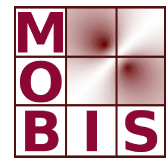




SpezialForschungsBereich F 32



Karl-Franzens Universität Graz  
Technische Universität Graz  
Medizinische Universität Graz



# Semi-smooth Newton Methods for Time-Optimal Control for a Class of ODEs

K. Ito      K. Kunisch

SFB-Report No. 2009-037

July 2009

A-8010 GRAZ, HEINRICHSTRASSE 36, AUSTRIA

Supported by the  
Austrian Science Fund (FWF)

**FWF** Der Wissenschaftsfonds.

SFB sponsors:

- **Austrian Science Fund (FWF)**
- **University of Graz**
- **Graz University of Technology**
- **Medical University of Graz**
- **Government of Styria**
- **City of Graz**



# Semi-smooth Newton Methods for Time-Optimal Control for a Class of ODEs

Kazufumi Ito<sup>1</sup>      Karl Kunisch<sup>2</sup>

July 11, 2009

<sup>1</sup>Department of Mathematics, North Carolina State University, Raleigh, North Carolina, 27695-8205, USA; research partially supported by the Army Research Office under DAAD19-02-1-0394

<sup>2</sup>Institut für Mathematik, Karl-Franzens-Universität Graz, A-8010 Graz, Austria, supported in part by the Fonds zur Förderung der wissenschaftlichen Forschung under SFB 32, "Mathematical Optimization and Applications in Biomedical Sciences".

## **Abstract**

Time optimal control problems for a class of linear multi-input systems are considered. The problems are regularized and the asymptotic and monotone behavior of the regularisation procedure is investigated. For the regularised problems the applicability of semi-smooth Newton methods is verified. First numerical tests are presented which show that the proposed approach, differently from other methods, does not rely a-priory information of the switching structure.

# 1 Introduction

This paper addresses time optimal control for a class of linear multi-input controls systems for ordinary differential equations. Due to their practical relevance and inherent structural difficulties, time optimal control has been receiving a considerable amount of attention for decades. Much of the literature up to the late sixties is covered in [HL]. Many recent results can be found or are referenced in [BPW, KLM, MO]. Time optimal control for infinite dimensional systems is considered in [Fa], for example.

The optimality system associated to time optimal control problems with pointwise constraints on the controls is complicated due to lack of smoothness of the optimal controls. In fact, the first order optimality system for time optimal control problems contains a multivalued operation which impedes the use of fast numerical methods. For this reason we introduce a regularization to the time optimal problem. In section 2 the behavior of the solutions of the regularized problems as the regularization parameter  $\varepsilon$  tends to zero is investigated. In particular monotonic structure of the solutions with respect to  $\varepsilon$  is shown. An optimality system for the regularized problems is derived under a condition which is stronger than controllability and weaker than normality. The optimal controls of the regularized problems are  $W^{1,\infty}$  regular and converge to a minimum norm solution of the original problem as the regularization parameter tends to zero.

The optimality system of the regularized problems is still not  $C^1$  so that second order methods with local quadratic convergence order are not directly applicable. However, sufficient conditions will be obtained in section 3 which imply that semi-smooth Newton methods [IK2] are wellposed and locally superlinearly convergent.

Section 4 contains a brief description of numerical results. We compare the chosen regularization to an alternative one, which has stronger regularization properties. Since the optimal controls of the original time optimal problems are typically not continuous, it appears that our choice of regularization which leads to  $W^{1,\infty}$  regularized controls is preferable over other regularization strategies which provide smoother controls. More detailed numerical tests are available in [XK].

Let us note that the approach that we propose for solving time optimal problems deviates from traditional approaches, which are frequently grouped into direct and indirect methods. Indirect methods based on multiple shooting techniques [Ke] solve the two point boundary value problem describing

first order necessary conditions. Equipped with a good initial guess for all unknowns, including the switching function, the shooting method is reported to converge fast and to generate very accurate solutions. The methods that we propose also originates from the first order condition, but differently from the shooting method it does not require accurate information on the switching structure in advance.

Direct methods on the other hand, consider time optimal problems as a genuine nonlinear programming problems. They are used in several variants, which frequently involve reparametrization of the controls as the unknowns. The new unknowns can be the switching times as in [MB] or the arc durations as in [KN].

## 2 The time-optimal problem and its regularization

Consider the time-optimal control problem for the linear multi-input system

$$(P) \quad \begin{cases} \min_{\tau \geq 0} \int_0^\tau dt \\ \text{subject to} \\ \frac{d}{dt}x(t) = Ax(t) + Bu(t), \quad |u(t)|_{\ell^\infty} \leq 1, \quad x(0) = x_0, \quad x(\tau) = x_1, \end{cases}$$

where  $A \in \mathbb{R}^{n \times n}$ ,  $B \in \mathbb{R}^{n \times m}$ ,  $x_0 \in \mathbb{R}^n$ ,  $x_1 \in \mathbb{R}^n$  are given,  $u(t) \in \mathbb{R}^m$ ,  $u$  is measurable, and  $|\cdot|_{\ell^\infty}$  denotes the infinity-norm on  $\mathbb{R}^m$ . The columns of  $B$  are denoted by  $b_i$ . It is assumed that  $x_1$  can be reached in finite time by an admissible control. Then (P) admits a solution with optimal time denoted by  $\tau^*$ , and associated state  $x^*$  and control  $u^*$ .

The first order optimality system for (P) can be expressed in terms of the adjoint  $p$  and the Hamiltonian

$$H(x, u, p_0, p) = p_0 + p^T(Ax + Bu),$$

as

$$(2.1) \quad \begin{cases} \dot{x} = Ax + Bu, & x(0) = x_0, & x(\tau) = x_1, \\ -\dot{p} = A^T p, \\ u = \operatorname{argmin}_{|v|_{\ell^\infty} \leq 1} H(x, v, p_0, p), \\ p_0 + p(\tau)^T (Ax(\tau) + Bu(\tau)) = 0, & p_0 \geq 0, \end{cases}$$

where the superscript  $T$  denotes transposition, see e.g. [MS], chapter V, pg. 109, 110. Further  $p$  is not identically 0, so that there exists a nontrivial vector  $q \in \mathbb{R}^n$  such that

$$p(t) = \exp(-A^T t) q.$$

Due to the special structure of  $H$  the optimal control can be expressed as

$$(2.2) \quad u_i = -\sigma(b_i^T p) = -\sigma(b_i^T \exp(-A^T t) q),$$

where  $\sigma$  denotes the coordinate-wise operation

$$(2.3) \quad \sigma(s) \in \begin{cases} -1 & \text{if } s < 0 \\ [-1, 1] & \text{if } s = 0 \\ 1 & \text{if } s > 0. \end{cases}$$

The last equation in (2.1) holds everywhere rather than a.e. on  $[0, \tau]$ . In fact,  $p$  and  $x$  are continuous and  $p(t)^T Bu(t) = -\sum_{i=1}^m |p(t)^T b_i|$ .

Let us recall the notions of controllability and normality, which will be referred to below.

$$(2.4) \quad \begin{cases} \text{The pair } (A, B) \text{ is called controllable i.e.} \\ \operatorname{rank} \{B, AB, \dots, A^{n-1}B\} = n \end{cases}$$

$$(2.5) \quad \begin{cases} \text{The pair } (A, B) \text{ is normal i.e. } (A, b_i) \\ \text{is controllable for all columns } b_i \text{ of } B. \end{cases}$$

Normality of  $(A, B)$  implies controllability. Moreover, if  $(A, B)$  is normal, then the optimal control  $u^*$  to (P) is unique, it is bang bang, and piecewise constant, see e.g. [MS, HL].

The requirement

$$(2.6) \quad p_0 > 0$$

is referred to as strict transversality. In this case it can be assumed that  $p_0 = 1$ , which can be achieved by scaling  $q$ . If strict transversality holds then  $(x^*, u^*, \tau^*)$  is a strict local minimum, in the sense that there exists  $\delta > 0$  such that  $x_1$  is not in the attainable set for  $t \in (\tau^* - \delta, \tau^*)$ , [HL], pg.89.

