbol{u}},\boldsymbol{\xi})\Delta\hat{\boldsymbol{u}} + \boldsymbol{\xi}\cdot(\operatorname{grad}\hat{\boldsymbol{u}})^{\mathrm{T}} - \hat{\boldsymbol{u}}\cdot\operatorname{grad}\boldsymbol{\xi} + \operatorname{grad}\sigma$$

$$= \hat{\boldsymbol{u}} - \boldsymbol{u}_0 \quad \text{in } \Omega, \tag{2.15}$$

$$\operatorname{div}\boldsymbol{\xi} = 0 \quad \text{in } \Omega, \tag{2.16}$$

$$\boldsymbol{\xi} = \boldsymbol{0} \quad \text{on } \Gamma, \tag{2.17}$$

and

$$\hat{\boldsymbol{g}} = -\boldsymbol{\xi} \quad \text{in } \Omega. \tag{2.18}$$

Note that in (2.15)

$$(\hat{\boldsymbol{u}}\cdot\operatorname{grad}\boldsymbol{\xi})_i = \sum_{j=1}^{d} \hat{u}_j \frac{\partial \xi_i}{\partial x_j} \text{ and } (\boldsymbol{\xi}\cdot(\operatorname{grad}\hat{\boldsymbol{u}})^{\mathrm{T}})_i = \sum_{j=1}^{d} \xi_j \frac{\partial \hat{u}_j}{\partial x_i} \qquad \text{for } i = 1,\dots,d.$$

The optimality system (2.12)–(2.18) includes of the Navier–Stokes system (2.12)–(2.14) and the system (2.15)–(2.17) whose left-hand side is the adjoint of the Navier–Stokes operator linearized about $(\hat{\boldsymbol{u}},\hat{p})$.

Remark. An equivalent weak formulation of the optimality system (2.7)–(2.9) and (2.11) that we use later is given by

$$[\nu_0 + \nu_1 a(\hat{\boldsymbol{u}},\hat{\boldsymbol{u}})]a(\hat{\boldsymbol{u}},\boldsymbol{v}) + c(\hat{\boldsymbol{u}},\hat{\boldsymbol{u}},\boldsymbol{v}) = (\boldsymbol{f} - \boldsymbol{\xi},\boldsymbol{v}) \quad \forall\, \boldsymbol{v} \in \boldsymbol{V} \tag{2.19}$$

and

$$[\nu_0 + \nu_1 a(\hat{\boldsymbol{u}},\hat{\boldsymbol{u}})]a(\boldsymbol{\xi},\boldsymbol{\omega}) + 2\nu_1 a(\boldsymbol{\xi},\hat{\boldsymbol{u}})a(\boldsymbol{u},\boldsymbol{\omega}) + c(\boldsymbol{\omega},\hat{\boldsymbol{u}},\boldsymbol{\xi}) + c(\hat{\boldsymbol{u}},\boldsymbol{\omega},\boldsymbol{\xi}) = (\hat{\boldsymbol{u}} - \boldsymbol{u}_0,\boldsymbol{\omega}) \quad \forall\, \boldsymbol{\omega} \in \boldsymbol{V}. \tag{2.20}$$

Once $\hat{\boldsymbol{u}}$ and $\boldsymbol{\xi}$ are determined, one can recover $\hat{p}$ and σ from (2.11) and (2.8).

Remark. Our notion of an optimal solution is a local one; see (2.4). Moreover, there is no reason to believe that, in general, optimal solutions are unique. This is to be expected since even the uncontrolled stationary Navier–Stokes equations are known to have multiple solutions for sufficiently large values of the Reynolds number. However, just as in the Navier–Stokes case (see, e.g., [6, 7, 12 or 13]), for sufficiently small values of the Reynolds number, i.e., for "small enough" data or "large enough" viscosity, one can guarantee that optimal solutions are unique.

In order to determine the rate of convergence of finite element approximations to the solutions of the optimality system, one must have knowledge about the smoothness of these solutions. Thus, our next main result concerns the regularity of solutions of the optimality system. The precise result is given in Theorem 4.9 (see also the remark that follows that theorem); here, we merely note that

the regularity of solutions of the optimality system (2.6)–(2.10) *is the same as that present for solutions of the analogous problem for the Navier–Stokes equations.*

Thus, $\hat{\boldsymbol{u}}$ and $\boldsymbol{\xi}$ are as smooth as the velocity field and $\hat{p}$ and σ are as smooth as the pressure field obtained from the Navier–Stokes equations posed over the same domain Ω and having the same data $\boldsymbol{f}$ as the optimality system (2.6)–(2.10).

3. Finite element approximations

We now define, using finite element methods, an approximate optimality system from which approximations to the optimal state, co-state, and control may be determined. It is important to emphasize at the beginning that the finite element methods that may be employed to this end are exactly those that may be used for determining approximate solutions of the Navier–Stokes equations. Thus, the same finite element spaces may be used for the pressure and velocity approximations (and for the corresponding adjoint variables) as those used for the corresponding variables in the Navier–Stokes equations. Thus, one may consult the vast literature and long catalog of stable finite element velocity–pressure pairs that are available for the Navier–Stokes equations; see, e.g., [6, 7].

A finite element discretization of the optimality system (2.7)–(2.9) and (2.11) is defined in the usual manner. First one chooses families of finite dimensional subspaces $\boldsymbol{X}^h \subset \boldsymbol{H}_0^1(\Omega)$ and $S^h \subset L_0^2(\Omega)$. These families are parametrized by a parameter h that tends to zero; commonly, h is chosen to be some measure of the grid size. It is natural to assume that as $h \to 0$,

$$\inf_{\boldsymbol{v}^h \in \boldsymbol{X}^h} \|\boldsymbol{v} - \boldsymbol{v}^h\|_1 \to 0 \quad \forall\, \boldsymbol{v} \in \boldsymbol{H}_0^1(\Omega) \tag{3.1}$$

and

$$\inf_{q^h \in S^h} \|q - q^h\|_0 \to 0 \quad \forall\, q \in L_0^2(\Omega). \tag{3.2}$$

Here we may choose any pair of subspaces $\boldsymbol{X}^h$ and S^h that can be used for finding finite element approximations of solutions of the Navier–Stokes equations and we make the same assumptions as are employed in that setting. Thus, we assume the *inf-sup condition*, or *Ladyzhenskaya–Babuška–Brezzi condition*: there exists a constant C, independent of h, such that

$$\inf_{0 \neq q^h \in S^h} \sup_{0 \neq \boldsymbol{v}^h \in \boldsymbol{X}^h} \frac{b(\boldsymbol{v}^h, q^h)}{\|\boldsymbol{v}^h\|_1 \, \|q^h\|_0} \geqslant C. \tag{3.3}$$

This condition assures the stability of finite element discretizations of the Navier–Stokes equations. We shall see that it also assures the stability of the approximation of the Ladyzhenskaya model and the optimality system. Similar discussions may be found in [5]. For thorough discussions of the approximation properties (3.1) and (3.2), see, e.g., [4] and for like discussions of the stability condition (3.3), see, e.g., [6 or 7]. These references may also be consulted for a catalog of finite element subspaces that meet the requirements of (3.1)–(3.3).

In the sequel we will use the following modified trilinear form $c(\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w})$: for any $(\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}) \in [\boldsymbol{H}^1(\Omega)]^3$,

$$c(\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}) = \frac{1}{2} \int_\Omega [\boldsymbol{u} \cdot \operatorname{grad} \boldsymbol{v} \cdot \boldsymbol{w} - \boldsymbol{u} \cdot \operatorname{grad} \boldsymbol{w} \cdot \boldsymbol{v}] \, \mathrm{d}\Omega.$$

Note that for $(\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}) \in [\boldsymbol{V}^3]$, this definition coincides with the original definition. Also, the modified $c(\cdot,\cdot,\cdot)$ satisfies

$$c(\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}) = 0 \quad \forall\, \boldsymbol{u}, \boldsymbol{v} \in \boldsymbol{H}^1(\Omega)$$

and, for some constant $C_0 > 0$,

$$|c(\boldsymbol{u},\boldsymbol{v},\boldsymbol{w})| \leqslant C_0 \,\|\operatorname{grad}\boldsymbol{u}\|_0 \,\|\operatorname{grad}\boldsymbol{v}\|_0 \,\|\operatorname{grad}\boldsymbol{w}\|_0 \quad \forall\, \boldsymbol{u},\boldsymbol{v},\boldsymbol{w} \in \boldsymbol{H}_0^1(\Omega). \tag{3.4}$$

Once the approximating subspaces have been chosen, we can define the approximate problem from which approximate states and co-states may be determined: seek $\boldsymbol{u}^h \in \boldsymbol{X}^h$, $p^h \in S^h$, $\xi^h \in \boldsymbol{X}^h$, and $\sigma^h \in S^h$ such that

$$[\nu_0 + \nu_1 a(\boldsymbol{u}^h,\boldsymbol{u}^h)]a(\boldsymbol{u}^h,\boldsymbol{v}^h) + c(\boldsymbol{u}^h,\boldsymbol{u}^h,\boldsymbol{v}^h) + b(\boldsymbol{v}^h,p^h) = (\boldsymbol{f} - \xi^h,\boldsymbol{v}^h) \quad \forall\, \boldsymbol{v}^h \in \boldsymbol{X}^h, \tag{3.5}$$

$$b(\boldsymbol{u}^h,q^h) = 0 \quad \forall\, q^h \in S^h, \tag{3.6}$$

$$\begin{aligned}&[\nu_0 + \nu_1 a(\boldsymbol{u}^h,\boldsymbol{u}^h)]a(\xi^h,\omega^h) + 2\nu_1 a(\boldsymbol{u}^h,\xi^h)a(\boldsymbol{u}^h,\omega^h) + c(\omega^h,\boldsymbol{u}^h,\xi^h) + c(\boldsymbol{u}^h,\omega^h,\xi^h) + b(\omega^h,\sigma^h)\\ &\quad = (\boldsymbol{u}^h - \boldsymbol{u}_0,\omega^h) \quad \forall\, \omega^h \in \boldsymbol{X}^h,\end{aligned} \tag{3.7}$$

and

$$b(\xi^h,\psi^h) = 0 \quad \forall\, \psi^h \in S^h. \tag{3.8}$$

Following (2.10) or (2.18), we define the approximate control by

$$\boldsymbol{g}^h = -\xi^h \quad \text{in } \Omega.$$

The main results concerning finite element approximations are proved in Sections 4.4 and 4.5. These may be summarized as follows:

under certain conditions on the data ν_0, ν_1, and $\boldsymbol{f}$, one can show that solutions of the approximate optimality system (3.5)–(3.8) *exist and converge to solutions of the optimality system* (2.7)–(2.9) *and* (2.11); *under more stringent restrictions on the data, one may also show that the solution of the approximate optimality system* (3.5)–(3.8) *is unique.*

One may also show that if one choose proper finite element subspaces such that the stability condition (3.3) is satisfied and such that the errors in (3.1) and (3.2) have the same asymptotic order of convergence in h as $h \to 0$, then the same asymptotic order of convergence holds for the finite element approximation $(\boldsymbol{u}^h, p^h, \xi^h, \sigma^h)$, provided that ν_0 is sufficiently large and that the regularity results presented in Section 4.3 hold. Then, for example, if one uses a Taylor–Hood element pair (see, e.g., [6 or 7]) consisting of continuous piecewise quadratic velocity approximations and continuous piecewise linear pressure approximations, both defined with respect to the same grid, then provided solutions are regular enough, we achieve $\mathrm{O}(h^2)$ convergence for the $H^1(\Omega)$-norm of the velocity error and the $L^2(\Omega)$-norm of the pressure error.

