HAL
open science

Pointwise second-order necessary optimality conditions and second-order sensitivity relations in optimal control

Hélène Frankowska, Daniel Hoehener

To cite this version:

Hélène Frankowska, Daniel Hoehener. Pointwise second-order necessary optimality conditions and second-order sensitivity relations in optimal control. Journal of Differential Equations, 2017, 262 (12), pp.5735-5772. 10.1016/j.jde.2017.02.013 . hal-03753942

HAL Id: hal-03753942

https://hal.science/hal-03753942

Submitted on 19 Aug 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Pointwise Second-Order Necessary Optimality Conditions and Second-Order Sensitivity Relations in Optimal Control*

Hélène Frankowska† Daniel Hoehener‡

December 5, 2016

Abstract

This paper is devoted to pointwise second-order necessary optimality conditions for the Mayer problem arising in optimal control theory. We first show that with every optimal trajectory it is possible to associate a solution $p(\cdot)$ of the adjoint system (as in the Pontryagin maximum principle) and a matrix solution $W(\cdot)$ of an adjoint matrix differential equation that satisfy a second-order transversality condition and a second-order maximality condition. These conditions seem to be a natural second-order extension of the maximum principle. We then prove a Jacobson like necessary optimality condition for general control systems and measurable optimal controls that may be only "partially singular" and may take values on the boundary of control constraints. Finally we investigate the second-order sensitivity relations along optimal trajectories involving both $p(\cdot)$ and $W(\cdot)$.

Key words: Optimal control; Second-order necessary optimality conditions; Singular control; Jacobson condition; Sensitivity relations; Second-order maximum principle.

AMS subject classification: 49K15.

1 Introduction

Consider the Mayer optimal control problem

$$\text{Minimize } \varphi(x(0), x(1)),$$

over trajectories $x(\cdot)$ of the control system

$$\begin{cases} \dot{x}(t) = f(t, x(t), u(t)), \quad u(t) \in U(t), \qquad \text{a.e. in } [0,1], \\ x(0) \in K_0, \end{cases}$$

where $f\colon [0,1] \times \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$, $\varphi : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$, $U(t) \subset \mathbb{R}^m$ for $t \in [0,1]$ and the set of initial point constraints $K_0 \subset \mathbb{R}^n$. Admissible controls are measurable selections of $U(\cdot)$. If $(\bar{x}, \bar{u})$ is an optimal trajectory-control pair, then it is well known that, under standard assumptions on f, the (unique) solution $\bar{p}(\cdot)$ of the adjoint system

$$-\dot{\bar{p}}(t) = f_x(t, \bar{x}(t), \bar{u}(t))^T \bar{p}(t), \ -\bar{p}(1) = \nabla_{x_2}\varphi(\bar{x}(0), \bar{x}(1)),$$

*The first author was partially supported by the Gaspard Monge Program for Optimisation and Operational Research, Jacques Hadamard Mathematical Foundation (FMJH). The second author was supported by the SNSF. Some results of this paper, without proofs, were announced at the 54th IEEE CDC Conference, [11].

†**Corresponding Author**, CNRS, Institut de Mathématiques de Jussieu - Paris Rive Gauche, UMR 7586, Sorbonne Universités, UPMC Univ Paris 06, Univ Paris Diderot, Sorbonne Paris Cité, Case 247, 4 Place Jussieu, 75252 Paris, France

‡Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA 02139-4307, USA

satisfies the maximality condition

$$\langle \bar{p}(t), f(t, \bar{x}(t), \bar{u}(t))\rangle = \max_{u\in U(t)} \langle \bar{p}(t), f(t, \bar{x}(t), u)\rangle, \qquad \text{a.e.}$$

and the transversality condition

$$\bar{p}(0) - \nabla_{x_1}\varphi(\bar{x}(0), \bar{x}(1)) \in N_{K_0}(\bar{x}(0)),$$

where $N_{K_0}(\bar{x}(0))$ is a normal cone to K_0 at $\bar{x}(0)$ and $\nabla_{x_1}\varphi$ denotes the partial derivative of φ with respect to the first variable (for a moment we do not make precise which normal cones we consider, to avoid technicalities in the introduction). This maximality condition however, in general, does not single out $\bar{u}(t)$ because the maximum may be attained by several $u \in U(t)$. Let us denote the set of all maximizers by

$$\overline{U}(t) := \left\{ z \in U(t) \;\middle|\; \langle \bar{p}(t), f(t, \bar{x}(t), z)\rangle = \max_{u\in U(t)} \langle \bar{p}(t), f(t, \bar{x}(t), u)\rangle \right\}.$$

An important ongoing research in optimal control theory concerns refining the set $\overline{U}(t)$ in order to restrict further the candidates for optimality. Several necessary optimality conditions can be found in the control literature of the 70ies including the Goh and the Jacobson conditions, [16, 21]. These results were usually assuming that U is time independent, that optimal control $\bar{u}$ takes values in the interior of U and that it is also piecewise continuous, which is often not the case. Both the Goh and the Jacobson conditions are pointwise conditions. The Goh necessary condition was generalized in various ways by many authors, see for instance [1, 14, 28] and the references therein.

There is also a rich literature on second-order optimality conditions in integral form which hold under less restrictive assumptions than the pointwise conditions, see for instance [4, 5, 9, 13, 15, 18, 23, 24, 25] and the references therein. However, conditions in integral form are not simple to check as they must be satisfied on a set of functions instead of pointwise.

The aim of our work is to prove a second-order maximum principle and the Jacobson inequality in a very general situation that we describe now. To simplify the discussion, let us assume that $\varphi(x_1, x_2) = \varphi_1(x_1) + \varphi_2(x_2)$ for all x_1, $x_2 \in \mathbb{R}^n$, where $\varphi_i : \mathbb{R}^n \to \mathbb{R}$ are twice continuously differentiable functions. Recall that the (unmaximized) *Hamiltonian* $\mathcal{H}\colon [0,1] \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ is defined by,

$$\mathcal{H}(t, x, p, u) = \langle p, f(t, x, u)\rangle . \tag{1.1}$$

Set $[t] := (t, \bar{x}(t), \bar{p}(t), \bar{u}(t))$ and denote by $\mathcal{H}_{xx}[t]$ the partial Hessian of $\mathcal{H}$ with respect to x at $[t]$. The notations $\mathcal{H}_{xp}[t]$ and $\mathcal{H}_{px}[t]$ are similarly defined. Consider the solution W of the matrix differential equation

$$\begin{cases} \dot{W}(t) = -\mathcal{H}_{px}[t]W(t) - W(t)\mathcal{H}_{xp}[t] - \mathcal{H}_{xx}[t], & \text{a.e.} \\ W(1) = -\varphi_2''(\bar{x}(1)). \end{cases}$$

Our results imply, in particular, that $\bar{p}$, W satisfy the following second-order maximality condition

$$\max_{u\in\overline{U}(t)} \left\langle \left(f_x(t, \bar{x}(t), u)^T - f_x[t]^T\right)\bar{p}(t) + W(t)\left(f(t, \bar{x}(t), u) - f[t]\right), f(t, \bar{x}(t), u) - f[t]\right\rangle = 0,$$

and the second-order transversality condition

$$W(0) - \varphi_1''(\bar{x}(0)) \in N^{(2)}_{K_0}\big(\bar{x}(0); \bar{p}(0) - \nabla\varphi_1(\bar{x}(0))\big),$$

where $[t] = (t, \bar{x}(t), \bar{u}(t))$ and $N^{(2)}_{K_0}$ denotes an appropriate second-order normal cone to K_0.

The above two relations are similar to those of the Pontryagin maximum principle but use second-order normals and second-order adjoints instead of the first-order ones. An earlier result of this nature was obtained in [22] for a Bolza problem with a fixed initial point and time independent control sets. However in [22] some ad hoc relaxation arguments and a special type linearization of control system are used, which would not allow to treat more complex problems involving state constraints. Our proof is more direct, since we use the classical Filippov-Ważewski relaxation theorem in the form of differential inclusion and a linearization of this differential inclusion. Such techniques allow to investigate state constrained problems as well, see [12, 18, 19]. Furthermore, thanks to the relaxation theorem, $(f_x(t,\bar{x}(t),u), f(t,\bar{x}(t),u))$ may be replaced by convex combinations of such elements for $u \in \overline{U}(t)$, see Theorem 3.1 in Section 3.

When, in addition, f is differentiable with respect to u, then the second-order necessary optimality condition from [13, 18] in the form of an integral inequality can be extended to problems with initial point constraints. Actually we improve [13, Theorem 4.1] by allowing a larger set of admissible second-order variations of $\bar{u}$, see Theorem 6.1 from Section 6. We would like to underline here that Theorem 3.1 is deduced from the integral inequality (6.2) that is different from the one in Theorem 6.1. More generally, Theorems 3.1 and 6.1 are not comparable between them : the first one holds true without the differentiability assumptions on data with respect to u and concerns controls maximizing the Hamiltonian, while the second one supposes differentiability of data with respect to u and concerns critical variations of controls.

In the difference with [13, 18], admissible second-order variations are defined here by using closed jets to $U(t)$ at $\bar{u}(t)$, instead of second-order tangents whose graph, being not closed in general, may be difficult to apply to deduce pointwise conditions.

The obtained integral inequality implies, in turn, a pointwise inequality, leading to a very general formulation of the Jacobson type necessary optimality condition for measurable optimal controls, see Theorem 3.6 below. Such inequality was previously derived for affine with respect to control systems, time independent (box-type) U, piecewise continuous optimal controls taking values in the interior of U and $\mathcal{H}_u[\cdot] = 0$. Since piecewise continuity of optimal controls does not hold in general, and typically optimal controls take values on the boundary of U, our extension is a major breakthrough avoiding to impose structural assumptions on optimal controls.

In particular, if the boundary of $U(t)$ is sufficiently smooth and $\bar{u}(\cdot)$ is singular in the sense that $\mathcal{H}_u[t] = 0$, $\mathcal{H}_{uu}[t] = 0$ a.e., our generalisation of the Jacobson inequality implies that for a.e. $t \in [0,1]$,

$$\left\langle f_u[t]^T \left(\mathcal{H}_{ux}[t] + W(t) f_u[t]\right) u, u \right\rangle \leq 0,$$

for every u tangent to $U(t)$ at $\bar{u}(t)$. Thus, whenever $\bar{u}(t)$ belongs to the interior of the set $U(t)$, we get the inequality from [20]: $f_u[t]^T \left(\mathcal{H}_{ux}[t] + W(t) f_u[t]\right) \leq 0$. Actually the second-order maximality condition and the above inequality look somewhat similar. Still Theorem 3.1 does not imply Theorem 3.6.

To complete our study, we also investigate the sensitivity relations for the value function $V : [0,1] \times \mathbb{R}^n \to \mathbb{R}$ of the Mayer problem. Similarly to the first-order case where $-\bar{p}(t)$ is an element of the superdifferential of $V(t,\cdot)$ at $\bar{x}(t)$, we show that $(-\bar{p}(t), -W(t))$ belongs to the second-order superjet of $V(t,\cdot)$ at $\bar{x}(t)$ for all $t \in [0,1]$. In the difference with [6], where the Riccati equation was used, our matrix differential equation does not contain a quadratic term. For this reason $W(\cdot)$ does not escape to infinity, unlike in [6], and the sensitivity relations derived here hold true on the whole interval $[0,1]$. These sensitivity relations are in turn less precise than those in [6] and do not allow to investigate the local C^2 regularity of the value function.

The outline of the paper is as follows. In Section 2 we provide some preliminaries and formulate the main assumptions. Section 3 presents our main results on second-order necessary optimality conditions, while Section 4 deals with sensitivity relations. Section 5 is devoted to the second-order variational equations. Proofs are given in Sections 6 and 7.

2 Problem statement and notations

2.1 Preliminaries

Throughout the paper $n, m, r, s \in \mathbb{N}$ are integers and $\mathbb{R}_+$, $\mathbb{Q}_+$ are nonnegative reals and rationals, respectively. We denote the norm in $\mathbb{R}^n$ by $|\cdot|$, $\langle \cdot, \cdot \rangle$ is the standard inner product and $B := \{x \in \mathbb{R}^n \mid |x| < 1\}$ stands for the open unit ball. For a set $K \subset \mathbb{R}^n$, $\overline{K}$ is its closure, $\overline{co}\, K$ is its closed convex hull, $K^\ominus$ is its (negative) polar cone, i.e. $K^\ominus := \{q \in \mathbb{R}^n \mid \langle q, k \rangle \leq 0 \quad \forall\, k \in K\}$ and $K^\perp$ is its orthogonal complement, i.e. $K^\perp := \{q \in \mathbb{R}^n \mid \langle q, k \rangle = 0 \quad \forall\, k \in K\}$. Given a matrix A of dimension $n \times m$ we denote by A^T its transpose, a matrix of dimension $m \times n$. By $\mathbf{S}(n)$ we denote the set of symmetric $n \times n$-matrices and by $\mathbf{S}^-(n)$ the subset of symmetric seminegative $n \times n$-matrices. Finally, for Banach spaces X and Y and a bounded linear map $L\colon X \to Y$, the norm of L is the operator norm, i.e. $\|L\| := \sup_{\|x\|=1} \|Lx\|$. Below $W^{1,1}([0,1];\mathbb{R}^n)$ stands for the space of absolutely continuous maps from $[0,1]$ to $\mathbb{R}^n$ and $\mathcal{C}^k(\mathbb{R}^n;\mathbb{R}^m)$ for the space of k-times continuously differentiable maps from $\mathbb{R}^n$ to $\mathbb{R}^m$. The space of continuous maps is denoted by $\mathcal{C}(\mathbb{R}^n;\mathbb{R}^m)$. For $1 \leq p \leq \infty$, $L^p([0,1];\mathbb{R}^n)$ denotes the Lebesgue space, in particular, for $u \in L^p([0,1];\mathbb{R}^n)$, $|u|^p$ is integrable if $p < \infty$, respectively u is essentially bounded if $p = \infty$. The usual norm of $L^p([0,1];\mathbb{R}^n)$ is denoted by $\|\cdot\|_p$.
For a mapping $f : \mathbb{R}^n \to \mathbb{R}^m$, partial derivatives (if they exist) are indicated by a subscript referring to the differentiation variable, hence for instance $f_x(x_0, u_0) := \frac{\partial}{\partial x} f(x_0, u_0)$. Second-order partial derivatives are indicated by a double subscript, i.e. $f_{xu}(x_0, u_0) := \frac{\partial^2}{\partial x \partial u} f(x_0, u_0)$. Moreover, we will simplify the notation for the bilinear form $f_{xu}(x_0, u_0)(y, u)$ by writing $f_{xu}(x_0, u_0)yu$.
Consider a subset $K \subset \mathbb{R}^n$. A map $\varphi\colon [0,1] \times K \to \mathbb{R}^m$ is a *Carathéodory map* if $\varphi(\cdot, x)$ is Lebesgue measurable for all $x \in K$ and $\varphi(t, \cdot)$ is continuous on K for almost all $t \in [0,1]$.
Next, we recall some definitions concerning tangent sets. The distance between a point $x \in \mathbb{R}^n$ and a subset $K \subset \mathbb{R}^n$ is defined by $\mathrm{dist}_K(x) := \inf_{k \in K} |x - k|$. Let $\mathcal{T}$ be a metric space and $\{K_\tau\}_{\tau \in \mathcal{T}}$ be a family of subsets of $\mathbb{R}^n$. The lower limit of $\{K_\tau\}$ at $\tau_0 \in \mathcal{T}$ in the Peano-Kuratowski sense is given by

$$\operatorname*{Liminf}_{\tau \to \tau_0} K_\tau := \left\{ v \in \mathbb{R}^n \;\middle|\; \lim_{\tau \to \tau_0} \mathrm{dist}_{K_\tau}(v) = 0 \right\}.$$

First- and second-order adjacent subsets are defined in the following way.

Definition. Let $K \subset \mathbb{R}^n$ and $x \in K$. The *adjacent cone* to K at x is the set,

$$T_K^\flat(x) := \operatorname*{Liminf}_{h \to 0+} \frac{K - x}{h}.$$

Further, let $u \in \mathbb{R}^n$. The *second-order adjacent subset* to K at (x, u) is the set,

$$T_K^{\flat(2)}(x, u) := \operatorname*{Liminf}_{h \to 0+} \frac{K - x - hu}{h^2}.$$

Observe that if $T_K^{\flat(2)}(x, u) \neq \emptyset$, then $u \in T_K^\flat(x)$. The set $T_K^{\flat(2)}(x, u)$ is closed for every u, but, in general, the set $\left\{(u, v) \in \mathbb{R}^n \times \mathbb{R}^n \,|\, v \in T_K^{\flat(2)}(x, u)\right\}$ is not closed. We refer to [27, pp. 638-641] for historical comments and a huge bibliography on second order tangents and second order derivatives of functions. See also [2] for various definitions of first and higher order tangent cones and their properties.
The dual notion to tangent cones are normal cones. Here we use the normal cone $N_K^\flat(x) := T_K^\flat(x)^\ominus$. It is closed and convex, and coincides with the normal cone of convex analysis when K is convex.
With every $q \in N_K^\flat(x)$ we associate the set of second-order "normals":

$$N_K^{\flat(2)}(x; q) := \left\{ Q \in \mathbf{S}(n) \;\middle|\; \langle q, w \rangle + \frac{1}{2} \langle Qy, y \rangle \leq 0, \quad \forall\, y \in T_K^\flat(x) \cap \{q\}^\perp, \quad \forall w \in T_K^{\flat(2)}(x, y) \right\}.$$

By convention we set $N_K^{\flat(2)}(x;q) = \mathbf{S}(n)$ whenever $T_K^{\flat(2)}(x,y) = \emptyset$ for all $y \in T_K^\flat(x) \cap \{q\}^\perp$. Observe that $N_K^{\flat(2)}(x;q)$ is a closed and convex, possibly empty, set. Clearly $N_K^{\flat(2)}(x;q) + \mathbf{S}^-(n) = N_K^{\flat(2)}(x;q)$. Observe that if $T_K^\flat(x) \cap \{q\}^\perp = \{0\}$, then $N_K^{\flat(2)}(x;q) = \mathbf{S}(n)$. Moreover $\mathbf{S}^-(n) \subset N_K^{\flat(2)}(x;0)$.

