**Control and Cybernetics**

vol. **40** (2011) No. 4

# Sensitivity analysis for state constrained optimal control problems*

by

**Kazimierz Malanowski**

Systems Research Institute of the Polish Academy of Sciences
ul. Newelska 6, 01-447 Warszawa, Poland

**Abstract:** A sensitivity result for cone-constrained optimization problem in abstract Hilbert spaces is obtained, using a slight modification of Haraux's theorem on differentiability of the metric projection onto polyhedric sets. This result is applied to sensitivity analysis for nonlinear optimal control problems subject to first order state constraints.

**Keywords:** sensitivity analysis, cone-constrained optimization problems, nonlinear optimal control, first order state constraints.

## 1. Introduction

In sensitivity analysis of optimization and optimal control problems local differentiability properties of the solutions treated as functions of a parameter are studied. The presence of inequality type constraints introduces an element of nonsmoothness. Therefore, to obtain continuous differentiability strong regularity assumptions must be imposed (see e.g., Bonnans and Hermant, 2008). Under weaker conditions one can expect only directional differentiability. The classical implicit function theorem cannot be applied here and the sensitivity analysis is performed in two steps. First, Lipschitz continuity of the solutions with respect to the parameter is proved and after that the limit of the difference quotient is investigated. In this second step the differentiability results for metric projection onto so called *polyhedric sets* (see Haraux, 1977, and Mignot, 1976) turned out to be useful. This approach was applied, in particular, to nonlinear optimal control problems subject to first order state constraints (see Malanowski, 1995). Recently, stability results for such optimal control problems were obtained under weakened second order condition, where coercivity of the Hessian of Lagrangian was assumed to be satisfied on the critical subspace, rather than on the whole space. This condition is too weak to allow a direct application of the differentiability results for the metric projection.

*Submitted: December 2010; Accepted: April 2011

In this paper we introduce a slight modification of the original Haraux's theorem, which allows for getting differentiability results under the weakened assumptions. The organization of the paper is the following.

In Section 2 we consider an abstract cone constrained optimization problem, depending on a parameter. The formulation of the problem enables overcoming the known difficulty connected with the so called *two-norm discrepancy*, which is typical for nonlinear optimal control problems. We find conditions under which the solutions of the problem are directionally differentiable at the reference value of the parameter. Polyhedricity of the cone is one of the basic assumptions. In Section 3 we formulate a simple model nonlinear optimal control problem subject to first order state constraints. The stability results for that problem were obtained in Malanowski (2007b). Applying the methodology of Section 2 and using the stability results of Malanowski (2007b), we show in Section 4 that the same weakened assumptions under which the solutions are Lipschitz stable ensure also directional differentiability. One should stress that, since the polyhedricity of the cone is one of the basic assumptions, the presented methodology is not applicable to optimal control problems with higher order state constraints.

## 2. Cone constrained optimization problems

Let $X, Y$ and $H$ be Hilbert spaces of arguments, constraints and parameters, respectively. In the space $Y$ there is given a closed convex cone $\mathcal{K}$. Furthermore, $\mathcal{G}: X \times H \to \mathbb{R}$ and $\phi: X \times H \to Y$ are given functions. Consider a family of the following optimization problems depending on the parameter $h$:

$$(\mathrm{P})_h \quad \min_{\xi \in X} \mathcal{G}(\xi, h) \qquad \text{subject to} \quad \phi(\xi, h) \in \mathcal{K}.$$

Assume:

(B1) For each $h \in H$ the functions $\mathcal{G}(\cdot, h)$ and $\phi(\cdot, h)$ are Fréchet differentiable on $X$.

Let us introduce the following standard Lagrangian for $(\mathrm{P})_h$:

$$L: X \times \mathcal{K}^+ \times H \to \mathbb{R}, \qquad L(\xi, \lambda, h) := \mathcal{G}(\xi, h) + (\lambda, \phi(\xi, h)), \tag{1}$$

where $\mathcal{K}^+ := \{\lambda \in Y^* \mid (\lambda, \xi) \leq 0 \text{ for all } \xi \in \mathcal{K}\}$ is the cone polar to $\mathcal{K}$.

The stationarity conditions for the Lagrangian can be written in the form

$$\begin{aligned} &D_\xi L(\xi, \lambda, h) := D_\xi \mathcal{G}(\xi, h) + D_\xi \phi^*(\xi, h)\lambda = 0, \\ &\phi(\xi, h) \in \mathcal{N}_{\mathcal{K}^+}(\lambda), \end{aligned} \tag{2}$$

where

$$\mathcal{N}_{\mathcal{K}^+}(\lambda) := \left\{ y \in Y \;\middle|\; \left\{ \begin{array}{lll} (\mu - \lambda, y) \leq 0 & \forall \mu \in \mathcal{K}^+ & \text{if } \lambda \in \mathcal{K}^+ \\ \emptyset & & \text{if } \lambda \notin \mathcal{K}^+ \end{array} \right. \right\}$$

is the cone normal to $\mathcal{K}^+$ at $\lambda$. For the sake of simplicity denote:

$$Z = X \times Y^*, \quad \zeta = (\xi, \lambda), \quad \mathcal{F} : Z \times H \to Z^*, \quad \mathcal{T} : Z \to 2^{Z^*},$$
$$\mathcal{F}(\zeta, h) = \begin{bmatrix} D_\xi L(\xi, \lambda, h) \\ \phi(\xi, h) \end{bmatrix}, \qquad \mathcal{T}(\zeta) = \begin{bmatrix} 0 \\ \mathcal{N}_{\mathcal{K}^+}(\lambda) \end{bmatrix}. \tag{3}$$

Then, (2) can be written in the form of the following inclusion (generalized equation):

$$\mathcal{F}(\zeta, h) \in \mathcal{T}(\zeta). \tag{4}$$

Solutions of (4) are called stationary points of $(\mathrm{P})_h$. In sensitivity analysis for $(\mathrm{P})_h$ we investigate conditions under which, in a neighborhood of a reference value $\widehat{h}$ of the parameter, there exists a locally unique stationary point of $(\mathrm{P})_h$, which is directionally differentiable function of $h$. Such an analysis is usually performed in two steps. First we prove that $\zeta(\cdot)$ is a Lipschitz continuous function of $h$ ( stability analysis ) and in the next step we pass to differentiability.

The main tool in stability analysis is Robinson's implicit function theorem for the so called *strongly regular generalized equations* (see Robinson, 1980 and 1991). Unfortunately, due to the phenomenon of *two-norm discrepancy* (see e.g., Maurer, 1981), Robinson's theorem cannot be directly applied to nonlinear optimal control problems subject to state constraints. However, a modification of that theorem due to A.L. Dontchev and W.W. Hager (Theorem 2.2 in Dontchev and Hager, 1998) can be used. In that theorem equation (4) is considered not on the whole space $Z$ but on a closed convex subset $\mathcal{Z} \subset Z$, which supplied with the metric induced by the norm of $Z$, can be treated as a complete nonlinear metric space. To formulate that theorem we assume:

(B2) There exists a closed convex set $\mathcal{Z} = \Xi \times \Lambda \subset Z$. For each $h \in H$, $\mathcal{G}(\cdot, h)$ and $\phi(\cdot, h)$, as well as $D_\zeta \mathcal{G}(\cdot, h)$ and $D_\zeta \phi(\cdot, h)$, are Fréchet differentiable on $\mathcal{Z}$, with the differentials continuous for all sequences $\{\zeta_n\} \subset \mathcal{Z}$ and $\{h_n\} \subset H$. For each $\zeta \in \mathcal{Z}$, $\mathcal{G}(\zeta, \cdot)$ and $\phi(\zeta, \cdot)$, as well as $D_\zeta \mathcal{G}(\zeta, \cdot)$ and $D_\zeta \phi(\zeta, \cdot)$, are Lipschitz continuous with modulus $l > 0$ and directionally differentiable in $h$.