4. Proofs of main results

We now provide proofs of the results given in Sections 3 and 4.

4.1. Existence of optimal solutions

We first show that an optimal solution exists. To this end we first present some useful results.

Lemma 4.1.

$$\begin{aligned} &a(\boldsymbol{u},\boldsymbol{u})a(\boldsymbol{u},\boldsymbol{u}-\boldsymbol{v}) - a(\boldsymbol{v},\boldsymbol{v})a(\boldsymbol{v},\boldsymbol{u}-\boldsymbol{v}) \\ &\quad \geqslant \tfrac{1}{4}(\|\operatorname{grad}\boldsymbol{u}\|_0^2 + \|\operatorname{grad}\boldsymbol{v}\|_0^2)\,\|\operatorname{grad}(\boldsymbol{u}-\boldsymbol{v})\|_0^2 + \tfrac{1}{4}[a(\boldsymbol{v},\boldsymbol{u}-\boldsymbol{v})]^2 \\ &\qquad + \tfrac{1}{4}[a(\boldsymbol{u},\boldsymbol{u}-\boldsymbol{v})]^2 \quad \forall\,\boldsymbol{u},\boldsymbol{v}\in \boldsymbol{H}^1(\Omega). \end{aligned}$$

Proof. By the Cauchy–Schwartz inequality, we have

$$\begin{aligned} &a(\boldsymbol{u},\boldsymbol{u})a(\boldsymbol{u},\boldsymbol{u}-\boldsymbol{v}) - a(\boldsymbol{v},\boldsymbol{v})a(\boldsymbol{v},\boldsymbol{u}-\boldsymbol{v}) \\ &\quad = a(\boldsymbol{u},\boldsymbol{u})a(\boldsymbol{u}-\boldsymbol{v},\boldsymbol{u}-\boldsymbol{v}) + [a(\boldsymbol{u},\boldsymbol{u}) - a(\boldsymbol{v},\boldsymbol{v})]a(\boldsymbol{v},\boldsymbol{u}-\boldsymbol{v}) \\ &\quad = a(\boldsymbol{u},\boldsymbol{u})a(\boldsymbol{u}-\boldsymbol{v},\boldsymbol{u}-\boldsymbol{v}) + a(\boldsymbol{u},\boldsymbol{u}-\boldsymbol{v})a(\boldsymbol{v},\boldsymbol{u}-\boldsymbol{v}) + a(\boldsymbol{v},\boldsymbol{u}-\boldsymbol{v})a(\boldsymbol{v},\boldsymbol{u}-\boldsymbol{v}) \\ &\quad \geqslant a(\boldsymbol{u},\boldsymbol{u})a(\boldsymbol{u}-\boldsymbol{v},\boldsymbol{u}-\boldsymbol{v}) - \tfrac{1}{2}[a(\boldsymbol{u},\boldsymbol{u}-\boldsymbol{v})]^2 - \tfrac{1}{2}[a(\boldsymbol{v},\boldsymbol{u}-\boldsymbol{v})]^2 + [a(\boldsymbol{v},\boldsymbol{u}-\boldsymbol{v})]^2 \\ &\quad \geqslant a(\boldsymbol{u},\boldsymbol{u})a(\boldsymbol{u}-\boldsymbol{v},\boldsymbol{u}-\boldsymbol{v}) + \tfrac{1}{2}[a(\boldsymbol{v},\boldsymbol{u}-\boldsymbol{v})]^2 - \tfrac{1}{2}[a(\boldsymbol{u},\boldsymbol{u})a(\boldsymbol{u}-\boldsymbol{v},\boldsymbol{u}-\boldsymbol{v})] \\ &\quad \geqslant \tfrac{1}{2}[a(\boldsymbol{v},\boldsymbol{u}-\boldsymbol{v})]^2 + \tfrac{1}{2}a(\boldsymbol{u},\boldsymbol{u})a(\boldsymbol{u}-\boldsymbol{v},\boldsymbol{u}-\boldsymbol{v}) \quad \forall\,\boldsymbol{u},\boldsymbol{v}\in \boldsymbol{H}^1(\Omega). \end{aligned}$$

Using the symmetry in $\boldsymbol{u}$ and $\boldsymbol{v}$, we obtain the result in the lemma. □

For convenience, let us define

$$a_0(\boldsymbol{w};\boldsymbol{u},\boldsymbol{v}) = \nu_0 a(\boldsymbol{u},\boldsymbol{v}) + \nu_1 a(\boldsymbol{w},\boldsymbol{w})a(\boldsymbol{u},\boldsymbol{v}) \quad \forall\,\boldsymbol{u},\boldsymbol{v},\boldsymbol{w}\in \boldsymbol{H}^1(\Omega). \tag{4.1}$$

From the previous lemma, we see that the nonlinear mapping $\boldsymbol{u}\mapsto a_0(\boldsymbol{u};\boldsymbol{u},\cdot)$ is of monotone type in $\boldsymbol{H}^1(\Omega)$. Thus, we have the following result.

Corollary 4.2. (i) *There exists a constant C such that*

$$a_0(\boldsymbol{u};\boldsymbol{u},\boldsymbol{u}) \geqslant C\,\|\boldsymbol{u}\|_1^2 \quad \forall\,\boldsymbol{u}\in \boldsymbol{H}_0^1(\Omega).$$

(ii) *For each $\boldsymbol{v}\in \boldsymbol{H}_0^1(\Omega)$, the mapping $\boldsymbol{u}\mapsto a_0(\boldsymbol{u};\boldsymbol{u},\boldsymbol{v})$ is sequentially weakly continuous on $\boldsymbol{H}_0^1(\Omega)$.*

The second part of the corollary follows from the sequential weak continuity of monotone operators; see, e.g., [6] for related discussions.

We next quote an abstract theorem on the existence of weak solutions for the problem

$$a_1(\boldsymbol{u};\boldsymbol{u},\boldsymbol{v}) = F(\boldsymbol{v}) \quad \forall\,\boldsymbol{v}\in \boldsymbol{X}, \tag{4.2}$$

where $\boldsymbol{X}$ is an separable Banach space equipped with norm $\|\cdot\|_X$ and the form $a_1(\cdot;\cdot,\cdot)$ is such that for each fixed $\boldsymbol{w}\in \boldsymbol{X}$, $a_1(\boldsymbol{w};\cdot,\cdot)$ is a continuous bilinear form.

Lemma 4.3. *Assume that there exists a constant C such that*

$$a_1(\boldsymbol{u};\boldsymbol{u},\boldsymbol{u}) \geqslant C\,\|\boldsymbol{u}\|_X^2 \quad \forall\,\boldsymbol{u}\in \boldsymbol{X}$$

and for each $\boldsymbol{v} \in \boldsymbol{X}$, the mapping $\boldsymbol{u} \mapsto a_1(\boldsymbol{u}; \boldsymbol{u}, \boldsymbol{v})$ is sequentially weakly continuous on $\boldsymbol{X}$. Then, for each $F \in \boldsymbol{X}^*$, (4.2) *has at least one solution $\boldsymbol{u} \in \boldsymbol{X}$. Furthermore, every solution $\boldsymbol{u}$ of* (4.2) *satisfies the estimate*

$$\|\boldsymbol{u}\|_X \leqslant \frac{1}{C} \|F\|_*.$$

Proof. See [6]. □

We are now in a position to establish the existence of a solution of (2.1).

Proposition 4.4. *If $\boldsymbol{f} + \boldsymbol{g} \in \boldsymbol{H}^{-1}(\Omega)$, then* (2.1) *has at least one solution $\boldsymbol{u} \in V$. Moreover, there exists a constant C, independent of ν_0, such that every solution $\boldsymbol{u}$ of* (2.1) *satisfies*

$$\|\boldsymbol{u}\|_1 \leqslant \frac{C}{\nu_0} \|\boldsymbol{f} + \boldsymbol{g}\|_{-1}. \tag{4.3}$$

Proof. We wish to apply Lemma 4.3 to (2.1). We choose the space X to be the Hilbert space V. We define

$$a_1(\boldsymbol{w}; \boldsymbol{u}, \boldsymbol{v}) = a_0(\boldsymbol{w}; \boldsymbol{u}, \boldsymbol{v}) + c(\boldsymbol{w}, \boldsymbol{u}, \boldsymbol{v}) \quad \forall\, \boldsymbol{w}, \boldsymbol{u}, \boldsymbol{v} \in V.$$

We immediately obtain the coercivity from Corollary 4.2 (i) and the fact that $c(\boldsymbol{u}, \boldsymbol{u}, \boldsymbol{u}) = 0$.