With (2.6) holding, we can express the optimality condition as

$$(2.7) \quad \begin{cases} \dot{x} = Ax + Bu, & x(0) = x_0, & x(\tau) = x_1, \\ -\dot{p} = A^T p, \\ u = \operatorname{argmin}_{|v|_{\ell^\infty} \leq 1} H(x, v, p), & \text{a.e. in } (0, \tau), \\ 1 + p(\tau)^T (Ax(\tau) + Bu(\tau)) = 0. \end{cases}$$

Here we eliminate the variable  $p_0$  from the notation for  $H$  since it was fixed to be 1.

Introducing the transformation  $\hat{t} = \frac{t}{\tau}$  and setting

$$\hat{x}(\hat{t}) = x(\tau\hat{t}) = x(t), \quad \hat{p}(\hat{t}) = p(\tau\hat{t}) = p(t), \quad \hat{u}(\hat{t}) = u(\tau\hat{t}) = u(t),$$

we obtain the following equivalent system to (2.7), where for the ease of presentation we omit the superscripts  $\wedge$ :

$$(2.8) \quad \begin{cases} \dot{x} = \tau(Ax + Bu), & x(0) = x_0, & x(1) = x_1 \\ -\dot{p} = \tau A^T p \\ u = \operatorname{argmin}_{|v|_{\ell^\infty} \leq 1} H(x, v, p) \\ 1 + p(1)^T (Ax(1) + Bu(1)) = 0. \end{cases}$$

The non-differentiable operation involved in characterizing the optimal control,

$$u = -\sigma(B^T p),$$

compare (2.2), prohibits the use of Newton-type methods for solving (2.8) numerically.



Therefore a family of regularized problems given by

$$(P_\varepsilon) \begin{cases} \min_{\tau \geq 0} \int_0^\tau (1 + \frac{\varepsilon}{2} |u(s)|^2) ds \\ \text{subject to} \\ \frac{d}{dt}x(t) = Ax(t) + Bu(t), |u(t)|_{\ell^\infty} \leq 1, x(0) = x_0, x(\tau) = x_1, \end{cases}$$

with  $\varepsilon > 0$  is considered. The norm  $|\cdot|$  used in the cost-functional denotes the Euclidean norm. It is straightforward to argue the existence of a solution  $(u_\varepsilon, x_\varepsilon, \tau_\varepsilon)$ .

Convergence of the solutions  $(x_\varepsilon, p_\varepsilon, u_\varepsilon, \tau_\varepsilon)$  of  $(P_\varepsilon)$  to a solution  $(x^*, p^*, u^*, \tau^*)$  of  $(P)$  is considered next. Note that  $\tau^*$  is unique.

**Proposition 2.1.** *For every  $0 < \varepsilon_0 < \varepsilon_1$  and any solution  $(\tau^*, u^*)$  of  $(P)$  we have*

$$(2.9) \quad \tau^* \leq \tau_{\varepsilon_0} \leq \tau_{\varepsilon_1} \leq \tau^* \left(1 + \frac{\varepsilon_1}{2}\right),$$

$$(2.10) \quad |u_{\varepsilon_1}|_{L^2(0, \tau_{\varepsilon_1})} \leq |u_{\varepsilon_0}|_{L^2(0, \tau_{\varepsilon_0})} \leq |u^*|_{L^2(0, \tau^*)}.$$

If  $u^*$  is a bang-bang solution, then

$$(2.11) \quad 0 \leq |u^*|_{L^2(0, \tau^*)}^2 - |u_\varepsilon|_{L^2(0, \tau_\varepsilon)}^2 \leq \text{meas} \{t \in [0, \tau^*] : |u_\varepsilon(t)| < 1\}$$

for every  $\varepsilon > 0$ .

*Proof.* From the definition of  $\tau^*$  and  $\tau_\varepsilon$  we have

$$\tau^* \leq \tau_\varepsilon \quad \text{for every } \varepsilon > 0,$$

and

$$\tau_\varepsilon + \frac{\varepsilon}{2} \int_0^{\tau_\varepsilon} |u_\varepsilon|^2 \leq \tau^* + \frac{\varepsilon}{2} \int_0^{\tau^*} |u^*|^2,$$

hence

$$|u_\varepsilon|_{L^2(0, \tau_\varepsilon)} \leq |u^*|_{L^2(0, \tau^*)}, \quad \text{and } \tau^* \leq \tau_\varepsilon \leq \tau^* \left(1 + \frac{\varepsilon}{2}\right).$$

For  $0 < \varepsilon_0 < \varepsilon_1$  we have

$$\int_0^{\tau_{\varepsilon_0}} \left(1 + \frac{\varepsilon_0}{2} |u_{\varepsilon_0}|^2\right) \leq \int_0^{\tau_{\varepsilon_1}} \left(1 + \frac{\varepsilon_0}{2} |u_{\varepsilon_1}|^2\right).$$

Adding  $\frac{1}{2}(\varepsilon_1 - \varepsilon_0) \int_0^{\tau_{\varepsilon_1}} |u_{\varepsilon_1}|^2$  on both sides implies that

$$(2.12) \quad \begin{aligned} \tau_{\varepsilon_0} + \frac{\varepsilon_1}{2} \int_0^{\tau_{\varepsilon_1}} |u_{\varepsilon_1}|^2 + \frac{\varepsilon_0}{2} \left( \int_0^{\tau_{\varepsilon_0}} |u_{\varepsilon_0}|^2 - \int_0^{\tau_{\varepsilon_1}} |u_{\varepsilon_1}|^2 \right) \\ \leq \int_0^{\tau_{\varepsilon_1}} \left( 1 + \frac{\varepsilon_1}{2} |u_{\varepsilon_1}|^2 \right) \leq \tau_{\varepsilon_0} + \frac{\varepsilon_1}{2} \int_0^{\tau_{\varepsilon_0}} |u_{\varepsilon_0}|^2. \end{aligned}$$

Estimating the first by the last expression in (2.12) implies that

$$\varepsilon_1 \left( \int_0^{\tau_{\varepsilon_1}} |u_{\varepsilon_1}|^2 - \int_0^{\tau_{\varepsilon_0}} |u_{\varepsilon_0}|^2 \right) \leq \varepsilon_0 \left( \int_0^{\tau_{\varepsilon_1}} |u_{\varepsilon_1}|^2 - \int_0^{\tau_{\varepsilon_0}} |u_{\varepsilon_0}|^2 \right),$$

and hence

$$(2.13) \quad |u_{\varepsilon_1}|_{L^2(0, \tau_{\varepsilon_1})} \leq |u_{\varepsilon_0}|_{L^2(0, \tau_{\varepsilon_0})}.$$

Estimating the first by the second expression in (2.12) we obtain

$$\tau_{\varepsilon_0} + \frac{\varepsilon_0}{2} |u_{\varepsilon_0}|_{L^2(0, \tau_{\varepsilon_0})}^2 \leq \tau_{\varepsilon_1} + \frac{\varepsilon_0}{2} |u_{\varepsilon_1}|_{L^2(0, \tau_{\varepsilon_1})}^2$$

and by (2.13)

$$\tau^* \leq \tau_{\varepsilon_0} \leq \tau_{\varepsilon_1}.$$

These estimates imply (2.9) and (2.10).

