PACIFIC JOURNAL OF MATHEMATICS
Vol. 33, No. 2, 1970

GENERALIZED HAMILTONIAN EQUATIONS FOR CONVEX PROBLEMS OF LAGRANGE

R. TYRRELL ROCKAFELLAR

Many nonclassical problems in the calculus of variations, arising for example from control theory, correspond in a sense to "Hamiltonian" functions which are not everywhere differentiable, but are convex in one vector argument and concave in the other. Optimal arcs in such problems satisfy generalized ordinary differential equations, defined in terms of subgradients of the "Hamiltonian." Such equations are treated in this paper by convexity methods. An existence theorem is derived from a result of Castaing, and various properties of solutions are established.

A *convex problem of Bolza*, according to our terminology in a preceding paper [6], is a variational problem of the form: minimize

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

over all absolutely continuous arcs $x: [0, T] \to R^n$ (T fixed), where l is a convex function from $R^n \times R^n$ to $R^1 \cup \{+\infty\}$, $L(t, \cdot, \cdot)$ is a convex function from $R^n \times R^n$ to $R^1 \cup \{+\infty\}$ for each $t \in [0, T]$, and l and L satisfy certain basic regularity conditions (given in [6]). Such a problem is called a *convex problem of Lagrange* in the special case where l is of the form

$$l(x(0), x(T)) = \begin{cases} 0 \text{ if } x(0) = a \text{ and } x(T) = b\,, \\ +\infty \text{ if } x(0) \neq a \text{ or } x(T) \neq b\,, \end{cases} \tag{1.2}$$

since then, in effect, one minimizes

$$\int_0^T L(t, x(t), \dot{x}(t))dt \tag{1.3}$$

subject to

$$x(0) = a \quad \text{and} \quad x(T) = b\,. \tag{1.4}$$

The class of convex problems of Bolza was introduced in [6] because of its duality properties, and because it could be studied extensively by convexity methods, without resorting to differentiability assumptions. In particular, we showed it was possible, by means of the theory of subgradients of convex functions, to define generalized Hamiltonian differential equations (with multivalued right-hand sides) which serve, along with certain transversality conditions, to characterize, the optimal arcs in such problems. In a convex problem of Lagrange, the transversality condition is trivial; it reduces to (1.4).

The *generalized Hamiltonian equation* for a convex problem of Lagrange is:

$$(1.5) \qquad (-\dot{p}(t), \dot{x}(t)) \in \partial H(t, x(t), p(t)) \text{ for almost every } t ,$$

where H, the *Hamiltonian* function in the problem, is defined in terms of the given Lagrangian function L by

$$(1.6) \qquad H(t, x, p) = \sup_{v \in R^n} \{\langle v, p\rangle - L(t, x, v)\} .$$

($\langle\cdot,\cdot\rangle$ denotes the inner product in R^n.) Here H is an extended-real-valued function on $[0, T] \times R^n \times R^n$ and, by virtue of the convexity of $L(t, x, v)$ in (x, v), $H(t, x, p)$ is concave as a function of x for every (t, p) and convex as a function of p for every (t, x) (see [7, Th. 33.1]). The symbol $\partial H(t, x, p)$ denotes the set of all *subgradients* of the concave-convex function $H(t, \cdot,\cdot)$ at the point (x, p). Thus $\partial H(t, x, p)$ is the set of all $(u, v) \in R^n \times R^n$ such that

$$(1.7a) \qquad H(t, x, p') \geqq H(t, x, p) + \langle v, p'-p\rangle \text{ for all } p' \in R^n ,$$

$$(1.7b) \qquad H(t, x', p) \leqq H(t, x, p) + \langle u, x'-x\rangle \text{ for all } x' \in R^n .$$

If $H(t, x, p)$ is differentiable with respect to x and p, this set reduces to the gradient of $H(t, \cdot,\cdot)$ at (x, p), and condition (1.5) reduces to the classical Hamiltonian system:

$$(1.8) \qquad \dot{x}(t) = H_p(t, x(t), p(t)) \quad \text{and} \quad \dot{p}(t) = -H_x(t, x(t), p(t)) .$$

We showed in particular in [6] that, if $x(t)$ and $p(t)$ satisfy (1.5), the arc x minimizes the integral (1.3) over the class of all arcs having the same endpoints as x, and the arc p has the same property with respect to a certain Lagrangian function M dual to L.

The purpose of this paper is to prove some theorems about solutions to the generalized Hamiltonian equation (1.5), where H is any extended-real-valued function on $[0, T] \times R^n \times R^n$ such that $H(t, x, p)$ is concave in x and convex in p. The first task is to deduce an existence theorem for (1.5) from a general result of C. Castaing [1] for differential equations with multivalued righthand sides. This involves an analysis of the regularity of the multifunction

$$(t, x, p) \longrightarrow \partial H(t, x, p) .$$

We then study properties of solution arcs $(x(t), p(t))$, showing in particular that, as in the classical case, $H(x(t), p(t))$ is constant along such an arc, if H is independent of t and finite in a neighborhood of the arc.

2. Existence theorem. As stated in the introduction, we shall assume throughout this paper that H is an extended-real-valued function on $[0, T] \times R^n \times R^n$ such that $H(t, x, p)$ is concave in x and convex in p. (A detailed account of the theory of extended-real-valued concave-convex functions may be found in [7].) It is not required that H correspond by formula (1.6) to a function L satisfying the conditions in [6].

We shall also assume that $H(t, x, p)$ is a (Lebesgue) measurable function of t for every $(x, p) \in R^n \times R^n$. This entails no loss of generality for applications to convex problems of Bolza or Lagrange, because of the following fact.

