Control and Cybernetics

vol. **27** (1998) No. 3

Application of the classical implicit function theorem in sensitivity analysis of parametric optimal control[1]

by

Kazimierz Malanowski

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

Abstract: A family $\{(\mathrm{O}_h)\}$ of parametric optimal control problems for nonlinear ODEs is considered. The problems are subject to pointwise *inequality* type state constraints. It is assumed that the reference solution is regular. The original problems (O_h) are substituted by problems $(\widetilde{\mathrm{O}}_h)$ subject to *equality* type constraints with the sets of activity depending on the parameter. Using the classical implicit function theorem, conditions are derived under which stationary points of $(\widetilde{\mathrm{O}}_h)$ are Fréchet differentiable functions of the parameter. It is shown that, under additional conditions, the stationary points of $(\widetilde{\mathrm{O}}_h)$ correspond to the solutions and Lagrange multipliers of (O_h).

Keywords: parametric optimal control, nonlinear ordinary differential equations, state constraints, differentiability of solutions, implicit function theorem.

1. Introduction

Optimal control problems subject to *inequality* type constraints on control and/or state are nonsmooth by their very nature. Therefore, in general it is not possible to use the classical implicit function theorem in sensitivity analysis for such problems. Instead of it Robinson's implicit function theorem for generalized equations Robinson (1980) is exploited, which allows to prove Lipschitz continuity of the solutions with respect to the parameters, under reasonable assumptions (see, e.g., Dontchev, Hager, Poore and Yang, 1995; Dontchev and Hager, 1998; Malanowski, 1992; 1995). Further analysis allows to show directional differantiability of the solutions in L^2 Malanowski (1995).

To get Fréchet differentiability, additional restrictive assumptions on regularity of the reference solution are needed. Results of this type were recently

[1] Supported by the Polish State Committee for Scientific Research (KBN) under grant No. [illegible]

obtained in Malanowski and Maurer (1996) and Malanowski and Maurer (1998), where nonlinear problems with mixed control-state and additional pure state space constraints of the first order were considered, respectively.

In these papers the idea based on the so called *shooting method* was used. This method was developed as a numerical procedure of solving nonlinear optimal control problems (see, Stoer and Bulirsch, 1980). In the shooting method the necessary optimality conditions are expressed in the form of a multipoint boundary value problem for nonlinear ODEs and solved using the Newton method in a finite dimensional space.

It was noticed in Malanowski and Maurer (1998) that the crucial role in the shooting method is played by the idea of substituting the original problems with *inequality* type constraints by problems with *equality* type constraints, having the same *structure of the optimal control* as the reference solution.

In the present paper we focused on this basic idea, forgetting about its origin connected with the shooting method. In particular, we use the necessary optimality conditions in the Karush-Kuhn-Tucker (KKT) form, rather than in the form of a multipoint boundary value problem. It seems that this approach is simpler and more natural. The use of the multipoint boundary value problem is justified in numerical calculations, since it allows to apply the Newton method in a finite dimensional space. However, in the theoretical sensitivity analysis, the use of the boundary value formulation introduces some unnecessary technical complications that obscure the very essence of the method.

In the paper, parametric optimal control problems subject to state constraints are considered. The assumptions are the same as in Malanowski and Maurer (1998) and the obtained results are identical. The basic difference is that we use optimality conditions in the (KKT) form and apply the implicit function theorem in a Banach, rather than in a finite-dimensional space. Accordingly, we will concentrate on that point, whereas the other steps of the used method will be only briefly described and refered to Malanowski and Maurer (1998).

The organization of the paper is the following. In Section 2 the considered parametric optimal control problem (O_h) is defined and the main assumptions are introduced. In Section 3 the idea of the used approach is described and the auxiliary optimal control problem $(\widetilde{\mathrm{O}}_h)$ with *equality* type constraints is introduced. In Section 4 the basic differentiability results for stationary points of $(\widetilde{\mathrm{O}}_h)$ are obtained using the classical implicit function theorem. In Section 5 the conditions are briefly discussed under which the stationary points of $(\widetilde{\mathrm{O}}_h)$ correspond to the solutions and Lagrange multipliers of (O_h).

We denote by $I\!R^n$ the n-dimensional Euclidean space, with the inner product $\langle \cdot, \cdot \rangle$ and the norm $|\cdot|$. For $q \in [1, \infty)$, $L^q(0, T; I\!R^n)$ denotes the space of functions $x : [0, T] \mapsto I\!R^n$ with $|x(\cdot)|^q$ Lebesgue integrable. $L^\infty(0, T; I\!R^n)$ is the space of Lebesgue measurable and essentially bounded functions. $W^{1,q}(0, T; I\!R^n)$ is the in these spaces are denoted by $\|\cdot\|_q$ and $\|\cdot\|_{1,q}$, respectively, whereas $(\cdot,\cdot)$ denotes the inner product in $L^2(0,T;\mathbb{R}^n)$. For $f: X \times Y \mapsto Z$, where X, Y, Z are Banach spaces, $D_x f(x,y), D_y f(x,y), D^2_{xy} f(x,y), \ldots$ denote the respective Fréchet derivatives in the corresponding arguments.

2. Preliminaries

In this section the considered problem is formulated and the basic assumptions are introduced.

Let H denotes a space of parameters. It may be a Banach space, but for the sake of simplicity we assume that it is finite-dimensional. $G \subset H$ denotes an open set of feasible parameters. For each feasible value of the parameter consider the following optimal control problem:

(O_h) Find $(x_h, u_h) \in Z^\infty$ such that

$$F(x_h, u_h, h) = \min\{F(x,u,h) := \int_0^T f^0(x(t), u(t), h)dt)\}$$

subject to

$$\dot{x}(t) - f(x(t), u(t), h) = 0, \qquad \text{for a.a. } t \in [0,T], \tag{1}$$

$$x(0) - \xi(h) = 0, \tag{2}$$

$$\vartheta(x(t), h) \leq 0, \qquad \text{for all } t \in [0,T], \tag{3}$$

where $Z^q := W^{1,q}(0,T;\mathbb{R}^n) \times L^q(0,T;\mathbb{R}^m)$, while $\vartheta : \mathbb{R}^n \times G \mapsto \mathbb{R}$.
Assume:

(I.1) All involved functions are of class C^3 in some open sets.

(I.2) For a given reference value $h_0 \in G$ of the parameter there exists a possibly local solution (x_0, u_0) of (O_{h_0}) and $u_0 \in C(0,T;\mathbb{R}^m)$.

Our purpose is to find conditions under which a neighborhood $G_0 \subset G$ of h_0 exists, such that for each $h \in G_0$ there exists a locally unique solution (x_h, u_h) of (O_h), which is a *Fréchet differentiable* function of h.

REMARK **2.1** *To avoid technicalities problem* (O_h) *is formulated as simple as possible. However, there is no difficulty to use the same approach for more complicated problems, e.g., with mixed boundary value conditions, vector-valued state constraints and additional mixed control-state constraints (see, Malanowski and Maurer, 1998).* ◇

In our analysis a crucial role is played by regularity of the reference solution. To get appropriate regularity we will need several other assumptions.

We introduce the set of active constraints

$$\Omega_0 := \{t \in [0,T] \mid \vartheta(x_0(t), h_0) = 0\}$$

and assume that the set Ω_0 has the regular structure. Namely:

(I.3) The set Ω_0 consists of J disjoint subintervals

$$\Omega_0 = \bigcup_{1 \le j \le J} [\omega_j'(0), \omega_j''(0)].$$

The isolated touch points are excluded.

