ANALYSIS AND FINITE ELEMENT APPROXIMATION OF OPTIMAL CONTROL PROBLEMS FOR THE STATIONARY NAVIER-STOKES EQUATIONS WITH DISTRIBUTED AND NEUMANN CONTROLS

M. D. GUNZBURGER, L. HOU, AND T. P. SVOBODNY

ABSTRACT. We examine certain analytic and numerical aspects of optimal control problems for the stationary Navier-Stokes equations. The controls considered may be of either the distributed or Neumann type; the functionals minimized are either the viscous dissipation or the L^4-distance of candidate flows to some desired flow. 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 the approximation, by finite element methods, of solutions of the optimality system and derive optimal error estimates.

1. INTRODUCTION

The optimization problem we study is to seek a state pair $(\mathbf{u}, p)$, i.e., velocity and pressure fields, and controls $\mathbf{g}_b$ and $\mathbf{g}_d$ such that a functional of $\mathbf{u}$, $\mathbf{g}_b$, and $\mathbf{g}_d$ is minimized subject to the constraint that the Navier-Stokes equations are satisfied. In other words, the state and controls are required to satisfy

$$(1.1) \quad -\nu \operatorname{div}((\operatorname{grad}\mathbf{u}) + (\operatorname{grad}\mathbf{u})^T) + \mathbf{u} \cdot \operatorname{grad}\mathbf{u} + \operatorname{grad} p = \mathbf{f} + \mathbf{g}_d \quad \text{in } \Omega,$$

$$(1.2) \quad \operatorname{div}\mathbf{u} = 0 \quad \text{in } \Omega,$$

and

$$(1.3) \quad -p\mathbf{n} + \nu(\operatorname{grad}\mathbf{u} + \operatorname{grad}\mathbf{u}^T) \cdot \mathbf{n} + \nu\alpha\mathbf{u} = \mathbf{h} + \mathbf{g}_b \quad \text{on } \Gamma,$$

where Ω denotes a bounded domain in $\mathbb{R}^d$, $d = 2$ or 3, with a boundary Γ, and ν is the kinematic viscosity. In (1.1)-(1.3), we have absorbed the constant density into the pressure and the body forces. If the variables in (1.1)-(1.3) are nondimensionalized, then ν is simply the inverse of the Reynolds number Re.

When finite element approximations are considered, we will assume that Ω is a convex polyhedral domain; otherwise, we will assume that Γ is either convex or is of class $C^{1,1}$. In (1.1)–(1.3), $\mathbf{u}$ and p denote the velocity and pressure fields, respectively, $\mathbf{f}$ a given body force, $\mathbf{h}$ a given stress type force, $\alpha \geq 0$ a given coefficient, $\mathbf{g}_d$ a distributed control, and $\mathbf{g}_b$ a boundary control.

A few words about the boundary condition (1.3) are in order. This boundary condition relates the stress vector and the velocity on the boundary. If $\alpha = 0$, then (1.3) is simply the stress vector on the boundary; with $\alpha > 0$, (1.3) can be also used in a variety of applications, e.g., compliant surfaces. Also, see [13, 16] for a discussion of the applicability of the boundary condition (1.3) with $\alpha > 0$. In any case, we do not specify the velocity field on the boundary. Velocity boundary conditions are sufficiently different, from both the analytic and algorithmic points of view, to warrant separate treatment (see [8, 9]). In practical situations it is likely that (1.3) is specified on only part of the boundary, with the velocity specified on other parts. Or, on some boundary segment, some component(s) of (1.3) may even be specified along with the complementary component(s) of the velocity. By combining the results of this paper with those of [8, 9], all of these alternative situations can, in principle, be handled. However, for some combinations of velocity and stress boundary conditions, some care must be exercised in defining finite element approximations (see [16]). In any case, the exposition is greatly simplified if we stick to the boundary condition (1.3), and treat the case of velocity boundary conditions separately.

The two functionals that we consider are given by

$$(1.4) \quad \mathscr{J}(\mathbf{u}, \mathbf{g}_b, \mathbf{g}_d) = \frac{1}{4} \sum_{j=1}^{d} \int_{\Omega} (u_j - u_{0j})^4 \, d\Omega + \frac{1}{2} \int_{\Gamma} |\mathbf{g}_b|^2 \, d\Gamma + \frac{1}{2} \int_{\Omega} |\mathbf{g}_d|^2 \, d\Omega$$

and

$$(1.5) \quad \begin{aligned} \mathscr{K}(\mathbf{u}, \mathbf{g}_b, \mathbf{g}_d) = & \frac{\nu}{2} \int_{\Omega} |(\text{grad}\,\mathbf{u}) + (\text{grad}\,\mathbf{u})^T|^2 \, d\Omega - \int_{\Omega} \mathbf{f} \cdot \mathbf{u} \, d\Omega \\ & + \frac{\nu}{2} \int_{\Gamma} \alpha \mathbf{u} \cdot \mathbf{u} \, d\Gamma + \frac{1}{2} \int_{\Gamma} |\mathbf{g}_b|^2 \, d\Gamma + \frac{1}{2} \int_{\Omega} |\mathbf{g}_d|^2 \, d\Omega. \end{aligned}$$

The first of these effectively measures the difference between the velocity field $\mathbf{u}$ and a prescribed field $\mathbf{u}_0$, while the second measures the drag due to viscosity. The inclusion in (1.5) of the boundary integral involving the coefficient α accounts for the force exerted on the fluid due to the absorption term $\nu \alpha \mathbf{u}$ appearing in (1.3). For a discussion of the relation between (1.5) and the viscous drag, see [13]. The appearance of the controls $\mathbf{g}_b$ and $\mathbf{g}_d$ in (1.4) and (1.5) is necessary, since we will not impose any a priori constraints on the size of these controls. Problems for which the controls are constrained to belong to closed, convex, bounded sets of the underlying control spaces, including cases in which the control may be omitted from the functional to be minimized, are treated in [8].

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 §2, we

give a precise statement of the optimization problem for the functional (1.4) and prove that an optimal solution exists. In §3, we prove the existence of Lagrange multipliers and then use the method of Lagrange multipliers to derive an optimality system. In §4, we consider finite element approximations and derive error estimates. In §5, we consider all of these issues in connection with four variations of the problem treated in §§2–4. These variations are: (1) optimizing the drag functional (1.5) instead of (1.4); (2) the case of having a distributed control acting in concert with specified boundary conditions; (3) cases wherein distributed controls act on only part of the flow domain; and (4) cases wherein Neumann controls act on only part of the boundary. We note that some of the material of §§2 and 3.1 may be found, in the context of distributed controls with homogeneous velocity boundary conditions, in [12]. Also, a preliminary announcement of some of the results of this paper may be found in [10].

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

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

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

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

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

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

Thus, the inner products in $L^2(\Omega)$ and $\mathbf{L}^2(\Omega)$ are both denoted by $(\cdot, \cdot)$ and those in $L^2(\Gamma)$ and $\mathbf{L}^2(\Gamma)$ by $(\cdot, \cdot)_\Gamma$. If X denotes a Banach space, X^* will denote its dual. Also, since in our context one of $\mathbf{L}^2(\Omega)$ or $\mathbf{L}^2(\Gamma)$ will play the role of the 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 two bilinear forms

$$a(\mathbf{u}, \mathbf{v}) = \frac{1}{2} \int_\Omega ((\text{grad}\,\mathbf{u}) + (\text{grad}\,\mathbf{u})^T) : ((\text{grad}\,\mathbf{v}) + (\text{grad}\,\mathbf{v})^T) \, d\Omega \quad \forall \mathbf{u}, \mathbf{v} \in \mathbf{H}^1(\Omega)$$

and

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

and the trilinear form

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

These forms are continuous in the sense that there exists a constant $C > 0$ such that

(1.6) $$|a(\mathbf{u}, \mathbf{v})| \le C\|\mathbf{u}\|_1 \|\mathbf{v}\|_1 \quad \forall \mathbf{u}, \mathbf{v} \in \mathbf{H}^1(\Omega),$$

(1.7) $$|b(\mathbf{v}, q)| \le C\|\mathbf{v}\|_1 \|q\|_0 \quad \forall \mathbf{v} \in \mathbf{H}^1(\Omega) \text{ and } q \in L^2(\Omega),$$

and

(1.8) $$|c(\mathbf{u}, \mathbf{v}, \mathbf{w})| \le C\|\mathbf{u}\|_1 \|\mathbf{v}\|_1 \|\mathbf{w}\|_1 \quad \forall \mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathbf{H}^1(\Omega).$$

Moreover, we have that

(1.9) $$\sup_{0 \ne \mathbf{v} \in \mathbf{H}^1(\Omega)} \frac{b(\mathbf{v}, q)}{\|\mathbf{v}\|_1} \ge C\|q\|_0 \quad \forall q \in L^2(\Omega).$$

For details concerning these forms, one may consult [6, 7, or 14].

One more useful property of the bilinear form $a(\cdot, \cdot)$ is given as follows. Suppose $\Gamma_1 \subset \Gamma$, such that Γ_1 has positive measure. Then, if $\alpha > 0$ on Γ_1, we have that

(1.10) $$C^{-1}\|\mathbf{v}\|_1^2 \le a(\mathbf{v}, \mathbf{v}) + (\alpha \mathbf{v}, \mathbf{v})_{\Gamma_1} \le C\|\mathbf{v}\|_1^2 \quad \forall \mathbf{v} \in \mathbf{H}^1(\Omega).$$

Thus, $a(\mathbf{v}, \mathbf{v}) + (\alpha \mathbf{v}, \mathbf{v})_{\Gamma_1}$ defines an equivalent norm on $\mathbf{H}^1(\Omega)$. The right inequality of (1.10) follows easily from (1.6) and the properties of the inner product $(\cdot, \cdot)_{\Gamma_1}$; the left inequality follows from the Korn inequalities.

2. THE OPTIMIZATION PROBLEM AND THE EXISTENCE OF OPTIMAL SOLUTIONS

We begin by giving a precise statement of the first optimization problem we consider. Let $\mathbf{g}_b \in \mathbf{L}^2(\Gamma)$ and $\mathbf{g}_d \in \mathbf{L}^2(\Omega)$ denote the boundary and distributed controls, respectively, and let $\mathbf{u} \in \mathbf{H}^1(\Omega)$ and $p \in L^2(\Omega)$ denote the state, i.e., the velocity and pressure fields, respectively. The state and control variables are constrained to satisfy the Navier-Stokes equations in the weak form (see, e.g., [6, 7, or 14])

(2.1)
$$\nu a(\mathbf{u}, \mathbf{v}) + c(\mathbf{u}, \mathbf{u}, \mathbf{v}) + b(\mathbf{v}, p) + \nu(\alpha \mathbf{u}, \mathbf{v})_\Gamma$$
$$= (\mathbf{f} + \mathbf{g}_d, \mathbf{v}) + (\mathbf{h} + \mathbf{g}_b, \mathbf{v})_\Gamma \quad \forall \mathbf{v} \in \mathbf{H}^1(\Omega)$$

and

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

where $\mathbf{f} \in \mathbf{L}^2(\Omega)$, $\mathbf{h} \in \mathbf{L}^2(\Gamma)$, and smooth $\alpha \ge 0$ are given functions. If $\alpha \equiv 0$, then here and throughout we should work with velocity test and trial functions

belonging to $\mathbf{H}^1(\Omega)/\{\mathscr{R}\}$, where $\mathscr{R}$ denotes the space of rigid body motions. Although all our results hold for the case $\alpha \equiv 0$, for the sake of simplicity, we will assume that $\alpha > 0$.

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

$$(2.3) \qquad \mathscr{J}(\mathbf{u}, \mathbf{g}_b, \mathbf{g}_d) = \tfrac{1}{4}\|\mathbf{u} - \mathbf{u}_0\|_{\mathbf{L}^4(\Omega)}^4 + \tfrac{1}{2}\|\mathbf{g}_b\|_{0,\Gamma}^2 + \tfrac{1}{2}\|\mathbf{g}_d\|_0^2,$$

where $\mathbf{u}_0 \in \mathbf{L}^4(\Omega)$ is a given function. The *admissibility set* $\mathscr{U}_{\mathrm{ad}}$ is defined by

$$\mathscr{U}_{\mathrm{ad}} = \{(\mathbf{u}, \mathbf{g}_b, \mathbf{g}_d) \in \mathbf{H}^1(\Omega) \times \mathbf{L}^2(\Gamma) \times \mathbf{L}^2(\Omega) : \mathscr{J}(\mathbf{u}, \mathbf{g}_b, \mathbf{g}_d) < \infty,$$

$$(2.4) \qquad\qquad\qquad \text{and there exists a } p \in L^2(\Omega)$$

$$\qquad\qquad\qquad\qquad \text{such that (2.1) and (2.2) are satisfied}\}.$$

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

$$(2.5) \qquad \mathscr{J}(\hat{\mathbf{u}}, \hat{\mathbf{g}}_b, \hat{\mathbf{g}}_d) \leq \mathscr{J}(\mathbf{u}, \mathbf{g}_b, \mathbf{g}_d) \quad \forall (\mathbf{u}, \mathbf{g}_b, \mathbf{g}_d) \in \mathscr{U}_{\mathrm{ad}} \text{ satisfying}$$

$$\|\mathbf{u} - \hat{\mathbf{u}}\|_1 + \|\mathbf{g}_b - \hat{\mathbf{g}}_b\|_{0,\Gamma} + \|\mathbf{g}_d - \hat{\mathbf{g}}_d\|_0 \leq \varepsilon.$$

We first show that an optimal solution exists.

Theorem 2.1. *There exists a* $(\hat{\mathbf{u}}, \hat{\mathbf{g}}_b, \hat{\mathbf{g}}_d) \in \mathscr{U}_{\mathrm{ad}}$ *such that* (2.3) *is minimized.*

Proof. We first claim that $\mathscr{U}_{\mathrm{ad}}$ is not empty. Let $\tilde{\mathbf{g}}_d \equiv \mathbf{0}$ and then let $(\tilde{\mathbf{u}}, \tilde{p})$ be a solution of (1.1)–(1.2) with $\tilde{\mathbf{u}} = \mathbf{0}$ on Γ. Then, since $\mathbf{f} \in \mathbf{L}^2(\Omega)$, it is well known [6 or 14] that $(\tilde{\mathbf{u}}, \tilde{p}) \in \mathbf{H}^2(\Omega) \times H^1(\Omega)$ and $\|\tilde{\mathbf{u}}\|_2 + \|\tilde{p}\|_1 \leq C\|\mathbf{f}\|_0$. Now, let

$$\tilde{\mathbf{g}}_b = [-\tilde{p}\mathbf{n} + ((\mathrm{grad}\,\tilde{\mathbf{u}}) + (\mathrm{grad}\,\tilde{\mathbf{u}})^T) \cdot \mathbf{n} + \alpha\tilde{\mathbf{u}}]_\Gamma - \mathbf{h}.$$

Clearly, if $\mathbf{h} \in \mathbf{L}^2(\Gamma)$, then $\tilde{\mathbf{g}}_b \in \mathbf{L}^2(\Gamma)$ and

$$\|\tilde{\mathbf{g}}_b\|_{0,\Gamma} \leq C(\|\mathbf{h}\|_{0,\Gamma} + \|\tilde{p}|_\Gamma\|_{0,\Gamma} + \|\tilde{\mathbf{u}}|_\Gamma\|_{1,\Gamma}),$$

so that $\|\tilde{\mathbf{g}}_b\|_{0,\Gamma} \leq C(\|\mathbf{f}\|_0 + \|\mathbf{h}\|_{0,\Gamma})$. Also, one easily finds that $(\tilde{\mathbf{u}}, \tilde{p}, \mathbf{0}, \tilde{\mathbf{g}}_b)$ satisfies (2.1)–(2.2). Moreover, we have

$$\mathscr{J}(\tilde{\mathbf{u}}, \mathbf{0}, \tilde{\mathbf{g}}_b) \leq C(\|\tilde{\mathbf{u}}\|_{L^4(\Omega)}^4 + \|\mathbf{u}_0\|_{L^4(\Omega)}^4 + \|\tilde{\mathbf{g}}_b\|_{0,\Gamma}^2) < \infty.$$

Thus, $(\tilde{\mathbf{u}}, \tilde{p}, \mathbf{0}, \tilde{\mathbf{g}}_b) \in \mathscr{U}_{\mathrm{ad}}$.

