OPTIMAL ARCS AND THE MINIMUM VALUE
FUNCTION IN PROBLEMS OF LAGRANGE

BY

R. TYRRELL ROCKAFELLAR[1]

ABSTRACT. Existence theorems are proved for basic problems of Lagrange in the calculus of variations and optimal control theory, in particular problems for arcs with both endpoints fixed. Emphasis is placed on deriving continuity and growth properties of the minimum value of the integral as a function of the endpoints of the arc and the interval of integration. Control regions are not required to be bounded. Some results are also obtained for problems of Bolza.
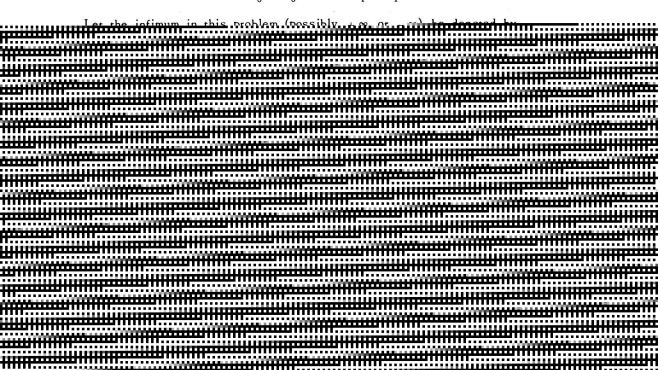
Conjugate convex functions and duality are used extensively in the development, but the problems themselves are not assumed to be especially "convex". Constraints are incorporated by the device of allowing the Lagrangian function to be extended-real-valued. This necessitates a new approach to the question of what technical conditions of regularity should be imposed that will not only work, but will also be flexible and general enough to meet the diverse applications. One of the underlying purposes of the paper is to present an answer to this question.

1. **Statement of main results.** Let $[a, b]$ be a real interval, and let L be a function on $[a, b] \times R^n \times R^n$ with values in $(-\infty, +\infty]$. For each subinterval $[t_0, t_1] \subset [a, b]$ and endpoint pair $(c_0, c_1) \in R^n \times R^n$, we consider the problem of Lagrange in which the integral

$$(1.1) \qquad \int_{t_0}^{t_1} L(t, x(t), \dot{x}(t))\, dt$$

is minimized over all absolutely continuous arcs $x: [t_0, t_1] \to R^n$ such that

$$(1.2) \qquad x(t_0) = c_0 \quad \text{and} \quad x(t_1) = c_1.$$

endpoints and variable time intervals. Certain existence theorems of Cesari [1] and Olech [7], among others, are thereby extended. The fact that L is allowed to have the value $+\infty$, and only very weak regularity conditions are imposed, is essential, of course, in achieving this degree of applicability.

We assume throughout that $L(t, x, v)$ is lower semicontinuous in (x, v) and measurable in (t, x, v) with respect to the σ-algebra in $[a, b] \times R^n \times R^n$ generated by products of Lebesgue sets in $[a, b]$ and Borel sets in $R^n \times R^n$. The function H on $[a, b] \times R^n \times R^n$ defined by

$$(1.3) \qquad H(t, x, p) = \sup\{p \cdot v - L(t, x, v) |\ v \in R^n\}$$

is called the Hamiltonian corresponding to the Lagrangian L. It is said to satisfy the *boundedness condition* if

$$(1.4) \qquad \sup_{|x| \le r} H(t, x, p) \le \phi(t, r, p),$$

where ϕ is some real-valued function on $[a, b] \times [0, +\infty) \times R^n$ such that $\phi(t, r, p)$ is summable in t for fixed r and p. This condition and the other assumptions on L imply in particular that the integral (1.1) is well defined and not $-\infty$; this is shown in Lemma 1.

In most of our results, we invoke the boundedness condition in a somewhat stronger form by applying it, not to H itself, but to a certain function H_0 which majorizes H. We define H_0 to be the least of all the extended-real-valued functions on $[a, b] \times R^n \times R^n$ which majorize H and are upper semicontinuous and concave in the x argument. Thus for each $t \in [a, b]$ and $p \in R^n$ we have

$$\{(x, a) \in R^n \times R^1 |\ a \le H_0(t, x, p)\} = \text{cl co}\{(x, a) \in R^n \times R^1 |\ a \le H(t, x, p)\}.$$

In the important case where $L(t, x, v)$ is convex in (x, v), the boundedness condition on H and the boundedness condition on H_0 are equivalent, and in fact $H_0 = H$ (Lemma 2). The two conditions are also equivalent obviously if there is an $r > 0$ such that $|x| > r$ implies $L(t, x, v) = +\infty$, or more generally, if the function ϕ in the boundedness condition on H can be chosen constant in r. Observe that if L does not depend on t, the boundedness condition is satisfied by H_0 provided only that H_0 is nowhere $+\infty$ (equivalently: for each p there is at least one affine function on R^n majorizing the function $x \to H(x, p)$).

To state the main theorems, we introduce some further notation and terminology. An extended-real-valued function f on $R^n \times R^n$ is said to satisfy the *growth condition* (G_r), where $0 < r \le +\infty$, if

$$(1.5) \qquad f(c_0, c_1) \ge \max\{\gamma(|c_1 - A_0 c_0| - a_0 \cdot c_0) + b_0 \cdot c_0,$$
$$\gamma(|c_0 - A_1 c_1| - a_1 \cdot c_1) + b_1 \cdot c_1\},$$

where A_0 and A_1 are nonsingular $n \times n$ matrices, a_0, a_1, b_0 and b_1 are vectors in R^n, and the function $\gamma: (-\infty, +\infty) \to (-\infty, +\infty]$ is nondecreasing, convex, bounded below, and satisfies

(1.6) $$\liminf_{\lambda \to +\infty} \gamma(\lambda)/\lambda \geq r.$$

(Observe that the right side of (1.5) is convex in (c_0, c_1).) We define

(1.7) $$S = \{(t_0, t_1) \mid a \leq t_0 \leq t_1 \leq b\},$$

(1.8) $$S_\epsilon = \{(t_0, t_1) \mid a \leq t_0 \leq t_1 \leq b, \ t_1 - t_0 \leq \epsilon\},$$

(1.9) $$H_1(t, x, p) = \liminf_{\mu \to +\infty} \left[\liminf_{\lambda \to +\infty} H_0(t, \bar{x}(t) + \lambda x, \mu p)/\lambda \mu \right].$$

In the latter formula, $\bar{x}(t)$ is any point of R^n such that the function $v \to L(t, \bar{x}(t), v)$ is not identically $+\infty$. To ensure the existence of such a point, and to avoid various trivialities, we henceforth make the harmless assumption that $L(t, \cdot, \cdot)$ is not identically $+\infty$ for any t, and in fact there exists at least one absolutely continuous $x: [a, b] \to R^n$ such that the function $t \to L(t, x(t), \dot{x}(t))$ is majorized by a summable, real-valued function on $[a, b]$. The definition of H_1 is actually independent of the particular choice of $\bar{x}(t)$, since $H_0(t, x, p)$ is by definition concave and upper semicontinuous in x and the condition on $\bar{x}(t)$ implies $H_0(t, \bar{x}(t), p) > -\infty$. We adopt the convention that

(1.10) $$\begin{aligned} F(t_0, t_1, c_0, c_1) &= 0 \quad \text{if } t_0 = t_1 \text{ and } c_0 = c_1, \\ &= +\infty \quad \text{if } t_0 = t_1 \text{ and } c_0 \neq c_1. \end{aligned}$$

Theorem 1. *Suppose that H_0 satisfies the boundedness condition and $L(t, x, v)$ is convex in v. Let $0 < r < +\infty$. Then there is an $\epsilon > 0$ such that the following is true.*

(a) *The infimum defining $F(t_0, t_1, c_0, c_1)$ is attained by at least one arc x for each (t_0, t_1, c_0, c_1) in $S_\epsilon \times R^n \times R^n$.*

(b) *F is lower semicontinuous relative to $S_\epsilon \times R^n \times R^n$.*

(c) *The function $F(t_0, t_1, \cdot, \cdot)$ satisfies the growth condition (G_r) for each $(t_0, t_1) \in S_\epsilon$.*

Theorem 2. *Suppose that H_0 and H_1 both satisfy the boundedness condition and $L(t, x, v)$ is convex in v. Then the following is true.*

(a) *The infimum defining $F(t_0, t_1, c_0, c_1)$ is attained by at least one arc x for each (t_0, t_1, c_0, c_1) in $S \times R^n \times R^n$.*

(b) *F is lower semicontinuous relative to $S \times R^n \times R^n$.*

(c) *The function $F(t_0, t_1, \cdot, \cdot)$ satisfies the growth condition (G_∞) for each $(t_0, t_1) \in S$.*

As we show in Lemma 5, $H_1(t, x, p)$ is concave in x, so that in the "autonomous" case the boundedness condition on H_1 is equivalent to $H_1(t, x, p) < +\infty$. Thus:

Corollary. *Suppose that $L(t, x, v)$ is independent of t and convex in v, and that H_0 and H_1 nowhere have the value $+\infty$. Then conclusions* (a), (b) *and* (c) *of Theorem 2 are valid.*

Theorem 3. *Suppose that H_0 satisfies the boundedness condition and $L(t, x, v)$ is convex in v. Assume further that there is an absolutely continuous function $p \colon [a, b] \to R^n$ (e.g. a constant function) such that*

$$(1.11) \qquad \sup_{x \in R^n} \{H(t, x, p(t)) + x \cdot \dot{p}(t)\} < +\infty \quad a.e.$$

Then conclusions (a) *and* (b) *of Theorem 2 are valid, and moreover:*

(c) *The function $F(t_0, t_1, \cdot, \cdot)$ majorizes at least one affine function on $R^n \times R^n$ for each $(t_0, t_1) \in S$.*

In the case of Theorem 3 with $p(t) \equiv p_0 \in R^n$, condition (1.11) requires simply that the function $(x, v) \to L(t, x, v) - p_0 \cdot v$ be bounded below for almost every $t \in [a, b]$.

Of course, it is not actually necessary to calculate H_0 and H_1 explicitly from H in order to apply Theorems 1, 2 and 3. It suffices to know a function $H_0' \geq H$ such that $H_0'(t, x, p)$ is concave and upper semicontinuous in x, and H_0' satisfies the boundedness condition. Then $H_0 \leq H_0'$ (by the definition of H_0), so that H_0 satisfies the boundedness condition. Similarly, if the function H_1' obtained by substituting H_0' for H_0 in (1.9) satisfies the boundedness condition, then so does H_1. For example, the boundedness conditions on H_0 and H_1 are both satisfied if

$$H(t, x, p) \leq \psi_1(t, x) + \psi_2(t, p) + p \cdot A(t)x,$$

where $\psi_1(t, x)$, $\psi_2(t, p)$ and the components of the matrix $A(t)$ are summable in t, and $\psi_1(t, x)$ is concave in x.