We now turn to the question of sequential weak continuity. As discussed in, e.g., [6], for each $\boldsymbol{v} \in V$, the weak continuity of the form $\boldsymbol{u} \mapsto c(\boldsymbol{u}, \boldsymbol{u}, \boldsymbol{v})$ may be verified as follows. From the relation

$$c(\boldsymbol{u}^{(n)}, \boldsymbol{u}^{(n)}, \boldsymbol{v}) = -\int_\Omega \boldsymbol{u}^{(n)} \cdot \operatorname{grad} \boldsymbol{v} \cdot \boldsymbol{u}^{(n)} \,\mathrm{d}\Omega \quad \forall\, \boldsymbol{v} \in C^\infty(\bar{\Omega}) \cap V$$

and the compact imbedding $\boldsymbol{H}_0^1(\Omega) \subset \boldsymbol{L}^2(\Omega)$ (i.e., if $\boldsymbol{u}^{(n)} \rightharpoonup \boldsymbol{u}$ in $\boldsymbol{H}_0^1(\Omega)$, then a subsequence $\boldsymbol{u}^{(n)} \to \boldsymbol{u}$ in $\boldsymbol{L}^2(\Omega)$), we obtain that

$$\lim_{n \to \infty} c(\boldsymbol{u}^{(n)}, \boldsymbol{u}^{(n)}, \boldsymbol{v}) = -\int_\Omega \boldsymbol{u} \cdot \operatorname{grad} \boldsymbol{v} \cdot \boldsymbol{u} \,\mathrm{d}\Omega = c(\boldsymbol{u}, \boldsymbol{u}, \boldsymbol{v}) \quad \forall\, \boldsymbol{v} \in C^\infty(\bar{\Omega}) \cap V.$$

Thus, the denseness of $C^\infty(\bar{\Omega})$ in $\boldsymbol{H}_0^1(\Omega)$ implies that

$$\lim_{n \to \infty} c(\boldsymbol{u}^{(n)}, \boldsymbol{u}^{(n)}, \boldsymbol{v}) = c(\boldsymbol{u}, \boldsymbol{u}, \boldsymbol{v}) \quad \forall\, \boldsymbol{v} \in V.$$

Combining this result with Corollary 4.2 (ii), we readily deduce, for each fixed $\boldsymbol{v}$, the sequential weak continuity of the mapping $\boldsymbol{u} \mapsto a_1(\boldsymbol{u}; \boldsymbol{u}, \boldsymbol{v})$. Hence the result of this proposition follows directly from Lemma 4.3. □

One can also prove the following regularity results:

Lemma 4.5. *If $\boldsymbol{f} + \boldsymbol{g} \in \boldsymbol{L}^2(\Omega)$, then $\boldsymbol{u} \in \boldsymbol{H}^2(\Omega)$. Moreover, there exists a constant C such that*

$$\|\boldsymbol{u}\|_2 \leqslant C \|\boldsymbol{f} + \boldsymbol{g}\|_0. \tag{4.4}$$

Proof. See [9]. □

We are now in a position to establish the existence of an optimal solution as defined in (2.4).

Theorem 4.6. *There exists a* $(\hat{\boldsymbol{u}},\hat{\boldsymbol{g}}) \in \mathscr{U}_{\mathrm{ad}}$ *such that* (2.2) *is minimized in the sense of* (2.4).

Proof. We first claim that $\mathscr{U}_{\mathrm{ad}}$ is not empty. The existence of a solution for (2.1) was given in Proposition 4.4 for any given right-hand side $\boldsymbol{f}+\boldsymbol{g} \in \boldsymbol{L}^2(\Omega)$. In particular, we deduce that there exists a $\tilde{\boldsymbol{u}} \in V$ that satisfies (2.1) with $\boldsymbol{g}=\boldsymbol{0}$. Moreover, we have $\mathscr{J}(\tilde{\boldsymbol{u}},\boldsymbol{0}) \leqslant C(\|\tilde{\boldsymbol{u}}\|_0^2 + \|\boldsymbol{u}_0\|_0^2) < \infty$. Thus, $(\tilde{\boldsymbol{u}},\boldsymbol{0}) \in \mathscr{U}_{\mathrm{ad}}$.

Now, let $\{\boldsymbol{u}^{(n)},\boldsymbol{g}^{(n)}\}$ be a sequence in $\mathscr{U}_{\mathrm{ad}}$ such that

$$\lim_{n\to\infty} \mathscr{J}(\boldsymbol{u}^{(n)},\boldsymbol{g}^{(n)}) = \inf_{(\boldsymbol{u},\boldsymbol{g})\in\mathscr{U}_{\mathrm{ad}}} \mathscr{J}(\boldsymbol{u},\boldsymbol{g}).$$

Then, by (2.2), $(\boldsymbol{u}^{(n)},\boldsymbol{g}^{(n)})$ is uniformly bounded in $\boldsymbol{L}^2(\Omega)\times\boldsymbol{L}^2(\Omega)$, and

$$[\nu_0 + \nu_1 a(\boldsymbol{u}^{(n)},\boldsymbol{u}^{(n)})]a(\boldsymbol{u}^{(n)},\boldsymbol{u}^{(n)}) + c(\boldsymbol{u}^{(n)},\boldsymbol{u}^{(n)},\boldsymbol{v}) = (\boldsymbol{f}+\boldsymbol{g}^{(n)},\boldsymbol{v}) \quad \forall\, \boldsymbol{v}\in V. \tag{4.5}$$

By (4.3), we have

$$\|\boldsymbol{u}^{(n)}\|_1 \leqslant C.$$

We may then extract subsequences such that

$$\boldsymbol{g}^{(n)} \rightharpoonup \hat{\boldsymbol{g}} \quad \text{in } \boldsymbol{L}^2(\Omega), \qquad \boldsymbol{u}^{(n)} \rightharpoonup \hat{\boldsymbol{u}} \quad \text{in } V, \qquad \boldsymbol{u}^{(n)} \to \hat{\boldsymbol{u}} \quad \text{in } \boldsymbol{L}^2(\Omega)$$

for some $(\hat{\boldsymbol{u}},\hat{\boldsymbol{g}}) \in V\times\boldsymbol{L}^2(\Omega)$. The last convergence result above follows from the compact imbedding $\boldsymbol{H}_0^1(\Omega) \subset \boldsymbol{L}^2(\Omega)$. We may then pass to the limit in (4.5). We may deduce that $(\hat{\boldsymbol{u}},\hat{\boldsymbol{g}})$ satisfies (2.1) using arguments similar to the ones given earlier to derive the result in Proposition 4.4.

Finally, by the weak lower semicontinuity of $\mathscr{J}(\cdot,\cdot)$, we conclude that $(\hat{\boldsymbol{u}},\hat{\boldsymbol{g}})$ is an optimal solution, i.e.,

$$\mathscr{J}(\hat{\boldsymbol{u}},\hat{\boldsymbol{g}}) = \inf_{(\boldsymbol{u},\boldsymbol{g})\in\mathscr{U}_{\mathrm{ad}}} \mathscr{J}(\boldsymbol{u},\boldsymbol{g}).$$

This proves the theorem. □

Remark. Because the optimal control $\hat{\boldsymbol{g}} \in \boldsymbol{L}^2(\Omega)$, we may deduce, using the regularity results in Lemma 4.5 for the Ladyzhenskaya equations, that $\hat{\boldsymbol{u}} \in \boldsymbol{H}^2(\Omega)$.

Remark. Using (1.7) and (1.10), one can show that, similar to the Navier–Stokes equations case, there exists a $\hat{p} \in L_0^2(\Omega)$ such that $(\hat{\boldsymbol{u}},\hat{p})$ is a weak solution of (1.1)–(1.3), i.e.,

$$[\nu_0 + \nu_1 a(\hat{\boldsymbol{u}},\hat{\boldsymbol{u}})]a(\hat{\boldsymbol{u}},\boldsymbol{v}) + c(\hat{\boldsymbol{u}},\hat{\boldsymbol{u}},\boldsymbol{v}) + b(\boldsymbol{v},\hat{p}) = (\boldsymbol{f}+\hat{\boldsymbol{g}},\boldsymbol{v}) \quad \forall\, \boldsymbol{v}\in\boldsymbol{H}_0^1(\Omega) \tag{4.6}$$

and

$$b(\hat{\boldsymbol{u}},q) = 0 \quad \forall\, q \in L_0^2(\Omega). \tag{4.7}$$

Since $\boldsymbol{f}+\hat{\boldsymbol{g}} \in \boldsymbol{L}^2(\Omega)$, we actually have $(\hat{\boldsymbol{u}},\hat{p}) \in \boldsymbol{H}^2(\Omega)\times H^1(\Omega)$.

4.2. Existence of Lagrange multipliers

We wish to use the method of Lagrange multipliers to turn the constrained optimization problem (2.4) into an unconstrained one. We begin by showing that suitable Lagrange multipliers exist.

We first quote the following abstract theorem concerning the existence of Lagrange multipliers for smooth constrained minimization problems on Banach spaces.

Proposition 4.7. *Let X and Y be two real Banach spaces, $\mathscr{J}$ a functional on X, and M a mapping from X to Y. Assume u is a solution of the following constrained minimization problem:*

$$\text{find } u \in X \text{ such that } \mathscr{J}(u) = \inf\{\mathscr{J}(v)\colon v \in X, M(v) = f_0\},$$

where f_0 is any fixed element of Y. Assume further the following conditions are satisfied:

(A) *$\mathscr{J}\colon \mathrm{Nbhd}(u) \subseteq X \to \mathbb{R}$ is Frechet-differentiable at u with Frechet derivative J',*

(B) *M is Frechet-differentiable in an open neighborhood of u and its Frechet derivative M' is continuous at u;*

(C) Range $(M'(u)) = Y$.

Then there exists a $\mu \in Y^*$ such that

$$-\langle \mathscr{J}'(u), w\rangle + \langle \mu, M'(u)w\rangle = 0 \quad \forall\, w \in X.$$

Proof. See [14]. □

Let $X = \boldsymbol{H}_0^1(\Omega) \times L_0^2(\Omega) \times \boldsymbol{L}^2(\Omega)$ and $Y = \boldsymbol{H}^{-1}(\Omega) \times L_0^2(\Omega)$ and let the functional $\mathscr{J}$ be defined as in (1.4) and the nonlinear mapping $M\colon X \to Y$ denote the (generalized) constraint equations, i.e., $M(\boldsymbol{u}, p, \boldsymbol{g}) = (\boldsymbol{f}, \phi)$ for $(\boldsymbol{u}, p, \boldsymbol{g}) \in X$ and $(\boldsymbol{f}, \phi) \in Y$ if and only if

$$[\nu_0 + \nu_1 a(\boldsymbol{u}, \boldsymbol{u})]a(\boldsymbol{u}, \boldsymbol{v}) + c(\boldsymbol{u}, \boldsymbol{u}, \boldsymbol{v}) + b(\boldsymbol{v}, p) - (\boldsymbol{g}, \boldsymbol{v}) = (\boldsymbol{f}, \boldsymbol{v}) \quad \forall\, \boldsymbol{v} \in \boldsymbol{H}_0^1(\Omega) \tag{4.8}$$

and

$$b(\boldsymbol{u}, q) = (\phi, q) \quad \forall\, q \in L_0^2(\Omega). \tag{4.9}$$

