Working Paper

Conjugate Points and Shocks in Nonlinear Optimal Control

H. Frankowska
N. Caroff

WP-94-56
July 1994

IIASA International Institute for Applied Systems Analysis □ A-2361 Laxenburg □ Austria
Telephone: +43 2236 71521 □ Telex: 079 137 iiasa a □ Telefax: +43 2236 71313

Conjugate Points and Shocks in Nonlinear Optimal Control

H. Frankowska
N. Caroff

WP-94-56
July, 1994

IIASA International Institute for Applied Systems Analysis ▫ A-2361 Laxenburg ▫ Austria
Telephone: +43 2236 71521 ▫ Telex: 079 137 iiasa a ▫ Telefax: +43 2236 71313

Conjugate Points and Shocks in Nonlinear Optimal Control

Hélène Frankowska
CNRS, CEREMADE,
Université Paris-Dauphine
75775 Paris Cedex 16, France

Nathalie Caroff
Department of Mathematics,
Université de Bordeaux II
Bordeaux, France

Foreword

In this paper the authors use the method of characteristics to extend the Jacobi conjugate points theory to the Bolza problem arising in nonlinear optimal control. This yields necessary and sufficient optimality conditions for weak and strong local minima stated in terms of the existence of a solution to a corresponding matrix Riccati differential equation. The same approach allows to investigate as well smoothness of the value function.

Key Words: Hamilton-Jacobi-Bellman equation, characteristics, conjugate point, necessary and sufficient conditions for optimality, Riccati differential equation, shock, value function, weak local minimum.

AMS (MOS) Subject Classification: 35B37, 35L67, 49K15, 49L05, 49L20

1 Introduction

Consider the Hamilton-Jacobi equation

$$-\frac{\partial V}{\partial t} + H\left(t, x, -\frac{\partial V}{\partial x}\right) = 0, \quad V(T,\cdot) = \varphi(\cdot) \tag{1}$$

The classical method of characteristics applied to this equation exhibits shocks, which justify that its solutions should be nonsmooth. Then different criteria are used to get continuous (or even discontinuous) solutions, by eliminating some "pieces" of characteristics (cf. the entropy and Rankine-Hugoniot conditions [15] or the properties of one sided limits [7]). In this paper we shall consider the Hamiltonian H associated to the Bolza problem in optimal control theory. Then, in the same way than [6], the solution to (1) is the value function of the Bolza problem, which may be nonsmooth. To study characteristics of (1) in the context of optimal control is particularly rewarding because the characteristic system

$$\left\{ \begin{array}{rcl} x' &=& \dfrac{\partial H}{\partial p}(t,x,p) \qquad\qquad x(T) = x_T \\ \\ -p' &=& \dfrac{\partial H}{\partial x}(t,x,p) \qquad\qquad p(T) = -\nabla\varphi(x_T) \end{array} \right. \tag{2}$$

is Pontryagin's first order necessary condition for optimality, which performs in the optimal control theory the same role as the Euler-Lagrange equation in the calculus of variations.

As long as there is no shock the value function remains smooth and characteristics are the optimal state-costate pairs. What happens when a shock does occur? We provide an answer based on the use of conjugate point along a solution (x,p) to (2).

To be more precise consider the Bolza problem

$$\text{minimize} \int_{t_0}^{T} L(t, x(t), u(t))\, dt + \varphi(x(T)) \tag{3}$$

over trajectory-control pairs (x, u) of the control system

$$x' = f(t, x, u(t)), \quad x(t_0) = x_0, \qquad u(t) \in U \tag{4}$$

It is well known that any optimal trajectory-control pair $(\overline{x}, \overline{u})$ of the above problem satisfies the maximum principle : There exists an absolutely continuous function $p : [t_0, T] \to \mathbf{R}^n$ such that $(\overline{x}, p)$, called optimal state-costate pair, solves the Hamiltonian system

$$\begin{cases} x' = \dfrac{\partial H}{\partial p}(t,x,p) & x(t_0) = x_0 \\ -p' = \dfrac{\partial H}{\partial x}(t,x,p) & p(T) = -\nabla\varphi(\overline{x}(T)) \end{cases} \tag{5}$$

where $H : [0,T] \times \mathbf{R}^n \times \mathbf{R}^n \to \mathbf{R}$ is given by

$$H(t,x,p) = \sup_{u\in U}(\langle p, f(t,x,u)\rangle - L(t,x,u)) \tag{6}$$

In general, the system (5) does not have an unique solution because the initial condition for $p(\cdot)$ at t_0 is not known. For this very reason, the necessary condition for optimality given by the maximum principle is not sufficient. In the other words, $(\overline{x},p)$ solves the characteristic system (2) for $x_T = \overline{x}(T)$. But since only the initial condition for $\overline{x}$ at t_0 is fixed and since a shock may happen, i.e. two different characteristics (x_i,p_i), $i = 1,2$ may verify $x_i(t_0) = x_0$, so that the necessary condition (5) is not sufficient.

