

SUFFICIENT CONDITIONS FOR THE GENERALIZED PROBLEM OF BOLZA

BY

VERA ZEIDAN

ABSTRACT. This paper presents sufficient conditions for strong local optimality in the generalized problem of Bolza. These conditions represent a unification, in the sense that they can be applied to both the calculus of variations and to optimal control problems, as well as problems with nonsmooth data. Also, this paper brings to light a new point of view concerning the Jacobi condition in the classical calculus of variations, showing that it can be considered as a condition which guarantees the existence of a canonical transformation which transforms the original Hamiltonian to a locally concave-convex Hamiltonian.

1. Introduction. Consider the generalized problem of Bolza:

$$\text{Minimize } \int_a^b L(t, x(t), \dot{x}(t)) dt$$

subject to

$$x(a) = A, \quad x(b) = B,$$

where x is an absolutely continuous function from $[a, b]$ to \mathbf{R}^n with derivative \dot{x} (almost everywhere), and where L is allowed to assume the value $+\infty$; i.e., $L: [a, b] \times \mathbf{R}^n \times \mathbf{R}^n \rightarrow \mathbf{R} \cup \{+\infty\}$.

The fact that L is extended real-valued implies that our generalized problem of Bolza unifies the calculus of variations and optimal control problems.

The Hamiltonian H is defined by the conjugacy formula

$$H(t, x, p) = \sup\{\langle p, v \rangle - L(t, x, v) : v \in \mathbf{R}^n\}.$$

The existence of a solution for the generalized problem of Bolza was studied by R. T. Rockafellar in [18] and necessary conditions were developed by F. H. Clarke in [4, 7]. This paper presents sufficient conditions and like the above articles focuses on the Hamiltonian H defined earlier. Thus, this paper completes the program of studying the generalized problem of Bolza from the point of view of the Hamiltonian.

In the classical calculus of variations a number of procedures are available to confirm the optimality of an extremal which satisfies the necessary conditions. These procedures evoke the field of extremals, the Hamilton-Jacobi equation and the

Received by the editors February 6, 1981 and, in revised form, October 9, 1981.

1980 *Mathematics Subject Classification.* Primary 49B10.

Key words and phrases. Sufficient conditions, strong local minimality, generalized problem of Bolza, canonical transformations, calculus of variations, conjugate points, Jacobi condition.

©1983 American Mathematical Society
0002-9947/81/0000-0097/\$07.25

Jacobi equation. The situation in optimal control theory is relatively undeveloped, since these procedures are largely unavailable—except in simplified contexts. For the generalized problem of Bolza, which is a general form of the calculus of variations and optimal control problems, the only sufficient condition existing in the literature requires the Hamiltonian $H(t, x, p)$ to be concave in x and convex in p .

In this paper sufficient conditions are developed for the generalized problem of Bolza where the Hamiltonian is not necessarily concave-convex. The method we employ, which we describe presently, can be used to treat problems in which endpoints constraints and/or functionals enter in a much more general way than described above. We shall present this extension in §6. Our method requires the use of a canonical transformation of “Hamiltonian inclusions”. This transformation takes the problem locally, near given arcs (\hat{x}, \hat{p}) satisfying the Hamiltonian equations, into a new problem having a concave-convex Hamiltonian to which the convex theory can be applied. The existence of such a transformation is guaranteed by assuming the existence of an auxiliary function satisfying a certain inequality as opposed to an equation.

In the classical setting, it can be shown that our inequality criterion is in fact equivalent to a well-known condition involving the Jacobi equation. This equivalence sheds new light on the Jacobi condition, since the latter can now be interpreted as being essentially a necessary and sufficient condition such that there exists a certain kind of canonical transformation for which the transformed Hamiltonian is locally concave-convex. However, there are examples in the classical setting where our form of the criterion is easier to apply, and of course it applies when others do not, such as in certain cases of nondifferentiable and/or extended real-valued data.

For certain problems in optimal control theory if we consider the special case treated by D. Mayne in [11], that is when the control $\hat{u}(t)$ belongs to the interior of the control set U , then our conditions reduce to his result in [11, Theorem 3.2].

2. Discussion of the problem and main results. A function $x(\cdot)$ from $[a, b]$ to \mathbf{R}^n is an *arc* if it is absolutely continuous. The arc $x(\cdot)$ is *admissible* if it satisfies $x(a) = A$ and $x(b) = B$. Suppose we are given an admissible arc \hat{x} from $[a, b]$ to \mathbf{R}^n . We are interested in finding conditions which guarantee the local optimality of \hat{x} for the generalized problem of Bolza. Hence, the study of the problem will focus on some neighborhoods of the arc \hat{x} and the arc \hat{p} such that (\hat{x}, \hat{p}) satisfy the “Hamiltonian inclusions”.

For given arcs \hat{x}, \hat{p} , and some positive numbers ϵ, δ , we define

$$N(\epsilon, \delta) = \{(t, x, p): t \in [a, b], |x - \hat{x}(t)| < \epsilon \text{ and } |p - \hat{p}(t)| < \delta\}$$

and

$$N(\epsilon, \infty) = \{(t, x, w): t \in [a, b], |x - \hat{x}(t)| < \epsilon \text{ and } w \in \mathbf{R}^n\}.$$

Consider the generalized Bolza problem:

$$(P) \quad \text{Minimize } J(x) = \int_a^b L(t, x(t), \dot{x}(t)) dt$$

subject to

$$x(a) = A, \quad x(b) = B,$$

where x is an arc from $[a, b]$ to \mathbf{R}^n , and where L is a given function, $L: N(\epsilon, \infty) \rightarrow \mathbf{R} \cup \{+\infty\}$ for some given positive number ϵ .

REMARK. Because L can take the value $+\infty$, the problem is much more general than it seems. For more information see [5 and 17].

The Hamiltonian of the problem,

$$(2.1) \quad H(t, z) = \sup\{\langle p, v \rangle - L(t, x, v) : v \in \mathbf{R}^n\},$$

where $z = (x, p)$, is then defined on $N(\epsilon, \infty)$ and it is convex in p .

Let \mathcal{L} be the collection of Lebesgue measurable subsets of $[a, b]$ and \mathcal{B} the Borel subsets of $\mathbf{R}^n \times \mathbf{R}^n$. We denote by $\mathcal{L} \times \mathcal{B}$ the σ -algebra of subsets of $[a, b] \times \mathbf{R}^n \times \mathbf{R}^n$ generated by products of sets in \mathcal{L} and \mathcal{B} .

Suppose we are also given an arc \hat{p} . The following hypothesis will be made:

(H₁) L is $\mathcal{L} \times \mathcal{B}$ measurable, and there exist an integrable function $\rho(\cdot)$ on $[a, b]$ and some positive real number δ such that, for $(t, z) \in N(\epsilon, \delta)$, $|H(t, z)| \leq \rho(t)$.

DEFINITION. The Hamiltonian H is said to be C^{1+} near (\hat{x}, \hat{p}) if there exist positive numbers ϵ and δ such that for each t in $[a, b]$, $H(t, \cdot)$ is C^1 with locally Lipschitz first derivatives on

$$\{z = (x, p) : |x - \hat{x}(t)| < \epsilon, |p - \hat{p}(t)| < \delta\}.$$

If the Hamiltonian H is C^{1+} then the gradient of H with respect to z , $H_z(t, \cdot)$, is locally Lipschitz and hence the generalized Jacobian $\partial_z H_z(t, \cdot)$ exists and it is defined at a point z as being the convex hull of all matrices M of the form

$$M = \lim_{i \rightarrow \infty} \{D_z H_z(t, z_i)\},$$

where z_i converges to z and the usual Jacobian $D_z H_z(t, z_i)$ exists for each i .

We will also make the following hypothesis.

(H₂) The Hamiltonian H is C^{1+} near (\hat{x}, \hat{p}) with associated (ϵ, δ) , and the map

$$(t, z) \rightarrow \partial_z H_z(t, z)$$

is upper semicontinuous on $N(\epsilon, \delta)$, $H_z(\cdot, z)$ is continuous for z near \hat{z} .

In the autonomous case the hypothesis (H₂) reduces to saying that the map $z \rightarrow H(z)$ is C^{1+} on

$$\{z = (x, p) : |x - \hat{x}(t)| < \epsilon, |p - \hat{p}(t)| < \delta\}.$$

DEFINITION. We say that L satisfies the *Weierstrass condition* at \hat{x} if there exists a function $\zeta(\cdot)$ on $[a, b]$ such that, for almost all t , for all v ,

$$L(t, \hat{x}(t), \dot{\hat{x}}(t) + v) - L(t, \hat{x}(t), \dot{\hat{x}}(t)) \geq \langle v, \zeta(t) \rangle.$$

REMARK. The Weierstrass condition is also a necessary condition for optimality of \hat{x} in the generalized problem of Bolza. For reference see [2 and 5].

DEFINITION. The admissible arc \hat{x} is a *strong local minimum* for the problem (P) if there exists a positive real number γ such that \hat{x} minimizes $J(x)$ over all admissible arcs x satisfying

$$|x(t) - \hat{x}(t)| < \gamma, \quad \text{for all } t \in [a, b].$$

When $H(t, \cdot)$ is locally Lipschitz and satisfies some additional technical hypotheses, a necessary condition for optimality of \hat{x} is that there exist an arc \hat{p} on $[a, b]$ such that $\hat{z} = (\hat{x}, \hat{p})$ satisfies the Hamiltonian inclusions, i.e.,

$$J\dot{\hat{z}}(t) \in \partial_z H(t, \hat{z}(t)) \quad \text{a.e.,}$$

where $J = \begin{bmatrix} 0 & -I \\ I & 0 \end{bmatrix}$, and I is the $n \times n$ -identity matrix [7].

The main result in this paper consists of sufficient conditions obtained by strengthening the necessary conditions, i.e., the Weierstrass condition and the Hamiltonian inclusions.

The proof of the following theorem will be postponed to §5.

THEOREM 1. *Let the arcs \hat{x}, \hat{p} be given such that \hat{x} is admissible. Assume the hypotheses (H_1) and (H_2) , and*

- (a) *L satisfies the Weierstrass condition at \hat{x} for $\xi = \hat{p}$,*
- (b) *the arc $\hat{z} = (\hat{x}, \hat{p})$ satisfies*

$$J\dot{\hat{z}}(t) = H_z(t, \hat{z}(t)) \quad \text{for } t \in [a, b],$$

- (c) *there exists a C^1 -matrix function $Q(\cdot)$ from $[a, b]$ to the space of $n \times n$ matrices such that, for all $t \in [a, b]$, $Q(t)$ is symmetric and satisfies*

$$(2.2) \quad \dot{Q}(t) - Q(t)\gamma(t)Q(t) + Q(t)\beta(t) + \delta(t)Q(t) - \alpha(t) > 0$$

(positive definite) for all $t \in [a, b]$, and for all matrices

$$\begin{pmatrix} \alpha(t) & \delta(t) \\ \beta(t) & \gamma(t) \end{pmatrix} \in \partial_z H_z(t, \hat{z}(t)).$$

Then \hat{x} provides a strong local minimum for (P).

