**PROGRAMS FOR ATTRACTING UNDER-REPRESENTED MINORITY
STUDENTS TO GRADUATE SCHOOL AND RESEARCH CAREERS IN
COMPUTATIONAL SCIENCE**

FINAL REPORT

for Period October 1, 1995 - September 30, 1997

James C. Turner Jr

*Florida A&M University
Tallahassee, FL 32307*

October 1997

Prepared for

THE U.S. DEPARTMENT OF ENERGY
AWARD NO. DE-FG02-96ER25302

1

We have no objection from a patent
standpoint to the publication or
dissemination of this material.

MP Driscoll 8/19/99
Office of Intellectual Date
Property Counsel
DOE Field Office, Chicago

DISCLAIMER

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, make any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

DISCLAIMER

Portions of this document may be illegible in electronic image products. Images are produced from the best available original document.

RECEIVED
AUG 26 1999
JST

**PROGRAMS FOR ATTRACTING UNDER-REPRESENTED MINORITY
STUDENTS TO GRADUATE SCHOOL AND RESEARCH CAREERS IN
COMPUTATIONAL SCIENCE**

ABSTRACT

Programs have been established within the Joint Science and Technology Alliance (JSTA) to attract minority students to research careers in mathematics and computational science. The primary goal of the program was to increase the number of such students attending graduate school. The longer range goal is, once these students receive their Ph.D. degree, to attract them to research careers in academic, industrial, and laboratory settings. One mechanism used for meeting these goals was to establish a Supercomputing Challenge that reached many students who were historically under-represented in scientific fields. This academic year program established at Florida A&M University Developmental Research School, introduced students to computational science projects using high-performance computers. Additional activities were conducted during the summer; those included workshops, meetings, and lectures. A unique aspect of our programs was follow-up interactions with student participants, during their undergraduate and graduate studies. Through the exposure provided by these programs, to scientific ideas and research, it is likely that their success rate in graduate schools will be high.

NOTICE

This report was prepared as an account of work sponsored by the United States Government. Neither the United States nor the United States Department of Energy, nor any of their employees, nor any of their contractors, subcontractors, or their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product or process disclosed or represents that its use would not infringe privately-owned rights.

INTRODUCTION

We have implemented programs that have the potential of being very successful for ameliorating what is an unquestioned and patently obvious fact: many minority groups are severely under-represented among Ph.D.'s in mathematics and computational science. We wish to make inroads towards the recruitment of minorities into Ph.D. programs in these disciplines, and subsequently into research careers. Our mechanisms for accomplishing these include the recruitment of minority undergraduate students into organized activities, both during the summer months and academic year. These activities directly address the causes for the lack of minorities in graduate schools and in research careers in mathematics or computational science. These programs are intended to ensure that students that have previously participated in our summer and academic year programs are successful in their graduate careers, and are well positioned to enter into research careers.

RESEARCH

During the first year of this award communication links between the Consortium members from Florida A&M University, Iowa State University, University of Minnesota, Carnegie-Mellon University, and the University of Maryland have been used to exchange information and ideas related to the study of smart materials and optimal control problems. Below we describe results of these collaborations to date, emphasizing our findings and their significance to computational science. We feel that our line of inquiry has had a beneficial effect on our understanding of various phenomena that arises in the study of smart materials, and our design of optimal control strategies.

• Control of the Chemical Vapor Deposition Processes

J. Turner, M. Gunzburger, L. Hou

In a chemical vapor deposition process, one is interested in maintaining a quasi-uniform temperature near the deposition site, so as to reduce the inhomogeneity of the deposition. In addition, in order to maximize the rate of deposition, one wishes to keep the concentration as high as possible near the deposition site. In our research two physical objectives have been mathematically formulated into one optimal control problem. As heat transfer is involved in the deposition process, we naturally chose heating and cooling as the control mechanism to achieve the objectives.

Let Ω denote the flow domain with boundary Γ . Also, denote $\mathbf{u}$, p , T , C the velocity field, pressure field, temperature field, and concentration field, respectively. The dimensionless equations governing the flow motion in question are given by

$$-\nu \nabla^2 \mathbf{u} + (\mathbf{u} \cdot \nabla) \mathbf{u} + \nabla p = \mathbf{f} + \frac{1}{L} g T \quad \text{in } \Omega, \quad (1)$$

$$\nabla \cdot \mathbf{u} = 0 \quad \text{in } \Omega, \quad (2)$$

$$-\kappa \Delta T + (\mathbf{u} \cdot \nabla) T = Q \quad \text{in } \Omega, \quad (3)$$

$$-\gamma \Delta c + (\mathbf{u} \cdot \nabla) c = Q \quad \text{in } \Omega. \quad (4)$$

We supplement the above system of equations with the appropriate boundary conditions. Thus far we have considered the following choice of boundary conditions:

$$\mathbf{u} = 0 \quad \text{on } \Gamma,$$

i.e. the flow is limited to inside Ω ,

$$T = T_0 \quad \text{on } \Gamma_d$$

i.e., the temperature at the deposition site is fixed at the solidification point, and

$$\frac{\partial T}{\partial n} = g \quad \text{on } \Gamma_c = \Gamma \setminus \Gamma_d$$

i.e., the heat flux occurs on this portion of the boundary.

Here, $\mathbf{f}$ is a given body force, m a given mass injection source, Q a given heat source, and T_0 the solidification temperature. The constants κ , γ , and ν depend on the thermal conductivity coefficient, density, specific heat at constant volume, and viscosity coefficient of the fluid. We assumed that buoyancy effects can be neglected, and thus the temperature variable does not appear in the equation of motion. The control g , i.e., the heat flux on the Γ_c portion of the boundary, is to be determined so as to maximize the concentration and minimize the temperature gradient in a region ω surrounding the deposition site Γ_d . This objective can be expressed mathematically as the following optimal control problem: seek a control g and a state $(\mathbf{u}, p, C, T)$ such that the functional

$$M(T, g) = \frac{\alpha}{2\varepsilon_1} \int_{\Omega \setminus \omega} |T|^2 dx + \frac{\alpha}{2\varepsilon_2} \int_{\Omega} |\nabla T|^2 dx \quad (5)$$

is minimized subject to

$$-\nu \nabla \mathbf{u} + (\mathbf{u} \cdot \nabla) \mathbf{u} + \nabla p = \mathbf{f} + \frac{1}{L} g T \quad \text{in } \Omega, \quad (6)$$

$$\nabla \cdot \mathbf{u} = 0 \quad \text{in } \Omega, \quad (7)$$

$$-\kappa \Delta T + (\mathbf{u} \cdot \nabla) T = Q \quad \text{in } \Omega, \quad (8)$$

$$-\gamma \Delta c + (\mathbf{u} \cdot \nabla) c = m \quad \text{in } \Omega, \quad (9)$$

with boundary conditions

$$u = 0 \quad \text{on } \Gamma, \quad (10)$$

$$T = T_0 \quad \text{on } \Gamma_d, \quad (11)$$

$$\frac{\partial T}{\partial n} = g \quad \text{on } \Gamma_c = \Gamma \setminus \Gamma_d \quad (12)$$

During the first year of this award we established the existence of an optimal solution. An optimality system of partial differential equations was derived, from which optimal controls and states may be determined. Next, finite element discretizations of the optimality system were defined and the corresponding error estimates obtained along with a solution algorithm for the discrete equations. We are currently developing computational code that will be used to demonstrate the effectiveness of our theory and methods. Finally, we are also seeking potential applications to research at DOE laboratories.

Optimal Control Problems for the Potential Equation

In this line of inquiry we considered an optimal control problem for impressed cathodic systems. A typical example of an impressed cathodic system is a metal container filled with an electrolyte. The painted portion of the container surface is usually treated as if it were insulated. The unpainted part is divided into cathode and anode, which are connected to the negative and positive poles of an electrical source, respectively. By adjusting the current density on the anode one can effectively alter the potential distribution on the entire bounding surface or in the entire flow domain. The potential distribution, of course, has a direct effect on the chemical reaction process occurring inside the flow domain, which in turn has an effect on the rate of corrosion of the metal container. Thus, the current density on the anode can be used as a practical control variable for generating a desired potential field. This idea can be conveniently formulated as optimal control problems for the potential equation with appropriate boundary conditions. In our research we have attempted to mathematically analyze optimal control problems with current density controls. The nonlinear model used involves an exponentially growing boundary condition.

Here we assume that the electrolyte occupies a physical domain $\Omega \subset R^2$ with boundary Γ . The electrical potential ϕ in Ω is governed by

$$- \operatorname{div}(\sigma \operatorname{grad} \phi) = 0 \quad \text{in } \Omega, \quad (13)$$

where the conductivity σ is a continuous function with a positive lower bound. The boundary Γ is divided into three components: the anode Γ_a , the cathode Γ_c and the insulated part Γ_o . On the cathode Γ_c , ϕ satisfies the relation

$$\sigma \frac{\partial \phi}{\partial n} = -f(\phi) \quad \text{on } \Gamma_c, \quad (14)$$

where f is the Butler-Volmer function defined by

$$f(\phi) = c_3 [e^{c_1 \phi} - e^{c_2 \phi}]. \quad (15)$$

On anode Γ_A , we have the boundary condition

$$\sigma \frac{\partial \phi}{\partial n} = u \quad \text{on } \Gamma_A, \quad (16)$$

which corresponds to the specification of the current density on the anode. Adjusting the current density on Γ_A amounts to treating u as a control variable. On the insulated part of Γ_o we have

$$\sigma \frac{\partial \phi}{\partial n} = 0 \quad \text{on } \Gamma_o. \quad (17)$$

In this setting we have considered the following optimal control problem: seek a state ϕ and a control u such that the functional

$$\mathfrak{J}(\phi, u) = \frac{1}{2\epsilon_0} \int_{\Omega} (\phi - \phi_0)^2 d\Omega + \frac{\delta_0}{2} \int_{\Gamma_A} u^2 d\Gamma, \quad (18)$$

is minimized subject to the constraint equations

$$-\text{div}(\sigma \text{grad} \phi) = 0 \quad \text{in } \Omega, \quad (19)$$

$$\sigma \frac{\partial \phi}{\partial n} = u \quad \text{on } \Gamma_A, \quad (20)$$

$$\sigma \frac{\partial \phi}{\partial n} = 0 \quad \text{on } \Gamma_o, \quad (21)$$

$$\sigma \frac{\partial \phi}{\partial n} = -f(\phi) \quad \text{on } \Gamma. \quad (22)$$

Here ϕ_0 is a desired potential distribution in Ω , ϵ_0 and δ_0 are positive constants.

We made use of an equivalent variational formulation of the nonlinear boundary value problem (19) - (22). This variational formulation takes the following form: seek $\phi \in H^1(\Omega)$ such that

$$\int_{\Omega} \sigma \text{grad } \phi \cdot \text{grad } \psi d\Omega + \int_{\Gamma_c} f(\phi) \psi d\Gamma = \int_{\Gamma_A} u \psi d\Gamma, \forall \psi \in H^1(\Omega). \quad (23)$$

We restate the minimization problem as follows:

seek a state $\phi \in H^1(\Omega)$ and a $u \in U$ such that the functional (18) is minimized subject to (23),

where

U is a non-empty, closed, convex subset of $L^2(\Gamma_A)$.

Our results for this optimal control problem for impressed cathodic systems in electrochemistry include the following:

existence and uniqueness of solutions to (23) have been established, so that the constraint equation is well-posed; establish the existence of an optimal pair (ϕ, u) that minimizes (18) subject to (23); justification of the use of Lagrange multiplier rules;

derivation of an optimality condition; regularity results for optimal solutions; development of a finite element algorithm for solving the optimality system;

derivation of error estimates.

Optimal Control Problems with Nonlinear Constraints

As is evidenced by the above research projects, the need to solve optimal control problems arises in many settings. Here our goal is to construct and analyze a framework for the approximate solution of many such problems. The components of our framework are as follows:

- define a general class of nonlinear control problems;
- establish the existence of optimal solutions;
- establish the existence of Lagrange multipliers, which may be used to enforce the constraints;
- derive an optimality system;
- define approximation algorithms;
- derive error estimates.

After developing and analyzing the general framework, we applied it to a specific problem, namely, the control of a magnetohydrodynamic flow. We used the general framework to analyze the concrete problem by merely showing that the latter fits into the former. We used the finite element method to obtain our numerical results.

In our abstract framework, we assumed that the functional to be minimized takes the form

$$\mathfrak{J}(v, z) = \lambda F(v) + \lambda \varepsilon(z) \quad \forall (v, z) \in X \times \Theta \quad (24)$$

Here, λ is a given parameter which is assumed to belong to a compact interval $\Omega \subset \mathbb{R}^+$, F is a functional on a reflexive Banach space X , ε a functional on Θ , where Θ is a closed and convex subset of another reflexive Banach space G .

The constraint equation $M(v, z) = 0$ relating the state variable v and the control variable z was defined as follows. Let N be a differentiable mapping from X to a third reflexive Banach space Y , κ a continuous linear operator from Θ to Y , and T a continuous linear operator from Θ to Y . Then, for any $\lambda \in \Omega$ and $(v, z) \in X \times \Theta$ the mapping from $X \times \Theta$ to Y , is given by

$$M(v, z) = v + \lambda TN(v) + \lambda Tk(z). \quad (25)$$

With these definitions we studied the constrained minimization problem:

$$\min_{(v, z) \in X \times \Theta} \mathfrak{J}(v, z) \quad \text{subject to} \quad M(v, z) = 0. \quad (26)$$

Even though we have shown that global minimizers as defined by (26) exist, in practice one can usually only locate local minimizers, i.e., points $(u, g) \in X \times \Theta$ such that for some $\varepsilon > 0$

$$\mathfrak{J}(u, g) \leq \mathfrak{J}(v, z) \quad \forall (v, z) \in X \times \Theta \quad (27)$$

such that

$$M(v, z) = 0 \quad \text{and} \quad \|u - v\|_x < \varepsilon \quad (28)$$

Using a fairly standard method of proof, we have proven the existence of optimal solutions.

Theorem 1 Assume that the functional $\mathfrak{S}$ and the mapping M defined by (24) and (25), respectively, satisfy suitable hypotheses. Then, there exists a solution to the minimization problem (26).

Next, using a result on an abstract Lagrange multiplier rule, we were able to show that the Lagrange multiplier rule may be used to turn the constrained minimization problem (27) and (28) into an unconstrained one.

Theorem 2 Let $\lambda \in \Omega$ be given and $(u, g) \in X \times G$ be a solution of problem (27) - (28). Then, under suitable conditions, there exists $\kappa \in \mathfrak{R}$ and $\mu \in X^*$, that are not both equal to zero, such that

$$\kappa < \mathfrak{S}_u(u, g), w > - < \mu, (I + \lambda TN'(u)) w > = 0 \quad \forall w \in X \quad (29)$$

$$\kappa < \varepsilon'(g), z > - < \mu, TKz > = 0 \quad \forall z \in G. \quad (30)$$

Note that if $\frac{1}{\lambda}$ is not in the spectrum of $-TN'(u)$, then we may choose $k = 1$. In addition, $TN'(u)$ is a compact operator from X to X , and therefore $\sigma(-TN'(u))$ is at most countable with zero being the only possible limit point. Thus (29) - (30) are valid for almost all $\lambda \in \Omega$. If the control g enters the constraint in a favorable manner, then we may in fact take $k = 1$ even if $\frac{1}{\lambda} \in \sigma(-TN'(u))$.

For purposes of numerical approximations, it turns out to be convenient to make the change of variables $\xi = T^* \mu$. Then, (29) and (30) with $k = 1$ can be written as

$$u + \lambda TN(u) + \lambda TKg = 0 \quad \text{in } X, \quad (31)$$

$$\xi + \lambda T^* [N'(u)]^* \xi - \lambda T^* F'(u) = 0 \quad \text{in } Y^*, \quad (32)$$

$$g - EK^* \xi = 0 \quad \text{in } G. \quad (33)$$

Clearly, the control g may be eliminated from the optimality system (31) - (33). However, given $\xi \in Y^*$, it is not always possible to evaluate g exactly from (33). For example, the application of the operator E may involve the solution of a partial differential equation.

A finite dimensional discretation of the optimality system (31) - (33) has been defined. Also, corresponding error estimates have been derived. These results imply that, under certain hypotheses, the error of approximation of solutions to our nonlinear problem, is basically the same as the error of approximation of solutions of related linear problems.

We have assumed that the solution $(u(\lambda), g(\lambda), \xi(\lambda))$ of the optimality system (31) - (33) that we are trying to approximate is a regular solution. In fact, the assumptions that we have made are sufficient to guarantee that for almost all values of λ , this is indeed the case.

Theorem 3 Let $(u(\lambda), g(\lambda), \xi(\lambda)) \in X$, for $\lambda \in \Omega$, be a branch of regular solutions of the optimality system (31) - (33). Then, under suitable conditions, there exists $\delta > 0$ and $h_0 > 0$ such that for $h < h_0$, the discrete optimality system:

Seek $u^h \in X^h$, $g^h \in G^h$, and $\xi^h \in (Y^*)^h$ such that

$$u^h + \lambda T^h N(u^h) + \lambda T^h K g^h = 0 \quad \text{in } X^h, \quad (34)$$

$$\xi^h + \lambda (T^*)^h [N'(u^h)]^* \xi^h - \lambda (T^*)^h F'(u^h) = 0 \quad \text{in } (Y^*)^h \quad (35)$$

$$g^h - E^h K^* \xi^h = 0 \quad \text{in } G^h, \quad (36)$$

has a unique solution $(u^h(\lambda), g^h(\lambda), \xi^h(\lambda))$ satisfying

$$\|(u(\lambda), g(\lambda), \xi(\lambda)) - (u^h(\lambda), g^h(\lambda), \xi^h(\lambda))\|_X < \delta. \quad (37)$$

Moreover,

$$\lim_{h \rightarrow 0} \|(u(\lambda), g(\lambda), \xi(\lambda)) - (u^h(\lambda), g^h(\lambda), \xi^h(\lambda))\|_X = 0 \quad (38)$$

uniformly in $\lambda \in \Omega$ and there exists a constant C, independent of h and λ , such that

$$\begin{aligned} & \lim_{h \rightarrow 0} \|(u(\lambda), g(\lambda), \xi(\lambda)) - (u^h(\lambda), g^h(\lambda), \xi^h(\lambda))\|_X \\ & \leq C \lambda \{ \|(T^h - T)(N(u(\lambda)) + K g(\lambda))\|_X \\ & \quad + \|(E^h - E) K^* \xi(\lambda)\|_G \\ & \quad + \|(T^*)^h - T^* \left([N'(u(\lambda))]^* \xi + F'(u(\lambda)) \right)\|_{Y^*} \} \end{aligned} \quad (39)$$

Finally, we applied the above theories to a special type of magnetohydrodynamic (MHD) flow. In this type of flow we assume that the externally applied magnetic field $\mathbf{B}$ is undisturbed by the flow. Here we denoted by Ω the flow domain which is bounded open set in $\mathbb{R}^2$ with a boundary Γ . The dimensionless equations governing the special type of MHD flow are given by

$$-\frac{1}{M^2} \Delta \mathbf{u} + \frac{1}{N} (\mathbf{u} \cdot \nabla) \mathbf{u} + \nabla p - (\mathbf{B} \times \nabla \phi) - (\mathbf{u} \times \mathbf{B}) \times \mathbf{B} = \mathbf{f} \quad \text{in } \Omega, \quad (40)$$

$$\operatorname{div} \mathbf{u} = 0 \quad \text{in } \Omega, \quad (41)$$

$$-\Delta \phi + \operatorname{div} (\mathbf{u} \times \mathbf{B}) = 0 \quad \text{in } \Omega, \quad (42)$$

where $\mathbf{u}$ denotes the velocity field, p the pressure field, and ϕ the electric potential. Also, N is the interaction number and M is the Hartmann number. $\mathbf{B}$ and $\mathbf{f}$ are given data. The system (36)-(38) is supplemented with the boundary conditions

$$\begin{aligned} \mathbf{u} &= 0 \quad \text{on } \Gamma, \\ \frac{\partial \phi}{\partial n} &= h \quad \text{on } \Gamma, \end{aligned}$$

where g denote the only control variable, namely, the normal electric current on Γ .

We verified that the hypotheses for the existence of optimal solutions hold in this setting. Having verified the hypotheses, we then applied Theorem 2 to conclude that there exists a

Lagrange multiplier $(v, \tau, s) \in X = H_0'(\Omega) \times L_0(\Omega) \times \widehat{H}'(\Omega)$ and a real number k such that (29)-(30) hold and for almost all values of λ , we may choose $k = 1$ in (29)-(30). We next defined a finite element discretation of the optimality system. We then used Theorem 3 to derive error estimates for the approximate solutions of the optimality system. We concluded this investigation by considering a computational example where we match the velocity field with a desired one by finding an appropriate boundary current density g .

• Controllability for Parabolic Equations

J. Turner, M. Gunzburger, Y. Cao

In many problems of interest to science and engineering there arises the need of controllability for parabolic equations. The purpose of this research was to investigate exact controllability problems for systems modeled by linear parabolic differential equations. Our aim of investigation has been to examine the limit behavior of the corresponding optimal control problem. Using this technique we obtained a new condition on the exact controllability of the system.

We now give a statement of the problem. Let Ω be a bounded domain in R^d . For fixed T , let $Q = [0, T] \times \Omega$ and $\Sigma = (0, T) \times \partial\Omega$. Let A be a second order elliptic differential operator defined by

$$A(y) = - \sum_{i,j=1}^d \frac{\partial}{\partial x_j} \left(a_{i,j}(x) \frac{\partial y}{\partial x_i}(t, x) \right) + c(x) y. \quad (43)$$

We assume that $c > 0$ on Ω and that the matrix $(a_{i,j}(x))$ is symmetric and positive definite.

With these definitions, we have studied the following initial boundary value problem

$$\begin{aligned} \frac{\partial y}{\partial t} + A(y) &= 0 & \text{in } Q, \\ \frac{\partial y}{\partial \nu} &= u(t, x) & \text{on } \Sigma, \\ (y(x, 0) &= y_0(x) & \text{in } \Omega, \end{aligned} \quad (44)$$

where

$$\frac{\partial}{\partial \nu} = \sum_{i,j=1}^d a_{i,j} n_i \frac{\partial}{\partial x_j}$$

with n_i being the i^{th} component of the outward unit normal to $\partial\Omega$. We say that the system is *exactly controllable* for a given state $\hat{y}$ if there exists $u \in L^2(\Sigma)$ such that the solution y of problem () also satisfies

$$y(T, \cdot) = \hat{y}(\cdot).$$

Next, we define the cost functional J_ε by

$$J_\varepsilon(u) = \frac{1}{2} \int_{\Omega} (y(T, x, u) - \hat{y}(x))^2 dx + \frac{\varepsilon}{2} \int_{\Sigma} u^2(t, x) dx dt,$$

where $y = y(t, x; u)$ satisfies (44). $u_\varepsilon(t, x)$ is said to be an *optimal control* if it satisfies

$$J_\varepsilon(u_\varepsilon) = \inf_{u \in L^2(Q)} J_\varepsilon(u).$$

Our weak formulation of problem (21) takes the form

$$\begin{aligned} \left(\frac{\partial y}{\partial t}, \phi \right) + a(y, \phi) &= \langle u, \phi \rangle & \text{for } \phi \in H^1(\Omega), \\ y(0, x) &= y_0(x), \end{aligned} \quad (45)$$

where the bilinear form $a: H^1(\Omega) \times H^1(\Omega) \rightarrow R$ is given by

$$a(\phi, \psi) = \int_{\Omega} \left\{ \sum_{i,j=1}^d a_{i,j}(x) \frac{\partial \phi}{\partial x_i} \frac{\partial \psi}{\partial x_j} + c(x) \phi \psi \right\} dx$$

Our main results of this research has been obtained by using an operator representation of the terminal state. That is, we define an operator $R: L^2(\Omega) \rightarrow L^2(\Omega)$ by

$$(\varepsilon + R)(y(T) - \hat{y}) = \varepsilon(\hat{y} - E(T)v)$$

where $E(t)$ is the solution operator for the problem

$$\frac{\partial w}{\partial t} + A(w) = 0 \quad \text{on } Q, \quad (46)$$

$$\frac{\partial w}{\partial \nu} = 0 \quad \text{on } \Sigma, \quad (47)$$

$$w(0, x) = z(x) \quad \text{in } \Omega, \quad (48)$$

i.e. $w(t, x) = E(t)z(x)$.

The following properties of the operator R have been proven: R is a compact operator in $L^2(\Omega)$; R is symmetric and semi-definite; if Σ is analytic, then $\ker R = 0$.

Note that system (46) - (48) is said to be *approximately controllable* if, for any given $\hat{y}$ and $\delta > 0$, there exists a control $u(t, x)$ and a function $y(t, x)$ such that $(y(t, x), u(t, x))$ is a solution of (44) and

$$\|y(T, x) - \hat{y}(x)\| < \delta.$$

Approximate controllability is well known for linear parabolic differential equations. However, the proof that we were able to obtain is different from standard proofs, in that it is constructive.

Finally, using properties of R we obtained the following result on exact controllability of system (44).

Theorem 4 Let $(\{e_j\}_{j=1}^\infty, \{\lambda_j\}_{j=1}^\infty)$ be the eigensystem of R . If

$$\hat{y} - E(T)y_0 \in D(R^{-1}),$$

then system (44) is exactly controllable.

- **Accident Induced Buoyancy Flows in Multilevel Nuclear Facilities**

J. Turner, K. Greenaugh (DoE)

Nuclear facilities consists of multilevel structures that may occur, resulting in a release of radioactive material to the public, two primary flow paths must be studied. These paths can be subdivided relative to near-field, and far-field. The near-field takes into account the energetic and material releases close to the source. This source may be provided by an accident or an initiating event such as a fire. Studying near-field phenomena consists of the mechanisms involved in transporting source material to upper levels of a facility, and to the facility's atmospheric boundary. When gases are not involved, there are two primary paths which dominate flows to upper levels and the atmospheric boundary from an internal source. These paths are formed by the heating, ventilation, and air condition (HVAC) systems, which are complicated by stack effects coupled with buoyancy flows up vertical shafts, such as stairwells or elevator openings. Several codes have been developed to study transport due to HVAC systems. The codes typically employ a network model to simulate ventilation system ducts, blowers, and dampers. A family of such codes have been developed and experimentally validated at Los Alamos National Laboratory. These codes include TORAC, which models the material transport through an HVAC system due to a tornado induced source; EXPAC which simulates flows from an explosion initiating event; and FIRAC, which simulates fire induced flows.

The focus of our research is to study buoyant flows up multi-level buildings, through vertical shafts due to accidents or initiating events. The information obtained and the model developed in this research provide tools for performing the deterministic analyses with a formalized safety analysis.

There are three primary approaches to developing a model of buoyant flows up vertical shafts. They are dimensional analysis (using scaling laws); numerical analysis which involves solving the nonlinear partial differential equations which govern buoyant flows; and experimental analysis, which requires the use of empirical data.

Our approach has been to develop a model based on a dimensional analysis technique, commonly referred to as *Froude number modeling*. We have also conducted a series of experiments to verify the scaling relationships of the Froude model. In addition, empirical data from these experiments have been used to validate the model. A numerical simulation has also been developed. This simulation will provide approximations which can be compared to the Froude model. Flow visualization have also been developed to aid in the determination of flow characterization. These flow visualizations have been used to assist in deciding which flow parameters to preserve and which may be omitted in the dimensional analysis.

ACADEMIC YEAR 1995-96

LECTURES

For a period of one week, invited speakers from various universities and industry participated in a lecture series. The lecturers included the following:

- Dr. Raymond Johnson, University of Maryland at College Park
- Dr. William Massey, AT&T Bell Laboratories
- Dr. Lee Lorch, York University (Canada)

A faculty member was invited to give a lecture at the following conference:

**Conference for African American Researchers in the Mathematical Sciences,
June 1995**

Mathematical Sciences Research Institute, Berkeley, California

Dr. James C. Turner Jr., Florida A&M University, *Mathematical Problems Arising in the Study of Smart Materials*

CONFERENCES AND MEETINGS

In order to enhance the their knowledge of mathematics and exposure to the mathematical community, both students and faculty attended the following mathematical conferences:

**Conference for African-American Researchers in the Mathematical Sciences,
June 1995**

Mathematical Sciences Research Institute, Berkeley, California

1 Student, 1 Faculty member attended

ACTIVITIES

A pilot High School Supercomputing Challenge (HS²C) was held at Florida A & M University Developmental Research School. Our goal was to have six teams. The teams were to work together during the school year to develop a project. The Challenge proceeded through the following

- Call for Participation
- Introductory Workshop
- Development of Projects
- Final Project Report
- Judging and Awards
- Evaluation and Feedback of HS²C

SUMMER 1996

LECTURES

Two formal lectures were given to students and faculty participating in a summer program at Florida A & M University. These lectures included:

- Dr. James Turner, Florida A & M University, *Computerized Tomography*

- Dr. Bruno Guerrieri, Florida A & M University, *Truck Backer-Upper & Travelling Salesman Problems*

RESEARCH PROJECTS

Students participating in our summer activities engaged, some on an individual basis and some in a team setting, in research or special study projects. These projects were tailored to the backgrounds and interests of the individual students, and allowed the participants to receive a detailed picture of one particular area of computational science. According to the students' interests and the availability of scientists who are experts in that area, each student was paired with a mentor. Each mentor devised a project for the student, gave the student adequate training to prepare the student for carrying out the project, and then closely guided the student through the actual performance of the project. In this way, a close student-mentor association was created; such associations, and the inherent trust built into them, can be used by the students as they seek advice about their futures. Near the end of the summer term, several students gave presentations of their work to other students and to interested faculty, and thus experience in that aspect of scientific research. These projects included:

- **Computerized Tomography**
Group Leader: J. Turner
Students: M. Curry, C. Zellars, T. Hodge
- **Truck Backer-Upper & Traveling Salesman Problems**
Group Leader: B. Guerrieri
Students: S. Carrington, A. Serieux, M. Aziz
- **Computational Science Educational Programs**
Group Leader: T. Mason
Students: K. Williams, D. Corey, T. Hodge
- **Wavelets**
Group Leader: J. Turner
Students: T. Hodge
- **Surface Representations with Deformable Splines**
Group Leader: J. Turner
Students: T. Hodge, D. Corey
- **A Phase-Field Model of Alloy Solidification**
Group Leader: J. Turner
Students: A. Livingston

ACTIVITIES

Since the field of computational science is new and not widely recognized by young minorities as an option for their future careers, we decided to accelerate the process of attracting minorities in this field by implementing a series of activities during the summer term. These activities included formal lectures, research projects, and participation in: the

National Association of Mathematicians Summer Institute in Computational Science; Undergraduate Computational Engineering and Sciences (UCES) Project; and the Calculus, Concepts, Computers and Cooperative Learning (C⁴L) Workshop.

