TRANSACTIONS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 348, Number 8, August 1996

# CONJUGATE POINTS AND SHOCKS IN NONLINEAR OPTIMAL CONTROL

N. CAROFF AND H. FRANKOWSKA

Abstract. We investigate characteristics of the Hamilton-Jacobi-Bellman equation arising in nonlinear optimal control and their relationship with weak and strong local minima. This leads to an extension of the Jacobi conjugate points theory to the Bolza control problem. Necessary and sufficient optimality conditions for weak and strong local minima are stated in terms of the existence of a solution to a corresponding matrix Riccati differential equation.

## 1. Introduction

Consider the Hamilton-Jacobi equation

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

It is well known that in general it does not have smooth solutions even when the data are smooth. The classical method of characteristics applied to this equation exhibits shocks, which justify that 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 [26] or the properties of one sided limits [14]). Smooth solutions to (1) were studied (under quite strong assumptions) in [12] using the Bolza problem of the classical calculus of variations. In this paper we shall consider the Hamiltonian $H$ associated to the Bolza problem in optimal control theory. Then, in the same way as [12], the solution to (1) is the value function of the Bolza problem, which may be nonsmooth. Solutions in this case can be described using extensions of the notion of gradient to nonsmooth functions (see [13]). The idea of viscosity solutions consists in taking super and subdifferentials $\partial_+ V$ and $\partial_- V$ respectively instead of the gradient and to use them to define super and subsolutions to (1). A continuous function $V$ which is simultaneously super and subsolution is uniquely defined. When $H(t,x,\cdot)$ is convex, then one can show that a lower semicontinuous function $V : [0,T]\times\mathbf{R}^n \mapsto \mathbf{R}\cup\{+\infty\}$

Received by the editors November 8, 1993 and, in revised form, May 8, 1995.
1991 *Mathematics Subject Classification.* Primary 35B37, 35L67, 49K15, 49L05, 49L20.
*Key words and phrases.* Hamilton-Jacobi-Bellman equation, characteristics, conjugate point, necessary and sufficient conditions for optimality, Riccati differential equation, shock, value function, weak local minimum.

satisfying

$$\forall\ (p_t, p_x) \in \partial_- V(t,x), \quad -p_t + H(t,x,-p_x) = 0, \quad V(T,\cdot) = \varphi(\cdot),$$

$$\liminf_{t\to 0+} V(t,x) = V(0,x), \ \ \liminf_{t\to T-} V(t,x) = \varphi(x)$$

is uniquely defined and is the value function of an associated Mayer optimal control problem (see [4, 17, 18]).

The Hamiltonians which are convex in the last variable arise in optimal control problems. To study characteristics of (1) in the context of optimal control is particularly rewarding because the characteristic system below

$$\text{(2)} \qquad \begin{cases} x' &= \dfrac{\partial H}{\partial p}(t,x,p), \qquad x(T) = x_T, \\ \\ -p' &= \dfrac{\partial H}{\partial x}(t,x,p), \qquad p(T) = -\nabla\varphi(x_T) \end{cases}$$

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? This paper provides an answer based on the use of conjugate point along a solution $(x,p)$ to (2). The conjugate point, as we explain later, is the lower bound of the time interval $]t_c, T]$ on which a solution to the associated Riccati matrix differential equation does exist. Then, for all $t_c < t_0 < T$, $x$ is at least locally optimal on $[t_0,T]$ and it is not even weakly locally optimal on the time interval $[t_0,T]$ for $t_0 < t_c$. This also brings some information about solutions to (1) by eliminating characteristics, which are no longer related to the gradient of $V$. Further analysis of characteristics and their relationships with the value function can be found in [9].

To be more precise consider the Bolza problem arising in optimal control

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

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

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

Under appropriate smoothness hypothesis, it can be shown 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

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

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

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

We would like to underline that, thanks to Proposition 3.2 below, the above system may be rewritten in a more familiar form of the Pontryagin principle involving an adjoint equation and a maximum condition. This fact under several additional assumptions was already observed in [21].

It may happen that even for smooth $f,\ L$ the Hamiltonian $H$ is non differentiable. In Section 2 we provide examples of smooth and nonsmooth Hamiltonians. Since our aim is to link characteristics of (1) to necessary and sufficient conditions for optimality, we shall assume in all results that $H$ is smooth enough around a reference trajectory. It is an interesting question to get an extension of results of Sections 4 and 5 to problems with nonsmooth Hamiltonians.

In general, the system (5) does not have a 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 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

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

When $\nabla H$ 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 [9]). 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 locally Lipschitz, 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 for instance [7]). Consequently, for all $t > t_0,\ p(t) = -\frac{\partial V}{\partial x}(t, \overline{x}(t))$.

The simplest situation occurs in the most investigated linear-convex control problems, i.e., when $f(t, x, u) = Ax + Bu,\ U$ is a subspace and $L, \varphi$ are convex continuously differentiable functions. Then it can be checked that whenever the Hamiltonian $H$ is smooth enough, then (2) has no shock and, consequently, necessary conditions (5) are also sufficient. Furthermore, in this case, the value function is $C^1$ and convex.

This result has only a slight extension to the nonlinear case, when the Hamiltonian is concave with respect to $x$, as it is shown in [6]. The most one may expect is a local result, saying that under appropriate boundedness conditions on the data, the value function $V$ is $C^1$ on $[t, T] \times \mathbf{R}^n$ for some $t < T$ (see [3]). When one investigates nonlinear and nonconvex problems, then the shocks may arise immediately.

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

(7)
$$\begin{cases} P' + \dfrac{\partial^2 H}{\partial p \partial x}(t, x(t), p(t))P \ + \ P \, \dfrac{\partial^2 H}{\partial x \partial p}(t, x(t), p(t)) \\[2ex] + \, P \, \dfrac{\partial^2 H}{\partial p^2}(t, x(t), p(t))P \ + \ \dfrac{\partial^2 H}{\partial x^2}(t, x(t), p(t)) \ = \ 0, \quad P(T) = -\varphi''(x(T)), \end{cases}$$

whose solution $P(\cdot)$ may escape to infinity in a finite time $t < T$. This equation was used in [6] 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 to the Jacobi conjugate point in the calculus of variations [19, 20]. 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

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

We underline that our notion of weak local minimum is different from those used in [22, 23, 33, 34]. We prefer it for several reasons. On one hand the maximum principle in this case is exactly (5), while in the above papers other (localized) necessary conditions, not related to characteristics, are given and convexity of the control set $U$ is often required. On the other hand it allows us to avoid the distinction between "interior" and "boundary" controls (see the "Important Remark" [36, p.1277] and [22]). Also in [33, 36] a different Hamiltonian is considered to state necessary conditions, while here we are interested by characteristics and, consequently, by the Hamiltonian defined by (6). But at the same time, we restrict ourselves to the free end point problems and we have to assume that the Hamiltonian is smooth enough to state our results.

We then show that if $\frac{\partial^2 H}{\partial p^2}$ is positively defined (which is the strengthened Legendre condition adapted to optimal control theory) and if a solution $(\overline{x}, \overline{p})$ of (2) is a weakly locally optimal state-costate pair for the problem (3), (4), then there is no point conjugate to $T$ in $]t_0, T]$ along $(\overline{x}, \overline{p})$. In this way we get the Jacobi type necessary condition for optimality. Furthermore, this condition allows us to conclude, that if a characteristic $(x, p)$ of (1) is so that some $0 < t_c < T$ is conjugate to $T$ along $(x, p)$, then for all $t < t_c$, $-p(t)$ *is not* the gradient of $V(t, \cdot)$ at $x(t)$ and even more

$$-p(t) \notin \mathrm{Limsup}_{x \to x(t)} \ \left\{ \frac{\partial V}{\partial x}(t, x) \right\}.$$

The second result of the same nature states that if a characteristic $(x, p)$ has no conjugate point in the time interval $[t_0, T]$, then $x$ restricted to $[t_0, T]$ provides a strong local minimum to the problem (3), (4). In this way we recover an analogue of the Jacobi sufficient condition. This result may be also deduced from [29, Corollary 4.1] but our proof is quite classical, based on construction of a local solution to (1) via characteristics (see, for instance, [21] for a similar result under a different set of assumptions). Observe that the solution to (7) is always defined on some time interval $[t_0, T]$ with $t_0 < T$. Hence every characteristic $(x, p)$ enjoys the property, that $x$ is strongly locally optimal on some time interval $[t_0, T]$, where $t_0 < T$.

In contrast with the classical calculus of variations, 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 [22, 23, 33, 34], where the interested reader can also get a further bibliography on this subject. Let us finally underline that the above Jacobi type sufficient conditions cannot distinguish between local and global minima. For this very reason further comparison, using the cost, is needed to pick up optimal solutions in the family of locally optimal ones. Some (quite restrictive) sufficient conditions for global minima can be found in [5, 6, 22, 23, 32].

