EXTREMAL ARCS AND EXTENDED
HAMILTONIAN SYSTEMS([1])

BY

FRANK H. CLARKE

ABSTRACT. A general variational problem is considered; it involves the minimization of an integrand L of a very general nature. The Lagrangian L is allowed to assume the value $+\infty$, and need satisfy no differentiability or convexity conditions. A Hamiltonian corresponding to the problem is defined via the conjugate function of convex analysis, and it is shown how one obtains necessary conditions in the form of an extended Hamiltonian system. This system is expressed in terms of certain "generalized gradients" previously developed by the author.

A further result is given which has the feature that the principal hypotheses required, as well as the ensuing conclusions, are entirely in terms of H. This allows the treatment of classes of problems in which H is more amenable to direct analysis than L. The approach also sheds light on the relation between existence theory and the theory of necessary conditions, since the results may easily be compared with R. T. Rockafellar's recent work on existence theory, in which H also plays a central role.

As an example of its application the main result is specialized to a differential inclusion problem. A specific example of its use is also given, an unorthodox optimal control problem with a discontinuous cost functional.

1. **Introduction.** The problem we consider is that of minimizing the functional

$$(1.1) \qquad l(x(0), x(1)) + \int_0^1 L(t, x(t), \dot{x}(t)) \, dt$$

over the absolutely continuous functions $x: [0, 1] \to R^n$ (we call such an x an *arc*), where

$$l: R^n \times R^n \to (-\infty, \infty], \qquad L: [0, 1] \times R^n \times R^n \to (-\infty, \infty].$$

The fact that the value $+\infty$ is permitted makes this *generalized problem of Bolza* quite versatile by allowing various constraints to be implicitly incorporated. For example, optimal control problems may be placed within the above

framework, as we shall see. More complete discussions may be found in [5] and [10].

In the calculus of variations, an optimal arc z solving the above problem in a classical setting satisfies the Euler-Lagrange equation

$$(d/dt)\nabla_v L(t, z(t), \dot{z}(t)) = \nabla_s L(t, z(t), \dot{z}(t)) \quad \text{a.e.,}$$

where L is thought of as a function of (t, s, v). This may be rephrased as follows: there exists an arc p such that

$$(1.2) \qquad\qquad (\dot{p}, p) = \nabla L(t, z, \dot{z}) \quad \text{a.e.,}$$

where the gradient is taken with respect to (s, v), and where the variable t has been suppressed to disencumber the notation. If for each t and s the Legendre transform of $L(t, s, \cdot)$ exists and is denoted $H(t, s, \cdot)$, then (1.2) is equivalent to the Hamiltonian system

$$(1.3) \qquad\qquad (-\dot{p}, \dot{z}) = \nabla H(t, z, p) \quad \text{a.e.,}$$

where H is a function of (t, s, p) and the gradient is with respect to (s, p).

R. T. Rockafellar [10], [11] obtained the following type of necessary condition for the generalized problem of Bolza in which l and $L(t, \cdot, \cdot)$ are assumed to be convex functions (not necessarily finite or differentiable): there exists an arc p such that

$$(1.4) \qquad\qquad (\dot{p}, p) \in \partial L(t, z, \dot{z}) \quad \text{a.e.,}$$

where the right-hand side is the set of subgradients at $(z, \dot{z})$ of the convex function $L(t, \cdot, \cdot)$ (see [9]). If $L(t, \cdot, \cdot)$ is differentiable at $(z, \dot{z})$, then (1.4) and (1.2) coincide. Furthermore, he showed that (1.4) is equivalent to

$$(1.5) \qquad\qquad (-\dot{p}, \dot{z}) \in \partial H(t, z, p) \quad \text{a.e.,}$$

where this time $H(t, s, \cdot)$ is the conjugate of $L(t, s, \cdot)$ in the sense of convex analysis, a transformation generalizing that of Legendre (the function $H(t, \cdot, \cdot)$ here is concave-convex). We note as before the similarity to (1.3).

In [5] the author developed necessary conditions for the generalized problem of Bolza, assuming neither differentiability nor convexity. These were stated in terms of certain *generalized gradients*. Because the symbol ∂ was chosen for these generalized gradients, the condition obtained is visually indistinguishable from (1.4). It was shown that these results subsume the classical and convex cases and have wide applicability.

It is natural to seek as well a counterpart to (1.5) in this generalized setting, which is the subject of this paper. Theorem 1 gives these necessary conditions as a consequence of hypotheses made directly on L, while Corollary 1 has the

interesting feature that both the hypotheses and the conclusion are in terms of the function H. These results sometimes allow us to analyze without much difficulty a problem in which L is "bad" but the associated Hamiltonian H is "nice"; an example involving a discontinuous cost functional in an optimal control problem is given in §2.

A further interesting facet of the Hamiltonian conditions is their relation to some recent work of Rockafellar [12]. He obtains conditions in terms of the Hamiltonian which assure the existence of minimizing arcs for the functional (1.1). As mentioned, our necessary conditions may be phrased entirely in terms of H as well. This common point of view sheds new light on the relation between existence theory and that of necessary conditions. We may also go a step further and couple the two theories, thereby obtaining a result (couched in terms of H) that simultaneously affirms the existence of a minimum and gives the necessary conditions which the minimizing arc must satisfy (see Corollary 2).

The plan of the paper is as follows. §2 presents and discusses the main results. §3 is devoted to subsidiary lemmas that are at the base of the main results, whose proofs are in §4.

2. The main results. We shall be dealing with the functions l and L of §1, and we shall assume throughout the lower-semicontinuity of l and, for each t, $L(t, \cdot, \cdot)$. The following unrestrictive concept of optimality will be used: the arc z is said to *solve* the generalized problem of Bolza (i.e. that of minimizing (1.1)) if the integral in (1.1) for $x = z$ is defined and finite, if $l(z(0), z(1))$ is finite, and if for any arc x such that $l(x(0), x(1))$ is finite and the integral in (1.1) is defined (finitely or otherwise), we have

$$l(x(0), x(1)) + \int_0^1 L(t, x, \dot{x})\, dt \geq l(z(0), z(1)) + \int_0^1 L(t, z, \dot{z})\, dt.$$

The above does not rule out the possibility that the integral equals $-\infty$ for some arc x. However, if there exists a solution z, this could only happen for an arc x such that $l(x(0), x(1))$ is $+\infty$.

For a given s in R^n and δ in $(0, \infty]$ we define

$$\Phi_\delta^0(s) = \inf\left\{ l(x(0) + s, x(1)) + \int_0^1 L(t, x, \dot{x})\, dt : \|x - z\| < \delta \right\},$$

where the infimum is over the class of arcs and $\|x\|$ is the supremum norm of x on $[0, 1]$. With the substitution of $l(x(0), x(1) + s)$, we define similarly $\Phi_\delta^1(s)$. Consistent with the above-mentioned use of the word "solve", the functional in this definition is assigned the value $+\infty$ whenever the integral is not defined, or whenever l is $+\infty$.