ACADEMIC YEAR 1996-97

LECTURES

In addition to inviting speakers to lecture in an Adventures in Computational Science lecture series at Florida A & M University, several faculty members were invited to universities to participate in their lecture series.

Iowa State University, Ames, Iowa

- Dr. Y. Cao, Florida A & M University, *A New Condition of the Exact Boundary Controllability for Parabolic Equations*
- Dr. K. Blayneh, Florida A&M University, *Population Models for Biological Mathematics*

North Carolina A & T, Greensboro, North Carolina

- Dr. James Turner, Florida A & M University, '
Florida A & M University
- Dr. Dawn Lott-Crumpler, University of Maryland at College Park, *The Nonlinear Propagation of Waves as Applied to Shear Band Formation: A Numerical Approach*
- Dr. David Kinderlehrer, Carnegie-Mellon University, *New Mathematics for New Materials*

PROJECTS

Throughout the school year, both graduate and undergraduate students were exposed to the following topics:

- **Computational Linear Algebra**
Group Leader: J. Turner
Students: M. Curry, A. Livingston, C. Zellars, D. Greene
- **Calculus with Variations**
Group Leader: J. Turner
Students: M. Curry, A. Livingston, C. Zellars, D. Greene
- **Applied Partial Differential Equations**
Group Leader: J. Turner
Students: M. Curry, A. Livingston, C. Zellars, D. Greene, T. Hodge, D. Corey
- **Introduction to the Finite Element Method**
Group Leader: J. Turner
Students: A. Livingston, D. Greene
- **Modeling the Effects of Ultraviolet Light on Organisms**

Group Leader: J. Turner
Students: M. Curry, C. Zellars, D. Greene

CONFERENCES AND MEETINGS

In order to enhance their knowledge of mathematics and exposure to the mathematical community, both students and faculty attended several mathematical conferences and meetings:

- **Institute for Mathematics and its Applications, October 1996**
University of Minnesota, Minneapolis, Minnesota
2 Students 1 Faculty member attended
- **Undergraduate MathFest 1996, October 1996**
Xavier University, New Orleans, Louisiana
5 Students, 3 Faculty members attended
- **Iowa State University, November 1996**
Ames, Iowa
3 Students, 5 Faculty members attended
- **North Carolina A & T State University, November 1996**
Greensboro, North Carolina
4 Students, 1 Faculty member attended
- **Joint Mathematics Meeting, January 1997**
San Diego, California
2 Students, 1 Faculty member attended
- **Optimal Control: Theory, Algorithms, and Applications, February 1997**
Center for Applied Optimization, University of Florida, Gainesville, Florida
1 Faculty member attended
- **NAM Regional Faculty for Research and Teaching Excellence Conference, April 1997**
Tuskegee University, Tuskegee, Alabama
2 Students, 2 Faculty members attended
- **ACTIVITIES**

Adventures in Computational Science was a program established to enhance the knowledge of undergraduate students in areas of emerging technology that are computational science based. The ACS activities included:

- **Supercomputing Challenge**
- **Opening Ceremony**
- **Workshops**
- **Summer Program**

SUMMER 1997

LECTURES

Both faculty and students, participated in summer activities by giving lectures. These lectures included:

- **University of New Orleans and McDonogh 35 High School, New Orleans, Louisiana**

Dr. James C. Turner Jr., Florida A & M University, *Mathematical Problems Arising in the Study of Smart Materials*

- **SIAM Meeting, Stanford University, Stanford, California**

Alice Livingston, Arizona State University, *Design and Control of the Cz Crystal Growth Process*

- **Arizona State University, Tempe, Arizona**

Dr. James C. Turner Jr., Florida A & M University, *Analysis and Approximation of Optimal Control Problems with Nonlinear Constraints*

CONFERENCES AND MEETINGS

In order to enhance the their knowledge of mathematics and exposure to the mathematical community, both students and faculty attended several mathematical conferences and meetings:

Conference for African American Researchers in the Mathematical Sciences, June 1997

Morgan State University, Baltimore, Maryland
3 Students, 1 Faculty member attended

Next Step Two, June 1997

University of New Orleans/Louisiana State University, New Orleans, Louisiana
1 Faculty member attended

Society for Industrial and Applied Mathematics Meeting, July 1997

Stanford University, Stanford, California
3 Students, 1 Faculty member attended

American Mathematical Society Summer Meeting, August 1997

Atlanta, Georgia
1 Faculty member attended

PUBLIC RELATIONS

Below is a list of several mediums in which various activities described in this report appeared:

- Tallahassee Democrat (local newspaper)
- Appearance on Channel 27 WTXL News Cast
- FAMU Cluster Group (a group of business sponsors for Florida A & M University)
- Black Collegiate Magazine

- **Florida A&M University Homecoming Brochure**
- **FAMU Today** (a quarterly publication from Florida A & M University)
- **SIAM (Society for Industrial and Applied Mathematics) News**
- **Minnesota Public Radio in conjunction with the Institute for Mathematics and its Applications**
- **Florida A&M University Radio**
- **Notices from American Mathematical Society**
- **Institute for Mathematics and its Applications Video**

RESEARCH PARTICIPANTS

The following faculty members, graduate students, undergraduate students, high school teachers, and high school students were supported for participation in the above mentioned activities:

Faculty

J. Turner
R. Andrews (5/96 - 5/97)
B. Guerrieri (5/96 - 8/96)
T. Mason (5/96 - 6/97)
M. Abdullah (5/97 - 8/97)
D. Hill (1/97 - 5/97)
R. E. Williams (5/96 - 4/97)
K. Blayneh (8/96 - 4/97)
Y. Cao (8/96 - 4/97)
I. Turner (5/96 - present)

Graduate Students

D. Corey (5/96 - 8/97)
T. Hodge (7/96 - 6/97)
A. Livingston (7/96 - present)
D. Greene (9/96 - 8/97)
M. Curry (5/97 - present)
C. Zellars (5/97 - present)

Undergraduate Students (Academic Year 1995-96)

M. Aziz (Engineering)
D. Berkeley (Engineering)
T. Coffee (Mathematics)
M. Curry (Mathematics)
N. Gardner (Engineering)
T. Smith (Mathematics)
D. Thomas (Mathematics)
S. Watkins (Mathematics)
C. Zellars (Mathematics)

High School Students (Academic Year 1995-96)

Althea Griffin
Andrew Smith
Kenneth Borders
Cecily Monroe
Jason MaCaskill

Undergraduate Students (Summer 1996)

M. Aziz (Engineering)
O. Burch (Mathematics)
S. Carrington (Mathematics)

M. Curry (Mathematics)
A. Serieux (Engineering)
K. Williams (Physics)
C. Zellars (Mathematics)

Undergraduate Students (Academic Year 1996-97)

M. Curry (Mathematics)
C. Zellars (Mathematics)
E. Hardison (Mathematics)
P. Adkins (Mathematics)
K. Williams (Physics)
T. Smith (Pharmacy)
L. Stephens (Mathematics)
D. Neal (Business Administration)
H. Davis (Business Administration)
F. Jones (Accounting/Finance)
J. Larry (Biology)
V. Davis (Political Science)
S. Hunt (Electrical Engineering)
A. Murray (Biology)
K. Cooks (Computer Science)
R. Powell (Engineering)
K. Nickerson (Engineering)
C. Lynch (Engineering)
J. Johnson (Pharmacy)
R. Buford (Engineering)
D. O'Neal (Civil Engineering Technology)
K. Stewart (Pharmacy)

High School Teachers (Academic Year 1996-97)

J. Saylor
J. Holder
C. Suluki
N. Adams
L. Lindsay
M. Rice
R. Ford
J. Wallace

Undergraduate Students (Summer 1997)

E. Hardison (Mathematics)
K. Williams (Physics)
P. Adkins (Mathematics)
K. Stewart (Pharmacy)

High School Students (Summer 1997)

V. Winchester

D. Duvugneaud

A. Green

A. Wells

C. Charles

A. Powell

High School Teachers (Summer 1997)

M. McGhee

APPENDIX B

This section contains abstracts for some lectures listed in the report.

ABSTRACT

Mathematical Problems Arising in the Study of Smart Materials

Dr. James C. Turner Jr., Florida A&M University

The smartness of a material is a consequence of its ability to form a flexible structure at one temperature or load while recognizing only a homogeneous equilibrium at a different temperature or load. These materials include shape-memory, magnetostrictive, and other active materials and are finding increase applications as actuators and sensors and inherently, they are highly nonlinear.

The mechanism leading to shape-memory behavior relies on a displacive transition or a temperature controlled change in shape. This is the hallmark of a martensitic transformation and it is frequently accompanied by a complicated microstructural pattern. A major challenge to materials science is the prediction and control of microstructure. In recent years, a theory of microstructure has been developed and put into practice that can describe the macroscopic properties of martensitic, shape-memory alloys. Mathematical models for the equilibrium of such materials give nonconvex variational problems that may fail to attain a minimum value for any admissible deformation.

It is my intention in this lecture to introduce several important nonlinear techniques developed in recent years for studying mathematical problems arising in materials science. We will discuss some of the first numerical simulations of equilibrium states for martensitic crystals. These numerical simulations have presented many new and challenging problems. In addition, we will describe a research program that has as its goal the development of a continuum theory and a computational model able to predict some of the behavior of shape-memory materials under the influence of loads, boundary displacements, and temperature changes.

ABSTRACT

The Nonlinear Propagation of Waves as Applied to Shear Band Formation: A Numerical Approach

Dr. Dawn A. Lott-Crumpler, University of Maryland at College Park

An efficient and effective numerical method to study the evolution of shear bands in viscoplastic material is presented. Shear bands, localized regions of intense strain, form when ductile solids are subject to large plastic deformation. The phenomenon of flow localization occurs in a wide variety of materials, can be a mechanism for further deformation and can be a precursor to material failure. Shear localization can be initiated by a material inhomogeneity, propagation of shear waves across a crack tip or mar form due to deformation gradients. We present a finite difference method based on Taylor series polynomial expansions and show how this numerical method is used to solve the partial differential equations which model the formation of the shear bands.

ABSTRACT

New Mathematics for New Materials

Dr. David Kinderlehrer, Carnegie Mellon University

The discovery and study of new materials presents exciting challenges to us, the mathematical scientists. Active materials, sometimes called "smart materials", are among these. Their structure transforms dramatically when subjected to heat, stress, electromagnetic fields, or other stimuli. Familiar examples are shape memory materials (to be explained) and pyroelectric materials which form the muscle of infrared remote controls. Additional and more dramatic examples will be presented as well with many demonstrations. The unifying characteristic of these material is their nonlinear behavior. We describe an adventure of discovery and the role of new mathematics in our attempt to understand how and why active material work. This is a high form of interdisciplinary science, involving modeling, analysis, and computational simulation. We invite you to join us in this exhilarating enterprise.

ABSTRACT

Design and Control of the Cz Crystal Growth Process

Alice Livingston, Arizona State University

The fundamental challenge in the modeling of the manufacturing of crystals is the resolution of the complete system of governing equations. These crystals are the critical first step in manufacturing electronic devices. The current industrial standard for growing crystals is the Czochralski (Cz) process. We will describe a Magnetohydrodynamic (MHD) model of the melt motion in the Cz process, and its resulting Finite Element algorithm. The use of the externally applied magnetic field to control the melt motion will also be described.

APPENDIX C

This section contains detailed description of projects mentioned in this report.

- **Computerized Tomography**

A FORTRAN 90 program was developed to implement the Iterative Reconstructive Technique for solving overdetermined systems. The program used a simple calculus formula for determining the orthogonal projection of a point onto a line. This program also was used to aid in the reproduction of a black and white picture. The first step in this process was to digitize the chosen picture. Digitization consists of producing a grid of the picture, assigning values to each box of the grid according to a gray scale, and putting it into matrix form. This matrix was then solved by the generate program.

- **Truck Backer-Upper Problem & Traveling Salesman Problem**

During the Summer of 1996, several students were introduced to two major areas of investigation, the Traveling Salesman Problem and the "Truck Backer-Upper" Problem. Students had the opportunity to come in contact with many different and seemingly unrelated areas of mathematics as they were directed to study papers presenting the different methods of solutions. The adjoining chart lists the different approaches used. All methods of solutions are extremely computer-intensive when applied to realistic size situations but initially we are prototyping these methods on small scale situations. One area that was of interest to students was the one referred to as the conventional approach and the accompanying computer program listing and sample graphical outputs give some sense of the learning process they followed. We are now in the process of porting this prototype to a more sophisticated environment using platforms such as Mathematica. In the immediate future, we will continue to prototype other and more recent methods of solution and then port them all on a common platform where they will become a library of UCES interactive modules in which the topics are fully described, analyzed, and solved.

- **Computational Science Educational Programs**

During the Summer of 1996, faculty members were joined by a group of both graduate and undergraduate students to begin the following:

- create an index of all relevant UCES modules;
- master the use of Yorick to author UCES modules;
- achieve fluency in Computer Symbolic Algebra systems such as MapleV and *Mathematica*;
- conduct educational experiments

Actions to date include reviewing UCES modules to determine those that are relevant for our purposes. Both faculty and students are fully participating in this exercise. Our method of review is to access each UCES module, exercise its capabilities, and abstract key parameters. This database will later be used by students in a Spring Workshop. In addition, our next step is to modify, extend, and create new UCES modules using Yorick.

Temperature Control In Polymer Extrusion Processes†

Max D. Gunzburger
Department of Mathematics

L. Steven Hou
Department of Mathematics and Statistics

James C. Turner

Abstract. An optimization problem is formulated motivated by the desire to obtain uniform extrudate temperature at the die exit in a polymer extrusion process. Control is effected by adjustments to the heat flux along the surface of the pipe. An optimality system of partial differential equations is derived from which optimal controls and states may be determined. Then, finite element discretizations of the optimality system are defined and error estimates are provided along with an efficient solution algorithm for the discrete equations. Finally, computational results are given for a model example with Oldroyd type fluid, demonstrating the effectiveness of our theory and method as well as their potential applicability to industrial problems.

1. Introduction

In polymer extrusion processes, one is often interested in maintaining a quasi-uniform temperature, so as to reduce material (extrusion product) inhomogeneity, throughout the extrudate cooling process. In this paper, we will study the somewhat simplified case: one tries to obtain a uniform temperature distribution at the exit under steady state situation. The means we use to achieve such a uniform temperature distribution at the exit is to adjust the heat flux on the surface of the pipe near the exit.

The extrudate in question will be assumed to be viscoelastic fluid of Oldroyd type such as polymer melts with a fast relaxation mode [10]. Let $\mathbf{u}$ denote the velocity field, p the pressure field, T the temperature field and $\boldsymbol{\tau}$ the purely elastic part of the extra stress. Let $\mathbf{D} = \frac{1}{2}(\text{grad } \mathbf{u} + \text{grad } \mathbf{u}^T)$, $\mathbf{W} = \frac{1}{2}(\text{grad } \mathbf{u} - \text{grad } \mathbf{u}^T)$ and $D_a \boldsymbol{\tau} = (\mathbf{u} \cdot \text{grad}) \boldsymbol{\tau} + \boldsymbol{\tau} \mathbf{W} - \mathbf{W} \boldsymbol{\tau} - a(\mathbf{D} \boldsymbol{\tau} + \boldsymbol{\tau} \mathbf{D})$ where $-1 \leq a \leq 1$. The parameters Re , We and ω are the Reynolds number, Weissenberg number and retardation parameter, respectively. The governing equations for the Oldroyd type fluid, in dimensionless form, is given by the Navier-Stokes equations

$$Re(\mathbf{u} \cdot \text{grad}) \mathbf{u} + \text{grad } p = (1 - \omega) \Delta \mathbf{u} + \text{div } \boldsymbol{\tau} + \mathbf{f}, \quad \text{in } \Omega, \quad (1.1)$$

the incompressibility constraint

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

the constitutive equation (Oldroyd model)

$$\boldsymbol{\tau} + We D_a \boldsymbol{\tau} = 2\omega \mathbf{D} \quad \text{in } \Omega, \quad (1.3)$$

and, for simplicity, the boundary condition

$$\mathbf{u} = \mathbf{h} \quad \text{on } \Gamma, \quad (1.4)$$

and also the energy equation

$$\begin{aligned} -\kappa \Delta T + (\mathbf{u} \cdot \text{grad}) T \\ = \bar{Q} + 2\mu(\text{grad } \mathbf{u} + \text{grad } \mathbf{u}^T) : (\text{grad } \mathbf{u} + \text{grad } \mathbf{u}^T) \quad \text{in } \Omega, \end{aligned} \quad (1.5)$$

with boundary conditions

$$T = 0 \quad \text{on } \Gamma_D, \quad (1.6)$$

$$\frac{\partial T}{\partial n} = H_N \quad \text{on } \Gamma_N \cup \Gamma_O, \quad (1.7)$$

$$\frac{\partial T}{\partial n} = g \quad \text{on } \Gamma_C. \quad (1.8)$$

The data functions $\mathbf{f}$, $\bar{Q}$, H_N , and $\mathbf{h}$ are assumed known; the control g is to be determined so that hot spots are avoided. The constants κ and μ depend on the thermal conductivity coefficient, density, specific heat at constant volume, and viscosity coefficient of the fluid. See [11] for details. We assumed that buoyancy effects can be neglected, and thus the temperature variable does not appear in (1.1).

Two means of obtaining a uniform temperature distribution come to mind. The first is to make the gradient of the temperature along the boundary Γ_o small. Thus, for example, given a velocity field $\mathbf{u}$, we would seek a temperature field T and a control field g such that the functional

$$\mathcal{M}(T, g) = \frac{\alpha}{2} \int_{\Gamma_o} |\text{grad}_s T|^2 d\Gamma + \frac{\kappa\delta}{2} \int_{\Gamma_C} |g|^2 d\Gamma \quad (1.9)$$

is minimized subject, of course, to the constraints imposed by the flow equations (1.5)-(1.8). Here, grad_s denotes the surface gradient operator, *e.g.*, in $\mathbb{R}^2$, the tangential derivative operator $\partial/\partial\tau$. The non-negative parameters α and δ can be used to change the relative importance of the two terms appearing in the definition of $\mathcal{M}$ as well as to act as penalty parameters. The appearance of the control g in the definition of $\mathcal{J}$ is necessary because we are not imposing any *a priori* limits on the size of this control. The minimization of (1.9) results in a quasi-uniform temperature distribution along the boundary segment Γ_o because the surface derivatives of the temperature are forced to be small. Another means of achieving the desired result is to try to directly force the temperature field itself to be quasi-uniform. Thus, now, given a velocity field $\mathbf{u}$, we would seek a temperature field T and a control field g such that the functional

$$\mathcal{J}(T, g) = \frac{1}{2\gamma} \int_{\Gamma_o} |T - T_d|^2 d\Gamma + \frac{\kappa\delta}{2} \int_{\Gamma_C} |g|^2 d\Gamma \quad (1.10)$$

is minimized subject to (1.1)-(1.8), where T_d is some desired temperature distribution, *e.g.*, something close to the average temperature along Γ_C for the uncontrolled system. The non-negative parameters γ and δ can be used to change the relative importance of the two terms appearing in the definition of $\mathcal{J}$ as well as to act as penalty parameters. As will be demonstrated by numerical examples in §6, a small γ is more useful in achieving quasi-uniform boundary temperature distributions, although it also reduces the accuracy of the

approximate solution. We will examine the latter issue in §4. Numerical experiments show that both (1.9) and (1.10) work effectively for the desired objective. We will focus our discussion on (1.10) throughout this paper.

Under the realistic assumption that $\mathbf{u} \cdot \mathbf{n} = 0$ on $\Gamma_C \cup \Gamma_N$, we may prove the existence and uniqueness of optimal solutions and derive an optimality system, *i.e.*, a set of equations from which the optimal control and state may be determined. Also, finite element methods are used to compute an approximate solution of the optimality system. Optimal error estimates are derived and numerical examples are presented. We have also developed an iterative algorithm to compute the approximate solution. The convergence of our algorithm is proved and a comparison with the direct method is made.

Control problems for the fully coupled problem as well as temperature matching for the entire extrudate will be addressed elsewhere.

We close this section by introducing some of the notation used in subsequent sections. Throughout, C will denote a positive constant whose meaning and value changes with context. Also, $H^s(\mathcal{D})$, $s \in \mathbb{R}$, denotes the standard Sobolev space of order s with respect to the set $\mathcal{D}$, where $\mathcal{D}$ is either the flow domain Ω , or its boundary Γ , or part of that boundary. Of course, $H^0(\mathcal{D}) = L^2(\mathcal{D})$. Dual spaces will be denoted by $(\cdot)^*$. Of particular interest will be the space

$$H_D^1(\Omega) = \{S \in H^1(\Omega) : S = 0 \text{ on } \Gamma_D\}.$$

Norms of functions belonging to $H^s(\Omega)$, $H^s(\Gamma)$ and $H^s(\Gamma_C)$ are denoted by $\|\cdot\|_s$, $\|\cdot\|_{s,\Gamma}$ and $\|\cdot\|_{s,\Gamma_C}$, respectively.

The inner product in $L^2(\Omega)$ is denoted by $(\cdot, \cdot)$, that in $L^2(\Gamma)$ by $(\cdot, \cdot)_\Gamma$, that in $L^2(\Gamma_O)$ by $(\cdot, \cdot)_{\Gamma_O}$, and that in $L^2(\Gamma_C)$ by $(\cdot, \cdot)_{\Gamma_C}$. Since, in general, we will use L^2 -spaces as pivot spaces, these notation will also be employed to denote pairings between Sobolev spaces and their duals.

We will use the bilinear form

$$a(T, S) = \int_{\Omega} \text{grad } T \cdot \text{grad } S \, d\Omega \quad \forall T, S \in H^1(\Omega)$$

and the trilinear form

$$c(\mathbf{u}, T, S) = \int_{\Omega} (\mathbf{u} \cdot \text{grad } T) S \, d\Omega \quad \forall \mathbf{u} \in \mathbf{H}^1(\Omega) \quad \text{and} \quad \forall T, S \in H^1(\Omega).$$

These forms are continuous in the sense that there exist constants c_a and $c_c > 0$ such that

$$|a(T, S)| \leq c_a \|T\|_1 \|S\|_1 \quad \forall T, S \in H^1(\Omega) \quad (1.11)$$

and

$$|c(\mathbf{u}, T, S)| \leq c_c \|\mathbf{u}\|_1 \|T\|_1 \|S\|_1 \quad \forall \mathbf{u} \in \mathbf{H}^1(\Omega) \quad \text{and} \quad \forall T, S \in H^1(\Omega). \quad (1.12)$$

Moreover, we have the coercivity property

$$a(T, T) \geq C_a \|T\|_1^2 \quad \forall T \in H_0^1(\Omega) \quad (1.13)$$

for some constant $C_a > 0$.

For details concerning the notation employed and the inequalities (1.11)-(1.13), one may consult, *e.g.*, [1] and [7].

2. The Optimization Problem, Existence of Solutions, And Optimality System

We begin by giving a precise statement of the optimization problem we consider. We will assume the domain Ω is a polygon in $\mathbf{R}^2$. We first recall that (1.1)-(1.4) uncouples from (1.5)-(1.8). We may solve for $(\mathbf{u}, p\tau)$ from (1.1)-(1.4) once and for all and then plug them into (1.5)-(1.8). Thus the only state variable is T , i.e., the temperature field, and the only boundary control variable is g . The state and control variables are constrained to satisfy the system (1.5)-(1.8), which we recast into the following weak form: find $T \in H_D^1(\Omega)$ such that

$$\kappa a(T, S) + c(\mathbf{u}, T, S) = (Q, S) + \kappa(g, S)_{\Gamma_C} + \kappa(H_N, S)_{\Gamma_N} \quad \forall S \in H_D^1(\Omega), \quad (2.1)$$

where we have introduced the simplifying notation

$$Q = \bar{Q} + 2\mu(\text{grad } \mathbf{u} + \text{grad } \mathbf{u}^T) : (\text{grad } \mathbf{u} + \text{grad } \mathbf{u}^T).$$

Note that since we seek $T \in H_D^1(\Omega)$,

$$T = 0 \quad \text{on } \Gamma_D. \quad (2.2)$$

Throughout, we will assume that the given velocity field $\mathbf{u}$ is smooth and satisfies

$$\text{div } \mathbf{u} = 0 \quad \text{in } \Omega \quad \text{and} \quad \mathbf{u} \cdot \mathbf{n} \geq 0 \quad \text{a.e. on } \Gamma_C \cup \Gamma_N. \quad (2.3)$$

Under these assumptions, we have the useful relation

$$c(\mathbf{u}, S, S) = \frac{1}{2} \int_{\Gamma} (\mathbf{u} \cdot \mathbf{n}) S^2 d\Gamma \geq 0 \quad \forall S \in H^1(\Omega), \quad (2.4)$$

which can be derived by setting $T = S$ in the following integration by parts formula:

$$c(\mathbf{u}, T, S) = \int_{\Gamma} (\mathbf{u} \cdot \mathbf{n}) TS d\Gamma - c(\mathbf{u}, S, T).$$

For each possible control function g , there exists a unique corresponding state function T .

Lemma 2.1— For every $g \in L^2(\Gamma_C)$, there exists a unique $T \in H_D^1(\Omega)$ such that (2.1) is satisfied. Moreover,

$$\|T\|_1 + \|T\|_{0, \Gamma_C} \leq C(\|g\|_{0, \Gamma_C} + \|Q\|_0 + \|H_N\|_{0, \Gamma_N}). \quad (2.5) \blacksquare$$

The admissibility set $\mathcal{U}_{ad}$ is defined by

$$\mathcal{U}_{ad} = \{(T, g) \in H_D^1(\Omega) \times L^2(\Gamma_C) : \mathcal{J}(T, g) < \infty, (2.1) \text{ is satisfied}\}. \quad (2.6)$$

Then, $(\hat{T}, \hat{g}) \in \mathcal{U}_{ad}$ is called an *optimal solution* if there exists $\epsilon > 0$ such that

$$\mathcal{J}(\hat{T}, \hat{g}) \leq \mathcal{J}(T, g) \quad \forall (T, g) \in \mathcal{U}_{ad} \text{ satisfying } \|T - \hat{T}\|_1 + \|g - \hat{g}\|_{0, \Gamma_C} \leq \epsilon. \quad (2.7)$$

Based on the previous lemma, we can show the existence and uniqueness of optimal solutions.

Theorem 2.2— There exists a unique optimal solution $(\hat{T}, \hat{g}) \in \mathcal{U}_{ad}$. \blacksquare

Using techniques in e.g. [9] we may obtain the optimality condition

$$g = -\frac{1}{\delta} \Phi|_{\Gamma_C} \quad (2.8)$$

where Φ is the solution of the adjoint state equation

$$\kappa a(R, \Phi) + c(\mathbf{u}, R, \Phi) - \frac{1}{\gamma} (R, T - T_d)_{\Gamma_C} = 0 \quad \forall R \in H_D^1(\Omega). \quad (2.9)$$

Eliminating g from (2.1) and combining with (2.9), we obtain the *optimality system*

$$\kappa a(T, S) + c(\mathbf{u}, T, S) + \frac{\kappa}{\delta} (\Phi, S)_{\Gamma_C} = (Q, S) + \kappa (H_N, S)_{\Gamma_N} \quad \forall S \in H_D^1(\Omega) \quad (2.10)$$

and

$$\kappa a(R, \Phi) + c(\mathbf{u}, R, \Phi) - \frac{1}{\gamma} (R, T - T_d)_{\Gamma_C} = 0 \quad \forall R \in H_D^1(\Omega). \quad (2.11)$$

Thus, the optimal state, *i.e.*, the temperature distribution T , can be found by solving the coupled system (2.10)-(2.11), which also provides the optimal co-state Φ . The optimal control g can then be deduced from (2.8).

3. Finite Element Approximation, Error Estimates, and Iterative Methods

In the usual manner, one may construct finite element subspaces $W^h \subset H_D^1(\Omega) \cap C(\overline{\Omega})$ parametrized by a parameter h that tends to zero. (In practice, h is, of course, related to a grid size.) We assume the approximation property (see [3]): there exist an integer k and a constant C such that

$$\inf_{S^h \in W^h} \|S - S^h\|_1 \leq Ch^m \|S\|_{m+1} \quad \forall S \in H_D^1(\Omega) \text{ and } 0 \leq m \leq k. \quad (3.1)$$

A finite element algorithm for determining approximations of the solution of the optimality system (2.9)-(2.10) is then defined as follows: seek $T^h \in W^h$ and $\Phi^h \in W^h$ such that

$$\begin{aligned} \kappa a(T^h, S^h) + c(\mathbf{u}, T^h, S^h) + \frac{\kappa}{\delta} (\Phi^h, S^h)_{\Gamma_C} \\ = (Q, S^h) + \kappa (H_N, S^h)_{\Gamma_N} \quad \forall S^h \in W^h \end{aligned} \quad (3.2)$$

and

$$\kappa a(R^h, \Phi^h) + c(\mathbf{u}, R^h, \Phi^h) - \frac{1}{\gamma} (R^h, T^h - T_d)_{\Gamma_C} = 0 \quad \forall R^h \in W^h. \quad (3.3)$$

Although the optimality system is linear, the coupling of $\mathbf{u}$ and Φ in the two equations make the derivation of error estimates nontrivial. It turns out to be convenient to apply the *Brezzi-Rappaz-Raviart* theory (see [2], [5], and [7]) to obtain error estimates.