The matrix Riccati equation (7) is related to the value function $V$ by the equality $P(t) = -\frac{\partial^2 V}{\partial x^2}(t, x(t))$ whenever $V(t, \cdot)$ is twice differentiable at $x(t)$. Roughly speaking, the first emergence of a conjugate point corresponds to the first time when $\frac{\partial V}{\partial x}(t, \cdot)$ stops to be locally Lipschitz. 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 [20, 22, 32]). However the global existence of a solution to the Riccati equation here is rather related to the preservation of regularity of the value function along optimal solutions, than with the Jacobi equation. Let us finally mention that for the well investigated in the literature linear quadratic regulator problem the second derivatives of the Hamiltonian in (7) do not depend on $(x(t), p(t))$. For this very reason in this case $P(t)$ does not depend on $x(t)$ and $\frac{\partial^2 V}{\partial x^2}(t, \cdot) = const$. This allows us to get an explicit expression for $V$ using $P$ and to solve as well the optimal synthesis problem in this case. However in general $\frac{\partial^2 V}{\partial x^2}(t, x)$ depends on $x$ and solution to the synthesis problem is not as straightforward.

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 concerns some preliminaries on local minima of the Bolza problem. In Section 4 we prove a sufficient condition for the strong local minimum and in Section 5 a necessary one for the weak local minimum.

## 2. Matrix Riccati Equations and Shocks

In this section we relate 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 $\psi : \mathbf{R}^n \mapsto \mathbf{R}^n$ define the compact set

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

where Limsup denotes the upper set-valued limit (see for instance [2]).

Consider $H : [0,T] \times \mathbf{R}^n \times \mathbf{R}^n \mapsto \mathbf{R}$, $\psi : \mathbf{R}^n \mapsto \mathbf{R}^n$ and the Hamiltonian system

$$\text{(8)} \qquad \begin{cases} x'(t) &= \dfrac{\partial H}{\partial p}(t, x(t), p(t)), \quad x(T) = x_T, \\[2ex] -p'(t) &= \dfrac{\partial H}{\partial x}(t, x(t), p(t)), \quad p(T) = \psi(x_T). \end{cases}$$

In general $H(t,\cdot,\cdot)$ may be non differentiable. Solutions $(x,p) : [t_0,T] \mapsto \mathbf{R}^n \times \mathbf{R}^n$ to (8) are understood in the usual way, assuming in addition that $(x(t), p(t))$ are in the domain of differentiability of $H(t,\cdot,\cdot)$ for all $t \in [t_0,T]$.

*Example* 1. 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

$$H(t,x,p) = \langle p, f(t,x)\rangle + \frac{1}{2}\left\langle R(t,x)^{-1} g(t,x)^\star p, \ g(t,x)^\star p\right\rangle - l(t,x).$$

An appropriate smoothness of $f(t,\cdot)$, $g(t,\cdot)$, $l(t,\cdot)$, $R(t,\cdot)^{-1}$ implies smoothness of $H(t,\cdot,\cdot)$. □

*Example* 2. Consider the same data as in Example 1 and let $B$ denote the closed unit ball in $U$. Define

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

Then it is not difficult to check that

$$H(t,x,p) = \begin{cases} \langle p, f(t,x)\rangle + \frac{1}{2}\|g(t,x)^\star p\|^2 - l(t,x) & \text{if } \ \|g(t,x)^\star p\| \le 1, \\[2ex] \langle p, f(t,x)\rangle + \|g(t,x)^\star p\| - l(t,x) - \frac{1}{2} & \text{otherwise.} \end{cases}$$

So if $f(t,\cdot)$, $g(t,\cdot)$, $l(t,\cdot)$ are smooth enough, then $H(t,\cdot,\cdot)$ is smooth at all $(x,p)$ such that $\|g(t,x)^\star p\| \neq 1$. □

**Definition 2.2.** The system (8) has a *shock* at time $t_0$ if there exist two solutions $(x_i, p_i)(\cdot)$, $i = 1, 2$, of (8) such that

$$x_1(t_0) = x_2(t_0) \quad \text{and} \quad p_1(t_0) \neq p_2(t_0).$$

**Theorem 2.3.** *Assume that $\psi$ is locally Lipschitz on an open set $\Omega$, that for all $x_T \in \Omega$ solutions to (8) are defined on $[0,T]$ and*

$$\text{(9)} \qquad \begin{cases} \forall\, r > 0, \ \exists\, \gamma_r \in L^1(0,T) \ \text{ such that for almost every } \ t \in [0,T], \\[2ex] \dfrac{\partial H}{\partial(x,p)}(t,\cdot,\cdot) \ \text{ is } \ \gamma_r(t)\text{-Lipschitz on} \\[2ex] \{(x(t), p(t)) \mid (x,p) \ \text{ solves } \ (8), \ \ x_T \in \Omega \cap B_r(0)\}. \end{cases}$$

*Define the sets*

$$M_t(\Omega) = \{(x(t), p(t)) \mid (x,p) \ \text{ solves } \ (8), \ \ x_T \in \Omega\}.$$

*We assume moreover that $H$ is measurable with respect to the first variable and for some $\varepsilon > 0$, $H(t,\cdot,\cdot) \in C^2$ on $\varepsilon-$neighborhood of $M_t(\Omega)$ for all $t \in [0,T]$.*

*Then the following two statements are equivalent:*

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

$$\mathcal{D}_t := \{x(t) \mid (x,p) \text{ solves } (8),\ x_T \in \Omega\}$$

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

ii) *$\forall$ $(x,p)$ solving (8) on $[0,T]$ with $x_T \in \Omega$ and all $P_T \in \partial^\star\psi(x_T)$, the matrix Riccati equation*

(10)
$$\left\{\begin{array}{l} P' + \dfrac{\partial^2 H}{\partial p \partial x}(t,x(t),p(t))P \ + \ P\ \dfrac{\partial^2 H}{\partial x \partial p}(t,x(t),p(t)) \\ \\ + \ P\ \dfrac{\partial^2 H}{\partial p^2}(t,x(t),p(t))P \ + \ \dfrac{\partial^2 H}{\partial x^2}(t,x(t),p(t)) \ = \ 0, \quad P(T) = P_T, \end{array}\right.$$

*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 2.4.** *Under all assumptions of Theorem* 2.3*, suppose that $\Omega = \mathbf{R}^n$ and that for every $(x,p)$ solving* (8) *on $[0,T]$ and $P_T \in \partial^\star\psi(x(T))$, the matrix Riccati equation* (10) *has a solution on $[0,T]$. Then the Hamiltonian system* (8) *has no shock in $[0,T]$.*

To prove the above theorem, the following technical result is needed.

**Lemma 2.5.** *Under all the assumptions of Theorem* 2.3 *consider a compact $K \subset \Omega$ and the subsets $M_t(K)$, $t \in [0,T]$ defined by*

$$M_t(K) = \{(x(t),p(t)) \mid (x,p) \text{ solves } (8),\ x_T \in K\}.$$

*Then there exists $\delta > 0$ such that for all $t \in [T-\delta,T]$, $M_t(K)$ is the graph of a Lipschitz function.*

*Proof.* Let $r > 0$ be such that $K \subset B_r(0)$. Set $k = \gamma_r$. We proceed by a contradiction argument. Assume for a moment that there exist $t_i \to T-$ such that $M_{t_i}(K)$ is not the graph of a Lipschitz function. Then for every $i$ we can find two distinct solutions $(x_j^i, p_j^i)$, $j = 1,2$, of the Hamiltonian system (8) such that $x_j^i(T) \in K, j = 1,2$, and

$$\varepsilon_i \ := \ \frac{\|x_1^i(t_i) - x_2^i(t_i)\|}{\|p_1^i(t_i) - p_2^i(t_i)\|} \ \to \ 0 \ \text{ as } \ i \to +\infty.$$

Since for every $s \in [t_i,T]$ we have

$$\begin{aligned} &\left\|x_1^i(s) - x_2^i(s)\right\| \\ &\qquad \le \varepsilon_i\|p_1^i(t_i) - p_2^i(t_i)\| + \textstyle\int_{t_i}^s k(\tau)(\left\|x_1^i(\tau) - x_2^i(\tau)\right\| + \left\|p_1^i(\tau) - p_2^i(\tau)\right\|)d\tau \end{aligned}$$

the Gronwall lemma implies that for some $C > 0$ independent from $i$ and for all $s \in [t_i, T]$

$$\left\|x_1^i(s) - x_2^i(s)\right\| \le C(\varepsilon_i \|p_1^i(t_i) - p_2^i(t_i)\| + \int_{t_i}^s k(\tau)\|p_1^i(\tau) - p_2^i(\tau)\|d\tau).$$