Theorem 1 implies that, for a large class of problems of Lagrange, optimal arcs exist when the t-interval is constrained to be sufficiently small. One is reminded of various theorems on the local existence of solutions to differential equations, and indeed there is a close relationship, apparent in the proof. Although no estimate is given for ϵ, so that direct application may be difficult, the result is nevertheless interesting theoretically, because it shows that the general question of existence can often be reduced to finite-dimensional questions about the growth properties of $F(t_0, t_1, c_0, c_1)$.

For example, suppose that $[t_0, t_1]$ is an interval of length between ϵ and 2ϵ, where ϵ has the properties in Theorem 1 for a certain $r > 0$. Let $t_2 = (t_0 + t_1)/2$. Then properties (a), (b) and (c) are applicable to $F(t_0, t_2, \cdot, \cdot)$ and $F(t_2, t_1, \cdot, \cdot)$. Furthermore, by definition,

$$(1.12) \qquad F(t_0, t_1, c_0, c_1) = \inf_{c_2 \in R^n} \{F(t_0, t_2, c_0, c_2) + F(t_2, t_1, c_2, c_1)\}.$$

Thus the question of the existence of optimal arcs over the interval $[t_0, t_1]$ is reduced in principal to the question of whether certain infima in $c_2 \in R^n$ are attained, and this in turn depends on the growth properties of $F(t_0, t_2, \cdot, \cdot)$ and $F(t_2, t_1, \cdot, \cdot)$. This idea can be extended, of course, to intervals $[t_0, t_1]$ whose length is an arbitrary multiple of ϵ by "concatenation".

There does not seem to be any direct counterpart in the literature to Theorem 1 or conclusions (c) of Theorems 2 and 3. However, conclusions (a) and (b) of Theorems 2 and 3 have been obtained under other assumptions. We have already proved Theorem 3 in [9] in the case where $L(t, x, v)$ is convex in (x, v) (so that $H_0 = H$). This previous version is essential to the present one. In fact, most of the results here rest on theorems in [9]. Thus it is demonstrated, in passing, that the special case where $L(t, x, v)$ is convex in (x, v), rather than just convex in v, is really fundamental to a much wider class of problems.

Aside from the case just mentioned, the closest result to Theorems 2 and 3 previously obtained is the following one of Olech [7]. This theorem is not stated as such in [7], but it is implicit in what is proved there. (We contribute the novel formulation in terms of an extended-real-valued Lagrangian L having an appropriate measurability property. Extended-real-valued functions have not been employed in this way by previous authors on the subject.)

Theorem 4 (Olech [7]). *Suppose that H satisfies the boundedness condition and $L(t, x, v)$ is convex in v. Assume further that one of the following two conditions is satisfied, where ϕ is the function in (1.4):*

(I) *$\phi(t, r, 0)$ is independent of r in (1.4), and there is a positive number s such that we can write*

$$(1.13) \qquad \phi(t, r, p) = \phi_0(t, p) + r\phi_1(t, p)$$

for all $p \in R^n$ with $|p| = s$, or

(II) *there is a positive number s such that (1.13) holds for all $p \in R^n$ with $|p| \geq s$, as well as for $p = 0$. Furthermore, there is a constant m such that*

$$(1.14) \qquad \int_a^b |\phi_1(t, p)| \, dt \leq m|p| \quad \text{if } |p| \geq s.$$

Then conclusions (a) *and* (b) *of Theorems 2 and 3 hold, and moreover,*

(c) *F has the growth property that for every positive* ρ *and real* α *the (closed) set*

$$(1.15) \quad \{(t_0, t_1, c_0, c_1) \in S \times R^n \times R^n \mid F(t_0, t_1, c_0, c_1) \leq \alpha \text{ and } \min\{|c_0|, |c_1|\} \leq \rho\}$$

is bounded.

The latter growth property is also implied obviously by conclusions (b) and (c) of Theorem 2, and similarly by conclusions (b) and (c) of Theorem 1 if S is replaced by S_ϵ.

The reader may wonder about the case, prominent in other papers on problems of Lagrange, where in minimizing the integral (1.1) there is the additional precondition that $x(t) \in X$ for all t, where X is a fixed compact set. Due to the fact that L is allowed to have the value $+\infty$, this case is fully covered by Theorems 2 and 4, as stated. Namely, in this case one can assume without loss of generality, redefining L if necessary, that whenever $x \notin X$ one has $L(t, x, v) = +\infty$, and hence $H(t, x, p) = -\infty$. Then minimizing (1.1) subject to (1.2) and $x(t) \in X$ for all t is equivalent to minimizing (1.1) subject only to (1.2). If H satisfies the boundedness condition, and $L(t, x, v)$ is convex in v, all the assumptions in Theorems 2 and 4 are trivially satisfied, and both theorems yield the same conclusions.

It should be mentioned that Olech's paper [7], from which we have distilled Theorem 4, treats a larger class of problems than those considered here, in particular certain problems with a vector-valued integral (1.1). The earlier work of Cesari [1], where a larger class of problems is also treated, likewise contains a version of Theorem 4(I), where in effect the stronger assumptions are made that the set

$$(1.16) \qquad\qquad \{(t, x, v) \mid L(t, x, v) < +\infty\}$$

is closed and has a closed projection in the (t, x)-space, L is continuous relative to (1.16), the function ϕ in the boundedness condition on H can be chosen constant in t, and for $|p|$ sufficiently small, constant as well in r. Cesari's result in turn generalizes classical theorems of McShane [5], Nagumo [6] and Tonelli [15] for free problems of the calculus of variations. At the same time, it has the advantage of yielding for control problems (with unilateral constraints) existence results more general than those of Filippov [3], Lee and Marcus [4], and Roxin [14], among others, in that unbounded control regions are permitted.

The results in this paper are likewise relevant to optimal control, as we explain in §4 pursuing a more general line of reasoning on the matter than Cesari or Olech, and thereby extending certain theorems of those authors in several respects.

A compactness result (Theorem 5) is proved in §3 which, as an obvious corollary, gives the existence of optimal arcs in various "nonconvex" problems of Bolza. Other such theorems can easily be derived from Theorems 1 and 2, for example, using the growth properties of F. Related theorems of a more detailed nature for "convex" problems of Bolza have been presented in [9] and [10].

2. **Preliminary facts.** The lemmas stated below are needed in proofs of Theorems 1, 2 and 3 in the next section.

Lemma 1. *Suppose that H satisfies the boundedness condition. Then the integral (1.1) is well defined for any absolutely continuous function $x: [t_0, t_1]$ $\rightarrow R^n$, in the sense that the integrand $L(t, x), \dot{x}(t))$ (where it exists, which is almost everywhere in $[t_0, t_1]$) is Lebesgue measurable in t and majorizes a summable function of t. (Thus the value of the integral is unambiguously a real number or $+ \infty$.)*

Proof. The Lebesgue measurability of $L(t, x(t), v(t))$ in t, where $x(t)$ and $v(t)$ are Lebesgue measurable in t, follows easily from the Lebesgue measurability of the mapping $t \rightarrow (t, x(t), v(t))$ and the special measurability property assumed for L (cf. [12]). (If L were only Lebesgue measurable, the composed function could fail to be measurable.) The boundedness condition on H implies further that

$$(2.1) \qquad L(t, x(t), \dot{x}(t)) \geq -\phi(t, r, 0)$$

if $|x(t)| \leq r$ for all t. Since $\phi(t, r, 0)$ is summable in t, the conclusion is apparent.

Lemma 2. *Suppose that H_0 nowhere has the value $+ \infty$. Define*

$$(2.2) \qquad L_0(t, x, v) = \sup\{p \cdot v - H_0(t, x, p)| \ p \in R^n\} > - \infty.$$

Then $L_0(t, \cdot, \cdot)$ is for each $t \in [a, b]$ the greatest, lower semicontinuous, convex (extended-real-valued) function on $R^n \times R^n$ majorized by $L(t, \cdot, \cdot)$, and L_0 is measurable on $[a, b] \times R^n \times R^n$ in the sense already described for L. Moreover, H_0 is the Hamiltonian corresponding to L_0, in other words

$$(2.3) \qquad H_0(t, x, p) = \sup\{p \cdot v - L_0(t, x, v)| \ v \in R^n\}.$$

Proof. Let

$$(2.4) \qquad \begin{aligned} M(t, p, w) &= \sup\{w \cdot x + H(t, x, p)| \ x \in R^n\} \\ &= \sup\{w \cdot x + p \cdot v - L(t, x, v)| \ (x, v) \in R^n \times R^n\}. \end{aligned}$$

Then M is convex in (p, w) by the second expression, and $M(t, p, \cdot)$ is the convex function on R^n conjugate to $-H(t, \cdot, p)$. It follows from the fundamental theorem on conjugate convex functions (see [13, Theorem 12.2]) that the conjugate of $M(t, p, \cdot)$ is in turn the greatest lower semicontinuous convex function on R^n majorized by $-H(t, \cdot, p)$, which is $-H_0(t, \cdot, p)$ by definition, unless the latter has the value $-\infty$ somewhere. Thus under our assumption that H_0 nowhere has the value $+\infty$ we have for all (t, x, p)

$$(2.5) \qquad H_0(t, x, p) = \inf\{M(t, p, w) - w \cdot x \mid w \in R^n\}.$$

Combining this with (2.2), we got

$$(2.6) \qquad L_0(t, x, v) = \sup\{w \cdot x + p \cdot v - M(t, p, w) \mid (w, p) \in R^n \times R^n\}.$$

Thus for each t the function $L_0(t, \cdot, \cdot)$ is conjugate to $M(t, \cdot, \cdot)$, while the latter is by (2.4) conjugate to $L(t, \cdot, \cdot)$. Applying once more the fundamental theorem on conjugate convex functions, we see that $L_0(t, \cdot, \cdot)$ is the greatest, lower semicontinuous, convex function on $R^n \times R^n$ majorized by $L(t, \cdot, \cdot)$. Finally, (2.2) defines $L_0(t, x, \cdot)$ as the conjugate of $H_0(t, x, \cdot)$. But $H_0(t, x, \cdot)$ is a convex function on R^n by virtue of (2.5) and the convexity of $M(t, \cdot, \cdot)$ [13, Theorem 33.1]. Since by hypothesis H_0 nowhere has the value $+\infty$, $H_0(t, x, \cdot)$ is either finite and continuous on R^n or identically $-\infty$ on R^n [13, Theorem 7.2]. This implies that the conjugate of $L_0(t, x, \cdot)$ is in turn $H_0(t, x, \cdot)$, so that (2.3) holds. The lemma is now proved.