Theorem 3.1- *Let (T, Φ) and (T^h, Φ^h) be the solutions of (2.9)-(2.10), and (3.2)-(3.3), respectively. Assume that $T, \Phi \in H^{m+1}(\Omega) \cap H_D^1(\Omega)$ for some $1 \leq m \leq k$; also assume that (3.1) holds. Then,*

$$\begin{aligned} \|T - T^h\|_1 + \|\Phi - \Phi^h\|_1 \\ \leq C \max \left\{ \frac{1}{\delta}, \frac{1}{\gamma}, 1 \right\} h^m (\|Q\|_{m-1} + \|H_N\|_{\Gamma_N, m-1/2} + \|T_d\|_{\Gamma_C, m-1/2}), \end{aligned}$$

where C is independent of h, δ, γ, T , and Φ . ■

A simple iterative algorithm for solving (2.10)-(2.11) can be defined as follows:

choose $\Phi^{(0)}$;

for $n = 1, 2, \dots$, solve for $T^{(n)}$ from

$$\begin{aligned} \kappa a(T^{(n)}, S) + c(u, T^{(n)}, S) \\ = -\frac{\kappa}{\delta}(\Phi^{(n-1)}, S)_{\Gamma_C} + (Q, S) + \kappa(H_N, S)_{\Gamma_N} \quad \forall S \in H_D^1(\Omega); \end{aligned} \quad (3.4)$$

then solve for $\Phi^{(n)}$ from

$$\kappa a(R, \Phi^{(n)}) + c(u, R, \Phi^{(n)}) = \frac{1}{\gamma}(R, T^{(n)} - T_d)_{\Gamma_C} \quad \forall R \in H_D^1(\Omega). \quad (3.5)$$

Of course, ultimately, this algorithm has to be carried out in a discretized version, such as one using a finite element method.

The convergence of this algorithm can be proved as a result of the observation that it is effectively a gradient method for the following minimization problem: find $g \in L^2(\Gamma_C)$ such that $\mathcal{K}(g) := \mathcal{J}(T(g), g)$ is minimized where $T(g) \in H_D^1(\Omega)$ is defined as the solution of (2.1).

Theorem 5.2- *Let $(T^{(n)}, \Phi^{(n)})$ be the solution of (3.4)-(3.5) and (T, Φ) the solution of (2.9)-(2.10). Then, $T^{(n)} \rightarrow T$ in $H_D^1(\Omega)$ and $\Phi^{(n)} \rightarrow \Phi$ in $H_D^1(\Omega)$ as $n \rightarrow \infty$. ■*

4. Computational Examples

Let $\Omega \subset \mathbb{R}^2$ be the unit square $(0, 1) \times (0, 1)$. Let $\Gamma = \Gamma_C \cup \Gamma_D \cup \Gamma_N \cup \Gamma_O$ be shown as in Figure 1.

Figure 1. Computational domain

The finite element spaces W^h are chosen to be piecewise linear elements on a triangular mesh. All the numerical results make use of the following parameters and data:

parameters: $\kappa = 1/0.73$; $Re = 1$; $We = 1$; $\omega = 1/2$;

boundary data: $T = 1$ on Γ_D , $\frac{\partial T}{\partial n} = 0$ on $\Gamma_N \cup \Gamma_O$;

heat source:
$$Q = -\frac{5\kappa}{2} \left(9\pi^2 \cos(3\pi x) \cos^2(\pi y) - 4\pi^2 \sin^2\left(\frac{3}{2}\pi x\right) \cos(2\pi y) \right) \\ + (1 - y^2) \frac{15\pi}{2} \sin(3\pi x) \cos^2(\pi y);$$

velocity profile:

$$u = (1 - y^2, 0),$$

elastic extra stress profile:

$$\tau = \begin{pmatrix} 0 & -y \\ -y & -4y^2 \end{pmatrix},$$

pressure profile:

$$p = 4(1 - y^2) - x.$$

In functional (1.10), we choose

$$T_d = 3.5.$$

For the data given above, the exact solution of the uncontrolled problem, i.e., for

$$\frac{\partial T}{\partial n} = 0 \quad \text{on } \Gamma_C,$$

is given by $T = 5 \sin^2(\frac{3}{2}\pi x) \cos^2(\pi y) + 1$.

We compare the temperature distribution in the uncontrolled case with the optimal temperature distribution in the controlled case for which

$$\frac{\partial T}{\partial n} = g \quad \text{on } \Gamma_C,$$

where g is the control such that (1.10) is minimized. Approximations to the optimal state and co-state are computed from (3.2)-(3.3); the approximate optimal control g^h is then obtained from (2.8), i.e., $g^h = -(1/\delta)\Phi^h|_{\Gamma_C}$. All of the computational results shown below were obtained with the use of a mesh size $h = \frac{1}{19}$. Of course, calculations with varying mesh sizes were performed. Since these merely verified the error estimates, we do not report on them here.

Specifically, Figures 2-4 deal with the following cases:

1. Exact, uncontrolled temperature and; Optimal temperature;
2. Exact, uncontrolled temperature and; Optimal temperature;
2. Optimal boundary control.

Figure 2. Temperature contours: uncontrolled and controlled.

Optimal boundary control on Γ_C .

(Γ_C is the top boundary segment.)

(Γ_O is the right boundary segment.)

Figure 3. Temperature surfaces: uncontrolled and controlled.
 (Γ_C is the top-right boundary segment.)
 (Γ_O is the lower-right boundary segment.)

Figure 4. Optimal boundary control on Γ_C .

References

- [1] R. ADAMS, *Sobolev Spaces*, Academic, New York, 1975.
- [2] 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.
- [3] P. CIARLET, *The Finite Element Method for Elliptic Problems*, North-Holland, Amsterdam, 1978.
- [4] P. CIARLET, *Introduction to Numerical Linear Algebra and Optimisation*, Cambridge, Cambridge, 1989.
- [5] M. CROUZIEUX and J. RAPPAZ, *On Numerical Approximation in Bifurcation Theory*, Masson, Paris, 1989.
- [6] P. CUVELIER, Optimal control of a system governed by the Navier-Stokes equations coupled with the heat equations, *New Developments in Differential equations* (Ed. by W. Eckhaus), North-Holland, Amsterdam, 1976, 81-98.
- [7] V. GIRAULT and P.-A. RAVIART, *Finite Element Methods for Navier-Stokes Equations*, Springer, Berlin, 1986.
- [8] M. GUNZBURGER, L. HOU and T. SVOBODNY, Control of temperature distributions along boundaries of engine components, *Numerical Methods for Laminar and Turbulent Flows - Proceedings of the 7th International Conference on Numerical Methods for Turbulent and Laminar Flows*, Stanford, July, 1991.
- [9] J.-L. LIONS *Contrôle Optimal de Systèmes Gouvernés par des Equations aux Dérivées Partielles*, Dunod, Paris, 1968.
- [10] J.G. OLDROYD, On the formation of rheological equations of state, *Proc. Roy. Soc. London*, A **200**, 1950, 523-541.
- [11] J. SERRIN, Mathematical principles of classical fluid mechanics, *Handbüch der Physik VIII/1* (ed. by S. Flügge and C. Truesdell), Springer, Berlin, 1959, 125-263.

ANALYSIS AND FINITE ELEMENT APPROXIMATION OF AN OPTIMAL CONTROL PROBLEM IN ELECTROCHEMISTRY WITH CURRENT DENSITY CONTROLS

L. STEVEN HOU[†] AND JAMES C. TURNER[‡]

ABSTRACT. An optimal control problem for impressed cathodic systems in electrochemistry is studied. The control in this problem is the current density on the anode. A matching objective functional is considered; (many other objective functionals can be similarly treated.) The existence of an optimal solution is proved. The use of Lagrange multiplier rules is justified and an optimality system of equations is established. Finally, a finite element algorithm is defined and optimal error estimates are derived.

Key words. Optimal control, impressed cathodic system, electrochemistry, nonlinear boundary condition, finite element method, error estimate

AMS(MOS) subject classifications. 49J20, 49K20, 65K10, 65N15, 65N30

1. INTRODUCTION

We consider an optimal control problem for impressed cathodic systems. A typical example of an impressed cathodic system is a metal container filled with an electrolyte. The painted portion of the container surface is usually treated as insulated. The unpainted part is divided into cathode and anode which are connected to the negative and positive poles of an electrical source, respectively. By adjusting the current density on the anode we could effectively alter the potential distribution on the entire bounding surface or in the entire flow domain. The potential distribution, of course, has a direct effect on the chemical reaction process occurring inside the flow domain, which in turn has an effect on the rate of corrosion of the metal container. Thus the current density on the anode can be used as a practical control variable for generating a desired potential field. This idea can be conveniently formulated as optimal control problems for the potential equation with appropriate boundary conditions. Optimal control problems of this sort has been studied in [20] and [21] where the goal was to match a desired potential distribution

[†]Department of Mathematics and Statistics,

on the cathode. The models analyzed in [20] and [21] are essentially linear. [14] discussed, mainly from an algorithmic point of view, several control mechanisms including adjusting the positions of anodes and/or the current density on the anodes in order to best match a desired potential on the structure surface; nonlinear models were employed as well as linear ones. [15] analyzed a "location control" problem, i.e., the control variable is the location of anodes, wherein nonlinear models with boundary conditions of polynomial or mixed polynomial-exponential growth type were considered. In this article, we will attempt to mathematically analyze optimal control problems with current density controls. The nonlinear model used involves an exponentially growing boundary condition.

We assume the electrolyte occupies a physical domain $\Omega \in \mathbb{R}^2$ with a boundary Γ . The domain is assumed to be finite in this paper, although infinite domain problems can be handled if appropriate decay rate at infinity is assumed. If $\Omega \subset \mathbb{R}^3$, we will need to work with a non-Hilbert space $W^{1,p}(\Omega)$ with, e.g., $p = 3$; similar results can still be obtained.

The electrical potential ϕ in Ω is governed by the differential equation

$$-\operatorname{div}(\sigma \operatorname{grad} \phi) = 0 \quad \text{in } \Omega,$$

where the conductivity σ is a continuous function with a positive lower bound.

The boundary Γ is divided into three components: the anode Γ_A , the cathode Γ_C and the insulated part Γ_0 . On the cathode Γ_C , ϕ satisfies the relation

$$\sigma \frac{\partial \phi}{\partial n} = -f(\phi) \quad \text{on } \Gamma_C,$$

where f is an empirical function that depends on the electrode materials (see [4]). In particular we will assume f is given by the Butler-Volmer function:

$$f(\phi) = C_3[e^{C_1\phi} - e^{-C_2\phi}] \quad (1.1)$$

where C_1 , C_2 and C_3 are positive constants (see [4]). Throughout this paper, f will be assumed to be defined by (1.1). For notational convenience, we will mainly use $f(\phi)$ rather than the explicit expression given in (1.1).

FIG. 1. A typical impressed cathodic system: an electrolyte container connected with an electrical current source

On the anode Γ_A , we have the boundary condition

$$\sigma \frac{\partial \phi}{\partial n} = u \quad \text{on } \Gamma_A,$$

which corresponds to the specification of the current density on the anode. Adjusting the current density on Γ_A amounts to treating u as a control variable. On the insulated part Γ_0 ,

$$\sigma \frac{\partial \phi}{\partial n} = 0 \quad \text{on } \Gamma_0.$$

We are concerned with the following optimal control problem: seek a state ϕ and a control u such that the functional

$$\mathcal{J}(\phi, u) = \frac{1}{2\epsilon_0} \int_{\Omega} (\phi - \phi_0)^2 d\Omega + \frac{\delta_0}{2} \int_{\Gamma_A} u^2 d\Gamma, \quad (1.2)$$

is minimized subject to the constraint equations

$$-\operatorname{div}(\sigma \operatorname{grad} \phi) = 0 \quad \text{in } \Omega \quad (1.3)$$

$$\sigma \frac{\partial \phi}{\partial n} = u \quad \text{on } \Gamma_A, \quad (1.4)$$

$$\sigma \frac{\partial \phi}{\partial n} = 0 \quad \text{on } \Gamma_0 \quad (1.5)$$

and

$$\sigma \frac{\partial \phi}{\partial n} = -f(\phi) \quad \text{on } \Gamma_C. \quad (1.6)$$

In (1.2) ϕ_0 is a desired potential distribution in Ω and ϵ_0 and δ_0 are positive constants.

We will make use of an equivalent variational formulation (1.7) of the nonlinear boundary value problem (1.3)-(1.6). We will utilize Sobolev spaces $H^m(\Omega)$, $H^s(\Gamma_A)$, $H^s(\Gamma_C)$, $H^s(\Gamma_0)$ and $H^s(\Gamma)$. The corresponding norms on these spaces will be denoted by, e.g., $\|\cdot\|_m$, $\|\cdot\|_{s,\Gamma_A}$, etc. For details, see [1] and [8]. A weak formulation of (1.3)-(1.6) is given as follows: seek a $\phi \in H^1(\Omega)$ such that

$$\int_{\Omega} \sigma \operatorname{grad} \phi \cdot \operatorname{grad} \psi d\Omega + \int_{\Gamma_C} f(\phi) \psi d\Gamma = \int_{\Gamma_A} u \psi d\Gamma, \quad \forall \psi \in H^1(\Omega). \quad (1.7)$$

Equation (1.7) can be formally shown to be equivalent to the nonlinear boundary value problem (1.3)-(1.6). We mention that second order elliptic differential equations with exponentially growing coefficients were studied in, among others, [9], [12], [13] and [16]. An elliptic equation with mixed Dirichlet-Neumann type boundary conditions that have an exponentially growing coefficient in the boundary condition was studied [7]. Some of the techniques in these articles are useful for the mathematical and numerical analysis of the state equation (1.7).

We restate the minimization problem as follows:

$$\begin{aligned} &\text{seek a state } \phi \in H^1(\Omega) \text{ and a } u \in U \text{ such that the} \\ &\text{functional (1.2) is minimized subject to (1.7),} \end{aligned} \quad (1.8)$$

where

$$U \text{ is a non-empty, closed, convex subset of } L^2(\Gamma_A). \quad (1.9)$$

Now we state a few useful facts. We set $\alpha = \frac{1}{2}C_3 \min\{C_1, C_2\}$. Then by Mean Value Theorem we obtain

$$f(\phi)\phi = C_3(C_1 + C_2)e^{\tilde{\phi}}\phi^2 \geq 2\alpha\phi^2 \quad \forall \phi \quad (1.10)$$

where $\tilde{\phi}$ is between $(C_1\phi)$ and $(-C_2\phi)$. It is also easy to see that

$$f'(\phi) \geq 2\alpha \quad \forall \phi. \quad (1.11)$$

The norm on $H^1(\Omega)$ defined by

$$|||\phi|||_1 = \left\{ \int_{\Omega} \sigma |\text{grad } \phi|^2 d\Omega + 2\alpha \int_{\Gamma_C} \phi^2 d\Gamma \right\}^{1/2} \quad \forall \phi \in H^1(\Omega)$$

is equivalent to the usual $H^1(\Omega)$ -norm $\|\cdot\|_1$, i.e., there exist constants $\rho > 0$ and $\gamma > 0$ such that

$$\rho \|\phi\|_1^2 \geq \int_{\Omega} \sigma |\text{grad } \phi|^2 d\Omega + 2\alpha \int_{\Gamma_C} \phi^2 d\Gamma \geq \gamma \|\phi\|_1^2 \quad \forall \phi \in H^1(\Omega). \quad (1.12)$$

A proof of (1.12) can found in, e.g., [17].

The rest of the paper is organized as follows. In §2 we prove the existence and uniqueness of solutions to (1.7) so that the constraint equation is well-posed. In §3 we show the existence of an optimal pair $(\hat{\phi}, \hat{u})$ that minimizes (1.1) subject to (1.7). In §4 we justify the use of Lagrange multiplier rules and derive an optimality condition. In §5 we discuss the regularity of optimal solutions. Finally in §6 we define a finite element algorithm for solving the optimality system and derive optimal error estimates.

2. EXISTENCE AND UNIQUENESS OF SOLUTIONS TO THE CONSTRAINT EQUATIONS

We first examine the existence of a solution to the nonlinear Neumann type boundary value problems (1.7).

Lemma 2.1 *X be a finite-dimensional Hilbert space whose scalar product is denoted by $(\cdot, \cdot)$ and the corresponding norm by $|\cdot|$. Let F be a continuous mapping from X into X with the following property: there exists an $r > 0$ such that*

$$(F(\phi), \phi) \geq 0 \quad \forall \phi \in X \text{ with } |\phi| = r.$$

Then there exists an $\phi \in X$ such that

$$F(\phi) = 0 \quad \text{and} \quad |\phi| \leq r.$$

Proof. See [Raviart], pp. 279. ■

Lemma 2.2 Assume $\phi \in H^1(\Omega)$ and $s > 0$. Then $e^{s|\phi|} \in L^1(\Gamma)$. Moreover, there exists a constant κ , independent of ϕ , such that

$$\int_{\Gamma} e^{s|\phi|} d\Gamma \leq 1 + |\Gamma| + e^{s^2 \kappa^2 \|\phi\|_1^2} |\Gamma| < \infty,$$

where $|\Gamma|$ is the measure of Γ .

Proof. Let $\phi \in H^1(\Omega)$ and $s > 0$ be given. A Sobolev embedding theorem implies $\phi \in H^{1/2}(\Gamma)$. Using embedding results for Orlicz-Sobolev spaces (see [1], [10] and [18]) (recall $\Omega \subset \mathbb{R}^2$), we have $H^{1/2}(\Gamma) \hookrightarrow L_A(\Gamma)$ where the N -function $A(t) = e^{t^2} - 1$. Thus there exists a constant $\kappa > 0$ such that

$$\|\phi\|_{L_A(\Gamma)} \equiv \inf \left\{ k : \int_{\Gamma} e^{\frac{|\phi|^2}{k^2}} d\Gamma \leq 1 \right\} \leq \kappa \|\phi\|_1 < \infty.$$

Hence for each sufficiently $\epsilon > 0$, the constant $k \equiv \|\phi\|_{L_A(\Gamma)} + \epsilon$ satisfies

$$\int_{\Gamma} [e^{\frac{|\phi|^2}{k^2}} - 1] d\Gamma \leq 1 \quad \text{so that} \quad \int_{\Gamma} e^{\frac{|\phi|^2}{k^2}} d\Gamma \leq 1 + |\Gamma|.$$

We set $M = sk^2$. By an elementary calculation we can show that

$$e^{sx} < e^{\frac{x^2}{k^2}} \quad \forall |x| > M.$$

We set $K = e^{sM} = e^{s^2 k^2} < \infty$. Then

$$\begin{aligned} \int_{\Gamma} e^{s|\phi|} d\Gamma &= \int_{\{\mathbf{x} \in \Gamma : |\phi(\mathbf{x})| \geq M\}} e^{s|\phi|} d\Gamma + \int_{\{\mathbf{x} \in \Gamma : |\phi(\mathbf{x})| < M\}} e^{s|\phi|} d\Gamma \\ &\leq \int_{\Gamma} e^{\frac{|\phi|^2}{k^2}} d\Gamma + K |\Gamma| \leq 1 + |\Gamma| + e^{s^2(\|\phi\|_{L_A(\Gamma)} + \epsilon)^2} |\Gamma| \\ &\leq 1 + |\Gamma| + e^{s^2(\kappa \|\phi\|_1 + \epsilon)^2} |\Gamma| < \infty. \end{aligned}$$

Letting $\epsilon \rightarrow 0$ yield the desired result. ■

Lemma 2.3 Assume $\{\phi_n\} \subset L^2(\Gamma_C)$ is a sequence such that $\phi_n \rightarrow \phi$ a.e. on Γ_C and

$$\int_{\Gamma_C} f(\phi_n) \phi_n d\Gamma \leq B \quad \forall n \quad (2.1)$$

where f is defined by (1.1) and $B > 0$ is a constant independent of n . Then

$$\int_{\Gamma_C} f(\phi) \phi d\Gamma \leq \liminf_{n \rightarrow \infty} \int_{\Gamma_C} f(\phi_n) \phi_n d\Gamma$$

and

$$\lim_{n \rightarrow \infty} \int_{\Gamma_C} |f(\phi_n) - f(\phi)| d\Gamma = 0.$$

Proof. The proof follows the ideas of [13], pp. 21-22. Since f is continuous and $\phi_n \rightarrow \phi$ a.e. on Γ_C , we deduce that $f(\phi_n) \rightarrow f(\phi)$ a.e. on Γ_C . Note that $f(\phi)\phi \geq 0$ on Γ_C so that we may use Fatou's Lemma to obtain

$$\int_{\Gamma_C} f(\phi)\phi \, d\Gamma \leq \liminf_{n \rightarrow \infty} \int_{\Gamma_C} f(\phi_n)\phi_n \, d\Gamma \leq B.$$

Hence $f(\phi)\phi \in L^1(\Omega)$. By setting $K = \sup_{|x| \leq 1} |f(x)|$ we easily conclude from the identity

$$|f(t)| = |t|^{-1} f(t)t \quad \forall t \neq 0 \quad (2.2)$$

that

$$|f(t)| \leq f(t)t + K \quad \forall t \in \mathbf{R}.$$

Thus

$$|f(\phi)| \leq |f(\phi)| |\phi| + K \quad \text{on } \Gamma_C,$$

i.e., $f(\phi) \in L^1(\Gamma_C)$. Utilizing (2.2) again, we deduce that for each $\delta > 0$ and for a.e. $x \in \Gamma_C$, we have either

$$|\phi_n| \leq \delta^{-1} \quad \text{or} \quad |f(\phi)| \leq \delta f(\phi_n)\phi_n$$

so that

$$|f(\phi)| \leq C_\delta + \delta f(\phi_n)\phi_n \quad \text{on } \Gamma_C,$$

where $C_\delta = \sup_{|x| \leq \delta^{-1}} |f(x)|$. For every measurable subset $S \subset \Gamma_C$ we have

$$\int_S |f(\phi_n)| \, d\Gamma \leq C_\delta |S| + \delta \int_{\Gamma_C} f(\phi_n)\phi_n \, d\Gamma.$$

Equation (2.1) implies

$$\int_S f(\phi_n)\phi_n \, d\Gamma \leq 2B$$

for n greater than some $N_0 > 0$. Thus

$$\int_S |f(\phi_n)| \, d\Gamma \leq C_\delta |S| + 2B\delta \quad \forall n > N_0,$$

where $|S|$ is the measure of S . Hence, the sequence of functions $\{f(\phi_n)\}$ has equi-absolutely continuous integrals. By Vitali's Convergence Theorem,

$$\lim_{n \rightarrow \infty} \int_{\Gamma_C} |f(\phi_n) - f(\phi)| \, d\Gamma = 0. \quad \blacksquare$$

Theorem 2.4 Assume $u \in L^2(\Gamma_A)$. Then there exists a unique $\phi \in H^1(\Omega)$ that satisfies (1.7). Furthermore, ϕ satisfies the estimate

$$\|\phi\|_1 \leq \frac{\beta}{\gamma} \{\|u\|_{0,\Gamma_A} + 1\}, \quad (2.3)$$

where β and γ are constants independent of ϕ .

Proof. We first establish an inequality

$$\begin{aligned} \int_{\Omega} \sigma \operatorname{grad} \phi \cdot \operatorname{grad} \phi \, d\Omega + \int_{\Gamma_C} f(\phi) \phi \, d\Gamma - \int_{\Gamma_A} u \phi \, d\Gamma &\geq \\ &\geq 0 \quad \forall \phi \in H^1(\Omega) \text{ with } \|\phi\|_1 = r \end{aligned}$$

for some $r > 0$. Using (1.10) we have that

$$\int_{\Gamma_C} f(\phi) \phi \, d\Gamma \geq \alpha \int_{\Gamma_C} \phi^2 \, d\Gamma \quad \forall \phi \in H^1(\Omega).$$

Hence it follows from (1.12) that

$$\int_{\Omega} \sigma \operatorname{grad} \phi \cdot \operatorname{grad} \phi \, d\Omega + \int_{\Gamma_C} f(\phi) \phi \, d\Gamma \geq \gamma \|\phi\|_1^2 \quad \forall \phi \in H^1(\Omega).$$

Using Cauchy-Schwartz inequality and trace theorems we obtain the estimate

$$\left| \int_{\Gamma_A} u \phi \, d\Gamma \right| \leq \beta \|\phi\|_1 \|u\|_{0,\Gamma_A} \quad \forall \phi \in H^1(\Omega),$$

where β is a positive constant. By combining the last two estimates we deduce that for $r = \frac{\beta}{\gamma} \{\|u\|_{0,\Gamma_A} + 1\} > 0$ we have

$$\begin{aligned} \int_{\Omega} \sigma \operatorname{grad} \phi \cdot \operatorname{grad} \phi \, d\Omega + \int_{\Gamma_C} f(\phi) \phi \, d\Gamma - \int_{\Gamma_A} u \phi \, d\Gamma &\geq \\ &\geq 0 \quad \forall \phi \in H^1(\Omega) \text{ with } \|\phi\|_1 = r. \end{aligned} \tag{2.4}$$

Since $H^1(\Omega)$ is separable, we may choose a countable basis of $H^1(\Omega)$: $\{\psi_i\}_{i=1}^{\infty}$. We set $X_n = \operatorname{span}\{\psi_1, \dots, \psi_n\}$. The inner product and norm on each X_n is defined by that of $H^1(\Omega)$ restricted to X_n . We introduce the mapping $F_n : X_n \rightarrow X_n$ as follows. For each $\phi \in H^1(\Omega)$, $F_n(\phi) \in X_n$ is defined by

$$(F_n(\phi), \psi_j) = \int_{\Omega} \sigma \operatorname{grad} \phi \cdot \operatorname{grad} \psi_j \, d\Omega + \int_{\Gamma_C} f(\phi) \psi_j \, d\Gamma - \int_{\Gamma_A} u \psi_j \, d\Gamma \quad 1 \leq j \leq n.$$

It follows from (2.4) and Lemma 2.1 that the finite dimensional problem

$$\int_{\Omega} \operatorname{grad} \phi_n \cdot \operatorname{grad} \psi \, d\Omega + \int_{\Gamma_C} f(\phi_n) \psi \, d\Gamma = \int_{\Gamma_A} u \psi \, d\Gamma, \quad \forall \psi \in X_n \tag{2.5}$$

has a solution $\phi_n \in X_n$ with a bound

$$\|\phi_n\|_1 \leq \frac{\beta}{\gamma} \{\|u\|_{0,\Gamma_A} + 1\}. \tag{2.6}$$

We can extract a subsequence of $\{\phi_n\}$, still denoted by $\{\phi_n\}$, that converges weakly to some $\phi \in H^1(\Omega)$ as $n \rightarrow \infty$. Then $\{\phi_n\}$ also converges weakly in $H^{1/2}(\Gamma)$ by a trace theorem. Thus $\{\phi_n\}$ converges strongly in $L^2(\Gamma)$ by compact imbedding. This in turn implies a subsequence satisfies $\phi_n \rightarrow \phi$ a.e. on Γ . By setting $\psi = \phi_n$ in (2.5) we obtain

$$\begin{aligned} \int_{\Gamma_C} f(\phi_n) \phi_n d\Gamma &\leq \|u\|_{0,\Gamma_A} \|\phi_n\|_{0,\Gamma_A} \\ &\leq C \|u\|_{0,\Gamma_A} \|\phi_n\|_1 \leq C \|u\|_{0,\Gamma_A} \frac{\beta}{\gamma} \{\|u\|_{0,\Gamma_A} + 1\}. \end{aligned}$$

By Lemma 2.3 we have that

$$\int_{\Gamma_C} f(\phi) \phi d\Gamma \leq \liminf_{n \rightarrow \infty} \int_{\Gamma_C} f(\phi_n) \phi_n d\Gamma$$

and

$$\lim_{n \rightarrow \infty} \int_{\Gamma_C} |f(\phi_n) - f(\phi)| d\Gamma = 0.$$

By passing to the limit in (2.5) for each $\psi \in C^\infty(\bar{\Omega})$ we see that

$$\int_{\Omega} \sigma \operatorname{grad} \phi \cdot \operatorname{grad} \psi d\Omega + \int_{\Gamma_C} f(\phi) \psi d\Gamma = \int_{\Gamma_A} u \psi d\Gamma, \quad \forall \psi \in C^\infty(\bar{\Omega}). \quad (2.7)$$

We next prove that this ϕ is a solution to (1.7). For each $\psi \in H^1(\Omega)$, we may choose a sequence $\{\psi_k\} \subset C^\infty(\bar{\Omega})$ such that $\|\psi - \psi_k\|_1 \rightarrow 0$ as $k \rightarrow \infty$. Using (2.7) we have

$$\int_{\Omega} \sigma \operatorname{grad} \phi \cdot \operatorname{grad} \psi_k d\Omega + \int_{\Gamma_C} f(\phi) \psi_k d\Gamma = \int_{\Gamma_A} u \psi_k d\Gamma, \quad \forall k. \quad (2.8)$$

Lemma 2.2 implies $e^{C_1 \phi} \in L^2(\Gamma_C)$ and $e^{-C_2 \phi} \in L^2(\Gamma_C)$ so that $\int_{\Gamma_C} |f(\phi)|^2 d\Gamma < \infty$. Hence,

$$\left| \int_{\Gamma_C} f(\phi) (\psi - \psi_k) d\Gamma \right| \leq \|f(\phi)\|_{0,\Gamma_C} \|\psi - \psi_k\|_{0,\Gamma_C} \leq C \|f(\phi)\|_{0,\Gamma_C} \|\psi - \psi_k\|_1$$

so that

$$\int_{\Gamma_C} f(\phi) (\psi - \psi_k) d\Gamma \rightarrow 0 \quad \text{as } k \rightarrow \infty.$$

Thus we may pass to the limit in (2.7) to show (1.7) holds. The estimate (2.3) follows easily from (2.6).