Hence for some $C_1 > 0$ and all $i$ large enough and $s \in [t_i, T]$,

$$\begin{aligned}
&\left\|p_1^i(s) - p_2^i(s)\right\| \\
&\quad \le \left\|p_1^i(t_i) - p_2^i(t_i)\right\| + \textstyle\int_{t_i}^s k(\tau)(\left\|x_1^i(\tau) - x_2^i(\tau)\right\| + \left\|p_1^i(\tau) - p_2^i(\tau)\right\|)d\tau \\
&\quad \le C_1 \left\|p_1^i(t_i) - p_2^i(t_i)\right\| + C_1 \textstyle\int_{t_i}^s k(\tau) \left\|p_1^i(\tau) - p_2^i(\tau)\right\| d\tau.
\end{aligned}$$

From the Gronwall lemma we deduce that for some $L > 0$ independent from $i$ and all $s \in [t_i, T]$,

$$\left\|p_1^i(s) - p_2^i(s)\right\| \le L \left\|p_1^i(t_i) - p_2^i(t_i)\right\|.$$

This implies that

$$\overline{\varepsilon}_i := \sup_{s \in [t_i, T]} \frac{\left\|x_1^i(s) - x_2^i(s)\right\|}{\left\|p_1^i(t_i) - p_2^i(t_i)\right\|} \text{ converge to zero.} \tag{11}$$

We next observe that for all $s \in [t_i, T]$,

$$\begin{aligned}
&\left\|p_1^i(s) - p_2^i(s)\right\| \\
&\quad \le \left\|p_1^i(T) - p_2^i(T)\right\| + \textstyle\int_s^T k(\tau)(\left\|x_1^i(\tau) - x_2^i(\tau)\right\| + \left\|p_1^i(\tau) - p_2^i(\tau)\right\|)d\tau \\
&\quad \le \left\|p_1^i(T) - p_2^i(T)\right\| + \textstyle\int_s^T k(\tau)(\left\|p_1^i(\tau) - p_2^i(\tau)\right\| + \overline{\varepsilon}_i \left\|p_1^i(t_i) - p_2^i(t_i)\right\|)d\tau.
\end{aligned}$$

Applying again the Gronwall lemma and taking $i$ large enough we get

$$\left\|p_1^i(t_i) - p_2^i(t_i)\right\| \le L_1 \left\|p_1^i(T) - p_2^i(T)\right\| + \frac{1}{2} \left\|p_1^i(t_i) - p_2^i(t_i)\right\|$$

for some $L_1$ independent from $i$. Hence for all large $i$

$$\left\|p_1^i(t_i) - p_2^i(t_i)\right\| \le 2L_1 \left\|p_1^i(T) - p_2^i(T)\right\|$$

and therefore, by (11),

$$\frac{\left\|x_1^i(T) - x_2^i(T)\right\|}{\left\|p_1^i(T) - p_2^i(T)\right\|} = \frac{\left\|x_1^i(T) - x_2^i(T)\right\|}{\left\|p_1^i(t_i) - p_2^i(t_i)\right\|} \times \frac{\left\|p_1^i(t_i) - p_2^i(t_i)\right\|}{\left\|p_1^i(T) - p_2^i(T)\right\|} \to 0.$$

Therefore

$$\frac{\|\psi(x_1^i(T)) - \psi(x_2^i(T))\|}{\left\|x_1^i(T) - x_2^i(T)\right\|} = \frac{\left\|p_1^i(T) - p_2^i(T)\right\|}{\left\|x_1^i(T) - x_2^i(T)\right\|} \to +\infty$$

which contradicts the Lipschitz continuity of $\psi$ on $K$. □ □

*Proof of Theorem* 2.3. Assume first that for all $t \in [0, T]$, $M_t(\Omega)$ is the graph of a locally Lipschitz function. Consider $x_T \in \Omega$, the solution $(x, p)$ of (8) and the linear system

(12)
$$\left\{\begin{array}{rcl} U' & = & \dfrac{\partial^2 H}{\partial x \partial p}(t,x(t),p(t))U \;+\; \dfrac{\partial^2 H}{\partial p^2}(t,x(t),p(t))V, \quad U(T)=Id, \\[2ex] -V' & = & \dfrac{\partial^2 H}{\partial x^2}(t,x(t),p(t))U \;+\; \dfrac{\partial^2 H}{\partial p \partial x}(t,x(t),p(t))V, \quad V(T)=P_T, \end{array}\right.$$

where $U,\ V : [0,T] \mapsto L(\mathbf{R}^n, \mathbf{R}^n)$ are matrix functions and $P_T \in \partial^\star \psi(x(T))$. Let $(x_n, p_n)$ be solutions to (8) such that

$$\lim_{n\to\infty} x_n(T) = x(T) \quad \text{and} \quad \lim_{n\to\infty} \psi'(x_n(T)) = P_T.$$

By our assumptions, $(x_n, p_n)$ converge uniformly to $(x,p)$.

The variational equation implies that for any $(w(\cdot), q(\cdot))$ solving

$$\text{(13)} \qquad \left\{\begin{array}{rl} w' & = \dfrac{\partial^2 H}{\partial x \partial p}(t,x_n(t),p_n(t))w \;+\; \dfrac{\partial^2 H}{\partial p^2}(t,x_n(t),p_n(t))q, \\[2ex] -q' & = \dfrac{\partial^2 H}{\partial x^2}(t,x_n(t),p_n(t))w \;+\; \dfrac{\partial^2 H}{\partial p \partial x}(t,x_n(t),p_n(t))q \end{array}\right.$$

and satisfying $w(T) = w_T,\ \ q(T) = \psi'(x_n(T))w_T$ we have

$$(w(t), q(t)) \in T_{M_t(\Omega)}(x_n(t), p_n(t))$$

(contingent cone to $M_t(\Omega)$ at $(x_n(t), p_n(t))$. See for instance [2, Chapter 4] for the definition of contingent cones). Because $M_t(\Omega)$ is the graph of a locally Lipschitz function, for some $l_t$ independent from $n$, $\|q(t)\| \le l_t \|w(t)\|$. Taking the limit in (13) we deduce that the solution $(w,q)$ of (13) with $(x_n, p_n)$ replaced by $(x,p)$ and $w(T) = w_T,\ q(T) = P_T w_T$ satisfies $\|q(t)\| \le l_t \|w(t)\|$. Thus, by uniqueness of solution, if $w_T \ne 0$, then $w(\cdot)$ never vanishes. Since $w(t) = U(t)w_T$ and $q(t) = V(t)w_T$ this implies that $U(t)$ is not singular for all $t \in [0,T]$. Setting $P(t) = V(t)U(t)^{-1}$, we check that $P$ solves (10).

Conversely let (10) have a solution on $[0,T]$ for all $(x,p)$ satisfying (8) with $x_T \in \Omega$ and let $P_T \in \partial^\star\psi(x_T)$. Consider a family of bounded open sets $\Omega_1 \subset \Omega_2 \subset \ldots \subset \Omega$ such that $\overline{\Omega}_i \subset \Omega_{i+1}$ and $\bigcup_{i\ge 1} \Omega_i = \Omega$. For every $i \ge 1,\ t \in [0,T]$ we introduce the compact sets

$$\Pi_t^i = \left\{(x(t),p(t)) \mid (x,p) \text{ solves (8)},\ x(T) \in \overline{\Omega}_i\right\}.$$

We claim that for every $i$ and $t_0$, $\Pi_{t_0}^i$ is the graph of a Lipschitz function. Indeed fix $i, t_0$ and assume for a moment that $\Pi_{t_0}^i$ is not the graph of a Lipschitz function. By Lemma 2.5 for all $s$ near $T$, $\Pi_s^i$ is still the graph of a Lipschitz function. Define

$$\bar t \;=\; \inf_{t\in[0,T]} \left\{\ \forall\, s \in [t,T],\ \ \Pi_s^i \ \text{ is the graph of a Lipschitz function}\right\}.$$

Then $\bar t < T$ and $\Pi_{\bar t}^i$ is not the graph of a Lipschitz function, because otherwise, using Lemma 2.5, we could make $\bar t$ smaller which would contradict its choice. Define the bounded sets

$$D_s^i = \{x(s) \mid (x,p) \text{ solves (8)},\ x_T \in \Omega_i\}.$$

Then for all $i \ge 1$ and $s \in ]\bar t, T]$, $D_s^i$ is open. Indeed consider $s > \bar t$ and define the map $\Phi$ from the open set $\Omega_i$ into $\mathbf{R}^n$ by $\Phi(x_T) = x(s)$, where $(x,p)$ solves (8). Since $\Pi_s^i$ is the graph of a Lipschitz function, $\Phi$ is one-to-one. The function $\Phi$

being continuous, by the Invariance of Domain Theorem (see for instance [27]), $D_s^i$ is open. Its closure is equal to the set

$$\overline{D_s^i} = \left\{ x(s) \mid (x,p) \text{ solves } (8),\ x(T) \in \overline{\Omega}_i \right\}.$$

Define next the Lipschitz function $\Phi_s^i : \overline{D_s^i} \mapsto \mathbf{R}^n$ by $\mathrm{Graph}(\Phi_s^i) = \Pi_s^i$. By the Rademacher theorem $\Phi_s^i$ is differentiable almost everywhere in $D_s^i$.