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

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

Then, by (2.3) and (2.4), $(\mathbf{u}^{(n)}, \mathbf{g}_b^{(n)}, \mathbf{g}_d^{(n)})$ is uniformly bounded in $\mathbf{L}^4(\Omega) \times \mathbf{L}^2(\Gamma) \times \mathbf{L}^2(\Omega)$, and for some $p^{(n)} \in L^2(\Omega)$ we have that

$$(2.6) \qquad \nu a(\mathbf{u}^{(n)}, \mathbf{v}) + b(\mathbf{v}, p^{(n)}) + \nu(\alpha\mathbf{u}^{(n)}, \mathbf{v})_\Gamma$$

$$= (\mathbf{f} + \mathbf{g}_d^{(n)}, \mathbf{v}) + (\mathbf{h} + \mathbf{g}_b^{(n)}, \mathbf{v})_\Gamma - c(\mathbf{u}^{(n)}, \mathbf{u}^{(n)}, \mathbf{v}) \quad \forall \mathbf{v} \in \mathbf{H}^1(\Omega)$$

and

$$(2.7) \qquad\qquad b(\mathbf{u}^{(n)}, q) = 0 \quad \forall q \in L^2(\Omega).$$

Now set $\mathbf{v} = \mathbf{u}^{(n)}$ in (2.6). Note that

$$|c(\mathbf{u}^{(n)}, \mathbf{u}^{(n)}, \mathbf{u}^{(n)})| = \frac{1}{2} \int_\Omega \mathbf{u}^{(n)} \cdot ((\operatorname{grad} \mathbf{u}^{(n)}) + (\operatorname{grad} \mathbf{u}^{(n)})^T) \cdot \mathbf{u}^{(n)} \, d\Omega$$

$$(2.8) \qquad\qquad \leq \frac{C}{2} \|(\operatorname{grad} \mathbf{u}^{(n)}) + (\operatorname{grad} \mathbf{u}^{(n)})^T\|_0 \|\mathbf{u}^{(n)}\|^2_{\mathbf{L}^4(\Omega)}$$

$$\leq \frac{\nu}{4} \|(\operatorname{grad} \mathbf{u}^{(n)}) + (\operatorname{grad} \mathbf{u}^{(n)})^T\|^2_0 + \frac{C^2}{4\nu} \|\mathbf{u}^{(n)}\|^4_{\mathbf{L}^4(\Omega)},$$

so that using (2.6) and (2.7), and the facts that $\|\mathbf{u}^{(n)}\|_{\mathbf{L}^4(\Omega)} \leq C \|\mathbf{u}^{(n)}\|_1$, $|(\mathbf{g}_d, \mathbf{u}^{(n)})|$ $\leq \|\mathbf{g}_d^{(n)}\|_0 \|\mathbf{u}^{(n)}\|_0$, and $|(\mathbf{g}_b, \mathbf{u}^{(n)})_\Gamma| \leq \|\mathbf{g}_b^{(n)}\|_{0,\Gamma} \|\mathbf{u}^{(n)}\|_{0,\Gamma}$, we have that

$$\frac{\nu}{4} \|(\operatorname{grad} \mathbf{u}^{(n)}) + (\operatorname{grad} \mathbf{u}^{(n)})^T\|^2_0 + \nu(\alpha \mathbf{u}^{(n)}, \mathbf{u}^{(n)})_\Gamma$$

$$\leq C(\|\mathbf{u}^{(n)}\|^3_{\mathbf{L}^4(\Omega)} + \|\mathbf{g}_b^{(n)}\|_{0,\Gamma} + \|\mathbf{g}_d^{(n)}\|_0 + \|\mathbf{h}\|_{0,\Gamma} + \|\mathbf{f}\|_0) \|\mathbf{u}^{(n)}\|_1.$$

Then, using (1.10) and the fact that $(\mathbf{u}^{(n)}, \mathbf{g}_b^{(n)}, \mathbf{g}_d^{(n)})$ is uniformly bounded in $\mathbf{L}^4(\Omega) \times \mathbf{L}^2(\Gamma) \times \mathbf{L}^2(\Omega)$, we have that

$$(2.9) \qquad \|\mathbf{u}^{(n)}\|_1 \leq C(\|\mathbf{u}^{(n)}\|^3_{\mathbf{L}^4(\Omega)} + \|\mathbf{g}_b^{(n)}\|_{0,\Gamma} + \|\mathbf{g}_d^{(n)}\|_0 + \|\mathbf{h}\|_{0,\Gamma} + \|\mathbf{f}\|_0) \leq K$$

for some K independent of n. Using (1.9) and (2.6), as well as this last result, one may obtain a similar estimate for $\|p^{(n)}\|_0$. We may then extract subsequences such that

$$\mathbf{g}_b^{(n)} \rightharpoonup \hat{g}_b \quad \text{in } \mathbf{L}^2(\Gamma),$$

$$\mathbf{g}_d^{(n)} \rightharpoonup \hat{g}_d \quad \text{in } \mathbf{L}^2(\Omega),$$

$$\mathbf{u}^{(n)} \rightharpoonup \hat{\mathbf{u}} \quad \text{in } \mathbf{H}^1(\Omega),$$

$$p^{(n)} \rightharpoonup \hat{p} \quad \text{in } L^2(\Omega),$$

$$\mathbf{u}^{(n)} \rightarrow \hat{\mathbf{u}} \quad \text{in } \mathbf{L}^2(\Omega),$$

$$\mathbf{u}^{(n)}|_\Gamma \rightarrow \hat{\mathbf{u}}|_\Gamma \quad \text{in } \mathbf{L}^2(\Gamma)$$

for some $(\hat{\mathbf{u}}, \hat{g}_b, \hat{g}_d) \in \mathbf{H}^1(\Omega) \times \mathbf{L}^2(\Gamma) \times \mathbf{L}^2(\Omega)$ and $\hat{p} \in L^2(\Omega)$. The last two convergence results above follow from the compact imbeddings $\mathbf{H}^1(\Omega) \subset \mathbf{L}^2(\Omega)$ and $\mathbf{H}^{1/2}(\Gamma) \subset \mathbf{L}^2(\Gamma)$. We may then pass to the limit in (2.6)–(2.7) to determine that $(\hat{\mathbf{u}}, \hat{p}, \hat{g}_b, \hat{g}_d)$ satisfies (2.1)–(2.2). Indeed, the only troublesome term, when one passes to the limit, is the nonlinearity $c(\cdot, \cdot, \cdot)$. However, note that

$$c(\mathbf{u}^{(n)}, \mathbf{u}^{(n)}, \mathbf{v}) = \int_\Gamma (\mathbf{u}^{(n)} \cdot \mathbf{n}) \mathbf{u}^{(n)} \cdot \mathbf{v} \, d\Gamma = \int_\Omega \mathbf{u}^{(n)} \cdot \operatorname{grad} \mathbf{v} \cdot \mathbf{u}^{(n)} \, d\Omega \quad \forall \mathbf{v} \in C^\infty(\overline{\Omega}).$$

Then, since $\mathbf{u}^{(n)} \to \hat{\mathbf{u}}$ in $\mathbf{L}^2(\Omega)$ and $\mathbf{u}^{(n)}|_\Gamma \to \hat{\mathbf{u}}|_\Gamma$ in $\mathbf{L}^2(\Gamma)$, we have that

$$\lim_{n\to\infty} c(\mathbf{u}^{(n)}, \mathbf{u}^{(n)}, \mathbf{v}) = \int_\Gamma (\hat{\mathbf{u}} \cdot \mathbf{n})\hat{\mathbf{u}} \cdot \mathbf{v} \, d\Gamma - \int_\Omega \hat{\mathbf{u}} \cdot \operatorname{grad}\mathbf{v} \cdot \hat{\mathbf{u}} \, d\Omega$$

$$= c(\hat{\mathbf{u}}, \hat{\mathbf{u}}, \mathbf{v}) \quad \forall \mathbf{v} \in C^\infty(\overline{\Omega}).$$

Then, since $C^\infty(\overline{\Omega})$ is dense in $\mathbf{H}^1(\Omega)$, we also have that

$$\lim_{n\to\infty} c(\mathbf{u}^{(n)}, \mathbf{u}^{(n)}, \mathbf{v}) = c(\hat{\mathbf{u}}, \hat{\mathbf{u}}, \mathbf{v}) \quad \forall \mathbf{v} \in \mathbf{H}^1(\Omega).$$

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

$$\mathcal{J}(\hat{\mathbf{u}}, \hat{\mathbf{g}}_b, \hat{\mathbf{g}}_d) = \inf_{(\mathbf{u}, \mathbf{g}_b, \mathbf{g}_d) \in \mathcal{U}_{\text{ad}}} \mathcal{J}(\mathbf{u}, \mathbf{g}_b, \mathbf{g}_d). \quad \square$$

Remark 1. The use of the $\mathbf{L}^4(\Omega)$-norm of $(\mathbf{u} - \mathbf{u}_0)$ in the functional (1.5), or equivalently (2.3), can now be explained. Note that as a result of this choice of norm we have that $\{\mathbf{u}^{(n)}\}$ is a bounded sequence in $\mathbf{L}^4(\Omega)$, a fact that was used to derive (2.9), i.e., that $\{\mathbf{u}^{(n)}\}$ is actually bounded in $\mathbf{H}^1(\Omega)$. The crucial step is included in the inequalities of (2.8). If, for example, we had used, in the functional (1.4), the $\mathbf{L}^2(\Omega)$-norm of $(\mathbf{u} - \mathbf{u}_0)$ instead, we would only know that $\{\mathbf{u}^{(n)}\}$ is a bounded sequence in $\mathbf{L}^2(\Omega)$ and the estimates in (2.8) would not hold.

Remark 2. Because the optimal control $\hat{\mathbf{g}}_b \in \mathbf{L}^2(\Gamma)$ and $\hat{\mathbf{g}}_d \in \mathbf{L}^2(\Omega)$, we may deduce, using regularity results for the Navier-Stokes equations, that $\hat{\mathbf{u}} \in \mathbf{H}^{3/2}(\Omega)$.

3. THE EXISTENCE OF LAGRANGE MULTIPLIERS AND AN OPTIMALITY SYSTEM

3.1. Existence of Lagrange multipliers. We wish to use the method of Lagrange multipliers to turn the constrained optimization problem (2.5) into an unconstrained one. We first show that suitable Lagrange multipliers exist.

Let $B_1 = \mathbf{H}^1(\Omega) \times L^2(\Omega) \times \mathbf{L}^2(\Omega) \times \mathbf{L}^2(\Gamma)$ and $B_2 = (\mathbf{H}^1(\Omega))^* \times L^2(\Omega)$, and let the nonlinear mapping $M: B_1 \to B_2$ denote the (generalized) constraint equations, i.e., $M(\mathbf{u}, p, \mathbf{g}_d, \mathbf{g}_b) = (\mathbf{f}, \varphi)$ for $(\mathbf{u}, p, \mathbf{g}_d, \mathbf{g}_b) \in B_1$ and $(\mathbf{f}, \varphi) \in B_2$ if and only if

$$\nu a(\mathbf{u}, \mathbf{v}) + c(\mathbf{u}, \mathbf{u}, \mathbf{v}) + b(\mathbf{v}, p) + \nu(\alpha\mathbf{u}, \mathbf{v})_\Gamma - (\mathbf{g}_d, \mathbf{v}) - (\mathbf{g}_b, \mathbf{v})_\Gamma = (\mathbf{f}, \mathbf{v}) \quad \forall \mathbf{v} \in \mathbf{H}^1(\Omega)$$

and

$$b(\mathbf{u}, q) = (\varphi, q) \quad \forall q \in L^2(\Omega).$$

Theorem 3.1. *Let* $(\hat{\mathbf{u}}, \hat{p}, \hat{\mathbf{g}}_d, \hat{\mathbf{g}}_b) \in \mathbf{H}^1(\Omega) \times L^2(\Omega) \times \mathbf{L}^2(\Omega) \times \mathbf{L}^2(\Gamma)$ *denote an optimal solution in the sense of* (2.5). *Then there exists a nonzero Lagrange multiplier* $(\boldsymbol{\xi}, \sigma) \in \mathbf{H}^1(\Omega) \times L^2(\Omega)$ *satisfying the Euler equations*

$$\mathcal{J}'(\hat{\mathbf{u}}, \hat{\mathbf{g}}_d, \hat{\mathbf{g}}_b) \cdot (\mathbf{w}, r, \mathbf{s}_d, \mathbf{s}_b)$$

(3.1)
$$+ \langle M'(\hat{\mathbf{u}}, \hat{p}, \hat{\mathbf{g}}_d, \hat{\mathbf{g}}_b) \cdot (\mathbf{w}, r, \mathbf{s}_d, \mathbf{s}_b), (\boldsymbol{\xi}, \sigma) \rangle = 0$$

$$\forall (\mathbf{w}, r, \mathbf{s}_d, \mathbf{s}_b) \in \mathbf{H}^1(\Omega) \times L^2(\Omega) \times \mathbf{L}^2(\Omega) \times \mathbf{L}^2(\Gamma),$$

where $\langle \cdot , \cdot \rangle$ denotes the duality pairing between $\mathbf{H}^1(\Omega) \times L^2(\Omega)$ and $(\mathbf{H}^1(\Omega))^* \times L^2(\Omega)$.

Proof. The operator $M'(\hat{\mathbf{u}}, \hat{p}, \hat{\mathbf{g}}_d, \hat{\mathbf{g}}_b) \in \mathscr{L}(B_1 ; B_2)$ is defined as follows:

$$M'(\hat{\mathbf{u}}, \hat{p}, \hat{\mathbf{g}}_d, \hat{\mathbf{g}}_b) \cdot (\mathbf{w}, r, \mathbf{s}_d, \mathbf{s}_b) = (\bar{\mathbf{f}}, \overline{\varphi})$$

for $(\mathbf{w}, r, \mathbf{s}_d, \mathbf{s}_b) \in B_1$ and $(\bar{\mathbf{f}}, \overline{\varphi}) \in B_2$, if and only if

$$\begin{aligned}
(3.2) \quad & \nu a(\mathbf{w}, \mathbf{v}) + c(\mathbf{w}, \hat{\mathbf{u}}, \mathbf{v}) + c(\hat{\mathbf{u}}, \mathbf{w}, \mathbf{v}) + b(\mathbf{v}, r) \\
& + \nu(\alpha \mathbf{w}, \mathbf{v})_\Gamma - (\mathbf{s}_d, \mathbf{v}) - (\mathbf{s}_b, \mathbf{v})_\Gamma = (\bar{\mathbf{f}}, \mathbf{v}) \quad \forall \mathbf{v} \in \mathbf{H}^1(\Omega)
\end{aligned}$$

and

$$(3.3) \qquad b(\mathbf{w}, q) = (\overline{\varphi}, q) \quad \forall q \in L^2(\Omega).$$

The operator $M'(\hat{\mathbf{u}}, \hat{p}, \hat{\mathbf{g}}_d, \hat{\mathbf{g}}_b)$ from B_1 into B_2 is onto. To see this, first note that (1.6), (1.7), (1.9), and (1.10) imply that there exists $(\mathbf{w}, r) \in \mathbf{H}^1(\Omega) \times L^2(\Omega)$ such that

$$\begin{aligned}
(3.4) \quad & \nu a(\mathbf{w}, \mathbf{v}) + \nu(\alpha \mathbf{w}, \mathbf{v})_\Gamma + b(\mathbf{v}, r) + c(\hat{\mathbf{u}}, \mathbf{w}, \mathbf{v}) - \tfrac{1}{2}((\hat{\mathbf{u}} \cdot \mathbf{n})\mathbf{w}, \mathbf{v})_\Gamma \\
& = (\bar{\mathbf{f}}, \mathbf{v}) \quad \forall \mathbf{v} \in \mathbf{H}^1(\Omega)
\end{aligned}$$

and

$$(3.5) \qquad b(\mathbf{w}, q) = (\overline{\varphi}, q) \quad \forall q \in L^2(\Omega).$$

This follows using well-known techniques for proving the existence of solutions of the Navier-Stokes equations [14], once one notes that $c(\hat{\mathbf{u}}, \mathbf{w}, \mathbf{w}) - \tfrac{1}{2}((\hat{\mathbf{u}} \cdot \mathbf{n})\mathbf{w}, \mathbf{w})_\Gamma = 0$ for all $\mathbf{w} \in \mathbf{H}^1(\Omega)$. Then, there is no difficulty in finding $\mathbf{s}_b \in \mathbf{L}^2(\Gamma)$ and $\mathbf{s}_d \in \mathbf{L}^2(\Omega)$ such that

$$(3.6) \qquad (\mathbf{s}_d, \mathbf{v}) + (\mathbf{s}_b, \mathbf{v})_\Gamma = c(\mathbf{w}, \hat{\mathbf{u}}, \mathbf{v}) + \tfrac{1}{2}((\hat{\mathbf{u}} \cdot \mathbf{n})\mathbf{w}, \mathbf{v})_\Gamma \quad \forall \mathbf{v} \in \mathbf{H}^1(\Omega).$$

Indeed, we may set $\mathbf{s}_b = \tfrac{1}{2}(\hat{\mathbf{u}} \cdot \mathbf{n})\mathbf{w}$ and $\mathbf{s}_d = \mathbf{w} \cdot \operatorname{grad} \hat{\mathbf{u}}$, so that, since $\mathbf{w} \in \mathbf{H}^1(\Omega)$ and $\hat{\mathbf{u}} \in \mathbf{H}^{3/2}(\Omega)$ (see Remark 2 at the end of §2), we have that $\mathbf{s}_d \in \mathbf{L}^2(\Omega)$, $\mathbf{s}_b \in \mathbf{L}^2(\Gamma)$ and (3.6) is satisfied.

Combining (3.4)–(3.6) then yields (3.2)–(3.3), showing that, for any $(\bar{\mathbf{f}}, \overline{\varphi}) \in B_2$ there exists $(\mathbf{w}, r, \mathbf{s}_d, \mathbf{s}_b)$ in B_1 satisfying (3.2)–(3.3).