Lemma 2 is basic to our whole approach. If H_0 satisfies the boundedness condition, Lemmas 1 and 2 imply that

$$(2.7) \qquad \int_a^b L(t, x(t), \dot{x}(t))\, dt \geq \int_a^b L_0(t, x(t), \dot{x}(t))\, dt > -\infty,$$

where both integrals are well defined. Furthermore, since $L_0(t, x, v)$ is convex in (x, v), the integral on the right is convex as a function of the arc x. Results in [9] yield compactness properties of the set of arcs satisfying

$$(2.8) \qquad \int_a^b L_0(t, x(t), \dot{x}(t))\, dt \leq \alpha$$

for a fixed α. Lower semicontinuity properties of the integral on the left in (2.7) enable us to transfer these compactness properties to the corresponding set of arcs satisfying

$$(2.9) \qquad \int_a^b L(t, x(t), \dot{x}(t))\, dt \leq \alpha.$$

The case of varying intervals $[t_0, t_1]$ is handled by reformulating the situation in such a manner that only the fixed interval $[a, b]$ appears.

Let $\mathfrak{A}$ denote the set of all absolutely continuous functions $x: [a, b] \to R^n$. We regard $\mathfrak{A}$ as a Banach space under the norm

$$(2.10) \qquad \|x\| = \int_a^b |\dot{x}(t)| \, dt + |x(a)|.$$

The lower semicontinuity we need is expressed by the following lemma. This lemma, essentially due to Olech [7], [8], is closely related to more recent results of Cesari [2] that are more general in many respects, but more restrictive in others. We furnish a proof based on the theory of conjugate convex functions. For a fresh discussion of the well-known fundamental connection between lower semicontinuity of the Lagrangian functional and the convexity of $L(t, x, v)$ in v, we refer the reader to the cited paper of Cesari.

Lemma 3. *Suppose that H satisfies the boundedness condition and $L(t, x, v)$ is convex in v. Then the functional*

$$(2.11) \qquad I_L(x) = \int_a^b L(t, x(t), \dot{x}(t)) \, dt$$

is lower semicontinuous sequentially in the weak topology on $\mathfrak{A}$.

Proof. Since the function $L(t, x, \cdot)$ from R^n to $(-\infty, +\infty]$ is lower semicontinuous and convex, it is the conjugate of its conjugate function. The latter is $H(t, x, \cdot)$ by definition. Thus we have

$$(2.12) \qquad L(t, x, v) = \sup\{p \cdot v - H(t, x, p)\mid p \in R^n\}.$$

We demonstrate first that this implies, for every $x \in \mathfrak{A}$,

$$(2.13) \qquad \int_a^b L(t, x(t), \dot{x}(t)) \, dt = \sup_{p \in \mathfrak{L}^\infty} \int_a^b [p(t) \cdot \dot{x}(t) - H(t, x(t), p(t))] \, dt,$$

where $\mathfrak{L}^\infty$ is the space of all bounded, measurable functions from $[a, b]$ to R^n. Formula (2.12) expresses I_L as the pointwise supremum of a certain collection of functionals, one for each $p \in \mathfrak{L}^\infty$, and only the lower semicontinuity of the latter functionals then needs to be proved to obtain the lower semicontinuity of I_L.

Fix $x \in \mathfrak{A}$ and let

$$(2.14) \qquad f(t, v) = L(t, x(t), v) \quad \text{and} \quad h(t, p) = H(t, x(t), p).$$

From the measurability of the mapping $(t, v) \to (t, x(t), v)$, it is clear that f inherits from L the property of being measurable with respect to the σ-algebra generated by products of Lebesgue sets in $[a, b]$ and Borel sets in R^n. The convex functions $f(t, \cdot)$ and $h(t, \cdot)$ are conjugate to each other, and hence h also has this measurability property [8, Proposition 1]. In particular, $h(t, p(t))$ is measurable in t for each $p \in \mathfrak{L}^\infty$. The boundedness condition on H implies that

for each $p \in R^n$ the function $b(\cdot, p)$ is majorized by at least one summable function. Therefore, by a fundamental theorem on conjugate integral functionals (see [8, Theorem 2]) the integral $\int_a^b b(t, p(t)) \, dt$ is well defined for each $p \in \mathfrak{L}^\infty$ and we have

$$(2.15) \qquad \sup_{p \in \mathfrak{L}^\infty} \int_a^b [p(t) \cdot v(t) - b(t, p(t))] \, dt = \int_a^b f(t, v(t)) \, dt$$

for every summable $v: [a, b] \to R^n$. (The cited theorem hypothesizes $b(t, p) \neq -\infty$, but the extension to allow the latter possibility is trivial.) Taking $v(t) = \dot{x}(t)$, we get (2.13).

We now fix $p \in \mathfrak{L}^\infty$ and proceed to investigate the lower semicontinuity of the functional

$$(2.16) \qquad x \to \int_a^b p(t) \cdot \dot{x}(t) \, dt - \int_a^b H(t, x(t), p(t)) \, dt.$$

The first integral describes a continuous linear functional on $\mathfrak{A}$, so the issue revolves around the weak upper semicontinuity of functional

$$(2.17) \qquad x \to \int_a^b H(t, x(t), p(t)) \, dt.$$

Given a sequence x_k in $\mathfrak{A}$ converging to x_0 in the weak topology, we can find an $r > 0$ such that $|x_k(t)| \leq r$ for all t and k. Let $\{p_i\}$ be a finite subset of R^n whose convex hull includes $p(t)$ for every t. The convexity of $H(t, x, \cdot)$ and the boundedness condition on H imply that, whenever $|x| \leq r$ and $q \in \text{co}\{p_i\}$, we have

$$(2.18) \qquad H(t, x, q) \leq \max_i H(t, x, p_i) \leq \max_i \phi(t, r, p_i) = \psi(t),$$

where ψ is a summable function. Therefore,

$$(2.19) \qquad H(t, x_k(t), p(t)) \leq \psi(t) \quad \text{for all } t \text{ and } k.$$

Since in particular $x_k(t)$ converges to $x_0(t)$ for each t, we have by Fatou's Lemma

$$(2.20) \qquad \limsup_{k \to \infty} \int_a^b H(t, x_k(t), p(t)) \, dt \leq \int_a^b H(t, x_0(t), p(t)) \, dt$$

as desired, provided that

$$(2.21) \qquad \limsup_{k \to \infty} H(t, x_k(t), p(t)) \leq H(t, x_0(t), p(t)).$$

Thus the proof is reduced to showing that for fixed $t \in [a, b]$ the function $H(t, x, p(t))$ is upper semicontinuous in x, at least as x ranges over the ball of radius r and center 0 in R^n. Equivalently, we need to show that for any $\alpha \in R^1$ the set

$$(2.22) \qquad \{(x, v) \in R^n \times R^n | \ L(t, x, v) - p(t) \cdot v \leq \alpha, \ |x| \leq r\}$$

has a closed image under the projection $(x, v) \to x$. The assumed lower semi-continuity of $L(t, \cdot, \cdot)$ ensures that (2.22) itself is closed. Hence boundedness of (2.22) will suffice for our conclusion. We may assume the points p_i in (2.18) chosen so that their convex hull includes a ball of radius $s > 0$ about $p(t)$. Then (2.18) holds whenever $|x| \leq r$ and $|q - p(t)| \leq s$, so that from (2.12) we have

$$(2.23) \qquad L(t, x, v) - p(t) \cdot v \geq s|v| - \psi(t).$$

Therefore, every (x, v) in the set (2.22) satisfies $|v| \leq [\alpha + \psi(t)]/s$, as well as $|x| \leq r$, and the lemma is proved.

Remark. A stronger property than described in Lemma 3 is actually true, although we do not need it here. It follows from Theorem 5, namely, that for every $r > 0$ and $\alpha \in R^1$ the set

$$\{x \in \mathcal{C} \mid I_L(x) \leq \alpha, \ |x(t)| \leq r \text{ for all } t\}$$

is compact in the weak topology. (One uses the device of redefinition mentioned in §1 following the statement of Theorem 4.) Almost the same fact was established tacitly by Olech [7]; see also [10] for refinements in the "convex" case.

Lemma 4. *Suppose that H_0 satisfies the boundedness condition. For each $t \in [a, b]$, define*

$$(2.24) \qquad E(t) = \mathrm{cl} \left\{ (p, w) \in R^n \times R^n \ \middle| \ \sup_{x \in R^n} [H(t, x, p) + w \cdot x] < +\infty \right\}.$$

Then $E(t)$ is nonempty and convex. Furthermore, given any real $s > 0$ there exist $n \times n$ matrices $N(t)$ and vectors $j(t)$ in R^n, components summable in t, such that

$$(2.25) \qquad (p, N(t)p + j(t)) \in E(t) \quad \text{whenever } |p| \leq s.$$

Proof. By definition we have

$$(2.26) \qquad E(t) = \mathrm{cl}\{(p, w) \in R^n \times R^n \mid M(t, p, w) < +\infty\},$$

where M is given by (2.4). Since $M(t, \cdot, \cdot)$ is convex by the latter formula, $E(t)$ is convex. From (2.4) and the definition of H_0 we have

$$(2.27) \qquad M(t, p, w) = \sup\{w \cdot x + H_0(t, x, p) \mid x \in R^n\},$$

and hence by (2.3)

$$(2.28) \qquad M(t, p, w) = \sup\{w \cdot x + p \cdot v - L_0(t, x, v) \mid (x, v) \in R^n \times R^n\}.$$

This, combined with (2.6), shows that L_0 and M are Lagrangians dual to each other in the sense of [9]. Then [9, Proposition 2a] asserts that the boundedness condition on H_0 is equivalent to a certain condition (C_0), according to which in particular there exists for each $p \in R^n$ at least one summable function $w: [a, b]$

$\to R^n$ such that $M(t,\ p,\ w(t)) < +\infty$. (In particular, $E(t)$ is nonempty.) Then [9, Proposition 3] applied to M furnishes for any $s > 0$ matrices $N(t)$ and vectors $j(t)$ with the specified property.

