Johann Radon Institute for
Computational and Applied Mathematics
Austrian Academy of Sciences (ÖAW)

Optimal Control for an Elliptic System with Pointwise Nonlinear Control Constraints

K. Kunisch, X. Lu

RICAM-Report 2010-03

OPTIMAL CONTROL FOR AN ELLIPTIC SYSTEM WITH POINTWISE NONLINEAR CONTROL CONSTRAINTS

KUNISCH KARL AND XILIANG LU

Abstract. Optimal control for an elliptic system with pointwise Euclidean norm constraints on the control variables is investigated. First order optimality conditions are derived in a manner that is amenable for numerical realisation. An efficient semi-smooth Newton algorithm is proposed based on this optimality system. Numerical examples are given to validate the superlinear convergence of the semismooth Newton algorithm.

1. Introduction

Optimal control with control constraints have been studied intensively in recent papers, see e.g. [3, 5, 6, 10, 12] and the references cited there. In most previous discussions, the control constraints are taken as linear inequalities (e.g. unilateral or bilateral constraints). But for systems, the constraints may have more complicated structure, see [9] for a discussion on affine constraints and [13] for general convex constraints. In this paper, we consider the treatment of nonlinear constraint with an efficient numerical method. A pointwise Euclidean norm constraint for control variables, i.e., the control variable $\vec{u} \in K$, where K is an ellipse is considered, and a semismooth Newton algorithm is analyzed. For optimal control problems with scalar-valued elliptic equation constraints and unilateral constraints on the controls, an algebraic manipulation linking a parameter of the complementarity system to the weight of the control cost, denoted by α below, was the key technical step for proving superlinear convergence of the semi-smooth Newton method in [4]. Here we cannot rely on this technique, but rather have to develop an alternate complementarity system, which lends itself to analyzing the Newton differentiability property.

We choose a quadratic tracking type cost functional

$$J(\vec{y},\vec{u}) = \frac{1}{2}\|\vec{y}-\vec{y}_d\|^2 + \frac{\alpha}{2}\|C\vec{u}\|^2,$$

where the target function $\vec{y}_d \in (L^2(\Omega))^m$. The state and control variables satisfy an elliptic system with zero Dirichlet boundary condition:

$$\Lambda\vec{y} = C\vec{u} + \vec{d}, \quad \vec{y}|_{\partial\Omega} = \vec{0}, \tag{1.1}$$

where Ω be an open bounded domain in $\mathbb{R}^n$, for $n = 2$ or 3 with smooth boundary $\partial\Omega$. The vectors $\vec{y}$ and $\vec{u}$ have m-components. The optimal control problem is given by

Date: January 25, 2010.
Key words and phrases. optimal control, control constraints, semi-smooth Newton method, nonlinear constraint.

Problem 1.1.

$$\min J(\vec{y},\vec{u}), \quad \text{such that equations (1.1) hold and } |B\vec{u}(\mathbf{x})-\vec{b}|\le 1,\ a.e.\ \mathbf{x}\in\Omega.$$

The control variable $\vec{u}$ has support in a sub-domain $\tilde{\Omega}$, and the operator $C : (L^2(\tilde{\Omega}))^m \mapsto (L^2(\Omega))^m$ denotes the extension-by-zero operator. The $m\times m$ matrix B is invertible, $\vec{b}$ is a given vector in $\mathbb{R}^m$, and $|\cdot|$ denotes the Euclidean norm in $\mathbb{R}^m$, .

The paper is organized as follows. In section 2 and 3, we consider a special case, where $\tilde{\Omega}=\Omega$, $B=I$ and $\vec{b}=0$ to avoid some tedious notation. For this case, we give the details for constructing the optimality system and super-linear convergence for a semi-smooth Newton algorithm. In section 4, we point out that our results can be generalized in several directions, i.e., for a general subdomain $\tilde{\Omega}$, a general invertible matrix B and translation vector $\vec{b}$. In the last section 5, numerical examples are given to depict the efficiency for the semismooth Newton Algorithm.

We will use the standard notations $W^{m,p}$ and H^m for the Sobolev spaces, and simplify the notation of the norm of H^m as $\|f\|_m = \|f\|_{H^m}$ and $\|f\| = \|f\|_{L^2}$. The vector function

$$\vec{z}\in (W^{m,p}(\Omega))^m = W^{m,p}(\Omega,\mathbb{R}^m)$$

if and only if each coordinate of $\vec{z}$ is an element in $W^{m,p}(\Omega)$. We use $(\cdot,\cdot)$ as inner product in $L^2(\Omega)$ (for scalar functions) or $(L^2(\Omega))^m$ (for vector functions).

2. Optimality System For a Special case

We start with a simple case, where the governing equation is an elliptic system with zero boundary condition

$$\Lambda\vec{y} = \vec{u}+\vec{d}, \quad \vec{y}|_{\partial\Omega} = \vec{0}, \tag{2.1}$$

where Λ is a strong elliptic operator with the a-priori estimation

$$\|\vec{y}\|_2 \le C\|\Lambda\vec{y}\|.$$

In this section, the control variable $\vec{u}$ has support in the whole domain Ω, and it lies in the unit ball $|\vec{u}(\mathbf{x})|\le 1$ for almost every $x\in\Omega$. The optimal control problems in this case is given by

Problem 2.1.

$$\min J(\vec{y},\vec{u}) = \frac{1}{2}\|\vec{y}-\vec{y}_d\|^2 + \frac{\alpha}{2}\|\vec{u}\|^2, \quad \text{s.t. equation (2.1) hold and } |\vec{u}(\mathbf{x})|\le 1,\ a.e.\ \mathbf{x}\in\Omega.$$

We first obtain the existence for the optimal solution and provide the first order optimality condition. Define the map f from $(L^\infty)^m$ to L^∞ by

$$f(\vec{u})(\mathbf{x}) = |\vec{u}(\mathbf{x})|^2 - 1.$$

Then let K be the subset of L^∞ defined by

$$K = \{g(\mathbf{x})\in L^\infty(\Omega) : g(\mathbf{x})\le 0, a.e.\mathbf{x}\in\Omega\}, \tag{2.2}$$

and I_K be the indicator functional of K:

$$I_K(g) = \begin{cases} +\infty, & g\notin K \\ 0, & g\in K. \end{cases} \tag{2.3}$$

It is clear that K is a convex closed set in L^∞, and that the indicator functional I_K is a convex lower semicontinuous functional.

By using the notation of the control to state mapping $T = \Lambda^{-1}$, the cost functional $J(\vec{y}, \vec{u})$ can equivalently be represented as

$$\hat{J}(\vec{u}) = J(T\vec{u}, \vec{u}).$$

Then Problem 2.1 can be rewritten in the following equivalent way:

$$\inf_{\vec{u} \in (L^\infty)^m} \hat{J}(\vec{u}) + I_K(f(\vec{u})). \tag{2.4}$$

Theorem 2.2. *There exists a unique solution $(\vec{y}^*, \vec{u}^*)$ for Problem 2.1.*

Proof. We can check that $\hat{J} + I_K \circ f$ is a lower semi-continuous convex functional. Standard arguments imply the existence of an optimal solution, and strict convexity implies its uniqueness. □

Next we establish the optimality system by convex analysis methods. The admissible set of control variables $\vec{u}$ is given by

$$U_{ad} = \{\vec{u} : \vec{u} \in (L^\infty(\Omega))^m, |\vec{u}(\mathbf{x})| \leq 1, \ a.e. \, \mathbf{x} \in \Omega\}.$$

First $\vec{0}$ is an interior point of U_{ad} and hence the Slater condition is satisfied (c.f. [2]). Then for the optimal solution $\vec{u}^*$, we have

$$0 \in \partial \hat{J}(\vec{u}^*) + f'(\vec{u}^*) \partial I_k(f(\vec{u}^*)).$$

Hence there exists $\lambda^* \in \partial I_k(f(\vec{u}^*))$ such that $\lambda^* \in (L^\infty)'$ and

$$0 \in \partial \hat{J}(\vec{u}^*) + f'(\vec{u}^*)\lambda^*.$$

Since $\hat{J}(\vec{u}^*) = \frac{1}{2}\|T\vec{u}^* - \vec{y}_d\|^2 + \frac{\alpha}{2}\|\vec{u}^*\|^2$ and $f'(\vec{u}^*) = 2\vec{u}^*$, we have

$$T^*(T\vec{u}^* - \vec{y}_d) + \alpha\vec{u}^* + 2\lambda^*\vec{u}^* = 0.$$

Denote the adjoint state $\vec{p}^*$ by $\vec{p}^* = -T^*(\vec{y}^* - \vec{y}_d)$. It satisfies the adjoint equation

$$\Lambda^T \vec{p}^* = \vec{y}_d - \vec{y}^*, \quad \vec{p}^*|_{\partial\Omega} = \vec{0}, \tag{2.5}$$

and the optimality condition in the weak sense

$$\vec{p}^* = (\alpha + 2\lambda^*)\vec{u}^*, \tag{2.6}$$

i.e.,

$$2\langle \lambda^* \vec{u}^*, \vec{v} \rangle_{(L^\infty)', L^\infty} = (\vec{p}^* - \alpha\vec{u}^*, \vec{v}), \quad \forall \vec{v} \in (L^\infty(\Omega))^m.$$

Moreover the following variational inequality holds:

$$\langle \lambda^*, g - f(\vec{u}^*) \rangle_{(L^\infty)', L^\infty} \leq 0, \forall g \in K. \tag{2.7}$$