One may show that M is a C^1-mapping and its Frechet derivative $M'(\hat{\boldsymbol{u}}, \hat{p}, \hat{\boldsymbol{g}}) \in \mathscr{L}(X; Y)$ is defined as follows: $M'(\hat{\boldsymbol{u}}, \hat{p}, \hat{\boldsymbol{g}}) \cdot (\boldsymbol{w}, r, \boldsymbol{s}) = (\bar{\boldsymbol{f}}, \bar{\phi})$ for $(\boldsymbol{w}, r, \boldsymbol{s}) \in X$, and $(\bar{\boldsymbol{f}}, \bar{\phi}) \in Y$, if and only if

$$\begin{aligned}[\nu_0 + \nu_1 a(\hat{\boldsymbol{u}}, \hat{\boldsymbol{u}})]a(\boldsymbol{w}, \boldsymbol{v}) + 2\nu_1 a(\boldsymbol{w}, \hat{\boldsymbol{u}})a(\hat{\boldsymbol{u}}, \boldsymbol{v}) + c(\hat{\boldsymbol{u}}, \boldsymbol{w}, \boldsymbol{v}) + c(\boldsymbol{w}, \hat{\boldsymbol{u}}, \boldsymbol{v}) + b(\boldsymbol{v}, r) - (\boldsymbol{s}, \boldsymbol{v})& \\ = (\bar{\boldsymbol{f}}, \boldsymbol{v}) \quad \forall\, \boldsymbol{v} \in \boldsymbol{H}_0^1(\Omega)&\end{aligned} \tag{4.10}$$

and

$$b(\boldsymbol{w}, q) = (\bar{\phi}, q) \quad \forall\, q \in L_0^2(\Omega). \tag{4.11}$$

Theorem 4.8. *Let $(\hat{\boldsymbol{u}}, \hat{p}, \hat{\boldsymbol{g}}) \in \boldsymbol{H}_0^1(\Omega) \times L_0^2(\Omega) \times \boldsymbol{L}^2(\Omega)$ denote an optimal solution in the sense of (2.4). Then, there exists a nonzero Lagrange multiplier $(\boldsymbol{\xi}, \sigma) \in \boldsymbol{H}_0^1(\Omega) \times L_0^2(\Omega)$ satisfying the Euler equations*

$$-\mathscr{J}'(\hat{\boldsymbol{u}}, \hat{\boldsymbol{g}}) \cdot (\boldsymbol{w}, r, \boldsymbol{s}) + \langle M'(\hat{\boldsymbol{u}}, \hat{p}, \hat{\boldsymbol{g}}) \cdot (\boldsymbol{w}, r, \boldsymbol{s}), (\boldsymbol{\xi}, \sigma)\rangle = 0 \quad \forall\, (\boldsymbol{w}, r, \boldsymbol{s}) \in \boldsymbol{H}_0^1(\Omega) \times L_0^2(\Omega) \times \boldsymbol{L}^2(\Omega), \tag{4.12}$$

where $\langle \cdot, \cdot \rangle$ denotes the duality pairing between $\boldsymbol{H}_0^1(\Omega) \times L_0^2(\Omega)$ and $\boldsymbol{H}^{-1}(\Omega) \times L_0^2(\Omega)$.

Proof. The operator $M'(\hat{\boldsymbol{u}},\hat{p},\hat{\boldsymbol{g}})$ from X into Y is onto. To see this, we let $(\bar{\boldsymbol{f}},\bar{\phi}) \in Y$ be given and we first examine the problem of seeking an element $(\boldsymbol{w},r) \in \boldsymbol{H}_0^1(\Omega) \times L_0^2(\Omega)$ such that

$$[\nu_0 + \nu_1 a(\hat{\boldsymbol{u}},\hat{\boldsymbol{u}})]a(\boldsymbol{w},\boldsymbol{v}) + 2\nu_1 a(\boldsymbol{w},\hat{\boldsymbol{u}})a(\hat{\boldsymbol{u}},\boldsymbol{v}) + c(\hat{\boldsymbol{u}},\boldsymbol{w},\boldsymbol{v}) + c(\boldsymbol{w},\hat{\boldsymbol{u}},\boldsymbol{v}) + b(\boldsymbol{v},r) = (\bar{\boldsymbol{f}},\boldsymbol{v}) \quad \forall\, \boldsymbol{v} \in \boldsymbol{H}_0^1(\Omega) \tag{4.13}$$

and

$$b(\boldsymbol{w},q) = (\bar{\phi},q) \quad \forall\, q \in L_0^2(\Omega), \tag{4.14}$$

which can be rewritten as

$$\tilde{a}(\boldsymbol{w},\boldsymbol{v}) + b(\boldsymbol{v},r) = (\bar{\boldsymbol{f}},\boldsymbol{v}) \quad \forall\, \boldsymbol{v} \in \boldsymbol{H}_0^1(\Omega) \tag{4.15}$$

and

$$b(\boldsymbol{w},q) = (\bar{\phi},q) \quad \forall\, q \in L_0^2(\Omega), \tag{4.16}$$

where $\tilde{a}(\boldsymbol{w},\boldsymbol{v}) = [\nu_0 + \nu_1 a(\hat{\boldsymbol{u}},\hat{\boldsymbol{u}})]a(\boldsymbol{w},\boldsymbol{v}) + 2\nu_1 a(\boldsymbol{w},\hat{\boldsymbol{u}})a(\hat{\boldsymbol{u}},\boldsymbol{v}) + c(\hat{\boldsymbol{u}},\boldsymbol{w},\boldsymbol{v})$. For each $\boldsymbol{w} \in \boldsymbol{H}_0^1(\Omega)$ we have $c(\hat{\boldsymbol{u}},\boldsymbol{w},\boldsymbol{w}) = 0$ such that

$$\tilde{a}(\boldsymbol{w},\boldsymbol{w}) = [\nu_0 + \nu_1 a(\hat{\boldsymbol{u}},\hat{\boldsymbol{u}})]a(\boldsymbol{w},\boldsymbol{w}) + 2\nu_1 a(\boldsymbol{w},\hat{\boldsymbol{u}})a(\hat{\boldsymbol{u}},\boldsymbol{w}) \geqslant \nu_0 \|\boldsymbol{w}\|_1^2.$$

Hence, using well-known results for proving the existence of solutions for the Stokes equations [12] we readily conclude that there exists a $(\boldsymbol{w},r) \in \boldsymbol{H}_0^1(\Omega) \times L_0^2(\Omega)$ such that (4.15) and (4.16), or (4.13) and (4.14) hold. Next, we set $\boldsymbol{s} = \boldsymbol{w} \cdot \operatorname{grad} \hat{\boldsymbol{u}}$ so that $c(\boldsymbol{w},\hat{\boldsymbol{u}},\boldsymbol{v}) = (\boldsymbol{s},\boldsymbol{v})$ for all $\boldsymbol{v} \in \boldsymbol{H}^1(\Omega)$. Since $\boldsymbol{w} \in \boldsymbol{H}^1(\Omega)$ and $\hat{\boldsymbol{u}} \in \boldsymbol{H}^2(\Omega)$ (see the remarks at the end of Section 4.1), we deduce from imbedding theorems that $\boldsymbol{s} \in \boldsymbol{L}^2(\Omega)$. By adding $c(\boldsymbol{w},\hat{\boldsymbol{u}},\boldsymbol{v}) - (\boldsymbol{s},\boldsymbol{v}) = 0$ to (4.13) we see that $(\boldsymbol{w},r,\boldsymbol{s}) \in X$ satisfies (4.10) and (4.11), i.e., we have proved that $M'(\hat{\boldsymbol{u}},\hat{p},\hat{\boldsymbol{g}})$ is onto.

Hence, by Proposition 4.7, we deduce that there exists a $(\boldsymbol{\xi},\sigma) \in \boldsymbol{H}_0^1(\Omega) \times L_0^2(\Omega)$ such that (4.12) hold. □

Using (4.10) and (4.11), we may rewrite (4.12) in the form

$$-(\hat{\boldsymbol{u}} - \boldsymbol{u}_0,\boldsymbol{w}) - (\hat{\boldsymbol{g}},\boldsymbol{s}) + [\nu_0 + \nu_1 a(\hat{\boldsymbol{u}},\hat{\boldsymbol{u}})]a(\boldsymbol{w},\boldsymbol{\xi}) + 2\nu_1 a(\boldsymbol{\xi},\hat{\boldsymbol{u}})a(\hat{\boldsymbol{u}},\boldsymbol{w}) + c(\boldsymbol{w},\hat{\boldsymbol{u}},\boldsymbol{\xi}) + c(\hat{\boldsymbol{u}},\boldsymbol{w},\boldsymbol{\xi}) + b(\boldsymbol{\xi},r) - (\boldsymbol{s},\boldsymbol{\xi}) + b(\boldsymbol{w},\sigma) = 0 \quad \forall\,(\boldsymbol{w},r,\boldsymbol{s}) \in \boldsymbol{H}_0^1(\Omega) \times L_0^2(\Omega) \times \boldsymbol{L}^2(\Omega).$$

Upon separation of the above equation, one obtains (2.8)–(2.10). Since the optimal solution $(\hat{\boldsymbol{u}},\hat{p},\hat{\boldsymbol{g}})$ also satisfies the constraint (2.6) and (2.7), we see that necessary conditions for an optimum are that the system (2.6)–(2.10), i.e., the optimality system, is satisfied.

4.3. Regularity of solutions of the optimality system

We now examine the regularity of solutions of the optimality system (2.7)–(2.9) and (2.11), or equivalently, (2.12)–(2.18).

Theorem 4.9. *Suppose the given data satisfies $\boldsymbol{f} \in \boldsymbol{L}^2(\Omega)$ and $\boldsymbol{u}_0 \in \boldsymbol{L}^2(\Omega)$. Suppose that Ω is of class $C^{1,1}$. Then, if $(\boldsymbol{u},p,\boldsymbol{\xi},\sigma) \in \boldsymbol{H}_0^1(\Omega) \times L_0^2(\Omega) \times \boldsymbol{H}_0^1(\Omega) \times L_0^2(\Omega)$ denotes a solution of the optimality system (2.7)–(2.9) and (2.11), or equivalently, (2.12)–(2.18), we have that $(\boldsymbol{u},p,\boldsymbol{\xi},\sigma) \in \boldsymbol{H}^2(\Omega) \times H^1(\Omega) \times \boldsymbol{H}^2(\Omega) \times H^1(\Omega)$.*

Proof. Since $\xi \in H^1(\Omega)$, we have that the right-hand side of (2.12) belongs to $\boldsymbol{L}^2(\Omega)$. Then, the additional regularity of $\boldsymbol{u}$ and p follows from Lemma 4.5 (or regularity results for the Navier–Stokes equations by noting that the term $[\nu_0 + \nu_1 a(\boldsymbol{u},\boldsymbol{u})]$ is a constant).