If  $u^*$  is bang-bang, then

$$\begin{aligned} 0 &\leq |u^*|_{L^2(0, \tau^*)}^2 - |u_\varepsilon|_{L^2(0, \tau_\varepsilon)}^2 \\ &\leq \int_{\{t \in (0, \tau^*) : |u_\varepsilon(t)| < 1\}} (1 - \sigma_\varepsilon^2) \leq \text{meas}\{t \in (0, \tau^*) : |u_\varepsilon(t)| < 1\}, \end{aligned}$$

so that (2.11) holds. □

**Theorem 2.1.** *For  $\varepsilon \rightarrow 0^+$  we have  $\tau_\varepsilon \rightarrow \tau^*$  and every convergent subsequence of solutions  $\{(u_\varepsilon, x_\varepsilon)\}_{\varepsilon > 0}$  to  $(P_\varepsilon)$  converges in  $L^2(0, \tau_\varepsilon; \mathbb{R}^m) \times W^{1,2}(0, \tau_\varepsilon; \mathbb{R}^n)$  to a solution  $(u^*, x^*)$  of  $(P)$ , where  $u^*$  is a minimum norm solution.*

Here convergence of  $u_\varepsilon$  to  $u^*$  is defined as

$$\int_0^1 |u_\varepsilon(\tau_\varepsilon t) - u^*(\tau^* t)|^2 dt \rightarrow 0$$

and analogously for  $\{x_\varepsilon\}$ , and for weak convergence.

*Proof.* The first claim follows from Proposition 2.1. Since  $\{u_\varepsilon(\tau_\varepsilon \cdot)\}_{\varepsilon>0}$  and  $\{x_\varepsilon(\tau_\varepsilon \cdot)\}_{\varepsilon>0}$  are bounded in  $L^2(0, 1; \mathbb{R}^m)$  and  $W^{1,2}(0, 1; \mathbb{R}^n)$ , there exist weak accumulation points  $u^* \in L^2(0, \tau^*; \mathbb{R}^m)$ , and  $x^* \in W^{1,2}(0, \tau^*; \mathbb{R}^n)$ . Subsequently we avoid subsequential indices. Passing to the limit in  $\dot{x}_\varepsilon(\tau_\varepsilon \cdot) = \tau_\varepsilon(Ax_\varepsilon(\tau_\varepsilon \cdot) + Bu_\varepsilon(\tau_\varepsilon \cdot))$  and  $x_\varepsilon(0) = x_0$ ,  $x_\varepsilon(\tau_\varepsilon) = x_1$  it follows that  $x^*$  is admissible. Due to weak closure of  $\{u \in L^2(0, 1; \mathbb{R}^m) : |u(x)|_{\ell^\infty} \leq 1 \text{ a.e.}\}$  we have that  $u^*$  is admissible as well. Since

$$\lim_{\varepsilon \rightarrow 0^+} \tau_\varepsilon + \frac{\varepsilon}{2} \int_0^{\tau_\varepsilon} |u_\varepsilon|^2 = \tau^*,$$

the triple  $(\tau^*, u^*, x^*)$  is optimal for (P). By Proposition 2.1 and weak lower semi-continuity of norms

$$(2.14) \quad \limsup_{\varepsilon \rightarrow 0} |u_\varepsilon|_{L^2(0, \tau_\varepsilon)} \leq |u^*|_{L^2(0, \tau^*)} \leq \liminf_{\varepsilon \rightarrow 0} |u_\varepsilon|_{L^2(0, \tau_\varepsilon)}$$

and hence  $\lim_{\varepsilon \rightarrow 0} |u_\varepsilon|_{L^2(0, \tau_\varepsilon; \mathbb{R}^m)} = |u^*|_{L^2(0, \tau^*; \mathbb{R}^m)}$ . As a consequence  $u_\varepsilon$  and  $x_\varepsilon$  converges strongly in  $L^2(0, \tau_\varepsilon)$ , respectively  $W^{1,2}(0, \tau_\varepsilon; \mathbb{R}^n)$ , to  $u^*$  and  $x^*$ . Let  $\hat{u}$  denote another optimal control for (P) with  $|\hat{u}| < |u^*|$ . Then by (2.10) and (2.14)

$$\limsup_{\varepsilon \rightarrow 0} |u_\varepsilon|_{L^2(0, \tau_\varepsilon; \mathbb{R}^m)} \leq |\hat{u}|_{L^2(0, \tau^*; \mathbb{R}^m)} < |u^*|_{L^2(0, \tau^*; \mathbb{R}^m)} \leq \liminf_{\varepsilon \rightarrow 0} |u_\varepsilon|_{L^2(0, \tau_\varepsilon; \mathbb{R}^m)},$$

which is a contradiction. Consequently (P) has a minimal norm control and the claimed strong convergence properties hold.  $\square$

**Corollary 2.1.** *If (2.5) holds, then the solution  $u^*$  to (P) is unique, it is bang-bang, and  $u_\varepsilon \rightarrow u^*$  in  $L^2$  as  $\varepsilon \rightarrow 0^+$ .*

*Proof.* (2.5) implies that the solution to (P) is unique and it is bang-bang. The remainder of the corollary follows from Theorem 2.1.  $\square$

We turn to the optimality condition for  $(P_\varepsilon)$ . Let

$$(2.15) \quad \sigma_\varepsilon(s) \in \begin{cases} -1 & \text{if } s \leq -\varepsilon \\ \frac{s}{\varepsilon} & \text{if } |s| < \varepsilon \\ 1 & \text{if } s \geq \varepsilon. \end{cases}$$

If  $\sigma_\varepsilon$  is applied to a vector, then it acts coordinate-wise.

We shall use a controllability assumption which is stronger than controllability and weaker than normality.

(H1) There exists  $i^*$  such that  $(A, b_{i^*})$  is controllable.

**Theorem 2.2.** *Assume that (H1) holds and let  $(x_\varepsilon, u_\varepsilon, \tau_\varepsilon)$  be a solution of  $(P_\varepsilon)$ . If there exist an interval  $I_{i^*} \subset (0, 1)$ ,  $\delta > 0$ ,  $\eta > 0$  such that*

$$(2.16) \quad |(\hat{u}_\varepsilon)_{i^*}(t)|_{\ell^\infty} \leq 1 - \eta \text{ for a.e. } t \in I_{i^*},$$

then there exists an adjoint state  $p_\varepsilon$  such that

$$(2.17) \quad \begin{cases} \dot{x}_\varepsilon = Ax_\varepsilon + Bu_\varepsilon, x_\varepsilon(0) = x_0, x_\varepsilon(\tau_\varepsilon) = x_1 \\ -\dot{p}_\varepsilon = A^T p_\varepsilon \\ u_\varepsilon = -\sigma_\varepsilon(B^T p_\varepsilon) \\ 1 + \frac{\varepsilon}{2} |u_\varepsilon(\tau_\varepsilon)|_{\mathbb{R}^m}^2 + p_\varepsilon(\tau_\varepsilon)^T (Ax_\varepsilon(\tau_\varepsilon) + Bu_\varepsilon(\tau_\varepsilon)) = 0. \end{cases}$$

*Proof.* We use a Lagrange multiplier argument for the reparameterized formulation of  $(P_\varepsilon)$  which is given by

$$(2.18) \quad \begin{cases} \min_{\tau \geq 0} \int_0^1 (\tau + \frac{\tau\varepsilon}{2} |\hat{u}(t)|^2) dt \\ \text{subject to} \\ \frac{d}{dt} \hat{x}(t) = \tau(A\hat{x}(t) + B\hat{u}(t)), \hat{u} \in C, \hat{x}(0) = x_0, \hat{x}(1) = x_1, \end{cases}$$

where  $C = \{\hat{u} \in L^2(0, 1; \mathbb{R}^m) : |\hat{u}|_{\ell^\infty} \leq 1\}$ . Here  $(\hat{u}, \tau)$  are treated as independent and  $\hat{x}$  as dependent variable. Further  $\hat{u} \in C$  is considered

as explicit constraint and a Lagrange multiplier  $\mu_0$  is introduced for the constrained  $e(\hat{u}, \tau) = \hat{x}(1) - x_1 = 0$ . The resulting Lagrangian is

$$\mathcal{L}(\hat{u}, \tau, \mu_0) = \int_0^1 \left( \tau + \frac{\tau \varepsilon}{2} |\hat{u}(t)|^2 \right) dt + \mu_0^T (\hat{x}(1) - x_1),$$

where  $\hat{x}(1)$  is defined through the differential equation and the initial condition.