We say the problem of Bolza is *calm* at z if, for some δ in $(0, \infty]$ and for

$i = 0$ or 1, we have

$$(2.1) \qquad \liminf_{s \to 0} [\Phi_\delta^i(s) - \Phi_\delta^i(0)]/|s| > -\infty.$$

Some simple criteria for calmness are given in [5]. We define the multifunction $E: [0,1] \times R^n \to R^{n+1}$ by

$$E(t,s) = \text{epi } L(t,s,\cdot) = \{(v,r) \in R^n \times R: L(t,s,v) \leqslant r\}.$$

Let ε be a positive number. L is said to be *epi-measurable* (in t) and *epi-Lipschitz* (in s) within ε of the arc z if E is Lebesgue measurable in t and Lipschitz in s (in the Hausdorff metric) for s within ε of $z(t)$. Letting $k(t)$ be the Lipschitz constant for each t, we impose further that $k(t)$ be integrable. The reader will find more detailed discussions of the above concepts in [5], including alternate definitions and special cases.

The *Hamiltonian* $H: [0,1] \times R^n \times R^n \to [-\infty, \infty]$ is defined by

$$(2.2) \qquad H(t,s,p) = \sup\{p \cdot v - L(t,s,v): v \text{ in } R^n\},$$

where $p \cdot v$ denotes the usual inner product. We shall say that H satisfies the *growth condition* within ε of the arc z if there exist a measurable function $\alpha(t)$ satisfying $|\alpha(t) - z(t)| < \varepsilon$ a.e., and a function $\beta(t,p)$ integrable in t for each p, such that for each t in [0, 1] and p in R^n,

$$(2.3) \qquad H(t, \alpha(t), p) \leqslant \beta(t, p).$$

It is worth pointing out that the necessary conditions obtained below are intrinsic to the problem in a way that is not valid for the usual multiplier rules, Pontryagin maximum principle, etc. This is because of the invariance of H with respect to different ways of posing the same problem or expressing constraints. For example, consider (as we do in Corollary 2) an optimal control problem

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

in which we seek to minimize a given function of $x(1)$ subject to endpoint constraints. This problem is equivalent to any similar one with (say) $\tilde{f}$ and $\bar{U}$, if $\tilde{f}(t,s,\bar{U}) = f(t,s,U)$. Yet any necessary condition in which f and its derivatives intervene is dependent on the particular representation of the problem.

THEOREM 1. *Let z solve the generalized problem of Bolza, which we assume calm at z. Suppose there exists a positive ε such that within ε of z, L is epi-measurable and epi-Lipschitz and H satisfies the growth condition. Then there exists an arc p such that:*

(2.4) $(-\dot{p}(t), \dot{z}(t)) \in \partial H(t, z(t), p(t))$ a.e.,

(2.5) $H(t, z(t), p(t)) = p(t) \cdot \dot{z}(t) - L(t, z(t), \dot{z}(t))$ a.e.,

(2.6) $(p(0), -p(1)) \in \partial l(z(0), z(1))$.

REMARKS. (a) ∂H signifies here the generalized gradient of the function $H(t, \cdot, \cdot)$; we shall not have occasion to consider derivatives with respect to t. The notation ∂l likewise refers to the generalized gradient of l. We must refer the reader to [3] for the relevant facts concerning this concept. We note only the following: when H is C^1, ∂H reduces to the gradient of H, and thus (2.4) reduces to the classical (1.3); when H is concave-convex (as in the case treated by Rockafellar),(2.4) reduces to (1.5), in which ∂H refers to the superdifferential-subdifferential of convex analysis. Since H is Lipschitz in s and p (unlike l), the following characterization of its generalized gradient holds [3, Definition 1.1]: for a Lipschitz function f, we have $\partial f(x) = \text{co}\{\lim_{i \to \infty} \nabla f(x_i)\}$ where we consider all sequences $\{x_i\}$ converging to x such that $\nabla f(x_i)$ and the indicated limit exist.

(b) Relation (2.6) is a "transversality condition" and is quite general. For example, if it were required to minimize the integral of L over the arcs satisfying $x(0)$ in C_0, $x(1)$ in C_1, where C_0 and C_1 are given closed sets, the function l defined by

$$l(x_0, x_1) = 0 \quad \text{if } (x_0, x_1) \text{ lies in } C_0 \times C_1,$$

$$= +\infty \quad \text{otherwise}$$

would take account of these constraints. Relation (2.6) would then yield: $p(0)$ is normal to C_0 at $z(0)$, $-p(1)$ is normal to C_1 at $z(1)$ (the use of the word "normal" is in a generalized sense [3]).

(c) Equation (2.5) is reminiscent of the maximum principle of optimal control theory; it affirms that the supremum in (2.2) is attained for $v = \dot{z}(t)$.

(d) The epi-Lipschitz hypothesis rules out "state constraints". That is, for each s near $z(t)$ there must be at least one v such that $L(t, s, v)$ is finite.

We shall now state a result that postulates, for each t and s, the convexity of the function $L(t, s, \cdot)$. We shall call this the *convexity condition*; its importance is well recognized in existence theory (see [12] and the references given there). There are general conditions under which an arc z solving an original problem continues to solve the "relaxed" version of the problem (see [4] and the references given there). In our present setting, the relaxed problem is the one in which $L(t, s, \cdot)$ is convexified (replaced by its convex hull [9]). Conditions assuring that this convexification preserves the optimality of z are given in [5, Theorem 2].

It is easy to see from (2.2) that L and its convexification as above give rise to the same function H. When L satisfies the convexity condition, we may

recover L from H via the formula

$$(2.7) \qquad L(t,s,v) = \sup\{\, p \cdot v - H(t,s,p) \colon p \text{ in } R^n \}.$$

We shall say that H satisfies the *strong Lipschitz condition* (in s) within ε of z if for some integrable function $k(t)$, for all t and p, and for all s_1 and s_2 within ε of $z(t)$ we have

$$(2.8) \qquad |H(t,s_1,p) - H(t,s_2,p)| \leqslant k(t)(1 + |p|)|s_1 - s_2|.$$

COROLLARY 1. *Let z solve the generalized problem of Bolza, where the problem is calm at z and L satisfies the convexity condition. Suppose that H satisfies the growth and strong Lipschitz conditions within ε of z, and that H is Lebesgue measurable in t. Then there exists an arc p such that (2.4)–(2.6) hold.*

Consider now the following problem: to minimize $l(x(0), x(1))$ over the arcs x satisfying the differential inclusion

$$(2.9) \qquad \dot{x}(t) \in E(t, x(t)) \quad \text{a.e.,}$$

where $E \colon [0,1] \times R^n \to R^n$ is a given multifunction. We note that this problem is a special case of the generalized problem of Bolza, obtained by setting $L(t,s,v)$ equal to 0 if v lies in $E(t,s)$ and equal to $+\infty$ otherwise. Formula (2.2) shows that in this case, $H(t,s,\cdot)$ is the support function of the set $E(t,s)$, i.e.

$$H(t,s,p) = \sup\{\, p \cdot v \colon v \text{ lies in } E(t,s)\}.$$

Condition (2.8) is then equivalent to E being Lipschitz in s, (if E is convex-valued) and H is measurable in t if E is.

