PREPRINT, Prepared on 5 February 1996

BOLZA PROBLEMS WITH GENERAL TIME CONSTRAINTS

P. D. LOEWEN† AND R. T. ROCKAFELLAR‡

Abstract. This work provides necessary conditions for optimality in problems of optimal control expressed as instances of the Generalized Problem of Bolza, with the added feature that the fundamental planning interval is allowed to vary. A central product of the analysis is a generalization of the conservation-of-Hamiltonian condition for problems on either fixed or variable intervals. The results allow for unprecedented generality in the problem data, and apply also in the presence of unilateral state constraints. They are derived from known results for fixed-interval problems under the hypothesis that the time-dependence of the objective integrand has the same modest level of regularity as the state-dependence.

Key words. Optimal control, calculus of variations, Bolza problem, free time, minimum-time problem, Erdmann condition, Euler-Lagrange condition, Hamiltonian condition, transversality condition, nonsmooth analysis.

AMS subject classifications. 49K05, 49K24, 49K15

1. Overview. This paper provides necessary conditions for optimality in a general optimal control problem where the endpoints of the underlying time interval are choice variables. The problem is to choose a nondegenerate interval $[a, b]$ and an absolutely continuous function (or *arc*) $x\colon [a, b] \to \mathbb{R}^n$ in such a way as to

$$\text{(P)} \qquad \begin{aligned} &\text{minimize} \quad l(a, x(a), b, x(b)) + \int_a^b L(t, x(t), \dot{x}(t))\, dt \\ &\text{subject to} \quad x(t) \in X(t) \cap \Omega_t \ \forall t \in [a, b]. \end{aligned}$$

The usefulness of this simple-looking model is directly correlated with the mildness of the hypotheses under which conclusive results can be obtained. In the current work, both the endpoint cost l and the integrand L are allowed to be nondifferentiable and even to take the value $+\infty$. These features allow for enormous flexibility in modelling applied problems—in particular, a wide range of differential and endpoint constraints can be introduced implicitly by encoding them in l and L. The only explicit constraint in problem (P) involves two time-varying sets: Ω_t, an open set, is introduced to allow us to make statements about *local* optimality; $X(t)$, a closed set, imposes a unilateral constraint on allowed state values at time t.

†loew@math.ubc.ca; Dept. of Mathematics, University of British Columbia, Vancouver, British Columbia, Canada. Research supported by Canada's Natural Science and Engineering Research Council..

‡rtr@math.washington.edu; Dept. of Mathematics, University of Washington, Seattle, WA, 98195. Research supported by the National Science Foundation under grant DMS-9200303..

The introduction of extended-real-valued functionals as practical modelling tools in dynamic optimization can be traced to the early work of Rockafellar [19] in the convex case. Many authors have addressed the technical challenges raised in this context since the 1970's: we mention in particular Rockafellar [20, 21] and Clarke [4, 5]. The hypotheses in this paper are in some respects weaker than those of Clarke [5], and moreover allow for a variable time interval. We handle this additional complication here by using the classical Erdmann transform to reduce our problem to a fixed-time problem in which the time plays the role of an additional state variable. This forces us to assume a time-dependence of the problem data whose regularity matches what we require on the state-dependence. Clarke, Loewen, and Vinter [8], using other methods, have discussed optimality conditions for related problems in which the time-dependence is merely measurable.

Our most general assertions about minimizers in problem (P) are presented in Theorem 3.1 below, which is based on the five hypotheses we state in subsection 1.2. In order to demonstrate the power and scope of the main result, however, we devote the paper's first two sections to a discussion of its consequences for some familiar model problems arising in applications.

1.1. Subgradients and Normals. The hypotheses under which we will present necessary conditions for optimality in problem (P) allow for problem data that lie well beyond the scope of classical first-order approximations. Indeed, even the formulation of these hypotheses relies on certain basic elements of nonsmooth analysis. In this work, the symbols $\partial f(x)$ and $\partial^\infty f(x)$ describe the sets of *limiting proximal subgradients* and *singular limiting proximal subgradients* associated with an extended-valued function $f\colon \mathbb{R}^m \to \mathbb{R} \cup \{\infty\}$ at a point x such that $f(x)$ is finite and the epigraph of f is locally closed near $(x, f(x))$. We write $N_S(s)$ for the set of *limiting proximal normals* associated with a set S at a point $s \in S$ near which S is locally closed. For simplicity, we refer to these objects simply as "subgradients," "singular subgradients," and "normals"; further discussion of these constructions and their relationships to other fundamental objects is now widely available ([6], [12], [17], [25]). However, we will later refer to the basic relationships

$$\begin{aligned} \partial f(x) &= \{\xi \,:\, (\xi, -1) \in N_{\operatorname{epi} f}(x, f(x))\}, \\ \partial^\infty f(x) &= \{\xi \,:\, (\xi, 0) \in N_{\operatorname{epi} f}(x, f(x))\}. \end{aligned} \tag{1.1}$$

Throughout the paper we write $\mathbb{B}$ for the *closed* unit ball centred at the origin in various Euclidean spaces distinguished by the context.

1.2. Hypotheses. An arc $\overline{x}$ is given, together with its associated interval of definition $[\overline{a}, \overline{b}]$; one has $\overline{b} - \overline{a} > 0$. For some $\rho > 0$, the open set

$$\Omega = \left\{(t, x) \,:\, |(t, x) - (r, \overline{x}(r))| < \rho \text{ for some } r \in [\overline{a}, \overline{b}]\right\}, \tag{1.2}$$

with sections $\Omega_t = \{x \,:\, (t, x) \in \Omega\}$, is one in which the data of problem (P) satisfy hypotheses (H1)–(H5) below. In these conditions, and throughout the paper, a key ingredient is the *Hamiltonian* $H\colon \Omega \times \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$, defined by

$$H(t,x,p) := \sup\{\langle p\,,\, v\rangle - L(t,x,v) \,:\, v \in \mathbb{R}^n\}\,. \tag{1.3}$$

In (H4) and throughout the sequel, we use the simplifying notation $\overline{L}(t) := L\big(t, \overline{x}(t), \dot{\overline{x}}(t)\big)$.

(H1) *The endpoint cost* $l\colon \mathbb{R} \times \mathbb{R}^n \times \mathbb{R} \times \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ *is lower semicontinuous on the set* $\big\{(a,x,b,y) \,:\, |(a,x) - (\overline{a}, \overline{x}(\overline{a}))| < \rho,\ \big|(b,y) - (\overline{b}, \overline{x}(\overline{b}))\big| < \rho\big\}$.

(H2) *For each fixed* (t,x) *in* Ω, *the function* $v \mapsto L(t,x,v)$ *is convex on* $\mathbb{R}^n$.

(H3) *The function* L *is lower semicontinuous on* $\Omega \times \mathbb{R}^n$ *and epi-continuous in* (t,x)*: that is, for any point* (t,x,v) *in* $\Omega \times \mathbb{R}^n$ *where* $L(t,x,v)$ *is finite, and any sequence* $(t_k, x_k) \to (t,x)$, *there exists a sequence* $v_k \to v$ *along which* $L(t_k, x_k, v_k) \to L(t,x,v)$.

(H4) *There are positive constants* δ *and* κ *such that the following statement is true for almost every* t *in* $\big[\overline{a}, \overline{b}\big]$*: for every point* (r,x,v) *in* $\Omega \times \mathbb{R}^n$ *satisfying the three conditions*

(i) $|(r,x) - (t, \overline{x}(t))\,| < \rho$,

(ii) $\big|v - \dot{\overline{x}}(t)\big| < \delta\big[1 + \big|\dot{\overline{x}}(t)\big|\big]$,

(iii) $\big|L(r,x,v) - L(t, \overline{x}(t), \dot{\overline{x}}(t))\big| < \delta\big[1 + \big|L(t, \overline{x}(t), \dot{\overline{x}}(t))\big|\big]$,

one has the subgradient inequality

$$|(u,w)| \le \kappa[1 + |p| + |H(r,x,p)|] \qquad \forall (u,w,p) \in \partial L(r,x,v).$$

(H5) *The state constraint multifunction* $X\colon \big[\overline{a} - \rho, \overline{b} + \rho\big] \rightrightarrows \mathbb{R}^n$ *has closed graph.*

The key to understanding hypotheses (H2)–(H4) lies in a geometric interpretation of problem (P) in terms of the multifunction $E\colon \Omega \rightrightarrows \mathbb{R}^n \times \mathbb{R}$ defined by

$$E(t,x) := \operatorname{epi} L(t,x,\,\cdot\,) = \{(v,\gamma) \,:\, \gamma \ge L(t,x,v)\}\,.$$

Note that a scalar arc y obeys $y(t) \ge y(a) + \int_a^t L(r, x(r), \dot{x}(r))\, dr$ if and only if $(\dot{x}(t), \dot{y}(t)) \in E(t, x(t))$ a.e. This observation furnishes a link between the objective values in (P) and the endpoint values of trajectories for the differential inclusion based on E: this reformulation, detailed in [15], makes the geometry easier to appreciate.

Hypothesis (H2), together with the lower semicontinuity part of (H3), ensures that the set $E(t,x)$ is closed and convex for every (t,x)—an important condition in the existence theory for problem (P).

Hypothesis (H3) is equivalent to the requirement that the multifunction E be continuous on Ω, in the sense that for every (t,x) in Ω,

$$\liminf_{(t',x') \to (t,x)} E(t',x') \supseteq E(t,x) \supseteq \limsup_{(t',x') \to (t,x)} E(t',x')\,. \tag{1.4}$$

As a consequence of Wijsman's theorem, which asserts that the correspondence between Lagrangian and Hamiltonian under the Legendre-Fenchel transform preserves epi-continuity, we can express (H3) in three equivalent ways:

(i) The set $\operatorname{epi} L(t, x, \cdot\,)$ depends continuously on (t, x) in Ω;

(ii) The set $\operatorname{epi} H(t, x, \cdot\,)$ depends continuously on (t, x) in Ω;

(iii) The function H is lower semicontinuous on $\Omega \times \mathbb{R}^n$, and epi-continuous in (t, x): that is, for any point (t, x, p) in $\Omega \times \mathbb{R}^n$ at which H is finite, and any sequence $(t_k, x_k) \to (t, x)$, there is a sequence $p_k \to p$ along which $H(t_k, x_k, p_k) \to H(t, x, p)$.

(Rockafellar [24, Prop. 2.1] provides details.) This equivalence reveals an appealing symmetry between the requirements on L and H, and is the key to establishing the continuity of the function $t \mapsto H(t, \overline{x}(t), p(t))$ implicit in conclusion (b) of Theorem 1.4.

Hypothesis (H4) is a subgradient sufficient condition for a graphical localization of E around the points $\big(t, \overline{x}(t), \dot{\overline{x}}(t), \overline{L}(t)\big)$ to display a type of Lipschitz continuity. The continuity property we require was introduced for general multifunctions by Aubin [3]; in our context, we require that the graph of $\overline{x}$ have a neighbourhood in which, for some constants R and K, one has

$$E(s, x) \cap \big(\big(\dot{\overline{x}}(t), \overline{L}(t)\big) + R\mathbb{B}\big) \subseteq E(t, y) + K|(s, x) - (t, y)|\mathbb{B}.$$

Mordukhovich [16] has shown that a necessary and sufficient condition for the validity of this "Aubin continuity property" to hold near a given point $\big(t, \overline{x}(t), \dot{\overline{x}}(t), \overline{L}(t)\big)$ is that $|(u, w)| \le K|(p, -\lambda)|$ for all vectors $(u, w, p, -\lambda)$ normal to the set $\operatorname{gph} E = \operatorname{epi} L$ at points near $\big(t, \overline{x}(t), \dot{\overline{x}}(t), \overline{L}(t)\big)$. The subgradient inequality in (H4) arises from just such considerations (compare (1.1)), weakened somewhat thanks to a change of variable we introduce in the proofs.

Another helpful perspective on (H4) is available if we consider the slightly stronger hypothesis obtained by omitting the Hamiltonian term from the right side of the central inequality. For any integrand L that satisfies the resulting hypothesis with respect to $\overline{x}$, the same hypothesis holds (with a larger coefficient κ) for any integrand $L + G_1$ involving a function G_1 that is Lipschitz with respect to (t, x). This insensitivity to Lipschitzian perturbations reveals that the real job of (H4) is to regulate the non-Lipschitz dependence on (t, x) near the nominal trajectory.

It is important to note that both (H3) and (H4) can be checked easily for several classes of integrand L. We describe some of these reductions in the discussion of special cases below, and treat this matter more generally in Section 4.

1.3. Regularity Issues. The minimization problem we state as (P) allows an open competition between all absolutely continuous functions x. With notation borrowed from distribution theory, we might call it "the $W^{1,1}$ problem," to distinguish it from "the $W^{1,\infty}$ problem," in which the same objective functional is to be minimized over all arcs x for which $\dot{x} \in L^\infty$. There are smooth classical examples in which these two problems have different infima, and both are attained. However, it often happens that the solution to the $W^{1,1}$ problem is indeed Lipschitzian, and consequently solves the $W^{1,\infty}$ problem too. In this situation, our results are particularly simple to state: our hypotheses simplify—the term $\left|\dot{\overline{x}}(t)\right|$ can be dropped from the right-hand side of (H4)(ii), and our conclusions improve—we can assert the identity in the first line of Theorem 1.4(b) *for all* t, rather than just for the set of t with full measure described somewhat more explicitly in the full statement of Theorem 3.1. The question of whether the necessary conditions presented here actually apply to all solutions of the $W^{1,\infty}$ problem (even when the $W^{1,1}$ problem has a smaller infimum) is beyond the scope of our current investigation. (It is addressed implicitly in Ioffe and Rockafellar [11].) A related line of research involves the search for general hypotheses which will guarantee that all solutions to the $W^{1,1}$ problem actually lie in $W^{1,\infty}$: here there remains much to be done, although significant progress under stronger hypotheses has been reported by Clarke and Vinter [9, 10] (see also Clarke and Loewen [7], Ambrosio, Ascenzi, and Buttazzo [1]).