We now argue that  $e : C \times \mathbb{R} \subset L^2(0, 1; \mathbb{R}^m) \times \mathbb{R} \rightarrow \mathbb{R}^n$  satisfies the regular point condition in the sense of Maurer-Zowe [MZ, IK2]. Thus we have to verify that

$$(2.19) \quad 0 \in \text{int} \{e'(\hat{u}_\varepsilon, \tau_\varepsilon)((C - \hat{u}_\varepsilon) \times \mathbb{R})\},$$

where  $e'(\hat{u}_\varepsilon, \tau_\varepsilon)$  denotes the linearisation of  $e$  at  $(u_\varepsilon(\cdot, \tau_\varepsilon), \tau_\varepsilon)$ .

Considering  $e'(\hat{u}_\varepsilon, \tau_\varepsilon)$  in directions  $\delta\tau = 0$  and  $\delta u$  satisfying  $(\delta u)_i = 0$  for  $i \neq i^*$  and  $(\delta u)_{i^*} = 0$  in  $(0, 1) \setminus (\alpha, \alpha + \delta)$ , with  $(\alpha, \alpha + \delta) := I_{i^*}$ , we find

$$\begin{aligned} e'(\hat{u}_\varepsilon, \tau_\varepsilon)(\delta u - \hat{u}_\varepsilon, 0) &= \int_\alpha^{\alpha+\delta} e^{\tau A(1-t)} \tau b_{i^*} (\delta u(t) - \hat{u}_\varepsilon(t))_{i^*} \\ &= \int_0^\delta e^{\tau A(\delta-t)} \tilde{b}_{i^*} (\delta u(t + \alpha) - \hat{u}_\varepsilon(t + \alpha))_{i^*}, \end{aligned}$$

where  $\tilde{b}_{i^*} = \tau e^{\tau A(1-\delta-\alpha)} b_{i^*}$ . Then

$$e'(\hat{u}_\varepsilon, \tau_\varepsilon)(\delta u - \hat{u}_\varepsilon, 0) = \int_0^\delta e^{\tau A(\delta-t)} \tilde{b}_{i^*} (\delta \tilde{u}(t) - \tilde{u}_\varepsilon(t))_{i^*} dt,$$

where  $\delta \tilde{u}_{i^*}(t) = \delta u_{i^*}(t + \alpha)$ ,  $\tilde{u}_{\varepsilon, i^*}(t) = \hat{u}_{\varepsilon, i^*}(t + \alpha)$ . Note that by (2.16)

$$\{(\delta \tilde{u} - \tilde{u}_\varepsilon)_{i^*} : [0, \delta] \rightarrow \mathbb{R}^1 \mid |(\delta \tilde{u})_{i^*}| \leq 1\} \supset S := \{v : [0, \delta] \rightarrow \mathbb{R}^1, |v| \leq \frac{\eta}{2}\}.$$

Observe that controllability of  $(A, b_{i^*})$  implies that  $(A, \tilde{b}_{i^*})$  is controllable as well. Controllability of the single input system  $(A, \tilde{b}_{i^*})$  implies that

$$(2.20) \quad 0 \in \text{int} \left\{ \int_0^\delta e^{\tau A(\delta-t)} \tilde{b}_{i^*} v \mid v \in S \right\}.$$

In fact the set on the right of (2.20) contains 0 and it has nonempty interior, see e.g. [LM], page 77,133. Moreover, if 0 was a boundary point of this set, then the corresponding control  $v = 0$  is an extremal control, which is impossible, e.g. [LM], page 133. Now (2.20) implies (2.19).

With the regular point condition satisfied, we can conclude the stationarity properties

$$(2.21) \quad \begin{aligned} \mathcal{L}_\tau(\hat{u}_\varepsilon, \tau_\varepsilon, \mu_0) &= 0, \\ \mathcal{L}_u(\hat{u}_\varepsilon, \tau_\varepsilon, \mu_0)(\delta u - \hat{u}_\varepsilon) &\geq 0 \text{ for all } \delta u \text{ with } |\delta u|_{\ell^\infty} \leq 1. \end{aligned}$$

From the second property in (2.21) we have,

$$\int_0^1 (\varepsilon \hat{u}_\varepsilon + B^T e^{\tau A^T(1-t)} \mu_0)(\delta u - \hat{u}_\varepsilon) \geq 0.$$

Setting

$$(2.22) \quad p(t) = e^{-\tau A^T t} q \text{ with } q = e^{\tau A^T} \mu_0,$$

this implies

$$\int_0^1 (\varepsilon \hat{u}_\varepsilon + B^T \hat{p}_\varepsilon)(\delta u - \hat{u}_\varepsilon) \geq 0,$$

for all  $\delta u$  as in (2.21). The second and third claim in (2.17) follow with  $p_\varepsilon(t) = \hat{p}_\varepsilon(\tau^{-1} t)$ .

Exploiting the first property in (2.21) implies that

$$\begin{aligned} \mathcal{L}_\tau(\hat{u}_\varepsilon, \tau_\varepsilon, \mu_0) &= 1 + \frac{\varepsilon}{2} \int_0^1 |\hat{u}_\varepsilon|^2 \\ &\quad + \mu_0^T \left( A e^{\tau A} x_0 + \int_0^1 e^{\tau A(1-t)} B \hat{u}_\varepsilon(t) + \int_0^1 A(1-t) e^{\tau A(1-t)} \tau B \hat{u}_\varepsilon(t) \right) \\ &= 1 + \frac{\varepsilon}{2} \int_0^1 |\hat{u}_\varepsilon|^2 \\ &\quad + \mu_0^T \left( A e^{(\tau(1-t)+\tau t)A} x_0 + \int_0^1 e^{\tau A(1-t)} B \hat{u}_\varepsilon(t) + \int_0^1 A e^{\tau A(1-t)} \int_0^t e^{\tau A(t-s)} \tau B \hat{u}_\varepsilon(s) \right) = 0. \end{aligned}$$

This implies

$$(2.23) \quad \mathcal{L}_\tau(\hat{u}_\varepsilon, \tau_\varepsilon, \mu_0) = 1 + \frac{\varepsilon}{2} \int_0^1 |\hat{u}_\varepsilon|^2 + \int_0^1 p^T(t) (A \hat{x}_\varepsilon(t) + B \hat{u}_\varepsilon(t)) = 0.$$

From  $u_\varepsilon = -\sigma_\varepsilon(B^T p_\varepsilon)$  we conclude that  $u_\varepsilon \in W^{1,\infty}(0, \tau_\varepsilon; \mathbb{R}^m)$ .

We introduce the Hamiltonian for  $(P_\varepsilon)$  as

$$H_\varepsilon(x, u, p) = 1 + \frac{\varepsilon}{2} |u|_{\mathbb{R}^m}^2 + p^T (Ax + Bu).$$

It is constant along the optimal solution. In fact we have almost everywhere on  $(0, 1)$

$$\begin{aligned} \frac{d}{dt} H_\varepsilon(\hat{x}_\varepsilon, \hat{u}_\varepsilon, \hat{p}_\varepsilon) &= \frac{\varepsilon}{2} \hat{u}_\varepsilon^T \frac{d}{dt} \hat{u}_\varepsilon + \frac{1}{\tau_\varepsilon} \frac{d}{dt} \hat{p}_\varepsilon^T \frac{d}{dt} \hat{x}_\varepsilon + \hat{p}_\varepsilon^T A \frac{d}{dt} \hat{x}_\varepsilon + \hat{p}_\varepsilon^T B \frac{d}{dt} \hat{u}_\varepsilon \\ &= \left( \frac{\varepsilon}{2} \hat{u}_\varepsilon + B^T \hat{p}_\varepsilon \right)^T \frac{d}{dt} \hat{u}_\varepsilon = 0. \end{aligned}$$

Combined with (2.23) this implies that

$$1 + \frac{\varepsilon}{2} |\hat{u}_\varepsilon|_{\mathbb{R}^m}^2 + \hat{p}_\varepsilon^T (A \hat{x}_\varepsilon + B \hat{u}_\varepsilon) = 0 \text{ on } [0, 1].$$

This implies the claim. □

The proof revealed extra regularity of  $u_\varepsilon$ :

**Corollary 2.2.** *Under the assumptions of Theorem 2.2 we have  $u_\varepsilon \in W^{1,\infty}(0, \tau_\varepsilon; \mathbb{R}^m)$ .*

**Remark 2.1.** Condition (2.16) requires that the modulus of at least one of the coordinates of  $\hat{u}_\varepsilon$  is not almost everywhere equal to 1. Once it is known from Corollary 2.1 that  $\hat{u}_\varepsilon$  is continuous this amounts to requiring that at least one of the coordinates of  $u^*$  switches from 1 to  $-1$  or vice versa.

