4[illegible]

SIAM J. CONTROL
Vol. 10, No. 4, November 1972

STATE CONSTRAINTS IN CONVEX CONTROL PROBLEMS OF BOLZA*

R. TYRRELL ROCKAFELLAR†

Abstract. Methods of convex analysis are applied to certain problems of Lagrange and Bolza in optimal control. Conditions characterizing optimal arcs are obtained without the usual differentiability assumptions on the data in the problem. Special existence theorems are proved. A dual control problem is formulated in terms of adjoint arcs which are not necessarily absolutely continuous, but of bounded variation, so as to allow for jumps caused by the presence of state-contraints in the primal problem.

1. Introduction. Convex problems of Bolza are problems of a basic type in the calculus of variations and optimal control which satisfy convexity conditions not only in the control or derivative variables, but also in the state variables. Such problems have many special properties not implied by standard theory, particularly as regards duality. Moreover, these properties can be deduced by convexity methods which do not require the customary assumptions of continuity and differentiability. A number of results in this vein have been obtained by the author in [10], [11], [12]. However, these results do not explicitly treat problems with bounded state variables, and indeed they often exclude state constraints, other than constraints on endpoints. The purpose of this paper is to show how the results can nevertheless be applied to many problems with bounded state variables by various devices. The effect of state constraints on duality, on the existence of solutions, and on necessary and sufficient conditions for optimality is analyzed in detail.

We take as our model a "convex" control problem of the form:

$$\text{(1.1)} \qquad \text{Minimize} \quad \int_0^1 f(t, x(t), u(t))\,dt + l(x(0), x(1))$$

subject to

$$\text{(1.2)} \qquad \dot{x}(t) = A(t)x(t) + u(t) \quad \text{for almost every } t,$$

$$\text{(1.3)} \qquad x(t) \in X(t) \quad \text{for every } t,$$

where $x:[0,1] \to R^n$ is absolutely continuous and $u:[0,1] \to R^n$ is measurable. The sets $X(t)$ in R^n are nonempty, closed and convex, while the functions $f(t,\cdot,\cdot)$ and l on $R^n \times R^n$ are lower semicontinuous, convex and extended-real-valued—they may take on $+\infty$, although not $-\infty$, as a possible value, but they are assumed not to be identically $+\infty$. Of course, $A(t)$ denotes an $n \times n$ real matrix. The dependence of $X(t)$, $f(t,\cdot,\cdot)$ and $A(t)$ on t is discussed below. No differentiability is assumed.

It should be emphasized that, despite appearances, our problem of Bolza includes as special cases many other types of problems, such as problems of Lagrange

* Received by the editors October 13, 1971.

† Department of Mathematics, University of Washington, Seattle, Washington 98105. This research was supported in part by the Air Force Office of Scientific Research under Grant AF-AFOSR-71-1994.

with fixed endpoints. In fact, other constraints, besides the abstract constraint (1.3), are implicitly incorporated into the problem through the use of $+\infty$. Thus there is the endpoint constraint

$$(x(0), x(1)) \in E, \tag{1.4}$$

where E is the convex subset of $R^n \times R^n$ which is the effective domain of l:

$$E = \{(c_0, c_1) | l(c_0, c_1) < +\infty\}. \tag{1.5}$$

Similarly, there is the implicit control constraint

$$u(t) \in U(t, x(t)) \quad \text{for almost every } t, \tag{1.6}$$

where

$$U(t, x) = \{u | f(t, x, u) < +\infty\} \subset R^n. \tag{1.7}$$

The reader unfamiliar with this method of representing constraints is referred to [10] for further discussion and examples.

The implicit control set $U(t, x)$ is convex, but possibly of less than full dimension in R^n; thus the absence of a matrix $B(t)$ in (1.2) does not mean a loss of generality. (Note that $U(t, x)$ is not necessarily bounded or even closed.) The functions $f(t, x, \cdot)$ or l might in particular vanish identically on the sets $U(t, x)$ or E, respectively. The latter sets and the sets $X(t)$ might be described by inequality constraints. However, such specific situations need not concern us in the development of the basic theory. They can be handled at a later stage by a routine application of standard theorems in convex analysis. This is discussed in [10] in the case of $U(t, x)$ and E, and the considerations are similar for $X(t)$. Thus we can concentrate on the main features and difficulties of the problem, relegating many distracting and notationally burdensome details to "computation" in particular examples.

If $U(t, x)$ were empty for certain values of t and x, such values would have to be avoided, and this could amount to an implicit state constraint in addition to (1.3). We eliminate this possibility through our assumptions in the next section, our aim here being to keep the state constraints explicit and separate from other aspects of the problem, so that their exact role can be seen. However, there are problems which cannot be treated adequately in this way, because there is no sense to $U(t, x)$ being nonempty for $x \notin X(t)$. For example, in mathematical economics one encounters the case where $X(t)$ is the nonnegative orthant of R^n, and there is no way to define $f(t, x, u)$ finitely outside of this orthant without violating our convexity assumptions. Such problems require a different approach, involving a generalization of the basic duality theory to problems where the trajectories are not absolutely continuous, but only of bounded variation.

The plan of the paper is as follows. In § 2 and § 3 we derive lower semicontinuity properties of the cost functional and a basic existence theorem. These results may be compared most closely perhaps with those of Cesari [1] and Olech [9]; however, there is not a large overlap, and certainly the methods used here are very different. A dual control problem, involving the costate variables, is introduced in § 4 and studied in relation to certain optimality conditions in § 5. The optimality conditions, in a slightly generalized form, are shown in § 6 to be necessary and sufficient under fairly mild hypotheses. The costate functions in these generalized conditions are characterized in § 7 as solutions to a generalized dual problem.

For comparison with other literature on necessary and sufficient conditions in problems with state constraints, papers of Neustadt [6], [7], [8] and Funk and Gilbert [3] may be cited. (Further references may be found in these papers.) For the most part, these authors focus on results applicable also to "nonconvex" problems, but which involve stronger regularity assumptions than ours (differentiability, etc.), even when specialized to the "convex" case. Neustadt's paper [8] does bring convexity to the fore, in methodology as well as hypothesis, and is thus the closest in spirit to the present work. However, also in [8] differentiability assumptions intervene, and the formulation of the optimality conditions is dependent on them. The conditions that we formulate do not even require the differentiability of the functions defining the state constraints, although, of course, this property could be exploited in analyzing the conditions in special cases.

No one has previously shown that the costate functions in the optimality conditions for problems with state constraints solve a general dual problem. The case of problems without state constraints was covered in our earlier papers [10], [12]. Some related results were also obtained under stronger regularity assumptions by Tsvetanov [16]. For an economic example not entirely covered by the results in this paper, for reasons mentioned above, see Makarov [17].

2. The Bolza functional. Let $\mathscr{A}$ denote the Banach space consisting of all absolutely continuous functions $x:[0, 1] \to R^n$ under the norm

$$\|x\|_{\mathscr{A}} = |x(0)| + \int_0^1 |\dot{x}(t)|\, dt, \tag{2.1}$$

and let $\mathscr{C}$ denote the larger Banach space consisting of all continuous functions $x:[0, 1] \to R^n$ under the usual norm

$$\|x\|_{\mathscr{C}} = \max_{0 \leqq t \leqq 1} |x(t)|. \tag{2.2}$$

(Here $|\cdot|$ denotes the Euclidean norm in R^n.) We have $\|x\|_{\mathscr{C}} \leqq \|x\|_{\mathscr{A}}$ for all $x \in \mathscr{A}$.

Let

$$S = \{x \in \mathscr{A} | x(t) \in X(t) \quad \text{for every } t\}. \tag{2.3}$$

Clearly S is a closed convex subject of $\mathscr{A}$. Our problem can be represented as:

$$\text{Minimize} \quad F(x) \quad \text{subject to} \quad x \in S, \tag{2.4}$$

where

$$F(x) = \int_0^1 f(t, x(t), \dot{x}(t) - A(t)x(t))\, dt + l(x(0), x(1)). \tag{2.5}$$

We call F a *Bolza functional* on the space $\mathscr{A}$. Of course, conditions must be imposed so that the integral in (2.5) makes sense.

INTEGRABILITY ASSUMPTION. *The components of the matrix $A(t)$ are summable as functions of $t \in [0, 1]$. Furthermore, $h(t, x, p)$ is (finite and) summable as a function* *of* $t \in [0, 1]$ *for each fixed* $x \in R^n$ *and* $p \in R^n$, *where*

$$h(t, x, p) = \sup \{u \cdot p - f(t, x, u) | u \in U(t, x)\}. \tag{2.6}$$

The function h is not only convex in p, but also concave in x, by virtue of the joint convexity of $f(t, x, u)$ in x and u [13, Thm. 33.1].

The constraint $u \in U(t, x)$ could be omitted from (2.6) without loss of generality, in view of (1.7). The condition that $h(t, x, p) > -\infty$ for all (t, x, p) amounts to the condition, mentioned earlier, that $U(t, x) \neq \emptyset$ for all (t, x). The rest of the assumption on $h(t, x, p)$ is a growth condition of Nagumo–Tonelli type on $f(t, x, u)$ as a function of $u \in U(t, x)$. This is essential in obtaining the existence of solutions to the control problem, as is seen in the next section. It is also convenient technically in a great many other respects.

THEOREM 1. *Under the integrability assumption, F is a well-defined, lower semicontinuous, convex functional from the Banach space $\mathscr{A}$ to $R^1 \cup \{+\infty\}$. Moreover, the (convex) level sets*

$$\{x \in \mathscr{A} | F(x) \leqq \alpha\}, \quad \alpha \text{ real}, \tag{2.7}$$

are locally compact relative to the weak topology on $\mathscr{A}$. These sets are also closed and locally compact as subsets of the Banach space $\mathscr{C}$, with respect to both the weak and the strong topologies on $\mathscr{C}$.