Dual to the issue of regularity for the minimizing arc $\overline{x}$ is the question of smoothness for the adjoint quantities (h,p) arising in the necessary conditions. To treat the full version of problem (P), these must be allowed to jump (see Section 3, and Rockafellar [22]); however, they turn out to be absolutely continuous in the important class of problems where state constraints are not involved. These are covered by the present formulation simply by setting $X(t) = \mathbb{R}^n$, which is perfectly consistent with (H5). More generally, a state constraint will be called *inactive* along the arc $\overline{x}$ if the graph of $\overline{x}$ lies in the interior of the graph of X, i.e.,

$$\left\{(t,\overline{x}(t))\,:\,t\in\left[\overline{a},\overline{b}\right]\right\}\subseteq \operatorname{int}\left\{(t,x)\,:\,t\in\left[\overline{a},\overline{b}\right],\ x\in X(t)\right\}. \tag{1.5}$$

Under this condition, the specialization of Theorem 3.1 to Lipschitzian minimizers takes the following form.

1.4. Theorem. *Let the arc $\overline{x}$ with interval $\left[\overline{a},\overline{b}\right]$ provide the minimum in problem* (P). *In addition to* (H1)–(H5)*, suppose that $\overline{x}$ is Lipschitzian, and that the state constraint is inactive along $\overline{x}$. Then some absolutely continuous function (h,p) with values in $\mathbb{R}\times\mathbb{R}^n$ satisfies either the normal conditions or the singular conditions below.*

[Normal Conditions]:

(a) $$\left(\dot h(t),\dot p(t)\right)\in \operatorname{co}\left\{(u,w)\,:\,(-u,w,p(t))\in\partial L\big(t,\overline{x}(t),\dot{\overline{x}}(t)\big)\right\}$$
$$=\operatorname{co}\left\{(u,w)\,:\,\big(u,-w,\dot{\overline{x}}(t)\big)\in\partial H\big(t,\overline{x}(t),p(t)\big)\right\}\ \text{ a.e. } t\in\left[\overline{a},\overline{b}\right].$$

(b) $h(t)=H(t,\overline{x}(t),p(t)) \quad \forall t\in\left[\overline{a},\overline{b}\right]$;
$h(t)=\left\langle p(t)\,,\,\dot{\overline{x}}(t)\right\rangle - L\big(t,\overline{x}(t),\dot{\overline{x}}(t)\big)\quad$ a.e. $t\in\left[\overline{a},\overline{b}\right]$.

(c) $\big(-h(\overline{a}),p(\overline{a}),h(\overline{b}),-p(\overline{b})\big)\in\partial l\big(\overline{a},\overline{x}(\overline{a}),\overline{b},\overline{x}(\overline{b})\big)$.

[Singular Conditions]: *One has $|(h(t),p(t))|>0$ for all t in $\left[\overline{a},\overline{b}\right]$, and*

(a^∞) $(-\dot{h}(t),\dot{p}(t)) \in \operatorname{co}\big\{(u,w) \,:\, (u,w,p(t)) \in \partial^\infty L(t,\overline{x}(t),\dot{\overline{x}}(t))\big\}$ a.e. $t \in [\overline{a},\overline{b}]$,

(b^∞) $h(t) = \big\langle p(t)\,,\,\dot{\overline{x}}(t)\big\rangle$ a.e. $t \in \big[\overline{a},\overline{b}\big]$,

(c^∞) $\big(-h(\overline{a}),p(\overline{a}),h(\overline{b}),-p(\overline{b})\big) \in \partial^\infty l\big(\overline{a},\overline{x}(\overline{a}),\overline{b},\overline{x}(\overline{b})\big)$.

In particular, if the only function pair (h,p) *satisfying conditions* (a^∞)–(c^∞) *is identically zero, then the normal conditions are satisfied.*

Before undertaking a technical justification of Theorem 1.4, whose proof is in Section 5, we pause to discuss some of its useful implications.

1.5. Conservation of the Hamiltonian. Although the general formulation of our problem allows for a variable time interval, the results also apply in situations where the fundamental interval is fixed in advance. Consider, for example, the problem of choosing an arc x in $AC([0,1];\mathbb{R}^n)$ with graph in Ω so as to minimize

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

The simple redefinition

$$l(a,x,b,y) := \widetilde{l}(x,y) + \Psi_{\{0\}}(a) + \Psi_{\{1\}}(b)$$

puts this problem into the form we are calling (P). The hypotheses imposed on $\widetilde{l}$ and L through (H1)–(H5) in this paper are slightly stronger than those of our related work [15] on the fixed-time problem, but here the conclusions are stronger too. Under the extra regularity conditions we are now assuming for the time-dependence of L, we obtain in the normal case of Theorem 1.4 a generalization of the classical equation

$$\dot{h}(t) = H_t(t,\overline{x}(t),p(t))\,, \text{ where } h(t) = H(t,\overline{x}(t),p(t))\,. \tag{1.6}$$

In particular, if the integrand L is free of explicit dependence on t, then the Hamiltonian must be constant along normal extremal trajectories—a conclusion that recalls the law of conservation of energy in classical mechanics, or the Second Weierstrass-Erdmann Condition in the calculus of variations. This condition is often useful in simplifying the solution of problems on a preassigned interval, and it remains valid even in the more general setting of problems with variable time. The difference is that in the fixed-time case, the transversality condition in Theorem 1.4(c) involves the right side

$$\partial l(0,\overline{x}(0),1,\overline{x}(1)) = \Big\{(r,\alpha,s,\beta) \,:\, r,s \in \mathbb{R},\ (\alpha,\beta) \in \partial\widetilde{l}(\overline{x}(0),\overline{x}(1))\Big\},$$

and therefore gives no information about the numerical values of $h(0)$ and $h(1)$. In variable-time problems, condition 1.4(c) is richer, and provides a nontrivial algebraic condition linking these values with other ingredients of the extremal system.

1.6. Maximization Conditions. Through the definition of H, conclusion (b) of Theorem 1.4 encodes the maximization condition

$$(1.7)\quad \big\langle p(t)\,,\,\dot{\overline{x}}(t)\big\rangle - L\big(t,\overline{x}(t),\dot{\overline{x}}(t)\big) = \max_{v\in\mathbb{R}^n}\{\langle p(t)\,,\,v\rangle - L(t,\overline{x}(t),v)\,\}\ \text{ a.e. } t\in\big[\overline{a},\overline{b}\big]\,.$$

Under the convexity hypothesis (H2), this is equivalent to $p(t)\in\partial_v L\big(t,\overline{x}(t),\dot{\overline{x}}(t)\big)$, whose dual formulation as $\dot{\overline{x}}(t)\in\partial_p H(t,\overline{x}(t),p(t))$ can be expressed as

$$(1.8)\quad \big\langle p(t)\,,\,\dot{\overline{x}}(t)\big\rangle - H(t,\overline{x}(t),p(t)) = \max_{q\in\mathbb{R}^n}\big\{\big\langle q\,,\,\dot{\overline{x}}(t)\big\rangle - H(t,\overline{x}(t),q)\big\}\ \text{ a.e. } t\in\big[\overline{a},\overline{b}\big]\,.$$

The equivalent assertions (1.7)–(1.8) can be viewed as consequences of the adjoint inclusion in Theorem 1.4(a). Indeed, that inclusion implicitly asserts the nonemptiness of the first set on the right side for almost all t. Thus, for almost all t, there are points (u,w) such that

$$(1.9)\qquad (-u,w,p(t))\in\partial L\big(t,\overline{x}(t),\dot{\overline{x}}(t)\big)\,.$$

Rockafellar [24, Prop. 2.2] shows that under our hypotheses on L, this inclusion implies the partial-subgradient relation $p(t)\in\partial_v L\big(t,\overline{x}(t),\dot{\overline{x}}(t)\big)$. (Clarke [4, Prop. 2.5.3] has the same result for Lipschitz L.) The definition of the convex subgradient allows this inclusion to be expressed as shown in (1.7). Thus inclusion (1.9) implies

$$(1.10)\qquad \big\langle p(t)\,,\,\dot{\overline{x}}(t)\big\rangle - L\big(t,\overline{x}(t),\dot{\overline{x}}(t)\big) = H(t,\overline{x}(t),p(t))\,.$$

This derivation of (1.10) from (1.9) proceeds equally well for more general evaluation points than those used here, and its more general form will be used repeatedly in the proofs presented in Section 5. As an illustration, we note that the subgradient inequality stated in (H4) is equivalent to

$$(1.11)\qquad |(u,w)|\le\kappa[1+|p|+|\langle p\,,\,v\rangle - L(r,x,v)|]\quad \forall (u,w,p)\in\partial L(r,x,v).$$

2. Problems with Explicit Velocity Constraints. The fully intrinsic formulation of (P) has the conceptual advantage of a simple statement and a rich heritage of classical antecedents. On the other hand, its importance in practice comes from its applicability to problems with velocity and endpoint constraints beyond the scope of its classical forebears. An example arising frequently in practice is the problem of choosing an interval $[a,b]$ and an arc x on $[a,b]$ in order to

$$(2.1)\qquad \begin{aligned} &\text{minimize } \Gamma[a,b;x] := g_1(a,x(a),b,x(b)) + \int_a^b G_1(t,x(t),\dot{x}(t))\,dt\\ &\text{subject to } \dot{x}(t)\in F(t,x(t)) \text{ a.e. } t\in[a,b],\\ &\qquad (a,x(a),b,x(b))\in S,\\ &\qquad x(t)\in\Omega(t)\ \forall t\in[a,b]. \end{aligned}$$

This problem can be recognized as an instance of (P) with endpoint cost $l = g_1 + \Psi_S$, integrand $L = G_1 + \Psi_{\mathrm{gph}\,F}$, and state constraint $X = \mathbb{R}^n$. (Nontrivial state constraints can be treated by the same methods.) Suitable hypotheses of Lipschitz continuity on g_1 and G_1, together with mild conditions on the multifunction F, will not only establish (H1)–(H5) for l and L, but also allow us to derive from the conclusions of Theorem 1.4 the usual dichotomy between the normal and abnormal forms of standard necessary conditions. To express this concisely, we write for any $\lambda \geq 0$

$$
\begin{aligned}
l_\lambda(a,x,b,y) &:= \lambda g_1(a,x,b,y) + \Psi_S(a,x,b,y),\\
L_\lambda(t,x,v) &:= \lambda G_1(t,x,v) + \Psi_{\mathrm{gph}\,F}(t,x,v),\\
H_\lambda(t,x,p) &:= \sup_{v\in\mathbb{R}^n} \{\langle p\,,\,v\rangle - L_\lambda(t,x,v)\}\\
&\qquad\quad = \sup\{\langle p\,,\,v\rangle - \lambda G_1(t,x,v)\,:\, v\in F(t,x)\}\,.
\end{aligned} \tag{2.2}
$$

The hypotheses in question are as follows. Again we state them in terms of a given arc $\overline{x}$ with associated interval $[\overline{a},\overline{b}]$ assumed to solve the given problem, and a fixed open set Ω containing the graph of $\overline{x}$. We assume that $\overline{x}$ is Lipschitzian.

(h_1) *The target set S is closed, and the endpoint cost g_1 is Lipschitzian on the set*

$$\big\{(a,x,b,y)\,:\, |(a,x)-(\overline{a},\overline{x}(\overline{a}))| < \rho,\ \big|(b,y)-(\overline{b},\overline{x}(\overline{b}))\big| < \rho\big\}\,.$$

(h_2) *For each fixed (t,x) in Ω, both the function $v\mapsto L(t,x,v)$ and the set $F(t,x)$ are convex.*

(h_3) *The function G_1 is finite-valued and continuous on $\Omega\times\mathbb{R}^n$. The multifunction F is continuous on Ω, in the sense that (1.4) holds for $E = F$, at every point (t,x) in Ω.*

(h_4) *There are positive constants δ and R for which almost all t in $(\overline{a}-\rho,\overline{b}+\rho)$ have this property: for every point (r,x,v) in $\Omega\times\mathbb{R}^n$ satisfying the three conditions*

$$|(r,x)-(t,\overline{x}(t))| < \rho, \qquad \big|v-\dot{\overline{x}}(t)\big| < \delta, \qquad v\in F(r,x),$$

both subgradient estimates below are valid:

$$
\begin{aligned}
|(u_1,w_1,p_1)| &\leq R &\qquad& \forall (u_1,w_1,p_1)\in\partial G_1(r,x,v),\\
|(u_\nu,w_\nu)| &\leq R[1+|p_\nu|] && \forall (u_\nu,w_\nu,p_\nu)\in N_{\mathrm{gph}\,F}(r,x,v)\,.
\end{aligned}
$$

Notice that the conditions on F imposed by (h_3)–(h_4) amount to nothing more than continuity in general together with a sort of uniform Aubin property near the minimizing arc of interest. Since the arc $\overline{x}$ is Lipschitzian, the first subgradient inequality in (h_4) will be satisfied (for some R) by any locally Lipschitzian integrand G_1, while a sufficient condition for the second is the qualification condition

$$(u,w,0)\in\partial L_0\big(t,\overline{x}(t),\dot{\overline{x}}(t)\big) \implies (u,w)=(0,0), \qquad \text{a.e. } t\in[\overline{a},\overline{b}]\,.$$

(This follows from Theorem 4.2 below, and the special structure of L.) In particular, there is no requirement that F be bounded or compact-valued—the next result applies even when $F \equiv \mathbb{R}^n$ and (2.1) reduces to a standard variational problem with Lipschitzian data and a general target set.