Lemma 5. *Suppose that H_0 satisfies the boundedness condition. Then $H_1(t,\ x,\ p)$ in (1.9) is concave in x and convex in p. The boundedness condition on H_1 implies the existence of $n \times n$ matrices $N(t)$ and vectors $j(t)$ and $k(t)$ in R^n, components summable in t, such that*

(2.29) $$(p,\ N(t)p + j(t) + sk(t)) \in E(t) \quad \text{if } |p| \le s < +\infty.$$

Proof. Let

(2.30) $$H_2(t,\ x,\ p) = \liminf_{\lambda \to +\infty} H_0(t,\ \bar{x}(t) + \lambda x,\ p)/\lambda,$$

where $\bar{x}(t)$ is chosen so that $L(t,\ \bar{x}(t),\ \cdot)$ is not the constant function $+\infty$ (and consequently $H_0(t,\ \bar{x}(t),\ p) > -\infty$). Thus $H_2(t,\ \cdot,\ p)$ is the recession function for the upper semicontinuous, concave function $H_0(t,\ \cdot,\ p)$ [13, p. 66]. Since $-H_0(t,\ \cdot,\ p)$ is by (2.5) conjugate to $M(t,\ p,\ \cdot)$, its recession function is the support function of the effective domain of $M(t,\ p,\ \cdot)$, in other words

(2.31) $$H_2(t,\ x,\ p) = \sup\{w \cdot x|\ w \in R^n,\ M(t,\ p,\ w) < +\infty\}$$

[13, Theorem 13.3]. Recall from the proof of Lemma 4 that the set

$$E_0(t) = \{(p,\ w)|\ M(t,\ p,\ w) < +\infty\}$$

is convex and (because of the boundedness condition on H_0) has all of R^n as its image under the projection $(p,\ w) \to p$. It follows easily then that

$$\text{cl}\{w|\ (p,\ w) \in E_0(t)\} = \{w|\ (p,\ w) \in \text{cl } E_0(t) = E(t)\}$$

(see [13, §6]). Thus, with $E(t)$ as in (2.24), we have

(2.32) $$H_2(t,\ w,\ p) = \sup\{w \cdot x|\ (p,\ w) \in E(t)\} < +\infty.$$

Since $E(t)$ is a convex set, we see from this formula that for each $(t,\ x)$ the function $H_2(t,\ x,\ \cdot)$ is convex and nowhere $+\infty$. Hence it is the conjugate of its conjugate: defining

(2.33)
$$L_2(t,\ x,\ v) = \sup\{p \cdot v - H_2(t,\ x,\ p)|\ p \in R^n\}$$
$$= \sup\{p \cdot v + x \cdot w|\ (p,\ w) \in E(t)\},$$

we have dually

(2.34) $$H_2(t,\ x,\ p) = \sup\{p \cdot v - L_2(t,\ x,\ v)|\ v \in R^n\}.$$

We note now that, by definition,

(2.35) $$H_1(t,\ x,\ p) = \liminf_{\mu \to +\infty} H_2(t,\ x,\ \mu p)/\mu.$$

Since $H_2(t, x, \cdot)$ is convex and nowhere $+\infty$, (2.35) says that $H_1(t, x, \cdot)$ is the recession function for $H_2(t, x, \cdot)$ (in fact, "lim inf" can be replaced by "lim"). Therefore, pursuing the same reasoning as above, $H_1(t, x, \cdot)$ is the support function of the effective domain of the conjugate of $H_2(t, x, \cdot)$, which is $L_2(t, x, \cdot)$:

$$(2.36) \qquad H_1(t, x, p) = \sup\{p \cdot v \mid v \in R^n, \; L_2(t, x, v) < +\infty\}$$

$$= \sup\{p \cdot v \mid (x, v) \in D(t)\};$$

where

$$(2.37) \qquad D(t) = \{(x, v) \mid L_2(t, x, v) < +\infty\}.$$

The convexity of $D(t)$ is apparent from (2.33), and, as a consequence, $H_1(t, x, p)$ is by (2.36) not only convex and lower semicontinuous in p, but concave in x. Next define

$$(2.38) \qquad M_1(t, p, w) = \sup\{w \cdot x + H_1(t, x, p) \mid x \in R^n\},$$

$$(2.39) \qquad -\overline{H}_1(t, x, p) = \sup\{w \cdot x - M_1(t, p, w) \mid w \in R^n\}.$$

In other words, $M_1(t, p, \cdot)$ and $-\overline{H}_1(t, \cdot, p)$ are the conjugate and biconjugate of $-\overline{H}_1(t, \cdot, p)$, respectively. In view of the concavity of $H_1(t, x, p)$ in x, the boundedness condition on H_1 is equivalent to the boundedness condition on $\overline{H}_1$, so that henceforth we work solely with the latter. Combining (2.38) and (2.36), we see that

$$(2.40) \qquad M_1(t, p, w) = \sup\{w \cdot x + p \cdot v \mid (x, v) \in D(t)\}.$$

But $D(t)$ consists of the vector pairs (x, v) such that the linear functional $(p, w) \to w \cdot x + p \cdot v$ is bounded above on $E(t)$. Hence [13, Corollary 14.2.1]:

$$(2.41) \qquad \begin{aligned} M_1(t, p, w) &= 0 && \text{if } (p, w) \in E_1(t), \\ &= +\infty && \text{if } (p, w) \notin E_1(t), \end{aligned}$$

where E_1 is the recession cone of $E(t)$:

$$(2.42) \qquad E_1(t) = \{(p, w) \mid E(t) + (p, w) \subset E(t)\}.$$

Note that $M_1(t, p, w)$ is convex and lower semicontinuous in (p, w), and $M_1(t, 0, 0) = 0$. We claim that M_1 is also measurable in (t, p, w) with respect to the σ-algebra in $[a, b] \times R^n \times R^n$ generated by products of Lebesgue sets in $[a, b]$ and Borel sets in $R^n \times R^n$. This property is needed so that M_1 satisfies condition (B), as well as (A), (C) and (D) of [9]. We can then apply [9, Proposition 2a], according to which (in view of (2.39) and (2.41)) $\overline{H}_1$ satisfies the boundedness condition if and only if there exists for each $p \in R^n$ a summable function $w: [a, b] \to R^n$ such that

(2.43) $(p, w(t)) \in E_1(t)$ for all t.

To prove the measurability property of M_1 in question, we argue first that M has this property, because L_0 has it (Lemma 2) and $M(t, \cdot, \cdot)$ is the conjugate of $L_0(t, \cdot, \cdot)$ (see (2.28)) [12, Proposition 1]. The measurability property of M implies for each $\alpha \in R^1$ the measurability of the set

(2.44) $\{(t, p, w) \mid M(t, p, w) \leq \alpha\}$

in the sense described. For each t let $E^\alpha(t)$ denote the (p, w)-section of (2.44), and let $d^\alpha(t, p, w)$ denote the distance of (p, w) from $E^\alpha(t)$ (distance from the empty set being $+\infty$). Since $M(t, \cdot, \cdot)$ is lower semicontinuous, $E^\alpha(t)$ is closed. It is known that in this event measurability of (2.44) (with respect to the σ-field generated by products of Lebesgue sets in $[a, b]$ and Borel sets in $R^n \times R^n$) is equivalent to the (Lebesgue) measurability of $d^\alpha(t, p, w)$ with respect to t for each (p, w) (see [12, Theorem 1]). Let $d(t, p, w)$ denote the infimum of $d^\alpha(t, p, w)$ over all real α (or equivalently, over all integers α). Then $d(t, p, w)$ is the distance of (p, w) from $E(t)$ (which is the closure of the union of the sets $E^\alpha(t)$), and $d(t, p, w)$ is again measurable in t. Therefore, by the fact just cited, the set

(2.45) $\{(t, p, w) \mid (p, w) \in E(t)\}$

is measurable in the desired sense in $[a, b] \times R^n \times R^n$. Now choose any summable function $j: [a, b] \rightarrow R^n$ such that

(2.46) $(0, j(t)) \in E(t)$ for all t.

Such a function exists by Lemma 4, since H_0 satisfies the boundedness condition. Since $E_1(t)$ is the recession cone of $E(t)$ by (2.42), we have

(2.47) $E_1(t) = \bigcap_m (1/m)[E(t) - (0, j(t))],$

where the intersection is over all natural numbers m [13, p. 61]. This shows that the set $\{(t, p, w) \mid (p, w) \in E_1(t)\}$ is the intersection of a countable collection of measurable sets and therefore measurable. The desired measurability property of M_1 is now verified.

Summarizing to this point, we have proved that the boundedness condition on H_1 is equivalent to the existence for each $p \in R^n$ of a summable function $w: [a, b] \rightarrow R^n$ for which (2.43) holds. We argue now from the latter property. Applying [9, Proposition 3] to M_1, we obtain $N(t) \in R^{n \times n}$ and $k(t) \in R^n$ with components summable in t, such that

(2.48) $(p, N(t)p + k(t)) \in E_1(t)$ for all t if $|p| = 1$.

Since $E_1(t)$ is a closed convex cone, we have accordingly

(2.49) $$(\bar{p}, \bar{N}(t)p + sk(t)) \in E_1(t) \quad \text{if } |p| \leq s < +\infty.$$

Then

(2.50) $$(0, j(t)) + (p, N(t)p + sk(t)) \in E(t) \quad \text{if } |p| \leq s < +\infty,$$

where j still denotes a summable function such that (2.46) holds. Since (2.50) and (2.29) are equivalent, Lemma 5 is proved.

3. **A compactness theorem and the proofs of the main results.** Our arguments are based on the following generalization of a result we established in [9, Theorem 1a] in the case where $L(t, x, v)$ is convex in (x, v).

Theorem 5. *Suppose that H_0 satisfies the boundedness condition and $L(t, x, v)$ is convex in v. Let $l: R^n \times R^n \to (-\infty, +\infty]$ be an arbitrary lower semicontinuous function. For the sets*

(3.1) $$\left\{ x \in \mathcal{Q} \,\middle|\, \int_a^b L(t, x(t), \dot{x}(t))\, dt + l(x(a), x(b)) \leq \alpha \right\}, \quad \alpha \text{ real},$$

all to be compact in the weak topology of $\mathcal{Q}$, it is sufficient that the following condition be satisfied. Let C_1 and C_2 be the convex subsets of $R^n \times R^n$ defined by

(3.2) $$C_1 = \left\{ (d_0, d_1) \,\middle|\, \sup_{(c_0, c_1)} [c_0 \cdot d_0 - c_1 \cdot d_1 - l(c_0; c_1)] < +\infty \right\},$$

(3.3) $$C_2 = \{(p(a), p(b)) \mid \exists\, p \in \mathcal{Q}, (p(t), \dot{p}(t)) \in E(t) \text{ a.e.}\},$$

where $E(t)$ is given by (2.24). The condition is that the relative interiors of C_1 and C_2 have a nonempty intersection, and the affine hull of $C_1 \cup C_2$ is all of $R^n \times R^n$. Here C_2 has the same relative interior and closure as the convex set $C_2' \subset C_2$, where

(3.4) $$C_2' = \left\{ (p(a), p(b)) \mid p \in \mathcal{Q}, \exists\, \beta \in L^1[a, b], \\ \sup_{x \in R^n} [H(t, x, p(t)) + \dot{p}(t) \cdot x] \leq \beta(t) \right\}.$$

Proof. Let l_0 be the greatest, lower semicontinuous, convex function on $R^n \times R^n$ majorized by l. Assuming that the condition in question holds, the set C_1 is in particular nonempty, so that l_0 nowhere has the value $-\infty$. The function L_0 in Lemma 2 satisfies the regularity assumptions imposed on L, so that by Lemma 1 the functional

(3.5) $$\Phi(x) = \int_a^b L_0(t, x(t), \dot{x}(t))\, dt + l_0(x(a), x(b)), \quad x \in \mathcal{Q},$$

is well defined. Since l_0 and $L_0(t, \cdot, \cdot)$ are convex, we can apply to Φ our compactness result in [9, Theorem 1a]: the boundedness condition on H_0 is, in view of (2.3) and [9, Proposition 2], equivalent to condition (C_0) of [9], and the other property required by [9, Theorem 1a] is precisely the condition we have stated on the sets C_1 and C_2. Thus the sets

(3.6) $\{x \in \mathcal{C} | \Phi(x) \le \alpha\}, \quad \alpha$ real,

are all weakly compact. These sets contain the corresponding sets (3.1), since $L_0 \le L$ and $l_0 \le l$. It follows then from Lemma 3 that the sets (3.1) are likewise weakly compact, since a relatively weakly compact subset of a Banach space is weakly closed if and only if weakly sequentially closed. The fact that C_2 and C_2' have the same relative interior and closure is asserted by [9, Corollary 4 to Theorem 3], applied to the function M in (2.4), (2.27) and (2.28). (In the cited corollary, condition (D_0) is mislabelled (C_0). In the context of [9], the function M is dual to L_0 and therefore satisfies conditions (A), (B), (C) and (D_0), since L_0 satisfies (A), (B), (C_0) and (D). Thus the corollary is applicable.) This completes the proof of Theorem 5.