Various kinds of results may be deduced from Theorem 1, depending on whether or not one wishes to hypothesize calmness and/or the convexity condition, couch the results in terms of L or H, or include existence in the conclusions. We limit ourselves here to one example, for the above differential inclusion problem. Problems of optimal control involving an equation such as $\dot{x}(t) = f(t, x(t), u(t))$ are equivalent to differential inclusion problems if we set $E(t,s) = f(t,s,U(t))$. However, there exist "nice" multifunctions E which have no such representation, and, hence, differential inclusion problems cannot be reduced to usual optimal control problems.

COROLLARY 2. *We assume that E is convex-valued, and that the set*

$$S = \{s_0 \in R^n \colon \text{ for some } s_1 \text{ in } R^n, l(s_0, s_1) < \infty\}$$

is bounded. Suppose there is a set W in R^n such that all solutions x of (2.9) with $x(0)$ in S lie in W, and that on a neighborhood of W, H satisfies the growth and

strong Lipschitz conditions and is measurable in t. Then if the set of admissible trajectories is nonempty, a solution z to the minimization problem exists. Furthermore, there is an arc p and a constant λ equal to 0 or 1 such that $|p(t)| + \lambda$ is not identically zero, p satisfies (2.4),

$$(2.10) \qquad\qquad H(t, z(t), p(t)) = p(t) \cdot \dot{z}(t) \quad a.e.$$

and

$(2.11) \quad (p(0), -p(1), -\lambda)$ *is normal to* epi *l at* $(z(0), z(1), l(z(0), z(1)))$.

REMARKS. Note that we have not assumed calmness, but have had to sacrifice the precision of the transversality condition (2.6). If λ equals 1, (2.6) and (2.11) are the same statement. Note that S will be bounded if l incorporates a constraint of the form $x(0) \in C$ for C bounded. The strong Lipschitz condition and the existence of W may be assured by supposing that H is locally Lipschitz and that for some integrable function $k(t)$, $E(t, s)$ is contained in the ball about 0 of radius $k(t)(1 + |s|)$.

It is in this case of a differential inclusion that this paper may most easily be compared with other results. Fedorenko [6] and Boltjanskiĭ [2] have obtained necessary conditions for a similar problem under strong regularity assumptions. Halkin [7] has studied the case in which E satisfies the strong hypothesis of "biconvexity"; Berliocchi and Lasry [1] use this same case as a step to more general problems. Warga [14] has treated the case where $E(t, s) = \text{co } f(t, s, U(t))$, f being Lipschitz in s and continuous in u; his results are couched in terms of "derivate containers" for f.

We conclude this section with an illustrative example: to minimize $cx(1) + \int_0^1 \{-|x(t)| + g(\dot{x}(t))\} dt$ subject to $x(0) = x_0$, where c in R and x_0 in R are given, and where g is the function defined by

$$\begin{aligned} g(v) &= +\infty & &\text{if } v < -1, \\ &= 0 & &\text{if } -1 \leqslant v \leqslant 0, \\ &= (1 + v^2)/10 & &\text{if } v > 0. \end{aligned}$$

This may be looked upon as an optimal control problem in which the control u equals $\dot{x}$, u is constrained to $[-1, \infty)$, and the cost functional is discontinuous at $u = 0$. (One might say that a velocity in $[-1, 0]$ is free, whereas any positive velocity entails a minimum rate of expense $+1/10$.) Note that the cost functional fails to be convex in x, and is neither differentiable nor continuous.

The above problem is a generalized problem of Bolza with $L(t, s, v) = -|s| + g(v)$ and $l(s_0, s_1) = cs_1$ if $s_0 = x_0$, and $+\infty$ otherwise. We deduce from [5, Propositions 3, 7] that L is epi-Lipschitz and epi-measurable, and from [5, Proposition 1] that the problem is calm. We calculate H via (2.2):

$$H(t,s,p) = |s| - p \qquad\qquad \text{if } p < 0,$$
$$= |s| \qquad\qquad\qquad \text{if } 0 \leqslant p \leqslant \tfrac{1}{5},$$
$$= |s| + 5p^2/2 - 1/10 \quad \text{if } p > \tfrac{1}{5}.$$

Upon applying Theorem 1, we find that if z is optimal there exists an arc p satisfying:

$$\dot{p}(t) = +1 \qquad\quad \text{if } z(t) < 0,$$
$$= -1 \qquad\qquad \text{if } z(t) > 0,$$
$$\in [-1, +1] \quad\;\; \text{if } z(t) = 0,$$

and

$$\dot{z}(t) = -1 \qquad\quad \text{if } p(t) < 0,$$
$$\in [-1, 0] \qquad \text{if } p(t) = 0,$$
$$= 0 \qquad\qquad \text{if } 0 < p(t) < \tfrac{1}{5},$$
$$\in [0, 1] \qquad\;\; \text{if } p(t) = \tfrac{1}{5},$$
$$= 5p(t) \qquad\; \text{if } p(t) > \tfrac{1}{5}.$$

Relation (2.6) implies $p(1) = -c$. One deduces from the above that $(z(t), p(t))$ moves along one of an infinite family of closed curves composed as follows (see Figure 1):
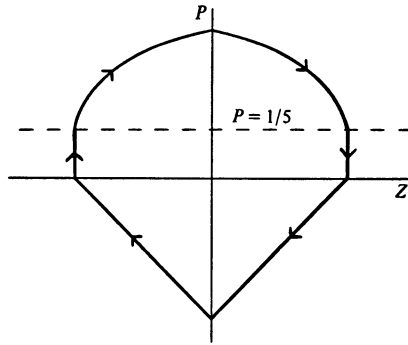

FIGURE 1

   (i) for $p > \tfrac{1}{5}$, parabolas of the form $2z = 5p^2 + K$ or $2z = -5p^2 + K$, according as z is negative or positive;
   (ii) for p in $(0, \tfrac{1}{5})$, the vertical line $z = K$;
   (iii) for $p < 0$, the lines $p = z + K$ or $p = -z + K$, according as z is positive or negative.
   Furthermore no "pause" occurs, and $|\dot{p}|$ equals 1 a.e. There is one exception to these rules: if ever z is 0 and p lies in $[0, \tfrac{1}{5}]$, then this is true for all t (note

then that for z identically zero, any arc p in $[0, \frac{1}{5}]$ with $|\dot{p}| \leqslant 1$ provides a solution to the Hamiltonian system.

Once given c and x_0, it is an easy matter to find the optimal arc from among the small number of possibilities. Suppose for example we have $x_0 = 1$ and $c = \frac{3}{5}$. We must find (z, p) moving along a curve as in Figure 1 according to the Hamiltonian equations, so that (z, p) begins at $t = 0$ on the vertical line $z = 1$, and ends at $t = 1$ on the horizontal line $p = -\frac{3}{5}$. We find that there is only one possibility: $p(0) = \frac{2}{5}$, and

$$
\begin{aligned}
\dot{z}(t) &= 2 - 5t & &\text{if } 0 < t < \tfrac{1}{5}, \\
&= 0 & &\text{if } \tfrac{1}{5} < t < \tfrac{2}{5}, \\
&= -1 & &\text{if } \tfrac{2}{5} < t < 1.
\end{aligned}
$$

We could, in fact, show that this arc z is the unique solution.