LEMMA 1. *If H is the Hamiltonian corresponding to a Lagrangian function L satisfying conditions* (A) *and* (B) *of* [6], *then $H(t, x(t), p(t))$ is a measurable function of $t \in [0, T]$ for any measurable functions $x: [0, T] \to R^n$ and $p: [0, T] \to R^n$.*

Proof. Let x be a measurable function from $[0, T]$ to R^n, and let

$$f(t, v) = L(t, x(t), v) \,. \tag{2.1}$$

Let D denote the set of all $t \in [0, T]$ such that, for at least one $v \in R^n$, $f(t, v) < +\infty$. Conditions (A) and (B) or [6] imply by [5, Corollary 4.5] that D is measurable, and that f is a (Lebesgue) normal convex integrand on $D \times R^n$ (see [4], [5] or [6] for the definition of "normal convex integrand"). Therefore the conjugate integrand

$$f^*(t, p) = \sup_{v \in R^n} \{\langle v, p\rangle - f(t, v)\} \tag{2.2}$$

is normal on $D \times R^n$ [4, Lemma 5], and this implies that $f^*(t, p(t))$ is a measurable function of $t \in D$ for any measurable function $p: D \to R^n$ [5, corollary to Lemma 5]. Since H is the Hamiltonian corresponding to L, we have

$$H(t, x(t), p) = f^*(t, p) \tag{2.3}$$

from (1.6), where $H(t, x(t), p) = -\infty$ for all $p \in R^n$ if $t \notin D$. In view of the measurability of D, it follows that $H(t, x(t), p(t))$ is a measurable function of $t \in [0, T]$ for any measurable function $p: [0, T] \to R^n$, and the proof of Lemma 1 is complete.

By a *solution* to the generalized Hamiltonian equation for H on an interval $I \subset [0, T]$, we shall mean, of course, a pair of absolutely continuous functions $x: I \to R^n$ and $p: I \to R^n$ such that (1.5) is satisfied for $t \in I$.

THEOREM 1. *Let U be an open subset of $R^n \times R^n$ such that, for* *every* $(x, p) \in U$, $H(t, x, p)$ *is finite and summable as a function of* $t \in [0, T]$. *Let* W *be any compact subset of* U. *Then there exists a positive* t_0 *such that, for every* $(a, c) \in W$, *the generalized Hamiltonian equation for* H *has at least one solution on* $[0, t_0]$ *satisfying*

$$x(0) = a \quad \text{and} \quad p(0) = c\ . \tag{2.4}$$

Furthermore, if $S(a, c)$ *denotes the set of all such solutions, regarded as a subset of the Banach space of continuous functions from* $[0, t_0]$ *to* $R^n \times R^n$, *then* $S(a, c)$ *is compact, and the multifunction* $S\colon (a, c) \rightarrow S(a, c)$ *is upper semicontinuous on* W.

Theorem 1 will be deduced by convexity arguments from the following result. Here we denote the euclidean norm of a vector w by $|w|$.

THEOREM (Castaing [1]). *Let* Ω *be an open subset of* R^m, *and let* F *be a multifunction from* $[0, T] \times \Omega$ *to* R^m *such that*

(a) $F(t, z)$ *is a nonempty compact convex set for every* $t \in [0, T]$ *and* $z \in \Omega$;

(b) *for every* $t \in [0, T]$, *the multifunction* $F(t, \cdot)$ *is upper semicontinuous from* Ω *to* R^m;

(c) *for every* $z \in \Omega$, *the multifunction* $F(\cdot, z)$ *is (Lebesgue) measurable from* $[0, T]$ *to* R^m;

(d) *there exists a summable, real-valued function* α *on* $[0, T]$ *such that* $|w| \leqq \alpha(t)$, *whenever* $w \in F(t, z)$, $t \in [0, T]$ *and* $z \in \Omega$.

Let W *be a compact subset of* Ω. *Then there is a positive* t_0 *such that, for each* $d \in \Omega$, *there exists at least one absolutely continuous function* $z\colon [0, t_0] \rightarrow R^m$ *satisfying* $z(0) = d$ *and* $\dot{z}(t) \in F(t, z(t))$ *for almost every* t. *Furthermore, if* $S(d)$ *denotes the set of all such functions* z, *regarded as a subset of the Banach space* $C_{R^m}[0, t_0]$, *then* $S(d)$ *is compact, and the multifunction* $S\colon d \rightarrow S(d)$ *is upper semicontinuous on* W.

The upper semicontinuity in condition (b) means, of course, that for every $t \in [0, T]$, every $z \in \Omega$ and every open set V containing $F(t, z)$, there exists a neighborhood N of z such that $F(t, z') \subset V$ for all $z' \in N$. The measurability in condition (c) means that, for every $z \in \Omega$ and every closed subset C of R^m, the set

$$\{t \in [0, T] \mid F(t, z) \cap C \neq \varnothing\}$$

is measurable.

To prove Theorem 1, we apply Castaing's Theorem to a suitable neighborhood Ω of W in U, with $z = (x, p)$ and

$$F(t, z) = \{(v, -u) \mid (u, v) \in \partial H(t, x, p)\} .$$

We need only show that conditions (a), (b), (c) and (d) are satisfied in this case under the assumptions in Theorem 1, and the following lemmas serve this purpose. (One may take Ω to be any open set containing W such that the closure K of Ω is a compact subset of U.)