(B3) For a reference value $\widehat{h}$ of the parameter there exists a solution $\widehat{\xi} := \xi(\widehat{h})$ of $(\mathrm{P})_{\widehat{h}}$ and an associated Lagrange multiplier $\widehat{\lambda} := \lambda(\widehat{h})$, such that $\widehat{\zeta} := (\widehat{\xi}, \widehat{\lambda})$ satisfies (4).

(B4) Define
$$\begin{aligned} \mathcal{A}(\zeta) &:= \mathcal{F}(\widehat{\zeta}, \widehat{h}) + D_\zeta \mathcal{F}(\widehat{\zeta}, \widehat{h})(\zeta - \widehat{\zeta}), \\ \Psi^h(\zeta) &:= \mathcal{F}(\zeta, h) - \mathcal{A}(\zeta). \end{aligned} \tag{5}$$
There exists a subset $\Delta \subset Z^*$ and a function $\eta : \Delta \to \mathcal{Z}$, such that
$$\delta + \mathcal{A}(\eta(\delta)) \in \mathcal{T}(\eta(\delta)) \qquad \text{for all } \delta \in \Delta, \tag{6}$$
and $\eta(\cdot)$ is Lipschitz continuous with the modulus $\ell > 0$.

(B5) There exist neighborhoods $\mathcal{O}_{\mathcal{Z}}(\widehat{\zeta})$ and $\mathcal{O}_H(\widehat{h})$ of $\widehat{\zeta}$ and $\widehat{h}$, in the spaces $\mathcal{Z}$ and $H$, respectively, such that $(\mathcal{F} - \mathcal{A})(\cdot, \cdot)$ maps $\mathcal{O}_{\mathcal{Z}}(\widehat{\zeta}) \times \mathcal{O}_H(\widehat{h})$ into $\Delta$.

THEOREM 1 *Let conditions* (B1) *and* (B5) *be satisfied. Then for any* $\ell^+ > \ell$ *there exist neighborhoods* $\widehat{\mathcal{O}}_{\mathcal{Z}}(\widehat{\zeta})$ *and* $\widehat{\mathcal{O}}_H(\widehat{h})$ *of* $\widehat{\zeta}$ *and* $\widehat{h}$*, as well as a function* $\zeta(\cdot):\widehat{\mathcal{O}}_H(\widehat{h}) \to \widehat{\mathcal{O}}_{\mathcal{Z}}(\widehat{\zeta})$*, such that*

$$\mathcal{F}(\zeta(h),h) \in \mathcal{T}(\zeta(h)) \qquad \text{for all } h \in \widehat{\mathcal{O}}_H(\widehat{h}). \tag{7}$$

*The function* $\zeta(\cdot)$ *is Lipschitz continuous with modulus* $l\ell^+$ *and satisfies the equation*

$$\zeta(h) = \eta(\Psi^h(\zeta(h))) := \eta(\mathcal{F}(\zeta(h),h) - \mathcal{A}(\zeta(h))). \tag{8}$$

Theorem 1 implies the following corollary (for similar results see Theorem 2.3 in Robinson, 1980, and Theorem 2.4 in Dontchev, 1995).

COROLLARY 1 *Let assumptions of Theorem* 1 *be satisfied. If the function* $\eta(\cdot)$ *is directionally differentiable at* $\delta = 0$*, with the differential* $d_\delta\eta(0;\Delta\delta)$ *for any* $\Delta\delta \in \Delta$*, then* $\zeta(\cdot)$ *is directionally differentiable at* $\widehat{h}$ *with the differential*

$$d_h\zeta(\widehat{h};\Delta h) = d_\delta\eta(0; d_h\mathcal{F}(\widehat{\zeta},\widehat{h})\Delta h). \tag{9}$$

*Proof.* Define the function $\varsigma : H \to \mathcal{Z}$ by

$$\varsigma(h) = \eta(\Psi^h(\widehat{\zeta})) := \eta(\mathcal{F}(\widehat{\zeta},h) - \mathcal{A}(\widehat{\zeta})). \tag{10}$$

In view of (8) we have $\varsigma(\widehat{h}) = \widehat{\zeta}$. By (B4), (8) and (10) we get

$$\|\zeta(h) - \varsigma(h)\|_Z = \|\eta(\Psi^h(\zeta(h))) - \eta(\Psi^h(\widehat{\zeta}))\|_Z \le \ell\|\Psi^h(\zeta(h)) - \Psi^h(\widehat{\zeta})\|_{Z^*}. \tag{11}$$

Note that by (B2) and (5) we obtain

$$\begin{aligned}
&\|\Psi^h(\zeta(h)) - \Psi^h(\widehat{\zeta})\|_{Z^*} = \|\mathcal{F}(\zeta(h),h) - \mathcal{F}(\widehat{\zeta},h) - D_\zeta\mathcal{F}(\widehat{\zeta},\widehat{h})(\zeta(h)-\widehat{\zeta})\|_{Z^*}\\
&\le \|\mathcal{F}(\zeta(h),h) - \mathcal{F}(\widehat{\zeta},h) - D_\zeta\mathcal{F}(\widehat{\zeta},h)(\zeta(h)-\widehat{\zeta})\|_{Z^*}\\
&+\|(D_\zeta\mathcal{F}(\widehat{\zeta},\widehat{h}) - D_\zeta\mathcal{F}(\widehat{\zeta},h))(\zeta(h)-\widehat{\zeta})\|_{Z^*} \le \epsilon\|\zeta(h)-\widehat{\zeta}\|_Z\\
&\le \epsilon\, l\, \ell^+\|h-\widehat{h}\|_H,
\end{aligned} \tag{12}$$

where $\epsilon \to 0$ as $h \to \widehat{h}$. From (11) and (12) we obtain

$$\|\zeta(h) - \varsigma(h)\| \le \epsilon\ell\, l\, \ell^+\|h-\widehat{h}\|_H,$$

which implies

$$\lim_{\alpha\to+0}\frac{1}{\alpha}\|\zeta(\widehat{h}+\alpha\Delta h) - \widehat{\zeta}\|_Z = \lim_{\alpha\to+0}\frac{1}{\alpha}\|\varsigma(\widehat{h}+\alpha\Delta h) - \widehat{\varsigma}\|_Z$$

and, in view of (10), (9) follows. ∎

Let us come back to the original problem $(\mathrm{P})_h$. Using (3) and (5) we get

$$\mathcal{A}(\zeta) = \begin{bmatrix} D_\xi L(\widehat{\xi}, \widehat{\lambda}, \widehat{h}) \\ \phi(\widehat{\xi}, \widehat{h}) \end{bmatrix} + \begin{bmatrix} D^2_{\xi\xi} L(\widehat{\xi}, \widehat{\lambda}, \widehat{h}) & D_\xi \phi^*(\widehat{\xi}, \widehat{h}) \\ D_\xi \phi(\widehat{\xi}, \widehat{h}) & 0 \end{bmatrix} \begin{bmatrix} \xi - \widehat{\xi} \\ \lambda - \widehat{\lambda} \end{bmatrix}$$