REMARK. As we shall see, the hypothesis (H_1) implies that the integral $J(x) = \int_a^b L(t, x(t), \dot{x}(t)) dt$ is well defined (possibly $+\infty$) near \hat{x} .

REMARK. If $L(t, x, v)$ is convex in v then the assumption (a) of the theorem is satisfied. Also, the measurability of H would imply the measurability of L and hence, in this case, all the hypotheses could be framed in terms of H .

3. The case of smooth Hamiltonians: H is C^2 . In this section we assume that $H(\cdot, \cdot, \cdot)$ is C^2 near some given arcs (\hat{x}, \hat{p}) . When we merely assume that H is C^2 , we can nonetheless reduce the condition (c) of Theorem 1 to a condition which, in the classical case, i.e., when L is also C^2 , will be equivalent to a familiar one involving the Jacobi equation. But, of course, the theorem also applies to functions L which are less than C^2 , in fact even nondifferentiable and extended-valued.

In the following proposition we will show that in the case when only $H(\cdot, \cdot, \cdot)$ is C^2 , but L can be nonsmooth and extended-valued, the inequality in (2.2) is equivalent to equality. Also we will prove that condition (c) of the theorem is

equivalent to a condition, which, in the case when $L(\cdot, \cdot, \cdot)$ is also C^2 , will be exactly the Jacobi condition in Hamiltonian form.

Suppose we are given arcs \hat{x} and \hat{p} with \hat{x} admissible. The notation $\hat{\phi}(t)$ will mean that the function ϕ is evaluated at $(t, \hat{x}(t), \hat{p}(t))$, and

$$\phi_{pp} = \frac{\partial^2}{\partial p \partial p} \phi, \quad \phi_{xp} = \frac{\partial^2}{\partial p \partial x} \phi, \quad \phi_{px} = \frac{\partial^2}{\partial x \partial p} \phi, \quad \text{and} \quad \phi_{xx} = \frac{\partial^2}{\partial x \partial x} \phi.$$

PROPOSITION 1. Assume that the Hamiltonian H , defined in (2.1), is such that $H(\cdot, \cdot, \cdot)$ is C^2 near (\hat{x}, \hat{p}) and that $\hat{H}_{pp}(t)$ is positive definite for t in $[a, b]$. Then the following are equivalent:

(i) the condition (c) of the theorem is satisfied, i.e., there exists a C^1 -matrix function $Q(\cdot)$ from $[a, b]$ to the space of $n \times n$ -matrices such that $Q(t)$ is symmetric and satisfies

$$(3.1) \quad \dot{Q}(t) - Q(t)\hat{H}_{pp}(t)Q(t) + \hat{H}_{xp}(t)Q(t) + Q(t)\hat{H}_{px}(t) - \hat{H}_{xx}(t) > 0$$

for all $t \in [a, b]$,

(ii) there exists a C^1 -matrix function $Q(\cdot)$ on $[a, b]$ such that $Q(t)$ is symmetric and satisfies

$$(3.2) \quad \hat{Q}(t) - Q(t)\hat{H}_{pp}(t)Q(t) + \hat{H}_{xp}(t)Q(t) + Q(t)\hat{H}_{px}(t) - \hat{H}_{xx}(t) = 0$$

for all $t \in [a, b]$,

(iii) there exists no nontrivial solution h from $[a, b]$ to \mathbb{R}^n of the equation

$$(3.3) \quad \frac{d}{dt} \left(\hat{H}_{pp}^{-1}(t)\dot{h}(t) - \hat{H}_{pp}^{-1}(t)\hat{H}_{px}(t)h(t) \right) + \hat{H}_{xp}(t)\hat{H}_{pp}^{-1}(t)\dot{h}(t) \\ + \hat{H}_{xx}(t)h(t) - \hat{H}_{xp}(t)\hat{H}_{pp}^{-1}(t)\hat{H}_{px}(t)h(t) = 0$$

such that $h(a) = 0$ and $h(c) = 0$ for some $c \in (a, b]$.

PROOF. It suffices to prove these implications: (iii) \rightarrow (ii), (ii) \rightarrow (i), and (i) \rightarrow (iii).

Suppose (iii) holds. By [9, Theorem 10.2] (iii) implies that there exists a solution $(U_0(t), V_0(t))$ of the matrix system

$$\begin{aligned} \dot{U}(t) &= \hat{H}_{px}(t)U(t) + \hat{H}_{pp}(t)V(t), \\ \dot{V}(t) &= -\hat{H}_{xx}(t)U(t) - \hat{H}_{xp}(t)V(t), \end{aligned}$$

with $\det U_0(t) \neq 0$ on $[a, b]$ and $U_0^T V_0 = V_0^T U_0$, the symbol $(^T)$ stands for the transpose of the matrix.

Let $Q_0(t) = -V_0(t)U_0^{-1}(t)$. Then $Q_0(\cdot)$ is C^1 , and $Q_0(t)$ is symmetric and solves (3.2), so that (ii) holds.

The implication (ii) \rightarrow (i) is obtained by using the embedding theorem of first order differential equations in [10].

Suppose (i) holds for some matrix function $Q_0(t)$. Define

$$Z(t) = \dot{Q}_0(t) - Q_0(t)\hat{H}_{pp}(t)Q_0(t) + \hat{H}_{xp}(t)Q_0(t) + Q_0(t)\hat{H}_{px}(t) - \hat{H}_{xx}(t).$$

Then $Z(t) > 0$ (positive definite) for all t in $[a, b]$.

Let $U_0(t)$ be the solution of the differential equation

$$\dot{U}(t) = (\hat{H}_{px}(t) - \hat{H}_{pp}(t)Q_0(t))U(t), \quad U(a) = I,$$

where I is the $n \times n$ -identity matrix, and let $V_0(t)$ be

$$V_0(t) = -Q_0(t)U_0(t) \quad \text{for } t \in [a, b].$$

Then we have that $\det U_0(t) \neq 0$ for t in $[a, b]$, and $U_0^T V_0 = V_0^T U_0$, and $(U_0(t), V_0(t))$ solves the matrix system

$$\begin{aligned} \dot{U}(t) &= \hat{H}_{px}(t)U(t) + \hat{H}_{pp}(t)V(t), \\ \dot{V}(t) &= -(Z(t) + \hat{H}_{xx}(t))U(t) - \hat{H}_{xp}(t)V(t), \end{aligned}$$

for $t \in [a, b]$. Thus, by [9, Theorem 10.2] we conclude that there exists no nontrivial solution $h(t)$ on $[a, b]$ solving

$$\begin{aligned} \frac{d}{dt} [\hat{H}_{pp}^{-1}(t)\dot{h}(t) - \hat{H}_{pp}^{-1}(t)\hat{H}_{px}(t)h(t)] + \hat{H}_{xp}(t)\hat{H}_{pp}^{-1}(t)\dot{h}(t) + \hat{H}_{xx}(t)h(t) \\ - \hat{H}_{xp}(t)\hat{H}_{pp}^{-1}(t)\hat{H}_{px}(t)h(t) + Z(t)h(t) = 0, \end{aligned}$$

with $h(a) = 0$ and $h(c) = 0$ for some $c \in (a, b]$. But that is equivalent by [9, Theorem 10.3] to saying that: for every interval $[\alpha, \beta] \subset [a, b]$ we have

$$\begin{aligned} I(\eta; \alpha, \beta) &= \int_{\alpha}^{\beta} \{ [\hat{H}_{pp}^{-1}(t)\dot{\eta} - \hat{H}_{pp}^{-1}(t)\hat{H}_{px}(t)\eta] \cdot \dot{\eta} \\ &\quad + [-\hat{H}_{xp}(t)\hat{H}_{pp}^{-1}(t)\dot{\eta} \\ &\quad - (\hat{H}_{xx}(t) - \hat{H}_{xp}(t)\hat{H}_{pp}^{-1}(t)\hat{H}_{px}(t) + Z(t))\eta] \cdot \eta \} dt \geq 0 \end{aligned}$$

for every absolutely continuous vector function η such that

$$\eta(\alpha) = \eta(\beta) = 0, \quad \text{and} \quad \dot{\eta} \text{ is of class } L^2 \text{ on } [\alpha, \beta],$$

and

$$(2) \quad I(\eta; \alpha, \beta) = 0 \quad \text{if and only if } \eta \equiv 0.$$

But $Z(t)$ is positive definite for each t , so for every interval $[\alpha, \beta] \subset [a, b]$

$$I_1(\eta; \alpha, \beta) = I(\eta; \alpha, \beta) + \int_{\alpha}^{\beta} Z(t)\eta \cdot \eta dt$$

satisfies the conditions (1) and (2), and hence by [9, Theorem 10.3] condition (iii) holds. Q.E.D.

Now, suppose that $L(\cdot, \cdot, \cdot)$ is also C^2 for x near \hat{x} . Then we are in the classical case. The following corollary will show the connection between our sufficient conditions and the one given in the literature which involves the Jacobi equation.

The notation $\hat{\phi}(t)$ will mean that the function ϕ is evaluated at $(t, \hat{x}(t), \dot{\hat{x}}(t))$.

DEFINITION. Given an arc \hat{x} on $[a, b]$. A point $c \in (a, b)$ is said to be *conjugate* to a , if there exists a nontrivial solution $h = (h_1, \dots, h_n)$ of the *Jacobi equation*

$$(3.4) \quad \frac{d}{dt} (\hat{L}_{vv}(t)\dot{h}(t) + \hat{L}_{vx}(t)h(t)) - \hat{L}_{xv}(t)\dot{h}(t) - \hat{L}_{xx}(t)h(t) = 0$$

with $h(a) = 0$ and $h(c) = 0$.

DEFINITION. The arc \hat{x} satisfies the *Jacobi condition* if there exists no conjugate point of a in $(a, b]$.

The *Weierstrass function* is defined by

$$E(t, x, v, w) = L(t, x, w) - L(t, x, v) - (w - v) \cdot L_v(t, x, v).$$

COROLLARY. Suppose we are given arcs \hat{x}, \hat{p} such that \hat{x} is admissible. Assume

- (1) $L(\cdot, \cdot, \cdot)$ is C^2 for x near \hat{x} , and $H(\cdot, \cdot, \cdot)$ is C^2 for (x, p) near (\hat{x}, \hat{p}) ,
- (2) the arc $\hat{z} = (\hat{x}, \hat{p})$ satisfies

$$J\dot{\hat{z}}(t) = H_z(t, \hat{z}(t)) \quad \text{for all } t \in [a, b],$$

where $J = \begin{bmatrix} 0 & -I \\ I & 0 \end{bmatrix}$ and I is the $n \times n$ -identity matrix,

- (3) the strengthened Legendre condition

$$L_{vv}(t, \hat{x}(t), \hat{x}(t)) > 0 \quad (\text{positive definite}),$$

- (4) for (x, v) near (\hat{x}, \hat{x}) and for all w in \mathbb{R}^n

$$E(t, x, v, w) \geq 0.$$

Then the conditions (i), (ii) and (iii) of Proposition 1 are equivalent to:

- (iv) the arc \hat{x} satisfies the Jacobi condition.