The proofs of Theorems 1, 2 and 3 also use a new device of reformulation that has wider theoretical applications. There are other methods, of course, for representing a problem with a variable t-interval in terms of a problem with a fixed t-interval, but these require that $L(t, x, v)$ be more than just measurable in t.

Reduction of variable interval case to fixed interval case. Each triple (x, t_0, t_1), where x is an absolutely continuous function from $[t_0, t_1]$ to R^n, $[t_0, t_1] \subset [a, b]$, may be identified with an absolutely continuous function $x' = (x, \tau_0, \tau_1)$ from $[a, b]$ to R^{n+2}, where $\tau_0(t) \equiv t_0, \tau_1(t) \equiv t_1, \dot{x}(t) = 0$ if $a \le t < t_0$ or $t_1 < t \le b$. Denote the set of all such functions x' by $\mathcal{C}'$. Define

(3.7) $L' : [a, b] \times R^{n+2} \times R^{n+2} \to (-\infty, +\infty]$

as follows, where $x' = (x, \tau_0, \tau_1)$ and $v' = (v, \sigma_0, \sigma_1)$:

(3.8) $L'(t, x', v') = +\infty$ unless $a \le \tau_0 \le \tau_1 \le b, \sigma_0 = 0, \sigma_1 = 0$,

and when the latter conditions are satisfied

$$L'(t, x', v') = L(t, x, v) \quad \text{if } \tau_0 < t < \tau_1,$$

(3.9) $= \delta_0(v) \qquad \text{if } t < \tau_0 \text{ or } t > \tau_1,$

$$= k(t, x, v) \quad \text{if } t = \tau_0 \text{ or } t = \tau_1,$$

where $\delta_0(v)$ is 0 if $v = 0$ and $+\infty$ if $v \ne 0$, and $k(t, x, \cdot)$ is the greatest l.s.c. convex function majorized by $L(t, x, \cdot)$ and δ_0. Then for an absolutely continuous function $x' : [a, b] \to R^{n+2}$ we have

(3.10) $$\int_a^b L'(t, x'(t), \dot{x}'(t))\, dt = \int_{t_0}^{t_1} L(t, x(t), \dot{x}(t))\, dt$$

if x' belongs to $\mathcal{Q}'$ and corresponds to $x: [t_0, t_1] \to R^n$, while

(3.11) $$\int_a^b L'(t, x'(t), \dot{x}'(t))\, dt = +\infty \quad \text{if } x' \notin \mathcal{Q}'.$$

Thus minimization problems for L with $[t_0, t_1]$ variable can be regarded as problems for L' over the fixed interval $[a, b]$.

To see by way of Lemma 1 that the integrals of L' make sense, observe that $L'(t, x', v')$ is again lower semicontinuous in (x', v') and measurable in (t, x', v') in the same sense that we have imposed on L. Furthermore, if H satisfies the boundedness condition, then so does the Hamiltonian H' corresponding to L', since for $p' = (p, \pi_0, \pi_1)$ we have

(3.12) $$H'(t, x', p') = -\infty \quad \text{unless } a \leq \tau_0 \leq \tau_1 \leq b,$$

and in the latter case

(3.13) $$\begin{aligned} H'(t, x', p') &= H(t, x, p) && \text{if } \tau_0 < t < \tau_1, \\ &= 0 && \text{if } t < \tau_0 \text{ or } t > \tau_1, \\ &= \max\{H(t, x, p), 0\} && \text{if } t = \tau_0 \text{ or } t = \tau_1. \end{aligned}$$

Note also that $L'(t, x', v')$ is convex in v' if $L(t, x, v)$ is convex in v.

Unfortunately, it is *not* necessarily true that if H_0 satisfies the boundedness condition, then the analogous function H_0' corresponding to L' satisfies the boundedness condition. The essential difficulty is that there may be no finite concave function majorizing both $H(t, \cdot, p)$ and 0, even though $H(t, \cdot, p)$ is majorized by such a function. However, this is a minor complication which, for our purposes, is easily obviated. We need only add to $\delta_0(v)$ in (3.9) the term

(3.14) $$\begin{aligned} \delta_B(x) &= 0 && \text{if } x \in B, \\ &= +\infty && \text{if } x \notin B, \end{aligned}$$

where B is some compact subset of $R^n \times R^n$. Then the identification (3.10) still holds provided that

(3.15) $$x(t_0) \in B \quad \text{and} \quad x(t_1) \in B.$$

The effect of this alteration on H' is to replace the 0 in (3.13) by $-\delta_B(x)$. It can then be verified that the boundedness condition on H_0 does imply the boundedness condition on H_0'.

(We sketch the proof of the last assertion. Fix $p \in R^n$. Assuming that H_0 satisfies the boundedness condition, we can actually find summable functions

$w: [a, b] \rightarrow R^n$ and $\alpha: [a, b] \rightarrow R^1$ such that

(3.16) $H_0(t, x, p) \leq w(t) \cdot x + \alpha(t)$ for all t and x

[9, Proposition 2]. Since B is compact, we can also find a summable function $\beta: [a, b] \rightarrow R^1$ such that

(3.17) $0 \leq w(t) \cdot x + \beta(t)$ for all t if $x \in B$.

We then have

(3.18) $\max\{H(t, x, p), -\delta_B(x)\} \leq w(t) \cdot x + |\alpha(t)| + |\beta(t)|$

for all x and t, and consequently

(3.19) $H'(t, x', p') \leq w(t) \cdot x + |\alpha(t)| + |\beta(t)|$.

Since the right side of (3.19) is affine in x, the left side can be replaced by $H_0'(t, x', p')$. Thus

(3.20) $H_0'(t, x', p') \leq |w(t)| \cdot r + |\alpha(t)| + |\beta(t)|$ if $|x'| \leq r$,

where the right side of (3.20) is summable in t. We conclude from this that H_0' satisfies the boundedness condition.)

Proof of Theorem 3. In establishing (a) and (b), it suffices to show that the assertions are valid for (t_0, t_1, c_0, c_1) in $S \times B \times B$, where B is a fixed compact subset of R^n. However, this case can be reduced by the reformulation device just described to the case of a fixed interval. Thus we need only prove that $F(a, b, \cdot, \cdot)$ is lower semicontinuous relative to $B \times B$, and that for each $(c_0, c_1) \in B \times B$ the infimum defining $F(a, b, c_0, c_1)$ is attained. To do this, we apply Theorem 5 with

$$l(c_0, c_1) = 0 \quad \text{if } c_0 \in B \text{ and } c_1 \in B,$$
(3.21)
$$= +\infty \quad \text{if } c_0 \notin B \text{ or } c_1 \notin B.$$

The compactness of B implies that the set C_1 in Theorem 5 is all of $R^n \times R^n$. On the other hand, C_2 is nonempty by the hypothesis of Theorem 3. We may therefore conclude from Theorem 5 that every set of the form

(3.22) $\left\{ x \in \mathcal{Q} \,\middle|\, \int_a^b L(t, x(t), \dot{x}(t)) \, dt \leq \alpha, \, x(a) \in B, \, x(b) \in B \right\}$

(α real) is weakly compact. The set

(3.23) $\{(c_0, c_1) \in B \times B| \, F(a, b, c_0, c_1) \leq \alpha\}$

is the image of (3.22) under the weakly continuous mapping $x \rightarrow (x(a), x(b))$. Hence every set of the form (3.23) is compact. In particular, $F(a, b, \cdot, \cdot)$ is

lower semicontinuous relative to $B \times B$. From the weak compactness of (3.22) we also see the weak compactness of

$$(3.24) \qquad \left\{ x \in \mathcal{C} \,\middle|\, \int_a^b L(t, x(t), \dot{x}(t))\, dt \le \alpha, \ x(a) = c_0, \ x(b) = c_1 \right\}$$

if $c_0 \in B$ and $c_1 \in B$. Since this is true for arbitrary real α, the infimum of $\int_a^b L(t, x(t), \dot{x}(t))\, dt$ subject to $x(a) = c_0$, $x(b) = c_1$, is attained by at least one arc x. This is of course the infimum defining $F(a, b, c_0, c_1)$, so assertions (a) and (b) are proved.

Assertion (c) of Theorem 3 is a consequence of the last part of Theorem 5, which implies in particular that, since the set C_2 is nonempty in the argument above, so is C_2'. Thus we can find functions $q \in \mathcal{C}$ and $\beta \in L^1[a, b]$ such that

$$(3.25) \qquad \beta(t) \ge H(t, x, q(t)) + \dot{q}(t) \cdot x \quad \text{for all } x \in R^n,$$

or in other words

$$(3.26) \qquad L(t, x, v) \ge \dot{q}(t) \cdot x + q(t) \cdot v - \beta(t) \quad \text{for all } (x, v) \in R^n \times R^n.$$

It follows that

$$(3.27) \qquad \begin{aligned} \int_{t_0}^{t_1} L(t, x(t), \dot{x}(t))\, dt &\ge \int_{t_0}^{t_1} [\dot{q}(t) \cdot x(t) + q(t) \cdot \dot{x}(t) - \beta(t)]\, dt \\ &= q(t_1) \cdot x(t_1) - q(t_0) \cdot x(t_0) - \int_{t_0}^{t_1} \beta(t)\, dt, \end{aligned}$$

and hence

$$(3.28) \qquad F(t_0, t_1, c_0, c_1) \ge q(t_1) \cdot c_1 - q(t_0) \cdot c_0 - \int_{t_0}^{t_1} \beta(t)\, dt.$$

Proof of Theorem 2. Applying Lemma 5, we consider the differential equation

$$(3.29) \qquad \dot{p}(t) = N(t) p(t) + j(t) + |p(t)| k(t) \quad \text{a.e.}$$

The summability of $N(t)$, $j(t)$ and $k(t)$ in t yields a global Lipschitz property of (3.29) which implies in particular the existence of a solution over the whole interval $[a, b]$. Any such solution p satisfies

$$(3.30) \qquad (p(t), \dot{p}(t)) \in E(t) \quad \text{a.e.}$$

by virtue of (2.29), and hence the hypothesis of Theorem 3 is satisfied. Therefore (a) and (b) of Theorem 2 are true.

