SIAM J. CONTROL AND OPTIMIZATION
Vol. 32, No. 5, pp. 1428–1446, September 1994

010

OPTIMAL CONTROLS OF NAVIER–STOKES EQUATIONS*

MIHIR DESAI† AND KAZUFUMI ITO†

Abstract. This paper studies optimal control problems of the fluid flow governed by the Navier–Stokes equations. Two control problems are formulated in the case of the driven cavity and flow through a channel with sudden expansion and solved successfully using a numerical optimization algorithm based on the augmented Lagrangian method. Existence and the first-order optimality condition of the optimal control are established. A convergence result on the augmented Lagrangian method for nonsmooth cost functional is obtained.

Key words. flow control, Navier–Stokes equation, augmented Lagrangian method

AMS subject classifications. 35Q10, 49B22, 49D29

1. Introduction. Considerable progress has been made in mathematical analysis and computation of Navier–Stokes equations. However, very little attention has been given to the question of controlling the Navier–Stokes equations with the exception of experiments (e.g., Nagib, Reisenthal, and Koga [NRK]) and the problem of optimal shape design for drag minimization (e.g., Pironneau [Pi]). An abstract formulation of a control problem does exist in literature (e.g., [Li]). However (to our knowledge[1]) there has been very little effort toward mathematical formulation and computation in the context of control of physical flow situations. As a step in this direction we formulate and solve computationally the steady-state control problem for two flows that have been investigated extensively in numerical computations; i.e., the driven cavity and flow through a channel with sudden expansion. Finding a suitable cost functional that is relevant to the physics of the flow is a very important step in formulating control problems in both the driven cavity and channel flow. The control problems, which are formulated in §2, involve only one-dimensional control input acting through a part of the boundary as Drichlet boundary control. The influence of a one-parameter control input is limited (which can be seen in our numerical calculations) and thus our numerical calculations have been carried out for relatively low Reynolds number flow. Also, some of the hypotheses for our analysis can be verified only for relatively low Reynolds number flow. To improve the performance of our control law we must consider a full boundary control through a part of the boundary. Our analysis can be extended to treat such a problem and will be discussed in a forthcoming paper.

The paper is organized as follows. In §2 two optimal control problems are described. In §3 the basic theory of Navier–Stokes equation is given. In §4 the existence and first-order optimality condition for optimal control problems are established. In §5 a solution technique based on the augmented Lagrangian method is described. Convergence of the augmented Lagrangian method is obtained. Finally some of our numerical findings are reported in §6.

* Received by the editors January 20, 1992; accepted for publication (in revised form) April 7, 1993. This research was supported in part by Air Force Office of Scientific Research grant AFOSR-90-0091 and by National Science Foundation grant DMS-8818530.

† Center for Applied Mathematical Sciences, University of Southern California, Los Angeles, CA 90089-1113.

[1] Since this paper was submitted, there has been increasing interest in the study of control of the Navier–Stokes equations, e.g., [AT], [FS], [GHS], [DG], and the references therein.

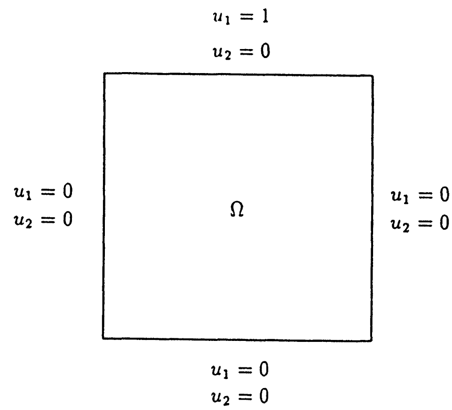

FIG. 1. *Driven cavity.*

2. Two flow problems. In this section two flow control problems will be formulated.

2.1. Driven cavity. Consider the two-dimensional motion of fluid modeled by the (stationary) Navier–Stokes equation,

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

$$\nabla \cdot u = 0, \tag{2.2}$$

confined in a square cavity Ω, depicted in Fig. 1. Here $u = (u_1, u_2)$ is the velocity field, p the pressure, ν the kinematic viscosity of the fluid ($\nu = 1/\text{Re}$, where Re is the Reynolds number), and f the density of external forces (in this example $f = 0$). The nonlinear term $(u \cdot \nabla)u$ in (2.1) (often called the convective term), is a symbolic notation for the vector

$$\left(u_1 \frac{\partial u_1}{\partial x_1} + u_2 \frac{\partial u_1}{\partial x_2},\ u_1 \frac{\partial u_2}{\partial x_1} + u_2 \frac{\partial u_2}{\partial x_2} \right).$$

The divergence-free condition (2.2) is the equation for law of conservation of mass.

Conventionally, the problem has been treated with boundary conditions as in Fig. 1; i.e., only the top surface moving with velocity U_{top}. However we observe (numerically) that if both the top and bottom surfaces move in the same direction, the flow separates into two distinct regions as shown in Fig. 2 (where the top and bottom velocities are .5 and the viscosity $\nu = 1/50$ is used). Hence, the control problem we consider is as follows.

PROBLEM. *Given the bottom velocity U_{bot}, find the top velocity U_{top} such that the separation of flow occurs at a desired horizontal line location Γ_L.*

We cast the problem as a minimization of cost functional defined for U_{top}

$$J(U_{\text{top}}) = \int_{\Gamma_L} |u_2|^2 \, ds, \tag{2.3}$$

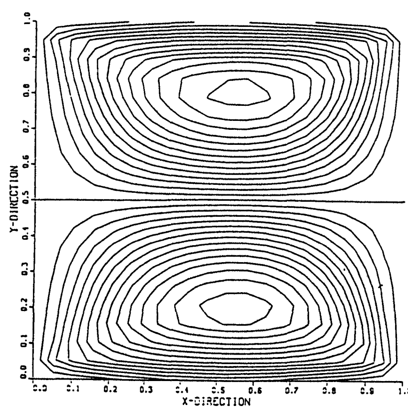

FIG. 2. *Seperation of flow in a cavity.*

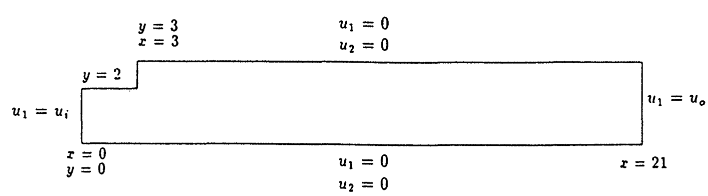

FIG. 3. *Channel with sudden expansion.*

subject to (2.1) and (2.2), where u_2 is the vertical component of the velocity field u and $u_2 = 0$ on Γ_L implies that no flow crosses the horizontal line Γ_L.

Remark 2.1. Note that the boundary condition in Fig. 1 is not in $H^{1/2}(\Gamma)$: i.e., it is not the trace of a function in $H^1(\Omega)$ on Γ. Thus, the existence theory (e.g., in [GR]; also see Theorem 3.5) cannot be applied in this case. However, we are able to show in [Di] that the Stokes equation (that is of the form (2.1), (2.2) without convective term) has a unique solution in $L^2(\Omega)^2$.