It can, however, be shown that $p(\cdot)$ may be chosen in such way that $-p(t_0)$ is equal to the gradient with respect to x of the cost function $V : [0,T] \times \mathbf{R}^n \to \mathbf{R}$ associated to the above problem provided $\frac{\partial V}{\partial x}(t_0,x_0)$ does exist. We may consider then the Cauchy problem

$$\begin{cases} x' = \dfrac{\partial H}{\partial p}(t,x,p) & x(t_0) = x_0 \\ -p' = \dfrac{\partial H}{\partial x}(t,x,p) & p(t_0) = -\dfrac{\partial V}{\partial x}(t_0,x_0) \end{cases}$$

When ∇H is locally Lipschitz, it has at most one solution and, in this way, the necessary condition (5) becomes a sufficient one. When $V(t_0,\cdot)$ is not differentiable at x_0, the gradient of V has to be replaced by any element from the Painlevé-Kuratowski upper limit $\mathrm{Limsup}_{x\to x_0,t\to t_0}\left\{\frac{\partial V}{\partial x}(t,x)\right\}$ to express sufficient conditions for optimality (see section 6). An easy consequence of the above is the following interesting behavior of solutions to (1): $V(t_0,\cdot)$ is differentiable at x_0 if and only if the optimal trajectory of the Bolza problem (3), (4) is unique.

Optimal solutions help to distinguish between "the good and the bad" characteristics. Indeed, when H is strictly convex in the last variable and V is semiconcave, which happens under an appropriate smoothness of data (see for instance [3,4]), then for all $t > t_0$, V is differentiable at $(t, \overline{x}(t))$, i.e. the optimal trajectory enters immediately into the domain of differentiability of V (see section 6). Consequently, for all $t > t_0$, $p(t) = -\frac{\partial V}{\partial x}(t, \overline{x}(t))$. The first results in this direction in the context of Mayer's problem were obtained in [3].

In this paper we go beyond the necessary condition (5), by further investigating characteristics of (2). Namely, we associate to a given solution (x,p) of (2) the matrix Riccati differential equation

$$\left\{ \begin{array}{l} P' + \dfrac{\partial^2 H}{\partial p \partial x}(t, x(t), p(t))P \; + \; P\, \dfrac{\partial^2 H}{\partial x \partial p}(t, x(t), p(t)) + \\ \\ + \; P\, \dfrac{\partial^2 H}{\partial p^2}(t, x(t), p(t))P \; + \; \dfrac{\partial^2 H}{\partial x^2}(t, x(t), p(t)) \; = \; 0, \quad P(T) = -\varphi''(x(T)) \end{array} \right. \tag{7}$$

whose solution $P(\cdot)$ may escape to infinity in a finite time $t < T$. This equation was used in [2] to investigate the global regularity of the value function and sufficiency of (5) to provide global minimum to the Bolza problem. We define the conjugate point (to T) along (x,p) by

$$t_c = \inf_{t \in [t_0, T]} \{P \text{ is defined on } [t, T]\}$$

If $t_c > t_0$, then $\|P(t)\| \to +\infty$ when $t \to t_c+$.

The conjugate point performs an identical role than the Jacobi conjugate point in the calculus of variations [11,12]. Namely, we introduce the notion of weak (respectively strong) local minimum of (3), (4) by saying that a trajectory-control pair $(\overline{x}, \overline{u})$ is a weak (resp. strong) local minimum if and only if there exists $\varepsilon > 0$ such that for every trajectory-control pair (x, u) of the control system (4) satisfying $\|x' - \overline{x}'\|_{L^1(t_0,T)} < \varepsilon$ (resp. $\|x - \overline{x}\|_\infty < \varepsilon$) we have

$$\varphi(\overline{x}(T)) + \int_{t_0}^{T} L(s, \overline{x}(s), \overline{u}(s))ds \; \leq \; \varphi(x(T)) + \int_{t_0}^{T} L(s, x(s), u(s))ds$$

and show that results similar to the Jacobi conjugate points theory hold true also in this context. We underline that our notion of weak local minimum is different from those used in [13,14,18,19]. We prefer it for several reasons. On one hand the maximum principle in this case is exactly (5), while in the above papers another (localized) necessary conditions, not related to characteristics, are given and it is often required that $\overline{u}$ is an interior control. Also in [16,20] two different Hamiltonians are considered, one to state sufficient conditions and a different one to formulate necessary ones, while here we use only the Hamiltonian defined by (6).

In contrast with the classical calculus of variations (and [18,19]), our results rely on the dynamic programming principle rather than the computation of second order variations (with respect to controls) and consideration of a Jacobi equation, as it was done in [13,14,18,19], where the interested reader can get as well a further bibliography on this subject. Relations between properties of solutions to the Jacobi and Riccati equations were often observed both in the calculus of variations and optimal control (see for instance [12,13,17]). However the global existence of a solution to the Riccati equation here is rather related to the preservation of the regularity of the value function along optimal solutions, than with the Jacobi equation.

The outline of the paper is as follows. Section 2 is devoted to the relationship between the matrix Riccati differential equations and shocks of characteristics. Section 3 provides necessary and sufficient conditions for local minima of the Bolza problem. Smoothness of the value function is investigated in section 4.

2 Matrix Riccati Equations and Shocks

We relate here the absence of shocks of the Hamilton-Jacobi-Bellman equation with the existence of solutions to matrix Riccati differential equations. For this aim we shall use the following tool:

Definition 2.1 *For a locally Lipschitz around $x_0 \in \mathbf{R}^n$ function $\psi : \mathbf{R}^n \mapsto \mathbf{R}$ define the compact set*

$$\partial^\star \psi(x_0) = \mathrm{Limsup}_{x \to x_0} \{\nabla \psi(x)\}$$

where Limsup *denotes the upper set-valued limit (see for instance* [1]*).*

Theorem 2.2 ([5]) *Consider a locally Lipschitz around $x_0 \in \mathbf{R}^n$ function $\psi : \mathbf{R}^n \mapsto \mathbf{R}$. If $\partial^\star \psi(x_0)$ is a singleton, then ψ is differentiable at x_0.*