2.1. THEOREM. *Let the arc $\overline{x}$ with interval $[\overline{a}, \overline{b}]$ provide the minimum in problem (2.1). Assume $\dot{\overline{x}} \in L^\infty([\overline{a}, \overline{b}]; \mathbb{R}^n)$, along with (h_1)–(h_4). Then there exist a scalar $\lambda \in \{0,1\}$ and an absolutely continuous pair $(h,p)\colon [\overline{a}, \overline{b}] \to \mathbb{R} \times \mathbb{R}^n$, with $\lambda + |(h(t), p(t))| > 0$ for all t in $[\overline{a}, \overline{b}]$, such that*

$$\text{(a)} \quad \left(\dot{h}(t), \dot{p}(t)\right) \in \operatorname{co}\left\{(u,w) \,:\, (-u, w, p(t)) \in \partial L_\lambda\big(t, \overline{x}(t), \dot{\overline{x}}(t)\big)\right\}$$
$$= \operatorname{co}\left\{(u,w) \,:\, (u, -w, \dot{\overline{x}}(t)) \in \partial H_\lambda(t, \overline{x}(t), p(t))\right\} \ \text{ a.e. } t \in [\overline{a}, \overline{b}] .$$

$$\text{(b)} \quad h(t) = H_\lambda(t, \overline{x}(t), p(t)) \quad \forall t \in [\overline{a}, \overline{b}] ;$$
$$h(t) = \langle p(t) \,,\, \dot{\overline{x}}(t)\rangle - \lambda G_1\big(t, \overline{x}(t), \dot{\overline{x}}(t)\big) \quad \text{ a.e. } t \in [\overline{a}, \overline{b}] .$$

$$\text{(c)} \quad \big(-h(\overline{a}), p(\overline{a}), h(\overline{b}), -p(\overline{b})\big) \in \partial l_\lambda\big(\overline{a}, \overline{x}(\overline{a}), \overline{b}, \overline{x}(\overline{b})\big).$$

Again, we stress that condition (b) in the statement of Theorem 2.1 expresses two (equivalent) maximization conditions valid almost everywhere:

$$\text{(2.3)} \quad \begin{aligned} \dot{\overline{x}}(t) &\in \partial_p H_\lambda(t, \overline{x}(t), p(t)) = \arg\max_{v \in \mathbb{R}^n} \left\{\langle p(t) \,,\, v\rangle - L_\lambda(t, \overline{x}(t), v)\right\}, \\ p(t) &\in \partial_v L_\lambda(t, \overline{x}(t), \dot{\overline{x}}(t)) = \arg\max_{q \in \mathbb{R}^n} \left\{\langle q \,,\, \dot{\overline{x}}(t)\rangle - H_\lambda(t, \overline{x}(t), q)\right\}. \end{aligned}$$

Proof. We reduce to an application of Theorem 1.4, by choosing $l = l_1$ and $L = L_1$. It is evident that (h_1) and (h_2) imply that these functions satisfy (H1) and (H2); checking that (h_3) implies (H3) is an easy exercise. To recognize (H4) as a consequence of (h_4), use the constant R provided by (h_4) to define

$$\kappa = R^2 + 2R.$$

Then consider any point (r, x, v) of the sort described in (H4)(i)–(iii) and any subgradient $(u, w, p) \in \partial L(r, x, v)$. Conditions (H4)(i)–(ii) imply that (h_4)(i)–(ii) hold for the evaluation point (r, x, v); condition (H4)(iii) certainly requires the finiteness of $L(r, x, v)$, so that (h_4)(iii) must follow. With the three prerequisites of (h_4) in place, we observe that the given subgradient has a decomposition as

$$(u, w, p) = (u_1, w_1, p_1) + (u_\nu, w_\nu, p_\nu)$$

for some $(u_1, w_1, p_1) \in \partial G_1(r, x, v)$ and some $(u_\nu, w_\nu, p_\nu) \in N_{\operatorname{gph} F}(r, x, v)$. Consequently the two estimates in (h_4) give the second inequality in the otherwise elementary estimate

$$\begin{aligned} |(u,w)| &\le |(u_1, w_1)| + |(u_\nu, w_\nu)| \\ &\le R + R[1 + |p_\nu|] \\ &\le R + R[1 + |p_1 + p_\nu| + |-p_1|] \\ &\le R + R[1 + |p|] + R^2 \\ &\le \kappa[1 + |p|] . \end{aligned}$$

In the last step we have used the choice $\kappa = R^2 + 2R$ mentioned above. The resulting inequality confirms (H4).

We may now apply Theorem 1.4 to $\overline{x}$ on $\big[\overline{a}, \overline{b}\big]$. If the conclusions of Theorem 1.4 hold in normal form, then conditions (a)–(c) follow immediately, with $\lambda = 1$. The nontriviality condition is evident. Suppose, therefore, that we have only the singular conditions of Theorem 1.4, satisfied by some nonvanishing pair (h, p). Again the nontriviality condition is immediate, but now the special form of l and L allows conclusions (a^∞)–(c^∞) to be simplified considerably. The key observation is that, since the endpoint cost g_1 and integrand G_1 are locally Lipschitz near the evaluation points of interest, we have

$$\partial^\infty l\big(\overline{a}, \overline{x}(\overline{a}), \overline{b}, \overline{x}(\overline{b})\big) = N_S\left(\overline{a}, \overline{x}(\overline{a}), \overline{b}, \overline{x}(\overline{b})\right) = \partial l_0\big(\overline{a}, \overline{x}(\overline{a}), \overline{b}, \overline{x}(\overline{b})\big)\,,$$
$$\partial^\infty L\big(t, \overline{x}(t), \dot{\overline{x}}(t)\big) = N_{\mathrm{gph}\, F}\left(t, \overline{x}(t), \dot{\overline{x}}(t)\right) = \partial L_0\big(t, \overline{x}(t), \dot{\overline{x}}(t)\big)\,.$$

Thus the transversality inclusion (c) with $\lambda = 0$ follows directly from condition (c^∞) of Theorem 1.4, while condition (a^∞) of that result gives

$$(*) \qquad \Big(\dot h(t), \dot p(t)\Big) \in \mathrm{co}\left\{(u, w) \,:\, (-u, w, p(t)) \in \partial L_0\big(t, \overline{x}(t), \dot{\overline{x}}(t)\big)\,\right\} \ \text{a.e. } t \in \big[\overline{a}, \overline{b}\big]\,.$$

Apply to this inclusion the results of Rockafellar [24, Thm. 1.1], as they pertain to the function $L_0 = \Psi_{\mathrm{gph}\, F}$. Hypothesis ($h_3$) implies that L_0 is lower semicontinuous and has the required epicontinuity property. Furthermore, any point $(u, w, 0) \in \partial^\infty L_0\big(t, \overline{x}(t), \dot{\overline{x}}(t)\big) = N_{\mathrm{gph}\, F}\left(t, \overline{x}(t), \dot{\overline{x}}(t)\right)$ must satisfy

$$\alpha(u, w, 0) \in N_{\mathrm{gph}\, F}\left(t, \overline{x}(t), \dot{\overline{x}}(t)\right) \quad \forall \alpha > 0,$$

whereupon (h_4) requires $\alpha|(u, w)| \le R$ for all $\alpha > 0$, i.e., $|(u, w)| = 0$. Thus the calculus qualification of [24] is in force, and we may conclude that

$$\mathrm{co}\left\{(u, w) \,:\, (-u, w, p(t)) \in \partial L_0\big(t, \overline{x}(t), \dot{\overline{x}}(t)\big)\,\right\}$$
$$= \mathrm{co}\left\{(u, w) \,:\, \big(u, -w, \dot{\overline{x}}(t)\big) \in \partial H_0(t, \overline{x}(t), p(t))\right\} \ \text{a.e. } t \in \big[\overline{a}, \overline{b}\big]\,.$$

In conjunction with $(*)$ above, this equation establishes conclusion (a) with $\lambda = 0$.

Turning finally to the maximum condition, we note that inclusion $(*)$ implies $p(t) \in \partial_v L_0\big(t, \overline{x}(t), \dot{\overline{x}}(t)\big) = N_{F(t,\overline{x}(t))}\left(\dot{\overline{x}}(t)\right)$. (See Remark 1.6.) In particular, $\big\langle p(t)\,,\, \dot{\overline{x}}(t)\big\rangle \ge \langle p(t)\,,\, v\rangle$ for all v in $F(t, \overline{x}(t))$, which can be restated as $\big\langle p(t)\,,\, \dot{\overline{x}}(t)\big\rangle = H_0(t, \overline{x}(t), p(t))$. We already have $h(t) = \big\langle p(t)\,,\, \dot{\overline{x}}(t)\big\rangle$ from conclusion (b^∞), so conclusion (b) holds for almost all t. The first line can be upgraded to an equation valid for all t by applying Corollary 3.6, below. ∎

2.2. Minimum-Time Problems. The problem of steering the state trajectory from the origin at time 0 to a moving target set C in least time, subject to given differential constraints, can be expressed as follows:

$$
\begin{aligned}
&\text{minimize } \Lambda[x] := b\\
&\text{subject to } b > 0,\ x \in AC([0,b],\mathbb{R}^n)\,,\\
&\qquad \dot{x}(t) \in F(t,x(t)) \text{ a.e. } t \in [a,b],\\
&\qquad x(0) = 0,\ x(b) \in C(b),\\
&\qquad x(t) \in \Omega_t\ \forall t \in [a,b].
\end{aligned}
\tag{2.4}
$$

This fits the pattern of the case just discussed, with

$$G_1 \equiv 0, \qquad g_1(a,x,b,y) = b, \qquad S = \{(0,0)\} \times \operatorname{gph} C.$$

Thus (h_1) will hold for any target multifunction C whose graph is closed, (h_2) reduces to a convexity requirement on the velocity sets $F(t,x)$, (h_3) changes only slightly, and the first of the subgradient estimates in (h_4) becomes self-evident. In addition, one has $L_\lambda = \Psi_{\operatorname{gph} F} = L_0$, and hence $H_\lambda = H_0$, for all $\lambda \geq 0$. Thus a Lipschitzian minimizer $\overline{x}$ with interval $\left[0,\overline{b}\right]$ must have an associated scalar $\lambda \in \{0,1\}$ and arc $(h,p)\colon\left[0,\overline{b}\right] \to \mathbb{R} \times \mathbb{R}^n$, not both zero, such that

(a) $$\left(\dot{h}(t),\dot{p}(t)\right) \in \operatorname{co}\left\{(u,w)\,:\,(-u,w,p(t)) \in N_{\operatorname{gph} F}\left(t,\overline{x}(t),\dot{\overline{x}}(t)\right)\right\}$$
$$= \operatorname{co}\left\{(u,w)\,:\,(u,-w,\dot{\overline{x}}(t)) \in \partial H_0(t,\overline{x}(t),p(t))\right\} \text{ a.e. } t \in \left[0,\overline{b}\right].$$

(b) $$h(t) = H_0(t,\overline{x}(t),p(t)) \quad \forall t \in \left[\overline{a},\overline{b}\right];$$
$$h(t) = \left\langle p(t)\,,\,\dot{\overline{x}}(t)\right\rangle \quad \text{a.e. } t \in \left[\overline{a},\overline{b}\right].$$

(c) $$\left(-h(0),p(0),h(\overline{b}),-p(\overline{b})\right) \in \lambda(0,0,1,0) + \mathbb{R} \times \mathbb{R}^n \times N_{\operatorname{gph} C}\left(\overline{b},\overline{x}(\overline{b})\right).$$

Here the normality indicator λ shows up only in the transversality condition, which reduces to

(c′) $$\left(h(\overline{b}) - \lambda, -p(\overline{b})\right) \in N_{\operatorname{gph} C}\left(\overline{b},\overline{x}(\overline{b})\right).$$

In the case where the target set multifunction $t \rightrightarrows C(t)$ is actually single-valued, and moves smoothly with time, this condition says

$$\left(h(\overline{b}) - \lambda, -p(\overline{b})\right) \in \left(1, C'(\overline{b})\right)^{\perp}, \quad \text{i.e.,} \quad h(\overline{b}) - \left\langle p(\overline{b})\,,\,C'(\overline{b})\right\rangle = \lambda,$$

in perfect harmony with the classical case.

Another possibility is that the target is stationary, so that $C(t) = C$ for all $t > 0$. In this case $\operatorname{gph} C(\cdot) = \mathbb{R} \times C$, so condition (c′) reduces to a pair of decoupled conditions relating the terminal value of the Hamiltonian to the normality indicator λ, and asserting that the final value of the adjoint arc is associated with a generalized normal to the fixed target set:

$$h(\overline{b}) = \lambda, \qquad -p(\overline{b}) \in N_C\left(\overline{x}(\overline{b})\right).$$

In the further special case where the velocity sets $F(t,x)$ actually have no explicit dependence on t, it follows from conditions (a)(b) that the Hamiltonian must be constant along extremal trajectories, with the fixed value $\lambda = 1$ for normal problems and $\lambda = 0$ for singular ones.

3. Problems with Explicit State Constraints. We return to problem (P) in its full generality, and concentrate on the case where the state constraint $x(t) \in X(t)$ is active at some times along the optimal trajectory. Loosely speaking, the natural space in which to impose this constraint is the space of continuous functions on $\left[\overline{a}-\rho, \overline{b}+\rho\right]$: thus we can expect the corresponding multipliers to lie in the dual of this space, which is a space of measures. An equivalent viewpoint, which harmonizes the notation nicely with the simpler cases studied above, is that we must now consider the adjoint quantities h and p to have *bounded variation,* rather than asserting their absolute continuity as we have done in the absence of state constraints.

The costate vector $(-h, p)$ is allowed to jump only at instants when $(t, \overline{x}(t))$ lies on the boundary of gph X; at such times, the jump direction must be an outward normal to gph X at the point $(t, \overline{x}(t))$. To formulate this principle precisely, we must use the Clarke normal cone [4] to gph X, characterized by

$$\overline{N}_{\operatorname{gph} X}(t,x) = \operatorname{cl}\operatorname{co} N_{\operatorname{gph} X}(t,x) \quad \forall (t,x) \in \operatorname{gph} X.$$

In what follows, we will deal only with points (t,x) in gph X where the cone $N_{\operatorname{gph} X}(t,x)$ is 'pointed', i.e., points for which any finite collection of nonzero normal vectors has a nonzero sum. At any such point, the closure operation on the right side above is superfluous and Clarke's cone $\overline{N}_{\operatorname{gph} X}(t,x)$ is pointed as well. This situation is known to arise if and only if the Clarke tangent cone to gph X at (t,x) has nonempty interior. (See [4] for this characterization and more discussion.)

The Lebesgue-Stieltjes theory establishes a correspondence between vector-valued measures on a real interval $[a,b]$ and (equivalence classes of) functions of bounded variation. (Two functions f and g in $BV([a,b]; \mathbb{R}^n)$ are equivalent when $f(t^-) = g(t^-)$ and $f(t^+) = g(t^+)$ for all t in $[a,b]$; here we define $f(a^-) = f(a)$ and $f(b^+) = f(b)$.) We recall that for any multifunction $\overline{N}\colon [a,b] \rightrightarrows \mathbb{R}^n$ whose values are closed convex cones, the statement that a given measure df is "$\overline{N}(t)$-valued" means that df is absolutely continuous with respect to some nonnegative measure μ on $[a,b]$, and the Radon-Nikodym derivative $\nu = df/d\mu$ satisfies $\nu(t) \in \overline{N}(t)$ for μ-almost all t in $[a,b]$.