This implies that

$$\langle \lambda^*, t \rangle_{(L^\infty)', L^\infty} \geq 0, \forall t \in L^\infty, t \geq 0$$

and

$$\langle \lambda^*, f(\vec{u}^*) \rangle_{(L^\infty)', L^\infty} = 0.$$

Define two disjoint subsets of Ω as following:

$$\tilde{\mathcal{A}} = \{\mathbf{x} \in \Omega : |\vec{u}^*(\mathbf{x})|^2 \geq \frac{1}{2}\},$$

$$\tilde{\mathcal{B}} = \{\mathbf{x} \in \Omega : |\vec{u}^*(\mathbf{x})|^2 < \frac{1}{2}\}.$$

The set $\tilde{\mathcal{A}}$ contains the active set $\mathcal{A}(\vec{u}^*)$ which is defined as

$$\mathcal{A}(\vec{u}^*) = \{\mathbf{x} \in \Omega : |\vec{u}^*(\mathbf{x})| = 1\}. \tag{2.8}$$

Clearly $\Omega = \tilde{\mathcal{A}} \cup \tilde{\mathcal{B}}$ and hence for any scalar valued function ϕ, we have $\phi = \phi\chi_{\tilde{\mathcal{A}}} + \phi\chi_{\tilde{\mathcal{B}}}$, where χ_X is a characteristic function on the set X. One can observe that

$$\langle \lambda^*, f(\vec{u}^*)\chi_{\tilde{\mathcal{B}}}\rangle_{(L^\infty)',L^\infty} = 0.$$

Now consider any nonnegative function $\phi \in L^\infty$ with $\|\phi\|_{L^\infty} \le \frac{1}{2}$. We have

$$0 \le \langle \lambda^*, \frac{\phi}{2}\chi_{\tilde{\mathcal{B}}}\rangle_{(L^\infty)',L^\infty} = \langle \lambda^*, (\frac{\phi}{2} + f(\vec{u}^*))\chi_{\tilde{\mathcal{B}}}\rangle_{(L^\infty)',L^\infty} \le 0,$$

and hence

$$\langle \lambda^*, \phi\rangle_{(L^\infty)',L^\infty} = 0, \text{ for all } \phi \text{ with support in } \tilde{\mathcal{B}}. \tag{2.9}$$

Define the function $\hat{\lambda}$ as

$$\hat{\lambda}(\mathbf{x}) = \begin{cases} 0 & \mathbf{x} \in \tilde{\mathcal{B}} \\ \frac{1}{2}(\frac{\vec{p}^*(\mathbf{x})\cdot\vec{u}^*(\mathbf{x})}{|\vec{u}^*(\mathbf{x})|^2} - \alpha) & \mathbf{x} \in \tilde{\mathcal{A}}. \end{cases}$$

We have $\hat{\lambda} \in L^\infty$, and for any $\phi \in L^\infty$, by (2.9)

$$\begin{aligned} &\langle \lambda^*, \phi\rangle_{(L^\infty)',L^\infty} = \langle \lambda^*, \phi\chi_{\tilde{\mathcal{A}}}\rangle_{(L^\infty)',L^\infty} = \langle \lambda^*\vec{u}^*, \phi\chi_{\tilde{\mathcal{A}}}\frac{\vec{u}^*}{|\vec{u}^*|^2}\rangle_{(L^\infty)',L^\infty} \\ &= \frac{1}{2}(\vec{p}^* - \alpha\vec{u}^*, \phi\chi_{\tilde{\mathcal{A}}}\frac{\vec{u}^*}{|\vec{u}^*|^2}) = \frac{1}{2}(\frac{\vec{p}^*\cdot\vec{u}^*}{|\vec{u}^*|^2} - \alpha, \phi\chi_{\tilde{\mathcal{A}}}) = (\hat{\lambda}, \phi). \end{aligned}$$

Hence

$$\lambda^* = \hat{\lambda} = \begin{cases} 0 & \mathbf{x} \in \tilde{\mathcal{B}} \\ \frac{1}{2}(\frac{\vec{p}^*(\mathbf{x})\cdot\vec{u}^*(\mathbf{x})}{|\vec{u}^*(\mathbf{x})|^2} - \alpha) & \mathbf{x} \in \tilde{\mathcal{A}}, \end{cases} \qquad \lambda^* \in L^\infty(\Omega).$$

Setting $\beta^* = \alpha + 2\lambda^*$, we have $\beta^* \in L^\infty$ and $\beta^* \ge \alpha$. The optimality condition (2.6) and the complimentary condition (2.7) can now be represented in pointwise form:

$$\begin{cases} \vec{p}^*(\mathbf{x}) = \beta^*(\mathbf{x})\vec{u}^*(\mathbf{x}), \quad \beta^*(\mathbf{x}) = \alpha + 2\lambda^*(\mathbf{x}), \\ \lambda^*(\mathbf{x}) \ge 0, \quad |\vec{u}^*(\mathbf{x})| \le 1, \quad \lambda^*(\mathbf{x})(|\vec{u}^*(\mathbf{x})| - 1) = 0, \end{cases} \tag{2.10}$$

for almost every $\mathbf{x} \in \Omega$. The first equation of (2.10) implies that at almost every point $\mathbf{x}$, the two vectors $\vec{p}^*(\mathbf{x})$ and $\vec{u}^*(\mathbf{x})$ are linearly dependent and

$$|\vec{p}^*(\mathbf{x})| = \beta^*(\mathbf{x})|\vec{u}^*(\mathbf{x})|. \tag{2.11}$$

It can be proved by contradiction

$$\begin{cases} |\vec{p}(\mathbf{x})| \le \alpha \Rightarrow \lambda^*(\mathbf{x}) = 0, \\ |\vec{p}(\mathbf{x})| \ge \alpha \Rightarrow |\vec{u}^*| = 1, \lambda^*(\mathbf{x}) = \frac{1}{2}(|\vec{p}(\mathbf{x})| - \alpha). \end{cases}$$

Therefore, the optimality condition and complementary condition (2.10) can equivalently be expressed as

$$\vec{p}^* = \beta^*\vec{u}^*, \quad \beta^* = \max(\alpha, |\vec{p}|).$$

Combining the above results, we have proved the following result:

Theorem 2.3. *The optimality system of Problem 2.1 is given by*

$$\begin{aligned} &\textit{Primal equation} && \Lambda \vec{y}^* = \vec{u}^* + \vec{d}, \quad \vec{y}^*|_{\partial\Omega} = \vec{0}, \\ &\textit{Adjoint equation} && \Lambda^T \vec{p}^* = \vec{y}_d - \vec{y}^*, \quad \vec{p}^*|_{\partial\Omega} = \vec{0}, \\ &\textit{Optimality condition} && \vec{p}^* = \beta^* \vec{u}^*, \quad \beta^* = \max(\alpha, |\vec{p}^*|). \end{aligned} \tag{2.12}$$

Remark 2.1. The optimality condition in Theorem 2.3 is equivalent but simpler than its original form (2.6) and (2.7). In this form, Lagrange multiplier $\lambda^* = \frac{1}{2}(\beta^*(\mathbf{x}) - \alpha)$ has been expressed explicitly as a function of the adjoint state $\vec{p}^*$, for which Newton differentiability will be shown later. This fact plays an essential role in the successful use of the semismooth Newton method.

From this optimality system, we deduce that the optimal solution $(\vec{y}^*, \vec{u}^*)$ enjoys more regularity, which is important to prove superlinear convergence of the semismooth Newton algorithm in the next section.

Corollary 2.4. Given $\vec{y}_d \in (L^2)^m$, $\vec{d} \in (L^2)^m$, the optimal solution $(\vec{y}^*, \vec{u}^*)$ of Problem 2.1 satisfies

$$\vec{y}^* \in (W^{1,p})^m, \quad \vec{u}^* \in (W^{1,p})^m,$$

where $p = 6$ for $n = 3$, and $p < \infty$ for $n = 2$ with $\Omega \subset \mathbb{R}^n$. In additional , if $\vec{d} \in (W^{1,p})^m$, then $\vec{y}^* \in (W^{3,p})^m$.

Proof. Let the index $p = 6$ for $n = 3$ and $p < \infty$ for $n = 2$. From the second equation in (2.12) and since $\vec{y}$, $\vec{y}_d \in (L^2)^m$, and by the Sobolev embedding theorem (c.f. [1]), we have

$$\vec{p}^* \in (H^2(\Omega))^m \hookrightarrow (W^{1,p}(\Omega))^m \hookrightarrow (C(\Omega))^m.$$

This implies that $|\vec{p}^*| \in C(\Omega)$ and $\beta^* = \max(\alpha, |\vec{p}^*|) \in C(\Omega)$. Therefore $\vec{u}^* = \frac{\vec{p}^*}{\beta^*} \in C(\Omega)$ as well. To check that the derivative is in L^p, we define two subsets of Ω by

$$\Omega_1 = \{\mathbf{x} \in \Omega : |\vec{p}^*(\mathbf{x})| > \frac{\alpha}{2}\}, \quad \Omega_2 = \{\mathbf{x} \in \Omega : |\vec{p}(\mathbf{x})| < \alpha\}.$$

Thanks to $|\vec{p}^*| \in C(\Omega)$, these two set Ω_i, $i = 1, 2$ are open. In Ω_1, we have

$$\nabla |\vec{p}^*| = \frac{\vec{p}^* \cdot \nabla \vec{p}^*}{|\vec{p}^*|^2}.$$

Therefore

$$\int_{\Omega_1} |\nabla |\vec{p}^*||^p = \int_{\Omega_1} \frac{|\vec{p}^* \cdot \nabla \vec{p}^*|^p}{|\vec{p}^*|^{2p}} \le \frac{|\vec{p}^*|^p_{L^\infty}}{(\frac{\alpha}{2})^{2p}} \int_{\Omega_1} |\nabla \vec{p}^*|^p < \infty.$$

