

# A SEMI-SMOOTH NEWTON METHOD FOR OPTIMAL BOUNDARY CONTROL OF A NONLINEAR REACTION-DIFFUSION SYSTEM

MSC-267

SIXTEENTH INTERNATIONAL SYMPOSIUM ON MATHEMATICAL THEORY OF NETWORKS  
AND SYSTEMS (MTNS2004), KATHOLIEKE UNIVERSITEIT LEUVEN, BELGIUM, JULY 5-9, 2004

R. GRIESSE AND S. VOLKWEIN

ABSTRACT. This paper is concerned with optimal boundary control of an in-stationary reaction-diffusion system in three spatial dimensions. This problem involves a coupled nonlinear system of parabolic differential equations with bilateral as well as integral control constraints. The integral constraint is included in the cost by a penalty term whereas the bilateral control constraints are handled explicitly. First-order conditions for the optimization problem are given. A semi-smooth Newton method is utilized to compute optimal solutions numerically.

## 1. Introduction

The subject matter of the present paper is an optimal control problem for a coupled system of semi-linear parabolic reaction-diffusion equations. The equations model a chemical or biological process where the species involved are subject to diffusion and reaction among each other. As an example, we consider the reaction  $A + B \rightarrow C$  which obeys the law of mass action. To simplify the discussion, we assume that the backward reaction  $C \rightarrow A + B$  is negligible and that the forward reaction proceeds with a constant (e.g., not temperature-dependent) rate. This leads to a coupled semilinear parabolic system for the respective concentrations; see (2.2) below.

The control function acts through the Neumann boundary values for one of the reaction components on some subset of the two-dimensional boundary manifold. It is natural to impose bilateral pointwise bounds on the control function: On the one hand, the substance can never be extracted through the boundary, i.e., the lower control bound should be nonnegative. On the other hand, only a limited amount may be added *at any given time*. In addition, we impose a constraint on the *total amount* of control action. This scalar integral constraint (see (2.4)) is very much in contrast with the usual pointwise bounds.

---

*Date:* September 14, 2005.

*1991 Mathematics Subject Classification.* 35Kxx, 49Lxx, 65Kxx.

*Key words and phrases.* Reaction-diffusion equations, optimal boundary control, bilateral and integral control constraints, semi-smooth Newton methods.

This work was supported in part by the *Fonds zur Förderung der wissenschaftlichen Forschung under the Special Research Center F003 "Optimization and Control"*.

The integral constraint is included into the cost functional by a penalty term, whereas the bilateral control constraints are treated explicitly by a primal-dual active set strategy for nonlinear problems.

Semi-smooth Newton methods have been studied by several authors. Utilizing the close relationship between the primal-dual active set strategy and semi-smooth Newton methods, local superlinear convergence of the former was shown in [7]. Semi-smooth Newton methods for general purpose nonlinear finite-dimensional optimal control problems are well studied, see, for instance, [8] and [2, Section 7.5]. Much less is known about such methods in infinite dimensions, and specifically in the context of optimal control problems. Let us refer here, e.g., to [5, 6, 9, 10].

The article is organized in the following manner: In Section 2, the optimal control problem is formulated and optimality conditions are presented. The semi-smooth Newton method is proposed in Section 3 and in the last section a numerical example is presented.

## 2. The optimal control problem

In this section, we introduce an optimal control problem for a reaction-diffusion system and review first-order necessary conditions. Let  $\Omega$  denote an open and bounded subset of  $\mathbb{R}^3$  with Lipschitz-continuous boundary  $\Gamma = \partial\Omega$  such that  $\Gamma$  is decomposed into two parts  $\Gamma = \Gamma_n \cup \Gamma_c$  with  $\Gamma_n \cap \Gamma_c = \emptyset$ . For terminal time  $T > 0$  let  $Q = (0, T) \times \Omega$ ,  $\Sigma = (0, T) \times \Gamma$  and  $\Sigma_c = (0, T) \times \Gamma_c$ .

By  $L^2(0, T; H^1(\Omega))$  we denote the space of all measurable functions  $\varphi : [0, T] \rightarrow H^1(\Omega)$ , which are square integrable, i.e.,

$$\int_0^T \|\varphi(t)\|_{H^1(\Omega)}^2 dt < \infty,$$

where  $\varphi(t)$  stands for the function  $\varphi(t, \cdot)$  considered as a function in  $\Omega$  only. The space  $W(0, T)$  is defined by

$$(2.1) \quad W(0, T) = \{\varphi \in L^2(0, T; H^1(\Omega)) : \varphi_t \in L^2(0, T; H^1(\Omega)')\}.$$

Here  $H^1(\Omega)'$  denotes the dual space of  $H^1(\Omega)$ . Recall that  $W(0, T)$  is a Hilbert space endowed with the common inner product and the induced norm; see, e.g., [1, pp. 286].