Now, since $\boldsymbol{u} \in \boldsymbol{H}^2(\Omega)$ and $\xi \in \boldsymbol{H}^1(\Omega)$, we have that $\Delta\boldsymbol{u}$, $\xi\cdot(\operatorname{grad}\boldsymbol{u})^{\mathrm{T}}$, and $\boldsymbol{u}\cdot\operatorname{grad}\xi$ belong to $\boldsymbol{L}^2(\Omega)$. Also, $\boldsymbol{u}-\boldsymbol{u}_0$ belongs to $\boldsymbol{L}^2(\Omega)$. Thus, by rewriting (2.15) in the form

$$-[\nu_0 + \nu_1 a(\boldsymbol{u},\boldsymbol{u})]\Delta\xi + \operatorname{grad}\sigma = 2\nu_1 a(\boldsymbol{u},\xi)\Delta\boldsymbol{u} - \xi\cdot(\operatorname{grad}\boldsymbol{u})^{\mathrm{T}} + \boldsymbol{u}\cdot\operatorname{grad}\xi + (\boldsymbol{u}-\boldsymbol{u}_0) \quad \text{in } \Omega, \tag{4.17}$$

we have a right-hand side that belongs to $\boldsymbol{L}^2(\Omega)$. Then, since Ω is of class $C^{1,1}$ and $a(\boldsymbol{u},\boldsymbol{u})$ and $a(\boldsymbol{u},\xi)$ are constants, well-known regularity results for the Stokes problem applied to (2.16), (2.17) and (4.17) yield that $\xi \in \boldsymbol{H}^2(\Omega)$ and $\sigma \in H^1(\Omega)$. □

Remark. The above result also holds for convex regions of $\mathbb{R}^2$. In general, we may show that if $\boldsymbol{f} \in \boldsymbol{H}^m(\Omega)$, $\boldsymbol{u}_0 \in \boldsymbol{H}^m(\Omega)$, and Ω is sufficiently smooth, then $(\boldsymbol{u},p,\xi,\sigma) \in \boldsymbol{H}^{m+2}(\Omega)\times H^{m+1}(\Omega)\times \boldsymbol{H}^{m+2}(\Omega)\times H^{m+1}(\Omega)$. In particular, if $\boldsymbol{f}$ and $\boldsymbol{u}_0$ are both of class $C^\infty(\bar{\Omega})$, and Ω is of class C^∞, then $\boldsymbol{u},p,\xi$, and σ are all $C^\infty(\bar{\Omega})$ functions as well.

4.4. Existence of finite element approximations

We now turn to the question of the existence of solutions of the discrete system (3.5)–(3.8). Note that here and in Section 4.5, we drop the (^) notation in denoting optimal solutions.

The discrete *inf-sup condition* (3.3) implies that the subspace

$$\boldsymbol{V}^h = \{\boldsymbol{v}^h \in \boldsymbol{X}^h \colon b(\boldsymbol{v}^h, q^h) = 0 \ \forall\, q^h \in S^h\} \tag{4.18}$$

of the finite element space $\boldsymbol{X}^h$ is nonempty and, along with the approximation property (3.1), that, as $h \to 0$,

$$\inf_{\boldsymbol{v}^h \in \boldsymbol{V}^h} \|\boldsymbol{v} - \boldsymbol{v}^h\|_1 \to 0 \quad \forall\, \boldsymbol{v} \in \boldsymbol{V};$$

see [6]. Thus, $(\boldsymbol{u}^h, p^h, \xi^h, \sigma^h) \in \boldsymbol{X}^h \times S^h \times \boldsymbol{X}^h \times S^h$ is a solution of (3.5)–(3.8) if and only if $(\boldsymbol{u}^h, \xi^h) \in \boldsymbol{V}^h \times \boldsymbol{V}^h$ is a solution of

$$[\nu_0 + \nu_1 a(\boldsymbol{u}^h,\boldsymbol{u}^h)]a(\boldsymbol{u}^h,\boldsymbol{v}^h) + c(\boldsymbol{u}^h,\boldsymbol{u}^h,\boldsymbol{v}^h) = (\boldsymbol{f} - \xi^h, \boldsymbol{v}^h) \quad \forall\, \boldsymbol{v}^h \in \boldsymbol{V}^h \tag{4.19}$$

and

$$\begin{aligned} &[\nu_0 + \nu_1 a(\boldsymbol{u}^h,\boldsymbol{u}^h)]a(\xi^h,\omega^h) + 2\nu_1 a(\boldsymbol{u}^h,\xi^h)\mathrm{a}(\boldsymbol{u}^h,\omega^h) + c(\omega^h,\boldsymbol{u}^h,\xi^h) + c(\boldsymbol{u}^h,\omega^h,\xi^h) \\ &\quad = (\boldsymbol{u}^h - \boldsymbol{u}_0, \omega^h) \quad \forall\, \omega^h \in \boldsymbol{V}^h. \end{aligned} \tag{4.20}$$

We define

$$\begin{aligned} a_1((\boldsymbol{u},\xi);(z,\eta),(\boldsymbol{v},\omega)) = {} & [\nu_0 + \nu_1 a(\boldsymbol{u},\boldsymbol{u})]a(z,\boldsymbol{v}) + c(\boldsymbol{u},z,\boldsymbol{v}) + (\eta,\boldsymbol{v}) \\ & + [\nu_0 + \nu_1 a(\boldsymbol{u},\boldsymbol{u})]a(\eta,\omega) + 2\nu_1 a(\boldsymbol{u},\eta)a(\boldsymbol{u},\omega) \\ & + c(\omega,\boldsymbol{u},\eta) + c(\boldsymbol{u},\omega,\eta) - (z,\omega) \quad \forall\,(z,\eta),(\boldsymbol{v},\omega) \in \boldsymbol{V}\times\boldsymbol{V} \end{aligned}$$

and

$$F((\boldsymbol{v},\omega)) = (\boldsymbol{f},\boldsymbol{v}) - (\boldsymbol{u}_0,\omega) \quad \forall (\boldsymbol{v},\omega) \in V \times V.$$

Clearly, $(\boldsymbol{u},\xi) \in V \times V$ is a solution of (2.19) and (2.20) if and only if

$$a_1((\boldsymbol{u},\xi);(\boldsymbol{u},\xi),(\boldsymbol{v},\omega)) = F((\boldsymbol{v},\omega)) \quad \forall (\boldsymbol{v},\omega) \in V \times V; \tag{4.21}$$

$(\boldsymbol{u}^h,\xi^h) \in V^h \times V^h$ is a solution of (4.19) and (4.20) if and only if

$$a_1((\boldsymbol{u}^h,\xi^h);(\boldsymbol{u}^h,\xi^h),(\boldsymbol{v}^h,\omega^h)) = F((\boldsymbol{v}^h,\omega^h)) \quad \forall (\boldsymbol{v}^h,\omega^h) \in V^h \times V^h. \tag{4.22}$$

We now prove the existence of solutions to Eqs. (4.19) and (4.20), or equivalently, (4.22), under suitable assumptions and derive a uniform bound for these solutions $(\boldsymbol{u}^h,\xi^h)$.

Theorem 4.10. *Assume* $4\nu_0\nu_1 > C_0^2$ *where* C_0 *is the constant in* (3.4). *Then, the discrete system* (4.22) *has at least one solution* $(\boldsymbol{u}^h,\xi^h) \in V^h \times V^h$. *Furthermore, there exists a positive* $\mu > 0$ *independent of* h *such that every solution* $(\boldsymbol{u}^h,\xi^h)$ *of* (4.22) *satisfies*

$$\|\boldsymbol{u}^h\|_1^2 + \|\xi^h\|_1^2 \leqslant \mu.$$

Proof. The assumption $4\nu_0\nu_1 > C_0^2$ implies that

$$\gamma = \gamma(\nu_0,\nu_1) \equiv \min_{-\infty < x < \infty} (\nu_0 + \nu_1 x^2 - C_0 x) > 0.$$

Then, for each $(\boldsymbol{u}^h,\xi^h) \in V^h \times V^h$ we have that

$$\begin{aligned}
&a_1((\boldsymbol{u}^h,\xi^h);(\boldsymbol{u}^h,\xi^h),(\boldsymbol{u}^h,\xi^h))\\
&\quad = [\nu_0 + \nu_1 a(\boldsymbol{u}^h,\boldsymbol{u}^h)]a(\boldsymbol{u}^h,\boldsymbol{u}^h) + c(\boldsymbol{u}^h,\boldsymbol{u}^h,\boldsymbol{u}^h) + (\xi^h,\boldsymbol{u}^h) + [\nu_0 + \nu_1 a(\boldsymbol{u}^h,\boldsymbol{u}^h)]a(\xi^h,\xi^h)\\
&\qquad + 2\nu_1 a(\boldsymbol{u}^h,\xi^h)a(\boldsymbol{u}^h,\xi^h) + c(\xi^h,\boldsymbol{u}^h,\xi^h) + c(\boldsymbol{u}^h,\xi^h,\xi^h) - (\boldsymbol{u}^h,\xi^h)\\
&\quad = [\nu_0 + \nu_1 \|\operatorname{grad}\boldsymbol{u}^h\|_0^2]\,\|\operatorname{grad}\boldsymbol{u}^h\|_0^2 + [\nu_0 + \nu_1 \|\operatorname{grad}\boldsymbol{u}^h\|_0^2\,\|\operatorname{grad}\xi^h\|_0^2]\\
&\qquad + 2\nu_1 a(\boldsymbol{u}^h,\xi^h)a(\boldsymbol{u}^h,\xi^h) + c(\xi^h,\boldsymbol{u}^h,\xi^h)\\
&\quad \geqslant [\nu_0 + \nu_1 \|\operatorname{grad}\boldsymbol{u}^h\|_0^2]\,\|\operatorname{grad}\boldsymbol{u}^h\|_0^2 + [\nu_0 + \nu_1 \|\operatorname{grad}\boldsymbol{u}^h\|_0^2 - C_0\|\operatorname{grad}\boldsymbol{u}^h\|_0]\,\|\operatorname{grad}\xi^h\|_0^2\\
&\quad \geqslant \gamma\|\operatorname{grad}\boldsymbol{u}^h\|_0^2 + \nu_0\|\operatorname{grad}\xi^h\|_0^2.
\end{aligned}$$

Note that on the finite dimensional space V^h, the mapping $\boldsymbol{u}^h \mapsto a_1(\boldsymbol{u}^h;\boldsymbol{u}^h,\boldsymbol{v}^h)$ (for each fixed $\boldsymbol{v}^h$) is automatically sequentially weakly continuous on V^h. Thus, by Lemma 4.2 we deduce that there exists at least one $(\boldsymbol{u}^h,\xi^h)$ satisfying (4.19) and (4.20).