2.2. Channel flow. The second problem is a control of channel flow illustrated in Fig. 3. We assume that the inflow (at $x = 0$) and outflow (at $x = 21$) are parabolic (Poiseuille flow assumption) with $u_{\text{in}} = x_2(2 - x_2)$ and $u_{\text{out}} = c\,x_2(3 - x_2)$ where c is chosen such that $\int_\Gamma u \cdot n\, ds = 0$. There is a recirculation region in the corner whose size increases with the Reynolds number. Figure 4 qualitatively illustrate the flow in the corner with Re= 50. The objective is to shape the flow in the recirculation region to a desired configuration by means of controlled injection (suction) along a portion Γ of the vertical boundary facing the recirculation flow (see the shaded line in Fig. 4). The key question is then: What is a "*desirable*" *flow*? The answer to this question clearly depends on the applications in which the flow situation occurs. We consider the following two cost functionals. The first one corresponds to the total vorticity in the flow given by

$$\int_\Omega \left| \frac{\partial u_2}{\partial x_1} - \frac{\partial u_1}{\partial x_2} \right|^2 dx, \tag{2.4}$$

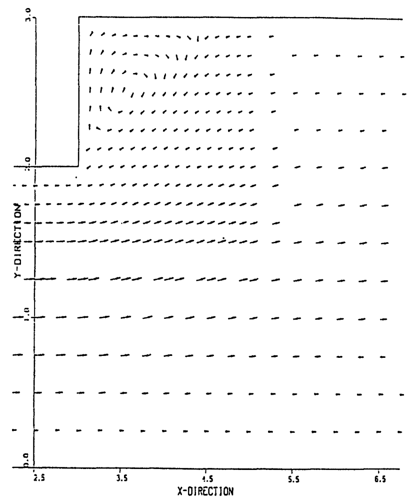

FIG. 4. *Channel flow for* Re $= 20$.

where the vorticity

$$\omega = \frac{\partial u_2}{\partial x_1} - \frac{\partial u_1}{\partial x_2}.$$

This cost is motivated by the fact that potential flows (zero vorticity) are frictionless and incur low energy dissipation. The second one is the "back flow" cost,

$$\int_\Omega (|\min(0,u_1)|^2 + |\min(0,u_2)|^2)\,dx, \tag{2.5}$$

which is motivated by the fact that if we do not have any recirculation, then $u_1 \geq 0$ and $u_2 \geq 0$ (i.e., fluid moving upward and to the right).

3. Basic theory for the Navier–Stokes equations. In this section we summarize the basic theory of the Navier–Stokes equation that we need for our discussions. We consider the boundary-value problems (2.1), (2.2) with the Dirichlet boundary condition $u|_\Gamma = g$ where $\int_\Gamma g \cdot n\,ds = 0$ since by Green's formula

$$\int_\Omega \nabla \cdot u\,dx = \int_\Gamma g \cdot n\,ds = 0. \tag{3.1}$$

Here n is the outward unit normal vector and it is assumed that Ω is bounded open set in R^n, $n = 2, 3$ and its boundary is at least Lipschitz continuous.

Define the function spaces and notation that will be used in what follows (our treatment and notation are along the lines presented in [GR] and [Te]):

$$V = \{u \in H_0^1(\Omega)^n \text{ with } \nabla \cdot u = 0\},$$

$$H = \{u \in L^2(\Omega)^n \text{ with } \nabla \cdot u = 0 \text{ and } u \cdot n = 0 \text{ on } \Gamma\}.$$

Let $a(u,v) : H^1(\Omega)^n \times H^1(\Omega)^n \to R$ be the symmetric sesquilinear form defined by

$$a(u,v) = \nu \int_\Omega \operatorname{grad} u \cdot \operatorname{grad} v \, dx \tag{3.2}$$

and define the bilinear form $c(u,p) : H^1(\Omega)^n \times L^2(\Omega) \to R$ by

$$c(u,p) = \int_\Omega (\nabla \cdot u) p \, dx \quad \text{for } u \in H^1(\Omega)^n \text{ and } p \in L^2(\Omega). \tag{3.3}$$

The trilinear form b on $H^1(\Omega)^n$ that corresponds to the convective term in (2.1) is defined by

$$b(u; v, w) = \int_\Omega \sum_{i,j=1}^n u_j \frac{\partial v_i}{\partial x_j} w_i \, dx. \tag{3.4}$$

Then the variational form for the Navier–Stokes equations (2.1), (2.2) with boundary condition $u|_\Gamma = g$ is given by

$$\begin{aligned} &a(u,v) + b(u;u,v) - c(v,p) = \langle f,\, v\rangle \quad \text{for all } v \in H_0^1(\Omega)^n, \\ &c(u,q) = 0 \quad \text{for all } q \in L^2(\Omega), \end{aligned} \tag{3.5}$$

where $\langle \cdot,\, \cdot \rangle$ denotes the dual product of $H^{-1}(\Omega)^n \times H_0^1(\Omega)^n$. Discarding the convective term in (2.1) results in the Stokes equation

$$-\nu \,\Delta u + \nabla p = f \quad \text{and} \quad \nabla u = 0 \tag{3.6}$$

with $u|_\Gamma = g$. Using our notation it can be written as

$$\begin{aligned} &a(u,v) - c(v,p) = \langle f,\, v\rangle \quad \text{for all } v \in H_0^1(\Omega)^n, \\ &c(u,q) = 0 \quad \text{for all } q \in L^2(\Omega). \end{aligned} \tag{3.7}$$

Note that (3.7) is the first-order necessary optimality condition for the following minimization problem:

$$\text{Minimize} \quad \tfrac{1}{2}\, a(u,u) - \langle f,\, u\rangle$$

$$\text{over } u \in H^1(\Omega)^n \text{ with } u|_\Gamma = g,$$

subject to $\nabla \cdot u = 0$. In fact, the Lagrangian corresponding to the constrained minimization stated above is

$$L(u,\lambda) = \tfrac{1}{2}\, a(u,u) - \langle f,\, u\rangle - c(u,\lambda),$$

where the Lagrange multiplier $\lambda \in L^2(\Omega)$ turns out to be the pressure p in (3.6). For the homogeneous boundary condition (i.e., $g = 0$) since

$$\operatorname{div} : H_0^1(\Omega) \to L_0^2(\Omega) = \left\{ \phi \in L^2(\Omega) : \int_\Omega \phi \, dx = 0 \right\}$$

is surjective, it follows from [GR] that there exists a unique solution $(u,p) \in V \times L_0^2(\Omega)$ of (3.7). Hence (3.7) is equivalently written as

$$a(u,\psi) = \langle f,\, \psi\rangle \quad \text{for all } \psi \in V. \tag{3.8}$$

We state the important property of the trilinear form b [GR].

LEMMA 3.1. *The trilinear form b defined by (3.4) is continuous on $(H^1(\Omega)^n)^3$. Let $v,\ w \in H^1(\Omega)^n$ and let $u \in H^1(\Omega)^n$ satisfy $\nabla \cdot u = 0$ and assume either $n \cdot u|_\Gamma = 0$ or $w \cdot v|_\Gamma = 0$. Then we have*

$$b(u; v, w) + b(u; w, v) = 0, \tag{3.9}$$

which implies in particular

$$b(u; v, v) = 0. \tag{3.10}$$

We now state the existence result and for the sake of completeness of our discussions we include a sketch of its proof.

THEOREM 3.2. *For $f \in H^{-1}(\Omega)^n$ and $g = 0$ there exists at least one solution $(u, p) \in V \times L^2_0(\Omega)$ of (3.5).*

Proof. It follows from (3.8) and [GR] that (3.5) is equivalent to an equation for $u \in V$:

$$a(u, \psi) + b(u; u, \psi) = \langle f, \psi \rangle \quad \text{for all } \psi \in V. \tag{3.11}$$

Thus, we argue the existence of solutions to (3.11). Define the map $T : V \to V$ by $z = T(u)$ where given $u \in V$, z is a unique solution to

$$a(z, \psi) + b(u; z, \psi) = \langle f, \psi \rangle \quad \text{for all } \psi \in V. \tag{3.12}$$