The points ω_j' and ω_j'' are called *entry* and *exit* points, respectively, whereas all these points are called *junction* points. For the sake of simplicity, it is assumed that $T \notin \Omega_0$.

Note that along a solution of the state equation we have

$$\frac{d}{dt}\vartheta(x(t),h) = D_x\vartheta(x(t),h)\dot{x}(t) = D_x\vartheta(x(t),h)f(x(t),u(t),h). \tag{4}$$

Define the function

$$\psi(x,u,h) := D_x\vartheta(x,h)f(x,u,h). \tag{5}$$

To simplify notation, the argument of functions evaluated at the reference point $(x_0(t), u_0(t), h_0)$ will be denoted by $[t]$, e.g., $\psi[t] := \psi(x_0(t), u_0(t), h_0)$. We need the following constraint qualifications:

(I.4) (*Linear independence condition*) There exists $\beta > 0$ such that

$$|D_u\psi[t]| \ge \beta \quad \text{for all } t \in \Omega_0.$$

(I.5) $\vartheta[0] < 0$.

In the same way as in Malanowski (1995) we introduce the following Lagrangian:

$$\begin{aligned}
&\mathcal{L} : W^{1,\infty}(0,T;\mathbb{R}^n) \times L^\infty(0,T;\mathbb{R}^m) \times W^{1,\infty}(0,T;\mathbb{R}^n) \times \\
&\times \mathbb{R}^n \times W^{1,\infty}(0,T;\mathbb{R}) \times G \mapsto \mathbb{R}, \\
&\mathcal{L}(x,u,p,\rho,\mu,h) = F(x,u,h) - (p, \dot{x} - f(x,u,h)) + \\
&+ \langle \rho, x(0) - \xi(h) \rangle + \mu(0)\vartheta(x(0),h) + (\dot{\mu}, \psi(x,u,h)),
\end{aligned} \tag{6}$$

as well as the Hamiltonian and augmented Hamiltonian

$$\begin{aligned}
&\mathcal{H} : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^n \times G \mapsto \mathbb{R}, \\
&\widetilde{\mathcal{H}} : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^n \times \mathbb{R} \times G \mapsto \mathbb{R}, \\
&\mathcal{H}(x,u,p,h) = f^0(x,u,h) + \langle p, f(x,u,h) \rangle, \\
&\widetilde{\mathcal{H}}(x,u,p,\dot{\mu},h) = \mathcal{H}(x,u,p,h) + \dot{\mu}\psi(x,u,h).
\end{aligned} \tag{7}$$

Note that the Lagrangian is in the so-called *Pontryagin form* with absolutely continuous function p (see, Section 7 in Hartl, Sethi and Vickson, 1995). The state constraints are considered in $W^{1,\infty}(0,T;\mathbb{R})$, where the general form of a linear functional is given by $\mu(0)y(0) + (\dot{\mu}, \dot{y})$. Accordingly the terms in Lagrangian (6) corresponding to the state constraints are obtained as follows:

$$\begin{aligned}
&\mu(0)\vartheta(x(0),h) + \left(\dot{\mu}, \tfrac{d}{dt}\vartheta(x,h)\right) = \\
&= \mu(0)\vartheta(x(0),h) + (\dot{\mu}, D_x\vartheta(x,h)\dot{x}) = \\
&= \mu(0)\vartheta(x(0),h) + (\dot{\mu}, D_x\vartheta(x,h)f(x,u,h)).
\end{aligned}$$

By (I.4) and (I.5) there exist (see, Malanowski, 1995) unique Lagrange multiare *more regular*) such that the following first order optimality conditions in the Karush-Kuhn-Tucker (KKT) form hold at $(x_0, u_0, p_0, \rho_0, \mu_0, h_0)$:

$$\dot{p}_0 + D_x\widetilde{\mathcal{H}}(x_0, u_0, p_0, \dot{\mu}_0, h_0) = 0, \quad p_0(T) = 0, \tag{8}$$

$$p_0(0) + \rho_0 + \mu_0(0)D_x\vartheta(x_0(0), h_0) = 0, \tag{9}$$

$$D_u\widetilde{\mathcal{H}}(x_0, u_0, p_0, \dot{\mu}_0, h_0) = 0, \tag{10}$$

$$\dot{\mu}_0(t) \geq 0 \text{ is nonincreasing a.e. on } [0,T], \tag{11}$$

$$\mu_0(0) = \dot{\mu}_0(0). \tag{12}$$

Conditions (11)-(12) follow from the form of the positive polar to the cone of nonnegative functions in $W^{1,\infty}(0,T;\mathbb{R})$, which is given (see, Outrata and Schindler, 1981) by the closure in $(W^{1,\infty}(0,T;\mathbb{R}))^*$ of the set

$$\{\mu \in W^{1,\infty}(0,T;\mathbb{R}) \mid \mu(0) \geq \dot{\mu}(0),\ \dot{\mu}(\cdot) \text{ is nonnegative and nonincreasing}\}.$$

Condition (I.5) implies *equality* in (12).

In addition to (I.1)-(I.5) we assume

(LC) *Legendre-Clebsch condition* A constant $\gamma > 0$ exists such that
$\langle v, D^2_{uu}\widetilde{\mathcal{H}}[t]v\rangle \geq \gamma|v|^2$
for all

$$v \in \begin{cases} \mathbb{R}^m & \text{if } t \in [0,T] \setminus \Omega_0, \\ \mathbb{R}^m \text{ such that } \langle D_u\psi[t], v\rangle = 0 & \text{if } t \in \Omega_0. \end{cases}$$

Using similar argument as in the proof of Lemma 7.2 in Malanowski and Maurer (1998) we obtain the following regularity result

LEMMA **2.2** *If (I.1)-(I.5) and (LC) hold, then u_0 and $\dot{\mu}_0$ are continuous and piecewise differentiable functions, with possible jumps of $\dot{u}_0$ and $\ddot{\mu}_0$ at the junction points.*

◇

3. The used approach

Using (4) and (5) we can write

$$\vartheta(x_0(t), h_0) = \vartheta(x_0(\omega'_j(0)), h_0) + \int_{\omega'_j(0)}^{t} \psi(x_0(s), u_0(s), h_0)ds.$$

Hence, for the reference solution, the *inequality* constraints (3) can be interpreted as the following *equality* type constraints

$$\begin{array}{ll} \psi(x(t), u(t), h_0) = 0, & \text{for } t \in \Omega_0 = \bigcup_{1\leq j\leq J}[\omega'_j(0), \omega''_j(0)], \\ \vartheta(x(\omega'_j(0)), h_0) = 0 & \text{for } 1 \leq j \leq J. \end{array} \tag{13}$$

This interpretation suggests to introduce modifications of the problems (O_h), in which constraints (3) are substituted by (13). For this new problems, denoted by

$(\widetilde{\mathrm{O}}_h)$, the structure of the solutions is imposed to be the same as for the reference one, i.e., *the number of subintervals of active constraints remains constant and equal to* J, but *the locations of the junction points can be changed.* The vector of these locations is treated as an additional argument of minimization. The main idea is that the sensitivity analysis for problems with equality type constraints is much easier than for those with inequality type constraints. It will be shown in Section 4. that the classical implicit function theorem can be used in this analysis. Subsequently, conditions will be derived under which the stationary points of $(\widetilde{\mathrm{O}}_h)$ correspond to the solutions and Lagrange multipliers of (O_h).