LEMMA 2. *Let U be an open subset of $R^n \times R^n$ such that, for every $t \in [0, T]$ and $(x, p) \in U$, $H(t, x, p)$ is finite. Then $\partial H(t, x, p)$ is a nonempty compact convex set for every $t \in [0, T]$ and $(x, p) \in U$. Furthermore, for every $t \in [0, T]$ the multifunction $\partial H(t, \cdot, \cdot)$ is upper semicontinuous from U to $R^n \times R^n$, and for every $(x, p) \in U$ the multifunction $\partial H(\cdot, x, p)$ is measurable from $[0, T]$ to $R^n \times R^n$.*

Proof. Since $H(t, \cdot, \cdot)$ is a concave-convex function which is finite on U, $\partial H(t, x, p)$ is a nonempty compact convex set [7, p. 374], and the multifunction $\partial H(t, \cdot, \cdot)$ is upper semicontinuous on U [7, Corollary 35.7.1]. To verify the measurability of the multifunction $\partial H(\cdot, x, p)$ for a fixed $(x, p) \in U$, it suffices according to [5, Corollary 3.2] (see also [2]) to demonstrate that $h(t, y, q)$ is a measurable function of $t \in [0, T]$ for each $(y, q) \in R^n \times R^n$, where

$$h(t, y, q) = \max \{\langle u, y\rangle + \langle v, q\rangle \mid (u, v) \in \partial H(t, x, p)\} . \tag{2.5}$$

We have show in [7, pp. 373–374] that this support function h is also given by the formula

$$h(t, y, q) = H'(t, x, p; 0, q) - H'(t, x, p; -y, 0) , \tag{2.6}$$

where

$$H'(t, x, p; 0, q) = \lim_{\lambda \downarrow 0} \frac{H(t, x, p + \lambda q) - H(t, x, p)}{\lambda} ,$$

$$H'(t, x, p; -y, 0) = \lim_{\lambda \downarrow 0} \frac{H(t, x - \lambda y, p) - H(t, x, p)}{\lambda} .$$

Therefore

$$h(t, y, q) = \lim_{\lambda \downarrow 0} \frac{H(t, x, p + \lambda q) - H(t, x - \lambda y, p)}{\lambda} . \tag{2.7}$$

(Note that the difference quotients make sense here, at least for λ sufficiently small, by our assumptions on U.) Since H is measurable in t, it follows from (2.7) that $h(\cdot, y, q)$ is the pointwise limit of a sequence of measurable functions on $[0, T]$ and hence is itself a measurable function on $[0, T]$.

LEMMA 3. *Let U be an open subset of $R^n \times R^n$ such that, for* *every* $(x, p) \in U$, $H(t, x, p)$ *is finite and summable as a function of* $t \in [0, T]$. *Let* K *be a compact subset of* U. *Then there exists a summable, real-valued function* α *on* $[0, T]$ *such that*

$$|H(t, x', p') - H(t, x, p)| \leqq \alpha(t)|(x', p') - (x, p)| \quad \text{for every } t \in [0, T],\ (x, p) \in K \text{ and } (x', p') \in K\,. \tag{2.8}$$

Furthermore, for such a function α *one has* $|(u, v)| \leqq \alpha(t)$, *whenever* $(u, v) \in \partial H(t, x, p)$, $t \in [0, T]$ *and* $(x, p) \in K$.

Proof. We shall show that, for every $(a, c) \in K$, there exist an open neighborhood V of (a, c) and a summable, real-valued function α_V on $[0, T]$ such that

$$|H(t, x', p') - H(t, x, p)| \leqq \alpha_V(t)|(x', p') - (x, p)| \quad \text{for every } t \in [0, T],\ (x, p) \in V \text{ and } (x', p') \in V\,. \tag{2.8$'$}$$

This will imply, by the following argument, that (2.8) holds for some α. Since K is compact, there exist points

$$(a_k, c_k) \in K, \qquad k = 1, \cdots, q\,,$$

together with bounded open neighborhoods V_k and summable functions α_{V_k} satisfying (2.8′) such that

$$K \subset V_1 \cup \cdots \cup V_q\,.$$

We have

$$|H(t, x, p)| \leqq |H(t, a_k, c_k)| + \alpha_{V_k}(t)|(x, p) - (a_k, c_k)|$$

by (2.8′), whenever $(x, p) \in V_k$, and this implies the existence of a summable function ρ on $[0, T]$ such that

$$|H(t, x, p)| \leqq \rho(t) \text{ for every } (x, p) \in K\,. \tag{2.9}$$

It is possible to choose a positive δ such that, whenever $(x, p) \in K$ and $(x', p') \in K$ satisfy

$$|(x', p') - (x, p)| \leqq \delta\,, \tag{2.10}$$

there is a V_k containing both (x, p) and (x', p'). Then

$$|H(t, x', p') - H(t, x, p)| \leqq \max_{k=1,\cdots,q} \alpha_{V_k}(t)|(x', p') - (x, p)|\,,$$

whenever $(x, p) \in K$ and $(x', p') \in K$ satisfy (2.10). On the other hand, we have

$$|H(t, x', p') - H(t, x, p)| \leqq 2\rho(t)\delta^{-1}|(x', p') - (x, p)|$$

by (2.9), whenever $(x, p) \in K$ and $(x', p') \in K$ do not satisfy (2.10). Therefore (2.8) holds for

$$\alpha(t) = \max\{2\rho(t)\delta^{-1}, \alpha_{V_1}(t), \cdots, \alpha_{V_q}(t)\} ;$$

this α is summable because the functions ρ and α_{V_k} are summable.

Let $(a, c) \in K$. Choose any subset of U of the form

$$\{(x_i, p_j) \mid i = 1, \cdots, r \quad \text{and} \quad j = 1, \cdots, s\}$$

whose convex hull, which we shall denote by P, is a (closed) neighborhood of (a, c) lying in U. Choose a positive ε sufficiently small that the set

$$V' = \{(x, p) \mid |x - a| < 2\varepsilon \quad \text{and} \quad |p - c| < 2\varepsilon\}$$

is contained in P, and let

$$V = \{(x, p) \mid |x - a| < \varepsilon \quad \text{and} \quad |p - c| < \varepsilon\} .$$

We shall construct a summable, real-valued function α_V on $[0, T]$ such that (2.8′) holds.