Now consider the nonlinear operator $N : B_1 \to \mathbb{R} \times B_2$ defined by

$$N(\mathbf{u}, p, \mathbf{g}_d, \mathbf{g}_b) = \begin{pmatrix} \mathscr{J}(\mathbf{u}, \mathbf{g}_d, \mathbf{g}_b) - \mathscr{J}(\hat{\mathbf{u}}, \hat{\mathbf{g}}_d, \hat{\mathbf{g}}_b) \\ M(\mathbf{u}, p, \mathbf{g}_d, \mathbf{g}_b) \end{pmatrix}.$$

The operator $N'(\hat{\mathbf{u}}, \hat{p}, \hat{\mathbf{g}}_d, \hat{\mathbf{g}}_b)$ from B_1 into $\mathbb{R} \times B_2$ may be defined as follows: $N'(\hat{\mathbf{u}}, \hat{p}, \hat{\mathbf{g}}_d, \hat{\mathbf{g}}_b) \cdot (\mathbf{w}, r, \mathbf{s}_d, \mathbf{s}_b) = (\tilde{\beta}, \tilde{\mathbf{f}}, \check{\varphi})$ for $(\mathbf{w}, r, \mathbf{s}_d, \mathbf{s}_b) \in B_1$ and $(\tilde{\beta}, \tilde{\mathbf{f}}, \check{\varphi}) \in \mathbb{R} \times B_2$, if and only if

$$(3.7) \qquad \sum_{j=1}^d ([(\hat{\mathbf{u}} - \mathbf{u}_0)_j]^3, (\mathbf{w})_j) + (\hat{\mathbf{g}}_d, \mathbf{s}_d) + (\hat{\mathbf{g}}_b, \mathbf{s}_b)_\Gamma = \tilde{\beta},$$

$$(3.8) \quad \begin{aligned} & \nu a(\mathbf{w}, \mathbf{v}) + c(\mathbf{w}, \hat{\mathbf{u}}, \mathbf{v}) + c(\hat{\mathbf{u}}, \mathbf{w}, \mathbf{v}) + b(\mathbf{v}, r) \\ & \quad + \nu(\alpha \mathbf{u}, \mathbf{v})_\Gamma - (\mathbf{s}_d, \mathbf{v}) - (\mathbf{s}_b, \mathbf{v})_\Gamma = (\tilde{\mathbf{f}}, \mathbf{v}) \quad \forall \mathbf{v} \in \mathbf{H}^1(\Omega) \end{aligned}$$

and

$$(3.9) \qquad b(\mathbf{w}, q) = (\tilde{\varphi}, q) \quad \forall q \in L^2(\Omega).$$

This operator has a closed range but is not onto. The fact that it has a closed range can be shown as follows. First note that $M'(\hat{\mathbf{u}}, \hat{p}, \hat{\mathbf{g}}_d, \hat{\mathbf{g}}_b)$ is onto B_2, and therefore has a closed range. Also, the continuity of the various bilinear and trilinear forms, i.e., (1.6)–(1.8), and of the inner products appearing in the definiton of $M'(\hat{\mathbf{u}}, \hat{p}, \hat{\mathbf{g}}_d, \hat{\mathbf{g}}_b)$ implies that this operator belongs to $\mathscr{L}(B_1, B_2)$, and therefore the kernel of $M'(\hat{\mathbf{u}}, \hat{p}, \hat{\mathbf{g}}_d, \hat{\mathbf{g}}_b)$ is a closed subspace. Now, $\mathscr{J}'(\hat{\mathbf{u}}, \hat{\mathbf{g}}_d, \hat{\mathbf{g}}_b)$ acting on the kernel of $M'(\hat{\mathbf{u}}, \hat{p}, \hat{\mathbf{g}}_d, \hat{\mathbf{g}}_b)$ is either identically zero or onto $\mathbb{R}$. This follows from the obvious result that whenever f is a linear functional on a Banach space X, then either $f \equiv 0$ or the range of f is $\mathbb{R}$. Thus, we have shown that $\mathscr{J}'(\hat{\mathbf{u}}, \hat{\mathbf{g}}_d, \hat{\mathbf{g}}_b)$ acting on the kernel of $M'(\hat{\mathbf{u}}, \hat{p}, \hat{\mathbf{g}}_d, \hat{\mathbf{g}}_b)$ has a closed range. Now, recall the following well-known result. Let X, Y, Z be Banach spaces and $A: X \to Y$ and $B: X \to Z$ linear continuous operators. Then, if the range of A is closed in Y and the subspace $B \ker(A)$ is closed in Z and further, if $Cx = (Ax, Bx)$, $C: X \to Y \times Z$, then the range of C is closed in $Y \times Z$. Thus, in our context, the operator $N'(\hat{\mathbf{u}}, \hat{p}, \hat{\mathbf{g}}_d, \hat{\mathbf{g}}_b)$ has a closed range in B_2.

The operator $N'(\hat{\mathbf{u}}, \hat{p}, \hat{\mathbf{g}}_d, \hat{\mathbf{g}}_b)$ is not onto because if it were, by the Implicit Function Theorem, we would have $(\tilde{\mathbf{u}}, \tilde{\mathbf{g}}_d, \tilde{\mathbf{g}}_b) \in \mathscr{U}_{\mathrm{ad}}$ such that $\|\hat{\mathbf{u}} - \tilde{\mathbf{u}}\|_{\mathbf{L}^4(\Omega)} + \|\hat{\mathbf{g}}_d - \tilde{\mathbf{g}}_d\|_0 + \|\hat{\mathbf{g}}_b - \tilde{\mathbf{g}}_b\|_{0,\Gamma} \leq \varepsilon$ and $\mathscr{J}(\tilde{\mathbf{u}}, \tilde{\mathbf{g}}_d, \tilde{\mathbf{g}}_b) < \mathscr{J}(\hat{\mathbf{u}}, \hat{\mathbf{g}}_d, \hat{\mathbf{g}}_b)$, contradicting the hypothesis that $(\hat{\mathbf{u}}, \hat{\mathbf{g}}_d, \hat{\mathbf{g}}_b)$ is an optimal solution. However, the Hahn-Banach Theorem implies that there exists a nonzero element of $(\mathbb{R} \times B_2)^* = \mathbb{R} \times \mathbf{H}^1(\Omega) \times L^2(\Omega)$ that annihilates the range of $N'(\hat{\mathbf{u}}, \hat{p}, \hat{\mathbf{g}}_d, \hat{\mathbf{g}}_b)$, i.e., there exists $(\beta, \boldsymbol{\xi}, \sigma) \in \mathbb{R} \times \mathbf{H}^1(\Omega) \times L^2(\Omega)$ such that

$$(3.10) \quad \begin{aligned} & \langle (\tilde{\beta}, \tilde{\mathbf{f}}, \tilde{\varphi}), (\beta, \boldsymbol{\xi}, \sigma) \rangle = 0 \\ & \quad \forall (\tilde{\beta}, \tilde{\mathbf{f}}, \tilde{\varphi}) \text{ in the range of } N'(\hat{\mathbf{u}}, \hat{p}, \hat{\mathbf{g}}_d, \hat{\mathbf{g}}_b), \end{aligned}$$

where $\langle \cdot, \cdot \rangle$ denotes the duality pairing between $\mathbb{R} \times B_2$ and its dual $(\mathbb{R} \times B_2)^*$. Note that $\beta \neq 0$, since otherwise we would have that $\langle (\tilde{\mathbf{f}}, \tilde{\varphi}), (\boldsymbol{\xi}, \sigma) \rangle = 0$ for all $(\tilde{\mathbf{f}}, \tilde{\varphi}) \in B_2$, so that $(\boldsymbol{\xi}, \sigma) \equiv 0$, contradicting the fact that $(\beta, \boldsymbol{\xi}, \sigma) \neq 0$. We may, without any loss of generality, set $\beta = -1$. Clearly, using the definition of the operator $N'(\hat{\mathbf{u}}, \hat{p}, \hat{\mathbf{g}}_d, \hat{\mathbf{g}}_b)$, (3.1) and (3.10) are equivalent. $\quad \square$

Using (3.7)–(3.9) and setting $\beta = -1$, we may rewrite (3.10) in the form

$$\begin{aligned} & -((\hat{\mathbf{u}} - \mathbf{u}_0)^3, \mathbf{w}) - (\hat{\mathbf{g}}_d, \mathbf{s}_d) - (\hat{\mathbf{g}}_b, \mathbf{s}_b)_\Gamma + \nu a(\mathbf{w}, \boldsymbol{\xi}) \\ & \quad + c(\mathbf{w}, \hat{\mathbf{u}}, \boldsymbol{\xi}) + c(\hat{\mathbf{u}}, \mathbf{w}, \boldsymbol{\xi}) + b(\boldsymbol{\xi}, r) + \nu(\alpha \mathbf{w}, \boldsymbol{\xi})_\Gamma - (\mathbf{s}_d, \boldsymbol{\xi}) - (\mathbf{s}_b, \boldsymbol{\xi})_\Gamma \\ & \quad + b(\mathbf{w}, \sigma) = 0 \quad \forall (\mathbf{w}, r, \mathbf{s}_d, \mathbf{s}_b) \in \mathbf{H}^1(\Omega) \times L^2(\Omega) \times \mathbf{L}^2(\Omega) \times \mathbf{L}^2(\Gamma), \end{aligned}$$

where $\mathbf{v}^r$, for any vector $\mathbf{v}$, denotes componentwise exponentiation, i.e., $\mathbf{v}^r = (v_1^r, \ldots, v_d^r)^T$. Upon separating the above equations, we have that

$$(3.11) \qquad \begin{aligned} \nu a(\mathbf{w}, \boldsymbol{\xi}) + c(\mathbf{w}, \hat{\mathbf{u}}, \boldsymbol{\xi}) + c(\hat{\mathbf{u}}, \mathbf{w}, \boldsymbol{\xi}) + \nu(\alpha\mathbf{w}, \boldsymbol{\xi})_\Gamma + b(\mathbf{w}, \sigma) \\ = ((\hat{\mathbf{u}} - \mathbf{u}_0)^3, \mathbf{w}) \quad \forall \mathbf{w} \in \mathbf{H}^1(\Omega), \end{aligned}$$

$$(3.12) \qquad b(\boldsymbol{\xi}, r) = 0 \quad \forall r \in L^2(\Omega),$$

$$(3.13) \qquad (\hat{\mathbf{g}}_d, \mathbf{s}_d) = -(\mathbf{s}_d, \boldsymbol{\xi}) \quad \forall \mathbf{s}_d \in \mathbf{L}^2(\Omega),$$

and

$$(3.14) \qquad (\hat{\mathbf{g}}_b, \mathbf{s}_b)_\Gamma = -(\mathbf{s}_b, \boldsymbol{\xi})_\Gamma \quad \forall \mathbf{s}_b \in \mathbf{L}^2(\Gamma).$$

Since the optimal solution $(\hat{\mathbf{u}}, \hat{p}, \hat{\mathbf{g}}_d, \hat{\mathbf{g}}_b)$ satisfies the constraint (2.1)–(2.2), we see necessary conditions for an optimum are that (2.1)–(2.2) and (3.11)–(3.14) are satisfied. This system of equations will be called the *optimality system*.

Remark. The existence of Lagrange multipliers could have been proven, and (different) optimality systems could have been derived, using weaker norms for the controls in the functional (1.4), or equivalently, (2.3). For example, instead of the $\mathbf{L}^2(\Omega)$-norm of $\mathbf{g}_d$, we could have used in (2.3) the $\mathbf{L}^{3/2}(\Omega)$- or $\mathbf{L}^{4/3}(\Omega)$-norms, or even the $\mathbf{L}^{6/5}(\Omega)$- or $(\mathbf{H}^1(\Omega))^*$-norms of $\mathbf{g}_d$. We did not use the first two norms since they complicate the relation (3.13) between the Lagrange multiplier $\boldsymbol{\xi}$ and the optimal control $\hat{\mathbf{g}}_d$, resulting in a more cumbersome numerical algorithm. We did not use the last two norms because they result in great difficulties in the considerations of §3.3, i.e., the derivation of regularity results.

3.2. The optimality system. We have just shown that we are justified in introducing Lagrange multipliers in order to turn the constrained optimization problem (2.5) into an unconstrained one. For the sake of clarity and unity, we repeat some of the steps carried out above in the following formal procedure.

We introduce the Lagrange multipliers $\boldsymbol{\xi} \in \mathbf{H}^1(\Omega)$ and $\sigma \in L^2(\Omega)$ and define the product space

$$V = \mathbf{H}^1(\Omega) \times L^2(\Omega) \times \mathbf{L}^2(\Gamma) \times \mathbf{L}^2(\Omega) \times \mathbf{H}^1(\Omega) \times L^2(\Omega)$$

and the Lagrangian

$$(3.15) \qquad \begin{aligned} \mathscr{M}(\mathbf{u}, p, \mathbf{g}_b, \mathbf{g}_d, \boldsymbol{\xi}, \sigma) \\ = \mathscr{J}(\mathbf{u}, \mathbf{g}_b, \mathbf{g}_d) - [\nu a(\mathbf{u}, \boldsymbol{\xi}) + c(\mathbf{u}, \mathbf{u}, \boldsymbol{\xi}) + b(\boldsymbol{\xi}, p) + b(\mathbf{u}, \sigma) \\ + \nu(\alpha\mathbf{u}, \boldsymbol{\xi})_\Gamma - (\mathbf{f} + \mathbf{g}_d, \boldsymbol{\xi}) - (\mathbf{h} + \mathbf{g}_b, \boldsymbol{\xi})_\Gamma] \\ \forall (\mathbf{u}, p, \mathbf{g}_b, \mathbf{g}_d, \boldsymbol{\xi}, \sigma) \in V. \end{aligned}$$

We now seek the minimum of $\mathscr{M}(\mathbf{u}, p, \mathbf{g}_b, \mathbf{g}_d, \boldsymbol{\xi}, \sigma)$ over V.

Formally using standard techniques of the calculus of variations, one may derive the Euler-Lagrange equations that correspond to rendering (3.15) stationary.

Variations in the Lagrange multipliers $\boldsymbol{\xi}$ and σ recover the constraints (2.1) and (2.2). Variations in the state variables $\mathbf{u}$ and p yield

$$(3.16) \quad \begin{aligned} \nu a(\boldsymbol{\omega}, \boldsymbol{\xi}) + c(\boldsymbol{\omega}, \mathbf{u}, \boldsymbol{\xi}) + c(\mathbf{u}, \boldsymbol{\omega}, \boldsymbol{\xi}) + b(\boldsymbol{\omega}, \sigma) + \nu(\alpha\boldsymbol{\omega}, \boldsymbol{\xi})_\Gamma \\ = ((\mathbf{u} - \mathbf{u}_0)^3, \boldsymbol{\omega}) \quad \forall \boldsymbol{\omega} \in \mathbf{H}^1(\Omega) \end{aligned}$$

and

$$(3.17) \qquad b(\boldsymbol{\xi}, \psi) = 0 \quad \forall \psi \in L^2(\Omega),$$

where, again, $(\mathbf{u} - \mathbf{u}_0)^3$ denotes a componentwise exponentiation. Finally, variations in the controls $\mathbf{g}_b$ and $\mathbf{g}_d$ yield that

$$(3.18) \qquad (\mathbf{g}_d + \boldsymbol{\xi}, \mathbf{w}) = 0 \quad \forall \mathbf{w} \in \mathbf{L}^2(\Omega)$$

and

$$(3.19) \qquad (\mathbf{g}_b + \boldsymbol{\xi}, \mathbf{r})_\Gamma = 0 \quad \forall \mathbf{r} \in \mathbf{L}^2(\Gamma).$$

Thus, the optimal solution necessarily satisfies the optimality system (2.1)–(2.2) and (3.16)–(3.19). Note that (3.18)–(3.19) enable us to eliminate the controls $\mathbf{g}_b$ and $\mathbf{g}_d$ from (2.1), resulting in

$$(3.20) \quad \begin{aligned} \nu a(\mathbf{u}, \mathbf{v}) + c(\mathbf{u}, \mathbf{u}, \mathbf{v}) + b(\mathbf{v}, p) + \nu(\alpha\mathbf{u}, \mathbf{v})_\Gamma \\ = (\mathbf{f} - \boldsymbol{\xi}, \mathbf{v}) + (\mathbf{h} - \boldsymbol{\xi}, \mathbf{v})_\Gamma \quad \forall \mathbf{v} \in \mathbf{H}^1(\Omega). \end{aligned}$$

Then the optimality system in terms of the variables $\mathbf{u}, p, \boldsymbol{\xi}$, and σ is given by (2.2), (3.16), (3.17), and (3.20). We will approximate solutions of the latter system by finite element methods. Once the state variables $\mathbf{u}$ and p and the Lagrange multipliers $\boldsymbol{\xi}$ and σ, or rather, approximations to these, are found, the optimal controls may be easily computed from the optimality conditions (3.18) and (3.19), i.e., we essentially have that $\mathbf{g}_d = -\boldsymbol{\xi}$ and $\mathbf{g}_b = -\boldsymbol{\xi}|_\Gamma$.