Let $H : [0,T] \times \mathbf{R}^n \times \mathbf{R}^n \mapsto \mathbf{R}$ be such that $H(t,\cdot,\cdot)$ is differentiable. We associate to it the Hamiltonian system

$$\begin{cases} x'(t) &= \dfrac{\partial H}{\partial p}(t,x(t),p(t)), \quad x(t_0) = x_0 \\[2ex] -p'(t) &= \dfrac{\partial H}{\partial x}(t,x(t),p(t)), \quad p(t_0) = p_0 \end{cases} \tag{8}$$

It is called *complete* if for every $(t_0, x_0, p_0) \in [0,T] \times \mathbf{R}^n \times \mathbf{R}^n$, the solution to (8) is unique and defined on $[0,T]$. The Hamiltonian system (8) is complete if for instance

$$\left\{ \begin{array}{l} \forall\, r > 0,\ \exists\, \gamma_r \in L^1(0,T) \text{ such that for almost every } t \in [0,T], \\ \\ \dfrac{\partial H}{\partial(x,p)}(t,\cdot,\cdot) \text{ is } \gamma_r(t) - \text{Lipschitz on } B_r(0) \times B_r(0) \end{array} \right. \tag{9}$$

and has a linear growth: for some $k \in L^1(0,T)$

$$\forall\, x,\ p \in \mathbf{R}^n, \quad \left\| \frac{\partial H}{\partial(x,p)}(t,x,p) \right\| \leq k(t)\left(\|x\| + \|p\| + 1\right)$$

Example — Consider

$$f : [0,T] \times \mathbf{R}^n \mapsto \mathbf{R}^n,\ g : [0,T] \times \mathbf{R}^n \mapsto L(U,\mathbf{R}^n),\ l : [0,T] \times \mathbf{R}^n \mapsto \mathbf{R}$$

where U is a finite dimensional space and let $R(t,x) \in L(U,U)$ be self-adjoint and positive for every $(t,x) \in [0,T] \times \mathbf{R}^n$. Define

$$H(t,x,p) = \langle p, f(t,x) \rangle + \sup_{u \in U} \left(\langle p, g(t,x)u \rangle - \frac{1}{2} \langle R(t,x)u, u \rangle \right) - l(t,x)$$

Then it is not difficult to check that

$$H(t,x,p) = \langle p, f(t,x) \rangle + \left\langle R(t,x)^{-1} g(t,x)^\star p,\ g(t,x)^\star p \right\rangle - l(t,x)$$

An appropriate smoothness of $f(t,\cdot),\ g(t,\cdot),\ l(t,\cdot),\ R(t,\cdot)^{-1}$ implies differentiability of $H(t,\cdot,\cdot)$ and completeness of the associated Hamiltonian system. □

Consider $\psi : \mathbf{R}^n \mapsto \mathbf{R}^n$ and the Hamiltonian system

$$\left\{ \begin{array}{rcl} x'(t) & = & \dfrac{\partial H}{\partial p}(t, x(t), p(t)),\ \ x(T) = x_T \\ \\ -p'(t) & = & \dfrac{\partial H}{\partial x}(t, x(t), p(t)),\ \ p(T) = \psi(x_T) \end{array} \right. \tag{10}$$

Definition 2.3 *The system* (10) *has a* shock *at time* t_0 *if there exist two solutions* $(x_i, p_i)(\cdot)$, $i = 1,2$ *of* (10) *such that*

$$x_1(t_0) = x_2(t_0) \quad \& \quad p_1(t_0) \neq p_2(t_0)$$

Theorem 2.4 *Assume that ψ is locally Lipschitz on an open set Ω, $H(t,\cdot,\cdot)$ is twice continuously differentiable, the Hamiltonian system (8) is complete and (9) holds true. Define the sets*

$$M_t(\Omega) = \{(x(t), p(t)) \mid (x,p) \text{ solves } (10), \ x_T \in \Omega\}$$

where $t \in [0,T]$. Then the following two statements are equivalent:

i) $\forall\, t \in [0,T]$, *the set*

$$\mathcal{D}_t = \{x(t) \mid (x,p) \text{ solves } (10), \ x_T \in \Omega\}$$

is open and $M_t(\Omega)$ is the graph of a locally Lipschitz function.

ii) $\forall\,(x,p)$ *solving* (10) *on* $[0,T]$ *and* $P_T \in \partial^\star\psi(x_T)$, *the matrix Riccati equation*

$$\begin{cases} P' + \dfrac{\partial^2 H}{\partial p \partial x}(t, x(t), p(t))P \ + \ P\,\dfrac{\partial^2 H}{\partial x \partial p}(t, x(t), p(t)) + \\ + \ P\,\dfrac{\partial^2 H}{\partial p^2}(t, x(t), p(t))P \ + \ \dfrac{\partial^2 H}{\partial x^2}(t, x(t), p(t)) \ = \ 0, \quad P(T) = P_T \end{cases} \tag{11}$$

has a solution on $[0,T]$.

Furthermore, if i) (or equivalently ii)) holds true, then

$$\psi \text{ is differentiable} \implies M_t(\Omega) \text{ is the graph of a differentiable function}$$

$$\psi \in C^1 \implies M_t(\Omega) \text{ is the graph of a } C^1 - \text{function}$$

Corollary 2.5 *Under all assumptions of Theorem 2.4, suppose that $\Omega = \mathbf{R}^n$ and that for every (x,p) solving* (10) *on* $[0,T]$ *and* $P_T \in \partial^\star\psi(x(T))$, *the matrix Riccati equation* (11) *has a solution on* $[0,T]$. *Then the Hamiltonian system* (10) *has no shock in* $[0,T]$.