We now state our main result. The formulation involves the block-structured $(1+n) \times (1+n)$ matrix

$$A = \begin{bmatrix} -1 & 0 \\ 0 & I \end{bmatrix}. \tag{3.1}$$

3.1. Theorem. *Let the arc $\overline{x}$ with interval $\left[\overline{a}, \overline{b}\right]$ provide the minimum in problem* (P). *Assume* (H1)–(H5). *Under the constraint qualification*

(CQ) the normal cone $N_{\operatorname{gph} X}(t, \overline{x}(t))$ is pointed for all t in $[\overline{a}, \overline{b}]$, *there is a function* (h,p) *of bounded variation on* $[\overline{a},\overline{b}]$, *with values in* $\mathbb{R}\times\mathbb{R}^n$, *for which the singular part of the measure* $(-dh,dp)$ *is* $\overline{N}_{\mathrm{gph}\,X}(t,\overline{x}(t))$*-valued, and in particular whose support is a subset of*

$$\{t\,:\,\overline{N}_{\mathrm{gph}\,X}(t,\overline{x}(t))\neq\{0\}\}=\{t\in[\overline{a},\overline{b}]\,:\,(t,\overline{x}(t))\in\mathrm{bdy}\,\mathrm{gph}\,X(t)\}\,.$$

This function satisfies either the normal conditions or the singular conditions below.

[Normal Conditions]:

$$\begin{aligned}\text{(a)}\quad \big(\dot h(t),\dot p(t)\big)&\in\mathrm{co}\,\{(u,w)\,:\,(-u,w,p(t))\in\partial L\big(t,\overline{x}(t),\dot{\overline{x}}(t)\big)\,\}+A\overline{N}_{\mathrm{gph}\,X}(t,\overline{x}(t))\\ &=\mathrm{co}\,\{(u,w)\,:\,\big(u,-w,\dot{\overline{x}}(t)\big)\in\partial H\big(t,\overline{x}(t),p(t)\big)\,\}+A\overline{N}_{\mathrm{gph}\,X}(t,\overline{x}(t))\\ &\qquad\text{a.e. } t\in[\overline{a},\overline{b}]\,.\end{aligned}$$

$$\begin{aligned}\text{(b)}\quad &h(t)=H(t,\overline{x}(t),p(t))\qquad\text{a.e. } t\in[\overline{a},\overline{b}]\ \text{(see also Prop. 3.5);}\\ &h(t)=\langle p(t)\,,\,\dot{\overline{x}}(t)\rangle-L\big(t,\overline{x}(t),\dot{\overline{x}}(t)\big)\quad\text{a.e. } t\in[\overline{a},\overline{b}]\,.\end{aligned}$$

$$\text{(c)}\quad \big(-h(\overline{a}),p(\overline{a}),h(\overline{b}),-p(\overline{b})\big)\in\partial l\big(\overline{a},\overline{x}(\overline{a}),\overline{b},\overline{x}(\overline{b})\big).$$

[Singular Conditions]: *The pair* (h,p) *is not identically zero, and*

$$(\mathrm{a}^\infty)\quad (\dot h(t),\dot p(t))\in\mathrm{co}\,\{(u,w)\,:\,(-u,w,p(t))\in\partial^\infty L(t,\overline{x}(t),\dot{\overline{x}}(t))\}+A\overline{N}_{\mathrm{gph}\,X}(t,\overline{x}(t))$$

a.e. $t\in[\overline{a},\overline{b}]$,

$$(\mathrm{b}^\infty)\quad h(t)=\langle p(t)\,,\,\dot{\overline{x}}(t)\rangle\ \text{ a.e. } t\in[\overline{a},\overline{b}].$$

$$(\mathrm{c}^\infty)\quad (-h(a),p(a),h(b),-p(b))\in\partial^\infty l(\overline{a},\overline{x}(a),\overline{b},\overline{x}(b)).$$

In particular, if the only function pair (h,p) *satisfying conditions* (a^∞)–(c^∞) *is identically zero, then the normal conditions are satisfied.*

3.2. Remarks.

(i) The equation in conclusion (a) conceals somewhat a sharper assertion that can be made in the normal case. In fact the same measurable selection of $\overline{N}_{\mathrm{gph}\,X}(t,\overline{x}(t))$ can be used on both sides of the equation—that is, there is some measurable $(f(t),y(t))\in\overline{N}_{\mathrm{gph}\,X}(t,\overline{x}(t))$ for which

$$\begin{aligned}\big(\dot h(t)+f(t),\dot p(t)-y(t)\big)&\in\mathrm{co}\,\{(u,w)\,:\,(-u,w,p(t))\in\partial L\big(t,\overline{x}(t),\dot{\overline{x}}(t)\big)\,\}\\ &=\mathrm{co}\,\{(u,w)\,:\,\big(u,-w,\dot{\overline{x}}(t)\big)\in\partial H\big(t,\overline{x}(t),p(t)\big)\,\}\\ &\qquad\text{a.e. } t\in[\overline{a},\overline{b}]\,.\end{aligned}\tag{3.2}$$

(ii) As before, conclusion (b) implies the following (equivalent) maximization conditions for almost all t:

$$\begin{aligned}\dot{\overline{x}}(t)&\in\partial_p H_\lambda(t,\overline{x}(t),p(t))=\arg\max_{v\in\mathbb{R}^n}\{\langle p(t)\,,\,v\rangle-L_\lambda(t,\overline{x}(t),v)\}\,,\\ p(t)&\in\partial_v L_\lambda(t,\overline{x}(t),\dot{\overline{x}}(t))=\arg\max_{q\in\mathbb{R}^n}\{\langle q\,,\,\dot{\overline{x}}(t)\rangle-H_\lambda(t,\overline{x}(t),q)\}\,.\end{aligned}\tag{3.3}$$

3.3. Costate Jumps. The appearance of the Clarke normal cone on the right sides of the adjoint inclusions of Theorem 3.1 signals the possibility that the costate vector $(-h(t), p(t))$ might jump—or, more generally, depart from absolute continuity—at instants when the state constraint is active: such jumps must drive the vector $(-h, p)$ in a direction outward normal to $\operatorname{gph} X$ at the point $(t, \overline{x}(t))$ as understood in the sense of Clarke [4].

Note that if the state constraint is inactive along the solution $\overline{x}$, i.e., if $(t, \overline{x}(t)) \in \operatorname{int}(\operatorname{gph} X)$ for all t in $[\overline{a}, \overline{b}]$, then we have

$$\overline{N}_{\operatorname{gph} X}\,(t, \overline{x}(t)) = \{0\} \qquad \forall t \in [\overline{a}, \overline{b}]\,.$$

This evidently confirms (CQ), but also forces the measure $(-dh, dp)$ to have a $\{0\}$-valued singular part. In other words, the function pair $(-h, p)$ must be indistinguishable from an absolutely continuous pair, and thus we recover the main conclusions of Theorem 1.4 as corollaries of this result, provided $\overline{x}$ is Lipschitz.

A second special case of interest arises when $X(t) = X$ is a fixed closed set independent of t. In this case $\operatorname{gph} X(\cdot) = \mathbb{R} \times X$, so we have

$$\overline{N}_{\operatorname{gph} X}\,(t, \overline{x}(t)) = \{0\} \times \overline{N}_X\,(\overline{x}(t)) \quad \forall t \in [\overline{a}, \overline{b}]\,.$$

In this setting (CQ) requires only that the convex cone $\overline{N}_X\,(\overline{x}(t))$ be pointed in $\mathbb{R}^n$ for all t; under this condition the function h must be absolutely continuous and the adjoint inclusions (in normal form) assert that the singular part of dp is $\overline{N}_X\,(\overline{x}(t))$-valued, while

$$\begin{aligned}\big(\dot h(t), \dot p(t)\big) &\in \operatorname{co}\big\{(u, w) \,:\, (-u, w, p(t)) \in \partial L\big(t, \overline{x}(t), \dot{\overline{x}}(t)\big)\,\big\} + \{0\} \times \overline{N}_X\,(\overline{x}(t)) \\ &= \operatorname{co}\big\{(u, w) \,:\, \big(u, -w, \dot{\overline{x}}(t)\big) \in \partial H\big(t, \overline{x}(t), p(t)\big)\,\big\} + \{0\} \times \overline{N}_X\,(\overline{x}(t)) \\ &\qquad\qquad \text{a.e. } t \in [\overline{a}, \overline{b}]\,.\end{aligned}$$

(A similar reformulation applies to the singular conclusion.)

3.4. Hamiltonian Continuity. Since the adjoint process $(-h, p)$ in Theorem 3.1 is a function of bounded variation, it has finite one-sided limits at every point t in $[\overline{a}, \overline{b}]$. (Recall that by convention, we insist on $h(\overline{a}^-) = h(\overline{a})$ and $h(\overline{b}^+) = h(\overline{b})$, for example.) This observation, together with the epi-continuity of H described in Remark 1.6, allows us to sharpen conclusion (b) of Theorem 3.1. This is the point of our next result, in which m denotes Lebesgue measure and we write

$$I^+ = \bigcup_{R>0} I_R^+, \quad I^- = \bigcup_{R>0} I_R^-, \quad \text{where}$$

$$I_R^+ = \big\{t \in [\overline{a}, \overline{b}) \,:\, \forall \varepsilon > 0,\ m\{s \in [t, t+\varepsilon] \,:\, \big|\dot{\overline{x}}(s)\big| \le R\} > 0\big\}\,,$$

$$I_R^- = \big\{t \in (\overline{a}, \overline{b}] \,:\, \forall \varepsilon > 0,\ m\{s \in [t-\varepsilon, t] \,:\, \big|\dot{\overline{x}}(s)\big| \le R\} > 0\big\}\,.$$

3.5. Proposition. *If the statements in Theorem* 3.1(b) *hold for an arc* $\overline{x}$ *and a pair* (h,p) *of bounded variation on* $\big[\overline{a},\overline{b}\big]$*, then one has*

(a) $h(t^+) = H(t,\overline{x}(t),p(t^+)) \quad \forall t \in I^+,$

(b) $h(t^-) = H(t,\overline{x}(t),p(t^-)) \quad \forall t \in I^-.$

Proof. We prove only (a), since (b) is similar.

Fix any t in I^+. Since h and p have bounded variation on $\big[\overline{a},\overline{b}\big]$, their right limits $h(t^+)$ and $p(t^+)$ exist finitely, and can moreover be realized along any sequence $s_k \to t^+$. But $t \in I_R^+$ for some $R > 0$, so each interval $[t, t+1/k]$ contains a set of positive measure in which $\big|\dot{\overline{x}}(s)\big| \le R$. We may therefore choose a sequence $s_k \to t^+$ along which $\dot{\overline{x}}(s_k)$ exists, the two equations in Theorem 3.1(b) hold, and $\sup_k \big|\dot{\overline{x}}(s_k)\big| \le R$.

Now H is lower semicontinuous on $\Omega \times \mathbb{R}^n$ by Remark 1.6, so we have the inequality

$$(*) \qquad h(t^+) = \liminf_{k\to\infty} H(s_k,\overline{x}(s_k),p(s_k)) \ge H(t,\overline{x}(t),p(t^+))\,.$$

On the other hand, H is epicontinuous on Ω. Line $(*)$ shows that $(t,\overline{x}(t),p(t^+))$ is a point where H is finite, and we have a sequence $(s_k,\overline{x}(s_k))$ converging to $(t,\overline{x}(t))$. Thus there must be a sequence $q_k \to p(t^+)$ along which $H(s_k,\overline{x}(s_k),q_k) \to H(t,\overline{x}(t),p(t^+))$. The maximum condition (3.3) then supplies the inequality

$$(**) \qquad \begin{aligned} h(t^+) &= \lim_{k\to\infty} H(s_k,\overline{x}(s_k),p(s_k)) \\ &\le \liminf_{k\to\infty}\big[H(s_k,\overline{x}(s_k),q_k) - \big\langle q_k - p(s_k)\,,\,\dot{\overline{x}}(s_k)\big\rangle\big] \\ &= H(t,\overline{x}(t),p(t^+)) - 0. \end{aligned}$$

Combining $(*)$ and $(**)$ gives $h(t^+) \le H(t,\overline{x}(t),p(t^+)) \le h(t^+)$, as required. □

The consequences of Proposition 3.4 are strongest when $\overline{x}$ is Lipschitzian. In this case, $I^+ = \big[\overline{a},\overline{b}\big)$ and $I^- = \big(\overline{a},\overline{b}\big]$, so the following corollary is immediate. We have applied this result in the statements of Theorems 1.4 and 2.1.

3.6. Corollary. *If* $\overline{x}$ *is a Lipschitzian arc for which the statements of Theorem 3.1(b) hold for some pair* (h,p) *of bounded variation on* $\big[\overline{a},\overline{b}\big]$*, then one has*

(a) $h(t^+) = H(t,\overline{x}(t),p(t^+)) \quad \forall t \in \big[\overline{a},\overline{b}\big),$

(b) $h(t^-) = H(t,\overline{x}(t),p(t^-)) \quad \forall t \in \big(\overline{a},\overline{b}\big].$

In particular, at any point t *in* $\big[\overline{a},\overline{b}\big]$ *where* $p(t^-) = p(t^+)$*, the redefinitions*

$$p(t) = p(t^-), \quad h(t) = H(t,\overline{x}(t),p(t))$$

render both p *and* h *continuous at* t.

4. Simplified Continuity Conditions. Both the epicontinuity condition (H3) and the Aubin continuity hypothesis (H4) follow from more elementary assumptions when the integrand L has suitable structure. We discuss some of these reductions in this section, paying particular attention to simplifications available when the function $L(t,x,v)$ is convex in (x,v) for each fixed t.

Our first result is a sufficient condition for the Aubin continuity property (H4). A simplified version appears as Theorem 4.2 below.