Suppose that  $d_1, d_2, d_3$  and  $k_1, k_2, k_3$  are positive constants. Moreover, let  $\alpha \in L^\infty(0, T; L^2(\Gamma_c))$  denote a shape function with  $\alpha \geq 0$  on  $\Sigma_c$  almost everywhere (a.e.). We consider the following system of semi-linear parabolic equations, where  $c_i$  denotes the concentration of the  $i$ -th substance:

$$(2.2a) \quad (c_1)_t(t, x) = d_1 \Delta c_1(t, x) - k_1 c_1(t, x) c_2(t, x) \quad \text{for all } (t, x) \in Q,$$

$$(2.2b) \quad (c_2)_t(t, x) = d_2 \Delta c_2(t, x) - k_2 c_1(t, x) c_2(t, x) \quad \text{for all } (t, x) \in Q,$$

$$(2.2c) \quad (c_3)_t(t, x) = d_3 \Delta c_3(t, x) + k_3 c_1(t, x) c_2(t, x) \quad \text{for all } (t, x) \in Q$$

together with the Neumann boundary conditions

$$(2.2d) \quad d_1 \frac{\partial c_1}{\partial n}(t, x) = 0 \quad \text{for all } (t, x) \in \Sigma,$$

$$(2.2e) \quad d_2 \frac{\partial c_2}{\partial n}(t, x) = u(t)\alpha(t, x) \quad \text{for all } (t, x) \in \Sigma_c,$$

$$(2.2f) \quad d_2 \frac{\partial c_2}{\partial n}(t, x) = 0 \quad \text{for all } (t, x) \in \Sigma_n = \Sigma \setminus \Sigma_c,$$

$$(2.2g) \quad d_3 \frac{\partial c_3}{\partial n}(t, x) = 0 \quad \text{for all } (t, x) \in \Sigma$$

and the initial conditions

$$(2.2h) \quad c_1(0, x) = c_{10}(x) \quad \text{for all } x \in \Omega,$$

$$(2.2i) \quad c_2(0, x) = c_{20}(x) \quad \text{for all } x \in \Omega,$$

$$(2.2j) \quad c_3(0, x) = c_{30}(x) \quad \text{for all } x \in \Omega,$$

where  $c_{i0} \in L^2(\Omega)$  for  $i = 1, 2, 3$ .

The control  $u \in L^2(0, T)$  enters the right-hand side of (2.2e) in the inhomogeneous Neumann condition. For instance, the function  $\alpha$  models a spray nozzle moving over the control part  $\Gamma_c$ , and  $u(t)$  denotes the intensity of the spray.

**Remark 2.1.** The parabolic problem for  $c_3$ , i.e., (2.2c) together with the Neumann boundary condition (2.2g) and initial condition (2.2j) can be solved independently of the problem for  $(c_1, c_2)$ . Therefore, we will focus on the computation of  $c_1$  and  $c_2$  and, in particular, we are interested in weak solutions for  $c_1$  and  $c_2$ .  $\diamond$

**Definition 2.2.** The two functions  $c_1$  and  $c_2$  in  $W(0, T)$  are called weak solutions to the system (2.2a), (2.2b), (2.2d)–(2.2f), (2.2h) and (2.2i) provided the initial conditions

$$(2.3a) \quad c_1(0) = c_{10} \quad \text{and} \quad c_2(0) = c_{20} \quad \text{in } L^2(\Omega)$$

hold and

$$(2.3b) \quad \langle (c_1)_t(t), \varphi \rangle_{H^1(\Omega)', H^1(\Omega)} + \int_{\Omega} d_1 \nabla c_1(t) \cdot \nabla \varphi + k_1 c_1(t) c_2(t) \varphi \, dx = 0,$$

$$(2.3c) \quad \begin{aligned} & \langle (c_2)_t(t), \varphi \rangle_{H^1(\Omega)', H^1(\Omega)} + \int_{\Omega} d_2 \nabla c_2(t) \cdot \nabla \varphi + k_2 c_1(t) c_2(t) \varphi \, dx \\ & = u(t) \int_{\Gamma_c} \alpha(t) \varphi \, dx \end{aligned}$$

for all  $\varphi \in H^1(\Omega)$  and almost all  $t \in [0, T]$ . In (2.3b) and (2.3c),  $\langle \cdot, \cdot \rangle_{H^1(\Omega)', H^1(\Omega)}$  denotes the duality pairing between  $H^1(\Omega)$  and its dual  $H^1(\Omega)'$ .

The following theorem ensures that (2.3) possesses a unique solution. For a proof, which is based on the Leray-Schauder's fixed point theorem and variational techniques, we refer to [3, Theorem 2.3].

**Theorem 2.3.** For every control  $u \in L^2(0, T)$ , there exists a unique pair  $(c_1, c_2) \in W(0, T) \times W(0, T)$  satisfying (2.3).