As the first step, we construct a summable, real-valued function β on $[0, T]$ such that

$$|H(t, x, p)| \leqq \beta(t) \text{ for all } (t, x, p) \in [0, T] \times V' . \tag{2.11}$$

Let (x, p) be any point of P such that $|p - c| < 2\varepsilon$. Then p belongs to the convex hull of $\{p_j \mid j = 1, \cdots, s\}$, and hence by the convexity of $H(t, x, \cdot)$ we have

$$H(t, x, p) \leqq \beta_1(t, x) , \tag{2.12}$$

where

$$\beta_1(t, x) = \max_{j=1,\cdots,s} H(t, x, p_j) . \tag{2.13}$$

Assuming that $p \neq c$, we have $c = (1 - \lambda)\bar{p} + \lambda p$ for

$$\bar{p} = c - (\varepsilon/|p - c|)(p - c)$$
$$\lambda = \varepsilon/(\varepsilon + |p - c|) .$$

Note that $|\bar{p} - c| = \varepsilon$, and hence $(x, \bar{p}) \in V'$ and

$$H(t, x, \bar{p}) \leqq \beta_1(t, x)$$

by the argument just given. Since $0 < \lambda < 1$, we have

$$H(t, x, c) \leqq (1 - \lambda)H(t, x, \bar{p}) + \lambda H(t, x, p)$$
$$\leqq |\beta_1(t, x)| + \lambda H(t, x, p)$$

and consequently

$$H(t, x, p) \geqq \lambda^{-1}(H(t, x, c) - |\beta_1(t, x)|) \tag{2.14}$$
$$= \varepsilon^{-1}(\varepsilon + |p - c|)(H(t, x, c) - |\beta_1(t, x)|) .$$

Of course, (2.14) also holds trivially if $p = c$. The last expression in (2.14) is concave as a function of p, inasmuch as

$$H(t, x, c) - |\beta_1(t, x)| \leqq 0$$

by (2.12). Therefore

$$H(t, x, p) \geqq \beta_2(t, x) , \tag{2.15}$$

where

$$\beta_2(t, x) = \min_{j=1,\cdots,s} \varepsilon^{-1}(\varepsilon + |p_j - c|)(H(t, x, c) - |\beta_1(t, x)|) . \tag{2.16}$$

Setting

$$\bar{\beta}(t, x) = |\beta_1(t, x)| + |\beta_2(t, x)| , \tag{2.17}$$

we have

$$|H(t, x, p)| \leqq \bar{\beta}(t, x) . \tag{2.18}$$

This holds for any $t \in [0, T]$ and $(x, p) \in P$ such that $|p - c| < 2\varepsilon$. Observe that $\bar{\beta}(t, x)$ is finite and summable in t for fixed x by formulas (2.13), (2.16), (2.17) and our assumptions on H.

We now reason similarly in the x argument. Let (x, p) be any point V'. Then x belongs to the convex hull of $\{x_i \mid i = 1, \cdots, r\}$, so that

$$H(t, x, p) \geqq \min_{i=1,\cdots,r} H(t, x_i, p) .$$

We have $|H(t, x_i, p)| \leqq \bar{\beta}(t, x_i)$ by (2.18), and hence

$$H(t, x, p) \geqq \bar{\beta}_1(t) , \tag{2.19}$$

where

$$\bar{\beta}_1(t) = -\max_{i=1,\cdots,r} \bar{\beta}(t, x_i) .$$

Assuming that $x \neq a$, we have $a = (1 - \mu)\bar{x} + \mu x$ for

$$\bar{x} = a - (\varepsilon/|x - a|)(x - a) ,$$
$$\mu = \varepsilon/(\varepsilon + |x - a|) .$$

Here $|\bar{x} - a| = \varepsilon$, so that $(\bar{x}, p) \in V'$; therefore

$$H(t, \bar{x}, p) \geqq \bar{\beta}_1(t)$$

holds, by the argument used to establish (2.19). Thus

$$H(t, a, p) \geqq (1-\mu)H(t, \bar{x}, p) + \mu H(t, x, p)$$
$$\geqq -|\bar{\beta}_1(t)| + \mu H(t, x, p) ,$$

and we have

$$H(t, x, p) \leqq \mu^{-1}(H(t, a, p) + |\bar{\beta}_1(t)|)$$
$$= \varepsilon^{-1}(\varepsilon + |x - a|)(H(t, a, p) + |\bar{\beta}_1(t)|) . \tag{2.20}$$

The latter also holds trivially if $x = a$. Since the final expression in (2.20) is convex as a function of x, and x belongs to the convex hull of $\{x_i \mid i = 1, \cdots, r\}$, it follows that

$$H(t, x, p) \leqq \max_{i=1,\cdots,r} \varepsilon^{-1}(\varepsilon + |x_i - a|)(H(t, a, p) + |\bar{\beta}_1(t)|) .$$

Thus

$$H(t, x, p) \leqq \bar{\beta}_2(t) \tag{2.21}$$

by inequality (2.18), where

$$\bar{\beta}_2(t) = \max_{i=1,\cdots,r} \varepsilon^{-1}(\varepsilon + |x_i - a|)(\bar{\beta}(t, a) + |\bar{\beta}_1(t)|) .$$

The function β on $[0, T]$ defined by

$$\beta(t) = |\bar{\beta}_1(t)| + |\bar{\beta}_2(t)|$$

is finite and summable, according to the construction of $\bar{\beta}_1$ and $\bar{\beta}_2$, and it satisfies (2.11) as desired.