Existence and uniqueness of a solution to (3.12) can be shown by the Lax–Milgram theory [Yo] since the bilinear form $a(\phi, \psi) + b(u; \phi, \psi)$ defined on $V \times V$ is continuous and V-coercive by Lemma 3.1. The fixed points of T are the solutions of (3.11). Taking $\psi = z$ in (3.12) and from (3.10), we obtain $\|z\|_V \le \frac{1}{\nu}\,\|f\|_{V^*}$. Hence $T : C \to C$, where the set $C = \{w \in V : \|w\|_V \le \frac{1}{\nu}\,\|f\|_{V^*}\}$is a bounded, closed convex subset of V. Since V is a Hilbert space, every bounded set in V contains a weak convergent sequence, $u_n \to u$ in V. Since V is compactly embedded into $L^4(\Omega)^n$, the sequence converges strongly in $L^4(\Omega)^n$. However, since for $z_n = T(u_n)$ and $z = T(u)$

$$a(z_n - z, \psi) + b(u_n - u; z_n, \psi) + b(u; z_n - z, \psi) = 0 \quad \text{for all } \psi \in V,$$

and $|b(u; v, w)| \le M\,\|u\|_{L^4}\|v\|_{H^1}\|w\|_{H^1}$ for $u,\ v,\ w \in H^1(\Omega)^n$,

$$\nu\,\|z_n - z\|_V \le M\,\|z_n\|_V\|u_n - u\|_{L^4}.$$

Thus z_n converges strongly to z in V. Hence T is compact and therefore there exists a $u \in C$ such that $T(u) = u$ by the Shauder fixed point theorem [Is]. □

COROLLARY 3.3. *The solution $u \in V$ of (3.11) is unique provided that $k\,\|f\|_{V^*} < \nu^2$ where for a constant $k > 0$*

$$|b(u; v, w)| \le k\,\|u\|_V\|v\|_V\|w\|_V. \tag{3.13}$$

For the case of nonhomogeneous Dirichlet boundary conditions we need the following technical lemma due to Hopf [GR].

LEMMA 3.4. *Suppose Ω is a bounded open domain with Lipschtz continuous boundary. Then, given $g \in H^{1/2}(\Gamma)^n$ satisfying $\int_\Gamma g \cdot n\, ds = 0$, for any $\varepsilon > 0$ there exists a function $u_\varepsilon \in H^1(\Omega)^n$ such that $\nabla \cdot u_\varepsilon = 0$, $u_\varepsilon|_\Gamma = g$ and*

$$|b(\phi; u_\varepsilon, \phi)| \le \varepsilon\,|\phi|^2_V \quad \textit{for all } \phi \in V. \tag{3.14}$$

Now, the nonhomogeneous boundary value problem can be transformed into the one with homogeneous boundary condition by using the transformation $u = w + u_\varepsilon$; substituting u into (3.5) we obtain the equation for $w \in V$:

$$\begin{aligned} a(w,\psi) + b(w;u_\varepsilon,\psi) + b(u_\varepsilon;w,\psi) + b(w;w,\psi) &= \langle \bar f,\, \psi\rangle \\ \text{for all } \psi \in V, \end{aligned} \tag{3.15}$$

where $\bar f \in V^*$ is given by

$$\langle \bar f,\, \psi\rangle = \langle f,\, \psi\rangle - a(u_\varepsilon,\psi) - b(u_\varepsilon;u_\varepsilon,\psi).$$

Define the sesquilinear form $\hat a$ on $V \times V$ by

$$\hat a(\phi,\psi) = a(\phi,\psi) + b(\phi;u_\varepsilon,\psi) + b(u_\varepsilon;\phi,\psi) \quad \text{for } \phi,\ \psi \in V.$$

Then, if $\varepsilon < \nu$ then $\hat a$ is V-coercive from (3.10) and (3.14). Thus, applying the Shauder fixed point theorem as in the proof of Theorem 3.2 to (3.15) we obtain the following result.

THEOREM 3.5. *Given $f \in H^{-1}(\Omega)$ and $g \in H^{1/2}(\Gamma)$ with $\int_\Gamma g\cdot n\, ds = 0$ there exists at least one pair $(u,p) \in H^1(\Omega)^n \times L_0^2(\Omega)$ satisfying* (3.5).

Remark. $u_\varepsilon \in H^1(\Omega)$ appearing in the Hopf's lemma is used to show the existence of a solution to the Navier–Stokes equations. However, it is not feasible to be used in computations. In our numerical calculations we use $\bar u \in H^1(\Omega)^n$ that satisfies the Stokes equation

$$a(\bar u, v) - c(v,p) = 0 \quad \text{and} \quad c(\bar u, q) = 0 \tag{3.16}$$

for all $(v,q) \in H_0^1(\Omega) \times L^2(\Omega)$ with boundary condition $\bar u = g$ on Γ. Note that $\bar u$ is unique [GR] but with $\bar u$, condition (3.14) is not necessarily satisfied for arbitrary $\varepsilon > 0$.

4. Existence and first-order necessary condition of optimal solutions.

Two control problems described in §2 can be formulated as a constrained minimization in a Hilbert sapce using the notation of §3:

$$\begin{aligned} &\text{Minimize} && J(u) \\ &\text{subject to} && a(u,\psi) + b(u;u,\psi) = 0 \quad \text{for all } \psi \in V \\ & && \nabla\cdot u = 0 \\ & && u = g_0 + \sum_{i=1}^m f_i\,\chi_i \quad \text{on } \Gamma \\ & && f \in U, \end{aligned} \tag{4.1}$$

where $g_0,\ \chi \in H^{1/2}(\Gamma)$ with $\int_\Gamma g\cdot n\, ds = \int_\Gamma \chi_i \cdot n\, ds = 0$ and U is a closed bounded set in R^m. We discuss the Dirichlet boundary control problem and thus the body force is discarded. The function $f\cdot\chi = \sum_{i=1}^m f_i\,\chi_i,\ f \in U$ is the control input and influences the equation only through a part of boundary Γ, and the functions χ_i represent distribution functions of control input at Γ. The specific form of g_0 and χ ($m = 1$) for the cavity and channel problems is as follows. For the cavity $g_0 = (U_{\text{bot}}, 0)$ at the bottom surface, zero otherwise and $\chi = (1,0)$ at the top surface, zero otherwise. For the channel

$$g_0 = \begin{cases} x_2(2-x_2) & \text{at } x_1 = 0, \\ c\,x_2(3-x_2) & \text{at } x_1 = 21, \\ \text{zero} & \text{otherwise,} \end{cases}$$

and

$$\chi = \begin{cases} (x_2 - 2.375)(2.625 - x_2) & \text{at } x_1 = 3, \\ \hat{c}\,x_2(3-x_2) & \text{at } x_1 = 21, \\ \text{zero} & \text{otherwise,} \end{cases}$$

where c, $\hat{c}$ are chosen such that $\int_\Gamma g_0 \cdot n\,ds = \int_\Gamma \chi \cdot n\,ds = 0$. As pointed out in Remark 2.1, for the driven cavity problem g_0, χ are not in $H^{1/2}(\Gamma)$. Hence in our discussion we consider the problem in which g_0 and χ are replaced by $C^\infty(\Gamma)$ function that approximates g_0 and χ in $L^\infty-$ norm, respectively.

Let $u = w + \bar{u}^{(0)} + \sum_{i=1}^m f_i\,\bar{u}^{(i)}$ with $w \in V$ where $\bar{u}^{(0)}$ and $\bar{u}^{(i)}$, $1 \le i \le m$ are the solution of the Stokes equation (3.16) with boundary condition g_0 and χ_i, $1 \le i \le m$, respectively. Let $\bar{u} = \text{col}(\bar{u}^{(1)}, \cdots, \bar{u}^{(m)})$. Then the problem (1.4) can be equivalently written as

$$\text{(4.2)} \qquad \begin{aligned} &\text{Minimize} \quad J(u) \\ &\text{subject to} \quad a(w,\psi) + b(u;u,\psi) = 0 \quad \text{for all } \psi \in V, \end{aligned}$$

where $u = w + \bar{u}^{(0)} + \sum_{i=1}^m f_i\,\bar{u}^{(i)}$ and the cost functional J is minimized over $(w,f) \in V \times U$. Here not only is the boundary control problem transformed into the distributed control problem but also the control f appears directly in the cost functional J.

In our example without loss of generality we assume that $U = [-1,1]$.

THEOREM 4.1. *The set S of solutions defined by*

$$S = \Big\{u \in H^1(\Omega)^n : u = w + \bar{u}^{(0)} + \textstyle\sum_{i=1}^m f_i\,\bar{u}^i,\ w \in V\ f \in U \textit{ and}$$
$$u \textit{ satisfies } a(u,\psi) + b(u;u,\psi) = 0 \textit{ for all } \psi \in V\ \Big\},$$

is bounded in $H^1(\Omega)^n$.