3. Preliminary lemmas. We shall say a closed subset C of R^n is *regular* if for each s in C the cone $N_C(s)$ of generalized normals to C at s (see [3]) consists of the nonnegative multiples of some given vector. We consider a multifunction E from R^n to R^n such that for each s, $E(s)$ is a closed, convex and regular set. We assume further that E is Lipschitz in the Hausdorff metric, and we set $\rho(s, v)$ equal to the Euclidean distance from v to $E(s)$. It follows from [4, Proposition 1] that ρ is a Lipschitz function. We define the function h as follows:

$$
h(s, p) = \sup\{ p \cdot \zeta \colon \zeta \text{ in } E(s) \}
$$

(where s and p are in R^n), and we remark that h is locally Lipschitz where it is finite. In the following, K is a positive number.

LEMMA 1. *If $|p| \leqslant K$, then $h(s, p) = \sup\{ p \cdot \zeta - K\rho(s, \zeta) \colon \zeta \text{ in } R^n \}$.*

PROOF. Upon writing $-\rho(s, \zeta)$ as $\sup\{-|\zeta - v| \colon v \text{ in } E(s)\}$, we note that the right-hand side above may be expressed as follows:

$$
\sup\{\sup\{ p \cdot \zeta - K|\zeta - v| \colon \zeta \text{ in } R^n \}\},
$$

where the first supremum is taken over all v in $E(s)$. It follows from the inequality $p \cdot \zeta - p \cdot v \leqslant K|\zeta - v|$ for all ζ that the expression in question reduces to $\sup\{ p \cdot v \colon v \text{ in } E(s) \}$, which is precisely $h(s, p)$.   Q.E.D.

LEMMA 2. *Let s, v, p and q in R^n satisfy $(q, p) = K\nabla\rho(s, v)$. Then*
(i) $h(s, p) = p \cdot v - K\rho(s, v)$,
(ii) *for any (α, β) in $R^n \times R^n$, we have*

$$\limsup_{\lambda \downarrow 0} \left[\frac{h(s + \lambda\alpha, p + \lambda\beta) - h(s,p)}{\lambda} \right] \geqslant (\alpha, \beta) \cdot (-q, v) - |\beta| \rho(s, v).$$

PROOF. The vector p is the gradient at v of the convex function $K\rho(s, \cdot)$, whence for all ζ in R^n, $K\rho(s, \zeta) - K\rho(s, v) \geqslant (\zeta - v) \cdot p$. Equivalently,

$$\sup\{ p \cdot \zeta - K\rho(s, \zeta): \zeta \text{ in } R^n \} = p \cdot v - K\rho(s, v).$$

Since the function $K\rho(s, \cdot)$ is Lipschitz with constant K, we have $|p| \leqslant K$, and (i) follows from Lemma 1.

Let us set K_λ equal to $|p + \lambda\beta|$. By Lemma 1 we have $h(s + \lambda\alpha, p + \lambda\beta) \geqslant (p + \lambda\beta) \cdot v - K_\lambda \rho(s + \lambda\alpha, v)$. We combine this with (i) to deduce

$$h(s + \lambda\alpha, p + \lambda\beta) - h(s, p) \geqslant \lambda\beta \cdot v - K_\lambda \rho(s + \lambda\alpha, v) + K\rho(s, v).$$

We may write this last expression in the form

$$\lambda\beta \cdot v - [K\rho(s + \lambda\alpha, v) - K\rho(s, v)] + (K - K_\lambda)\rho(s + \lambda\alpha, v).$$

We divide both sides by λ, we use the inequality $K - K_\lambda \geqslant -\lambda|\beta|$, and upon taking the lim sup we arrive at (ii).   Q.E.D.

LEMMA 3. *Let h be finite near (i.e. in a neighborhood of) (s, p), for p different from 0, and suppose (q, v) belongs to $\partial h(s, p)$. Then for any positive ε, (q, v) belongs to $\partial h(s, \varepsilon p)$.*

PROOF. Since $h(s, \varepsilon p') = \varepsilon h(s, p')$, we see that h is finite (and Lipschitz) near (s, p) and near $(s, \varepsilon p)$. In view of [3, Definition 1.1], we see that $\partial h(s, p)$ is "generated" by points of the form $[\nabla_s h(s', p'), \nabla_p h(s', p')]$. The above is identical to $[\nabla_s h(s', \varepsilon p')/\varepsilon, \nabla_p h(s', \varepsilon p')]$, and the conclusion follows.   Q.E.D.

LEMMA 4. *Let h be finite near (s, p) for p different from 0, and suppose (q, p) belongs to $\partial K\rho(s, v)$, where v lies in $E(s)$. Then $(-q, v)$ belongs to $\partial h(s, p)$.*

PROOF. As a preliminary step, suppose that (q_0, p_0) is the limit of points (q_i, p_i), where h is finite near (s, p_0), $(q_i, p_i) = \nabla K\rho(s_i, v_i)$ and (s_i, v_i) converges to (s, v). Let (α, β) in $R^n \times R^n$ be given. By Lemma 2 we may find for each i a positive λ_i in $(0, 1/i)$ such that

$$[h(s_i + \lambda_i\alpha, p_i + \lambda_i\beta) - h(s_i, p_i)]/\lambda_i \geqslant (\alpha, \beta) \cdot (-q_i, v_i) - |\beta|\rho(s_i, v_i) - 1/i.$$

It is clear from this that for any (α, β) we have

$$\limsup[h(s' + \lambda\alpha, p' + \lambda\beta) - h(s', p')]/\lambda \geqslant (\alpha, \beta) \cdot (-q_0, v),$$

where the lim sup is as (s', p') tends to (s, p_0) and λ decreases to 0. Since h is

Lipschitz near (s, p_0), it follows from [3, Proposition 1.4] that $(-q_0, v)$ belongs to $\partial h(s, p_0)$.

If (s_i, v_i) converges to (s, v), where v_i is not in $E(s_i)$, if $\nabla \rho(s_i, v_i)$ and $\lim \nabla \rho(s_i, v_i)$ exist, then $\nabla_v \rho(s_i, v_i)$ is a unit vector for each i [3, Proposition 2.4] and converges to the unique unit vector ζ that generates $N_{E(s)}(v)$ (this uses the fact that $E(s)$ is regular). When v_i lies in $E(s_i)$, $\nabla \rho(s_i, v_i)$ must be the zero vector if it exists (since ρ attains a minimum at (s_i, v_i)). We conclude [3, Definition 1.1] that $\partial K \rho(s, v)$ is the convex hull of the point $(0, 0)$ and of all points of the form

$$(3.1) \qquad\qquad \lim K \nabla \rho(s_i, v_i),$$

where v_i lies outside $E(s_i)$ and (s_i, v_i) converges to (s, v). It follows that the point (q, p) may be expressed as follows: $(q, p) = \sum \lambda_j (q_j, K\zeta)$, where each point $(q_j, K\zeta)$, $0 \leqslant j \leqslant 2n$, is of the form (3.1), and where the positive numbers λ_j sum to t in $(0, 1]$. We note that p equals $tK\zeta$, and that since h is positively homogeneous in p, h is finite near $(s, K\zeta)$.