The first order necessary optimality condition for  $(P_\varepsilon)$  after the transformation  $t \rightarrow \frac{t}{\tau}$  is given by

$$(2.24) \quad \begin{cases} \dot{x} = \tau(Ax + Bu), & x(0) = x_0, & x(1) = x_1 \\ -\dot{p} = \tau A^T p \\ u = -\sigma_\varepsilon(B^T p) \\ 1 + \frac{\varepsilon}{2} |u(1)|^2 + p(1)^T (Ax(1) + Bu(1)) = 0, \end{cases}$$

where for convenience of notation the dependence on  $\varepsilon$  and the superscript hat were dropped.

In the following section we shall investigate semi-smooth Newton methods for solving (2.24).

We close this section with a simple example which illustrates some of the features of the regularization approach.

**Example 2.1.** Consider the two-dimensional time optimal problem for the simple control system

$$\begin{cases} \dot{x}_1 = u_1 \\ \dot{x}_2 = u_2, \end{cases}$$

with so that  $A$  is the zero, and  $B$  the identity matrix, with initial condition  $(1, \frac{1}{2})$  and terminal condition the origin. This system is controllable but it is not normal. The optimal time is  $\tau^* = 1$ , the first coordinate of an optimal control is uniquely determined  $u_1^* = -1$ , with associate state  $x_1 = 1 - t$ . There are infinitely many choices for an optimal solutions  $u_2^*$  of bang-bang and non bang-bang type. The associated constant adjoints are  $(p_1, p_2) = (1, 0)$ . They satisfy

$$u = -\sigma(p) \in \begin{pmatrix} -1 \\ [-1, 1] \end{pmatrix}.$$

The transversality condition  $1 + p^T B u = 0$  is satisfied.

For the regularized problem we find  $\tau_\varepsilon = 1$ . Differently from the unregularized problem the solution to the regularized problem is unique. The optimal control and trajectory are given by

$$(u_1, u_2) = (-1, -.5) \text{ with } (x_1, x_2) = (1 - t, .5(1 - t)).$$

In this particular example the solution of the regularized problem does not depend on  $\varepsilon$ . Note that this solution is also one of the minimum norm solutions of the unregularized problem. The adjoint is  $p_\varepsilon = (1 + \frac{3\varepsilon}{8}, \frac{\varepsilon}{2})$ . It satisfies

$$u = -\sigma_\varepsilon(p_\varepsilon^T) = \begin{pmatrix} -1 \\ -\frac{1}{2} \end{pmatrix}$$

and the transversality condition  $1 + \frac{\varepsilon}{2}|u|^2 + p_\varepsilon^T B u = 0$ .

### 3 Semi-smooth Newton method

In this section the semi-smooth Newton method for solving the regularized optimality system (2.24) is described and analyzed. It will allow that (2.24) can be solved efficiently inspite of the fact that  $\sigma_\varepsilon$  is not differentiable.



Throughout we fix  $\varepsilon > 0$  and denote by  $(x_\varepsilon, u_\varepsilon, \tau_\varepsilon) \in W^{1,2}(0, 1) \times L^2(0, 1) \times \mathbb{R}$  a solution to  $(P_\varepsilon)$  with associated adjoint  $p_\varepsilon \in W^{1,2}(0, 1)$ . It is assumed that

$$(H2) \quad \text{there exists } \bar{s} \in (0, 1) \text{ such that } \left| \frac{1}{\varepsilon} b_{i^*}^T p_\varepsilon(\bar{s}) \right| = |(u_\varepsilon)_{i^*}(\bar{s})| < 1,$$

and

$$(H3) \quad |b_i^T p_\varepsilon(1)| \neq \varepsilon, \text{ for all } i = 1, \dots, m.$$

Assumption (H2) corresponds to (2.16), where we now use the fact that  $u_\varepsilon$  is continuous. With (H2) and (H3) holding there exists a neighborhood  $\mathcal{U}_{p_\varepsilon}$  of  $p_\varepsilon$  in  $W^{1,2}(0, 1; \mathbb{R}^n)$ ,  $\bar{t} \in (0, 1)$ , and a nontrivial interval  $(\alpha, \alpha + \delta) \subset (0, 1)$  such that for  $p \in \mathcal{U}_{p_\varepsilon}$  we have

$$|b_i^T p(t)| \neq \varepsilon \text{ for all } t \in [\bar{t}, 1], \text{ and } i = 1, \dots, m$$

and

$$(3.1) \quad |b_{i^*}^T p(t)| < \varepsilon \text{ for } t \in (\alpha, \alpha + \delta).$$

We set  $U = \{u \in L^2(0, 1; \mathbb{R}^m) : u|_{[\bar{t}, 1]} \in W^{1,2}(\bar{t}, 1; \mathbb{R}^m)\}$  endowed with the norm

$$|u|_U = (|u|_{L^2(0,1)}^2 + |\dot{u}|_{L^2(\bar{t},1)}^2)^{\frac{1}{2}},$$

and introduce

$$F : D_F \subset X \rightarrow L^2(0, 1; \mathbb{R}^n) \times L^2(0, 1; \mathbb{R}^n) \times U \times \mathbb{R} \times \mathbb{R}$$

where

$$D_F = W^{1,2}(0, 1) \times \mathcal{U}_{p_\varepsilon} \times U \times \mathbb{R},$$

$$X = W^{1,2}(0, 1; \mathbb{R}^n) \times W^{1,2}(0, 1; \mathbb{R}^n) \times U \times \mathbb{R},$$

and

$$(3.2) \quad F(x, p, u, \tau) = \begin{pmatrix} \dot{x} - \tau Ax - \tau Bu \\ -\dot{p} - \tau A^T p \\ u + \sigma_\varepsilon(B^T p) \\ x(1) - x_1 \\ 1 + \frac{\varepsilon}{2}|u(1)|^2 + p(1)^T (Ax(1) + Bu(1)) \end{pmatrix}.$$

Note that  $F = (F_1, \dots, F_5)$  is well-defined. This is obvious for  $F_1, F_2$  and  $F_3$ . For  $F_4, F_5$  it follows from the fact that  $W^{1,2}(0, 1)$  embeds continuously into  $C(0, 1)$ . Moreover  $F(x_\varepsilon, p_\varepsilon, u_\varepsilon, \tau_\varepsilon) = 0$ . We shall keep  $x_\varepsilon(0) = x_0$  as explicit constraint.

**Remark 3.1.** The need for introducing  $U$  in such a way that its elements are more regular at 1 is due to the fact that we use here the point-wise transversality condition rather than the integrated form (2.23). - Concerning (H3) note that  $F_3$  is welldefined with values in  $U$  also without assumption, but extra regularity of the mapping  $p \rightarrow \sigma_\varepsilon(B^T p)$  will be needed to prove superlinearity of the Newton step. For this purpose (H3) will be required.