To answer the question of uniqueness, we assume ϕ and $\bar{\phi}$ are two solutions to (1.7). Then we have

$$\int_{\Omega} \sigma \operatorname{grad}(\phi - \bar{\phi}) \cdot \operatorname{grad} \psi d\Omega + \int_{\Gamma_C} [f(\phi) - f(\bar{\phi})] \psi d\Gamma = 0, \quad \forall \psi \in H^1(\Omega).$$

Setting $\psi = \phi - \bar{\phi}$, we see that

$$\int_{\Omega} \sigma |\text{grad}(\phi - \bar{\phi})|^2 d\Omega + \int_{\Gamma_C} [f(\phi) - f(\bar{\phi})](\phi - \bar{\phi}) d\Gamma = 0.$$

Using Mean Value Theorem,

$$\int_{\Omega} \sigma |\text{grad}(\phi - \bar{\phi})|^2 d\Omega + \int_{\Gamma_C} f'(\tilde{\phi})(\phi - \bar{\phi})^2 d\Gamma = 0$$

for some $\tilde{\phi}$ between ϕ and $\bar{\phi}$. Using (1.11), we see that

$$\int_{\Omega} \sigma |\text{grad}(\phi - \bar{\phi})|^2 d\Omega + \alpha \int_{\Gamma_C} (\phi - \bar{\phi})^2 d\Gamma \leq 0.$$

Hence we deduce that $\text{grad}(\phi - \bar{\phi}) = 0$ in Ω and $(\phi - \bar{\phi}) = 0$ on Γ_C . This in turn implies $(\phi - \bar{\phi}) = 0$ in Ω , i.e., uniqueness holds. ■

3. EXISTENCE OF AN OPTIMAL SOLUTION

Having shown that the constraint equation (1.7) is well-posed, we are now prepared to study the existence of an optimal solution $(\hat{\phi}, \hat{u})$ that minimizes the functional (1.2) subject to (1.7). We introduce the admissible set

$$\mathcal{U}_{ad} = \{(\phi, u) \in H^1(\Omega) \times U : (\phi, u) \text{ satisfies (1.7)}\},$$

where U is given by (1.9).

Theorem 3.1 *There exists a $(\hat{\phi}, \hat{u}) \in H^1(\Omega) \times U$ that minimizes (1.2) subject to (1.7).*

Proof. Theorem 2.4 implies an element $(\phi, u) \in \mathcal{U}_{ad}$ exists such that $\mathcal{J}(\phi, u) < \infty$.

Let $\{(\phi_n, u_n)\} \subset \mathcal{U}_{ad}$ be a minimizing sequence, i.e.,

$$\lim_{n \rightarrow \infty} \mathcal{J}(\phi_n, u_n) = \inf_{(\phi, u) \in \mathcal{U}_{ad}} \mathcal{J}(\phi, u) \quad (3.1)$$

and

$$\int_{\Omega} \sigma \text{grad} \phi_n \cdot \text{grad} \psi d\Omega + \int_{\Gamma_C} f(\phi_n) \psi d\Gamma = \int_{\Gamma_A} u_n \psi d\Gamma, \quad \forall \psi \in H^1(\Omega). \quad (3.2)$$

Using (1.2) and (3.1) we deduce $\{u_n\}$ is bounded in $L^2(\Gamma_A)$. Then (2.3) implies $\{\|\phi_n\|_1\}$ is bounded. Hence we may extract a subsequence $\{(\phi_n, u_n)\}$ such that

$$\phi_n \rightharpoonup \hat{\phi} \text{ in } H^1(\Omega) \quad \text{and} \quad u_n \rightharpoonup \hat{u} \text{ in } L^2(\Gamma_A).$$

Furthermore, trace theorems implies $\phi_n \rightarrow \phi$ in $L^2(\Gamma_C)$; this in turn implies $\phi_n \rightarrow \phi$ a.e. on Γ_C (after extracting subsequences if necessary). By setting $\psi = \phi_n$ in (3.2) we obtain

$$\int_{\Omega} \sigma |\text{grad} \phi_n|^2 d\Omega + \int_{\Gamma_C} f(\phi_n) \phi_n \leq \|u_n\|_{0, \Gamma_A} \|\phi_n\|_{0, \Gamma_A} \leq \beta \|u_n\|_{0, \Gamma_A} \|\phi_n\|_1.$$

Hence we deduce

$$\int_{\Gamma_C} f(\phi_n) \phi_n \leq M$$

where M is a constant independent of n . By Lemma 2.3,

$$\int_{\Gamma_C} f(\hat{\phi}) \hat{\phi} d\Gamma \leq \liminf_{n \rightarrow \infty} \int_{\Gamma_C} f(\phi_n) \phi_n d\Gamma$$

and

$$\lim_{n \rightarrow \infty} \int_{\Gamma_C} |f(\phi_n) - f(\hat{\phi})| d\Gamma = 0.$$

By passing to the limit in (3.2) for $\psi \in C^\infty(\bar{\Omega})$ we obtain

$$\int_{\Omega} \sigma \operatorname{grad} \hat{\phi} \cdot \operatorname{grad} \psi d\Omega + \int_{\Gamma_C} f(\hat{\phi}) \psi d\Gamma = \int_{\Gamma_A} \hat{u} \psi d\Gamma, \quad \forall \psi \in C^\infty(\bar{\Omega}).$$

Then using the denseness of $C^\infty(\bar{\Omega})$ in $H^1(\Omega)$ and the fact that $f(\hat{\phi}) \in L^2(\Gamma_C)$, we obtain

$$\int_{\Omega} \sigma \operatorname{grad} \hat{\phi} \cdot \operatorname{grad} \psi d\Omega + \int_{\Gamma_C} f(\hat{\phi}) \psi d\Gamma = \int_{\Gamma_A} \hat{u} \psi d\Gamma, \quad \forall \psi \in H^1(\Omega).$$

Thus $(\hat{\phi}, \hat{u}) \in \mathcal{U}_{ad}$. Finally using the weak lower semi-continuity of $\mathcal{J}(\cdot, \cdot)$, we conclude that $(\hat{\phi}, \hat{u})$ is indeed an optimal solution, i.e.,

$$\mathcal{J}(\hat{\phi}, \hat{u}) = \inf_{(\phi, u) \in \mathcal{U}_{ad}} \mathcal{J}(\phi, u). \quad \blacksquare$$

4. LAGRANGE MULTIPLIER RULES

In this section we will attempt to characterize optimal solutions whose existence has been established in §3. Since the constraint equation has a unique solution for each given control u , the state ϕ is a well-defined function of u . Since U given in (1.9) is not necessarily an open set, the mapping $u \mapsto \phi$ is in general not differentiable. Although other approaches available, it turns out to be convenient to use the Lagrange multiplier rule to turn the constrained minimization problem (1.8) into an unconstrained one and derive an optimality condition.

We first quote the following abstract theorem concerning the existence of Lagrange multipliers for minimization problems on Banach spaces (see, e.g., [19]):

Lemma 4.1. *Let B_1 and B_2 be two Banach spaces, and U an arbitrary set. Suppose $\mathcal{J}$ is a functional on $B_1 \times U$, $\mathcal{K}$ a mapping from $B_1 \times U$ to B_2 . Assume $(\hat{\phi}, \hat{u}) \in B_1 \times U$ is a local minimum for the constrained minimization problem:*

$$\min \mathcal{J}(\phi, u) \quad \text{subject to} \quad \mathcal{K}(\phi, u) = 0, \quad (4.1)$$

i.e., there exists an open neighborhood Θ of $\hat{\phi}$ in B_1 such that

$$\mathcal{J}(\hat{\phi}, \hat{u}) \leq \mathcal{J}(\phi, u) \quad \forall (\phi, u) \in \Theta \times U \text{ satisfying } \mathcal{K}(\phi, u) = 0.$$

Assume further the following conditions are satisfied:

- (A) for each $u \in U$, $\phi \mapsto \mathcal{J}(\phi, u)$ and $\phi \mapsto \mathcal{K}(\phi, u)$ are Frechet-differentiable in Θ ;
 (B) for every $\phi \in \Theta$, $u_1, u_2 \in U$ and $\alpha \in [0, 1]$, there exists a $u_\alpha = u_\alpha(\phi, u_1, u_2) \in U$ such that

$$\mathcal{K}(\phi, u_\alpha) = \alpha \mathcal{K}(\phi, u_1) + (1 - \alpha) \mathcal{K}(\phi, u_2)$$

and

$$\mathcal{J}(\phi, u_\alpha) \leq \alpha \mathcal{J}(\phi, u_1) + (1 - \alpha) \mathcal{J}(\phi, u_2);$$

- (C) the algebraic sum of $\mathcal{K}_\phi(\hat{\phi}, \hat{u})B_1 + \mathcal{K}(\hat{\phi}, U)$ contains a neighborhood of 0.
 Then there exists a $\lambda \in B_2^*$ such that

$$\langle \mathcal{J}_\phi(\hat{\phi}, \hat{u}), \psi \rangle - \langle \lambda, \mathcal{K}_\phi(\hat{\phi}, \hat{u})\psi \rangle = 0 \quad \forall \psi \in B_1$$

and

$$\min_{u \in U} \mathcal{L}(\hat{\phi}, u, \lambda) = \mathcal{L}(\hat{\phi}, \hat{u}, \lambda),$$

where $\mathcal{L}(\phi, u, \lambda) \equiv \mathcal{J}(\phi, u) - \langle \lambda, \mathcal{K}(\phi, u) \rangle$ is the Lagrangian of the constrained minimization problem (4.1).

Proof: See [19]. ■

We will fit our optimization problem (1.8) into the above abstract framework. We define the Banach spaces $B_1 \equiv H^1(\Omega)$ and $B_2 \equiv H^1(\Omega)^*$. Let U be given by (1.9). The (generalized) nonlinear constraint $\mathcal{K} : B_1 \times U \rightarrow B_2$ is defined as follows: $\mathcal{K}(\bar{\phi}, \bar{u}) = \bar{l}$ for $(\bar{\phi}, \bar{u}) \in B_1 \times U$ and $\bar{l} \in B_2$ if and only if

$$\begin{aligned} \langle \bar{l}, \psi \rangle_{\Gamma_A} &= \int_{\Omega} \sigma \operatorname{grad} \bar{\phi} \cdot \operatorname{grad} \psi \, d\Omega + \int_{\Gamma_C} f(\bar{\phi}) \psi \, d\Gamma \\ &\quad - \int_{\Gamma_A} \bar{u} \psi \, d\Gamma \quad \forall \psi \in H^1(\Omega). \end{aligned} \quad (4.2)$$

We easily see that (1.8) is equivalent to

$$\begin{aligned} &\text{find } (\phi, u) \in B_1 \times U \text{ such that} \\ &\mathcal{J}(\phi, u) = \inf \{ \mathcal{J}(\bar{\phi}, \bar{u}) : (\bar{\phi}, \bar{u}) \in B_1 \times U, \mathcal{K}(\bar{\phi}, \bar{u}) = 0 \}. \end{aligned}$$

Let the Frechet derivative of $\mathcal{J}$ and $\mathcal{K}$ with respect to ϕ be denoted by $D_\phi \mathcal{J}$ and $D_\phi \mathcal{K}$, respectively. Let the Frechet derivative of $\mathcal{J}$ and $\mathcal{K}$ with respect to u be denoted by $D_u \mathcal{J}$ and $D_u \mathcal{K}$, respectively. $D_\phi \mathcal{K}(\phi, u) \in \mathcal{L}(B_1, B_2)$ is defined as follows. $D_\phi \mathcal{K}(\hat{\phi}, \hat{u}) \cdot \bar{\phi} = \bar{l}$ for $\bar{\phi} \in B_1$ and $\bar{l} \in B_2$ if and only if

$$\langle \bar{l}, \psi \rangle_{\Gamma_A} = \int_{\Omega} \sigma \operatorname{grad} \bar{\phi} \cdot \operatorname{grad} \psi \, d\Omega + \int_{\Gamma_C} f'(\phi) \bar{\phi} \psi \, d\Gamma \quad \forall \psi \in H^1(\Omega). \quad (4.3)$$

In order to apply Lemma 4.1, we need to verify conditions (A)–(C). We begin with the verification of (C).

Lemma 4.2. Assume $(\hat{\phi}, \hat{u}) \in \mathcal{U}_{ad}$ is an optimal solution to (1.7). Then the operator $D_{\phi}\mathcal{K}(\hat{\phi}, \hat{u})$ is onto from B_1 to B_2 .

Proof: Using (1.11)-(1.12) we easily obtain the coercivity for the bilinear form

$$(\bar{\phi}, \psi) \mapsto \int_{\Omega} \sigma \operatorname{grad} \bar{\phi} \cdot \operatorname{grad} \psi \, d\Omega + \int_{\Gamma_C} f'(\phi) \bar{\phi} \psi \, d\Gamma \quad \forall \bar{\phi}, \psi \in H^1(\Omega).$$

Then Lax-Milgram Lemma implies that for any $\bar{l} \in H^1(\Omega)^*$ there exists a unique $\bar{\phi} \in H^1(\Omega)$ that solves (4.3). ■

Now we are prepared to derive an optimality condition.

Theorem 4.3. Assume $(\hat{\phi}, \hat{u}) \in \mathcal{U}_{ad}$ is an optimal solution to the minimization problem (1.8). Then there exists a $\lambda \in H^1(\Omega)$ such that

$$\int_{\Omega} \sigma \operatorname{grad} \psi \cdot \operatorname{grad} \lambda \, d\Omega + \int_{\Gamma_C} f'(\hat{\phi}) \psi \lambda \, d\Gamma = \frac{1}{\epsilon_0} \int_{\Omega} (\hat{\phi} - \phi_0) \psi \, d\Omega \quad \forall \psi \in H^1(\Omega). \quad (4.4)$$

and

$$\int_{\Gamma_A} (\delta_0 \hat{u} + \lambda)(u - \hat{u}) \, d\Gamma \geq 0 \quad \forall u \in U. \quad (4.5)$$

Proof: Condition (A) in Lemma 4.1 is obviously satisfied for the present setting. The convexity in u for the constraint is readily verified since the control variable u enters the constraint equation in a linear manner and the control set U is convex. The convexity in u for the functional is also easily seen from the convexity of the mapping $u \mapsto \|u\|_{0,\Gamma_C}^2$. Thus Condition (B) is verified. The validity of Condition (C) is established in Lemma 4.2. Hence, by Lemma 4.1, there exists a $\lambda \in B_2^* = H^1(\Omega)$ that satisfies

$$\langle D_{\phi}\mathcal{J}(\hat{\phi}, \hat{u}), \psi \rangle - \langle \lambda, D_{\phi}\mathcal{K}(\hat{\phi}, \hat{u}) \cdot \psi \rangle = 0 \quad \forall \psi \in H^1(\Omega) \quad (4.6)$$

and

$$\min_{u \in U} \mathcal{L}(\hat{\phi}, u, \lambda) = \mathcal{L}(\hat{\phi}, \hat{u}, \lambda), \quad (4.7)$$

where

$$\mathcal{L}(\phi, u, \lambda) \equiv \mathcal{J}(\phi, u) - \left[\int_{\Omega} \sigma \operatorname{grad} \phi \cdot \operatorname{grad} \lambda \, d\Omega + \int_{\Gamma_C} f(\phi) \lambda \, d\Gamma - \int_{\Gamma_A} u \lambda \, d\Gamma \right]. \quad (4.8)$$

Using (1.2) and (4.3), which are the defining equations for $\mathcal{J}(\phi, u)$ and $D_{\phi}\mathcal{K}(\hat{\phi}, \hat{u})$, respectively, (4.6) can be rewritten as

$$\frac{1}{\epsilon_0} \int_{\Omega} (\hat{\phi} - \phi_0) \psi \, d\Omega - \int_{\Omega} \sigma \operatorname{grad} \psi \cdot \operatorname{grad} \lambda \, d\Omega - \int_{\Gamma_C} f'(\hat{\phi}) \psi \lambda \, d\Gamma = 0 \quad \forall \psi \in H^1(\Omega)$$

which is clearly equivalent to (4.4). Equations (4.7)–(4.8) implies that, for all $u \in U$,

$$\begin{aligned} & \frac{1}{2\epsilon_0} \int_{\Omega} (\hat{\phi} - \phi_0)^2 d\Omega + \frac{\delta_0}{2} \int_{\Gamma_C} \hat{u}^2 d\Gamma \\ & - \left[\int_{\Omega} \sigma \operatorname{grad} \hat{\phi} \cdot \operatorname{grad} \lambda d\Omega + \int_{\Gamma_C} f(\hat{\phi}) \lambda d\Gamma - \int_{\Gamma_A} \hat{u} \lambda d\Gamma \right] \\ & \leq \frac{1}{\epsilon_0} \int_{\Omega} (\hat{\phi} - \phi_0)^2 d\Omega + \frac{\delta_0}{2} \int_{\Gamma_C} u^2 d\Gamma \\ & - \left[\int_{\Omega} \sigma \operatorname{grad} \hat{\phi} \cdot \operatorname{grad} \lambda d\Omega + \int_{\Gamma_C} f(\hat{\phi}) \lambda d\Gamma - \int_{\Gamma_A} u \lambda d\Gamma \right], \end{aligned}$$

i.e.,

$$\frac{\delta_0}{2} \int_{\Gamma_C} (u - \hat{u})(u + \hat{u}) d\Gamma + \int_{\Gamma_A} (u - \hat{u}) \lambda d\Gamma \geq 0 \quad \forall u \in U. \quad (4.9)$$

Given any $w \in U$ and $\epsilon \in (0, 1)$, we set $u = (1 - \epsilon)\hat{u} + \epsilon w$ in (4.9) (note $u \in U$ due to the convexity of U) and obtain

$$\frac{\delta_0}{2} \int_{\Gamma_C} \epsilon(w - \hat{u})(\epsilon w - \epsilon \hat{u} + 2\hat{u}) d\Gamma + \int_{\Gamma_A} \epsilon(w - \hat{u}) \lambda d\Gamma \geq 0 \quad \forall w \in U.$$

Thus (4.5) follows by dividing the last inequality by ϵ and then letting $\epsilon \rightarrow 0^+$. ■

5. AN OPTIMALITY SYSTEM AND THE REGULARITY OF ITS SOLUTIONS

In the sequel we will treat the special case $U = L^2(\Gamma_A)$. From (4.5) we easily obtain

$$\hat{u} = -\frac{1}{\delta_0} \lambda. \quad (5.1)$$

From (4.4), (5.1) and the original constraint equation (1.7), we form the following system of equations (dispensing with the hat notations to denote optimal solutions):

$$\int_{\Omega} \sigma \operatorname{grad} \phi \cdot \operatorname{grad} \psi d\Omega + \int_{\Gamma_C} f(\phi) \psi d\Gamma = -\frac{1}{\delta_0} \int_{\Gamma_A} \lambda \psi d\Gamma, \quad \forall \psi \in H^1(\Omega) \quad (5.2)$$

and

$$\begin{aligned} & \int_{\Omega} \sigma \operatorname{grad} \lambda \cdot \operatorname{grad} \omega d\Omega + \int_{\Gamma_C} f'(\phi) \lambda \omega d\Gamma \\ & = \frac{1}{\epsilon_0} \int_{\Omega} (\phi - \phi_0) \omega d\Omega, \quad \forall \omega \in H^1(\Omega). \end{aligned} \quad (5.3)$$

This system of equations will be called the *optimality system*.

Integrations by parts may be used to show that the system (5.2)–(5.3) constitutes a weak formulation of the problem

$$-\operatorname{div}(\sigma \operatorname{grad} \phi) = 0 \quad \text{in } \Omega, \quad (5.4)$$

$$\sigma \frac{\partial \phi}{\partial n} = -\frac{1}{\delta_0} \lambda \quad \text{on } \Gamma_A, \quad \sigma \frac{\partial \phi}{\partial n} = 0 \quad \text{on } \Gamma_0 \quad \text{and} \quad \sigma \frac{\partial \phi}{\partial n} = -f(\phi) \quad \text{on } \Gamma_C, \quad (5.5)$$

$$-\operatorname{div}(\sigma \operatorname{grad} \lambda) = \frac{1}{\epsilon_0}(\phi - \phi_0) \quad \text{in } \Omega, \quad (5.6)$$

$$\sigma \frac{\partial \lambda}{\partial n} = 0 \quad \text{on } \Gamma_A \cup \Gamma_0 \quad \text{and} \quad \sigma \frac{\partial \lambda}{\partial n} = -f'(\phi)\lambda \quad \text{on } \Gamma_C, \quad (5.7)$$

Now we examine the regularity of solutions of the optimality system (5.2)–(5.3), or equivalently, (5.4)–(5.7).

Theorem 5.1 *Suppose that $\Omega \subset \mathbb{R}^2$ is convex or of class $C^{1,1}$. Assume $(\phi, \lambda) \in H^1(\Omega) \times H^1(\Omega)$ is a solution to the optimality system (5.2)–(5.3), or equivalently, (5.4)–(5.7), then we have that $(\phi, \lambda) \in W^{3/2,r}(\Omega) \times W^{3/2,r}(\Omega)$ for $r \in [1, \infty)$.*

Proof: Since $\phi, \lambda \in H^1(\Omega)$, Lemma 2.3 implies $f(\phi) \in L^r(\Gamma_C)$ and $f'(\phi) \in L^r(\Gamma_C)$ for all $r \in [1, \infty)$. We infer from trace theorems that $\lambda \in L^q(\Gamma_C)$ for all $q > 1$. Hence we have $\sigma \frac{\partial \phi}{\partial n} \in L^q(\Gamma)$ and $\sigma \frac{\partial \lambda}{\partial n} \in L^q(\Gamma)$ for each $q > 1$. By applying elliptic regularity results to equations (5.4)–(5.7), we obtain $\phi \in W^{3/2,q}(\Omega)$ and $\lambda \in W^{3/2,q}(\Omega)$ for each $q > 1$.

Remark *In general the possible discontinuity of the normal derivative on the intersection of Γ_C , Γ_0 and Γ_A prohibits us from obtaining further regularity. However, if ϕ and λ vanish on the entire intersection of Γ_C , Γ_0 and Γ_A , then we could in fact show that $\phi \in C^2(\Omega) \cap C(\bar{\Omega})$ and $\lambda \in C^2(\Omega) \cap C(\bar{\Omega})$, i.e., ϕ and λ are in fact classical solutions of the optimality system. Also, $H^2(\Omega)$ -regularity for ϕ and λ is expected.*

6. FINITE ELEMENT APPROXIMATIONS

6.1 Finite Element Discretizations. A finite element discretization of the optimality system (5.2)–(5.3) is defined in the usual manner. For simplicity we assume the domain Ω is a convex polygon. We first choose families of finite dimensional subspaces $V^h \subset H^1(\Omega)$ satisfying the approximation property: there exists a constant C and an integer k such that

$$\|v - v^h\|_1 \leq Ch^m \|v\|_{m+1}, \quad \forall v \in H^{m+1}(\Omega), \quad 1 \leq m \leq k. \quad (6.1)$$

One may consult, e.g., [3] or [6] for a catalogue of finite element spaces satisfying (6.1). Then, we may formulate the approximate problem for the optimality system (5.2)–(5.3): seek $\phi^h \in V^h$ and $\lambda^h \in V^h$ such that

$$\begin{aligned} \int_{\Omega} \sigma \operatorname{grad} \phi^h \cdot \operatorname{grad} \psi^h d\Omega + \int_{\Gamma_C} f(\phi^h) \psi^h d\Gamma \\ = -\frac{1}{\delta_0} \int_{\Gamma_A} \lambda^h \psi^h d\Gamma, \quad \forall \psi^h \in V^h \end{aligned} \quad (6.2)$$

and

$$\begin{aligned} \int_{\Omega} \sigma \operatorname{grad} \lambda^h \cdot \operatorname{grad} \omega^h d\Omega + \int_{\Gamma_C} f'(\phi^h) \lambda^h \omega^h d\Gamma \\ = \frac{1}{\epsilon_0} \int_{\Omega} (\phi^h - \phi_0) \omega^h d\Omega, \quad \forall \omega^h \in V^h. \end{aligned} \quad (6.3)$$

6.2. Quotation of Brezzi-Rappaz-Raviart Approximation Theory.

The error estimate to be derived in Section 6.3 makes use of results developed by Brezzi, Rappaz and Raviart (see [5], also [7] and [8]) 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 considered in [5] (also [7] and [8]) are of the type

$$F(\psi) \equiv \psi + TG(\psi) = 0 \quad (6.4)$$

where X and Y are Banach spaces and $T \in \mathcal{L}(Y; X)$, G is a C^2 mapping from X into Y . A solution ψ to the equation $F(\psi) = 0$ is called a *nonsingular solution* if we have that $F'(\psi)$ is an isomorphism from X into X . (Here, $F'(\cdot)$ denotes the Frechet derivative of $F(\cdot)$.)

Approximations are defined by introducing a family of finite dimensional subspaces $X^h \subset X$ and for each $h > 0$ an approximating operator $T^h \in \mathcal{L}(Y; X^h)$. Then, we seek $\psi^h \in X^h$ such that

$$F^h(\psi^h) \equiv \psi^h + T^h G(\psi^h) = 0. \quad (6.5)$$

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

$$G'(\psi) \in \mathcal{L}(X; Z) \quad \forall \psi \in X. \quad (6.6)$$

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

$$\lim_{h \rightarrow 0} \|(T^h - T)y\|_X = 0 \quad \forall y \in Y \quad (6.7)$$

and

$$\lim_{h \rightarrow 0} \|(T^h - T)\|_{\mathcal{L}(Z; X)} = 0. \quad (6.8)$$

Note that (6.6) and (6.8) imply that the operator $G'(\psi) \in \mathcal{L}(X; X)$ is compact. Moreover, (6.8) follows from (6.7) whenever the imbedding $Z \subset Y$ is compact.

We can now state the first result that will be used in the sequel. In the statement of the theorem, G'' represents the second order Frechet derivative of G .

Theorem 6.1 *Let X and Y be Banach spaces. Assume that G is a second order Frechet differentiable mapping from X into Y and that G'' is bounded on all bounded sets of X . Assume that (6.6)-(6.8) hold and that ψ is a nonsingular solution of (6.4). Then, there exists a $\delta > 0$ and an $h_0 > 0$ such that for $h \leq h_0$, there exists a unique $\psi^h \in X^h$ satisfying ψ^h is a nonsingular solution of (6.5) and $\|\psi^h - \psi\|_X \leq \delta$. Moreover, there exists a constant $C > 0$, independent of h , such that*

$$\|\psi^h - \psi\|_X \leq C \|(T^h - T)G(\psi)\|_X \quad \blacksquare \quad (6.9)$$

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

for all $w \in W$, the operator $G'(w)$ may be extended as a linear operator of $\mathcal{L}(H; Y)$, the mapping $w \rightarrow G'(w)$ being continuous from W onto $\mathcal{L}(H; Y)$. (6.10)

We also suppose that

$$\lim_{h \rightarrow 0} \|T^h - T\|_{\mathcal{L}(Y;H)} = 0. \quad (6.11)$$

Then we may state the following additional result.

Theorem 6.2 *Assume that the hypotheses of Theorem 6.1 hold and that (6.10) and (6.11) hold. Assume further that*

$$F'(\psi) \text{ is an isomorphism of } H. \quad (6.12)$$

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

$$\|\psi^h - \psi\|_H \leq C\|(T^h - T)G(\psi)\|_H + \|\psi^h - \psi\|_X^2. \quad \blacksquare \quad (6.13)$$

6.3 Error Estimates for the Approximations of Solutions of the Optimality System.

In order to derive error estimates, we begin by recasting the optimality system (5.2)–(5.3) and its discretization (6.2)–(6.3) into a form that fits into the framework of *Brezzi-Rappaz-Raviart* theory summarized in §6.2.

We define

$$\begin{aligned} X &= H^1(\Omega) \times H^1(\Omega), \\ Y &= H^{-1/2}(\Gamma) \times H^1(\Omega)^* \times H^{-1/2}(\Gamma), \\ Z &= L^2(\Gamma) \times L^2(\Omega) \times L^2(\Gamma) \end{aligned}$$

and

$$X^h = V^h \times V^h,$$

where $H^1(\Omega)^*$ denotes the dual space of $H^1(\Omega)$. Note that using Sobolev imbedding theorems, $Z \subset Y$ with a compact imbedding.

Let the operator $T \in \mathcal{L}(Y; X)$ be defined in the following manner: $T(\zeta, \eta, \theta) = (\phi, \lambda)$ for $(\zeta, \eta, \theta) \in Y$ and $(\phi, \lambda) \in X$ if and only if

$$\int_{\Omega} \sigma \operatorname{grad} \phi \cdot \operatorname{grad} \psi \, d\Omega + \alpha \int_{\Gamma_C} \phi \psi \, d\Gamma = \langle \zeta, \psi \rangle_{\Gamma}, \quad \forall \psi \in H^1(\Omega) \quad (6.14)$$

and

$$\int_{\Omega} \sigma \operatorname{grad} \lambda \cdot \operatorname{grad} \omega \, d\Omega + \alpha \int_{\Gamma_C} \lambda \omega \, d\Gamma = \langle \eta, \omega \rangle + \langle \theta, \omega \rangle_{\Gamma}, \quad \forall \omega \in H^1(\Omega). \quad (6.15)$$

Clearly, (6.14)–(6.15) consists of two *uncoupled* elliptic equations with mixed Robin-Neumann type boundary conditions and T is its solution operator.