4.1. PROPOSITION. *Suppose $\dot{\overline{x}}$ and $\overline{L}$ are essentially bounded. Then, upon reducing $\rho > 0$ if necessary, hypothesis* (H4) *holds whenever there exists a multifunction $\Gamma\colon \mathbb{R}\times\mathbb{R}^n\times\mathbb{R}^n \rightrightarrows \mathbb{R}$ with these three properties:*

(a) *The graph of Γ is a compact subset of* epi L*; the images of Γ are compact convex sets.*

(b) *If $\Gamma(t,x,v)\neq\emptyset$ then $L(t,x,v)\in\Gamma(t,x,v)$; moreover, $\Gamma\big(t,\overline{x}(t),\dot{\overline{x}}(t)\big)\neq\emptyset$ for almost all t in $[\overline{a},\overline{b}]$.*

(c) *One has*

$$\forall\gamma\in\Gamma(t,x,v),\qquad (u,w,0,0)\in N_{\operatorname{epi} L}(t,x,v,\gamma)\implies (u,w)=(0,0). \tag{4.1}$$

Conditions (a)–(b) require, among other things, that the graph of Γ be a compact superset of the 'curve' $\big\{(t,\overline{x}(t),\dot{\overline{x}}(t),L\big(t,\overline{x}(t),\dot{\overline{x}}(t)\big)\ :\ t\in[\overline{a},\overline{b}]\big\}$. This curve will admit a compact superset if and only if both $\dot{\overline{x}}$ and $\overline{L}$ are essentially bounded, which we have emphasized by making it an explicit hypothesis in the statement above. In the case where both $\dot{\overline{x}}$ and $\overline{L}$ are continuous, the curve in question is itself compact, and conditions (a)–(b) are satisfied by the multifunction

$$\Gamma(t,x,v)=\begin{cases}\{L(t,x,v)\}, & \text{if } (t,x,v)=\big(t,\overline{x}(t),\dot{\overline{x}}(t)\big),\ t\in[\overline{a},\overline{b}],\\ \emptyset, & \text{otherwise.}\end{cases} \tag{4.2}$$

Somewhat more generally, suppose there is a compact set Γ_0 in $\mathbb{R}\times\mathbb{R}^n\times\mathbb{R}^n$ containing almost all points $\big(t,\overline{x}(t),\dot{\overline{x}}(t)\big)$ and enjoying the additional property that the restriction $L|_{\Gamma_0}$ is continuous. In this case a multifunction satisfying (a)–(b) is

$$\Gamma(t,x,v)=\begin{cases}\{L(t,x,v)\}, & \text{if } (t,x,v)\in\Gamma_0,\\ \emptyset, & \text{otherwise.}\end{cases} \tag{4.3}$$

In both the simple cases just mentioned, and more generally at any point where $\Gamma(t,x,v)=\{L(t,x,v)\}$, the key condition (4.1) can be rewritten in terms of singular subgradients using (1.1):

$$(u,w,0)\in\partial^\infty L(t,x,v)\implies (u,w)=(0,0). \tag{4.4}$$

The discrepancy between (4.1) and (4.4) shows up only when (4.1) refers to normals to epi L at points (r,x,v,γ) where $\gamma > L(r,x,v)$. For convex integrands, the two conditions are equivalent, and can indeed be simplified further, as we will show later in this section.

Proof of Proposition 4.1. The essential boundedness of $\dot{\overline{x}}$ and $\overline{L}$ allow the right sides in (H4)(ii)–(iii) to be replaced by δ. This is the form of (H4) we will confirm.

The first step is local analysis. Fix any point $(t,\overline{x},\overline{v},\overline{\gamma})$ in $\operatorname{gph}\Gamma$. We claim that there exists $M>0$ so large that the following geometrical inequality holds near the given point:

$$\left.\begin{aligned}&|(r,x,v)-(t,\overline{x},\overline{v})|<1/M,\\ &\gamma<\overline{\gamma}+1/M\end{aligned}\right\}\implies$$
$$|(u,w)|\le M[|p|+|q|]\qquad\forall (u,w,p,q)\in N_{\operatorname{epi}L}\,(r,x,v,\gamma)\,.\qquad(*)$$

Indeed, if this claim were false, then it would have to fail for every positive integer m. In this case each m would give rise to a point (r_m,x_m,v_m,γ_m) satisfying the antecedent inequalities in $(*)$, but associated with a normal vector (u_m,w_m,p_m,q_m) in $N_{\operatorname{epi}L}\,(r_m,x_m,v_m,\gamma_m)$ for which

$$(\dagger)\qquad \frac{1}{R}|(u_m,w_m)|>[|p_m|+|q_m|]\,,\qquad |(u_m,w_m,p_m,q_m)|=1.$$

In particular, we have $(r_m,x_m,v_m)\to(t,\overline{x},\overline{v})$ and $\gamma_m\ge L(r_m,x_m,v_m)$, whereupon

$$\begin{aligned}\overline{\gamma}\ge\limsup_{m\to\infty}\gamma_m&\ge\limsup_{m\to\infty}L(r_m,x_m,v_m)\\&\ge\liminf_{m\to\infty}L(r_m,x_m,v_m)\\&\ge L(t,\overline{x},\overline{v})\quad\text{(by lower semicontinuity).}\end{aligned}$$

Conditions (a)–(b) imply that all the limit points of the sequence γ_m lie in the compact interval $\Gamma(t,\overline{x},\overline{v})$; by passing to a subsequence if necessary, we may assume that $\gamma_m\to\widehat{\gamma}$, where $L(t,\overline{x},\overline{v})\le\widehat{\gamma}\le\overline{\gamma}$, and $\widehat{\gamma}\in\Gamma(t,\overline{x},\overline{v})$. Along a further subsequence, the given normals converge to a unit vector (u,w,p,q) with the property that

$$0\ge|p|+|q|,\qquad (u,w,p,q)\in N_{\operatorname{epi}L}\,(t,\overline{x},\overline{v},\widehat{\gamma})\,.$$

Thus we arrive at a contradiction to condition (4.1). This proves the claim involving $(*)$.

The second step is to harness the compact-graph condition (a). Fix any point $(t,\overline{x},\overline{v},\overline{\gamma})$ in $\operatorname{gph}\Gamma$. The previous paragraph provides a constant $M=M(t,\overline{x},\overline{v},\overline{\gamma})$ with the properties specified in $(*)$. Use this constant to define an open set in $\mathbb{R}^{1+n+n+1}$:

$$\Omega(t,\overline{x},\overline{v},\overline{\gamma})=\{(r,x,v,\gamma)\,:\,|(r,x)-(t,\overline{x})|<1/M,\ |v-\overline{v}|<1/M,\ \gamma<\overline{\gamma}+1/M\}\,.$$

Now as $(t,\overline{x},\overline{v},\overline{\gamma})$ runs through $\operatorname{gph}\Gamma$, the open sets $\Omega(t,\overline{x},\overline{v},\overline{\gamma})$ cover $\operatorname{gph}\Gamma$. Thus we can extract a finite list of points $(t_i,\overline{x}_i,\overline{v}_i,\overline{\gamma}_i)$, $i=1,\ldots,N$, in $\operatorname{gph}\Gamma$, such that

$$(**)\qquad \operatorname{gph}\Gamma \subseteq \bigcup_{i=1}^{N} \Omega(t_i,\overline{x}_i,\overline{v}_i,\overline{\gamma}_i).$$

Since the left side is compact and the right side is open, we can enlarge the left side by adding an open ball of small radius while maintaining the indicated inclusion. Thus there is a positive constant δ such that any point (r,x,v,γ) satisfying the three conditions below for some point $(t,\overline{x},\overline{v},\overline{\gamma})$ in $\operatorname{gph}\Gamma$ will lie in the right side of $(**)$:

$$|(r,x)-(t,\overline{x})|<\delta, \qquad |v-\overline{v}|<\delta, \qquad \gamma<\overline{\gamma}+\delta.$$

In particular, fix any time t in $\left[\overline{a},\overline{b}\right]$ at which the inclusion in (b) holds. Then the point $(t,\overline{x},\overline{v},\overline{\gamma})=\big(t,\overline{x}(t),\dot{\overline{x}}(t),L\big(t,\overline{x}(t),\dot{\overline{x}}(t)\big)\big)$ lies in $\operatorname{gph}\Gamma$, so the three hypotheses below are enough to situate the point $(r,x,v,L(r,x,v))$ in the right side of $(**)$:

(i) $|(r,x)-(t,\overline{x}(t))|<\delta$,

(ii) $\left|v-\dot{\overline{x}}(t)\right|<\delta$,

(iii) $L(r,x,v)<L(t,\overline{x}(t),\dot{\overline{x}}(t))+\delta$

This means that there is some index i for which $(r,x,v,L(r,x,v))$ lies in $\Omega(t_i,\overline{x}_i,\overline{v}_i,\overline{\gamma}_i)$, and hence that any vector (u,w,p,q) in $N_{\operatorname{epi}L}\,(r,x,v,L(r,x,v))$ obeys

$$|(u,w)|\le M_i[|p|+|q|]\le \widehat{M}[|p|+|q|]\,,$$

where $\widehat{M}=\max\{M_1,\ldots,M_N\}$. To summarize, conditions (i)–(iii) imply

$$(u,w,p)\in\partial L(r,x,v) \implies |(u,w)|\le \widehat{M}[1+|p|]\,.$$

This completes the proof of (H4)—in fact it gives a sharper estimate than we need; the required constants are δ and $\widehat{M}$. □

4.2. THEOREM. *Suppose both $\dot{\overline{x}}$ and $\overline{L}$ are essentially bounded. Then, upon reducing $\rho>0$ if necessary, the following condition implies* (H4)*: there exists $\delta>0$ so small that for almost all $t\in\left[\overline{a},\overline{b}\right]$, the three inequalities*

$$|(r,x)-(t,\overline{x}(t))|<\delta, \qquad \left|v-\dot{\overline{x}}(t)\right|<\delta, \qquad \left|\gamma-\overline{L}(t)\right|<\delta$$

imply the geometrical condition

$$(4.5)\qquad (u,w,0,0)\in N_{\operatorname{epi}L}\,(t,x,v,\gamma) \implies (u,w)=(0,0).$$

Proof. Since the geometrical condition in (4.5) is identical to the one in (4.1), it suffices to construct a multifunction Γ satisfying conditions (a)–(b) of Proposition 4.1, such that any quadruple (r,x,v,γ) satisfying the three inequalities above automatically lies in $\operatorname{gph}\Gamma$. Intermediate figures in our definition of Γ are the 'essential value' multifunctions

$$V(t) := \left\{v \in \mathbb{R}^n \,:\, \forall\varepsilon>0,\ 0<m\left\{s\in[t-\varepsilon,t+\varepsilon]\cap\left[\overline{a},\overline{b}\right] \,:\, \left|v-\dot{\overline{x}}(s)\right|<\varepsilon\right\}\right\},$$
$$I(t) := \left\{\gamma \in \mathbb{R} \,:\, \forall\varepsilon>0,\ 0<m\left\{s\in[t-\varepsilon,t+\varepsilon]\cap\left[\overline{a},\overline{b}\right] \,:\, \left|\gamma-\overline{L}(s)\right|<\varepsilon\right\}\right\}.$$

(See [8].) Elementary measure theory implies that both $\operatorname{gph} V$ and $\operatorname{gph} I$ are compact, while $\dot{\overline{x}}(t) \in V(t)$ and $\overline{L}(t) \in I(t)$ for almost all t. It follows that for any constant $\delta_0 \in (0,\delta)$, the set

$$\begin{aligned} G := \{(r,x,v,\gamma) \,:\, & \text{for some } t \text{ in } \left[\overline{a},\overline{b}\right], \text{ one has} \\ & |(r,x)-(t,\overline{x}(t))| \le \delta_0, \\ & v \in V(t)+\delta_0\mathbb{B}, \\ & \gamma \in \operatorname{co} I(t)+\delta_0\mathbb{B}\} \end{aligned}$$

is compact, contains almost all the points $(t,\overline{x}(t),\dot{\overline{x}}(t),\overline{L}(t))$ for t in $\left[\overline{a},\overline{b}\right]$, and has convex sections in the last variable. Thus we can define Γ by $\operatorname{gph}\Gamma = G$ to obtain a multifunction satisfying conditions 4.1(a)–(b). For almost all t, any quadruple (r,x,v,γ) obeying the three inequalities in the theorem statement will satisfy $\gamma \in \Gamma(r,x,v)$. Thus condition (4.5) implies (4.1), and the result follows. □

The Convex Case. Necessary (and sufficient) conditions for the fixed-time case of problem (P) in which the Lagrangian is a convex function of (x,v) for each fixed t have been known for some time: see [19, 20, 21] for example. These early results have a role not only in the direct solution of applied problems, but also as test cases for the correctness of further generalizations. Thus we seek to confirm that our results in this paper represent faithful generalizations of the convex theory. The key issue, of course, is the extent to which convex problems can be expected to satisfy our standing hypotheses (H3)–(H4). Let us label our convexity condition, and incorporate an assumption that requires any state constraints to be introduced explicitly through the multifunction X, rather than implicitly through extended values in L:

(H2)$_c$ For each fixed t in $\left(\overline{a}-\rho,\overline{b}+\rho\right)$, the function $(x,v)\mapsto L(t,x,v)$ is convex on $\Omega_t\times\mathbb{R}^n$. Furthermore, one has $\operatorname{dom} L(t,x,\cdot)\neq\emptyset$ for each (t,x) in Ω_t.

We will show that for autonomous problems, both (H3) and (H4) follow from (H2)$_c$, and more generally, that assuming (H2)$_c$ allows (H3)–(H4) to be derived from their weakened analogues below, in which only the t-dependence is involved:

(H3)$_c$ The function L is lower semicontinuous on $\Omega\times\mathbb{R}^n$ and epicontinuous in t: that is, for any point (t,x,v) in $\Omega\times\mathbb{R}^n$ where $L(t,x,v)$ is finite, and any sequence $t_k\to t$, there exists a sequence $(x_k,v_k)\to(x,v)$ along which $L(t_k,x_k,v_k)\to L(t,x,v)$.

$(\mathrm{H4})_c$ Both $\dot{\overline{x}}$ and $\overline{L}$ are essentially bounded, and there are positive constants δ and κ such that for almost all t in $\left[\overline{a},\overline{b}\right]$, every point (r,x,v,γ) in $\Omega\times\mathbb{R}^n\times\mathbb{R}$ obeying the three inequalities

$$|(r,x)-(t,\overline{x}(t))|<\rho, \qquad \left|v-\dot{\overline{x}}(t)\right|<\delta, \qquad \left|\gamma-\overline{L}(t)\right|<\delta,$$

satisfies the normal inequality condition

$$(u,0,0,0)\in N_{\operatorname{epi}L}\,(r,x,v,\gamma) \implies u=0.$$

Both $(\mathrm{H3})_c$ and $(\mathrm{H4})_c$ hold trivially if L has no explicit dependence on t, provided both $\dot{\overline{x}}$ and $\overline{L}$ are essentially bounded. Notice that $(\mathrm{H4})_c$ is a geometrical sufficient condition for the uniform Aubin continuity of the multifunction $t \rightrightarrows \operatorname{epi} L(t,\cdot,\cdot)$ near the optimal trajectory.

We deal first with the epicontinuity conditions (H3) and $(\mathrm{H3})_c$, starting from a technical lemma.