Applying Newton's method to  $F = 0$  is impeded by the non-differentiability of  $\sigma_\varepsilon$ . We use

$$(3.3) \quad G\sigma_\varepsilon(s) := \begin{cases} \frac{1}{\varepsilon} & \text{if } |s| < \varepsilon \\ 0 & \text{if } |s| \geq \varepsilon \end{cases}$$

as a generalized derivative and argue that the resulting Newton iteration is semi-smooth and hence locally superlinearly convergent. The Newton iteration step is given by

$$(3.4) \quad DF(x, p, u, \tau)(\delta x, \delta p, \delta u, \delta \tau) = -F(x, p, u, \tau)$$

where  $\delta x(0) = 0$  and  $DF$  denotes the Frechet-derivative in all terms of  $F$  except for  $p \rightarrow \sigma_\varepsilon(B^T p)$ , for which the generalized derivative is taken according to (3.3). For further reference we give the detailed form of (3.4):

$$(3.5) \quad \begin{cases} \frac{d}{dt}\delta x - \tau A\delta x - \tau B\delta u - \delta\tau(Ax + Bu) = -F_1, & \delta x(0) = 0 \\ -\frac{d}{dt}\delta p - \tau A^T\delta p - \delta\tau A^T p = -F_2 \\ \delta u + G\sigma_\varepsilon(B^T p)B^T\delta p = -F_3 \\ \delta x(1) = -F_4 \\ p(1)^T(-AF_4 + B\delta u(1)) + \delta p(1)^T(Ax(1) + Bu(1)) \\ \quad + \varepsilon u(1)^T\delta u(1) = -F_5, \end{cases}$$

where the coordinates of  $G\sigma_\varepsilon(B^T p)B^T\delta p$  are given by  $G\sigma_\varepsilon((B^T p)_i)(B^T\delta p)_i$ .

A possible initialization may consist in choosing  $((u)_0, \tau_0)$ , setting  $(x)_0$  as the linear interpolation between  $x_0$  and  $x_1$ , and determining  $(p)_0$  such that the transversality condition and the adjoint equation are satisfied.

We now briefly summarize those facts from semi-smooth Newton methods which are relevant for this paper. Let  $X$  and  $Z$  be Banach spaces and let  $F : D_F \subset X \rightarrow Z$  be a nonlinear mapping with open domain  $D_F$ .

**Definition 3.1.** *The mapping  $F : D_F \subset X \rightarrow Z$  is called Newton-differentiable on an open subset  $\mathcal{U} \subset D_F$ , if for each  $\mathbf{x} \in \mathcal{U}$  there exists a generalized derivative  $DF(\mathbf{x}) \in \mathcal{L}(X, Z)$  and*

$$(3.6) \quad \lim_{h \rightarrow 0} \frac{1}{|h|_X} |F(\mathbf{x} + h) - F(\mathbf{x}) - DF(\mathbf{x})h|_Z = 0.$$

**Theorem 3.1.** *Suppose that  $\mathbf{x}^* \in \mathcal{U}$  is a solution to  $F(\mathbf{x}) = 0$  and that  $F$  is Newton-differentiable in an open set  $\mathcal{U}$  containing  $\mathbf{x}^*$ . If further  $\{\|DF(\mathbf{x})^{-1}\| : \mathbf{x} \in \mathcal{U}\}$  is bounded, then the Newton-iteration  $\mathbf{x}_{k+1} = \mathbf{x}_k - DF(\mathbf{x}_k)^{-1} F(\mathbf{x}_k)$  converges  $q$ -superlinearly to  $\mathbf{x}^*$ , provided that  $|\mathbf{x}_0 - \mathbf{x}^*|_X$  is sufficiently small.*

For the statement and proof of superlinear convergence of the time-optimal control problem, some further notation is required. For  $(x, p, u, \tau) \in D_F$  we define  $\mathcal{A} \in \mathbb{R}^{(n+1) \times (n+1)}$  by

$$\mathcal{A} = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & 0 \end{pmatrix},$$

where

$$(3.7) \quad A_{11} = \varepsilon^{-1} \tau \int_0^1 e^{\tau A(1-t)} B \chi_I B^T e^{\tau A^T(1-t)} \in \mathbb{R}^{n \times n},$$

$$(3.8) \quad \begin{aligned} A_{12} &= \varepsilon^{-1} \tau \int_0^1 e^{\tau A(1-t)} B \chi_I B^T \int_t^1 e^{-\tau A^T(t-s)} A^T p(s) ds dt \\ &\quad - \int_0^1 e^{\tau A(1-t)} (Ax + Bu) dt \in \mathbb{R}^n, \end{aligned}$$

$$(3.9) \quad A_{21} = (Ax(1) + Bu(1))^T - (p^T(1)B + \varepsilon u^T(1)) G \sigma_\varepsilon (B^T p(1)) B^T \in (\mathbb{R}^n)^T,$$

with  $\chi_I = \text{diag}(\chi_{I_1}, \dots, \chi_{I_m})$  and  $\chi_{I_i}$  the characteristic function of the set

$$I_i = I_i(p) = \{t : |(B^T p)_i| < \frac{1}{\varepsilon}\}, \quad i = 1, \dots, m$$

which is nonempty for  $p \in \mathcal{U}_{p_\varepsilon}$  and  $i = i^*$ . The controllability assumption (H1) together with (H2) imply that the symmetric matrix  $A_{11}$  is invertible with uniformly bounded inverse with respect to  $p \in \mathcal{U}_{p_\varepsilon}$  and  $\tau$  in compact subsets of  $(0, \infty)$ . In fact, since  $I_{i^*}(p) \supset (\alpha, \alpha + \delta)$  we obtain for some  $\bar{c} > 0$

$$\begin{aligned} A_{11} &= \varepsilon^{-1} \tau \int_0^1 e^{\tau A(1-t)} \sum_{i=1}^m b_i \chi_{I_i} b_i^T e^{\tau A^T(1-t)} \\ &\geq \varepsilon^{-1} \tau \int_\alpha^{\alpha+\delta} e^{\tau A(1-t)} b_{i^*} b_{i^*}^T e^{\tau A^T(1-t)} \\ &= \varepsilon^{-1} \tau e^{\tau A(1-\alpha)} \int_0^\delta e^{-\tau A t} b_{i^*} b_{i^*}^T e^{-\tau A^T t} dt e^{\tau A^T(1-\alpha)} > \bar{c}, \end{aligned}$$

where we use that the controllability Grammian

$$\int_0^\delta e^{-\tau A t} b_{i^*} b_{i^*}^T e^{-\tau A^T t} dt,$$

is positive definite for  $\tau$  in compact subsets of  $(0, \infty)$ .

For our analysis we shall utilize the fact that the Schur complement

$$A_{21} A_{11}^{-1} A_{12} \in \mathbb{R}$$

of  $\mathcal{A}$  for  $(x, p, u, \tau)$  in a neighborhood of  $(x_\varepsilon, p_\varepsilon, u_\varepsilon, \tau_\varepsilon)$  is nontrivial. If  $A_{21} A_{11}^{-1} A_{12} \neq 0$  at  $(x_\varepsilon, p_\varepsilon, u_\varepsilon, \tau_\varepsilon)$ , we cannot conclude that  $A_{21} A_{11}^{-1} A_{12} \neq 0$  for  $(x, p, u, \tau)$  in a neighborhood of  $(x_\varepsilon, p_\varepsilon, u_\varepsilon, \tau_\varepsilon)$ , since, while with (H2) holding,  $A_{21}$  is continuous with respect to  $(x, p, u, \tau) \in X$ , this is not the case for  $A_{11}^{-1}$  and  $A_{12}$  due to the term  $\chi_I$ . We therefore assume that

$$(H4) \quad \begin{cases} \text{there exists a bounded neighborhood} \\ \mathcal{U} \subset D_F \subset X \text{ of } (x_\varepsilon, p_\varepsilon, u_\varepsilon, \tau_\varepsilon) \text{ and } c > 0 \text{ such that} \\ |A_{21} A_{11}^{-1} A_{12}| \geq c \text{ for all } (x, p, u, \tau) \in \mathcal{U}. \end{cases}$$

**Theorem 3.2.** *If (H1)–(H4) hold and  $(x_\varepsilon, u_\varepsilon, \tau_\varepsilon)$  denotes a solution to  $(P_\varepsilon)$  with associated adjoint  $p_\varepsilon$ , then the semi-smooth Newton algorithm converges superlinearly, provided that the initialization is sufficiently close to  $(x_\varepsilon, u_\varepsilon, \tau_\varepsilon)$ .*

For the proof we require the following lemma.