**Remark 2.4.** Theorem 2.3 also implies the unique solvability of the partial differential equation for the reaction product (2.2c), (2.2g), and (2.2j), see, for more details, [4, Corollary 2.4].  $\diamond$

Our goal is to drive the reaction-diffusion system from the given initial state near a desired terminal state. Since only a limited amount may be added at any given time, we have the following integral constraint:

$$(2.4) \quad \int_0^T u \, dt \leq u_c$$

with  $u_c > 0$ . To handle (2.4) numerically we introduce the penalized cost functional

$$J_\varepsilon(c_1, c_2, u) = \frac{1}{2} \int_\Omega \beta_1 |c_1(T) - c_{1T}|^2 + \beta_2 |c_2(T) - c_{2T}|^2 \, dx + \frac{\gamma}{2} \int_0^T |u - u_d|^2 \, dt \\ + \frac{1}{\varepsilon} \max \left\{ 0, \int_0^T u \, dt - u_c \right\}^3,$$

where  $\beta_1, \beta_2 \geq 0$ ,  $\beta_1 + \beta_2, \gamma > 0$ ,  $c_{1T}, c_{2T} \in L^2(\Omega)$  are given desired terminal states and  $u_d \in L^2(0, T)$  denotes some nominal (or expected) control.

The closed and bounded convex set of admissible control parameters involves bilateral control constraints:

$$U_{\text{ad}} = \left\{ u \in L^2(0, T) : u_a \leq u \leq u_b \text{ in } [0, T] \right\} \subset L^\infty(0, T),$$

where  $u_a$  and  $u_b$  are given functions in  $L^\infty(0, T)$  satisfying  $u_a \leq u_b$  in  $[0, T]$  almost everywhere (a.e.).

$$(2.5) \quad \min J_\varepsilon(x) \quad \text{s.t.} \quad x = (c_1, c_2, u) \in W(0, T) \times W(0, T) \times U_{\text{ad}} \text{ solves (2.3).}$$

It was proved in [4, Theorem 2.8] that (2.5) has at least one optimal solution  $\hat{x}_\varepsilon$ .

Problem (2.5) is a non-convex programming problem so that different local minima might occur. Numerical methods will produce a local minimum close to their starting point. Therefore, we do not direct our investigation to global solutions of (2.5). We will assume that a fixed reference solution  $\hat{x}^\varepsilon = (\hat{c}_1^\varepsilon, \hat{c}_2^\varepsilon, \hat{u}^\varepsilon) \in W(0, T) \times W(0, T) \times U_{\text{ad}}$  is given satisfying first-order optimality conditions. Let us define the active sets at  $\hat{x}^\varepsilon$  by  $\hat{A}^\varepsilon = \hat{A}_-^\varepsilon \cup \hat{A}_+^\varepsilon$ , where

$$\hat{A}_-^\varepsilon = \{t \in [0, T] : \hat{u}^\varepsilon(t) = u_a(t)\} \quad \text{and} \quad \hat{A}_+^\varepsilon = \{t \in [0, T] : \hat{u}^\varepsilon(t) = u_b(t)\}.$$

The corresponding inactive set at  $\hat{x}^\varepsilon$  is given by  $\hat{I}^\varepsilon = [0, T] \setminus \hat{A}^\varepsilon$ . First-order necessary optimality conditions are presented in the next theorem, which was proved in [4, Theorem 3.4].

**Theorem 2.5.** *Let  $\hat{x}^\varepsilon = (\hat{c}_1^\varepsilon, \hat{c}_2^\varepsilon, \hat{u}^\varepsilon) \in W(0, T) \times W(0, T) \times U_{\text{ad}}$  be a local solution to (2.5). Then there exists a unique Lagrange multiplier  $\hat{p}^\varepsilon = (\hat{\lambda}_1^\varepsilon, \hat{\lambda}_2^\varepsilon) \in W(0, T) \times W(0, T)$  such that the pair  $(\hat{\lambda}_1^\varepsilon, \hat{\lambda}_2^\varepsilon)$  are weak solutions to the adjoint (or dual)*

equations

$$(2.6a) \quad -(\hat{\lambda}_1^\varepsilon)_t - d_1 \Delta \hat{\lambda}_1^\varepsilon = -k_1 \hat{c}_2^\varepsilon \hat{\lambda}_1^\varepsilon - k_2 \hat{c}_2^\varepsilon \hat{\lambda}_2^\varepsilon \quad \text{in } Q,$$

$$(2.6b) \quad -(\hat{\lambda}_2^\varepsilon)_t - d_2 \Delta \hat{\lambda}_2^\varepsilon = -k_1 \hat{c}_1^\varepsilon \hat{\lambda}_1^\varepsilon - k_2 \hat{c}_1^\varepsilon \hat{\lambda}_2^\varepsilon \quad \text{in } Q,$$

$$(2.6c) \quad d_1 \frac{\partial \hat{\lambda}_1^\varepsilon}{\partial n} = 0 \quad \text{on } \Sigma,$$

$$(2.6d) \quad d_2 \frac{\partial \hat{\lambda}_2^\varepsilon}{\partial n} = 0 \quad \text{on } \Sigma,$$

$$(2.6e) \quad \hat{\lambda}_1^\varepsilon(T) = -\beta_1(\hat{c}_1^\varepsilon(T) - c_1 T) \quad \text{in } \Omega,$$

$$(2.6f) \quad \hat{\lambda}_2^\varepsilon(T) = -\beta_2(\hat{c}_2^\varepsilon(T) - c_2 T) \quad \text{in } \Omega.$$