Proof. Let $u^{(0)}$ and $u^{(i)}$, $1 \le i \le m$ are the Hopf's function (see Lemma 3.4) corresponding to g_0 and χ_i, $1 \le i \le m$, respectively. Then for any $f \in U$ we can write $u = w + u^{(0)} + \sum_{i=1}^m f_i\,u^{(i)}$ where $w \in V$ satisfies

$$\text{(4.3)} \quad a(w,\psi) + b(z;w,\psi) + b(w;z,\psi) + b(w;w,\psi) = -a(z,\psi) - b(z;z,\psi) \quad \text{for all } \psi \in V,$$

where $z = u^{(0)} + \sum_{i=1}^m f_i\,u^{(i)}$. It follows from Theorems 3.2 and 3.5 that there exists a solution $w \in V$ of (4.3) and taking $\psi = w$ in (4.3) and using Lemma 3.1, we obtain

$$(\nu - (m+1)\varepsilon)\|w\|_V \le \alpha\,\|z\|_{H^1} \quad \text{for some } \alpha > 0.$$

Thus, there exists a constant $\gamma > 0$ such that $\|u\|_{H^1} \leq \gamma$. ☐

THEOREM 4.2. *Suppose the cost functional J is weakly, sequentially lower semicontinuous. Then, the control problem (4.2) has at least one solution. In particular, each control problem described in §2 has at least one solution.*

Proof. Let (w_n, f_n) be a minimizing sequence. Since U is compact and the solution set S is bounded in $H^1(\Omega)^n$ there exists a subsequence (w_n, f_n) such that $f_n \to f^*$ in U and $w_n \to w^*$ weakly in V. Note that $(w_n, f_n) \in V \times U$ satisfies

$$a(u_n, \psi) + b(u_n; u_n, \psi) = 0 \quad \text{for all } \psi \in V$$

$$u_n = w_n + \bar{u}^{(0)} + f_n \cdot \bar{u}$$

and that $|b(w; u, \psi)| \leq M \|w\|_{L^4} \|u\|_{H^1} \|\psi\|_{H^1}$ for $w,\ u,\ \psi \in H^1(\Omega)^n$. Since $H^1(\Omega)$ is compactly embedded int $L^4(\Omega)$, $\|u_n - u^*\|_{L^4} \to 0$ and therefore it follows from (3.4) and Lemma 3.1 that $(w^*, f^*) \in S$. Hence if the cost functional $J : H^1(\Omega)^n \to R$ is weakly, sequentially lower semicontinuos, then (w^*, f^*) minimizes J. It is not difficult to show that each cost functional J is weakly, sequentially lower semicontinuous. In fact, for (2.3) the claim follows from the fact that the trace operator of $H^1(\Omega)$ on Γ_L is compact in $L^2(\Gamma_L)$. The cost functional (2.4) is the square of a norm on $H^1(\Omega)^2$. For (2.5) note that the cost functional is continuous on $L^2(\Omega)^2$. Hence each control problem has at least one solution. ☐

Remark. It is not difficult to extend Theorems 4.1 and 4.2 to the case when the control input g belongs to a compact subset of $H^{1/2}(\Gamma)^n$.

Next we discuss the first-order necessary optimality condition. Assume that $u^* = (w^*, f^*) \in V \times U$ is a local solution of (4.2) and that $f^* \in \text{int}(U)$. Let $G : V \times U \to V^*$ be defined by

$$\langle G(w, f),\, \psi \rangle = a(w, \psi) + b(u; u, \psi) \quad \text{for } \psi \in V$$

where $u = w + \bar{u}^{(0)} + \sum_{i=1}^m f_i\, \bar{u}^{(i)}$. In what follows we identify u with the pair (w, f) whenever $u = w + \bar{u}^{(0)} + \sum_{i=1}^m f_i\, \bar{u}^{(i)}$. It follows from [MZ] that if the Fréchet derivative of G at (w^*, f^*) is surjective, then the regular point condition is satisfied and hence there exists a Lagrange multiplier $\lambda \in V$ such that

$$J'(u^*)(v + h \cdot \bar{u}) + \langle G'(u^*)(v, h),\, \lambda \rangle = 0 \quad \text{for all } (v, h) \in V \times R^m.$$

LEMMA 4.3. *$G'(u^*)$ is given by*

$$\begin{aligned} \langle G'(u^*)(v, h),\, \psi \rangle &= a(v, \psi) + b(v; u^*, \psi) + b(u^*; v, \psi) \\ &\quad + h \cdot (b(\bar{u}; u^*, \psi) + b(u^*, \bar{u}, \psi)) \quad \textit{for } \psi \in V \end{aligned} \tag{4.6}$$

and $G'(u^*)$ is surjective if and only if the equation for $\psi \in V$

$$a(v, \psi) + b(v, u^*, \psi) + b(u^*; v, \psi) = 0 \quad \textit{for all } v \in V \tag{4.7}$$

$$b(\bar{u}^{(i)}; u^*, \psi) + b(u^*; \bar{u}^{(i)}, \psi) = 0 \quad \textit{for all } 1 \leq i \leq m \tag{4.8}$$

implies $\psi = 0$.

Proof. It is easy to show that the Fréchet derivative G at (w^*, f^*) is given by (4.6). It thus follows from (4.6) that (4.7), (4.8) are equivalent to the fact that

$\ker(G'(u^*)^*) = \{0\}$. Note that $G'(u^*)$ is surjective if and only if there exists a $(v, h) \in V \times R^m$ such that

$$v + Cv + Bh = T_0 f \quad \text{for arbitrary } f \in V^*, \tag{4.9}$$

where $T_0 f = (-\nu\,\Delta_S)^{-1} f,\ f \in V^*$ is the unique solution to the Stokes equation (3.8), the linear operator $C : V \to V$, defined by $Cv = T_0((v \cdot \nabla)u^* + (u^* \cdot \nabla)v)$, and the linear operator $B : R^m \to V$ is defined by $Bh = h \cdot T_0((\bar{u} \cdot \nabla)u^* + (u^* \cdot \nabla)\bar{u})$. Then since $H^1(\Omega)$ is embedded compactly into $L^4(\Omega)$, C is compact. Since B is of finite rank, V admits the orthogonal decomposition $V = \text{range}(B) \oplus \ker(B^*)$. Let Q be the orthogonal projection onto $\ker(B^*)$. Then (4.9) is equivalent to

$$v + QCv = QT_0 f \quad v \in \ker(B^*).$$

Since QC is compact it follows from the Riesz–Schauder theory [Yo] that range$(I + QC)$ on $\ker(B^*)$ is closed, which implies that range$(T_0 G'(u^*))$ is closed. Since $T_0 : V^* \to V$ is isometric isomorphism it follows from the Banach closed range theory [Yo] that $G'(u^*)$ is surjective if and only if $\ker(G'(u^*)^*) = \{0\}$. □

Remark. Lemma 4.3 implies that if the only solution of $v + Cv = 0$ is the zero solution (i.e., -1 is not an eigenvalue of C) then the surjectivity of $G'(u^*)$ is satisfied. Such a case occurs when u^* satisfies

$$|b(\phi; u^*, \phi)| < \nu\, \|\phi\|_V^2. \tag{4.10}$$

This inequality is, for example, satisfied when ν is sufficiently large since from Theorem 4.1 $\|u^*\|_{H^1(\Omega)^n}$ is uniformly bounded in $\nu \ge \nu_0 > 0$. Moreover, if C has the eigenvalue -1 with multiplicity less than $m + 1$, then the assumption of Lemma 4.3 still holds when (4.8) has only trivial solution on the eigenmanifold corresponding to the eigenvalue -1 of C.