To prove (c) of Theorem 2, we note first that for each $p_0 \in R^n$ there is actually a unique solution $p \in \mathcal{C}$ to (3.29) with $p(a) = p_0$. Furthermore, in view of (3.30) each solution to (3.29) satisfies $(p(a), p(b)) \in C_2$, where C_2 is the convex set introduced in (3.3). Thus C_2 has all of R^n as its image under the projection

$(p_0, p_1) \rightarrow p_0$, and hence ri C_2 (the relative interior of C_2) likewise has all of R^n as its image under this projection. Therefore, we can select $b_1 \in R^n$ with $(0, b_1) \in$ ri C_2.

We now consider the differential equation

$$(3.31) \qquad \dot{q}(t) = N(t)q(t) + |q(t)|k(t),$$

which similarly has for each initial point $q_0 \in R^n$ a unique solution $q \in \mathcal{C}$ with $q(a) = q_0$. Denote the value of $q(b)$ for this solution by $Q(q_0)$; observe that

$$(3.32) \qquad Q(\lambda q_0) = \lambda Q(q_0) \quad \text{if } \lambda \geq 0.$$

If p satisfies (3.29) and q satisfies (3.31), we have for any $\lambda \geq 0$

$$(3.33) \qquad \dot{p}(t) + \lambda \dot{q}(t) = N(t)[p(t) + \lambda q(t)] + j(t) + \sigma(t)k(t),$$

where

$$\sigma(t) = |p(t)| + \lambda|q(t)| \geq |p(t) + \lambda q(t)|,$$

so that

$$(3.34) \qquad (p(t) + \lambda q(t), \dot{p}(t) + \lambda \dot{q}(t)) \in E(t) \quad \text{a.e.}$$

by (2.29). Then $(p(a) + \lambda q(a), p(b) + \lambda q(b)) \in C_2$. In other words, for each $q_0 \in R^n$ the convex set C_2 contains half-lines of the form $\{(p_0, p_1) + \lambda(q_0, Q(q_0)) \mid \lambda \geq 0\}$. Each pair $(q_0, Q(q_0))$ therefore belongs to the recession cone of cl C_2, which is the same as the recession cone of ri C_2 [13, Corollary 8.3.1]. In particular,

$$(3.35) \qquad (0, b_1) + (q_0, Q(q_0)) \in \text{ri } C_2 \quad \text{for all } q_0 \in R^n.$$

We recall now from Theorem 5 that C_2 has the same relative interior as the convex set C_2' in (3.4). This enables us to infer from (3.35) that

$$(3.36) \qquad (q_0, Q(q_0) + b_1) \in \text{ri } C_2' \quad \text{for all } q_0 \in R^n.$$

Let $\{q_0^i \mid i = 0, \cdots, n\}$ be the vertex set of a simplex in R^n which includes the unit ball. There exist a unique matrix $\overline{A}_1 \in R^{n \times n}$ and vector $a_1 \in R^n$ such that

$$(3.37) \qquad \overline{A}_1 q_0^i + a_1 = Q(q_0^i), \quad i = 0, 1, \cdots, m.$$

If q_0 is any nonzero vector in R^n, the vector $|q_0|^{-1}q_0$ belongs to the unit ball and hence is expressible as a convex combination of the q_0^i:

$$(3.38) \qquad |q_0|^{-1}q_0 = \sum_{i=0}^{m} \lambda_i q_0^i.$$

We then have

(3.39) $$\bar{A}_1(|q_0|^{-1}q_0) + a_1 = \sum_{i=0}^{m} \lambda_i(\bar{A}_1 q_0^i + a_1),$$

so that by (3.37) and (3.32)

(3.40) $$\bar{A}_1(q_0) + |q_0|a_1 = |q_0| \sum_{i=0}^{m} \lambda_i Q(q_0^i) = \sum_{i=0}^{m} \lambda_i Q(|q_0|q_0^i).$$

It follows from (3.36) and the convexity of ri C_2' that

(3.41) $$(q_0, \bar{A}_1 q_0 + |q_0|a_1 + b_1) = \sum_{i=0}^{m} \lambda_i(|q_0|q_0^i, Q(|q_0|q_0^i) + b_1) \in \text{ri } C_2'.$$

We have shown this for $q_0 \neq 0$, but it also holds trivially for $q_0 = 0$ by (3.36). Thus

(3.42) $$(q_0, \bar{A}_1 q_0 + |q_0|a_1 + b_1) \in \text{ri } C_2' \quad \text{for all } q_0 \in R^n.$$

Now if $(p_0, p_1) \in C_2'$, there exist by definition functions $p \in \mathcal{Q}$ and $\beta \in L^1[a, b]$ such that $p(a) = p_0$, $p(b) = p_1$, and (3.25) holds. This implies, as we have already seen in the proof of Theorem 3, that (3.28) holds, and in particular

(3.43) $$F(a, b, c_0, c_1) \geq p_1 \cdot c_1 - p_0 \cdot c_0 - \int_a^b \beta(t)\, dt.$$

Therefore

(3.44) $$G(p_0, p_1) < +\infty \quad \text{for all } (p_0, p_1) \in C_2',$$

where

(3.45) $$G(p_0, p_1) = \sup\{p_1 \cdot c_1 - p_0 \cdot c_0 - F(a, b, c_0, c_1)\} > -\infty.$$

(The inequality $G(p_0, p_1) > -\infty$ results from our basic assumption that for at least one $x \in \mathcal{Q}$ the function $t \to L(t, x(t), \dot{x}(t))$ is majorized by a summable function. In other words, there is at least one pair (c_0, c_1) such that $F(a, b, c_0, c_1) < +\infty$.) The function G is convex according to its definition (3.45). Since C_2' is a convex set, (3.44) implies that G is continuous relative to ri C_2'. Combining this fact with (3.42), we see that the function

(3.46) $$p_0 \to G(p_0, \bar{A}_1 p_0 + |p_0|a_1 + b_1)$$

is finite and continuous throughout R^n. Let

(3.47) $$g(\mu) = \max_{|p_0| \leq \mu} G(p_0, \bar{A}_1 p_0 + |p_0|a_1 + b_1) < +\infty.$$

Reasoning from (3.45), we deduce that for every $p_0 \in R^n$

$$(3.48) \qquad F(a, b, c_0, c_1) \geq (\overline{A}_1 p_0 + |p_0| a_1 + b_1) \cdot c_1 - p_0 \cdot c_0 - g(|p_0|)$$

and consequently, letting A_1 denote the transpose of $\overline{A}_1$,

$$F(a, b, c_0, c_1) \geq b_1 \cdot c_1 + \sup_{p_0 \in R^n} \{(A_1 c_1 - c_0) \cdot p_0 + |p_0| a_1 \cdot c_1 - g(|p_0|)\}$$

$$(3.49) \qquad = b_1 \cdot c_1 + \sup_{\mu \geq 0} \{\mu |c_0 - A_1 c_1| + \mu a_1 \cdot c_1 - g(\mu)\}$$

$$= b_1 \cdot c_1 + \gamma_1(|c_0 - A_1 c_1| - a_1 \cdot c_1),$$

with

$$(3.50) \qquad \gamma_1(\lambda) = \sup_{\mu \geq 0} \{\mu \lambda - g(\mu)\}.$$

The finiteness of $g(\mu)$ for all $\mu \in [0, +\infty)$ allows us to conclude from (3.50) that γ_1 is a nondecreasing, convex function which is bounded below by $-g(0)$ and satisfies

$$(3.51) \qquad \lim_{\lambda \to +\infty} \gamma_1(\lambda)/\lambda = +\infty.$$

By a parallel argument (with the above roles of p_0 and p_1 reversed), we can find a similar function γ_0, matrix A_0, and vectors a_0 and b_0 such that

$$(3.52) \qquad F(a, b, c_0, c_1) \geq b_0 \cdot c_0 + \gamma_0(|c_1 - A_0 c_0| + a_0 \cdot c_0).$$

A function γ can be constructed such that $\gamma \leq \gamma_0$, $\gamma \leq \gamma_1$, and γ too is non-decreasing, convex, bounded below and satisfies (3.52). Then $F(a, b, \cdot, \cdot)$ satisfies (1.5) and the growth condition (G_∞) holds as asserted.

Proof of Theorem 1. Choosing any s such that $2r < s < +\infty$, we consider the differential equation

$$(3.53) \qquad \dot{p}(t) = N(t)p(t) + j(t) \quad \text{a.e.,}$$

where $N(t)$ and $j(t)$ have the property in Lemma 4. Let B_r denote the closed ball in R^n with center 0 and radius r. Since the components of $N(t)$ and $j(t)$ are summable in t, there is an $\epsilon > 0$ such that, over interval $[\overline{a}, \overline{b}] \subset [a, b]$ of length $\leq 2\epsilon$, the differential equation has a (unique) solution in B_s emanating from each point of B_{2r}, as well as a unique solution in B_s terminating at each point of B_{2r}. Each solution p to (3.53) over $[\overline{a}, \overline{b}]$ that stays inside B_s satisfies

$$(3.54) \qquad (p(t), \dot{p}(t)) \in E(t) \quad \text{for almost every } t \in [\overline{a}, \overline{b}]$$

by (2.25). In particular, therefore, Theorem 3 can be applied with $[a, b]$ replaced by any subinterval $[\bar{a}, \bar{b}]$ of length $\leq 2\epsilon$. This shows that conclusions (a) and (b) of Theorem 1 are valid.