Proof. This follows mainly from results in [12]. In the notation of [12], we have

$$F(x) = \Phi_{l,L}(x), \tag{2.8}$$

where

$$L(t, x, v) = f(t, x, v - A(t)x). \tag{2.9}$$

Condition (A) of [12] is satisfied by l and L, in view of our convexity and lower semicontinuity assumptions on l and $f(t, \cdot, \cdot)$. The Hamiltonian function corresponding to L is

$$\begin{aligned} H(t, x, p) &= \sup \{v \cdot p - L(t, x, v) | v \in R^n\} \\ &= h(t, x, p) + p \cdot A(t)x. \end{aligned} \tag{2.10}$$

By duality, we also have [10, p. 211]

$$L(t, x, v) = \sup \{v \cdot p - H(t, x, p) | p \in R^n\}. \tag{2.11}$$

The integrability assumptions on h and A imply that $H(t, x, p)$ is finite and summable in $t \in [0, 1]$ for each fixed $x \in R^n$ and $p \in R^n$. Hence L also satisfies conditions (B), (C_0) and (D_0) of [12] by the corollary and remark after Proposition 4 of [12]. These conditions guarantee in particular that F is well-defined, convex and lower semicontinuous [10, Thm. 1]. The local weak compactness of the level sets of F in $\mathscr{A}$ is a consequence of (C_0), as noted in [12, discussion following Thm. 1]. The last assertion of the theorem is obtained from the following fact.

LEMMA 1. *If a convex set K in $\mathscr{A}$ is closed and weakly locally compact, then it is also closed and locally compact as a subset of $\mathscr{C}$, both weakly and strongly.*

Proof. We observe first that if a set K_0 is weakly compact in $\mathscr{A}$, then it is strongly compact as a subset of $\mathscr{C}$. In fact, the set of function $\dot{x}$, as x ranges over K_0, is weakly compact in the L^1-space of R^n-valued functions on [0, 1]. Hence, by the Dunford–Pettis criterion for weak compactness in L^1-spaces, there exists for every $\varepsilon > 0$ some $\delta > 0$ such that

$$\int_T |\dot{x}(t)|\,dt < \varepsilon \quad \text{whenever} \quad x \in K_0 \quad \text{and} \quad \text{mes}\, T < \delta. \tag{2.12}$$

In particular, (2.12) implies that

$$|x(t_1) - x(t_2)| \leqq \int_{t_1}^{t_2} |\dot{x}(t)|\,dt < \varepsilon \quad \text{for all } x \in K_0 \tag{2.13}$$

if $0 \leqq t_1 < t_2 \leqq 1, t_2 - t_1 < \delta$. The functions in K_0 are thus equicontinuous on [0, 1], as well as, of course, uniformly bounded pointwise. The strong compactness of K_0 in $\mathscr{C}$ then follows from the theorem of Ascoli–Arzela.

Now let K be a convex subset of $\mathscr{A}$ which is closed and locally compact with respect to $w_{\mathscr{A}}$, the weak topology on $\mathscr{A}$. Let $n_{\mathscr{C}}$ and $w_{\mathscr{C}}$ denote the norm topology and weak topology on $\mathscr{C}$, respectively. Let x be an arbitrary point of K, and let U_1 be a $w_{\mathscr{A}}$-closed, convex $w_{\mathscr{A}}$-neighborhood of x such that $K \cap U_1$ is $w_{\mathscr{A}}$-compact. Certainly $K \cap U_1$ is then also $n_{\mathscr{C}}$-compact and $w_{\mathscr{C}}$-compact, as just pointed out. Thus, to prove that K is locally compact in $n_{\mathscr{C}}$ and $w_{\mathscr{C}}$ it suffices to demonstrate the existence of a $w_{\mathscr{C}}$-closed $w_{\mathscr{C}}$-neighborhood U_2 of x in $\mathscr{C}$ such that

$$K \cap U_2 \subset K \cap U_1. \tag{2.14}$$

For notational simplicity, we can suppose that $x = 0$. Let

$$U_1' = \tfrac{1}{2}U_1 \subset w_{\mathscr{A}}\text{-int } U_1, \tag{2.15}$$

$$W = K \cap [U_1 \setminus w_{\mathscr{A}}\text{-int } U_1']. \tag{2.16}$$

Then W is a $w_{\mathscr{A}}$-closed subset of $K \cap U_1$, hence $w_{\mathscr{A}}$-compact and consequently $w_{\mathscr{C}}$-compact. Also, $0 \notin W$. Therefore it is possible to select a $w_{\mathscr{C}}$-closed, convex $w_{\mathscr{C}}$-neighborhood U_2 of 0 such that $W \cap U_2 = \varnothing$. Then (2.14) must hold, for if U_2 contained a point y in K but not in U_1, the line segment joining 0 and y would contain points of $U_1 \setminus \text{int } U_1'$. Such points would lie in $W \cap U_2$ by convexity, contradicting $W \cap U_2 = \varnothing$.

It remains to show that K is also $n_{\mathscr{C}}$-closed, and therefore, by convexity, $w_{\mathscr{C}}$-closed. Let α be any positive real number, and let

$$K_\alpha = \{x \in K \mid \|x\|_{\mathscr{C}} \leqq \alpha\}. \tag{2.17}$$

Then K_α is a $w_{\mathscr{A}}$-closed subset of K, hence $w_{\mathscr{A}}$-locally compact. Moreover, K_α is convex and contains no half-lines. Therefore K_α is actually $w_{\mathscr{A}}$-compact [2]. It follows that K_α is also $n_{\mathscr{C}}$-compact and in particular $n_{\mathscr{C}}$-closed. Since this is true for arbitrary α, K itself is $n_{\mathscr{C}}$-closed.

Remark. The converse of Lemma 1 fails, at least without convexity, since a sequence in $\mathscr{A}$ converging in the $\mathscr{C}$-norm need not even be bounded as a subset of $\mathscr{A}$, much less weakly compact. Thus the properties of F asserted by Theorem 1 relative to the weak topology on $\mathscr{A}$ are considerably stronger than the properties asserted relative to the weak or strong topologies on $\mathscr{C}$.

3. Existence of optimal arcs. By a *feasible arc* for our control problem, we mean an $x \in S$ such that $F(x) < +\infty$. An *optimal arc* is a feasible arc for which the infimum of F over S is attained. Theorem 1 implies that the convex sets

$$\{x \in S | F(x) \leqq \alpha\}, \quad \alpha \text{ real}, \tag{3.1}$$

are not only closed, but locally compact in the weak topologies of $\mathscr{A}$ and $\mathscr{C}$ and the strong topology of $\mathscr{C}$. Thus any slight additional condition which ensures that these sets are actually compact (and not all empty) is enough to give us a theorem on the existence of optimal arcs.

It is well known that *a locally compact convex set is compact if and only if it does not contain any half-lines* [2]. We have already made use of this fact in proving Lemma 1. Thus the sets (3.1) are compact (in all the topologies mentioned) if S does not contain any half-line along which F is (finitely) bounded above (and hence, by convexity, "nonincreasing"). Therefore, the latter condition guarantees the existence of an optimal arc, provided there is at least one feasible arc. This criterion for existence is geometrically appealing, but not specific enough for most applications.

We proceed to formulate the half-line condition equivalently as an assumption on the sets $X(t)$ and functions l and $f(t, \cdot, \cdot)$ appearing in the control problem. For each t we denote by $\hat{X}(t)$ the *recession cone* (asymptotic cone) of $X(t)$:

$$\hat{X}(t) = \{z \in R^n | X(t) + z \subset X(t)\}. \tag{3.2}$$

We denote by $\hat{l}$ the *recession function* of l. Thus

$$\hat{l}(c_0, c_1) = \lim_{\lambda \to +\infty} l(\bar{c}_0 + \lambda c_0, \bar{c}_1 + \lambda c_1)/\lambda, \tag{3.3}$$

where $(\bar{c}_0, \bar{c}_1)$ is any element of the set E in (1.5). (The limit is independent of the particular choice of $(\bar{c}_0, \bar{c}_1)$ [13, p. 66].) Similarly, we let $\hat{f}(t, \cdot, \cdot)$ denote the recession function of $f(t, \cdot, \cdot)$.

BOUNDEDNESS ASSUMPTION. *There does not exist a nonzero arc $z \in \mathscr{A}$ such that*

$$z(t) \in \hat{X}(t) \quad \textit{for every } t, \tag{3.4}$$

$$\int_0^1 \hat{f}(t, z(t), \dot{z}(t) - A(t)z(t))\, dt + \hat{l}(z(0), z(1)) \leqq 0. \tag{3.5}$$

The integral in (3.5) is well-defined under our previous integrability assumption, and in fact it is a lower semicontinuous convex function of $z \in \mathscr{A}$ [12, Prop. 6].

THEOREM 2. *Suppose there is at least one feasible arc, and the integrability assumption is satisfied. Then the boundedness assumption is necessary for any nonempty level set of the form* (3.1) *to be weakly compact in $\mathscr{A}$ or strongly compact in $\mathscr{C}$, and it is sufficient for them all to be both weakly compact in $\mathscr{A}$ and strongly compact in $\mathscr{C}$. In particular, the boundedness assumption ensures the existence of an optimal arc. Indeed, every minimizing sequence of feasible arcs has a subsequence which converges to an optimal arc, not only in the uniform norm $\|\cdot\|_{\mathscr{C}}$, but also in the weak topology of $\mathscr{A}$.*

Proof. Let $\hat{S}$ denote the set of all arcs $z \in \mathscr{A}$ satisfying (3.4). Let $\hat{F}(z)$ denote the left side of (3.5). Then $\hat{S}$ is the recession cone of S and, as shown in [12, Prop. 6], $\hat{F}$

is the recession function of F, since for the function L in (2.9) we have

$$\hat{L}(t, z, \dot{z}) = \hat{f}(t, z, \dot{z} - A(t)z). \tag{3.6}$$

Therefore, a half-line

$$\{x + \lambda z | 0 \leqq \lambda < +\infty\}, \quad z \neq 0,$$

is contained in a set of the form (3.1) if and only if x belongs to this set and $z \in \hat{S}$, $\hat{F}(z) \leqq 0$. The conclusion of the theorem is immediate from this and Theorem 1.