Under the condition in Lemma 4.3 we obtain the first-order necessary condition for optimality:

$$a(u^*, \psi) + b(u^*; u^*, \psi) = 0 \quad \text{for all } \psi \in V, \tag{4.11a}$$

$$a(\lambda, \phi) + b(u^*; \lambda, \phi) + b(\lambda; u^*, \phi) + J'(u^*)(\phi) = 0 \quad \text{for all } \phi \in V, \tag{4.11b}$$

$$b(u^*; \lambda, \bar{u}^{(i)}) + b(\lambda; u^*, \bar{u}^{(i)}) + J'(u^*)(\bar{u}^{(i)}) = 0 \quad \text{for } 1 \le i \le m. \tag{4.11c}$$

5. Augmented Lagrangian method and convergence analysis. We solve the constrained minimization problem (4.2) (or equivalently (4.1)) using the augmented Lagragian method [He], [Po]. In our approach the divergence free constraint is imposed explicitly (without augmentation) but the Navier–Stokes constraint $G(w, f) = 0$ in V^* will be treated by the augmented Lagrangian method. We consider the augmented Lagrange functional

$$L_c(u, \lambda) = J(u) + \langle \lambda,\, G(w, f)\rangle_{V,V^*} + \frac{c}{2}\, a(z, z) \tag{5.1}$$

where $z \in V$ satisfies

$$a(z, \psi) = a(u, \psi) + b(u; u, \psi) \quad \text{for all } \psi \in V. \tag{5.2}$$

Note that $a(z, z)$ represents the square of the norm $\|G(w, f)\|_{V^*}$. Then, the augmented Lagrangian method applied to (4.2) is the following iterative scheme.

AUGMENTED LAGRANGIAN METHOD

Step 1. Choose the starting $\lambda_0 \in V$, a nondecreasing sequence of positive numbers c_k and set $k = 0$.

Step 2. Given λ_k, c_k find $u_k = (w_k, f_k) \in V \times U$ by

$$L_{c_k}(u_k, \lambda_k) = \min L_{c_k}(u, \lambda_k) \quad \text{over } (w, f) \in V \times U.$$

Step 3. Update λ_k by $\lambda_{k+1} = \lambda_k + c_k\, z_k$ where $z_k \in V$ satisfies

$$a(z_k, \psi) = \langle G(w_k, f_k), \psi\rangle \quad \text{for all } \psi \in V. \tag{5.3}$$

Step 4. If convergence criterion is not satisfied then $k = k + 1$ and go to Step 2.

Let $u^* = (w^*, f^*)$ be a local solution of (4.2). Assume that the assumption in Lemma 4.3 is satisfied. Then there exists a Lagrange multiplier $\lambda^* \in V$ such that (4.5) holds, where J is assumed to be continuously Fréchet differentiable in a neighborhood of u^*. Augmentabilty defined as below is of central importance in showing the convergence of the augmented Lagrangian method [He], [IK1].

DEFINITION. *The problem* (4.2) *is augmentable at* u^* *if there exist a neighborhood* $\tilde{U}(u^*)$ *of* (w^*, f^*) *in* $V \times U$ *and positive constants* $\bar{\sigma}$, $\bar{c}$ *such that*

$$L_c(u, \lambda^*) - L_c(u^*, \lambda^*) \geq \bar{\sigma}\,(\|w - w^*\|_V^2 + |f - f^*|^2) \tag{5.4}$$

for all $(w, f) \in \tilde{U}$ *and* $c \geq \bar{c}$.

Then the following theorem follows from [IK1], [IK2].

THEOREM 5.1. *Assume that the augmentability* (5.4) *holds. Given* $\lambda_0 \in V$ *for each* k *assume that* (w_k, f_k) *in the neighborhood* $\tilde{U}$ *satisfies*

$$L_c(u_k, \lambda_k) \leq L_c(u^*, \lambda_k) = J(u^*) \tag{5.5}$$

and update $\lambda_{k+1} = \lambda_k + (c_k - \bar{c})\, z_k$ *with* z_k *satisfying* (5.3). *Then we have*

(1) $\quad \bar{\sigma}\,(\|w_k - w^*\|_V^2 + |f_k - f^*|^2) + (c_k - \bar{c})\, \|\lambda_k - \lambda^*\|_V^2 \leq (c_k - \bar{c})\, \|\lambda_{k-1} - \lambda^*\|_V^2,$

(2) $\quad \displaystyle \bar{\sigma} \sum_{k-1}^{\infty} (\|w_k - w^*\|_V^2 + |f_k - f^*|^2) \leq \frac{1}{c_0 - \bar{c}}\, \|\lambda_0 - \lambda^*\|_V^2.$

Remark. (i) The condition (w_k, f_k) being in the neighborhood of u^* can be satisfied either by taking c_0 sufficiently large or λ_0 sufficiently close to λ^*.

(ii) The statement (2) implies the strong convergence of u_k to u^* in $H^1(\Omega)^n$.

(ii) The condition (5.5) means the sufficient reduction of successive cost functional.

Next we discuss a sufficient condition for the augmentability of (4.2). Let L be the Lagrangian corresponding to (4.2) defined by

$$L(u, \lambda) = J(u) + \langle \lambda, G(w, f)\rangle.$$

Suppose $J : H^1(\Omega) \to R$ is twice continuously differentiable in a neighborhood of u^*. The second derivative of $L(u, \lambda^*)$ at $u^* = (w^*, f^*)$ is given by

$$L''(u^*, \lambda^*)((v, h), (v, h)) = J''(u^*)(\xi, \xi) + b(\xi; \xi, \lambda^*), \tag{5.6}$$

where $\xi = v+h\cdot\bar{u}$. Then it follows from [IK1] that the augmentability (5.4) is achieved if the following second-order sufficient optimality condition is satisfied.

$$
\begin{aligned}
&L''(u^*,\lambda^*)((v,h),(v,h)) \geq \sigma\,(\|v\|_V^2 + |h|^2) \\
&\text{for all } (v,h) \in V \times R^m \text{ satisfying } G'(u^*)(v,h) = 0,
\end{aligned} \tag{5.7}
$$

where $\sigma > 0$. Moreover, it follows from [PT], [IK1] that $(u_k, \lambda_k) \in H^1(\Omega)^n \times V$ converges q-linearly to (u^*, λ^*); i.e.,

$$
\|\lambda_n - \lambda^*\|_V \leq \frac{K^n}{c_0 \cdots c_{n-1}}\,\|\lambda_0 - \lambda\|_V,
$$

$$
\|u_n - u^*\|_{H^1(\Omega)^n} \leq \frac{K^n}{c_0 \cdots c_{n-1}}\,\|\lambda_0 - \lambda\|_V
$$

for $n \geq 1$ and some constant K, provided that c_0 is sufficiently large or λ_0 sufficiently close to λ^*.

The following lemma gives an algebraic characterization of the second order sufficient optimality (5.7).

LEMMA 5.2. *Assume that for each i equation for V*

$$
a(v,\psi)+b(v;u^*,\psi)+b(u^*;v,\psi) = -b(u^*;\bar{u}^{(i)},\psi)-b(\bar{u}^{(i)};u^*,\psi)) \quad \textit{for all } \psi \in V \tag{5.8}
$$

has a unique solution $v_i \in V$ Then the condition (5.7) *is satisfied if and only if the matrix M on R^m, defined by*

$$
M_{i,j} = J''(u^*)(\phi_i,\phi_j) + b(\phi_i;\phi_j,\lambda^*) \quad \textit{with } \phi_i = v_i + \bar{u}^{(i)} \tag{5.9}
$$

is positive definite; *i.e., $h^tMh \geq \alpha\,|h|^2$ for all $h \in R^m$ and some $\alpha > 0$.*