Analogously, the operator $T^h \in \mathcal{L}(Y; X^h)$ is defined as follows: $T^h(\zeta, \eta, \theta) = (\phi^h, \lambda^h)$ for $(\zeta, \eta, \theta) \in Y$ and $(\phi^h, \lambda^h) \in X^h$ if and only if

$$\int_{\Omega} \sigma \operatorname{grad} \phi^h \cdot \operatorname{grad} \psi^h \, d\Omega + \alpha \int_{\Gamma_C} \phi^h \psi^h \, d\Gamma = \langle \zeta, \psi^h \rangle_{\Gamma}, \quad \forall \psi^h \in V^h \quad (6.16)$$

and

$$\int_{\Omega} \sigma \operatorname{grad} \lambda^h \cdot \operatorname{grad} \omega^h d\Omega + \alpha \int_{\Gamma_C} \lambda^h \omega^h d\Gamma = \langle \eta, \omega^h \rangle + \langle \theta, \omega^h \rangle_{\Gamma}, \quad \forall \omega^h \in V^h. \quad (6.17)$$

Clearly, (6.16)–(6.17) consists of two discrete Poisson-type equations that are discretizations of the equations (6.14)–(6.15); also, T^h is the solution operator for these two discrete equations.

By the well-known results concerning the approximation of elliptic equations (see, e.g., [3] or [6]), we obtain:

$$\|(T - T^h)(\zeta, \eta, \theta)\|_X \rightarrow 0 \quad \text{as } h \rightarrow 0, \quad (6.18)$$

for all $(\zeta, \eta, \theta) \in Y$ and, in addition, if $T(\zeta, \eta, \theta) \in H^{m+1}(\Omega) \times H^{m+1}(\Omega)$, then

$$\|(T - T^h)(\zeta, \eta, \theta)\|_X \leq Ch^m \|T(\zeta, \eta, \theta)\|_{H^{m+1}(\Omega) \times H^{m+1}(\Omega)}. \quad (6.19)$$

Also, because $Z \subset Y$ with a compact imbedding, we have that

$$\|(T - T^h)\|_{\mathcal{L}(Z; X)} \rightarrow 0 \quad \text{as } h \rightarrow 0. \quad (6.20)$$

Next, we define the *nonlinear* mapping $G : X \rightarrow Y$ as follows: $G(\phi, \lambda) = (\zeta, \eta, \theta)$ for $(\phi, \lambda) \in X$ and $(\zeta, \eta, \theta) \in Y$ if and only if

$$\langle \zeta, \pi \rangle_{\Gamma} = \frac{1}{\delta_0} \int_{\Gamma_A} \lambda \pi d\Gamma + \int_{\Gamma_C} (f(\phi) - \alpha \phi) \pi d\Gamma \quad \forall \pi \in H^{1/2}(\Gamma), \quad (6.21)$$

$$\langle \eta, \omega \rangle = -\frac{1}{\epsilon_0} \int_{\Omega} (\phi - \phi_0) \omega d\Omega \quad \forall \omega \in H^1(\Omega) \quad (6.22)$$

and

$$\langle \theta, \tau \rangle_{\Gamma} = \int_{\Gamma_C} (f'(\phi) - \alpha) \lambda \tau d\Gamma \quad \forall \tau \in H^{1/2}(\Gamma). \quad (6.23)$$

(6.21)–(6.23) is equivalent to

$$\zeta = \begin{cases} \frac{1}{\delta_0} \lambda & \text{on } \Gamma_A; \\ f(\phi) - \alpha \phi & \text{on } \Gamma_C; \\ 0 & \text{on } \Gamma_0, \end{cases} \quad (6.24)$$

$$\eta = -\frac{1}{\epsilon_0} (\phi - \phi_0) \quad \text{in } \Omega \quad (6.25)$$

and

$$\theta = \begin{cases} (f'(\phi) - \alpha) \lambda & \text{on } \Gamma_C; \\ 0 & \text{on } \Gamma_0 \cup \Gamma_A. \end{cases} \quad (6.26)$$

Recall $f(\phi) = C_3(e^{C_1\phi} - e^{C_2\phi})$ so that $f'(\phi) = C_3(C_1e^{C_1\phi} + C_2e^{C_2\phi})$. Using Lemma 2.2 and trace theorems we infer that if $(\phi, \lambda) \in H^1(\Omega) \times H^1(\Omega)$, then for all $q > 1$,

$\phi|_{\Gamma} \in L^q(\Gamma)$, $\lambda|_{\Gamma} \in L^q(\Gamma)$, $f(\phi) \in L^q(\Gamma)$ and $f'(\phi) \in L^q(\Gamma)$. Hence we see that the triplet (ζ, η, θ) defined by (6.24)–(6.26) is indeed in Y , i.e., G is well-defined.

It is easily seen that the optimality system (5.2)–(5.3) is equivalent to

$$(\phi, \lambda) + TG(\phi, \lambda) = 0 \quad (6.27)$$

and that the discrete optimality system (6.2)–(6.3) is equivalent to

$$(\phi^h, \lambda^h) + T^h G(\phi^h, \lambda^h) = 0. \quad (6.28)$$

We have thus recast our continuous and discrete optimality problems into a form that enables us to apply the theories of §6.2. It remains to verify the hypotheses in Theorem 6.1. This will be the task of the next two propositions.

Proposition 6.3 *The operator $G : X \rightarrow Y$ defined by (6.21)–(6.23) is second order Frechet differentiable. Furthermore, (6.6) holds and G'' is bounded on all bounded sets of X .*

Proof. In showing the differentiability of G , the linear terms appearing in the definition of G does not pose any difficulty. Furthermore the nonlinear terms in (6.21) and (6.23) can be dealt with in a similar way. For clarity, we will only analyse the differentiability of the nonlinear term $\tau \mapsto \int_{\Gamma_C} f'(\phi) \lambda \tau d\Gamma$. We define a mapping $Q : X \rightarrow H^{-1/2}(\Gamma)$ by $\langle Q(\phi, \lambda), \tau \rangle \equiv \int_{\Gamma_C} f'(\phi) \lambda \tau d\Gamma$ for all $(\phi, \lambda) \in X$ and $\tau \in H^{1/2}(\Gamma)$. For each given $(\phi, \lambda) \in X$ we have that

$$\begin{aligned} & \langle Q(\phi + \delta\phi, \lambda + \delta\lambda) - Q(\phi, \lambda), \tau \rangle \\ &= \int_{\Gamma_C} f''(\phi)(\delta\phi) \lambda \tau d\Gamma - \int_{\Gamma_C} f'(\phi)(\delta\lambda) \tau d\Gamma \\ &= \int_{\Gamma_C} [f'(\phi + \delta\phi) - f'(\phi) - f''(\phi)(\delta\phi)] \lambda \tau d\Gamma \\ & \quad + \int_{\Gamma_C} [f'(\phi + \delta\phi) - f'(\phi)](\delta\lambda) \tau d\Gamma \\ &= \int_{\Gamma_C} \int_0^1 [f''((1-t)\phi + t(\phi + \delta\phi)) - f''(\phi)] dt (\delta\phi) \lambda \tau d\Gamma + \\ & \quad + \int_{\Gamma_C} \int_0^1 f''((1-t)\phi + t(\phi + \delta\phi)) dt (\delta\phi)(\delta\lambda) \tau d\Gamma \\ &= \int_{\Gamma_C} \int_0^1 \int_0^1 t f'''(s(1-t)\phi + st(\phi + \delta\phi) + (1-s)\phi) ds dt |\delta\phi|^2 \lambda \tau d\Gamma \\ & \quad + \int_{\Gamma_C} \int_0^1 f''((1-t)\phi + t(\phi + \delta\phi)) dt (\delta\phi)(\delta\lambda) \tau d\Gamma \quad \forall (\delta\phi, \delta\lambda) \in X. \end{aligned} \quad (6.29)$$

Note that $f(\phi) = C_3(e^{C_1\phi} - e^{-C_2\phi})$, $f'(\phi) = C_3(C_1e^{C_1\phi} + C_2e^{-C_2\phi})$, $f''(\phi) = C_3(C_1^2e^{C_1\phi} - C_2^2e^{-C_2\phi})$ and $f'''(\phi) = C_3(C_1^3e^{C_1\phi} + C_2^3e^{-C_2\phi})$. By Lemma 2.2 we have that for all real number $m > 1$,

$$\|f'(\phi)\|_{L^m(\Gamma_C)} \leq C \left\{ 1 + |\Gamma| + e^{m^2 C \|\phi\|_1^2} |\Gamma| \right\}^{\frac{1}{m}},$$

$$\|f''(\phi)\|_{L^m(\Gamma_C)} \leq C \left\{ 1 + |\Gamma| + e^{m^2 C \|\phi\|_1^2} |\Gamma| \right\}^{\frac{1}{m}}$$

and

$$\|f'''(\phi)\|_{L^m(\Gamma_C)} \leq C \left\{ 1 + |\Gamma| + e^{m^2 C \|\phi\|_1^2} |\Gamma| \right\}^{\frac{1}{m}},$$

where C is a generic constant independent of ϕ . Trace theorems for $\Omega \subset \mathbb{R}^2$ implies that for all $p > 1$, $q > 1$ and $r > 1$,

$$\|\delta\phi\|_{L^p(\Gamma_C)} \leq C \|\delta\phi\|_1 \quad \forall \delta\phi \in H^1(\Omega),$$

$$\|\delta\lambda\|_{L^q(\Gamma_C)} \leq C \|\delta\lambda\|_1 \quad \forall \delta\lambda \in H^1(\Omega)$$

and

$$\|\tau\|_{L^r(\Gamma_C)} \leq C \|\tau\|_{1/2, \Gamma} \quad \forall \tau \in H^{1/2}(\Gamma).$$

We fix some $m > 1$, $p > 1$, $q > 1$ and $r > 1$ with $\frac{1}{m} + \frac{1}{p} + \frac{1}{q} + \frac{1}{r} = 1$. Then for every $(\delta\phi, \delta\lambda) \in X$ we have that

$$\begin{aligned} & \int_{\Gamma_C} \int_0^1 f''((1-t)\phi + t(\phi + \delta\phi)) dt (\delta\phi)(\delta\lambda)\tau d\Gamma \leq \\ & \leq \sup_{0 \leq t \leq 1} \|f''((1-t)\phi + t(\phi + \delta\phi))\|_{L^m(\Gamma_C)} \|\delta\phi\|_{L^p(\Gamma_C)} \|\delta\lambda\|_{L^q(\Gamma_C)} \|\tau\|_{L^r(\Gamma_C)} \\ & \leq C \left\{ 1 + |\Gamma| + e^{m^2 2C \|\phi\|_1^2} |\Gamma| \right\}^{\frac{1}{m}} \|\delta\phi\|_1 \|\delta\lambda\|_1 \|\tau\|_{1/2, \Gamma} \quad \forall \tau \in H^{1/2}(\Gamma). \end{aligned}$$

Similarly, we have that for every $\delta\phi \in H^1(\Omega)$ and every $\tau \in H^{1/2}(\Gamma)$,

$$\begin{aligned} & \int_{\Gamma_C} \int_0^1 \int_0^1 t f'''(s(1-t)\phi + st(\phi + \delta\phi) + (1-s)\phi) ds dt |\delta\phi|^2 \lambda \tau d\Gamma \leq \\ & \leq C \left\{ 1 + |\Gamma| + e^{m^2 3C \|\phi\|_1^2} |\Gamma| \right\}^{\frac{1}{m}} \|\delta\phi\|_1^2 \|\lambda\|_1 \|\tau\|_{1/2, \Gamma}. \end{aligned}$$

Returning to (6.29) we obtain that for all $\tau \in H^{1/2}(\Gamma)$ and $(\delta\phi, \delta\lambda) \in X$,

$$\begin{aligned} & \left\langle Q(\phi + \delta\phi, \lambda + \delta\lambda) - Q(\phi, \lambda), \tau \right\rangle - \int_{\Gamma_C} f''(\phi)(\delta\phi) \lambda \tau d\Gamma - \int_{\Gamma_C} f'(\phi)(\delta\lambda) \tau d\Gamma \leq \\ & \leq C \left\{ 1 + |\Gamma| + e^{m^2 3C \|\phi\|_1^2} |\Gamma| \right\}^{\frac{1}{m}} \{ \|\delta\phi\|_1 \|\delta\lambda\|_1 + \|\delta\phi\|_1^2 \|\lambda\|_1 \} \|\tau\|_{1/2, \Gamma}, \end{aligned}$$

so that we conclude Q as a mapping from X to $H^{-1/2}(\Gamma)$ is Frechet differentiable and its derivative $Q'(\phi, \lambda)$ is given by

$$\left\langle Q'(\phi, \lambda)(\tilde{\psi}, \tilde{\omega}), \tau \right\rangle = \int_{\Gamma_C} [f''(\phi)\tilde{\psi}\lambda\tau + f'(\phi)\tilde{\omega}\tau] d\Gamma \quad \forall \tau \in H^{1/2}(\Gamma).$$

Hence, taking into account the remarks in the beginning of the proof, we have justified that G is Frechet differentiable and its Frechet derivative $G'(\phi, \lambda)$ is defined

as follows. For each $(\phi, \lambda) \in X$, $G'(\phi, \lambda)(\tilde{\psi}, \tilde{\omega}) = (\tilde{\zeta}, \tilde{\eta}, \tilde{\theta})$ for $(\tilde{\psi}, \tilde{\omega}) \in X$ and $(\tilde{\zeta}, \tilde{\eta}, \tilde{\theta}) \in Y$ if and only if

$$\langle \tilde{\zeta}, \pi \rangle_{\Gamma} = \frac{1}{\delta_0} \int_{\Gamma_A} \tilde{\omega} \pi d\Gamma + \int_{\Gamma_C} (f'(\phi) \tilde{\psi} - \alpha \tilde{\psi}) \pi d\Gamma \quad \forall \pi \in H^{1/2}(\Gamma), \quad (6.30)$$

$$\langle \tilde{\eta}, \omega \rangle = -\frac{1}{\epsilon_0} \int_{\Omega} \tilde{\psi} \omega d\Omega \quad \forall \omega \in H^1(\Omega) \quad (6.31)$$

and

$$\langle \tilde{\theta}, \tau \rangle_{\Gamma} = \int_{\Gamma_C} f''(\phi) \tilde{\psi} \lambda \tau d\Gamma + \int_{\Gamma_C} (f'(\phi) - \alpha) \tilde{\omega} \tau d\Gamma \quad \forall \tau \in H^{1/2}(\Gamma); \quad (6.32)$$

or, equivalently,

$$\tilde{\zeta} = \begin{cases} \frac{1}{\delta_0} \tilde{\omega} & \text{on } \Gamma_A; \\ f'(\phi) \tilde{\psi} - \alpha \tilde{\psi} & \text{on } \Gamma_C; \\ 0 & \text{on } \Gamma_0, \end{cases} \quad (6.33)$$

$$\tilde{\eta} = -\frac{1}{\epsilon_0} \tilde{\psi} \quad \text{in } \Omega \quad (6.34)$$

and

$$\tilde{\theta} = \begin{cases} f''(\phi) \lambda \tilde{\psi} + (f'(\phi) - \alpha) \tilde{\omega} & \text{on } \Gamma_C; \\ 0 & \text{on } \Gamma_0 \cup \Gamma_A. \end{cases} \quad (6.35)$$

(These defining equations can be formally derived by differentiating (6.21)–(6.23).) It is easy to verify from the above equations that for each $(\tilde{\psi}, \tilde{\omega}) \in X$, we have $(\tilde{\zeta}, \tilde{\eta}, \tilde{\theta}) \in Z$, i.e., $G'(\phi, \lambda)$ maps X into Z ; furthermore, using trace theorems and Lemma 2.2 we obtain that

$$\begin{aligned} \|\tilde{\zeta}\|_{0,\Gamma} &\leq \frac{1}{\delta_0} \|\tilde{\omega}\|_{0,\Gamma} + \|f'(\phi)\|_{L^4(\Gamma_C)} \|\tilde{\psi}\|_{L^4(\Gamma)} + \alpha \|\tilde{\psi}\|_{0,\Gamma} \\ &\leq \frac{C}{\delta_0} \|\tilde{\omega}\|_1 + C \left\{ 1 + |\Gamma| + e^{C\|\phi\|_1^2} |\Gamma| \right\}^{1/4} \|\tilde{\psi}\|_1 + C \|\tilde{\psi}\|_1, \end{aligned}$$

$$\|\tilde{\eta}\|_0 \leq \frac{1}{\epsilon} \|\tilde{\psi}\|_1$$

and

$$\begin{aligned} \|\tilde{\theta}\|_{0,\Gamma} &\leq \|f''(\phi)\|_{L^6(\Gamma_C)} \|\tilde{\psi}\|_{L^6(\Gamma)} \|\lambda\|_{L^6(\Gamma)} + \|f'(\phi)\|_{L^4(\Gamma_C)} \|\tilde{\omega}\|_{L^4(\Gamma)} + \alpha \|\tilde{\omega}\|_{0,\Gamma} \\ &\leq C \left\{ 1 + |\Gamma| + e^{C\|\phi\|_1^2} |\Gamma| \right\}^{1/4} \{ \|\tilde{\psi}\|_1 \|\lambda\|_1 + \|\tilde{\omega}\|_1 \} + C \|\tilde{\omega}\|_1. \end{aligned}$$

Thus $G'(\phi, \lambda) \in \mathcal{L}(X; Z)$, i.e., we have shown that (6.6) hold.

To show the second order differentiability of G , again for clarity we will examine only one nonlinear term appearing in the definition of G' , e.g., the term $\tau \mapsto \int_{\Gamma_C} f''(\phi) \tilde{\psi} \lambda \tau d\Gamma$. We define a mapping $R : X \rightarrow \mathcal{L}(X; H^{-1/2}(\Gamma))$ by

$\langle R(\phi, \lambda)(\tilde{\psi}, \tilde{\omega}), \tau \rangle \equiv \int_{\Gamma_C} f''(\phi) \tilde{\psi} \lambda \tau d\Gamma$ for all $(\phi, \lambda), (\tilde{\psi}, \tilde{\omega}) \in X$ and $\tau \in H^{1/2}(\Gamma)$.
For each given $(\phi, \lambda) \in X$ we have that for all $(\delta\phi, \delta\lambda) \in X$,

$$\begin{aligned} & \langle [R(\phi + \delta\phi, \lambda + \delta\lambda) - R(\phi, \lambda)](\tilde{\psi}, \tilde{\omega}), \tau \rangle \\ &= \int_{\Gamma_C} f'''(\phi)(\delta\phi) \tilde{\psi} \lambda \tau d\Gamma - \int_{\Gamma_C} f''(\phi) \tilde{\psi}(\delta\lambda) \tau d\Gamma \\ &= \int_{\Gamma_C} [f''(\phi + \delta\phi) - f''(\phi) - f'''(\phi)(\delta\phi)] \tilde{\psi} \lambda \tau d\Gamma \\ & \quad + \int_{\Gamma_C} [f''(\phi + \delta\phi) - f''(\phi)] \tilde{\psi}(\delta\lambda) \tau d\Gamma \\ &= \int_{\Gamma_C} \int_0^1 \int_0^1 t f''''(s(1-t)\phi + st(\phi + \delta\phi) + (1-s)\phi) ds dt \tilde{\psi} |\delta\phi|^2 \lambda \tau d\Gamma \\ & \quad + \int_{\Gamma_C} \int_0^1 f'''((1-t)\phi + t(\phi + \delta\phi)) dt (\delta\phi)(\delta\lambda) \tilde{\psi} \tau d\Gamma. \end{aligned}$$

Thus similar to the analysis ensuing (6.29), we can show that the operator R is Frechet differentiable and its derivative $R'(\phi, \lambda)$ is defined by:

$$R'(\phi, \lambda) \cdot ((\tilde{\psi}, \tilde{\omega}), (\tilde{\psi}, \tilde{\omega})) = \int_{\Gamma_C} f'''(\phi) \tilde{\phi} \tilde{\psi} \lambda \tau d\Gamma + \int_{\Gamma_C} f''(\phi) \tilde{\psi} \tilde{\lambda} \tau d\Gamma.$$

Hence, G is second order Frechet differentiable and $G''(\phi, \lambda)$ is defined as follows. For each $(\phi, \lambda) \in X$, $G''(\phi, \lambda) \cdot ((\tilde{\psi}, \tilde{\omega}), (\tilde{\psi}, \tilde{\omega})) = (\tilde{\zeta}, \tilde{\eta}, \tilde{\theta})$ for $((\tilde{\psi}, \tilde{\omega}), (\tilde{\psi}, \tilde{\omega})) \in X \times X$ and $(\tilde{\zeta}, \tilde{\eta}, \tilde{\theta}) \in Y$ if and only if

$$\langle \tilde{\zeta}, \pi \rangle_{\Gamma} = \int_{\Gamma_C} f''(\phi) \tilde{\psi} \tilde{\psi} \pi d\Gamma \quad \forall \pi \in H^{1/2}(\Gamma),$$

$$\langle \tilde{\eta}, \omega \rangle = 0 \quad \forall \omega \in H_0^1(\Omega)$$

and

$$\langle \tilde{\theta}, \tau \rangle_{\Gamma} = \int_{\Gamma_C} [f'''(\phi) \tilde{\psi} \tilde{\psi} \lambda + f''(\phi) \tilde{\psi} \tilde{\lambda}] \tau d\Gamma + \int_{\Gamma_C} f''(\phi) \tilde{\psi} \tilde{\omega} \tau d\Gamma \quad \forall \tau \in H^{1/2}(\Gamma);$$

or, equivalently,

$$\tilde{\zeta} = \begin{cases} 0 & \text{on } \Gamma_A; \\ f'''(\phi) \tilde{\psi} \tilde{\psi} & \text{on } \Gamma_C; \\ 0 & \text{on } \Gamma_0, \end{cases} \quad (6.36)$$

$$\tilde{\eta} = 0 \quad \text{in } \Omega \quad (6.37)$$

and

$$\tilde{\theta} = \begin{cases} [f'''(\phi) \tilde{\psi} \tilde{\psi} \lambda + f''(\phi) \tilde{\psi} \tilde{\lambda}] + f''(\phi) \tilde{\psi} \tilde{\omega} & \text{on } \Gamma_C; \\ 0 & \text{on } \Gamma_0 \cup \Gamma_A. \end{cases} \quad (6.38)$$

Furthermore, using Lemma 2.2, (6.36)-(6.38) and trace theorems, we may derive a bound for $G''(\phi, \lambda)$ for each given (ϕ, λ) :

$$\|G''(\phi, \lambda)\|_Y \leq C \left\{ 1 + |\Gamma| + e^{C\|\phi\|_1^2} |\Gamma| \right\} (1 + \|\lambda\|_1)$$

for some constant C , so that G'' is bounded on every bounded subset of X . ■

A solution (ϕ, λ) of the problem (5.2)-(5.3), or equivalently, of (6.27), is nonsingular if the *linear system*

$$\begin{aligned} \int_{\Omega} \sigma \operatorname{grad} \tilde{\phi} \cdot \operatorname{grad} \psi \, d\Omega + \int_{\Gamma_C} f'(\phi) \tilde{\phi} \psi \, d\Gamma \\ + \frac{1}{\delta_0} \int_{\Gamma_A} \tilde{\lambda} \psi \, d\Gamma = \langle \tilde{\zeta}, \psi \rangle \quad \forall \psi \in H^1(\Omega) \end{aligned} \quad (6.39)$$

and

$$\begin{aligned} \int_{\Omega} \sigma \operatorname{grad} \tilde{\lambda} \cdot \operatorname{grad} \omega \, d\Omega + \int_{\Gamma_C} f''(\phi) \tilde{\phi} \lambda \omega \, d\Gamma + \int_{\Gamma_C} f'(\phi) \tilde{\lambda} \omega \, d\Gamma \\ - \frac{1}{\epsilon_0} \int_{\Omega} \tilde{\phi} \omega \, d\Omega = \langle \tilde{\eta}, \omega \rangle \quad \forall \omega \in H^1(\Omega) \end{aligned} \quad (6.40)$$

has a unique solution $(\tilde{\phi}, \tilde{\lambda}) \in X$ for every $\tilde{\zeta}, \tilde{\eta} \in H^1(\Omega)^*$.

An analogous definition holds for nonsingular solutions of the discrete optimality system (6.2)-(6.3), or equivalently, (6.28).

It is evident that (6.39)-(6.40) has a unique solution for large enough σ , e.g.,

$$\sigma > \max \left\{ \frac{C}{\delta_0}, \frac{C}{\epsilon_0}, C \|\lambda\|_{L^4(\Gamma_C)} \|f''(\phi)\|_{L^4(\Gamma_C)} \right\}.$$

It is reasonable to assume that (6.39)-(6.40) has a unique solution generically with respect to σ , i.e., the optimal solutions are almost always nonsingular. Thus Theorem 6.1 and Proposition 6.3 lead to the following:

Theorem 6.4 *Assume (ϕ, λ) is a nonsingular solution of the optimality system (5.2)-(5.3). Assume that the finite element spaces V^h satisfy the condition (6.1). Then, there exists a $\delta > 0$ and $h_0 > 0$ such that for $h \leq h_0$, there exists a unique nonsingular solution (ϕ^h, λ^h) of the discrete optimality system (6.2)-(6.3) satisfying $\|\phi^h - \phi\|_1 + \|\lambda^h - \lambda\|_1 \leq \delta$. Moreover,*

$$\|\phi^h - \phi\|_1 + \|\lambda^h - \lambda\|_1 \rightarrow 0 \quad \text{as } h \rightarrow 0. \quad (6.41)$$

If, in addition, the solution of the optimality system satisfies $(\phi, \lambda) \in H^{m+1}(\Omega) \times H^{m+1}(\Omega)$, then there exists a constant C , independent of h , such that

$$\|\phi - \phi^h\|_1 + \|\lambda - \lambda^h\|_1 \leq Ch^m (\|\phi\|_{m+1} + \|\lambda\|_{m+1}). \quad \blacksquare \quad (6.42)$$

A consequence of Theorems 6.4 is the following corollary that gives error estimates for the approximation of the controls.

Corollary 6.5 *Assume (ϕ, λ) is a nonsingular solution of the optimality system (5.2)-(5.3). Assume that the finite element spaces V^h satisfy the condition (6.1). Define the approximate control by*

$$u^h = -\frac{1}{\delta_0} \lambda^h \quad \text{on } \Gamma_A.$$

Then

$$\|u^h - u\|_{1/2, \Gamma_A} \rightarrow 0 \quad \text{as } h \rightarrow 0. \quad (6.43)$$

If, in addition, the solution of the optimality system satisfies $(\phi, \lambda) \in H^{m+1}(\Omega) \times H^{m+1}(\Omega)$, then there exists a constant C , independent of h , such that for $h \leq h_0$,

$$\|u^h - u\|_{1/2, \Gamma_A} \leq \frac{C}{\delta_0} h^m (\|\phi\|_{m+1} + \|\lambda\|_{m+1}). \quad (6.44)$$

Proof: Recall that $u = -\frac{1}{\delta_0} \lambda$ on Γ_A ; see (5.1). Then (6.43) and (6.44) follow trivially from (6.41)-(6.42) and the inequalities (see [1])

$$\|u - u^h\|_{1/2, \Gamma_A} = \frac{1}{\delta_0} \|\lambda - \lambda^h\|_{1/2, \Gamma_C} \leq \frac{1}{\delta_0} \|\lambda - \lambda^h\|_{1/2, \Gamma} \leq \frac{C}{\delta_0} \|\lambda - \lambda^h\|_1. \quad \blacksquare$$

Now we wish to apply Theorem 6.2 to derive $L^2(\Gamma_C)$ -error estimates for the approximations of u . To this end, we assume the domain Ω is convex and for each given $\epsilon \in (0, 1/4)$, we introduce spaces

$$H = H^{1/2+\epsilon}(\Omega) \times H^{1/2+\epsilon}(\Omega) \quad \text{and} \quad W = H^{3/2+\epsilon}(\Omega) \times H^{3/2+\epsilon}(\Omega).$$

Note that $X \subset H$ with a compact imbedding so that (6.18) implies

$$\|(T - T^h)\|_{\mathcal{L}(Y; H)} \rightarrow 0 \quad \text{as } h \rightarrow 0.$$

Again using finite element approximation results in [6] we have that if Ω is convex and $T(\zeta, \eta, \theta) \in H^{m+1}(\Omega) \times H^{m+1}(\Omega)$, then

$$\|(T - T^h)(\zeta, \eta, \theta)\|_H \leq C h^{m-\epsilon+1/2} \|T(\zeta, \eta, \theta)\|_{H^{m+1}(\Omega) \times H^{m+1}(\Omega)}.$$

Proposition 6.6 *For each $(\phi, \lambda) \in W$, the operator $G'(\phi, \lambda) : X \rightarrow Y$ defined by (6.33)-(6.35) can be extended as a linear operator of $\mathcal{L}(H; Y)$. Furthermore, the mapping $w \rightarrow G'(w)$ is continuous from W onto $\mathcal{L}(H; Y)$.*

Proof. Note that $W \subset L^\infty(\Omega) \times L^\infty(\Omega)$ and $H^{1/2+\epsilon}(\Omega)|_\Gamma \subset L^2(\Gamma)$ with continuous imbeddings. For each $(\phi, \lambda) \in W$, we can easily verify from (6.33)-(6.35) that

$$\begin{aligned} & \|G'(\phi, \lambda)(\tilde{\psi}, \tilde{\omega})\|_Y \\ & \leq A(\phi) \{ \|\tilde{\omega}\|_{0, \Gamma_A} + \|\tilde{\psi}\|_{0, \Gamma_C} + \|\tilde{\psi}\|_0 + \|\tilde{\omega}\|_{0, \Gamma_C} \} \\ & \leq C_T A(\phi) \{ \|\tilde{\omega}\|_{1/2+\epsilon} + \|\tilde{\psi}\|_{1/2+\epsilon} \} \quad \forall (\tilde{\psi}, \tilde{\omega}) \in H, \end{aligned}$$

where

$$A(\phi) = C \max \left\{ \frac{1}{\delta_0}, \frac{1}{\epsilon_0}, \max_{|\mathbf{x}| \leq \|\phi\|_{3/2+\epsilon}} (|f'(\mathbf{x})| + \alpha), \max_{|\mathbf{x}| \leq \|\phi\|_{3/2+\epsilon}} (|f''(\mathbf{x})| \|\lambda\|_{3/2+\epsilon}) \right\}$$

and C_T is a constant such that $\|\psi\|_{0,\Gamma} \leq C_T \|\psi\|_{1/2+\epsilon}$ for all $\psi \in H^{1/2+\epsilon}(\Omega)$. The desired results follow easily from this estimate. ■

If (ϕ, λ) is a nonsingular solution of (5.2)-(5.3), using the denseness of $H^1(\Omega)$ in $H^{1/2+\epsilon}(\Omega)$ and regularity theories for (6.39)-(6.40), we infer that (6.12) holds.