Fix a sequence $t_n \to \overline{t}+$ and observe that the family $\left\{\Phi_{t_n}^i\right\}_{n\geq 1}$ cannot be equilipschitz, because otherwise, using that $\Pi_{\overline{t}}^i = \mathrm{Lim}_{n\to\infty}\Pi_{t_n}^i$, (set-valued limit, see [2]), we would deduce that $\Pi_{\overline{t}}^i$ is the graph of a Lipschitz function. Thus there exists a sequence $\overline{x}_n \in D_{t_n}^i$ such that $\left\|(\Phi_{t_n}^i)'(\overline{x}_n)\right\| \to \infty$, i.e.

$$\exists\, (u_n, v_n) \in \mathbf{R}^n \times \mathbf{R}^n \text{ satisfying } (\Phi_{t_n}^i)'(\overline{x}_n)u_n = v_n,\ \|v_n\| = 1,\ u_n \to 0.$$

Let $(x_n, p_n)$ be a solution to (8) such that $x_n(t_n) = \overline{x}_n$ and $p_n(t_n) = \Phi_{t_n}^i(\overline{x}_n)$. Since $\Phi_{t_n}^i$ is differentiable at $\overline{x}_n$, using variational equation, we deduce that $\psi$ is differentiable at $x_n(T)$. Taking a subsequence and keeping the same notations, we may assume that $(x_n, p_n)$ converge uniformly to a solution $(x,p)$ of (8), $v_n \to v$ and $\psi'(x_n(T)) \to P_T \in \partial^\star\psi(x(T))$. Consider next the solutions $(w_n, q_n)$ to (13) such that $w_n(t_n) = u_n$, $q_n(t_n) = v_n$. The variational equation yields $q_n(T) = \psi'(x_n(T))w_n(T)$. Passing to the limit we deduce that the Hamiltonian system

$$\left\{\begin{array}{rcl} w' &=& \dfrac{\partial^2 H}{\partial x \partial p}(s, x(s), p(s))w \;+\; \dfrac{\partial^2 H}{\partial p^2}(s, x(s), p(s))q, \\[2ex] -q' &=& \dfrac{\partial^2 H}{\partial x^2}(s, x(s), p(s))w \;+\; \dfrac{\partial^2 H}{\partial p \partial x}(s, x(s), p(s))q \end{array}\right. \tag{14}$$

has a solution $(w,q)$ satisfying $w(\overline{t}) = 0$, $q(\overline{t}) \neq 0$, $q(T) = P_T w(T)$. In particular $w(T) \neq 0$ and $U(\overline{t})w(T) = 0$. On the other hand, by the first part of the proof, $P(t) = V(t)U(t)^{-1}$ solves (10) on $]\overline{t}, T]$. If $P$ is well defined on $[\overline{t}, T]$, then $V(\overline{t}) = P(\overline{t})U(\overline{t})$ and $q(\overline{t}) = V(\overline{t})w(T) = 0$, which leads to a contradiction and proves our claim.

Define next the open sets $\mathcal{D}_t = \bigcup_{i\geq 1} D_t^i$, where $t \in [0,T]$. Then

$$\mathcal{D}_t = \{x(t) \mid \exists\, p \text{ such that } (x,p) \text{ solves } (8),\ x_T \in \Omega\} \text{ and } M_t(\Omega) = \bigcup_{i\geq 1} \Pi_t^i.$$

Since $\{\Pi_t^i\}_{i\geq 1}$ is a nondecreasing family of graphs of Lipschitz functions, $M_t(\Omega)$ is the graph of a function from $\mathcal{D}_t$ into $\mathbf{R}^n$. To check that it is the graph of a locally Lipschitz function, fix $\overline{x} \in \mathcal{D}_t$, $\rho > 0$ such that $B_\rho(\overline{x}) \subset \mathcal{D}_t$. Since $B_\rho(\overline{x})$ is compact and the family of open sets $\{D_t^i\}_{i\geq 1}$ is nondecreasing, for some $i \geq 1$, $B_\rho(\overline{x}) \subset D_t^i$. But we already know that $\Pi_t^i$ is the graph of a Lipschitz function. The last two statements of theorem follow from the variational equation. □

## 3. Bolza Optimal Control Problem

Consider the Bolza minimization problem

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

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

$$x'(t) = f(t, x(t), u(t)), \quad x(t_0) = x_0, \quad u(t) \in U, \tag{15}$$

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

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

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 (15) 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 (15) is called *weakly locally optimal* for the problem $(P)$ if there exists $\varepsilon > 0$ such that for every trajectory-control pair $(x, u)$ of (15)

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

It is called *strongly locally optimal* if there exists $\varepsilon > 0$ such that for every trajectory-control pair $(x, u)$ of (15)

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

It is optimal if $\varepsilon$ 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.** *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*

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

The proof is comparable to the one given in [16] and is omitted.

*Remark* 1. In general even for smooth data $H(t, \cdot, \cdot)$ is merely locally Lipschitz and $H(t, x, \cdot)$ is convex. In particular, $H(t, x, \cdot)$ is differentiable at some $p$ if and only if the set

$$A(t, x) := \{ u \in U \mid \langle p, f(t, x, u) \rangle - L(t, x, u) = H(t, x, p) \}$$

is so that $f(t, x, A(t, x))$ is a singleton. In Example 1 of Section 2 the Hamiltonian is smooth. In Example 2 it is smooth outside of a set described in this example. $\square$

Throughout the paper we will have call for the following (global) hypothesis concerning the dynamics, although in all the results below such assumptions are needed only around a reference trajectory.

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

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

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

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

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

**Theorem 3.3** (First Order Necessary Conditions)**.** *Assume* $\mathbf{H}_1)-\mathbf{H}_3)$ *and let* $(\overline{x},\overline{u})$ *be a weakly locally optimal trajectory-control pair of* $(P)$*. Let* $p(\cdot)$ *be the solution to*

(16)
$$-p' = \frac{\partial f}{\partial x}(t,\overline{x}(t),\overline{u}(t))^\star p - \frac{\partial L}{\partial x}(t,\overline{x}(t),\overline{u}(t)), \quad p(T) = -\nabla\varphi(\overline{x}(T)).$$

*If* $H(t,\cdot,\cdot)$ *is differentiable at* $(\overline{x}(t),p(t))$ *for all* $t\in[t_0,T]$*, then* $(\overline{x},p)$ *solves the Hamiltonian system* (5) *on* $[t_0,T]$*.*

*Proof.* We introduce the set-valued map $F:[0,T]\times\mathbf{R}^{n+1}\rightsquigarrow\mathbf{R}^{n+1}$ by

$$F(t,(x,x^0)) = \{(f(t,x,u), L(t,x,u)+r)\,|\,u\in U,\ r\ge 0\}\,.$$

It is $k_r(t)$-Lipschitz on $B_r(0)$ for all $r>0$ and, by $\mathbf{H}_3$), it has closed convex images. Furthermore, the mapping

$$\overline{y}(t) := \left(\overline{x}(t),\ \overline{x}^0(t) := \int_{t_0}^t L(s,\overline{x}(s),\overline{u}(s))ds\right)$$

provides the weak local minimum to the Mayer problem

$$\text{minimize } \ \varphi(x(T)) + x^0(T) \tag{17}$$

over absolutely continuous solutions $y(\cdot)=(x,x^0)(\cdot)$ of the differential inclusion

$$y'\in F(t,y),\ \ y(t_0)=(x_0,0), \tag{18}$$

i.e. for some $\varepsilon>0$ and all $y(\cdot)=(x,x^0(\cdot))$ solving (18) with $\|y'-\overline{y}'\|_{L^1(t_0,T)}\le\varepsilon$

$$\varphi(x(T)) + x^0(T) \ge \varphi(\overline{x}(T)) + \overline{x}^0(T).$$

Set

$$V(t) = \bigcup_{u\in U,\ r\ge 0} \{(f(t,\overline{x}(t),u)-\overline{x}'(t),\ L(t,\overline{x}(t),u)-L(t,\overline{x}(t),\overline{u}(t))+r)\}$$

and consider the linearized control system

(19)
$$\begin{cases} (w,w^0)' = \left(\dfrac{\partial f}{\partial x}(t,\overline{x}(t),\overline{u}(t))w, \dfrac{\partial L}{\partial x}(t,\overline{x}(t),\overline{u}(t))w^0\right) + v(t),\ v(t)\in V(t), \\ (w(t_0),w^0(t_0)) = 0. \end{cases}$$

By [2, p.193] we know that for almost all $t$ and all $(w,w^0)\in\mathbf{R}^{n+1}$,

$$\left(\frac{\partial f}{\partial x}(t,\overline{x}(t),\overline{u}(t))w,\ \frac{\partial L}{\partial x}(t,\overline{x}(t),\overline{u}(t))w^0\right) + V(t) \subset D^\flat_y F(t;\overline{y}(t),\overline{y}'(t))(w,w^0)$$

where $D^\flat_y F(t;x,y)(\cdot)$ denotes the *adjacent derivative* of $F(t,\cdot)$ at $(x,y) \in \mathrm{Graph}(F(t,\cdot))$ (see [2, p.189] for the definition and properties).