and the linearized equation (6) takes the form

$$\begin{bmatrix} \delta^1 \\ \delta^2 \end{bmatrix} + \begin{bmatrix} \widehat{\delta}^1 \\ \widehat{\delta}^2 \end{bmatrix} + \begin{bmatrix} Q & \Phi^* \\ \Phi & 0 \end{bmatrix} \begin{bmatrix} \xi \\ \lambda \end{bmatrix} \in \begin{bmatrix} 0 \\ \mathcal{N}_{\mathcal{K}^+}(\lambda) \end{bmatrix}, \tag{13}$$

where $Q := D^2_{\xi\xi} L(\widehat{\xi}, \widehat{\lambda}, \widehat{h}), \Phi := D_\xi \phi(\widehat{\xi}, \widehat{h})$ and

$$\begin{bmatrix} \widehat{\delta}^1 \\ \widehat{\delta}^2 \end{bmatrix} = \begin{bmatrix} -D^2_{\xi\xi} L(\widehat{\xi}, \widehat{\lambda}, \widehat{h})\widehat{\xi} + D_\xi \mathcal{G}(\widehat{\xi}, \widehat{h}) \\ \phi(\widehat{\xi}, \widehat{h}) - D_\xi \phi(\widehat{\xi}, \widehat{h})\widehat{\xi} \end{bmatrix}. \tag{14}$$

Note that (13) can be interpreted as the first order optimality condition for the following linear-quadratic accessory problem for $(\mathrm{P})_h$:

$$(\mathrm{AP})_\delta \quad \min_{\eta \in X} \left\{ \frac{1}{2}(\eta, Q\eta) + (\delta^1 + \widehat{\delta}^1, \eta) \right\}, \quad \text{subject to} \quad \Phi\eta + \delta^2 + \widehat{\delta}^2 \in \mathcal{K}.$$

Thus, Theorem 1 and Corollary 1 allow to reduce the sensitivity analysis for the nonlinear problem $(\mathrm{P})_h$ to such an analysis for the linear-quadratic accessory problem $(\mathrm{AP})_\delta$. To show directional differentiability of the stationary points of $(\mathrm{AP})_\delta$ we introduce new variables, in which the constraints in $(\mathrm{AP})_\delta$ are independent of $\delta^2$. To this end we need the following assumption:
(B6) The map $\Phi : X \to Y$ is surjective.
Without loss of generality we can assume that $X^* = X$ and $Y^* = Y$. Then by (B6) the map $(\Phi\Phi^*) : Y \to Y$ is invertible. Let us define the new variable

$$\chi = \eta + \Phi^*(\Phi\Phi^*)^{-1}(\delta^2 + \widehat{\delta}^2). \tag{15}$$

In terms of $\chi$ problem $(\mathrm{AP})_\delta$ takes the form

$$(\mathrm{AP})'_\beta \qquad \min_{\chi \in X} \left\{ \frac{1}{2}(\chi, Q\chi) + (\beta + \widehat{\beta}, \chi) \right\}, \quad \text{subject to} \quad \Phi\chi \in \mathcal{K},$$

where

$$\beta = \delta^1 - 2Q\Phi^*(\Phi\Phi^*)^{-1}\delta^2, \quad \widehat{\beta} = \widehat{\delta}^1 - 2Q\Phi^*(\Phi\Phi^*)^{-1}\widehat{\delta}^2. \tag{16}$$

Define the set

$$\mathcal{B} := \{\beta = \delta^1 - 2Q\Phi^*(\Phi\Phi^*)^{-1}\delta^2 \mid \delta = (\delta^1, \delta^2) \in \Delta\}. \tag{17}$$

Optimality condition for $(\mathrm{AP})'_\beta$ can be expressed in the form

$$\begin{aligned} &(Q\chi_\beta + \beta + \widehat{\beta}, \chi - \chi_\beta) \geq 0 \qquad \text{for all feasible } \chi, \\ \text{or} \quad &Q\chi_\beta + \beta + \widehat{\beta} + \Phi^*\mu_\beta = 0, \end{aligned} \tag{18}$$

where $\mu_\beta \in \mathcal{K}^+$ is the Lagrange multiplier. Denote $\widehat{\chi} := \chi_0$, $\widehat{\mu} := \mu_0$. In view of (18)

$$[\widehat{\mu}]^\perp := \{y \in Y \mid (\widehat{\mu}, y) = 0\} = \{\Phi\chi \in Y \mid (Q\widehat{\chi} + \widehat{\beta}, \chi) = 0\}. \tag{19}$$

Define the sets

$$\left.\begin{array}{l} \Pi(\widehat{\chi}) = \bigcup_{r>0} r(\mathcal{K} - \Phi\widehat{\chi}), \\ \Gamma(\widehat{\chi}) = \Pi(\widehat{\chi}) \cap [\widehat{\mu}]^\perp, \\ \Upsilon(\widehat{\chi}) = \overline{\Pi(\widehat{\chi})} \cap [\widehat{\mu}]^\perp, \\ \Xi(\widehat{\chi}) = \{\chi \in X \mid \chi \in \Upsilon(\widehat{\chi})\}. \end{array}\right\} \tag{20}$$

In addition to the previous assumptions we assume:

(B7) There exists $\gamma > 0$ such that

$$(\chi, Q\chi) \geq \gamma\|\chi\|_X^2 \qquad \text{for all } \chi \in \Xi(\widehat{\chi}) - \Xi(\widehat{\chi}), \tag{21}$$

$$(\chi_\beta - \widehat{\chi}, Q(\chi_\beta - \widehat{\chi})) \geq \gamma\|\chi_\beta - \widehat{\chi}\|_X^2 \qquad \text{for all } \beta \in \mathcal{B}. \tag{22}$$

Remark 1 Coercivity conditions (B7) are not strong enough to ensure that the quadratic form $(\chi, Q\,\chi)$ induce a norm in the space $X$. That is the reason why in sensitivity analysis we cannot directly use the differentiability results for metric projections onto closed convex sets in a Hilbert space derived in Haraux (1977) and Mignot (1976).

Let $\beta \in \mathcal{B}$ be fixed. For $s > 0$ denote $\chi(s) = \chi_{s\beta}$, $\mu(s) = \mu_{s\beta}$ and

$$\varpi(s) = \frac{\chi(s) - \widehat{\chi}}{s}, \qquad \nu(s) = \frac{\mu(s) - \widehat{\mu}}{s} \tag{23}$$

In view of (B4), $\varpi(s)$ and $\nu(s)$ are uniformly bounded for $s \to 0$. Let $(\bar{\varpi}, \bar{\nu})$ be a weak cluster point of the sequence $\{\varpi(s), \nu(s)\}$. We will show that $(\bar{\varpi}, \bar{\nu})$ is defined uniquely. To this end we need the following lemma.