PROOF. The hypotheses (1), (2), (3) and (4) allow us to write the Jacobi equation (3.4) in terms of the Hamiltonian. Now (3.4) turns out to be the equation (3.3), and hence condition (iv) is equivalent to condition (iii). The hypothesis (3) implies that $\hat{H}_{pp}(t) > 0$, so apply Proposition 1 to complete the proof. Q.E.D.

REMARK. For the verification of the Jacobi condition, condition (i) of Proposition 1 suggests a method which can be more convenient (see Example 1), and which does not appear to have been noted in the calculus of variations.

4. Concave-convex Hamiltonians. A criterion for local concavity. In studying the generalized problem of Bolza sufficient conditions were obtained in [15, 16 and 19] when the Hamiltonian is concave in x and convex in p . The main technique of the present paper is based on using canonical transformations [8] to transform the original Hamiltonian $H(t, x, p)$, which is not necessarily concave-convex, even locally in x , to a new Hamiltonian $H^*(t, X, P)$ such that H^* is locally concave in X around an arc \hat{X} and convex in P .

For the new Hamiltonian H^* we need our own version of the sufficiency theorem obtained for concave-convex Hamiltonians.

We are given arcs \hat{X}, \hat{P} from $[a, b]$ to \mathbb{R}^n . For given positive numbers α^* and δ^* we define

$$N^*(\alpha^*, \delta^*) = \{(t, X, P): t \in [a, b], |X - \hat{X}(t)| < \alpha^*, |P - \hat{P}(t)| < \delta^*\},$$

and

$$N^*(\alpha^*, \infty) = \{(t, X, P): t \in [a, b], |X - \hat{X}(t)| < \alpha^*, P \in \mathbb{R}^n\}.$$

Let $H^*(t, Z)$ be a given function defined in $N^*(\alpha^*, \infty)$ for some given positive number α^* .

DEFINITION. $H^*(t, \cdot, \hat{P}(t))$ is said to be *locally concave around the arc \hat{X}* if there exists a positive number γ^* such that, for any $t \in [a, b]$, and X_1, X_2 satisfying

$$|X_1 - \hat{X}(t)| \leq \gamma^*, \quad |X_2 - \hat{X}(t)| \leq \gamma^*,$$

and for any $\lambda, 0 \leq \lambda \leq 1$, we have

$$H^*(t, \lambda X_1 + (1 - \lambda)X_2, \hat{P}(t)) \geq \lambda H^*(t, X_1, \hat{P}(t)) + (1 - \lambda)H^*(t, X_2, \hat{P}(t)).$$

Let us now define

$$L^*(t, X, V) = \sup\{\langle P, V \rangle - H^*(t, X, P) : P \in \mathbf{R}^n\}$$

as a function on $N^*(\alpha^*, \infty)$, and consider the problem

$$(P^*) \quad \text{minimize } J^*(X) = \int_a^b L^*(t, X(t), \dot{X}(t)) dt$$

subject to

$$X(a) = A^*, \quad X(b) = B^*,$$

where A^* and B^* are given constants in \mathbf{R}^n .

PROPOSITION 2. Assume that \hat{X} is admissible, and

(1) H^* is $\mathbb{E} \times \mathbb{B}$ measurable, and there exist an integrable function $\rho^*(\cdot)$ on $[a, b]$ and a positive number δ^* such that, for all $(t, Z) \in N^*(\alpha^*, \delta^*)$

$$|H^*(t, Z)| \leq \rho^*(t),$$

(2) $H^*(t, \cdot)$ is locally Lipschitz on the set

$$\{Z = (X, P) : |X - \hat{X}(t)| < \alpha^*, |P - \hat{P}(t)| < \delta^*\},$$

(3) the arc $\hat{Z} = (\hat{X}, \hat{P})$ satisfies

$$J\dot{\hat{Z}}(t) \in \partial_Z H^*(t, \hat{Z}(t)) \quad a.e.,$$

where $\partial_Z H^*$ is the generalized gradient of $H^*(t, \cdot)$,

(4) $H^*(t, \cdot, \hat{P}(t))$ is locally concave around \hat{X} , and $H^*(t, X, \cdot)$ is convex for (t, X) in the set

$$\{(t, X) : t \in [a, b], |X - \hat{X}(t)| < \alpha^*\}.$$

Then $J^*(X)$ is defined (possibly $+\infty$) for X near \hat{X} , and \hat{X} provides a strong local minimum for (P^*) .

PROOF. Condition (1) implies that L^* is $\mathbb{E} \times \mathbb{B}$ measurable and that, for $|X - \hat{X}(t)| < \alpha^*$

$$L^*(t, X, V) \geq \langle \hat{P}(t), V \rangle - \rho^*(t).$$

Then, $J^*(X)$ is defined (possibly $+\infty$) for all arcs X in the set

$$\{X : [a, b] \rightarrow \mathbf{R}^n, |X(t) - \hat{X}(t)| < \alpha^*\}.$$

We have that $H^*(t, \cdot, \hat{P}(t))$ is locally concave on a γ -neighborhood of \hat{X} , where $\gamma^* < \alpha^*$, then define the following function

$$(4.1) \quad \bar{H}(t, X, P) = \begin{cases} H^*(t, X, P) & \text{if } |X - \hat{X}(t)| \leq \gamma^* \text{ and } P = \hat{P}(t), \\ +\infty & \text{if } |X - \hat{X}(t)| \leq \gamma^* \text{ and } P \neq \hat{P}(t), \\ -\infty & \text{if } |X - \hat{X}(t)| > \gamma^* \end{cases}$$

from $[a, b] \times \mathbf{R}^n \times \mathbf{R}^n$ to $[-\infty, +\infty]$.

The hypothesis (4) implies that \bar{H} is concave in X and convex in P . Consider the function

$$(4.2) \quad \bar{L}(t, X, V) = \sup\{\langle P, V \rangle - \bar{H}(t, X, P) : P \in \mathbf{R}^n\} \\ = \begin{cases} +\infty & \text{if } |X - \hat{X}(t)| > \gamma^*, \\ \langle \hat{P}(t), V \rangle - H^*(t, X, \hat{P}(t)) & \text{if } |X - \hat{X}(t)| \leq \gamma^*. \end{cases}$$

Condition (1) implies that $\int_a^b \bar{L}(t, X(t), \dot{X}(t)) dt$ is finite for all arcs X such that $|X(t) - \hat{X}(t)| \leq \gamma^*$, for all $t \in [a, b]$.

Let us now define the problem

$$(\bar{P}) \quad \text{minimize} \left\{ \int_a^b \bar{L}(t, X(t), \dot{X}(t)) dt : X(a) = A^*, X(b) = B^* \right\}.$$

From hypothesis (3) and the definition of \bar{H} it follows that

$$J\dot{Z}(t) \in \partial \bar{H}(t, \hat{Z}(t)) \quad \text{a.e.}$$

Also we have that \bar{H} is a proper and closed function, then by [14, Theorem 37.5] we conclude that

$$(\dot{\hat{P}}(t), \hat{P}(t)) \in \partial \bar{L}(t, \hat{X}(t), \dot{\hat{X}}(t)) \quad \text{a.e.}$$

So, the convexity of $\bar{L}(t, \cdot, \cdot)$ on $\mathbf{R}^n \times \mathbf{R}^n$ will then imply that, for any admissible arc X

$$(4.3) \quad \int_a^b \{ \bar{L}(t, X(t), \dot{X}(t)) - \bar{L}(t, \hat{X}(t), \dot{\hat{X}}(t)) \} dt \\ \geq \int_a^b \langle (\dot{\hat{P}}(t), \hat{P}(t)), (X(t) - \hat{X}(t), \dot{X}(t) - \dot{\hat{X}}(t)) \rangle dt = 0.$$

Thus, the arc \hat{X} solves (\bar{P}) .

On the other hand, hypothesis (3), the convexity of $H^*(t, X, \cdot)$ when X is near \hat{X} , and the definition of \bar{L} lead to the following equalities:

$$(4.4) \quad L^*(t, \hat{X}(t), \dot{\hat{X}}(t)) = \bar{L}(t, \hat{X}(t), \dot{\hat{X}}(t)) = \hat{P}(t) \cdot \dot{\hat{X}}(t) - H^*(t, \hat{X}(t), \hat{P}(t)).$$

Also, for any arc X such that $|X(t) - \hat{X}(t)| \leq \gamma^*$ for all $t \in [a, b]$ we have

$$(4.5) \quad \bar{L}(t, X(t), \dot{X}(t)) \leq L^*(t, X(t), \dot{X}(t)).$$

Therefore, \hat{X} provides a strong local minimum for (P^*) . Q.E.D.

The following result establishes a criterion for local concavity.

PROPOSITION 3. Suppose that for some positive number γ^* and for all t in $[a, b]$ the function $H^*(t, \cdot, \hat{P}(t))$ is C^{1+} on $N_t(\gamma^*) = \{X \in \mathbf{R}^n: |X - \hat{X}(t)| < \gamma^*\}$. Assume in addition that $H_{XX}^*(t, X, \hat{P}(t)) \leq 0$ whenever H_{XX}^* exists on $\{(t, X, \hat{P}(t)): t \in [a, b], |X - \hat{X}(t)| < \gamma^*\}$. Then $H^*(t, \cdot, \hat{P}(t))$ is locally concave around \hat{X} .

PROOF. Consider t_0 in $[a, b]$, X_0 in $N_{t_0}(\gamma^*)$, and d in \mathbf{R}^n ; $d \neq 0$. Then we can find real numbers T_1, T_2 such that

$$X_0 + \tau d \in N_{t_0}(\gamma^*) \quad \text{for } \tau \in (T_1, T_2),$$

and

$$X_0 + \tau d \notin N_{t_0}(\gamma^*) \quad \text{for } \tau \notin (T_1, T_2).$$

Define $g(\tau) = H^*(t_0, X_0 + \tau d, \hat{P}(t_0))$ as a function from (T_1, T_2) to \mathbf{R} . Since $H^*(t_0, \cdot, \hat{P}(t_0))$ is C^{1+} on $N_{t_0}(\gamma^*)$ then $g(\cdot)$ is C^{1+} on (T_1, T_2) , and hence by the chain rule in [6] we obtain

$$\dot{g}(\tau) = H_X^*(t_0, X_0 + \tau d, \hat{P}(t_0)) \cdot d$$

and

$$\partial \dot{g}(\tau) \subset d \cdot \partial_X H_X^*(t_0, X_0 + \tau d, \hat{P}(t_0))d,$$

where $\partial \dot{g}$ is the generalized gradient of \dot{g} and $\partial_X H_X$ is the generalized Jacobian of H_X . But from the hypothesis on H_{XX}^* we get that, for all $\tau \in (T_1, T_2)$, for all $v \in \partial_X H_X^*(t_0, X_0 + \tau d, \hat{P}(t_0))$, $d \cdot vd \leq 0$. Thus, $\dot{g}(\tau)$ is nonincreasing, and hence $g(\cdot)$ is concave on (T_1, T_2) . Since the argument is valid for any choice of t_0 in $[a, b]$, X_0 in $N_{t_0}(\gamma^*)$, and d in \mathbf{R}^n , we conclude that $H^*(t, \cdot, \hat{P}(t))$ is locally concave around \hat{X} . Q.E.D.