Integrations by parts may be used to show that the system (2.2), (3.16), (3.17), and (3.20) constitutes a weak formulation of the problem

$$(3.21) \quad -\nu\,\mathrm{div}((\mathrm{grad}\,\mathbf{u}) + (\mathrm{grad}\,\mathbf{u})^T) + \mathbf{u} \cdot \mathrm{grad}\,\mathbf{u} + \mathrm{grad}\,p = \mathbf{f} - \boldsymbol{\xi} \quad \text{in } \Omega,$$

$$(3.22) \qquad \mathrm{div}\,\mathbf{u} = 0 \quad \text{in } \Omega,$$

$$(3.23) \quad -p\mathbf{n} + \nu((\mathrm{grad}\,\mathbf{u}) + (\mathrm{grad}\,\mathbf{u})^T) \cdot \mathbf{n} + \nu\alpha\mathbf{u} = \mathbf{h} - \boldsymbol{\xi} \quad \text{on } \Gamma,$$

$$(3.24) \quad \begin{aligned} -\nu\,\mathrm{div}((\mathrm{grad}\,\boldsymbol{\xi}) + (\mathrm{grad}\,\boldsymbol{\xi})^T) + \boldsymbol{\xi} \cdot (\mathrm{grad}\,\mathbf{u})^T \\ - \mathbf{u} \cdot \mathrm{grad}\,\boldsymbol{\xi} + \mathrm{grad}\,\sigma = (\mathbf{u} - \mathbf{u}_0)^3 \quad \text{in } \Omega, \end{aligned}$$

$$(3.25) \qquad \mathrm{div}\,\boldsymbol{\xi} = 0 \quad \text{in } \Omega,$$

and

$$(3.26) \quad -\sigma\mathbf{n} + \nu((\mathrm{grad}\,\boldsymbol{\xi}) + (\mathrm{grad}\,\boldsymbol{\xi})^T) \cdot \mathbf{n} + \nu\alpha\boldsymbol{\xi} = -(\mathbf{u} \cdot \mathbf{n})\boldsymbol{\xi} \quad \text{on } \Gamma.$$

Note that in (3.24)

$$(\mathbf{u} \cdot \mathrm{grad}\, \boldsymbol{\xi})_i = \sum_{j=1}^d u_j \frac{\partial \xi_i}{\partial x_j}, \quad (\boldsymbol{\xi} \cdot (\mathrm{grad}\, \mathbf{u})^T)_i = \sum_{j=1}^d \xi_j \frac{\partial u_j}{\partial x_i} \quad \text{for } i = 1, \ldots, d.$$

The optimality system (3.21)–(3.26) consists of the Navier-Stokes system (3.21)–(3.23) and the system (3.24)–(3.26) whose left-hand side is the adjoint of the Navier-Stokes operator linearized about $(\mathbf{u}, p)$.

Remark. Our notion of an optimal solution is a local one; see (2.5). Moreover, there is no reason to believe that, in general, optimal solutions are unique. This is to be expected, since 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 [6, 7, 14, 15], 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.

3.3. Regularity of solutions of the optimality system. We now examine the regularity of solutions of the optimality system (2.2), (3.16), (3.17), and (3.20), or equivalently, (3.21)–(3.26).

Theorem 3.2. *Suppose the given data satisfies* $\mathbf{h} \in \mathbf{H}^{1/2}(\Gamma)$, $\mathbf{f} \in \mathbf{L}^2(\Omega)$, *and* $\mathbf{u}_0 \in \mathbf{L}^6(\Omega)$. *Suppose that* Ω *is of class* $C^{1,1}$. *Then, if* $(\mathbf{u}, p, \boldsymbol{\xi}, \sigma) \in \mathbf{H}^1(\Omega) \times L^2(\Omega) \times \mathbf{H}^2(\Omega) \times L^2(\Omega)$ *denotes a solution of the optimality system* (2.2), (3.16), (3.17), *and* (3.20), *or equivalently,* (3.21)–(3.26), *we have that* $(\mathbf{u}, p, \boldsymbol{\xi}, \sigma) \in \mathbf{H}^2(\Omega) \times H^1(\Omega) \times \mathbf{H}^2(\Omega) \times H^1(\Omega)$.

Proof. Since $\boldsymbol{\xi} \in \mathbf{H}^1(\Omega)$, we have that $\boldsymbol{\xi}|_\Gamma \in \mathbf{H}^{1/2}(\Gamma)$, and $\boldsymbol{\xi} \in \mathbf{L}^2(\Omega)$. Then, using the hypotheses, the right-hand side of (3.21) belongs to $\mathbf{L}^2(\Omega)$ and the right-hand side of (3.23) belongs to $\mathbf{H}^{1/2}(\Gamma)$. Then, the additional regularity of $\mathbf{u}$ and p follow from well-known theories concerning the Navier-Stokes equations (see [6, 14]).

Now, since $\mathbf{u} \in \mathbf{H}^2(\Omega)$ and $\boldsymbol{\xi} \in \mathbf{H}^1(\Omega)$, we have that $\boldsymbol{\xi} \cdot (\mathrm{grad}\, \mathbf{u})^T$ and $\mathbf{u} \cdot \mathrm{grad}\, \boldsymbol{\xi}$ belong to $\mathbf{L}^2(\Omega)$. Moreover, $(\mathbf{u} - \mathbf{u}_0)^3$ certainly belongs to $\mathbf{L}^2(\Omega)$ as well, since $(\mathbf{u} - \mathbf{u}_0) \in \mathbf{L}^6(\Omega)$. Thus, if we rewrite (3.24) in the form

$$(3.27) \quad \begin{aligned} &-\nu\, \mathrm{div}((\mathrm{grad}\, \boldsymbol{\xi}) + (\mathrm{grad}\, \boldsymbol{\xi})^T) + \mathrm{grad}\, \sigma \\ &= -\boldsymbol{\xi} \cdot (\mathrm{grad}\, \mathbf{u})^T + \mathbf{u} \cdot \mathrm{grad}\, \boldsymbol{\xi} + (\mathbf{u} - \mathbf{u}_0)^3 \quad \text{in } \Omega, \end{aligned}$$

we have a right-hand side that belongs to $\mathbf{L}^2(\Omega)$. Moreover, we have that $\mathbf{u}|_\Gamma \in \mathbf{H}^{3/2}(\Gamma)$ and $\boldsymbol{\xi}|_\Gamma \in \mathbf{H}^{1/2}(\Gamma)$. Then, since Ω is of class $C^{1,1}$, one may conclude that $\mathbf{u} \cdot \mathbf{n} \in H^1(\Gamma)$ and that $(\mathbf{u} \cdot \mathbf{n})\boldsymbol{\xi}|_\Gamma \in \mathbf{L}^2(\Gamma)$, i.e., the right-hand side of (3.26) belongs to $\mathbf{L}^2(\Gamma)$. Then, well-known results for the Stokes problem applied to (3.25), (3.26), and (3.27) yield that $\boldsymbol{\xi} \in \mathbf{H}^{3/2}(\Omega)$. With this new information, we can deduce that actually the right-hand side of (3.26) belongs

to $\mathbf{H}^{1/2}(\Gamma)$, so that again, results for the Stokes problem yield that $\boldsymbol{\xi} \in \mathbf{H}^2(\Omega)$ and $\sigma \in H^1(\Omega)$. $\square$

Remark. The above result also holds for convex regions of $\mathbb{R}^2$. In general, we may show that if $\mathbf{f} \in \mathbf{H}^m(\Omega)$, $\mathbf{u}_0 \in \mathbf{W}^{m,6}(\Omega)$, $\mathbf{h} \in \mathbf{H}^{m+1/2}(\Gamma)$, and Ω is sufficiently smooth, then $(\mathbf{u}, p, \boldsymbol{\xi}, \sigma) \in \mathbf{H}^{m+2}(\Omega) \times H^{m+1}(\Omega) \times \mathbf{H}^{m+2}(\Omega) \times H^{m+1}(\Omega)$. In particular, if $\mathbf{f}$, $\mathbf{u}_0$, and $\mathbf{h}$ are all of class $C^\infty(\overline{\Omega})$, and Ω is of class C^∞, then $\mathbf{u}, p, \boldsymbol{\xi}$, and σ are all $C^\infty(\overline{\Omega})$ functions as well.

4. FINITE ELEMENT APPROXIMATIONS

4.1. Finite element discretizations. A finite element discretization of the optimality system (2.2), (3.16), (3.17), and (3.20) is defined in the usual manner. First one chooses families of finite-dimensional subspaces $\mathbf{V}^h \subset \mathbf{H}^1(\Omega)$ and $S^h \subset L^2(\Omega)$. These families are parametrized by a parameter h that tends to zero; commonly, h is chosen to be some measure of the grid size. Here we may choose any pair of subspaces $\mathbf{V}^h$ and S^h that can be used for finding finite element approximations of solutions of the Navier-Stokes equations. Thus, we make the following standard assumptions, which are exactly those employed in well-known finite element methods for the Navier-Stokes equations. First, we have the approximation properties: there exist an integer k and a constant C, independent of h, $\mathbf{v}$, and q, such that

$$(4.1) \qquad \inf_{\mathbf{v}^h \in \mathbf{V}^h} \|\mathbf{v} - \mathbf{v}^h\|_1 \leq Ch^m \|\mathbf{v}\|_{m+1} \quad \forall \mathbf{v} \in \mathbf{H}^{m+1}(\Omega), \ 1 \leq m \leq k,$$

and

$$(4.2) \qquad \inf_{q^h \in S^h} \|q - q^h\|_0 \leq Ch^h \|q\|_m \quad \forall q \in H^m(\Omega), \ 1 \leq m \leq k.$$

Next, we assume the *inf-sup condition*, or *Ladyzhenskaya-Babuška-Brezzi condition*: there exists a constant C, independent of h, such that

$$(4.3) \qquad \inf_{0 \neq q^h \in S^h} \sup_{0 \neq \mathbf{v}^h \in \mathbf{V}^h} \frac{b(\mathbf{v}^h, q^h)}{\|\mathbf{v}^h\|_1 \|q^h\|_0} \geq C.$$

This condition assures the stability of finite element discretizations of the Navier-Stokes equations. For thorough discussions of the approximation properties (4.1)–(4.2), see, e.g., [4] and for like discussions of the stability condition (4.3), see, e.g., [6, 7]. These references may also be consulted for a catalogue of finite element subspaces that meet the requirements of (4.1)–(4.3).

Once the approximating subspaces have been chosen, we seek $\mathbf{u}^h \in \mathbf{V}^h$, $p^h \in S^h$, $\boldsymbol{\xi}^h \in \mathbf{V}^h$, and $\sigma^h \in S^h$ such that

$$(4.4) \qquad \begin{aligned} \nu a(\mathbf{u}^h, \mathbf{v}^h) + c(\mathbf{u}^h, \mathbf{u}^h, \mathbf{v}^h) + b(\mathbf{v}^h, p^h) + \nu(\alpha \mathbf{u}^h, \mathbf{v}^h)_\Gamma \\ = (\mathbf{f} - \boldsymbol{\xi}^h, \mathbf{v}^h) + (\mathbf{h} - \boldsymbol{\xi}^h, \mathbf{v}^h)_\Gamma \quad \forall \mathbf{v}^h \in \mathbf{V}^h, \end{aligned}$$

$$(4.5) \qquad b(\mathbf{u}^h, q^h) = 0 \quad \forall q^h \in S^h,$$

$$(4.6) \quad \nu a(\boldsymbol{\omega}^h, \boldsymbol{\xi}^h) + c(\boldsymbol{\omega}^h, \mathbf{u}^h, \boldsymbol{\xi}^h) + c(\mathbf{u}^h, \boldsymbol{\omega}^h, \boldsymbol{\xi}^h) + b(\boldsymbol{\omega}^h, \sigma^h)$$
$$+ \nu(\alpha\boldsymbol{\omega}^h, \boldsymbol{\xi}^h)_\Gamma = ((\mathbf{u}^h - \mathbf{u}_0)^3, \boldsymbol{\omega}^h) \quad \forall \boldsymbol{\omega}^h \in \mathbf{V}^h,$$

and

$$(4.7) \quad b(\boldsymbol{\xi}^h, \psi^h) = 0 \quad \forall \psi^h \in S^h,$$

where, as always, $(\mathbf{u}^h - \mathbf{u}_0)^3$ denotes componentwise exponentiations. From a computational standpoint, this is a formidable system. In three dimensions, we have a coupled system of *eight* nonlinear equations involving *eight* unknown discrete scalar fields. Therefore, how one solves this system is a rather important question. However, in this paper we concern ourselves only with questions about the accuracy of finite element approximations; questions about efficient solution methods and implementation techniques, as well as computational examples, will be addressed in another paper.

4.2. Quotation of some results concerning the approximation of a class of nonlinear problems.
The error estimates to be derived in §4.3 make use of results of [3] and [5] (see also [6]) concerning the approximation of a class of nonlinear problems. Here, for the sake of completeness, we will state the relevant results, specialized to our needs. The nonlinear problems to be considered are of the type

$$(4.8) \quad F(\lambda, \varphi) \equiv \varphi + TG(\lambda, \varphi) = 0,$$

where $T \in \mathscr{L}(Y; X)$, G is a C^2 mapping from $\Lambda \times X$ into Y, X and Y are Banach spaces, and Λ is a compact interval of $\mathbb{R}$. We say that $\{(\lambda, \varphi(\lambda)): \lambda \in \Lambda\}$ is a branch of solutions of (4.8) if $\lambda \to \varphi(\lambda)$ is a continuous function from Λ into X such that $F(\lambda, \varphi(\lambda)) = 0$. The branch is called a *nonsingular branch* if we also have that $D_\varphi F(\lambda, \varphi(\lambda))$ is an isomorphism from X into X for all $\lambda \in \Lambda$. Here, D_φ denotes the Fréchet derivative with respect to φ.

Approximations are defined by introducing a subspace $X^h \subset X$ and an approximating operator $T^h \in \mathscr{L}(Y; X^h)$. Then, we seek $\varphi^h \in X^h$ such that

$$(4.9) \quad F^h(\lambda, \varphi^h) \equiv \varphi^h + T^h G(\lambda, \varphi^h) = 0.$$

We will assume that there exists another Banach space Z, contained in Y, with continuous imbedding, such that

$$(4.10) \quad D_\varphi G(\lambda, \varphi) \in \mathscr{L}(X; Z) \quad \forall \lambda \in \Lambda \text{ and } \varphi \in X.$$

Concerning the operator T^h, we assume the approximation properties

$$(4.11) \quad \lim_{h \to 0} \|(T^h - T)r\|_X = 0 \quad \forall r \in Y$$

and

$$(4.12) \quad \lim_{h \to 0} \|(T^h - T)\|_{\mathscr{L}(Z; X)} = 0.$$

Note that (4.12) follows from (4.11) whenever the imbedding $Z \subset Y$ is compact.

We now may state the first result that will be used in the sequel. In the statement of the theorem, $D^2 G$ represents any and all second Fréchet derivatives of G.

Theorem 4.1. *Let X and Y be Banach spaces and Λ a compact subset of $\mathbb{R}$. Assume that G is a C^2 mapping from $\Lambda \times X$ into Y and that $D^2 G$ is bounded on all bounded sets of $\Lambda \times X$. Assume that (4.10)–(4.12) hold and that $\{(\lambda, \varphi(\lambda)); \lambda \in \Lambda\}$ is a branch of nonsingular solutions of (4.8). Then there exists a neighborhood $\mathscr{O}$ of the origin in X and, for $h \leq h_0$ small enough, a unique C^2 function $\lambda \to \varphi^h(\lambda) \in X^h$ such that $\{(\lambda, \varphi^h(\lambda)); \lambda \in \Lambda\}$ is a branch of nonsingular solutions of (4.9) and $\varphi^h(\lambda) - \varphi(\lambda) \in \mathscr{O}$ for all $\lambda \in \Lambda$. Moreover, there exists a constant $C > 0$, independent of h and λ, such that*

$$(4.13) \qquad \|\varphi^h(\lambda) - \varphi(\lambda)\|_X \leq C\|(T^h - T)G(\lambda, \varphi(\lambda))\|_X \quad \forall \lambda \in \Lambda.$$

For the second result, we have to introduce two other Banach spaces H and W, such that $W \subset X \subset H$, with continuous imbeddings, and assume that

$$(4.14) \qquad \begin{array}{l} \text{for all } w \in W, \text{ the operator } D_\varphi G(\lambda, w) \text{ may be extended as} \\ \text{a linear operator of } \mathscr{L}(H; Y), \text{ the mapping } w \to D_\varphi G(\lambda, w) \\ \text{being continuous from } W \text{ into } \mathscr{L}(H; Y). \end{array}$$

We also suppose that

$$(4.15) \qquad \lim_{h \to 0} \|T^h - T\|_{\mathscr{L}(Y; H)} = 0.$$

Then we may state the following additional result.

Theorem 4.2. *Assume the hypotheses of Theorem 4.1 and also assume that (4.14) and (4.15) hold. Assume in addition that*

$$(4.16) \qquad \begin{array}{l} \text{for each } \lambda \in \Lambda, \ \varphi(\lambda) \in W \text{ and the function } \lambda \to \varphi(\lambda) \text{ is} \\ \text{continuous from } \Lambda \text{ into } W \end{array}$$

and

$$(4.17) \qquad \text{for each } \lambda \in \Lambda, \ D_\varphi F(\lambda, \varphi(\lambda)) \text{ is an isomorphism of } H.$$

Then, for $h \leq h_1$ sufficiently small, there exists a constant C, independent of h and λ, such that

$$(4.18) \quad \|\varphi^h(\lambda) - \varphi(\lambda)\|_H \leq C\|(T^h - T)G(\lambda, \varphi(\lambda))\|_H + \|\varphi^h(\lambda) - \varphi(\lambda)\|_X^2 \quad \forall \lambda \in \Lambda.$$

4.3. Error estimates. We begin by recasting the optimality system (2.2), (3.16), (3.17), and (3.20) and its discretization (4.4)–(4.7) into a form that fits into the framework of §4.2. Let $\lambda = 1/\nu$; thus, if our governing system has been

nondimensionalized, λ is the Reynolds number. Let

$$X = \mathbf{H}^1(\Omega) \times L^2(\Omega) \times \mathbf{H}^1(\Omega) \times L^2(\Omega),$$

$$Y = (\mathbf{H}^1(\Omega))^* \times \mathbf{H}^{-1/2}(\Gamma) \times (\mathbf{H}^1(\Omega))^*,$$

$$Z = \mathbf{L}^{3/2}(\Omega) \times \mathbf{L}^2(\Gamma) \times \mathbf{L}^{3/2}\Omega,$$

$$X^h = \mathbf{V}^h \times S^h \times \mathbf{V}^h \times S^h,$$

where $(\mathbf{H}^1(\Omega))^*$ denotes the dual space of $\mathbf{H}^1(\Omega)$. Note that $Z \subset Y$ with a compact imbedding.