Let $(w,w_0)$ be a solution to (19). By the variational inclusion from [2, p.404] there exist solutions $(x_h,x_h^0)$ to (18) such that

$$\frac{(x_h,x_h^0)'-\overline{y}'}{h} \to (w',w_0') \text{ in } L^1(t_0,T) \text{ when } h\to 0+.$$

Since $\overline{y}$ is weakly locally optimal for the problem (17), (18) we deduce that

$$\langle \nabla\varphi(\overline{x}(T)), w(T)\rangle + w^0(T) \geq 0. \tag{20}$$

Denote by $Y$ the fundamental solution to

$$Y' = \begin{pmatrix} \frac{\partial f}{\partial x}(t,\overline{x}(t),\overline{u}(t)) & 0 \\ \\ \frac{\partial L}{\partial x}(t,\overline{x}(t),\overline{u}(t)) & 0 \end{pmatrix} Y, \quad Y(t_0)=Id.$$

Then for every integrable selection $v(t)\in V(t)$ the solution $(w,w^0)(\cdot)$ to (19) satisfies $(w,w^0)(T)=\int_{t_0}^T Y(T)Y(t)^{-1}v(t)dt$. Consequently, by (20), for every integrable selection $v(t)\in V(t)$,

$$\int_{t_0}^T \left\langle Y(t)^{\star^{-1}}Y(T)^\star\left(\nabla\varphi(\overline{x}(T)),1\right), v(t)\right\rangle dt \geq 0.$$

Setting $(p(t),p^0(t)) = -Y(t)^{\star^{-1}}Y(T)^\star\left(\nabla\varphi(\overline{x}(T)),1\right)$, we check that

$$p^0(t)\equiv -1, \;\; \langle p(t), f(t,\overline{x}(t),\overline{u}(t))\rangle - L(t,\overline{x}(t),\overline{u}(t)) = H(t,\overline{x}(t),p(t))$$

for almost all $t\in[t_0,T]$ and that $p$ solves the system (16). Proposition 3.2 ends the proof. $\square$

## 4. Jacobi Sufficient Condition for Strong Local Minimum

From now on we restrict our attention to the Hamiltonian defined in Section 3. We first introduce the notion of conjugate point.

**Definition 4.1.** 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

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

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

**Theorem 4.2.** *Let $(\overline{x},\overline{p})$ be a solution to (2) defined on $[t_0,T]$ and $\overline{u}$ be a corresponding control. We assume that $H$ is continuous, $\varphi\in C^2$, that for some $\varepsilon>0,\ k\in L^1(t_0,T)$,*

i) $\frac{\partial H}{\partial p}$ *is continuous on* $\{t\times B_\varepsilon(\overline{x}(t),\overline{p}(t)) \mid t\in[t_0,T]\}$,

ii) $\frac{\partial H}{\partial(x,p)}(t,\cdot,\cdot)$ *is $k(t)$-Lipschitz and $H(t,\cdot,\cdot)\in C^2$ on $B_\varepsilon(\overline{x}(t),\overline{p}(t))$.*

*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 $(P)$.*

*Proof.* We set $\psi = -\nabla\varphi$ in Theorem 2.3. By our assumptions and the proof of Theorem 2.3, there exists $\Omega \subset \mathbf{R}^n$ such that for every $x_T \in \Omega$ the solution $(x, p)$ to the Hamiltonian system (2) is defined at least on $[t_0, T]$ and there is no conjugate point along $(x, p)$ in the time interval $[t_0, T]$. We may choose $\Omega$ in such way that the sets $M_t(\Omega)$ defined as in Theorem 2.3 verify $M_t(\Omega) \subset B_{\frac{\varepsilon}{2}}(\overline{x}(t), \overline{p}(t))$ for all $t \in [t_0, T]$. By Theorem 2.3, $M_t(\Omega)$ is the graph of a $C^1$ function from an open set $\mathcal{D}_t$ into $\mathbf{R}^n$. Let $(x_0, p_0) \in M_t(\Omega)$ and $(x, p)$ be the corresponding solution to (2). By assumptions of theorem, $x \in C^1$. Define $W : \{t \times \mathcal{D}_t \mid t \in [t_0, T]\} \mapsto \mathbf{R}$ by

$$W(t, x_0) = \varphi(x(T)) + \int_t^T (\langle x'(s), p(s)\rangle - H(s, x(s), p(s)))\, ds.$$

We first check that $-p_0$ is the derivative (with respect to $x$) of $W(t, \cdot)$ at $x_0$ and then show that $W \in C^1$ on $\{t \times \mathcal{D}_t \mid t \in [t_0, T]\}$, and

$$-\frac{\partial W}{\partial t}(t, x) + H\left(t, x, -\frac{\partial W}{\partial x}(t, x)\right) = 0 \text{ on } \{t \times \mathcal{D}_t \mid t \in [t_0, T]\}. \tag{21}$$

Let $w_0 \in \mathbf{R}^n$ and $(w, q)$ be the solution to the system (14) such that $w(t) = w_0,\ q(t) = P(t)w_0$. By the variational equation and the definition of $W$, the directional derivative of $W(t, \cdot)$ in the direction $w_0$ satisfies

$$\begin{aligned}
\frac{\partial W}{\partial w_0}(t, x_0) &= \langle \nabla\varphi(x(T)), w(T)\rangle \\
&\quad + \int_t^T \Bigg( \langle w'(s), p(s)\rangle + \langle x'(s), q(s)\rangle \\
&\qquad - \left\langle \frac{\partial H}{\partial x}(s, x(s), p(s)), w(s)\right\rangle - \left\langle \frac{\partial H}{\partial p}(s, x(s), p(s)), q(s)\right\rangle \Bigg) ds \\
&= \langle -p(T), w(T)\rangle + \int_t^T (\langle w'(s), p(s)\rangle + \langle x'(s), q(s)\rangle \\
&\qquad + \langle p'(s), w(s)\rangle - \langle x'(s), q(s)\rangle)\, ds \\
&= \langle -p(T), w(T)\rangle + \langle w(T), p(T)\rangle - \langle p(t), w_0\rangle = \langle -p(t), w_0\rangle = \langle -p_0, w_0\rangle.
\end{aligned}$$

Thus, the Gâteaux derivative of $W(t, \cdot)$ at $x_0$ is equal to $-p_0$. Using that for all $x_0 \in \mathcal{D}_t,\ (x_0, p_0) \in M_t(\Omega)$ and $M_t(\Omega)$ is the graph of a $C^1-$function, we obtain that the partial derivative $\frac{\partial W}{\partial x}(t, x_0) = -p_0$ and $W(t, \cdot) \in C^2$ on $\mathcal{D}_t$. Furthermore,

$$\lim_{\substack{h \to 0,\ x \to x_0, \\ t+h \in [t_0, T],\ x \in \mathcal{D}_{t+h}}} \frac{\partial W}{\partial x}(t + h, x) = \frac{\partial W}{\partial x}(t, x_0). \tag{22}$$

Let $h > 0$ be so small that the interval $[x(t), x(t+h)] \subset \mathcal{D}_{t+h}$. By the mean value theorem, for some $\Theta_h \in [0,1]$,

$$\begin{aligned}
&W(t+h, x_0) - W(t, x_0)\\
&\quad = W(t+h, x_0) - W(t+h, x(t+h))\\
&\qquad - \int_t^{t+h} (\langle p(s), x'(s)\rangle - H(s, x(s), p(s)))\, ds\\
&\quad = -\left\langle \frac{\partial W}{\partial x}(t+h, x_0) + \Theta_h(x(t+h) - x_0), x(t+h) - x_0 \right\rangle\\
&\qquad - \int_t^{t+h} (\langle p(s), x'(s)\rangle - H(s, x(s), p(s)))\, ds
\end{aligned}$$

Dividing by $h \neq 0$ and taking the limit, using (22), we get

$$\begin{aligned}
\frac{\partial W}{\partial t}(t, x_0) &= \left\langle -\frac{\partial W}{\partial x}(t, x_0), x'(t) \right\rangle - \langle p_0, x'(t)\rangle + H(t, x_0, p_0)\\
&= H\left(t, x_0, -\frac{\partial W}{\partial x}(t, x_0)\right).
\end{aligned}$$

This implies (21) and, by (22), that $W \in C^1$ on $\{t \times \mathcal{D}_t \mid t \in [t_0, T]\}$.