Lemma 1 *If assumptions* (B1) - (B7) *hold, then*

$$Q\bar{\varpi} + \beta + \Phi^*\bar{\nu} = 0, \tag{24}$$

$$(Q\bar{\varpi} + \beta, \bar{\varpi}) \leq 0, \tag{25}$$

$$\Phi\bar{\varpi} \in \Upsilon(\widehat{\chi}), \tag{26}$$

$$(Q\bar{\varpi} + \beta, \varpi) \geq 0 \quad \textit{for all } \varpi \in \{\varpi \in X \mid \Phi\varpi \in \Gamma(\widehat{\chi})\}. \tag{27}$$

*Proof.* The proof is a slight modification of the proof of Proposition 1 in Haraux (1977). Equation (24) follows immediately from (18). Using notation (23) we get $\chi(s) = \widehat{\chi} + s\varpi(s)$. Thus, the optimality condition (18) for $(\mathrm{AP})'_{s\beta}$ yields

$$(Q(\widehat{\chi} + s\varpi(s)) + (\widehat{\beta} + s\beta), s\varpi(s)) \leq 0.$$

Hence

$$s^2(Q\varpi(s)+\beta,\varpi(s)) \le -s(Q\widehat{\chi}+\widehat{\beta},\varpi(s)) = (Q\widehat{\chi}+\widehat{\beta},\widehat{\chi}-\chi(s)) \le 0. \tag{28}$$

In view of (B7) we have $(Q\varpi,\varpi) \le \lim_{s\to 0}\ \inf\,(Q\varpi(s),\varpi(s))$. Hence (28) implies (25). On the other hand, since $\varpi(s)$ is bounded, the estimate

$$s(Q\varpi(s)+\beta,\varpi(s)) \le -(Q\widehat{\chi}+\widehat{\beta},\varpi(s)) \le 0$$

yields $(Q\widehat{\chi}+\widehat{\beta},\bar{\varpi})=0$, i.e., by (19)

$$\Phi\bar{\varpi} \in [\widehat{\mu}]^{\perp}. \tag{29}$$

Clearly, by (23) $\Phi\bar{\varpi} \in \overline{\Pi(\widehat{\chi})}$, so (26) follows.

To prove (27) let us choose a sequence $\{s_n\}$ such that $s_n \to 0$ and $\varpi(s_n) \rightharpoonup \bar{\varpi}$. Denote $\epsilon_n = \varpi(s_n) - \bar{\varpi}$. Hence, $\chi(s_n) = \widehat{\chi} + s_n\bar{\varpi} + s_n\epsilon_n$. Let $\varpi$ be such that $\Phi\varpi \in \Gamma(\widehat{\chi})$. Then, in view of (19) and (20), as well as of (B6), there exist $r>0$ and $\pi \in X$ feasible for $(\mathrm{AP})'_\beta$ such that

$$\varpi = r(\pi - \widehat{\chi}) \tag{30}$$

and

$$(Q\widehat{\chi}+\widehat{\beta},\pi-\widehat{\chi}) = 0. \tag{31}$$

From the optimality condition (18) at $s_n\beta$, as well as from (29) and (31) we get

$$\begin{aligned} 0 &\ge (Q\chi(s_n) + (\widehat{\beta}+s_n\beta),\chi(s_n)-\pi) \\ &= ((Q\widehat{\chi}+\widehat{\beta}) + s_n(Q\bar{\varpi}+\beta) + s_nQ\epsilon_n, (\widehat{\chi}-\pi) + s_n\bar{\varpi} + s_n\epsilon_n) \\ &= s_n(Q\widehat{\chi}+\widehat{\beta},\epsilon_n) + s_n(Q\bar{\varpi}+\beta,\widehat{\chi}-\pi) + s_n^2(Q\bar{\varpi}+\beta,\bar{\varpi}) \\ &\quad + s_n^2(Q\bar{\varpi}+\beta,\epsilon_n) + s_n(Q\epsilon_n,\widehat{\chi}-\pi) + s_n^2(Q\epsilon_n,\bar{\varpi}) + s_n^2(Q\epsilon_n,\epsilon_n), \end{aligned}$$

which implies

$$\begin{aligned} (Q\bar{\varpi}+\beta,\widehat{\chi}-\pi) \;&\le -(Q\widehat{\chi}+\widehat{\beta},\epsilon_n) - (Q\epsilon_n,\widehat{\chi}-\pi) \\ &\quad - s_n[(Q\bar{\varpi}+\beta,\bar{\varpi}) + (Q\bar{\varpi}+\beta,\epsilon_n) + (Q\epsilon_n,\bar{\varpi}) + (Q\epsilon_n,\epsilon_n)]. \end{aligned}$$

Since the right-hand side of this inequality tends to zero as $n\to\infty$, we obtain

$$(Q\bar{\varpi}+\beta,\pi-\widehat{\chi}) \ge 0.$$

Thus, (27) follows from (30), what completes the proof of the lemma. ∎

We assume that the following condition, stronger than (27), holds

(B8)

$$(Q\bar{\varpi}+\beta,\varpi) \ge 0 \quad \text{for all } \varpi \in \Xi(\widehat{\chi})$$

or equivalently

$$(\Phi\varpi,\bar{\nu}) \le 0 \quad \text{for all } \varpi \in \Xi(\widehat{\chi}). \tag{32}$$

Remark 2 It follows from (20) and (27) that condition (B8) is satisfied if $\Upsilon(\widehat{\chi})] = \overline{\Gamma(\widehat{\chi})}$. This last condition is called *polyhedricity* and was introduced in Haraux (1977) and Mignot (1976).

Proposition 1 *If* (B1) - (B8) *hold, then the solutions* $\chi_\beta$ *and the Lagrange multipliers* $\mu_\beta$ *of* $(\mathrm{AP})'_\beta$ *are directionally differentiable at* 0 *and the differentials* $d_\beta\chi(0;\beta)$ *and* $d_\beta\mu(0;\beta)$ *in a direction* $\beta \in \mathcal{B}$ *are given by the solution and Lagrange multiplier of the following linear-quadratic optimization problem:*

$$(\mathrm{DA})'_\beta \qquad \min_{\varpi\in X}\left\{\frac{1}{2}(\varpi, Q\varpi) + (\beta,\varpi)\right\}, \quad \textit{subject to} \quad \Phi\varpi \in \Upsilon(\widehat{\chi}).$$

*Proof.* Note that by (B6) the feasible set of $(\mathrm{DA})'_\beta$ is nonempty for any $\beta \in \mathcal{B}$. Hence, in view of (21) there exists a unique solution $\varpi_\beta$ of $(\mathrm{DA})'_\beta$. Thus, in view of (B6), it follows from (24) that $\bar{\nu}$ is defined uniquely. By (24), (25) and (B8) we find that $(\Phi\bar{\varpi}, \bar{\nu}) = 0$ and $(\Phi\varpi, \bar{\nu}) \le 0$ for all $\Phi\varpi \in \Upsilon(\widehat{\chi})$, so $\bar{\nu}$ is the unique multiplier of $(\mathrm{DA})'_\beta$. ■

By (15) and (16) we get the following corollary:

Corollary 2 *If* (B1) - (B8) *hold, then the solutions* $\eta(\delta)$ *and the Lagrange multipliers* $\mu(\delta)$ *of* $(\mathrm{AP})_\delta$ *are directionally differentiable at* $\delta = (0,0)$ *and the differentials* $d_\delta\eta(0;\delta)$ *and* $d_\delta\mu(0;\delta)$ *in a direction* $\delta \in \Delta$ *are given by the solution and Lagrange multiplier of the following linear-quadratic optimization problem:*

$$(\mathrm{DA})_\delta \; \min_{\theta\in X}\left\{\frac{1}{2}(\theta, Q\theta)+(\delta^1,\theta)\right\} \; \textit{subject to}\; \Phi\theta+\delta^2 \in \overline{\bigcup_{r>0} r(\mathcal{K}-(\Phi\widehat{\eta}+\widehat{\delta^2}))} \cap [\widehat{\mu}]^\perp.$$

Using definitions (3) and (14) and combining Corollaries 1 and 2 we obtain the following main result of this section:

Theorem 2 *If* (B1) - (B8) *hold, then the stationary points* $(\xi(h),\lambda(h))$ *of* $(\mathrm{P})_h$ *are directionally differentiable at* $\widehat{h}$ *and the differentials* $d_h\xi(\widehat{h};\Delta h)$ *and* $d_h\lambda(\widehat{h};\Delta h)$ *in a direction* $\Delta h \in H$ *are given by the solution and Lagrange multiplier of the following linear-quadratic optimization problem:*

$$(\mathrm{DP})_{\Delta h} \quad \min_{\vartheta\in X}\left\{\tfrac{1}{2}(\vartheta, D_{xx}L(\widehat{x},\widehat{\lambda},\widehat{h})\vartheta) + (\vartheta, D_{xh}L(\widehat{x},\widehat{\lambda},\widehat{h})\Delta h)\right\},$$
*subject to*
$$D_x\phi(\widehat{x},\widehat{\lambda})\vartheta + D_h\phi(\widehat{x},\widehat{\lambda})\Delta h \in \overline{\textstyle\bigcup_{r>0} r(\mathcal{K}-\phi(\widehat{x},\widehat{\lambda}))} \cap [\widehat{\lambda}]^\perp.$$

## 3. Optimal control problem

In this section a simple parameter dependent optimal control problem $(\mathrm{O})_h$ subject to state constraints is introduced and some basic properties of its solutions are recalled. Stability results for that problem were obtained in Malanowski

(2007b) under weakened second order optimality conditions. The results of that paper, together with those of Section 2, will be used in Section 4 to perform sensitivity analysis for $(\mathrm{O})_h$.

For the sake of simplicity assume that the space of parameters $H = \mathbb{R}^k$ is finite-dimensional, and for each $h \in H$ consider the following optimal control problem.

$$(\mathrm{O})_h \qquad \begin{array}{l} \text{Find } (x_h, u_h) \in X^2 \quad \text{such that} \\ F(x_h, u_h, h) = \min_{(x,u)\in X^2} \left\{ F(x, u, h) := \int_0^1 g(x(t), u(t), h)dt \right\} \\ \text{subject to} \\ \dot{x}(t) - f(x(t), u(t), h) = 0, \\ x(0) = 0, \\ \varphi(x(t), h) \le 0, \end{array}$$

where $g : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^k \to \mathbb{R},\ f : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^k \to \mathbb{R}^n,\ \varphi : \mathbb{R}^n \times \mathbb{R}^k \to \mathbb{R}$ and

$$X^p := W_0^{1,p}(0,1;\mathbb{R}^n) \times L^p(0,1;\mathbb{R}^m), \qquad p \in [1, \infty] \tag{33}$$

with $W_0^{1,p}(0,1;\mathbb{R}^n) := \{x \in W_0^{1,p}(0,1:\mathbb{R}^n) \,|\, x(0) = 0\}$.
Assume:

(A1) The functions $g(\cdot,\cdot,\cdot)$, $D_x g(\cdot,\cdot,\cdot)$ and $D_u g(\cdot,\cdot,\cdot)$, as well as $f(\cdot,\cdot,\cdot)$, $D_x f(\cdot,\cdot,\cdot)$ and $D_u f(\cdot,\cdot,\cdot)$ are Fréchet differentiable in $(x, u, h)$. The functions $\varphi(\cdot,\cdot)$, $D_x\varphi(\cdot,\cdot)$ and $D^2_{xx}\varphi(\cdot,\cdot)$ are Fréchet differentiable in $(x, h)$.

(A2) For a given reference value $\widehat{h} \in H$ of the parameter there exists a local solution $(\widehat{x}, \widehat{u})$ of $(\mathrm{O})_{\widehat{h}}$, where $\widehat{u} \in C(0,1;\mathbb{R}^m)$.

Note that the assumption $\widehat{u} \in C(0,1;\mathbb{R}^m)$ is used in Lemmas 2 and 4 below.

Let us define the spaces of multipliers

$$Y^p := L^p(0,1;\mathbb{R}^n) \times W^{1,p}(0,1;\mathbb{R}), \qquad p \in [1, \infty]. \tag{34}$$

Denote by $K := \{d \in W^{1,2}(0,1;\mathbb{R}) \,|\, d(t) \le 0 \ \text{ for all } t \in [0,1]\}$ the cone of nonpositive functions in $W^{1,2}(0,1;\mathbb{R})$. The cone polar to $K$ is given (see e.g., Outrata and Schindler, 1980) by

$$K^+ = \left\{ \kappa \in W^{1,2}(0,1;\mathbb{R}) \;\middle|\; \begin{cases} \kappa(0) - \dot\kappa(0+) \ge 0, \quad \dot\kappa(t) \ge 0 \\ \text{and } \dot\kappa(\cdot) \text{ is nonincreasing} \end{cases} \right\}. \tag{35}$$

Introduce the following Lagrangian $\mathcal{L} : X^2 \times K^+ \times H \to \mathbb{R}$ for $(\mathrm{O})_h$:

$$\begin{aligned} \mathcal{L}(x, u, p, \kappa, h) \quad &= F(x, u, h) - (p, \dot{x} - f(x, u, h)) \\ &\quad + \kappa(0)\varphi(x(0), h) + (\dot\kappa, D_x\varphi(x, h) f(x, u, h)), \end{aligned} \tag{36}$$

where $(\cdot,\cdot)$ denotes the inner product in $L^2$. The Lagrangian is in the so called indirect or Pontryagin form (see e.g., Section 5 in Hartl, Sethi and Vickson, 1995).

The standard first order optimality conditions for $(\mathrm{O})_h$ can be expressed in the form

$$\left.\begin{array}{l} D_x g(x,u,h) + \dot{p} + D_x f^*(x,u,h)p + (D_x f^*(x,u,h) D_x \varphi^*(x,h) \\ \qquad + D^2_{xx}\varphi(x,h) f(x,u,h))\dot{\kappa} = 0, \\ p(1) = 0, \\ D_u g(x,u,h) + D_u f^*(x,u,h)p + D_u f^*(x,u,h) D_x \varphi^*(x,h)\dot{\kappa} = 0 \\ \varphi(x,h) \in \mathcal{N}_{K^+}(\kappa), \end{array}\right\} \tag{37}$$

where $\mathcal{N}_{K^+}(\kappa)$ is the cone normal to $K^+$ at $\kappa$.