We now demonstrate that (2.8′) holds for

$$\alpha_V(t) = 4\varepsilon^{-1}\beta(t) .$$

Let (x, p) and (x', p') be points of V. Since

$$|H(t, x', p') - H(t, x, p)| \leqq |H(t, x', p') - H(t, x, p')|$$
$$+ |H(t, x, p') - H(t, x, p)| ,$$

it is enough to show that

$$|H(t, x', p') - H(t, x, p')| \leqq \alpha_V(t)|x' - x|/2 , \tag{2.22}$$

$$|H(t, x, p') - H(t, x, p)| \leqq \alpha_V(t)|p' - p|/2 . \tag{2.23}$$

The argument is the same for both inequalities, except for the substitution of concavity for convexity, and for this reason we treat only (2.23). Assuming that $p \neq p'$, we have $p' = (1-\theta)p + \theta p''$ for

$$p'' = p' + (\varepsilon/|p' - p|)(p' - p) ,$$
$$\theta = |p' - p|/(\varepsilon + |p' - p|) .$$

The points (x, p) and (x, p'') belong to V', so that

$$|H(t, x, p)| \leqq \beta(t) \quad \text{and} \quad |H(t, x, p'')| \leqq \beta(t)$$

by (2.11). Since $H(t, x, \cdot)$ is a convex function and $0 < \theta < 1$, we have

$$H(t, x, p') \leqq (1 - \theta)H(t, x, p) + \theta H(t, x, p'') .$$

Therefore

$$\begin{aligned} H(t, x, p') - H(t, x, p) &\leqq \theta[H(t, x, p'') - H(t, x, p)] \leqq 2\theta\beta(t) \\ &= 2\beta(t)|p' - p|/(\varepsilon + |p' - p|) \leqq 2\varepsilon^{-1}\beta(t)|p' - p| . \end{aligned}$$

By the same argument with the roles of p' and p reversed, we also have

$$H(t, x, p) - H(t, x, p') \leqq 2\varepsilon^{-1}\beta(t)|p - p'| ,$$

and hence (2.23) is valid.

Only the final assertion of Lemma 3 remains to be proved. Let $t \in [0, T]$, $(x, p) \in K$ and $(u, v) \in \partial H(t, x, p)$. We have

$$|(u, v)|^2 = \langle u, u\rangle + \langle v, v\rangle \leqq h(t, u, v) , \tag{2.24}$$

where h is defined by (2.5). As observed in the proof of Lemma 1, h is also given by (2.7), so that from inequality (2.8) we have

$$\begin{aligned} |h(t, u, v)| &= \lim_{\lambda \downarrow 0} \frac{|H(t, x, p + \lambda v) - H(t, x - \lambda u, p)|}{\lambda} \\ &\leqq \lim_{\lambda \downarrow 0} \lambda^{-1}\alpha(t)|(x, p + \lambda v) - (x - \lambda u, p)| \\ &= \alpha(t)|(u, v)| . \end{aligned} \tag{2.25}$$

Combining (2.24) and (2.25), we see that $|(u, v)| \leqq \alpha(t)$.

3. Properties of solutions.

THEOREM 2. *Let U be an open subset of $R^n \times R^n$ such that, for every $(x, p) \in U$, $H(t, x, p)$ is finite and bounded as a function of $t \in [0, T]$. If $x(t)$ and $p(t)$ satisfy the generalized Hamiltonian equation for H over a closed interval $I \subset [0, T]$ and*

$$(x(t), p(t)) \in U \text{ for every } t \in I , \tag{3.1}$$

then $\dot{x}(t)$ and $\dot{p}(t)$ are essentially bounded as functions of $t \in I$.

Proof. Let

$$K = \{(x(t), p(t)) \mid t \in I\} . \tag{3.2}$$

Then K is a compact subset of U. The theorem is obtained by applying to K the last assertion of the following lemma, which is a stronger version of Lemma 3.

LEMMA 4. *Let U be an open subset of $R^n \times R^n$ such that, for every $(x, p) \in U$, $H(t, x, p)$ is finite and bounded as a function of $t \in [0, T]$. Let K be a compact subset of U. Then there exists a constant α such that*

$$(3.3) \qquad \begin{gathered} |H(t, x', p') - H(t, x, p)| \leqq \alpha |(x', p') - (x, p)| \\ \text{for every } t \in [0, T],\ (x, p) \in K \text{ and } (x', p') \in K\,. \end{gathered}$$

Furthermore, for such a constant α one has $|(u, v)| \leqq \alpha$, whenever $(u, v) \in \partial H(t, x, p)$, $t \in [0, T]$ and $(x, p) \in K$.

Proof. The compactness argument in the first paragraph of the proof of Lemma 3 shows that (3.3) holds for some α, if for every $(a, c) \in K$ and every sufficiently small neighborhood V of (a, c), there exists a constant α_V such that

$$(3.3') \qquad \begin{gathered} |H(t, x', p') - H(t, x, p)| \leqq \alpha_V |(x', p') - (x, p)| \\ \text{for every } t \in [0, T],\ (x, p) \in V \text{ and } (x', p') \in V\,. \end{gathered}$$

Thus we need only establish (3.3) for sets K of sufficiently small diameter, and we may assume without loss of generality that there exist open convex subsets C and D in R^n such that

$$K \subset C \times D \subset U\,.$$

The fact that (3.3) holds in the latter case for some α has already been proved in [7, Th. 35.2]. The last sentence of the lemma is established by the argument in the last paragraph of the proof of Lemma 3.