Moreover, there is a Lagrange multiplier  $\hat{\xi}^\varepsilon \in L^2(0, T)$  associated with the bilateral inequality constraint satisfying

$$(2.7) \quad \hat{\xi}^\varepsilon|_{\hat{A}_-^\varepsilon} \leq 0, \quad \hat{\xi}^\varepsilon|_{\hat{A}_+^\varepsilon} \geq 0$$

and the optimality condition

$$(2.8) \quad \gamma(\hat{u}^\varepsilon(t) - u_d(t)) + \frac{3}{\varepsilon} \max \left\{ 0, \int_0^T \hat{u}^\varepsilon dt - u_c \right\}^2 - \int_{\Gamma_c} \alpha(t) \hat{\lambda}_2^\varepsilon(t) dx + \hat{\xi}^\varepsilon(t) = 0$$

holds for almost all  $t \in [0, T]$ .

Due to Theorem 2.3 we can define the solution operator

$$\mathcal{S} : L^2(0, T) \rightarrow W(0, T) \times W(0, T)$$

by  $(c_1, c_2) = \mathcal{S}(u)$  for  $u \in L^2(0, T)$ , where the pair  $(c_1, c_2) \in W(0, T) \times W(0, T)$  is the solution to (2.3). Introducing the reduced cost functional

$$\hat{J}_\varepsilon(u) = J_\varepsilon(\mathcal{S}(u), u),$$

problem (2.5) can be expressed as

$$(2.9) \quad \min \hat{J}_\varepsilon(u) \quad \text{s.t.} \quad u \in U_{\text{ad}}.$$

Notice that (2.9) is a minimization problem with bilateral control constraints but with no equality constraints. The gradient of  $\hat{J}_\varepsilon$  at a point  $u \in L^2(0, T)$  is given by

$$(2.10) \quad \nabla \hat{J}_\varepsilon(\hat{u}^\varepsilon) = \gamma(\hat{u}^\varepsilon - u_d) + \frac{3}{\varepsilon} \max \left\{ 0, \int_0^T \hat{u}^\varepsilon dt - u_c \right\}^2 - \int_{\Gamma_c} \alpha \hat{\lambda}_2^\varepsilon dx \in L^2(0, T),$$

where  $(\hat{\lambda}_1^\varepsilon, \hat{\lambda}_2^\varepsilon) \in W(0, T) \times W(0, T)$  solves (2.6) for the state pair  $(\hat{c}_1^\varepsilon, \hat{c}_2^\varepsilon)$ , which in turn is the solution to (2.3) for the control input  $\hat{u}^\varepsilon$ . This dependence of the adjoint variable on the control via the map  $u \mapsto \lambda$  will be emphasized by writing  $\lambda(u)$  in the sequel.

From Theorem 2.5 we derive that the first-order necessary optimality conditions

$$\langle \nabla \hat{J}_\varepsilon(\hat{u}^\varepsilon), u - \hat{u}^\varepsilon \rangle_{L^2(0, T)} \geq 0 \quad \text{for all } u \in U_{\text{ad}}$$

are equivalent to

$$(2.11a) \quad \gamma(\hat{u}^\varepsilon - u_d) + \frac{3}{\varepsilon} \max \left\{ 0, \int_0^T \hat{u}^\varepsilon dt - u_c \right\}^2 - \int_{\Gamma_c} \alpha \hat{\lambda}_2^\varepsilon(\hat{u}^\varepsilon) dx + \hat{\xi}^\varepsilon = 0 \quad \text{in } L^2(0, T),$$

where the Lagrange multiplier  $\hat{\xi}^\varepsilon \in L^2(0, T)$  associated to the bilateral control constraints satisfies (compare (2.7))

$$(2.11b) \quad \hat{\xi}^\varepsilon = \max \{0, \hat{\xi}^\varepsilon + c(\hat{u}^\varepsilon - u_b)\} + \min \{0, \hat{\xi}^\varepsilon + c(\hat{u}^\varepsilon - u_a)\} \quad \text{in } L^2(0, T)$$

for some  $c > 0$ . Note that if (2.11b) holds for any  $c > 0$ , then it holds for all  $c > 0$ . In (2.11b) the functions max and min are interpreted as pointwise a.e. operations.

### 3. The semi-smooth Newton method

In this section we describe in detail the semi-smooth Newton method for the solution of the first order optimality system (2.11). The goal is to solve (2.5) by applying a Newton-type algorithm to the first-order necessary conditions (2.11). However, the max and min functions which appear in (2.11b) are not Fréchet differentiable. For this reason we recall the notion of Newton differentiability:

**Definition 3.1.** *Let  $X, Y$  be two Banach spaces,  $S \subset X$  a nonempty open set and  $f : S \rightarrow Y$  a given mapping. If there exists a neighborhood  $N(x^*) \subset S$  and a family of mappings  $g : N(x^*) \rightarrow L(X, Y)$  such that*

$$\|f(x^* + h) - f(x^*) - g(x^* + h)(h)\|_Y = o(\|h\|_X) \quad \text{for } \|h\|_X \rightarrow 0,$$

then  $f$  is called Newton-differentiable at  $x^*$  and  $g(x^*)$  is said to be a generalized derivative for  $f$  at  $x^*$ .