Proof. Suppose $(v,h) \in V \times R^m$ satisfies $G'(u^*)(v,h) = 0$. Since for each i (5.8) has a unique solution v_i it thus follows from (4.6) and (5.8) that $v = \sum_{i=1}^m h_i v_i$ and therefore $\xi = \sum_{i=1}^m h_i(v_i + \bar{u}^{(i)})$ in (5.6). Then from (5.6) and (5.9) we have

$$
L''(u^*)((v,h),(v,h)) = h^tMh \quad \text{for } (v,h) \in \ker(G'(u^*)).
$$

Since $\|v\|_V \leq |h|\sqrt{\sum_{i=0}^m \|v_i\|_V^2}$ the positivity of the matrix M defined by (5.9) is equivalent to (5.7). $\square$

COROLLARY 5.3. *Suppose the matrix $\bar{M}$ on R^m, defined by $\bar{M}_{i,j} = J''(\bar{u}^{(i)},\bar{u}^{(j)})$, is positive definite. Then, if ν is sufficiently large then* (5.8) *has a unique solution and* (5.9) *holds.*

Proof. It follows from (3.16) and Theorem 4.1 that $\|u^*\|_{H^1(\Omega)^n}$, $\|\bar{u}^{(i)}\|_{H^1(\Omega)^n}$, $1 \leq i \leq m$ are uniformly bounded. Thus, if ν is sufficiently large then (4.10) holds and hence (5.8) has a unique solution $v_i \in V$ for each i. It follows from (4.11b) and (5.8) that

$$
\|v_i\|_V \leq \frac{M_1}{\nu}\|u^*\|\|\bar{u}^{(i)}\| \quad \text{and} \quad \|\lambda^*\|_V \leq \frac{M_2}{\nu}\,\|J'(u^*)\|_{V^*}
$$

for some constants M_1, M_2. Thus, from Lemma 3.1 and (5.9) if the matrix $\bar{M}$ is positive definite then the matrix M is positive definite provided that ν is sufficiently large. Hence, the corollary follows from Lemma 5.2. $\square$

Note that the cost functional (2.5) is not twice Fréchet differentiable. Motivated by this example we consider the following minimization problem in a Hilbert space X:

$$\text{Minimize } f(x) \text{ subject to } g(x) = 0,$$

where $g : X \to Y$ and y is a Hilbert space. Assume the following standard hypotheses:

(H1) $x^* \in X$ is a local solution;

(H2) g is twice continuously F-differentiable and f is (only) continuously $F-$ differentiable in a convex neighborhood of x^*;

(H3) $g'(x^*) : X \to Y$ is surjective.

Then there exists a unique Lagrange multiplier $\lambda^* \in Y^*$ such that

$$f'(x^*)(h) + \langle \lambda^*,\, g'(x^*)(h)\rangle = 0 \quad \text{for all } h \in X. \tag{5.10}$$

Define the functional F_t defined on $X \times X$ for $t > 0$ by

$$F_t(x)(h,h) = \frac{1}{t^2}\,(f(x+t\,h) - f(x) - t\,f'(x)h). \tag{5.11}$$

We make the following hypotheses on F_t:

(H4) There exists $\delta > 0$ and $0 < c_1 < 1 < c_2$ such that for $h = h_1 + h_2,\ h_1,\ h_2 \in X$ and $0 < t < \delta$

$$F_t(x^*)(h,h) \geq c_1\,F_t(x^*)(h_1,h_1) - c_2\,\|h_2\|_X^2.$$

(H5) There exists a $\sigma > 0$ such that

$$c_1\,F_t(x^*)(h,h) + \tfrac{1}{2}\,\langle \lambda^*, g''(h,h)\rangle \geq \sigma\,\|h\|_X^2$$

for all $h \in X$ satisfying $g'(x^*)h = 0$ and t sufficiently small.

Remark. If is f is twice differentiable the hypothesis (H5) reduces to the second sufficient optimality condition (5.7). Moreover, assuming $f''(x^*)$ is nonnegative definite then (H4) holds with $c_1 = 1/2$ and $c_2 = \|f''(x^*)\|$.

We have the following result.

THEOREM 5.4. *Assuming* (H1)–(H5), *we have the augmentability*; *i.e., there exists a neighborhood* $\tilde{U}(x^*)$ *and* $\bar{\sigma}$, $\bar{c} > 0$ *such that*

$$L_c(x,\lambda^*) - L_c(x^*,\lambda^*) \geq \bar{\sigma}\,\|x - x^*\|_X^2$$

for all $x \in \tilde{U}(x^*)$ *and* $c \geq \bar{c}$ *where for* $c \geq 0$ *and* $\lambda \in Y^*$

$$L_c(x,\lambda) = f(x) + \langle \lambda,\, g(x)\rangle_{Y^*,Y} + \tfrac{1}{2}\,\|g(x)\|_Y^2.$$

Proof. For each $h \in X$ with $\|h\| = 1$ we have

$$
\begin{aligned}
L_c(x + t\,h, \lambda^*) - L_c(x^*, \lambda^*) = {} & t\,(f'(x^*)h + \langle \lambda^*,\, g'(x^*)h\rangle) \\
& + t^2\,(F_t(x^*)(h,h) + \frac{1}{2}\,\langle \lambda^*,\, g''(x^*)(h,h)\rangle) \\
& + \frac{1}{2}t^2\,\langle \lambda^*,\, g''(\zeta(t))(h,h) - g''(x^*)(h,h)\rangle) \\
& + \frac{c}{2}t^2\,\|g'(x^*)h + \frac{t}{2}\,g''(\zeta(t))(h,h)\|_Y^2,
\end{aligned}
$$

where $\zeta(t) = x^* + \hat{t}\,h$, $\hat{t} \in (0,t)$ (which depends upon $(t,h) \in R \times X$). Since $g'(x^*) : X \to Y$ is surjective there exists a constant $\beta > 0$ such that

$$
\|g'(x^*)h\|_Y^2 \geq \beta\,\|h\|_X^2 \tag{5.12}
$$

for $h \in \mathrm{range}(g'(x^*)^*)$ where by the closed range theory [Yo] we have

$$
X = \mathrm{range}(g'(x^*)^*) \oplus \ker(g'(x^*)).
$$

Thus, for each $h \in X$ we have a unique representation $h = h_1 + h_2$, $h_1 \in \ker(g'(x^*))$ and $h_2 \in \mathrm{range}(g'(x^*)^*)$. It then follows from (5.10) – (5.12) and (H4) that for $x = x^* + t\,h$

$$
\begin{aligned}
I = {} & \frac{1}{t^2}\,(L_c(x, \lambda^*) - L_c(x^*, \lambda^*)) \\
\geq {} & c_1\,F_t(x^*)(h_1,h_1) + \frac{1}{2}\,\langle \lambda^*, g''(x^*)(h_1,h_1)\rangle + \left(\frac{c\beta}{4} - c_2\right)\beta\,\|h_2\|^2 \\
& + \langle \lambda^*, g''(x^*)(h_1,h_2)\rangle + \frac{1}{2}\,\langle \lambda^*, g''(x^*)(h_2,h_2)\rangle - Mt(1+ct)
\end{aligned}
$$

for some $M > 0$, where we used the fact that g'' is Lipschitz continuous in a neighborhood of x^*. It then follows from (H5) that

$$
\begin{aligned}
I \geq {} & \sigma\,\|h_1\|^2 + \left(\frac{c\beta}{4} - c_2\right)\,\|h_2\|^2 - Mt(1+ct) - \frac{\varepsilon}{2}\,\|h_1\|^2 \\
& - \frac{1}{2}\left(1 + \frac{\|\lambda^*\|\,\|g''(x^*)\|}{\varepsilon}\right)\,\|\lambda^*\|\,\|g''(x^*)\|\,\|h_2\|^2,
\end{aligned}
$$

for all $\varepsilon > 0$. Thus, for $t > 0$ sufficiently small (say, $0 < t < \delta$) we can choose $\bar{c} > 0$ such that

$$
I \geq \bar{\sigma}\,\|h\|_X^2 \quad \text{with } \bar{\sigma} = \frac{\sigma}{2},
$$

for all $c \geq \bar{c}$ and $h \in X$ satisfying $\|h\| = 1$ which completes the proof. □