The proof uses the variational equation of ODE to express the tangent space to $M_t(\Omega)$ at $(x(t), p(t))$.

3 Bolza Optimal Control Problem

Consider the Bolza minimization problem

$$(P) \qquad \min \int_{t_0}^{T} L(t, x(t), u(t))dt \; + \; \varphi(x(T))$$

over solution-control pairs (x, u) of the control system

$$x'(t) = f(t, x(t), u(t)), \quad x(t_0) = x_0, \quad u(t) \in U \tag{12}$$

where $t_0 \in [0, T]$, $x_0 \in \mathbf{R}^n$, U is a complete separable metric space,

$$\varphi : \mathbf{R}^n \mapsto \mathbf{R}, \;\; L : [0, T] \times \mathbf{R}^n \times U \mapsto \mathbf{R}, \;\; f : [0, T] \times \mathbf{R}^n \times U \mapsto \mathbf{R}^n$$

are continuous functions. We denote by $\mathcal{U}$ the set of all measurable controls $u : [0, T] \mapsto U$ and by $x(\cdot; t_0, x_0, u)$ the solution to (12) starting at time t_0 from the initial condition x_0 and corresponding to the control $u(\cdot) \in \mathcal{U}$ (the assumptions we shall impose below imply that it is at most unique). In general not to every $u \in \mathcal{U}$ corresponds such a solution. For all $(t_0, x_0, u) \in [0, T] \times \mathbf{R}^n \times \mathcal{U}$ set

$$\Phi(t_0, x_0, u) \; = \; \int_{t_0}^{T} L(t, x(t; t_0, x_0, u), u(t))dt + \varphi(x(T; t_0, x_0, u))$$

if this expression is well defined and $\Phi(t_0, x_0, u) = +\infty$ otherwise.

The value function associated to the Bolza problem (P) is given by

$$V(t_0, x_0) \; = \; \inf_{u \in \mathcal{U}} \Phi(t_0, x_0, u)$$

when (t_0, x_0) range over $[0, T] \times \mathbf{R}^n$.

Definition 3.1 *A trajectory-control pair* $(\overline{x}, \overline{u})$ *of* (12) *is called* weakly locally optimal *for the problem* (P) *if there exists* $\varepsilon > 0$ *such that for every trajectory-control pair* (x, u) *of* (12)

$$\|x' - \overline{x}'\|_{L^1(t_0, T)} < \varepsilon \quad \Longrightarrow \quad +\infty \neq \Phi(t_0, x_0, \overline{u}) \leq \Phi(t_0, x_0, u)$$

It is called strongly locally optimal *if there exists* $\varepsilon > 0$ *such that for every trajectory-control pair* (x, u) *of* (12)

$$\|x - \overline{x}\|_{\infty} < \varepsilon \quad \Longrightarrow \quad +\infty \neq \Phi(t_0, x_0, \overline{u}) \leq \Phi(t_0, x_0, u)$$

It is optimal if ε *can be taken equal to* $+\infty$.

To express necessary conditions for optimality we use the maximum principle in its Hamiltonian form with the *Hamiltonian* H defined by (6).

Proposition 3.2 **([9])** *Assume that $H(t,\cdot,\cdot)$ is differentiable. Then*

$$\frac{\partial H}{\partial p}(t,x,p) = \{f(t,x,u) \mid \langle p, f(t,x,u)\rangle - L(t,x,u) = H(t,x,p)\}$$

and

$$\frac{\partial H}{\partial x}(t,x,p) = \left\{\frac{\partial f}{\partial x}(t,x,u)^{\star}p - \frac{\partial L}{\partial x}(t,x,u) \mid \langle p, f(t,x,u)\rangle - L(t,x,u) = H(t,x,p)\right\}$$

Throughout the paper we will use the following (global) hypothesis concerning the dynamics and the Hamiltonian, although in many theorems below such assumptions are needed only around a reference trajectory.

H$_1$) $\forall\, r > 0,\ \exists\, k_r \in L^1(0,T)$ such that for almost every $t \in [0,T]$,

$$\forall\, u \in U,\ \ (f(t,\cdot,u), L(t,\cdot,u)) \text{ is } k_r(t) - \text{Lipschitz on } B_r(0)$$

H$_2$) The functions $\varphi,\ f(t,\cdot,u),\ L(t,\cdot,u)$ are differentiable for all $u \in U$

H$_3$) For all $(t,x) \in [0,T] \times \mathbf{R}^n$, the set

$$\{(f(t,x,u), L(t,x,u) + r) \mid u \in U,\ r \geq 0\} \text{ is closed and convex}$$

H$_4$) The Lipschitz condition (9) holds true

H$_5$) The Hamiltonian system (8) is complete

H$_6$) The Hamiltonian H is continuous in all variables

H$_7$) The partial derivative $\frac{\partial H}{\partial p}$ is continuous in all variables

From Proposition 3.2 it follows that if the assumptions **H**$_4$) and **H**$_5$) are satisfied and $H(\cdot,0,0)$ is integrable, then there exists at least one trajectory-control pair of (12) such that $t \mapsto L(t,x(t),u(t))$ is integrable. Thus, if in addition L and φ are bounded from below, then $V(t_0,x_0)$ is finite for all (t_0,x_0).