This implies that $|\vec{p}^*| \in W^{1,p}(\Omega_1)$. By $\max(0, \cdot) : W^{1,p} \mapsto W^{1,p}$ (see [8]), and the third equation in (2.12), we obtain $\beta^* \in W^{1,p}(\Omega_1)$. On the other hand, $\beta^* = \alpha$ in Ω_2. Since Ω_i, $i = 1, 2$ are two open sets, with $\Omega_1 \cup \Omega_2 = \Omega$ we find

$$\beta^* \in W^{1,p}(\Omega_1 \cup \Omega_2) = W^{1,p}(\Omega).$$

Using $\vec{u}^* = \frac{\vec{p}^*}{\beta^*}$, we find

$$\begin{aligned} \int_\Omega |\nabla \vec{u}^*|^p &= \int_\Omega |\tfrac{\beta^* \nabla \vec{p}^* - \vec{p}^* \otimes \nabla \beta^*}{(\beta^*)^2}|^p \le \tfrac{2^{p-1}}{\alpha^{2p}} \int_\Omega (|\beta^* \nabla \vec{p}^*|^p + |\vec{p}^* \otimes \nabla \beta^*|^p) \\ &\le \tfrac{2^{p-1} \|\beta^*\|^p_{L^\infty}}{\alpha^{2p}} \int_\Omega \|\nabla \vec{p}^*\|^p + \tfrac{2^{p-1} |\vec{p}|^p_{L^\infty}}{\alpha^{2p}} \int_\Omega \|\nabla \beta^*\|^p < \infty, \end{aligned}$$

where the tensor product $(\cdot \otimes \cdot)$ for two column vectors is defined by

$$\vec{v}_1 \otimes \vec{v}_2 = \vec{v}_1 \vec{v}_2^T. \tag{2.13}$$

Therefore $\vec{u}^* \in (W^{1,p})^m$. If $\vec{d} \in (W^{1,p})^m$ standard elliptic regularity theory implies that $\vec{y}^* \in (W^{3,p})^m$ (c.f. [7]). □

Proposition 2.5. *The solution to the optimality system (2.12) is unique.*

Proof. We consider the original optimality system, i.e., primal equation (2.1), adjoint equation (2.5), the optimality condition (2.6) and complementary condition (2.7). Since $\lambda^* \in L^\infty(\Omega)$, we can express these equations as:

$$\begin{aligned} &\Lambda \vec{y}^* = \vec{u}^* + \vec{d}, \quad \vec{y}^*|_{\partial\Omega} = \vec{0}, \\ &\Lambda^T \vec{p}^* = \vec{y}_d - \vec{y}^*, \quad \vec{p}^*|_{\partial\Omega} = \vec{0}, \\ &\vec{p}^* = (\alpha + 2\lambda^*)\vec{u}^*, \quad |\vec{u}^*| \le 1, \quad (\lambda^*, t - f(\vec{u^*})) \le 0, \forall t \in K. \end{aligned} \tag{2.14}$$

From the arguments in Theorem 2.3, the optimality system (2.12) is equivalent to system (2.14). Suppose that system (2.14) admits two solutions $(\vec{y}_1, \vec{u}_1, \vec{p}_1, \lambda_1)$ and $(\vec{y}_2, \vec{u}_2, \vec{p}_2, \lambda_2)$, the difference of these two solutions by $(\delta\vec{y}, \delta\vec{u}, \delta\vec{p}, \delta\lambda) = (\vec{y}_1, \vec{u}_1, \vec{p}_1, \lambda_1) - (\vec{y}_2, \vec{u}_2, \vec{p}_2, \lambda_2)$. We observe that

$$(\lambda_1, |\vec{u}_1|^2 - |\vec{u}_2|^2) \ge 0, \quad (\lambda_2, |\vec{u}_1|^2 - |\vec{u}_2|^2) \le 0.$$

Hence $(\delta\lambda, |\vec{u}_1|^2 - |\vec{u}_2|^2) \ge 0$. By (2.14) we find that

$$\begin{aligned} &\Lambda \delta\vec{y} = \delta\vec{u}, \quad \delta\vec{y}|_{\partial\Omega} = \vec{0}, \\ &\Lambda^T \delta\vec{p} = -\delta\vec{y}, \quad \delta\vec{p}|_{\partial\Omega} = \vec{0}, \\ &\delta\vec{p} = \alpha\delta\vec{u} + 2(\lambda_1\vec{u}_1 - \lambda_2\vec{u}_2). \end{aligned}$$

Taking the inner product of the first equation with $\delta\vec{p}$ and the second equation with $\delta\vec{y}$ and adding the resulting expression, we have

$$\alpha\|\delta\vec{u}\|^2 + \|\delta\vec{y}\|^2 + 2(\lambda_1\vec{u}_1 - \lambda_2\vec{u}_2, \delta\vec{u}) = 0.$$

The last term in the above equation satisfies

$$\begin{aligned} &2(\lambda_1\vec{u}_1 - \lambda_2\vec{u}_2, \delta\vec{u}) = 2(\lambda_1, |\vec{u}_1|^2) + 2(\lambda_2, |\vec{u}_2|^2) - 2(\lambda_1 + \lambda_2, \vec{u}_1 \cdot \vec{u}_2) \\ &= (\delta\lambda, |\vec{u}_1|^2 - |\vec{u}_2|^2) + (\lambda_1 + \lambda_2, |\vec{u}_1|^2 + |\vec{u}_2|^2 - 2\vec{u}_1 \cdot \vec{u}_2) \ge 0. \end{aligned}$$

This implies that $\delta\vec{y} = \delta\vec{u} = 0$, hence the uniqueness follows. □

3. Semismooth Newton Algorithm

We will use a semismooth Newton method to solve the nonlinear system (2.12). First $\vec{u}^*$ can be replaced by $\frac{\vec{p}^*}{\beta^*}$. Let $x = (\vec{y}, \vec{p}, \beta)^t$, and

$$F(x) = \begin{pmatrix} \Lambda\vec{y} - \frac{\vec{p}}{\beta} - \vec{d} \\ \Lambda^T\vec{p} + \vec{y} - \vec{y}_d \\ \beta - \max(\alpha, |\vec{p}|) \end{pmatrix}, \tag{3.1}$$

where $\max(\alpha, |\vec{p}|)$ is defined pointwise. Associated to $\vec{p}$ define the active set and inactive set as

$$\mathcal{A} = \{\mathbf{x} : |\vec{p}(\mathbf{x})| > \alpha\}, \quad \mathcal{I} = \{\mathbf{x} : |\vec{p}(\mathbf{x})| \le \alpha\}. \tag{3.2}$$

The active set we defined here is different from the previous definition (2.8) which was used to derive the optimality system. The Newton derivative of F at x can be written as

$$D_N F(x) = \begin{pmatrix} \Lambda & -\frac{1}{\beta} & \frac{\vec{p}}{\beta^2} \\ I & \Lambda^T & 0 \\ 0 & -\frac{\vec{p}^t}{|\vec{p}|}\chi_{\mathcal{A}} & I \end{pmatrix}, \tag{3.3}$$

see e.g. [4]. Now we introduce the semismooth Newton iteration from x^k to x^{k+1}. Define the active sets and inactive sets at iterative level k as

$$\mathcal{A}^k = \{\mathbf{x} : |\vec{p}^k(\mathbf{x})| > \alpha\}, \quad \mathcal{I}^k = \{\mathbf{x} : |\vec{p}^k(\mathbf{x})| \leq \alpha\}. \tag{3.4}$$

Then the Newton step

$$D_N F(x^k)(x^{k+1} - x^k) = -F(x^k)$$

is equivalent to the system of equations,

$$\begin{cases} \Lambda \vec{y}^{k+1} = \frac{1}{\beta^k}\vec{p}^{k+1} + \frac{\vec{p}^k}{(\beta^k)^2}(\beta^k - \beta^{k+1}) + \vec{d}, \quad \vec{y}^{k+1}|_{\partial\Omega} = \vec{0}, \\ \Lambda^T \vec{p}^{k+1} = \vec{y}_d - \vec{y}^{k+1}, \quad \vec{p}^{k+1}|_{\partial\Omega} = \vec{0}, \\ \beta^{k+1} = \alpha\chi_{\mathcal{I}^k} + \frac{\vec{p}^k \cdot \vec{p^{k+1}}}{|\vec{p}^k|}\chi_{\mathcal{A}^k}. \end{cases} \tag{3.5}$$

We give the semismooth Newton algorithm in Algorithm 1. Step 4 ensures that $\beta^k \geq \alpha$, and hence the right hand side of first equation in (3.5) is well-defined.

Algorithm 1 Semismooth Newton Algorithm

1: Set k=0, initialize $\vec{p}^0$, $\beta^0 > 0$.
2: Let $\mathcal{A}^k = \{\mathbf{x} : |\vec{p}^k| > \alpha\}$, $\mathcal{I}^k = \{\mathbf{x} : |\vec{p}^k| \leq \alpha\}$.
3: Solve $(\vec{y}^{k+1}, \vec{p}^{k+1}, \beta^{k+1})$ from system (3.5).
4: Replace β^{k+1} by $\max(\alpha, \beta^{k+1})$.
5: stop or update $k = k + 1$, and go to 2.

For the analysis of Algorithm 1, we require the following technical lemma.