4.3. Lemma. *Let $\overline{x}_0,\ldots,\overline{x}_n$ be points of $\mathbb{R}^n$ such that*

$$0\in\operatorname{int}\operatorname{co}\{\overline{x}_0,\ldots,\overline{x}_n\}. \tag{4.6}$$

For each $j=0,\ldots,n$, let $\left\{\overline{x}_j^k\right\}_k$ be a sequence converging to $\overline{x}_j$. Given any sequence $w_k\to 0$ in $\mathbb{R}^n$, there exists for each index k sufficiently large a collection $\lambda_0^k,\lambda_1^k,\ldots,\lambda_n^k\geq 0$ such that

$$\text{(i) } w_k=\sum_{j=0}^n \lambda_j^k\overline{x}_j^k, \qquad \text{(ii)} \left(\sum_{j=0}^n \lambda_j^k\right)\to 0 \text{ as } k\to\infty.$$

Proof. Condition (4.6) allows us to fix $\sigma>0$ so small that $\operatorname{co}\{\overline{x}_0,\ldots,\overline{x}_n\}$ contains $2\sigma\mathbb{B}$. It follows that for all k sufficiently large, the set $S^k:=\operatorname{co}\left\{\overline{x}_0^k,\ldots,\overline{x}_n^k\right\}$ contains the smaller ball $\sigma\mathbb{B}$. In particular, $\sigma w_k/|w_k|$ lies in S^k for all such k, and therefore has a representation in terms of scalars $\mu_j^k\geq 0$, $\sum_{j=0}^n \mu_j^k=1$:

$$\sigma\frac{w_k}{|w_k|}=\sum_{j=0}^n \mu_j^k\overline{x}_j^k, \text{ i.e., } w_k=\sum_{j=0}^n\left(\frac{\mu_j^k|w_k|}{\sigma}\right)\overline{x}_j^k.$$

Choosing $\lambda_j^k=\sigma^{-1}\mu_j^k|w_k|$ gives the representation asserted by conclusion (i), and also satisfies (ii) because

$$\sum_{j=0}^n\lambda_j^k=\sigma^{-1}|w_k|\sum_{j=0}^n\mu_j^k=|w_k|/\sigma\to 0 \text{ as } k\to\infty. \qquad \square$$

Here is an abstract result concerning epicontinuity, phrased in notation that will make its application to our problem particularly straightforward.

4.4. THEOREM. *Let* $L: \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ *be lower semicontinuous at every point in* $\{(\bar{t}, \bar{x})\} \times \mathbb{R}^n$*, and suppose that for some* $\eta > 0$*, the function* $(x, v) \mapsto L(t, x, v)$ *is convex for each fixed* t *obeying* $|t - \bar{t}| \leq \eta$*. Assume further that*

(i) $\operatorname{dom} L(\bar{t}, x, \cdot) \neq \emptyset$ *for all* x *where* $|x - \bar{x}| \leq \eta$*;*

(ii) L *is epicontinuous in* t *at the point* $t = \bar{t}$*.*

Then L *is epicontinuous in* (t, x) *at the point* $(\bar{t}, \bar{x})$*.*

Proof. Without loss of generality, take $\bar{t} = 0$ and $\bar{x} = 0$. Fix unit vectors $\widehat{u}_0, \ldots, \widehat{u}_n$ such that $0 \in \operatorname{int} \operatorname{co} \{\widehat{u}_0, \ldots, \widehat{u}_n\}$. Then let $\bar{x}_j = \eta \widehat{u}_j$ $(j = 0, \ldots, n)$: note that $\operatorname{dom} L(0, \bar{x}_j, \cdot) \neq \emptyset$ by (i), so there exists $\bar{v}_j$ where $L(0, \bar{x}_j, \bar{v}_j) < \infty$.

Now fix any v in $\operatorname{dom} L(0, 0, \cdot)$, and let any sequence $(t_k, x_k) \to (0, 0)$ be given. We must construct a sequence $v_k \to v$ along which $L(t_k, x_k, v_k) \to L(0, 0, v)$. To do this, we apply the epicontinuity property (ii) $n+2$ times: once to generate a sequence

$$(*) \qquad (x_k^*, v_k^*) \to (0, v) \text{ along which } L(t_k, x_k^*, v_k^*) \to L(0, 0, v),$$

and $n + 1$ more times to find for each $j = 0, \ldots, n$ a sequence

$$(**) \qquad (\bar{x}_j^k, \bar{v}_j^k) \to (\bar{x}_j, \bar{v}_j) \text{ along which } L(t_k, \bar{x}_j^k, \bar{v}_j^k) \to L(0, \bar{x}_j, \bar{v}_j),$$

Now apply Lemma 4.3 to the null sequence $w_k = x_k - x_k^*$, using the moving simplex $\{\bar{x}_j^k - x_k^* : j = 0, \ldots, n\}$. This produces sequences of scalars $\lambda_j^k \geq 0$ $(j = 0, \ldots, n)$ such that both

$$x_k - x_k^* = \sum_{j=0}^{n} \lambda_j^k \left(\bar{x}_j^k - x_k^*\right), \text{ i.e., } x_k = \sum_{j=0}^{n} \lambda_j^k \bar{x}_j^k + \left(1 - \sum_{j=0}^{n} \lambda_j^k\right) x_k^*,$$

and $\sum_{j=0}^{n} \lambda_j^k \to 0$ as $k \to \infty$. We use these scalars to define

$$v_k = \sum_{j=0}^{n} \lambda_j^k \bar{v}_j^k + \left(1 - \sum_{j=0}^{n} \lambda_j^k\right) v_k^*.$$

For each k, the convexity of the function $L(t_k, \cdot, \cdot)$ now yields

$$\begin{aligned} L(t_k, x_k, v_k) &= L\Big(t_k, \textstyle\sum_{j=0}^{n} \lambda_j^k (\bar{x}_j^k, \bar{v}_j^k) + \Big(1 - \sum_{j=0}^{n} \lambda_j^k\Big) (x_k^*, v_k^*)\Big) \\ &\leq \sum_{j=0}^{n} \lambda_j^k L(t_k, \bar{x}_j^k, \bar{v}_j^k) + \left(1 - \sum_{j=0}^{n} \lambda_j^k\right) L(t_k, x_k^*, v_k^*). \end{aligned}$$

On the right side of this estimate, each of the sequences $L(t_k, \overline{x}_j^k, \overline{v}_j^k)$ in the first term is bounded, by $(**)$. The sequence $L(t_k, x_k^*, v_k^*)$ in the second term converges to $L(0,0,v)$, by $(*)$. The construction of the coefficient sequences λ_j^k therefore implies that the right side above converges to $L(0,0,v)$. This gives the first inequality in the estimate

$$\begin{aligned} L(0,0,v) &\geq \limsup_{k\to\infty} L(t_k, x_k, v_k) \\ &\geq \liminf_{k\to\infty} L(t_k, x_k, v_k) \geq L(0,0,v). \end{aligned}$$

The second inequality is obvious, and the third follows from the lower semicontinuity of L. Taken together, they show that $L(t_k, x_k, v_k) \to L(0,0,v)$, as required. □

4.5. Corollary. *Together* $(\mathrm{H2})_c$ *and* $(\mathrm{H3})_c$ *imply* (H3).

Proof. It suffices to show that $(\mathrm{H2})_c$ and $(\mathrm{H3})_c$ imply conditions (i)–(ii) of Theorem 4.4 at every point $(\bar{t}, \overline{x})$ in Ω. This is obvious. □

The autonomous case of Theorem 4.4 is of independent interest.

4.6. Corollary. *Let* $g\colon \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ *be lower semicontinuous and jointly convex. Let* $(\overline{x}, \overline{v})$ *be a point where* g *is finite. Then the following statements are equivalent:*

(a) *There exists* $\eta > 0$ *such that* $\operatorname{dom} g(x,\cdot) \neq \emptyset$ *for all* x *with* $|x - \overline{x}| < \eta$.

(b) g *is epicontinuous in* x *at the point* $x = \overline{x}$.

Proof. If (a) holds, then Theorem 4.4 applies to $L(t,x,v) = g(x,v)$ to establish (b). Conversely, if (b) holds, then the definition of epicontinuity implies that for every sequence $x_k \to \overline{x}$, one has $\operatorname{dom} g(x_k,\cdot) \neq \emptyset$. This clearly implies (a). □

Now we turn from epicontinuity to the Aubin property.

4.7. Theorem. *Together* $(\mathrm{H2})_c$*–*$(\mathrm{H4})_c$ *imply* (H4).

Proof. We set up an application of Theorem 4.2. The central point is a simple sufficient condition for (4.5) available in the presence of convexity. Consider the indicator function $\psi(t,x,v,\gamma) := \Psi_{\operatorname{epi} L}(t,x,v,\gamma)$. It follows from $(\mathrm{H3})_c$ that the function ψ is both lower semicontinuous and epicontinuous in t. Meanwhile, $(\mathrm{H2})_c$ implies that for each fixed t, the function $(x,v,\gamma) \mapsto \psi(t,x,v,\gamma)$ is convex. Thus we have (Rockafellar [24, Prop. 2.2]) the inclusion

$$\begin{aligned} N_{\operatorname{epi} L}\,(t,x,v,\gamma) &= \partial\psi(t,x,v,\gamma) \\ &\subseteq \mathbb{R} \times \partial_{x,v,\gamma}\psi(t,x,v,\gamma) = \mathbb{R} \times N_{\operatorname{epi} L(t,\cdot,\cdot)}\,(x,v,\gamma)\,. \end{aligned}$$

It follows that any point $(u,w,0,0)$ normal to $\operatorname{epi} L$ at (t,x,v,γ) will have a projection $(w,0,0)$ normal to $\operatorname{epi} L(t,\cdot,\cdot)$ at (x,v,γ). But the latter relation concerns a closed convex set, for which normality has a simple characterization by inequalities:

$$0 \geq \langle (w,0,0)\,,\, (x',v',\gamma') - (x,v,\gamma) \rangle = \langle w\,,\, x' - x \rangle \qquad \forall (x',v',\gamma') \in \operatorname{epi} L(t,\cdot,\cdot).$$

Evidently any x' for which $\operatorname{dom} L(t, x', \cdot) \neq \emptyset$ must satisfy the inequality above; in view of $(H2)_c$, this includes all x' in the neighbourhood Ω_t of x. Thus $w = 0$. It follows that the key condition (4.5) of Theorem 4.2 can be rewritten as

$$\forall \gamma \in \Gamma(t, x, v), \qquad (u, 0, 0, 0) \in N_{\operatorname{epi} L}(t, x, v, \gamma) \implies u = 0.$$

This is precisely the condition supplied by $(H4)_c$. ◻

Lipschitzian Perturbations of Convex Problems. The various formulations of (H3) and (H4) described above are all designed to regulate non-Lipschitz behaviour in the (t, x)-dependence of the integrand L. Perhaps the easiest way to see this is to note that if (H3) and (H4) hold for a given integrand L, then they also hold for the integrand $L + G_1$, under the sole hypothesis that G_1 is Lipschitzian of constant rank on some neighbourhood of $\operatorname{gph}(\overline{x}, \dot{\overline{x}})$. It follows that the simplified conditions outlined in Corollary 4.5 and Theorem 4.7 apply not just to convex integrands, but equally well to Lipschitzian perturbations of convex integrands. An interesting family of such functions has the form $L(t, x, v) = G_1(t, x, v) + k(t, v - A(t)x)$, where G_1 is locally Lipschitzian on $\mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^n$, and $k(t, \cdot)$ is a lower semicontinuous, proper convex function for each fixed t. If we deal with a Lipschitzian arc $\overline{x}$ along which $\overline{L}$ is essentially bounded, conditions (H3)–(H4) for L are consequences of $(H3)_c$–$(H4)_c$ for the integrand $K(t, x, v) = k(t, v - A(t)x)$. Here Corollary 4.5 and Theorem 4.7 apply directly, under the additional assumption that the matrix-valued function A is Lipschitzian, to show that (H3)–(H4) hold whenever k is epi-continuous in t and one has the Aubin continuity condition

$$(u, 0, 0) \in N_{\operatorname{epi} k}(t, v - A(t)x, \gamma) \implies u = 0$$

on a suitable neighbourhood of the points $(t, \overline{x}(t), \dot{\overline{x}}(t), \overline{L}(t))$.

5. Proofs of the Main Results. The proofs of Theorems 1.4 and 3.1 take up this whole section; in particular, hypotheses (H1)–(H5) are in force throughout. The unifying idea is a classical one. We formulate an autonomous fixed-time problem for which a relative of $\overline{x}$ provides the solution, and then deduce the main result from the fixed-time theory of Loewen and Rockafellar [15].

The Erdmann Transform. There is a standard method for transforming problem (P) and its solution $\overline{x}$ on $[\overline{a}, \overline{b}]$ into a fixed-time problem solved by an arc related to $\overline{x}$. (See, for example, Clarke [4, Section 3.6].) We use Greek characters to describe states and costates in the latter problem, for which the underlying time interval is $[0, 1]$, and the state is a vector (θ, ξ) in $\mathbb{R} \times \mathbb{R}^n$. Here ξ is a transformed version of the original state x, while θ is a new state which keeps track of the variable time in problem (P).

Choose any constant m such that $0 < m < \overline{b} - \overline{a}$. The transformed problem is to

$$(\Pi) \qquad \begin{aligned} &\text{minimize} \quad l(\theta(0), \xi(0), \theta(1), \xi(1)) + \int_0^1 \widetilde{L}(\theta(\tau), \xi(\tau), \theta'(\tau), \xi'(\tau))\, d\tau \\ &\text{subject to} \quad (\theta(\tau), \xi(\tau)) \in \Xi(\tau) \cap \widetilde{\Omega}_\tau \ \forall \tau \in [0, 1]. \end{aligned}$$

Here the appropriate integrand $\widetilde{L}$, domain $\widetilde{\Omega}$, and state constraint multifunction Ξ are given by

$$\tag{5.1} \begin{aligned} \widetilde{L}(\theta,\xi,\theta',\xi') &= \begin{cases} \theta' L(\theta,\xi,\xi'/\theta'), & \text{if } \theta' \geq m, \\ +\infty, & \text{otherwise,} \end{cases} \\ \widetilde{\Omega} &:= [0,1]\times\Omega,\ \ \Xi(\tau) := \operatorname{gph} X. \end{aligned}$$

Comparing the objective functional in problem (Π) with that in problem (P) provides the key to understanding this reduction, as the following Lemma reveals.