Let the operator $T \in \mathscr{L}(Y; X)$ be defined in the following manner: $T(\boldsymbol{\zeta}, \boldsymbol{\theta}, \boldsymbol{\eta}) = (\mathbf{u}, p, \boldsymbol{\xi}, \sigma)$ for $(\boldsymbol{\zeta}, \boldsymbol{\theta}, \boldsymbol{\eta}) \in Y$ and $(\mathbf{u}, p, \boldsymbol{\xi}, \sigma) \in X$, if and only if

$$(4.19) \qquad a(\mathbf{u}, \mathbf{v}) + b(\mathbf{v}, p) + (\alpha \mathbf{u}, \mathbf{v})_\Gamma = (\boldsymbol{\zeta}, \mathbf{v}) + (\boldsymbol{\theta}, \mathbf{v})_\Gamma \quad \forall \mathbf{v} \in \mathbf{H}^1(\Omega),$$

$$(4.20) \qquad\qquad b(\mathbf{u}, q) = 0 \quad \forall q \in L^2(\Omega),$$

$$(4.21) \qquad a(\boldsymbol{\omega}, \boldsymbol{\xi}) + b(\boldsymbol{\omega}, \sigma) + (\alpha \boldsymbol{\omega}, \boldsymbol{\xi})_\Gamma = (\boldsymbol{\eta}, \boldsymbol{\omega}) \quad \forall \boldsymbol{\omega} \in \mathbf{H}^1(\Omega),$$

and

$$(4.22) \qquad\qquad b(\boldsymbol{\xi}, \psi) = 0 \quad \forall \psi \in L^2(\Omega).$$

Clearly, (4.19)–(4.22) consist of two *uncoupled* Stokes problems and T is their solution operator.

Analogously, the operator $T^h \in \mathscr{L}(Y; X)$ is defined as follows: $T^h(\boldsymbol{\zeta}, \boldsymbol{\theta}, \boldsymbol{\eta}) = (\mathbf{u}^h, p^h, \boldsymbol{\xi}^h, \sigma^h)$ for $(\boldsymbol{\zeta}, \boldsymbol{\theta}, \boldsymbol{\eta}) \in Y$ and $(\mathbf{u}^h, p^h, \boldsymbol{\xi}^h, \sigma^h) \in X^h$, if and only if

$$(4.23) \qquad a(\mathbf{u}^h, \mathbf{v}^h) + b(\mathbf{v}^h, p^h) + (\alpha \mathbf{u}^h, \mathbf{v}^h)_\Gamma = (\boldsymbol{\zeta}, \mathbf{v}^h) + (\boldsymbol{\theta}, \mathbf{v}^h)_\Gamma \quad \forall \mathbf{v}^h \in \mathbf{V}^h,$$

$$(4.24) \qquad\qquad b(\mathbf{u}^h, q^h) = 0 \quad \forall q^h \in S^h,$$

$$(4.25) \qquad \nu a(\boldsymbol{\omega}^h, \boldsymbol{\xi}^h) + b(\boldsymbol{\omega}^h, \sigma^h) + (\alpha \boldsymbol{\omega}^h, \boldsymbol{\xi}^h)_\Gamma = (\boldsymbol{\eta}, \boldsymbol{\omega}^h) \quad \forall \boldsymbol{\omega}^h \in \mathbf{V}^h,$$

and

$$(4.26) \qquad\qquad b(\boldsymbol{\xi}^h, \psi^h) = 0 \quad \forall \psi^h \in S^h.$$

Clearly, (4.23)–(4.26) consist of two discrete Stokes problems that are discretizations of the Stokes problems (4.19)–(4.22); also, T^h is the solution operator for these two discrete Stokes problems.

Let Λ denote a compact subset of $\mathbb{R}_+$. Next, we define the *nonlinear* mapping $G: \Lambda \times X \to Y$ as follows: $G(\lambda, (\mathbf{u}, p, \boldsymbol{\xi}, \sigma)) = (\boldsymbol{\zeta}, \boldsymbol{\theta}, \boldsymbol{\eta})$ for $\lambda \in \Lambda$, $(\mathbf{u}, p, \boldsymbol{\xi}, \sigma) \in X$, and $(\boldsymbol{\zeta}, \boldsymbol{\theta}, \boldsymbol{\eta}) \in Y$, if and only if

$$(4.27) \qquad (\boldsymbol{\zeta}, \mathbf{v}) = \lambda c(\mathbf{u}, \mathbf{u}, \mathbf{v}) - \lambda(\mathbf{f} - \boldsymbol{\xi}, \mathbf{v}) \quad \forall \mathbf{v} \in \mathbf{H}^1(\Omega),$$

$$(4.28) \qquad (\boldsymbol{\theta}, \mathbf{w})_\Gamma = -\lambda(\mathbf{h} - \boldsymbol{\xi}, \mathbf{w})_\Gamma \quad \forall \mathbf{w} \in \mathbf{H}^1(\Omega),$$

and

$$(4.29) \quad (\pmb{\eta}, \pmb{\omega}) = \lambda c(\pmb{\omega}, \mathbf{u}, \pmb{\xi}) + \lambda c(\mathbf{u}, \pmb{\omega}, \pmb{\xi}) - \lambda((\mathbf{u} - \mathbf{u}_0)^3, \pmb{\omega}) \quad \forall \pmb{\omega} \in \mathbf{H}^1(\Omega).$$

It is easily seen that the optimality system (2.2), (3.16), (3.17), and (3.20) is equivalent to

$$(4.30) \qquad (\mathbf{u}, \lambda p, \pmb{\xi}, \lambda \sigma) + TG(\lambda, (\mathbf{u}, \lambda p, \pmb{\xi}, \lambda \sigma)) = 0,$$

and that the discrete optimality system (4.4)–(4.7) is equivalent to

$$(4.31) \qquad (\mathbf{u}^h, \lambda p^h, \pmb{\xi}^h, \lambda \sigma^h) + TG(\lambda, (\mathbf{u}^h, \lambda p^h, \pmb{\xi}^h, \lambda \sigma^h)) = 0.$$

We have thus recast our continuous and discrete optimality problems into a form that enables us to apply Theorems 4.1 and 4.2.

A solution $(\mathbf{u}(\lambda), p(\lambda), \pmb{\xi}(\lambda), \sigma(\lambda))$ of the problem (2.2), (3.16), (3.17), and (3.20), or equivalently, of (4.30), is nonsingular if the linear system

$$a(\tilde{\mathbf{u}}, \mathbf{v}) + \lambda c(\tilde{\mathbf{u}}, \mathbf{u}, \mathbf{v}) + \lambda c(\mathbf{u}, \tilde{\mathbf{u}}, \mathbf{v}) + \lambda b(\mathbf{v}, \tilde{p}) + (\alpha \tilde{\mathbf{u}}, \mathbf{v})_\Gamma + \lambda(\tilde{\pmb{\xi}}, \mathbf{v})_\Gamma + \lambda(\tilde{\pmb{\xi}}, \mathbf{v})$$

$$= (\hat{\mathbf{f}}, \mathbf{v}) + (\hat{\mathbf{h}}, \mathbf{v})_\Gamma \quad \forall \mathbf{v} \in \mathbf{H}^1(\Omega),$$

$$b(\tilde{\mathbf{u}}, q) = (\hat{\varphi}, q) \quad \forall q \in L^2(\Omega),$$

$$a(\pmb{\omega}, \tilde{\pmb{\xi}}) + \lambda c(\pmb{\omega}, \tilde{\mathbf{u}}, \pmb{\xi}) + \lambda c(\pmb{\omega}, \mathbf{u}, \tilde{\pmb{\xi}}) + \lambda c(\mathbf{u}, \pmb{\omega}, \tilde{\pmb{\xi}}) + \lambda c(\tilde{\mathbf{u}}, \pmb{\omega}, \pmb{\xi}) + \lambda b(\pmb{\omega}, \tilde{\sigma})$$

$$+ (\alpha \tilde{\pmb{\xi}}, \pmb{\omega})_\Gamma - 3\lambda((\mathbf{u} - \mathbf{u}_0)^2 \cdot \tilde{\mathbf{u}}, \pmb{\omega}) = (\hat{\pmb{\eta}}, \pmb{\omega}) \quad \forall \pmb{\omega} \in \mathbf{H}^1\Omega,$$

and

$$b(\tilde{\pmb{\xi}}, \psi) = (\hat{\mu}, \psi) \quad \forall \psi \in L^2(\Omega)$$

has a unique solution $(\tilde{\mathbf{u}}, \tilde{p}, \tilde{\pmb{\xi}}, \tilde{\sigma}) \in X$ for every $\hat{\mathbf{f}}, \hat{\pmb{\eta}} \in (\mathbf{H}^1(\Omega))^*$, $\hat{\varphi}, \hat{\mu} \in L^2(\Omega)$, and $\hat{\mathbf{h}} \in \mathbf{L}^2(\Gamma)$. In the above equations we have used the notation $(\mathbf{u} - \mathbf{u}_0)^2 \cdot \tilde{\mathbf{u}}$ to denote a vector having components given by $(\mathbf{u} - \mathbf{u}_0)_j^2(\tilde{\mathbf{u}})_j$, $j = 1, \ldots, d$, where there is no implied summation.

An analogous definition holds for nonsingular solutions of the discrete optimality system (4.4)–(4.7), or equivalently, (4.31). We will assume throughout that the optimality system (2.2), (3.16), (3.17), and (3.20), or equivalently, (4.30), has a branch of nonsingular solutions for λ belonging to a compact interval of $\mathbb{R}_+$.

Remark. It can be shown, using techniques similar to those employed for the Navier-Stokes equations (see [15] and the references cited therein) that for almost all values of the Reynolds number, i.e., for almost all data and values of the viscosity ν, the optimality system (2.2), (3.16), (3.17), and (3.20), or equivalently, (4.30), is nonsingular, i.e., has locally unique solutions. Thus, it is reasonable to assume that the optimality system has branches of nonsingular solutions. (However, we note that, just as in the Navier-Stokes case, it is impossible to *predict*, except in very simple settings, exactly at what values of the Reynolds number singularities, e.g., bifurcations, appear.)

Using Theorem 4.1, we are led to the following error estimate.

Theorem 4.3. *Assume that Λ is a compact interval of $\mathbb{R}_+$ and that there exists a branch $\{(\lambda, \varphi(\lambda) = (\mathbf{u}, p, \boldsymbol{\xi}, \sigma)) : \lambda \in \Lambda\}$ of nonsingular solutions of the optimality system (2.2), (3.16), (3.17), and (3.20). Assume that the finite element spaces $\mathbf{V}^h$ and S^h satisfy the conditions (4.1)–(4.3). Then, there exists a neighborhood $\mathscr{O}$ of the origin in $X = \mathbf{H}^1(\Omega) \times L^2(\Omega) \times \mathbf{H}^1(\Omega) \times L^2(\Omega)$ and, for $h \leq h_0$ small enough, a unique branch $\{(\lambda, \varphi^h(\lambda) = (\mathbf{u}^h, p^h, \boldsymbol{\xi}^h, \sigma^h)) : \lambda \in \Lambda\}$ of solutions of the discrete optimality system (4.4)–(4.7) such that $\varphi^h(\lambda) - \varphi(\lambda) \in \mathscr{O}$ for all $\lambda \in \Lambda$. Moreover,*

$$
(4.32) \quad \begin{aligned}
\|\varphi^h(\lambda) - \varphi(\lambda)\|_X &= \|\mathbf{u}(\lambda) - \mathbf{u}^h(\lambda)\|_1 + \|p(\lambda) - p^h(\lambda)\|_0 \\
&\quad + \|\boldsymbol{\xi}(\lambda) - \boldsymbol{\xi}^h(\lambda)\|_1 + \|\sigma(\lambda) - \sigma^h(\lambda)\|_0 \to 0
\end{aligned}
$$

as $h \to 0$, uniformly in $\lambda \in \Lambda$.

If, in addition, the solution of the optimality system satisfies $(\mathbf{u}(\lambda), p(\lambda), \boldsymbol{\xi}(\lambda), \sigma(\lambda)) \in \mathbf{H}^{m+1}(\Omega) \times \mathbf{H}^m(\Omega) \times \mathbf{H}^{m+1}(\Omega) \times \mathbf{H}^m(\Omega)$ for $\lambda \in \Lambda$, then there exists a constant C, independent of h, such that

$$
(4.33) \quad \begin{aligned}
&\|\mathbf{u}(\lambda) - \mathbf{u}^h(\lambda)\|_1 + \|p(\lambda) - p^h(\lambda)\|_0 + \|\boldsymbol{\xi}(\lambda) - \boldsymbol{\xi}^h(\lambda)\|_1 + \|\sigma(\lambda) - \sigma^h(\lambda)\|_0 \\
&\leq Ch^m(\|\mathbf{u}(\lambda)\|_{m+1} + \|p(\lambda)\|_m + \|\boldsymbol{\xi}(\lambda)\|_{m+1} + \|\sigma(\lambda)\|_m)
\end{aligned}
$$

uniformly in $\lambda \in \Lambda$.

Proof. Clearly, we may write $G = \lambda \widetilde{G}(\mathbf{u}, \boldsymbol{\xi})$, where $\widetilde{G}$ is a C^∞ polynomial map from X into Y. Therefore, using (1.8), $D^2\widetilde{G}(\mathbf{u}, \boldsymbol{\xi})$ is easily shown to be bounded on all bounded sets of X. Now, given $(\mathbf{u}, p, \boldsymbol{\xi}, \sigma) \in X$, a direct computation yields that $(\tilde{\boldsymbol{\zeta}}, \tilde{\boldsymbol{\theta}}, \tilde{\boldsymbol{\eta}}) \in Y$ satisfies

$$
(\tilde{\boldsymbol{\zeta}}, \tilde{\boldsymbol{\theta}}, \tilde{\boldsymbol{\eta}}) = D_\varphi G(\lambda, (\mathbf{u}, p, \boldsymbol{\xi}, \sigma))(\mathbf{v}, q, \boldsymbol{\omega}, \psi)
$$

for $(\mathbf{v}, q, \boldsymbol{\omega}, \psi) \in X$, if and only if

$$
(\tilde{\boldsymbol{\zeta}}, \overline{\mathbf{v}}) = \lambda c(\mathbf{u}, \mathbf{v}, \overline{\mathbf{v}}) + \lambda c(\mathbf{v}, \mathbf{u}, \overline{\mathbf{v}}) + (\boldsymbol{\omega}, \overline{\mathbf{v}}) \quad \forall \overline{\mathbf{v}} \in \mathbf{H}^1\Omega,
$$

$$
(\tilde{\boldsymbol{\theta}}, \overline{\mathbf{w}})_\Gamma = \lambda(\boldsymbol{\omega}, \overline{\mathbf{w}})_\Gamma \quad \forall \overline{\mathbf{w}} \in \mathbf{H}^1\Omega,
$$

and

$$
\begin{aligned}
(\tilde{\boldsymbol{\eta}}, \overline{\boldsymbol{\omega}}) &= \lambda c(\overline{\boldsymbol{\omega}}, \mathbf{v}, \boldsymbol{\xi}) + \lambda c(\overline{\boldsymbol{\omega}}, \mathbf{u}, \boldsymbol{\omega}) + \lambda c(\mathbf{v}, \overline{\boldsymbol{\omega}}, \boldsymbol{\xi}) + \lambda c(\mathbf{u}, \overline{\boldsymbol{\omega}}, \boldsymbol{\omega}) \\
&\quad - 3\lambda(\mathbf{u} - \mathbf{u}_0)^2 \cdot \mathbf{v}, \overline{\boldsymbol{\omega}}) \quad \forall \overline{\boldsymbol{\omega}} \in \mathbf{H}^1\Omega,
\end{aligned}
$$

where again $(\mathbf{u} - \mathbf{u}_0)^2 \cdot \mathbf{v}$ denotes the vector having a jth component given by $(\mathbf{u} - \mathbf{u}_0)_j^2 (\mathbf{v})_j$, with no implied summation. Thus, for given $(\mathbf{u}, p, \boldsymbol{\xi}, \sigma) \in X$, it follows from (1.8) that $D_\varphi G(\lambda, (\mathbf{u}, p, \boldsymbol{\xi}, \sigma)) \in \mathscr{L}(X; Y)$. On the other hand, since $(\mathbf{u}, p, \boldsymbol{\xi}, \sigma) \in X$ and $(\mathbf{v}, q, \boldsymbol{\omega}, \psi) \in X$, by the Sobolev imbedding theorem, $\mathbf{u}, \mathbf{v}, \boldsymbol{\xi}$, and $\boldsymbol{\omega} \in \mathbf{L}^6(\Omega)$; $\partial\mathbf{u}/\partial x_j$, $\partial\mathbf{v}/\partial x_j$, $\partial\boldsymbol{\xi}/\partial x_j$, and $\partial\boldsymbol{\omega}/\partial x_j \in \mathbf{L}^2(\Omega)$ for $j = 1, \ldots, d$; and $\boldsymbol{\omega}|_\Gamma \in \mathbf{L}^2(\Gamma)$. Then, it follows that $(\tilde{\boldsymbol{\zeta}}, \tilde{\boldsymbol{\theta}}, \tilde{\boldsymbol{\eta}}) \in Z$ and that, for $(\mathbf{u}, p, \boldsymbol{\xi}, \sigma) \in X$, $D_\varphi G(\lambda, (\mathbf{u}, p, \boldsymbol{\xi}, \sigma)) \in \mathscr{L}(X; Z)$. Of course,

Z is continuously imbedded into Y; moreover, the imbedding $Z \subset Y$ is compact.