COROLLARY 1. *Under the integrability assumption, an optimal arc x exists if a feasible arc exists and the sets $X(t)$ are all bounded (since then the boundedness assumption is satisfied).*

Proof. In this case (3.4) holds only for the zero arc, because $\hat{X}(t) = \{0\}$ for every t.

It is interesting to note that the boundedness of every $X(t)$ in Corollary 1 does not necessarily entail the boundedness of S, even in the norm $\|\cdot\|$, since no assumption has been made on the behavior of $X(t)$ with respect to t.

COROLLARY 2. *Under the integrability assumption, every (convex) set of the form*

$$\{x \in S | F(x) \leqq \alpha, \|x\|_{\mathscr{C}} \leqq \beta\}, \quad \alpha \text{ and } \beta \text{ real}, \tag{3.7}$$

is weakly compact in $\mathscr{A}$ and strongly compact in $\mathscr{C}$.

Proof. Apply Corollary 1 to

$$X_\beta(t) = \{x \in X(t) | \, |x| \leqq \beta\}. \tag{3.8}$$

COROLLARY 3. *Suppose that*

$$f(t, x, u) = f_0(t, x) + f_1(t, u), \tag{3.9}$$

$$l(x(0), x(1)) = l_0(x(0)) + l_1(x(1)), \tag{3.10}$$

and denote the recession functions of $f_i(t, \cdot, \cdot)$ and l_i by $\hat{f}_i(t, \cdot, \cdot)$ and $\hat{l}_i$, respectively. Under the integrability assumption, an optimal arc x exists if a feasible arc exists and every solution $z \in \mathscr{A}$ to the differential equation

$$\dot{z}(t) = A(t)z(t) \quad a.e. \tag{3.11}$$

satisfying (3.4) *and*

$$\int_0^1 \hat{f}_0(t, z(t))\, dt + \hat{l}_0(z(0)) + \hat{l}_1(z(1)) \leqq 0 \tag{3.12}$$

has $z(t) = 0$ for at least one $t \in [0, 1]$. (The boundedness assumption is satisfied in this case.)

Proof. If f and l have the structure in (3.9) and (3.10), their recession functions also have this structure:

$$\hat{f}(t, z, y) = \hat{f}_0(t, z) + \hat{f}_1(t, y), \tag{3.13}$$

$$\hat{l}(z(0), z(1)) = \hat{l}_0(z(0)) + \hat{l}_1(z(1)). \tag{3.14}$$

Furthermore, the integrability assumption implies that

$$\sup_{u\in R^n}\{u\cdot p - f_1(t,u)\} < +\infty \quad \text{for all } p\in R^n, \tag{3.15}$$

and consequently [13, p. 116] that

$$\hat{f}_1(t,y) = \delta_0(y) = \begin{cases} 0 & \text{if } y = 0, \\ +\infty & \text{if } y \neq 0. \end{cases} \tag{3.16}$$

Thus in this case condition (3.5) is equivalent to (3.11) and (3.12). The result now follows from Theorem 2 and the fact that a solution to (3.11) which vanishes for some $t \in [0, 1]$ must be the zero arc.

Remark. A simple but common case where the condition in Corollary 3 is satisfied occurs when one of the (convex) endpoint sets

$$E_i = \{c \in R^n | l_i(c) < +\infty\}, \quad i = 0, 1, \tag{3.17}$$

is bounded (so that $\hat{l}_i$ is finite only at 0), or one of the sets $X(t)$ is bounded (so that $\hat{X}(t) = \{0\}$). A more general result resembling Corollary 3 can be derived from [12, Cor. 3 to Thm. 3].

4. The dual control problem. The question of necessary and sufficient conditions for the optimality of an arc x is closely tied in with duality, a topic of interest in its own right. In this section we describe the basic duality briefly, to set the stage for later developments.

The dual control problem is:

$$\text{Minimize} \quad \int_0^1 g(t, p(t), w(t))\,dt + m(p(0), p(1)) \tag{4.1}$$

subject to

$$\dot{p}(t) = -A^*(t)p(t) + w(t) \quad \text{for almost every } t, \tag{4.2}$$

where $A^*(t)$ is the transpose of $A(t)$ and

$$g(t, p, w) = \sup\{u\cdot p + x\cdot w - f(t, x, u) | u \in U(t, x), x \in X(t)\}, \tag{4.3}$$

$$m(d_0, d_1) = \sup\{c_0\cdot d_0 - c_1\cdot d_1 - l(c_0, c_1) | c_0 \in R^n, c_1 \in R^n\}. \tag{4.4}$$

In terms of the Bolza functional

$$G(p) = \int_0^1 g(t, p(t), \dot{p}(t) + A^*(t)p(t))\,dt + m(p(0), p(1)), \tag{4.5}$$

we can express this problem as:

$$\text{Minimize} \quad G(p) \quad \text{over all} \quad p \in \mathscr{A}. \tag{4.6}$$

Before discussing the circumstances under which the integral in the Bolza functional G is well-defined, we remark that the dual problem, like the original problem, involves implicit constraints on the controls and endpoints. The implicit dual control set is

$$W(t, p) = \{w \in R^n | g(t, p, w) < +\infty\}, \tag{4.7}$$

while the endpoint pair $(p(0), p(1))$ must belong to the set where the function m is finite-valued. However, there are no *state* constraints in the dual problem, even implicit ones. In particular,

$$W(t, p) \neq \varnothing \quad \text{for every } t \in [0, 1] \text{ and } p \in R^n. \tag{4.8}$$

This is evident from the following result, which may also be helpful in calculating $g(t, p, w)$ in specific cases.

LEMMA 2. *In terms of the function h in* (2.6) *and the functions*

$$s(t, w) = \sup \{x \cdot w | x \in X(t)\}, \tag{4.9}$$

$$k(t, p, w) = \sup \{x \cdot w + h(t, x, p) | x \in R^n\}, \tag{4.10}$$

one has, under the integrability assumption,

$$\begin{aligned} g(t, p, w) &= \sup \{x \cdot w + h(t, x, p) | x \in X(t)\} \\ &= \min \{k(t, p, w - z) + s(t, z) | z \in R^n\}. \end{aligned} \tag{4.11}$$

Furthermore, for every measurable, essentially bounded function $p:[0, 1] \to R^n$ it is possible to find a summable function $w:[0, 1] \to R^n$ and a summable function $\alpha:[0, 1] \to R^1$ such that

$$g(t, p(t), w(t)) \leqq \alpha(t) \quad \textit{for all } t. \tag{4.12}$$

Proof. The first equality in (4.11) is immediate from the definitions of g and h. The second equality results from a basic theorem of convex analysis expressing the conjugate of the sum of two convex functions in terms of infimal convolution of the conjugate functions [13, Thm. 16.4]: according to the first equality in (4.11), $g(t, p, \cdot)$ is the conjugate of the sum of $-h(t, \cdot, p)$ and the indicator of the set $X(t)$, while, by definition, $k(t, p, \cdot)$ is the conjugate of $-h(t, \cdot, p)$, and $s(t, \cdot)$ is the conjugate of the indicator of $X(t)$. The theorem in question is applicable, because $-h(t, \cdot, p)$ is a finite function under the integrability assumption. (The convexity of $-h(t, x, p)$ in x has already been noted in § 2.) In the notation and terminology of [10], [12], the function

$$M(t, p, r) = k(t, p, r + A^*(t)p) \tag{4.13}$$

is the Lagrangian dual to the function L in (2.9). We have shown in the proof of Theorem 1 that L satisfies conditions (A), (B), (C_0) and (D_0) of [12] under the integrability assumption, and this implies that M satisfies the same conditions [10, Thm. 2], [12, § 1]. Then for every measurable, essentially bounded function $p:[0, 1] \to R^n$ it is possible to find a summable function $r:[0, 1] \to R^n$ and a summable function $\alpha:[0, 1] \to R^1$ such that

$$M(t, p(t), r(t)) \leqq \alpha(t) \quad \text{for every } t. \tag{4.14}$$

Setting $w(t) = r(t) + A^*(t)p(t)$, we obtain a summable function w for which (4.12) holds, since $g \leqq k$. The lemma is thereby proved.

We now introduce a further condition from which it will be deduced, in particular, that the integral in (4.5) is well-defined.

INTERIORITY ASSUMPTION. *The multifunction $X:t \to X(t)$ satisfies*

$$\operatorname{int} X(t) \neq \varnothing \quad \textit{for every } t \tag{4.15}$$

and

$$\{(t, x) | x \in \text{int } X(t)\} = \text{int cl } \{(t, x) | x \in X(t)\}. \tag{4.16}$$

For an equivalent form of this assumption, see [14, Lemma 2].

LEMMA 3. *Under the interiority assumption, the constraint set S has a nonempty interior in $\mathscr{A}$ consisting of the functions x such that*

$$x(t) \in \text{int } X(t) \quad \textit{for every } t, \tag{4.17}$$

and this is the same as the interior of S relative to the norm $\|\cdot\|_{\mathscr{C}}$. Furthermore, any $x \in \mathscr{A}$ satisfying

$$x(t) \in X(t) \quad \textit{for almost every } t \tag{4.18}$$

actually satisfies (1.3). *Thus any $x \in \mathscr{A}$ satisfying* (4.18) *belongs to S.*

Proof. If $\mathscr{A}$ were replaced by $\mathscr{C}$ here (and in the definition of S), this would be the special case of [14, Thm. 5 and Lemma 2], where the $D(t)$ and $f(t, \cdot)$ in the notation of the latter results are taken to be $X(t)$ and the indicator of $X(t)$, respectively. The lemma is also valid in the present form, because $\mathscr{A}$ is dense in $\mathscr{C}$.