THEOREM 3. *Let U be an open subset of $R^n \times R^n$ such that, on $[0, T] \times U$, H is finite, and $\partial H/\partial t$ exists and is continuous. If $x(t)$ and $p(t)$ satisfy the generalized Hamiltonian equation for H over an interval $I \subset [0, T]$, and*

$$(x(t), p(t)) \in U \text{ for every } t \in I\,,$$

then $H(t, x(t), p(t))$ is a continuously differentiable function of $t \in I$, and one has

$$(3.4) \qquad \frac{d}{dt} H(t, x(t), p(t)) = \frac{\partial H}{\partial t}(t, x(t), p(t))\,, \qquad t \in I\,.$$

Proof. It is enough to show that $H(t, x(t), p(t))$ is an absolutely continuous function of $t \in I$, and that (3.4) holds almost everywhere. Since the properties to be established are local in t, there is no loss of generality if we suppose I to be closed (and nontrivial) and sufficiently small that $(x(t), p(t')) \in U$ for every $t \in I$ and $t' \in I$. The set K defined by (3.2) is then compact, so that (3.3) holds for some constant α by Lemma 4. The continuous differentiability of H in t on $[0, T] \times U$ implies the existence of a constant α' such that

$$|H(t', x, p) - H(t, x, p)| \leqq \alpha' |t' - t| ,$$

wherever $(x, p) \in K$, $t \in [0, T]$ and $t' \in [0, T]$. On the other hand, Theorem 2 implies the existence of a constant α'' such that

$$|(x(t'), p(t')) - (x(t), p(t))| \leqq \alpha'' |t' - t|$$

for every $t \in I$ and $t' \in I$. We have

$$\begin{aligned} &|H(t', x(t'), p(t')) - H(t, x(t), p(t))| \\ &\quad \leqq |H(t', x(t'), p(t')) - H(t', x(t), p(t))| + |H(t', x(t), p(t)) \\ &\qquad - H(t, x(t), p(t))| \\ &\quad \leqq \alpha |(x(t'), p(t')) - (x(t), p(t))| + \alpha' |t' - t| \\ &\quad \leqq (\alpha\alpha'' + \alpha')|t' - t| \end{aligned}$$

for every $t \in I$ and $t' \in I$, and this implies in particular that $H(\cdot, x(\cdot), p(\cdot))$ is absolutely continuous on I.

Since $x(t)$ and $p(t)$ are absolutely continuous functions of $t \in I$, there exists a subset D of I with the following properties: the complement of D in I has measure zero, the derivative functions $\dot{x}$ and $\dot{p}$ are defined throughout D, and for every $t \in D$ there is a decreasing sequence $\{t_i\}$ in D such that

$$(3.5) \qquad t_i \longrightarrow t,\ \dot{x}(t_i) \longrightarrow \dot{x}(t) \quad \text{and} \quad \dot{p}(t_i) \longrightarrow \dot{p}(t) .$$

(This may be deduced easily from Lusin's Theorem [3, p. 243] and the fact that $\dot{x}(t)$ and $\dot{p}(t)$ are almost-everywhere-defined, measurable functions of t.) Deleting a set of measure zero from D if necessary, we can arrange also that the derivative $(d/dt)H(t, x(t), p(t))$ exists for every $t \in D$, and that

$$(3.6) \qquad (-\dot{p}(t), \dot{x}(t)) \in \partial H(t, x(t), p(t)) \text{ for every } t \in D .$$

We shall verify that (3.4) holds for every $t \in D$, and this will prove Theorem 3.

Let $t \in D$, and let $\{t_i\}$ be a decreasing sequence in D such that (3.5) holds. We have

$$\frac{d}{dt}H(t, x(t), p(t)) = \lim_{i\to\infty} \frac{H(t_i, x(t_i), p(t_i)) - H(t, x(t), p(t))}{t_i - t} \tag{3.7}$$

by our assumptions. We shall show that $(\partial H/\partial t)(t, x(t), p(t))$ is a lower bound to this limit, and that it is also an upper bound, thereby proving equality. The numerator in the difference quotient in (3.7) can be expressed as

$$\begin{aligned} &H(t_i, x(t_i), p(t_i)) - H(t_i, x(t), p(t_i)) \\ &\quad + H(t_i, x(t), p(t_i)) - H(t, x(t), p(t_i)) \\ &\quad + H(t, x(t), p(t_i)) - H(t, x(t), p(t)) \,. \end{aligned} \tag{3.8}$$

Since

$$(-\dot{p}(t_i), \dot{x}(t_i)) \in \partial H(t_i, x(t_i), p(t_i))$$

by (3.6), we see from inequality (1.7b) in the definition of ∂H that

$$H(t_i, x(t), p(t_i)) \leqq H(t_i, x(t_i), p(t_i)) + \langle -\dot{p}(t_i), x(t) - x(t_i)\rangle ,$$

or in other words

$$H(t_i, x(t_i), p(t_i)) - H(t_i, x(t), p(t_i)) \geqq -\langle \dot{p}(t_i), x(t_i) - x(t)\rangle \,. \tag{3.9}$$

Similarly, we see from (3.6) and inequality (1.7a) in the definition of ∂H that

$$H(t, x(t), p(t_i)) - H(t, x(t), p(t)) \geqq \langle \dot{x}(t), p(t_i) - p(t)\rangle \,. \tag{3.10}$$

The difference quotient in (3.7) is therefore not less than

$$\begin{aligned} &\frac{H(t_i, x(t), p(t_i)) - H(t, x(t), p(t_i))}{t_i - t} \\ &\quad - \left\langle \dot{p}(t_i), \frac{x(t_i) - x(t)}{t_i - t}\right\rangle + \left\langle \dot{x}(t), \frac{p(t_i) - p(t)}{t_i - t}\right\rangle \,. \end{aligned} \tag{3.11}$$

The limit of expression (3.11) as $i \to \infty$ is

$$\frac{\partial H}{\partial t}(t, x(t), p(t)) - \langle \dot{p}(t), \dot{x}(t)\rangle + \langle \dot{p}(t), \dot{x}(t)\rangle ,$$

and we may therefore conclude that

$$\frac{d}{dt}H(t, x(t), p(t)) \geqq \frac{\partial H}{\partial t}(t, x(t), p(t)) \,.$$

The proof of the opposite inequality is parallel; in place of (3.8), one considers the expression

$$\begin{aligned} &H(t_i, x(t_i), p(t_i)) - H(t_i, x(t_i), p(t)) \\ &\quad + H(t_i, x(t_i), p(t)) - H(t, x(t_i), p(t)) \\ &\quad + H(t, x(t_i), p(t)) - H(t, x(t), p(t)) \end{aligned} \tag{3.12}$$

for the numerator of the difference quotient in (3.7).