5. Proof of Theorem 1. The proof will follow these steps: Condition (c) of the theorem will allow us to find a canonical transformation transforming the original problem (P) to a new problem (P*) which has a locally concave-convex Hamiltonian H^* around the transformed arc \hat{X} of the given arc \hat{x} . Then, we will apply the sufficiency theory of the previous section to the new problem. Conditions (a) and (b) will imply the optimality of \hat{x} for the original problem, knowing that the arc \hat{X} solves the transformed problem (P*).

Suppose we are given C^1 -functions $h(\cdot)$, $F(\cdot, \cdot)$ and a positive number α ; $\alpha \leq \varepsilon$, such that

$$h: [a, b] \rightarrow \mathbf{R}^n, \quad F: \{(t, x): t \in [a, b], |x - \hat{x}(t)| < \alpha\} \rightarrow \mathbf{R}^n,$$

and $F(t, \cdot)$ is invertible with inverse $g(t, \cdot) = F^{-1}(t, \cdot)$. Consider the canonical transformation whose generating function has the form

$$(5.1) \quad \Psi(t, X, p) = -(p - h(t)) \cdot g(t, X).$$

This transformation transforms the original variables (x, p) and the original Hamiltonian H to new variables (X, P) and a new Hamiltonian H^* in the following way:

$$(5.2) \quad \begin{cases} x = -\Psi_p(t, X, p) = g(t, X), \\ P = -\Psi_X(t, X, p) = (g_X(t, X))^T (p - h(t)), \end{cases}$$

where A^T denotes the transpose of the matrix A ,

$$(5.3) \quad H^*(t, X, P) = H(t, x, p) + \dot{h}(t) \cdot g(t, X) - g_t(t, X) \cdot (p - h(t))$$

and

$$(5.4) \quad P \cdot dX - H^*(t, X, P) = p \cdot dx - H(t, x, p) - d(h(t) \cdot g(t, X)).$$

Let \hat{X}, \hat{P} be the transformed arcs of the given arcs \hat{x}, \hat{p} . Then \hat{X}, \hat{P} are defined as

$$\hat{X}(t) = F(t, \hat{x}(t)) \quad \text{and} \quad \hat{P}(t) = F_x^{T^{-1}}(t, \hat{x}(t))(\hat{p}(t) - h(t)).$$

Since the transformation is defined for $(t, x, p) \in N(\alpha, \infty)$, then there exists a positive number α^* such that the original variables $(t, x, p) \in N(\alpha, \infty)$ whenever the transformed variables $(t, X, P) \in N^*(\alpha^*, \infty)$.

From the definition of H we have that $H(t, x, \cdot)$ is convex. Then the equations (5.2) and (5.3) imply that $H^*(t, X, \cdot)$ is convex.

Let us define

$$(5.5) \quad L^*(t, X, V) = \sup\{\langle P, V \rangle - H^*(t, X, P) : P \in \mathbb{R}^n\},$$

then L^* is defined on $N^*(\alpha^*, \infty)$.

Now consider the transformed problem

$$(P^*) \quad \text{minimize } J^*(X) = \int_a^b L^*(t, X(t), \dot{X}(t)) dt$$

subject to

$$X(a) = F(a, A), \quad X(b) = F(b, B).$$

The following lemma shows the connection between the original problem (P) and the transformed problem (P*).

LEMMA 1. Assume the hypothesis (H_1) and

- (i) L satisfies the Weierstrass condition,
- (ii) $H(t, \cdot, \cdot)$ is locally Lipschitz near (\hat{x}, \hat{p}) , and $\hat{z} = (\hat{x}, \hat{p})$ satisfies the Hamiltonian inclusions

$$J\dot{\hat{z}}(t) \in \partial_z H(t, \hat{z}(t)) \quad a.e.$$

Then $J^*(X)$ is well defined for X near \hat{X} , and if \hat{X} provides a strong local minimum for (P*), \hat{x} provides a strong local minimum for (P).

PROOF. The hypothesis (H_1) implies that L^* is $\mathbb{L} \times \mathbb{B}$ measurable and that there exist a positive number δ^* and an integrable function $\rho^*(\cdot)$ on $[a, b]$ such that the new Hamiltonian H^* defined in (5.3) satisfies

$$|H^*(t, X, P)| \leq \rho^*(t) \quad \text{for } (t, X, P) \in N^*(\alpha^*, \delta^*).$$

Then, from (5.5) it follows that

$$L^*(t, X, V) \geq \langle \hat{P}(t), V \rangle - \rho^*(t),$$

and hence $J^*(X)$ is well defined (possibly $+\infty$) for X near \hat{X} .

Consider an admissible arc X for the problem (P^*) such that $|X(t) - \hat{X}(t)| < \alpha^*$ for all $t \in [a, b]$. Then $x(t) = g(t, X(t))$ is an admissible arc for (P) , and $|x(t) - \hat{x}(t)| < \alpha$ for all $t \in [a, b]$. Now by using equations (5.5), (5.3) and (2.1) we obtain

$$\begin{aligned} L^*(t, X(t), \dot{X}(t)) &= \sup \{ \langle P, \dot{X}(t) \rangle - H^*(t, X(t), P) : P \in \mathbb{R}^n \} \\ &= \sup_P \inf_v \left\{ \langle P, \dot{X}(t) \rangle - \left(h(t) + (g_X(t, X(t)))^{T^{-1}} P \right) \cdot v \right. \\ &\quad \left. + L(t, g(t, X(t)), v) + g_t(t, X(t)) \cdot (g_X(t, X(t)))^{T^{-1}} P \right\} \\ &\quad - \dot{h}(t) \cdot g(t, X(t)) \\ &\leq \inf_v \sup_P \left\{ \langle P, \dot{X}(t) \rangle - \left(h(t) + (g_X(t, X(t)))^{T^{-1}} P \right) \cdot v \right. \\ &\quad \left. + L(t, g(t, X(t)), v) + g_t(t, X(t)) \cdot (g_X(t, X(t)))^{T^{-1}} P \right\} \\ &\quad - \dot{h}(t) \cdot g(t, X(t)) \\ &= \inf_v K(v), \end{aligned}$$

where

$$K(v) = \begin{cases} +\infty & \text{if } v \neq g_t(t, X(t)) + g_X(t, X(t))\dot{X}(t), \\ L(t, g(t, X(t)), v) - h(t) \cdot v - \dot{h}(t) \cdot g(t, X(t)) & \\ & \text{if } v = g_t(t, X(t)) + g_X(t, X(t))\dot{X}(t). \end{cases}$$

But we have $x(t) = g(t, X(t))$, and $\dot{x}(t) = g_t(t, X(t)) + g_X(t, X(t))\dot{X}(t)$, then we conclude that

$$L^*(t, X(t), \dot{X}(t)) \leq L(t, x(t), \dot{x}(t)) - \frac{d}{dt} (h(t) \cdot x(t)),$$

and hence

$$(5.6) \quad \int_a^b L^*(t, X(t), \dot{X}(t)) dt \leq \int_a^b L(t, x(t), \dot{x}(t)) dt + h(a) \cdot A - h(b) \cdot B.$$

On the other hand, we know from [12] that canonical transformations conserve the “Hamiltonian inclusions”, then condition (ii) implies that the transformed arc $\hat{Z} = (\hat{X}, \hat{P})$ satisfies

$$J\dot{\hat{Z}}(t) \in \partial_z H^*(t, \hat{Z}(t)) \quad \text{a.e.}$$

Thus, the convexity of $H^*(t, X, \cdot)$ and equation (5.5) imply that

$$L^*(t, \hat{X}(t), \dot{\hat{X}}(t)) = \langle \hat{P}(t), \dot{\hat{X}}(t) \rangle - H^*(t, \hat{X}(t), \hat{P}(t)).$$

Also conditions (i), (ii), and equation (2.1) imply that

$$L(t, \hat{x}(t), \dot{\hat{x}}(t)) = \langle \hat{p}(t), \dot{\hat{x}}(t) \rangle - H(t, \hat{x}(t), \hat{p}(t)).$$

By using the equalities above in equation (5.4) we obtain

$$(5.7) \quad \int_a^b L^*(t, \hat{X}(t), \dot{\hat{X}}(t)) dt = \int_a^b L(t, \hat{x}(t), \dot{\hat{x}}(t)) dt + h(a) \cdot A - h(b) \cdot B.$$

Since \hat{X} provides a strong local minimum for (P^*) , we conclude from (5.6) and (5.7) that \hat{x} provides a strong local minimum for (P) . Q.E.D.

To prove the theorem it is sufficient, by Lemma 1, to prove that the transformed arc \hat{X} provides a strong local minimum for (P^*) . Proposition 2 will be used to prove the optimality of \hat{X} . In fact, given the hypotheses of the theorem all the conditions of Proposition 2 are satisfied except the local concavity of $H^*(t, \cdot, \hat{P}(t))$ around \hat{X} . Using hypothesis (c) of the theorem we are going to find particular functions $h(\cdot)$ and $F(\cdot, \cdot)$ such that the corresponding canonical transformation in (5.1) guarantees the local concavity of $H^*(t, \cdot, \hat{P}(t))$ around \hat{X} , and then Proposition 2 will complete the proof.

Consider the C^1 -matrix function $Q(t)$ given in the condition (c). Then we can find a C^1 -function \bar{P} from $[a, b]$ to \mathbf{R}^n , a positive number α , and a function $F(t, x)$ from $\{(t, x): t \in [a, b], |x - \hat{x}(t)| < \alpha\}$ to \mathbf{R}^n such that the following hold:

(1) $F(\cdot, x)$ is C^1 , $F_t(t, \cdot)$ is C^2 , $F(t, \cdot)$ is C^3 and invertible, $F_x(\cdot, x)$ is C^1 , $F_{xx_i}(\cdot, x)$ is C^1 for $i = 1, \dots, n$.

(2) $F_{xt}(t, \hat{x}(t)) = F_{tx}(t, \hat{x}(t))$, $F_{xx_i t}(t, \hat{x}(t)) = F_{tx x_i}(t, \hat{x}(t))$ for $i = 1, \dots, n$, and $F_x(t, \hat{x}(t)) = I$.

(3) For $g(t, X) = F^{-1}(t, x)$, and for $\hat{X}(t) = F(t, \hat{x}(t))$, the functions g and \bar{P} satisfy

$$(5.8) \quad Q(t) = D_X((g_X(t, X))^T \bar{P}(t))|_{X=\hat{X}(t)},$$

where D_X denotes the Jacobian operator with respect to X .