We apply the preceding argument to deduce that for each j, $(-q_j, v) \in \partial h(s, K\zeta)$, and since $\{\lambda_j / t\}$ is a convex combination, we derive $(-\sum \lambda_j q_j / t, v) \in \partial h(s, K\zeta)$, and hence by Lemma 3, $(-q, v) \in \partial h(s, tK\zeta) = \partial h(s, p)$. Q.E.D.

LEMMA 5. *For some $\delta > 0$, let the arc z solve the problem of minimizing $f(x(1))$ subject to*

$$x(0) \in C_0, \quad \|x - z\| < \delta, \quad \dot{x}(t) \in E(t, x(t)) \quad a.e.,$$

where C_0 is a closed set and f is Lipschitz. We suppose that E is closed, convex, and regular-valued, that $E(t, s)$ is measurable in t, and that $E(t, \cdot)$ is Lipschitz (for each t) on the set $|s - z(t)| < \delta$, the Lipschitz constant $k(t)$ being integrable. Then there exist a positive number K and an arc p such that:

(a) *$(\dot{p}(t), p(t)) \in \partial K\rho(t, z(t), \dot{z}(t))$ a.e.,*
(b) *$p(0)$ is normal to C_0 at $z(0)$,*
(c) *$-p(1) \in \partial f(z(1))$.*

PROOF. An elementary proof by contradiction [4, Lemma 2] shows that for some positive K and δ', z solves the problem of minimizing

$$f(x(1)) + \int_0^1 K\rho(t, x, \dot{x}) \, dt$$

subject to $x(0) \in C_0$, $\|x - z\| < \delta'$. We apply [5, Theorem 1] to deduce the existence of an arc p satisfying (a)–(c).   Q.E.D.

REMARK. If we can proceed to deduce that a.e. $h(t, \cdot, \cdot)$ is finite near

$(z(t), p(t))$, then by Lemma 4 and (a) above we have

$$(-\dot{p}(t), \dot{z}(t)) \in \partial h(t, z(t), p(t)) \quad \text{a.e.}$$

We now suppose that L and H are as in §2 and we define, for a fixed t that we suppress henceforth,

$$E(s) = \text{epi } L(t, s, \cdot).$$

LEMMA 6. *Let E be Lipschitz in the Hausdorff metric for $|s - s_0| < \varepsilon$ with constant k, and suppose H is finite. Then H satisfies (2.8) for all p and for all such s. Conversely, if H satisfies (2.8) for all p and for all s within ε of s_0, and if L satisfies the convexity condition, then E is Lipschitz for $|s - s_0| < \varepsilon$ with constant $2k$.*

PROOF. Let E be Lipschitz as described, and set $h(s, p, q)$ equal to $\sup\{(p, q) \cdot w : w \text{ in } E(s)\}$, for (p, q) in $R^n \times R$. Then $h(s, p, -1) = H(s, p)$ and h is locally Lipschitz in s within ε of s_0 (with constant $k(|p| + |q|)$) for $q < 0$. Let any p be given, and any s_1, s_2 within ε of s_0. We define (p_0, q_0) by

$$p_0 = p/(1 + |p|), \quad q_0 = -1 + |p|/(1 + |p|)$$

and we note the equality $p = p_0/|q_0|$. We have, since $|p_0| + |q_0| = 1$,

$$|h(s_1, p_0, q_0) - h(s_2, p_0, q_0)| \leqslant k|s_1 - s_2|,$$

or equivalently,

$$|h(s_1, p_0/|q_0|, -1) - h(s_2, p_0/|q_0|, -1)| \leqslant k|s_1 - s_2|/|q_0|.$$

This is (2.8), since $1/|q_0| = 1 + |p|$.

For the converse, it suffices to show that $\rho(s_2, w_1) \leqslant 2k|s_1 - s_2|$ whenever s_1, s_2 within ε of s_0 and w_1 in $E(s_1)$ are given. The point w_1 is of the form $(v_1, L(s_1, v_1) + \delta_1)$ for $\delta_1 \geqslant 0$. If the above inequality were false, since E is convex-valued, there would be a unit vector (p, q) and a positive δ such that, for all w_2 in $E(s_2)$,

$$(3.2) \qquad (p, q) \cdot (w_1 - w_2) \geqslant 2k|s_1 - s_2| + \delta.$$

Necessarily then q is negative or zero. Suppose first that q is less than zero; then (3.2) yields, for all v_2 in R^n,

$$(3.3) \quad \bar{p} \cdot v_1 - L(s_1, v_1) \geqslant \bar{p} \cdot v_2 - L(s_2, v_2) + (2k|s_1 - s_2| + \delta)/|q|$$

where $\bar{p} = p/|q|$. Since $2 > |q| + |p|$, it follows that $2/|q|$ is no less than $1 + |\bar{p}|$, and hence (3.3) implies

$$H(s_1, \bar{p}) > H(s_2, \bar{p}) + k(1 + |\bar{p}|)|s_1 - s_2|,$$

a contradiction.

Suppose now that q were zero. Since $L(s_2, v_2) \geqslant -H(s_2, 0)$ for all v_2, we may choose a λ in $(0, 1)$ such that $\lambda L(s_2, v_2) > -\delta/3$ for all v_2 and such that $\lambda L(s_1, v_1) < \delta/3$. Thus for all v_2,

$$p \cdot v_1 - \lambda L(s_1, v_1) \geqslant p \cdot v_2 - \lambda L(s_2, v_2) + 2k|s_1 - s_2| + \delta/3,$$

and we obtain

$$H(s_1, p/\lambda) - H(s_2, p/\lambda) \geqslant 2k|s_1 - s_2|/\lambda + \delta/(3\lambda)$$
$$> k(1 + |p|/\lambda)|s_1 - s_2|,$$

which is contrary to the hypothesis.   Q.E.D.

B denotes the unit ball centered at 0.

LEMMA 7. *We assume the conditions of Theorem* 1, *and we let any positive number r be given. Then there exist integrable functions* $\alpha_1(t)$ *and* $\alpha_2(t)$ *such that, for all p in rB, for all s within* ε *of* $z(t)$, *we have* $\alpha_1(t) \leqslant H(t, s, p) \leqslant \alpha_2(t)$.