For notational simplicity in the proof of (c), it suffices to assume that the interval $[a, b]$ itself is of length $\leq 2\epsilon$ and to demonstrate that then $F(a, b, \cdot, \cdot)$ satisfies the growth condition (G_r). Thus we can assume that for each $p_0 \in B_{2r}$ the solution p to (3.53) with $p(a) = p_0$ satisfies .

$$(3.55) \qquad (p(t), \dot{p}(t)) \in E(t) \quad \text{for almost every } t \in [\bar{a}, \bar{b}].$$

Let $P_0(p_0)$ denote the value of $p(b)$ for this solution; of course, P_0 is an affine transformation. For the set C_2 in (3.3) we have

$$(3.56) \qquad (p_0, P_0(p_0)) \in C_2 \quad \text{if } p_0 \in B_{2r}.$$

We see in particular from (3.32) that the image of the convex set C_2 under the projection $(p_0, p_1) \to p_0$ includes B_{2r}. Hence the image of ri C_2 under the same projection (where "ri" denotes relative interior) includes the interior of B_{2r}. Select an arbitrary $d \in R^n$ such that $(0, d) \in$ ri C_2. Then

$$(3.57) \qquad (2/3)(p_0, P_0(p_0)) + (1/3)(0, d) \in \text{ri } C_2 \quad \text{if } p_0 \in B_{2r},$$

so that if we define

$$(3.58) \qquad P(p_0) = (2/3)P_0((3/2)p_0) + (1/3)d \cdot$$

we have

$$(3.59) \qquad (p_0, P(p_0)) \in \text{ri } C_2 \quad \text{if } p_0 \in B_r.$$

Here P is again affine. Once more we use the fact provided by Theorem 5 that C_2 has the same relative interior as the set C_2' in (3.4), and that the convex function G in (3.45) is finite and continuous on ri C_2'. This implies in view of (3.59) that the function $p_0 \to G(p_0, P(p_0))$ is bounded above on B_r, say by α_0. We see then from (3.45) that

$$(3.60) \qquad F(a, b, c_0, c_1) \geq P(p_0) \cdot c_1 - p_0 c_0 - \alpha_0 \quad \text{if } p_0 \in B_r.$$

Since P is an affine mapping, we have $P(p_0) = \bar{A}_1 p_0 + b_1$, where $\bar{A}_1 \in R^{n \times n}$ and $b_1 \in R^n$. Substituting this expression in (3.60) and letting A_1 denote the transpose of $\bar{A}_1$, we obtain

$$(3.61) \qquad F(a, b, c_0, c_1) \geq \sup_{|p_0| \leq r} \{(\bar{A}_1 p_0 + b_1) \cdot c_1 - p_0 \cdot c_0 - \alpha_0\}$$

$$= -\alpha_0 + b_1 \cdot c_1 + r|c_0 - A_1 c_1|.$$

By a parallel argument, we have

(3.62) $F(a, b, c_0, c_1) \geq -\alpha_1 + b_0 \cdot c_0 + r|c_1 - A_0' c_0|.$

Therefore growth condition (G_r) is satisfied by $F(a, b, \cdot, \cdot)$ for

(3.63) $\gamma(\lambda) = r \max\{\lambda, 0\} - \max\{\alpha_0, \alpha_1\}.$

4. **Application to control problems.** The preceding sections have dealt with
minimization problems which are of the same form as the classical "free" prob-
lems of the calculus of variations, except that the Lagrangian function L is
allowed to have the value $+\infty$. Actually, the exception makes a world of differ-
ence, because it enables us to apply the same results to problems like those in
control theory, where there may be complicated constraints. This application
can be carried out much in the manner of Cesari [1], and Olech [7]. However, we
take a different approach which yields somewhat sharper conclusions.

It is helpful at the outset to formulate the general control problem more
abstractly than is customary. We assume we are given a function

(4.1) $K : [a, b] \times R^n \times R^n \times R^m \rightarrow (-\infty, +\infty]$

such that $K(t, x, v, u)$ is lower semicontinuous in (x, v, u) and measurable in
(t, x, v, u) (with respect to the σ-algebra generated by products of Lebesgue
sets in $[a, b]$ and Borel sets in $R^n \times R^n \times R^m$). The basic Lagrange problem is
to minimize the integral

(4.2) $\int_{t_0}^{t_1} K(t, x(t), \dot{x}(t), u(t))\, dt$

over all measurable control functions $u : [t_0, t_1] \rightarrow R^m$ and all absolutely con-
tinuous functions $x : [t_0, t_1] \rightarrow R^n$ satisfying

(4.3) $x(t_0) = c_0 \quad \text{and} \quad x(t_1) = c_1.$

Here $[t_0, t_1]$ is a fixed subinterval of $[a, b]$. (To avoid trivial technical nui-
sances, we assume that $K(t, \cdot, \cdot, \cdot)$ is not identically $+\infty$ for any $t \in [a, b]$,
and that it is possible to find at least one absolutely continuous $x : [a, b] \rightarrow R^n$
and measurable $u : [a, b] \rightarrow R^n$ such that $K(t, x(t), \dot{x}(t), u(t))$ is majorized by a
summable function of t.)

Of course, endpoint conditions more general than (4.3) could be considered
or a term $k(t_0, t_1, x(t_0), x(t_1))$ could be added to (4.2). But this would be
irrelevant for our present purposes. Our aim is rather to deduce under certain
assumptions the equivalence of the above problem and the problem of minimizing
$\int_{t_0}^{t_1} L(t, x(t), \dot{x}(t))\, dt$ subject to the same constraints (4.3), where

(4.4) $$L(t, x, v) = \inf \{ K(t, x, v, u) \mid u \in R^m \}.$$

Once this is accomplished, the theorems in this paper can be applied to the basic control problem, and it is a simple matter to extend the coverage to other model problems. For example, the compactness theorem in §3 for Bolza functionals, as well as the device explained in §3 for reformulating the variable interval case, can be brought to bear.

Before going further, we discuss the relationship between our abstract control problem and the usual one. In the usual problem, one seeks to minimize

(4.5) $$\int_{t_0}^{t_1} f_0(t, x(t), u(t)) \, dt$$

over all measurable functions $u: [t_0, t_1] \to R^m$ and absolutely continuous functions $x: [t_0, t_1] \to R^n$ satisfying (4.3) and the constraints

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

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

(4.8) $$x(t) \in X(t) \quad \text{a.e.}$$

To put this in the above form, we simply define

(4.9) $$K(t, x, v, u) = f_0(t, x, u) \quad \text{if } x \in X(t), \ u \in U(t, x),$$
$$\text{and } v = f(t, x, u),$$
$$= +\infty \quad \text{in all other cases.}$$

It is clear that then the integral (4.2) (assuming it is well defined—see below) has the value $+\infty$ unless the constraints (4.6), (4.7) and (4.8) are satisfied, so that minimizing (4.2) subject to (4.3) is identical to the control problem just stated.

Lemma 6. *Assume in the forgoing that the set*

(4.10) $$G = \{(t, x, u) \in [a, b] \times R^n \times R^m \mid x \in X(t), \ u \in U(t, x)\}$$

is measurable (in the specified sense) and the functions $f_0: G \to R^1$ *and* $f: G \to R^n$ *are likewise measurable. Assume that for each* $t \in [a, b]$ *and* $\alpha \in R^1$, *the set*

(4.11) $$G_{t, \alpha} = \{(x, u) \in R^n \times R^m \mid x \in X(t), \ u \in U(t, x), \ f_0(t, x, u) \leq \alpha\}$$

is closed, and relative to it the mapping $f(t, \cdot, \cdot)$ *is continuous. Then* $K(t, x, v, u)$ *in (4.9) is indeed lower semicontinuous in* (x, v, u) *and measurable (in the specified sense) in* (t, x, u, v).

Proof. Fix $\alpha \in R^1$ and consider the sets

$$(4.12) \qquad \{(t, x, v, u)|\ (t, x, u) \in G,\ f_0(t, x, u) \leq \alpha,\ f(t, x, u) - v = 0\},$$

$$(4.13) \qquad \{(x, v, u)|\ (x, u) \in G_{t, \alpha},\ f(t, x, u) - v = 0\}, \qquad t \in [a, b].$$

The question is whether the sets (4.13) are closed and (4.12) is measurable. The fact that the sets (4.13) are closed is obvious from the closedness of $G_{t, \alpha}$ and continuity of $f(t, \cdot, \cdot)$ relative to $G_{t, \alpha}$. Our assumptions imply the measurability of the mapping $(t, x, v, u) \to f(t, x, u) - v$, so that the measurability of (4.12) is also immediate.

The regularity assumptions in Lemma 6 are in several respects weaker than those made by other authors in treating the existence of solutions to control problems, most recently Cesari [1] and Olech [7]. The sets $U(t, x)$ and $X(t)$ are not required to depend "upper semicontinuously" on t, and, as a matter of fact, they do not even have to be closed. Furthermore, $f_0(t, \cdot, \cdot)$ is not required to be continuous, but only to have a certain lower semicontinuity property. Thus, for example, cases are covered where $f_0(t, x, u)$ grows to $+\infty$ as u approaches the boundary of $U(t, x)$, or as x approaches certain boundary points of $X(t)$. Such cases are excluded by the assumptions of Cesari and Olech.

We should mention in connection with Lemma 6 that, if $g: [a, b] \times W \to R^k$ has the property that $g(t, w)$ is Lebesgue measurable in t and continuous in w (W being a subset of R^d), then g is measurable with respect to the σ-algebra in $[a, b] \times W$ generated by products of Lebesgue sets in $[a, b]$ and Borel sets in W. This is apparent from the representation in the case of $R^k = R^1$ that

$$\{(t, w)|\ g(t, w) \leq \alpha\} = \bigcap_{j=1}^{\infty} \bigcup_{i=1}^{\infty} T_{ij} \times W_{ij},$$

where for a dense sequence $\{w_i\}$ in W one defines

$$T_{ij} = \{t \in [a, b]|\ g(t, w_i) \leq \alpha + (1/j)\},$$
$$W_{ij} = \{w \in W|\ |w - w_i| \leq 1/j\}.$$

In the example where K satisfies (4.9), the function L (which throughout this section is given by (4.4)) can be expressed by

$$(4.14) \qquad L(t, x, v) = \inf\{\alpha \in R^1|\ (v, \alpha) \in \tilde{Q}(t, x)\},$$

where

$$(4.15) \qquad \tilde{Q}(t, x) = \{(f(t, x, u), f_0(t, x, u))|\ x \in X(t),\ u \in U(t, x)\}.$$

Thus L is convex in v if $\widetilde{Q}(t, x)$ is convex, or more generally if the set

$$(4.16) \qquad \widetilde{Q}(t, x) = \{(f(t, x, u), f_0(t, x, u) + s) \mid x \in X(t), u \in U(t; x), s \geq 0\}$$

is convex. Both Cesari and Olech deal directly with (4.16), imposing continuity assumptions on its behavior with respect to both t and x, etc., in addition to stronger forms of the assumptions in Lemma 6.

We avoid any such direct assumptions below. This is achieved through a boundedness condition on

$$(4.17) \qquad J(t, x, p, q) = \sup\{p \cdot v + q \cdot u - K(t, x, v, u) \mid (v, u) \in R^n \times R^m\},$$

rather than just on

$$(4.18) \qquad H(t, x, p) = \sup\{p \cdot v - L(t, x, v) \mid v \in R^n\} = J(t, x, p, 0).$$

The function J is said, of course, to satisfy the boundedness condition if

$$(4.19) \qquad \sup_{|x| \leq r} J(t, x, p, q) \leq \phi(t, r, p, q),$$

where ϕ is some real-valued function on $[a, b] \times [0, +\infty) \times R^n \times R^m$ such that $\phi(t, r, p, q)$ is summable in t for fixed (r, p, q).