THEOREM 3. *Under the assumptions of integrability and interiority, G in* (4.5) *is a well-defined, lower semicontinuous, convex functional from $\mathscr{A}$ to $R^1 \cup \{+\infty\}$. If the boundedness assumption also holds, then*

$$-\min_{x \in S} F(x) = \inf_{p \in \mathscr{A}} G(p) < +\infty. \tag{4.19}$$

Proof. We deduce this from the main theorem of [12]. Let L be as in (2.9), and let

$$L_0(t, x, v) = \begin{cases} L(t, x, v) & \text{if } x \in X(t); \\ +\infty & \text{if } x \notin X(t). \end{cases} \tag{4.20}$$

The measurability properties of L_0, to be described in a moment, ensure that the Bolza functional

$$F_0(x) = \int_0^1 L_0(t, x(t), \dot{x}(t))\, dt + l(x(0), x(1)) \tag{4.21}$$

is well-defined. The given control problem can be regarded as that of minimizing F_0 over all of $\mathscr{A}$, since from the second assertion of Lemma 3 we have

$$F_0(x) = \begin{cases} F(x) & \text{if } x \in S, \\ +\infty & \text{if } x \notin S. \end{cases} \tag{4.22}$$

The corresponding dual problem, in the terminology of [10] and [12], is that of minimizing

$$\int_0^1 M_0(t, p(t), \dot{p}(t))\, dt + m(p(0), p(1)) \tag{4.23}$$

over all $p \in \mathscr{A}$, where

$$M_0(t, p, r) = \sup_{x, v} \{x \cdot r + v \cdot p - L_0(t, x, v)\}, \tag{4.24}$$

and m is given by (4.4). It is easily seen from the definitions that

$$M_0(t, p, r) = g(t, p, r + A^*(t)p), \tag{4.25}$$

so that this dual problem is indeed the problem presented here as the dual.

We shall verify that conditions (A), (B), (C_0) and (D) of [12] hold for L_0. The assertions about the nature of G will then be valid by [10, Thm. 2]. The duality relation (4.19) will follow from Theorem 1(a) of [12], since our boundedness assumption is equivalent by [12, Cor. 1 of Thm. 3] to the condition of dual attainability in the hypothesis of that theorem.

Since $X(t)$ is a closed, convex set, the function $L_0(t, \cdot, \cdot)$ is lower semicontinuous and convex on $R^n \times R^n$. Thus condition (A) of [12] is satisfied by L_0, provided that $L_0(t, \cdot, \cdot)$ is not identically $+\infty$ for any t. The latter will be seen in a moment, in connection with condition (D).

We have already verified in the proof of Theorem 1 that conditions (A), (B), (C_0) and (D_0) of [12] are satisfied by the function L, and in the case of (B) this means that L is measurable with respect to the σ-field in $[0, 1] \times R^{2n}$ generated by products of the Lebesgue sets in $[0, 1]$ and Borel sets in R^{2n}. The corresponding measurability property of L_0 then follows from the measurability of the set

$$\{(t, x, v) | x \in X(t)\} \subset [0, 1] \times R^{2n}, \tag{4.26}$$

which is implied by our interiority assumption. Indeed, since $X(t)$ is a closed, convex set with nonempty interior, we have $x \in X(t)$ if and only if $x + B_j$ meets $\operatorname{int} X(t)$ for every natural number j, where B_j is the open ball of radius $1/j$ and center 0 in R^n. Furthermore, the openness of the set on the left in (4.16) implies the openness of the set of all (t, x) such that $x + B_j$ meets $\operatorname{int} X(t)$. Thus the set (4.26) can be expressed as the intersection of a countable collection of open sets and in particular is Borel measurable.

The fact that L satisfies (C_0) trivially implies that L_0 satisfies (C_0), since $L_0 \geqq L$. As for condition (D), we must demonstrate the existence of a bounded, measurable function $x:[0, 1] \to R^n$, a summable function $v:[0, 1] \to R^n$, and a summable function $\alpha:[0, 1] \to R^1$ such that

$$L_0(t, x(t), v(t)) \leqq \alpha(t) \quad \text{for every } t, \tag{4.27}$$

or, in other words, such that (1.3) holds and

$$L(t, x(t), v(t)) \leqq \alpha(t) \quad \text{for every } t. \tag{4.28}$$

Since L is known to satisfy condition (D_0), corresponding functions v and α satisfying (4.28) exist for any bounded, measurable function x [12, Prop. 3]. Therefore, it is enough to show there is at least one bounded, measurable function x for which (1.3) holds. In fact, Lemma 3 asserts the existence of such a function x belonging to $\mathscr{A}$.

COROLLARY. *Under the assumptions of integrability, boundedness, and interiority, a feasible arc x exists in the original control problem if and only if, in the dual problem,*

$$\inf_{p \in \mathscr{A}} G(p) > -\infty. \tag{4.29}$$

Proof. A feasible arc exists if and only if the infimum of F over S in (4.19) is not $+\infty$.

5. Optimality conditions without state constraints. In the case where $X(t) = R^n$ for every t, it is possible to derive from the results in [10] and [12] a companion theorem to Theorem 3, furnishing conditions that are necessary and sufficient for the optimality of an arc x. We carry this out as a step in the derivation of more general optimality conditions in § 6 for the case where (nontrivial) state constraints are present. The conditions we formulate at this stage do in fact have some bearing on state constraints, and they help to clarify the relationship between the given control problem and its dual.

As usual, we denote by $\partial\varphi(z)$ the (closed, convex) set of all *subgradients* of a convex function φ on R^n at the point z, that is, the set of all vectors y such that

$$\varphi(z') \geqq \varphi(z) + y' \cdot (z' - z) \quad \text{for all } z' \in R^n. \tag{5.1}$$

Subgradients of a concave function are defined analogously, with the reversed inequality. The relationship between subgradients and ordinary gradients is discussed at length in [13, § 25]. For the function h in (2.6), we denote by $\partial_p h(t, x, p)$ the set of subgradients of the convex function $h(t, x, \cdot)$ at p and by $\partial_x h(t, x, p)$ the set of subgradients of the concave function $h(t, \cdot, p)$ at x. These sets are described further in Lemma 4 below.

We denote by $N(t, x)$ the *cone of normals* [13] to $X(t)$ at the point x. This is the closed, convex cone defined by

$$N(t, x) = \begin{cases} \{w \mid w \cdot (z - x) \leqq 0 \text{ for all } z \in X(t)\} & \text{if } x \in X(t), \\ \varnothing \quad \text{if } x \notin X(t). \end{cases} \tag{5.2}$$

The conditions to be analyzed may now be stated.

OPTIMALITY CONDITIONS. *The functions $x \in \mathscr{A}$ and $p \in \mathscr{A}$ satisfy*

$$u(t) \in \partial_p h(t, x(t), p(t)) \quad a.e., \tag{5.3}$$

$$w(t) \in -\partial_x h(t, x(t), p(t)) + N(t, x(t)) \quad a.e., \tag{5.4}$$

where

$$\dot{x}(t) = A(t)x(t) + u(t) \quad \text{and} \quad \dot{p}(t) = -A^*(t)p(t) + w(t) \quad a.e. \tag{5.5}$$

Furthermore,

$$(p(0), -p(1)) \in \partial l(x(0), x(1)) \quad (\textit{transversality}). \tag{5.6}$$

The state constraint (1.3) on x is embedded in (5.4) by virtue of (5.2) and Lemma 3, if the interiority condition holds.

The relationship between these optimality conditions and the familiar "maximum principle" is made clearer by the following result. (Note the *normality*: the multiplier of the cost function is taken to be -1.)

LEMMA 4. *Let the integrability assumption be satisfied. Then the set $\partial_p h(t, x, p)$ consists of the control vectors $u \in U(t, x)$ for which the supremum of $u \cdot p - f(t, x, u)$ is attained. On the other hand, in terms of the dual cost function g and dual control set W, the set $-\partial_x h(t, x, p) + N(t, x)$ consists of the dual control vectors $w \in W(t, p)$ for which the supremum of $x \cdot w - g(t, p, w)$ is attained.*

Moreover, if the function H_0 *is defined by*

$$(5.7)\qquad H_0(t,x,p) = \begin{cases} h(t,x,p) + p\cdot A(t)x & \text{if } x \in X(t), \\ -\infty & \text{if } x \notin X(t), \end{cases}$$

the optimality conditions, except for transversality, can be expressed in the Hamiltonian form

$$(5.8)\qquad \dot{x} \in \partial_p H_0(t,x,p) \quad \text{and} \quad -\dot{p} \in \partial_x H_0(t,x,p) \quad \text{a.e.}$$

Proof. The assertion concerning $\partial_p h(t,x,p)$ is immediate from the fact that $h(t,x,\cdot)$ is by definition the conjugate of the convex function $f(t,x,\cdot)$, whose effective domain is $U(t,x)$ (see [13, Thm. 23.5]). On the other hand, let

$$(5.9)\qquad h_0(t,x,p) = \begin{cases} h(t,x,p) & \text{if } x \in X(t), \\ -\infty & \text{if } x \notin X(t). \end{cases}$$

Then $g(t,p,\cdot)$ is by (4.11) the conjugate of the convex function $-h_0(t,\cdot,p)$, and the effective domain of $g(t,p,\cdot)$ is $W(t,p)$. If the integrability assumption is satisfied, so that $h(t,x,p)$ is finite and hence continuous as a concave function of x, $-h_0(t,\cdot,p)$ is lower semicontinuous ($X(t)$ being closed) and consequently is in turn the conjugate of $g(t,p,\cdot)$. Thus, by the same reasoning as for $\partial_p h(t,x,p)$, the set $-\partial h_0(t,x,p)$ consists of the vectors $w \in W(t,p)$ for which the supremum of $x\cdot w - g(t,p,w)$ is attained. We note now that, since $-h_0(t,\cdot,p)$ is the sum of the finite convex function $-h(t,\cdot,p)$ and the indicator of the nonempty convex set $X(t)$, we have

$$(5.10)\qquad -\partial h_0(t,x,p) = -\partial h(t,x,p) + N(t,x).$$

This is a special case of a basic formula for subdifferentiation (see [13, p. 215 and Thm. 23.8]). The same formula also yields the Hamiltonian form of the optimality conditions.

The generalized Hamiltonian equations (5.8) have been studied in [10] and [11]. They are very similar to ordinary differential equations, despite the "multivaluedness". As a matter of fact, if the integrability and interiority assumptions are satisfied, the subgradient sets in (5.8) reduce to single elements, the ordinary gradients of H_0, for almost all choices of (t,x,p) such that $x \in \operatorname{int} X(t)$. This result may be deduced from [13, Thm. 35.9].