To show that conditions (37) hold at the reference point $\widehat{h}$ we need some constraint qualifications to be satisfied at $(\widehat{x},\widehat{u})$. To formulate these conditions define the following sets of $\alpha$-active constraints:

$$M_\alpha = \{t \in [0,1] \mid \varphi(\widehat{x}(t),h) \geq -\alpha\}, \tag{38}$$

where $\alpha \geq 0$. We assume:

(A3) $0 \notin M_0$ and there exists $\rho > 0$ such that
$$|D_u f^*(\widehat{x}(t),\widehat{u}(t),\widehat{h}) D_x \varphi^*(\widehat{x}(t),\widehat{h})| \geq \rho \qquad \text{for all } t \in M_0.$$

Note that assumption (A3) eliminates higher order state constraints (see Bryson and Ho, 1975). Clearly, in view of (A2) there exists $\alpha > 0$ such that (A3) holds for all $t \in M_\alpha$. Hence, one gets (see, e.g., Lemma 4.1 and Theorem 4.3 in Malanowski, 2003) the following lemmas.

LEMMA 2 *If assumptions* (A1) - (A3) *hold, then the map*

$$\mathcal{C}_\alpha : X^2 \times W^{1,2}(M_\alpha;\mathbb{R}) \to L^2(0,1;\mathbb{R}^n) \times W^{1,2}(M_\alpha;\mathbb{R})$$
$$\mathcal{C}_\alpha \begin{bmatrix} y \\ v \end{bmatrix} = \begin{bmatrix} \dot{y} - D_x f(\widehat{x},\widehat{u},\widehat{h}) - D_u f(\widehat{x},\widehat{u},\widehat{h})v \\ D_x\varphi(\widehat{x},\widehat{h}) \end{bmatrix}$$

*is surjective.*

LEMMA 3 *If assumptions* (A1) - (A3) *are satisfied, then there exists a unique Lagrange multiplier* $\widehat{\lambda} := (\widehat{p},\widehat{\kappa}) \in (Y^2)^*$, *such that the first order optimality conditions* (37) *hold at* $(\widehat{x},\widehat{u},\widehat{p},\widehat{\kappa},\widehat{h})$.

For the sake of simplicity the functions evaluated at the reference points $(\widehat{x},\widehat{u},\widehat{p},\widehat{\kappa},\widehat{h})$ will be denoted by "hat", e.g., $\widehat{\mathcal{L}}(t) := \mathcal{L}(\widehat{x}(t),\widehat{u}(t),\widehat{p}(t),\widehat{\kappa}(t),\widehat{h})$.

In addition to constraint qualifications we will need some coercivity conditions. Assume that the following Legendre - Clebsch condition holds:

(A4) There exists $\gamma > 0$ such that
$$\langle v, D^2_{uu}\widehat{\mathcal{L}}(t)\, v\rangle \geq \gamma\, |v|^2 \quad \text{for all } v \in \mathbb{R}^m \text{ and } t \in [0,1],$$
where $\langle\cdot,\cdot\rangle$ denotes the inner product in $\mathbb{R}^m$.

The following regularity result follows from Theorem 2.1 in Hager (1979) (see Proposition 6.6 in Malanowski, 1995).

LEMMA 4 *If assumptions* (A1) - (A4) *hold, then the functions* $\dot{\widehat{x}}, \widehat{u}, \dot{\widehat{p}}$ *and* $\dot{\widehat{\kappa}}$ *are Lipschitz continuous, i.e., there exists* $\widehat{\varsigma} > 0$ *such that*

$$\|\ddot{\widehat{x}}\|_\infty, \|\dot{\widehat{u}}\|_\infty, \|\ddot{\widehat{p}}\|_\infty, \|\ddot{\widehat{\kappa}}\|_\infty \leq \widehat{\varsigma}.$$

In view of the uniqueness and regularity of $\widehat{\kappa}$ we can introduce the following sets depending on the parameter $\alpha > 0$

$$N_\alpha = [0,1] \setminus \overline{\{t \in [0,1] \mid -\ddot{\widehat{\kappa}}(t) \leq \alpha\}}, \quad \text{as well as } N_0 = \bigcup_{\alpha > 0} N_\alpha. \tag{39}$$

Define the following subspace of $X^2$:

$$\mathcal{E}_\alpha = \left\{ (y,v) \in X^2 \;\middle|\; \left\{ \begin{array}{l} \dot{y}(t) - D_x\widehat{f}(t)y(t) - D_u\widehat{f}(t)v(t) = 0 \\ \langle D_x\widehat{\varphi}(t), y(t)\rangle = 0 \quad \text{for all } t \in N_\alpha \\ \langle D_x\widehat{\varphi}(1), y(1)\rangle = 0 \quad \text{if } \dot{\widehat{\kappa}}(1) > 0 \end{array} \right. \right\}. \tag{40}$$

Moreover, we assume:

(A5) There exist constants $\alpha > 0$ and $\gamma > 0$ such that

$$((y,v), D^2\widehat{\mathcal{L}}(y,v)) \geq \gamma(\|y\|_{1,2}^2 + \|v\|_2^2) \quad \text{for all } (y,v) \in \mathcal{E}_\alpha, \tag{41}$$

where $D^2\widehat{\mathcal{L}} := D^2_{(x,u)(x,u)}\mathcal{L}(\widehat{x}, \widehat{u}, \widehat{p}, \widehat{\kappa}, \widehat{h})$.

Note that (41) constitutes a second order sufficient optimality condition for $(\mathrm{O})_{\widehat{h}}$.

## 4. Sensitivity analysis for optimal control problem

To apply sensitivity results of Section 2 to the optimal control problem $(\mathrm{O})_h$ we have to represent it in the form of a cone-constrained problem. To this end we set:

$$\left.\begin{array}{l} \mathbb{X} = X^2, \quad \mathbb{Y} = Y^2, \quad \mathbb{Z} = X^2 \times Y^2, \quad \mathcal{K} = \{0\} \times K, \\ (X^2)^* = L^2(0,1;\mathbb{R}^n) \times L^2(0,1;\mathbb{R}^m), \quad (Y^2)^* = Y^2, \\ \xi = (x,u), \quad \lambda = (p,\kappa), \quad \zeta = (\xi,\lambda) \\ \mathcal{G}(\xi,h) = F(x,u,h), \quad \phi(\xi,h) = \left[ \begin{array}{c} \dot{x} - f(x,u,h) \\ \varphi(x,h) \end{array} \right]. \end{array} \right\} \tag{42}$$

As in (4), the optimality conditions (37) can be represented in the form $\mathcal{F}(\zeta,h) \in \mathcal{T}(\zeta)$, where $\mathcal{F}$ and $\mathcal{T}$ are given by the left and right-hand side of (37), respectively. Let $\delta = (\delta^1, \delta^2, \delta^3, \delta^4) \in Z^*$. The accessory problem $(\mathrm{AP})_\delta$, which plays the crucial role in Theorem 1, takes the form:

$$\begin{array}{ll} (\mathrm{AO})_\delta & \min_{x,u} J(x,u,\delta) \qquad \text{subject to} \\ & \dot{x} - Ax - Bu + (\delta^3 + \widehat{\delta}^3) = 0, \quad x(0) = 0, \\ & Gx + (\delta^4 + \widehat{\delta}^4) \leq 0, \end{array}$$

where

$$\left.\begin{aligned}
&J(x,u,\delta)=\tfrac{1}{2}((x,u),D^2\widehat{\mathcal{L}}\,(x,u))+(\delta^1+\widehat{\delta}^1,x)+(\delta^2+\widehat{\delta}^2,u),\\
&A=D_x\widehat{f},\quad B=D_u\widehat{f},\quad G=D_x\widehat{\varphi}\\
&\widehat{\delta}^1=-D^2_{xx}\widehat{\mathcal{L}}\,\widehat{x}-D^2_{xu}\widehat{\mathcal{L}}\,\widehat{u}+D_xF(\widehat{x},\widehat{u},\widehat{h}),\\
&\widehat{\delta}^2=-D^2_{ux}\widehat{\mathcal{L}}\,\widehat{u}-D^2_{uu}\widehat{\mathcal{L}}\,\widehat{u}+D_uF(\widehat{x},\widehat{u},\widehat{h}),\\
&\widehat{\delta}^3=D_x\widehat{f}\,\widehat{x}+D_u\widehat{f}\,\widehat{u}-\widehat{f},\\
&\widehat{\delta}^4=\widehat{\varphi}-D_x\widehat{\varphi}\,\widehat{x}.
\end{aligned}\right\}\qquad(43)$$