**Remark 3.2.** The function  $\max : L^p(\Omega) \rightarrow L^q(\Omega)$  is Newton differentiable for  $1 \leq q < p \leq \infty$  [5, 10]. If  $f : L^r(\Omega) \rightarrow L^p(\Omega)$  is Fréchet differentiable for some  $1 \leq r \leq \infty$ , then the function

$$x \mapsto \chi_A(x) \cdot f(u(x))$$

is a generalized derivative for  $\max\{0, f(\cdot)\}$ . Here,  $\chi_A$  denotes the characteristic function of the set  $A$  where  $f(u(\cdot))$  is nonnegative, i.e.,  $\chi_A(x) = 1$  where  $f(u(x)) \geq 0$  and  $\chi_A(x) = 0$  otherwise. From  $\min\{0, f(\cdot)\} = -\max\{0, -f(\cdot)\}$ , we see that the same differentiation formula holds for the min function.  $\diamond$

Using this concept of generalized differentiability, we can formulate the semi-smooth Newton step for the system (2.11). According to Remark 3.2, we need to distinguish subsets of  $[0, T]$  according to the sign of the arguments in (2.11b) under the max and min functions. Hence, we define the active and inactive sets

$$\begin{aligned} A_+ &= \{t \in [0, T] : \xi + c(u - u_b) \geq 0\} & A &= A_+ \cup A_- \\ A_- &= \{t \in [0, T] : \xi + c(u - u_a) \leq 0\} & I &= [0, T] \setminus (A_+ \cup A_-). \end{aligned}$$

From the requirement  $u_a \leq u_b$  on  $[0, T]$  it follows that  $A_+, A_-$  and  $I$  form a partition of the interval  $[0, T]$ , possibly after the removal of ambiguities where  $u_a = u_b$  holds. Carrying out the differentiation at the current iterate  $(u, \xi)$ , the Newton update  $(\delta u, \delta \xi)$  is given by

$$\begin{aligned} \gamma \delta u + \frac{6}{\varepsilon} \max \left\{ 0, \int_0^T u \, dt - u_c \right\} \int_0^T \delta u \, dt - \int_{\Gamma_c} \alpha \delta \lambda_2(u) \, dx + \delta \xi \\ = -\gamma(u - u_d) - \frac{3}{\varepsilon} \max \left\{ 0, \int_0^T u \, dt - u_c \right\}^2 + \int_{\Gamma_c} \alpha \lambda_2(u) \, dx - \xi \quad (3.1a) \\ -\delta \xi + \chi_{A_+} \cdot (\delta \xi + c \delta u) + \chi_{A_-} \cdot (\delta \xi + c \delta u) \\ = \xi - \max\{0, \xi + c(u - u_b)\} - \min\{0, \xi + c(u - u_a)\}. \quad (3.1b) \end{aligned}$$

Both equations hold on the time interval  $[0, T]$ . We point out that the terms involving the max function in (3.1a) are constant on  $[0, T]$ . In formula (3.1a),  $\delta\lambda_2(u)$  is the second component of the derivative of the map  $u \mapsto \lambda$ , as described after (2.10), in the direction of  $\delta u$ . By the chain rule,  $(\delta\lambda_1(u), \delta\lambda_2(u))$  is given by the solution to the adjoint system (2.6) with terminal data  $(-\beta_1\delta c_1(T), -\beta_2\delta c_2(T))$  where in turn  $(\delta c_1, \delta c_2)$  solves the linearized forward equation

$$(3.2a) \quad (\delta c_1)_t - d_1\Delta\delta c_1 = -k_1c_2\delta c_1 - k_1c_2\delta c_2 \quad \text{in } Q,$$

$$(3.2b) \quad (\delta c_2)_t - d_2\Delta\delta c_2 = -k_2c_1\delta c_1 - k_2c_1\delta c_2 \quad \text{in } Q,$$

$$(3.2c) \quad d_1\frac{\partial\delta c_1}{\partial n} = 0 \quad \text{on } \Sigma,$$

$$(3.2d) \quad d_2\frac{\partial\delta c_2}{\partial n} = 0 \quad \text{on } \Sigma \setminus \Sigma_c,$$

$$(3.2e) \quad d_2\frac{\partial\delta c_2}{\partial n} = \delta u \alpha \quad \text{on } \Sigma_c,$$

$$(3.2f) \quad \delta c_1(0) = 0 \quad \text{in } \Omega,$$

$$(3.2g) \quad \delta c_2(0) = 0 \quad \text{in } \Omega.$$

Of course, it is possible to consider the non-symmetric linear system (3.1a)–(3.1b) and solve it using an iterative solution scheme. However, it appears advantageous to decompose the two equations (3.1) by considering them separately on the active and inactive sets,  $A_\pm$  and  $I$ . Let us start with equation (3.1b). On the active sets, the  $\xi$  terms cancel and we obtain

$$\delta u = -(u - u_b) \quad \text{on } A_+ \quad \delta u = -(u - u_a) \quad \text{on } A_-. \quad (3.3)$$