The next lemma explains the fundamental connection between the optimality conditions and the dual control problem (cf. Thm. 3).

LEMMA 5. *Let the assumptions of integrability and interiority hold. Then, in order that* $x \in S$ *and* $p \in \mathscr{A}$ *be arcs such that*

$$(5.11)\qquad -\min_S F = -F(x) = G(p) = \min_{\mathscr{A}} G,$$

it is necessary and sufficient that the preceding optimality conditions be satisfied.

Proof. As we have observed at the beginning of the proof of Theorem 3, the given control problem corresponds under our assumptions to the problem of minimizing the Bolza functional F_0 in (4.21) over all of $\mathscr{A}$. Thus this is just the special case of [10, Thm. 5] corresponding to l and the function L in (4.20).

COROLLARY. *Let the assumptions of integrability and interiority hold, and let* $x \in \mathscr{A}$. *In order that* x *be an optimal arc in the given control problem, it is sufficient* *that there exist an arc $p \in \mathscr{A}$ such that the optimality conditions are satisfied. These conditions are also necessary, if the boundedness assumption holds and the infimum of G in the dual problem is attained.*

Proof. The sufficiency is obvious from the lemma, and the necessity is a consequence of Theorem 3.

The preceding corollary indicates the need of some criterion for the attainment of the infimum in the dual control problem. This is where state constraints in the original problem cause difficulties. A dual existence theorem ought to involve some Nagumo–Tonelli-type growth condition on the convex functions $g(t, p, \cdot)$: at the very least,

$$\lim_{\lambda \to +\infty} g(t, p, \bar{w} + \lambda w)/\lambda = +\infty \tag{5.12}$$

for every $\bar{w}$ and nonzero w. However, this growth property is not present if $X(t) \neq R^n$. Indeed, since $g(t, p, \cdot)$ is the conjugate of the convex function $-h_0(t, \cdot, p)$, where h_0 is given by (5.9), the property holds if and only if $h_0(t, x, p)$ is finite as a function of x [13, Cor. 13.3.1], and this is not true if $X(t) \neq R^n$.

These observations show that we cannot hope to prove the *necessity* of the above optimality conditions in the case of nontrivial state constraints. However, this is not surprising. It is well known that state constraints can cause "jumps" in the costate variables $p(t)$, yet the optimality conditions, as we have stated them, do not allow for such discontinuities. There is a dual interpretation of the situation: the reason we cannot prove necessity is that the dual problem is formulated too narrowly for solutions to exist. If the dual cost function g does not have the growth properties which guarantee that the Bolza functional G attains its infimum over $\mathscr{A}$, then perhaps G can be extended to a larger space of functions p for which we do have attainment. We shall reconcile these two interpretations in § 7. For the present, we state a result applicable to the case where $X(t) = R^n$.

ATTAINABILITY ASSUMPTION. *The relative interiors of the convex sets D and E in $R^n \times R^n$ have a nonempty intersection, where E is the set of all "admissible" end-point pairs defined in* (1.5), *and D is the set of all pairs* $(x(0), x(1))$ *arising from arcs* $x \in \mathscr{A}$ *such that* $\dot{x}(t) = A(t)x(t) + u(t)$ *with* $u(t) \in U(t, x(t))$ *for almost every t.*

The relative interior of a convex set is the interior relative to the affine hull of the set [13, § 6]. Note that one does at least have $D \cap E \neq \emptyset$ if a feasible arc x exists. Thus, if there are not state constraints, the attainability assumption says that "feasible arcs exist, and not just marginally."

THEOREM 4 [12]. *Suppose that $X(t) = R^n$ for every $t \in [0, 1]$, so that S is all of $\mathscr{A}$ and the normal cone $N(t, x)$ is just $\{0\}$ for every $(t, x) \in [0, 1] \times R^n$. Let the assumptions of integrability and attainability hold. Then*

$$-\inf_{\mathscr{A}} F = \min_{\mathscr{A}} G > -\infty. \tag{5.13}$$

In order that $x \in \mathscr{A}$ be an optimal arc in the original control problem, it is necessary and sufficient that there exist an arc $p \in \mathscr{A}$ for which the optimality conditions hold. The arcs p obtained in this way are precisely the optimal arcs in the dual problem.

Proof. This combines Lemma 4 with [12, Thm. 1(b)] for the function L in (2.9). (Condition (D_0) of [12] holds for L, as observed in the proof of Theorem 1 above. Furthermore, the set C_L in [12] has the same relative interior as the set D

here, because C_L has the same relative interior as

$$\left\{(x(0), x(1)) | x \in \mathscr{A}, \int_0^1 L(t, x(t), \dot{x}(t))\, dt < \infty\right\} \tag{5.14}$$

according to [12, Cor. 4 of Thm. 3], and D lies between these two sets.)

6. Generalized optimality conditions for state constraints. We proceed to reduce the general case to the case of no state constraints by means of an abstract multiplier principle.

Every continuous linear functional on the Banach space C corresponds, of course, to an R^n-valued measure on [0, 1], which can be expressed as dp for a certain function $p:[0, 1] \to R^n$ of *bounded variation.* We denote by $\mathscr{B}$ the space of all such functions under the norm

$$\|p\|_{\mathscr{B}} = |p(0)| + \|dp\|. \tag{6.1}$$

Thus $\mathscr{B}$ is a Banach space isomorphic to $R^n \times C^*$, and $\mathscr{A}$ is isometrically embedded in $\mathscr{B}$. It should be said that $\mathscr{B}$ actually consists of equivalence classes: we identify two functions p and q if $p(t^-) = q(t^-)$ and $p(t^+) = q(t^+)$ at all the (countably many) places where these functions have jumps, since then $dp = dq$. In this context we regard $p(0)$ as $p(0^-)$ and $p(1)$ as $p(1^+)$.

STRICT FEASIBILITY ASSUMPTION. *There is at least one arc $x \in \mathscr{A}$ satisfying $F(x) < +\infty$ such that $x(t) \in \operatorname{int} X(t)$ for every t.*

LEMMA 6. *Let the assumptions of integrability, interiority and strict feasibility hold. Then an arc $x \in \mathscr{A}$ is optimal in the original control problem (that is, F attains its minimum over S at x), if and only if there is a function $p_0 \in \mathscr{B}$ for which the linear functional*

$$\begin{aligned} \Lambda(z) &= -\int_0^1 p_0(t) \cdot \dot{z}(t)\, dt - p_0(0) \cdot z(0) + p_0(1) \cdot z(1) \\ &= \int_0^1 z(t)\, dp_0(t), \quad z \in \mathscr{A}, \end{aligned} \tag{6.2}$$

has the following properties:

(a) *Λ attains its maximum over S at x;*

(b) *$F + \Lambda$ attains its minimum over $\mathscr{A}$ at x.*

Proof. The sufficiency is immediate from the fact that (a) and (b) imply, for arbitrary $z \in S$,

$$F(z) = (F + \Lambda)(z) - \Lambda(z) \geqq (F + \Lambda)(x) - \Lambda(x) = F(x). \tag{6.3}$$

To prove the necessity, let x denote any optimal arc. In the space $\mathscr{A} \times R^1$, we consider two convex sets, the epigraph of F (i.e., the set of all pairs (z, α) such that $\alpha \geqq F(z)$) and $S \times (-\infty, F(x)]$. The latter has a nonempty interior which does not meet the former, so that the two sets can be separated by a closed hyperplane. The hyperplane cannot be "vertical," because of our strict feasibility assumption, and therefore it is the graph of a certain continuous linear functional Λ on $\mathscr{A}$. Since both sets contain the point $(x, F(x))$, properties (a) and (b) hold for Λ. Furthermore, since S has a nonempty interior relative to the norm $\|\cdot\|_{\mathscr{C}}$ (Lemma 3), Λ must actually be continuous relative to $\|\cdot\|_{\mathscr{C}}$, due to (a), and hence Λ can be represented as in (6.2). (For the integration by parts, see [4].)

The virtue of Lemma 6 is that the minimization in (b) corresponds to a control problem without state constraints. Namely, one has

$$(6.4)\qquad (F + \Lambda)(x) = \int_0^1 f_1(t, x(t), \dot{x}(t) - A(t)x(t))\,dt + l_1(x(0), x(1)),$$

where

$$(6.5)\qquad f_1(t, x, u) = f(t, x, u) - p_0(t)\cdot(A(t)x + u),$$

$$(6.6)\qquad l_1(x(0), x(1)) = l(x(0), x(1)) - p_0(0)\cdot x(0) + p_0(1)\cdot x(1).$$

The functions f_1 and l_1 satisfy the same assumptions as f and l. Thus we can use Theorem 4 of the preceding section to characterize (b).

At the same time, the situation in (a) of Lemma 6 can be characterized by results in [14]. These results make use of the following concept.

Let $K(t)$ denote a convex cone in R^n (containing the origin) for each $t \in [0, 1]$. An R^n-valued measure μ on $[0, 1]$ is said to be *$K(t)$-valued* if the Radon–Nikodym derivative $d\mu/d\theta$ satisfies

$$(6.7)\qquad \frac{d\mu}{d\theta}(t) \in K(t), \quad \theta\text{-a.e.},$$

where θ is any positive measure on $[0, 1]$ with respect to which μ is absolutely continuous. (This property is independent of the particular θ.)

LEMMA 7. *Under the interiority assumption, a functional of the form (6.2) attains its maximum over S at x if and only if ($x \in S$ and) the measure dp_0 is $N(t, x(t))$-valued (where $N(t, x(t))$ is the cone of normals to $X(t)$ at $x(t)$).*

Proof. If $\mathscr{A}$ were replaced by $\mathscr{C}$ in the definition of S this would be Corollary 6A of [14], since our interiority assumption on the multifunction $X : t \to X(t)$ implies lower semicontinuity [14, Lemma 2]. The result follows in the present case because Λ is continuous in the norm $\|\cdot\|_{\mathscr{C}}$, and S as a subset of $\mathscr{A}$ is dense in the corresponding subset of $\mathscr{C}$ (Lemma 3).