The construction of the functions F and \bar{P} is possible. First notice that condition (b) of the theorem implies that \hat{x} is C^1 . Let $n > 1$, and let $q_1(t)$ be the first column of $Q(t) = (q_1(t) \cdots q_n(t))$. Define $d(t, x) = \langle (x - \hat{x}(t)), q_1(t) \rangle + 1$. Then we can find a positive number α such that $d(t, x) \neq 0$ on the set $\{(t, x): t \in [a, b], |x - \hat{x}(t)| < \alpha\}$.

Define the functions \bar{P} and

$$F = \begin{pmatrix} F_1 \\ \vdots \\ F_n \end{pmatrix}$$

as

$$\bar{P}(t) \equiv -e_1 = \begin{pmatrix} -1 \\ 0 \\ \vdots \\ 0 \end{pmatrix},$$

$$F_1(t, x) = \frac{1}{2} \langle (x - \hat{x}(t)), Q(t)(x - \hat{x}(t)) \rangle + x_1,$$

$$F_i(t, x) = x_i \quad \text{for } i = 2, 3, \dots, n.$$

It is clear that \bar{P} and $F(\cdot, x)$ are C^1 , $F(t, \cdot)$ is C^3 , $F_t(t, \cdot)$ is C^2 . Also we have

$$F_x(t, x) = I + \begin{pmatrix} (x - \hat{x}(t))^T Q(t) \\ 0 & \cdots & 0 \\ \vdots & & \vdots \\ 0 & \cdots & 0 \end{pmatrix},$$

and $\det F_x(t, x) = d(t, x)$. Then $F(t, \cdot)$ is invertible for $|x - \hat{x}(t)| < \alpha$, $F_x(\cdot, x)$ is C^1 , $F_x(t, \hat{x}(t)) = I$,

$$F_{xt}(t, x) = F_{tx}(t, x) = \begin{pmatrix} (x - \hat{x}(t))^T \dot{Q}(t) - \dot{\hat{x}}(t)^T Q(t) \\ 0 & \cdots & 0 \\ \vdots & & \vdots \\ 0 & \cdots & 0 \end{pmatrix}$$

and

$$F_{xx_i t}(t, x) = F_{txx_i}(t, x) = \begin{pmatrix} \dot{q}_i^T(t) \\ 0 & \cdots & 0 \\ \vdots & & \vdots \\ 0 & \cdots & 0 \end{pmatrix}.$$

So conditions (1) and (2) hold.

Let $\hat{X}(t) = F(t, \hat{x}(t))$, and $g(t, X) = F^{-1}(t, X)$. It is easy to see that $g_x(t, \hat{X}(t)) = I$,

$$\left. \frac{\partial}{\partial X_i} g_x(t, X) \right|_{X=\hat{X}(t)} = - \left. \frac{\partial}{\partial X_i} F_x(t, x) \right|_{x=\hat{x}(t)} = \begin{pmatrix} -\dot{q}_i^T(t) \\ 0 & \cdots & 0 \\ \vdots & & \vdots \\ 0 & \cdots & 0 \end{pmatrix}.$$

So (5.8) holds.

For the case $n = 1$ the construction of F and \bar{P} is simpler. In fact, take $\bar{P}(t) = -Q(t)$, and $F(t, x)$ a function on $[a, b] \times R$ such that $F_x(t, \hat{x}(t)) = 1$, $F_x(t, x) \neq 0$, and $F_{xx}(t, \hat{x}(t)) = 1$. Then, the function $g(t, X) = F^{-1}(t, X)$ satisfies

$$g_{xx}(t, \hat{X}(t)) = -F_{xx}(t, \hat{x}(t)) = -1 \quad \text{and} \quad g_{xx}(t, \hat{X}(t))\bar{P}(t) = Q(t).$$

Having found functions F and \bar{P} satisfying conditions (1), (2) and (3) define $h(t) = \hat{p}(t) - \bar{P}(t)$. Condition (b) of the theorem implies that \hat{p} is C^1 , and hence h is C^1 . Consider the canonical transformation of the form (5.1) corresponding to the functions $h(\cdot)$ and $F(\cdot, \cdot)$ already constructed. For this transformation (5.2) implies that

$$\hat{P}(t) = (g_x(t, \hat{X}(t)))^T (\hat{p}(t) - h(t)) = \bar{P}(t).$$

The following result will be used to prove that the transformed Hamiltonian $H^*(t, \cdot, \hat{P}(t))$ is locally concave around \hat{X} .

LEMMA 2. If the condition (c) of the theorem holds then there exists a positive number β such that for all $t \in [a, b]$ and for $|x - \hat{x}(t)| < \beta$, $|p - \hat{p}(t)| < \beta$, $|E - I| < \beta$, $|W + Q(t)| < \beta$ and $|S + \dot{Q}(t)| < \beta$ we have

$$(E^T, W^T)D_z H_z(t, x, p) \begin{pmatrix} E \\ W \end{pmatrix} + S \leq 0 \quad (\text{negative semidefinite})$$

whenever $D_z H_z(t, x, p)$ exists.

PROOF. Suppose the result is false. Then for every positive integer k such that $k \geq \max(\varepsilon, \delta)$, where (ε, δ) is given in (H_2) , there exist $(t_k, x_k, p_k) \in [a, b] \times \mathbf{R}^n \times \mathbf{R}^n$, $r_k \in \mathbf{R}^n$ and $n \times n$ -matrices E_k, W_k, S_k satisfying

$$(t_k, x_k, p_k) \in N(\tfrac{1}{k}, \tfrac{1}{k}), \quad |E_k - I| < \tfrac{1}{k}, \quad |W_k + Q(t_k)| < \tfrac{1}{k}, \\ |S_k + \dot{Q}(t_k)| < \tfrac{1}{k}, \quad \|r_k\| = 1,$$

$D_z H_z(t_k, x_k, p_k)$ exists, and

$$r_k \cdot \left[(E_k^T, W_k^T)D_z H_z(t_k, x_k, p_k) \begin{pmatrix} E_k \\ W_k \end{pmatrix} + S_k \right] r_k > 0.$$

When $k \rightarrow +\infty$, then $t_k \rightarrow t \in [a, b]$, $x_k \rightarrow \hat{x}(t)$, $p_k \rightarrow \hat{p}(t)$, $E_k \rightarrow I$, $W_k \rightarrow -Q(t)$, $S_k \rightarrow -\dot{Q}(t)$, and $r_k \rightarrow r$ with $\|r\| = 1$. On the other hand the hypothesis (H_2) implies that there exists a positive constant M such that

$$|\partial_z H_z(t, z)| \leq M \quad \text{for } (t, z) \in N(\varepsilon, \delta).$$

Thus $D_z H_z(t_k, x_k, p_k)$ converges to some matrix v . But the map $(t, z) \rightarrow \partial_z H_z(t, z)$ is upper semicontinuous, then $v \in \partial_z H_z(t, \hat{x}(t), \hat{p}(t))$.

At the limit the above inequality becomes

$$r \left[(I, -Q(t))v \begin{pmatrix} I \\ -Q(t) \end{pmatrix} - \dot{Q}(t) \right] r \geq 0,$$

which contradicts the condition (c) of the theorem. Q.E.D.

To complete the proof of the theorem it remains to establish the local concavity of $H^*(t, \cdot, \hat{P}(t))$ around \hat{X} . For, it suffices by Proposition 3 to prove that there exists a positive number γ^* such that $H_{XX}^*(t, X, \hat{P}(t)) \leq 0$ whenever H_{XX}^* exists on $\{(t, X, \hat{P}(t)): t \in [a, b], |X - \hat{X}(t)| < \gamma^*\}$.

The hypothesis (H_2) and equation (5.3) imply that there exist positive numbers α^*, δ^* such that $H^*(t, \cdot, \cdot)$ is C^{1+} near (\hat{X}, \hat{P}) , i.e., $H_{XX}^*(t, X, P)$ exists almost everywhere on $N^*(\alpha^*, \delta^*)$, and then H_{XX}^* can be computed from (5.3).

We define the matrix function $G(t, X) = (g_X(t, X))^{T^{-1}}$. For a given vector P in \mathbf{R}^n , $D_X(GP)$ denotes the derivative with respect to X of the vector function $G(t, X)P$. Let $A(t, X)$ be any matrix function and v a vector in \mathbf{R}^n . We define the matrix

$$(D_X A(t, X))v = \left(\frac{\partial A(t, X)}{\partial X_1} v \cdots \frac{\partial A(t, X)}{\partial X_n} v \right).$$

When H_{XX}^* exists on $N^*(\alpha^*, \delta^*)$ it is then given by the following expression.

$$(5.9) \quad H_{XX}^*(t, X, P) = \left((g_X)^T, (D_X(GP))^T \right) D_z H_z \begin{pmatrix} g_X \\ D_X(GP) \end{pmatrix} \\ + \left(D_X(D_X(GP))^T \right) (H_p - g_t) + \left(D_X(g_X)^T \right) (H_x + \dot{h}) \\ - (D_X(GP))^T g_{tx} - (D_X(g_{tx})^T) (GP) - (g_{tx})^T D_X(GP),$$

where g, h, H are evaluated at $(t, X), t, (t, g(t, X), h(t) + G(t, X)P)$ respectively.

Since there exists a one-to-one correspondence between the original variables (x, p) and their transformed variables (X, P) , then H_{XX}^* in (5.9) can be written as a function $R(t, x, p)$ of the original variables by using equations (5.2).

Define $E(t, x) = g_X(t, X)$, $W(t, x, p) = D_X(G(t, X)P)$, and $S(t, x, p)$ is the last five terms in (5.9). We have $E(t, \hat{x}(t)) = g_X(t, \hat{X}(t)) = I$. Also, since $\bar{P} \equiv \hat{P}$ and $g_X(t, \hat{X}(t)) = I$, then equation (5.8) implies that $W(t, \hat{x}(t), \hat{p}(t)) = -Q(t)$. On the other hand by using hypothesis (b) of the theorem and equation (5.8) and that $\hat{x}(t) = g(t, \hat{X}(t))$, $g_X(t, \hat{X}(t)) = I$, we obtain

$$\left[D_X(D_X(GP))^T \right] (H_p - g_t) \big|_{X=\hat{x}(t), P=\hat{P}(t)} = Q(t) (D_X g_X(t, X)) \big|_{X=\hat{x}(t)} \dot{\hat{X}}(t) \\ + \left((D_X g_X(t, X)) \big|_{X=\hat{x}(t)} \dot{\hat{X}}(t) \right)^T Q(t) \\ - \left(D_X \left(D_X \left((g_X(t, X))^T P \right) \right) \big|_{X=\hat{x}(t), P=\hat{P}(t)} \dot{\hat{X}}(t) \right).$$

Since $h(t) = \hat{p}(t) - \hat{P}(t)$ and $(d/dt)(g_X(t, \hat{X}(t))) = 0$, then

$$S(t, \hat{x}(t), \hat{p}(t)) = - \left(D_X(g_X(t, X))^T \right) \big|_{X=\hat{x}(t)} \dot{\hat{P}}(t) \\ - \left(D_X(g_{X_t}(t, X))^T \right) \big|_{X=\hat{x}(t)} \hat{P}(t) \\ - \left(D_X \left(D_X \left((g_X(t, X))^T P \right) \right) \big|_{X=\hat{x}(t), P=\hat{P}(t)} \dot{\hat{X}}(t) \right).$$

By computing $\dot{Q}(t)$ from equation (5.8) we conclude that $S(t, \hat{x}(t), \hat{p}(t)) = -\dot{Q}(t)$.