COROLLARY. *Let H be independent of t, and let U be an open subset of $R^n \times R^n$ on which H is finite. If $x(t)$ and $p(t)$ satisfy the generalized Hamiltonian equation for H over an interval I, and $(x(t), p(t))$ belongs to U for every $t \in I$, then*

$$H(x(t), p(t)) = \text{const.}, \qquad t \in I .$$

Note that Theorem 2 is also applicable under the assumptions in this corollary.

Following the terminology of [7], we shall say that the concave-convex function $H(t, \cdot, \cdot)$ is *proper* for a given t, if there exists some $(a, c) \in R^n \times R^n$ such that

$$H(t, a, p) > -\infty \text{ for every } p \in R^n , \tag{3.13a}$$

$$H(t, x, c) < +\infty \text{ for every } x \in R^n . \tag{3.13b}$$

This condition is satisfied for every $t \in [0, T]$, if H is the Hamiltonian function corresponding to a Lagrangian function L satisfying condition (A) of [6] (see [7, Th. 34.2]). If $H(t, \cdot, \cdot)$ is proper and $\partial H(t, x, p)$ is nonempty, one has $-\infty < H(t, x, p) < +\infty$, as may easily be deduced from relations (3.13) and the definition of $\partial H(t, x, p)$.

THEOREM 4. *Let $(x_i(t), p_i(t))$ satisfy the generalized Hamiltonian equation for H over the interval $I \subset [0, T]$, $i = 1, 2$. If $H(t, \cdot, \cdot)$ is proper for every $t \in I$, the function*

$$f(t) = \langle x_1(t) - x_2(t), p_1(t) - p_2(t) \rangle \tag{3.14}$$

is nondecreasing on I. Moreover, if f is constant on a subinterval I' of I, then for any $\lambda \in [0, 1]$ and $\mu \in [0, 1]$ the arcs

$$\begin{aligned} x(t) &= (1 - \lambda)x_1(t) + \lambda x_2(t) , \\ p(t) &= (1 - \mu)p_1(t) + \mu p_2(t) , \end{aligned} \tag{3.15}$$

satisfy the generalized Hamiltonian equation over I', and one has

$$\begin{aligned} H(t, x(t), p(t)) = {} & (1 - \lambda)(1 - \mu)H(t, x_1(t), p_1(t)) \\ & + \lambda(1 - \mu)H(t, x_2(t), p_1(t)) \\ & + (1 - \lambda)\mu H(t, x_1(t), p_2(t)) + \lambda\mu H(t, x_2(t), p_2(t)) . \end{aligned} \tag{3.16}$$

for every $t \in I'$.

Proof. The function f is absolutely continuous. For almost every $t \in I$, $\dot{x}_i(t)$ and $\dot{p}_i(t)$ are defined, and one has

$$(3.17) \qquad (-\dot{p}_i(t), \dot{x}_i(t)) \in \partial H(t, x_i(t), p_i(t)) , \quad i = 1, 2 .$$

Furthermore, relation (3.17) implies by definition that

$$(3.18a) \qquad H(t, x_2(t), p_1(t)) \leqq H(t, x_1(t), p_1(t)) + \langle -\dot{p}_1(t), x_2(t) - x_1(t)\rangle ,$$

$$(3.18b) \qquad -H(t, x_1(t), p_2(t)) \leqq -H(t, x_1(t), p_1(t)) - \langle \dot{x}_1(t), p_2(t) - p_1(t)\rangle ,$$

$$(3.18c) \qquad H(t, x_1(t), p_2(t)) \leqq H(t, x_2(t), p_2(t)) + \langle -\dot{p}_2(t), x_1(t) - x_2(t)\rangle ,$$

$$(3.18d) \qquad -H(t, x_2(t), p_1(t)) \leqq -H(t, x_2(t), p_2(t)) - \langle \dot{x}_2(t), p_1(t) - p_2(t)\rangle .$$

Here $H(t, x_i(t), p_i(t))$ is finite for $i = 1, 2$ by (3.17) and the remark preceding the theorem, and hence $H(t, x_2(t), p_1(t))$ and $H(t, x_1(t), p_2(t))$ are also finite, in view of the inequalities (3.18). Adding the inequalities (3.18), we see that for almost every t

$$(3.19) \qquad \begin{aligned} 0 \leqq \langle \dot{x}_1(t) - \dot{x}_2(t), p_1(t) - p_2(t)\rangle \\ + \langle \dot{p}_1(t) - \dot{p}_2(t), x_1(t) - x_2(t)\rangle = f'(t) . \end{aligned}$$

Thus the derivative f' is nonnegative almost everywhere on I, and f is consequently nondecreasing.

If f is constant on a subinterval I', then f' vanishes on I' and, as seen from the derivation of (3.19), equality must hold for $t \in I'$ in (3.18 a–d). Suppose that this is true, and that $x(t)$ and $p(t)$ are given by (3.15). Since $H(t, \cdot, p_1(t))$ is a concave function, equality in (3.18a) and (3.18c) implies

$$(3.20a) \qquad \begin{aligned} &H(t, x(t), p_i(t)) \\ &= (1 - \lambda)H(t, x_1(t), p_i(t)) + \lambda H(t, x_2(t), p_i(t)),\ i = 1, 2 . \end{aligned}$$

Similarly, equality in (3.18b) and (3.18d) implies

$$(3.20b) \qquad \begin{aligned} &H(t, x_i(t), p(t)) \\ &= (1 - \mu)H(t, x_i(t), p_1(t)) + \mu H(t, x_i(t), p_2(t)),\ i = 1, 2 . \end{aligned}$$

From (3.20a) and the concavity-convexity of $H(t, \cdot, \cdot)$, we have

$$\begin{aligned} H(t, x(t), p(t)) &\leqq (1 - \mu)H(t, x(t), p_1(t)) + \mu H(t, x(t), p_2(t)) \\ &= (1 - \lambda)(1 - \mu)H(t, x_1(t), p_1(t)) + \lambda(1 - \mu)H(t, x_1(t), p_2(t)) \\ &\quad + (1 - \lambda)\mu H(t, x_2(t), p_1(t)) + \lambda\mu H(t, x_2(t), p_2(t)) . \end{aligned}$$

The opposite inequality is obtained similarly from (3.20b), and this proves (3.16).