Remark 2.1. To the best of our knowledge this definition of second-order normals never appeared in the literature before [11], where we used it to express the second-order transversality conditions.
A second-order normal cone was defined in [3] (without using second-order tangents) by

$$\mathcal{N}_K^{(2)}(x) := \left\{ (q,Q) \in \mathbb{R}^n \times \mathbf{S}(n) \;\middle|\; \langle q, k-x\rangle + \frac{1}{2}\langle Q(k-x), k-x\rangle \le o(|k-x|^2),\ \forall k \in K \right\},$$

where $\lim_{h\to 0+} o(h^2)/h^2 = 0$. For every $(q,Q) \in \mathcal{N}_K^{(2)}(x)$, the vector q is a proximal normal to K at x. We would like to underline that for every proximal normal q to K at x there exists $Q \in \mathbf{S}(n)$ such that $(q,Q) \in \mathcal{N}_K^{(2)}(x)$.
Observe that if $(q,Q) \in \mathcal{N}_K^{(2)}(x)$, then $q \in N_K^\flat(x)$ and $Q \in N_K^{\flat(2)}(x;q)$. It is not difficult to construct an example of K, x and $q \neq 0$ such that $N_K^{\flat(2)}(x;q)$ is nonempty and, at the same time, $\mathcal{N}_K^{(2)}(x) = \{0\}$. The set $N_K^{\flat(2)}(x;q)$ is better adapted to express the second-order optimality conditions below.

Example 2.2. Consider twice continuously differentiable functions $\psi_1, \dots, \psi_r \colon \mathbb{R}^n \to \mathbb{R}$ and

$$K = \bigcap_{j=1}^r \{x \in \mathbb{R}^n \mid \psi_j(x) \le 0\}.$$

Let $x \in K$ and denote by $I(x)$ the set of all active indices, i.e. $j \in I(x)$ if and only if $\psi_j(x) = 0$. We assume that $\{\nabla\psi_j(x) : j \in I(x)\}$ are positively independent or, equivalently, $0 \notin \operatorname{co}\{\nabla\psi_j(x) : j \in I(x)\}$. Then it is well known that

$$T_K^\flat(x) = \{y \in \mathbb{R}^n \mid \langle \nabla\psi_j(x), y\rangle \le 0 \quad \forall\, j \in I(x)\}, \quad N_K^\flat(x) = \sum_{j\in I(x)} \mathbb{R}_+ \nabla\psi_j(x).$$

Furthermore,

$$T_K^{\flat(2)}(x,y) = \left\{ z \in \mathbb{R}^n \mid \langle \nabla\psi_j(x), z\rangle + \frac{1}{2}\langle \psi_j''(x)y, y\rangle \le 0,\ \forall\, j \in I_1(y) \right\}, \tag{2.1}$$

where $I_1(y) = \{j \in I(x) : \langle \nabla\psi_j(x), y\rangle = 0\}$. Moreover $T_K^{\flat(2)}(x,y) = \mathbb{R}^n$ whenever $y \in T_K^\flat(x)$ and $I_1(y) = \emptyset$. Notice that this implies that under our assumptions, $T_K^{\flat(2)}(x,y) \neq \emptyset$ for all $y \in T_K^\flat(x)$.
Fix any $0 \neq q \in N_K^\flat(x)$. In the same way as in [13, Section 2] it follows that for every $y \in T_K^\flat(x) \cap \{q\}^\perp$ there exist $z_y \in T_K^{\flat(2)}(x,y)$ and $\alpha_j(y) \ge 0$, in general depending on y, such that

$$q = \sum_{j=1}^r \alpha_j(y)\nabla\psi_j(x) = \sum_{j\in\Pi(y)} \alpha_j(y)\nabla\psi_j(x), \quad \langle q, z_y\rangle = -\frac{1}{2}\sum_{j=1}^r \alpha_j(y)\psi_j''(x)yy,$$

where $\Pi(y) \subset I_1(y)$ is the set of all $j \in I_1(y)$ such that $\alpha_j(y) > 0$. Then for any $Q \in N_K^{\flat(2)}(x;q)$ we have

$$\langle q, z_y\rangle + \frac{1}{2}Qyy \le 0 = \langle q, z_y\rangle + \frac{1}{2}\sum_{j=1}^r \alpha_j(y)\psi_j''(x)yy.$$

Consequently $Qyy \le \sum_{j=1}^r \alpha_j(y)\psi_j''(x)yy$.

More can be said if $\{\nabla\psi_j(x) : j \in I(x)\}$ are linearly independent. Then $\{\alpha_j\}_{j=1}^r$ are uniquely defined and therefore they do not depend on y. Hence from (2.1) we deduce that

$$\sum_{j=1}^{r} \alpha_j \psi_j''(x) \in N_K^{\flat(2)}(x;q).$$

Consequently, $Q \in N_K^{\flat(2)}(x;q)$ *if and only if* $Q \leq \sum_{j=1}^r \alpha_j \psi_j''(x)$ on $T_K^\flat(x) \cap \{q\}^\perp$, in the sense that *for every* $y \in T_K^\flat(x) \cap \{q\}^\perp$, we have $Qyy \leq \left(\sum_{j=1}^r \alpha_j \psi_j''(x)\right) yy$.
In particular, if $r = 1$, then $N_K^\flat(x) = \mathbb{R}_+ \nabla\psi_1(x)$, $T_K^\flat(x) \cap \{\nabla\psi_1(x)\}^\perp = \{\nabla\psi_1(x)\}^\perp$ and

$$N_K^{\flat(2)}(x; \nabla\psi_1(x)) = \psi_1''(x) + S^-(\nabla\psi_1(x)),$$

where $S^-(\nabla\psi_1(x))$ is the set of symmetric matrices that are seminegative on $\{\nabla\psi_1(x)\}^\perp$.

Definition. Let $f\colon \mathbb{R}^n \to [-\infty, +\infty]$ be an extended real-valued function and $x \in \operatorname{dom}(f)$. A pair $(q, Q) \in \mathbb{R}^n \times \mathbf{S}(n)$ is said to be a *superjet* of f at x if for some $\delta > 0$ and for all $y \in x + \delta B$,

$$f(y) \leq f(x) + \langle q, y - x\rangle + \frac{1}{2}\langle Q(y-x), y-x\rangle + o(|y-x|^2). \tag{2.2}$$

The set of all superjets of f at x is denoted by $J^{2,+}f(x)$. Similarly, $(q, Q) \in \mathbb{R}^n \times \mathbf{S}(n)$ is a *subjet* of f at x if (2.2) holds with $\leq$ replaced by $\geq$. The set of all subjets of f at x is denoted by $J^{2,-}f(x)$.

If f is equal to the indicator function of $K \subset \mathbb{R}^n$ and $x \in K$, then $J^{2,-}f(x) = \mathcal{N}_K^{(2)}(x)$. If f is differentiable at x and semiconcave on a neighborhood of x with the semiconcavity constant c, then $(\nabla f(x), c\,\mathrm{Id}) \in J^{2,+}f(x)$, see [6, Prop. 2.6, Rem. 2.5]. Notice that for any $f \in \mathcal{C}^1(\mathbb{R}^n;\mathbb{R})$ that is twice differentiable at $x \in \mathbb{R}^n$, $(\nabla f(x), f''(x)) \in J^{2,+}f(x)$.
We end this subsection by recalling the definitions of directional derivatives of set-valued maps.

Definition. Let $F\colon \mathbb{R}^n \rightsquigarrow \mathbb{R}^m$ be a set-valued map, locally Lipschitz around some $x \in \mathbb{R}^n$ and let $y \in F(x)$. The *adjacent derivative* $dF(x,y)$ is the set-valued map defined by,

$$dF(x,y)(u) \coloneqq \operatorname*{Liminf}_{h\to 0+} \frac{F(x+hu) - y}{h}, \quad \forall\, u \in \mathbb{R}^n.$$

For $v_1 \in dF(x,y)(u_1)$, define the set-valued map $d^2F(x,y,u_1,v_1)$ by

$$d^2F(x,y,u_1,v_1)(u_2) \coloneqq \operatorname*{Liminf}_{h\to 0+} \frac{F(x+hu_1+h^2u_2) - y - hv_1}{h^2}, \quad \forall\, u_2 \in \mathbb{R}^n.$$

Below, for a set-valued map $[0,1]\times\mathbb{R}^n \ni (t,x) \rightsquigarrow F(t,x)$ and $t_0 \in [0,1]$ such that $F(t_0,\cdot)$ is Lipschitz on a neighborhood of some $x_0 \in \mathbb{R}^n$, we denote the partial derivatives with respect to the second variable by a subscript x. That is $d_xF(t_0,x_0,y_0)$ is equal to the adjacent derivative of $F(t_0,\cdot)$ at (x_0,y_0) for any $y_0 \in F(t_0,x_0)$.
Similarly for a function $[0,1]\times\mathbb{R}^n \ni (t,x) \mapsto f(t,x)$, $J_x^{2,+}f(t,x)$ and $J_x^{2,-}f(t,x)$ denote respectively the sets of superjets and subjets of $f(t,\cdot)$ at x.

2.2 Problem statement and assumptions

Consider a deterministic control system of the form

$$\begin{cases} \dot{x}(t) = f(t, x(t), u(t)), \quad u(t) \in U(t), \qquad \text{a.e. in } [0,1], \\ x(0) \in K_0, \end{cases} \tag{C}$$

where $f\colon [0,1] \times \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$, the set-valued map $U\colon [0,1] \rightsquigarrow \mathbb{R}^m$ and $K_0 \subset \mathbb{R}^n$ are given. Later on it will be convenient to use the following notation:

$$F(t,x) := \overline{co}\,\{f(t,x,u) \mid u \in U(t)\}.$$

A Lebesgue measurable map $u\colon [0,1] \to \mathbb{R}^m$ such that $u(t) \in U(t)$ a.e. in $[0,1]$ is called *control*. The set of all controls is denoted by

$$\mathcal{U} := \{u\colon [0,1] \to \mathbb{R}^m \text{ is Lebesgue measurable} \mid u(t) \in U(t) \text{ a.e.}\}.$$

Under standard assumptions on f, for any control $u \in \mathcal{U}$ and any initial state $x_0 \in K_0$, there exists a unique solution $x \in W^{1,1}([0,1];\mathbb{R}^n)$ of the ordinary differential equation in (C) satisfying $x(0) = x_0$, called a *state trajectory*. For a fixed initial state $x_0 \in K_0$, the tuple (x,u) consisting of a control $u \in \mathcal{U}$ and the corresponding state trajectory $x \in W^{1,1}([0,1];\mathbb{R}^n)$ is called a process. The set of processes and the set of state trajectories are given by,

$$\begin{aligned} \mathcal{P}_s(x_0) &:= \left\{(x,u) \in W^{1,1}([s,1];\mathbb{R}^n) \times \mathcal{U} \;\middle|\; x(s) = x_0 \text{ and } \dot{x}(t) = f(t,x(t),u(t)) \text{ for a.e. } t \in [s,1]\right\}, \\ \mathcal{S}_s(x_0) &:= \left\{x \in W^{1,1}([s,1];\mathbb{R}^n) \;\middle|\; \exists u \in \mathcal{U} \text{ such that } \; (x,u) \in \mathcal{P}_s(x_0)\right\}. \end{aligned}$$

To simplify the notation we set $\mathcal{P}(x_0) := \mathcal{P}_0(x_0)$ and $\mathcal{S}(x_0) := \mathcal{S}_0(x_0)$. Let X, Y be vector spaces, $V \subset X$ and h be a single- or set-valued map from X to Y. Then $h(V) := \bigcup_{v \in V} h(v)$. Thus for instance $\mathcal{P}(K_0) := \bigcup_{x_0 \in K_0} \mathcal{P}(x_0)$. The objective of the present work is to study second-order necessary optimality conditions for the Mayer optimal control problem:

$$\underset{x \in \mathcal{S}(K_0)}{\text{Minimize}} \;\; \varphi(x(0), x(1)), \tag{P}$$

where the function $\varphi\colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ that associates to any $(x_1, x_2) \in \mathbb{R}^n \times \mathbb{R}^n$ a real is given. There are several notions of local minimizers of problem (P). Here we are interested by strong local minimizers.

Definition. A process $(\bar{x}, \bar{u}) \in \mathcal{P}(K_0)$ is a *strong local minimizer* if there exists $\bar{\rho} > 0$ such that for all $(x,u) \in \mathcal{P}(K_0)$,

$$\|x - \bar{x}\|_\infty \le \bar{\rho} \implies \varphi(x(0), x(1)) \ge \varphi(\bar{x}(0), \bar{x}(1)). \tag{2.3}$$

It is a *weak local minimizer* if the left-hand side of (2.3) can be replaced by $|x(0) - \bar{x}(0)| + \|u - \bar{u}\|_\infty \le \bar{\rho}$.

Given a strong local minimizer $(\bar{x}, \bar{u}) \in \mathcal{P}(K_0)$ of problem (P) we impose the following assumptions:

Assumptions: There exists $\rho > 0$ such that the following properties are satisfied :

(A1) (a) For all $(x,u) \in \mathbb{R}^n \times \mathbb{R}^m$, $f(\cdot, x, u)$ is measurable. For all $(t,x) \in [0,1] \times \mathbb{R}^n$, $f(t,x,\cdot)$ is continuous, $f(t,x,U(t))$ is closed and there exists $a_1 > 0$ such that $\sup_{u \in U(t)} |f(t,x,u)| \le a_1(|x| + 1)$. Moreover, for every $R > 0$, there exists an integrable map $k_R : [0,1] \to \mathbb{R}_+$ such that for a.e. $t \in [0,1]$,

$$|f(t,x,u) - f(t,y,u)| \le k_R(t)\,|x - y|, \qquad \forall\, x,\, y \in RB, \quad \forall\, u \in U(t);$$

(b) For all $t \in [0,1]$ and $u \in U(t)$, $f(t,\cdot,u)$ is twice continuously differentiable on $\bar{x}(t)+\rho B$, $f_x(t,\bar{x}(t),\cdot)$ is continuous on $U(t)$, and there exists $a_2 > 0$ such that $\sup_{u\in U(t)} \|f_x(t,\bar{x}(t),u)\| \leq a_2$ for a.e. $t \in [0,1]$. Furthermore, there exists an integrable map $l : [0,1] \to \mathbb{R}_+$ such that for a.e. $t \in [0,1]$

$$\|f_x(t,x,u) - f_x(t,y,u)\| \leq l(t)\,|x-y|\,, \qquad \forall\, x,y \in \bar{x}(t)+\rho B, \quad \forall\, u \in U(t);$$

(c) $U\colon [0,1] \rightsquigarrow \mathbb{R}^m$ is measurable with closed nonempty images;

(d) φ is differentiable.

3 Second-order necessary optimality conditions

We state here pointwise second-order necessary optimality conditions for the Mayer problem (P).

3.1 General case

For optimal control problems, first-order necessary optimality conditions are given by the celebrated maximum principle [26] which states that if $(\bar{x},\bar{u}) \in \mathcal{P}(K_0)$ is a strong local minimizer, then the (unique) solution $\bar{p} \in W^{1,1}([0,1];\mathbb{R}^n)$ of the so-called *adjoint equation*

$$\begin{cases} -\dot{p}(t) = f_x(t,\bar{x}(t),\bar{u}(t))^T p(t), & \text{a.e.} \\ -p(1) = \nabla_{x_2}\varphi(\bar{x}(0),\bar{x}(1)), \end{cases} \tag{3.1}$$

satisfies the transversality condition

$$\bar{p}(0) - \nabla_{x_1}\varphi(\bar{x}(0),\bar{x}(1)) \in N^\flat_{K_0}(\bar{x}(0)), \tag{3.2}$$

and the maximality condition

$$\mathcal{H}(t,\bar{x}(t),\bar{p}(t),\bar{u}(t)) = \max_{u\in U(t)} \mathcal{H}(t,\bar{x}(t),\bar{p}(t),u), \qquad \text{a.e.,} \tag{3.3}$$

where the *Hamiltonian* $\mathcal{H}$ is defined by (1.1). Observe that the maximality condition can be equivalently formulated in the following, less traditional way:

$$\langle \bar{p}(t), v\rangle \leq 0, \qquad \forall\, v \in T_{F(t,\bar{x}(t))}(\dot{\bar{x}}(t)), \ \text{ a.e.} \tag{3.4}$$

Let us also mention that using $N^\flat_{K_0}(\bar{x}(0))$ in the transversality condition (3.2) is not very common: in [29, Prop. 6.4.4] a larger limiting normal cone is used. However, in the absence of final-point constraints, it is not difficult to show that the above stronger transversality condition (3.2) holds true. Moreover in (3.2), $N^\flat_{K_0}(\bar{x}(0))$ may be replaced by the negative polar of the contingent cone to K at $\bar{x}(0)$. Not to complicate the discussion of second-order conditions, we stick to the adjacent tangents.
Fix a triple $(\bar{x},\bar{u},\bar{p}) \in \mathcal{P}(K_0) \times W^{1,1}([0,1];\mathbb{R}^n)$ such that $(\bar{x},\bar{u})$ is a strong local minimizer and $\bar{p}$ is the corresponding adjoint state, i.e. $\bar{p}$ satisfies (3.1) - (3.3). To simplify the notation $[t]$ replaces $(t,\bar{x}(t),\bar{p}(t),\bar{u}(t))$ when evaluating the Hamiltonian $\mathcal{H}$ or $(t,\bar{x}(t),\bar{u}(t))$ when evaluating the dynamics f. The closed set of *maximizing controls* at $t \in [0,1]$ and the set of *critical initial directions* are defined respectively by

$$\begin{aligned} \overline{U}(t) &:= \left\{ z \in U(t) \;\middle|\; \mathcal{H}(t,\bar{x}(t),\bar{p}(t),z) = \max_{u\in U(t)} \mathcal{H}(t,\bar{x}(t),\bar{p}(t),u) \right\}, \\ \Gamma_0 &:= \left\{ y_0 \in T^\flat_{K_0}(\bar{x}(0)) \;\middle|\; \langle \bar{p}(0) - \nabla_{x_1}\varphi(\bar{x}(0),\bar{x}(1)), y_0\rangle = 0 \right\}. \end{aligned} \tag{3.5}$$

By [2, Thm. 8.2.9] the set-valued map $\overline{U}(\cdot)$ is measurable and (3.3) implies that $\bar{u}(t) \in \overline{U}(t)$ a.e. It will be convenient to have the following notation:

$$\overline{D}(t) := \text{co}\left\{(\Delta f[t,u], \Delta f_x[t,u]) \;\middle|\; u \in \overline{U}(t)\right\},$$

where $\Delta f[t,u] := f(t,\bar{x}(t),u) - f[t]$ and $\Delta f_x[t,u] := f_x(t,\bar{x}(t),u) - f_x[t]$. Finally, let $Y(\cdot)$ denote the (fundamental) solution of the linear system

$$\begin{cases} \dot{Y}(t) = f_x[t]Y(t), & \text{a.e. in } [0,1], \\ Y(0) = \text{Id}. \end{cases} \tag{3.6}$$

We are ready to state the main result of this section:

Theorem 3.1. *Let $(\bar{x},\bar{u}) \in \mathcal{P}(K_0)$ be a strong local minimizer of problem* (P). *Assume (A1) holds true and $\bar{p} \in W^{1,1}([0,1];\mathbb{R}^n)$ solves* (3.1). *Then for every $\Psi \in \boldsymbol{S}(2n)$ satisfying $(\nabla\varphi(\bar{x}(0),\bar{x}(1)),\Psi) \in J^{2,+}\varphi(\bar{x}(0),\bar{x}(1))$ and $\Psi_i \in \boldsymbol{S}(n)$, $i=1,2$ such that*

$$\Psi = \begin{pmatrix} \Psi_1 & \Psi_0 \\ \Psi_0^T & \Psi_2 \end{pmatrix},$$

the solution $W \in W^{1,1}([0,1];\boldsymbol{S}(n))$ of the matrix differential equation

$$\begin{cases} \dot{W}(t) = -\mathcal{H}_{px}[t]W(t) - W(t)\mathcal{H}_{xp}[t] - \mathcal{H}_{xx}[t], & \text{a.e.} \\ W(1) = -\Psi_2, \end{cases} \tag{3.7}$$

satisfies the second-order transversality condition

$$W(0) - \Psi_1 - 2\Psi_0 Y(1) \in N_{K_0}^{\flat(2)}\big(\bar{x}(0); \bar{p}(0) - \nabla_{x_1}\varphi(\bar{x}(0),\bar{x}(1))\big), \tag{3.8}$$

and the second-order maximality condition

$$\max_{(v,M)\in\overline{D}(t)} \left\langle M^T\bar{p}(t) + W(t)v, v\right\rangle = 0, \quad \text{a.e. in } [0,1]. \tag{3.9}$$

Observe that since $\bar{u}(t) \in \overline{U}(t)$ a.e. we have $(0,0) \in \overline{D}(t)$. Hence the maximum in the above expression is attained on optimal control $\bar{u}$. Therefore the above theorem can be seen as a second-order maximum principle. Its proof is postponed to Section 6 below and is based on a second-order variational equation which is studied in Section 5.