On the inactive set, the terms involving  $u$  cancel and we have

$$-\delta\xi = \xi \quad \text{on } I. \quad (3.4)$$

Hence, equation (3.1b) determines the control update  $\delta u$  on the active sets and the multiplier update  $\delta\xi$  on the inactive set. Plugging (3.3) into equation (3.1a) and choosing  $c = \gamma$ , we obtain

$$\begin{aligned} \delta\xi = & -\gamma(u_b - u_d) - \frac{3}{\varepsilon} \max\left\{0, \int_0^T u \, dt - u_c\right\}^2 + \int_{\Gamma_c} \alpha(\lambda_2(u) + \delta\lambda_2(u)) \, dx \\ & - \frac{6}{\varepsilon} \max\left\{0, \int_0^T u \, dt - u_c\right\} \int_0^T \delta u \, dt \quad \text{on } A_+ \end{aligned} \quad (3.5)$$

and the same formula with  $u_b$  replaced by  $u_a$  holds on  $A_-$ . Finally, the core equation is (3.1a) evaluated on the inactive set. Using (3.4) yields

$$\begin{aligned} \gamma\delta u + \frac{6}{\varepsilon} \max\left\{0, \int_0^T u \, dt - u_c\right\} \int_0^T \delta u \, dt - \int_{\Gamma_c} \alpha \delta\lambda_2(u) \, dx \\ = -\gamma(u - u_d) - \frac{3}{\varepsilon} \max\left\{0, \int_0^T u \, dt - u_c\right\}^2 + \int_{\Gamma_c} \alpha \lambda_2(u) \, dx \quad \text{on } I. \end{aligned} \quad (3.6)$$

Note that the second and third term on the left hand side involve contributions from  $\delta u$  on the active set. Hence, when solving (3.6) for  $\delta u$  on the inactive set  $I$ , these contributions need to be transferred to the right hand side, see (3.8) below. Altogether, we have the following semi-smooth Newton algorithm, written as an active set method:

- 1: Choose  $(u^0, \xi^0) \in L^2(0, T) \times L^2(0, T)$  and let  $k := 0$ ;
- 2: Set

$$\begin{aligned} A_+^k &:= \{t \in [0, T] : \xi^k + \gamma(u^k - u_b) \geq 0\} \\ A_-^k &:= \{t \in [0, T] : \xi^k + \gamma(u^k - u_a) \leq 0\} \\ I^k &:= [0, T] \setminus (A_+^k \cup A_-^k); \end{aligned}$$

- 3: Set  $\delta u := -(u^k - u_b)$  on  $A_+^k$  and  $\delta u := -(u^k - u_a)$  on  $A_-^k$ ;
- 4: Solve (3.6) for  $\delta u$  on  $I^k$  using, e.g., the conjugate gradient method;
- 5: Set  $\delta \xi := -\xi^k$  on  $I^k$  and  $\delta \xi$  according to (3.5), evaluated at  $u^k$  and  $\lambda_2^k(u^k)$ , on  $A_+^k$  and  $A_-^k$ ;
- 6: Set  $u^{k+1} := u^k + \delta u$  and  $\xi^{k+1} := \xi^k + \delta \xi$ ;
- 7: Set  $k := k + 1$  and go to 2.

In practice, we can replace  $\gamma$  in step 2 by  $c \neq \gamma$ , see [9]. Clearly, the core step of the algorithm is step 4. Let us describe it in more detail. The left hand side of (3.6) is called the reduced Hessian operator or the reduced Hessian matrix in the discretized case. Let us split the control update  $\delta u = (\delta u|_A, \delta u|_I)^\top$  into its active and inactive components. Likewise, we can split the reduced Hessian

$$H = \begin{pmatrix} H_{II} & H_{IA} \\ H_{AI} & H_{AA} \end{pmatrix}.$$

Typically, the matrix–vector product with  $H$  is implemented as a subroutine. In order to evaluate  $H_{II}\delta u|_I$ , one simply forms the product of  $H$  with the vector  $(\delta u|_I, 0)^\top$  and sets the result equal to zero on the active set. Now we can rewrite (3.3) and (3.6) as

$$\begin{pmatrix} H_{II} & H_{IA} \\ 0 & \text{Id} \end{pmatrix} \begin{pmatrix} \delta u|_I \\ \delta u|_A \end{pmatrix} = \begin{pmatrix} b_I \\ -u + u_b \cdot \chi_{A+} + u_a \cdot \chi_{A-} \end{pmatrix} \quad (3.7)$$

where  $b_I$  is the right hand side of (3.6), restricted to the inactive set  $I$ . Hence, as mentioned above, (3.6) is equivalent to

$$H_{II}\delta u|_I = b_I - H_{IA}(-u + u_b \cdot \chi_{A+} + u_a \cdot \chi_{A-}) \quad (3.8)$$

which we solve using the conjugate gradient method.