The continuity of the functions $E(t, x), W(t, x, p), S(t, x, p)$, and Lemma 2 imply that $R(t, x, p) \leq 0$ for (x, p) near (\hat{x}, \hat{p}) and hence there exists a positive number γ^* such that $H_{XX}^*(t, X, \hat{P}(t)) \leq 0$ when it exists on $\{(t, X, \hat{P}(t)): t \in [a, b], |X - \hat{X}(t)| < \gamma^*\}$. Q.E.D.

6. General boundary constraints and objective functions. Consider the generalized problem of Bolza with general boundary constraints

$$(P_1) \quad \text{minimize } J_1(x) = l(x(a), x(b)) + \int_a^b L(t, x(t), \dot{x}(t)) dt$$

where x is an arc from $[a, b]$ to \mathbf{R}^n , and where L is as before;

$$L: N(\varepsilon, \infty) \rightarrow \mathbf{R} \cup \{+\infty\},$$

and

$$l: \mathbf{R}^n \times \mathbf{R}^n \rightarrow \mathbf{R} \cup \{+\infty\}.$$

The problem (P_1) is a more general form of the Bolza problem (P) of the earlier sections, where we considered $x(a), x(b)$ as being fixed. Specifically, the previous case is the one in which

$$l(x(a), x(b)) = \begin{cases} 0 & \text{if } x(a) = A \text{ and } x(b) = B, \\ +\infty & \text{if } x(a) \neq A \text{ or } x(b) \neq B. \end{cases}$$

This section presents sufficient conditions for the problem (P_1) based on the technique used for the problem (P) with some modifications resulting from the fact that (P_1) has a more general form. In fact, these conditions, as we shall see, are the ones given in Theorem 1 and a new condition involving, of course, the function l . In the case of fixed boundary values, i.e., when the problem (P_1) reduces to (P) , this condition is automatically satisfied.

DEFINITION. An arc x from $[a, b]$ to \mathbf{R}^n is said to be *admissible* for (P_1) if we have $l(x(a), x(b)) < +\infty$.

THEOREM 2. Suppose we are given arcs \hat{x}, \hat{p} from $[a, b]$ to \mathbf{R}^n with \hat{x} admissible. Assume that all the hypotheses of Theorem 1 are satisfied and in addition assume

(d) for $c \in \mathbf{R}^n, d \in \mathbf{R}^n$, with $|c| < \epsilon, |d| < \epsilon$, we have

$$(6.1) \quad l(\hat{x}(a) + c, \hat{x}(b) + d) - l(\hat{x}(a), \hat{x}(b)) \\ \geq \langle \hat{p}(a), c \rangle - \langle \hat{p}(b), d \rangle + \frac{1}{2} \langle d, Q(b)d \rangle - \frac{1}{2} \langle c, Q(a)c \rangle.$$

Then \hat{x} provides a strong local minimum for (P_1) .

PROOF OF THEOREM 2. The proof is similar to the proof of Theorem 1 with some modifications. So, consider the canonical transformation, described in §5 and given by (5.1). Define the transformed problem of (P_1) to be

$$(P_1^*) \quad \text{minimize } J_1^*(X) = \langle h(b), g(b, X(b)) \rangle - \langle h(a), g(a, X(a)) \rangle \\ + l(g(a, X(a)), g(b, X(b))) + \int_a^b L^*(t, X(t), \dot{X}(t)) dt,$$

where L^* is given by (5.5).

The following lemma will play here the role played by Lemma 1 for the case of given boundary values.

LEMMA 3. Assume all the hypotheses of Lemma 1. Then $J_1^*(X)$ is well defined near \hat{X} , and if \hat{X} is a strong local minimum for (P_1^*) , \hat{x} is a strong local minimum for (P_1) .

PROOF. The proof is similar to the proof of Lemma 1, where (5.6) is now

$$\int_a^b L^*(t, X(t), \dot{X}(t)) dt \leq \int_a^b L(t, x(t), \dot{x}(t)) dt \\ + \langle h(a), g(a, X(a)) \rangle - \langle h(b), g(b, X(b)) \rangle,$$

and hence

$$(6.2) \quad J_1^*(X) \leq J_1(x).$$

On the other hand equation (5.7) becomes

$$\int_a^b L^*(t, \hat{X}(t), \dot{\hat{X}}(t)) dt = \int_a^b L(t, \hat{x}(t), \dot{\hat{x}}(t)) dt \\ + \langle h(a), g(a, \hat{X}(a)) \rangle - \langle h(b), g(b, \hat{X}(b)) \rangle,$$

so that

$$(6.3) \quad J_1^*(\hat{X}) = J_1(\hat{x}).$$

By using (6.2) and (6.3) and the fact that \hat{X} is a strong local minimum for (P_1^*) , we then conclude the result. Q.E.D.

The proof of Theorem 2 will then be completed if we prove that the transformed arc \hat{X} provides a strong local minimum for (P_1^*) . For, let $Q(\cdot)$ be the matrix function satisfying condition (c) of Theorem 1. As in §5, choose functions F and \bar{P} satisfying conditions (1), (2) and (3) of §5, and

$$(4) \quad \langle F(t, x), \bar{P}(t) \rangle = -\frac{1}{2} \langle (x - \hat{x}(t)), Q(t)(x - \hat{x}(t)) \rangle + \langle \bar{P}(t), x \rangle.$$

This choice of F and \bar{P} is possible. In fact, the functions

$$\bar{P}(t) \equiv \begin{pmatrix} -1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad F(t, x) = \begin{pmatrix} \frac{1}{2} \langle x - \hat{x}(t), Q(t)(x - \hat{x}(t)) \rangle + x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix},$$

satisfy the conditions required. Then define, as in §5,

$$(6.4) \quad h(t) = \hat{p}(t) - \bar{P}(t).$$

Now, consider the canonical transformation of the form (5.1), where $g(t, \cdot) = F^1(t, \cdot)$, and h is given in (6.4). For this transformation equation (5.2) implies that the arcs \hat{x} , \hat{p} are transformed to the arcs \hat{X} , \hat{P} , given by

$$\hat{X}(t) = F(t, \hat{x}(t)) \quad \text{and} \quad \hat{P}(t) = \hat{p}(t) - h(t) = \bar{P}(t).$$

From §5 we have that $H^*(t, X, P)$, defined in (5.3), is convex in P and locally concave in X around the transformed arc \hat{X} . Let \bar{H} , \bar{L} be the functions defined by (4.1) and (4.2) respectively, and define

$$(6.5) \quad \bar{l}(X_1, X_2) = \begin{cases} +\infty & \text{if } |X_1 - \hat{X}(a)| > \gamma^*, \text{ or } |X_2 - \hat{X}(b)| > \gamma^*, \\ l^*(X_1, X_2) & \text{otherwise,} \end{cases}$$

where $l^*(X_1, X_2) = \langle h(b), g(b, X_2) \rangle - \langle h(a), g(a, X_1) \rangle + l(g(a, X_1), g(b, X_2))$.

Consider then the problem

$$(\bar{P}_1) \quad \text{minimize} \left\{ \bar{l}(X(a), X(b)) + \int_a^b \bar{L}(t, X(t), \dot{X}(t)) dt \right\}.$$

Let X be any arc from $[a, b]$ to \mathbf{R}^n . From (4.3) we derive

$$(6.6) \quad \int_a^b \{ \bar{L}(t, X(t), \dot{X}(t)) - \bar{L}(t, \hat{X}(t), \dot{\hat{X}}(t)) \} dt \\ \geq \int_a^b \langle \dot{\hat{P}}(t), \hat{P}(t) \rangle, (X(t) - \hat{X}(t), \dot{X}(t) - \dot{\hat{X}}(t)) \rangle dt \\ = \langle \hat{P}(b), X(b) - \hat{X}(b) \rangle - \langle \hat{P}(a), X(a) - \hat{X}(a) \rangle.$$

If $|X(a) - \hat{X}(a)| > \gamma^*$ or $|X(b) - \hat{X}(b)| > \gamma^*$, then (6.5) gives

$$\begin{aligned} \bar{l}(X(a), X(b)) - \bar{l}(\hat{X}(a), \hat{X}(b)) \\ \geq \langle \hat{P}(a), X(a) - \hat{X}(a) \rangle - \langle \hat{P}(b), X(b) - \hat{X}(b) \rangle. \end{aligned}$$

But, if $|X(a) - \hat{X}(a)| \leq Y^*$ and $|X(b) - \hat{X}(b)| \leq Y^*$, (6.5) implies that, for $x(t) = g(t, X(t))$ we have

$$\begin{aligned} \bar{l}(X(a), X(b)) - \bar{l}(\hat{X}(a), \hat{X}(b)) &= \langle h(b), x(b) - \hat{x}(b) \rangle - \langle h(a), x(a) - \hat{x}(a) \rangle \\ &\quad + l(x(a), x(b)) - l(\hat{x}(a), \hat{x}(b)). \end{aligned}$$

By using (6.1), condition (4), and the fact that $\hat{P} \equiv \bar{P} = \hat{p} - h$ we obtain

$$\begin{aligned} \bar{l}(X(a), X(b)) - \bar{l}(\hat{X}(a), \hat{X}(b)) &\geq \langle \bar{P}(a), F(a, x(a)) \rangle - \langle \bar{P}(b), F(b, x(b)) \rangle \\ &\quad - \langle \bar{P}(a), \hat{x}(a) \rangle + \langle \bar{P}(b), \hat{x}(b) \rangle. \end{aligned}$$

But, condition (4) gives that $\langle \bar{P}(t), \hat{x}(t) \rangle = \langle \bar{P}(t), F(t, \hat{x}(t)) \rangle$, and hence, since $\bar{P} \equiv \hat{P}$ we conclude

$$\begin{aligned} (6.7) \quad \bar{l}(X(a), X(b)) - \bar{l}(\hat{X}(a), \hat{X}(b)) \\ \geq \langle \hat{P}(a), X(a) - \hat{X}(a) \rangle - \langle \hat{P}(b), X(b) - \hat{X}(b) \rangle. \end{aligned}$$

Thus, for any arc X from $[a, b]$ to \mathbf{R}^n the function \bar{l} satisfies (6.7). Adding the inequalities (6.6) and (6.7) we obtain that \hat{X} solves (\bar{P}_1) .