Lemma 3.1. *Consider the elliptic system with zero Dirichlet boundary condition as*

$$\begin{cases} \Lambda \vec{z} = (A + B)\vec{q} + \vec{h}, \\ \Lambda^T \vec{q} = -\vec{z}. \end{cases} \tag{3.6}$$

Let A be a nonnegative linear operator in $\mathcal{L}((L^2)^m, (L^2)^m)$, i.e. $(\vec{q}, A\vec{q}) \geq 0$, for all $\vec{q} \in (L^2(\Omega))^m$, and let $B \in \mathcal{L}((L^2)^m, (L^2)^m)$ be a perturbation operator which satisfies $|(\vec{q}, B\vec{q})| \leq \frac{1}{4\kappa}\|\vec{q}\|_{L^4}^2$, where κ is a constant such that

$$\|u\|^2 + \|u\|_{L^4}^2 \leq \kappa\|\Lambda^T u\|^2, \quad \forall\, u \in H^2 \cap H_0^1.$$

Then system (3.6) has a unique solution $(\vec{z}, \vec{q})$ for any $h \in (L^2(\Omega))^m$, and we have the a-priori estimation

$$\|\vec{q}\|_2 + \|\vec{z}\| \leq C\|\vec{h}\|. \tag{3.7}$$

Proof. We only need to check inequality (3.7), the existence being guaranteed by the Fredholm alternative (c.f. [7]). Taking the inner product of (3.6) with $(\vec{q}, \Lambda^T \vec{q})$ we have

$$(A\vec{q}, \vec{q}) + (B\vec{q}, \vec{q}) + (\vec{h}, \vec{q}) = (\Lambda \vec{z}, \vec{q}) = -\|\Lambda^T \vec{q}\|^2,$$

and hence

$$\|\Lambda^T \vec{q}\|^2 + (A\vec{q}, \vec{q}) = -(B\vec{q} + \vec{h}, \vec{q}) \le \frac{1}{4\kappa}\|\vec{q}\|_{L^4}^2 + \kappa\|\vec{h}\|^2 + \frac{1}{4\kappa}\|\vec{q}\|^2 \le \frac{1}{2}\|\Lambda^T \vec{q}\|^2 + \kappa\|\vec{h}\|^2.$$

Together with $\|\vec{q}\|_2 \le C\|\Lambda^T \vec{q}\|$, this gives estimate (3.7). □

Setting

$$(\vec{z}^k, \vec{q}^k, \gamma^k) = (\vec{y}^k, \vec{p}^k, \beta^k) - (\vec{y}^*, \vec{p}^*, \beta^*),$$

we find

$$\begin{cases} \Lambda \vec{z}^{k+1} = \frac{\beta^* \vec{q}^{k+1} - \vec{p}^* \gamma^k}{\beta^* \beta^k} + \frac{\vec{p}^k}{(\beta^k)^2}(\gamma^k - \gamma^{k+1}), \quad \vec{z}^{k+1}|_{\partial\Omega} = \vec{0}, \\ \Lambda^T \vec{q}^{k+1} = -\vec{z}^{k+1}, \quad \vec{q}^{k+1}|_{\partial\Omega} = \vec{0}, \\ \gamma^{k+1} = \frac{\vec{p}^k \cdot \vec{q}^{k+1}}{|\vec{p}^k|}\chi_{\mathcal{A}^k} + R^k, \end{cases} \tag{3.8}$$

where

$$R^k = \max(\alpha, |\vec{p}^k|) - \max(\alpha, |\vec{p}^*|) - \frac{\vec{p}^k \cdot \vec{q}^k}{|\vec{p}^k|}\chi_{\mathcal{A}^k}.$$

By the chain rule, $\vec{p} \to \max(\alpha, |\vec{p}|)$ is Newton differentiable from L^∞ to L^4 (c.f. [4]), and $D_N \max(\alpha, |\vec{p}|) = \frac{\vec{p}}{|\vec{p}|}\chi_{\mathcal{A}}$. Hence

$$\|R^k\|_{L^4} = o(\|\vec{q}^k\|_{L^\infty}).$$

From the first equation of (3.8), we have

$$\Lambda \vec{z}^{k+1} = \frac{1}{\beta^k}\left[\vec{q}^{k+1} - \frac{1}{\beta^k}\left(\frac{\vec{p}^k \cdot \vec{q^{k+1}}}{|\vec{p}^k|}\chi_{\mathcal{A}^k} + R^k\right)\vec{p}^k + \frac{\gamma^k}{\beta^k}\vec{p}^k - \frac{\gamma^k}{\beta^*}\vec{p}^*\right] = M^k \vec{q}^{k+1} + h^k,$$

where, using $\gamma^k = \beta^k - \beta^*$, $\vec{q}^k = \vec{p}^k - \vec{p}^*$,

$$M^k = \frac{1}{\beta^k}\left(I - \frac{1}{\beta^k}\frac{\vec{p}^k \otimes \vec{p^k}}{|\vec{p}^k|}\chi_{\mathcal{A}^k}\right), \quad h^k = \frac{1}{\beta^k}\left(\frac{\gamma^k}{\beta^k\beta^*}(\beta^*\vec{q}^k - \gamma^k\vec{p}^*) - \frac{R^k}{\beta^k}\vec{p}^k\right). \tag{3.9}$$

Then we decompose M^k into two parts

$$M^k = \underbrace{\frac{1}{\beta^k}\left(I - \frac{1}{\beta^*}\mathcal{H}(\vec{p}^*)\chi_{\mathcal{A}^k}\right)}_{D^k} + \underbrace{\frac{1}{\beta^k}\left(\frac{1}{\beta^*}\mathcal{H}(\vec{p}^*)\chi_{\mathcal{A}^k} - \frac{1}{\beta^k}\frac{\vec{p}^k \otimes \vec{p^k}}{|\vec{p}^k|}\chi_{\mathcal{A}^k}\right)}_{E^k},$$

where

$$\mathcal{H}(\vec{p}^*) = \begin{cases} \frac{\vec{p}^* \otimes \vec{p}^*}{|\vec{p}^*|} & \text{if } |\vec{p}^*| > 0, \\ 0_{m\times m} & \text{if } \vec{p}^* = 0. \end{cases}$$

Lemma 3.2. *If $\beta^k \ge \alpha$, then D^k is a nonnegative operator in $\mathcal{L}((L^2)^m, (L^2)^m)$, i.e. $(\vec{q}, D^k\vec{q}) \ge 0$ for all $\vec{q} \in (L^2(\Omega))^m$. Moreover if $\|\vec{p}^k - \vec{p}^*\|_{L^\infty} + \|\beta^k - \beta^*\|_{L^4} \le \epsilon$, for sufficient small ϵ, then we have*

$$|(\vec{q}, E^k\vec{q})| \le \frac{1}{4\kappa}\|\vec{q}\|_{L^4}^2. \tag{3.10}$$

Proof. We first check that D^k is nonnegative. From the optimality system (2.12), we find $\vec{p}^* = 0 \Leftrightarrow \vec{u}^* = 0$. Then defining

$$\mathcal{Z} = \{\mathbf{x} : |\vec{p}^*(\mathbf{x})| > 0\},$$

we have

$$\frac{1}{\beta^*}\mathcal{H}(\vec{p}^*)\chi_{\mathcal{A}^k} = \frac{1}{\beta^*}\frac{\vec{p}^* \otimes \vec{p}^*}{|\vec{p}^*|}\chi_{\mathcal{Z}}\chi_{\mathcal{A}^k} = \frac{\vec{u}^* \otimes \vec{u}^*}{|\vec{u}^*|}\chi_{\mathcal{Z}}\chi_{\mathcal{A}^k}, a.e. \text{ in } \Omega. \tag{3.11}$$

Hence for any $\vec{q} \in (L^2)^m$, we obtain

$$(\vec{q}, D^k\vec{q}) = (\vec{q}, \frac{\vec{q}}{\beta^k}) - (\frac{\vec{q}}{\beta^k}, \frac{\vec{u}^* \otimes \vec{u}^*}{|\vec{u}^*|}\chi_{\mathcal{Z}}\chi_{\mathcal{A}^k}\vec{q}).$$

Let $\vec{w} = \frac{\vec{q}}{\sqrt{\beta^k}}$ and $y = \vec{w}^t\vec{u}^*$, and compute

$$(\vec{q}, D^k\vec{q}) = \|\vec{w}\|^2 - \int_\Omega \frac{(\vec{w}\cdot\vec{u}^*)^2}{|\vec{u}^*|}\chi_{\mathcal{Z}}\chi_{\mathcal{A}^k}dx \geq \|\vec{w}\|^2 - \int_\Omega |\vec{w}|^2|\vec{u}^*|dx \geq 0.$$

The last step is due to the constraint $|\vec{u}^*| \leq 1$. Next we check that E^k satisfies (3.10). By (3.11),

$$\begin{aligned}
E^k &= \frac{1}{\beta^k\beta^*}\frac{\vec{p}^* \otimes \vec{p}^*}{|\vec{p}^*|}\chi_{\mathcal{Z}}\chi_{\mathcal{A}^k} - \frac{1}{(\beta^k)^2}\frac{\vec{p}^k \otimes \vec{p}^k}{|\vec{p}^k|}\chi_{\mathcal{A}^k} \\
&= \left(\frac{1}{\beta^k\beta^*}\frac{\vec{p}^* \otimes \vec{p}^*}{|\vec{p}^*|} - \frac{1}{(\beta^k)^2}\frac{\vec{p}^k \otimes \vec{p}^k}{|\vec{p}^k|}\right)\chi_{\mathcal{Z}\cap\mathcal{A}^k} - \frac{1}{(\beta^k)^2}\frac{\vec{p}^k \otimes \vec{p}^k}{|\vec{p}^k|}\chi_{\mathcal{A}^k\setminus\mathcal{Z}} \\
&= \underbrace{\frac{(|\vec{p}^k| - |\vec{p}^*|)\vec{p}^* \otimes \vec{p}^*}{\beta^k\beta^*|\vec{p}^k||\vec{p}^*|}\chi_{\mathcal{Z}\cap\mathcal{A}^k}}_{N_1} + \underbrace{\frac{(\beta^k - \beta^*)\vec{p}^* \otimes \vec{p}^*}{(\beta^k)^2\beta^*|\vec{p}^k|}\chi_{\mathcal{Z}\cap\mathcal{A}^k}}_{N_2} + \underbrace{\frac{(\vec{p}^* - \vec{p}^k) \otimes \vec{p}^*}{(\beta^k)^2|\vec{p}^k|}\chi_{\mathcal{Z}\cap\mathcal{A}^k}}_{N_3} \\
&\quad + \underbrace{\frac{\vec{p}^k \otimes (\vec{p}^* - \vec{p}^k)}{(\beta^k)^2|\vec{p}^k|}\chi_{\mathcal{Z}\cap\mathcal{A}^k}}_{N_4} - \underbrace{\frac{1}{(\beta^k)^2}\frac{\vec{p}^k \otimes \vec{p}^k}{|\vec{p}^k|}\chi_{\mathcal{A}^k\setminus\mathcal{Z}}}_{N_5},
\end{aligned}$$

hence

$$|(\vec{q}, E^k\vec{q})| \leq |(\vec{q}, N_1\vec{q})| + |(\vec{q}, N_2\vec{q})| + |(\vec{q}, N_3\vec{q})| + |(\vec{q}, N_4\vec{q})| + |(\vec{q}, N_5\vec{q})|.$$