Remark 3.2. Taking the maximum in (3.9) only over $\left\{(\Delta f[t,u], \Delta f_x[t,u]) \;\middle|\; u \in \overline{U}(t)\right\}$, we obtain

$$\max_{u\in\overline{U}(t)} \left\langle (\mathcal{H}_x(t,\bar{x}(t),u) - \mathcal{H}_x[t])^T + W(t)\left(f(t,\bar{x}(t),u) - f[t]\right), f(t,\bar{x}(t),u) - f[t]\right\rangle = 0, \quad \text{a.e.} \tag{3.10}$$

which is a necessary optimality condition similar to the one from [22] derived for a Bolza optimal control problem with a fixed initial state. Moreover, for a fixed $t \in [0,1]$, $k = n + n^2$ and fixed $u_1,\ldots,u_{k+1} \in \overline{U}(t)$ condition (3.9) implies that

$$\max_{\lambda\in\Lambda^k} \left\langle \sum_{i=1}^{k+1} \lambda_i \Delta f_x[t,u_i]^T\bar{p}(t) + W(t)\sum_{i=1}^{k+1}\lambda_i\Delta f[t,u_i], \sum_{i=1}^{k+1}\lambda_i\Delta f[t,u_i]\right\rangle = 0, \tag{3.11}$$

where $\Lambda^k := \left\{(\lambda_1, \dots, \lambda_{k+1}) \;\middle|\; \sum_{i=1}^{k+1} \lambda_i = 1, \lambda_i \geq 0 \quad \forall i\right\}$. Therefore (3.11) is equivalent to testing the copositivity of the matrix $-(M + V^T W(t))V$ where

$$M := \begin{pmatrix} \bar{p}(t)^T \Delta f_x[t, u_1] \\ \vdots \\ \bar{p}(t)^T \Delta f_x[t, u_{k+1}] \end{pmatrix}, \quad V^T := \begin{pmatrix} \Delta f[t, u_1]^T \\ \vdots \\ \Delta f[t, u_{k+1}]^T \end{pmatrix}.$$

We refer to [17] and the bibliography contained therein for a survey of this important problem of optimisation theory.

Corollary 3.3. *In Theorem 3.1 assume in addition that the state and control variables are separated, that is there exist $f_1 : [0,1] \times \mathbb{R}^n \to \mathbb{R}^n$ and $f_2 : [0,1] \times \mathbb{R}^m \to \mathbb{R}^n$ satisfying assumptions (A1) (a), (b) such that*

$$f(t, x, u) = f_1(t, x) + f_2(t, u) \;\; \forall\, x \in \mathbb{R}^n,\, u \in \mathbb{R}^m.$$

Then for a.e. $t \in [0,1]$ the matrix $W(t)$ is seminegative on the closed convex cone spanned by the set $f_2(t, \bar{U}(t)) - f_2(t, \bar{u}(t))$.

We end this section with a few comments on the mapping $W(\cdot)$ which we call the *second-order adjoint matrix*. Let $\mathcal{R}(t, t_0) := Y(t)Y^{-1}(t_0)$ denote the *resolvent*. Then it is not difficult to check that

$$W(t) = -\mathcal{R}(1,t)^T \Psi_2 \mathcal{R}(1,t) + \int_t^1 \mathcal{R}(s,t)^T \mathcal{H}_{xx}[s] \mathcal{R}(s,t) ds, \;\; \forall\, t \in [0,1].$$

The next result follows directly from this formula and the first-order maximality condition.

Corollary 3.4. *If Ψ_2 is positive semidefinite and $\mathcal{H}_{xx}[t]$ is negative semidefinite for almost all $t \in [0,1]$, then $W(t)$ is negative semidefinite for all $t \in [0,1]$. Consequently, $(\bar{p}(t), W(t)) \in \mathcal{N}^{(2)}_{F(t,\bar{x}(t))}(\dot{\bar{x}}(t))$,* a.e.

Example 3.5. Consider a Mayer problem where the dynamics and control constraints are given by

$$f(x,u) = \begin{cases} \begin{pmatrix} u_1 & u_1 - u_2 x_1 \end{pmatrix}^T & \text{if } u_1 \in [0,1] \\ \begin{pmatrix} u_1 & \frac{1}{2}u_1 - u_2 x_1 \end{pmatrix}^T & \text{if } u_1 \in [-1,1], \end{cases} \qquad U = [-1,1] \times [0,1],$$

where subscripts indicate the components of vectors. Denoting the initial and terminal state by x^S and x^E respectively, initial constraints and cost function are defined by

$$K_0 = \{(0,0)\} \qquad \text{and} \qquad \varphi(x^S, x^E) = -(x_1^E)^2 \operatorname{sign}(x_1^E) + \frac{1}{2}(x_2^E)^2 \operatorname{sign}(x_2^E).$$

The Hamiltonian corresponding to this problem is

$$\mathcal{H}(x,p,u) = \begin{cases} p_1 u_1 + p_2(u_1 - u_2 x_1) & \text{if } u_1 \in [0,1] \\ p_1 u_1 + p_2\left(\frac{1}{2}u_1 - u_2 x_1\right) & \text{otherwise.} \end{cases}$$

Let us consider a candidate of optimality $(\bar{x}, \bar{u})$ with $\bar{u}_1 \equiv 0$. Then it is easy to see that $\bar{x}_1 \equiv \bar{x}_2 \equiv 0$ and it follows directly from the adjoint equation (3.1) that $\bar{p}_1 \equiv \bar{p}_2 \equiv 0$. Consequently,

$$\mathcal{H}(\bar{x}(t), \bar{p}(t), u) = 0 \qquad \forall u \in U,\; \forall t \in [0,1].$$

Thus the maximality condition (3.3) is trivially satisfied and $\overline{U}(t) = U$ for all $t \in [0,1]$.

Noticing that $\left(\begin{pmatrix} -2\,|\bar{x}_1(1)| \\ |\bar{x}_2(1)| \end{pmatrix}, \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -2 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}\right) \in J^{2,+}\varphi(x^S, x^E)$, the corresponding second-order adjoint equation is

$$\dot{W}(t) = \begin{pmatrix} 2\bar{u}_2(t)W_{12}(t) & \bar{u}_2(t)W_{22}(t) \\ \bar{u}_2(t)W_{22}(t) & 0 \end{pmatrix}, \quad W(1) = \begin{pmatrix} 2 & 0 \\ 0 & -1 \end{pmatrix}.$$

It follows that $W_{22} \equiv -1$, $W_{12}(t) = W_{21}(t) = \int_t^1 \bar{u}_2(s)ds \geq 0$ for all $t \in [0,1]$. Similarly,

$$0 \leq 2\bar{u}_2(t)W_{12}(t) = \dot{W}_{11}(t) \leq 2, \qquad W_{11}(1) = 2. \tag{3.12}$$

The second-order maximality condition (3.9) requires in particular that for almost all $t \in [0,1]$,

$$0 = \max_{u \in U} \langle W(t)f(\bar{x}(t), u), f(\bar{x}(t), u)\rangle = \begin{cases} W_{11}(t)u_1^2(t) + 2W_{12}(t)u_1^2 - u_1^2 & \text{if } u_1 \in [0,1], \\ W_{11}(t)u_1^2 + W_{12}(t)u_1^2 - \frac{1}{4}u_1^2 & \text{otherwise.} \end{cases}$$

However, from (3.12) it is clear that $W_{11}(t) > 1$ on a set of positive measure. Hence the second-order condition is violated for all $u \in \overline{U}(t) = U$ with $u_1 \neq 0$ and the candidate $(\bar{x}, \bar{u})$ cannot be optimal.

3.2 Case of dynamics differentiable in the control variable

Let $(\bar{x}, \bar{u}) \in \mathcal{P}(K_0)$ be a strong local minimizer of (P) and $\bar{p} \in W^{1,1}([0,1]; \mathbb{R}^n)$ solve (3.1). When f is differentiable also with respect to u we can derive pointwise optimality conditions involving tangent cones to $U(t)$ at $\bar{u}(t)$, where $t \in [0,1]$. We would like to underline that results of this section are also valid for weak local minimizers. Indeed, if $(\bar{x}, \bar{u})$ is a weak local minimizer, we can always replace $U(t)$ by $U_1(t) := U(t) \cap (\bar{u}(t) + \epsilon\overline{B})$ with appropriately chosen $\epsilon > 0$. Then $(\bar{x}, \bar{u})$ becomes a strong local minimizer for the Mayer problem with $U(t)$ replaced by $U_1(t)$.
Here we impose, in addition to (A1), the following regularity assumptions for $\rho > 0$ as in (A1).

(A2) (a) For almost all $t \in [0,1]$, $f(t, \cdot, \cdot)$ is twice differentiable on $(\bar{x}(t) + \rho B) \times (\bar{u}(t) + \rho B)$;

(b) There exists $c_1 > 0$ s.t. for a.e. $t \in [0,1]$, $\forall\, x, y \in \bar{x}(t) + \rho B$, $\forall\, u, v \in \bar{u}(t) + \rho B$,

$$\|f_u[t]\| \leq c_1 \quad \text{and} \quad \left\|f'(t,x,u) - f'(t,y,v)\right\| \leq c_1(|x-y| + |u-v|),$$

where for every t, $f'(t,x,u)$ denotes the derivative of the map $(x,u) \mapsto f(t,x,u)$;

(c) There exists an integrable $c_2 : [0,1] \to \mathbb{R}_+$ such that for a.e. $t \in [0,1]$, $\forall\, x, y \in \bar{x}(t) + \rho B$, $\forall\, u, v \in \bar{u}(t) + \rho B$,

$$\left\|f''(t,x,u) - f''(t,y,v)\right\| \leq c_2(t)(|x-y| + |u-v|),$$

where for every t, $f''(t,x,u)$ denotes the Hessian of the map $(x,u) \mapsto f(t,x,u)$;

(d) φ is twice Fréchet differentiable.

We introduce the set of "non-singular" times :

$$A := \{t \in [0,1] \mid \mathcal{H}_u[t] \neq 0\}, \tag{3.13}$$

and observe that, by (3.3), for a.e. $t \in A$, $\bar{u}(t)$ belongs to the boundary of $U(t)$.

Theorem 3.6. *Let $(\bar{x},\bar{u}) \in \mathcal{P}(K_0)$ be a strong local minimizer of* (P)*, (A1), (A2) hold true and $\bar{p} \in W^{1,1}([0,1];\mathbb{R}^n)$ solve* (3.1)*. Set*

$$\Psi_0 := \frac{\partial^2\varphi}{\partial x_1 x_2}(\bar{x}(0),\bar{x}(1)), \quad \Psi_1 := \frac{\partial^2\varphi}{\partial x_1^2}(\bar{x}(0),\bar{x}(1)), \quad \Psi_2 := \frac{\partial^2\varphi}{\partial x_2^2}(\bar{x}(0),\bar{x}(1)).$$

Then the solution $W \in W^{1,1}([0,1];\boldsymbol{S}(n))$ of (3.7) *satisfies the transversality condition* (3.8)*. Furthermore, for a.e. $t \in [0,1]$ and for every $u \in T^{\flat}_{U(t)}(\bar{u}(t))$ such that either*
(i) $t \in A$, $\mathcal{H}_u[t]u = 0$ and $\mathcal{H}_u[t]v + \frac{1}{2}u^T\mathcal{H}_{uu}[t]u = 0$ for some $v \in T^{\flat(2)}_{U(t)}(\bar{u}(t),u)$,
or
(ii) $t \in [0,1]\backslash A$ and $u^T\mathcal{H}_{uu}[t]u = 0$,

we have

$$\left\langle f_u[t]^T\left(\mathcal{H}_{ux}[t] + W(t)f_u[t]\right)u, u\right\rangle \le 0. \tag{3.14}$$

Observe that, by the maximality condition (3.3), for a.e. $t \in [0,1]$, and all $u \in T^{\flat}_{U(t)}(\bar{u}(t)) \cap \mathcal{H}_u[t]^{\perp}$,

$$\mathcal{H}_u[t]v + \frac{1}{2}u^T\mathcal{H}_{uu}[t]u \le 0, \quad \forall\, v \in T^{\flat(2)}_{U(t)}(\bar{u}(t),u). \tag{3.15}$$

Hence (i) says that, for the given $u \in T^{\flat}_{U(t)}(\bar{u}(t)) \cap \mathcal{H}_u[t]^{\perp}$ the maximum over the left-hand side of the above inequality is attained and is equal to 0.

Remark 3.7. a) If f is affine in u and U is equal to a (time independent) polyhedron in $\mathbb{R}^m$, then $\mathcal{H}_{uu} = 0$ and $0 \in T^{\flat(2)}_{U(s)}(\bar{u}(s),u)$ for any $u \in T^{\flat}_{U(s)}(\bar{u}(s))$ and a.e. $s \in [0,1]$. Therefore inequality (3.14) holds true for any $u \in T^{\flat}_{U(t)}(\bar{u}(t)) \cap \mathcal{H}_u[t]^{\perp}$. With every $t \in [0,1]$ let us associate the largest subspace $P(t)$ contained in $T^{\flat}_{U(t)}(\bar{u}(t))$. Then the above theorem implies that $f_u[t]^T\mathcal{H}_{ux}[t] + f_u[t]^TW(t)f_u[t]$ is seminegative definite on $P(t)$.

b) The proof of Theorem 3.6 provided below actually allows to use a larger subset of tangents. Namely, (i) can be replaced by
$t \in A$, $\mathcal{H}_u[t]u = 0$ and $\mathcal{H}_u[t]v + \frac{1}{2}u^T\mathcal{H}_{uu}[t]u = 0$ for some v satisfying $(u,v) \in \overline{\{(\alpha,\beta) \,|\, \beta \in T^{\flat(2)}_{U(t)}(\bar{u}(t),\alpha)\}}$.
Moreover, using the relaxation theorem, (see for instance [2, Thm. 10.4.4]), the statement could be generalized by replacing u in (3.14) with $\sum_{i=1}^{k+1}\lambda_i u_i$ where $k \in \mathbb{N}$, $\lambda \in \Lambda^k$ with Λ^k defined as in Remark 3.2 and u_i satisfying the same assumptions as u in Theorem 3.6 for all $i \in \{1,\dots,k+1\}$.

c) To the best of our knowledge, the above result is the first extension of the Jacobson inequality from [20], originally stated for $U(\cdot) \equiv [a,b]$ for some $-\infty < a < b < +\infty$, a system affine with respect to controls and $\mathcal{H}_u[\cdot] = 0$, to such a general framework. The Jacobson inequality was generalized by Jacobson and Speyer in [21] to multidimensional $U \equiv [a,b] \times .... \times [a,b]$. These authors have shown that $f_u[t]^T\mathcal{H}_{ux}[t] + f_u[t]^TW(t)f_u[t]$ is seminegative definite assuming that the control system is affine, $\mathcal{H}_u[\cdot] = 0$ and that the optimal control is piecewise continuous and takes values in the interior of U. Let us underline again that the existence theorems in optimal control theory do not guarantee such structural properties of optimal controls.

The above Theorem and Remark b) imply the following corollary.

Corollary 3.8. *In Theorem 3.6 assume in addition that the state and control variables are separated, that is there exist $f_1 : [0,1] \times \mathbb{R}^n \to \mathbb{R}^n$ and $f_2 : [0,1] \times \mathbb{R}^m \to \mathbb{R}^n$ satisfying assumptions (A1) (a), (b) and (A2) (a), (b), (c) such that*

$$f(t,x,u) = f_1(t,x) + f_2(t,u) \;\; \forall\, x \in \mathbb{R}^n,\, u \in \mathbb{R}^m.$$

If $\bar{u}$ is singular in the classical sense, that is $\mathcal{H}_u[t] = 0$ and $\mathcal{H}_{uu}[t] = 0$ a.e., then for a.e. $t \in [0,1]$ the matrix $f_u[t]^T W(t) f_u[t]$ is seminegative on the closed convex cone $\overline{co}\, T^{\flat}_{U(t)}(\bar{u}(t))$.