#### 4. Numerical Example

In this section, we describe the behavior of the semi-smooth Newton method by means of one example for problem (2.5). All coding is done in MATLAB using routines from the FEMLAB 2.2 package concerning the finite element implementation. The given CPU times were obtained on a standard 1700 MHz desktop PC. They include only the run time for the core algorithm, excluding the generation of the mesh, pre-computing integrals and incomplete Cholesky decompositions. The three-dimensional geometry of the problem is given by the annular cylinder between the planes  $z = 0$  and  $z = 0.5$  with inner radius 0.4 and outer radius 1.0 whose rotational axis is the  $z$ -axis (Figure 4.2). The control boundary  $\Gamma_c$  is the upper annulus, and we use the control shape function

$$(4.1) \quad \alpha(t, x) = \exp\left(-5 \left[(x - 0.7 \cos(2\pi t))^2 + (y - 0.7 \sin(2\pi t))^2\right]\right),$$

see Figure 4.1. Note that  $\alpha$  corresponds to a nozzle circling for  $t \in [0, 1]$  once around in counter-clockwise direction at a radius of 0.7. For fixed  $t$ ,  $\alpha$  is a function which

decays exponentially with the square of the distance from the current location of the nozzle.

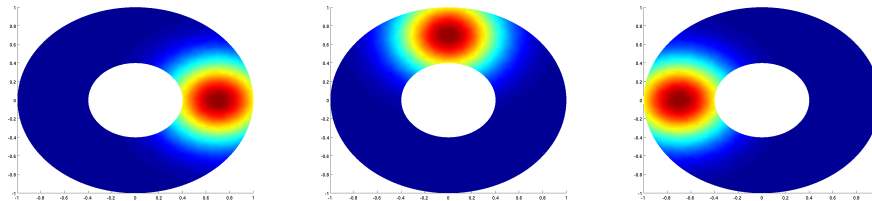


FIGURE 4.1. Control shape function  $\alpha(t, x)$  at  $t = 0.0$ ,  $t = 0.25$ , and  $t = 0.5$ .

The 'triangulation' of the domain  $\Omega$  by tetrahedra is also shown in Figure 4.2. It was created using an initial mesh obtained from `meshinit(fem, 'Hmax', 0.4)`. As the geometry suggests that much of the reaction will take place near the top surface  $\Gamma_c$  of the annular cylinder, we refine this initial mesh near the top using the command `meshinit(fem, 'Hexpr', '0.4*(0.5-z)+0.10', ...)`. The final mesh consists of 1797 points and 7519 tetrahedra. In the time direction, we use  $T = 1$

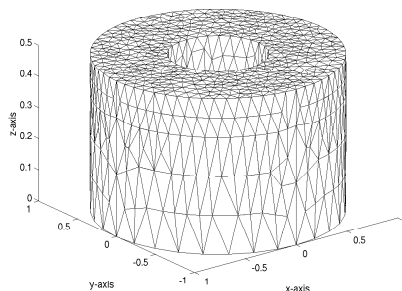


FIGURE 4.2. Domain  $\Omega \subset \mathbb{R}^3$  and its triangulation with tetrahedra

and partition the interval into 100 subintervals of equal lengths. The values of the control  $u$  at the interval endpoints serve as optimization variables. For more details concerning the problem, we refer to [4]. In particular, we choose  $u_c = 1.8$  as the upper limit in the integral constraint (2.4). We use the semi-implicit Euler time integration scheme for the state equation (2.2a)–(2.2b), where the nonlinearities are treated as explicit terms, i.e., they are always taken from the previous time step. In the adjoint equation, we use the same semi-implicit scheme, i.e., the right hand side terms in (2.6a)–(2.6b) are taken from the previously computed time step. The elliptic problem which arises on each time level in the state and adjoint equations is solved using the conjugate gradient method with incomplete Cholesky preconditioning. Note that the preconditioner needs to be computed only once since the coefficient matrices are the same in each time step, provided the time step lengths are all identical.

The reduced Hessian system arising in step 4 of the semi-smooth Newton algorithm was solved inexactly. The conjugate gradient iteration was terminated when

the residual in (3.8) satisfied  $\|r\| \leq 3 \cdot 10^{-4}$ . The overall algorithm was terminated in step 2 as soon as the active sets coincided for two consecutive iterations and the norm of the right hand side  $\|RHS\|$  in (3.8) was small. The convergence behavior is shown in Table 4.1 and the optimal solution is depicted in Figure 4.3. In the true solution, one expects that the multiplier  $\hat{\xi}^\varepsilon$  is negative on both parts of  $\hat{A}_-^\varepsilon$ . However, in the numerical solution, the multiplier is zero on the first part of the active set. In order to resolve it more accurately, a shaper convergence criterion can be used.