Theorem 3.3 (First Order Necessary Conditions) *Assume* **H**$_1$)–**H**$_3$) *and let $(\overline{x},\overline{u})$ be a weakly locally optimal trajectory-control pair of (P). If $H(t,\cdot,\cdot)$ is differentiable, then there exists $p : [t_0,T] \mapsto \mathbf{R}^n$ such that $(\overline{x},p)$ solves the Hamiltonian system* (5).

The proof uses the ideas similar to the one from [9] but adapted to the weak minima and Proposition 3.2.

Definition 3.4 (Conjugate Point) *Let (x, p) be a solution to the Hamiltonian system* (2) *and P be the solution to the matrix Riccati differential equation (7). A point $t_c \in [0, T]$ is called* conjugate to T along (x, p) *if and only if P is well defined on $]t_c, T]$ and can not be extended (by continuity) on $[t_c, T]$.*

From Proposition 3.2 it follows that, for every solution (x, p) of the Hamiltonian system (2) if there exist two controls u_1, u_2 corresponding to x, then

$$f(s, x(s), u_1(s)) = f(s, x(s), u_2(s)) \quad \& \quad L(s, x(s), u_1(s)) = L(s, x(s), u_2(s)) \quad \text{a.e.}$$

Thus the cost associated to (x, p) does not depend on the choice of the corresponding control.

Theorem 3.5 *Assume* $\mathbf{H}_4)-\mathbf{H}_7)$*, that $\varphi \in C^2$ and $H(t, \cdot, \cdot)$ is twice continuously differentiable. Let $(\overline{x}, \overline{p})$ be a solution to (2) and $\overline{u}$ be a corresponding control. If there is no conjugate to T along $(\overline{x}, \overline{p})$ in the time interval $[t_0, T]$, then $(\overline{x}, \overline{u})$ provides a strong local minimum to the problem (P).*

The proof uses the method of characteristics and the dynamic programming principle associated to (1).

Corollary 3.6 *Assume* $\mathbf{H}_4)-\mathbf{H}_7)$*, that $\varphi \in C^2$ and $H(t, \cdot, \cdot)$ is twice continuously differentiable. Let $(\overline{x}, \overline{p})$ be a solution to (2) and $\overline{u}$ be a corresponding control. If $\varphi''(\overline{x}(T)) \geq 0$ and $\frac{\partial^2 H}{\partial x^2}(t, \overline{x}(t), \overline{p}(t)) \leq 0$ for all $t \in [t_0, T]$, then $(\overline{x}, \overline{u})$ provides a strong local minimum to the problem (P).*

Since a trajectory-control pair providing a strong local minimum is a weak local minimum as well, the sufficient condition can be applied to study weak local minima. We next give a necessary condition for a trajectory-control pair to be a weak local minimum, which (of course) is also necessary for strong local minima.

Theorem 3.7 *Assume* $\mathbf{H}_1)$, $\mathbf{H}_4)$*, that $H(t, x, \cdot)$ is strictly convex and $\frac{\partial^2 H}{\partial p^2}$ is continuous. Further suppose that φ'' is locally Lipschitz and for every*

$r > 0$ *there exists* $l_r \in L^1(0,T)$ *such that for all* $u \in U$ *and almost all* $t \in [0,T]$,

$$\frac{\partial f}{\partial x}(t,\cdot,u),\ \ \frac{\partial L}{\partial x}(t,\cdot,u),\ \ \frac{\partial^2 H}{\partial (x,p)^2}(t,\cdot,\cdot)\ \text{are}\ l_r(t) - \text{Lipschitz}$$

on the ball of center zero and radius r.

Consider a solution (x,p) *to (2) and a corresponding control* $\overline{u}$. *If there exists a conjugate point* $t_c > t_0$, *along* (x,p), *then* $(x,\overline{u})$ *is not weakly locally optimal for the problem (P).*

The proof uses several technical lemmas given below and the Taylor decomposition of the cost functional Φ.

Consider the system

$$\begin{cases} U' &= \dfrac{\partial^2 H}{\partial x \partial p}(t,x(t),p(t))U + \dfrac{\partial^2 H}{\partial p^2}(t,x(t),p(t))V,\ U(T) = Id \\ \\ -V' &= \dfrac{\partial^2 H}{\partial x^2}(t,x(t),p(t))U + \dfrac{\partial^2 H}{\partial p \partial x}(t,x(t),p(t))V,\ V(T) = -\varphi''(x(T)) \end{cases} \tag{13}$$

Then $P(s) = V(s)U(s)^{-1}$ for all $s \in]t_c,T]$ and thus $U(t_c)$ is singular. Fix $w_T \in \mathbf{R}^n$ of norm one such that $U(t_c)w_T = 0$ and let (w,q) be the solution to

$$\begin{cases} w' &= \dfrac{\partial^2 H}{\partial x \partial p}(s,x(s),p(s))w + \dfrac{\partial^2 H}{\partial p^2}(s,x(s),p(s))q,\ w(T) = w_T \\ \\ -q' &= \dfrac{\partial^2 H}{\partial x^2}(s,x(s),p(s))w + \dfrac{\partial^2 H}{\partial p \partial x}(s,x(s),p(s))q,\ q(T) = -\varphi''(x(T))w_T \end{cases} \tag{14}$$

Lemma 3.8 *There exists* $\gamma > 0$ *such that for all* $t < t_c$ *sufficiently close to* t_c

$$\langle q(t), w(t)\rangle \leq -\gamma \|w(t)\|$$

Consider $t_0 \leq t < t_c$ sufficiently close to t_c and denote by (x_h, p_h) the solution to the Hamiltonian system (8) with $t_0 = t$, $x_0 = x(t) + hw(t)$ and $p_0 = p(t) + hq(t)$. From Proposition 3.2 there exists $u_h \in \mathcal{U}$ such that x_h solves the system

$$y' = f(s,y,u_h(s)) \tag{15}$$

and p_h solves the linear system

$$-p' = \frac{\partial f}{\partial x}(s, x_h(s), u_h(s))^\star p - \frac{\partial L}{\partial x}(s, x_h(s), u_h(s)) \tag{16}$$

Denote by $\overline{p}_h$ the solution to (16) satisfying $\overline{p}_h(T) = -\nabla\varphi(x_h(T))$.