Example 3.9. Let the set-valued map $U\colon [0,1] \rightsquigarrow \mathbb{R}^m$ be given by the inequality constraints

$$U(t) := \bigcap_{j=1}^{s} \left\{ u \in \mathbb{R}^m \;\middle|\; c^j(t,u) \le 0 \right\},$$

where $c^1, \dots, c^s\colon [0,1] \times \mathbb{R}^m \to \mathbb{R}$ are Carathéodory functions. Assume that for a.e. $t \in [0,1]$ and for every $j \in \{1, \dots, s\}$, $c^j(t,\cdot)$ is twice continuously differentiable and $\{\nabla_u c^j(t,\bar{u}(t)) \,|\, j = 1, ..., s\}$ are positively independent.
If $(\bar{x}, \bar{u})$, $\bar{p}$ are as in Theorem 3.6, then, by (3.4), for a.e. $t \in [0,1]$ and for every $u \in T^{\flat}_{U(t)}(\bar{u}(t))$ we have $\mathcal{H}_u[t]u \le 0$. That is $\nabla_u \mathcal{H}[t] \in N^{\flat}_{U(t)}(\bar{u}(t))$. By Example 2.2, for a.e. $t \in A$ and for every $u \in T^{\flat}_{U(t)}(\bar{u}(t))$ satisfying $\mathcal{H}_u[t]u = 0$, there exist $\alpha_j(t,u) \ge 0$, $j = 1, ..., s$ and $\bar{v} \in T^{\flat(2)}_{U(t)}(\bar{u}(t), u)$ such that

$$\mathcal{H}_u[t] = \sum_{j=1}^{s} \alpha_j(t,u) c^j_u(t,\bar{u}(t)), \quad \mathcal{H}_u[t]\bar{v} = -\frac{1}{2} \sum_{j=1}^{s} \alpha_j(t,u) c^j_{uu}(t,\bar{u}(t)) uu,$$

and

$$\mathcal{H}_u[t]v + \frac{1}{2} \sum_{j=1}^{s} \alpha_j(t,u) c^j_{uu}(t,\bar{u}(t)) uu \le 0, \quad \forall\, v \in T^{\flat(2)}_{U(t)}(\bar{u}(t), u).$$

If $t \notin A$, then set $\alpha_j(t,u) = 0$ for all $u \in T^{\flat}_{U(t)}(\bar{u}(t))$. Thus

$$\sup_{v \in T^{\flat(2)}_{U(t)}(\bar{u}(t),u)} \mathcal{H}_u[t]v = -\frac{1}{2} \sum_{j=1}^{s} \alpha_j(t,u) c^j_{uu}(t,\bar{u}(t)) uu.$$

If in addition $\{\nabla_u c^j(t,\bar{u}(t))\}_{j=1}^{s}$ are linearly independent, then the choice of $\alpha_j(t,u)$ is unique (up to a set of measure zero) and independent from u. Consequently we get the following corollary:

Corollary 3.10. *Under the assumptions of Theorem 3.6, suppose that the set-valued map $U\colon [0,1] \rightsquigarrow \mathbb{R}^m$ is as in Example 3.9 and for almost all $t \in [0,1]$, $\{\nabla_u c^j(t,\bar{u}(t))\}_{j=1}^{s}$ are linearly independent. Then the solution $W \in W^{1,1}([0,1]; \boldsymbol{S}(n))$ of (3.7) satisfies the transversality condition (3.8) and there exist measurable, uniquely defined (up to a set of measure zero) $\alpha_j : [0,1] \to \mathbb{R}_+$, $j = 1, ..., r$ such that for almost all $t \in [0,1]$,*

(i) $\alpha_j(t) c^j(t,\bar{u}(t)) = 0$ *for all* $j \in \{1, \dots, s\}$;

(ii) $\mathcal{H}_u[t] = \sum_{j=1}^{s} \alpha_j(t) c^j_u(t,\bar{u}(t))$;

(iii) $\max_{u \in U_0(t)} \left\langle f_u[t]^T \left(\mathcal{H}_{ux}[t] + W(t) f_u[t] \right) u, u \right\rangle = 0$, *where*

$$U_0(t) := \left\{ u \in T^{\flat}_{U(t)}(\bar{u}(t)) \;\middle|\; \mathcal{H}_u[t]u = 0,\ u^T \mathcal{H}_{uu}[t]u - u^T \left(\sum_{j=1}^{s} \alpha_j(t) c^j_{uu}(t,\bar{u}(t)) \right) u = 0 \right\}. \quad (3.16)$$

Remark 3.11. As mentioned earlier, by the maximum principle, for a.e. $t \in [0,1]$ and for any $u \in T^{\flat}_{U(t)}(\bar{u}(t))$ satisfying $\mathcal{H}_u[t]u = 0$ also (3.15) holds true. The set $U_0(t)$ is therefore the natural choice for the set of maximizing directions.

Proof of Corollary 3.10. As recalled before, whenever $\{\nabla_u c^j(t,\bar{u}(t))\}_{j=1}^s$ are linearly independent, $\alpha_j(t,u)$ from Example 3.9 are unique and independent of u. The existence of measurable $\alpha_j : [0,1] \to \mathbb{R}_+$ such that (i) and (ii) of Corollary 3.10 are satisfied follows. Theorem 3.6 yields (iii). □

Example 3.12. Consider the Mayer problem with the following dynamics and cost function

$$f(x,u) = \begin{pmatrix} u_1 & u_1 - u_2 & -x_1^2 + \frac{1}{2}x_2^2 + 8u_2^2 \end{pmatrix}^T, \quad \varphi(x^S, x^E) = x_3^E,$$

where subscripts indicate the components of vectors and x^S, x^E denote the start and end point. Inital and control constraints are given by

$$K_0 = \{(0,0,0)\}, \quad U = \{(u_1,u_2) \in [0,1] \times [0,1] \mid u_2 \le u_1\}.$$

Notice that standard results guarantee the existance of a global minimizer of this problem.
The corresponding adjoint equation (3.1) for a candidate $(\bar{x},\bar{u}) \in \mathcal{P}(0)$ is

$$\left\{ -\begin{pmatrix} \dot{\bar{p}}_1(t) \\ \dot{\bar{p}}_2(t) \\ \dot{\bar{p}}_3(t) \end{pmatrix} = \begin{pmatrix} 0 & 0 & -2\bar{x}_1(t) \\ 0 & 0 & \bar{x}_2(t) \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \bar{x}_1(t) \\ \bar{x}_2(t) \\ \bar{x}_3(t) \end{pmatrix}, \qquad -\begin{pmatrix} \bar{p}_1(1) \\ \bar{p}_2(1) \\ \bar{p}_3(1) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}. \right.$$

Taking the candidate minimizer $\bar{u} \equiv 0$ we find readily that in this case $\bar{x}_1 \equiv \bar{x}_2 \equiv \bar{p}_1 \equiv \bar{p}_2 \equiv 0$ and $\bar{p}_3 \equiv -1$. It follows then that $\mathcal{H}(\bar{x}(t),\bar{p}(t),u) = -8u_2^2$ which is maximized for $u_2 = \bar{u}_2(t) = 0$ for all $t \in [0,1]$. Therefore $\bar{u}$ satisfies the maximum principle (3.3). Notice also that since $\mathcal{H}_{uu}[t] \neq 0$, Goh type necessary conditions cannot be applied.
Simple computations show that $\mathcal{H}_u[t] \equiv \mathcal{H}_{ux}[t] \equiv 0$. Moreover, noticing that

$$u^T \mathcal{H}_{uu} u = -16u_2^2 = 0 \qquad \forall u \in \left\{ (v_1,v_2) \in T_U^\flat((0,0)) \;\middle|\; v_2 = 0 \right\},$$

the second-order optimality condition from Theorem 3.6 reads

$$\left\langle f_u[t]^T W(t) f_u[t] u, u \right\rangle \le 0 \qquad \forall u \in \left\{ (v_1,v_2) \in T_U^\flat((0,0)) \;\middle|\; v_2 = 0 \right\}.$$

Solving the second-order adjoint equation (3.7) we find that

$$\left\langle f_u[t]^T W(t) f_u[t] u, u \right\rangle = (1-t)u_1^2 > 0 \qquad \forall t \in [0,1[,\ u \in \left\{ (v_1,v_2) \in T_U^\flat((0,0)) \;\middle|\; v_1 > 0, v_2 = 0 \right\},$$

violating the second-order necessary condition. Hence $(\bar{x},\bar{u})$ cannot be optimal.

4 Second-order sensitivity relations

In this section we assume that φ does not depend on the first variable, i.e. $\varphi : \mathbb{R}^n \to \mathbb{R}$. Using the super/subjets of the value function along optimal trajectories, we derive sensitivity relations for the first- and second-order adjoints $\bar{p}$ and W, solving respectively (3.1) and (3.7). Proofs of the results of this section are deferred to Section 7.
The *value function* $V\colon [0,1] \times \mathbb{R}^n \to \mathbb{R}$ associated with the Mayer problem (P) is defined by

$$V(t,x) \coloneqq \inf \{\varphi(z(1)) \mid z \in \mathcal{S}_t(x)\}. \tag{4.1}$$

This function was introduced by Bellman and Isaacs in 50ies and has been studied by many researchers. The interested reader can find in [7, 29] extended discussions on value functions. It is well known that if $\bar{x} \in \mathcal{S}(x)$ is an optimal solution of (P), then the adjoint state $\bar{p}$ solving (3.1) satisfies

$$-\bar{p}(t) \in \partial_x^+ V(t, \bar{x}(t)), \qquad \forall t \in [0,1],$$

where $\partial_x^+ V(t, \bar{x}(t))$ denotes the superdifferential of $V(t, \cdot)$ at $\bar{x}(t)$, see for instance [2, Def. 6.5.7] for the definition of the superdifferential. The main result of this section is an analogous second-order sensitivity relation involving first- and second-order adjoints.

Theorem 4.1. *Let $\bar{x} \in \mathcal{S}_{t_0}(x_0)$ satisfy $V(t_0, x_0) = \varphi(\bar{x}(1))$. Assume (A1) and let $\Psi_2 \in \boldsymbol{S}(n)$ be such that $(\nabla\varphi(\bar{x}(1)), \Psi_2) \in J^{2,+}\varphi(\bar{x}(1))$, $\bar{p} \in W^{1,1}([t_0,1];\mathbb{R}^n)$ and $W \in W^{1,1}([t_0,1];\boldsymbol{S}(n))$ be solutions of* (3.1) *and* (3.7) *respectively. Then*

$$(-\bar{p}(t), -W(t)) \in J_x^{2,+} V(t, \bar{x}(t)), \qquad \forall t \in [t_0, 1].$$

The statement of Theorem 4.1 can be interpreted as follows. If $(-\bar{p}_T, -W_T) \in J_x^{2,+} V(1, \bar{x}(1))$, then this property propagates backward in time along the trajectory $\bar{x}$ for the solutions of the first- and second-order adjoint equations with terminal conditions given by $\bar{p}_T$ and W_T. Actually, if we replace the superjet by the subjet then a related sensitivity inclusion propagates forward in time.

Theorem 4.2. *Let $\bar{x} \in \mathcal{S}_{t_0}(x_0)$ be such that $V(t_0, x_0) = \varphi(\bar{x}(1))$. Assume (A1) and let $\bar{p} \in W^{1,1}([t_0,1];\mathbb{R}^n)$ be the corresponding solution of the adjoint equation* (3.1) *defined on $[t_0, 1]$. If for some $W_0 \in \boldsymbol{S}(n)$ we have $(-\bar{p}(t_0), -W_0) \in J_x^{2,-} V(t_0, x_0)$ then for the solution $W \in W^{1,1}([t_0,1];\boldsymbol{S}(n))$ of*

$$\begin{cases} \dot{W}(t) + \mathcal{H}_{px}[t]W(t) + W(t)\mathcal{H}_{xp}[t] + \mathcal{H}_{xx}[t] = 0, \\ W(t_0) = W_0, \end{cases}$$

the following sensitivity relation holds true :

$$(-\bar{p}(t), -W(t)) \in J_x^{2,-} V(t, \bar{x}(t)), \qquad \forall t \in [t_0, 1].$$

Remark 4.3. Let $\bar{x}$, $\bar{p}$, and W be as in Theorem 4.1. By Theorem 4.2, if there exists $t_0 \in [0,1]$ such that $(-\bar{p}(t_0), -W(t_0)) \in J_x^{2,-} V(t_0, \bar{x}(t_0))$, then $(-\bar{p}(t), -W(t)) \in J_x^{2,-} V(t, \bar{x}(t))$ for all $t \geq t_0$. In particular, from [6, Remark 2.8] it follows that in this case the value function is differentiable with respect to x along the optimal trajectory and $(-\bar{p}(t), -W(t))$ is the jet of $V(t, \cdot)$ at $\bar{x}(t)$ for all $t \in [t_0, 1]$.

5 Variational equations

Second-order variational equations have been recently studied in [12, 18, 19] for control systems, resp. differential inclusions with state constraints. In [18, 19] differentiability of f with respect to the control is required. In [12] a second-order variational inclusion is given in a very general context by using second-order jets to the velocity set of the differential inclusion. In this section we are going to use this abstract result to provide second-order variational equations which are well adapted for the derivation of our second-order necessary optimality conditions.
We recall first, for the convenience of the reader, [12, Thm. 3.3] in the case when state constraints are absent. The hypotheses of Theorem 3.1 are imposed throughout this section.

5.1 Second-order variational inclusions

Consider a reference trajectory $\bar{x} \in \mathcal{S}(x_0)$ where $x_0 \in K_0$. The set of *admissible first-order variations* $\mathcal{V}^{(1)}(\bar{x})$ at $\bar{x}$ is the set of all absolutely continuous maps $y \in W^{1,1}([0,1];\mathbb{R}^n)$ satisfying,

(i) $\dot{y}(t) \in d_x F(t, \bar{x}(t), \dot{\bar{x}}(t))(y(t)), \quad$ for a.e. $t \in [0,1]$;

(ii) $y(0) \in T^{\flat}_{K_0}(x_0)$;

(iii) there exists an integrable function $\ell_1 : [0,1] \to \mathbb{R}_+$ and $h_0 > 0$ such that for a.e. $t \in [0,1]$,

$$\operatorname{dist}_{F(t,\bar{x}(t)+hy(t))}(\dot{\bar{x}}(t) + h\dot{y}(t)) \le \ell_1(t)h^2, \quad \forall\, h \in [0, h_0].$$

For a given admissible first-order variation $y \in \mathcal{V}^{(1)}(\bar{x})$, we abbreviate $(t, \bar{x}(t), \dot{\bar{x}}(t), y(t), \dot{y}(t))$ by $[t]$ (for $t \in [0,1]$ such that the derivatives $\dot{\bar{x}}(t)$, $\dot{y}(t)$ do exist) and define the set of *admissible second-order variations* $\mathcal{V}^{(2)}(\bar{x}, y)$ as the set of absolutely continuous maps $w \in W^{1,1}([0,1];\mathbb{R}^n)$ satisfying,

(i) $\dot{w}(t) \in d_x^2 F[t](w(t)), \quad$ for a.e. $t \in [0,1]$;

(ii) $w(0) \in T^{\flat(2)}_{K_0}(x_0, y(0))$.

Theorem 5.1 ([12])**.** *Assume (A1) (a). Let $\bar{x} \in \mathcal{S}(x_0)$ for some $x_0 \in K_0$, $y \in \mathcal{V}^{(1)}(\bar{x})$ and $w \in \mathcal{V}^{(2)}(\bar{x}, y)$. Consider any sequences $h_i \to 0+$, $w_i^0 \to w(0)$ such that $x_0 + h_i y(0) + h_i^2 w_i^0 \in K_0$. Then there exist $x_i \in \mathcal{S}(x_0 + h_i y(0) + h_i^2 w_i^0)$, such that $\frac{1}{h_i^2}(x_i - \bar{x} - h_i y)$ converge to w uniformly on $[0,1]$ when $i \to \infty$.*

We work below with some subsets of the admissible first- and second-order variations. For the first-order variations we use the following classical linearization of control system (C):

$$\begin{cases} \dot{y}(t) = f_x[t]y(t) + v(t), \quad v(t) \in T_{F(t,\bar{x}(t))}(\dot{\bar{x}}(t)), \qquad \text{a.e. in } [0,1], \\ y(0) \in T^{\flat}_{K_0}(\bar{x}(0)). \end{cases} \tag{5.1}$$

To simplify the notations define,

$$\mathcal{V}_1 := \left\{ v \in L^1([0,1];\mathbb{R}^m) \;\middle|\; v(t) \in T_{F(t,\bar{x}(t))}(\dot{\bar{x}}(t)) \text{ a.e.} \right\},$$

and the solution map $\operatorname{Lin}\colon T^{\flat}_{K_0}(\bar{x}(0)) \times \mathcal{V}_1 \to W^{1,1}([0,1];\mathbb{R}^n)$, associating with any $y_0 \in T^{\flat}_{K_0}(\bar{x}(0))$ and $v \in \mathcal{V}_1$ the unique solution of (5.1) satisfying $y(0) = y_0$. By [10] (see also [10, Section 5]) for a.e. $t \in [0,1]$ and all $\xi \in \mathbb{R}^n$,

$$f_x[t]\xi + T_{F(t,\bar{x}(t))}(\dot{\bar{x}}(t)) \subset d_x F(t, \bar{x}(t), \dot{\bar{x}}(t))(\xi). \tag{5.2}$$

Our next aim is to provide a second-order approximation similar to (5.1) which is convenient for the derivation of second-order necessary optimality conditions.