We need to estimate each term:

$$\begin{aligned}
|(\vec{q}, N_1\vec{q})| &\leq \frac{1}{\alpha^2}\int_\Omega \frac{(|\vec{p}^k| - |\vec{p}^*|)|\vec{p}^*|}{|\vec{p}^k|}|\vec{q}|^2\chi_{\mathcal{Z}\cap\mathcal{A}^k}dx \leq \frac{1}{\alpha^3}\|\vec{p}^*\|_{L^\infty}\|\vec{p}^k - \vec{p}^*\|_{L^\infty}\|\vec{q}\|^2 \leq C\epsilon\|\vec{q}\|^2_{L^4}, \\
|(\vec{q}, N_2\vec{q})| &\leq \frac{1}{\alpha^3}\int_\Omega \frac{|\vec{p}^*|^2}{|\vec{p}^k|}|\beta^k - \beta^*||\vec{q}|^2\chi_{\mathcal{Z}\cap\mathcal{A}^k}dx \leq \frac{C}{\alpha^4}\|\vec{p}^*\|^2_{L^\infty}\|\beta^k - \beta^*\|_{L^4}\|\vec{q}\|^2_{L^4} \leq C\epsilon\|\vec{q}\|^2_{L^4}, \\
|(\vec{q}, N_3\vec{q})| &\leq \frac{1}{\alpha^2}\int_\Omega \frac{(|\vec{p}^k| - |\vec{p}^*|)|\vec{p}^*|}{|\vec{p}^k|}|\vec{q}|^2\chi_{\mathcal{Z}\cap\mathcal{A}^k}dx \leq \frac{1}{\alpha^3}\|\vec{p}^*\|_{L^\infty}\|\vec{p}^k - \vec{p}^*\|_{L^\infty}\|\vec{q}\|^2 \leq C\epsilon\|\vec{q}\|^2_{L^4}, \\
|(\vec{q}, N_4\vec{q})| &\leq \frac{1}{\alpha^2}\int_\Omega \frac{(|\vec{p}^k| - |\vec{p}^*|)|\vec{p}^*|}{|\vec{p}^k|}|\vec{q}|^2\chi_{\mathcal{Z}\cap\mathcal{A}^k}dx \leq \frac{1}{\alpha^3}\|\vec{p}^*\|_{L^\infty}\|\vec{p}^k - \vec{p}^*\|_{L^\infty}\|\vec{q}\|^2 \leq C\epsilon\|\vec{q}\|^2_{L^4}, \\
|(\vec{q}, N_5\vec{q})| &\leq \frac{1}{\alpha^2}\int_\Omega |\vec{p}^k||\vec{q}|^2\chi_{\mathcal{A}^k\setminus\mathcal{Z}}dx \leq \frac{\epsilon}{\alpha^2}\|\vec{q}\|^2.
\end{aligned}$$

The last inequality follows from the fact:

$$\vec{p}^k = \vec{p}^k - \vec{p}^*, \forall \mathbf{x} \in \mathcal{A}^k\setminus\mathcal{Z}.$$

If ϵ is sufficiently small, we obtain (3.10). □

Lemma 3.3. *Let $\beta^k \geq \alpha$, and $\|\vec{p}^k - \vec{p}^*\|_{L^\infty} + \|\beta^k - \beta^*\|_{L^4} \leq \epsilon$, then for sufficiently small ϵ, we have $\|h^k\| = o(\epsilon)$.*

Proof. Recall the definition (3.9),

$$h^k = \frac{1}{\beta^k}\left(\frac{\gamma^k}{\beta^k\beta^*}(\beta^*\vec{q}^k - \gamma^k\vec{p}^*) - \frac{R^k}{\beta^k}\vec{p}^k\right),$$

and $r^k = \beta^k - \beta^*$. Then

$$\begin{aligned}
\|\frac{\gamma^k}{(\beta^k)^2}\vec{q}^k\| &\leq \frac{1}{\alpha^2}\|\vec{q}^k\|_{L^\infty}\|\gamma^k\| = o(\epsilon),\\
\|\frac{(\gamma^k)^2}{(\beta^k)^2\beta^*}\vec{p}^*\| &\leq \frac{1}{\alpha^3}\|\vec{p}^*\|_{L^\infty}\|\gamma^k\|_{L^4}^2 = o(\epsilon),\\
\|\frac{R^k}{(\beta^k)^2}\vec{p}^k\| &\leq \frac{1}{\alpha^2}\|\vec{p}^k\|_{L^\infty}\|R^k\| = o(\epsilon),
\end{aligned}$$

which completes the estimate. □

Now we move to our convergence result for the semismooth Newton method.

Theorem 3.4. *Superlinear Convergence for Algorithm 1.*
If we assume $\beta^k \geq \alpha$, and $\|\beta^k - \beta^*\|_{L^4} + \|\vec{p}^k - \vec{p}^*\|_{L^\infty} \leq \epsilon$, then the Newton iteration in Algorithm 1 is well defined and $\|\beta^{k+1} - \beta^*\|_{L^4} + \|\vec{p}^{k+1} - \vec{p}^*\|_\infty = o(\epsilon)$.

Proof. We denote the solution to (3.5) by $(\vec{y}^{k+1}, \vec{p}^{k+1}, \vec{u}^{k+1}, \tilde{\beta}^{k+1})$ and

$$\beta^{k+1} = \max(\alpha, \tilde{\beta}^{k+1}).$$

We apply Lemma 3.2 to the first two equations in (3.8) by setting

$$D^k = A, \quad E^k = B, \quad \vec{z}^{k+1} = \vec{z}, \quad \vec{q}^{k+1} = \vec{q}, \quad \vec{h}^k = \vec{h},$$

where D^k, E^k and h^k are defined in (3.9). By Lemma 3.2, it is noticed that D^k and E^k satisfy the conditions in Lemma 3.1. Combined with Lemma 3.3, and $\vec{z}^{k+1} = \vec{p}^{k+1} - \vec{p}^*$, this implies that

$$\|\vec{p}^{k+1} - \vec{p}^*\|_{L^\infty} \leq C\|\vec{p}^{k+1} - \vec{p}^*\|_2 \leq C\|\vec{h}^k\| = o(\epsilon).$$

Then we notice that

$$\|\tilde{\beta}^{k+1} - \beta^*\|_{L^4} \leq \|\frac{\vec{p}^k\cdot\vec{q}^{k+1}}{|\vec{p}^k|}\chi_{\mathcal{A}^k}\|_{L^4} + \|R^k\|_{L^4} \leq C\|\vec{p}^{k+1} - \vec{p}^*\|_{L^\infty} + \|R^k\|_{L^4} = o(\epsilon).$$

Thanks to $\beta^* \geq \alpha$, the projection step does not increase error, i.e.

$$|\beta^{k+1}(\mathbf{x}) - \beta^*(\mathbf{x})| = |\max(\alpha, \tilde{\beta}^{k+1}(\mathbf{x})) - \beta^*(\mathbf{x})| \leq |\tilde{\beta}^{k+1}(\mathbf{x}) - \beta^*(\mathbf{x})|, a.e.\mathbf{x} \in \Omega.$$

Then we obtain

$$\|\beta^{k+1} - \beta^*\|_{L^4} + \|\vec{p}^{k+1} - \vec{p}^*\|_\infty = o(\epsilon).$$

□

Remark 3.1. We eliminated the control variable $\vec{u}$ from the optimality system using $\vec{p} = \beta\vec{u}$. The semismooth Newton algorithm can also be applied is we treat $\vec{u}$ as an independent variable.

4. Generalization

We will discuss a few generalizations which are related to Problem 1.1. The optimality system corresponding to Problem 1.1 can be obtained by a similar argument as in section 2 and is given by

(4.1)

Primal equation	$\Lambda\vec{y} = C\vec{u} + \vec{d}, \quad \vec{y}\|_{\partial\Omega} = \vec{0},$
Adjoint equation	$\Lambda^T\vec{p} = \vec{y}_d - \vec{y}, \quad \vec{p}\|_{\partial\Omega} = \vec{0},$
Optimality condition	$C^*\vec{p} = \alpha C^*C\vec{u} + 2\lambda B^T(B\vec{u} + \vec{b})$ in the weak sense,
Complementary condition	$\lambda \geq 0,\ \|B\vec{u} - \vec{b}\| \leq 1,\ \langle \lambda, \|B\vec{u} - \vec{b}\|^2 - 1\rangle_{(L^\infty)', L^\infty} = 0.$

To apply the semismooth Newton algorithm, we need to rewrite the optimality system in terms of a system of nonlinear equations which are Newton differentiable. Since the primal and adjoint equations are both linear, we only focus on the optimality and the complementary conditions.