Thus we have verified all the requirements in Theorem 6.2 so that we can draw the following conclusion:

Theorem 6.7 *Assume Ω is convex and (ϕ, λ) is a nonsingular solution of the optimality system (5.2)-(5.3). Assume that the finite element spaces V^h satisfy the condition (6.1). Then, there exists a $\delta > 0$ and $h_0 > 0$ such that for $h \leq h_0$, there exists a unique nonsingular solution (ϕ^h, λ^h) of the discrete optimality system (6.2)-(6.3) satisfying $\|\phi^h - \phi\|_1 + \|\lambda^h - \lambda\|_1 \leq \delta$. If, in addition, the solution of the optimality system satisfies $(\phi, \lambda) \in H^{m+1}(\Omega) \times H^{m+1}(\Omega)$, then there exists a constant C , independent of h , such that*

$$\|\phi - \phi^h\|_{\epsilon+1/2} + \|\lambda - \lambda^h\|_{\epsilon+1/2} \leq C h^{m-\epsilon+1/2} (\|\phi\|_{m+1} + \|\lambda\|_{m+1}). \quad \blacksquare \quad (6.45)$$

A consequence of Theorem 6.7 is the following corollary that gives the $L^2(\Gamma_A)$ -error estimates for the the approximation of the controls.

Corollary 6.8 *Assume Ω is convex and (ϕ, λ) is a nonsingular solution of the optimality system (5.2)-(5.3). Assume that the finite element spaces V^h satisfy the condition (6.1). Define the approximate control by*

$$u^h = -\frac{1}{\delta_0} \lambda^h \quad \text{on } \Gamma_A.$$

If the solution of the optimality system satisfies $(\phi, \lambda) \in H^{m+1}(\Omega) \times H^{m+1}(\Omega)$, then for each $\epsilon \in (0, 1/4)$ there exists a constant C , independent of h , such that for $h \leq h_0$,

$$\|u^h - u\|_{0,\Gamma_A} \leq \frac{C}{\delta_0} h^{m-\epsilon+1/2} (\|\phi\|_{m+1} + \|\lambda\|_{m+1}). \quad (6.46)$$

Proof: Recall that $u = -\frac{1}{\delta_0} \lambda$ on Γ_A ; see (5.1). Then (6.46) follows trivially from (6.45) and the trace theorems (see [1])

$$\|u - u^h\|_{0,\Gamma_A} = \frac{1}{\delta_0} \|\lambda - \lambda^h\|_{0,\Gamma_C} \leq \frac{1}{\delta_0} \|\lambda - \lambda^h\|_{0,\Gamma} \leq \frac{C}{\delta_0} \|\lambda - \lambda^h\|_{\epsilon+1/2}. \quad \blacksquare$$

Acknowledgements. The authors thank Dr. W. Sun for useful discussions related to this article.

REFERENCES

- [1] R. Adams, *Sobolev Spaces*, Academic Press, New York (1975).
- [2] K. Amaya and S. Aoki, Optimum design of cathodic protection system by 3-D BEM. *Boundary Element Technology VII*, 375-388 (1992).
- [3] I. Babuška and A. Aziz, Survey lectures on the mathematical foundations of the finite element method, in *The Mathematical Foundations of the Finite Element Method with Applications to Partial Differential Equations* (ed. A. Aziz), Academic, New York, 3 (1972).
- [4] J.O.M. Bockris and A.K.M. Reddy, *Modern Electrochemistry*, Plenum Press, New York (1974).
- [5] F. Brezzi, J. Rappaz and P.-A. Raviart, Finite-dimensional approximation of nonlinear problems. Part I: Branches of nonsingular solutions, *Numer. Math.* **36**, 1 (1980).
- [6] P. Ciarlet, *The Finite Element Method for Elliptic Problems*, North-Holland, Amsterdam (1978).
- [7] M. Crouzeix and J. Rappaz, *On Numerical Approximation in Bifurcation Theory*, Masson, Paris (1989).
- [8] V. Girault and P.-A. Raviart, *Finite Element Methods for Navier-Stokes Equations*. Springer, Berlin (1986).
- [9] J.P. Gossez, Nonlinear elliptic boundary value problems for equations with rapidly or slowly increasing coefficients. *Trans. Amer. Math. Soc.* **190**, 163-205 (1974).
- [10] Grisvard, Commutativité de deux foncteurs d'interpolation et applications. *J. Math. Pures Appl.* **45**, 143-290 (1966).
- [11] Grisvard, *Elliptic Problems on Nonsmooth Domains*. Pitman, Boston, 1985.
- [12] P. Hess, A strongly nonlinear elliptic boundary value problem. *J. Math. Anal. Appl.* **43**, 241-249 (1973).
- [13] P. Hess, On nonlinear mappings of monotone type with respect to two Banach spaces. *J. Math. Pures Appl.* **52**, 13-26 (1973).
- [14] L.S. Hou and W. Sun, Numerical methods for optimal control of impressed cathodic protection systems, to appear.
- [15] L.S. Hou and W. Sun, Optimal positioning of anodes for cathodic protection, to appear.
- [16] J. Kačur, On boundedness of the weak solution for some class of quasilinear partial differential equations. *Čas. Pěst. Mat.* **98**, 43-55 (1973).
- [17] J. Nečas, *Les Méthodes Directes en Théorie des Équations Elliptiques*. Academia, Prague (1967).
- [18] Peetre, Espaces d'interpolation et théorème de Soboleff. *Ann. Inst. Fourier (Grenoble)* **16**, 279-317 (1966).
- [19] V. Tikhomorov, *Fundamental Principles of the Theory of Extremal Problems*. Wiley, Chichester, 1982.
- [20] N.G. Zamani and J.M. Chuang, Optimal control of current in a cathodic protection system: a numerical investigation. *Optimal Control Appl. & Meth.*, **8**, 339-350 (1987).
- [21] N.G. Zamani, J.F. Porter and A.A. Mufti, A survey of computational efforts in the field of corrosion engineering. *Int. J. Numer. Meth. Eng.*, **23**, 1295-1311 (1986).

ANALYSIS AND APPROXIMATION OF OPTIMAL CONTROL PROBLEMS WITH NONLINEAR CONSTRAINTS[†]

Max D. Gunzburger

L. Steven Hou

James C. Turner, Jr.

Abstract

A general framework for treating nonlinearly constrained optimal control and optimization problems is given. Based on a set of hypotheses, optimal solutions are shown to exist and the use of Lagrange multipliers to enforce the constraints is justified. An optimality system is derived whose solutions provide the optimal states and controls. Finite dimensional approximations are then considered; an approximate problem is defined and optimal error estimates are derived. The general framework has been applied to numerous concrete settings; here we illustrate its use in the context of a magnetohydrodynamics control problem.

1. Introduction

The need to solve nonlinear optimization and control problems arises in many settings. Here, our goal is to construct and analyze a framework for the approximate solution of many such problems. The components of our framework are as follows:

- define an general class of nonlinear control or optimization problems;
- show that optimal solutions exist;
- show that Lagrange multipliers exist that may be used to enforce the constraints;
- use the Lagrange multiplier technique to derive an optimality system from which optimal states and controls may be deduced;
- define algorithms for the approximation, in finite dimensional spaces, of optimal states and controls; and
- derive estimates for the error in the approximations to the optimal states and controls.

After developing and analyzing the general framework, we will apply it to a specific, concrete problem,

namely, the control of a magnetohydrodynamic flow. We use the general framework to analyze the concrete problem by merely showing that the latter fits into the former. We will also provide the results of a numerical implementation, using finite element methods, of our theory. Due to space limitations, we cannot provide detailed proofs of our results; these are given in [4].

2. The general problem

We assume that the functional to be minimized takes the form

$$\mathcal{J}(v, z) = \lambda \mathcal{F}(v) + \lambda \mathcal{E}(z) \quad \forall (v, z) \in X \times \Theta. \quad (1)$$

Here, λ is a given parameter which is assumed to belong to a compact interval $\Lambda \subset \mathbb{R}_+$, $\mathcal{F}$ is a functional on a reflexive Banach space X , and $\mathcal{E}$ a functional on Θ , where Θ is closed and convex subset of another reflexive Banach space G .

The constraint equation $M(v, z) = 0$ relating the state variable v and the control variable z is defined as follows. Let N be a differentiable mapping from X to a third reflexive Banach space Y , K a continuous linear operator from Θ to Y , and T a continuous linear operator from Y to X . Then, for any $\lambda \in \Lambda$ and $(v, z) \in X \times \Theta$, the mapping M from $X \times \Theta$ to X , is given by

$$M(v, z) = v + \lambda TN(v) + \lambda TK(z). \quad (2)$$

With these definitions we now consider the constrained minimization problem:

$$\min_{(v, z) \in X \times \Theta} \mathcal{J}(v, z) \quad \text{subject to} \quad M(v, z) = 0. \quad (3)$$

Although we will show that global minimizers as defined by (3) exist, in practice one can usually only locate local minimizers, i.e., points $(u, g) \in X \times \Theta$ such that for some $\epsilon > 0$

$$\begin{aligned} \mathcal{J}(u, g) &\leq \mathcal{J}(v, z) \quad \forall (v, z) \in X \times \Theta \\ \text{such that } M(v, z) &= 0 \text{ and } \|u - v\|_X \leq \epsilon. \end{aligned} \quad (4)$$

The following hypotheses will be invoked in the analyses that follow:

- (H1) $\inf_{v \in X} \mathcal{F}(v) > -\infty$;
- (H2) there exist constants $\alpha, \beta > 0$ such that $\mathcal{E}(z) \geq \alpha \|z\|^\beta \forall z \in \Theta$;
- (H3) there exists a $(v, z) \in X \times \Theta$ satisfying $M(v, z) = 0$;
- (H4) if $u^{(n)} \rightarrow u$ in X and $g^{(n)} \rightarrow g$ in G where $\{(u^{(n)}, g^{(n)})\} \subset X \times \Theta$, then $N(u^{(n)}) \rightarrow N(u)$ in Y and $K(g^{(n)}) \rightarrow K(g)$ in Y ;
- (H5) $\mathcal{J}(\cdot, \cdot)$ is weakly lower semicontinuous on $X \times \Theta$; and
- (H6) if $\{(u^{(n)}, g^{(n)})\} \subset X \times \Theta$ is such that $\{\mathcal{F}(u^{(n)})\}$ is a bounded set in $\mathbb{R}$ and $M(u^{(n)}, g^{(n)}) = 0$, then $\{u^{(n)}\}$ is a bounded set in X .
- (H7) for each $z \in \Theta$, $v \mapsto \mathcal{J}(v, z)$ and $v \mapsto M(v, z)$ are Fréchet differentiable;
- (H8) $z \mapsto \mathcal{E}(z)$ is convex;
- (H9) for $v \in X$, $N'(v)$ maps X into Z ;
- (H10) $\Theta = G$, and the mapping $z \mapsto \mathcal{E}(z)$ is Fréchet differentiable on G ;
- (H11) if $v^* \in X^*$ satisfies $(I + \lambda [N'(u)]^* T^*) v^* = 0$ and $K^* T^* v^* = 0$, then $v^* = 0$ or
- (H11)' the mapping $(v, z) \mapsto v + \lambda TN'(u)v + \lambda TKz$ is onto from $X \times G$ to Y ; and
- (H12) $\mathcal{E}'(g) = E^{-1}g$, where E is an invertible linear operator from G^* to G and g is an optimal control for the constrained minimization problem (4).

In (H9), N' denotes the Fréchet derivative of N and Z is a subspace of Y with a compact imbedding.

Assumptions (H1)-(H6) are used to establish the existence of optimal solutions; (H7)-(H9) will be needed to justify the use of the Lagrange multiplier rule and to derive an optimality system from which, for almost all $\lambda \in \Lambda$, optimal states and controls may be determined; (H10) will allow for a simplification of the optimality system; either or (H11) or (H11)' can be used to extend the validity of the optimality system to all values of $\lambda \in \Lambda$; finally, (H12) can be used to render the optimality system more amenable to approximation and computation.

2.1. Existence results and an optimality system

A fairly standard method of proof yields the existence of optimal solutions.

Theorem 1 Assume that the functional $\mathcal{J}$ and mapping M defined by (1) and (2), respectively, satisfy (H1)-(H6). Then, there exists a solution to the minimization problem (3).

Using a proof found in [8] of an abstract Lagrange multiplier rule, one can show that the Lagrange multiplier rule may be used to turn the constrained minimization problem (4) into an unconstrained one.

Theorem 2 Let $\lambda \in \Lambda$ be given. Assume that (H1)-(H10) hold. Let $(u, g) \in X \times G$ be a solution of the problem (4). Then, there exists a $k \in \mathbb{R}$ and a $\mu \in X^*$ that are not both equal to zero such that

$$k \langle \mathcal{J}_u(u, g), w \rangle - \langle \mu, (I + \lambda TN'(u))w \rangle = 0 \quad \forall w \in X \quad (5)$$

$$k \langle \mathcal{E}'(g), z \rangle - \langle \mu, TKz \rangle = 0 \quad \forall z \in G. \quad (6)$$

If $(1/\lambda) \notin \sigma(-TN'(u))$, i.e., if $(1/\lambda)$ is not in the spectrum of $-TN'(u)$, we may choose $k = 1$.

Actually, only (H1)-(H9) are needed to apply the result of [8] and show that the Lagrange multiplier rule may be used; (H10) simply simplifies the optimality system.

It can be shown that $TN'(u)$ is a compact operator from X to X , and therefore $\sigma(-TN'(u))$ is at most countable with zero being the only possible limit point. Thus, (5)-(6) are valid for almost all $\lambda \in \Lambda$. If the control g enters the constraint in a favorable manner, then we may take $k = 1$ even when $(1/\lambda) \in \sigma(-TN'(u))$. Specifically, we invoke one of the assumptions (H11) and (H11)'.

Theorem 3 Assume that the hypotheses of Theorem 2 hold. Assume that if $(1/\lambda) \in \sigma(-TN'(u))$, then either (H11) or (H11)' holds. Then, for all $\lambda \in \Lambda$, there exists a $\mu \in X^*$ such that (5) and (6) hold with $k = 1$.

Under the assumptions of Theorem 3, an optimal state $u \in X$, an optimal control $g \in G$, and the corresponding Lagrange multiplier $\mu \in X^*$ satisfy the optimality system of equations formed by (2), (5), and (6) with $k = 1$. From (1) we have that $\mathcal{J}_u = \lambda \mathcal{F}'$ and $\mathcal{J}_g = \lambda \mathcal{E}'$, where $\mathcal{F}'$ denotes the obvious Fréchet derivative. For purposes of numerical approximations, it turns out to be convenient to make the change of variable $\xi = T^* \mu$. It will also be convenient to invoke the additional simplifying assumption (H12) concerning the dependence of the objective functional on the control. Then, (2), (5), and (6) with $k = 1$ can be rewritten as

$$u + \lambda TN(u) + \lambda TKg = 0 \quad \text{in } X, \quad (7)$$

$$\xi + \lambda T^*[N'(u)]^* \xi - \lambda T^* \mathcal{F}'(u) = 0 \quad \text{in } Y^*, \quad (8)$$

$$g - EK^* \xi = 0 \quad \text{in } G. \quad (9)$$

Clearly, the control g may be eliminated from the optimality system (7)-(9). However, given a $\xi \in Y^*$, it is not always possible to evaluate g exactly from (9). For example, the application of the operator E may involve the solution of a partial differential equation. In many applications we have that $X^* = Y$. Since these spaces are assumed to be reflexive, we also have that $Y^* = X$. In this case, we have that both u and ξ belong to X .

3. Finite dimensional approximations

A finite dimensional discretization of the optimality system (7)-(9) is defined as follows. First, one chooses families of finite dimensional subspaces $X^h \subset X$, $(Y^*)^h \subset Y^*$, and $G^h \subset G$. These families are parameterized by a parameter h that tends to zero. (For example, this parameter can be chosen to be some measure of the grid size in a subdivision of Ω into finite elements.) Next, we define approximate operators $T^h : Y \rightarrow X^h$, $E^h : G^* \rightarrow G^h$, and $(T^*)^h : X^* \rightarrow (Y^*)^h$. Of course, one views T^h , E^h , and $(T^*)^h$ as approximations to the operators T , E , and T^* , respectively. Note that $(T^*)^h$ is not necessarily the same as $(T^h)^*$. The former is a discretization of an adjoint operator while the latter is the adjoint of a discrete operator.

Once the approximating subspaces and operators have been chosen, an approximate problem is defined as

follows. We seek $u^h \in X^h$, $g^h \in G^h$, and $\xi^h \in (Y^*)^h$ such that

$$u^h + \lambda T^h N(u^h) + \lambda T^h K g^h = 0 \quad \text{in } X^h, \quad (10)$$

$$\begin{aligned} \xi^h + \lambda (T^*)^h [N'(u^h)]^* \xi^h \\ - \lambda (T^*)^h \mathcal{F}'(u^h) = 0 \quad \text{in } (Y^*)^h, \end{aligned} \quad (11)$$

$$g^h - E^h K^* \xi^h = 0 \quad \text{in } G^h. \quad (12)$$

We make the following hypotheses concerning the approximate operators T^h , $(T^*)^h$, and E^h and the operators appearing in the definition of the abstract problem (4):

- (H13) $\lim_{h \rightarrow 0} \|(T - T^h)y\|_X = 0 \quad \forall y \in Y$;
(H14) $\lim_{h \rightarrow 0} \|(T^* - (T^*)^h)v\|_{Y^*} = 0 \quad \forall v \in X^*$;
(H15) $\lim_{h \rightarrow 0} \|(E - E^h)s\|_G = 0 \quad \forall s \in G^*$;
(H16) $N \in C^3(X; Y)$ and $\mathcal{F} \in C^3(X; \mathbb{R})$;
(H17) N'' , N''' , $\mathcal{F}''$, and $\mathcal{F}'''$ are locally bounded, i.e., they map bounded sets to bounded sets;
(H18) for $v \in X$, in addition to (H9), i.e., $N'(v) \in \mathcal{L}(X; Z)$ where $Z \hookrightarrow Y$, we have that $[N'(v)]^* \in \mathcal{L}(Y^*; \tilde{Z})$ where $\tilde{Z} \hookrightarrow X^*$, that for $\eta \in Y^*$, $[N''(v)]^* \cdot \eta \in \mathcal{L}(Y^*; \tilde{Z})$, and that for $w \in X$, $\mathcal{F}''(v) \cdot w \in \mathcal{L}(X; \tilde{Z})$; and

(H19) K maps G into Z .

Here, $(\cdot)''$ and $(\cdot)'''$ denote second and third Fréchet derivatives, respectively.

The proof of the error estimate given below makes use of results of [1] and [2] (see also [3]) concerning the approximation of a class of nonlinear problems. These results imply that, under certain hypotheses, the error of approximation of solutions of certain nonlinear problems is basically the same as the error of approximation of solutions of related linear problems.

We assume that the solution $(u(\lambda), g(\lambda), \xi(\lambda))$ of the optimality system (7)-(9) that we are trying to approximate is a regular solution. The assumptions we have made, in particular (H9) and (H18)-(H19), are sufficient to guarantee that for almost all values of λ , this is indeed the case.

Theorem 4 Let $(u(\lambda), g(\lambda), \xi(\lambda)) \in \mathcal{X}$, for $\lambda \in \Lambda$, be a branch of regular solutions of the optimality system (7)-(9). Assume that the hypotheses (H13)-(H19) hold. Then, there exists a $\delta > 0$ and an $h_0 > 0$ such that for $h < h_0$, the discrete optimality system (10)-(12) has a unique solution $(u^h(\lambda), g^h(\lambda), \xi^h(\lambda))$ satisfying

$$\|(u(\lambda), g(\lambda), \xi(\lambda)) - (u^h(\lambda), g^h(\lambda), \xi^h(\lambda))\|_{\mathcal{X}} < \delta.$$

Moreover,

$$\lim_{h \rightarrow 0} \|(u(\lambda), g(\lambda), \xi(\lambda)) - (u^h(\lambda), g^h(\lambda), \xi^h(\lambda))\|_{\mathcal{X}} = 0$$

uniformly in $\lambda \in \Lambda$ and there exists a constant C , independent of h and λ , such that

$$\begin{aligned} \lim_{h \rightarrow 0} \|(u(\lambda), g(\lambda), \xi(\lambda)) - (u^h(\lambda), g^h(\lambda), \xi^h(\lambda))\|_{\mathcal{X}} \\ \leq C \lambda \left\{ \|(T^h - T)(N(u(\lambda)) + K g(\lambda))\|_X \right. \\ \left. + \|(E^h - E)K^* \xi(\lambda)\|_G \right. \\ \left. + \|((T^*)^h - T^*)([N'(u(\lambda))]^* \xi - \mathcal{F}'(u(\lambda)))\|_{Y^*} \right\}. \end{aligned}$$

If the operator T from Y to X is invertible, it is easily seen that the last relation is equivalent to

$$\begin{aligned} \|u(\lambda) - u^h(\lambda)\|_X + \|g(\lambda) - g^h(\lambda)\|_G + \|\xi(\lambda) - \xi^h(\lambda)\|_{Y^*} \\ \leq C \left\{ \|(T^h T^{-1} - I)u(\lambda)\|_X + \|(E^h E^{-1} - I)g(\lambda)\|_G \right. \\ \left. + \|((T^*)^h (T^*)^{-1} - I)\xi(\lambda)\|_{Y^*} \right\}. \end{aligned}$$

4. Application to an MHD control problem

We now apply the theories presented in the previous sections to an optimal control problem for a special type of MHD (magnetohydrodynamic) flow in which we assume that the externally applied magnetic field $\mathbf{B}$ is undisturbed by the flow. We denote by Ω the flow domain which is a bounded open set in $\mathbb{R}^2$ with a boundary Γ . The dimensionless equations governing this special type of MHD flow are given by (see [9] and [10])

$$\begin{aligned} -\frac{1}{M^2} \Delta \mathbf{u} + \frac{1}{N} (\mathbf{u} \cdot \nabla) \mathbf{u} + \nabla p - (\mathbf{B} \times \nabla \phi) \\ - (\mathbf{u} \times \mathbf{B}) \times \mathbf{B} = \mathbf{f} \quad \text{in } \Omega, \end{aligned} \quad (13)$$

$$\operatorname{div} \mathbf{u} = 0 \quad \text{in } \Omega, \quad (14)$$

$$-\Delta \phi + \operatorname{div} (\mathbf{u} \times \mathbf{B}) = 0 \quad \text{in } \Omega. \quad (15)$$

Here $\mathbf{u}$ denotes the velocity field, p the pressure field, and ϕ the electric potential. Also, N is the interaction number and M is the Hartmann number. $\mathbf{B}$ and $\mathbf{f}$ are given data. The system (13)-(15) is supplemented with the boundary conditions

$$\mathbf{u} = \mathbf{0} \quad \text{on } \Gamma, \quad (16)$$

$$\frac{\partial \phi}{\partial n} = g \quad \text{on } \Gamma, \quad (17)$$

where g denotes the only control variable, namely, the normal electric current on Γ . Such a control can be effected by attaching electric sources with adjustable resistors to the electrode along the flow boundary.

Our goal is to try to obtain a desired flow field by appropriately choosing the control - the normal electric current on Γ . Mathematically, this task can be described by the following optimal control setting: minimize the cost functional

$$\mathcal{J}(\mathbf{u}, \phi, p, g) = \frac{1}{2\epsilon N} \int_{\Omega} |\mathbf{u} - \mathbf{u}_d|^2 d\Omega + \frac{\delta}{2N} \int_{\Gamma} |g|^2 d\Gamma \quad (18)$$

subject to the constraints (13)-(17). Here $\epsilon > 0$ and $\delta > 0$ are positive parameters, and $\mathbf{u}_d$ is a desired velocity field.

We introduce the function space $H^1(\Omega) = \{v \in L^2(\Omega) : \frac{\partial v}{\partial x_i} \in L^2(\Omega) \text{ for } i = 1, 2\}$; $H_0^1(\Omega) = \{v \in H^1(\Omega) : v|_{\Gamma} = 0\}$; $L_0^2(\Omega) = \{q \in L^2(\Omega) : \int_{\Omega} q d\Omega = 0\}$; $\tilde{H}^1(\Omega) = H^1(\Omega) \cap L_0^2(\Omega)$; and $H^m(\Omega) = \{v \in L^2(\Omega) : \frac{\partial^{|\alpha|} v}{\partial x_1^{\alpha_1} \partial x_2^{\alpha_2}} \in L^2(\Omega), \text{ for all } \alpha = (\alpha_1, \alpha_2) \text{ with } |\alpha| \leq m\}$. Vector-valued counterparts of these spaces are denoted by bold-face symbols, e.g., $\mathbf{H}^1(\Omega) = [H^1(\Omega)]^2$.

Introducing the forms

$$a_1(\mathbf{u}, \mathbf{v}) = \frac{1}{M^2} \int_{\Omega} \nabla \mathbf{u} : \nabla \mathbf{v} d\Omega + \int_{\Omega} (\mathbf{u} \times \mathbf{B}) \cdot (\mathbf{v} \times \mathbf{B}) d\Omega$$

$$\begin{aligned}
a_2(\phi, \psi) &= \int_{\Omega} \nabla \phi : \nabla \psi \, d\Omega, \\
b_1(\mathbf{v}, q) &= - \int_{\Omega} q \operatorname{div} \mathbf{v} \, d\Omega, \\
b_2(\mathbf{v}, \phi) &= - \int_{\Omega} (\nabla \phi) \cdot (\mathbf{v} \times \mathbf{B}) \, d\Omega, \\
c(\mathbf{u}, \mathbf{v}, \mathbf{w}) &= \int_{\Omega} (\mathbf{u} \cdot \nabla) \mathbf{v} \cdot \mathbf{w} \, d\Omega,
\end{aligned}$$

for all $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathbf{H}^1(\Omega)$, $q \in H^1(\Omega)$ and $\phi, \psi \in \tilde{H}^1(\Omega)$, and introducing the parameter $\lambda = \frac{1}{N}$, and respectively rescaling $\mathbf{f}$ and g to $\lambda \mathbf{f}$ and λg , we may define, in a standard manner, weak solutions for (13)-(17): a triplet $(\mathbf{u}, p, \phi)$ is called a weak solution of (13)-(17) if $\mathbf{u} \in \mathbf{H}_0^1(\Omega)$, $p \in L_0^2(\Omega)$ and $\phi \in \tilde{H}^1(\Omega)$ satisfies

$$a_1(\mathbf{u}, \mathbf{v}) + b_1(\mathbf{v}, p) + b_2(\mathbf{v}, \phi) + \lambda c(\mathbf{u}, \mathbf{u}, \mathbf{v}) = \lambda(\mathbf{f}, \mathbf{v}) \quad \forall \mathbf{v} \in \mathbf{H}_0^1(\Omega), \quad (19)$$

$$a_2(\phi, \psi) + b_2(\mathbf{u}, \psi) = \lambda(g, \psi)_{\Gamma} \quad \forall \psi \in \tilde{H}^1(\Omega), \quad (20)$$

$$b_1(\mathbf{u}, q) = 0 \quad \forall q \in L_0^2(\Omega). \quad (21)$$

Here, the colon notation stands for the scalar product on $\mathbb{R}^{2 \times 2}$, and the notations $(\cdot, \cdot)$ and $(\cdot, \cdot)_{\Gamma}$ denote, respectively, the $L^2(\Omega)$ and $L^2(\Gamma)$ inner products.

We define the spaces $X = \mathbf{H}_0^1(\Omega) \times L_0^2(\Omega) \times \tilde{H}^1(\Omega)$, $Y = \mathbf{H}^{-1}(\Omega) \times L_0^2(\Omega) \times H^{-1/2}(\Gamma)$, $G = L^2(\Gamma)$, and $Z = L^{3/2}(\Omega) \times \{0\} \times L^2(\Gamma)$. By compact imbedding results, Z is compactly imbedded into Y . For the time being, we assume that the admissible set Θ for the control g is a closed, convex subset of $G = L^2(\Gamma)$.

The precise mathematical statement of the optimal control problem (3) in this concrete setting is given as follows:

$$\begin{aligned}
&\text{seek } a(\mathbf{u}, p, \phi, g) \in \mathbf{H}_0^1(\Omega) \times L_0^2(\Omega) \times \tilde{H}^1(\Omega) \times \\
&\Theta \text{ such that the functional (18) is} \\
&\text{minimized subject to the constraints (19)-(21).}
\end{aligned} \quad (22)$$

With the obvious definitions for $\mathcal{F}(\cdot)$ and $\mathcal{E}(\cdot)$, the functional (18) can be expressed as

$$\mathcal{J}(\mathbf{u}, p, \phi, g) = \lambda \mathcal{F}(\mathbf{u}, p, \phi) + \lambda \mathcal{E}(g),$$

i.e., in the form of (1).