Note that the Lagrangian $\mathcal{L}$ is not twice continuously differentiable on $X^2$, but it is on $X^\infty$. So, as in Theorem 1, we introduce the set $\mathcal{Z}\subset Z=X^2\times Y^2$ of more regular functions, on which this differentiability property of $\mathcal{L}$ is satisfied. Namely, we choose $\varsigma>\widehat{\varsigma}$, where $\widehat{\varsigma}$ is given in Lemma 4, and we define

$$\mathcal{Z}:=\{(x,u,p,\kappa)\in Z\mid \|\ddot{x}\|_\infty,\|\dot{u}\|_\infty,\|\ddot{p}\|_\infty,\|\ddot{\kappa}\|_\infty\le\varsigma\},\qquad(44)$$

Clearly, on this set conditions (B2) are satisfied. On the other hand, by (43) and by Lemma 4, there exists $\widehat{\sigma}>0$ such that

$$\|\dot{\widehat{\delta}}^1\|_\infty,\|\dot{\widehat{\delta}}^2\|_\infty,\|\dot{\widehat{\delta}}^3\|_\infty,\|\ddot{\widehat{\delta}}^4\|_\infty\le\widehat{\sigma}.$$

Accordingly, the set $\Delta\subset Z^*$ needed in (B4) is chosen as

$$\Delta=\{(\delta^1,\delta^2,\delta^3,\delta^4)\in X^*\times Y\mid \|\dot{\delta}^1\|_\infty,\|\dot{\delta}^2\|_\infty,\|\dot{\delta}^3\|_\infty,\|\ddot{\delta}^4\|_\infty\le\sigma\},\qquad(45)$$

where $\sigma>\widehat{\sigma}$. It was proved in Theorem 4.6 in Malanowski (2007b) that with the above definitions all assumptions (B1) -(B5) of Theorem 1 are satisfied. Thus, by that theorem, if assumptions (A1) -(A5) hold, then there is a constant $\ell>0$ such that

$$\|x_{h'}-x_{h''}\|_{1,2},\|u_{h'}-u_{h''}\|_2,\|p_{h'}-p_{h''}\|_{1,2},\|\kappa_{h'}-\kappa_{h''}\|_{1,2}\le\ \ell\,|h'-h''|\qquad(46)$$

for all $h',h''$ in a neighborhood of $\widehat{h}$. Moreover, by Corollary 4.8 in Malanowski (2007b) $(x_h,u_h)$ is a local solution of $(\mathrm{O})_h$.

To use Theorem 2 we still have to verify assumptions (B6) - (B8). Note that, in view of (46), $|\varphi(x_h(t),h)-\varphi(\widehat{x}(t),\widehat{h})|<\alpha$ for all $t\in[0,1]$ and all $h$ sufficiently close to $\widehat{h}$. Hence, for those $h$, $\varphi_h(t)<0$ for all $t\notin M_\alpha$, where $M_\alpha$ is given in (38). Thus, in sensitivity analysis it is enough to consider inequality constraints only on the set $M_\alpha$. Hence, (B6) is satisfied by Lemma 2. The condition (B7) was shown to be satisfied in the proof of Theorem 4.3 in Malanowski (2007a). The following lemma shows that condition (B8) holds.

Lemma 5 *If conditions* (A1) - (A5) *are satisfied, then* (B8) *holds.*

*Proof.* Using Lemma 3 we can introduce the change of variables as in (15), so that locally the constraints in $(\mathrm{AO})_\delta$ become independent of $\delta$. In the new variables, the problem $(\mathrm{AO})_\delta$ takes the form $(\mathrm{AO})'_\beta$, where $\beta$ is a linear function of $(\delta^1, \delta^2, \delta^3, \delta^4)$, as in $(\mathrm{AP})'_\beta$. Let $(q_\beta, \imath_\beta))$ denote the Lagrange multiplier of $(\mathrm{AO})'_\beta$. Choose any direction $\beta$ in the space of perturbations and a sequence of positive numbers $s \to +0$. Denote $q(s) = q_{s\beta}$, $\imath(s) = \imath_{s\beta}$,

$$r(s) = \frac{q(s) - \widehat{q}}{s}, \qquad \jmath(s) = \frac{\imath(s) - \widehat{\imath}}{s}, \tag{47}$$

and let $(\bar{r}, \bar{\jmath})$ be a weak cluster point of $\{r(s), \jmath(s)\}$.

Note that the set $\Xi(\widehat{\chi})$, defined in (20), is given by

$$\Xi(\widehat{\chi}) = \left\{ (y,v) \in X \;\middle|\; \left\{ \begin{array}{l} \dot{y} - Ay - Bv = 0 \\ \\ G(t)y(t) \left\{ \begin{array}{ll} = 0 & \text{for } t \in N_0 \\ \leq 0 & \text{for } t \in M_0 \setminus N_0 \\ \text{free} & \text{for } t \in [0,1] \setminus M_0 \end{array} \right. \end{array} \right. \right\}, \tag{48}$$

where the sets $M_0$ and $N_0$ are given in (38) and (39), respectively. In view of (48), (B8) holds if

$$\left. \begin{array}{l} \bar{\jmath}(0)y(0) + \displaystyle\int_0^1 \frac{d}{dt}\bar{\jmath}(t)\frac{d}{dt}(G(t)y(t))dt = \int_0^1 \frac{d}{dt}\bar{\jmath}(t)\frac{d}{dt}(G(t)y(t))dt \leq 0 \\ \text{for all } G(t)y(t) \left\{ \begin{array}{ll} = 0 & \text{for } t \in N_0 \\ \leq 0 & \text{for } t \in M_0 \setminus N_0 \\ \text{free} & \text{for } t \in [0,1] \setminus M_0 \end{array} \right. \end{array} \right\}. \tag{49}$$

Note that $\frac{d}{dt}(G(t)y(t)) = 0$ for all $t \in N_0$, so

$$\int_0^1 \frac{d}{dt}\bar{\jmath}(t)\frac{d}{dt}(G(t)y(t))dt = \int_{[0,1]\setminus \overline{N}_0} \frac{d}{dt}\bar{\jmath}(t)\frac{d}{dt}(G(t)y(t))dt \tag{50}$$