PROOF. For any s and p as above we have, by Lemma 6,

$$H(t, s, p) \geqslant H(t, z(t), p) - k(t)(1 + |p|)|s - z(t)|$$
$$\geqslant p \cdot \dot{z}(t) - L(t, z, \dot{z}) - k(t)(1 + r)\varepsilon$$
$$\geqslant -r|\dot{z}(t)| - L(t, z, \dot{z}) - k(t)(1 + r)\varepsilon = \alpha_1(t).$$

Now let $\{p_i\}$ be a finite set whose convex hull contains rB, and set $c = \max|p_i|$. By the growth condition, for each i, $H(t, \alpha(t), p_i) \leqslant \beta(t, p_i)$, so

$$H(t, s, p_i) \leqslant \beta(t, p_i) + k(t)(1 + |p_i|)|s - \alpha(t)|$$
$$\leqslant \beta(t, p_i) + k(t)(1 + c)2\varepsilon.$$

Put $\alpha_2(t) = \max \beta(t, p_i) + k(t)(1 + c)2\varepsilon$. Then any p in rB may be written as a convex combination of the points p_i: $p = \sum \lambda_i p_i$, and hence by convexity,

$$H(t, s, p) \leqslant \sum \lambda_i H(t, s, p_i) \leqslant \sum \lambda_i \alpha_2(t) = \alpha_2(t).   \text{Q.E.D.}$$

REMARK. It follows from a standard argument (see, for example, [8]) that for any s within ε of $z(t)$, the convex function $H(t, s, \cdot)$ satisfies a Lipschitz condition on $(r/2)B$ with constant $2(\alpha_2(t) - \alpha_1(t))/r$. Hence $\partial H(t, s, p)$ is integrably bounded for s within ε of $z(t)$ and p in a bounded set.

The following result, which we state here for convenience, is a simple corollary of [13].

LEMMA 8. *Let $F(t, s)$ be a multifunction from $[0, 1] \times R^n$ to R^n assuming convex compact values, measurable in t and u.s.c. in s. We suppose r is a positive number and $k(t)$ an integrable function such that $F(t, 2rB)$ is contained in $k(t)B$. Then, for any sequence of arcs x_i and sequence of positive numbers r_i satisfying $\dot{x}_i(t) \in F(t, x_i(t)) + r_i B$ a.e., $\|x_i\| \leq r$, $\lim r_i = 0$, there is a subsequence of $\{x_i\}$ which converges uniformly to an arc x satisfying $\dot{x}(t) \in F(t, x(t))$ a.e.*

4. Proof of the Theorem. Let C be a subset of R^n. The *indicator function* of C, denoted $\delta(\cdot \mid C)$, is defined as follows: $\delta(s \mid C)$ equals 0 if s lies in C, and $+\infty$ otherwise.

We make the following (temporary) assumption: l is the indicator of $C_0 \times C_1$, where C_0 and C_1 are given closed subsets of R^n. Thus the problem of Bolza consists of minimizing the integral of L subject to $x(0)$ in C_0, $x(1)$ in C_1.

We shall denote by $\bar{s}$ a point (s, s^0) in $R^n \times R$. We define a multifunction E from $[0, 1] \times R^{n+1}$ to R^{n+1} by

$$E(t, \bar{s}) = \text{cl co epi } L(t, s, \cdot) + (0, |s - z(t)|^2),$$

and we note that E is measurable in t and Lipschitz in $\bar{s}$ within ε of the arc $\bar{z}$ defined by

$$\bar{z}(t) = \left[z(t), \int_0^t L(r, x, \dot{x}) \, dr \right].$$

We invoke [4, Theorem 2] to deduce that $\bar{z}$ minimizes for some positive δ_1, over the arcs $\bar{x} = (x, x^0)$, the value of $x^0(1)$ subject to

$$\dot{\bar{x}} \in E(t, \bar{x}) \text{ a.e.}, \quad \bar{x}(0) \in C_0 \times \{0\}, \quad x(1) \in C_1, \quad \|x - z\| < \delta_1.$$

A simple argument by contradiction (see [5, Lemma 3]) shows that for some positive δ and m, $\bar{z}$ minimizes

(4.1) $$x^0(1) + m d_1(x(1))$$

(where $d_1(s)$ is the Euclidean distance from s to C_1) under the constraints

(4.2) $$\dot{\bar{x}} \in E(t, \bar{x}) \text{ a.e.}, \quad \bar{x}(0) \in C_0 \times \{0\}, \quad \|x - z\| < \delta,$$

and furthermore is the unique solution (this results from the term $|s - z(t)|^2$).

We now define, for j a positive integer, a multifunction E_j as follows:

$$E_j(t, \bar{s}) = \{\bar{v} \in R^{n+1} : d(\bar{v}, E(t, \bar{s})) \leq 1/j\},$$

and we leave to the reader the verification of the fact that for each $(t, \bar{s})$, $E_j(t, \bar{s})$ is a closed, convex, regular set (see §3). We let $L_j(t, s, \cdot)$ be the convex function whose epigraph is $E_j(t, \bar{s})$, and we define (for $\bar{p} = (p, p^0)$)

$$h_j(t,\bar{s},\bar{p}) = \sup\{\bar{p}\cdot\bar{v}:\bar{v}\text{ in }E_j(t,\bar{s})\}.$$

We define similarly h for E, and we note the relationships

$$h(t,\bar{s},p,-1) = H(t,s,p) - |s - z(t)|^2,$$

(4.3) $h_j(t,\bar{s},p,-1) = H_j(t,s,p) = h(t,\bar{s},p,-1) + (1/j)|(p,-1)|$

$$= H(t,s,p) - |s - z(t)|^2 + (1/j)|(p,-1)|,$$

where H_j is the Hamiltonian corresponding to L_j.

CLAIM 1. The problem of minimizing (4.1) subject to

(4.4) $\dot{\bar{x}} \in E_j(t,\bar{x})$ a.e., $\quad \bar{x}(0) \in C_0 \times \{0\}, \quad \|x - z\| \leqslant \delta$

admits a solution $\bar{z}_j$ (the minimum being finite).

PROOF. It suffices to show that the following problem of Bolza admits a solution z_j: to minimize over the arcs satisfying $\|x - z\| \leqslant \delta$ the functional $md_1(x(1)) + \delta(x(0)|C_0) + \int_0^1 L_j(t,x,\dot{x})\,dt$, for then the arc $\bar{z}_j(t) = [z_j(t), \int_0^t L_j(r,z_j,\dot{z}_j)\,dr]$ would be the required solution.

We may suppose δ no greater than the ε intervening in the hypotheses of Theorem 1. With this in mind, and in view of (4.3), we have, for any s within δ of $z(t)$ and for any p,

$$H_j(t,s,p) \leqslant H(t,s,p) + (1/j)|(p,-1)|$$

$$\leqslant H(t,\alpha(t),p) + k(t)(1 + |p|)|s - \alpha(t)| + (1/j)|(p,-1)|$$

$$\leqslant \beta(t,p) + (1 + |p|)(1 + \alpha(t)k(t)) + k(t)(1 + |p|)|s|.$$

If we redefine $L_j(t,s,v)$ as $+\infty$ for $|s - z(t)| > \delta$, or equivalently, $H_j(t,s,p)$ as $-\infty$ for the same s, then the preceding condition on H_j allows us to apply [12, Theorem 1] to conclude that the required solution z_j does exist, and hence the claim is justified (it follows from [4, Proposition 2] that since the problem (4.1), (4.2) admits a finite minimum, so does the problem (4.1), (4.4)).

CLAIM 2. A subsequence of $\{z_j\}$ converges uniformly to z.