To formulate problems $(\widetilde{\mathrm{O}}_h)$ we have to choose the appropriate spaces of arguments and constraints. Regularity of the elements of these spaces is motivated by the regularity of the reference solutions and Lagrange multipliers. To this end, let us denote by $\omega = (\omega_1, \omega_2, \ldots, \omega_{2J})$, a $2J$-dimensional vector of junction points such that

$$0 < \omega_1, \quad \omega_j < \omega_{j+1}, \quad \omega_{2J} < T, \tag{14}$$

and put $\omega_0 = 0,\ \omega_{2J+1} = T$.

Introduce functions u that are uniformly continuously differentiable on each subinterval (ω_j, ω_{j+1}), continuous, together with their derivatives, on each boundary subinterval $[\omega_j', \omega_j'']$, $1 \le j \le J$. Locally, the functions u will be identified with elements of a Banach space. To do that, let us introduce the space

$$PC^1(0,T;I\!R) = C^1(0,T;I\!R) \times I\!R^{2J} \times I\!R^{2J} \tag{15}$$

of elements $\xi = (\alpha, \beta, \gamma)$. Endowed with the norm

$$\|\xi\|_{PC^1} = \max\{\|\alpha\|_{C^1}, |\beta|, |\gamma|\},$$

$PC^1(0,T;I\!R)$ is a Banach space. For a fixed ω satisfying (14), to any piece-wise differentiable function u we can assign an element $\xi \in PC^1(0,T;I\!R)$ putting

$$\begin{aligned} &\beta^j = \Delta u^j, \qquad\qquad \gamma^j = \Delta \dot{u}^j, \\ &\dot{\alpha}(t) = \dot{u}(t) - \Sigma_{j=1}^{2J} \Delta \dot{u}^j \mathbf{1}^j(t), \\ &\alpha(t) = u(t) - \Sigma_{j=1}^{2J} (\Delta u^j \mathbf{1}^j(t) + \Delta \dot{u}^j \mathbf{R}^j(t)) \quad , \end{aligned} \tag{16}$$

where

$$\Delta u^j = \lim_{t\to\omega_j+} u(t) - \lim_{t\to\omega_j-} u(t),$$

$$\Delta \dot{u}^j = \lim_{t\to\omega_j+} \dot{u}(t) - \lim_{t\to\omega_j-} \dot{u}(t),$$

$$\mathbf{1}^j = \begin{cases} 0 & \text{for } t \in [0,\omega_j), \\ 1 & \text{for } t \in [\omega_j, T], \end{cases} \quad , \quad \mathbf{R}^j = \begin{cases} 0 & \text{for } t \in [0,\omega_j), \\ (t-\omega_j) & \text{for } t \in [\omega_j, T]. \end{cases}$$

Reciprocally, using (16), to any pair (ξ, ω), with ω satisfying (14), we can assign

As in (15), we introduce the Banach space

$$PC^2(0,T;\mathbb{R}) = C^2(0,T;\mathbb{R}) \times \mathbb{R}^{2J} \times \mathbb{R}^{2J}, \tag{17}$$

of elements $\eta = (\alpha, \beta, \gamma)$ endowed with the norm

$$\|\eta\|_{PC^2} = \max\{\|\alpha\|_{C^2}, |\beta|, |\gamma|\},$$

Let x be a continuous function defined on $[0,T]$, with $\dot{x}$ being differentiable on (ω_j, ω_{j+1}). In the way analogous to (16), we can assign to x the element $\eta \in PC^2(0,T;\mathbb{R})$ putting

$$\begin{aligned} &\beta^j = \Delta\dot{x}^j, \qquad\qquad \gamma^j = \Delta\ddot{x}^j, \\ &\ddot{\alpha}(t) = \ddot{x}(t) - \Sigma_{j=1}^{2J}\Delta\ddot{x}^j\mathbf{1}^j(t), \\ &\dot{\alpha}(t) = \dot{x}(t) - \Sigma_{j=1}^{2J}(\Delta\dot{x}^j\mathbf{1}_j(t) + \Delta\ddot{x}^j\mathbf{R}^j(t)). \end{aligned} \tag{18}$$

Finally, define the Banach space

$$X = PC^2(0,T;\mathbb{R}^n) \times PC^1(0,T;\mathbb{R}^m) \times \mathbb{R}^{2J} \tag{19}$$

of elements $\chi = (\eta, \xi, \omega)$ endowed with the norm

$$\|\chi\|_X = \max\{\|\eta\|_{PC^2}, \|\xi\|_{PC^1}, |\omega|\}.$$

In view of (16) and (18), for any ω satisfying (14), we can identify the elements $\chi = (\eta, \xi, \omega)$ with the pairs (x, u) of functions piece-wise differentiable on (ω_j, ω_{j+1}).

Using the above identification, we can formulate the modified problems $(\widetilde{\mathrm{O}}_h)$ as follows:

$(\widetilde{\mathrm{O}}_h)$ Find $(\widetilde{x}_h, \widetilde{u}_h, \omega(h)) \in X$ such that

$$F(\widetilde{x}_h, \widetilde{u}_h, h) = \min_{x,u,\omega} F(x, u, h)$$

subject to

$$\dot{x}(t) - f(x(t), u(t), h) = 0 \quad \text{for } t \in [0.T], \tag{20}$$

$$x(0) - \xi(h) = 0, \tag{21}$$

$$\vartheta(x(\omega_j'), h) = 0 \quad \text{for } 1 \le j \le J, \tag{22}$$

$$\psi(x(t), u(t), h) = 0 \quad \text{for } t \in \Omega := \bigcup_{1\le j\le J} [\omega_j', \omega_j'']. \tag{23}$$

In general, an element feasible for $(\widetilde{\mathbf{O}}_h)$ is *not feasible* for $(\mathbf{O}_h)$. In order to ensure such feasibility of $(\widetilde{x}_h, \widetilde{u}_h, \omega(h))$ the additional *nontangential junction* condition (I.8) will be imposed in Section 4.

First, we are going to show that $(x_0, u_0, \omega(0))$ is a stationary point of $(\widetilde{\mathrm{O}}_{h_0})$. To this end, let us introduce the following Lagrangian for $(\widetilde{\mathrm{O}}_h)$:

$$\begin{array}{l}\widetilde{\mathcal{L}} : PC^2(0,T;I\!R^n) \times PC^1(0,T;I\!R^m) \times I\!R^{2J} \times \\ \times PC^2(0,T;I\!R^n) \times I\!R^n \times I\!R^J \times PC^1(0,T;I\!R) \times G \mapsto I\!R, \\ \widetilde{\mathcal{L}}(x,u,\omega,p,\rho,\phi,\lambda,h) = F(x,u,h) - (p, \dot{x} - f(x,u,h)) + \\ + \langle \rho, x(0) - \xi(h) \rangle + \Sigma_{j=1}^J \phi^j \vartheta(x(\omega_j'), h) + \\ + \Sigma_{j=1}^J \int_{\omega_j'}^{\omega_j''} \lambda(t) \psi(x(t), u(t), h) dt,\end{array} \tag{24}$$

where ϕ^j and λ represent multipliers corresponding to (22) and (23), respectively.