We define the bounded linear operator $T \in \mathcal{L}(Y; X)$ as follows: for each $(\zeta, \eta, \kappa) \in Y$, $T(\zeta, \eta, \kappa) = (\tilde{\mathbf{u}}, \tilde{p}, \tilde{\phi}) \in X$ is the unique solution of

$$a_1(\tilde{\mathbf{u}}, \mathbf{v}) + b_1(\mathbf{v}, \tilde{p}) + b_2(\mathbf{v}, \tilde{\phi}) = (\zeta, \mathbf{v}) \quad \forall \mathbf{v} \in \mathbf{H}_0^1(\Omega),$$

$$b_1(\tilde{\mathbf{u}}, q) = (\eta, q), \quad \forall q \in L_0^2(\Omega)$$

$$a_2(\tilde{\phi}, \psi) + b_2(\tilde{\mathbf{u}}, \psi) = (\kappa, \psi)_{\Gamma} \quad \forall \psi \in \tilde{H}^1(\Omega).$$

It can be easily verified that T is self-adjoint. We respectively define the (differentiable) nonlinear mapping $N : X \rightarrow Y$ and the linear mapping $K : H^{-1/2}(\Gamma) \rightarrow Y$ by

$$N(\mathbf{u}, p, \phi) = -(\mathbf{f} - (\mathbf{u} \cdot \nabla) \mathbf{u}, 0, 0)^T$$

and

$$Kg = -(0, 0, g)^T.$$

Clearly, the constraint equations (19)-(21) can be expressed as $(\mathbf{u}, p, \phi) + \lambda TN(\mathbf{u}, p, \phi) + \lambda TKg = 0$, i.e., in the form (2).

We first briefly verify, for the minimization problem (22), all the hypotheses of §2. Then we define finite element approximations and briefly verify the hypotheses in §3. Details are similar to those found in [4] and [5] for optimal control problems in some different settings.

4.1. Verification of existence hypotheses

We verify that the hypotheses for the existence of optimal solutions, i.e., the hypotheses (H1)-(H6), hold in the current setting. (H1) is obviously satisfied with a lower bound 0. (H2) holds with $\alpha = 1$ and $\beta = 2$. (H3) follows from results in [7] on the existence of solutions for the constraint equations. (H4) follows from the properties of weak limits and compact imbeddings. (H5) is a simple consequence of the observation that the mappings $(\mathbf{u}, p, \phi) \mapsto \mathcal{F}(\mathbf{u}, p, \phi) = (1/2) \|\mathbf{u} - \mathbf{u}_d\|_0^2$ and $g \mapsto \mathcal{E}(g) = (1/2) \|g\|_{0,\Gamma}^2$ are convex. (H6) can be easily verified through a combination of the a priori estimates for the constraint equations and the estimates obtained from the functional.

It is now just a matter of citing Theorem 1 to conclude the existence of an optimal solution that minimizes (18) subject (19)-(21).

Theorem 5 *There exists a solution $(\mathbf{u}, p, \phi, g) \in \mathbf{H}_0^1(\Omega) \times L_0^2(\Omega) \times \tilde{H}^1(\Omega) \times \Theta$ for the minimization problem (22).*

We now assume that $(\mathbf{u}, p, \phi, g)$ is an optimal solution and turn to the verification of hypotheses for the existence of Lagrange multipliers, i.e., the hypotheses (H7)-(H9). (H7) and (H8) can be easily verified. (H9) can be verified as follows. For any $(\mathbf{u}, p, \phi) \in X$, the operator $N'(\mathbf{u}, p, \phi) : X \rightarrow Y$ is given by $N'(\mathbf{u}, p, \phi) \cdot (\mathbf{v}, q, \psi) = -(\mathbf{u} \cdot \nabla \mathbf{v} + \mathbf{v} \cdot \nabla \mathbf{u}, 0, 0)^T$ for all $(\mathbf{v}, q, \psi) \in \mathbf{H}_0^1(\Omega) \times L_0^2(\Omega) \times \tilde{H}^1(\Omega)$. Thus we obtain $N'(\mathbf{u}, p, \phi) \cdot (\mathbf{v}, q, \psi) \in L^{3/2}(\Omega) \times \{0\} \times \{0\} = Z$.

Having verified the hypotheses (H7)-(H9), we may apply Theorem 2 to conclude that there exists a Lagrange multiplier $(\nu, \tau, s) \in X = \mathbf{H}_0^1(\Omega) \times L_0^2(\Omega) \times \tilde{H}^1(\Omega)$ and a real number k such that (5)-(6) hold and that for almost all values of λ , we may choose $k = 1$ in (5)-(6).

We now specialize to the case $\Theta = L^2(\Gamma)$, then (H10) is obviously satisfied; we can also verify (H11) with some effort. Thus, we can use Theorem 3 to show that (5)-(6) hold for all $\lambda \in \Lambda$ with $k = 1$.

Theorem 6 *There exists a $(\mathbf{u}, p, \phi, g, \nu, \tau, s) \in \mathbf{H}_0^1(\Omega) \times L_0^2(\Omega) \times \tilde{H}^1(\Omega) \times L^2(\Gamma) \times \mathbf{H}_0^1(\Omega) \times L_0^2(\Omega) \times \tilde{H}^1(\Omega)$ such that $(\mathbf{u}, p, \phi, g)$ is a solution for (22) with $\Theta = L^2(\Gamma)$ and (5)-(6) hold with $k = 1$.*

(H12) holds with $E = I$, the identity mapping on $L^2(\Gamma)$. Hence, we may express (5)-(6) as

$$\begin{aligned}
&a_1(\mathbf{w}, \nu) + \lambda c(\mathbf{w}, \mathbf{u}, \nu) + \lambda c(\mathbf{u}, \mathbf{w}, \nu) + b_1(\mathbf{w}, \tau) \\
&+ b_2(\mathbf{w}, s) = \frac{\lambda}{\epsilon} (\mathbf{u} - \mathbf{u}_0, \mathbf{w}) \quad \forall \mathbf{w} \in \mathbf{H}_0^1(\Omega),
\end{aligned} \quad (23)$$

$$b_1(\nu, \sigma) = 0 \quad \forall \sigma \in L_0^2(\Omega), \quad (24)$$

$$a_2(r, s) + b_2(\nu, r) = 0 \quad \forall r \in \tilde{H}^1(\Omega), \quad (25)$$

and

$$\delta(g, z)_\Gamma - (s, z)_\Gamma = 0 \quad \forall z \in \Theta. \quad (26)$$

The system formed by (19)-(21) and (23)-(26) is called an *optimality system* of equations. Note that we may use (26) to eliminate the control g from the problem formulation.

4.2. Finite element approximations

A finite element discretization of the optimality system (19)-(21) and (23)-(26) is defined in the usual manner. First one chooses families of finite dimensional subspaces $X^h \subset H^1(\Omega)$, $S^h \subset L^2(\Omega)$ and $G^h \subset L^2(\Gamma)$. These finite dimensional function spaces are defined on an approximate domain Ω^h . For simplicity we will state our results in this section by assuming $\Omega^h = \Omega$, which is the case when Ω is a convex polygon. We set $\tilde{X}^h = X^h \cap L_0^2(\Omega)$, $X_0^h = [X^h]^2$, $X_0^h = X^h \cap H_0^1(\Omega)$ and $S_0^h = S^h \cap L_0^2(\Omega)$. Here we may choose any pair of subspaces X^h and S^h such that X_0^h and S_0^h 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 , $\mathbf{v}$, v and g , such that

$$\inf_{\mathbf{v}^h \in X_0^h} \|\mathbf{v} - \mathbf{v}^h\|_1 \leq Ch^m \|\mathbf{v}\|_{m+1} \quad \forall \mathbf{v} \in H^{m+1}(\Omega) \cap H_0^1(\Omega),$$

$$\inf_{v^h \in S_0^h} \|v - v^h\|_1 \leq Ch^m \|v\|_{m+1} \quad \forall v \in H^{m+1}(\Omega) \cap L_0^2(\Omega),$$

$$\inf_{q^h \in S_0^h} \|q - q^h\|_0 \leq Ch^m \|q\|_m \quad \forall q \in H^m(\Omega) \cap L_0^2(\Omega),$$

$$\inf_{z^h \in G^h} \|g - g^h\|_{0,\Gamma} \leq Ch^m \|g\|_m \quad \forall g \in H^m(\Gamma),$$

for $1 \leq m \leq k$. The last approximation property is redundant if we choose $G^h = X^h|_\Gamma$. Next, we assume the *inf-sup condition*: there exists a constant C , independent of h , such that

$$\inf_{0 \neq q^h \in S_0^h} \sup_{0 \neq \mathbf{v}^h \in X_0^h} \frac{-\int_\Omega q^h \operatorname{div} \mathbf{v}^h d\Omega}{\|\mathbf{v}^h\|_1 \|q^h\|_0} \geq C.$$

This condition assures the stability of finite element discretizations. For thorough discussions of the above approximation properties and the inf-sup condition, see, e.g., [3]. The reference [3] may also be consulted for a catalog of finite element subspaces that meet the requirements of the foregoing approximation properties and inf-sup condition.

Once the approximating subspaces have been chosen, we seek $\mathbf{u}^h \in X_0^h$, $p^h \in S_0^h$, $\phi^h \in \tilde{X}^h$, $g^h \in G^h$, $\nu^h \in X_0^h$, $\tau^h \in S_0^h$ and $s^h \in \tilde{X}^h$ such that

$$\begin{aligned} a_1(\mathbf{u}^h, \mathbf{v}^h) + b_1(\mathbf{v}^h, p^h) + b_2(\mathbf{v}^h, \phi^h) + \lambda c(\mathbf{u}^h, \mathbf{u}^h, \mathbf{v}^h) \\ = \lambda(\mathbf{f}, \mathbf{v}^h) \quad \forall \mathbf{v}^h \in X_0^h, \end{aligned} \quad (27)$$

$$a_2(\phi^h, \psi^h) + b_2(\mathbf{u}^h, \psi^h) = \lambda(g^h, \psi^h)_\Gamma \quad \forall \psi^h \in \tilde{X}^h, \quad (28)$$

$$b_1(\mathbf{u}^h, q^h) = 0 \quad \forall q^h \in S_0^h, \quad (29)$$

$$\begin{aligned} a_1(\mathbf{w}^h, \nu^h) + \lambda c(\mathbf{w}^h, \mathbf{u}^h, \nu^h) + \lambda c(\mathbf{u}^h, \mathbf{w}^h, \nu^h) + b_1(\mathbf{w}^h, \tau^h) \\ + b_2(\mathbf{w}^h, s^h) = \lambda(\mathbf{u}^h - \mathbf{u}_0, \mathbf{w}^h) \quad \forall \mathbf{w}^h \in X_0^h, \end{aligned} \quad (30)$$

$$a_2(r^h, s^h) + b_2(\nu^h, r^h) = 0 \quad \forall r^h \in \tilde{X}^h, \quad (31)$$

$$b_1(\nu^h, \sigma^h) = 0 \quad \forall \sigma^h \in S_0^h, \quad (32)$$

$$\delta(g^h, z^h)_\Gamma - (s^h, z^h)_\Gamma = 0 \quad \forall z^h \in G^h. \quad (33)$$

The operator $T^h \in \mathcal{L}(Y, X^h)$ is defined as the solution operator for

$$\begin{aligned} a_1(\tilde{\mathbf{u}}^h, \mathbf{v}^h) + b_1(\mathbf{v}^h, \tilde{p}^h) + b_2(\mathbf{v}^h, \tilde{\phi}^h) \\ = (\zeta, \mathbf{v}^h) \quad \forall \mathbf{v}^h \in X_0^h, \end{aligned}$$

$$b_1(\tilde{\mathbf{u}}^h, q^h) = (\eta, q^h) \quad \forall q^h \in S_0^h,$$

$$a_2(\tilde{\phi}^h, \psi^h) + b_2(\tilde{\mathbf{u}}^h, \psi^h) = (\kappa, \psi^h)_\Gamma \quad \forall \psi^h \in \tilde{X}^h$$

i.e., for each $(\zeta, \eta, \kappa) \in Y$, $T^h(\zeta, \eta, \kappa) = (\tilde{\mathbf{u}}^h, \tilde{p}^h, \tilde{\phi}^h) \in X^h$ is the solution of the above system of equations. Since $T = T^*$, we define $(T^*)^h = T^h$.

We define the operator $E^h : G^* \rightarrow G^h$ as follows. For each $s \in G^*$, $g^h = E^h s$ if and only if

$$(g^h, z^h)_\Gamma = (s, z^h)_\Gamma \quad \forall z^h \in G^h.$$

The operator E^h is clearly well-defined and is, in fact, the projection operator from $G^* = L^2(\Gamma)$ onto G^h .

By results concerning the approximation of the linearized problem (see [7]), we obtain

$$\|(T - T^h)f\|_X \rightarrow 0$$

as $h \rightarrow 0$, for all $f = (\zeta, \eta, \kappa) \in Y$. This is simply a restatement of (H13). (H14) follows trivially from (H13), the fact that T is self-adjoint, and the choice $(T^*)^h = T^h$. It follows from interpolation results and the best approximation properties of the projection operator that

$$\|(E - E^h)s\|_{0,\Gamma} \rightarrow 0 \quad \text{as } h \rightarrow 0,$$

i.e., (H15) is verified. (H16) and (H17) follow from imbedding theorems, Cauchy inequalities and the fact that N and $\mathcal{F}$ are polynomials. We set $\tilde{Z} = L^{3/2}(\Omega) \times \{0\} \times \{0\}$. For each $(\mathbf{v}, q, \psi) \in X = H_0^1(\Omega) \times L_0^2(\Omega) \times \tilde{H}^1(\Omega)$ and $(\mathbf{w}, r, k) \in X = H_0^1(\Omega) \times L_0^2(\Omega) \times \tilde{H}^1(\Omega)$, we may use Sobolev imbedding theorems to verify (H18). From the definition of the operator K we see that K maps $L^2(\Gamma)$ into $L^{3/2}(\Omega) \times \{0\} \times L^2(\Gamma)$, i.e., K maps G into Z . Thus (H19) is verified.

Hence, we are now in a position to apply Theorem 4 to derive error estimates for the approximate solutions of the optimality system. (It should be noted that for almost all values of λ , the solutions of the optimality system are regular.)

Theorem 7 Assume that Λ is a compact interval of $\mathbb{R}_+$ and that there exists a sufficiently smooth branch $\{(\lambda, \mathbf{u}(\lambda), p(\lambda), \phi(\lambda), g(\lambda), \nu(\lambda), \tau(\lambda), s(\lambda)) : \lambda \in \Lambda\}$ of regular solutions of the optimality system (19)-(21) and (23)-(26). Assume that the finite element spaces X^h and G^h satisfy the approximation and stability hypotheses given at the beginning of this section. Then, there

- (H1) $\inf_{v \in X} \mathcal{F}(v) > -\infty$;
- (H2) there exist constants $\alpha, \beta > 0$ such that $\mathcal{E}(z) \geq \alpha \|z\|^\beta \quad \forall z \in \Theta$;
- (H3) there exists a $(v, z) \in X \times \Theta$ satisfying $M(v, z) = 0$;
- (H4) if $u^{(n)} \rightarrow u$ in X and $g^{(n)} \rightarrow g$ in G where $\{(u^{(n)}, g^{(n)})\} \subset X \times \Theta$, then $N(u^{(n)}) \rightarrow N(u)$ in Y and $K(g^{(n)}) \rightarrow K(g)$ in Y ;
- (H5) $\mathcal{J}(\cdot, \cdot)$ is weakly lower semicontinuous on $X \times \Theta$; and
- (H6) if $\{(u^{(n)}, g^{(n)})\} \subset X \times \Theta$ is such that $\{\mathcal{F}(u^{(n)})\}$ is a bounded set in $\mathbb{R}$ and $M(u^{(n)}, g^{(n)}) = 0$, then $\{u^{(n)}\}$ is a bounded set in X .
- (H7) for each $z \in \Theta$, $v \mapsto \mathcal{J}(v, z)$ and $v \mapsto M(v, z)$ are Fréchet differentiable;
- (H8) $z \mapsto \mathcal{E}(z)$ is convex;
- (H9) for $v \in X$, $N'(v)$ maps X into Z ;
- (H10) $\Theta = G$, and the mapping $z \mapsto \mathcal{E}(z)$ is Fréchet differentiable on G ;
- (H11) if $v^* \in X^*$ satisfies $(I + \lambda[N'(u)]^* T^*)v^* = 0$ and $K^* T^* v^* = 0$, then $v^* = 0$ or
- (H11)' the mapping $(v, z) \mapsto v + \lambda TN'(u)v + \lambda TKz$ is onto from $X \times G$ to Y ; and
- (H12) $\mathcal{E}'(g) = E^{-1}g$, where E is an invertible linear operator from G^* to G and g is an optimal control for the constrained minimization problem (4).

In (H9), N' denotes the Fréchet derivative of N and Z is a subspace of Y with a compact imbedding.

Assumptions (H1)-(H6) are used to establish the existence of optimal solutions; (H7)-(H9) will be needed to justify the use of the Lagrange multiplier rule and to derive an optimality system from which, for almost all $\lambda \in \Lambda$, optimal states and controls may be determined; (H10) will allow for a simplification of the optimality system; either or (H11) or (H11)' can be used to extend the validity of the optimality system to all values of $\lambda \in \Lambda$; finally, (H12) can be used to render the optimality system more amenable to approximation and computation.

2.1. Existence results and an optimality system

A fairly standard method of proof yields the existence of optimal solutions.

Theorem 1 Assume that the functional $\mathcal{J}$ and mapping M defined by (1) and (2), respectively, satisfy (H1)-(H6). Then, there exists a solution to the minimization problem (3).

Using a proof found in [8] of an abstract Lagrange multiplier rule, one can show that the Lagrange multiplier rule may be used to turn the constrained minimization problem (4) into an unconstrained one.

Theorem 2 Let $\lambda \in \Lambda$ be given. Assume that (H1)-(H10) hold. Let $(u, g) \in X \times G$ be a solution of the problem (4). Then, there exists a $k \in \mathbb{R}$ and a $\mu \in X^*$ that are not both equal to zero such that

$$k \langle \mathcal{J}_u(u, g), w \rangle - \langle \mu, (I + \lambda TN'(u))w \rangle = 0 \quad \forall w \in X \quad (5)$$

$$k \langle \mathcal{E}'(g), z \rangle - \langle \mu, TKz \rangle = 0 \quad \forall z \in G. \quad (6)$$

If $(1/\lambda) \notin \sigma(-TN'(u))$, i.e., if $(1/\lambda)$ is not in the spectrum of $-TN'(u)$, we may choose $k = 1$.

Actually, only (H1)-(H9) are needed to apply the result of [8] and show that the Lagrange multiplier rule may be used; (H10) simply simplifies the optimality system.

It can be shown that $TN'(u)$ is a compact operator from X to X , and therefore $\sigma(-TN'(u))$ is at most countable with zero being the only possible limit point. Thus, (5)-(6) are valid for almost all $\lambda \in \Lambda$. If the control g enters the constraint in a favorable manner, then we may take $k = 1$ even when $(1/\lambda) \in \sigma(-TN'(u))$. Specifically, we invoke one of the assumptions (H11) and (H11)'.

Theorem 3 Assume that the hypotheses of Theorem 2 hold. Assume that if $(1/\lambda) \in \sigma(-TN'(u))$, then either (H11) or (H11)' holds. Then, for all $\lambda \in \Lambda$, there exists a $\mu \in X^*$ such that (5) and (6) hold with $k = 1$.

Under the assumptions of Theorem 3, an optimal state $u \in X$, an optimal control $g \in G$, and the corresponding Lagrange multiplier $\mu \in X^*$ satisfy the optimality system of equations formed by (2), (5), and (6) with $k = 1$. From (1) we have that $\mathcal{J}_u = \lambda \mathcal{F}'$ and $\mathcal{J}_g = \lambda \mathcal{E}'$, where $\mathcal{F}'$ denotes the obvious Fréchet derivative. For purposes of numerical approximations, it turns out to be convenient to make the change of variable $\xi = T^* \mu$. It will also be convenient to invoke the additional simplifying assumption (H12) concerning the dependence of the objective functional on the control. Then, (2), (5), and (6) with $k = 1$ can be rewritten as

$$u + \lambda TN(u) + \lambda TKg = 0 \quad \text{in } X, \quad (7)$$

$$\xi + \lambda T^*[N'(u)]^* \xi - \lambda T^* \mathcal{F}'(u) = 0 \quad \text{in } Y^*, \quad (8)$$

$$g - EK^* \xi = 0 \quad \text{in } G. \quad (9)$$

Clearly, the control g may be eliminated from the optimality system (7)-(9). However, given a $\xi \in Y^*$, it is not always possible to evaluate g exactly from (9). For example, the application of the operator E may involve the solution of a partial differential equation. In many applications we have that $X^* = Y$. Since these spaces are assumed to be reflexive, we also have that $Y^* = X$. In this case, we have that both u and ξ belong to X .

3. Finite dimensional approximations

A finite dimensional discretization of the optimality system (7)-(9) is defined as follows. First, one chooses families of finite dimensional subspaces $X^h \subset X$, $(Y^*)^h \subset Y^*$, and $G^h \subset G$. These families are parameterized by a parameter h that tends to zero. (For example, this parameter can be chosen to be some measure of the grid size in a subdivision of Ω into finite elements.) Next, we define approximate operators $T^h : Y \rightarrow X^h$, $E^h : G^* \rightarrow G^h$, and $(T^*)^h : X^* \rightarrow (Y^*)^h$. Of course, one views T^h , E^h , and $(T^*)^h$ as approximations to the operators T , E , and T^* , respectively. Note that $(T^*)^h$ is not necessarily the same as $(T^h)^*$. The former is a discretization of an adjoint operator while the latter is the adjoint of a discrete operator.

Once the approximating subspaces and operators have been chosen, an approximate problem is defined as

exists an $h_0 > 0$ such that for $h \leq h_0$, the discrete optimality system (27)-(33) has a unique branch of solutions $\{(\lambda, u^h(\lambda), p^h(\lambda), \phi^h(\lambda), g^h(\lambda), \nu^h(\lambda), \tau^h(\lambda), s^h(\lambda)) : \lambda \in \Lambda\}$ satisfying, for sufficiently smooth solutions,

$$\begin{aligned} & \left(\|u(\lambda) - u^h(\lambda)\|_1 + \|p(\lambda) - p^h(\lambda)\|_0 + \|\phi(\lambda) - \phi^h(\lambda)\|_1 \right. \\ & \quad + \|g(\lambda) - g^h(\lambda)\|_{0,\Gamma} + \|\nu(\lambda) - \nu^h(\lambda)\|_1 \\ & \quad \left. + \|\tau(\lambda) - \tau^h(\lambda)\|_0 + \|s(\lambda) - s^h(\lambda)\|_1 \right) \\ & \leq Ch^m \left(\|u(\lambda)\|_{m+1} + \|p(\lambda)\|_m + \|\phi\|_{m+1} \right. \\ & \quad \left. + \|\nu(\lambda)\|_{m+1} + \|\tau(\lambda)\|_m + \|s\|_{m+1} \right), \end{aligned}$$

uniformly in h and $\lambda \in \Lambda$.

Theorem 7 applies in particular to the case where $G^h = X^h|_\Gamma$. In this case, g^h can be eliminated from the discrete system (27)-(33) and we subsequently recover an approximate control by the relation $g^h = -\frac{1}{\delta} s^h$.

4.3. A computational example

We conclude this section with a computational example. We consider the problem of minimizing (18) subject to (19)-(21), i.e., we attempt to match the velocity field with a desired one by finding an appropriate boundary current density g . We assume the physical domain is a unit square, $M = N = 1$, $\mathbf{B} = (0, 0, y)^T$, and we then choose the body force $\mathbf{f} = (f_1, f_2)^T$ and an electric source e so that the exact solution of the uncontrolled problem, i.e., the solution for (13)-(17) with $g = 0$ and a non-zero right hand side e (electric source) in (15), is given by

$$u_0 = \begin{pmatrix} [\cos(2\pi x) - 1] \sin(2\pi y) \\ \sin(2\pi x) [1 - \cos(2\pi y)] \end{pmatrix},$$

$p_0 = C$, where C is a constant, and $\phi_0 = \cos(\pi x) \cos(\pi y)$. The desired velocity field is chosen to be

$$u_d = \begin{pmatrix} \cos(2\pi y) [\cos(2\pi x) - 1] \\ \sin(2\pi x) \sin(2\pi y) \end{pmatrix}.$$

The proper choice of the constants ϵ and δ in the functional plays an important role in obtaining a best velocity matching. For the computational results shown below, our choice of these two constants were $\epsilon = 0.00001$ and $\delta = 0.01$.

In Figure 1 we give, respectively, the uncontrolled velocity field u_0 (under zero electrical current on the boundary) and the desired velocity field u_d . In Figure 2 we give the optimal control solutions obtained by solving (27)-(33) with g^h eliminated: velocity field u^h and the adjoint potential field s^h . The optimal control g^h can be gleaned from the adjoint potential field s^h in Figure 2 and the relation $g^h = -\frac{1}{\delta} s^h$. All the computational results shown in Figure 1 and Figure 2 were obtained with a 10 by 10 triangulation of the unit square. A nonuniform grid with corner refinements was used. We see from the figures that optimal control does a very good job in matching the desired velocity field which has double circulations.

5. References

- [1] F. Brezzi, J. Rappaz, and P.-A. Raviart, "Finite-dimensional approximation of nonlinear problems. Part I: Branches of nonsingular solutions," *Numer. Math.* **36** 1980, pp. 1-25.
- [2] M. Crouzeix and J. Rappaz, *On Numerical Approximation in Bifurcation Theory*, Masson, Paris, 1990.
- [3] V. Girault and P.-A. Raviart, *Finite Element Methods for Navier-Stokes Equations*, Springer, Berlin, 1986.
- [4] M. Gunzburger and L. Hou, "Finite dimensional approximation of a class of constrained nonlinear optimal control problems," to appear.
- [5] L.S. Hou and J. Peterson, "Boundary optimal control for an electrically conducting fluid using boundary electrical potential controls," *Nonlinear Analysis*, in press.
- [6] L.S. Hou and S.S. Ravindran, "Computations of boundary optimal control problems for an electrically conducting fluid," to appear.
- [7] J. Peterson, "On the finite element approximation of incompressible flows of an electrically conducting fluid," *Numer. Meth. P.D.E.s* **4**, 1988, 57-68.
- [8] V. Tichomorov, *Fundamental Principles of the Theory of Extremal Problems*, Wiley, Chichester, 1982.
- [9] J.S. Walker, "Large interaction parameter magnetohydrodynamics and applications in fusion reactor technology," in *Fluid Mechanics in Energy Conversion*, J. Buchmaster (Ed.), SIAM, Philadelphia, 1980.
- [10] N. Winowich and W. Hughes, "A finite element analysis of two dimensional MHD flows," in *Liquid-Metal Flows and Magnetohydrodynamics*, H. Branover, P.S. Lykoudis and A. Yakhot (Ed.), AIAA, New York, 1983.

Figure 1. uncontrolled velocity contours (left)
desired velocity contours (right)

Figure 2. optimal velocity contours (left)
optimal adjoint potential surface (right)

exists an $h_0 > 0$ such that for $h \leq h_0$, the discrete optimality system (27)-(33) has a unique branch of solutions $\{(\lambda, \mathbf{u}^h(\lambda), p^h(\lambda), \phi^h(\lambda), g^h(\lambda), \boldsymbol{\nu}^h(\lambda), \tau^h(\lambda), s^h(\lambda)) : \lambda \in \Lambda\}$ satisfying, for sufficiently smooth solutions,

$$\begin{aligned} & \left(\|\mathbf{u}(\lambda) - \mathbf{u}^h(\lambda)\|_1 + \|p(\lambda) - p^h(\lambda)\|_0 + \|\phi(\lambda) - \phi^h(\lambda)\|_1 \right. \\ & \quad + \|g(\lambda) - g^h(\lambda)\|_{0,\Gamma} + \|\boldsymbol{\nu}(\lambda) - \boldsymbol{\nu}^h(\lambda)\|_1 \\ & \quad \left. + \|\tau(\lambda) - \tau^h(\lambda)\|_0 + \|s(\lambda) - s^h(\lambda)\|_1 \right) \\ & \leq Ch^m \left(\|\mathbf{u}(\lambda)\|_{m+1} + \|p(\lambda)\|_m + \|\phi\|_{m+1} \right. \\ & \quad \left. + \|\boldsymbol{\nu}(\lambda)\|_{m+1} + \|\tau(\lambda)\|_m + \|s\|_{m+1} \right), \end{aligned}$$

uniformly in h and $\lambda \in \Lambda$.

Theorem 7 applies in particular to the case where $G^h = X^h|_\Gamma$. In this case, g^h can be eliminated from the discrete system (27)-(33) and we subsequently recover an approximate control by the relation $g^h = -\frac{1}{\delta} s^h$.

4.3. A computational example

We conclude this section with a computational example. We consider the problem of minimizing (18) subject to (19)-(21), i.e., we attempt to match the velocity field with a desired one by finding an appropriate boundary current density g . We assume the physical domain is a unit square, $M = N = 1$, $\mathbf{B} = (0, 0, y)^T$, and we then choose the body force $\mathbf{f} = (f_1, f_2)^T$ and an electric source e so that the exact solution of the uncontrolled problem, i.e., the solution for (13)-(17) with $g = 0$ and a non-zero right hand side e (electric source) in (15), is given by

$$\mathbf{u}_0 = \begin{pmatrix} [\cos(2\pi x) - 1] \sin(2\pi y) \\ \sin(2\pi x) [1 - \cos(2\pi y)] \end{pmatrix},$$

$p_0 = C$, where C is a constant, and $\phi_0 = \cos(\pi x) \cos(\pi y)$. The desired velocity field is chosen to be

$$\mathbf{u}_d = \begin{pmatrix} \cos(2\pi y) [\cos(2\pi x) - 1] \\ \sin(2\pi x) \sin(2\pi y) \end{pmatrix}.$$

The proper choice of the constants ϵ and δ in the functional plays an important role in obtaining a best velocity matching. For the computational results shown below, our choice of these two constants were $\epsilon = 0.00001$ and $\delta = 0.01$.