5.2 Second-order variational equation for control systems

For an arbitrary control $u \in \mathcal{U}$ we define similarly to Section 3 the map $\delta f[u]\colon [0,1] \to \mathbb{R}^n$ by

$$\delta f[u](t) = \Delta f[t, u(t)] \in T_{F(t,\bar{x}(t))}(\dot{\bar{x}}(t)).$$

The map $\delta f_x[u]\colon [0,1] \to \mathbb{R}^{n\times n}$ is defined analogously. With any $x_0 \in K_0$, $(\bar{x}, \bar{u}) \in \mathcal{P}(x_0)$, $u \in \mathcal{U}$, $y_0 \in T^{\flat}_{K_0}(\bar{x}(0))$ and $y = \operatorname{Lin}(y_0, \delta f[u])$ we associate the second-order approximation,

$$\begin{cases} \dot{w}(t) = f_x[t]w(t) + \Delta f_x[t, u(t)]y(t) + \frac{1}{2} f_{xx}[t]y(t)y(t), \qquad \text{a.e. in } [0,1], \\ w(0) \in T^{\flat(2)}_{K_0}(x_0, y_0). \end{cases} \tag{5.3}$$

The set of solutions to this linear equation is denoted by :

$$\mathrm{Lin}^2(\delta f_x[u], y) := \left\{ w \in W^{1,1}([0,1];\mathbb{R}^n) \;\middle|\; w \text{ is solution of (5.3)} \right\}.$$

As we show below, for every $u \in \mathcal{U}$, $y = \mathrm{Lin}(y_0, \delta f[u])$ and $w \in \mathrm{Lin}^2(\delta f_x[u], y)$ we have $y \in \mathcal{V}^{(1)}(\bar{x})$ and $w \in \mathcal{V}^{(2)}(\bar{x}, y)$ which leads to the following Corollary of Theorem 5.1:

Corollary 5.2. *Assume (A1)(a)-(c). Let $y_0 \in T^{\flat}_{K_0}(\bar{x}(0))$, $u \in \mathcal{U}$, $y = \mathrm{Lin}(y_0, \delta f[u])$ and $w \in \mathrm{Lin}^2(\delta f_x[u], y)$. Then for any sequences $h_i \to 0+$, $w_i^0 \to w(0)$ such that $\bar{x}(0) + h_i y(0) + h_i^2 w_i^0 \in K_0$ there exist $x_i \in \mathcal{S}(\bar{x}(0) + h_i y(0) + h_i^2 w_i^0)$ satisfying,*

$$\lim_{i\to\infty} \frac{1}{h_i^2} \left\| \bar{x} + h_i y + h_i^2 w - x_i \right\|_\infty = 0.$$

Remark 5.3. It is possible to show that in the above statement $\delta f[u]$ and $\delta f_x[u]$ could be replaced by any v and M such that $(v, M) \in \left\{ (w, N) \colon [0,1] \to \mathbb{R}^n \times \mathbb{R}^{n\times n} \text{ measurable} \;\middle|\; (w(t), N(t)) \in \overline{D}(t) \text{ a.e.} \right\}$. However, for the proof of Theorem 3.1 the statement of Corollary 5.2 is sufficient.

Proof of Corollary 5.2. By Theorem 5.1 we have to show that $y \in \mathcal{V}^{(1)}(\bar{x})$ and $w \in \mathcal{V}^{(2)}(\bar{x}, y)$. Define $R = \|\bar{x}\|_\infty + 1$.
Step 1: We start by showing that $y \in \mathcal{V}^{(1)}(\bar{x})$. By (5.2) it suffices to prove that y satisfies property (iii) of the definition of admissible first-order variations. We know that for a.e. t and for all $h > 0$,

$$\dot{\bar{x}}(t) + h\dot{y}(t) = f[t] + f_x[t]hy(t) + h\Delta f[t, u(t)].$$

On the other hand, since

$$f(t, \bar{x}(t) + hy(t), \bar{u}(t)) - f[t] = \int_0^h f_x(t, \bar{x}(t) + sy(t), \bar{u}(t))y(t)ds,$$

by the Lipschitz continuity of f_x, we get for almost every $t \in [0,1]$ and for all $h > 0$ sufficiently small,

$$\begin{aligned} |f(t, \bar{x}(t) + hy(t), \bar{u}(t)) - f[t] - f_x[t]hy(t)| &\le \int_0^h |f_x(t, \bar{x}(t) + sy(t), \bar{u}(t)) - f_x[t]|\, |y(t)|\, ds \\ &\le \frac{1}{2} l(t) \|y\|_\infty^2 h^2. \end{aligned} \tag{5.4}$$

Similarly, taking $h > 0$ small enough, we get

$$|f(t, \bar{x}(t) + hy(t), u(t)) - f(t, \bar{x}(t) + hy(t), \bar{u}(t)) - \Delta f[t, u(t)]| \le 2k_R(t)h \|y\|_\infty. \tag{5.5}$$

Finally, by the convexity of $F(t, \bar{x}(t) + hy(t))$, for all small $h > 0$,

$$F(t, \bar{x}(t) + hy(t)) \ni (1-h) f(t, \bar{x}(t) + hy(t), \bar{u}(t)) + hf(t, \bar{x}(t) + hy(t), u(t)),$$

which together with (5.4)-(5.5) implies (iii) for $\ell_1(\cdot) := \|y\|_\infty^2 l(\cdot) + 2\|y\|_\infty k_R(\cdot)$.
Step 2: It remains to show that for a.e. $t \in [0,1]$,

$$f_x[t]w + \Delta f_x[t, u(t)]y(t) + \frac{1}{2} f_{xx}[t]y(t)y(t) \subset d_x^2 F[t](w), \qquad \forall\, w \in \mathbb{R}^n. \tag{5.6}$$

Fix t such that the last inequality in (A1)(a) holds true, $w \in \mathbb{R}^n$ and let $h > 0$ be small enough. By the convexity of $F(t, x)$,

$$F(t, \bar{x}(t) + hy(t) + h^2 w) \ni (1-h) f(t, \bar{x}(t) + hy(t) + h^2 w, \bar{u}(t)) + hf(t, \bar{x}(t) + hy(t) + h^2 w, u(t)). \tag{5.7}$$

On the other hand, by Taylor's expansion,

$$\begin{aligned}(1-h)f(t,\bar{x}(t)+hy(t)+h^2w,\bar{u}(t)) &= f[t]+hf_x[t]y(t)+h^2f_x[t]w+\frac{h^2}{2}f_{xx}[t]y(t)y(t)+o(h^2)\\ &\quad -h\left[f(t,\bar{x}(t)+hy(t),\bar{u}(t))+\big(f(t,\bar{x}(t)+hy(t)+h^2w,\bar{u}(t))-f(t,\bar{x}(t)+hy(t),\bar{u}(t))\big)\right],\end{aligned}$$

which, by the Lipschitz continuity of f with respect to x, similarly to (5.5), leads to

$$\begin{aligned}&(1-h)f(t,\bar{x}(t)+hy(t)+h^2w,\bar{u}(t))\\ &\quad= f[t]+hf_x[t]y(t)+h^2f_x[t]w+\frac{h^2}{2}f_{xx}[t]y(t)y(t)-hf(t,\bar{x}(t)+hy(t),\bar{u}(t))+o(h^2)\\ &\quad= f[t]+hf_x[t]y(t)+h^2f_x[t]w+\frac{h^2}{2}f_{xx}[t]y(t)y(t)-hf[t]-h^2f_x[t]y(t)+o(h^2).\end{aligned} \tag{5.8}$$

Similarly,

$$\begin{aligned}f(t,\bar{x}(t)+hy(t)+h^2w,u(t)) &= f(t,\bar{x}(t)+hy(t),u(t))+o(h)\\ &= (f(t,\bar{x}(t),u(t))+hf_x(t,\bar{x}(t),u(t))y(t))+o(h).\end{aligned} \tag{5.9}$$

Consequently, (5.7)-(5.9) together imply that for a.e. $t\in[0,1]$

$$\begin{aligned}&F(t,\bar{x}(t)+hy(t)+h^2w)\\ &\quad\ni f[t]+h\left(f_x[t]y(t)+\Delta f[t,u(t)]\right)+h^2\left(f_x[t]w+\frac{1}{2}f_{xx}[t]y(t)y(t)+\Delta f_x[t,u(t)]y(t)\right)+o(h^2)\\ &\quad= \dot{\bar{x}}(t)+h\dot{y}(t)+h^2\left(f_x[t]w+\Delta f_x[t,u(t)]y(t)+\frac{1}{2}f_{xx}[t]y(t)y(t)\right)+o(h^2),\end{aligned}$$

which proves that (5.6) is satisfied a.e. in $[0,1]$. □

6 Proofs of the second-order optimality conditions

In this section we provide the proofs of the second-order necessary optimality conditions from Section 3. Most of them are based on variational equations and a second-order adjoint equation.

6.1 Proof of Theorem 3.1

To simplify notations we introduce the set

$$\overline{\mathcal{U}} := \left\{u\in\mathcal{U} \mid u(t)\in\overline{U}(t) \text{ for a.e. } t\in[0,1]\right\}.$$

Step 1: Let $y_0\in\Gamma_0$, $u\in\overline{\mathcal{U}}$ and $y=\mathrm{Lin}(y_0,\delta f[u])$. We claim that,

$$\langle\nabla\varphi(\bar{x}(0),\bar{x}(1)),(y(0),y(1))\rangle = 0. \tag{6.1}$$

Indeed, using integration by parts and the properties of the adjoint state we find that,

$$\begin{aligned}\langle\nabla\varphi(\bar{x}(0),\bar{x}(1)),(y(0),y(1))\rangle &= -\int_0^1\langle\dot{\bar{p}}(t),y(t)\rangle+\langle\bar{p}(t),\dot{y}(t)\rangle\,dt+\langle\nabla_{x_1}\varphi(\bar{x}(0),\bar{x}(1))-\bar{p}(0),y_0\rangle\\ &= -\int_0^1\langle\bar{p}(t),\Delta f[t,u(t)]\rangle\,dt+\langle\nabla_{x_1}\varphi(\bar{x}(0),\bar{x}(1))-\bar{p}(0),y_0\rangle.\end{aligned}$$

Since $(y_0,u)\in\Gamma_0\times\overline{\mathcal{U}}$, the right-hand side of the above equality is zero and the conclusion follows.

Step 2: Fix $\Psi \in \mathbf{S}(2n)$ as in the statement of our theorem and let $u \in \overline{\mathcal{U}}$. We show next that the unique matrix solution $W \in W^{1,1}([0,1];\mathbb{R}^{n\times n})$ of the second-order adjoint equation (3.7) satisfies

$$\int_0^1 \left\langle Y(t)^T \left(W(t)\Delta f[t,u(t)] + \Delta f_x[t,u(t)]^T \bar{p}(t)\right), \int_0^t Y^{-1}(s)\Delta f[s,u(s)]ds \right\rangle dt \leq 0. \tag{6.2}$$

Let $y_0 \in \Gamma_0$, $y = \mathrm{Lin}(y_0, \delta f[u])$, $w \in \mathrm{Lin}^2(\delta f_x[u], y)$ be arbitrary and let the sequences $h_i \to 0+$, $w_i^0 \to w(0)$ and $x_i \in \mathcal{S}(\bar{x}(0) + h_i y(0) + h_i^2 w_i^0)$ be as in Corollary 5.2. Then for all i large enough

$$0 \leq \varphi(x_i(0), x_i(1)) - \varphi(\bar{x}(0), \bar{x}(1)).$$

From the uniform convergence of x_i to $\bar{x}$ and the definition of the superjet we deduce that

$$0 \leq \varphi'(\bar{x}(0),\bar{x}(1)) \begin{pmatrix} x_i(0)-\bar{x}(0) \\ x_i(1)-\bar{x}(1) \end{pmatrix} + \frac{1}{2} \begin{pmatrix} x_i(0)-\bar{x}(0) \\ x_i(1)-\bar{x}(1) \end{pmatrix}^T \begin{pmatrix} \Psi_1 & \Psi_0 \\ \Psi_0^T & \Psi_2 \end{pmatrix} \begin{pmatrix} x_i(0)-\bar{x}(0) \\ x_i(1)-\bar{x}(1) \end{pmatrix} + o(\|x_i - \bar{x}\|_\infty^2),$$

which, by Corollary 5.2, implies that

$$0 \leq \varphi'(\bar{x}(0),\bar{x}(1)) \begin{pmatrix} h_i y(0) + h_i^2 w(0) \\ h_i y(1) + h_i^2 w(1) \end{pmatrix} + \frac{h_i^2}{2} \begin{pmatrix} y(0) \\ y(1) \end{pmatrix}^T \begin{pmatrix} \Psi_1 & \Psi_0 \\ \Psi_0^T & \Psi_2 \end{pmatrix} \begin{pmatrix} y(0) \\ y(1) \end{pmatrix} + o(h_i^2).$$

Thanks to (6.1) we can divide both sides of the last inequality by h_i^2 and pass to the limit as $i \to +\infty$ which leads to

$$0 \leq \varphi'(\bar{x}(0),\bar{x}(1)) \begin{pmatrix} w(0) \\ w(1) \end{pmatrix} + \frac{1}{2} \begin{pmatrix} y(0) \\ y(1) \end{pmatrix}^T \begin{pmatrix} \Psi_1 & \Psi_0 \\ \Psi_0^T & \Psi_2 \end{pmatrix} \begin{pmatrix} y(0) \\ y(1) \end{pmatrix}. \tag{6.3}$$

Next, consider the (unique) solution W of the matrix differential equation (3.7). One readily checks that if W solves (3.7), then so does W^T. Hence $W(t)$ is a symmetric matrix for all $t \in [0,1]$. Using this, the properties of the adjoint state and assumption (A1), we can rewrite inequality (6.3) as follows:

$$\begin{aligned} &\langle \bar{p}(0), w(0)\rangle - \langle \bar{p}(1), w(1)\rangle + \langle \nabla_{x_1}\varphi(\bar{x}(0),\bar{x}(1)) - \bar{p}(0), w(0)\rangle + \langle y(0), \Psi_0 y(1)\rangle \\ &\qquad + \frac{1}{2}\left(\langle \Psi_1 y(0), y(0)\rangle - \langle W(1)y(1), y(1)\rangle\right) \geq 0. \end{aligned} \tag{6.4}$$

Similarly to step 1 we note that,

$$\begin{aligned} \langle \bar{p}(1), w(1)\rangle - \langle \bar{p}(0), w(0)\rangle &= \int_0^1 \left(\langle \dot{\bar{p}}(t), w(t)\rangle + \langle \bar{p}(t), \dot{w}(t)\rangle\right) dt \\ &= \int_0^1 \left\langle \Delta f_x[t,u(t)]^T \bar{p}(t) + \frac{1}{2}\bar{p}(t)^T f_{xx}[t] y(t), y(t) \right\rangle dt. \end{aligned} \tag{6.5}$$

By the definition of W,

$$\begin{aligned} &-\langle W(1)y(1), y(1)\rangle \\ &\quad = -\langle W(1)y(1), y(1)\rangle + \int_0^1 \left\langle \left(\dot{W}(t) + f_x[t]^T W(t) + W(t) f_x[t] + \bar{p}(t)^T f_{xx}[t]\right) y(t), y(t) \right\rangle dt. \end{aligned} \tag{6.6}$$

Using the integration by parts and that $W(t)$ is symmetric for any t, we find

$$\begin{aligned} &-\langle W(1)y(1), y(1)\rangle + \int_0^1 \left\langle \dot{W}(t)y(t), y(t)\right\rangle dt \\ &\quad = -\langle W(1)y(1), y(1)\rangle + \int_0^1 \left(\frac{d}{dt}\langle Wy, y\rangle(t) - \langle W(t)\dot{y}(t), y(t)\rangle - \langle W(t)y(t), \dot{y}(t)\rangle\right) dt \\ &\quad = -\langle W(0)y(0), y(0)\rangle - \int_0^1 \left(\langle W(t)y(t), \dot{y}(t)\rangle + \langle W(t)\dot{y}(t), y(t)\rangle\right) dt. \end{aligned} \tag{6.7}$$

By the very definition of y we know that

$$\begin{aligned}
&\int_0^1 \left\langle \left(f_x[t]^T W(t) + W(t) f_x[t]\right) y(t), y(t)\right\rangle dt - \int_0^1 \left(\langle W(t)y(t), \dot{y}(t)\rangle + \langle W(t)\dot{y}(t), y(t)\rangle\right) dt \\
&\quad = -\int_0^1 \left(\langle W(t)y(t), \Delta f[t,u(t)]\rangle + \langle W(t)\Delta f[t,u(t)], y(t)\rangle\right) dt \\
&\quad = -2\int_0^1 \langle W(t)\Delta f[t,u(t)], y(t)\rangle\, dt.
\end{aligned} \tag{6.8}$$

Combining equations (6.6)-(6.8) we obtain,

$$-\langle W(1)y(1), y(1)\rangle = -\langle W(0)y_0, y_0\rangle - \int_0^1 2\,\langle W(t)\Delta f[t,u(t)], y(t)\rangle\, dt + \int_0^1 \left\langle \bar{p}(t)^T f_{xx}[t]y(t), y(t)\right\rangle dt.$$

This and (6.4)-(6.5) lead to the inequality

$$\begin{aligned}
&-\langle \bar{p}(0) - \nabla_{x_1}\varphi(\bar{x}(0), \bar{x}(1)), w(0)\rangle - \frac{1}{2}\langle (W(0) - \Psi_1)\, y_0, y_0\rangle + \langle \Psi_0 y(1), y_0\rangle \\
&\qquad - \int_0^1 \left\langle W(t)\Delta f[t,u(t)] + \Delta f_x[t,u(t)]^T \bar{p}(t), y(t)\right\rangle dt \geq 0.
\end{aligned} \tag{6.9}$$

Let $u = \bar{u}$. Then $\delta f[u] \equiv 0$, $\delta f_x[u] \equiv 0$ and $y(1) = Y(1)y_0$. Hence it follows from (6.9) that for all $y_0 \in \Gamma_0$ and all $w_0 \in T^{\flat(2)}_{K_0}(\bar{x}(0), y_0)$

$$\langle \bar{p}(0) - \nabla_{x_1}\varphi(\bar{x}(0), \bar{x}(1)), w_0\rangle + \frac{1}{2}\langle (W(0) - \Psi_1 - 2\Psi_0 Y(1))\, y_0, y_0\rangle \leq 0, \tag{6.10}$$

which allows to conclude that $W(0) - \Psi_1 - 2\Psi_0 Y(1) \in N^{\flat(2)}_{K_0}\left(\bar{x}(0); \bar{p}(0) - \nabla_{x_1}\varphi(\bar{x}(0), \bar{x}(1))\right)$.
Let $y_0 = 0 \in \Gamma_0$, $w(0) = 0 \in T^{\flat(2)}_{K_0}(\bar{x}(0), 0)$. Since for any $u \in \overline{\mathcal{U}}$, $y = \mathrm{Lin}(0, \delta f[u])$ is given by

$$y(t) = \int_0^t Y(t)Y^{-1}(s)\Delta f[s,u(s)]ds,$$

it follows from (6.9) that inequality (6.2) is verified for all $u \in \overline{\mathcal{U}}$.