5.1. Lemma. *The following arc solves the problem* (Π)*:*

$$\tag{5.2} \overline{\theta}(\tau) = \overline{a} + \tau(\overline{b}-\overline{a}), \qquad \overline{\xi}(\tau) = \overline{x}\left(\overline{\theta}(\tau)\right).$$

Proof. Consider any pair $(x,[a,b])$ admissible for (P), with $b-a \geq m$. Define $\theta(\tau) := a+\tau(b-a)$ and $\xi(\tau) = x(\theta(\tau))$. Notice that $(\theta(\tau),\xi(\tau)) \in \widetilde{\Omega}_\tau = \Omega$ and $(\theta(\tau),\xi(\tau)) \in \Xi(\tau)$ for all $\tau \in [0,1]$, so this pair satisfies the explicit constraints in (Π). The chain rule implies that on some subset of $[0,1]$ with full Lebesgue measure, $\xi'(\tau) = \dot{x}(\theta(\tau))\theta'(\tau)$. Thus the substitution $t = \theta(\tau)$ yields both

$$\int_{t=a}^{b} L(t,x(t),\dot{x}(t))\,dt = \int_{\tau=0}^{1} L(\theta(\tau),\xi(\tau),\xi'(\tau)/\theta'(\tau))\,\theta'(\tau)\,d\tau$$

and

$$(a,x(a),b,x(b)) = (\theta(0),\xi(0),\theta(1),\xi(1)).$$

It follows that the cost of $(x,[a,b])$ in problem (P) equals the cost of (θ,ξ) in problem (Π). If we apply these observations to the optimal pair $(\overline{x},[\overline{a},\overline{b}])$ in (P), we find that

$$\tag{5.3} \begin{aligned} \inf(\Pi) &\leq l(\overline{\theta}(0),\overline{\xi}(0),\overline{\theta}(1),\overline{\xi}(1)) + \int_0^1 \widetilde{L}(\overline{\theta}(\tau),\overline{\xi}(\tau),\overline{\theta}'(\tau),\overline{\xi}'(\tau))\,d\tau \\ &= l(\overline{a},\overline{x}(\overline{a}),\overline{b},\overline{x}(\overline{b})) + \int_{\overline{a}}^{\overline{b}} L(t,\overline{x}(t),\dot{\overline{x}}(t))\,dt = \inf\,(\mathrm{P}). \end{aligned}$$

Conversely, consider any arc (θ,ξ) admissible for (Π). One has $\theta'(\tau) \geq m$ for all $\tau \in [0,1]$, so the function θ is strictly increasing, hence invertible. Define $a = \theta(0)$, $b = \theta(1)$, and let $x(t) = \xi(\theta^{-1}(t))$ for $t \in [a,b]$. Evidently $x(t) \in X(t)\cap\Omega_t$ for all t in $[a,b]$, while the same substitution as before gives both

$$\int_{\tau=0}^{1} \widetilde{L}(\theta(\tau),\xi(\tau),\theta'(\tau),\xi'(\tau))\,d\tau = \int_{t=a}^{b} L(t,x(t),\dot{x}(t))\,dt.$$

and

$$(\theta(0),\xi(0),\theta(1),\xi(1)) = (a,x(a),b,x(b)).$$

Thus the pair $(x, [a, b])$ is admissible for (P), where it is assigned the same objective value as the pair (θ, ξ) receives in (Π). In particular, $\inf(\mathrm{P}) \le \inf(\Pi)$. Thus equality holds throughout in (5.3) above, and the lemma follows. ∎

Verification of Hypotheses. The fixed-time problem (Π), with its solution $(\overline{\theta}, \overline{\xi})$, has the form to which the necessary conditions of Loewen and Rockafellar [15] may be applied. The first step is to confirm the hypotheses. It is evident that the endpoint cost in problem (Π) is lower semicontinuous, and that the (autonomous) integrand $\widetilde{L}$ is lower semicontinuous, hence Borel measurable. Also, for each fixed (θ, ξ), the mapping $(\theta', \xi') \mapsto \widetilde{L}(\theta, \xi, \theta', \xi')$ is convex on the set $(0, \infty) \times \mathbb{R}^n$—the proof is a standard exercise in convex analysis. Finally, the state constraint multifunction $\tau \rightrightarrows \operatorname{gph} X$ is constant with a closed image, so the lower semicontinuity condition of [15, (H6)] is obvious. Only two of the requirements of [15] remain to check.

First, we must verify the epi-continuity of $\widetilde{L}$. To do this, let τ in $[0, 1]$, and fix a point $(\widehat{\theta}, \widehat{\xi}, \widehat{\theta}', \widehat{\xi}')$ in $\Omega \times \mathbb{R}^{1+n}$ where $\widetilde{L}$ is finite and $\left|(\widehat{\theta}, \widehat{\xi}) - (\overline{\theta}(\tau), \overline{\xi}(\tau))\right| < \rho$. Consider any sequence $(\theta_k, \xi_k) \to (\widehat{\theta}, \widehat{\xi})$. Then the finiteness of $\widetilde{L}$ implies that $\widehat{\theta}' > m$, so $L(\widehat{\theta}, \widehat{\xi}, \widehat{\xi}'/\widehat{\theta}')$ is finite; also $\left|\left(\widehat{\theta}, \widehat{\xi}\right) - (t, \overline{x}(t))\right| < \rho$ for $t = \overline{\theta}(\tau)$. Taking $\widehat{v} = \widehat{\xi}'/\widehat{\theta}'$, we recognize this as a situation to which (H3) applies: that hypothesis provides a sequence $v_k \to \widehat{v}$ along which $L(\theta_k, \xi_k, v_k) \to L(\widehat{\theta}, \widehat{\xi}, \widehat{v})$. Upon defining $\widehat{\theta}'_k = \widehat{\theta}'$ and $\xi'_k = \widehat{\theta}' v_k$, we deduce that $\widetilde{L}(\theta_k, \xi_k, \theta'_k, \xi'_k) \to \widetilde{L}(\widehat{\theta}, \widehat{\xi}, \widehat{\theta}', \widehat{\xi}')$, as required.

Second, we must verify the differential inequality in [15, (H5)]. This requires us to produce nonnegative functions $\widetilde{\delta}$ and $\widetilde{\kappa}$ in $L^1[0, 1]$, with $\widetilde{\kappa}/\widetilde{\delta}$ in $L^\infty[0, 1]$, such that for almost all τ in $[0, 1]$, the three conditions

(i) $\left|(\theta, \xi) - (\overline{\theta}(\tau), \overline{\xi}(\tau))\right| < \rho$,

(ii) $\left|(\theta', \xi') - (\overline{\theta}'(\tau), \overline{\xi}'(\tau))\right| < \widetilde{\delta}(\tau)$,

(iii) $\left|\widetilde{L}(\theta, \xi, \theta', \xi') - \widetilde{L}(\overline{\theta}(\tau), \overline{\xi}(\tau), \overline{\theta}'(\tau), \overline{\xi}'(\tau))\right| < \widetilde{\delta}(\tau)$,

imply the subgradient inequality

$$|(u, w)| \le \widetilde{\kappa}(\tau)[1 + |p| + |q|] \qquad \forall (u, w, q, p) \in \partial \widetilde{L}(\theta, \xi, \theta', \xi'). \tag{5.4}$$

Recalling that

$$\widetilde{L}(\theta, \xi, \theta', \xi') = \theta' L(\theta, \xi, \xi'/\theta') + \Psi_{[m, +\infty)}(\theta'),$$

we can estimate the subgradient set appearing in (5.4) using the calculus rules of Rockafellar [23, Cor. 7.1.2]: for evaluation points where $\theta' > m$, this gives

$$\partial \widetilde{L}(\theta, \xi, \theta', \xi') \subseteq \Big\{(u\theta', w\theta', -H(\theta, \xi, p), p) \,:\, (u, w, p) \in \partial L(\theta, \xi, \xi'/\theta')\Big\}. \tag{5.5}$$

(A direct application of the cited chain rule produces a third component of the form $L(\theta, \xi, \xi'/\theta') - \langle p\,, \xi'/\theta'\rangle$ on the right side. However, the inclusion $(u, w, p) \in$

$\partial L(\theta,\xi,\xi'/\theta')$ implies that this expression equals $-H(\theta,\xi,p)$, as explained in Remark 1.6.)

It follows that a sufficient condition for (5.4) (at least when $\theta' > m$) is the inequality

$$|(u,w)| \le \frac{\widetilde{\kappa}(\tau)}{\theta'}[1+|p|+|H(\theta,\xi,p)|] \quad \forall (u,w,p)\in\partial L(\theta,\xi,\xi'/\theta'). \tag{5.6}$$

Let us prove that under our assumption (H4), conditions (i)–(iii) imply both $\theta' > m$ and (5.6), using the constant functions

$$\widetilde{\kappa} = \left(\overline{\theta}'(\tau)+m\right)\kappa, \qquad \widetilde{\delta} = \min\left\{m\delta,\, \overline{\theta}'(\tau)-m,\, m\right\}. \tag{5.7}$$

(Recall that $\overline{\theta}'(\tau) = \overline{b}-\overline{a}$ is constant, so this proof will establish [15, (H5)].)

Indeed, fix τ in $[0,1]$ and a point $(\theta,\xi,\theta',\xi')$ obeying (i)–(iii). Inequality (ii) and our choice of $\widetilde{\delta}$ imply

$$m \le \overline{\theta}'(\tau)-\widetilde{\delta} < \theta' < \overline{\theta}'(\tau)+\widetilde{\delta} < \overline{\theta}'(\tau)+m, \tag{5.8}$$

so we have $\theta' > m$ as required. Next, using the triangle inequality with condition (ii), we have

$$\begin{aligned}\left|\frac{\xi'}{\theta'}-\frac{\overline{\xi}'(\tau)}{\overline{\theta}'(\tau)}\right| &= \left|\frac{\xi'-\overline{\xi}'(\tau)}{\theta'}+\overline{\xi}'(\tau)\left(\frac{\overline{\theta}'(\tau)-\theta'}{\overline{\theta}'(\tau)\theta'}\right)\right| \\ &\le \frac{1}{\theta'}\left|\xi'-\overline{\xi}'(\tau)\right| + \left|\frac{\overline{\xi}'(\tau)}{\overline{\theta}'(\tau)}\right|\frac{\left|\overline{\theta}'(\tau)-\theta'\right|}{\theta'} \\ &\le \frac{1}{m}\left[1+\left|\frac{\overline{\xi}'(\tau)}{\overline{\theta}'(\tau)}\right|\right]\widetilde{\delta}.\end{aligned} \tag{5.9}$$

Likewise, consider condition (iii). Expanding the definition of $\widetilde{L}$, and using the shorthand $L = L(\theta,\xi,\xi'/\theta')$ and $\overline{L} = L(\overline{\theta}(\tau),\overline{\xi}(\tau),\overline{\xi}'(\tau)/\overline{\theta}'(\tau))$, we derive

$$\begin{aligned}\left|\theta' L-\overline{\theta}'\overline{L}\right| &\ge \theta'\left|L-\overline{L}\right| - \left|\left(\overline{\theta}'(\tau)-\theta'\right)\overline{L}\right| \\ &\ge m\left|L-\overline{L}\right| - \left|\theta'-\overline{\theta}'(\tau)\right|\left|\overline{L}\right|.\end{aligned}$$

This implies

$$\left|L-\overline{L}\right| \le \frac{\widetilde{\delta}}{m}\left[1+\left|\overline{L}\right|\right]. \tag{5.10}$$

Now for $\widetilde{\kappa}$ and $\widetilde{\delta}$ as shown in (5.7) above, we have $\widetilde{\delta}/m \le \delta$. Thus, with the change of variable $t = \overline{\theta}(\tau)$, under which

$$\tau = \overline{\theta}^{-1}(t), \quad \overline{\xi}(\tau) = \overline{x}(t), \quad \overline{\xi}'(\tau)/\overline{\theta}'(\tau) = \dot{\overline{x}}(t), \tag{5.11}$$

and the parallel change of notation

$$r = \theta, \quad \xi = x, \quad \xi'/\theta' = v, \tag{5.12}$$

condition (i) states

(i′) $|(r,x) - (t, \overline{x}(t))| < \rho.$

Meanwhile, condition (ii) implies, through (5.9), that

(ii′) $\left|v - \dot{\overline{x}}(t)\right| \le \delta\left[1 + \left|\dot{\overline{x}}(t)\right|\right].$

Finally, estimate (5.10)—a consequence of (iii)—becomes

(iii′) $\left|L(r,x,v) - L(t, \overline{x}(t), \dot{\overline{x}}(t))\right| \le \delta\left[1 + \left|L(t, \overline{x}(t), \dot{\overline{x}}(t))\right|\right].$

These are precisely the conditions under which (H4) implies

$$|(u,w)| \le \kappa[1 + |p| + |H(r,x,p)|] \qquad \forall (u,w,p) \in \partial L(r,x,v).$$

Our choice of $\widetilde{\kappa}$ and inequality (5.8) imply that $\kappa = \widetilde{\kappa}/\Big(\overline{\theta}'(\tau) + m\Big) < \widetilde{\kappa}/\theta'$ for all points of interest, so inequality (5.6) follows. This completes the verification of [15, (H5)].

Proof of Theorem 1.4. Having confirmed the hypotheses of [15], we may now use its conclusions. In the case of Theorem 1.4, where we have $\big(\overline{\theta}(\tau), \overline{\xi}(\tau)\big) \in \operatorname{int} \Xi(\tau)$ for all τ, these are expressed most clearly in [15, Theorem 2.1]. They assert the existence of an absolutely continuous function $(\eta, \pi)\colon [0,1] \to \mathbb{R} \times \mathbb{R}^n$ satisfying adjoint equations of either normal or singular type.