**Lemma 3.1.** *Suppose that (H1) – (H4) hold. Then there exists a constant  $C$  such that for every  $(x, p, u, \tau) \in \mathcal{U}$ , and  $F \in L^2(0, 1) \times L^2(0, 1) \times U \times \mathbb{R}$*

$$DF(x, p, u, \tau)(\delta x, \delta p, \delta u, \delta \tau) = -F$$

*admits a unique solution  $(\delta x, \delta p, \delta u, \delta \tau) \in X$  and*

$$(3.10) \quad |(\delta x, \delta p, \delta u, \delta \tau)|_X \leq C |F|_{L^2 \times L^2 \times U \times \mathbb{R}}.$$

*Proof.* Let  $(x, p, u, \tau) \in \mathcal{U}$  and note that system (3.5) is equivalent to

$$(3.11) \quad \begin{cases} \delta x(t) = \int_0^t e^{\tau A(t-s)} (-F_1 + \tau b \delta u + \delta \tau (Ax + Bu)) ds \\ \delta p(t) = e^{\tau A^T(1-t)} \delta p(1) + \int_t^1 e^{-\tau A^T(t-s)} (\delta \tau A^T p - F_2) ds \\ \delta u + G \sigma_\varepsilon(B^T p) B^T \delta p = -F_3 \\ \delta x(1) = -F_4 \\ p(1)^T (-A F_4 + B \delta u(1)) + \delta p(1)^T (Ax(1) + Bu(1)) + \varepsilon \delta u(1) = -F_5. \end{cases}$$

On  $I_i$  we have  $(G \sigma_\varepsilon(B^T p))_i = \varepsilon^{-1}$ . The third equation in (3.11) can be expressed as

$$(3.12) \quad \delta u = -F_3 - \varepsilon^{-1} \chi_I B^T \delta p \quad \text{a.e. in } (0, 1).$$

Let us set

$$\hat{F}_1 = - \int_0^1 e^{\tau A(1-t)} F_1(t) dt, \quad \hat{F}_2 = \varepsilon^{-1} \tau \int_0^1 \int_t^1 e^{\tau A(1-t)} B \chi_I B^T e^{-\tau A^T(t-s)} F_2(s) ds dt,$$

$$\hat{F}_3 = -\tau B \int_0^1 e^{\tau A(1-t)} F_3(t) dt.$$

From (3.12), and the first and fourth equation in (3.11) we have

$$(3.13) \quad \begin{aligned} -F_4 &= \delta x(1) \\ &= -\varepsilon^{-1} \tau \int_0^1 e^{\tau A(1-t)} B \chi_I B^T \delta p + \delta \tau \int_0^1 e^{\tau A(1-t)} (Ax + Bu) dt + \hat{F}_1 + \hat{F}_3 \end{aligned}$$

Replacing  $\delta p$  by the second equation in (3.11) we find

$$\begin{aligned} -F_4 &= -\varepsilon^{-1} \tau \int_0^1 e^{\tau A(1-t)} B \chi_I B^T e^{\tau A^T(1-t)} \delta p(1) \\ &\quad - \varepsilon^{-1} \tau \delta \tau \int_0^1 e^{\tau A(1-t)} B \chi_I B^T \int_t^1 e^{-\tau A^T(t-s)} A^T p(s) ds dt + \hat{F}_2 \\ &\quad + \delta \tau \int_0^1 e^{\tau A(1-t)} (Ax + Bu) dt + \hat{F}_1 + \hat{F}_3, \end{aligned}$$

which involves  $\delta p(1)$  and  $\delta \tau$  as unknowns, and can be expressed as

$$(3.14) \quad A_{11} \delta p(1) + A_{12} \delta \tau = \hat{F}_1 + \hat{F}_2 + \hat{F}_3 + F_4 =: r_1.$$

Eliminating  $\delta u(1)$  from the last equation in (3.11) by means of the third equation implies

$$(3.15) \quad A_{21} \delta p(1) = -F_5 + F_3(1)^T (B^T p(1) + \varepsilon u(1)) + p(1)^T A F_4 =: r_2.$$

Combining (3.14) and (3.15) we obtain the following linear system for  $(\delta p(1), \delta \tau)$  :

$$(3.16) \quad \mathcal{A} \begin{pmatrix} \delta p(1) \\ \delta \tau \end{pmatrix} = \begin{pmatrix} r_1 \\ r_2 \end{pmatrix}.$$

By (H1), (H2), and (H4) its unique solution is given by

$$(3.17) \quad \delta \tau = (A_{21} A_{11}^{-1} A_{12})^{-1} (A_{21} A_{11}^{-1} r_1 - r_2), \quad \delta p(1) = A_{11}^{-1} (r_1 - A_{12} \delta \tau).$$

Moreover there exists a constant  $C = C(\tau, |x|_{C(0,1)}, |p|_{C(0,1)}, |u|_{L^2(0,1)}, |u(1)|)$ , such that

$$|\delta p(1)| + |\delta \tau| \leq C |F|_{L^2 \times L^2 \times U \times \mathbb{R}}.$$

From (3.5) and (3.11),  $C$  can also be chosen such that

$$|(\delta x, \delta p, \delta u, \delta \tau)|_X \leq C |F|_{L^2 \times L^2 \times U \times \mathbb{R}}.$$

□

*Proof of Theorem 3.2.* We apply Theorem 3.1 with  $\mathbf{x}^* = (x_\varepsilon, p_\varepsilon, u_\varepsilon, \tau_\varepsilon)$ . Lemma 3.1 implies the required uniform bound of the generalized inverses  $DF$  in the neighborhood  $\mathcal{U} \subset D_F$  of  $(x_\varepsilon, p_\varepsilon, u_\varepsilon, \tau_\varepsilon)$ . Therefore it suffices to argue Newton-differentiability of  $F$  in  $D_F$ . This is obvious for all coordinates of  $F$  except for  $F_3$ , and specifically for the mapping

$$\mathcal{F} : p \rightarrow \sigma_\varepsilon(B^T p)$$

from  $\mathcal{U}_{p_\varepsilon} \subset W^{1,2}(0, 1; \mathbb{R}^n) \rightarrow U$ . Utilizing the definitions of  $\mathcal{U}_{p_\varepsilon}$  and  $\sigma_\varepsilon$  it suffices to consider the restriction of  $\mathcal{F}$  from  $W^{1,2}(0, \bar{t}; \mathbb{R}^n)$  to  $L^2(0, \bar{t}; \mathbb{R}^1)$  which we again denote by  $\mathcal{F}$ . Note that  $\mathcal{F}$  can be decomposed as

$$\mathcal{F} = \mathcal{F}_3 \circ \mathcal{F}_2 \circ \mathcal{F}_1,$$

where

$$\begin{aligned} \mathcal{F}_1 : W^{1,2}(0, \bar{t}; \mathbb{R}^n) &\rightarrow W^{1,2}(0, \bar{t}; \mathbb{R}), & \mathcal{F}_2 : W^{1,2}(0, \bar{t}; \mathbb{R}) &\rightarrow L^4(0, \bar{t}; \mathbb{R}), \\ \mathcal{F}_3 : L^4(0, \bar{t}; \mathbb{R}) &\rightarrow L^2(0, \bar{t}; \mathbb{R}), \end{aligned}$$

are given by

$$\mathcal{F}_1(u) = B^T u, \quad \mathcal{F}_2(v) = \max(-1, \frac{v}{\varepsilon}), \quad \mathcal{F}_3(v) = \min(1, v).$$