Now, let $(\boldsymbol{u}^h,\xi^h)$ be an arbitrary solution of (4.19) and (4.20). It can be verified that

$$\gamma\|\operatorname{grad}\boldsymbol{u}^h\|_0^2 + \nu_0\|\operatorname{grad}\xi^h\|_0^2 \leqslant C,$$

where

$$C = \frac{1}{\alpha^2\gamma}\|\boldsymbol{f}\|_{-1}^2 + \frac{1}{\alpha^2\nu_0}\|\boldsymbol{u}_0\|_{-1}^2$$

and $\alpha > 0$ is the constant in (1.9). By setting

$$\mu = \frac{\|\boldsymbol{f}\|_{-1}^2 + \|\boldsymbol{u}_0\|_{-1}^2}{\alpha^3 \min\{\nu_0, \gamma\}}$$

the desired estimate follows. □

4.5. Convergence of finite element approximations

With the uniform bound on the finite element approximations, we may pass to the limit to show the convergence to the solution of the continuous optimality system. In doing so, we will need the following result.

Lemma 4.11. *For each $\boldsymbol{v} \in \boldsymbol{H}_0^1(\Omega)$, the mapping*

$$(\boldsymbol{u}, \xi) \in \boldsymbol{H}_0^1(\Omega) \times \boldsymbol{H}_0^1(\Omega) \mapsto a(\boldsymbol{u}, \boldsymbol{u})a(\xi, \boldsymbol{v}) + 2a(\boldsymbol{u}, \xi)a(\boldsymbol{u}, \boldsymbol{v})$$

is weakly sequentially continuous in $\boldsymbol{H}_0^1(\Omega)$.

Proof. Using the identity

$$\begin{aligned} &a(\boldsymbol{u} + \xi, \boldsymbol{u} + \xi)a(\boldsymbol{u} + \xi, \boldsymbol{v}) - a(\boldsymbol{u} - \xi, \boldsymbol{u} - \xi)a(\boldsymbol{a} - \xi, \boldsymbol{v}) \\ &\quad = 4a(\boldsymbol{u}, \xi)a(\boldsymbol{u}, \boldsymbol{v}) + 2[a(\boldsymbol{u}, \boldsymbol{u}) + a(\xi, \xi)]a(\xi, \boldsymbol{v}), \end{aligned}$$

we obtain

$$\begin{aligned} &a(\boldsymbol{u}, \boldsymbol{u})a(\xi, \boldsymbol{v}) + 2a(\boldsymbol{u}, \xi)a(\boldsymbol{u}, \boldsymbol{v}) \\ &\quad = \tfrac{1}{2}\{a(\boldsymbol{u} + \xi, \boldsymbol{u} + \xi)a(\boldsymbol{u} + \xi, \boldsymbol{v}) - a(\boldsymbol{u} - \xi, \boldsymbol{u} - \xi)a(\boldsymbol{u} - \xi, \boldsymbol{v}) - 2a(\xi, \xi)a(\xi, \boldsymbol{v})\}. \end{aligned}$$

Thus, the weak sequential continuity of the mapping

$$(\boldsymbol{u}, \xi) \mapsto a(\boldsymbol{u}, \boldsymbol{u})a(\xi, \boldsymbol{v}) + 2a(\boldsymbol{u}, \xi)a(\boldsymbol{u}, \boldsymbol{v})$$

follows from the weak sequential continuity of the mapping (for each fixed $\boldsymbol{v}$)

$$\boldsymbol{u} \mapsto a(\boldsymbol{u}, \boldsymbol{u})a(\boldsymbol{u}, \boldsymbol{v}),$$

while the latter is a consequence of Lemma 4.1. □

Remark. It is interesting to note that the above result implies the weak continuity (for each fixed $\boldsymbol{v}$) of the mapping $(\boldsymbol{u}, \xi) \mapsto a(\boldsymbol{u}, \boldsymbol{u})a(\xi, \boldsymbol{v}) + 2a(\boldsymbol{u}, \xi)a(\boldsymbol{u}, \boldsymbol{v})$, even though one may not have the weak continuity of mappings $(\boldsymbol{u}, \xi) \mapsto a(\boldsymbol{u}, \boldsymbol{u})a(\xi, \boldsymbol{v})$ and $(\boldsymbol{u}, \xi) \mapsto a(\boldsymbol{u}, \xi)a(\boldsymbol{u}, \boldsymbol{v})$ separately.

Proposition 4.12. *Assume the hypotheses of Theorem* 4.10. *Then, there exists a subsequence $(\boldsymbol{u}^{h_n}, \xi^{h_n})$ that converges weakly in $\boldsymbol{H}_0^1(\Omega)$ to a solution of the optimality system* (4.21) *as $h_n \to 0$ $(n \to \infty)$.*

Proof. By the uniform bounds in Theorem 4.10, we can extract a subsequence $(\boldsymbol{u}^{h_n}, \xi^{h_n})$ which converges weakly to $(\boldsymbol{u}, \xi)$ in $\boldsymbol{H}_0^1(\Omega) \times \boldsymbol{H}_0^1(\Omega)$. It is easy to verify that F is a bounded linear functional on $\boldsymbol{H}_0^1(\Omega)\times\boldsymbol{H}_0^1(\Omega)$. Combining Lemma 4.2 with the weak continuity of the mappings

$$\boldsymbol{u}\mapsto a(\boldsymbol{u},\boldsymbol{u})a(\boldsymbol{u},\boldsymbol{v}),\qquad \boldsymbol{u}\mapsto c(\boldsymbol{u},\boldsymbol{u},\boldsymbol{v}),$$

$$(\boldsymbol{u},\boldsymbol{\xi})\mapsto c(\boldsymbol{u},\boldsymbol{\xi},\boldsymbol{v}),\qquad (\boldsymbol{u},\boldsymbol{\xi})\mapsto c(\boldsymbol{\xi},\boldsymbol{u},\boldsymbol{v}),$$

for each fixed $\boldsymbol{v}\in\boldsymbol{H}_0^1(\Omega)$, we apply standard procedures to pass to the limit in (4.22) and use the approximation property (3.1) and *inf-sup* condition (3.3), which imply that, as $h\to 0$,

$$\inf_{\boldsymbol{v}^h\in\boldsymbol{V}^h}\|\boldsymbol{v}-\boldsymbol{v}^h\|_1\to 0\quad\forall\,\boldsymbol{v}\in\boldsymbol{V},$$

to obtain

$$a_1((\boldsymbol{u},\boldsymbol{\xi});(\boldsymbol{u},\boldsymbol{\xi}),(\boldsymbol{v},\boldsymbol{\omega}))=F((\boldsymbol{v},\boldsymbol{\omega}))\quad\forall\,(\boldsymbol{v},\boldsymbol{\omega})\in\boldsymbol{V}\times\boldsymbol{V},$$

i.e., $(\boldsymbol{u},\boldsymbol{\xi})$ is a weak solution of the optimality system (4.21). □

Theorem 4.13. *Assume the hypothesis of Theorem* 4.10. *Then, the subsequence* $(\boldsymbol{u}^{h_n},\boldsymbol{\xi}^{h_n})$ *in Theorem* 4.12, *which converges weakly in* $\boldsymbol{H}_0^1(\Omega)\times\boldsymbol{H}_0^1(\Omega)$ *to a solution* $(\boldsymbol{u},\boldsymbol{\xi})$ *of the optimality system* (4.22) *as* $h_n\to 0$, *converges strongly to* $(\boldsymbol{u},\boldsymbol{\xi})$ *in* $\boldsymbol{H}_0^1(\Omega)\times\boldsymbol{H}_0^1(\Omega)$.

Proof. Imbedding theorems imply the following strong convergence: $\boldsymbol{u}^{h_n}\to\boldsymbol{u}$ in $\boldsymbol{L}^4(\Omega)$, $\boldsymbol{\xi}^{h_n}\to\boldsymbol{\xi}$ in $\boldsymbol{L}^4(\Omega)$. Hence $(\boldsymbol{\xi}_{h_n},\boldsymbol{u}_{h_n})\to(\boldsymbol{u},\boldsymbol{\xi})$ as $h_n\to 0$. Now using (2.19) and (4.19) we obtain

$$a_0(\boldsymbol{u};\boldsymbol{u},\boldsymbol{u})=(\boldsymbol{f}-\boldsymbol{\xi},\boldsymbol{u})=\lim_{h_n\to 0}(\boldsymbol{f}-\boldsymbol{\xi}^{h_n},\boldsymbol{u}^{h_n})=\lim_{h_n\to 0}a_0(\boldsymbol{u}^{h_n},\boldsymbol{u}^{h_n},\boldsymbol{u}^{h_n}),$$

i.e.,

$$\lim_{h_n\to 0}(\nu_0\|\operatorname{grad}\boldsymbol{u}^{h_n}\|_0^2+\nu_1\|\operatorname{grad}\boldsymbol{u}^{h_n}\|_0^4)=\nu_0\|\operatorname{grad}\boldsymbol{u}\|_0^2+\nu_1\|\operatorname{grad}\boldsymbol{u}\|_0^4.$$

We set $y_n=\nu_0\|\operatorname{grad}\boldsymbol{u}^{h_n}\|_0^2+\nu_1\|\operatorname{grad}\boldsymbol{u}^{h_n}\|_0^4$ and $y=\nu_0\|\operatorname{grad}\boldsymbol{u}\|_0^2+\nu_1\|\operatorname{grad}\boldsymbol{u}\|_0^4$. Then $\lim_{n\to\infty}y_n=y$. It can be easily verified that

$$\|\operatorname{grad}\boldsymbol{u}^{h_n}\|_0^2=\frac{-\nu_0+\sqrt{\nu_0^2+4\nu_1y_n}}{2\nu_1}\quad\text{and}\quad\|\operatorname{grad}\boldsymbol{u}\|_0^2=\frac{-\nu_0+\sqrt{\nu_0^2+4\nu_1y}}{2\nu_1}.$$

Thus we readily deduce that

$$\lim_{n\to\infty}\|\operatorname{grad}\boldsymbol{u}^{h_n}\|_0^2=\|\operatorname{grad}\boldsymbol{u}\|_0^2.$$

This implies that $\|\boldsymbol{u}^{h_n}\|_1$ converges to $\|\boldsymbol{u}\|_1$. Therefore, the sequence $\{\boldsymbol{u}^{h_n}\}$ converges strongly to $\boldsymbol{u}$ in $\boldsymbol{H}_0^1(\Omega)$.