Now we apply Theorems 5.1 and 5.4 to the control problem (2.5). An elementary calculation in [Di] shows that the function defined by

$$
f_t(x,k) = \tfrac{1}{2}\,(|\min(0, x+tk)|^2 - |\min(0,x)|^2) - tk\,\min(0,x)
$$

for $x,\ k \in R$, satisfies

$$f_t(x, k_1 + k_2) \geq \tfrac{1}{2}\, f_t(x, k_1) - t^2\, |k_2|^2 \tag{5.14}$$

for $k_1,\ k_2 \in R$. Note that for the control problem (2.5), F_t defined by (5.11) is given by

$$F_t(u^*)(h, h) = \frac{1}{t^2} \int_\Omega (f_t(u_1^*(x), h_1(x)) + f_t(u_2^*(x), h_2(x)))\, dx, \tag{5.15}$$

where $u^* = (u_1^*, u_2^*)$ and $h = (h_1, h_2) \in L^2(\Omega)^2$. It thus follows from (5.14) and the dominated convergence theorem that

$$F_t(u^*)(\xi, \xi) \geq \tfrac{1}{2}\, F_t(\xi^{(1)}, \xi^{(1)}) - \|\xi^{(2)}\|_{L^2}^2, \tag{5.16}$$

where $h = v + k\bar{u}^{(1)}$, $(v, k) \in V \times R$ and $\xi = \xi^{(1)} + \xi^{(2)} \in H^1(\Omega)^2$. Assume henceforth that (5.8) has a unique solution $v_1 \in V$. Then, as argued in the proof of Lemma 5.2, the hypothesis (H5) is equivalent to the following condition:

$$F_t(u^*)(\phi, \phi) + b(\phi; \phi, \lambda^*) > 0,$$

where $\phi = v_1 + \bar{u}^{(1)}$, for the control (2.5). Let $\Omega^* = \{x \in \Omega : u_1^*(x) \leq 0$ and $u_2^*(x) \leq 0\}$. Then it follows from (5.14), (5.15) that

$$F_t(u^*)(\phi, \phi) \geq \frac{1}{t^2} \int_{\Omega^*} (f_t(u_1^*(x), h_1(x)) + f_t(u_2^*(x), h_2(x)))\, dx, \tag{5.17}$$

since the right-hand side of (5.17) is monotonically nonincreasing in t. Thus, assuming

$$\frac{1}{t^2} \int_{\Omega^*} (f_t(u_1^*(x), h_1(x)) + f_t(u_2^*(x), h_2(x)))\, dx$$

is positive for some $t > 0$, it can be shown as in the proof of Corollary 5.3 that for ν sufficiently large (5.16) holds. Hence Theorems 5.1 and 5.4 apply to the control problem (2.5).

6. Numerical results. In this section we discuss numerical solution of the optimal control problems formulated in §2. The solution to (4.2) (equivalently (4.1)) is determined using the augmented Lagrangian method described in §5. The method involves the successive minimization of the cost functional of form

$$L_c(u, \lambda) = J(u) + a(u, \lambda) + b(u; u, \lambda) + \frac{c}{2}\, a(z, z) \tag{6.1}$$

over $(w, f) \in V \times U$ where $c > 0$ and $\lambda \in V$ are given, $u = w + \bar{u}^{(0)} + h \cdot \bar{u}$ and $z \in v$ satisfies (5.2). We use the projected conjugate gradient method (e.g., see [Gl]) to solve the constrained (i.e., $w \in V$ involves the divergence free condition $\nabla \cdot w = 0$) minimization problem (6.1).

PROJECTED CONJUGATE GRADIENT METHOD

Step 1. Choose the start-up $(w_0, f_0) \in V \times R^m$ and set $k = 0$.

Step 2. Compute the gradient $(g_k, r_k) \in V^* \times R^m$ by

$$\langle g_k, \psi \rangle_{V^*, V} = J'(u_k)(\psi) + a(\psi, \tilde{\lambda}) + b(\psi; u_k, \tilde{\lambda}) + b(u_k; \psi, \tilde{\lambda}) \quad \text{for all } \psi \in V$$

$$r_k = b(u_k; \bar{u}, \tilde{\lambda}) + b(u_k; \bar{u}, \tilde{\lambda}),$$

where $\tilde{\lambda} = \lambda + c\, z_k$ and $z_k \in V$ satisfies (5.3).

Step 3. The gradient $g_k \in V^*$ is projected onto V by the Stokes projection [Gl]; i.e., the projection $h_k \in V$ of g_k is given by

$$a(h_k, \psi) = \langle g_k,\, \psi \rangle \quad \text{for all } \psi \in V. \tag{6.2}$$

Set $d_k = (h_k, r_k) \in V \times R^m$ as the search direction.

Step 4. Set $\eta_k = h_k + r_k \cdot \bar{u}$. Compute $a_k =$ Argmin $L_c(u_k - \alpha\, \eta_k, \lambda)$ over $\alpha > 0$ (line search) and set

$$w_{k+1} = w_k - \alpha_k\, h_k,$$

$$f_{k+1} = f_k - \alpha_k\, r_k.$$

Step 5. Find gradients $(h_{k+1}, r_{k+1}) \in V \times R^m$ as in Step 2, 3 and compute

$$\beta_k = \frac{a(h_k, h_k) + |r_k|^2}{a(h_{k+1}, h_{k+1}) + |r_{k+1}|^2}$$

and set the search direction as

$$d_{k+1} = \begin{pmatrix} h_{k+1} \\ r_{k+1} \end{pmatrix} + \beta_k\, d_k.$$

Step 6. If the convergence criterion is not satisfied, then set $k = k+1$ and go to Step 4.

Remark. Note that for $\alpha > 0$ if $z(\alpha) \in V$ satisfies

$$a(z(\alpha), \psi) = a(u(\alpha), \psi) + b(u(\alpha); u(\alpha), \psi) \quad \text{for all } \psi \in V$$

with $u(\alpha) = u_k - \alpha \eta_k$, then $z(\alpha)$ can be written as

$$z(\alpha) = z_k + \alpha\, z_k^{(1)} + \alpha^2\, z_k^{(2)},$$

where $z_k^{(1)}$, $z_k^{(2)} \in V$ satisfy

$$a(z_k^{(1)}, \psi) = -a(\eta_k, \psi) - b(\eta_k; u_k, \psi) - b(u_k; \eta_k, \psi) \tag{6.3}$$

$$a(z_k^{(2)}, \psi) = b(\eta_k; \eta_k, \psi) \tag{6.4}$$

for all $\psi \in V$. Thus, $L_c(u_k - \alpha\, \eta_k, \lambda)$ is the polynomial of degree four in α and one can carry out the line search in Step 4 exactly provided that J is quadratic in u. Moreover, once the value of α_k in Step 4 is determined, then $z_{k+1} \in V$ is given by

$$z_{k+1} = z_k + \alpha_k\, z_k^{(1)} + \alpha_k^2\, z_k^{(2)}.$$

Hence our algorithm is reduced to solving a series of Stokes problem. Each inner iteration of the augmented Lagrangian algorithm (Step 2) requires solution of three Stokes problems; two (i.e., (6.3), (6,4)) in the line search and one ((6.2)) for the projection of the gradient onto V.

To carry out the computation we discretized the problem using the mixed finite element method [GR], [Te]. In our calculations the quadrilateral element for the

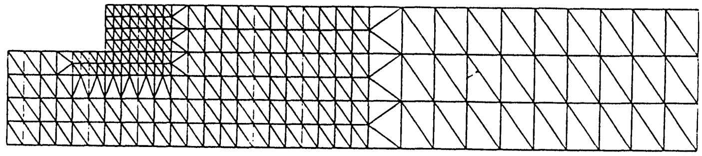

FIG. 5. *The triangulation for the channel*

velocity and the bilinear element for the pressure defined on the uniform rectangular grid with meshsize $h = .1$, are used in the cavity problem. The piecewise quadratic element for the velocity and the linear element for the pressure over the triangular grid shown Fig. 5 are used for the channel flow.