Let $\varepsilon > 0$ be such that for every $t \in [t_0, T]$, $\overline{x}(t) + \varepsilon B \subset \mathcal{D}_t$. Consider any solution-control pair $(y(\cdot), u(\cdot))$ to the control system (15) such that $\|\overline{x} - y\|_\infty < \varepsilon$. Then $t \mapsto \psi(t) := W(t, y(t))$ is absolutely continuous and for every $t \in [t_0, T]$ such that $y'(t) = f(t, y(t), u(t))$ we have

$$\begin{aligned}
\psi'(t) &= \frac{\partial W}{\partial t}(t, y(t)) + \left\langle \frac{\partial W}{\partial x}(t, y(t)), y'(t) \right\rangle\\
&= H\left(t, y(t), -\frac{\partial W}{\partial x}(t, y(t))\right) + \left\langle \frac{\partial W}{\partial x}(t, y(t)), y'(t) \right\rangle \geq -L(t, y(t), u(t)).
\end{aligned}$$

Hence

$$\psi(T) \geq \psi(t_0) - \int_{t_0}^T L(t, y(t), u(t))dt$$

and therefore, using Proposition 3.2, we obtain

$$\varphi(y(T)) + \int_{t_0}^T L(t, y(t), u(t))dt \geq W(t_0, x_0) = \varphi(\overline{x}(T)) + \int_{t_0}^T L(t, \overline{x}(t), \overline{u}(t))dt. \qquad \square$$

The above result and [25] yield

**Corollary 4.3.** *Under all assumptions of Theorem* 4.2*, if*

$$\varphi''(\overline{x}(T)) \geq 0 \ \text{ and } \ \frac{\partial^2 H}{\partial x^2}(t, \overline{x}(t), \overline{p}(t)) \leq 0$$

*for all* $t \in [t_0, T]$*, then* $(\overline{x}, \overline{u})$ *provides a strong local minimum to the problem* $(P)$.

## 5. Jacobi Necessary Condition for Weak Local Minimum

Since a trajectory-control pair providing a strong local minimum is a weak local minimum as well, the sufficient condition of Section 4 can be applied to study weak local minima. In this section we 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 5.1.** *Consider a solution $(x,p)$ to (2) defined on $[t_0,T]$ and a corresponding control $\overline{u}$. Assume $\mathbf{H}_1$), that $\varphi''$ is locally Lipschitz, that for some $\varepsilon > 0$, $k \in L^1(t_0,T)$, for all $u \in U$ and almost all $t \in [t_0,T]$,*

$$\frac{\partial f}{\partial x}(t,\cdot,u) \text{ and } \frac{\partial L}{\partial x}(t,\cdot,u) \text{ are } k(t)\text{-Lipschitz on } B_\varepsilon(x(t))$$

*and*

$$\frac{\partial H}{\partial(x,p)}(t,\cdot,\cdot) \text{ and } \frac{\partial^2 H}{\partial(x,p)^2}(t,\cdot,\cdot) \text{ are } k(t)\text{-Lipschitz on } B_\varepsilon(x(t),p(t)).$$

*If the conjugate point $t_c > t_0$ (along $(x,p)$) and for some $\lambda > 0$, $\frac{\partial^2 H}{\partial p^2}(t,x(t),p(t)) \geq \lambda$ for all $t < t_c$ near $t_c$, $\sup_{t\in[t_c-\lambda,t_c]} \left\| \frac{\partial^2 H}{\partial(x,p)^2}(t,x(t),p(t)) \right\| < \infty$, then $(x,\overline{u})$ is not weakly locally optimal for the problem $(P)$.*

The proof is preceded by several lemmas.

Consider the system (12) where $P_T = -\varphi''(x(T))$. From the proof of Theorem 2.3 we know that $P(s) = V(s)U(s)^{-1}$ 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 (14) with $w(T) = w_T,\ q(T) = -\varphi''(x(T))w_T$.

**Lemma 5.2.** *There exists $\gamma > 0$ such that for all $t < t_c$ sufficiently close to $t_c$*

$$\langle q(t), w(t)\rangle \leq -\gamma \|w(t)\| \quad \text{and} \quad \|w(t)\| \geq \gamma(t_c - t)$$

*Proof.* Since $w(s) = U(s)w_T$ we have $w(t_c) = 0$ and $q(t_c) \neq 0$. Multiplying the first equation in (14) by $q$ and the second one by $-w$ and adding them yields

$$\begin{aligned}\langle q(t), w(t)\rangle &= \int_{t_c}^t \left(\langle w'(s), q(s)\rangle + \langle q'(s), w(s)\rangle\right) ds \\ &= \int_{t_c}^t \left( \left\langle \frac{\partial^2 H}{\partial p^2}(s,x(s),p(s))q(s), q(s) \right\rangle - \left\langle \frac{\partial^2 H}{\partial x^2}(s,x(s),p(s))w(s), w(s) \right\rangle \right) ds\end{aligned}$$

From this and our assumptions we deduce that for some $\gamma, \rho > 0$ and all $t < t_c$ near $t_c$

$$\rho(t_c - t) \geq \|w(t)\| \geq \gamma(t_c - t). \tag{23}$$

On the other hand, since $t < t_c$, the very same equality implies

$$\langle q(t), w(t)\rangle \leq (t - t_c)\lambda \|q(t_c)\|^2 + o(t_c - t).$$

Hence, by (23), for a constant $\delta > 0$ and all $t < t_c$ sufficiently close to $t_c$ we have

$$\langle q(t), w(t)\rangle \leq -\delta(t_c - t) \leq -\frac{\delta}{\rho} \|w(t)\|. \quad \square$$

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

$$\left\{ \begin{array}{rcl} x_h'(s) &=& \dfrac{\partial H}{\partial p}(s, x_h(s), p_h(s)), \quad x_h(t) = x(t) + hw(t), \\ \\ -p_h'(s) &=& \dfrac{\partial H}{\partial x}(s, x_h(s), p_h(s)), \quad p_h(t) = p(t) + hq(t). \end{array} \right.$$

By the Gronwall lemma and assumptions of theorem, there exists $M_1$ independent from $t$ such that for all small $h > 0$

$$\|x_h - x\|_\infty \;+\; \|p_h - p\|_\infty \;\le\; M_1 h\left(\|w(t)\| + \|q(t)\|\right). \tag{24}$$

From Proposition 3.2 there exists $u_h \in \mathcal{U}$ such that $x_h$ solves the system

$$y' = f(s, y, u_h(s)) \tag{25}$$

and $p_h$ solves the linear system

$$-\,p' = \frac{\partial f}{\partial x}(s, x_h(s), u_h(s))^\star p - \frac{\partial L}{\partial x}(s, x_h(s), u_h(s)). \tag{26}$$

Denote by $\overline{p}_h$ the solution to (26) satisfying $\overline{p}_h(T) = -\nabla\varphi(x_h(T))$.