Next, we turn to the approximation properties of the operator T^h. By extending well-known results for the Stokes problem with Dirichlet boundary conditions (see, e.g., [6]), one can show, provided that (4.1)–(4.3) hold, that if $(\bar{\mathbf{u}}, \bar{p}, \bar{\boldsymbol{\xi}}, \bar{\sigma})$ and $(\bar{\mathbf{u}}^h, \bar{p}^h, \bar{\boldsymbol{\xi}}^h, \bar{\sigma}^h)$ are solutions of (4.19)–(4.22) and (4.23)–(4.26), respectively, then

$$\|\bar{\mathbf{u}} - \bar{\mathbf{u}}^h\|_1 + \|\bar{p} - \bar{p}^h\|_0 + \|\bar{\boldsymbol{\xi}} - \bar{\boldsymbol{\xi}}^h\|_1 + \|\bar{\sigma} - \bar{\sigma}^h\|_0 \to 0,$$

so that (4.11), and therefore (4.32), hold. Since the imbedding of Z into Y is compact, (4.12) follows from (4.11). We also have the following results concerning the approximation of the Stokes problem. Suppose $(\mathbf{u}, p, \boldsymbol{\xi}, \sigma) = -TG(\lambda, (\mathbf{u}, p, \boldsymbol{\xi}, \sigma))$ satisfies $(\mathbf{u}, p, \boldsymbol{\xi}, \sigma) \in \mathbf{H}^{m+1}(\Omega) \times \mathbf{H}^m(\Omega) \times \mathbf{H}^{m+1}(\Omega) \times \mathbf{H}^m(\Omega)$. Then, there exists a constant C, independent of h, such that

$$\begin{aligned}
&\|\bar{\mathbf{u}} - \bar{\mathbf{u}}^h\|_1 + \|\bar{p} - \bar{p}^h\|_0 + \|\bar{\boldsymbol{\xi}} - \bar{\boldsymbol{\xi}}^h\|_1 + \|\bar{\sigma} - \bar{\sigma}^h\|_0 \\
&\leq Ch^m(\|\mathbf{u}\|_{m+1} + \|p\|_m + \|\boldsymbol{\xi}\|_{m+1} + \|\sigma\|_m),
\end{aligned}$$

i.e., in our notation we have that

$$\|(T - T^h)G(\lambda, \varphi(\lambda))\|_X \leq Ch^m(\|\mathbf{u}\|_{m+1} + \|p\|_m + \|\boldsymbol{\xi}\|_{m+1} + \|\sigma\|_m).$$

Then (4.33) follows from (4.13). $\square$

Using Theorem 4.2, we now derive an estimate for the error of $\mathbf{u}^h$ and $\boldsymbol{\xi}^h$ in the $\mathbf{L}^2(\Omega)$-norm. At this point it is convenient to examine (4.27)–(4.29) and note that $G(\lambda, (\mathbf{u}, p, \boldsymbol{\xi}, \sigma))$ does not depend on p or σ. Therefore, we now define $X = \mathbf{H}^1(\Omega) \times \mathbf{H}^1(\Omega)$, $Y = (\mathbf{H}^1(\Omega))^* \times (\mathbf{H}^1(\Omega))^*$, $Z = \mathbf{L}^{3/2}(\Omega) \times \mathbf{L}^{3/2}(\Omega)$, and $X^h = \mathbf{V}^h \times \mathbf{V}^h$ and restrict our view of the various mappings to these spaces. We introduce the spaces $H = \mathbf{L}^2(\Omega) \times \mathbf{L}^2(\Omega)$ and $W = \mathbf{H}^2(\Omega) \times \mathbf{H}^2(\Omega)$.

Theorem 4.4. *Assume the hypotheses of Theorems 3.2 and 4.3. Then, for $h \leq h_1$ sufficiently small, there exists a constant C, independent of h, such that*

$$(4.34) \quad \begin{aligned}
&\|\mathbf{u}(\lambda) - \mathbf{u}^h(\lambda)\|_0 + \|\boldsymbol{\xi}(\lambda) - \boldsymbol{\xi}^h(\lambda)\|_0 \\
&\leq Ch^{m+1}(\|\mathbf{u}(\lambda)\|_{m+1} + \|p(\lambda)\|_m + \|\boldsymbol{\xi}(\lambda)\|_{m+1} + \|\sigma(\lambda)\|_m).
\end{aligned}$$

Proof. We need only verify that (4.14)–(4.17) hold in our setting; then, well-known results [6] concerning $\mathbf{L}^2(\Omega)$-norm estimates for the velocity approximations in the Stokes equations setting and the results (4.18) and (4.33) easily lead to (4.34).

From Theorem 3.2 we have that $\mathbf{u}, \boldsymbol{\xi} \in \mathbf{H}^2(\Omega)$; then one can easily show that

$$(4.35) \quad \begin{aligned}
&|c(\mathbf{u}, \mathbf{v}, \bar{\mathbf{v}}) + c(\mathbf{v}, \mathbf{u}, \bar{\mathbf{v}})| \leq C\|\mathbf{u}\|_2 \|\mathbf{v}\|_0 \|\bar{\mathbf{v}}\|_1 \\
&\qquad \forall \mathbf{u} \in \mathbf{H}^2(\Omega), \ \mathbf{v} \in \mathbf{H}_0^1\Omega, \ \bar{\mathbf{v}} \in \mathbf{H}^1(\Omega)
\end{aligned}$$

and

$$|c(\overline{\boldsymbol{\varpi}}, \mathbf{v}, \boldsymbol{\xi}) + c(\overline{\boldsymbol{\varpi}}, \mathbf{u}, \boldsymbol{\omega}) + c(\mathbf{v}, \overline{\boldsymbol{\varpi}}, \boldsymbol{\xi}) + c(\mathbf{u}, \overline{\boldsymbol{\varpi}}, \boldsymbol{\omega})|$$

(4.36)
$$\leq C(\|\mathbf{u}\|_2\|\boldsymbol{\omega}\|_0 + \|\boldsymbol{\xi}\|_2\|\mathbf{v}\|_0)\|\overline{\boldsymbol{\varpi}}\|_1$$

$$\forall \mathbf{u}, \boldsymbol{\xi} \in \mathbf{H}^2(\Omega), \quad \mathbf{v}, \boldsymbol{\omega} \in \mathbf{H}_0^1(\Omega), \quad \overline{\mathbf{v}}, \overline{\boldsymbol{\varpi}} \in \mathbf{H}^1\Omega.$$

Then, we have that for $(\mathbf{u}, \boldsymbol{\xi}) \in W$, $(\mathbf{v}, \boldsymbol{\omega}) \to D_\varphi G(\lambda, (\mathbf{u}, \boldsymbol{\xi}))(\mathbf{v}, \boldsymbol{\omega})$ belongs to $\mathscr{L}(\mathbf{H}_0^1(\Omega) \times \mathbf{H}_0^1(\Omega), Y)$, where $\mathbf{H}_0^1(\Omega) \times \mathbf{H}_0^1(\Omega)$ is considered as a dense subspace of H. Then we may uniquely extend $D_\varphi G(\lambda, (\mathbf{u}, \boldsymbol{\xi}))$ as an operator in $\mathscr{L}(H, Y)$. The continuity of the mapping $D_\varphi G(\lambda, (\cdot)) \in \mathscr{L}(W, \mathscr{L}(H, Y))$ is an easy consequence of (4.35) and (4.36). Thus we have verified (4.14). Next, (4.15) follows from (4.11) and the fact that X is compactly imbedded into H. The results of §3.3 and the fact that λ belongs to Λ, a compact interval of $\mathbb{R}_+$, easily yield (4.16), where $\varphi(\lambda) = (\mathbf{u}(\lambda), \boldsymbol{\xi}(\lambda))$. Finally, (4.17) follows from the well-known properties of the solution operator for the Stokes problem, the continuity of the mapping $D_\varphi G(\lambda, (\mathbf{u}, \boldsymbol{\xi}))$, and the fact that we have assumed that $(\mathbf{u}(\lambda), \boldsymbol{\xi}(\lambda))$, $\lambda \in \Lambda$, defines a nonsingular branch of solutions. □

A consequence of Theorems 4.3 and 4.4 is the following corollary that gives estimates for the error in the approximation of the controls.

Corollary 4.5. *Define the approximate controls by*

(4.37)
$$\mathbf{g}_d^h = -\boldsymbol{\xi}^h \quad and \quad \mathbf{g}_b^h = -\boldsymbol{\xi}^h|_\Gamma$$

and assume the hypotheses of Theorem 4.3. Then, for $h \leq h_0$ sufficiently small, there exists a constant C, independent of h, such that

(4.38)
$$\|\mathbf{g}_d^h - \mathbf{g}_d\|_1 \leq Ch^m(\|\mathbf{u}(\lambda)\|_{m+1} + \|p(\lambda)\|_m + \|\boldsymbol{\xi}(\lambda)\|_{m+1} + \|\sigma(\lambda)\|_m)$$

and

(4.39)
$$\|\mathbf{g}_b^h - \mathbf{g}_b\|_{1/2, \Gamma} \leq Ch^m(\|\mathbf{u}(\lambda)\|_{m+1} + \|p(\lambda)\|_m + \|\boldsymbol{\xi}(\lambda)\|_{m+1} + \|\sigma(\lambda)\|_m).$$

If in addition the hypotheses of Theorem 4.4 hold, then, for $h \leq h_1$ sufficiently small, there exists a constant C, independent of h, such that

(4.40)
$$\|\mathbf{g}_d^h - \mathbf{g}_d\|_0 \leq Ch^{m+1}(\|\mathbf{u}(\lambda)\|_{m+1} + \|p(\lambda)\|_m + \|\boldsymbol{\xi}(\lambda)\|_{m+1} + \|\sigma(\lambda)\|_m)$$

and

(4.41) $\|\mathbf{g}_b^h - \mathbf{g}_b\|_{0, \Gamma} \leq Ch^{m+1/2}(\|\mathbf{u}(\lambda)\|_{m+1} + \|p(\lambda)\|_m + \|\boldsymbol{\xi}(\lambda)\|_{m+1} + \|\sigma(\lambda)\|_m).$

Proof. Recall that $\mathbf{g}_d = -\boldsymbol{\xi}$ and $\mathbf{g}_b = -\boldsymbol{\xi}|_\Gamma$ (see (3.18) and (3.19)). Then (4.38) and (4.40) easily follow from (4.33) and (4.34), respectively; (4.39) follows from (4.33) and the trace theorem (see [1]), i.e., $\|\mathbf{g}_b^h - \mathbf{g}_b\|_{1/2, \Gamma} \leq \|\boldsymbol{\xi}^h - \boldsymbol{\xi}\|_1$. Finally, (4.41) follows from (4.33), (4.34) and the well-known inequality $\|\mathbf{g}_b^h - \mathbf{g}_b\|_{0, \Gamma} \leq C[\varepsilon\|\boldsymbol{\xi}^h - \boldsymbol{\xi}\|_1 + (1/\varepsilon)\|\boldsymbol{\xi}^h - \boldsymbol{\xi}\|_0]$, with the choice $\varepsilon = h^{1/2}$. □

5. The drag functional, partial controls, and distributed controls

We now consider four variations on the problem considered in §§2–4. A substantial portion of the analyses and results of those sections that apply to the case that uses the functional (2.3) with distributed and Neumann controls will also apply to the variations considered in this section. Therefore, here we will merely point out the differences. Details about all the topics considered in this section may be found in [11].

5.1. Optimizing the drag functional.

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

$$(5.1) \quad \mathscr{K}(\mathbf{u}, \mathbf{g}_b, \mathbf{g}_d) = \tfrac{\nu}{2} a(\mathbf{u}, \mathbf{u}) - (\mathbf{f}, \mathbf{u}) + \tfrac{\nu}{2}(\alpha \mathbf{u}, \mathbf{u})_\Gamma + \tfrac{1}{2}\|\mathbf{g}_b\|_{0,\Gamma}^2 + \tfrac{1}{2}\|\mathbf{g}_d\|_0^2.$$

The admissibility set is now defined by

$$\mathscr{V}_{\mathrm{ad}} = \{(\mathbf{u}, \mathbf{g}_b, \mathbf{g}_d) \in \mathbf{H}^1(\Omega) \times \mathbf{L}^2(\Gamma) \times \mathbf{L}^2(\Omega): \mathscr{K}(\mathbf{u}, \mathbf{g}_b, \mathbf{g}_d) < \infty,$$

and there exists a $p \in L^2(\Omega)$ such that (2.1) and (2.2) are satisfied$\}$.

The optimization problem at hand is to minimize (5.1) over $\mathscr{V}_{\mathrm{ad}}$. The existence of optimal solutions may be shown as in Theorem 2.1; this task is somewhat easier in the present context, since the fact that $\mathscr{K}(\mathbf{u}^{(n)}, \mathbf{g}_b^{(n)}, \mathbf{b}_d^{(n)})$ is bounded for $(\mathbf{u}^{(n)}, \mathbf{g}_b^{(n)}, \mathbf{g}_d^{(n)}) \in \mathscr{V}_{\mathrm{ad}}$ immediately yields that $\|\mathbf{u}^{(n)}\|_1$ is bounded. Also, Theorem 3.1 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 (2.1)–(2.2), (3.17)–(3.19), and, instead of (3.16),

$$(5.2) \quad \begin{aligned} \nu a(\boldsymbol{\omega}, \boldsymbol{\xi}) + c(\boldsymbol{\omega}, \mathbf{u}, \boldsymbol{\xi}) + c(\mathbf{u}, \boldsymbol{\omega}, \boldsymbol{\xi}) + b(\boldsymbol{\omega}, \sigma) + \nu(\alpha \boldsymbol{\omega}, \boldsymbol{\xi})_\Gamma \\ = \nu a(\mathbf{u}, \boldsymbol{\omega}) + \nu(\alpha \mathbf{u}, \boldsymbol{\omega})_\Gamma - (\mathbf{f}, \boldsymbol{\omega}) \quad \forall \boldsymbol{\omega} \in \mathbf{H}^1\Omega. \end{aligned}$$

We may substitute (2.1) in the right-hand side of (5.2) to yield

$$(5.3) \quad \begin{aligned} \nu a(\boldsymbol{\omega}, \boldsymbol{\xi}) + c(\boldsymbol{\omega}, \mathbf{u}, \boldsymbol{\xi}) + c(\mathbf{u}, \boldsymbol{\omega}, \boldsymbol{\xi}) + b(\boldsymbol{\omega}, \sigma) + \nu(\alpha \boldsymbol{\omega}, \boldsymbol{\xi})_\Gamma \\ = -b(\boldsymbol{\omega}, p) - c(\mathbf{u}, \mathbf{u}, \boldsymbol{\omega}) + (\mathbf{g}_d, \boldsymbol{\omega}) + (\mathbf{h} + \mathbf{g}_b, \boldsymbol{\omega})_\Gamma \quad \forall \boldsymbol{\omega} \in \mathbf{H}^1(\Omega). \end{aligned}$$

Using the optimality conditions (3.18) and (3.19), we can eliminate the controls from (2.1) and (5.3) to yield (3.20) and

$$(5.4) \quad \begin{aligned} \nu a(\boldsymbol{\omega}, \boldsymbol{\xi}) + c(\boldsymbol{\omega}, \mathbf{u}, \boldsymbol{\xi}) + c(\mathbf{u}, \boldsymbol{\omega}, \boldsymbol{\xi}) + b(\boldsymbol{\omega}, \check{\sigma}) + \nu(\alpha \boldsymbol{\omega}, \boldsymbol{\xi})_\Gamma \\ = -c(\mathbf{u}, \mathbf{u}, \boldsymbol{\omega}) - (\boldsymbol{\xi}, \boldsymbol{\omega}) + (\mathbf{h} - \boldsymbol{\xi}, \boldsymbol{\omega})_\Gamma \quad \forall \boldsymbol{\omega} \in \mathbf{H}^1(\Omega), \end{aligned}$$

where $\check{\sigma} = \sigma + p$. In the sequel we will dispense with the $(\check{\ })$ 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 (2.2), (3.17), (3.20), 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:

$$-\nu \operatorname{div}((\operatorname{grad}\mathbf{u}) + (\operatorname{grad}\mathbf{u})^T) + \mathbf{u}\cdot\operatorname{grad}\mathbf{u} + \operatorname{grad}p = \mathbf{f} - \boldsymbol{\xi} \quad \text{in } \Omega,$$

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

$$-p\mathbf{n} + \nu((\operatorname{grad}\mathbf{u}) + (\operatorname{grad}\mathbf{u})^T)\cdot\mathbf{n} + \nu\alpha\mathbf{u} = \mathbf{h} - \boldsymbol{\xi} \quad \text{on } \Gamma,$$

$$(5.5) \quad -\nu\operatorname{div}((\operatorname{grad}\boldsymbol{\xi}) + (\operatorname{grad}\boldsymbol{\xi})^T) + \boldsymbol{\xi}\cdot(\operatorname{grad}\mathbf{u})^T - \mathbf{u}\cdot\operatorname{grad}\boldsymbol{\xi} + \operatorname{grad}\sigma$$
$$= -\boldsymbol{\xi} - \mathbf{u}\cdot\operatorname{grad}\mathbf{u} \quad \text{in } \Omega,$$

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

and

$$-\sigma\mathbf{n} + \nu((\operatorname{grad}\boldsymbol{\xi}) + (\operatorname{grad}\boldsymbol{\xi})^T)\cdot\mathbf{n} + \nu\alpha\boldsymbol{\xi} = \mathbf{h} - \boldsymbol{\xi} - (\mathbf{u}\cdot\mathbf{n})\boldsymbol{\xi} \quad \text{on } \Gamma.$$

Existence and regularity results for this optimality system may be derived in the same manner as that employed in §§3.1 and 3.3. Finite element approximations are defined exactly as in §4.1.