The set $[0,1] \setminus \overline{N}_0$ consists of at most countable number of open subsets $(t', t'')$. On each $(t', t'')$ we have $\frac{d}{dt}\widehat{\imath}(\cdot) = \text{const.}$ Hence, in view of (47), $\frac{d}{dt}\bar{\jmath}(\cdot)$ is non-increasing on $(t', t'')$. Moreover, it is easy to see that $\frac{d}{dt}\bar{\jmath}(\cdot)$=const. on each subinterval $(\tau', \tau'') \subset (t', t'') \cap M_0$. Therefore,

$$\int_{\tau'}^{\tau''} \frac{d}{dt}\bar{\jmath}(t)\frac{d}{dt}(G(t)y(t))dt = \bar{\jmath}(\tau'+)[G(\tau'')y(\tau'') - G(\tau')y(\tau')].$$

So, the value of that integral remains unchanged if we substitute $G(t)y(t)$ by $\frac{G(\tau'')y(\tau'') - G(\tau')y(\tau')}{\tau'' - \tau'}(t - \tau')$. Define

$$z(t) = \left\{ \begin{array}{ll} G(t)y(t) & \text{for } t \in (t', t'') \setminus M_0, \\ \frac{G(\tau'')y(\tau'') - G(\tau')y(\tau')}{\tau'' - \tau'}(t - \tau') & \text{for } t \in (t', t'') \cap M_0. \end{array} \right.$$

It follows from the construction that

$$\int_{t'}^{t''} \frac{d}{dt}\bar{\jmath}(t)\frac{d}{dt}(G(t)y(t))dt = \int_{t'}^{t''} \frac{d}{dt}\bar{\jmath}(t)\frac{d}{dt}z(t)dt. \tag{51}$$

Moreover, $z(t') = z(t'') = 0$ and $z(t) \leq 0$ for all $t \in (t', t'')$. Let us approximate $\frac{d}{dt}\bar{\jmath}(\cdot)$ by a sequence of piecewise constant and nonincreasing functions $j_\tau(\cdot)$. It is easy to see that

$$\int_{t'}^{t''} j_\tau(t)\frac{d}{dt}z(t)dt \leq 0.$$

Passing to the limit with $j_\tau$, and using (51) we finally obtain

$$\int_{t'}^{t''} \frac{d}{dt}\bar{\jmath}(t)\frac{d}{dt}(G(t)y(t)) \leq 0. \tag{52}$$

Combining (50) with (52) we get (49), which completes the proof of the lemma. ■

Thus, all assumptions of Theorem 2 are satisfied and by that theorem we obtain the following principal result of this paper.

THEOREM 3 *If assumptions* (A1) - (A5) *are satisfied, then the solutions* $\xi(h) := (x(h), u(h))$ *and the Lagrange multipliers* $\lambda(h) := (p(h), \kappa(h))$ *of* $(\mathrm{O})_h$ *are directionally differentiable at* $\widehat{h}$ *and the differentials in a direction* $\Delta h \in H$ *are given by the solutions and Lagrange multipliers of the following linear-quadratic optimal control problem:*

$$(\mathrm{DO})_h \quad \min_{(y,v)\in X^2} \left\{ \frac{1}{2}\left( (y,v), \begin{pmatrix} D^2_{xx}\widehat{\mathcal{L}} & D^2_{xu}\widehat{\mathcal{L}} \\ D^2_{xx}\widehat{\mathcal{L}} & D^2_{xu}\widehat{\mathcal{L}} \end{pmatrix} (y,v) \right) + \left( (y,v), \begin{pmatrix} D^2_{xh}\widehat{\mathcal{L}} \\ D^2_{uh}\widehat{\mathcal{L}} \end{pmatrix} \Delta h \right) \right\}$$

*subject to*

$$\dot{y}(t) - D_x\widehat{f}(t)y(t) - D_u\widehat{f}(t)v(t) - D_h\widehat{f}(t)\Delta h = 0,$$

$$y(0) = 0,$$

$$D_x\widehat{\varphi}(t)y(t) + D_h\widehat{\varphi}(t)\Delta h \begin{cases} = 0 & \text{for all } t \in N_0, \\ \leq 0 & \text{for all } t \in M_0 \setminus N_0. \end{cases}$$

Thus, the same assumptions, which are needed to obtain Lipschitz stability of the solutions to $(\mathrm{O})_h$, ensure also sensitivity results.

One should stress here that in the case of higher order state constraints, the optimal trajectories are more regular - they belong to higher order Sobolev spaces (see, e.g., Hermant, 2009). For those spaces we cannot repeat the proof of Lemma 5. Accordingly, the results presented in Section 2 cannot be applied to such a class of optimal control problems. In sensitivity analysis for those problems the curvature of the respective cones has to be taken into account. To the knowledge of the author that is still an open problem.

**Acknowledgment**

The author would like to express his gratitude to two anonymous referees for their valuable comments and corrections.

## References

BONNANS, J.F. and HERMANT, A. (2008) Stability and sensitivity analysis for optimal control problems with a first-order state constraint and application to continuation methods. *ESAIM: Control, Optimisation and Calculus of Variations* **14**, 825–863.

BRYSON, A.E. and HO, Y.C. (1975) *Applied Optimal Control*. Wiley, New York.

DONTCHEV, A.L. (1995) Implicit function theorems for generalized equations. *Math. Programming* **70**, 91–106.

DONTCHEV, A.L. and HAGER, W.W. (1998) Lipschitz stability for state constrained nonlinear control and optimization. *SIAM J. Control Optim.* **36**, 698–718.

HAGER, W.W. (1979) Lipschitz continuity for constrained processes. *SIAM J. Control Optim.* **17**, 321–337.

HARAUX, A. (1977) How to differentiate the projection on a convex set in Hilbert space. Some applications to variational inequalities. *J. Math. Soc. Japan* **29**, 615–631.

HARTL, R.F., SETHI, S.P. and VICKSON, R.G. (1995) A survey of maximum principle for optimal control problems with state constraints. *SIAM Review* **37**, 181–218.

HERMANT, A. (2009) Stability analysis of optimal control problems with a second-order state constraint. *SIAM J. Optim.* **20**, 104–129.

MALANOWSKI, K. (1995) Stability and sensitivity of solutions to nonlinear optimal control problems. *Appl. Math. Optim.* **32**, 111–141.

MALANOWSKI, K. (2003) On normality of Lagrange multipliers for state constrained optimal control problems. *Optimization* **52**, 75–91.

MALANOWSKI, K. (2007a) Sufficient optimality conditions in stability analysis for state-constrained optimal control. *Appl. Math. Optim.* **55**, 255–271.

MALANOWSKI, K. (2007b) Stability analysis for nonlinear optimal control problems subject to state constraints. *SIAM J. Optim.* **18**, 926–945.

MAURER, H. (1981) First- and second-order sufficient optimality conditions in mathemathical programming and optimal control. *Math. Program. Studies* **14**, 163–177.

MIGNOT, F. (1976) Contrôle dans les inéquations variationelles. *J. Funct. Anal.* **22**, 130–185.

OUTRATA, J.V. and SCHINDLER, Z. (1981) An augmented Lagrangian method for a class of convex optimal control problems. *Problems Contr. Inform. Theory* **10**, 67–81.

Robinson, S.M. (1980) Strongly regular generalized equations. *Math. Oper. Res.* **5**, 43–62.

Robinson, S.M. (1991) An implicit-function theorem for a class of nonsmooth functions. *Math. Oper. Res.* **16**, 292–309.