In this section we generalize our previous work in three directions, namely the case where the control $\vec{u}$ is supported in a subdomain $\tilde{\Omega} \subset \Omega$, the case of an arbitrary translative vector $\vec{b}$, and the case of an invertible linear transformation B. For simplicity of discussion, we take each generalization individually, but these generalization can be combined.

Our goal in the following subsections is to rewrite the optimality system into equivalent nonlinear equations and to prove that these nonlinear equations are actually Newton differentiable, and hence that the semismooth Newton method provides an efficient algorithm with locally superlinear convergence.

4.1. Extension to a control domain $\tilde{\Omega} \subset \Omega$.

Recall the definition of the bounded linear operator $C : (L^2(\tilde{\Omega}))^m \mapsto (L^2(\Omega))^m$ as the extension-by-zero operator. We first note that the dual operator $C^* : (L^2(\Omega))^m \mapsto (L^2(\tilde{\Omega}))^m$ is the restriction operator $C^*\vec{f} = \vec{f}\chi_{\tilde{\Omega}}$, and the operator $C^*C : (L^2(\tilde{\Omega}))^m \mapsto (L^2(\tilde{\Omega}))^m$ is identity operator. Then the optimality condition can be equivalently expressed as

$$\vec{p}\chi_{\tilde{\Omega}} = \alpha\vec{u} + 2\lambda\vec{u}.$$

Using the same arguments as in Theorem 2.3, the optimality and complementary conditions can be expressed in the following nonlinear equation

$$\vec{p}\chi_{\tilde{\Omega}} = \beta\vec{u}, \quad \beta = \max(\alpha, |\vec{p}|).$$

For a given $\vec{y}_d \in L^2(\Omega)$, the adjoint equation provides a solution $\vec{p} \in H^2 \cap H_0^1(\Omega) \hookrightarrow C_0(\Omega)$. Then the mapping $\vec{p} \mapsto \max(\alpha, |\vec{p}|)$ is Newton differentiable (c.f. [4]) from L^∞ to L^p, for any $p < \infty$. Hence this nonlinear equation can be solved by a semismooth Newton algorithm as in section 3.

4.2. Extension to a general translation vector b.

In this case, the optimality system reads

$$\vec{p} = \alpha\vec{u} + 2\lambda(\vec{u} - \vec{b}),$$

and it can be equivalently expressed as

$$\vec{p} - \alpha\vec{b} = (\alpha + 2\lambda)(\vec{u} - \vec{b}).$$

We can then rewrite the optimality condition and complementary condition as

$$\vec{p} - \alpha\vec{b} = \beta\vec{u}, \quad \beta = \max(\alpha, |\vec{p} - \alpha\vec{b}|)$$

by applying the same argument as in Theorem 2.3. This is again a Newton differentiable function and the semismooth Newton algorithm is applicable.

The first two generalizations are straightforward, but the next case which involves a general linear transformation B, is rather complicated.

4.3. **Extension to a general invertible matrix B.**

For a general $m \times m$ invertible matrix B, the argument in Theorem 2.3 on improving regularity of the Lagrange multiplier (from a measure to a function) needs to be slightly modified as follows. Recall the optimality and complementary conditions

$$\vec{p}^* = (\alpha I + 2\lambda B^T B)\vec{u}^* \text{ in the weak sense,}$$

$$\lambda^* \geq 0, \, |B\vec{u}^* - \vec{b}| \leq 1, \, \langle \lambda^*, |B\vec{u}^* - \vec{b}|^2 - 1\rangle_{(L^\infty)', L^\infty} = 0.$$

Define two disjoint subsets of Ω as follows:

$$\tilde{\mathcal{A}} = \{\mathbf{x} \in \Omega : |B\vec{u}^*(\mathbf{x})| \geq \frac{1}{2}\},$$
$$\tilde{\mathcal{B}} = \{\mathbf{x} \in \Omega : |B\vec{u}^*(\mathbf{x})| < \frac{1}{2}\}.$$

Then $\Omega = \tilde{\mathcal{A}} \cup \tilde{\mathcal{B}}$ and hence for any scalar valued function ϕ, $\phi = \phi\chi_{\tilde{\mathcal{A}}} + \phi\chi_{\tilde{\mathcal{B}}}$. Now as in Theorem 2.3, we have $\langle \lambda^*, \phi\rangle_{(L^\infty)', L^\infty} = 0$ for all $\phi \in L^\infty$ with support in $\tilde{\mathcal{B}}$. Define a function $\hat{\lambda}$ as

$$\hat{\lambda}(\mathbf{x}) = \begin{cases} 0 & \mathbf{x} \in \tilde{\mathcal{B}}, \\ \frac{1}{2}\frac{(B\vec{u}^*(\mathbf{x}))\cdot(B^{-T}(\vec{p}^*(\mathbf{x}) - \alpha\vec{u}^*(\mathbf{x}))}{|B\vec{u}^*(\mathbf{x})|^2} & \mathbf{x} \in \tilde{\mathcal{A}}. \end{cases}$$

We have $\hat{\lambda} \in L^\infty$. As in Theorem 2.3, for any $\phi \in L^\infty$, we find $\langle \lambda^*, \phi\rangle_{(L^\infty)', L^\infty} = \hat{\lambda}, \phi)$, and hence $\lambda^* = \hat{\lambda} \in L^\infty(\Omega)$. Thus the optimality and the complementary conditions hold in a pointwise a.e. sense,

$$\vec{p}^*(\mathbf{x}) = (\alpha I + 2\lambda B^T B)\vec{u}^*(\mathbf{x}), \tag{4.2}$$

$$\lambda^*(\mathbf{x}) \geq 0, \, |B\vec{u}^*(\mathbf{x})| \leq 1, \, \lambda^*(\mathbf{x}) \cdot (|B\vec{u}^*(\mathbf{x})| - 1) = 0, a.e. \mathbf{x} \in \Omega. \tag{4.3}$$

Equation (4.2) is equivalent to

$$B^{-T}\vec{p}^* = (\alpha D + 2\lambda I)(B\vec{u}^*), \tag{4.4}$$

where $D = (BB^T)^{-1}$ is a symmetric, positive definite (SPD) matrix. The eigenvalues of D satisfy

$$0 < \xi_1 \leq \xi_2 \leq \cdots \leq \xi_m < \infty. \tag{4.5}$$

Lemma 4.1. *Given a nonzero vector $\vec{p}$, the real valued function*

$$h_{\vec{p}}(\lambda) : \lambda \mapsto |(\alpha D + 2\lambda I)^{-1} B^{-T} \vec{p}|$$

is well-defined and strictly monotonically decreasing in $(-\frac{1}{2}\alpha\xi_1, \infty)$. Moreover, $\lim_{\lambda\to\infty} h_{\vec{p}}(\lambda) = 0$.

Proof. First we observe for every $\lambda > -\frac{1}{2}\alpha\xi_1$, the matrix $\alpha D + 2\lambda I$ is also a SPD matrix, hence it is invertible, and the function $h_{\vec{p}}$ is well-defined. Now we check the monotone property. Let

$$-\frac{1}{2}\alpha\xi_1 < \lambda_1 < \lambda_2, \quad \vec{q}_1 = (\alpha D + 2\lambda_1 I)^{-1} B^{-T}\vec{p}, \quad \vec{q}_2 = (\alpha D + 2\lambda_2 I)^{-1} B^{-T}\vec{p},$$

and define

$$\tilde{q} = (\alpha D + 2\lambda_1 I)^{-1}(\alpha D + 2\lambda_2 I)^{-1} B^{-T}\vec{p}.$$

Clearly $\tilde{q}$ is not the zero vector. Since matrices $\alpha D+2\lambda_1 I$ and $\alpha D+2\lambda_2 I$ commute, we have

$$\vec{q}_1 = (\alpha D + 2\lambda_2 I)\tilde{q}, \quad \vec{q}_2 = (\alpha D + 2\lambda_1 I)\tilde{q}.$$

Therefore,

$$\begin{aligned} |\vec{q}_1|^2 &= |\vec{q}_2 + 2(\lambda_2 - \lambda_1)\tilde{q}|^2 = |\vec{q}_2|^2 + 4(\lambda_2 - \lambda_1)^2|\tilde{q}|^2 + 4(\lambda_2 - \lambda_1)(\vec{q}_2, \tilde{q}) \\ &> |\vec{q}_2|^2 + 4(\lambda_2 - \lambda_1)((\alpha D + 2\lambda_1 I)\tilde{q}, \tilde{q}) > |\vec{q}_2|^2, \end{aligned}$$

which implies the strict monotonicity. The last claim $\lim_{\lambda\to\infty} h_{\vec{p}}(\lambda) = 0$ is straightforward. □

If $h_{\vec{p}}(0) = |(\alpha D)^{-1}B^{-T}\vec{p}| \geq 1$, then there exists by Lemma 4.1 a unique $\lambda \geq 0$, such that $h_{\vec{p}}(\lambda) = 1$. With the help of this fact, we can define a mapping $g : \vec{p} \mapsto \lambda$ as

(4.6)

$$g(\vec{p}) = \begin{cases} 0 & \text{if } |(\alpha D)^{-1}B^{-T}\vec{p}| < 1, \\ \text{the unique nonnegative solution to } h_{\vec{p}}(\lambda) = 1 & \text{if } |(\alpha D)^{-1}B^{-T}\vec{p}| \geq 1. \end{cases}$$

Lemma 4.2. *Using the function $g(\vec{p})$ as above, we can rewrite the optimality and the complementary conditions (4.2) - (4.3) in one equation:*

$$\vec{p}^*(\mathbf{x}) = \alpha\vec{u}^*(\mathbf{x}) + 2g(\vec{p}^*(\mathbf{x}))B^T B\vec{u}^*(\mathbf{x}), \ a.e. \mathbf{x} \in \Omega.$$

Proof. Throughout the proof we evaluate the operators at a fixed point $\mathbf{x} \in \Omega$ which is not indicated in the notation. We first assume that (4.2) - (4.3) hold. If $h_{\vec{p}^*}(0) = |(\alpha D)^{-1}B^{-T}\vec{p}^*| < 1$, we have $|B\vec{u}^*| = h_{\vec{p}^*}(\lambda^*) \leq h_{\vec{p}^*}(0) < 1$ and hence $\lambda^* = 0 = g(\vec{p}^*)$. On the other hand, if $h_{\vec{p}^*}(0) = |(\alpha D)^{-1}B^{-T}\vec{p}^*| \geq 1$, then a proof by contradiction implies that it must be an active point, i.e., $|B\vec{u}^*| = 1$. In this case, λ^* satisfies $h_{\vec{p}^*}(\lambda^*) = 1$, this implies $\lambda^* = g(\vec{p}^*)$.