Our main result on necessary and sufficient conditions, Theorem 5 below, concerns the following conditions.

GENERALIZED OPTIMALITY CONDITIONS. *These are the same as the optimality conditions in § 5, except that $p \in \mathscr{B}$ rather than $p \in \mathscr{A}$, and*

$$(6.8)\qquad \textit{the singular part of } dp \textit{ is } N(t, x(t))\textit{-valued}.$$

Of course, if p is of bounded variation, the derivative $\dot{p}(t)$ in condition (5.5) does exist for almost every t, although it is not necessarily true that p is the integral of $\dot{p}$. The singular part of the measure dp may be regarded as the "singular dual control contribution," in the sense that one has

$$(6.9)\qquad dp(t) = -A^*(t)p(t)\,dt + d\mu(t),$$

where

$$(6.10)\qquad d\mu(t) = w(t)\,dt + (\text{singular part}).$$

The generalized optimality conditions reduce to the previous ones if there are no state constraints, since then $N(t, x(t)) \equiv \{0\}$. (To say that the singular part of dp is $\{0\}$-valued is to say that p is absolutely continuous.) More generally, condition (6.8) implies that the singular part of dp is concentrated in the set of t values for which $x(t)$ lies on the boundary of $X(t)$. If p is discontinuous at t, we get the jump condition

(6.11) $$p(t^+) - p(t^-) \in N(t, x(t)).$$

The generalized optimality conditions, except for transversality, may be regarded as the natural extension of the Hamiltonian "equations" (5.8) to allow for $p \in \mathscr{B}$, instead of just $p \in \mathscr{A}$.

THEOREM 5. *Let the assumptions of integrability, interiority, attainability and strict feasibility be satisfied. Then, in order that an arc $x \in \mathscr{A}$ be optimal in the given control problem, it is necessary and sufficient that there exist a function $p \in \mathscr{B}$ for which the generalized optimality conditions are satisfied.*

Proof. We are in the situation of Lemma 6. Property (a) of Lemma 6 has already been characterized in Lemma 7. On the other hand, we have observed that property (b) characterizes x as an optimal arc for a certain control problem without state constraints, corresponding to the functions f_1 and l_1 in (6.5) and (6.6). The integrability and attainability assumptions on f and l carry over to f_1 and l_1, as may easily be checked. Thus Theorem 4 is valid for the unconstrained problem with f_1 and l_1. The function h_1 which corresponds to f_1, as h does to f in (2.6), is

(6.12) $$h_1(t, x, p) = h(t, x, p + p_0(t)) + p_0(t) \cdot A(t)x.$$

Thus the optimality conditions for the unconstrained problem, expressed in terms of a function $p_1 \in \mathscr{A}$, are:

(6.13) $$\dot{x}(t) = A(t)x(t) + u(t) \quad \text{and} \quad \dot{p}_1(t) = -A^*(t)p_1(t) + w_1(t) \quad \text{a.e.},$$

(6.14) $$u(t) \in \partial_p h_1(t, x(t), p_1(t)) = \partial_p h(t, x(t), p_0(t) + p_1(t)) \quad \text{a.e.},$$

(6.15) $$w_1(t) \in -\partial_x h_1(t, x(t), p_1(t)) = -\partial_x h(t, x(t), p_0(t) + p_1(t)) - A^*(t)p_0(t) \quad \text{a.e.},$$

(6.16) $$(p_1(0), -p_1(1)) \in \partial l_1(x(0), x(1)) = \partial l(x(0), x(1)) - (p_0(0), -p_0(1)).$$

We may conclude therefore, as an intermediate step, that x is an optimal arc if and only if $x \in S$ and there exist functions $p_0 \in \mathscr{B}$ and $p_1 \in \mathscr{A}$ such that dp_0 is $N(t, x(t))$-valued and conditions (6.13) through (6.16) are satisfied.

We can write dp_0 as

(6.17) $$dp_0(t) = \dot{p}_0(t)\, dt + d\mu(t),$$

where μ is a certain singular measure. Then dp_0 is $N(t, x(t))$-valued, and x belongs to S if and only if μ is $N(t, x(t))$-valued and $\dot{p}_0(t) \in N(t, x(t))$ a.e. (The latter implies that $N(t, x(t)) \neq \varnothing$ a.e., so that $x(t) \in X(t)$ a.e.; then $x \in S$ by Lemma 3.)

Suppose now that the preceding conditions are satisfied by x, p_0 and p_1, and let

(6.18) $$p(t) = p_0(t) + p_1(t),$$

(6.19) $$w(t) = \dot{p}_0(t) + A^*(t)p_0(t) + w_1(t).$$

Then $\dot{p}(t) = \dot{p}_0(t) + \dot{p}_1(t)$, and the singular part of dp is μ. A simple check shows that the conditions at hand reduce to the generalized optimality conditions for x and p.

Conversely, suppose that x and p satisfy the generalized optimality conditions. It would be possible to show that p can be written in the form (6.18) in such a way that the preceding intermediate conditions are satisfied. However, this approach requires a complicated measurability argument. We therefore proceed more directly, via the theory of subgradients and the fact (Lemma 4) that the optimality conditions (5.3), (5.4) and (5.5) can be expressed in the Hamiltonian form (5.8). Since H_0 is the Hamiltonian corresponding to the Lagrangian L_0 in (4.20), we can express these conditions equivalently in the Lagrangian form

$$(\dot{p}(t), p(t)) \in \partial L_0(t, x(t), \dot{x}(t)) \quad \text{a.e.,} \tag{6.20}$$

where $\partial L_0(t, \cdot, \cdot)$ denotes the set of subgradients (in $R^n \times R^n$) of the convex function $L_0(t, \cdot, \cdot)$ [10, p. 212]. As observed in the proof of Theorem 3, the given control problem amounts to minimizing the functional F_0 in (4.21) over all of $\mathscr{A}$. Thus we need only show that (6.20), (5.6) and (6.8) imply

$$F_0(z) \geqq F_0(x) \quad \text{for every } z \in \mathscr{A}. \tag{6.21}$$

Fixing $z \in \mathscr{A}$, we observe from (6.20), (5.6) and the definition of "subgradient" that

$$\begin{aligned} L_0(t, z(t), \dot{z}(t)) \geqq L(t, x(t), \dot{x}(t)) &+ (z(t) - x(t)) \cdot \dot{p}(t) \\ &+ (\dot{z}(t) - \dot{x}(t)) \cdot p(t) \quad \text{a.e.,} \end{aligned} \tag{6.22}$$

while

$$l(z(0), z(1)) \geqq l(x(0), x(1)) + (z(0) - x(0)) \cdot p(0) - (z(1) - x(1)) \cdot p(1). \tag{6.23}$$

Integrating both sides of (6.22) and adding to (6.23), we get

$$F_0(z) \geqq F_0(x) - \int_0^1 (z(t) - x(t))\, d\mu(t), \tag{6.24}$$

where μ is the singular part of dp. Unless $F_0(z) = +\infty$ (in which event $F_0(z) \geqq F_0(x)$ trivially), $z(t)$ must belong to $X(t)$ for every t (cf. Lemma 3). Then, since μ is $N(t, x(t))$-valued, the integral in (6.24) is nonpositive, so that $F_0(z) \geqq F_0(x)$.

Remark. The preceding argument shows that the generalized optimality conditions are sufficient even without the assumptions of attainability and strict feasibility, as long as at least one feasible arc exists. (The existence of a feasible arc was used in concluding from (6.21) that x is itself feasible, i.e., $F_0(x) \neq +\infty$. Actually, this assumption is unnecessary, in view of Lemma 8 in the next section.)

7. The generalized dual problem. Theorem 5 is incomplete in comparison with Theorem 4, because the meaning of the functions p that appear in the generalized optimality conditions is unexplained. Presumably such functions are generalized solutions to the dual problem. We show now that this is true in a certain precise sense.

Returning to the function s in (4.9), we define for an R^n-valued measure μ:

$$\int_0^1 s(t, d\mu(t)) = \int_0^1 s(t, (d\mu/d\theta)(t))\, d\theta(t), \tag{7.1}$$

where θ is any positive measure with respect to which μ is absolutely continuous. (Since $s(t, w)$ is positively homogeneous as a function of w, this formula is independent of the particular θ.) Under our interiority assumption, s has the measurability needed for this definition, and the integral is well-defined (possibly $+\infty$, but not $-\infty$) [14, §4].

It will be recalled that the dual control problem in § 4 consists of minimizing the Bolza functional G in (4.5) over the space $\mathscr{B}$. We take the *generalized dual problem* to be that of minimizing over the space $\mathscr{B}$ the functional

$$\bar{G}(p) = G(p) + \int_0^1 s(t, d\mu(t)), \tag{7.2}$$

where μ is the singular part of dp. Since $s(t, 0) \equiv 0$, it is clear that $\bar{G}$ agrees with G on $\mathscr{A}$. Thus $\bar{G}$ is a certain extension of G from $\mathscr{A}$ to the larger space $\mathscr{B}$. The topological nature of this extension is described below (Theorem 6). It may be ascertained that $\bar{G}$ is again convex. (This can be deduced using (4.11) and the definitions. It is also shown by the last part of the proof of Theorem 6, which assumes only "integrability" and "interiority.")