We will rearrange (24) to get Lagrangian $\widetilde{\mathcal{L}}$ in the form analogous to (6). Let us extend λ to $[0,T]$ putting

$$\lambda(t) = 0 \text{ for } t \notin \Omega. \tag{25}$$

Moreover, note that, in view of (4) and (5),

$$\phi^j \vartheta(x(\omega_j'), h) = \phi^j \vartheta(x(0), h) + \int_0^{\omega_j'} \phi^j \psi(x(t), u(t), h) dt.$$

Denote

$$\varphi = \Sigma_{j=1}^J \phi^j \tag{26}$$

and define the functions

$$\begin{array}{l}\sigma^j(t) = \left\{ \begin{array}{ll} \phi^j & \text{for } t \in [0, \omega_j'), \\ 0 & \text{for } t \in [\omega_j', T], \end{array} \right. \\ \sigma(t) = \Sigma_{j=1}^J \sigma^j(t), \qquad \nu(t) = \sigma(t) + \lambda(t).\end{array} \tag{27}$$

Note that (25)-(27) establish a one-to-one correspondence between (λ, ϕ) and (ν, φ), where

$$\begin{array}{l}\phi^j = \nu(\omega_j'-) - \nu(\omega_j''+), \quad 1 \le j \le J, \\ \lambda^j(t) = \left\{ \begin{array}{ll} \nu(t) - \nu(\omega_j'') & \text{for } t \in [\omega_j', \omega_j''], \\ 0 & \text{for } t \notin [\omega_j', \omega_j'']. \end{array} \right.\end{array} \tag{28}$$

Using (26)-(28) we can rewrite Lagrangian (24) in the form

$$\begin{array}{l}\widetilde{\mathcal{L}}(z,\zeta,h) = \widetilde{\mathcal{L}}(x,u,\omega,p,\rho,\varphi,\nu,h) = F(x,u,h) - (p, \dot{x} - f(x,u,h)) + \\ + \langle \rho, x(0) - \xi(h) \rangle + \varphi \vartheta(x(0), h) + \int_0^T \nu(t) \psi(x(t), u(t), h) dt,\end{array} \tag{29}$$

where, for the sake of simplicity, we put $z := (x, u, \omega)$, $\zeta := (p, \rho, \varphi, \nu)$. Lagrangians (6) and (29) coincide, with $\varphi = \mu(0)$ and $\nu = \dot{\mu}$. Denote

$$\varphi_0 = \mu_0(0), \; \nu_0 = \dot{\mu}_0, \; z_0 = (x_0, u_0, \omega(0)), \; \zeta_0 = (p_0, \rho_0, \varphi_0, \nu_0). \tag{30}$$

LEMMA **3.1** *The element z_0 is a stationary point of $\widetilde{\mathcal{L}}(z,\zeta_0,h_0)$.* ◇

Proof Using (7) we obtain the following stationarity conditions of $\widetilde{\mathcal{L}}$ with respect to x, u, ω, respectively

$$\dot{p} + D_x\widetilde{\mathcal{H}}(x,u,p,\nu,h) = 0 \quad p(T) = 0, \tag{31}$$

$$p(0) + \rho + \varphi D_x\vartheta(x(0),h) = 0, \tag{32}$$

$$D_u\widetilde{\mathcal{H}}(x,u,p,\nu,h) = 0, \tag{33}$$

$$\begin{aligned} &f^0(x(\tau),u(\tau+),h) + \nu(\tau+)\psi(x(\tau),u(\tau+),h) = \\ &= f^0(x(\tau),u(\tau-),h) + \nu(\tau-)\psi(x(\tau),u(\tau-),h) \\ &\text{for } \tau = \omega_j', \omega_j'',\ 1 \le j \le J. \end{aligned} \tag{34}$$

In view of (8)-(10) and (30), conditions (31)-(33) hold at the reference point. On the other hand, by continuity of u_0 and of $\dot{\mu}_0 = \nu_0$, condition (34) is also satisfied. □

4. Differentiability of stationary points

In this section we are going to investigate differentiability with respect to the parameter of the stationary points for $(\widetilde{\mathrm{O}}_h)$. To this end, we need the optimality system for $(\widetilde{\mathrm{O}}_h)$, which consists of the stationary conditions (31)-(34) and the constrains (20)-(23).

Note that conditions (34) are satisfied if functions u and ν are continuous at the junction points. So, in particular they are satisfied at the reference solution. Since, in view of (23),

$$\psi(x(\omega_j'),u(\omega_j'+),h) = 0,\ \psi(x(\omega_j''),u(\omega_j''-),h) = 0,\ 1 \le j \le J, \tag{35}$$

in case of continuity of u we must have

$$\psi(x(\omega_j'),u(\omega_j'-),h) = 0,\ \psi(x(\omega_j''),u(\omega_j''+),h) = 0,\ 1 \le j \le J. \tag{36}$$

Actually, in our analysis, we will require that (36), rather than (34) holds. Later on it will be shown that, for stationary points, (36) implies (34).

Let us introduce the following spaces

$$\begin{aligned} &PC^2_T(0,T;I\!\!R^n) = \{p \in PC^2(0,T;I\!\!R^n) \mid p(T) = 0\}, \\ &U = PC^2(0,T;I\!\!R^n) \times PC^1(0,T;I\!\!R^m) \times I\!\!R^{2J} \times \\ &\times PC^2_T(0,T;I\!\!R^n) \times I\!\!R^n \times I\!\!R \times PC^1(0,T;I\!\!R), \\ &V = PC^1(0,T;I\!\!R^n) \times I\!\!R^n \times PC^1(0,T;I\!\!R^m) \times PC^1(0,T;I\!\!R^n) \times \\ &\times I\!\!R^n \times I\!\!R^J \times PC^1(0,T;I\!\!R) \times I\!\!R^J \times I\!\!R^J, \end{aligned} \tag{37}$$

and define the mapping given by the left-hand sides of (31)-(33), (20)-(23) and (36):

$$\mathcal{F} : U \times G \mapsto V,$$

$$\mathcal{F}(z,\zeta,h) = \begin{pmatrix} \mathcal{F}'(z,\zeta,h) \\ \ldots \end{pmatrix} \tag{38}$$

where

$$\mathcal{F}'(z,\zeta,h) := \begin{bmatrix} \dot{p} + D_x\widetilde{\mathcal{H}}(x,u,p,\nu,h) & \\ p(0) + \rho + \varphi D_x\vartheta(x(0),h) & \\ D_u\widetilde{\mathcal{H}}(x,u,p,\nu,h) & \\ \dot{x} - f(x,u,h) & \\ x(0) - \xi(h) & \\ \vartheta(x(\omega_j'),h), & 1 \le j \le J \\ \psi(x(t),u(t),h), & t \in \Omega \end{bmatrix}, \tag{39}$$

$$\mathcal{F}''(z,\zeta,h) := \begin{bmatrix} \psi(x(\omega_j'),u(\omega_j'-),h), & 1 \le j \le J \\ \psi(x(\omega_j'),u(\omega_j'+),h), & 1 \le j \le J \end{bmatrix}. \tag{40}$$

We are going to apply the classical implicit function theorem to the equation

$$\mathcal{F}(z,\zeta,h) = 0 \tag{41}$$

at the reference point (z_0,ζ_0,h_0). To this end, we have to show that the Jacobian of $\mathcal{F}$ with respect to (z,ζ) evaluated at the reference point is regular. Using (38), we find that the Jacobian is regular if and only if, for any

$$\delta := (a,b) := (a_1,a_2,a_3,a_4,a_5,a_6,a_7,b_1,b_2) \in V,$$

there exists a unique solution $(w,\eta) := (y,v,\tau,q,\varrho,\theta,\kappa) \in U$ of the following linear equation:

$$\begin{bmatrix} D_{(z,\zeta)}\mathcal{F}'(z_0,\zeta_0,h_0) \\ D_{(z,\zeta)}\mathcal{F}''(z_0,\zeta_0,h_0) \end{bmatrix} \begin{bmatrix} w \\ \eta \end{bmatrix} = \begin{bmatrix} a \\ b \end{bmatrix}. \tag{42}$$