**Lemma 5.3.** *There exists $M_2 \ge 0$ independent from $t$ such that for all small $h > 0$*

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

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

*Proof.* Set $z_h = x_h - x - hw, \quad r_h = p_h - p - hq$. Then $z_h(t) = r_h(t) = 0$. By our assumptions there exists $c$ independent from $t$ such that for all small $h > 0$ and every $s \in [t, T]$ we have

$$\begin{aligned}
&\|z_h(s)\| + \|r_h(s)\| \\
&\quad\le \left|\int_t^s \left(\left\|\frac{\partial^2 H}{\partial x \partial p}(\tau, x(\tau), p(\tau))\right\| \|z_h(\tau)\| + \left\|\frac{\partial^2 H}{\partial p^2}(\tau, x(\tau), p(\tau))\right\| \|r_h(\tau)\|\right) d\tau\right| \\
&\qquad + \left|\int_t^s \left(\left\|\frac{\partial^2 H}{\partial x^2}(\tau, x(\tau), p(\tau))\right\| \|z_h(\tau)\| + \left\|\frac{\partial^2 H}{\partial p \partial x}(\tau, x(\tau), p(\tau))\right\| \|r_h(\tau)\|\right) d\tau\right| \\
&\qquad + c\left(\|x_h - x\|_\infty^2 \;+\; \|p_h - p\|_\infty^2\right).
\end{aligned}$$

The first inequality of our lemma follows from the Gronwall lemma and (24). On the other hand, for all small $h > 0$

$$\begin{aligned}
x_h'(s) - x'(s) - hw'(s) \quad &\in \quad \tfrac{\partial^2 H}{\partial x \partial p}(s, x(s), p(s)) z_h(s) + \tfrac{\partial^2 H}{\partial p^2}(s, x(s), p(s)) r_h(s) \\
&\quad + \quad 2nk(s)(\|x_h - x\|_\infty^2 \;+\; \|p_h - p\|_\infty^2) B
\end{aligned}$$

and, by (24), for some $c_1 > 0$ independent from $t$

$$\overline{p}_h(T) = -\varphi'(x(T)) - \varphi''(x(T))(x_h(T) - x(T)) + \varepsilon_h$$

where $\|\varepsilon_h\| \le c_1 h^2 \left(\|w(t)\|^2 + \|q(t)\|^2\right)$. Consequently,

$$\overline{p}_h(T) = p(T) - h\varphi''(x(T)) w(T) + \varepsilon_h^1 = p(T) + hq(T) + \varepsilon_h^1 = p_h(T) + \varepsilon_h^2$$

where $\left\|\varepsilon_h^1\right\| + \left\|\varepsilon_h^2\right\| \le c_2 h^2 \left(\|w(t)\|^2 + \|q(t)\|^2\right)$ and $c_2$ does not depend on $t$. Applying Gronwall's inequality we end the proof. □

**Lemma 5.4.** *Define*

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

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

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

*Proof.* Set $w_h = (x_h - x)/h,\ q_h = (p_h - p)/h$. By Lemma 5.3,

$$\|w_h(\tau) - w(\tau)\| + \|q_h(\tau) - q(\tau)\| \le M_2 h \left( \|w(t)\|^2 + \|q(t)\|^2 \right).$$

Thus for all small $h > 0$,

$$\varphi(x_h(T)) - \varphi(x(T)) = h \langle \varphi'(x(T)), w_h(T) \rangle + \frac{h^2}{2} \langle \varphi''(x(T)) w(T), w(T) \rangle + \varepsilon_h,$$

$$\begin{aligned}&\textstyle\int_t^T \left( \langle p_h(\tau), x_h'(\tau) \rangle - \langle p(\tau), x'(\tau) \rangle \right) d\tau \\ &\quad = \textstyle\int_t^T \left( h \langle q_h(\tau), x'(\tau) \rangle + h \langle p(\tau), w_h'(\tau) \rangle + h^2 \langle q(\tau), w'(\tau) \rangle \right) d\tau + \varepsilon_h^1\end{aligned}$$

and

$$\begin{aligned}&\textstyle\int_t^T (-H(\tau, x_h(\tau), p_h(\tau)) + H(\tau, x(\tau), p(\tau)))\, d\tau \\ &\quad = \textstyle\int_t^T \Big( -h \left\langle \frac{\partial H}{\partial x}(\tau, x(\tau), p(\tau)), w_h(\tau) \right\rangle - h \left\langle \frac{\partial H}{\partial p}(\tau, x(\tau), p(\tau)), q_h(\tau) \right\rangle \\ &\qquad - \tfrac{h^2}{2} \left\langle \frac{\partial^2 H}{\partial x^2}(\tau, x(\tau), p(\tau)) w(\tau), w(\tau) \right\rangle - \tfrac{h^2}{2} \left\langle \frac{\partial^2 H}{\partial p \partial x}(\tau, x(\tau), p(\tau)) q(\tau), w(\tau) \right\rangle \\ &\qquad - \tfrac{h^2}{2} \left\langle \frac{\partial^2 H}{\partial x \partial p}(\tau, x(\tau), p(\tau)) w(\tau), q(\tau) \right\rangle \\ &\qquad - \tfrac{h^2}{2} \left\langle \frac{\partial^2 H}{\partial p^2}(\tau, x(\tau), p(\tau)) q(\tau), q(\tau) \right\rangle \Big)\, d\tau + \varepsilon_h^2\end{aligned}$$

where $\|\varepsilon_h\| + \|\varepsilon_h^1\| + \|\varepsilon_h^2\| \le ch^3 \left( \|w(t)\|^2 + \|q(t)\|^2 \right)$ and $c > 0$ is independent from $t$.

Recalling that $(x, p)$ solves (2), that $(w, q)$ solves (14) and

$$q(T) = -\varphi''(x(T)) w(T)$$

we obtain, by adding the above expressions,

$$\begin{aligned}I_h =\,& h \langle -p(T), w_h(T) \rangle - \frac{h^2}{2} \langle q(T), w(T) \rangle \\ &+ h \int_t^T \left( \langle q_h(\tau), x'(\tau) \rangle + \langle p(\tau), w_h'(\tau) \rangle + \langle p'(\tau), w_h(\tau) \rangle - \langle x'(\tau), q_h(\tau) \rangle \right) d\tau \\ &+ h^2 \int_t^T \left( \langle q(\tau), w'(\tau) \rangle + \frac{1}{2} \langle q'(\tau), w(\tau) \rangle - \frac{1}{2} \langle w'(\tau), q(\tau) \rangle \right) + \varepsilon_h^3\end{aligned}$$

where $\|\varepsilon_h^3\| \le ch^3 \left( \|w(t)\|^2 + \|q(t)\|^2 \right)$. Hence

$$\begin{aligned}I_h \quad &= -\tfrac{h^2}{2} \langle q(T), w(T) \rangle - h \langle p(t), w_h(t) \rangle + \tfrac{h^2}{2} \left( \langle q(T), w(T) \rangle - \langle q(t), w(t) \rangle \right) + \varepsilon_h^3 \\ &= -h \langle p(t), w_h(t) \rangle - \tfrac{h^2}{2} \langle q(t), w(t) \rangle + \varepsilon_h^3. \quad \square\end{aligned}$$

*Proof of Theorem* 5.1. Denote by $y_h$ the solution to (25) with $y_h(t) = x(t)$. Then for some $c$ independent from $t$ and all small $h > 0$ we have

$$\|y_h - x_h\|_\infty + \|y_h' - x_h'\|_{L^1} \le ch \|w(t)\|. \tag{27}$$

Let $X_h$ denote the fundamental solution to

$$X' = \frac{\partial f}{\partial x}(s, x_h(s), u_h(s)) X, \quad X(t) = Id,$$

and $v_h$ be the solution to

$$v' = \frac{\partial f}{\partial x}(s, x_h(s), u_h(s))v, \quad v(t) = -hw(t).$$

Then $v_h(s) = -hX_h(s)w(t)$. From the Gronwall lemma and (27), it is not difficult to deduce that there exists $c_1$ independent from $t$ such that for all small $h > 0$

$$\forall\, s, \quad \|y_h(s) - x_h(s) - v_h(s)\|_\infty \le c_1 h^2 \|w(t)\|^2 .$$

Set

$$II_h := \varphi(y_h(T)) + \int_t^T L(\tau, y_h(\tau), u_h(\tau))d\tau - \varphi(x_h(T)) - \int_t^T L(\tau, x_h(\tau), u_h(\tau))d\tau.$$

Then for a constant $c_2$ independent from $t$ and for all small $h > 0$ we have

$$\begin{aligned} II_h &= \langle \nabla\varphi(x_h(T)), v_h(T)\rangle + \textstyle\int_t^T \left\langle \frac{\partial L}{\partial x}(\tau, x_h(\tau), u_h(\tau)), v_h(\tau)\right\rangle d\tau + \varepsilon_h \\ &= -h\left\langle X_h(T)^\star \nabla\varphi(x_h(T)) + \textstyle\int_t^T X_h(\tau)^\star \frac{\partial L}{\partial x}(\tau, x_h(\tau), u_h(\tau))d\tau, w(t)\right\rangle + \varepsilon_h \\ &= h\,\langle \overline{p}_h(t), w(t)\rangle + \varepsilon_h \end{aligned}$$

where $\|\varepsilon_h\| \le c_2 h^2 \|w(t)\|^2$. This and Lemma 5.3 imply that for some $c_3$ independent from $t$

$$\|II_h - h\,\langle p(t) + hq(t), w(t)\rangle\| \le c_2 h^2 \|w(t)\|^2 + c_3 h^3 \left(\|w(t)\|^2 + \|q(t)\|^2\right).$$