In [HIK, IK2] it was shown that  $v \rightarrow \max(0, v)$  is Newton differentiable from  $L^p(\Omega)$  to  $L^q(\Omega)$  if  $\infty \geq p > q \geq 1$ , if  $\Omega$  is a bounded domain with Lipschitz boundary. Since  $\min(1, v) = 1 + \min(0, v - 1)$  this implies that  $\mathcal{F}_3$  and similarly that  $\mathcal{F}_2$  are Newton differentiable. From the chain rule for Newton differentiable mapping in [HK] it follows that  $\mathcal{F}_3 \circ \mathcal{F}_2$  is Newton differentiable. The chain rule for a linear mapping, here  $\mathcal{F}_1$ , followed by the Newton differentiable mappings  $\mathcal{F}_3 \circ \mathcal{F}_2$ , [IK1], implies that  $\mathcal{F}$  is Newton differentiable in  $D_F$ . □

## 4 A numerical example

The semi-smooth Newton method is used to solve a classical time optimal problem related to the harmonic oscillator with three switching points. We consider

$$(4.1) \quad \begin{cases} \min_{\tau \geq 0} \int_0^\tau dt \\ \text{subject to} \\ \frac{d}{dt}x(t) = Ax(t) + Bu(t), |u(t)| \leq 1, x(0) = x_0, x(\tau) = x_1, \end{cases}$$

where

$$A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad x_0 = \begin{pmatrix} -5 \\ 5 \end{pmatrix}, \quad x_1 = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

The optimal minimal time for the continuous problem is known to be  $\tau^* = 10.5871$ . To solve (4.1) numerically a time discretization based on the Crank Nicolson method with equidistant grid points was applied to (3.5). The initialisation for the state was chosen as a semicircle connecting  $x_0$  and  $x_1$ . Then  $u(1)$  was chosen to be active, and  $p$  was chosen so that the transversality condition and the adjoint equation hold. With respect to the choice of the parameter  $c = \frac{1}{\varepsilon}$  we utilized a continuation procedure, starting with a small value and increasing it, using the solution from the smaller value of  $c$  as initialization for the next larger  $c$ -value. Certainly this procedure can be automated as has been done elsewhere, but this was not the focus of this paper. In Table 1 we show the number of iterates of the Newton iteration (outer loop) that was required for this continuation procedure with respect to  $c$ . The Newton iteration was stopped when the residual of the optimality system in the  $L^2$ -norm was below  $10^{-8}$ . Also in Table 1 we depict the optimal minimal times  $\tau^*(c)$ . These results are obtained for meshsize  $h = \frac{1}{32}$ . Then

c	1	5	10	20
No. of iterations	8	8	4	7
Final Time	11.26515	10.84455	10.82977	10.81781

Table 1

the results for  $c = 1$  are interpolated to the finer grid  $h = \frac{1}{128}$  and the



continuation procedure with respect to  $c$  is repeated. The results are depicted in Table 2. The graphs for the corresponding controls are given in Figure 1.

$c$	1	10	50	100	200
No. of iterations	5	46	4	4	3
Final Time	11.1088	10.6092	10.6034	10.6033	10.6031

Table 2

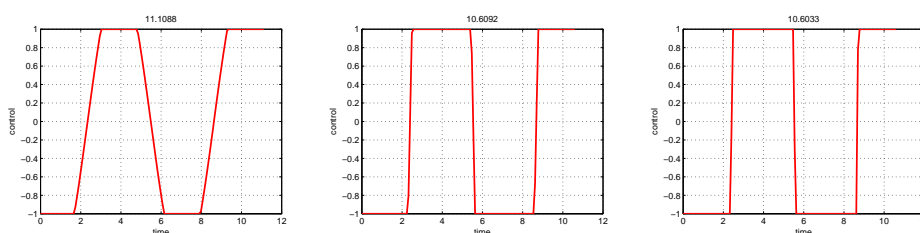


Figure 1:  $N = 128$  and  $c = 1$ (left),  $10$ (middle) and  $100$ (right)

The same procedure with  $h = 1/512$  and  $c = 100$  gives the optimal time 10.588. In some cases, typically at the beginning of the iterations and for the lowest values of  $c$  the full Newton step was too large. Therefore we used a one-dimensional line search based on a quadratic polynomial interpolation for the  $L^2$ - norm of the residual combined with an Armijo rule.

Table 3 depicts the quotients  $\frac{|u^{k+1} - u^*(c)|_{L^2}}{|u^k - u^*(c)|_{L^2}}$ , where  $u^*(c)$  is the solution to the discretized version of (2.17) for  $c = 50$ . It shows that the algorithm is in fact superlinearly convergent.

No. of iterations	1	2	3	4
$c_k$	0.94138	0.00037	0.00001	0.00000

Table 3

In this paper we chose to regularize  $\sigma$  by the ramp functions  $\sigma_\varepsilon$  with increasing slopes as  $\varepsilon \rightarrow 0^+$ . Certainly other alternatives are possible as for instance  $\tilde{\sigma}_c(s) = \frac{2}{\pi} \operatorname{atan}(cs)$ . This family of  $C^\infty$ - functions also has the property that it converges to  $\sigma$  as  $c \rightarrow 0$ , but it appears to be less apt for

the purpose of approximating the discontinuous switching structure of the optimal controls since  $c$  has to be taken significantly larger for  $\tilde{\sigma}_c$  than for  $\sigma_{\frac{1}{c}}$  to obtain comparable results.

*Acknowledgement:* We thank Mrs. J. Rubesa for providing us with the numerical example.

## References

- [BPW] Ch. Büskens, H.J. Pesch, and S. Winderl: Real-time solutions for bang-bang and singular optimal control problems, in: M. Grtschel, et al. (Eds), *Online Optimization of Large Scale Systems*, Springer-Verlag, Berlin, 2001, 129-142.
- [Fa] H.O. Fattorini: *Infinite Dimensional Linear Control Systems: The Time Optimal and Norm Optimal Problems*, Elsevier, Amsterdam, 2005.
- [HL] H. Hermes and J. LaSalle: *Functional Analysis and Time Optimal Control*, Academic Presse, 1969.
- [HK] M. Hintermüller and K. Kunisch: PDE-constrained optimization subject to pointwise control and zero- and first-order state constraints, preprint.
- [HIK] M. Hintermüller, K. Ito and K. Kunisch: The primal–dual active set strategy as a semi–smooth Newton method, *SIAM Journal on Optimization*, 13(2002), 865-888.
- [IK1] K. Ito and K. Kunisch: The primal-dual active set method for non-linear optimal control problems with bilateral constraints, *SIAM J. on Control and Optimization*, 43(2004), 357-376.
- [IK2] K. Ito and K. Kunisch: *On the Lagrange Multiplier Approach to Variational Problems and Applications*, SIAM, Philadelphia, 2008.
- [KN] C.Y. Kaya and J.L.Noakes: Computational methods for time-optimal switching controls, *JOTA*, 117(2003), 69-92.
- [Ke] H.B. Keller: *Numerical Methods for Two-Point Boundary Value Problems*: Blaisdell, London, 1968.

- [KLM] J.-H.R.Kim, G.L. Lippi, and H. Maurer: Minimizing the transition time in lasers by optimal control methods. Single-mode semiconductor lasers with homogenous transverse profile: *Physica D*, 191(2004), 238-260.
- [LM] E.B. Lee and L. Markus: *Foundations of Optimal Control Theory*: John Wiley & Sons, New York, 1967.
- [MS] J. Macki and A. Strauss: *Introduction to Optimal Control Theory*, Springer Verlag, New York, 1982.
- [MO] H. Maurer and N.P. Osmolovskii: Second order sufficient optimality conditions for time-optimal bang-bang control, *SIAM J. Control and Optimization*, 42(2004), 2239-2263.
- [MZ] H. Maurer and J. Zowe: *First and second-order necessary and sufficient optimality conditions for infinite-dimensional programming problems*, *Mathematical Programming* 16(1979), 98–110.
- [MB] E.Meier and A.E.Bryson: Efficient algorithms for time-optimal rigid spacecraft reorientation problem, *J. Guidance*, 13(1990), 859-866.
- [XK] L. Xie and K. Kunisch: Numerical methods for time-optimal control problem, SFB-Report, Dep. Math. University of Graz.