In view of (39)-(40), this equation takes on the form

$$\dot{q} + D_xf[t]^*q(t) + D^2_{xx}\widetilde{\mathcal{H}}[t]y(t) + D^2_{xu}\widetilde{\mathcal{H}}[t]v(t) + D_x\psi[t]^*\kappa(t) = a_1(t), \tag{43}$$

$$q(0) + \varrho + \varphi_0 D^2_{xx}\vartheta[0]y(0) + \theta D_x\vartheta[0] = a_2, \tag{44}$$

$$D^2_{ux}\widetilde{\mathcal{H}}[t]y(t) + D^2_{uu}\widetilde{\mathcal{H}}[t]v(t) + D_uf[t]^*q(t) + D_u\psi[t]^*\kappa(t) = a_3(t), \tag{45}$$

$$\dot{y}(t) - D_xf[t]y(t) - D_uf[t]v(t) = a_4(t), \tag{46}$$

$$y(0) = a_5, \tag{47}$$

$$D_x\vartheta[\omega_j'(0)]y(\omega_j'(0)) = a_6^j, \quad 1 \le j \le J, \tag{48}$$

$$D_x\psi[t]y(t) + D_u\psi[t]v(t) = a_7(t), \quad t \in \Omega_0, \tag{49}$$

and

$$\begin{aligned} &D_x\psi[\omega_j'(0)]y(\omega_j'(0)) + D_u\psi[\omega_j'(0)]v(\omega_j'(0)-) + \tfrac{d}{dt}\psi[\omega_j(0)-]\tau_j' = b_1^j, \\ &D_x\psi[\omega_j''(0)]y(\omega_j''(0)) + D_u\psi[\omega_j''(0)]v(\omega_j''(0)+) + \tfrac{d}{dt}\psi[\omega_j(0)+]\tau_j'' = b_2^j, \\ &\text{for } 1 \le j \le J. \end{aligned} \tag{50}$$

Note that (43)-(49) do not depend on τ. It follows from the fact that, in view of (40),

d ... 0 for $t = \omega_j'(0), \omega_j''(0)$, $1 \le j \le J$

Thus, we can solve (42) in two steps. In the first step we find $(y, v, q, \varrho, \theta, \kappa)$ that solve (43)-(49) and, in the second step, we substitute this solution to (50) and calculate τ. This fact simplifies substentially the further analysis. Note that such a situation takes place *only for problems with the first order* state constraints. For higher order constraints, the structure of $\mathcal{F}(z, \zeta, h)$ is more complicated and we are not able to repeat this simple two-step procedure (see, Malanowski and Maurer, 1997).

An inspection shows that (43)-(49) constitute an optimality system for the following accessory optimal control problem:

$$
\begin{array}{lll}
(\mathrm{LO}_\delta) & \text{Find } (y_\delta, v_\delta) \in Z^\infty \text{ such that} & \\
& I(y_\delta, v_\delta, \delta) = \min_{(y,v)} I(y, v, \delta) & \\
& \text{subject to} & \\
& \dot{y}(t) - D_x f[t] y(t) - D_u f[t] v(t) - a_4(t) = 0, & \\
& y(0) - a_5 = 0, & \\
& D_x \vartheta[\omega_j'(0)] y(w_j'(0)) - a_6^j = 0, & 1 \le j \le J, \\
& D_x \psi[t] y(t) + D_u \psi[t] v(t) - a_7(t) = 0, & t \in \Omega_0,
\end{array}
$$

where

$$
I(y, v, \delta) = \int_0^T \left\{ \frac{1}{2} \begin{bmatrix} y(t) \\ v(t) \end{bmatrix}^* \begin{bmatrix} D^2_{xx}\widetilde{\mathcal{H}}[t] & D^2_{xu}\widetilde{\mathcal{H}}[t] \\ D^2_{ux}\widetilde{\mathcal{H}}[t] & D^2_{uu}\widetilde{\mathcal{H}}[t] \end{bmatrix} \begin{bmatrix} y(t) \\ v(t) \end{bmatrix} - \right.
$$

$$
\left. - \begin{bmatrix} y(t) \\ v(t) \end{bmatrix}^* \begin{bmatrix} a_1(t) \\ a_3(t) \end{bmatrix} \right\} dt.
$$

Thus, (43)-(49) have a solution for any $\delta \in V$ if and only if (LO_δ) has a unique stationary point, which is more regular, with the regularity corresponding to U. We will use conditions that ensure existence and uniqueness of the solutions and Lagrange multipliers of (LO_δ). To this end, we will require that the mapping given by the linear part of the constraints is surjective and the cost functional is coercive on the kernel of this mapping. To meet these properties we need the following assumptions (see, Malanowski and Maurer, 1998):

(I.6) (*Controllability condition*) For any $e \in I\!R^J$ there exists $(z, w) \in Z^\infty$ such that

$$
\begin{array}{ll}
\dot{z}(t) - D_x f[t] z(t) - D_u f[t] w(t) = 0, & \\
z(0) = 0, & \\
D_x \vartheta[\omega_j'(0)] z(\omega_j'(0)) = e^j, & 1 \le j \le J, \\
D_x \psi[t] z(t) + D_u \psi[t] w(t) = 0, & \text{for a.a. } t \in \Omega_0.
\end{array}
$$

(I.7) (*Coercivity condition*) There exists $\gamma > 0$ such that

$$
\left((z, w), \begin{pmatrix} D^2_{xx}\widetilde{\mathcal{L}}_0 & D^2_{xu}\widetilde{\mathcal{L}}_0 \\ D^2_{ux}\widetilde{\mathcal{L}}_0 & D^2_{uu}\widetilde{\mathcal{L}}_0 \end{pmatrix} (z, w) \right) \ge \gamma(\|z\|^2_{1,2} + \|w\|^2_2), \tag{51}
$$

for all $(z, w) \in Z^\infty$ such that

$$
\dot{z}(t) - D_x f[t] z(t) - D_u f[t] w(t) = 0, \tag{52}
$$

$$D_x\psi[t]z(t) + D_u\psi[t]w(t) = 0, \quad \text{for a.a. } t \in \Omega_0, \tag{53}$$

where $\widetilde{\mathcal{L}}_0 := \widetilde{\mathcal{L}}(x_0, u_0, \omega(0), p_0, \rho_0, \varphi_0, \nu_0, h_0)$.

REMARK **4.1** *By Lemma 2 in Dontchev, Hager, Poore and Yang (1995), if (I.4) holds then* Coercivity condition *(I.7) implies* Legendre-Clebsch condition *(LC), so that under assumption (I.7) Lemma 2.2 holds.* □

LEMMA **4.2** *If (I.1)-(I.7) hold then, for any $\delta \in V$, problem* (LO_δ) *has a unique solution $(y_\delta, v_\delta) \in PC^2(0,T;\mathbb{R}^n) \times PC^1(0,T;\mathbb{R}^m)$ and unique associated Lagrange multiplier $(q_\delta, \varrho_\delta, \chi_\delta, \kappa_\delta) \in PC^2(0,T;\mathbb{R}^n) \times \mathbb{R}^n \times \mathbb{R}^J \times PC^1(0,T;\mathbb{R})$.* ◇

Proof If (I.1)-(I.7) are satisfied then, by Proposition 5.3 in Malanowski and Maurer (1998) there exists a unique solution and a unique Lagrange multiplier of (LO_δ). Hence, to complete the proof of the lemma, it is enough to show the appropriate regularity of the primal and dual variables.