Step 3: We deduce from (6.2) that for a.e. $t \in [0,1]$,

$$\left\langle W(t)\Delta f[t,u] + \Delta f_x[t,u]^T \bar{p}(t), \Delta f[t,u]\right\rangle \leq 0, \qquad \forall\, u \in \overline{U}(t). \tag{6.11}$$

This part of the proof uses the standard ideas of needle variations of controls, as in [22]. We provide the details for the reader convenience. To simplify the notations we introduce the following maps:

$$\mathcal{G}(t,u) := Y(t)^T \left(W(t)\Delta f[t,u] + \Delta f_x[t,u]^T \bar{p}(t)\right) \qquad \text{and} \qquad G(t,u) := Y^{-1}(t)\Delta f[t,u].$$

Let $u \in \overline{\mathcal{U}}$. By our assumptions $\text{ess-sup}_{t\in[0,1]}\left\{|\mathcal{G}(t,u(t))| + |G(t,u(t))|\right\} =: C < +\infty$. Note that the set

$$E := \left\{ t \in [0,1[\;\middle|\; \lim_{\alpha\to+0} \frac{1}{\alpha}\int_t^{t+\alpha} \begin{pmatrix} \mathcal{G}(s,u(s)) \\ |\mathcal{G}(s,u(s))|^2 \\ G(s,u(s)) \end{pmatrix} ds = \begin{pmatrix} \mathcal{G}(t,u(t)) \\ |\mathcal{G}(t,u(t))|^2 \\ G(t,u(t)) \end{pmatrix} \right\},$$

is of full measure in $[0,1]$. We then fix an arbitrary $\beta \in E$ such that $|G(\beta, u(\beta))| \leq C$, $\alpha \in]0, 1-\beta[$ and define,

$$u^\alpha(t) := \begin{cases} u(t) & \text{if } t \in [\beta, \beta+\alpha], \\ \bar{u}(t) & \text{otherwise.} \end{cases}$$

It is clear that $u^\alpha \in \overline{\mathcal{U}}$. It follows from (6.2) that

$$0 \geq \int_0^1 \left(\int_0^t \langle \mathcal{G}(t, u^\alpha(t)), G(s, u^\alpha(s)) \rangle \, ds \right) dt = \int_\beta^{\beta+\alpha} \left(\int_\beta^t \langle \mathcal{G}(t, u(t)), G(s, u(s)) \rangle \, ds \right) dt. \tag{6.12}$$

Observe that

$$\begin{aligned}
&\left| \frac{2}{\alpha^2} \int_\beta^{\beta+\alpha} \left(\int_\beta^t \langle \mathcal{G}(t, u(t)), G(s, u(s)) \rangle \, ds \right) dt - \langle \mathcal{G}(\beta, u(\beta)), G(\beta, u(\beta)) \rangle \right| \\
&= \left| \frac{2}{\alpha^2} \left[\int_\beta^{\beta+\alpha} \left(\int_\beta^t \left\langle (t-\beta)\mathcal{G}(t, u(t)), \frac{1}{t-\beta} G(s, u(s)) \right\rangle ds - \langle (t-\beta)\mathcal{G}(t, u(t)), G(\beta, u(\beta)) \rangle \right) dt \right.\right. \\
&\qquad \left.\left. + \left\langle \int_\beta^{\beta+\alpha} (t-\beta)\mathcal{G}(t, u(t)) dt, G(\beta, u(\beta)) \right\rangle - \left\langle \int_\beta^{\beta+\alpha} (t-\beta)\mathcal{G}(\beta, u(\beta)) dt, G(\beta, u(\beta)) \right\rangle \right] \right| \\
&\leq \left| \frac{2}{\alpha^2} \int_\beta^{\beta+\alpha} \left\langle (t-\beta)\mathcal{G}(t, u(t)), \frac{1}{t-\beta} \int_\beta^t G(s, u(s)) ds - G(\beta, u(\beta)) \right\rangle dt \right| \\
&\qquad + \left| \frac{2}{\alpha^2} \left\langle \int_\beta^{\beta+\alpha} (t-\beta)\big(\mathcal{G}(t, u(t)) - \mathcal{G}(\beta, u(\beta))\big) dt, G(\beta, u(\beta)) \right\rangle \right| \\
&\leq \frac{2C}{\alpha^2} \int_\beta^{\beta+\alpha} \left((t-\beta) \left| \frac{1}{t-\beta} \int_\beta^t G(s, u(s)) ds - G(\beta, u(\beta)) \right| \right) dt \\
&\qquad + \frac{2C}{\alpha^2} \int_\beta^{\beta+\alpha} \left| (t-\beta)\big(\mathcal{G}(t, u(t)) - \mathcal{G}(\beta, u(\beta))\big) \right| dt.
\end{aligned}$$

From Hölder's inequality we obtain

$$\begin{aligned}
&\left| \frac{2}{\alpha^2} \int_\beta^{\beta+\alpha} \left(\int_\beta^t \langle \mathcal{G}(t, u(t)), G(s, u(s)) \rangle \, ds \right) dt - \langle \mathcal{G}(\beta, u(\beta)), G(\beta, u(\beta)) \rangle \right| \\
&\quad \leq \frac{2C}{\alpha^2} \sup_{r \in]\beta, \beta+\alpha]} \left\{ \left| \frac{1}{r-\beta} \int_\beta^r G(s, u(s)) ds - G(\beta, u(\beta)) \right| \right\} \int_\beta^{\beta+\alpha} (t-\beta) dt \\
&\qquad + \frac{2C}{\alpha^2} \left(\int_\beta^{\beta+\alpha} (t-\beta)^2 dt \right)^{\frac{1}{2}} \left(\int_\beta^{\beta+\alpha} |\mathcal{G}(t, u(t)) - \mathcal{G}(\beta, u(\beta))|^2 \, dt \right)^{\frac{1}{2}} \\
&\quad = C \sup_{r \in]\beta, \beta+\alpha]} \left\{ \left| \frac{1}{r-\beta} \int_\beta^r G(s, u(s)) ds - G(\beta, u(\beta)) \right| \right\} + 2C \left(\frac{1}{\alpha^3} \int_\beta^{\beta+\alpha} (t-\beta)^2 dt \right)^{\frac{1}{2}} \cdot \\
&\qquad \cdot \left(\frac{1}{\alpha} \int_\beta^{\beta+\alpha} \left(|\mathcal{G}(t, u(t))|^2 + |\mathcal{G}(\beta, u(\beta)|^2 - 2 \langle \mathcal{G}(t, u(t)), \mathcal{G}(\beta, u(\beta)) \rangle \right) dt \right)^{\frac{1}{2}} \\
&\quad = C \sup_{r \in]\beta, \beta+\alpha]} \left\{ \left| \frac{1}{r-\beta} \int_\beta^r G(s, u(s)) ds - G(\beta, u(\beta)) \right| \right\} \\
&\qquad + \frac{2C}{\sqrt{3}} \left(\frac{1}{\alpha} \int_\beta^{\beta+\alpha} \left(|\mathcal{G}(t, u(t))|^2 + |\mathcal{G}(\beta, u(\beta))|^2 - 2 \langle \mathcal{G}(t, u(t)), \mathcal{G}(\beta, u(\beta)) \rangle \right) dt \right)^{\frac{1}{2}}.
\end{aligned}$$

Consequently, from the choice of β we conclude that

$$\lim_{\alpha\to+0}\frac{2}{\alpha^2}\int_\beta^{\beta+\alpha}\left(\int_\beta^t\langle\mathcal{G}(t,u(t)),G(s,u(s))\rangle\,ds\right)dt=\langle\mathcal{G}(\beta,u(\beta)),G(\beta,u(\beta))\rangle\,. \tag{6.13}$$

It follows therefore from (6.12) and (6.13) that for all $u\in\overline{\mathcal{U}}$ we have

$$\langle\mathcal{G}(t,u(t)),G(t,u(t))\rangle\le 0,\qquad \text{for a.e. } t\in[0,1]. \tag{6.14}$$

To complete the proof we show that (6.14) implies (6.11) using a contradiction argument. If (6.11) is not satisfied then there exist $\varepsilon>0$ and a set of positive measure $E_1\subset[0,1]$ such that for all $t\in E_1$,

$$\exists\, u\in\overline{U}(t)\quad\text{satisfying}\quad\langle\mathcal{G}(t,u),G(t,u)\rangle\ge\varepsilon. \tag{6.15}$$

Define the set-valued map $\Omega:[0,1]\rightsquigarrow\mathbb{R}^m$ by $\Omega(t)=\{\bar u(t)\}$ for every $t\in[0,1]\backslash E_1$ and

$$\Omega(t):=\left\{u\in\overline{U}(t)\;\middle|\;\langle\mathcal{G}(t,u),G(t,u)\rangle\ge\varepsilon\right\},$$

for $t\in E_1$. It follows from [2, Thm. 8.2.9] that $\Omega(\cdot)$ is measurable. Furthermore $\Omega(t)$ is nonempty and closed a.e. Hence there exists a measurable selection $\tilde u(t)\in\Omega(t)$ for all $t\in[0,1]$. It is clear that $\tilde u\in\overline{\mathcal{U}}$ but (6.14) is violated. This yields the desired contradiction and completes Step 3.

Step 4: It follows directly from (6.11) that for a.e. $t\in[0,1]$

$$\max_{(v,M)\in\left\{(\Delta f[t,u],\Delta f_x[t,u])\;\middle|\;u\in\overline{U}(t)\right\}}\left\langle M^T\bar p(t)+W(t)v,v\right\rangle=0.$$

Set $k=n+n^2$ and consider the optimisation problem

$$\text{Minimize }\ \varphi(x(0),x(1)),$$

over trajectories $x(\cdot)$ of the control system

$$\begin{cases}\dot x(t)=\Sigma_{i=1}^{k+1}\lambda_i(t)f(t,x(t),u_i(t)), & \forall\, i,\ u_i\in\mathcal{U},\quad(\lambda_1(t),...,\lambda_{k+1}(t))\in\Lambda^k\ \text{ a.e. in }[0,1],\\ x(0)\in K_0.\end{cases}$$

In the above we consider only measurable functions λ_i.
By the relaxation Theorem, see for instance [2, Thm. 10.4.4], the process

$$(\bar x,\underbrace{\bar u,....,\bar u}_{k+1},1,\underbrace{0,...,0}_{k})$$

is a local minimizer for this new problem. Furthermore the corresponding first and second-order adjoint states are still $\bar p$ and W. Moreover the set corresponding to $\left\{(\Delta f[t,u],\Delta f_x[t,u])\;\middle|\;u\in\overline{U}(t)\right\}$ for this new problem is equal to $\overline{D}(t)$. Therefore the maximality condition (3.9) follows. □

6.2 Necessary optimality conditions in integral form

To prove Theorem 3.6, we shall use a second-order necessary optimality condition in integral form that we describe next. Define

$$\begin{aligned}J^2_{K_0}(\bar x(0))=\big\{(y_0,w_0)\in\mathbb{R}^n\times\mathbb{R}^n\;\big|\;\forall\, h_i\to 0+,\ \exists\,(y_0^i,w_0^i)\to(y_0,w_0),\\ \langle\bar p(0)-\nabla_{x_1}\varphi(\bar x(0),\bar x(1)),y_0^i\rangle=0,\ \bar x(0)+h_iy_0^i+h_i^2w_0^i\in K_0\big\}\,,\end{aligned}$$

$$J^2(t) = \{(u,v) \in \mathbb{R}^n \times \mathbb{R}^n \,|\, \forall\, h_i \to 0+,\ \exists\,(u_i,v_i) \to (u,v),\ \mathcal{H}_u[t]u_i = 0,\ \bar{u}(t) + h_i u_i + h_i^2 v_i \in U(t)\}.$$

The above sets are closed, contain zero and for every $(u,v) \in J^2(t)$ we have $u \in T^{\flat}_{U(t)}(\bar{u}(t))$ and $\mathcal{H}_u[t]u = 0$. Moreover, if $u \in T^{\flat}_{U(t)}(\bar{u}(t))$ is such that $\mathcal{H}_u[t]u = 0$ and $v \in T^{\flat(2)}_{U(t)}(\bar{u}(t), u)$, then $(u,v) \in J^2(t)$. Finally, as we prove in Proposition 6.2 below, $J^2(\cdot)$ is measurable. This will allow us to apply a measurable selection theorem to deduce pointwise necessary conditions from the integral necessary conditions. Define the *local critical cone*, at $\bar{u} \in \mathcal{U}$ by

$$\mathcal{C}_{loc}(\bar{u}) := \{u(\cdot) \in L^1([0,1];\mathbb{R}^m) \mid u(t) \in T^{\flat}_{U(t)}(\bar{u}(t)) \text{ and } \mathcal{H}_u[t]u(t) = 0 \text{ a.e. in } [0,1]\}, \tag{6.16}$$

and the set of admissible second-order variations of $\bar{u}$ by

$$\mathcal{M}^{(2)}(\bar{u}) = \Big\{(u,v) \in L^2([0,1];\mathbb{R}^m) \times L^1([0,1];\mathbb{R}^m) \mid u \in \mathcal{C}_{loc}(\bar{u}),\ \ (u(t),v(t)) \in J^2(t) \text{ a.e. in } A\Big\}.$$

Consider the *linearized system*

$$\begin{cases} \dot{y}(t) = f_x[t]y(t) + f_u[t]u(t), & \text{a.e. in } [0,1], \\ y(0) = y_0, \end{cases} \tag{6.17}$$

and define the quadratic functional Φ : for any $(u(\cdot),v(\cdot)) \in L^2([0,1];\mathbb{R}^m) \times L^1([0,1];\mathbb{R}^m)$ and any $y_0 \in \mathbb{R}^n,\ w_0 \in \mathbb{R}^n$,

$$\begin{aligned}\Phi(u,v,y_0,w_0) := \langle \nabla_{x_1}\varphi(\bar{x}(0),\bar{x}(1)) - \bar{p}(0), w_0\rangle + \frac{1}{2}(y_0,y(1))^T\varphi''(\bar{x}(0),\bar{x}(1))(y_0,y(1)) \\ - \int_0^1 \Big(\mathcal{H}_u[t]v(t) + \frac{1}{2}y(t)^T\mathcal{H}_{xx}[t]y(t) + u(t)^T\mathcal{H}_{xu}[t]y(t) + \frac{1}{2}u(t)^T\mathcal{H}_{uu}[t]u(t)\Big)\,dt,\end{aligned}$$

where $y(\cdot)$ is the solution of (6.17). The second-order necessary optimality condition below generalizes [13, Thm 4.1] to problems with initial state constraints. It uses a larger set $\mathcal{M}^{(2)}(\bar{u})$ of admissible second-order variations of $\bar{u}$ than in [13]. For this reason the next result is stronger than [13, Theorem 4.1].

Theorem 6.1. *Let $(\bar{x},\bar{u}) \in \mathcal{P}(K_0)$ be a strong local minimizer of* (P)*, (A1), (A2) hold true and $\bar{p} \in W^{1,1}([0,1];\mathbb{R}^n)$ solve* (3.1)*. Then*

$$\Phi(u,v,y_0,w_0) \geq 0 \quad \text{for all} \quad (u,v) \in \mathcal{M}^{(2)}(\bar{u}) \ \text{ and } (y_0,w_0) \in J^2_{K_0}(\bar{x}(0)).$$

We first prove two useful propositions. It is not restrictive to assume that $\bar{u}(t) \in U(t)$ for every $t \in [0,1]$. For any $i \in \mathbb{N}$ and $t \in [0,1]$ define

$$V^i(t) := \{(u,v) \in \mathbb{R}^m \times \mathbb{R}^m \mid |(u,v)| \leq i,\ \mathcal{H}_u[t]u = 0\}.$$

By [2, Thm. 8.2.9], $t \rightsquigarrow V^i(t)$ is measurable and has closed images. Then, by the same theorem, for all $q \in \mathbb{Q}_+$, the set-valued map

$$Q^i_q(t) := \big\{(u,v) \in V^i(t) \,\big|\, \bar{u}(t) + qu + q^2v \in U(t)\big\},$$

is also measurable and has closed images. Consequently, for every integer $s \geq 1$ the set valued map $\widehat{Q}^i_s : [0,1] \rightsquigarrow \mathbb{R}^m$ defined by

$$\widehat{Q}^i_s(t) := \overline{\bigcup_{r\in\mathbb{Q}_+} \bigcap_{q\in\mathbb{Q}_+\cap(0,r)} \Big(Q^i_q(t) + \frac{1}{s}\overline{B}\Big)},$$

is measurable and therefore the set-valued map

$$Q(t) := \overline{\bigcup_{i\geq 1}\bigcap_{s\geq 1}\widehat{Q}^i_s(t)},$$

is also measurable and has closed images containing zero.