PROOF. We define the functional Φ_j by

$$\Phi_j(x) = md_1(x(1)) + \delta(x(0)|C_0) + \int_0^1 [L_j(t,x,\dot{x}) + \delta(x - z|\delta B)]\,dt,$$

and similarly Φ_0 for L_j replaced by L. Then z_j minimizes Φ_j and z minimizes Φ_0. Since, in general,

$$L_j(t,s,v) \leqslant L(t,s,v) + |s - z(t)|^2,$$

we have for each j, $-\infty < \Phi_j(z_j) \leqslant \Phi_0(z)$. Now let T be any positive integer. Since Φ_j increases with j, we have for all $j > T$, $-\infty < \Phi_T(z_j) \leqslant \Phi_0(z)$. It follows from [12, Semicontinuity Theorem] that a subsequence of $\{z_j\}$ converges uniformly to an arc z_0 satisfying $-\infty < \Phi_T(z_0) \leqslant \Phi_0(z)$. We obtain the same z_0 for each T, since a subsequence of a convergent sequence converges to the same limit, and we may assume without loss of generality that the original sequence converges. One may show that $L_T(t,s,v)$ increases with T to the limit co $L(t,s,v) + |s - z(t)|^2$, where co $L(t,s,\cdot)$ is the function whose epigraph is cl co epi $L(t,s,\cdot)$.

We have for each T,

$$-\infty < md_1(z_0(1)) + \int_0^1 L_T(t,z_0,\dot{z}_0)\,dt \leqslant \Phi_0(z),$$

whence by monotone convergence,

$$-\infty < md_1(z_0(1)) + \int_0^1 [\text{co } L(t,z_0,\dot{z}_0) + |z_0 - z|^2]\,dt \leqslant \Phi_0(z).$$

The fact that $\bar{z}$ is the unique solution to (4.1), (4.2), combined with this last inequality implies that $z_0 = z$, and hence the claim is proved.

We now note that in the problem of minimizing (4.1) subject to (4.4), the constraint $\|x - z\| \leqslant \delta$ may as well be $\|x - z\| < \delta$, since z_j converges uniformly to z. The stage is set for applying Lemma 5 (and the subsequent remark) (the fact that E_j is Lipschitz near $\bar{z}_j$ follows from (4.3) and Lemma 6). We derive the fact that for all j large there is an arc $\bar{p}_j = (p_j, p_j^0)$ such that

(4.5) $(-\dot{p}_j, -\dot{p}_j^0, \dot{z}_j, L(t,z_j,\dot{z}_j)) \in \partial h_j(t,z_j,z_j^0,p_j,p_j^0)$ a.e.,

(4.6) $p_j(0) \in N_{C_0}(z_j(0))$,

(4.7) $-p_j(1) \in \partial md_1(z_j(1))$,

(4.8) $-p_j^0(1) = 1$.

Since h_j is independent of z_j^0, we deduce from (4.5) and (4.8) that p_j^0 is identically -1. Using (4.3) and arguing as in Lemma 3, we deduce from (4.5):

(4.9) $\qquad\qquad (-\dot{p}_j, \dot{z}_j) \in \partial H_j(t,z_j,p_j)$ a.e.

This relation yields, in particular, $|\dot{p}_j| \leqslant k(t)(1 + |p_j|)$, and (4.7) shows that $|p_j(1)|$ is bounded above by m. It follows that for some constant r, we have for all t and for all j that $(z_j(t),p_j(t))$ lies in rB. From (4.3) we obtain $\partial H_j(t,z_j,p_j) \subset \partial H(t,z_j,p_j) + r_j B$, where $r_j = 2\|z_j - z\| + 1/j$. Note that r_j converges to 0, and that (by the remark following Lemma 7) $\partial H(t,s,p)$ is integrably bounded in a neighborhood of (z_j,p_j). Finally we note that ∂H is compact and convex-valued, measurable in t and u.s.c. in (s,p). An application of Lemma 8 shows that a subsequence of $\{(z_j,p_j)\}$ converges uniformly to a limit (z,p) such that

(4.10) $(-\dot{p}, \dot{z}) \in \partial H(t, z, p)$ a.e.

In view of (4.6) and (4.7) we have

(4.11) $p(0) \in N_{C_0}(z(0)), \quad -p(1) \in \partial md_1(z(1))$.

This last inclusion implies, by [3, Definition 3.1],

(4.12) $-p(1) \in N_{C_1}(z(1))$.

Relations (4.11) and (4.12) are equivalent to (2.6) when l is the indicator of $C_0 \times C_1$. Finally, (4.10) implies that $\dot{z}(t)$ is a.e. in the subdifferential of the convex function $H(t, z, \cdot)$ at $p(t)$. This in turn implies [9, Theorem 23.5]

$$H(t, z, p) = p \cdot \dot{z} - \text{co } L(t, z, \dot{z})   \text{a.e.}$$

But by [5, Theorem 2], co $L(t, z, \dot{z}) = L(t, z, \dot{z})$ a.e., so we obtain (2.5).

This completes the proof of the Theorem in the case $l = \delta(\cdot \,|C_0 \times C_1)$. We now show how the reduction to this case is made. We use the notation s' to denote points (s^1, s^2, s^3) in $R^n \times R^n \times R$, and we define
$z'(t) = [z(t), z(1), l(z(0), z(1))]$,
$L'(t, s', v') = L(t, s^1, v^1) + \delta((v^2, v^3)|(0, 0)) + s^3$,
$C_0 = \{s': l(s^1, s^2) \leqslant s^3\} = \text{epi } l$,
$C_1 = \{s': s^1 = s^2\}$.
We may verify the following without difficulty:

(a) z' solves the problem of minimizing $\int_0^1 L'(t, x', \dot{x}')\,dt$ subject to $x'(0)$ in C_0, $x'(1)$ in C_1.

(b) L' is epi-measurable and epi-Lipschitz within ε of z', L' is l.s.c. in (s', v'), and the problem is calm at z'.

(c) $H'(t, s', p') = H(t, s^1, p^1) - s^3$, and H' satisfies the growth condition within ε of z' with $\alpha'(t) = (\alpha(t), z(1), l(z(0), z(1)))$, $\beta'(t, p') = \beta(t, p^1) - l(z(0),$ $z(1))$.

We now apply to this transformed problem the case of the Theorem that has been proved, obtaining the conclusions for z', H', with an arc p'. We leave to the reader the straightforward proof that p^1 satisfies the conclusions of the Theorem.   Q.E.D.

PROOF OF COROLLARY 1. Lemma 6 implies that L is epi-Lipschitz within ε of z. The fact that H is measurable in t and continuous in (s, p) (near z) implies that H is measurable with respect to the product of measurable sets in $[0, 1]$ and Borel sets in $R^n \times R^n$. The same is then true for L [12, Proposition 4]; i.e. L is epi-measurable near z. We have thus demonstrated the applicability of the Theorem.   Q.E.D.