By (45) and (49) we have

$$\begin{bmatrix} D^2_{uu}\widetilde{\mathcal{H}}[t] & D_u\psi[t]^* \\ D_u\psi[t] & 0 \end{bmatrix} \begin{bmatrix} v_\delta(t) \\ \kappa_\delta(t) \end{bmatrix} =$$
$$= -\begin{bmatrix} D^2_{ux}\widetilde{\mathcal{H}}[t]y_\delta(t) + D_u f[t]^* q(t) - a_3(t) \\ D_x\psi[t]y_\delta(t) - a_7(t) \end{bmatrix}$$

for $t \in (\omega'_j, \omega''_j),\ 1 \le j \le J$.

By (I.4) and (I.6) the matrix

$$K(t) := \begin{bmatrix} D^2_{uu}\widetilde{\mathcal{H}}[t] & D_u\psi[t]^* \\ D_u\psi[t] & 0 \end{bmatrix} \tag{54}$$

is nonsingular (see, e.g., Lemma 3.2 in Hager, 1979). Hence

$$\begin{bmatrix} v_\delta(t) \\ \kappa_\delta(t) \end{bmatrix} = -[K(t)]^{-1} \begin{bmatrix} D^2_{ux}\widetilde{\mathcal{H}}[t]y_\delta(t) + D_u f[t]^* q(t) - a_3(t) \\ D_x\psi[t]y_\delta(t) - a_7(t) \end{bmatrix},$$

and, in view of (I.1) as well as of the regularity of a_3 and a_7, we find that v_δ and κ_δ are of class C^1 on (ω'_j, ω''_j).

Similarly, by (I.6)

$$v_\delta(t) = -\left(D^2_{uu}\widetilde{\mathcal{H}}[t]\right)^{-1}\left[D^2_{ux}\widetilde{\mathcal{H}}[t]y_\delta(t) + D_u f[t]^* q(t) - a_3(t)\right],$$
$$\kappa_\delta = 0,$$

for all $t \in (0,T) \setminus \Omega_0$, i.e., v_δ and κ_δ are of class C^1 on any open subinterval of $(0,T) \setminus \Omega_0$.

Thus, $v_\delta \in PC^1(0,T;\mathbb{R}^m)$, $\kappa_\delta \in PC^1(0,T;\mathbb{R})$. These regularity results, together with equations (46) and (43), show that $y_\delta, q_\delta \in PC^2(0,T;\mathbb{R}^n)$. That completes the proof of the lemma. □

To ensure existence and uniqueness of τ satisfying (50) for any given

(I.8) (*Nontangential junction*)
$$\frac{d}{dt}\psi[\omega_j'(0)-] \neq 0, \quad \frac{d}{dt}\psi[\omega_j''(0)+] \neq 0, \quad 1 \leq j \leq J.$$

REMARK **4.3** *The role played by the* nontangential junction *condition is two-folded:*

1) *by this condition the Jacobian of $\mathcal{F}$ is regular (see, Proposition 4.4, below),*
2) *it ensures that the stationary points of $(\widetilde{\mathrm{O}}_h)$ are feasible for (O_h) (see, comments in Section 5. and Lemma 7.3 in Malanowski and Maurer, 1998).* ◇

PROPOSITION **4.4** *If (I.1)-(I.8) hold, then there exist neighborhoods $\widetilde{G} \subset G \subset H$ and $\mathcal{U} \subset U$ of h_0 and (z_0, ζ_0), respectively, such that for each $h \in \widetilde{G}$ there is a unique in $\mathcal{U}$ stationary point $(\widetilde{z}_h, \widetilde{\zeta}_h)$ of $(\widetilde{\mathrm{O}}_h)$ such that $\widetilde{u}_h, \widetilde{\nu}_h$ are continuous functions on $[0,T]$.*

The perturbed stationary point at $h = h_0 + d \in \widetilde{G}$ can be expressed by the following Taylor expansion

$$\begin{aligned}
\widetilde{x}_h &= x_0 + D_h x_0 d + o_{1,q}(d),\\
\widetilde{u}_h &= u_0 + D_h u_0 d + o_q(d),\\
\widetilde{\omega}_h &= \omega_0 + D_h \omega_0 d + o(d),\\
\widetilde{p}_h &= p_0 + D_h p_0 d + o_{1,q}(d),\\
\widetilde{\rho}_h &= \rho_0 + D_h \rho_0 d + o(d),\\
\widetilde{\varphi}_h &= \varphi_0 + D_h \varphi_0 d + o(d),\\
\widetilde{\nu}_h &= \nu_0 + D_h \nu_0 d + o_q(d),
\end{aligned} \tag{55}$$

where

$$\frac{\|o_{1,q}(d)\|_{1,q}}{\|d\|_H}, \ \frac{\|o_q(d)\|_q}{\|d\|_H}, \ \frac{|o(d)|}{\|d\|_H} \to 0 \quad \text{as } \|d\|_H \to 0 \quad \text{for } q \in [1,\infty). \tag{56}$$

The Fréchet differentials $(D_h x_0 d, D_h u_0 d)$ and $(D_h p_0 d, D_h \rho_0 d, D_h \varphi_0 d, D_h \nu_0 d)$ are given as the solution and Lagrange multipliers of the following linear-quadratic optimal control problem:

$$(\mathrm{LQ}_d) \qquad \begin{array}{ll}
\textit{Find } (y_d, v_d) \in PC^2(0,T;\mathbb{R}^n) \times PC^1(0,T;\mathbb{R}^m) & \textit{such that}\\
J(y_d, v_d, d) = \min_{(y,v)} J(y,v,d) & \\
\textit{subject to} & \\
\dot{y}(t) - D_x f[t]y(t) - D_u f[t]v(t) - D_h f[t]d = 0, & \\
y(0) - D_h\xi(h_0)d = 0, & \\
D_x\vartheta[\omega_j'(0)]y(w_j'(0)) + D_h\vartheta[\omega_j'(0)]d = 0, & 1 \leq j \leq J,\\
D_x\psi[t]y(t) + D_u\psi[t]v(t) + D_h\psi[t]d = 0, & t \in \Omega_0,
\end{array}$$

where

$$J(y,v,d) = \int_0^T \left\{ \frac{1}{2} \begin{bmatrix} y(t) \\ v(t) \end{bmatrix}^* \begin{bmatrix} D_{xx}^2\widetilde{\mathcal{H}}[t] & D_{xu}^2\widetilde{\mathcal{H}}[t] \\ D_{ux}^2\widetilde{\mathcal{H}}[t] & D_{uu}^2\widetilde{\mathcal{H}}[t] \end{bmatrix} \begin{bmatrix} y(t) \\ v(t) \end{bmatrix} - \right.$$
$$\left. - \begin{bmatrix} y(t) \\ v(t) \end{bmatrix}^* \begin{bmatrix} D_{xh}^2\widetilde{\mathcal{H}}[t] \\ D_{uh}^2\widetilde{\mathcal{H}}[t] \end{bmatrix} d \right\} dt.$$