Proposition 6.2. *For every $t \in [0,1]$ we have $J^2(t) = Q(t)$.*

Proof. Fix $(u,v) \in J^2(t)$. Then for every $q \in \mathbb{Q}_+$ there exist $(u_q, v_q) \in \mathbb{R}^m \times \mathbb{R}^m$ such that $\lim_{q\to 0+}(u_q,v_q) = (u,v)$, $\mathcal{H}_u[t]u_q = 0$ and $\bar{u}(t) + qu_q + q^2 v_q \in U(t)$. The sequence of sets $\{\widehat{Q}^i_s(t)\}_i$ is nondecreasing (with respect to i). Let $i \geq 1$ be such that $|(u,v)| < i$. Then for all $q \in \mathbb{Q}_+$ sufficiently small we have $|(u_q,v_q)| \leq i$ and therefore $(u_q,v_q) \in Q^i_q(t)$ for all small $q \in \mathbb{Q}_+$. Fix an integer $s \geq 1$. Then for all small $r > 0$ we have $(u,v) \in \bigcap_{q\in\mathbb{Q}_+\cap(0,r)} \left(Q^i_q(t) + \frac{1}{s}\overline{B}\right)$. Since $s \geq 1$ is arbitrary, we deduce that $(u,v) \in \bigcap_{s\geq 1}\widehat{Q}^i_s(t)$. Thus $(u,v) \in Q(t)$.
Conversely, let $(u,v) \in Q(t)$ and $h_k \to 0+$. We have to show that there exist $(u_k,v_k) \to (u,v)$ such that $\mathcal{H}_u[t]u_k = 0$ and $\bar{u}(t) + h_k u_k + h_k^2 v_k \in U(t)$. It is not restrictive to assume that $h_k \in \mathbb{Q}_+$. Fix $s \geq 1$. Then there exist i and $(\tilde{u}_k,\tilde{v}_k) \in Q^i_{h_k}(t)$ such that $(u,v) \in (\tilde{u}_k,\tilde{v}_k) + \frac{2}{s}\overline{B}$ for all large k. In particular $\mathcal{H}_u[t]\tilde{u}_k = 0$ and $\bar{u}(t) + h_k\tilde{u}_k + h_k^2\tilde{v}_k \in U(t)$. Letting s tend to ∞, we construct a sequence (u_k,v_k) as desired. □

Proposition 6.3. *Consider $(u,v) \in L^\infty([0,1];\mathbb{R}^m\times\mathbb{R}^m)$ such that $(u(t),v(t)) \in J^2(t)$* a.e. in *A. Then for every sequence $h_k \to 0+$ there exists a bounded sequence $(u_k,v_k) \in L^\infty([0,1];\mathbb{R}^m\times\mathbb{R}^m)$ converging to (u,v) a.e. in A such that $\mathcal{H}_u[t]u_k(t) = 0$ and $\bar{u}(t) + h_k u_k(t) + h_k^2 v_k(t) \in U(t)$ for a.e. $t \in A$.*

Proof. We may assume that $(u(t),v(t)) = 0$ for $t \notin A$. Observe that for any fixed $t \in [0,1]$ the sequence of sets $P^i(t) := \bigcap_{s\geq 1}\widehat{Q}^i_s(t)$ is increasing with respect to i. Furthermore, by Proposition 6.2, $\lim_{i\to\infty}$ $\mathrm{dist}_{P^i(t)}((u(t),v(t))) = 0$. This and [2, Cor. 8.2.13] imply that we can find measurable selections $(u_i(t),v_i(t)) \in P^i(t)$ converging to (u,v) a.e. such that

$$\mathrm{dist}_{P^i(t)}((u(t),v(t))) = |(u(t),v(t)) - (u_i(t),v_i(t))| \leq |(u(t),v(t))|.$$

Hence the sequence $\|(u_i,v_i)\|_\infty$ is bounded by $2\|(u,v)\|_\infty$. Observe next that if the claim of our proposition is valid for every such (u_i,v_i), then, given a sequence $h_k \to 0+$ we can construct (u_k,v_k) for (u,v) as in the claim.
It remains to prove our proposition for every $i \geq 1$ and any bounded measurable selection $(u(t),v(t)) \in P^i(t)$. Consider such i and (u,v). Fix $s \geq 1$ and for every rational $r > 0$ and $t \in [0,1]$ define

$$R^s_r(t) := \bigcap_{q\in\mathbb{Q}_+\cap(0,r)}\left(Q^i_q(t) + \frac{1}{s}\overline{B}\right).$$

Notice that for every integer $s \geq 1$ the family of sets $R^s_r(t)$ is decreasing with respect to r. Hence

$$\lim_{r\to 0+,\, r\in\mathbb{Q}_+}\mathrm{dist}_{R^s_r(t)}((u(t),v(t))) = 0 \quad \text{a.e.},$$

and we deduce that

$$\lim_{k\to\infty}\mathrm{dist}_{Q^i_{h_k}(t)+\frac{1}{s}\overline{B}}((u(t),v(t))) = 0.$$

Since the set valued map $t \rightsquigarrow Q^i_{h_k}(t) + \frac{1}{s}\overline{B}$ is measurable and has closed nonempty images containing zero, by [2, Cor. 8.2.13 and Thm. 8.2.9] there exists a measurable selection $(\tilde{u}_k(t),\tilde{v}_k(t)) \in Q^i_{h_k}(t)$ such that

$\lim_{k\to\infty} \operatorname{dist}_{(\tilde{u}_k(t),\tilde{v}_k(t))+\frac{1}{s}\overline{B}}((u(t),v(t))) = 0$ a.e. and $|(\tilde{u}_k(t),\tilde{v}_k(t)) - (u(t),v(t))| \le |(u(t),v(t))| + 1/s$. Moreover, $\mathcal{H}_u[t]\tilde{u}_k(t) = 0$ and $\bar{u}(t) + h_k\tilde{u}_k(t) + h_k^2\tilde{v}_k(t) \in U(t)$. Making s to converge to ∞ we construct a bounded in L^∞ sequence (u_k, v_k) of measurable functions converging pointwise to (u,v) such that $\mathcal{H}_u[t]u_k(t) = 0$ and $\bar{u}(t) + h_k u_k(t) + h_k^2 v_k(t) \in U(t)$ a.e. □

We are now ready for the proof of Theorem 6.1 which, in its remaining part, is similar to the one of [13, Thm. 4.1] (but in the presence of initial point constraints), and so we skip some details. It is also similar to [18, Proof of Thm. 3.2], but for the Mayer problem instead of the Bolza one. Proposition 6.3 allows us to avoid additional assumptions imposed on u in [18, Thm. 3.2].

Proof of Theorem 6.1. Fix $(u(\cdot),v(\cdot)) \in \mathcal{M}^{(2)}(\bar{u})$, $(y_0,w_0) \in J^2_{K_0}(\bar{x}(0))$ and let y be the solution of (6.17). We claim that it is not restrictive to assume that $(u(\cdot),v(\cdot))$ is essentially bounded. Indeed, define

$$u_i(t) = \begin{cases} u(t) & \text{if } |u(t)| + |v(t)| \le i, \\ 0 & \text{otherwise,} \end{cases} \qquad v_i(t) = \begin{cases} v(t) & \text{if } |u(t)| + |v(t)| \le i, \\ 0 & \text{otherwise.} \end{cases}$$

Then $(u_i,v_i) \in \mathcal{M}^2(\bar{u})$ converge to (u,v) almost everywhere. Let y_i be the solution of (6.17) with u replaced by u_i. Then y_i converge to y uniformly on $[0,1]$. Observe that if $\Phi(u_i,v_i,y_0,w_0) \ge 0$ for all large i, then, using the Lebesgue dominated convergence theorem and taking the limit we get $\Phi(u,v,y_0,w_0) \ge 0$.
We continue the proof assuming that u, v are essentially bounded. Fix any sequence $h_i \to 0+$. By Proposition 6.3 there exists a bounded in $L^\infty([0,1];\mathbb{R}^m \times \mathbb{R}^m)$ family $\{(\tilde{u}_i,\tilde{v}_i)\}_{i\in\mathbb{N}}$ such that

$$\mathcal{H}_u[t]\tilde{u}_i(t) = 0, \quad \bar{u}(t) + h_i\tilde{u}_i(t) + h_i^2\tilde{v}_i(t) \in U(t), \quad \lim_{i\to\infty}(\tilde{u}_i(t),\tilde{v}_i(t)) = (u(t),v(t)) \quad \text{for a.e. } t \in A.$$

By [18, Prop. 4.1] there exists a bounded in $L^\infty([0,1];\mathbb{R}^m)$ family $\{\hat{u}_i\}_{i\in\mathbb{N}}$ such that for all i we have $\bar{u}(t) + h_i\hat{u}_i(t) \in U(t)$ a.e. in $[0,1]$ and $\lim_{i\to\infty}\hat{u}_i(t) = u(t)$. Define the new controls

$$u_i(t) = \begin{cases} \tilde{u}_i(t) & \text{if } t \in A, \\ \hat{u}_i(t) & \text{otherwise,} \end{cases} \qquad v_i(t) = \begin{cases} \tilde{v}_i(t) & \text{if } t \in A, \\ 0 & \text{otherwise.} \end{cases}$$

Let (y_0^i, w_0^i) be as in the definition of $J^2_{K_0}(\bar{x}(0))$ for (y_0,w_0) and our sequence h_i. For every $i \in \mathbb{N}$ large enough consider the solution $x_i : [0,1] \to \mathbb{R}^n$ of the system

$$\dot{x}_i(t) = f(t, x_i(t), \bar{u}(t) + h_i u_i(t) + h_i^2 v_i(t)), \quad x_i(0) = \bar{x}(0) + h_i y_0^i + h_i^2 w_0^i.$$

By the variational equation $(x_i - \bar{x})/h_i$ converge uniformly to y. Moreover, for all large $i \in \mathbb{N}$ we have $\varphi(x_i(0),x_i(1)) \ge \varphi(\bar{x}(0),\bar{x}(1))$. By the Taylor expansion and the properties of the adjoint state, setting $y_i = (x_i - \bar{x})/h_i$ we obtain similarly to [13, Proof of Thm. 4.1],

$$\begin{aligned} 0 &\le \varphi(x_i(0),x_i(1)) - \varphi(\bar{x}(0),\bar{x}(1)) \\ &= h_i^2\langle \nabla_{x_1}\varphi(\bar{x}(0),\bar{x}(1)) - \bar{p}(0), w_0\rangle + \int_0^1 \mathcal{H}_x[t]h_i y_i(t)\,dt - \int_0^1 \langle \bar{p}(t), \dot{x}_i(t) - \dot{\bar{x}}(t)\rangle\,dt \\ &\quad + \frac{h_i^2}{2}\langle \varphi''(\bar{x}(0),\bar{x}(1))(y_0,y(1)),(y_0,y(1))\rangle + o(h_i^2), \end{aligned} \tag{6.18}$$

and

$$\begin{aligned} \int_0^1 \langle \bar{p}(t), \dot{x}_i(t) - \dot{\bar{x}}(t)\rangle dt = \int_0^1 \big(&\mathcal{H}_x[t]h_i y_i(t) + \mathcal{H}_u[t]h_i^2 v(t) + \frac{h_i^2}{2}\langle \mathcal{H}_{xx}[t]y(t), y(t)\rangle \\ &+ h_i^2\langle \mathcal{H}_{ux}[t]u(t), y(t)\rangle + \frac{h_i^2}{2}\langle \mathcal{H}_{uu}[t]u(t), u(t)\rangle\big)\,dt + o(h_i^2). \end{aligned} \tag{6.19}$$

From (6.18) and (6.19) we deduce that

$$0 \leq h_i^2 \langle \nabla_{x_1} \varphi(\bar{x}(0), \bar{x}(1)) - \bar{p}(0), w_0 \rangle + \frac{h_i^2}{2} \langle \varphi''(\bar{x}(0), \bar{x}(1))(y_0, y(1)), (y_0, y(1)) \rangle$$
$$- h_i^2 \int_0^1 \left(\mathcal{H}_u[t]v(t) + \frac{1}{2} \langle \mathcal{H}_{xx}[t]y(t), y(t) \rangle + \langle \mathcal{H}_{xu}[t]y(t), u(t) \rangle + \frac{1}{2} \langle \mathcal{H}_{uu}[t]u(t), u(t) \rangle \right) dt + o(h_i^2). \tag{6.20}$$

Dividing by h_i^2 and passing to the limit we end the proof. □

6.3 Proof of Theorem 3.6

By [2, Thm. 8.2.9 and 8.5.1] and Proposition 6.2 the set-valued maps $[0,1] \ni s \rightsquigarrow \widehat{F}(s)$ and $[0,1] \ni s \rightsquigarrow G(s)$ given by

$$\widehat{F}(s) := \left\{ (u, v) \in J^2(s) \;\middle|\; \mathcal{H}_u[s]v + \frac{1}{2} u^T \mathcal{H}_{uu}[s]u = 0 \right\},$$
$$G(s) := \left\{ u \in T^{\flat}_{U(t)}(\bar{u}(t)) \;\middle|\; u^T \mathcal{H}_{uu}[s]u = 0 \right\},$$

are measurable and have closed images. Define

$$F(s) := \begin{cases} \widehat{F}(s) & \text{if } s \in A, \\ \{(u, 0) \mid u \in G(s)\} & \text{if } s \in [0,1] \backslash A. \end{cases}$$

Then $F(\cdot)$ is also measurable and has closed images. Moreover $0 \in F(s)$ for a.e. $s \in [0,1]$. This and the Castaing representation theorem, see for instance [2, Chapter 8], imply that it is sufficient to prove (3.8) and that for every bounded measurable selection $(u(s), v(s)) \in F(s)$ a.e. the following inequality holds true for a.e. $t \in [0,1]$

$$\left\langle f_u[t]^T \left(\mathcal{H}_{ux}[t] + W(t) f_u[t] \right) u(t), u(t) \right\rangle \leq 0. \tag{6.21}$$

Let us consider such a measurable selection $(u(\cdot), v(\cdot))$, $y_0 \in \Gamma_0$, $w_0 \in T^{\flat(2)}_{K_0}(\bar{x}(0), y_0)$ and let y be the solution of (6.17). By Theorem 6.1,

$$\langle \nabla_{x_1} \varphi(\bar{x}(0), \bar{x}(1)) - \bar{p}(0), w_0 \rangle + \frac{1}{2} (y_0, y(1))^T \varphi''(\bar{x}(0), \bar{x}(1))(y_0, y(1))$$
$$- \int_0^1 \left(\frac{1}{2} y(t)^T \mathcal{H}_{xx}[t]y(t) + u(t)^T \mathcal{H}_{xu}[t]y(t) \right) dt \geq 0.$$

In the same way as in the proof of Theorem 3.1, taking $u(\cdot) = 0$, we deduce from the above inequality that W satisfies the second-order transversality condition (3.8) with $\Psi = -\varphi''(\bar{x}(0), \bar{x}(1))$. Let $y_0 = 0 \in \Gamma_0$ and $w_0 = 0 \in T^{\flat(2)}_{K_0}(\bar{x}(0), 0)$. Then

$$y(t) = \int_0^t Y(t) Y^{-1}(s) f_u[s] u(s) ds,$$

where Y is the fundamental solution of system (3.6). By the last inequality, integrating by parts as in (6.6) and (6.7), we obtain

$$\int_0^1 \left\langle Y(t)^T \left(\mathcal{H}_{ux}[t] + W(t) f_u[t] \right) u(t), \int_0^t Y^{-1}(s) f_u[s] u(s) ds \right\rangle dt \leq 0.$$

Finally, defining $\mathcal{G}(t, u) := Y(t)^T \left(\mathcal{H}_{ux}[t] + W(t) f_u[t] \right) u$ and $G(t, u) := Y^{-1}(t) f_u[t] u$ we can repeat the arguments of Step 3 of the proof of Theorem 3.1 to deduce the pointwise statement (6.21) which completes the proof. □

7 Proofs of the second-order sensitivity relations

Define $\overline{\mathcal{H}}\colon [0,1]\times\mathbb{R}^n\times\mathbb{R}^n\to\mathbb{R}$ by

$$\overline{\mathcal{H}}(t,x,p)=\langle p, f(t,x,\bar{u}(t))\rangle\,.$$

Clearly assumption (A1) implies that $\overline{\mathcal{H}}(\cdot,x,p)$ is measurable for all $(x,p)\in\mathbb{R}^n\times\mathbb{R}^n$, $\overline{\mathcal{H}}(t,\cdot,\cdot)$ is twice differentiable on $(\bar{x}(t)+\rho B)\times\mathbb{R}^n$ for all $t\in[0,1]$ and $\overline{\mathcal{H}}_{pp}(t,x,p)\equiv 0$. Since $(\nabla\varphi(\bar{x}(1)),\Psi_2)\in J^{2,+}\varphi(\bar{x}(1))$, it follows from [6, Prop. 2.6] that there exists $\phi\in\mathcal{C}^2(\mathbb{R}^n;\mathbb{R})$ such that

$$\varphi\le\phi,\qquad \varphi(\bar{x}(1))=\phi(\bar{x}(1))\qquad\text{and}\qquad \big(\nabla\varphi(\bar{x}(1)),\Psi_2)=(\nabla\phi(\bar{x}(1)),\phi''(\bar{x}(1))\big)\,. \tag{7.1}$$

Consider the Hamiltonian system

$$\begin{cases} \dot{x}(t) &= \nabla_p\overline{\mathcal{H}}(t,x(t),p(t)), \quad x(1)=x_T,\\ -\dot{p}(t) &= \nabla_x\overline{\mathcal{H}}(t,x(t),p(t)), \quad p(1)=-\nabla\phi(x_T),\end{cases} \tag{7.2}$$

where $x_T\in\mathbb{R}^n$. Notice that for $x_T=\bar{x}(1)$, $(\bar{x},\bar{p})$ is a solution of (7.2). Moreover, it follows from the definition of $\overline{\mathcal{H}}$ and assumption (A1) that for all $x_T\in\mathbb{R}^n$ in a sufficiently small neighborhood of $\bar{x}(1)$ there exists a unique solution of (7.2). We associate with $\bar{x}$ and $\bar{p}$ the matrix differential equation

$$\begin{cases}\dot{R}(t)+\overline{\mathcal{H}}_{px}[t]R(t)+R(t)\overline{\mathcal{H}}_{xp}[t]+\overline{\mathcal{H}}_{xx}[t]=0,\\ R(1)=-\phi''(\bar{x}(1)),\end{cases} \tag{7.3}$$

where $[t]:=(t,\bar{x}(t),\bar{p}(t))$. Notice that its unique solution is the second-order adjoint state $W\in W^{1,1}([0,1];\mathbf{S}(n))$ from Theorem 3.1. Finally, for an open set $\mathcal{O}$ such that $\bar{x}(1)\in\mathcal{O}$ and $t\in[0,1]$, set

$$\Omega_t(\mathcal{O}):=\{(x(t),p(t))\mid (x,p)\text{ solves (7.2) and } x_T\in\mathcal{O}\}\,.$$

Remark 7.1. Using standard arguments, one can show that it follows from (A1) that there exists $\tilde{\rho}>0$, such that $\overline{\mathcal{H}}(t,\cdot,\cdot)$ is twice continuously differentiable in a $\rho/2$-neighborhood of $\Omega_t(\bar{x}(1)+\tilde{\rho}B)$ for all $t\in[0,1]$. Similarly one can show that there exists an integrable function $\tilde{k}:[0,1]\to\mathbb{R}_+$ such that for almost all $t\in[0,1]$, $\frac{\partial\overline{\mathcal{H}}}{\partial(x,p)}(t,\cdot,\cdot)$ is $\tilde{k}(t)$-Lipschitz on $\Omega_t(\bar{x}(1)+\tilde{\rho}B)$.

By [8, Thm. 2.3] and Remark 7.1 for all $t\in[t_0,1]$, the set $\mathcal{D}_t:=\{x(t)\mid (x,p)\text{ solves (7.2), } x_T\in\bar{x}(1)+\tilde{\rho}B\}$, is open and $\Omega_t(\bar{x}(1)+\tilde{\rho}B)$ is the graph of a $\mathcal{C}^1$-map.