PROOF OF COROLLARY 2. We continue the notational device of denoting

points (s^1, s^2, s^3) in $R^n \times R^n \times R$ by s', and we define C_0 as in the proof of the Theorem. For each positive integer j we consider the problem of minimizing

$$(4.13) \qquad x^3(1) + j|x^1(1) - x^2(1)| + \delta(x^1(1) - x^2(1)|B)$$

subject to

$$(4.14) \qquad x'(0) \text{ in } C_0, \quad \dot{x}' \in E'(t, x') \quad \text{a.e.,}$$

where $E'(t, s') = E(t, s^1) \times \{0\} \times \{0\}$. We may apply [12, Theorem 1] to deduce that the problem admits a (finite) minimum attained at an arc z'_j. It follows from known results (see, for example, [13]) that the set of attainable points $x^1(1)$ is compact, and consequently that $x^3(1)$ is bounded below for admissible arcs (independently of j). Suppose that y is a trajectory for E with $l(y(0), y(1))$ finite, and set $y'(t) = [y(t), y(1), l(y(0), y(1))]$. We have $z_j^3(1) + j|z_j^1(1) - z_j^2(1)| \leqslant y^3(1)$, and it follows that $z_j^1(1) - z_j^2(1)$ converges to 0 as j tends to infinity, and that $z_j^3(1) = l(z_j^1(0), z_j^2(1))$ is bounded above. The trajectories for E' being sequentially compact, we may assume that z'_j converges to an arc z'. We set $z = z^1$. Since $(z_j^1(0), z_j^2(1))$ converges to $(z(0), z(1))$, it follows that $l(z(0), z(1))$ is finite; of course $z^3(1) \geqslant l(z(0), z(1))$.

CLAIM. z is optimal for the original problem, and $z^3(1)$ equals $l(z(0), z(1))$.

PROOF. Suppose that z is not optimal. Then an admissible trajectory x exists such that $l(x(0), x(1))$ is strictly less than $l(z(0), z(1))$. Let us define $x'(t) = [x(t), x(1), l(x(0), x(1))]$ and note that x' is feasible for the problem (4.13), (4.14). We have $\liminf l(z_j^1(0), z_j^2(0)) \geqslant l(z(0), z(1))$, and hence for j large, $x^3(1) < l(z_j^1(0), z_j^2(0)) = z_j^3(1)$, which contradicts the optimality of z'_j. A similar argument demonstrates the second assertion of the claim.

For j large, the term $\delta(x^1(1) - x^2(1)|B)$ is inoperative, and hence the problem (4.13), (4.14) is calm at z'_j by [5, Proposition 1]. We now apply the Theorem to deduce the existence of an arc p'_j such that

(4.15) $(-\dot{p}_j^1, \dot{z}_j^1) \in \partial H(t, z_j^1, p_j^1)$, $\dot{p}_j^2 = \dot{p}_j^3 = 0$ a.e.,
(4.16) $p'_j(0)$ is normal to C_0 at $z'(0)$,
(4.17) $-p_j^3(1) = 1$,
(4.18) $(-p_j^1(1), -p_j^2(1)) \in \partial j|z_j^1(1) - z_j^2(1)|$.

Relation (4.18) yields $p_j^1(1) + p_j^2(1) = 0$, and from (4.15) and (4.17) we derive

$$(4.19) \qquad (p_j^1(0), -p_j^1(1), -1) \text{ is normal to}$$
$$C_0 = \text{epi } l \text{ at the point } z'_j(0) = [z_j^1(0), z_j^2(1), l(z_j^1(0), z_j^2(1))].$$

Suppose for some j we had $|z_j^1(1) - z_j^2(1)| = 0$. It would then follow that z_j^1 is optimal for the original problem, and in view of (4.15) and (4.19) the corollary would be proved, for $p = p_j^1$ and $\lambda = 1$. We may thus suppose the

contrary for each j, so that from (4.18), $|p_j^1(1)| = j$. We set $\tilde{p}_j$ equal to p_j^1/j, and we note by Lemma 3,

$$(4.20) \qquad\qquad (-\dot{\tilde{p}}_j, \dot{z}_j^1) \in \partial H(t, z_j^1, \tilde{p}_j) \quad \text{a.e.,}$$

$$(4.21) \qquad \begin{gathered} (\tilde{p}_j(0), -\tilde{p}_j(1), -1/j) \text{ is normal to epi } l \text{ at the point} \\ (z_j^1(0), z_j^2(1), l(z_j^1(0), z_j^2(1))). \end{gathered}$$

As in the proof of the Theorem, $\{(\tilde{p}_j, z_j^1)\}$ admits a subsequence converging uniformly to a limit arc (p, z), where z was defined above, and where p and z satisfy (2.4), and (2.11) with $\lambda = 0$. Since $|\tilde{p}_j(1)| = 1$ for each j, $p(1)$ is not zero; (2.10) is as before a consequence of (2.4).   Q.E.D.

REFERENCES

1. H. Berliocchi and J.-M. Lasry, *Principe de Pontryagin pour des systèmes régis par une équation différentielle multivoque*, C. R. Akad. Sci. Paris Sér. A–B **277** (1973), A1103–A1105. MR **50** #5580.

2. V. G. Boltjanskiĭ, *The maximum principle for problems of optimal steering*, Differencial'nye Uravnenija **9** (1973), 1363–1370.

3. F. H. Clarke, *Generalized gradients and applications*, Trans. Amer. Math. Soc. **205** (1975), 247–262.

4. ———, *Admissible relaxation in variational and control problems*, J. Math. Anal. Appl. **51** (1975), 557–576.

5. ———, *The generalized problem of Bolza*, SIAM J. Control Optimization **14** (1976), 682–699.

6. R. P. Fedorenko, *A maximum principle for differential inclusions*, Ž. Vyčisl. Mat. i Mat. Fiz. **10** (1970), 1385–1393 = USSR Comput. Math. and Math. Phys. **10** (1970), 57–68. MR **44** #2127.

7. H. Halkin, *Extremal properties of biconvex contingent equations*, Ordinary Differential Equations (NRL–MRC Conf.), Academic Press, New York, 1972, pp. 109–120.

8. A. W. Roberts and D. E. Varberg, *Another proof that convex functions are locally Lipschitz*, Amer. Math. Monthly **81** (1974), 1014–1016. MR **50** #4858.

9. R. T. Rockafellar, *Convex analysis*, Princeton Univ. Press, Princeton, N. J., 1970. MR **43** #445.

10. ———, *Conjugate convex functions in optimal control and the calculus of variations*, J. Math. Anal. Appl. **32** (1970), 174–222. MR **42** #929.

11. ———, *Existence and duality theorems for convex problems of Bolza*, Trans. Amer. Math. Soc. **159** (1971), 1–40. MR **43** #7995.

12. ———, *Existence theorems for general control problems of Bolza and Lagrange*, Advances in Math. **15** (1975), 312–333.

13. M. Valadier, *Existence globale pour les équations différentielles multivoques*, C. R. Acad. Sci. Paris Sér. A –B **272** (1971), A474–A477. MR **43** #5351.

14. J. Warga, *Necessary conditions without differentiability assumptions in optimal control*, J. Differential Equations **18** (1975), 41–62.

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF BRITISH COLUMBIA, VANCOUVER, BRITISH COLUMBIA V6T 1W5, CANADA