The differentials of the junction points are given by:

$$\begin{aligned}
&D_h\omega_j'(0)d = -\left\{\tfrac{d}{dt}\psi[\omega_j'(0)-]\right\}^{-1}\times\\
&\times\left\{D_x\psi[\omega_j'(0)]y_d(\omega_j'(0)) + D_u\psi[\omega_j'(0)]v_d(\omega_j'(0)) + D_h\psi[\omega_j'(0)]d\right\},\\
&D_h\omega_j''(0)d = -\left\{\tfrac{d}{dt}\psi[\omega_j''(0)+]\right\}^{-1}\times\\
&\times\left\{D_x\psi[\omega_j''(0)]y_d(\omega_j''(0)) + D_u\psi[\omega_j''(0)]v_d(\omega_j''(0)) + D_h\psi[\omega_j''(0)]d\right\}.
\end{aligned}$$

◇

Proof By Lemma 4.2 and assumption (I.8) the Jacobian $D_{(z,\zeta)}\mathcal{F}(z_0,\zeta_0,h_0)$ is regular. Hence, by the implicit function theorem, there exists a neighborhood $\widetilde{G}$ of h_0 such that for each $h\in\widetilde{G}$ there is a locally unique solution of (41), which is a differentiable function of h in U and

$$D_h\begin{pmatrix} z_0\\ \zeta_0\end{pmatrix} = -\left(D_{(z,\zeta)}\mathcal{F}(z_0,\zeta_0,h_0)\right)^{-1}D_h\mathcal{F}(z_0,\zeta_0,h_0).$$

Using the above formula, together with definitions (39) and (40) of $\mathcal{F}$, and performing straightforward calculations, we find that D_hz_0d and $D_h\zeta_0d$ are characterized as in the formulation of the proposition.

We will show that $(\widetilde{z}_h,\widetilde{\zeta}_h)$ are the stationary points of $(\widetilde{\mathrm{O}}_h)$, i.e., that conditions (34) hold. To this end, it is enough to show that $\widetilde{u}_h$ and $\widetilde{\nu}_h$ are continuous at the junction points ω_j',ω_j''. Note that by (33), (35) and (36), at any junction point $\omega_j(h)=\omega_j'(h),\omega_j''(h)$, the following equations hold

$$\begin{aligned}
&\mathcal{G}(\widetilde{x}_h(\omega_j(h)),\widetilde{u}_h(\omega_j(h)-),\widetilde{p}_h(\omega_j(h)),\widetilde{\nu}_h(\omega_j(h)-),h)=0,\\
&\mathcal{G}(\widetilde{x}_h(\omega_j(h)),\widetilde{u}_h(\omega_j(h)+),\widetilde{p}_h(\omega_j(h)),\widetilde{\nu}_h(\omega_j(h)+),h)=0,
\end{aligned}\tag{57}$$

where

$$\mathcal{G}(x,u,p,\nu,h):=\begin{pmatrix} D_u\widetilde{\mathcal{H}}(x,u,p,\nu,h)\\ \psi(x,u,h)\end{pmatrix}.$$

Hence, both pairs $(\widetilde{u}_h(\omega_j(h)-),\widetilde{\nu}_h(\omega_j(h)-))$ and $(\widetilde{u}_h(\omega_j(h)+),\widetilde{\nu}_h(\omega_j(h)+))$ are solutions of the parametric equation

$$\mathcal{G}(\widetilde{x}_h(\omega_j(h)),u,\widetilde{p}_h(\omega_j(h)),\nu,h)=0 \tag{58}$$

corresponding to the same value $(\widetilde{x}_h(\omega_j(h)),\widetilde{p}_h(\omega_j(h)),h)$ of the parameter. At the reference point $(x_0(\omega_j(0)),u_0(\omega_j(0)),p_0(\omega_j(0)),\nu_0(\omega_j(0)),h_0)$, the Jacobian of (58) is given by nonsingular matrix $K(\omega_j(0))$ defined in (54). Hence, in a neighborhood of the reference point, (58) has a locally unique solution. On the other hand, $(\widetilde{z}_h,\widetilde{\zeta}_h)$ is a continuous function of h, so for $(h-h_0)$ sufficiently small, (58) must have a locally unique solution, i.e., $\widetilde{u}_h(\omega_j(h)-)=\widetilde{u}_h(\omega_j(h)+)$ and $\widetilde{\nu}_h(\omega_j(h)-)=\widetilde{\nu}_h(\omega_j(h)+)$. Thus, $(\widetilde{z}_h,\widetilde{\zeta}_h)$ is a stationary point of $(\widetilde{\mathrm{O}}_h)$. On the other hand, let us note that any stationary point $(\widetilde{z}_h,\widetilde{\zeta}_h)$ of $(\widetilde{\mathrm{O}}_h)$, such that

To complete the proof of the proposition it remains to show that the remainder terms in (55) satisfy conditions (56). Let us confine ourselves to $\widetilde{u}_h$. By (14) and by continuity of $\widetilde{u}_h(\cdot)$ we have

$$\widetilde{u}_h(t) = \alpha_h(t) + \Sigma_{j=1}^{2J} \gamma_h^j \mathbf{R}_h^j(t),$$

where α_h and γ_h as well as the junction points $\omega_j(h)$ are differentiable functions of h. Hence

$$\widetilde{u}_h(t) - u_0(t) = (\alpha_h(t) - \alpha_0(t)) + \left[\Sigma_{j=1}^{2J} \gamma_h^j \mathbf{R}_h^j(t) - \Sigma_{j=1}^{2J} \gamma_0^j \mathbf{R}_0^j(t)\right]. \tag{59}$$

In view of differentiability of α_h in $C^2(0, T; I\!R^m)$, we have

$$\alpha_h - \alpha_0 = D_h \alpha_0 (h - h_0) + o_{1,\infty}(h - h_0). \tag{60}$$

Consider the second term on the right-hand side of (59). Without loss of generality we can assume that $\omega_j(h_0) < \omega_j(h)$. In view of the definition of $\mathbf{R}_h^j$,

$$\begin{aligned} &\gamma_h^j \mathbf{R}_h^j(t) - \gamma_0^j \mathbf{R}_0^j(t) = \\ &= \begin{cases} 0 & \text{for } t \in [0, \omega_j(h_0)], \\ -\gamma_0^j(t - \omega_j(h_0)) & \text{for } t \in (\omega_j(h_0), \omega_j(h)), \\ \gamma_h^j(t - \omega_j(h)) - \gamma_0^j(t - \omega_j(h_0)) & \text{for } t \in [\omega_j(h), T]. \end{cases} \end{aligned} \tag{61}$$

Since γ_h^j and $\omega_j(h)$ are differentiable functions of h, we have

$$\begin{aligned} &\gamma_h^j(t - \omega_j(h)) - \gamma_0^j(t - \omega_j(h_0)) = \\ &= (t - \omega_j(h_0))(\gamma_h^j - \gamma_0^j) - \gamma_h^j(\omega_j(h) - \omega_j(h_0)) = \\ &= \left[(t - \omega_j(h_0)) D_h \gamma_0 - \gamma_0^j D_h \omega_j(h_0)\right](h - h_0) + \varrho(t, h - h_0), \end{aligned} \tag{62}$$

where $\frac{|\varrho(t,h-h_0)|}{|h-h_0|} \to 0$ as $|h - h_0| \to 0$ uniformly in $t \in [\omega_j(h), T]$.