Conversely, if

$$\vec{p}^* = \alpha\vec{u}^* + 2g(\vec{p}^*)B^T B\vec{u}^*,$$

we can then define $\lambda^* = g(\vec{p}^*)$. Equation (4.2) is clearly satisfied. For the complementary condition (4.3), we only need to check $\lambda^*(|B\vec{u}^*| - 1) = 0$. Suppose $\lambda^* = g(\vec{p}^*) > 0$. From (4.6) we have

$$|(\alpha D)^{-1}B^{-T}\vec{p}^*| \geq 1 \text{ and } h_{\vec{p}^*}(\lambda^*) = 1.$$

The later equality implies by (4.4) that

$$1 = h_{\vec{p}^*}(\lambda^*) = |(\alpha D + 2\lambda^* I)^{-1}B^{-T}\vec{p}^*| = |B\vec{u}^*|.$$

□

By singular value decomposition, we can express D as

$$D = Q^{-1}\hat{D}Q,$$

where Q is an orthonormal matrix and $\hat{D}$ is a diagonal matrix of the form $\mathrm{diag}(\xi_1, \xi_2, ..., \xi_m)$. Then the equation

$$|(\alpha D + 2\lambda I)^{-1} B^{-T}\vec{p}| = 1$$

can be simplified as

$$|(\alpha\hat{D} + 2\lambda I)^{-1} Q B^{-T}\vec{p}| = 1.$$

Let

$$\hat{p} = QB^{-T}\vec{p} = (\hat{p}_1, ..., \hat{p}_m).$$

Then equation $|(\alpha\hat{D} + 2\lambda I)^{-1} Q B^{-T}\vec{p}| = 1$ becomes

$$\frac{\hat{p}_1^2}{(\alpha\xi_1 + 2\lambda)^2} + \cdots + \frac{\hat{p}_m^2}{(\alpha\xi_m + 2\lambda)^2} = 1.$$

We define two sets:

$\mathcal{O}_\lambda = \{\lambda : -\frac{\alpha}{2}\xi_1 < \lambda < \infty\}$, $\quad \mathcal{O}_p = \{\vec{q} \in \mathbb{R}^m : |(\alpha\hat{D}+2\lambda I)^{-1}\vec{q}| = 1, \text{ for some } \lambda \in \mathcal{O}_\lambda\}$,

and the mapping $\mathcal{F} : \mathcal{O}_p \mapsto \mathcal{O}_\lambda$ by

$$\lambda \in \mathcal{F}(\vec{q}) \Leftrightarrow |(\alpha\hat{D} + 2\lambda I)^{-1}\vec{q}| = 1.$$

Lemma 4.3. *The following properties hold.*

(1) *The map $\mathcal{F}$ is a well defined single valued function.*
(2) *$\mathcal{O}_p$ is open, and the function $\mathcal{F}$ is continuously differentiable in $\mathcal{O}_p$.*

Proof. **(1):** By definition of $\mathcal{O}_p$, for any $\vec{p} \in \mathcal{O}_p$, $\mathcal{F}(\vec{p})$ is not empty. We notice $\vec{0} \notin \mathcal{O}_p$. Applying the same argument as in Lemma 4.1, for any $\vec{p} \in \mathcal{O}_p$, the function $\lambda \mapsto |(\alpha\hat{D} + 2\lambda I)^{-1}\vec{p}|$ is strictly monotonically decreasing from $\mathcal{O}_\lambda$ to $\mathbb{R}$, and $\mathcal{F}$ is a single valued.
(2): Let $G : \mathbb{R}^m \times \mathcal{O}_\lambda \mapsto \mathbb{R}$ be given by

$$G(q_1, ..., q_m, \lambda) = \frac{q_1^2}{(\alpha\xi_1 + 2\lambda)^2} + \cdots + \frac{q_m^2}{(\alpha\xi_m + 2\lambda)^2} - 1.$$

Then G is a smooth function in the domain $\mathbb{R}^m \times \mathcal{O}_\lambda$. At any point $\vec{p} = (p_1, ..., p_m) \in \mathcal{O}_p$, there existes a unique $\lambda \in \mathcal{O}_\lambda$, such that $\lambda = \mathcal{F}(\vec{p})$. This implies that $G(p_1, ..., p_m, \lambda) = 0$, and

$$\frac{\partial G}{\partial \lambda} = -4\left(\frac{p_1^2}{(\alpha\xi_1 + 2\lambda)^3} + \cdots + \frac{p_m^2}{(\alpha\xi_m + 2\lambda)^3}\right) \le \frac{-4}{(\alpha\xi_m + 2\lambda)^3}|\vec{p}|^2 \ne 0.$$

By the implicit function theorem, (c.f. pp.224 in [11]): there exists an open neighborhood $U(\vec{p}) \subset \mathbb{R}^m$, such that for any $\vec{q} \in U(\vec{p})$, we can find a unique $\tau \in \mathcal{O}_\lambda$, such that $G(\vec{q}, \tau) = 0$. This defines a smooth map $\tau = \tau(\vec{q})$ from $U(\vec{p})$ to $\mathbb{R}$.
From this result, we conclude that $U(\vec{p}) \subset \mathcal{O}_p$, which implies that the set $\mathcal{O}_p$ is open. By local uniqueness of the implicit function, we find $\tau = \mathcal{F}(\vec{q})$ for any $\vec{q} \in U(\vec{p})$. Therefore, the function $\mathcal{F}$ is continuously differentiable in $\mathcal{O}_p$. □

Now we define

$$\mathcal{O}_{p,1} = \{\vec{p} \in \mathbb{R}^m : QB^{-T}\vec{p} \in \mathcal{O}_p\}, \quad \mathcal{O}_{p,2} = \{\vec{p} \in \mathbb{R}^m : |(\alpha D)^{-1}B^{-T}\vec{p}| < 1\}.$$

Lemma 4.4. *The two sets $\mathcal{O}_{p,i}$, $i = 1, 2$ are both open. Moreover $\mathcal{O}_{p,1} \cup \mathcal{O}_{p,2} = \mathbb{R}^m$.*

Proof. Since $\mathcal{O}_p$ is open, the set $\mathcal{O}_{p,1} = B^T Q^{-1}\mathcal{O}_p$ is open as well. A similar argument implies that $\mathcal{O}_{p,2}$ is also open. Now assume $\vec{p} \notin \mathcal{O}_{p,2}$, which gives $|(\alpha D)^{-1}B^{-T}\vec{p}| \geq 1$. Then there exists a unique $\lambda \geq 0$, such that

$$|(\alpha D + 2\lambda I)^{-1}B^{-T}\vec{p}| = 1.$$

Since

$$|(\alpha D + 2\lambda I)^{-1}B^{-T}\vec{p}| = |(\alpha\hat{D} + 2\lambda I)^{-1}QB^{-T}\vec{p}|,$$

we have $\vec{p} \in \mathcal{O}_{p,1}$. □

Now we can characterize the map $g(\vec{p}) : \vec{p} \mapsto \lambda$ by

$$\lambda = g(\vec{p}) = \begin{cases} 0 & \vec{p} \in \mathcal{O}_{p,2} \\ \max(0, \mathcal{F}(QB^{-T}\vec{p})) & \vec{p} \in \mathcal{O}_{p,1}, \end{cases} \tag{4.7}$$

where $\mathcal{F}$ is defined in Lemma 4.3.

Proposition 4.5. *The Lagrange multiplier $\lambda^* = g(\vec{p}^*)$ and control variable $\vec{u}^*$ are continuous functions in Ω.*

Proof. Since the adjoint state $\vec{p}^* \in C(\Omega)$, the following two subsets in the domain Ω are also open:

$$\mathcal{O}_1 = \{x : \vec{p}^*(x) \in \mathcal{O}_{p,1}\}, \quad \mathcal{O}_2 = \{x : \vec{p}^*(x) \in \mathcal{O}_{p,2}\}.$$

And these two open sets cover the whole domain Ω, hence the composite function $\lambda^* = g(\vec{p}^*)$ is also continuous. The continuity of $\vec{u}^*$ follows from (4.2) and the fact that $B^T B$ is positive definite. □

Proposition 4.6. *The map $g(\vec{p}) : \vec{p} \mapsto \lambda$ is Newton differentiable from L^∞ to L^q.*

Proof. Newton differentiability of $g(\vec{p}) : \vec{p} \mapsto \lambda$ follows from Lemma 4.3, Newton differentiability for $u \mapsto \max(0, u)$ (c.f. [4]) and the chain rule. □

The optimality and the complementary conditions in this case can be expressed as

$$\vec{p}^* = \alpha\vec{u}^* + 2g(\vec{p}^*)B^T B\vec{u}^*,$$

where the Newton differentiable function $g(\vec{p})$ is defined as in (4.7). Even though the function $g(\vec{p})$ is well defined and serves for verifying the local superlinear convergence of the semismooth Newton method, it may be difficult to realize it in practice since we need to solve the $2m$-order polynomial equation in the algorithm,

$$\frac{p_1^2}{(\xi_1 + \lambda)^2} + \cdots + \frac{p_m^2}{(\xi_m + \lambda)^2} = 1.$$

When $m = 2$, this can be solved explicitly. If $m \geq 3$, the explicit solution does not exist in general. Therefore the implementation is still a challenge problem.