We also have that (4.4) holds, and (4.5) is satisfied for all arcs X such that $|X(t) - \hat{X}(t)| \leq \gamma^*$. Thus, it follows that \hat{X} provides a strong local minimum for (P_1^*) . Q.E.D.

7. Application to an optimal control problem. Suppose we are given a subset U of \mathbf{R}^m , an arc \hat{x} from $[a, b]$ to \mathbf{R}^n , two constants A, B , a positive number ε and functions g, f such that

$$g: \{(x, u): |x - \hat{x}(t)| < \varepsilon, u \in U\} \rightarrow \mathbf{R},$$

and

$$f: \{x \in \mathbf{R}^n: |x - \hat{x}(t)| < \varepsilon\} \rightarrow M_{n \times m},$$

where $M_{n \times m}$ is the space of $n \times m$ -matrices.

Consider the autonomous control problem:

$$(C) \quad \text{Minimize } J(x, u) = \int_a^b g(x(t), u(t)) dt$$

over all measurable functions $u: [a, b] \rightarrow \mathbf{R}^m$ and absolutely continuous functions $x: [a, b] \rightarrow \mathbf{R}^n$ satisfying the constraints

$$(7.1) \quad \dot{x}(t) = f(x(t))u(t) \quad \text{a.e.},$$

$$(7.2) \quad u(t) \in U \quad \text{a.e.},$$

$$(7.3) \quad x(a) = A, \quad x(b) = B.$$

Suppose that x is an arc satisfying $|x(t) - \hat{x}(t)| < \varepsilon$ for all $t \in [a, b]$, and u is a measurable function from $[a, b]$ to \mathbf{R}^m . If (x, u) satisfies (7.1), (7.2) and (7.3) we say that (x, u) is *admissible*.

DEFINITION. An admissible function (\hat{x}, \hat{u}) is a *strong local minimum* for (C) if there exists a positive real number γ such that (\hat{x}, \hat{u}) minimizes $J(x, u)$ over all admissible functions (x, u) satisfying

$$|x(t) - \hat{x}(t)| < \gamma \quad \text{for all } t \in [a, b].$$

The Hamiltonian $H(x, p)$ of the problem (C) is defined on $\{(x, p): |x - \hat{x}(t)| < \epsilon, p \in \mathbb{R}^n\}$ by

$$(7.4) \quad H(x, p) = \sup\{\langle p, f(x)u \rangle - g(x, u) : u \in U\}.$$

If the supremum in (7.4) is attained for each (x, p) at a unique point u in U , then we write the solution $u(x, p)$ as a function of (x, p) .

THEOREM 3. Assume that \hat{x} satisfies (7.3), and

- (1) U is a nonempty convex compact polyhedron in \mathbb{R}^m ,
- (2) g and f are C^2 , $g(x, \cdot)$ is strictly convex for x near \hat{x} , and there exists an arc \hat{p} such that, for $\hat{u}(t) = u(\hat{x}(t), \hat{p}(t))$ we have $g_{uu}(\hat{x}(t), \hat{u}(t)) > 0$ for $t \in [a, b]$,
- (3) the arc $\hat{z} = (\hat{x}, \hat{p})$ satisfies

$$J\dot{\hat{z}}(t) = H_z(\hat{z}(t)) \quad \text{for } t \in [a, b],$$

- (4) the condition (c) of Theorem 1 is satisfied for H given in (7.4).

Then the function (\hat{x}, \hat{u}) is admissible and provides a strong local minimum for the problem (C).

PROOF. The compactness of U , the hypothesis (2), and [1, Theorem 2.1] imply that for each (x, p) the supremum in (7.4) is attained at a unique point $u(x, p)$, the function $u(\cdot, \cdot)$ is continuous, and $H(\cdot, \cdot)$ is C^1 with gradient given by

$$(7.5) \quad H_z(z) = ((Df(x))u(x, p))^T p - g_x(x, u(x, p), f(x)u(x, p)).$$

Since \hat{x} satisfies (7.3) and since condition (3) holds, then (\hat{x}, \hat{u}) is admissible for the problem (C).

The proof of optimality of the function (\hat{x}, \hat{u}) is based on converting the optimal control problem (C) to a generalized Bolza problem (P_C) , and then applying Theorem 1.

Define the function

$$(7.6) \quad L(x, v) = \inf\{g(x, u) : v = f(x)u \text{ and } u \in U\}.$$

Then L is an extended real-valued function on $\{(x, v) : |x - \hat{x}(t)| < \epsilon, v \in \mathbb{R}^n\}$.

The Hamiltonian \tilde{H} corresponding to the function L is exactly the Hamiltonian H defined in (7.4). In fact, equations (2.1) and (7.6) give

$$\begin{aligned} \tilde{H}(x, P) &= \sup\{\langle p, v \rangle - L(x, v) : v \in \mathbb{R}^n\} \\ &= \sup\{\langle p, v \rangle - \inf\{g(x, u) : v = f(x)u, u \in U\} : v \in \mathbb{R}^n\} \\ &= \sup\{\langle p, f(x)u \rangle - g(x, u) : u \in U\} \\ &= H(x, p). \end{aligned}$$

Define the generalized Bolza problem corresponding to the problem (C) by

$$(P_C) \quad \text{minimize } \int_a^b L(x(t), \dot{x}(t)) dt$$

subject to

$$x(a) = A, \quad x(b) = B,$$

where L is the function defined in (7.6).

Since U is compact, and f, g are C^2 , then by [17, Lemma 6 and Theorem 6] we conclude that L defined in (7.6) is measurable, and

$$\min(P_C) = \min(C).$$

Thus, to prove that (\hat{x}, \hat{u}) is a strong local minimum for (C) it suffices to prove that \hat{x} is a strong local minimum for the problem (P_C) .

The compactness of U , and the measurability of L imply that the hypothesis (H_1) holds. The hypothesis (H_2) in this case reduces to saying that $H(\cdot)$ is C^{1+} near $\hat{z} = (\hat{x}, \hat{p})$. Assume for the moment that H is C^{1+} (we will prove it later). The arc \hat{x} is admissible for the generalized Bolza problem (P_C) , since it satisfies (7.3). Also the convexity of $g(x, \cdot)$ imply the convexity of $L(x, \cdot)$ defined in (7.6). So, to apply Theorem 1 to the problem (P_C) , we only need to show that $H(\cdot)$ is C^{1+} near $\hat{z} = (\hat{x}, \hat{p})$, and hence the proof of Theorem 3 will be completed.

We will use the following notation:

$$N_\alpha(\hat{z}) = \{z \in \mathbf{R}^n \times \mathbf{R}^n: |z - \hat{z}(t)| < \alpha \text{ for } t \in [a, b]\}.$$

From hypothesis (2) we have that

$$g_{uu}(\hat{x}(t), \hat{u}(t)) > 0 \quad \text{for } t \in [a, b].$$

Since we also have that the functions \hat{x}, \hat{p} , and $u(\cdot, \cdot)$ are continuous, and $[a, b]$ is compact, it follows that there exists a positive number α such that for $z = (x, p) \in N_\alpha(\hat{z})$ we have

$$g_{uu}(x, u(z)) > 0.$$

Let $\lambda_1(z), \dots, \lambda_m(z)$ be the eigenvalues of the matrix function $g_{uu}(x, u(z))$. Define $\lambda(z) = \min_i \lambda_i(z)$. Since $g_{uu}(x, u(z))$ is positive definite we conclude that for $z \in N_\alpha(\hat{z})$, and for $h \in \mathbf{R}^n$,

$$h \cdot g_{uu}(x, u(z))h \geq \lambda(z)\|h\|^2.$$

By using the continuity of $\lambda_i(\cdot)$, and then of $\lambda(\cdot)$, it follows that there exist positive numbers γ and μ such that for $z \in N_\gamma(\hat{z})$, and for $h \in \mathbf{R}^n$,

$$(7.7) \quad h \cdot g_{uu}(x, u(z))h \geq \mu\|h\|^2.$$

Since U is a compact polyhedron, f, g are C^2 , and (7.7) holds then by [13, Theorem 4.2 and Corollary 4.3] we have the following.

There exists a positive constant λ such that, for each $y_0 \in N_\gamma(\hat{z})$ there exists a neighborhood $N(y_0)$ of y_0 with

$$(7.8) \quad |u(z) - u(y_0)| \leq \lambda|z - y_0| \quad \text{for } z \in N(y_0).$$

This proves that $u(\cdot)$ is uniformly point-Lipschitz. The point-Lipschitz property is quite different from the Lipschitz property. However, as we shall see, the uniformly point-Lipschitz condition will imply that $u(\cdot)$ is Lipschitz.

So, it remains to show that for $v, w \in N_\gamma(\hat{z})$ we have

$$|u(v) - u(w)| \leq \lambda |v - w|.$$

For, let $v, w \in N_\gamma(\hat{z})$, and let $d = (w - v)/|w - v|$. Define

$$[v, w] = \{z = v + td : 0 \leq t \leq |v - w|\}.$$

So, $[v, w] \subset N_\gamma(\hat{z})$. Then, for each $y_0 \in [v, w]$ we can find a neighborhood $N(y_0)$ such that, for $z \in N(y_0)$, (7.8) holds. Since $[v, w]$ is compact then there exist elements $z_0 = v, z_1, \dots, z_{r+1} = w$ in $[v, w]$, and neighborhoods $N(z_0), \dots, N(z_{r+1})$ of z_0, \dots, z_{r+1} such that $\bigcup_{i=0}^{r+1} N(z_i) \supset [v, w]$, and for each $y \in N(z_i)$ we have

$$|u(y) - u(z_i)| \leq \lambda |y - z_i|.$$

Choose elements v_0, v_1, \dots, v_r in $[v, w]$ such that, for $i = 0, \dots, r$, $v_i \in N(z_i) \cap N(z_{i+1})$. Thus, we have

$$\begin{aligned} |u(v) - u(w)| &\leq |u(v) - u(v_0)| + |u(v_0) - u(v_1)| + \dots + |u(v_r) - u(w)| \\ &\leq \lambda(|v - v_0| + |v_0 - v_1| + \dots + |v_r - w|) \\ &= \lambda |v - w|. \end{aligned}$$

Therefore, $H(\cdot)$ is C^{1+} on $N_\gamma(\hat{z})$. Q.E.D.

REMARK. For the class of optimal control problems considered in this section we already showed that H is C^{1+} . However, for a large class of problems (see Example 2), H is not necessarily C^2 . This fact demonstrates the utility of considering H to be C^{1+} which is less restrictive than C^2 .

REMARK. In the special case when $\hat{u}(t) \in \text{int } U$, the polyhedral condition on U , which was needed to prove that $H(\cdot)$ is C^{1+} , could be omitted, since $H(\cdot)$ would be C^2 . This special case, in which it is required that the control \hat{u} take values in the interior of U , has also been considered by Mayne [11, Theorem 3.2].