On the other hand

$$-\gamma_0^j(t - \omega_j(h_0)) = [\gamma_h^j(t - \omega_j(h)) - \gamma_0^j(t - \omega_j(h_0))] - \gamma_h^j(t - \omega_j(h)). \tag{63}$$

We have

$$\left[\int_{\omega_j(h_0)}^{\omega_j(h)} |\gamma_h^j(t - \omega_j(h))|^q dt\right]^{\frac{1}{q}} = |\gamma_h^j|(q+1)^{-\frac{1}{q}} |\omega_j(h) - \omega_j(h_0)|^{1+\frac{1}{q}},$$

and, in view of uniform boundedness of γ_h^j and differentiability of $\omega_j(h)$,

$$\frac{\left[\int_{\omega_j(h_0)}^{\omega_j(h)} |\gamma_h^j(t - \omega_j(h))|^q dt\right]^{\frac{1}{q}}}{|h - h_0|} \to 0 \quad \text{as } |h - h_0| \to 0, \tag{64}$$

for any $q \in [1, \infty)$.

Combining (60) through (64) we obtain from (59)

$$\widetilde{u}_h = u_0 + D_h u_0 (h - h_0) + o_q(h - h_0) \quad \text{for all } q \in [1, \infty).$$

REMARK **4.5** *It was incorrectly stated in Malanowski and Maurer (1998) that convergence in (56) holds for $q = \infty$.* ◇

5. Differentiability of solutions

In this section we are going to show that, under additional *strict complementarity* conditions and a strengthened version of (I.7), the stationary points of $(\widetilde{\mathrm{O}}_h)$ become the solutions and Lagrange multipliers of (O_h), so by Proposition 4.4, they are Fréchet differentiable functions of the parameter. Both, the obtained results and the used procedure, are virtually the same as those in Sections 7 and 8 of Malanowski and Maurer (1998), so we confine ourselves to a short recollection of these results.

In Section 7 of Malanowski and Maurer (1998) the conditions are discussed under which the stationary points of $(\widetilde{\mathrm{O}}_h)$ become the (KKT) points of (O_h). In Lemma 7.3 therein, it is proved that

$$\vartheta(\widetilde{x}_h(t), h) < 0, \quad \text{for } t \in [0, T] \setminus \Omega_h,$$

i.e., $(\widetilde{x}_h, \widetilde{u}_h)$ is feasible for (O_h). In the proof, nontangential junction condition (I.8) plays the crucial role.

To show that the Lagrange multipliers φ_h, ν_h correspond to $\mu_h(0), \dot{\mu}_h$ an additional *strict complementarity* condition is introduced:

(I.9) (*Strict complementarity*)
$\dot{\mu}_0(\cdot) = \nu_0(\cdot)$ is positive and decreasing on each subinterval $(\omega_j'(0), \omega_j''(0))$, $1 \leq j \leq J$.

It is shown in Lemma 7.5 in Malanowski and Maurer (1998) that, for $(h - h_0)$ sufficiently small, condition (I.9) is also satisfied by ν_h, so that $\dot{\mu}_h = \nu_h$ is a Lagrange multiplier for (O_h). Thus, $(\widetilde{z}_h, \widetilde{\zeta}_h)$ corresponds to a (KKT)-point of (O_h).

To complete sensitivity analysis it remains to show that the (KKT) points of (O_h) are actually the solutions and Lagrange multipliers. To do that we have to strengthen coercivity conditions. Namely, instead of (I.7) we assume:

(I.7') Condition (51) holds for all $(z, w) \in Z^\infty$ satisfying (52).

Using (I.7') and Proposition 4.4 we easily find that there exists a neighborhood G_0 of h_0 such that for all $h \in G_0$ we have

$$\left((z, w), \begin{pmatrix} D^2_{xx}\mathcal{L}_h & D^2_{xu}\mathcal{L}_h \\ D^2_{ux}\mathcal{L}_h & D^2_{uu}\mathcal{L}_h \end{pmatrix}(z, w)\right) \geq \frac{\gamma}{2}(\|z\|^2_{1,2} + \|w\|^2_2), \tag{65}$$

for all $(z, w) \in Z^\infty$ such that

$$\dot{z}(t) - D_x f(x_h(t), u_h(t), h)z(t) - D_u f(x_h(t), u_h(t), h)w(t) = 0,$$

where $\mathcal{L}_h := \mathcal{L}(x_h, u_h, p_h, \rho_h, \mu_h, h)$. It is well known that by (65) there exist positive constants $\rho(h)$ and $c(h)$ such that

$$F(x, u, h) \geq F(x_h, u_h, h) + c(h)\|(x, u) - (x_h, u_h)\|_{Z^2}^2$$
$$\text{for all feasible } (x, u) \text{ such that } \|(x, u) - (x_h, u_h)\|_{Z^\infty} \leq \rho(h),$$

i.e., (x_h, u_h) is a locally unique local solution of (O_h). By (65) and by Proposition 4.4, the constants $\rho(h) = \rho$ and $c(h) = c$ can be chosen independent of $h \in G_0$. In view of Proposition 4.4 we can shrink G_0 so that, for all $h \in G_0$ we have

$$(x_h, u_h) \in \mathcal{Z}_0 := \{(x, u) \in Z^\infty \mid \|(x, u) - (x_0, u_0)\|_{Z^\infty} \leq \frac{\rho}{2}\}.$$

Thus, we arrive at our principal differentiability result:

THEOREM **5.1** *If assumptions (I.1)-(I.6), (I.7'), (I.8) and (I.9) hold, then there exist neighborhoods $G_0 \subset G \subset H$ and $\mathcal{Z}_0 \subset Z^\infty$ of h_0 and of (x_0, u_0), respectively, such that, for each $h \in G_0$, there exists a unique in $\mathcal{Z}_0$ solution (x_h, u_h) of (O_h) and unique associated Lagrange multipliers (p_h, ρ_h, μ_h) that are Fréchet differentiable functions of h in Z^q and W^q, respectively, for any $q \in [1, \infty)$. The respective differentials are given by the stationary points of the problem* (LQ_d).
◇

Acknowledgment

The author would like to express his gratitude to two anonimous referees, whose comments allowed to improve the text of the paper.

References

DONTCHEV, A.L., HAGER, W.W., POORE, A.B. and YANG, B. (1995) Optimality, stability and covergence in nonlinear control. *Appl. Math. Optim.*, **31**, 297-326.

DONTCHEV, A.L. and HAGER, W.W. (1998) Lipschitzian stability for state constrained nonlinear optimal control. To appear.

HAGER, W.W. (1979) Lipschitz continuity for constrained processes. *SIAM J. Contr. Optim.*, **17**, 321-338.

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

MALANOWSKI, K. (1992) Second order conditions and constraint qualifications in stability and sensitivity analysis of solutions to optimization problems in Hilbert spaces. *Appl. Math. Optim.*, **25**, 51-79.

MALANOWSKI, K. (1995) Stability and sensitivity of solutions to nonlinear

MALANOWSKI, K. and MAURER, H. (1996) Sensitivity analysis for parametric optimal control problems with control-state constraints. *Comp. Optim. Appl.*, **5**, 253-283.

MALANOWSKI, K. and MAURER, H. (1997) Sensitivity analysis for optimal control problems subject to higher order state constraints. *Angewandte Mathematik und Informatik*, Preprint 16/97-N, Universität Münster.

MALANOWSKI, K. and MAURER, H. (1998) Sensitivity analysis for state constrained optimal control problems. *Discrete and Continuous Dynamical Systems*, **4**, 241-272.

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

ROBINSON, S.M. (1980) Strongly regular generalized equations. *Math. Oper. Research*, **5**, 43-62.

STOER, J. and BULIRSCH, R. (1980) *Introduction to Numerical Analysis.* Springer-Verlag, New York.