We again want to apply the results of §4.2. We define the spaces X and X^h as in §4.3. However, we now define

$$Y = (\mathbf{H}^1(\Omega))^* \times \mathbf{H}^{-1/2}(\Gamma) \times (\mathbf{H}^1(\Omega))^* \times \mathbf{H}^{-1/2}(\Gamma),$$
$$Z = \mathbf{L}^{3/2}(\Omega) \times \mathbf{L}^2(\Gamma) \times \mathbf{L}^{3/2}(\Omega) \times \mathbf{L}^2(\Gamma).$$

Also, the operator $T \in \mathscr{L}(Y; X)$ is now defined in the following manner: $T(\boldsymbol{\zeta}, \boldsymbol{\theta}, \boldsymbol{\eta}, \boldsymbol{\delta}) = (\mathbf{u}, p, \boldsymbol{\xi}, \sigma)$ for $(\boldsymbol{\zeta}, \boldsymbol{\theta}, \boldsymbol{\eta}, \boldsymbol{\delta}) \in Y$, and $(\mathbf{u}, p, \boldsymbol{\xi}, \sigma) \in X$, if and only if

$$a(\mathbf{u}, \mathbf{v}) + b(\mathbf{v}, p) + (\alpha\mathbf{u}, \mathbf{v})_\Gamma = (\boldsymbol{\zeta}, \mathbf{v}) + (\boldsymbol{\theta}, \mathbf{v})_\Gamma \quad \forall\mathbf{v} \in \mathbf{H}^1(\Omega),$$

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

$$a(\boldsymbol{\omega}, \boldsymbol{\xi}) + b(\boldsymbol{\omega}, \sigma) + (\alpha\boldsymbol{\omega}, \boldsymbol{\xi})_\Gamma = (\boldsymbol{\eta}, \boldsymbol{\omega}) + (\boldsymbol{\delta}, \boldsymbol{\omega})_\Gamma \quad \forall\boldsymbol{\omega} \in \mathbf{H}^1(\Omega),$$

and

$$b(\boldsymbol{\xi}, \psi) = 0 \quad \forall\psi \in L^2(\Omega).$$

The definition of the operator T^h is changed from that given by (4.23)–(4.26) in an analogous manner. The mapping G is now defined as follows: $G(\lambda, (\mathbf{u}, p, \boldsymbol{\xi}, \sigma)) = (\boldsymbol{\zeta}, \boldsymbol{\theta}, \boldsymbol{\eta}, \boldsymbol{\delta})$ for $\lambda \in \Lambda$, $(\mathbf{u}, p, \boldsymbol{\xi}, \sigma) \in X$, and $(\boldsymbol{\zeta}, \boldsymbol{\theta}, \boldsymbol{\eta}, \boldsymbol{\delta}) \in Y$, if and only if

$$(\boldsymbol{\zeta}, \mathbf{v}) = \lambda c(\mathbf{u}, \mathbf{u}, \mathbf{v}) - \lambda(\mathbf{f} - \boldsymbol{\xi}, \mathbf{v}) \quad \forall\mathbf{v} \in \mathbf{H}^1(\Omega),$$

$$(5.6) \quad (\boldsymbol{\theta}, \mathbf{w})_\Gamma = \lambda(\boldsymbol{\xi}, \mathbf{w})_\Gamma \quad \forall\mathbf{w} \in \mathbf{H}^1(\Omega),$$

$$(\boldsymbol{\eta}, \boldsymbol{\omega}) = \lambda c(\boldsymbol{\omega}, \mathbf{u}, \boldsymbol{\xi}) + \lambda c(\mathbf{u}, \mathbf{u}, \boldsymbol{\omega}) + \lambda(\boldsymbol{\xi}, \boldsymbol{\omega}) \quad \forall\boldsymbol{\omega} \in \mathbf{H}^1(\Omega),$$

and

$$(\boldsymbol{\delta}, \boldsymbol{\chi})_\Gamma = -\lambda(\mathbf{h} - \boldsymbol{\xi}, \boldsymbol{\chi})_\Gamma \quad \forall \boldsymbol{\chi} \in \mathbf{H}^1(\Omega).$$

With these changes in definitions, Theorems 4.3 and 4.4 and Corollary 4.5 can be shown to be applicable to the case of the optimization of the functional (5.1).

The main effect of the substitution of (2.1) into the right-hand side of (5.2) is to replace $\nu a(\mathbf{u}, \boldsymbol{\omega}) - (\mathbf{f}, \boldsymbol{\omega})$ in favor of the form $-c(\mathbf{u}, \mathbf{u}, \boldsymbol{\omega})$, i.e., to have, in the right-hand side of (5.5), $-(\mathbf{u} \cdot \operatorname{grad} \mathbf{u})$ plus some additional benign terms instead of $-\nu \operatorname{div}((\operatorname{grad} \mathbf{u}) + (\operatorname{grad} \mathbf{u})^T) - \mathbf{f}$. This replacement is necessary in order to validate the analyses of §4.2 for the present case.

Let us see what would happen if we did not substitute (2.1). The definition of the mapping G would then have

$$(5.7) \quad \begin{aligned} (\boldsymbol{\eta}, \boldsymbol{\omega}) &= \lambda c(\boldsymbol{\omega}, \mathbf{u}, \boldsymbol{\xi}) + \lambda c(\mathbf{u}, \boldsymbol{\omega}, \boldsymbol{\xi}) + \lambda a(\mathbf{u}, \boldsymbol{\omega}) + \lambda(\alpha \mathbf{u}, \boldsymbol{\omega})_\Gamma \\ &\qquad\qquad\qquad\qquad\qquad\qquad\qquad \forall \boldsymbol{\omega} \in \mathbf{H}^1(\Omega), \end{aligned}$$

replacing (5.6). In this case, one cannot show that the mapping G, which is defined as a mapping from $\Lambda \times X$ into Y, actually maps $\Lambda \times X$ into $Z \subset Y$, with a compact imbedding. To see this, note that the right-hand side of (5.6) becomes, if one uses (5.7),

$$-\nu \operatorname{div}((\operatorname{grad} \mathbf{u}) + (\operatorname{grad} \mathbf{u})^T)$$

and, for $\mathbf{u} \in \mathbf{H}^1(\Omega)$, this right-hand side is merely in $(\mathbf{H}^1(\Omega))^*$, and thus $\boldsymbol{\eta}$ computed by using the mapping G will only belong to $(\mathbf{H}^1(\Omega))^*$, i.e., G does not map $\Lambda \times X$ into Z, where Z is compactly imbedded into Y.

5.2. Optimizing with distributed controls with specified boundary data. Consider the problem of minimizing either of the functionals

$$(5.8) \quad \widetilde{\mathscr{F}}(\mathbf{u}, \mathbf{g}_d) = \tfrac{1}{4}\|\mathbf{u} - \mathbf{u}_0\|^4_{\mathbf{L}^4(\Omega)} + \tfrac{1}{2}\|\mathbf{g}_d\|^2_0$$

or

$$(5.9) \quad \widetilde{\mathscr{K}}(\mathbf{u}, \mathbf{g}_d) = \tfrac{\nu}{2}a(\mathbf{u}, \mathbf{u}) - (\mathbf{f}, \mathbf{u}) + \tfrac{1}{2}\|\mathbf{g}_d\|^2_0,$$

where $(\mathbf{u}, p, \mathbf{g}_d)$ are constrained to satisfy the Navier-Stokes equations

$$(5.10) \quad \nu a(\mathbf{u}, v) + c(\mathbf{u}, \mathbf{u}, v) + b(v, p) = (\mathbf{f} + \mathbf{g}_d, v) \quad \forall v \in \mathbf{H}^1_0(\Omega)$$

and

$$(5.11) \quad b(\mathbf{u}, q) = 0 \quad \forall q \in L^2_0(\Omega)$$

and boundary condition

$$(5.12) \quad \mathbf{u} = \mathbf{h} \quad \text{on } \Gamma \text{ with } \int_\Gamma \mathbf{h} \cdot \mathbf{n} \, d\Gamma = 0,$$

where

$$\mathbf{H}^1_0(\Omega) = \{\mathbf{v} \in \mathbf{H}^1(\Omega) : \mathbf{v} = 0 \text{ on } \Gamma\}$$

and

$$L_0^2(\Omega) = \left\{ q \in L^2(\Omega) \colon \int_\Omega q \, d\Omega = 0 \right\}.$$

The system (5.10)–(5.12) is a weak formulation of the problem

$$-\nu \operatorname{div}((\operatorname{grad} \mathbf{u}) + (\operatorname{grad} \mathbf{u})^T) + \mathbf{u} \cdot \operatorname{grad} \mathbf{u} + \operatorname{grad} p = \mathbf{f} + \mathbf{g}_d \quad \text{in } \Omega,$$
$$\operatorname{div} \mathbf{u} = 0 \quad \text{in } \Omega,$$

and

$$\mathbf{u} = \mathbf{h} \quad \text{on } \Gamma \text{ with } \int_\Gamma \mathbf{h} \cdot \mathbf{n} \, d\Gamma = 0.$$

Thus, our only control is of the distributed type and our boundary data is the specification of the velocity field. If the given data satisfies

$$\mathbf{h} \in \mathbf{H}_n^{1/2}(\Gamma) = \left\{ \mathbf{g} \in \mathbf{H}^{1/2}(\Gamma) \colon \int_\Gamma \mathbf{h} \cdot \mathbf{n} \, d\Gamma = 0 \right\}$$

and $\mathbf{f} \in \mathbf{L}^2(\Omega)$, we can proceed to show the existence of optimal solutions for either of the functionals (5.8) or (5.9). The control $\mathbf{g}_d$ may be chosen from $\mathbf{L}^2(\Omega)$.

With some changes, the approach of §3.1 for showing that Lagrange multipliers exist can be used in the present context. Specifically, we now define the spaces $B_1 = \mathbf{H}^1(\Omega) \times L_0^2(\Omega) \times \mathbf{L}^2(\Omega)$ and $B_2 = (\mathbf{H}_0^1(\Omega))^* \times L_0^2(\Omega) \times \mathbf{H}_n^{1/2}(\Gamma)$; the nonlinear mapping $M : B_1 \to B_2$ is now defined as follows: $M(\mathbf{u}, p, \mathbf{g}_d) = (\mathbf{f}, \varphi, \mathbf{h})$ for $(\mathbf{u}, p, \mathbf{g}_d) \in B_1$ and $(\mathbf{f}, \varphi, \mathbf{h}) \in B_2$, if and only if

$$\nu a(\mathbf{u}, \mathbf{v}) + c(\mathbf{u}, \mathbf{u}, \mathbf{v}) + b(\mathbf{v}, p) - (\mathbf{g}_d, \mathbf{v}) = (\mathbf{f}, \mathbf{v}) \quad \forall \mathbf{v} \in \mathbf{H}_0^1(\Omega),$$
$$b(\mathbf{u}, q) = (\varphi, q) \quad \forall q \in L_0^2\Omega,$$

and

$$\mathbf{u} = \mathbf{h} \quad \text{on } \Gamma \text{ with } \int_\Gamma \mathbf{h} \cdot \mathbf{n} \, d\Gamma = 0.$$

Similarly, the operator $M'(\hat{\mathbf{u}}, \hat{p}, \hat{\mathbf{g}}_d) \in \mathscr{L}(B_1; B_2)$ is now defined as follows: $M'(\hat{\mathbf{u}}, \hat{p}, \hat{\mathbf{g}}_d) \cdot (\mathbf{w}, r, \mathbf{s}_d) = (\bar{\mathbf{f}}, \bar{\varphi}, \bar{\mathbf{h}})$ for $(\mathbf{w}, r, \mathbf{s}_d) \in B_1$ and $(\bar{\mathbf{f}}, \bar{\varphi}, \bar{\mathbf{h}}) \in B_2$, if and only if

$$\nu a(\mathbf{w}, \mathbf{v}) + c(\mathbf{w}, \hat{\mathbf{u}}, \mathbf{v}) + c(\hat{\mathbf{u}}, \mathbf{w}, \mathbf{v}) + b(\mathbf{v}, r) - (\mathbf{s}_d, \mathbf{v}) = (\bar{\mathbf{f}}, \mathbf{v}) \quad \forall \mathbf{v} \in \mathbf{H}_0^1(\Omega),$$
$$b(\mathbf{w}, q) = (\bar{\varphi}, q) \quad \forall q \in L_0^2(\Omega),$$

and

$$\mathbf{w} = \bar{\mathbf{h}} \quad \text{on } \Gamma \text{ with } \int_\Gamma \bar{\mathbf{h}} \cdot \mathbf{n} \, d\Gamma = 0.$$

To show that the operator $M'(\hat{\mathbf{u}}, \hat{p}, \hat{\mathbf{g}}_d)$ from B_1 into B_2 is onto, first note that there exists (see [6]) $(\mathbf{w}, r) \in \mathbf{H}^1(\Omega) \times L_0^2(\Omega)$ such that

$$\nu a(\mathbf{w}, \mathbf{v}) + b(\mathbf{v}, r) + c(\hat{\mathbf{u}}, \mathbf{w}, \mathbf{v}) = (\bar{\mathbf{f}}, \mathbf{v}) \quad \forall \mathbf{v} \in \mathbf{H}_0^1(\Omega),$$
$$b(\mathbf{w}, q) = (\bar{\varphi}, q) \quad \forall q \in L_0^2(\Omega),$$

and

$$\mathbf{w} = \overline{\mathbf{h}} \quad \text{on } \Gamma \text{ with } \int_{\Gamma} \overline{\mathbf{h}} \cdot \mathbf{n} \, d\Gamma = 0.$$

Then, there is no difficulty in finding $\mathbf{s}_d \in \mathbf{L}^2(\Omega)$ such that

$$(\mathbf{s}_d, \mathbf{v}) = c(\mathbf{w}, \hat{\mathbf{u}}, \mathbf{v}) \quad \forall \mathbf{v} \in H_0^1(\Omega).$$

Indeed, we merely set $\mathbf{s}_d = \mathbf{w} \cdot \text{grad}\, \hat{\mathbf{u}}$, so that, since $\mathbf{w} \in \mathbf{H}^1(\Omega)$ and $\hat{\mathbf{u}} \in \mathbf{H}^{3/2}(\Omega)$ (see Remark 2 at the end of §2), we have that $\mathbf{s}_d \in \mathbf{L}^2(\Omega)$, and the last equation is satisfied. Collecting the defining relations for $\mathbf{w}$, r, and $\mathbf{s}_d$ implies that the operator $M'(\hat{\mathbf{u}}, \hat{p}, \hat{\mathbf{g}}_d)$ from B_1 into B_2 is onto. The remainder of the proof of the existence of Lagrange multipliers proceeds in a similar manner to the proof of Theorem 3.1.

The optimality system for the problem of minimizing (5.8) subject to (5.10)-(5.12) is given by

(5.13) $-\nu \, \text{div}((\text{grad}\, \mathbf{u}) + (\text{grad}\, \mathbf{u})^T) + \mathbf{u} \cdot \text{grad}\, \mathbf{u} + \text{grad}\, p = \mathbf{f} - \boldsymbol{\xi} \quad \text{in } \Omega,$

(5.14) $\text{div}\, \mathbf{u} = 0 \quad \text{in } \Omega,$

(5.15) $\mathbf{u} = \mathbf{h} \quad \text{on } \Gamma,$

(5.16) $-\nu \, \text{div}((\text{grad}\, \boldsymbol{\xi}) + (\text{grad}\, \boldsymbol{\xi})^T) - \mathbf{u} \cdot \text{grad}\, \boldsymbol{\xi} + \boldsymbol{\xi} \cdot (\text{grad}\, \mathbf{u})^T + \text{grad}\, \sigma$
$= (\mathbf{u} - \mathbf{u}_0)^3 \quad \text{in } \Omega$

(5.17) $\text{div}\, \boldsymbol{\xi} = 0 \quad \text{in } \Omega,$

and

(5.18) $\boldsymbol{\xi} = \mathbf{0} \quad \text{on } \Gamma.$

If instead we minimize (5.9), the optimality system is given by these same equations except that (5.16) is replaced by

(5.19) $-\nu \, \text{div}((\text{grad}\, \boldsymbol{\xi}) + (\text{grad}\, \boldsymbol{\xi})^T) - \mathbf{u} \cdot \text{grad}\, \boldsymbol{\xi} + \boldsymbol{\xi} \cdot (\text{grad}\, \mathbf{u})^T + \text{grad}\, \sigma$
$= -\boldsymbol{\xi} - \mathbf{u} \cdot \text{grad}\, \mathbf{u} \quad \text{in } \Omega.$

As always, the control may be determined from the condition $\mathbf{g}_d = -\boldsymbol{\xi}$ in Ω.