In the normal conditions, we have the transversality relation

$$(\eta(0), \pi(0), -\eta(1), -\pi(1)) \in \partial l\big(\overline{\theta}(0), \overline{\xi}(0), \overline{\theta}(1), \overline{\xi}(1)\big) \tag{5.13}$$

and the Euler-Lagrange inclusion

$$(\eta'(\tau), \pi'(\tau)) \in \operatorname{co}\Big\{(\alpha, \beta) \,:\, (\alpha, \beta, \eta(\tau), \pi(\tau)) \in \partial \widetilde{L}(\overline{\theta}(\tau), \overline{\xi}(\tau), \overline{\theta}'(\tau), \overline{\xi}'(\tau))\Big\} \tag{5.14}$$

for almost all τ in $[0,1]$. The subgradient estimate (5.5) reveals that the inclusion $(\alpha, \beta, \eta(\tau), \pi(\tau)) \in \partial \widetilde{L}(\overline{\theta}(\tau), \overline{\xi}(\tau), \overline{\theta}'(\tau), \overline{\xi}'(\tau))$ in the latter condition implies that for some (u,w,p) in $\partial L(\overline{\theta}(\tau), \overline{\xi}(\tau), \overline{\xi}'(\tau)/\overline{\theta}'(\tau))$, one has $\alpha = u\overline{\theta}'(\tau)$, $\beta = w\overline{\theta}'(\tau)$, $\eta(\tau) = -H\big(\overline{\theta}(\tau), \overline{\xi}(\tau), p\big)$, and $\pi(\tau) = p$. Consequently the Euler-Lagrange inclusion implies

$$\begin{aligned}(\eta'(\tau), \pi'(\tau)) \in \operatorname{co}\Big\{\Big(u\overline{\theta}'(\tau), w\overline{\theta}'(\tau)\Big) \,:\, &(u, w, \pi(\tau)) \in \partial L\Big(\overline{\theta}(\tau), \overline{\xi}(\tau), \overline{\xi}'(\tau)/\overline{\theta}'(\tau)\Big), \\ &\eta(\tau) = -H\big(\overline{\theta}(\tau), \overline{\xi}(\tau), \pi(\tau)\big)\Big\}.\end{aligned}$$

In particular, the definitions $h(t) = -\eta\Big(\overline{\theta}^{-1}(t)\Big)$ and $p(t) = \pi\Big(\overline{\theta}^{-1}(t)\Big)$ produce absolutely continuous functions satisfying, for almost all t in $[\overline{a}, \overline{b}]$,

$$\Big(-\dot{h}(t), \dot{p}(t)\Big) \in \operatorname{co}\{(u,w) : (u,w,p(t)) \in \partial L(t, \overline{x}(t), \dot{\overline{x}}(t)),$$
$$h(t) = H(t, \overline{x}(t), p(t))\}.$$

The second condition defining the set on the right side places no restrictions on the points (u, w) in this set, but it does serve to make the set empty at every time t where the condition fails. Thus the indicated inclusion can be split apart to give the Euler-Lagrange inclusion stated in Theorem 1.4(a) and the identity

$$h(t) = H(t, \overline{x}(t), p(t)) \text{ a.e. } t \in \big[\overline{a}, \overline{b}\big]\,. \tag{5.15}$$

The Hamiltonian form of the set on the right in Theorem 1.4(a) is provided by Rockafellar [24]. Conclusion (b) of Theorem 1.4 follows from (5.15), as explained in Remark 1.6.

In the singular case, the transversality relation states

$$(\eta(0), \pi(0), -\eta(1), -\pi(1)) \in \partial^\infty l\big(\overline{\theta}(0), \overline{\xi}(0), \overline{\theta}(1), \overline{\xi}(1)\big) \tag{5.16}$$

and the Euler-Lagrange inclusion is replaced by

$$(\eta'(\tau), \pi'(\tau)) \in \operatorname{co}\Big\{(\alpha, \beta) \,:\, (\alpha, \beta, \eta(\tau), \pi(\tau)) \in \partial^\infty \widetilde{L}(\overline{\theta}(\tau), \overline{\xi}(\tau), \overline{\theta}'(\tau), \overline{\xi}'(\tau))\Big\}$$

for almost all τ in $[0, 1]$. The analysis of this statement parallels the developments in the normal case line by line, starting with an estimate of the singular subgradients of $\widetilde{L}$ again furnished by [23, Cor. 7.1.2]:

$$\partial^\infty \widetilde{L}(\theta, \xi, \theta', \xi') \subseteq \Big\{(u\theta', w\theta', -\langle p\,,\, \xi'/\theta'\rangle\,, p) \,:\, (u, w, p) \in \partial^\infty L(\theta, \xi, \xi'/\theta')\Big\}. \tag{5.17}$$

As before, the subgradient estimate leads to the conclusion that $h(t) = -\eta\Big(\overline{\theta}^{-1}(t)\Big)$ and $p(t) = \pi\Big(\overline{\theta}^{-1}(t)\Big)$ are arcs satisfying, for almost all t in $[\overline{a}, \overline{b}]$,

$$\Big(-\dot{h}(t), \dot{p}(t)\Big) \in \operatorname{co}\{(u,w) : (u,w,p(t)) \in \partial^\infty L(t, \overline{x}(t), \dot{\overline{x}}(t)),$$
$$h(t) = \big\langle p(t)\,,\, \dot{\overline{x}}(t)\big\rangle\, \}.$$

The latter inclusion implies both the Euler-Lagrange inclusion stated as conclusion (a^∞) of Theorem 1.4 and the identity

$$h(t) = \big\langle p(t)\,,\, \dot{\overline{x}}(t)\big\rangle \text{ a.e. } t \in \big[\overline{a}, \overline{b}\big] \tag{5.18}$$

of conclusion (b^∞).

Nontriviality in Theorem 1.4. A direct application of [15, Theorem 2.1] yields an apparently weaker nontriviality assertion than the one made in Theorem 1.4, in that the singular conditions of [15] refer to an adjoint arc that is *not the zero function,* whereas those in the present paper assert that it is a function that *never takes the value zero.* Under our hypotheses, these two properties are actually equivalent—although the second, being more explicit, is clearly preferable. We justify this claim in the context of our earlier paper [15], so as to sharpen the results in that work at the same time that we note their consequences for our current investigation. No generality is lost in working on the fixed time interval $[0,1]$.

Assume [15, (H5)], which provides two positive-valued functions $\kappa, \delta \in L^1[0,1]$ such that $\kappa/\delta \in L^\infty[0,1]$ and the three inequalities

$$|x-\overline{x}(t)|<\rho, \quad \left|v-\dot{\overline{x}}(t)\right|<\delta(t), \quad \left|L(t,x,v)-L\big(t,\overline{x}(t),\dot{\overline{x}}(t)\big)\right|<\delta(t) \tag{5.19}$$

imply the subgradient inequality

$$|w| \le \kappa(t)[1+|p|] \quad \forall (w,p)\in\partial L(t,x,v). \tag{5.20}$$

(Here the subgradient is taken only with respect to the pair (x,v).) We claim that the same three inequalities also imply the *singular* subgradient estimate

$$|w| \le \kappa(t)|p| \quad \forall (w,p)\in\partial^\infty L(t,x,v). \tag{5.21}$$

Indeed, fix (t,x,v) according to (5.19) and $(w,p)\in\partial^\infty L(t,x,v)$. By definition, there are sequences $s_k \to 0^+$, $(x_k,v_k)\xrightarrow[L(t,\cdot,\cdot)]{}(x,v)$, and $(w_k,p_k)\in\partial L(t,x_k,v_k)$ such that $(w,p)=\lim_{k\to\infty} s_k(w_k,p_k)$. Since the three inequalities concerning the chosen point (x,v) in (5.19) are strict, they are valid also for all the points (x_k,v_k) with k sufficiently large. Thus (5.20) gives an estimate relating w_k to p_k. Multiplying this by s_k yields

$$|s_k w_k| \le \kappa(t)[s_k+|s_k p_k|]\,.$$

Now taking the limit as $k\to\infty$ produces (5.21), as required.

Now suppose the arc p satisfies the singular Euler-Lagrange inclusion of [15, Theorem 2.1], i.e.,

$$\dot{p}(t)\in \operatorname{co}\left\{w \,:\, (w,p(t))\in\partial^\infty L\big(t,\overline{x}(t),\dot{\overline{x}}(t)\big)\right\} \text{ a.e. } t\in[0,1].$$

Then for almost all t, one has scalars $\lambda_j\ge 0$, $\sum\lambda_j=1$, such that

$$\dot{p}(t)=\sum\lambda_j w_j, \qquad (w_j,p(t))\in\partial^\infty L\big(t,\overline{x}(t),\dot{\overline{x}}(t)\big)\,.$$

From (5.21) we get

$$|\dot{p}(t)| \le \sum\lambda_j|w_j| \le \sum\lambda_j[\kappa(t)|p(t)|] = \kappa(t)|p(t)|. \tag{5.22}$$

Under (5.22), the statements, "$p(t) \neq 0$ for *all* $t \in [0,1]$" and "$p(t) \neq 0$ for *some* $t \in [0,1]$" are known to be equivalent, in consequence of Gronwall's lemma.

Proof of Theorem 3.1. Theorem 3.1, like Theorem 1.4, follows from a corresponding result for fixed-time problems through a careful application of the Erdmann transformation. The reductions described above apply without change to the state-constrained case, until we reach the subsection headed, "Known Necessary Conditions." At this point it is necessary to appeal to [15, Thm. 6.1] instead of to [15, Thm. 2.1], but then the adjoint inclusions, maximum condition, and transversality condition follow from the corresponding statements in [15] very much as described above. The additional complication in the state-constrained case concerns the management of the jump direction cone. The main thing to notice in this regard is that [15, (6.1)] introduces special notation for this cone, which, when applied to the data of problem (Π), amounts to

$$\overline{N}_\Xi(\tau,\theta,\xi) = \operatorname{cl}\operatorname{co}\Big\{\nu \in \mathbb{R}^n \;: \nu = \lim_{k\to\infty} \nu_k \text{ for some sequences } \nu_k \in \hat{\hat{N}}_{\Xi(\tau_k)}(\theta_k,\xi_k),\ (\tau_k,\theta_k,\xi_k) \xrightarrow[\operatorname{gph}\Xi]{} (\tau,\theta,\xi)\Big\}.$$

In the current context, the multifunction Ξ is constant, with the value $\operatorname{gph} X$. Thus

$$\overline{N}_\Xi(\tau,\theta,\xi) = \operatorname{cl}\operatorname{co}\Big\{\nu \in \mathbb{R}^n \;: \nu = \lim_{k\to\infty} \nu_k \text{ for some sequences } \nu_k \in \hat{\hat{N}}_{\operatorname{gph} X}(\theta_k,\xi_k),\ (\theta_k,\xi_k) \xrightarrow[\operatorname{gph} X]{} (\theta,\xi)\Big\}.$$

The latter expression reveals $\overline{N}_\Xi(\tau,\theta,\xi)$ as a closed convex cone independent of τ, formed by taking the closed convex hull of all limiting proximal normals associated with the closed set $\operatorname{gph} X$ at the point (θ,ξ). This is precisely the Clarke normal cone—see [4, Prop. 2.5.7]. Treating functions η and π of bounded variation instead of their absolutely continuous namesakes is essentially the same as what we have done above, and the linear change of variables relating t and τ allows the statements about the support of the indicated measures to be drawn immediately from the corresponding assertions of [15, Thm. 6.1].

Nontriviality in Theorem 3.1. It seems difficult to come up with a pointwise-in-t condition of nontriviality comparable to the one asserted in the state-constraint-free context of Theorem 1.4. This is especially true in problems where the state constraint is active at one of the optimal endtimes $\overline{a}$ or $\overline{b}$: here the recent work of Arutyunov and Aseev [2] is highly pertinent.

REFERENCES

[1] Ambrosio, L., O. Ascenzi, and G. Buttazzo, *Lipschitz regularity for the minimizers of integral functionals with highly discontinuous integrands,* JMAA **142**(1985), 301–316.

[2] Arutyunov, A. V., and S. M. Aseev, *Investigation of the degeneracy phenomenon of the maximum principle for optimal control problems with state constraints,* preprint, 1994.

[3] Aubin, Jean-Pierre, *Lipschitz behavior of solutions to convex minimization problems,* Math. Oper. Res. **9**(1984), 87–111.

[4] Clarke, F. H., *Optimization and Nonsmooth Analysis.* New York: John Wiley, 1983.

[5] ———, *Hamiltonian analysis of the Generalized Problem of Bolza,* Trans. Amer. Math. Soc. **301**(1987), 385–400.

[6] ———, *Methods of Dynamic and Nonsmooth Optimization,* CBMS-NSF Regional Conference Series, vol. 57. Philadelphia: SIAM, 1989.

[7] Clarke, F. H., and Philip D. Loewen, *An intermediate existence theory in the calculus of variations,* Ann. Sc. Norm. Sup. Pisa (Fis. e Mat.) **16**(1989), 487–526.

[8] Clarke, F. H., P. D. Loewen, and R. B. Vinter, *Differential inclusions with free time.* Annales de l'Institut Henri Poincaré—Analyse non linéaire **5**(1988), 573–593.

[9] Clarke, F. H., and R. B. Vinter, *Regularity properties of solutions ot the basic problem in the calculus of variations*, Trans. Amer. Math. Soc. **289**(1985), 73–98.

[10] ———, *Existence and regularity in the small in the calculus of variations,* J. Differential Equations **59**(1985), 336–354.

[11] Ioffe, A. D., and R. T. Rockafellar, *The Euler and Weierstrass conditions for nonsmooth variational problems*, Calculus of Variations and Partial Differential Equations, to appear.

[12] Loewen, Philip D., *Optimal Control via Nonsmooth Analysis,* CRM-AMS Lecture Notes in Mathematics **2**(1993).

[13] Loewen, Philip D., and R. T. Rockafellar, *The adjoint arc in nonsmooth optimization*, Trans. Amer. Math. Soc. **325**(1991), 39–72.

[14] ———, *Optimal control of unbounded differential inclusions.* SIAM J. Control and Optimization, **32**(1994), 442–470.

[15] ———, *New necessary conditions for the Generalized Problem of Bolza,* SIAM J. Control and Optimization, to appear.

[16] Mordukhovich, Boris S. *Complete characterization of openness, metric regularity, and Lipschitzian properties of multifunctions*, Trans. Amer. Math. Soc. **340**(1993), 1–35.

[17] Mordukhovich, Boris S., *Generalized differential calculus for nonsmooth and set-valued mappings*, JMAA **183**(1994), 250–288.

[18] ———, *Discrete approximations and refined Euler-Lagrange conditions for nonconvex differential inclusions*, SIAM J. Control Optim. **33**(1995), 882–915.

[19] Rockafellar, R. T., *Conjugate convex functions in optimal control and the calculus of variations*, J. Math. Analysis Appl., 23(1970), pp. 174–222.

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

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

[22] ———, *Dual problems of Lagrange for arcs of bounded variation*, in Calculus of Variations and Control Theory, D. L. Russell, ed., New York: Academic Press, 1976, pp. 155–192.

[23] ———, *Extensions of subgradient calculus with applications to optimization*, Nonlinear Analysis, Theory Methods & Applications, **9**(1985), 665–698.

[24] ———, *Equivalent subgradient versions of Hamiltonian and Euler-Lagrange equations in variational analysis,* preprint, 1994.

[25] Rockafellar, R. T., and R. J-B Wets, *Variational Analysis,* Book in preparation.

[26] Vinter, R. B., and H. Zheng, *The extended Euler-Lagrange condition for nonconvex variational problems*, SIAM J. Control Optim., to appear.