Lemma 3.9 *There exists $M_2 \geq 0$ independent from t such that for all small $h > 0$*

$$\|x_h - x - hw\|_\infty + \|p_h - p - hq\|_\infty \leq M_2 h^2 \left(\|w(t)\|^2 + \|q(t)\|^2\right)$$

$$\|x_h' - x' - hw'\|_{L^1(0,T)} + \|p_h - \overline{p}_h\|_\infty \leq M_2 h^2 \left(\|w(t)\|^2 + \|q(t)\|^2\right)$$

Lemma 3.10 *Define*

$$\begin{aligned} I_h := \ & \varphi(x_h(T)) + \textstyle\int_t^T \left(\langle p_h(\tau), x_h'(\tau)\rangle - H(\tau, x_h(\tau), p_h(\tau))\right) d\tau - \\ & -\varphi(x(T)) - \textstyle\int_t^T \left(\langle p(\tau), x'(\tau)\rangle - H(\tau, x(\tau), p(\tau))\right) d\tau \end{aligned}$$

There exists $M_3 > 0$ independent from t such that for all small $h > 0$

$$\left\| I_h + h\langle p(t), w(t)\rangle + \frac{h^2}{2}\langle q(t), w(t)\rangle \right\| \leq M_3 h^3 \left(\|w(t)\|^2 + \|q(t)\|^2\right)$$

4 Smoothness of the Value Function

We shall use the following generalization of the derivative.

Definition 4.1 ([1]) *Consider an extended function $g : \mathbf{R}^n \to \mathbf{R} \cup \{+\infty\}$. The contingent hypoderivative of g at $x_0 \in \mathrm{Dom}(g)$ in the direction $v \in \mathbf{R}^n$ is defined by*

$$D_\downarrow g(x_0)(v) = \limsup_{h \to 0+,\ v' \to v} \frac{g(x_0 + hv') - g(x_0)}{h}$$

The superdifferential of g at x_0 is the closed, convex, possibly empty set

$$\partial_+ g(x_0) = \{p \in \mathbf{R}^n |\ \forall\, v \in \mathbf{R}^n,\quad D_\downarrow g(x_0)(v) \leq \langle p, v\rangle\}$$

For globally optimal solutions we have an extension of Theorem 3.3:

Theorem 4.2 (Costate and Gradients of the Value Function) *Assume* $\mathbf{H}_1) - \mathbf{H}_3)$ *and let* $(\overline{x}, \overline{u})$ *be an optimal trajectory-control pair of* (P). *If* $H(t, \cdot, \cdot)$ *is differentiable, then there exists* $p : [t_0, T] \mapsto \mathbf{R}^n$ *such that* $(\overline{x}, p)$ *solves the Hamiltonian system* (5) *and*

$$\forall\, t \in [t_0, T], \quad -p(t) \in \partial_+ V_x(t, \overline{x}(t))$$

where $\partial_+ V_x(t, x)$ *denotes the superdifferential of* $V(t, \cdot)$ *at* x. *Consequently, if* $V(t_0, \cdot)$ *is differentiable at* x_0, *then the optimal trajectory of* (P) *is unique.*

If in addition V *is locally Lipschitz around* Graph$(\overline{x})$, *then for a.e.* $t \in [t_0, T]$,

$$(H(t, \overline{x}(t), p(t)), -p(t)) \in \partial_+ V(t, \overline{x}(t))$$

The proof is similar to [10], where the Mayer problem was considered. The above theorem and corollary 2.5 imply the following result.

Theorem 4.3 *Assume* $\mathbf{H}_1) - \mathbf{H}_7)$, *that* $\varphi \in C^2$ *and* $H(t, \cdot, \cdot)$ *is twice continuously differentiable. Further assume that for every* $(t_0, x_0) \in [0, T] \times \mathbf{R}^n$ *the problem* (P) *has an optimal solution. If for every solution* $(\overline{x}, \overline{p})$ *of* (2) *there is no conjugate to* T *in the time interval* $[t_0, T]$ *along* $(\overline{x}, \overline{p})$, *then* $V \in C^1([t_0, T] \times \mathbf{R}^n)$, $V(t, \cdot) \in C^2$ *and*

$$\text{Graph}\left(-\frac{\partial V}{\partial x}(t, \cdot)\right) = \{\, (x(t), p(t)) \mid \ (x, p) \ \text{solves} \ (2), \ x_T \in \mathbf{R}^n \,\}$$

Corollary 4.4 *If all the assumptions of Theorem 4.2 hold true and* V *is locally Lipschitz around* Graph$(\overline{x})$, *then for almost all* $t \in [t_0, T]$, $\partial_+ V(t, \overline{x}(t)) \neq \emptyset$ *and*

$$\forall\, (p_t, p_x) \in \partial_+ V(t, \overline{x}(t)), \quad -p_t + H(t, \overline{x}(t), -p_x) = 0$$

The proof proceeds as in [3] where a similar result was proved for the Mayer problem.