In Figure 1 we give, respectively, the uncontrolled velocity field $\mathbf{u}_0$ (under zero electrical current on the boundary) and the desired velocity field $\mathbf{u}_d$. In Figure 2 we give the optimal control solutions obtained by solving (27)-(33) with g^h eliminated: velocity field $\mathbf{u}^h$ and the adjoint potential field s^h . The optimal control g^h can be gleaned from the adjoint potential field s^h in Figure 2 and the relation $g^h = -\frac{1}{\delta} s^h$. All the computational results shown in Figure 1 and Figure 2 were obtained with a 10 by 10 triangulation of the unit square. A nonuniform grid with corner refinements was used. We see from the figures that optimal control does a very good job in matching the desired velocity field which has double circulations.

5. References

- [1] F. Brezzi, J. Rappaz, and P.-A. Raviart, "Finite-dimensional approximation of nonlinear problems. Part I: Branches of nonsingular solutions," *Numer. Math.* **36** 1980, pp. 1-25.
- [2] M. Crouzeix and J. Rappaz, *On Numerical Approximation in Bifurcation Theory*, Masson, Paris, 1990.
- [3] V. Girault and P.-A. Raviart, *Finite Element Methods for Navier-Stokes Equations*, Springer, Berlin, 1986.
- [4] M. Gunzburger and L. Hou, "Finite dimensional approximation of a class of constrained nonlinear optimal control problems," to appear.
- [5] L.S. Hou and J. Peterson, "Boundary optimal control for an electrically conducting fluid using boundary electrical potential controls," *Nonlinear Analysis*, in press.
- [6] L.S. Hou and S.S. Ravindran, "Computations of boundary optimal control problems for an electrically conducting fluid," to appear.
- [7] J. Peterson, "On the finite element approximation of incompressible flows of an electrically conducting fluid," *Numer. Meth. P.D.E.s* **4**, 1988, 57-68.
- [8] V. Tichomorov, *Fundamental Principles of the Theory of Extremal Problems*, Wiley, Chichester, 1982.
- [9] J.S. Walker, "Large interaction parameter magnetohydrodynamics and applications in fusion reactor technology," in *Fluid Mechanics in Energy Conversion*, J. Buchmaster (Ed.), SIAM, Philadelphia, 1980.
- [10] N. Winowich and W. Hughes, "A finite element analysis of two dimensional MHD flows," in *Liquid-Metal Flows and Magnetohydrodynamics*, H. Branover, P.S. Lykoudis and A. Yakhot (Ed.), AIAA, New York, 1983.

Figure 1. uncontrolled velocity contours (left)
desired velocity contours (right)

Figure 2. optimal velocity contours (left)
optimal adjoint potential surface (right)

A New Condition of the Exact Boundary Controllability for Parabolic Equations

Yanzhao Cao, Max Gunzburger and James Turner

April 26, 1997

1 Introduction

The purpose of this paper is to investigate the exact controllability problems for systems modeled by linear parabolic differential equations. This problem has been studied extensively in the last 30 years. A very significant early contribution is due to Egorov ([2], [3]). Another early paper is due to Gal'chuk ([9]). Many developments in the controllability theory of the linear parabolic equations and hyperbolic equations are due to Fattorini and Russell ([4], [5], [19], [20]). Using a harmonic analysis method, they obtained results such as approximate controllability and conditions on $\hat{y}$ for the exact controllability of systems governed by hyperbolic and parabolic equations. We refer to [7], [8] for more recent development by Furskov and Imanuvilov.

In this paper we attempt to study the exact controllability problem by examining the limit behavior of the corresponding optimal control problem. Using this method we obtained a new condition on the exact controllability of the system. We believe that this condition is at least comparable condition obtained in [18]. The paper is organized as follows. In §2, we state the problem and some results concerning linear parabolic differential equations. In §3 we give a constructive proof for the approximate controllability of the system using our method. Finally in §4 we give a condition on the exact controllability of the system.

2 Statement of the Problem, Notation and Preliminaries

Let Ω be a bounded domain in $\mathbb{R}^d$. For a fixed T , let $Q = [0, T] \times \Omega$ and $\Sigma = (0, T) \times \partial\Omega$. On the domain Ω , let A be the second order elliptic differential operator

$$Ay = - \sum_{i,j=1}^d \frac{\partial}{\partial x_i} (a_{i,j}(x) \frac{\partial y}{\partial x_j}) + c(x)y.$$

We will assume that $c > 0$ on $\bar{\Omega}$ and that the matrix $(a_{i,j}(x))$ is symmetric and positive definite.

Consider now the parabolic initial boundary value problem

$$\begin{aligned} \frac{\partial y}{\partial t} + Ay &= 0, & \text{in } Q, \\ \frac{\partial y}{\partial \nu} &= u, & \text{on } \Sigma, \\ y(0) &= y_0, & \text{in } \Omega \text{ for } t = 0. \end{aligned} \quad (2.1)$$

Here $\frac{\partial}{\partial \nu} = \sum_{i,j=1}^d a_{i,j} n_i (\frac{\partial}{\partial x_j})$ where n_i is the i -th component of the outward unit normal on $\partial\Omega$. We say that the system is exactly controllable for a given state $\hat{y}$ if there exists $u \in L^2(\Sigma)$ such that the solution y of problem (2.1) also satisfies

$$y(T, \cdot) = \hat{y}.$$

Define a cost functional

$$J_\epsilon(u) = \frac{1}{2} \int_{\Omega} \left(y(T, x; u) - \hat{y}(x) \right)^2 dx + \frac{\epsilon}{2} \int_{\Sigma} u^2(t, x) dx dt,$$

where $y = y(t, x; u)$ satisfies (2.1). $u_\epsilon(t, x)$ is said to be an optimal control if it satisfies

$$J_\epsilon(u_\epsilon) = \inf_{u \in L^2(Q)} J_\epsilon(u). \quad (2.2)$$

For a Banach space X , define

$$L^2(0, T; X) = \{f : (0, T) \rightarrow X; \int_0^T \|f(t)\|_X^2 dt < \infty\}.$$

It is a Hilbert space with the norm

$$\|f\|_{L^2(0, T; X)} = \left(\int_0^T \|f(t)\|_X^2 dt \right)^{\frac{1}{2}}.$$

Let $L^2(\Omega)$ and $L^2(\partial\Omega)$ be the usual L^2 function spaces. We shall use the notation

$$(\phi, \psi) = \int_{\Omega} \phi \psi dx$$

and

$$\langle \phi, \psi \rangle = \int_{\partial\Omega} \phi \psi d\omega$$

for the inner products and the associated norms will be denoted by $\|\cdot\|$ and $|\cdot|$, respectively.

We also define

$$W(0, T) = \{f; f \in L^2(0, T; H^1(\Omega)), \frac{df}{dt} \in L^2(0, T; H^{-1}(\Omega))\},$$

where $\frac{d}{dt}$ is taken in the sense of distributions on $(0, T)$ with values in $H^1(\Omega)$. $W(0, T)$ is a Hilbert space with the norm

$$\|f\|_{W(0, T)} = (\|f\|_{L^2(0, T; H^1(\Omega))}^2 + \|\frac{df}{dt}\|_{L^2(0, T; H^{-1}(\Omega))}^2)^{\frac{1}{2}}.$$

For $p, q > 0$, let also

$$H^{p, q}(Q) = L^2(0, T; H^q(\Omega)) \cap H^p(0, T; L^2(\Omega)).$$

These spaces are described in [15], and their norms are defined by

$$\|f\|_{H^{p, q}(Q)} = (\|f\|_{L^2(0, T; H^q(\Omega))}^2 + \|f\|_{H^p(0, T; L^2(\Omega))}^2)^{\frac{1}{2}}.$$

The spaces $H^{p, q}(\Sigma)$ with associated norms $\|\cdot\|_{H^{p, q}(\Sigma)}$ are defined similarly by replacing Ω by Σ above.

We recall that if $q > \frac{1}{2}$ and $p = q - \frac{1}{2}$, then there is a constant c such that for any $f \in H^{p, \frac{p}{2}}(Q)$,

$$\|f\|_{H^{p, \frac{p}{2}}(\Sigma)} \leq C \|f\|_{H^{q, \frac{q}{2}}(Q)},$$

and if $q > 1$, $g \in H^{q, \frac{q}{2}}(\Sigma)$ and $t \in [0, T]$ then

$$\|f(t)\|_{H^{p-1}(\Omega)} \leq c \|f\|_{H^{p, \frac{p}{2}}(Q)}$$

and

$$|g|_{H^{p-1}(\partial\Omega)} \leq c \|g\|_{H^{q, \frac{q}{2}}(\Sigma)}.$$

Define the bilinear form $a: H^1(\Omega) \times H^1(\Omega) \rightarrow R$ by

$$a(\phi, \psi) = \int_{\Omega} \left\{ \sum_{i, j=1}^d a_{ij}(x) \frac{\partial \phi}{\partial x_i} \frac{\partial \psi}{\partial x_j} + c(x) \phi \psi \right\} dx.$$

Note that from the properties of the operator A we have that there exist constants $c_1, c_2 > 0$ such that

$$|a(\phi, \psi)| \leq c_1 \|\phi\|_1 \|\psi\|_1$$

and

$$a(\phi, \phi) \geq c_2 \|\phi\|_1^2$$

for all $\phi, \psi \in H^1(\Omega)$.

We now define the operators for weak solutions of elliptic boundary value problems. For $f \in H^{-1}(\Omega)$, let $Tf \in H^1(\Omega)$ denote the unique solution of the problem

$$a(Tf, \phi) = (f, \phi) \quad \text{for } \phi \in H^1(\Omega).$$

T is a linear operator on $H^{-1}(\Omega)$. Also since

$$(Tf, \phi) = a(Tf, T\phi) = (f, T\phi)$$

for any $f, \phi \in L^2(\Omega)$, T is self adjoint on $L^2(\Omega)$. Next we introduce the eigenvalue problem

$$A\varphi = \Lambda\varphi, \quad \text{in } \Omega,$$

$$\frac{\partial \varphi}{\partial \nu} = 0, \quad \text{on } \partial\Omega.$$

It is well known that this problem has a system of eigenfunctions $\{\varphi_j\}_{j=1}^\infty$, forming a complete orthonormal set in $L^2(\Omega)$, with eigenvalues $0 < \lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_j \rightarrow \infty$ as $j \rightarrow \infty$. Note that if we let $\nu_j = \lambda_j^{-1}$, then

$$T\phi_j = \nu_j \phi_j \quad j = 1, 2, \dots$$

We now consider the weak formulation of problem (2.1)

$$\begin{aligned} \left(\frac{dy}{dt}, \phi\right) + a(y, \phi) &= \langle u, \phi \rangle, \quad \text{for } \phi \in H^1(\Omega), \\ y(0, \cdot) &= y_0. \end{aligned} \tag{2.3}$$

It is well known that if $y_0 \in L^2(\Omega)$ and $u \in L^2(\Sigma)$ then (2.3) has a unique solution y in $W(0, T)$, and there is a constant C , independent of y_0 and u , such that

$$\|y\|_{W(0, T)} \leq C(\|v\| + \|u\|_{L^2(\Sigma)}).$$

If $u = 0$, then the exact solution of (2.3) can be represented by

$$y(t, x) = \sum_{j=1}^{\infty} (v, \varphi_j) e^{-\lambda_j t} \varphi_j(x)$$

The following lemma can be found in [21].

Lemma 2.1 *Let $\phi \in H^p(\Omega)$ for $-1 \leq p \leq 1$. Then*

$$\|\phi\| = \left(\sum_{j=1}^{\infty} (\phi, \varphi_j)^2 \lambda_j^p \right)^{\frac{1}{2}}$$

is an equivalent norm in $H^p(\Omega)$.

3 Approximate Controllability

3.1 An operator representation of the terminal state. In the following lemma we state the optimality system for the problem (2.2). It will play a central role in obtaining the main results of this section and the next section. From now on, we suppress the explicit dependence on x , e.g., we will use $u(t)$ to mean $u(t, x)$.

Lemma 3.1 ([12]) *The optimality system of problem (2.1) is given by*

$$\begin{aligned} \frac{\partial p_\epsilon}{\partial t} - Ap_\epsilon &= 0 && \text{in } Q, \\ \frac{\partial p_\epsilon}{\partial \nu} &= 0 && \text{on } \Sigma, \\ p_\epsilon &= y_\epsilon - \hat{y} && \text{in } \Omega \text{ for } t = T, \\ p_\epsilon &= -\epsilon u_\epsilon && \text{on } \Sigma. \end{aligned} \tag{3.1}$$

The weak formulation of the optimality system can be written as follows.

$$\begin{aligned} \left(\frac{dy_\epsilon}{dt}, \phi\right) + a(y_\epsilon, \phi) &= \frac{1}{\epsilon} \langle p_\epsilon, \phi \rangle, \quad \text{for } \phi \in H^1(\Omega), \\ -\left(\frac{dp_\epsilon}{dt}, \phi\right) + a(p_\epsilon, \phi) &= 0, \quad \text{for } \phi \in H^1(\Omega), \\ y_\epsilon(0) &= y_0, \\ p_\epsilon(T) &= \hat{y} - y_\epsilon(T). \end{aligned} \tag{3.2}$$

To represent the terminal state $y_\epsilon(T)$ by an operator, we consider, for a given $z \in L^2(\Omega)$, the system

$$\begin{aligned} \left(\frac{dy_1}{dt}, \phi\right) + a(y_1, \phi) &= \langle p_1, \phi \rangle, \quad \text{for } \phi \in H^1(\Omega), \\ -\left(\frac{dp_1}{dt}, \phi\right) + a(p_1, \phi) &= 0, \quad \text{for } \phi \in H^1(\Omega), \\ y_1(0) &= 0, \\ p_1(T) &= z. \end{aligned} \tag{3.3}$$

This system has a unique solution $(y_1(t), p_1(t)) \in W(0, T)$.

Let $E(t)$ denote the solution operator for the problem

$$\begin{aligned} \frac{\partial w}{\partial t} + Aw &= 0 && \text{on } Q, \\ \frac{\partial w}{\partial \nu} &= 0 && \text{on } \Sigma, \\ w &= z && \text{in } \Omega \text{ for } t = 0, \end{aligned}$$

i.e.,

$$w(t) = E(t)z.$$

We now define an operator $R: L^2(\Omega) \rightarrow L^2(\Omega)$ by

$$Rz = y_1(T). \quad (3.4)$$

Comparing (3.2) with (3.3), we have that

$$\frac{1}{\epsilon} R(y(T) - \hat{y}) = -E(T)v + y(T) \quad (3.5)$$

or

$$(\epsilon + R)(y(T) - \hat{y}) = \epsilon(\hat{y} - E(T)v) \quad (3.6)$$

3.2. Properties of the operator R .

Lemma 3.2 $R \in L(L^2(\Omega), H^1(\Omega))$. Therefore R is a compact operator in $L^2(\Omega)$.

Proof: Consider the mapping

$$z \rightarrow p_1(t, \cdot) = E(T-t)z.$$

We prove that this mapping is continuous from $L^2(\Omega)$ into $H^{\frac{1}{2},1}(Q)$. By standard theory for parabolic differential equations ([15]) this mapping is continuous from $H^1(\Omega)$ into $H^{1,2}(Q)$. If $z \in H^{-1}(\Omega)$ then

$$\begin{aligned} \int_0^T \|p_1(t, \cdot)\|^2 dt &= \int_0^T \sum_{j=1}^{\infty} (z, \phi_j)^2 e^{-2\lambda_j(T-t)} dt \\ &= \frac{1}{2} \sum_{j=1}^{\infty} (z, \phi_j)^2 \lambda_j^{-1} (1 - e^{-2\lambda_j T}) \\ &\leq C \|z\|_{-1}. \end{aligned}$$

Hence the mapping is continuous from $H^{-1}(\Omega)$ into $L^2(Q)$. By interpolation ([1]) we have that R is a continuous mapping from $L^2(\Omega)$ to $H^{\frac{1}{2},1}(Q)$. By (3.3), we have that $\frac{\partial y_1}{\partial \nu}|_{\Sigma} = p_1|_{\Sigma}$. Thus

$$\|Rz\|_1 = \|y_1(T)\|_1 \leq C \|y_1\|_{H^{1,2}(Q)} \leq C \|z\|.$$

The proof is complete.

Lemma 3.3 R is symmetric and semi-definite.

Proof: For given $z, z_0 \in L^2(\Omega)$, let $(y, p), (y_0, p_0)$ be the corresponding solutions of (3.3), respectively. Then

$$\left(\frac{dy}{ds}, p_0\right) + a(y, p_0) = \langle p, p_0 \rangle,$$

$$(\frac{dy_0}{ds}, p) + a(y_0, p) = \langle p_0, p \rangle,$$

$$a(y, p_0) = (y, \frac{d}{ds} p_0),$$

$$a(y_0, p) = (y_0, \frac{d}{ds} p).$$

We therefore have that

$$(\frac{d}{ds} y(s), p_0(s)) + (y(s), \frac{d}{ds} p_0(s)) = (\frac{d}{ds} y_0(s), p(s)) + (y_0(s), \frac{d}{ds} p(s))$$

or

$$\frac{d}{ds} (y_0(s), p(s)) = \frac{d}{ds} (y(s), p_0(s)).$$

Integrating the above equality from 0 to T , noting that $y(0) = y_0(0) = 0$, we have that

$$\begin{aligned} (y(T), p_0(T)) &= \int_0^T \frac{d}{ds} (y(s), p_0(s)) ds \\ &= \int_0^T \frac{d}{ds} (y_0(s), p(s)) ds \\ &= (y_0(T), p(T)) \end{aligned}$$

or

$$(Rz, z_0) = (Rz_0, z).$$

Thus R is self-adjoint. We also have that

$$\begin{aligned} |p|^2 &= (\frac{d}{ds} y, p) + a(y, p) \\ &= (\frac{d}{ds} y, p) + (y, \frac{d}{ds} p) \\ &= \frac{d}{ds} (y, p). \end{aligned}$$

Integrating the above equality, we have that

$$(Rz, z) = \int_0^T |p|^2 ds \geq 0.$$

So R is semi-definite. The proof is complete.

Lemma 3.4 *Assume that Σ is analytic. Then $\text{Ker} R = 0$.*

Proof: Under the given assumption we know that p is analytic in Q and on Σ . Assume that $Rz = 0$. Then from the proof of Lemma 3.3 we have that

$$0 = (Rz, z) = \int_0^T |p|^2 ds.$$

Thus $p = 0$ on Σ . Hence the data corresponding to the Cauchy problem is zero on Σ which implies according to the Cauchy-Kowaleska theorem (since p is analytic) that $p = 0$ on Q . In particular $z = p(T) = 0$. The proof is complete.

3.3 Approximate Controllability

Theorem 3.5 *Let $\hat{y} \in L^2(\Omega)$ and $(y_\epsilon, u_\epsilon) = (y_\epsilon(t), u_\epsilon(t))$ be the solution of the optimality system. Then*

$$\lim_{\epsilon \rightarrow 0} \|y_\epsilon(T) - \hat{y}\|_{L^2(\Omega)} = 0.$$

Proof: By Lemma 3.2, Lemma 3.3 and Lemma 3.4, and the Gilbert-Schmidt theorem, the operator R has a system of eigenfunctions $\{e_j\}_{j=1}^\infty$ with eigenvalues $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_j \rightarrow 0$ as $j \rightarrow \infty$. Moreover, $\{e_j\}_{j=1}^\infty$ forms an orthonormal basis in $L^2(\Omega)$. For $\hat{y} \in L^2(\Omega)$ let

$$\hat{y} - E(T)v = \sum_{j=1}^{\infty} y_j e_j.$$

Then

$$R(\hat{y} - E(T)v) = \sum_{j=1}^{\infty} \lambda_j y_j e_j.$$

Using (3.6) we have that

$$y_\epsilon(T) - \hat{y} = \sum_{j=1}^{\infty} \frac{\epsilon y_j e_j}{\epsilon + \lambda_j}.$$

Thus

$$\begin{aligned} \|y_\epsilon(T) - \hat{y}\|_{L^2(\Omega)}^2 &= \sum_{j=1}^{\infty} \frac{\epsilon^2 y_j^2}{(\epsilon + \lambda_j)^2} \\ &\leq \sum_{j=1}^N \frac{\epsilon^2 y_j^2}{(\epsilon + \lambda_j)^2} + \sum_{j=N+1}^{\infty} y_j^2. \end{aligned}$$

For any given $\delta > 0$, there exists N such that

$$\sum_{j=N+1}^{\infty} y_j^2 < \frac{\delta}{2}.$$

For this fixed N there exists $\beta > 0$ such that

$$\sum_{j=1}^N \frac{\epsilon^2 y_j^2}{(\epsilon + \lambda_j)^2} < \frac{\delta}{2}$$

for $0 < \epsilon < \beta$. Thus

$$\|y_\epsilon(T) - \hat{y}\|_{L^2(\Omega)}^2 < \delta$$

for $0 < \epsilon < \beta$. The proof is complete.

Definition System (2.1) is said to be approximately controllable if, for any given $\hat{y}$ and $\delta > 0$, there exists a control $u(t)$ and a function $y(t)$ such that $(y(t), u(t))$ is a solution of (2.1) and

$$\|y(T) - \hat{y}\| < \delta.$$

The following theorem is an immediate consequence of Theorem 3.5.

Theorem 3.6 *System (2.1) is approximately controllable.*

Remark Approximate controllability is well known for linear parabolic differential equations. Our proof is different in that it is a constructive proof.

4 A Condition for Exact Controllability

In this section we give a condition on $\hat{y}$ for the exact controllability of system (2.1). The main tool is the operator R introduced in §2.2.

Theorem 4.1 *Let R be the operator defined in §2.2 and $(\{e_j\}_{j=1}^\infty, \{\lambda_j\}_{j=1}^\infty)$ be the eigensystem of R . Assume that R^{-1} is the inverse of operator R . If*

$$\hat{y} - E(T)y_0 \in D(R^{-1})$$

then system (2.1) is exactly controllable.

Proof: Assume that

$$\hat{y} - E(T)v = \sum_{j=1}^{\infty} y_j e_j.$$

Then

$$\sum_{j=1}^{\infty} \frac{y_j^2}{\lambda_j^2} < \infty. \quad (4.1)$$

By the proof of Theorem 2.6 we have that

$$\|y_\epsilon(T) - \hat{y}\|_{L^2(\Omega)}^2 = \sum_{j=1}^{\infty} \frac{\epsilon^2 y_j^2}{(\epsilon + \lambda_j)^2}.$$

Hence

$$\begin{aligned}\|y_\epsilon(T) - \hat{y}\|_{L^2(\Omega)} &= \left(\sum_{j=1}^{\infty} \frac{\epsilon^2 y_j^2}{(\epsilon + \lambda_j)^2} \right)^{\frac{1}{2}} \\ &\leq \epsilon \left(\sum_{j=1}^{\infty} \frac{y_j^2}{\lambda_j^2} \right)^{\frac{1}{2}} \leq C\epsilon.\end{aligned}$$

Let $\{\epsilon_k\}_{k=1}^{\infty}$ be a sequence such that $\epsilon_k \rightarrow 0$ as $k \rightarrow \infty$. From the optimality system (3.1) and the proof of Lemma 2.3 we have that

$$\begin{aligned}\|y_{\epsilon_k}(T)\|_{H^{1,2}(Q)} &\leq C(\|u_{\epsilon_k}\|_{H^{\frac{1}{4},\frac{1}{2}}(\Sigma)} + \|y_0\|_{H^1(\Omega)}) \\ &= C(\|\frac{p_{\epsilon_k}}{\epsilon_k}\|_{H^{\frac{1}{4},\frac{1}{2}}(\Sigma)} + \|y_0\|_{H^1(\Omega)}) \\ &\leq C(\|\frac{p_{\epsilon_k}}{\epsilon_k}\|_{H^{\frac{1}{2},1}(Q)} + \|y_0\|_{H^1(\Omega)}) \\ &\leq C(\|\frac{y_{\epsilon_k}(T) - \hat{y}}{\epsilon_k}\|_{L^2(\Omega)} + \|y_0\|_{H^1(\Omega)}) \\ &\leq C_1\|y_0\|_{H^1(\Omega)} + C_2.\end{aligned}$$

Thus there exists a subsequence of $\{\epsilon_k\}$, still denoted as $\{\epsilon_k\}$, and $y \in H^{1,2}(Q)$ such that

$$\lim_{k \rightarrow \infty} y_{\epsilon_k} = y \quad \text{weakly in } H^{1,2}(Q).$$

Using the trace theorem ([1]) and the compact embedding theorem ([1]) in Sobolev spaces we also have that there exists $u \in L^2(\Sigma)$ such that

$$\lim_{k \rightarrow \infty} u_{\epsilon_k} = u \quad \text{in } L^2(\Sigma).$$

Also by the trace theorem ([1]) we have that $y(T) = \hat{y}$ and $y(0) = y_0$. We now prove that (y, u) is a solution of (2.1). Since $(y_{\epsilon_k}(t), u_{\epsilon_k}(t))$ is a solution of (2.2)

$$\left(\frac{dy_{\epsilon_k}}{dt}, \phi\right) + a(y_{\epsilon_k}, \phi) = \langle u_{\epsilon_k}, \phi \rangle, \quad \text{for } \phi \in H^1(\Omega).$$

Passing to the limit in the above equality we obtain

$$\left(\frac{dy}{dt}, \phi\right) + a(y, \phi) = \langle u, \phi \rangle, \quad \text{for } \phi \in H^1(\Omega).$$

Thus $(y, u) \in H^{1,2}(Q) \times L^2(\Sigma)$ is a solution of (2.1). This proves that u is an exact control. The proof is complete.

Remark 4.1 From the proof of Theorem 4.1 we see that if (4.1) holds, then

$$\|y_\epsilon(T) - \hat{y}\|_{L^2(\Omega)} \leq C\epsilon.$$

It is easy to show that (4.1) is actually a necessary condition for the above inequality to hold. In fact, assume that the above inequality holds. Then by the proof of Theorem 2.8 we have that

$$\|y(\epsilon, T) - \hat{y}\|_{L^2(\Omega)}^2 = \sum_{j=1}^{\infty} \frac{\epsilon^2 y_j^2}{(\epsilon + \lambda_j)^2} \leq C\epsilon^2.$$

Thus

$$\sum_{j=1}^{\infty} \frac{y_j^2}{(\epsilon + \lambda_j)^2} \leq C.$$

Letting $\epsilon \rightarrow 0$, we obtain (4.1).

Remark 4.2. In [18], Russell gave a condition on exact controllability of system (2.1) using the harmonic analysis method. The condition is

$$\hat{y} - E(T)v \in e^{C\sqrt{A}}$$

where $C > 0$ is a constant and A is the elliptic operator that appears in (2.1). For $d = 1$, this condition implies that $\hat{y} - E(T)v$ is at least a C^∞ function. At this moment it is not clear if our condition is weaker. However, if the control is distribution we can show that

$$D(R^{-1}) = D(A) = H^2(\Omega)$$

which is much weaker than $D(e^{C\sqrt{A}})$.

REFERENCES

1. R. Adams, *Sobolev Spaces*, Academic, New York, 1975.
2. Yu. V. Egorov, Some problems in the theory of optimal control, *Soviet Math.* **3** 1962, 1080-1084.
3. Yu. V. Egorov, *Ž. Vyčisl. Mat. i Mat. Fiz.* **5** 1963, 887-904
4. H. Fattorini, Control in finite time of differential equations in Banach space, *Comm. Pure Appl. Math.*, **19** 1996, 7-34.
5. H. Fattorini and D. Russell, Exact controllability theorems for linear parabolic equations in one space dimension, *Arch. Rational Mech. Anal.* **49** 1971, 272-292.

6. A. Friedman, *Partial Differential Equations*, Holt, Rinehart, and Winston, New York, 1960.
7. A. Fursikov and O. Imanuvilov, On exact boundary zero controllability of two-dimensional Navier-Stokes equations, *Acta Appl. Math.* **36** 1994, 1-10.
8. A. Fursikov and O. Imanuvilov, On controllability of certain systems stimulating a fluid flow, in: *Flow Control*, IMA Volumes in Mathematics and Its Applications **68**, Springer, New York, 1995, 149-184.
9. L. Gal'chuk, Optimal control of systems described parabolic equations, *SIAM. J. Control* **7** 1969, 546-558.
10. R. Glowinski, C. H. Lee and J. -L. Lions, A numerical approach to the exact boundary controllability of the wave equation (I) Dirichlet controls: Description of the numerical methods, *Japan J. Appl. Math.* **7** 1990, 1-76.
11. M. Gunzburger, L. Hou, and T. Svobodny, Analysis and finite element approximation of optimal control problems for the stationary Navier-Stokes equations with distributed and Neumann controls, *Math. Comput.* **57** 1991, 123-151.
12. J.-L. Lions, *Optimal Control of Systems Governed by Partial Differential Equations*, Springer, Berlin, 1971.
13. J.-L. Lions, *Some Methods in the Mathematical Analysis of Systems and Their Controls*, Gordon and Breach, New York, 1981.
14. J.-L. Lions, *Some Aspect of the Optimal Control of Distributed Parameter Systems*, SIAM, Philadelphia, 1972.
15. J.-L. Lions and E. Magenes, *Nonhomogeneous Boundary Value Problems and Applications*, Springer, Berlin, 1972.
16. J.-L. Lions, Exact controllability, stabilization and perturbations for distributed systems, *SIAM Review* **30** 1988, 1-68.
17. J. Oden and L. Demkowicz, *Applied Functional Analysis*, CRC Press, Boca Raton-New York-London-Tokyo, 1996.

18. D. Russell, A unified boundary controllability theory for hyperbolic and parabolic partial differential equations, *Studies in Appl. Math.* **52** 1973, 189-211.
19. D. Russell, Controllability and stabilizability theory for linear partial differential equations: recent progress and open questions, *SIAM Review* **20** 1978, 639-739.
20. D. Russell, Some remarks on numerical aspects of coefficient identification in elliptic systems, in: *Optimal Control of Partial Differential Equations*, Birkhauser Verlag- Boston 1984, 210-228.
21. R. Winther, Error estimates for a Galerkin approximation of a parabolic control problem, *Ann. Mat. Pura. Appl.* **4** 1978, 173-206.