To end the proof we proceed by a contradiction argument. Assume for a moment that $(x, \overline{u})$ provides a weak local minimum to $(P)$. Then for all small $h > 0$ we have

$$\varphi(x(T)) + \int_t^T L(\tau, x(\tau), \overline{u}(\tau))d\tau \le \varphi(y_h(T)) + \int_t^T L(\tau, y_h(\tau), u_h(\tau))d\tau.$$

By Proposition 3.2,

$$\begin{aligned} L(\tau, x(\tau), \overline{u}(\tau)) &= \langle p(\tau), x'(\tau)\rangle - H(\tau, x(\tau), p(\tau)), \\ L(\tau, x_h(\tau), u_h(\tau)) &= \langle p_h(\tau), x_h'(\tau)\rangle - H(\tau, x_h(\tau), p_h(\tau)). \end{aligned}$$

Thus, $I_h + II_h \ge 0$. From Lemma 5.4,

$$\begin{aligned} 0 \le & I_h + II_h \le -h\,\langle p(t), w(t)\rangle - \frac{h^2}{2}\langle q(t), w(t)\rangle + h\,\langle p(t) + hq(t), w(t)\rangle \\ & + (c_3 + M_3)h^3 \left(\|w(t)\|^2 + \|q(t)\|^2\right) + c_2 h^2 \|w(t)\|^2 \\ \le & \frac{h^2}{2}\langle q(t), w(t)\rangle + (c_3 + M_3)h^3 \left(\|w(t)\|^2 + \|q(t)\|^2\right) + c_2 h^2 \|w(t)\|^2 \end{aligned}$$

Dividing the above by $h^2$ and taking the limit when $h \to 0$ yields

$$0 \le \tfrac{1}{2}\langle q(t), w(t)\rangle + c_2 \|w(t)\|^2 .$$

Hence, from Lemma 5.2 we obtain $-\gamma \ge -2c_2 \|w(t)\|$. Since $w(t_c) = 0$, taking $t$ sufficiently close to $t_c$ we derive a contradiction. $\square$

## References


1. J.-P. Aubin and A. Cellina (1984) *Differential Inclusions*, Springer-Verlag, Gründlehren der Math. Wiss. MR **85j:**49010
2. J.-P. Aubin and H. Frankowska (1990) *Set-Valued Analysis*, Birkhäuser, Boston, Basel, Berlin. MR **91d:**49001
3. V. Barbu and G. Da Prato (1983) *Hamilton-Jacobi equations and synthesis of nonlinear control processes in Hilbert space*, J. Diff. Eqs., 48, 350-372. MR **84i:**49048
4. E. N. Barron and R. Jensen (1990) *Semicontinuous viscosity solutions for Hamilton-Jacobi equations with convex hamiltonians*, Comm. Partial Diff. Eqs., 15, 113-174. MR **91h:**35069
5. A. Bensoussan (1988) *Perturbation methods in optimal control*, Wiley/Gautier-Villars Series in Modern Applied Mathematics, Bordas, Paris and John Wiley & Sons Ltd. MR **89m:**93002
6. Ch. Byrnes and H. Frankowska (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. MR **93h:**49092
7. Ch. Byrnes and H. Frankowska (submitted) *Uniqueness of optimal trajectories and the nonexistence of shocks for Hamilton-Jacobi-Bellman and Riccati partial differential equations*.
8. P. Cannarsa and H. Frankowska (1991) *Some characterizations of optimal trajectories in control theory*, SIAM J. Control and Optimiz., 29, 1322-1347. MR **92j:**49025
9. N. Caroff (1994) *Caractéristiques de l'équation d'Hamilton-Jacobi et conditions d'optimalité en contrôle optimal non linéaire*, Thèse de Doctorat, Université Paris-Dauphine.
10. N. Caroff and H. Frankowska (1992) *Optimality and characteristics of Hamilton-Jacobi-Bellman equations,* in *Proceedings of Premier Colloque Franco-Roumain sur l'Optimisation, Contrôle Optimal, Equations aux Dérivées Partielles*, 7-11 Septembre, 1992, International Series of Numerical Mathematics, vol. 107, Birkhäuser Verlag, Basel, 169-180. MR **94g:**49071
11. L. Cesari (1993) *Optimization Theory and Applications,* Springer-Verlag, New-York, Heidelberg, Berlin. MR **85c:**49001
12. E. D. Conway and E. Hopf (1964) *Hamilton's Theory and Generalized Solutions of the Hamilton-Jacobi Equation*, J. Math. Mech., 13, 939-986. MR **32:**243
13. M. G. Crandall and P.-L. Lions (1983) *Viscosity solutions of Hamilton-Jacobi equations*, Transactions of A.M.S., 277, 1-42. MR **85g:**35029
14. C. M. Dafermos (1977) *Generalized Characteristics and the Structure of Solutions of Hyperbolic Conservation Laws*, Indiana University Math. J., 26, 1097-1119. MR **56:**16151
15. W. H. Fleming and R. W. Rishel (1975) *Deterministic and Stochastic Optimal Control*, Springer-Verlag, New York. MR **56:**13016
16. H. Frankowska (1989) *Contingent cones to reachable sets of control systems*, SIAM J. on Control and Optimization, 27, 170-198. MR **89m:**49034
17. H. Frankowska (1991) *Lower semicontinuous solutions of Hamilton-Jacobi-Bellman equation*, Proceedings of 30th IEEE Conference on Decision and Control, Brighton, GB, December 11-13.
18. H. Frankowska (1993) *Lower semicontinuous solutions of Hamilton-Jacobi-Bellman equation*, SIAM J. on Control and Optimization, 31, 257-272. MR **94c:**49037
19. M. R. Hestenes (1966) *Calculus of Variations and Optimal Control Theory*, Wiley. MR **34:**3390
20. A. D. Ioffe and V. Tichomirov (1979) *Theory of Extremal Problems*, North-Holland, Amsterdam, New York, Oxford. MR **80d:**49001b
21. E. B. Lee and L. Markus (1967) *Foundations of Optimal Control Theory*, J. Wiley and Sons. MR **36:**3596
22. D. Q. Mayne (1977) *Sufficient Conditions for a Control to be a Strong Minimum*, J. Opt. Th. Appl., 21, 339-351. MR **56:**16486
23. D. Orell and V. Zeidan (1988) *Another Jacobi Sufficiency Criterion for Optimal Control with Smooth Constraints*, J. Opt. Th. Appl., 58, 283-300. MR **89h:**49023

24. Count J. F. Riccati (1724) *Animadversationes in aequationes differentiales secundi gradus*, Actorum Eruditorum quae Lipsiae Publicantur, Supplementa 8, 66-73.
25. W. T. Reid (1972) *Riccati Differential Equations*, Academic Press. MR **50:**10401
26. J. Smoller (1960) *Shock Waves and Reaction-Diffusion Equations*, Springer-Verlag, Gründlehren der Math. Wiss. MR **95g:**35002
27. M. Spivak (1975) *A Comprehensive Introduction to Differential Geometry*, Publish or Perish, Inc., Boston, Mass. MR **52:**15254b
28. V. Zeidan (1983) *Sufficient Conditions for the generalized problem of Bolza*, Transactions of A.M.S., 275, 561-583. MR **84c:**49033
29. V. Zeidan (1984) *A Modified Hamilton-Jacobi Approach in the Generalized Problem of Bolza*, Appl. Math. Opt., 11, 97-109. MR **85m:**49056
30. V. Zeidan (1984) *First and Second Order Sufficient Conditions for Optimal Control and the Calculus of Variations*, Appl. Math. Opt., 11, 209-226. MR **86c:**49034
31. V. Zeidan (1984) *Extended Jacobi Sufficiency Criterion for Optimal Control*, SIAM J. Control and Optim., 22, 294-300. MR **85f:**49035
32. V. Zeidan (1984) *Sufficiency Conditions with Minimal Regularity Assumptions*, Appl. Math. Opt., 20, 19-31. MR **90d:**49016
33. V. Zeidan and P. Zezza (1988) *Necessary Conditions for Optimal Control Problems: conjugate points*, SIAM J. Control and Optim., 26, 592-608. MR **89h:**49015
34. V. Zeidan and P. Zezza (1988) *The Conjugate Point Condition for Smooth Control Sets*, J. Math. An. Appl., 132, 572-589. MR **89j:**49014
35. V. Zeidan and P. Zezza (1989) *Coupled Points in the Calculus of Variations and Applications to Periodic Problems*, Transactions of A.M.S., 315, 323-335. MR **90b:**49037
36. V. Zeidan and P. Zezza (1991) *Coupled Points in Optimal Control Theory*, IEEE, Transactions on Automatic Control, 36, 1276-1281. MR **92k:**49040

CEREMADE, Université Paris-Dauphine, 75775 Paris Cedex 16, France