We show now that $x(t)$ and $p(t)$ satisfy the generalized Hamiltonian equation for H over I', in other words that the inequalities

$$(3.21a) \qquad H(t, x(t), p') \geqq H(t, x(t), p(t)) + \langle \dot{x}(t), p' - p(t)\rangle ,$$

$$(3.21b) \qquad H(t, x', p(t)) \leqq H(t, x(t), p(t)) + \langle -\dot{p}(t), x' - x(t)\rangle ,$$

hold for almost every $t \in I'$, if x' and p' are arbitrary vectors in R^n. We have

$$H(t, x(t), p') \geqq (1 - \lambda)H(t, x_1(t), p') + \lambda H(t, x_2(t), p') . \tag{3.22}$$

On the other hand, by (3.16) we have (for almost every t)

$$\begin{aligned} H(t, x_1(t), p') &\geqq H(t, x_1(t), p_1(t)) + \langle \dot{x}_1(t), p' - p_1(t) \rangle , \\ H(t, x_2(t), p') &\geqq H(t, x_2(t), p_2(t)) + \langle \dot{x}_2(t), p' - p_2(t) \rangle . \end{aligned}$$

Furthermore, equality in (3.17d) implies

$$\begin{aligned} H(t, x_2(t), p_2(t)) + \langle \dot{x}_2(t), p' - p_2(t) \rangle = H(t, x_2(t), p_1(t)) \\ + \langle \dot{x}_2, p' - p_1(t) \rangle . \end{aligned}$$

Therefore, from (3.22) and (3.20a)

$$\begin{aligned} H(t, x(t), p') \geqq (1 - \lambda)&[H(t, x_1(t), p_1(t)) + \langle \dot{x}_1(t), p' - p_1(t) \rangle] \\ &+ \lambda[H(t, x_2(t), p_1(t)) + \langle \dot{x}_2(t), p' - p_1(t) \rangle] \\ &= H(t, x(t), p_1(t)) + \langle \dot{x}(t), p' - p_1(t) \rangle . \end{aligned} \tag{3.23}$$

The same argument, with the roles of p_1 and p_2 reversed, also yields the inequality

$$H(t, x(t), p') \geqq H(t, x(t), p_2(t)) + \langle \dot{x}(t), p' - p_2(t) \rangle . \tag{3.24}$$

Thus

$$\begin{aligned} H(t, x(t), p') \geqq (1 - \mu)&[H(t, x(t), p_1(t)) + \langle \dot{x}(t), p' - p_1(t) \rangle] \\ &+ \mu[H(t, x(t), p_2(t)) + \langle \dot{x}(t), p' - p_2(t) \rangle] \\ &\geqq H(t, x(t), p(t)) + \langle \dot{x}(t), p' - p(t) \rangle , \end{aligned}$$

and (3.21a) is established. The proof of (3.21b) is by a parallel argument.

COROLLARY. *Let H be independent of t and proper, and let $x(t)$ and $p(t)$ satisfy the generalized Hamiltonian equation for H over an interval I. Then the function*

$$g(t) = \langle \dot{x}(t), \dot{p}(t) \rangle$$

is (essentially) nondecreasing on I.

Proof. Let I' be any closed, bounded subinterval of the interior of I. If h is a sufficiently small nonzero real number, the arcs

$$\begin{aligned} (x_1(t), p_1(t)) &= (x(t + h), p(t + h)) , \\ (x_2(t), p_2(t)) &= (x(t), p(t)) \end{aligned}$$

are defined over I' and in fact satisfy the generalized Hamiltonian equation over I'. The function

$$g_h(t) = \langle x(t+h) - x(t),\, p(t+h) - p(t)\rangle / h^2$$

is then nondecreasing on I' by Theorem 4. For almost every $t \in I'$, the limit of $g_h(t)$ as h tends to 0 exists and equals $g(t)$; therefore $g(t)$ is essentially nondecreasing on I'.

REFERENCES

1. C. Castaing, *Sur les equations différentielles multivoques*, C. R. Acad. Sci. Paris **263** (1966), 63-66.

2.———, *Sur les multi-applications mésurables*, Thèse, Caën, 1967. This has been published in part in Révue Française d'Informatique et de Recherche Opérationelle **1** (1967), 3-34.

3. P. Halmos, *Measure theory*, Van Nostrand, Princeton, 1950.

4. R. T. Rockafellar, *Integrals which are convex functionals*, Pacific J. Math. **24** (1968), 525-539.

5. ———, *Measurable dependence of convex sets and functions on parameters*, J. Math. Anal. Appl. (1969), **28** (1969), 4-25.

6. ———, *Conjugate convex functions in optimal control and the calculus of variations*, J. Math. Anal. Appl. (1970).

7. ———, *Convex analysis*, Princeton Univ. Press, 1969.

Received November 14, 1969. This work was supported in part by the Air Force Office of Scientific Research under grant no. AF-AFOSR-1202-67A.

UNIVERSITY OF WASHINGTON
SEATTLE, WASHINGTON