7.1 Proof of Theorem 4.1

Let $\bar{x}$, $\bar{p}$ and W be as in the statement of Theorem 4.1 and ϕ be a function satisfying (7.1). Fix an arbitrary $t\in[t_0,1]$. Since $\mathcal{D}_t$ is an open neighborhood of $\bar{x}(t)$, it suffices to show that for all $x_t\in\mathbb{R}^n$ such that $(x_t,p_t)\in\Omega_t(\bar{x}(1)+\tilde{\rho}B)$ for some $p_t\in\mathbb{R}^n$ we have

$$V(t,x_t)-V(t,\bar{x}(t))\le\langle-\bar{p}(t),x_t-\bar{x}(t)\rangle+\frac{1}{2}\langle -W(t)(x_t-\bar{x}(t)),x_t-\bar{x}(t)\rangle+o(|x_t-\bar{x}(t)|^2). \tag{7.4}$$

Let (x,p) be a solution of (7.2) such that $(x_t,p_t):=(x(t),p(t))\in\Omega_t(\bar{x}(1)+\tilde{\rho}B)$. Using that the value function is non-decreasing along feasible trajectories and constant along optimal trajectories, we get

$$V(t,x_t)-V(t,\bar{x}(t))\le V(1,x(1))-V(1,\bar{x}(1))\le\phi(x(1))-\phi(\bar{x}(1)),$$

where we used (7.1) and that $V(1,x)=\varphi(x)$ for all $x\in\mathbb{R}^n$. Hence (7.1) implies

$$\begin{aligned}&V(t,x_t)-V(t,\bar{x}(t))\\&\quad\le\langle\nabla\varphi(\bar{x}(1)),x(1)-\bar{x}(1)\rangle+\frac{1}{2}\langle\Psi_2(x(1)-\bar{x}(1)),x(1)-\bar{x}(1)\rangle+o(|x(1)-\bar{x}(1)|^2).\end{aligned}\tag{7.5}$$

By Gronwall's lemma and the Lipschitz continuity of f with respect to x, there exists a constant $c>0$ such that for all $s\in[t,1]$,

$$|x(s)-\bar{x}(s)|\le c\,|x_t-\bar{x}(t)|\,.$$

This implies that for all $s\in[t,1]$, we can replace rest terms of order $o(|x(s)-\bar{x}(s)|^2)$ by rest terms of order $o(|x_t-\bar{x}(t)|^2)$. Using this and the properties of $\bar{p}$ and W we deduce from (7.5) that

$$\begin{aligned}&V(t,x_t)-V(t,\bar{x}(t))\\&\quad\le\langle-\bar{p}(1),x(1)-\bar{x}(1)\rangle+\frac{1}{2}\langle-W(1)(x(1)-\bar{x}(1)),x(1)-\bar{x}(1)\rangle+o(|x_t-\bar{x}(t)|^2)\\&\quad\le\langle-\bar{p}(t),x_t-\bar{x}(t)\rangle+\frac{1}{2}\langle-W(t)(x_t-\bar{x}(t)),x_t-\bar{x}(t)\rangle+o(|x_t-\bar{x}(t)|^2)\\&\qquad+\int_t^1\frac{d}{ds}\left(\langle-\bar{p}(s),x(s)-\bar{x}(s)\rangle+\frac{1}{2}\langle-W(s)(x(s)-\bar{x}(s)),x(s)-\bar{x}(s)\rangle\right)ds.\end{aligned}\tag{7.6}$$

We show next that the integral on the right-hand side of (7.6) is of order $o(|x_t-\bar{x}(t)|^2)$. Since (x,p) and $(\bar{x},\bar{p})$ are solutions of (7.2) we find that for every $t\in[t_0,1]$

$$\begin{aligned}&\int_t^1\frac{d}{ds}\langle-\bar{p}(s),x(s)-\bar{x}(s)\rangle\,ds\\&\qquad=\int_t^1\left(\left\langle\nabla_x\overline{\mathcal{H}}[s],x(s)-\bar{x}(s)\right\rangle+\left\langle-\bar{p}(s),\nabla_p\overline{\mathcal{H}}(s,x(s),p(s))-\nabla_p\overline{\mathcal{H}}[s]\right\rangle\right)ds.\end{aligned}\tag{7.7}$$

Similarly, using that W solves (7.3) we obtain

$$\begin{aligned}&-\frac{1}{2}\int_t^1\frac{d}{ds}\langle W(s)(x(s)-\bar{x}(s)),x(s)-\bar{x}(s)\rangle\,ds\\&\quad=-\int_t^1\left\langle W(s)\left(\nabla_p\overline{\mathcal{H}}(s,x(s),p(s))-\nabla_p\overline{\mathcal{H}}[s]\right),x(s)-\bar{x}(s)\right\rangle ds\\&\qquad-\frac{1}{2}\int_t^1\left\langle\left(-\overline{\mathcal{H}}_{px}[s]W(s)-W(s)\overline{\mathcal{H}}_{xp}[s]-\overline{\mathcal{H}}_{xx}[s]\right)(x(s)-\bar{x}(s)),x(s)-\bar{x}(s)\right\rangle ds.\end{aligned}\tag{7.8}$$

Consider the variational system of (7.2) obtained by differentiating the solution map of (7.2) with respect to x_T,

$$\begin{cases}\dot{X}(s)=\overline{\mathcal{H}}_{xp}[s]X(s), & X(1)=\mathrm{Id},\\ -\dot{P}(s)=\overline{\mathcal{H}}_{xx}[s]X(s)+\overline{\mathcal{H}}_{px}[s]P(t), & -P(1)=\phi''(x_T).\end{cases}$$

Then $W(s)=P(s)X(s)^{-1}$, and therefore,

$$W(s)\,(x(s)-\bar{x}(s))=p(s)-\bar{p}(s)+o_s(|x(1)-\bar{x}(1)|),\qquad\forall s\in[t,1].$$

For more details see for instance [6, 8]. With this observation (7.8) becomes

$$
\begin{aligned}
&-\frac{1}{2}\int_t^1 \frac{d}{ds}\langle W(s)(x(s)-\bar{x}(s)),x(s)-\bar{x}(s)\rangle\,ds\\
&=\int_t^1\left(\frac{1}{2}\left\langle\overline{\mathcal{H}}_{xx}[s](x(s)-\bar{x}(s)),x(s)-\bar{x}(s)\right\rangle+\left\langle\overline{\mathcal{H}}_{px}[s](p(s)-\bar{p}(s)),x(s)-\bar{x}(s)\right\rangle\right)ds\\
&\quad-\int_t^1\left(\left\langle\nabla_p\overline{\mathcal{H}}(s,x(s),p(s))-\nabla_p\overline{\mathcal{H}}[s],p(s)-\bar{p}(s)\right\rangle+o_s(|x_t-\bar{x}(t)|^2)\right)ds.
\end{aligned}
\tag{7.9}
$$

Finally, since $\left\langle p(s),\nabla_p\overline{\mathcal{H}}(s,x(s),p(s))\right\rangle=\overline{\mathcal{H}}(s,x(s),p(s))$ and similarly $\left\langle\bar{p}(s),\nabla_p\overline{\mathcal{H}}[s]\right\rangle=\overline{\mathcal{H}}[s]$, (7.7) and (7.9) lead to

$$
\begin{aligned}
&\int_t^1\frac{d}{ds}\left(\langle-\bar{p}(s),x(s)-\bar{x}(s)\rangle+\frac{1}{2}\langle-W(s)(x(s)-\bar{x}(s)),x(s)-\bar{x}(s)\rangle\right)ds\\
&=\int_t^1\left(\frac{1}{2}\left\langle\overline{\mathcal{H}}_{xx}[s](x(s)-\bar{x}(s)),x(s)-\bar{x}(s)\right\rangle+\left\langle\overline{\mathcal{H}}_{px}[s](p(s)-\bar{p}(s)),x(s)-\bar{x}(s)\right\rangle-\overline{\mathcal{H}}(s,x(s),p(s))\right.\\
&\quad\left.+\overline{\mathcal{H}}[s]+\left\langle\nabla_x\overline{\mathcal{H}}[s],x(s)-\bar{x}(s)\right\rangle+\left\langle\nabla_p\overline{\mathcal{H}}[s],p(s)-\bar{p}(s)\right\rangle+o_s(|x_t-\bar{x}(t)|^2)\right)ds.
\end{aligned}
$$

Since $\overline{\mathcal{H}}(t,\cdot,\cdot)$ is of class $\mathcal{C}^2$ in a neighborhood of $\Omega_t(\bar{x}(1)+\tilde{\rho}B)$, it is not difficult to verify via Taylor's formula and the dominated convergence theorem that

$$
\int_t^1\frac{d}{ds}\left(\langle-\bar{p}(s),x(s)-\bar{x}(s)\rangle+\frac{1}{2}\langle-W(s)(x(s)-\bar{x}(s)),x(s)-\bar{x}(s)\rangle\right)ds=o(|x_t-\bar{x}(t)|^2).
$$

This, together with (7.6) completes the proof. □

7.2 Proof of Theorem 4.2

The proof of Theorem 4.2 follows along the same lines as the proof of Theorem 4.1. We therefore just provide the main ideas.
Since $(-\bar{p}(t_0),-W_0)\in J_x^{2,-}V(t_0,\bar{x}(t_0))$, we can use the characterization of the subjets, see for instance [6, Prop. 2.6], to deduce the existence of a function $g\in\mathcal{C}^2(\mathbb{R}^n;\mathbb{R})$ such that

$$
g(\cdot)\le V(t_0,\cdot),\qquad g(x_0)=V(t_0,x_0),\qquad(\nabla g(x_0),g''(x_0))=(-\bar{p}(t_0),-W_0).
\tag{7.10}
$$

Consider next the Hamiltonian system

$$
\begin{cases}\dot{x}(t)=\nabla_p\overline{\mathcal{H}}(t,x(t),p(t)), & x(t_0)=x_S,\\ -\dot{p}(t)=\nabla_x\overline{\mathcal{H}}(t,x(t),p(t)), & -p(t_0)=\nabla g(x_S).\end{cases}
\tag{7.11}
$$

Clearly $(\bar{x},\bar{p})$ solves (7.11) with $x_S=\bar{x}(t_0)=x_0$. Introducing the time transformation $\tau\colon t\mapsto 1-t+t_0$, notice that if (x,p) solves (7.11) for some $x_S\in\mathbb{R}^n$, then $(x_\tau,p_\tau):=(x(\tau(t)),p(\tau(t)))$ solves (7.2) with $x_\tau(1)=x_S$ and Hamiltonian $\overline{\mathcal{H}}_\tau(t,x,p):=-\overline{\mathcal{H}}(\tau(t),x,p)$. Moreover, $W_\tau(t):=W(\tau(t))$ is a solution of equation (7.3) with $\overline{\mathcal{H}}[t]$ replaced by $\overline{\mathcal{H}}_\tau(t,\bar{x}_\tau(t),\bar{p}_\tau(t))$ and $\phi''(\bar{x}(1))$ replaced by $g''(\bar{x}_\tau(1))$. As in the proof of Theorem 4.1 we deduce that for some $\tilde{\rho}>0$ and all $t\in[t_0,1]$ the set

$$
\tilde{\mathcal{D}}_t:=\{x(t)\mid(x,p)\text{ solves (7.11)},\ x_S\in\bar{x}(t_0)+\tilde{\rho}B\},
$$

is open and $W(t)(x(t)-\bar{x}(t)) = p(t) - \bar{p}(t) + o_t(|x(t_0) - \bar{x}(t_0)|)$. It suffices therefore to prove that for all (x,p) solving (7.11) with $x_S \in \bar{x}(t_0) + \tilde{\rho}B$ and all $t \in [t_0, 1]$ we have

$$V(t,x(t)) - V(t,\bar{x}(t)) \geq \langle -\bar{p}(t), x(t) - \bar{x}(t)\rangle + \frac{1}{2}\langle -W(t)(x(t)-\bar{x}(t)), x(t)-\bar{x}(t)\rangle + o(|x(t)-\bar{x}(t)|^2).$$

Fix $t \in [t_0, 1]$ and let (x,p) be a solution of (7.11) with $x_S \in \bar{x}(t_0) + \tilde{\rho}B$. Then it follows from the dynamic programming principle and (7.10) that

$$V(t,x(t)) - V(t,\bar{x}(t)) \geq V(t_0, x_S) - V(t_0, \bar{x}(t_0)) \geq g(x_S) - g(\bar{x}(t_0)).$$

Still using (7.10) this leads to

$$\begin{aligned}
&V(t,x(t)) - V(t,\bar{x}(t)) \\
&\quad \geq \langle -\bar{p}(t_0), x_S - \bar{x}(t_0)\rangle + \frac{1}{2}\langle -W(t_0)(x_S - \bar{x}(t_0)), x_S - \bar{x}(t_0)\rangle + o(|x_S - \bar{x}(t_0)|^2) \\
&\quad = \langle -\bar{p}(t), x(t) - \bar{x}(t)\rangle + \frac{1}{2}\langle -W(t)(x(t)-\bar{x}(t)), x(t)-\bar{x}(t)\rangle + o(|x_S - \bar{x}(t_0)|^2) \\
&\qquad - \int_{t_0}^{t} \frac{d}{ds}\left(\langle -\bar{p}(s), x(s) - \bar{x}(s)\rangle + \frac{1}{2}\langle -W(s)(x(s)-\bar{x}(s)), x(s)-\bar{x}(s)\rangle\right) ds.
\end{aligned}$$

Similar arguments as those used in the proof of Theorem 4.1 can be employed to show that

$$\begin{aligned}
\int_{t_0}^{t} \frac{d}{ds}\left(\langle -\bar{p}(s), x(s) - \bar{x}(s)\rangle + \frac{1}{2}\langle -W(s)(x(s)-\bar{x}(s)), x(s)-\bar{x}(s)\rangle\right) ds + o(|x_S - \bar{x}(t_0)|^2) \\
= o(|x(t)-\bar{x}(t)|^2),
\end{aligned}$$

which completes the proof. □

References

[1] M. Aronna, J. Bonnans, A. V. Dmitruk, and P. Lotito, *Quadratic order conditions for bang-singular extremals*, Numer Algebra Contr Optim, special issue in honor of Helmut Maurer, 2 (2012), pp. 511–546.

[2] J.-P. Aubin and H. Frankowska, *Set-Valued Analysis, Systems Control Found. Appl.*, Birkhäuser Boston, Inc., Boston, MA, 1990.

[3] M. Bardi and P. Goatin, *Invariant sets for controlled degenerate diffusions: a viscosity solutions approach*, in Stochastic Analysis, Control, Optimization and Applications, Birkhäuser, 1999, pp. 191–208.

[4] J. F. Bonnans, X. Dupuis, and L. Pfeiffer, *Second-order necessary conditions in Pontryagin form for optimal control problems*, SIAM J. Control Optim., 52 (2014), pp. 3887–3916.

[5] J. F. Bonnans and A. Hermant, *No-gap second-order optimality conditions for optimal control problems with a single state constraint and control*, Math. Program., 117 (2009), pp. 21–50.

[6] P. Cannarsa, H. Frankowska, and T. Scarinci, *Second-order sensitivity relations and regularity of the value function for Mayer's problem in optimal control*, SIAM J. Control Optim., 53 (2015), pp. 3642–3672.

[7] P. Cannarsa and C. Sinestrari, *Semiconcave Functions, Hamilton-Jacobi Equations and Optimal Control*, Birkhäuser Boston, 2004.

[8] N. Caroff and H. Frankowska, *Conjugate points and shocks in nonlinear optimal control*, Trans. Amer. Math. Soc., 348 (1996).

[9] A. V. Dmitruk, *Quadratic order conditions of a local minimum for abnormal extremals*, Nonlinear Anal., 30 (1997), pp. 2439–2448.

[10] H. Frankowska, *The maximum principle for an optimal solution to a differential inclusion with end point constraints*, SIAM J. Control Optim., 25 (1987), pp. 145–157.

[11] H. Frankowska and D. Hoehener, *Jacobson type necessary optimality conditions for general control systems*, in 54th IEEE Conference on Decision and Control, December 15-18, Osaka, Japan, 2015, pp. 1304–1309.

[12] H. Frankowska, D. Hoehener, and D. Tonon, *A second-order maximum priniciple in optimal control under state constraints*, Serdica Math. J, 39 (2013), pp. 233–270.

[13] H. Frankowska and N. Osmolovskii, *Second-order necessary optimality conditions for the Mayer problem subject to a general control constraint*, in Analysis and Geometry in Control Theory and its Applications, vol. 12 of Springer INDAM series, Springer Verlag, 2015.

[14] H. Frankowska and D. Tonon, *Pointwise second-order necessary optimality conditions for the Mayer problem with control constraints*, SIAM J. Control Optim., 51 (2013), pp. 3814–3843.

[15] E. G. Gilbert and D. S. Bernstein, *Second-order necessary conditions in optimal control: Accessory-problem results without normality conditions*, J. Optim. Theory Appl., 41 (1983), pp. 75–106.

[16] B. S. Goh, *Necessary conditions for singular extremals involving multiple control variables*, SIAM J. Control, 4 (1966), pp. 716–731.

[17] J.-B. Hiriart-Urruty and A. Seeger, *A variational approach to copositive matrices*, SIAM Rev., 52 (2010), pp. 593–629.

[18] D. Hoehener, *Variational approach to second-order optimality conditions for control problems with pure state constraints*, SIAM J. Control Optim., 50 (2012), pp. 1139–1173.

[19] ——, *Feasible perturbations of control systems with pure state constraints and applications to second-order optimality conditions*, Appl. Math. Optim., 68 (2013), pp. 219–253.

[20] D. H. Jacobson, *A new necessary condition of optimality for singular control problems.*, SIAM J. Control, 7 (1969), pp. 578–595.

[21] D. H. Jacobson and J. L. Speyer, *Necessary and sufficient conditions for optimality for singular control problems: A limit approach*, J. Math. Anal. Appl., 34 (1971), pp. 239–266.

[22] H. Lou, *Second-order necessary/sufficient optimality conditions for optimal control problems in the absence of linear structure*, Discrete Contin. Dyn. Syst. Ser. B, 14 (2010), pp. 1445–1464.

[23] K. Malanowski, H. Maurer, and S. Pickenhain, *Second-order sufficient conditions for state-constrained optimal control problems*, J. Optim. Theory Appl., 123 (2004), pp. 595–617.

[24] A. A. Milyutin and N. P. Osmolovskii, *Calculus of Variations and Optimal Control*, American Mathematical Society, Providence, RI, 1998.

[25] Z. Páles and V. Zeidan, *First- and second-order necessary conditions for control problems with constraints*, Trans. Amer. Math. Soc., 346 (1994), pp. 421–453.

[26] L. S. Pontryagin, V. G. Boltyanskii, R. V. Gamkrelidze, and E. F. Mishchenko, *The Mathematical Theory of Optimal Processes*, Interscience Publishers, Inc., New York, 1962.

[27] R. Rockafellar and R. Wets, *Variational Analysis*, Springer-Verlag, 1998.

[28] H. Schättler and U. Ledzewicz, *Geometric optimal control*, vol. 38 of Interdisciplinary Applied Mathematics, Springer, New York, 2012. Theory, methods and examples.

[29] R. B. Vinter, *Optimal Control*, Birkhäuser, Boston, MA, 2000.