LEMMA 8. *Let the assumptions of integrability and interiority hold. Then the generalized optimality conditions are satisfied by $x \in \mathscr{A}$ and $p \in \mathscr{B}$ if and only if $x \in S$ and*

$$-\min_S F = -F(x) = \bar{G}(p) = \min_{\mathscr{B}} \bar{G}. \tag{7.3}$$

Proof. Define L_0 as in (4.20) and M_0 as in (4.25); thus L_0 and M_0 are the Lagrangians dual to each other that correspond to the Hamiltonian H_0 given by (5.7). The Hamiltonian "equations" (5.8), which according to Lemma 4 express the optimality conditions (5.3), (5.4) and (5.5), can then [10, p. 212] be expressed as

$$L_0(t, x(t), \dot{x}(t)) + M_0(t, p(t), \dot{p}(t)) = x(t) \cdot \dot{p}(t) + x(t) \cdot p(t) \quad \text{a.e.}, \tag{7.4}$$

where for arbitrary $x \in \mathscr{A}$ and $p \in \mathscr{B}$ it would be true that

$$L_0(t, x(t), \dot{x}(t)) + M_0(t, p(t), \dot{p}(t)) \geqq x(t) \cdot \dot{p}(t) + \dot{x}(t) \cdot p(t) \quad \text{a.e.} \tag{7.5}$$

On the other hand, if θ is any positive measure on $[0, 1]$ with respect to which both Lebesgue measure on $[0, 1]$ and the singular part μ of the measure dp are absolutely continuous, we can write the conditions (6.8) and $x \in S$ as

$$r(t, x(t)) + s(t, (d\mu/d\theta)(t)) = x(t) \cdot (d\mu/d\theta)(t), \quad \theta\text{-a.e.}, \tag{7.6}$$

where

$$r(t, x(t)) = \begin{cases} 0 & \text{if } x(t) \in X(t), \\ +\infty & \text{if } x(t) \notin X(t). \end{cases} \tag{7.7}$$

(In view of Lemma 3, we have $x \in S$ if $r(t, x(t))$ is finite θ-almost everywhere, as implied by (7.6).) For arbitrary $x \in \mathscr{A}$ and $p \in \mathscr{B}$ (with θ depending on p as above), it would be true that

$$r(t, x(t)) + s(t, (d\mu/d\theta)(t)) \geqq x(t) \cdot (d\mu/d\theta)(t) \quad \theta\text{-a.e.} \tag{7.8}$$

Multiplying both sides of (7.5) by $(dt/d\theta)(t)$, adding this inequality to (7.8) and integrating with respect to θ, we obtain, since

$$L_0(t, x, \dot{x}) + r(t, x) = L_0(t, x, \dot{x}), \tag{7.9}$$

the inequality

$$\begin{aligned} &\int_0^1 L_0(t, x(t), \dot{x}(t))\,dt + \int_0^1 M_0(t, p(t)\dot{p}(t))\,dt + \int_0^1 s(t, d\mu(t)) \\ &\geqq \int_0^1 x(t)\,dp(t) + \int_0^1 \dot{x}(t)\cdot p(t)\,dt. \end{aligned} \tag{7.10}$$

Thus (7.10) holds for arbitrary $x \in \mathscr{A}$ and $p \in \mathscr{B}$, with equality if and only if (7.4) and (7.6) are satisfied. We note next that the transversality condition (5.6) can be expressed as

$$l(x(0), x(1)) + m(p(0), p(1)) = x(0)\cdot p(0) - x(1)\cdot p(1), \tag{7.11}$$

where for arbitrary $x \in \mathscr{A}$ and $p \in \mathscr{B}$ we would have

$$l(x(0), x(1)) + m(p(0), p(1)) \geqq x(0)\cdot p(0) - x(1)\cdot p(1). \tag{7.12}$$

(This is immediate from the definitions.) Adding (7.12) to (7.10), we get the inequality

$$F_0(x) + \bar{G}(p) \geqq 0 \tag{7.13}$$

(F_0 as in (4.21) and (4.22)) for arbitrary $x \in \mathscr{A}$ and $p \in \mathscr{B}$, with equality if and only if (7.4), (7.6) and (7.11) hold. Since the latter conditions are equivalent to the generalized optimality conditions, Lemma 8 is proved.

COROLLARY. *The functions $p \in \mathscr{B}$ in Theorem 5 are precisely the solutions to the generalized dual problem.*

It remains only to show that the solutions to the generalized dual problem can be construed as limits of minimizing (generalized) sequences in the previous dual problem. This is a corollary of the following theorem. Here, by the *weak* topology* on $\mathscr{B}$, we mean the topology induced by the linear functionals

$$p \to a\cdot p(0) + \int_0^1 y(t)\,dp(t), \quad (a, y) \in R^n \times \mathscr{C}. \tag{7.14}$$

THEOREM 6. *Let the assumptions of integrability, interiority and boundedness be satisfied. Then $\bar{G}$ is the lower semicontinuous extension of G to $\mathscr{B}$ in the weak* topology. In other words, for each $p \in \mathscr{B}$ one has*

$$\bar{G}(p) = \liminf G(p_i), \quad p_i \in \mathscr{A}, \quad p_i \to p, \tag{7.15}$$

where the limit is taken over all weak-convergent generalized sequences.*

Proof. For each $(a, y) \in R^n \times \mathscr{C}$, let $\varphi(a, y)$ denote the infimum of

$$\int_0^1 f(t, x(t) + y(t), \dot{x}(t) - A(t)[x(t) + y(t)])\,dt + l(x(0) + a, x(1)) \tag{7.16}$$

over all $x \in \mathscr{A}$ satisfying

$$x(t) + y(t) \in X(t) \quad \text{for every } t. \tag{7.17}$$

We shall demonstrate that

$$\varphi(a, y) = \sup_{p \in \mathscr{A}} \left\{ a \cdot p(0) + \int_0^1 y(t)\, dp(t) - G(p) \right\}, \tag{7.18}$$

while on the other hand, for every $p \in \mathscr{B}$,

$$\bar{G}(p) = \sup_{a \in R^n} \sup_{y \in \mathscr{C}} \left\{ a \cdot p(0) + \int_0^1 y(t)\, dp(t) - \varphi(a, y) \right\}. \tag{7.19}$$

This will imply by the fundamental theorem on conjugate convex functions (see [5]) that $\bar{G}$ is the lower semicontinuous extension of G to $\mathscr{B}$ in the weak* topology.

As a matter of fact, (7.18) is a case of Theorem 3. For the functions

$$f^y(t, x, u) = f(t, x + y(t), u - A(t)y(t)), \tag{7.20}$$

$$l^a(x(0), x(1)) = l(x(0) + a, x(1)), \tag{7.21}$$

and sets

$$X^y(t) = X(t) - y(t) \quad \text{(translation)}, \tag{7.22}$$

$\varphi(a, y)$ denotes the infimum in the control problem, where

$$\int_0^1 f^y(t, x(t), \dot{x}(t) - A(t)x(t))\, dt + l^a(x(0), x(1)) \tag{7.23}$$

is minimized over all $x \in \mathscr{A}$ satisfying

$$x(t) \in X^y(t) \quad \text{for every } t. \tag{7.24}$$

In the corresponding dual control problem, we have

$$g^y(t, p, w) = g(t, p, w) - (w - A^*(t)p) \cdot y(t), \tag{7.25}$$

$$m^a(p(0), p(1)) = m(p(0), p(1)) - a \cdot p(0). \tag{7.26}$$

Thus the dual problem consists of minimizing

$$G(p) - a \cdot p(0) - \int_0^1 y(t) \cdot \dot{p}(t)\, dt \tag{7.27}$$

over all $p \in \mathscr{A}$, so that (7.18) is just equation (4.19) in Theorem 3 in the case of f^y, l^a and $X^y(t)$. The interiority assumption in Theorem 3 is satisfied by $X^y(t)$, since it is satisfied by $X(t)$ and the function y is continuous. The boundedness assumption is likewise satisfied: the recession functions of $f^y(t, \cdot, \cdot)$ and l^a and the recession cones of the sets $X^y(t)$ are actually identical with those of $f(t, \cdot, \cdot)$, l and $X(t)$. As for the integrability assumption, the question is whether the function

$$\begin{aligned} h^y(t, x, p) &= \sup_u \{p \cdot u - f^y(t, x, u)\} \\ &= h(t, x + y(t), p) + p \cdot A(t)y(t) \end{aligned} \tag{7.28}$$

is summable in t for each fixed $x \in R^n$ and $p \in R^n$. The last term in (7.28) is summable in t, because $A(t)$ is summable in t and $y(t)$ is continuous in t. According to our integrability assumption on h, $h(t, x, p)$ is finite and summable in t for each x and p. Since the function $x \to -h(t, x, p)$ is convex, this implies [14, Cor. 2A] that $h(t, x(t), p(t))$ is summable in t for all bounded, measurable functions x and p. Hence, in particular, $h(t, x + y(t), p)$ is summable in t for $x \in R^n$, $y \in \mathscr{C}$ and $p \in R^n$, and it follows that $h^y(t, x, p)$ is summable in t. This verifies that the hypothesis of Theorem 3 is satisfied for f^y, l^a and $X^y(t)$, and equation (7.18) is thereby established.

We now turn to the proof of (7.19), which is by direct calculation employing results in [14]. The supremum in (7.19) is

$$\sup_{a \in R^n} \sup_{y \in \mathscr{C}} \sup_{x \in \mathscr{A}} \left\{ a \cdot p + \int_0^1 y(t)\, dp(t) + \int_0^1 L_0(t, x(t) + y(t), \dot{x}(t))\, dt + l(x(0) + a, x(1)) \right\} \tag{7.29}$$

(cf. proof of Theorem 3), where L_0, as before, is given by (4.20). This can also be written as

$$\begin{aligned} &\sup_{c_0 \in R^n} \sup_{z \in \mathscr{C}} \sup_{x \in \mathscr{A}} \left\{ (c_0 - x(0)) \cdot p(0) + \int_0^1 (z(t) - x(t))\, dp(t) - \int_0^1 L_0(t, z(t), \dot{x}(t))\, dt - l(c_0, x(1)) \right\} \\ &= \sup_{c_0 \in R^n} \sup_{z \in \mathscr{C}} \sup_{x \in \mathscr{A}} \left\{ c_0 \cdot p(0) - x(1) \cdot p(1) + \int_0^1 z(t)\, dp(t) + \int_0^1 p(t) \cdot \dot{x}(t)\, dt - \int_0^1 L_0(t, z(t), y(t))\, dt - l(c_0, c_1) \right\} \\ &= m(p(0), p(1)) + \sup_{z \in \mathscr{C}} \left\{ \int_0^1 z(t)\, dp(t) + Q(z) \right\}, \end{aligned} \tag{7.30}$$

where

$$Q(z) = \sup_{x \in \mathscr{A}} \left\{ \int_0^1 p(t) \cdot \dot{x}(t)\, dt - \int_0^1 L_0(t, z(t), \dot{x}(t))\, dt \right\}. \tag{7.31}$$