Theorem 4.5 (Sufficient Condition for Global Optimality) *Assume* $\mathbf{H}_1) - \mathbf{H}_5)$, *that* $H(\cdot, 0, 0)$ *is integrable,* V *is locally Lipschitz around* $(\overline{t}_0, \overline{x}_0) \in [0, T] \times \mathbf{R}^n$ *and for every* (t_0, x_0) *near* $(\overline{t}_0, \overline{x}_0)$ *the problem* (P) *has an optimal solution. Then for every*

$$\overline{p}_0 \in -\partial_x^\star V(\overline{t}_0, \overline{x}_0) \ := \ -\text{Limsup}_{x \to \overline{x}_0,\, t \to \overline{t}_0} \left\{ \frac{\partial V}{\partial x}(t, x) \right\}$$

the solution (x, p) *to* (8) *with* (t_0, x_0, p_0) *replaced by* $(\overline{t}_0, \overline{x}_0, \overline{p}_0)$ *is so that* x *is optimal for the problem* (P).

To prove this result we use Theorem 4.2 and the fact that the limit of optimal solutions is again an optimal solution.

Remark — Sufficient conditions for local Lipschitz continuity of the value function and for the existence of optimal solutions for (P) can be found in [2,4,8].

Since the value function satisfies the Hamilton-Jacobi equation (1) at points of differentiability Theorems 4.5 and 2.2 yield

Corollary 4.6 (Uniqueness and Regularity) *Under all the assumptions of Theorem 4.5 suppose that (P) has a unique optimal trajectory $z(\cdot)$ for the initial time $\bar{t}_0$ and the initial condition $\bar{x}_0$. Then $V(\bar{t}_0, \cdot)$ is differentiable at $\bar{x}_0$ and the set $\partial_x^\star V(\bar{t}_0, \bar{x}_0)$ is a singleton. Moreover if* $\mathbf{H}_6$) *is satisfied and for every x near $\bar{x}_0$ the set-valued maps $L(\cdot, x, U)$ and $f(\cdot, x, U)$ are upper semicontinuous, then V is differentiable at $(\bar{t}_0, \bar{x}_0)$ and the set $\partial^\star V(\bar{t}_0, \bar{x}_0)$ is a singleton.*

Furtermore we deduce from Theorem 4.5 and the variational equation of ODE the following

Corollary 4.7 (Preservation of Smoothness of Value Function) *Assume* $\mathbf{H}_1$)
$-$ $\mathbf{H}_5$), *that $H(\cdot, 0, 0)$ is integrable, V is locally Lipschitz and for every $(t_0, x_0) \in [0, T] \times \mathbf{R}^n$ the problem (P) has an optimal solution. Let $\bar{x}$ be an optimal trajectory of (P) for the initial time $\bar{t}_0$ and the initial condition $\bar{x}_0$.*

If $V(\bar{t}_0, \cdot)$ is differentiable (resp. twice differentiable) at $\bar{x}_0$, then for all $t \geq t_0$, $V(t, \cdot)$ is differentiable (resp. twice differentiable) at $\bar{x}(t)$. Furthermore, if $V(\bar{t}_0, \cdot)$ is C^1 (resp. C^2) around $\bar{x}_0$, then for all $t \geq t_0$ also $V(t, \cdot)$ is C^1 (resp. C^2) around $\bar{x}(t)$.

When the Hamiltonian H is strictly convex in the last variable, then the sufficient condition of Theorem 4.5 is necessary as well.

Theorem 4.8 *Assume* $\mathbf{H}_1) - \mathbf{H}_6)$, *that V is locally Lipschitz and (P) has an optimal solution for all $(t_0, x_0) \in [0, T] \times \mathbf{R}^n$. Further suppose that $H(t, x, \cdot)$ is strictly convex and for every x the set-valued maps $L(\cdot, x, U)$ and $f(\cdot, x, U)$ are upper semicontinuous. Let $(\bar{x}, \bar{u})$ be a trajectory-control pair of the system (12).*

Then $(\bar{x}, \bar{u})$ is optimal if and only if there exists $p_0 \in -\partial_x^\star V(t_0, x_0)$ such that for the solution (x, p) to the Hamiltonian system (8) we have $x = \bar{x}$.

Proof — The implication $\Longleftarrow$ follows from Theorem 4.5. Assume next that $\overline{x}$ is optimal. By Corollary 4.4 for almost all $t > t_0$, $\partial_+ V(t, \overline{x}(t)) \neq \emptyset$ and

$$\forall\ (p_t, p_x) \in \partial_+ V(t, \overline{x}(t)), \quad -p_t + H(t, \overline{x}(t), -p_x) = 0$$

If $H(t, \overline{x}(t), \cdot)$ is strictly convex, then from the last equality it follows that for almost all $t > t_0$, $\partial_+ V(t, \overline{x}(t))$ is a singleton. By the Hamilton-Jacobi equality satisfied by the value function and $\mathbf{H}_6$), for all (t, x)

$$\forall\ (p_t, p_x) \in \partial^\star V(t, x), \quad -p_t + H(t, x, -p_x) = 0$$