Let $\{\phi_h^i\}$, $\{\chi_h^i\}$ be the linearly independent basis functions of $H_h \subset H^1(\Omega)^2$ and $Q_h \subset L^2(\Omega)$ for velocity and pressure, respectively. Then, for example, the discretization of the Stokes equation (3.6) is given by

$$\begin{aligned} &a(u_h, \psi_h) - c(\psi_h, p_h) = \langle f, \psi_h \rangle \quad \text{for all } \psi_h \in H_h^0 \\ &c(u_h, \chi_h) = 0 \quad \text{for all } \chi_h \in Q_h \end{aligned} \tag{6.5}$$

where $u_h = \sum u_h^i \phi_h^i \in H_h$, $p_h = \sum p_h^i \chi_h^i \in Q_h$ and H_h^0 is the subspace of H_h that consists of all functions ψ_h in H_h satisfying $\psi_h \in H_0^1(\Omega)^2$. The solution $(u_h, p_h) \in H_h \times Q_h$ of (6.5) can be obtained by solving the following system of linear equations:

$$\begin{aligned} &A_h x + B_h^t y = f_h, \\ &B_h x = c_h \end{aligned} \tag{6.6}$$

where the (i, j)th element of the square matrix A_h is given by $a(\psi_h^i, \psi_h^j)$, ψ_h^i, $\psi_h^j \in H_h^0$ and the one of B_h is given by $c(\psi_h^i, \chi_h^j)$, $\chi_h^j \in Q_h$. We refer to [Di] for the detailed discussion of the discretization procedure and solution techniques for (6.6).

We now present numerical results for the control problems. For the problem (2.3) we take the Reynolds number to be 50 ($\nu = 1/50$) and Γ_L to be the horizontal line 0.4 units from the bottom. Given the bottom velocity as 0.5 we obtain the top velocity $U_{\text{top}}^{\text{opt}} = 1.16$ after four iterations of the augmented Lagrangian method with the value of c's equal to 20. The resulting flow field is shown in Fig. 6.

For the channel flow problem with Reynolds number 20 ($\nu = 1/20$) we obtain the following results. For the vorticity cost (2.4), using $c_k = 20$ in Step 2, the optimal control $f^{\text{opt}} = -.77$ (suction) and the optimal cost functional 23.13 are attained after five updates of the Lagrange multiplier. The resulting flow field is shown in Fig. 7. For the back flow cost (2.5), using $c_k = 0.05$ the optimal control $f^{\text{opt}} = 0.11$ (injection) is obtained and the optimal field flow is shown in Fig. 8. It must be noted, however, that a larger injection than f^{opt} decreases the size of the "bubble," but the strength (velocity of recirculation) is higher.

We also simulate the flow corresponding to the optimal control input using the ADI scheme (e.g., see [Gl]) with a finer discretization. Such simulations are in good agreement with the resulting flows we obtain in all three problems.

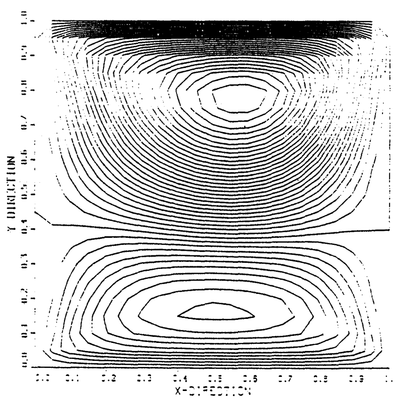

FIG. 6. *Locating seperation line at* $x_2 = 0.4$.

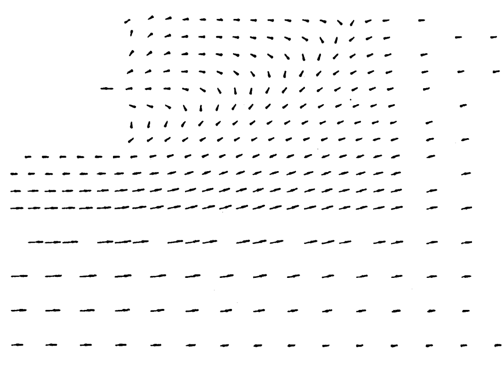

FIG. 7. *Control problem for channel with vorticity cost function* (Re $= 20$).

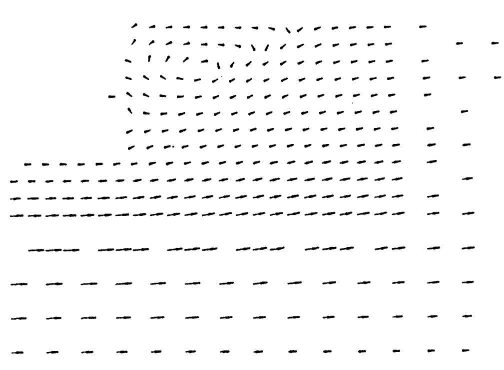

FIG. 8. *Control problem for channel with back flow cost function* (Re $= 20$).

REFERENCES

[AT] F. Abergel and R. Temam, *On some control problems in fluid mechanics*, Theoret. Comput. Fluid Mech., 1 (1990), pp. 303–325.

[Di] M. C. Desai, *Computations in Optimal Control of Navier–Stokes Equations*, Ph.D. Thesis, Division of Applied Mathematics, Brown University, May, 1990.

[DG] E. Dean and P. Gubernatis, *Pointwise control of Burgers' equation -a numerical approach*, Computer and Mathematics, 22 (1991), pp. 93–100.

[FS] H. O. Fattorini and S. S. Sritharan, *Necessary and sufficient conditions for optimal controls in viscous flow problems*, Proc. Roy. Soc. of Edinburgh Sect. A, to appear.

[Gl] R. Glowinski, *Numerical Methods for Nonlinear Variational Problems*, Springer-Verlag, Berlin, 1984.

[GHS] M. Gunzburger, L. Hou, and T. Svobodny, *Boundary velocity control of incompressible flow with an application to drag reduction*, SIAM J. Control Optim., 30 (1992), pp. 167–181.

[GR] V. Girault and P. A. Raviart, *Finite Element Methods for Navier–Stokes Equations*, Springer-Verlag, Berlin, 1984.

[He] M. R. Hestenes, *Multiplier and gradient methods*, J. Optim. Theory Appl., 4 (1968), pp. 303–320.

[Is] V. Istratescu, *Fixed Point Theory*, D. Riedel Publishing Company, Holland, 1981.

[IK1] K. Ito and K. Kunisch, *The augmented Lagrangian method for equality and inequality constraints in Hilbert spaces*, Math. Programming, 46 (1990), pp. 341–360.

[IK2] ———, *The augmented Lagrangian method for parameter estimation in elliptic systems*, SIAM J. Control Optim., 28 (1990), pp. 113–136.

[Li] J. L. Lions, *Optimal Control of Systems Governed by Partial Differential Equations*, Springer-Verlag, Berlin, 1971.

[MZ] H. Maurer and J. Zowe, *First and second-order necessary and sufficient optimality conditions for infinite-dimensional programming problems*, Math Programming, 16 (1979), pp. 98–110.

[NRK] H. M. Nagib, P. H. Reisenthal, and D. J. Koga, *Control of separated flows using forced unsteadiness*, AIAA Shear Flow Conference, Boulder, CO, 1985.

[Pi] O. Pironneau, *On optimal design in fluid mechanics*, J. Fluid Mech., 64 (1974), pp. 97–110.

[Po] M. J. D. Powell, *A method for nonlinear constraints in minimization problems*, in Optimization, R. Fletcher, ed., Academic Press, New York, 1969.

[PT] V. T. Polyak and N. Y. Tret'yakov, *The method of penalty estimates for conditional extremum problems*, Z. Vychisl. Mat. i Mat. Fiz., 13 (1973), pp. 34–46.

[Te] R. Temam, *Navier-Stokes Equations: Theory and Numerical Analysis*, North–Holland, Amsterdam, 1971.

[Yo] K. Yosida, *Functional Analysis*, Springer-Verlag, New York, 1980.