8. Examples. This section consists of two numerical examples.

EXAMPLE 1. Consider the classical variational problem

$$(P) \quad \text{minimize } \int_0^{\pi/8} \left(\dot{x}^2(t) - x^2(t) + \frac{1}{x^2(t)} \right) dt$$

subject to

$$x(0) = 1, \quad x\left(\frac{\pi}{8}\right) = \sqrt{2}.$$

The Hamiltonian of the problem is

$$H(x, p) = \frac{p^2}{4} + x^2 - \frac{1}{x^2}$$

So, $H(\cdot, \cdot)$ and $L(\cdot, \cdot)$ are C^2 for $x \neq 0$, and

$$(8.1) \quad L_{vv}(x, v) = 2 > 0.$$

The Hamiltonian equations of the problem are

$$(8.2) \quad \begin{cases} -\dot{\hat{p}}(t) = H_x(\hat{z}(t)) = 2\hat{x}(t) + \frac{2}{\hat{x}^3(t)}, \\ \dot{\hat{x}}(t) = \hat{H}_p(\hat{z}(t)) = \frac{\hat{p}(t)}{2}. \end{cases}$$

Hence, by solving (8.2), we obtain

$$(8.3) \quad \begin{cases} \hat{x}(t) = \sqrt{\sqrt{2} \sin 2t + 1}, \\ \hat{p}(t) = \frac{2\sqrt{2} \cos 2t}{\sqrt{\sqrt{2} \sin 2t + 1}} \end{cases}$$

with \hat{x} admissible for (P).

To be able to apply the classical sufficient conditions using the Jacobi condition we need to solve the Jacobi equation (3.4). In fact the Jacobi equation for this problem is

$$(8.4) \quad \ddot{h}(t) + \left(1 - \frac{3}{(\sqrt{2} \sin 2t + 1)^2}\right) h = 0.$$

It does not seem simple to find a solution for (8.4). So, we cannot check whether or not we have conjugate point of 0 in $(0, \frac{\pi}{8}]$, and hence, we are not able to use the sufficiency theorem existing in the literature and involving the Jacobi equation to check the optimality of \hat{x} .

But, we are going to show that condition (c) of Theorem 1 is easily satisfied and then the optimality of \hat{x} will be obtained.

For this problem, equation (3.1) is

$$(8.5) \quad \dot{Q}(t) - \frac{1}{2} Q^2(t) - 2 + \frac{6}{(\sqrt{2} \sin 2t + 1)^2} > 0.$$

Take for $Q(\cdot)$ the following function

$$Q_0(t) = t.$$

Since for $t \in [0, \frac{\pi}{8}]$ we have $\sqrt{2} \sin 2t + 1 \leq 2$, then (8.5) is satisfied by $Q_0(\cdot)$. So condition (c) holds. By using (8.1) and (8.2), we conclude from Theorem 1 that \hat{x} , given in (8.3), provides a strong local minimum for (P).

REMARK. For Example 1 condition (c) of Theorem 1 is equivalent, by the Corollary of §3, to the Jacobi condition. However, as we have seen, while the Jacobi condition was hard to check, condition (c) was possible to verify.

EXAMPLE 2. Consider the optimal control problem

$$(C) \quad \text{minimize } \int_0^1 \left(2u_1^2 + 2u_2^2 - \frac{x^3}{24} \right) dt$$

subject to

$$\begin{aligned}\dot{x}(t) &= x(t)(u_1(t) + u_2(t)), \\ (u_1(t), u_2(t)) &\in U = [-1, 0] \times [-1, 0] \quad \text{for } t \in [0, 1], \\ x(0) &= x(1) = 1.\end{aligned}$$

From (7.4) the Hamiltonian H of this problem is

$$\begin{aligned}(8.6) \quad H(x, p) &= \sup \left\{ px(u_1 + u_2) - 2u_1^2 - 2u_2^2 + \frac{x^3}{24} : (u_1, u_2) \in U \right\} \\ &= \begin{cases} p^2x^2/4 + x^3/24 & \text{if } px \in [-4, 0], \\ x^3/24 & \text{if } px \geq 0, \\ -2px - 4 + x^3/24 & \text{if } px \leq -4. \end{cases}\end{aligned}$$

So, $H(\cdot, \cdot)$ is C^1 with gradient H_z given by

$$(8.7) \quad H_z(x, p) = \begin{cases} (p^2x/2 + x^2/8, px^2/2) & \text{for } px \in [-4, 0], \\ (x^2/8, 0) & \text{for } px \geq 0, \\ (-2p + x^2/8, -2x) & \text{for } px \leq -4. \end{cases}$$

Take $\hat{x}(t) \equiv 1, \hat{p}(t) = (1 - t)/8$. Then we have that for $\hat{z} = (\hat{x}, \hat{p})$

$$J\hat{z}(t) = H_z(\hat{z}(t)) \quad \text{for } t \in [0, 1].$$

We also have that $g_{uu}(x, u) = \begin{bmatrix} 4 & 0 \\ 0 & 4 \end{bmatrix} > 0$, and $U = [-1, 0] \times [-1, 0]$ is a convex compact polyhedron.

From (8.6), it follows that $\hat{u}(t) = u(\hat{x}(t), \hat{p}(t)) \equiv (0, 0)$. Equation (8.7) implies that

$$H_{xx}(\hat{x}(t), \hat{p}(t)) = \frac{1}{4}, \quad H_{xp}(\hat{x}(t), \hat{p}(t)) = H_{px}(\hat{x}(t), \hat{p}(t)) \equiv 0,$$

and

$$H_{pp}(\hat{x}(t), \hat{p}(t)) = \begin{cases} 0 & \text{for } t \neq 1, \\ \text{does not exist} & \text{for } t = 1. \end{cases}$$

Then $H(\cdot)$ is not C^2 near $\hat{z} = (\hat{x}, \hat{p})$. But, since U is a compact polyhedron and $g_{uu} > 0$, then from the proof of Theorem 3 we have that $H(\cdot)$ is C^{1+} near \hat{z} .

To apply Theorem 3 it remains to check condition (4) which is condition (c) of Theorem 1.

From [6, §15] we have the generalized Jacobian inclusion

$$\partial_z H_z(z) \subset A(z) = \begin{pmatrix} \partial_x H_x(z) & \partial_p H_x(z) \\ \partial_x H_p(z) & \partial_p H_p(z) \end{pmatrix}.$$

So, to check condition (c) for the elements of $\partial_z H_z(\hat{z}(t))$ it suffices to check it for the elements of the bigger set $A(\hat{z}(t))$.

From (8.7) it follows that $\partial_p H_p(\hat{z}(1)) = [0, \frac{1}{2}]$, and hence

$$A(\hat{z}(t)) = \begin{pmatrix} \frac{1}{4} & 0 \\ 0 & \partial_p H_p(\hat{z}(t)) \end{pmatrix},$$

where

$$\partial_p H_p(\hat{z}(t)) = \begin{cases} 0 & \text{for } t \neq 1, \\ [0, \frac{1}{2}] & \text{for } t = 1. \end{cases}$$

Take $Q_0(t) = t$. Then for $(\frac{\alpha}{\beta}, \frac{\delta}{\gamma}) \in A(\hat{z}(t))$ we have $\dot{Q}_0(t) - \gamma Q_0^2(t) + Q_0(t)\beta + \delta Q_0(t) - \alpha = \frac{3}{4} - \gamma t^2$. But, we have $\gamma \in [0, \frac{1}{2}]$. Then

$$\frac{3}{4} - \gamma t^2 \geq \frac{1}{4} > 0 \quad \text{for } t \in [0, 1].$$

Thus, condition (4) of Theorem 3 holds, and hence (\hat{x}, \hat{u}) provides a strong local minimum for (C).

REMARK. In Example 2 the control function $\hat{u} \equiv 0$ is on the boundary of the control set. So, the sufficiency theorem of [11] cannot be applied to our problem.

REMARK. Using §5 with some minor changes, it is not hard to show that Theorem 1, and hence Theorems 2 and 3, remain valid if the function Q of condition (c) is only Lipschitz. In this case, the matrix function $\dot{Q}(t)$ is replaced in (2.2) by the generalized Jacobian $\partial Q(t)$ of the function $Q(t)$.

ACKNOWLEDGEMENT. It is my pleasant task to acknowledge with gratitude Professor F. H. Clarke for discussion of the results and his many valuable suggestions and encouragement.

REFERENCES

1. F. H. Clarke, *Generalized gradients and applications*, Trans. Amer. Math. Soc. **205** (1975), 247–262.
2. ———, *Admissible relaxation in variational and control problems*, J. Math. Anal. Appl. **51** (1975), 557–576.
3. ———, *L'Hamiltonien en optimisation*, Mimeo, Univ. British Columbia, Vancouver, B.C., 1975.
4. ———, *Necessary conditions for a general control problem*, Proc. Sympos. Calculus of Variations and Optimal Control, D. L. Russell, Ed., Mathematics Research Center, Univ. of Wisconsin-Madison, Academic Press, New York, 1976.
5. ———, *The generalized problem of Bolza*, SIAM J. Control Optim. **14** (1976), 682–699.
6. ———, *Generalized gradients of Lipschitz functionals*, Adv. in Math. **40** (1981), pp. 52–67.
7. ———, *Extremal arcs and extended Hamiltonian systems*, Trans. Amer. Math. Soc. **231** (1977), 349–367.
8. I. M. Gelfand and S. V. Fomin, *Calculus of variations*, (translated by R. Silverman), Prentice-Hall, Englewood Cliffs, N.J., 1963.
9. P. Hartman, *Ordinary differential equations*, Wiley, New York, 1964.
10. M. R. Hestenes, *Calculus of variations and optimal control theory*, Wiley, New York, 1966.
11. D. Q. Mayne, *Sufficient conditions for a control to be a strong minimum*, J. Optim. Theory Appl. **21** (1977), 339–351.
12. D. C. Offin, *A Hamilton-Jacobi approach to the differential inclusion problem*, Thesis, Univ. British Columbia, Vancouver, B.C., 1979.
13. S. M. Robinson, *Generalized equations and their solutions. II. Applications to nonlinear programming*, Univ. of Wisconsin-Madison, Madison, 1980 (Preprint).
14. R. T. Rockafellar, *Convex analysis*, Princeton Univ. Press, Princeton, N.J., 1970.
15. ———, *Conjugate convex functions in optimal control and the calculus of variations*, J. Math. Anal. Appl. **32** (1970), 174–222.

16. ———, *Generalized Hamiltonian equations for convex problems of Lagrange*, Pacific J. Math. **33** (1970), 411–427.
17. ———, *Optimal arcs and the minimum value function in problems of Lagrange*, Trans. Amer. Math. Soc. **180** (1973), 53–83.
18. ———, *Existence theorems for general control problems of Bolza and Lagrange*, Adv. in Math. **15** (1975), 312–333.
19. A. Seierstad and K. Sydsaeter, *Sufficient conditions in optimal control theory*, Internat. Econom. Rev. **18** (1977), 367–391.

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF BRITISH COLUMBIA, VANCOUVER, BRITISH COLUMBIA, CANADA

Current address: Centre de Recherche des Mathématiques Appliquées, Université de Montréal, Montréal, Québec, Canada