Regularity results for solutions of these optimality systems may be proved exactly as in §§3.3 and 5.1, provided, of course, that $\mathbf{f}$ and $\mathbf{h}$ have the appropriate smoothness. Likewise, finite element error estimates may be obtained as in §4, and the results of these sections apply to the cases in hand. The only changes are that the velocity test space is now $\mathbf{H}_0^1(\Omega)$ and the velocity trial set must be constrained so that (5.12) is satisfied; also, the pressure is now determined only up to an additive constant so that, e.g., we should constrain the pressure space so that its members have zero mean over Ω. Thus, the velocity finite element test space must be chosen to be a subspace of $\mathbf{H}_0^1(\Omega)$, and the velocity finite element trial set must be constrained so that some approximation to the boundary condition (5.12) is satisfied. Likewise, the pressure finite element test and

trial functions are constrained to have zero mean over Ω. This does not pose any theoretical or practical difficulties, since such procedures are well known in the context of Navier-Stokes calculations (see [6, 7]).

One may also consider the case of a distributed control acting in concert with specified Neumann data. Here new difficulties are encountered in showing the existence of Lagrange multipliers. In fact, we can no longer use the $\mathbf{L}^2(\Omega)$-norm of $\mathbf{g}_d$ in the definition of the functionals (5.8) and (5.9); we must employ a weaker norm. For example, if we use

$$\tfrac{3}{4}\|\mathbf{g}_d\|_{\mathbf{L}^{4/3}(\Omega)}^{4/3} \quad \text{instead of} \quad \tfrac{1}{2}\|\mathbf{g}_d\|_0^2$$

in (5.8) or (5.9), then we may show that suitable Lagrange multipliers exist. In the optimality system, the relation between the control $\mathbf{g}_d$ and the Lagrange multiplier $\boldsymbol{\xi}$ changes to $\mathbf{g}_d = -\boldsymbol{\xi}^3$, where the notation $\boldsymbol{\xi}^3$ denotes component-wise exponentiation. The differential equations in the optimality system are analogously changed. For example, in (5.13) we would replace the term $\boldsymbol{\xi}$ on the right-hand side with $\boldsymbol{\xi}^3$. Of course, the boundary conditions (5.15) and (5.18) are changed to Neumann conditions. The error estimates of Theorems 4.4 and 4.5 are still valid; however, for the approximate control $\mathbf{g}_d^h = -(\boldsymbol{\xi}^h)^3$, we can only show that $\|\mathbf{g}_d^h - \mathbf{g}_d\|_0 = O(h^m)$. See [11] for details.

5.3. Distributed controls acting on only part of the flow domain.

We now turn to the case where the distributed control acts on only a subset of the flow domain Ω. We separate this case from that of the Neumann control acting on only part of the boundary owing to the fact that in these cases we do not always achieve the same results as we have obtained so far. The results that we are about to present apply equally well to all three settings that have been treated above, i.e., that of §§2–4, of §5.1, or of §5.2. However, for the sake of simplicity, we will present them in the setting of §5.2.

Let ω be an open set of Ω whose closure is strictly contained in Ω. Consider the functionals $\widetilde{\mathcal{F}}(\mathbf{u}, \mathbf{g}_d)$ and $\widetilde{\mathcal{K}}(\mathbf{u}, \mathbf{g}_d)$ defined in §5.2 and the sets

$$\widetilde{\mathcal{U}}_{\text{ad}} = \{(\mathbf{u}, \mathbf{g}_d) \in \mathbf{H}^1(\Omega) \times \mathbf{L}_\omega^2(\Omega) : \widetilde{\mathcal{F}}(\mathbf{u}, \mathbf{g}_d) < \infty, \text{ and there exists}$$
$$\text{a } p \in L_0^2(\Omega) \text{ such that (5.10)–(5.12) are satisfied}\}$$

and

$$\widetilde{\mathcal{V}}_{\text{ad}} = \{(\mathbf{u}, \mathbf{g}_d) \in \mathbf{H}^1(\Omega) \times L_0^2(\Omega) \times \mathbf{L}_\omega^2(\Omega) : \widetilde{\mathcal{K}}(\mathbf{u}, \mathbf{g}_d) < \infty, \text{ and there exists}$$
$$\text{a } p \in L_0^2(\Omega) \text{ such that (5.10)–(5.12) are satisfied}\},$$

where

$$\mathbf{L}_\omega^2(\Omega) = \{\mathbf{v} \in \mathbf{L}^2(\Omega) : \mathbf{v} = \mathbf{0} \text{ in } \Omega\backslash\omega\}.$$

Now consider the problem of minimizing $\widetilde{\mathcal{F}}(\cdot, \cdot)$ over $\widetilde{\mathcal{U}}_{\text{ad}}$. The existence of optimal solutions can be shown in the same manner as was used in §2. Using

Lagrange multipliers, one may derive the optimality system (5.13)–(5.18), with (5.13) replaced by

$$(5.20) \qquad -\nu \operatorname{div}((\operatorname{grad} \mathbf{u}) + (\operatorname{grad} \mathbf{u})^T) + \mathbf{u} \cdot \operatorname{grad} \mathbf{u} + \operatorname{grad} p = \mathbf{f} - \overline{\boldsymbol{\xi}} \quad \text{in } \Omega,$$

where

$$\overline{\boldsymbol{\xi}} = \begin{cases} \boldsymbol{\xi} & \text{in } \boldsymbol{\omega}, \\ 0 & \text{in } \Omega \backslash \boldsymbol{\omega}. \end{cases}$$

The control is given by $\mathbf{g}_d = -\boldsymbol{\xi}$ in $\boldsymbol{\omega}$.

For the problem of minimizing $\widetilde{\mathcal{K}}(\cdot, \cdot)$ over $\widetilde{\mathcal{V}}_{\mathrm{ad}}$, the optimality system is given by (5.14), (5.15), (5.17), (5.18), (5.20), and

$$-\nu \operatorname{div}((\operatorname{grad} \boldsymbol{\xi}) + (\operatorname{grad} \boldsymbol{\xi})^T) - \mathbf{u} \cdot \operatorname{grad} \boldsymbol{\xi} + \boldsymbol{\xi} \cdot (\operatorname{grad} \mathbf{u})^T + \operatorname{grad} \sigma$$
$$= -\overline{\boldsymbol{\xi}} - \mathbf{u} \cdot \operatorname{grad} \mathbf{u} \quad \text{in } \Omega.$$

The proof of Theorem 3.1 showing the existence of Lagrange multipliers cannot be extended to the present case. The difficulty arises in showing that an $\mathbf{s}_d$ exists satisfying (3.6), since now $\mathbf{s}_d$ vanishes outside $\boldsymbol{\omega}$. However, using methods similar to those used in [9], one can show that indeed nonzero Lagrange multipliers exist. We also have to worry about a loss of regularity owing to the fact that the control $\mathbf{g}_d$ cannot be any smoother than an $\mathbf{H}^{1/2-\varepsilon}(\Omega)$-function. Thus, one can in general only show that, regardless of how smooth is the domain or the data, $\mathbf{u} \in \mathbf{H}^{5/2-\varepsilon}(\Omega)$, $p \in H^{3/2-\varepsilon}(\Omega)$, $\boldsymbol{\xi} \in \mathbf{H}^{5/2-\varepsilon}(\Omega)$, and $\sigma \in H^{3/2-\varepsilon}(\Omega)$, where $\varepsilon \in (0, 1/2)$ is arbitrary. As a result, the best finite element error estimate that can be obtained, using, for example, piecewise polynomials of degree two or higher for the velocity, and linear polynomials for the pressure, is

$$\|\mathbf{u}(\lambda) - \mathbf{u}^h(\lambda)\|_1 + \|p^h(\lambda) - p^h(\lambda)\|_0 + \|\boldsymbol{\xi}(\lambda) - \boldsymbol{\xi}^h(\lambda)\|_1 + \|\sigma(\lambda) - \sigma^h(\lambda)\|_0$$
$$= O(h^{3/2-\varepsilon}).$$

It is possible, by using special techniques that explicitly take into account the boundary of the subdomain $\boldsymbol{\omega}$ in the definition of the method, that this estimate can be improved.

5.4. Neumann controls acting on only part of the boundary. The final variation we consider is having the boundary control (1.3) acting on only part of the boundary Γ. For simplicity we assume that this is the only control acting on the problem; all we have to say holds equally well if there is also available a distributed control.

We divide Γ into two parts, Γ_o and Γ_c, such that $\overline{\Gamma}_o \cup \overline{\Gamma}_c = \overline{\Gamma}$ and $\Gamma_o \cap \Gamma_c = \varnothing$. Then our minimization problems will require the minimization of either of the functionals

$$(5.21) \qquad \mathcal{F}(\mathbf{u}, \mathbf{g}_d) = \tfrac{1}{4} \|\mathbf{u} - \mathbf{u}_0\|_{\mathbf{L}^4(\Omega)}^4 + \tfrac{1}{2} \|\mathbf{g}_b\|_{0,\Gamma_c}^2$$

or

$$(5.22) \qquad \overline{\mathcal{K}}(\mathbf{u}, \mathbf{g}_d) = \tfrac{\nu}{2} a(\mathbf{u}, \mathbf{u}) + \tfrac{\nu}{2}(\alpha \mathbf{u}, \mathbf{u})_\Gamma - (\mathbf{f}, \mathbf{u}) + \tfrac{1}{2} \|\mathbf{g}_b\|_{0,\Gamma_c}^2,$$

where $\|\mathbf{g}_b\|_{0,\Gamma_c}^2$ has the obvious definition. Candidate minimizers are required to satisfy

$$(5.23) \quad -\nu \operatorname{div}((\operatorname{grad}\mathbf{u}) + (\operatorname{grad}\mathbf{u})^T) + \mathbf{u}\cdot\operatorname{grad}\mathbf{u} + \operatorname{grad}p = \mathbf{f} \quad \text{in } \Omega,$$

$$(5.24) \quad \operatorname{div}\mathbf{u} = 0 \quad \text{in } \Omega,$$

$$(5.25) \quad -p\mathbf{n} + \nu(\operatorname{grad}\mathbf{u} + \operatorname{grad}\mathbf{u}^T)\cdot\mathbf{n} + \nu\alpha\mathbf{u} = \mathbf{h} + \mathbf{g}_b \quad \text{on } \Gamma_c,$$

and

$$(5.26) \quad -p\mathbf{n} + \nu(\operatorname{grad}\mathbf{u} + \operatorname{grad}\mathbf{u}^T)\cdot\mathbf{n} + \nu\alpha\mathbf{u} = \mathbf{h} \quad \text{on } \Gamma_o,$$

where, as always, $\mathbf{f}$, $\mathbf{h}$, and $\mathbf{u}_0$ are given functions in appropriate function spaces. The control is to be chosen from $\mathbf{L}^2(\Gamma_c)$.

The existence of optimal solutions can be shown as in §2. Using Lagrange multipliers, one easily derives an optimality system. For the minimization of the functional (5.21), this optimality system is given by

$$-\nu \operatorname{div}((\operatorname{grad}\boldsymbol{\xi}) + (\operatorname{grad}\boldsymbol{\xi})^T) - \mathbf{u}\cdot\operatorname{grad}\boldsymbol{\xi} + \boldsymbol{\xi}\cdot(\operatorname{grad}\mathbf{u})^T + \operatorname{grad}\sigma$$

$$(5.27) \qquad = (\mathbf{u} - \mathbf{u}_0)^3 \quad \text{in } \Omega,$$

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

$$(5.28) \quad -\sigma\mathbf{n} + \nu(\operatorname{grad}\boldsymbol{\xi} + \operatorname{grad}\boldsymbol{\xi}^T)\cdot\mathbf{n} + \nu\alpha\boldsymbol{\xi} = -(\mathbf{u}\cdot\mathbf{n})\boldsymbol{\xi} \quad \text{on } \Gamma$$

and (5.23)–(5.26), where in (5.25) we may eliminate the control $\mathbf{g}_b$ through the use of the relation $\mathbf{g}_b = -\boldsymbol{\xi}|_{\Gamma_c}$. For the minimization of the functional (5.22), we obtain the same optimality system except that (5.27) and (5.28) are respectively replaced by

$$-\nu \operatorname{div}((\operatorname{grad}\boldsymbol{\xi}) + (\operatorname{grad}\boldsymbol{\xi})^T) - \mathbf{u}\cdot\operatorname{grad}\boldsymbol{\xi} + \boldsymbol{\xi}\cdot(\operatorname{grad}\mathbf{u})^T + \operatorname{grad}\sigma$$

$$= -\mathbf{u}\cdot\operatorname{grad}\mathbf{u} \quad \text{in } \Omega$$

and

$$-\sigma\mathbf{n} + \nu(\operatorname{grad}\boldsymbol{\xi} + \operatorname{grad}\boldsymbol{\xi}^T)\cdot\mathbf{n} + \nu\alpha\boldsymbol{\xi} = -(\mathbf{u}\cdot\mathbf{n})\boldsymbol{\xi} + \mathbf{h} - \bar{\boldsymbol{\xi}} \quad \text{on } \Gamma,$$

where

$$\bar{\boldsymbol{\xi}} = \begin{cases} \boldsymbol{\xi} & \text{on } \Gamma_c, \\ \mathbf{0} & \text{on } \Gamma_o. \end{cases}$$

If the boundary segments Γ_o and Γ_c are disjoint, e.g., they are the separate parts of the boundary of a doubly connected region, then there is no difficulty in extending the existence, regularity, and approximation results of §§3 and 4 to the present setting, again with the exception that Theorem 3.1 must be proved by the methods of [9]. However, if these boundary segments are not disjoint, we have to worry about a loss of regularity at the interfaces adjoining the boundary segments. In fact, one can in general only show that, regardless of how smooth is the domain or the data, $\mathbf{u} \in \mathbf{H}^{2-\varepsilon}(\Omega)$, $p \in H^{1-\varepsilon}(\Omega)$, $\boldsymbol{\xi} \in \mathbf{H}^{2-\varepsilon}(\Omega)$, and

$\sigma \in H^{1-\varepsilon}(\Omega)$, where $\varepsilon \in (0, 1/2)$ is arbitrary. As a result, the best finite element error estimate that can be obtained is

$$\|\mathbf{u}(\lambda) - \mathbf{u}^h(\lambda)\|_1 + \|p(\lambda) - p^h(\lambda)\|_0 + \|\boldsymbol{\xi}(\lambda) - \boldsymbol{\xi}^h(\lambda)\|_1 + \|\sigma(\lambda) - \sigma^h(\lambda)\|_0 = O(h^{1-\varepsilon}).$$

We repeat that if the boundary segments Γ_o and Γ_c are disjoint, then we obtain the same type of estimates as in §4.

BIBLIOGRAPHY

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*, The Mathematical Foundations of the Finite Element Method with Applications to Partial Differential Equations (A. Aziz, ed.), Academic Press, New York, 1972, pp. 3–359.

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

4. P. Ciarlet, *The finite element method for elliptic problems*, North-Holland, Amsterdam, 1978.

5. M. Crouzeix and J. Rappaz, *On numerical approximation in bifurcation theory*, Masson, Paris, 1990.

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, 1989.

8. M. Gunzburger, L. Hou, and T. Svobodny, *Boundary velocity control of incompressible flow with an application to viscous drag reduction*, SIAM J. Control Optim. (to appear).

9. ____, *Analysis and finite element approximation of optimal control problems for the stationary Navier-Stokes equations with Dirichlet controls*, Math. Modelling Numer. Anal. (to appear).

10. ____, *Numerical approximation of an optimal control problem associated with the Navier-Stokes equations*, Appl. Math. Lett. **2** (1989), 29–31.

11. L. Hou, *Analysis and finite element approximation of some optimal control problems associated with the Navier-Stokes equations*, Ph.D. Thesis, Carnegie Mellon University, Pittsburgh, 1989.

12. J.-L. Lions, *Control of distributed singular systems*, Bordas, Paris, 1985.

13. J. Serrin, *Mathematical principles of classical fluid mechanics*, Handbuch der Physik VIII/1 (S. Flügge and C. Truesdell, eds.), Springer, Berlin, 1959, pp. 125–263.

14. R. Temam, *Navier-Stokes equations*, North-Holland, Amsterdam, 1979.

15. ____, *Navier-Stokes equations and nonlinear functional analysis*, SIAM, Philadelphia, PA, 1983.

16. R. Verfürth, *Finite element approximation of incompressible Navier-Stokes equations with slip boundary condition*, Numer. Math. **50** (1987), 697–721.

DEPARTMENT OF MATHEMATICS, VIRGINIA POLYTECHNIC INSTITUTE AND STATE UNIVERSITY, BLACKSBURG, VIRGINIA 24061
E-mail address: gunzburger@vtcc1.bitnet

DÉPARTEMENT DE MATHÉMATIQUES ET DE STATISTIQUE, UNIVERSITÉ LAVAL, QUÉBEC, G1K 7P4, CANADA

DEPARTMENT OF MATHEMATICS AND STATISTICS, WRIGHT STATE UNIVERSITY, DAYTON, OHIO 45435