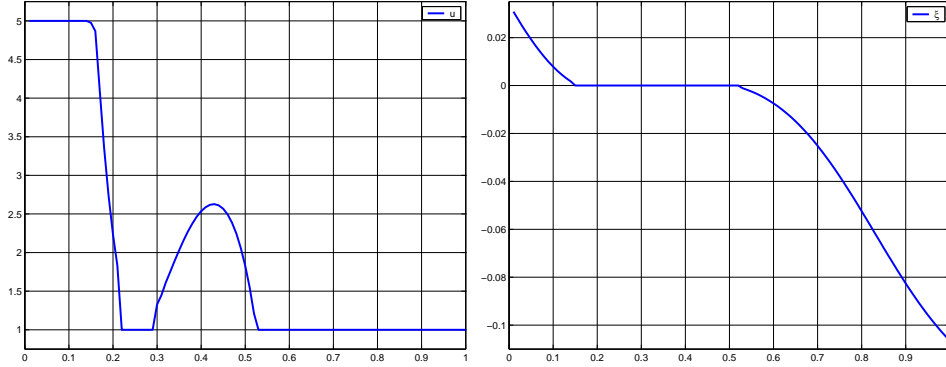


FIGURE 4.3. Optimal control  $u$  (left) and control constraint multiplier  $\xi$  (right)

$n$	$ A_+^n $	$ A_-^n $	$\ RHS\ $	#CG	$\ r\ $	objective	penalty
1	0	0	4.2473E+00	9	2.3501E-04	4.4024E-01	2.1999E-01
2	55	19	2.9243E+00	4	1.8966E-04	3.9442E-01	3.4150E-02
3	13	35	2.1628E+00	7	3.1059E-05	3.7371E-01	8.9096E-03
4	16	33	2.9777E-01	5	8.3128E-05	3.8231E-01	6.1422E-03
5	11	40	2.0084E-01	5	9.0991E-05	3.8367E-01	6.4499E-03
6	15	54	4.1048E-02	2	1.7892E-04	3.8410E-01	6.5784E-03
7	14	56	5.0219E-03	1	1.7539E-04	3.8411E-01	6.5979E-03
8	14	52	1.8502E-04				
9	14	48	1.8568E-04				
Run time: 1549 seconds						Objective: 0.3841	
						$\int_0^T u(t) dt = 1.9876$	

TABLE 4.1. Performance of the semi-smooth Newton method for  $\varepsilon = 1$ .

In [4, Proposition 2.9], it was proved that as the penalization parameter  $\varepsilon \searrow 0$ , the solution  $\hat{x}^\varepsilon$  converges in a weak sense to a solution which satisfies the integral constraint (2.4). This convergence can be observed from Table 4.2. One also notices that the problem becomes numerically more challenging as  $\varepsilon$  approaches zero.

$\varepsilon$	$\int_0^T \hat{u}^\varepsilon(t) dt$	#CG
100	3.2259	20
10	2.3451	30
1	1.9876	33
0.1	1.8612	37
0.01	1.8195	41

TABLE 4.2. Solving the problem for various values of  $\varepsilon$  and  $u_c = 1.8$ .

## REFERENCES

- [1] L. C. Evans. *Partial Differential Equations*. Graduate Studies in Mathematics, vol. 19, American Mathematical Society, Providence, Rhode Island, 1998.
- [2] C. Geiger and C. Kanzow. *Theorie und Numerik restringierter Optimierungsaufgaben*. Springer-Verlag, Berlin, 2002.
- [3] R. Griesse and S. Volkwein. Analysis for optimal boundary control for a three-dimensional reaction-diffusion system. Technical Report No. 286, Special Research Center F 003 *Optimization and Control*, Project area *Continuous Optimization and Control*, University of Graz & Technical University of Graz, Dezember 2003 (see <http://www.uni-graz.at/imawww/volkwein/publist.html>).
- [4] R. Griesse and S. Volkwein. A primal-dual active set strategy for optimal boundary control of a nonlinear reaction-diffusion system. Submitted, 2003.
- [5] M. Hintermüller, K. Ito, and K. Kunisch. The primal-dual active set strategy as a semi-smooth Newton method. *SIAM J. Optimization*, 13:865-888, 2003.
- [6] M. Hintermüller and G. Stadler. A semi-smooth Newton method for constrained linear-quadratic control problems. *Z. Angew. Math. Mech.*, 83:219-237, 2003.
- [7] K. Ito and K. Kunisch. The primal-dual active set method for nonlinear problems with bilateral constraints. Technical Report No. 214, Special Research Center F 003 *Optimization and Control*, Project area *Continuous Optimization and Control*, University of Graz & Technical University of Graz, 2001, to appear in *SIAM J. Control and Optimization*.
- [8] Z. Q. Luo, J. S. Pang, and D. Ralph. *Mathematical Programs with Equilibrium Constraints*. Cambridge University Press, Cambridge, 1996.
- [9] G. Stadler. *Infinite-Dimensional Semi-Smooth Newton and Augmented Lagrangian Methods for Friction and Contact Problems in Elasticity*. PhD thesis, University of Graz, 2004.
- [10] M. Ulbrich. Semismooth Newton methods for operator equations in function spaces. *SIAM J. Optimization*, 13:805-842, 2003.

R. GRIESSE AND S. VOLKWEIN, KARL-FRANZENS-UNIVERSITÄT GRAZ, INSTITUT FÜR MATHEMATIK, HEINRICHSTRASSE 36, A-8010 GRAZ, AUSTRIA

*E-mail address:* {roland.griesse,stefan.volkwein}@uni-graz.at