Now, using (4.20) and (2.20), we have that

$$\lim_{n\to\infty}[\nu_0+\nu_1a(\boldsymbol{u}^{h_n},\boldsymbol{u}^{h_n})]a(\boldsymbol{\xi}^{h_n},\boldsymbol{\xi}^{h_n})+2\nu_1[a(\boldsymbol{u}^{h_n},\boldsymbol{\xi}^{h_n})]^2+c(\boldsymbol{\xi}^{h_n},\boldsymbol{u}^{h_n},\boldsymbol{\xi}^{h_n})+c(\boldsymbol{u}^{h_n},\boldsymbol{\xi}^{h_n},\boldsymbol{\xi}^{h_n})$$

$$=\lim_{n\to\infty}(\boldsymbol{u}^{h_n}-\boldsymbol{u}_0,\boldsymbol{\xi}^{h_n})=(\boldsymbol{u}-\boldsymbol{u}_0,\boldsymbol{\xi})$$

$$=[\nu_0+\nu_1a(\boldsymbol{u},\boldsymbol{u})]a(\boldsymbol{\xi},\boldsymbol{\xi})+2\nu_1[a(\boldsymbol{u},\boldsymbol{\xi})]^2+c(\boldsymbol{\xi},\boldsymbol{u},\boldsymbol{\xi})+c(\boldsymbol{u},\boldsymbol{\xi},\boldsymbol{\xi}).$$

Using the strong convergence of $\boldsymbol{u}^{h_n} \to \boldsymbol{u}$ in $\boldsymbol{H}_0^1(\Omega)$ and $\boldsymbol{u}^{h_n} \to \boldsymbol{u}$ in $\boldsymbol{L}^4(\Omega)$, as well as the estimate $|c(\boldsymbol{v},\boldsymbol{w},\boldsymbol{z})| \leqslant \|\boldsymbol{v}\|_{L^4(\Omega)} \|\boldsymbol{w}\|_1 \|\boldsymbol{z}\|_{L^4(\Omega)}$ for all $\boldsymbol{w},\boldsymbol{v},\boldsymbol{z} \in \boldsymbol{H}_0^1(\Omega)$, we deduce that

$$\lim_{n\to\infty} (2\nu_1[a(\boldsymbol{u}^{h_n},\xi^{h_n})]^2 + c(\xi^{h_n},\boldsymbol{u}^{h_n},\xi^{h_n}) + c(\boldsymbol{u}^{h_n},\xi^{h_n},\xi^{h_n}))$$

$$= 2\nu_1[a(\boldsymbol{u},\xi)]^2 + c(\xi,\boldsymbol{u},\xi) + c(\boldsymbol{u},\xi,\xi).$$

Thus

$$\lim_{n\to\infty} [\nu_0 + \nu_1 a(\boldsymbol{u}^{h_n},\boldsymbol{u}^{h_n})]\, a(\xi^{h_n},\xi^{h_n}) = [\nu_0 + \nu_1 a(\boldsymbol{u},\boldsymbol{u})]\, a(\xi,\xi),$$

which, together with the fact that $\lim_{n\to\infty} a(\boldsymbol{u}^{h_n},\boldsymbol{u}^{h_n}) = a(\boldsymbol{u},\boldsymbol{u})$, implies that

$$\lim_{n\to\infty} [\nu_0 + \nu_1 a(\boldsymbol{u},\boldsymbol{u})]\, a(\xi^{h_n},\xi^{h_n}) = [\nu_0 + \nu_1 a(\boldsymbol{u},\boldsymbol{u})]\, a(\xi,\xi),$$

or, equivalently,

$$a(\xi,\xi) = \lim_{n\to\infty} a(\xi^{h_n},\xi^{h_n}).$$

Hence, ξ^{h_n} converges strongly in $\boldsymbol{H}_0^1(\Omega)$ to ξ as $n \to \infty$. □

Recall that the optimal control is given by $\boldsymbol{g} = -\xi$. Thus, we have the following result for the approximation of the control.

Corollary 4.14. *Define $\boldsymbol{g}^h = -\xi^h$. Then, there exists a subsequence $\boldsymbol{g}^{h_n}$ that converges weakly in $\boldsymbol{H}^1(\Omega)$ and strongly in $\boldsymbol{L}^2(\Omega)$ to an optimal control $\boldsymbol{g}$.*

Remark. In general, the solution of the optimality system is not unique. However, under suitable assumptions on $\nu_0, \nu_1, \|\boldsymbol{f}\|_{-1}$ and $\|\boldsymbol{u}_0\|_0$, one can show the solution is unique. In this case, the convergence of the subsequence $\{(\boldsymbol{u}^{h_n},\xi^{h_n})\}$ actually implies the convergence of the entire sequence $\{(\boldsymbol{u}^h,\xi^h)\}$ as $h \to 0$.

For the sake of completeness, we present a uniqueness result as follows.

Proposition 4.15. *Assume $(\nu_0 - 2C_0\mu - 2r_0\mu^2) > 0$, where $r_0 = (1 + \sqrt{5})/2$. Then, the solution to* (4.19) *and* (4.20) *is unique.*

Proof. Assume $(\boldsymbol{u}_1,\xi_1)$ and $(\boldsymbol{u}_2,\xi_2)$ are two solutions of (4.19) and (4.20). Theorem 4.10 provides the following bounds:

$$\|\boldsymbol{u}_1\|_1 + \|\xi_1\|_1 \leqslant \mu \quad \text{and} \quad \|\boldsymbol{u}_2\|_1 + \|\xi_2\|_1 \leqslant \mu,$$

where $\mu = [\|\boldsymbol{f}\|_{-1}^2 + \|\boldsymbol{u}_0\|_{-1}^2]/[\alpha^3 \min(\nu_0,\gamma)]$.

We denote $e_u = u_1 - u_2$ and $e_\xi = \xi_1 - \xi_2$. From Eqs. (4.19) and (4.20) we deduce that

$$\begin{aligned}0 &= \nu_0 a(e_u, e_u) + \nu_0 a(e_\xi, e_\xi) + \nu_1 [a(u_1, u_1) a(u_1, e_u) - a(u_2, u_2) a(u_2, e_u)] \\ &\quad + [c(u_1, u_1, e_u) - c(u_2, u_2, e_u)] \\ &\quad + \nu_1 [a(u_1, u_1) a(\xi_1, e_\xi) - a(u_2, u_2) a(\xi_2, e_\xi) + 2a(u_1, \xi_1) a(u_1, e_\xi) - 2\nu_1 a(u_2, \xi_2) a(u_2, e_\xi)] \\ &\quad + [c(u_1, e_\xi, \xi_1) + c(e_\xi, u_1, \xi_1) - c(u_2, e_\xi, \xi_2) - c(e_\xi, u_2, \xi_2)].\end{aligned}$$

Note that

$$\begin{aligned}& a(u_1, u_1) a(u_1, e_u) - a(u_2, u_2) a(u_2, e_u) \geqslant 0, \\ & |c(u_1, u_1, e_u) - c(u_2, u_2, e_u)| = |c(u_1, e_u, e_u) + c(e_u, u_2, e_u)| = |c(e_u, u_2, e_u)| \leqslant C_0 \mu \|e_u\|_1^2, \\ & |c(u_1, e_\xi, \xi_1) + c(e_\xi, u_1, \xi_1) - c(u_2, e_\xi, \xi_2) - c(e_\xi, u_2, \xi_2)| \\ & \quad = |c(e_u, e_\xi, \xi_1) + c(e_\xi, e_u, \xi_1) + c(e_\xi, u_2, e_\xi)| \leqslant C_0 \mu \|e_u\|_1^2 + 2C_0 \mu \|e_\xi\|_1^2,\end{aligned}$$

and

$$\begin{aligned}& |a(u_1, u_1) a(\xi_1, e_\xi) - a(u_2, u_2) a(\xi_2, u_\xi) + 2a(u_1, \xi_1) a(u_1, e_\xi) - 2a(u_2, \xi_2) a(u_2, e_\xi)| \\ & \quad = |a(u_1, e_u) a(\xi_2, e_\xi) + a(e_u, u_2) a(\xi_2, e_\xi) + a(u_1, u_2) a(e_\xi, e_\xi) + 2a(e_\xi, \xi_1) a(u_1, e_\xi) \\ & \qquad + 2a(u_2, e_\xi) a(u_1, e_\xi) + 2a(u_2, \xi_2) a(e_u, e_\xi)| \leqslant 4\mu^2 \|e_u\|_1 \|e_\xi\|_1 + 2\mu^2 \|e_\xi\|_1^2 \\ & \quad \leqslant 2\mu^2 r_0 \|e_u\|_1 + 2\mu^2 r_0 \|e_\xi\|_1^2.\end{aligned}$$

Hence, we have that

$$0 \geqslant (\nu_0 - 2C_0\mu - 2r_0\mu^2) \|e_u\|_1^2 + (\nu_0 - 2C_0\mu - 2r_0\mu^2) \|e_\xi\|_1^2$$

so that under the hypothesis $(\nu_0 - 2C_0\mu - 2r_0\mu^2) > 0$,

$$\|e_u\|_1^2 = \|e_\xi\|_1^2 = 0,$$

i.e., the solution to (4.19)–(4.20) is unique. □

Remark. Similar results may be obtained for finite element solutions of (3.5)–(3.8).

5. The drag functional

We now consider a variation on the problem considered in Sections 2–4. A substantial portion of the analyses and results of those sections that apply to the minimization of the functional (1.4) with distributed controls will also apply to the variation considered in this section. Therefore, here we will merely point out the differences.

Consider flow control problems wherein the functional (1.5) involving the viscous drag dissipation is to be minimized, subject, of course, to the Ladyzhenskaya equations (1.1)–(1.3) (or (2.1)) as constraints. Using the notation of Section 1, we rewrite (1.5) as

$$\mathscr{K}(u, g) = \tfrac{1}{4}[2\nu_0 + \nu_1 a(u, u)] a(u, u) - (f + g, u) + \frac{1}{2\delta} \|g\|_0^2. \qquad (5.1)$$

The parameter δ will be chosen below. The admissibility set is now defined by

$$\mathscr{V}_{\mathrm{ad}} = \{(\boldsymbol{u}, \boldsymbol{g}) \in V \times \boldsymbol{L}^2(\Omega) \colon (2.1) \text{ is satisfied}\}.$$