5. Numerical test

Here we validate the superlinear convergence for a simple test problem involving the vector Laplace equations on the unit square $[0, 1] \times [0, 1]$ with homogenous Dirichlet boundary condition. The optimality system from section 2 is given by

$$\begin{cases} -\triangle\vec{y}^* = \vec{u}^*, \quad \vec{y}^*|_{\partial\Omega} = \vec{0}, \\ -\triangle\vec{p}^* = \vec{y}_d - \vec{y}^*, \quad \vec{p}^*|_{\partial\Omega} = \vec{0}, \\ \vec{p}^* = \beta^*\vec{u}^*, \quad \beta^* = \max(\alpha, |\vec{p}^*|). \end{cases}$$

Table 5.1. super-linear convergence

iter. number	1	2	3	4	5	6
$\|y^k - y^*\|$	1.981617	0.41951	0.05183	0.00537	0.0000002	0
$\frac{\|y^k - y^*\|}{\|y^{k-1} - y^*\|}$		0.2117	0.12356	0.10366	0.000039	0
$\|p^k - p^*\|$	3.575191	0.387668	0.025340	0.003437	0.00000006	0
$\frac{\|p^k - p^*\|}{\|p^{k-1} - p^*\|}$		0.10843	0.065367	0.135646	0.000019	0
$\|\beta^k - \beta^*\|$	0.198249	0.309420	0.162607	0.001718	0.0000004	0
$\frac{\|\beta^k - \beta^*\|}{\|\beta^{k-1} - \beta^*\|}$		1.56076	0.52552	0.010568	0.000258	0

Table 5.2. number of iteration for Algorithm 1

	α =0.1	0.01	0.001	0.0001
$p^0 = 0.1\alpha$ rand	5	6	6	7
$p^0 = 1\alpha$ rand	5	6	6	8
$p^0 = 10\alpha$rand	5	7	8	26
$p^0 = 100\alpha$ rand	6	7	9	37
$p^0 = 1000\alpha$ rand	6	8	12	NA

Our first example involves the construction of an the exact solution as follows: given a function $\vec{p}^*$ with enough regularity, define $\beta^* = \max(\alpha, |\vec{p}^*|)$, $\vec{u}^* = \frac{1}{\beta^*}\vec{p}^*$, $\vec{y}^* = (-\triangle)^{-1}\vec{u}^*$ and let $\vec{y}_d = \vec{y}^* - \triangle\vec{p}^*$. Then for any given initial guess $\vec{p}^0, \beta^0$, we can apply Algorithm 1 to find the solution to this problem, and we choose the residual level as a stop criterion.

In the numerical test, we choose $\vec{p}^* = \alpha(\sin(4\pi xy), \sin(8\pi xy) + x(1-x)y(1-y))$, and compute β^*, $\vec{u}^*$, $\vec{y}^*$ and $\vec{y}_d$ accordingly. For this example, we observe superlinear convergence in Table 5.1 for $\alpha = 0.001$ and random initial guess $\vec{p}^0, \beta^0$. It is also noted from Table 5.2, that for a decreasing sequence of α values and large random initial guesses, the algorithm always convergence in a few steps except in the most extreme case.

Our second example is slightly different. This time we do not construct an exact solution to compare with our approximation. Instead we let $\vec{y}_d = c\alpha(\sin(\pi xy) + x + 3y, \sin(2\pi x) + \cos(2\pi y))$, where c is a constant between 10 to 100 to make the constraint active at the optimal solution. We observe that for a relatively large α ($\alpha > 0.02$), the convergence is not sensitive to the initial guess. After decreasing α, the convergence region becomes smaller. With randomly chosen initial data the algorithm may not converge. We therefore introduce a continuation technique

Table 5.3. number of iteration for Algorithm 2

α^k	0.1	0.01	0.001	0.0001
$\alpha = 0.001$	3	3	10	
$\alpha = 0.0001$	3	3	4	9

which is explained in Algorithm 2. With this procedure the algorithm is robust. We use it with $\alpha^0 = 0.1$ and $\rho = 0.1$ for the following two test problems:

(1) $\alpha = 0.001$, $c = 30$, i.e., $\vec{y}_d = 0.03(\sin(\pi xy) + x + 3y, \sin(2\pi x) + \cos(2\pi y))$.
(2) $\alpha = 0.0001$, $c = 80$, i.e., $\vec{y}_d = 0.008(\sin(\pi xy) + x + 3y, \sin(2\pi x) + \cos(2\pi y))$.

In Table 5.3, we record the number of iterations at each sub-problem as α^k becomes small with $(\vec{p}^k, \beta^k)$ as the initial guess. The total number of iterations is less than 20. Figure 5.1 depicts the Lagrange multiplier for each of the two cases (we plot $\beta = \alpha + 2\lambda$). As expected these quantities are $W^{1,p}$ regular.

Algorithm 2 Continuation Algorithm

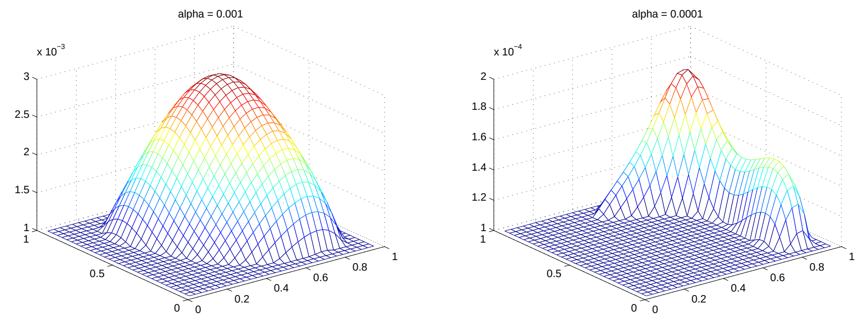
1: Choose a large α^0 (e.g. $\alpha^0 = 0.1$), and random initial guess $(\vec{p}^0, \beta^0)$, apply Algorithm 1, let it converges to $(\vec{y}^1, \vec{p}^1, \beta^1)$.

2: For $k = 1, 2, ...$, let $\alpha^k = \rho\alpha^{k-1}$, where ρ is a constant less than 1 (e.g. $\rho = \frac{1}{3}$).

3: Apply algorithm 1 for α^k, with $(\vec{p}^k, \beta^k)$ as our initial guess, let the algorithm converges to $(\vec{y}^{k+1}, \vec{p}^{k+1}, \beta^{k+1})$.

4: Stop (if the desired α value is achieved) or go to 2.

Figure 5.1. Lagrange multiplier



References

[1] R. Adams, J. Fournier, *Sobolev Spaces*, Academic Press, 2003.
[2] V. Barbu, T. Precupanu, *Convexity and Optimization in Banach Spaces*, Springer, 1986.
[3] M. Hintermüller, M. Hinze, *A SQP-semismooth Newton-type algorithm applied to control of the instationary Navier-Stokes system subject to control constraints*, SIAM J. Optim. 16 (2006), pp. 1177-1200.

[4] M. Hintermüller, K. Ito, and K. Kunisch, *The Primal-Dual Active Set Strategy as a Semi-Smooth Newton Method*, SIAM J. Optim., 13 (2003), 865-888.
[5] M. Hinze, *A variational discretization concept in control constrained optimization: the linear-quadratic case*, Comput. Optim. Appl. 30 (2005), pp. 45-61.
[6] K. Ito, and K. Kunisch, *The primal-dual active set method for nonlinear optimal control problems with bilateral constraints*, SIAM J. Control Optim. 43 (2004), pp. 357-376.
[7] D. Gilbarg, N. Trudinger, *Elliptic Partial Differential Equations of Second Order*, Springer-Verlag, Berlin, 2001.
[8] E. Lieb and M. Loss, *Analysis*, volume 14 of *Graduate Studies in Mathematics.* American Mathematical Society, Providence, RI, second edition, 2001.
[9] J. C. De Los Reyes and K. Kunisch, *Optimal control of partial differential equations with affine control constraints*, to appear in Control and Cybernet.
[10] A. Rösch, *Error estimates for parabolic optimal control problems with control constraints*, Z. Anal. Anwendungen 23 (2004), pp. 353-376.
[11] W. Rudin, *Principles of Mathematical Analysis*, McGraw-Hill, 1976.
[12] B. Vexler and W. Wollner, *Adaptive finite elements for elliptic optimization problems with control constraints*, SIAM J. Control Optim. 47 (2008), pp. 509-534.
[13] D. Wachsmuth, *Sufficient second-order optimality conditions for convex control constraints*, J. Math. Anal. Appl. 319 (2006), pp. 228-247.

Institute of Mathematics and Scientific Computing, University of Graz, Heinrichstrasse 36 A-8010 Graz, Austria
E-mail address: karl.kunisch@uni-graz.at

Johann Radon Institute for Computational and Applied Mathematics (RICAM), Austrian Academy of Sciences, Altenbergerstrasse 69, A-4040 Linz, Austria
E-mail address: xiliang.lu@ricam.oeaw.ac.at