If K is given by (4.9), as in Lemma 6, one has

$$(4.20) \qquad J(t, x, p, q) = \sup_{u \in U(t,x)} \{q \cdot u + p \cdot f(t, x, u) - f_0(t, x, u)\}$$

if $x \in X(t)$, whereas $J(t, x, p, q) = -\infty$ if $x \notin X(t)$. (If $U(t, x)$ is empty in (4.20), the convention is used that the supremum of the empty set of real numbers is $-\infty$.) It should be emphasized, however, that we are not limiting ourselves to this special case. The following theorem applies equally to the more abstract control problem formulated above.

Theorem 6. *Assume that J satisfies the boundedness condition. Then the integral (4.2) is well defined (in the sense of Lemma 1) for any absolutely continuous* $x: [t_0, t_1] \to R^n$ *and measurable* $u: [t_0, t_1] \to R^m$, *and if its value is not* $+\infty$ *the function u must be summable. Furthermore,*

(a) *$L(t, x, v)$ is lower semicontinuous in (x, v), measurable (in the sense described) in (t, x, v), nowhere $-\infty$, and H satisfies the boundedness condition;*

(b) *for every absolutely continuous* $x: [t_0, t_1] \to R^m$, *one has (where the minimum is attained):*

$$(4.21) \qquad \int_{t_0}^{t_1} L(t, x(t), \dot{x}(t)) \, dt = \min\left\{ \int_{t_0}^{t_1} K(t, x(t), \dot{x}(t), u(t)) \, dt \mid u \text{ measurable} \right\}.$$

Proof. We can regard K as a Lagrangian in (t, x, y, v, u) that happens to be independent of y, with the u variables playing the role of the derivatives of the y variables. Then J is the Hamiltonian corresponding to K, and Lemma 1 is directly applicable. This verifies that the integral of K is always well defined. Next, fix any $r > 0$ and $s > 0$ and choose a finite subset (p_i, q_i) of $R^n \times R^m$ whose convex hull contains

$$(4.22) \qquad \{(p, q) \in R^n \times R^n| \; |p| \leq 1 \text{ and } |q| \leq 1\}.$$

For the function ϕ in the boundedness condition on J, define

$$(4.23) \qquad \psi_r(t) = \max_i \phi(t, r, p_i, q_i).$$

Then, since $J(t, x, p, q)$ is convex in (p, q), we have

$$(4.24) \qquad J(t, x, p, q) \leq \psi_r(t) \quad \text{if } |x| \leq r, \; |p| \leq 1, \; |q| \leq 1,$$

where ψ_r is summable. This implies from the definition of J that

$$(4.25) \qquad K(t, x, v, u) \geq |v| + |u| - \psi_r(t) \quad \text{if } |x| \leq r.$$

The summability of $u(t)$ in t, whenever the integral (4.2) is finite, is then obvious. It is also obvious from (4.25) that for each $\alpha \in R^1$ and $t \in [a, b]$ the set

$$(4.26) \qquad \{(x, v, u)| \; K(t, x, v, u) \leq \alpha, \; |x| \leq r\}$$

is bounded. This set is also closed, in view of the lower semicontinuity assumption on K, and hence it is compact (for any real α and $r > 0$). Therefore the infimum defining $L(t, x, v)$ in (4.4) is attained (not $-\infty$) by at least one u for each (t, x, v), and the set

$$(4.27) \qquad \{(x, v)| \; L(t, x, v) \leq \alpha, \; |x| \leq r\},$$

being the image of (4.26) under the projection $(x, v, u) \rightarrow (x, v)$, is compact (for any real α and $r > 0$). In particular, $L(t, x, v)$ is lower semicontinuous in (x, v).

The desired measurability property of L is more difficult to verify, and it seems most efficient to invoke the theory of measurable, closed-valued multifunctions, as this will also assist in proving (b). (By definition, a multifunction $Z: [a, b] \rightarrow R^k$ is measurable if for each closed set $B \subset R^k$ the set of $t \in [a, b]$ such that $Z(t) \cap B \neq \emptyset$ is Lebesgue measurable in $[a, b]$.) For each $t \in [a, b]$, let

$$(4.28) \qquad Z_K(t) = \{(x, v, u, \alpha) \in R^n \times R^n \times R^m \times R^1| \; \alpha \geq K(t, x, v, u)\},$$

$$(4.29) \qquad Z_L(t) = \{(x, v, u, \alpha) \in R^n \times R^n \times R^m \times R^1| \; \alpha \geq L(t, x, v)\}.$$

Thus $Z_K(t)$ is the epigraph of K, while $Z_L(t)$ is the epigraph, not precisely of L, but of L regarded also as a (constant) function of u. The lower semicontinuity of K and L is equivalent to the closedness of $Z_K(t)$ and $Z_L(t)$. Moreover, the measurability of K (with respect to the σ-algebra generated by products of Lebesgue sets in $[a, b]$ and Borel sets in $R^n \times R^n \times R^m$) is equivalent to the measurability of the multifunction

(4.30) $$Z_K : [a, b] \to R^n \times R^n \times R^m$$

(see [11] or [12, Theorem 1]). Similarly for the desired measurability property of L; thus to obtain this property we need only verify that the multifunction Z_L is measurable. To do this, we observe that, since the infimum in (4.4) is always attained, as already demonstrated, we have

(4.31) $$Z_L(t) = Z_K(t) + Z(t) \quad \text{for all } t,$$

where

(4.32) $$Z(t) \equiv \{(0, 0, u, 0) \in R^n \times R^n \times R^m \times R^1 \mid u \text{ arbitrary}\}.$$

Trivially, the multifunction $Z: t \to Z(t)$ is measurable. It is known that the (closure of) the sum of two measurable multifunctions is measurable [11, Corollary 1.2], and hence Z_L is measurable. Therefore L is measurable in the sense described. The fact that H satisfies the boundedness condition, since J does, is immediate from (4.18). Thus (a) is proved.

The proof of (b) is relatively easy, using what has already been established. Let $x: [t_0, t_1] \to R^n$ be absolutely continuous. We know that for each t there is at least one choice of $u(t) \in R^n$ such that

(4.33) $$L(t, x(t), \dot{x}(t)) = K(t, x(t), \dot{x}(t), u(t)).$$

The problem is to show that $u(t)$ can be chosen measurably, and we need only be concerned with the case where the integral on the left of (4.21) is finite. In this case $L(t, x(t), \dot{x}(t))$ is summable, hence finite almost everywhere, say for t in a certain subset T of $[a, b]$. Let

(4.34) $$Z_x(t) = \{(x(t), \dot{x}(t), u, \alpha(t)) \mid u \text{ arbitrary}\}, \quad t \in T,$$

where

(4.35) $$\alpha(t) = L(t, x(t), \dot{x}(t)).$$

Then Z_x is a measurable, closed-valued multifunction from T to $R^n \times R^n \times R^m \times R^1$, as is Z_K. The intersection $Z_K \cap Z_x: t \to Z_K(t) \cap Z_x(t)$, $t \in T$, is therefore likewise a measurable, closed-valued multifunction [9, Corollary 1.3], and it is also nonempty-valued. It follows (from a result proved independently

by Rokhlin, Castaing, and Kuratowski and Ryll-Nardzewski—see [11, Corollary 1.1] or [12, Theorem 1]) that $Z_K \cap Z_x$ possesses a measurable selection. In other words, there is at least one measurable function on T whose value for each t lies in $Z_K(t) \cap Z_x(t)$. But this means precisely that $u(t)$ can be chosen measurably in (4.33) for $t \in T$, and the choice can be made arbitrarily $t \in [a, b]/T$. Theorem 6 is thereby proved.

   Remarks. When Theorem 6 is combined with Theorem 4 (Olech [7]) and Lemma 6, one obtains a result more general than that described in Olech's paper for the same problem. (The regularity assumptions are weaker.) It is interesting that, when Theorem 6 is combined with Theorems 1, 2, 3 and 5, the assumptions on H_0 and H_1 (constructed from H in (4.18)) suffice—there is no need to introduce the analogous functions J_0 and J_1 corresponding to J and depending on q as well as p. The existence of optimal arcs in problems where $[t_0, t_1]$ is allowed to vary (as in the work of Cesari, Olech, and others) is obtained immediately from the lower semicontinuity of the minimum value function $F(t_0, t_1, c_0, c_1)$ in (t_0, t_1). Similarly for problems where c_0 and c_1 vary, except that there the established growth properties of $F(t_0, t_1, c_0, c_1)$ with respect to (c_0, c_1) enter in.

REFERENCES

   1. L. Cesari, *Existence theorems for weak and usual optimal solutions in Lagrange problems with unilateral constraints*. I, Trans. Amer. Math. Soc. 124 (1966), 369–412. MR 34 #3392.

   2. ———, *Closure, lower closure, and semicontinuity theorems in optimal control*, SIAM J. Control 9 (1971), 287–315.   MR 44 #4607.

   3. A. F. Filippov, *On certain questions in the theory of optimal control*, Vestnik Moskov. Univ. Ser. Mat. Meh. Astronom. Fiz. Him. 1959, no. 2, 25–32; English transl., J. SIAM Control Ser. A 1 (1962), 76–84.   MR 22 #13373;   MR 26 #7469.

   4. E. B. Lee and L. Markus, *Optimal control for nonlinear processes*, Arch. Rational Mech. Anal. 8 (1961), 36–58.   MR 23 #B1610.

   5. E. J. McShane, *Existence theorems for ordinary problems of the calculus of variation*, Ann. Scuola Norm. Sup. Pisa 3 (1934), 181–211.

   6. M. Nagumo, *Über die gleichmässige Summierbarkeit und ihre Anwendung auf ein Variationsproblem*, Japan. J. Math. 6 (1929), 178–182.

   7. C. Olech, *Existence theorems for optimal problems with vector-valued cost functions*, Trans. Amer. Math. Soc. 136 (1969), 159–180.   MR 38 #2655.

   8. ———, *Existence theorems for optimal control problems involving multiple integrals*, J. Differential Equations 6 (1969), 512–526.   MR 39 #7488.

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

   10. ———, *State constraints in convex problems of Bolza*, SIAM J. Control 10 (1972), 691–715.

   11. ———, *Measurable dependence of convex sets and functions on parameters*, J. Math. Anal. Appl. 28 (1969), 4–25.   MR 40 #288.

12. R. T. Rockafellar, *Convex integral functionals and duality*, Contributions to Nonlinear Functional Analysis, Academic Press, New York, 1971, pp. 215–236.

13. ———, *Convex analysis*, Princeton, Univ. Press, Princeton Math. Series, no 28, Princeton, N. J., 1970. MR 43 #445.

14. E. O. Roxin, *The existence of optimal controls*, Michigan Math. J. 9 (1962), 109–119. MR 25 #305.

15. L. Tonneli, *Sugli integrali del calcolo delle variazioni in forma ordinaria*, Ann. Scuola Norm. Sup. Pisa 3 (1934), 401–450.

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF WASHINGTON, SEATTLE, WASHINGTON 98195