The optimization problem at hand is to minimize (5.1) over $\mathscr{V}_{\mathrm{ad}}$. The existence of optimal solutions may be shown exactly as in Theorem 4.6. Also, Theorem 4.8 on the existence of Lagrange multipliers is easily amended to apply to the context of this section. An optimality system, which may be derived using the method of Lagrange multipliers, is given by (2.7), (2.9), and, instead of (2.11) and (2.8),

$$[\nu_0 + \nu_1 a(\boldsymbol{u},\boldsymbol{u})]a(\boldsymbol{u},\boldsymbol{v}) + c(\boldsymbol{u},\boldsymbol{u},\boldsymbol{v}) + b(\boldsymbol{v},p) = (\boldsymbol{f} - \delta\boldsymbol{\xi} + \delta\boldsymbol{u},\boldsymbol{v}) \quad \forall\, \boldsymbol{v} \in \boldsymbol{H}_0^1(\Omega) \tag{5.2}$$

and

$$\begin{aligned}&[\nu_0 + \nu_1 a(\boldsymbol{u},\boldsymbol{u})]a(\boldsymbol{\xi},\boldsymbol{\omega}) + 2\nu_1 a(\boldsymbol{\xi},\boldsymbol{u})a(\boldsymbol{u},\boldsymbol{\omega}) + c(\boldsymbol{\omega},\boldsymbol{u},\boldsymbol{\xi}) + c(\boldsymbol{u},\boldsymbol{\omega},\boldsymbol{\xi}) + b(\boldsymbol{\omega},\sigma)\\ &\quad = [\nu_0 + \nu_1 a(\boldsymbol{u},\boldsymbol{u})]a(\boldsymbol{u},\boldsymbol{\omega}) - (\boldsymbol{f} + \boldsymbol{g},\boldsymbol{\omega}) \quad \forall\, \boldsymbol{\omega} \in \boldsymbol{H}_0^1(\Omega),\end{aligned} \tag{5.3}$$

respectively. In (5.2) we have used the optimality condition which, instead of (2.10), is now given by

$$(\boldsymbol{g},\boldsymbol{s}) = \delta(\boldsymbol{s},\boldsymbol{u} - \boldsymbol{\xi}) \quad \forall\, \boldsymbol{s} \in \boldsymbol{L}^2(\Omega).$$

We may substitute (5.2) on the right-hand side of (5.3) to yield

$$\begin{aligned}&[\nu_0 + \nu_1 a(\boldsymbol{u},\boldsymbol{u})]a(\boldsymbol{\xi},\boldsymbol{\omega}) + 2\nu_1 a(\boldsymbol{\xi},\boldsymbol{u})a(\boldsymbol{u},\boldsymbol{\omega}) + c(\boldsymbol{\omega},\boldsymbol{u},\boldsymbol{\xi}) + c(\boldsymbol{u},\boldsymbol{\omega},\boldsymbol{\xi}) + b(\boldsymbol{\omega},\check{\sigma})\\ &\quad = -c(\boldsymbol{u},\boldsymbol{u},\boldsymbol{\omega}) \quad \forall\, \boldsymbol{\omega} \in \boldsymbol{H}_0^1(\Omega)\end{aligned} \tag{5.4}$$

where $\check{\sigma} = \sigma + p$. In the sequel we will dispense with the $(\check{\cdot})$ notation. Thus, the optimality system for the problem of minimizing (5.1) over $\mathscr{V}_{\mathrm{ad}}$ is given, in a form not explicitly involving the controls, by (2.7), (2.9), (5.2), and (5.4).

By integration by parts one easily finds that the optimality system is a weak formulation of the following system of partial differential equations and boundary conditions:

$$\begin{aligned}&-[\nu_0 + \nu_1 a(\boldsymbol{u},\boldsymbol{u})]\Delta\boldsymbol{u} + \boldsymbol{u}\cdot\operatorname{grad}\boldsymbol{u} + \operatorname{grad} p = \boldsymbol{f} + \delta(\boldsymbol{u} - \boldsymbol{\xi}) \quad \text{in } \Omega,\\ &\operatorname{div}\boldsymbol{u} = 0 \quad \text{in } \Omega, \qquad \boldsymbol{u} = 0 \quad \text{on } \Gamma,\\ &-[\nu_0 + \nu_1 a(\boldsymbol{u},\boldsymbol{u})]\Delta\boldsymbol{\xi} - 2\nu_1 a(\boldsymbol{u},\boldsymbol{\xi})\Delta\boldsymbol{u} + \boldsymbol{\xi}\cdot(\operatorname{grad}\boldsymbol{u})^{\mathrm{T}} - \boldsymbol{u}\cdot\operatorname{grad}\boldsymbol{\xi} + \operatorname{grad}\sigma = -(\boldsymbol{u}\cdot\operatorname{grad})\boldsymbol{u}\\ &\quad \text{in } \Omega,\\ &\operatorname{div}\boldsymbol{\xi} = 0 \quad \text{in } \Omega, \qquad \boldsymbol{\xi} = \boldsymbol{0} \quad \text{on } \Gamma.\end{aligned} \tag{5.5}$$

By choosing δ sufficiently small, the existence and regularity results for this optimality system may be derived in the same manner as that employed in Section 4. Finite element approximations are defined exactly as in Section 3, which are given by (3.6), (3.8) and, instead of (3.5) and (3.7).

$$[\nu_0 + \nu_1 a(\boldsymbol{u}^h,\boldsymbol{u}^h)]a(\boldsymbol{u}^h,\boldsymbol{v}^h) + c(\boldsymbol{u}^h,\boldsymbol{u}^h,\boldsymbol{v}^h) + b(\boldsymbol{v}^h,{}^hp) = (\boldsymbol{f} - \delta\boldsymbol{\xi}^h + \delta\boldsymbol{u}^h,\boldsymbol{v}^h) \quad \forall\, \boldsymbol{v}^h \in \boldsymbol{X}^h \tag{5.6}$$

and

$$\begin{aligned}&[\nu_0 + \nu_1 a(\boldsymbol{u}^h,\boldsymbol{u}^h)]a(\boldsymbol{\xi}^h,\boldsymbol{\omega}^h) + 2\nu_1 a(\boldsymbol{\xi}^h,\boldsymbol{u}^h)a(\boldsymbol{u}^h,\boldsymbol{\omega}^h) + c(\boldsymbol{\omega}^h,\boldsymbol{u}^h,\boldsymbol{\xi}^h) + c(\boldsymbol{u}^h,\boldsymbol{\omega}^h,\boldsymbol{\xi}^h) + b(\boldsymbol{\omega}^h,\sigma^h)\\ &\quad = -(\boldsymbol{u}^h,\boldsymbol{u}^h,\boldsymbol{\omega}^h) \quad \forall\, \boldsymbol{\omega}^h \in \boldsymbol{X}^h,\end{aligned} \tag{5.7}$$

respectively.

In extending the result of Theorem 4.10 we note that we now have two more terms to estimate, i.e.,

$$|(\xi^h, \boldsymbol{u}^h)| \leqslant \|\xi^h\|_1 \|\boldsymbol{u}^h\|_1 \leqslant \frac{1}{\delta^2} \|\operatorname{grad} \xi^h\|_0 \|\operatorname{grad} \boldsymbol{u}^h\|_0 \leqslant \frac{1}{2\delta^2} \|\operatorname{grad} \xi^h\|_0^2 + \frac{1}{2\delta^2} \|\operatorname{grad} \boldsymbol{u}^h\|_0^2$$

and

$$|c(\boldsymbol{u}^h, \boldsymbol{u}^h, \omega^h)| \leqslant C_0 \|\operatorname{grad} \boldsymbol{u}^h\|_0^2 \|\operatorname{grad} \xi^h\|_0 \leqslant \frac{C_0}{2} \|\operatorname{grad} \boldsymbol{u}^h\|_0^2 + \frac{C_0}{2} \|\operatorname{grad} \xi^h\|_0^2.$$

Thus, similar to Theorem 4.10, we have the following result.

Theorem 5.1. *For ν_0 sufficiently large and δ sufficiently small, the discrete system* (3.6), (3.8), (5.6), *and* (5.7) *has at least one solution $(\boldsymbol{u}^h, p^h, \xi^h, \sigma^h) \in V^h \times S^h \times V^h \times S^h$. Furthermore, there exists a positive $\mu > 0$ independent of h such that every solution $(\boldsymbol{u}^h, \xi^h)$ of* (3.6), (3.8), (5.6), *and* (5.7) *satisfies*

$$\|\boldsymbol{u}^h\|_1^2 + \|\xi^h\|_1^2 \leqslant \mu.$$

Again, at least for ν_0 sufficiently large, the results of Propositions 4.12 and 4.15, Theorem 4.13, and Corollary 4.14 can be shown to be applicable to the present case.

The main effect of making the substitution of (5.2) into the right-hand side of (5.3) is to replace $[\nu_0 + \nu_1 a(\boldsymbol{u}, \boldsymbol{u})] a(\boldsymbol{u}, \omega) - (\boldsymbol{f} + \boldsymbol{g}, \omega)$ in favor of the term $-c(\boldsymbol{u}, \boldsymbol{u}, \omega)$, i.e., to have, on the right-hand side of (5.5), $(\boldsymbol{u} \cdot \operatorname{grad} \boldsymbol{u})$ instead of $-[\nu_0 + \nu_1 a(\boldsymbol{u}, \boldsymbol{u})] \Delta \boldsymbol{u} - \boldsymbol{f} - \boldsymbol{g}$. This replacement is necessary in order to validate the analyses of Sections 4.4 and 4.5 for the present case.

References

[1] R. Adams, *Sobolev Spaces* (Academic Press, New York, 1975).

[2] I. Babuška and A. Aziz, Survey lectures on the mathematical foundations of the finite element method, in: A. Aziz, Ed., *The Mathematical Foundations of the Finite Element Method with Applications to Partial Differential Equations* (Academic Press, New York, 1972) 3–359.

[3] H. Bellout, F. Bloom and J. Necas, Young measure valued solutions for non-newtonian incompressible fluids, to appear.

[4] P. Ciarlet, *The Finite Element Method for Elliptic Problems* (North-Holland, Amsterdam, 1978).

[5] Q. Du and M. Gunzburger, Analysis of a Ladyzhenskaya model for incompressible viscous flow, *J. Math. Anal. Appl.* **155** (1991) 21–45.

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

[7] M. Gunzburger, *Finite Element Methods for Incompressible Viscous Flows: A Guide to Theory, Practice and Algorithms* (Academic Press, Boston, MA, 1989).

[8] S. Kaniel, On the initial value problem for an incompressible fluid with nonlinear viscosity, *J. Math. Mech.* **19** (1992) 681–706.

[9] O. Ladyzhenskaya, New equations for the description of motion of viscous incompressible fluids and solvability in the large of boundary value problems for them, *Proc. Steklov Math. Inst.* **102** (1967) 95–108.

[10] W. Layton, A nonlinear, sub-gridscale model for incompressible viscous flow problems, to appear.

[11] J. Serrin, Mathematical principles of classical fluid mechanics, in: S. Flügge and C. Truesdell, Eds., *Handbüch der Physik* VIII/1 (Springer, Berlin, 1959) 125–263.

[12] R. Temam, *Navier–Stokes Equations* (North-Holland, Amsterdam, 1979).

[13] R. Temam, *Navier–Stokes Equations and Nonlinear Functional Analysis* (SIAM, Philadelphia, PA, 1983).

[14] E. Zeidler, *Nonlinear Functional Analysis and Its Applications*, Vol. III (Springer, New York, 1985).