Let H_0 be the function in (5.7), so that

$$H_0(t, z(t), p(t)) = \sup_{v \in R^n} \{ p(t) \cdot v - L_0(t, z(t), v) \}. \tag{7.32}$$

We claim that

$$Q(z) = \int_0^1 H_0(t, z(t), p(t))\, dt, \qquad z \in \mathscr{C}, \quad p \in \mathscr{B}. \tag{7.33}$$

The verification uses the fact, already remarked earlier in the proof, that $h(t, z(t), p(t))$ is summable in t for all bounded, measurable functions z and p (and hence for all $z \in \mathscr{C}$ and $p \in \mathscr{B}$). This implies, first of all, that if $z \in \mathscr{C}$ and

$$z(t) \in X(t) \quad \text{for every } t, \tag{7.34}$$

the function $(t, p) \to H(t, z(t), p)$ is summable in $t \in [0, 1]$ as well as finite and convex in $p \in R^n$. It follows then from [14, Cor. 2A] and (7.32) that

$$\int_0^1 H_0(t, z(t), p(t))\, dt = \sup \left\{ \int_0^1 [p(t) \cdot v(t) - L_0(t, z(t), v(t))]\, dt \right\} \tag{7.35}$$

for every bounded, measurable function p, where the supremum is taken over all *summable* functions $v:[0, 1] \to R^n$. Therefore, (7.33) is valid if (7.34) holds. On the other hand, if (7.34) does not hold, then the set of t values for which $z(t) \notin X(t)$ is of positive measure, due to our interiority assumption (Lemma 3). Then the integral in (7.33) is unambiguously $-\infty$, while the integral of L_0 in (7.31) is unambiguously $+\infty$ for every $x \in \mathscr{A}$. The latter implies that the supremum defining $Q(z)$ is $-\infty$. Thus (7.33) is valid even if (7.34) does not hold.

Summarizing to this point, we have shown that the supremum in (7.19) can be written as

$$m(p(0), p(1)) + \sup_{z \in \mathscr{C}} \left\{ \int_0^1 z(t)\, dp(t) + \int_0^1 H_0(t, z(t), p(t))\, dt \right\} \tag{7.36}$$

for arbitrary $p \in \mathscr{B}$. Now let

$$q(t, z) = -H_0(t, z, p(t)), \tag{7.37}$$

so that

$$\{z \in R^n | q(t, z) < +\infty\} = X(t). \tag{7.38}$$

The function q is convex in $z \in R^n$ (since $h(t, x, p)$ is concave in x), and the supremum in (7.36) is

$$\sup_{z \in \mathscr{C}} \left\{ \int_0^1 z(t)\, dp(t) - \int_0^1 q(t, z(t))\, dt \right\}. \tag{7.39}$$

To calculate the latter, we use [14, Thm. 5]. The hypothesis of this theorem requires q to be lower semicontinuous ($\not\equiv +\infty$) as a function of z for each t, measurable on $[0, 1] \times R^n$ with respect to the σ-field generated by products of Lebesgue sets in $[0, 1]$ and Borel sets in R^n, and

$$\int_V |q(t, z)|\, dt < +\infty \quad \text{for } z \in Z \tag{7.40}$$

whenever $V \subset [0, 1]$ and $Z \subset R^n$ are open sets such that $Z \subset X(t)$ for all $t \in T$. (In addition, the set (7.38) is required to satisfy conditions equivalent to our interiority assumption here; see [14, Lemma 2].)

Postponing for a moment the verification of these properties of q, we note that the theorem in question asserts

$$\begin{aligned} &\sup_{z \in \mathscr{C}} \left\{ \int_0^1 z(t)\, dp(t) - \int_0^1 q(t, z(t))\, dt \right\} \\ &\quad = \int_0^1 q^*(t, \dot{p}(t))\, dt + \int_0^1 \hat{q}^*(t, d\mu(t)), \end{aligned} \tag{7.41}$$

where $q^*(t, \cdot)$ is the convex function conjugate to $q(t, \cdot)$, $\hat{q}^*(t, \cdot)$ denotes the recession function of $q^*(t, \cdot)$, and $d\mu$ is the singular part of dp. In fact,

$$\hat{q}^*(t, w) = \sup_{x \in X(t)} w \cdot x = s(t, w) \tag{7.42}$$

in view of (7.38) [13, Thm. 13.3], while by definition,

$$\begin{aligned} q^*(t, w) &= \sup_{z \in R^n} \{z \cdot w - q(t, s)\} \\ &= g^*(t, p(t), w + A^*(t)p(t)) \end{aligned} \tag{7.43}$$

(see (5.7) and (4.11)). Thus (7.41), inserted in (7.36), yields the expression

$$m(p(0), p(1)) + \int_0^1 g^*(t, p(t), \dot{p}(t) + A^*(t)p(t))\, dt + \int_0^1 s(t, d\mu(t)) \tag{7.44}$$

for the supremum in (7.19), and this is $\bar{G}(p)$, as desired.

We complete the proof of Theorem 6 by verifying that q does have the properties listed above (preceding (7.40)). The lower semicontinuity of $q(t, z)$ in $z \in R^n$ follows from (7.32): since $L_0(t, \cdot, \cdot)$ is a lower semicontinuous, convex function on $R^n \times R^n$, this formula implies that $H_0(t, \cdot, \cdot)$ is a "lower closed" concave-convex function on $R^n \times R^n$ [13, p. 357]. This "closedness" of $H_0(t, \cdot, \cdot)$ entails upper semicontinuity in the concave argument, because H_0 nowhere has the value $+\infty$ [13, pp. 356–357]. To see the measurability of q in the required sense, we represent

$$q(t, z) = q_1(t, z) + q_2(t, z), \tag{7.45}$$

where

$$q_1(t, z) = -h(t, z, p(t)) - p(t) \cdot A(t)z \tag{7.46}$$

and

$$q_2(t, z) = \begin{cases} 0 & \text{if } z \in X(t), \\ +\infty & \text{if } z \notin X(t). \end{cases} \tag{7.47}$$

Recalling that $h(t, z, p(t))$ is summable in z (since p, being a function in $\mathscr{B}$, is measurable and bounded), we observe that $q_1(t, z)$ is certainly measurable in t for fixed $z \in R^n$, as well as finite and convex in z for fixed $t \in [0, 1]$. It follows that q_1 is measurable on $[0, 1] \times R^n$ [15, Prop. 1, ff]. On the other hand, the set

$$\{(t, z) | z \in X(t)\} \subset [0, 1] \times R^n \tag{7.48}$$

is measurable in the specified sense, because, under our interiority assumption, it is a countable intersection of open sets—this has already been shown in the proof of Theorem 3, following display (4.26). Thus the function q_2 is also measurable, and the measurability of $q = q_1 + q_2$ may be concluded. Finally, we remark that the summability of $h(t, z, p(t))$ in t for each $z \in R^n$ trivially implies the summability property required of q. Theorem 6 is now proved.

COROLLARY. *Let the assumptions of integrability, interiority and boundedness be satisfied. Then a function $p \in \mathscr{B}$ solves the generalized dual problem if and only if p is the weak* limit of a generalized sequence (p_i) in $\mathscr{A}$ which is a minimizing sequence* *in the earlier dual problem; that is,*

$$\lim_i G(p_i) = \inf_{\mathscr{A}} G. \tag{7.49}$$

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), pp. 369–412.

[2] J. DIEUDONNÉ, *Sur la separation des ensembles convexes*, Math. Ann., 163 (1966), pp. 1–3.

[3] J. E. FUNK AND E. G. GILBERT, *Some sufficient conditions for optimality in control problems with state space constraints*, this Journal, 8 (1970), pp. 498–504.

[4] L. M. GRAVES, *The Theory of Functions of a Real Variable*, McGraw-Hill, New York, 1956.

[5] J. J. MOREAU, *Fonctionelles convexes*, mimeographed lecture notes, Séminaire sur les equations aux derivées partielles, Collège de France, 1966–1967.

[6] L. W. NEUSTADT, *An abstract variational theory with applications to a broad class of problems. II: Applications*, this Journal, 5 (1967), pp. 90–137.

[7] ———, *A general theory of extremals*, J. Comput. System Sci., 3 (1969), pp. 57–92.

[8] ———, *Sufficiency conditions and a duality theory for mathematical programming in arbitrary linear spaces*, Nonlinear Programming, J. B. Rosen et al., eds., Academic Press, New York, 1970, pp. 323–348.

[9] C. OLECH, *Existence theorems for optimal problems with vector-valued cost function*, Trans. Amer. Math. Soc., 136 (1969), pp. 159–180.

[10] R. T. ROCKAFELLAR, *Conjugate convex functions in optimal control and the calculus of variations*, J. Math. Anal. Appl., 32 (1970), pp. 174–222.

[11] ———, *Generalized Hamiltonian equations for convex problems of Lagrange*, Pacific J. Math., 33 (1970), pp. 411–427.

[12] ———, *Existence and duality theorems for convex problems of Bolza*, Trans. Amer. Math. Soc., 159 (1971), pp. 1–40.

[13] ———, *Convex Analysis*. Princeton Univ. Press, Princeton, N.J., 1970.

[14] ———, *Integrals which are convex functionals, II*, Pacific J. Math., 39 (1971), pp. 439–469.

[15] ———, *Convex integral functionals and duality*, Contributions to Nonlinear Functional Analysis, E. Zarantonello, ed., Academic Press, New York, 1971, pp. 215–235.

[16] M. M. TSVETANOV, *Duality in problems of the calculus of variations and optimal control*, Doctoral thesis, Moscow State University, Moscow, 1970.

[17] V. L. MAKAROV, *A property of solutions of a problem of continuous linear and convex programming*, Dokl. Akad. Nauk SSSR, 176 (1967) = Soviet Math. Dokl., 8 (1967), pp. 1246–1247.