But $\partial_+ V(t, \overline{x}(t)) \subset \overline{co}\ \partial^\star V(t, \overline{x}(t))$ (see for instance [3]). Using again that $H(t, \overline{x}(t), \cdot)$ is strictly convex we get

$$\partial_+ V(t, \overline{x}(t)) \subset \partial^\star V(t, \overline{x}(t))$$

for all $t > t_0$. Consider p as in Theorem 4.2. Thus for almost all $t > t_0$, $-p(t) \in \partial_x^\star V(t, \overline{x}(t))$. To end the proof it is enough to consider a sequence $t_i \to t_0+$. Since $p(t_i) \to p(t_0)$ we obtain that $-p(t_0) \in \partial_x^\star V(t_0, \overline{x}(t_0))$. □

Corollary 4.9 *Under all the assumptions of Theorem 4.8 suppose in addition that $\partial_+ V(t, x) = \overline{co}\, \partial^\star V(t, x)$ for all $(t, x) \in]t_0, T[\times \mathbf{R}^n$. If $\overline{x}$ is an optimal solution to the problem (P), then for all $t \in]t_0, T[$, V is differentiable at $(t, \overline{x}(t))$.*

Remark — The above assumption about superdifferentials of V holds true in particular whenever V is semiconcave. Definition and sufficient conditions for semiconcavity of V (which are just smoothness assumptions on the data) can be found in [4] and for the Mayer problem in [3].

Proof — Since $\partial_+ V(t, x) = \overline{co}\, \partial^\star V(t, x)$ for all $(t, x) \in]t_0, T[\times \mathbf{R}^n$, by the proof of Theorem 4.8 for almost every $t > t_0$, $\overline{co}\ \partial^\star V(t, \overline{x}(t))$ is a singleton. This and Theorem 2.2 imply that V is differentiable at $(t, \overline{x}(t))$ for a.e. $t \in [t_0, T]$. From Corollary 4.7 we deduce that for all $t > t_0$, $V(t, \cdot)$ is differentiable at $\overline{x}(t)$. Hence, by Theorem 4.2, for all $t > t_0$, the restriction of $\overline{x}$ to the time interval $[t, T]$ is the unique optimal trajectory of problem (P) with (t_0, x_0) replaced by $(t, \overline{x}(t))$. Corollary 4.6 ends the proof. □

References

[1] AUBIN J.-P. & FRANKOWSKA H. (1990) SET-VALUED ANALYSIS, Birkhäuser, Boston, Basel, Berlin

[2] BYRNES Ch. & FRANKOWSKA H. (1992) *Unicité des solutions optimales et absence de chocs pour les équations d'Hamilton-Jacobi-Bellman et de Riccati,* Comptes-Rendus de l'Académie des Sciences, t. 315, Série 1, Paris, 427-431

[3] CANNARSA P. & FRANKOWSKA H. (1991) *Some characterizations of optimal trajectories in control theory*, SIAM J. Control and Optimiz., 29, 1322-1347

[4] CAROFF N. (1994) Caractéristiques de l'équation d'Hamilton-Jacobi et conditions d'optimalité en contrôle optimal non linéaire, Thése, Université Paris-Dauphine

[5] CLARKE F.H. (1983) OPTIMIZATION AND NONSMOOTH ANALYSIS, Wiley-Interscience

[6] CONWAY E. D. & HOPF E. (1964) *Hamilton's Theory and Generalized Solutions of the Hamilton-Jacobi Equation*, J. Math. Mech., 13, 939-986

[7] DAFERMOS C.M. (1977) *Generalized Characteristics and the Structure of Solutions of Hyperbolic Conservation Laws*, Indiana University Math. J., 26, 1097-1119

[8] FLEMING W.H. & RISHEL R.W. (1975) DETERMINISTIC AND STOCHASTIC OPTIMAL CONTROL, Springer-Verlag, New York

[9] FRANKOWSKA H. (1989) *Contingent cones to reachable sets of control systems*, SIAM J. on Control and Optimization, 27, 170-198

[10] FRANKOWSKA H. (1989) *Optimal trajectories associated to a solution of contingent Hamilton-Jacobi equations*, Applied Mathematics and Optimization, 19, 291-311

[11] HESTENES M. R. (1966) CALCULUS OF VARIATIONS AND OPTIMAL CONTROL THEORY, Wiley

[12] IOFFE A. D. & TICHOMIROV V. (1979) THEORY OF EXTREMAL PROBLEMS, North-Holland, Amsterdam, New York, Oxford

[13] MAYNE D. Q. (1977) *Sufficient Conditions for a Control to be a Strong Minimum*, J. Opt. Th. Appl., 21, 339-351

[14] ORELL D. & ZEIDAN V. (1988) *Another Jacobi Sufficiency Criterion for Optimal Control with Smooth Constraints*, J. Opt. Th. Appl., 58, 283-300

[15] SMOLLER J. (1980) SHOCK WAVES AND REACTION-DIFFUSION EQUATIONS, Springer-Verlag, Gründlehren der Math. Wiss.

[16] ZEIDAN V. (1984) *Extended Jacobi Sufficiency Criterion for Optimal Control*, SIAM J. Control and Optim., 22, 294-300

[17] ZEIDAN V. (1989) *Sufficiency Conditions with Minimal Regularity Assumptions*, Appl. Math. Opt., 20, 19-31

[18] ZEIDAN V. & ZEZZA P. (1988) *Necessary Conditions for Optimal Control Problems: conjugate points*, SIAM J. Control and Optim., 26, 592-608

[19] ZEIDAN V. & ZEZZA P. (1988) *The Conjugate Point Condition for Smooth Control Sets*, J. Math. An. Appl., 132, 572-589

[20] ZEIDAN V. & ZEZZA P. (1991) *Coupled Points in Optimal Control Theory*, IEEE, Transactions on Automatic Control, 36, 1276-1281