NECESSARY CONDITIONS FOR OPTIMAL IMPULSIVE CONTROL PROBLEMS

G. N. SILVA† AND R. B. VINTER‡

Abstract. Necessary conditions of optimality, in the form of a maximum principle, are derived for a class of optimal control problems, certain of whose controls are represented by measures and whose state trajectories are functions of bounded variation. State trajectories are interpreted as robust solutions of the dynamic equations, a concept of solutions which takes account of the interaction between the measure control and the state variables during the jumps. The maximum principle which is derived improves on earlier optimality conditions for problems of this nature, by allowing nonsmooth data, measurable time dependence, and a possibly time-varying constraint set for the conventional controls.

Key words. impulsive control, necessary conditions, nonsmooth analysis, maximum principle

AMS subject classifications. 65F10, 65F15

PII. S0363-0129(95)28185-7

1. Introduction. In this paper, optimal control problems in which certain control variables are represented by measures (“impulsive” control variables) and the state trajectories are functions of bounded variation are studied. Necessary conditions of optimality in the form of a maximum principle are derived for these problems. Specifically, we consider the following problem:

\[
(P) \begin{align*}
\text{Minimize } & \ h(x(0), x(1)) \\
\text{subject to } & \ dx(t) = f(t, x(t), u(t)) \, dt + g(t, x(t))\,\mu(dt), \ t \in [0, 1], \\
& \ (x(0), x(1)) \in C, \\
& \ u(t) \in U_t, \ \mathcal{L} - \text{a.e. } t \in [0, 1], \ \text{and } \mu \geq 0.
\end{align*}
\]

Here \( h : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}, f : [0, 1] \times \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n, \) and \( g : [0, 1] \times \mathbb{R}^n \to \mathbb{R}^n \) are given functions. \( U \) is a Borel subset of \([0, 1] \times \mathbb{R}^m\) (\( U_t \) denotes the “section” \( \{x : (t, x) \in U\}\)), and \( C \) is a closed subset of \( \mathbb{R}^n \times \mathbb{R}^n \).

A control policy is taken to be a pair of elements \((u : [0, 1] \to \mathbb{R}, \mu)\), in which the “conventional control” component \( u \) is a Lebesgue measurable function satisfying \( u(t) \in U_t \) a.e. with respect to Lebesgue measure and the “impulsive” control \( \mu \) is a regular, Borel, nonnegative valued measure. A process is a triple \((x, u, \mu)\) and a corresponding state trajectory \( x \); that is, \( x \) is a function of bounded variation which is a solution (appropriately defined) of the dynamical equation of problem (P). The control problem is to minimize the cost function \( h(x(0), x(1)) \) over processes \((x, u, \mu)\) for which \((x(0), x(1)) \in C \).

We particularly stress that, in our formulation, the coefficient \( g(t, x) \) associated with the impulsive control is allowed to be \( x \)-dependent. This raises at the outset...
questions of how we should interpret state trajectories to take account of the inter-
action between the evolving state trajectory and the impulsive control at times when
jumps occur.

It is natural to regard the state equation as shorthand for the integral equation
\begin{equation}
(1.1) \quad x(t) = x(0) + \int_0^t f(\tau, x(\tau), u(\tau))d\tau + \int_{[0,t]} g(\tau, x(\tau))\mu(d\tau) \quad \forall \ t \in (0,1].
\end{equation}

The second term on the right is affected by the value of the integral at an atom \( \tau \)
of the impulsive control \( \mu \). But here an ambiguity arises, because \( x \) can be expected
to jump at \( \tau \) and it is not obvious how we should interpret \( g(t, x(t)) \): a left limit,
\( g(t, x(t^-)) \) as in [23], some averaged value, or what?

Our approach to defining state trajectories is in the spirit of recent work by
Bressan, Dal Maso, and Rampazzo [1, 2, 6] and Miller [13], which in turn has
its origins in the reparameterization techniques of Rishel [16] and refinements due to
Warga [22]. The control problem studied here arises in the calculation of optimal flight
trajectories, “midcourse guidance problems” [8], [10], in which context an impulsive
control \( \mu \) is the idealization of a “conventional” control which, at the “atoms” of \( \mu \),
takes large values over a small interval of time. To be consistent with this view of
an impulsive control, we need to be sure that the state trajectory corresponding to
the idealized \( \mu \) approximates the conventional control with which it is associated. We
might require, say,
\[ x_i(t) \to x(t) \quad \forall \ t \in C_\mu \cup \{0,1\}, \]
where \( C_\mu \) denotes the points in \([0,1]\) which are not atoms of \( \mu \) (the “continuity points”
of \( \mu \)). In this relationship, \( \{x_i(t)\} \) is a sequence of state trajectories, with initial value
\( x(0) \), corresponding to a sequence of conventional controls \( \{\mu_i(t)\} \subset L^1 \) (think of \( \mu_i \)
as defining a Borel measure \( \mu_i(t)dt \), with \( \mu_i(t) \geq 0 \) a.e. for each \( i \), for which
\[ \mu_i(t)dt \to \mu(dt) \quad \text{weak*.} \]

As shown by Dal Maso and Rampazzo [6] there is a concept of solutions with this
continuous dependence property; we call them robust solutions to the state equation.
A robust solution \( x \) is a solution of the integral equation
\begin{equation}
(1.2) \quad x(t) = x(0) + \int_0^t f(\tau, x(\tau), u(\tau))d\tau + \int_{[0,t]} \tilde{g}(\tau, x(\tau^-); \mu(\{t\}))\mu(d\tau)
\end{equation}
\begin{equation}
\quad \text{for} \quad t \in (0,1],
\end{equation}
(in which \( x(\tau^-) \) denotes limit from the left). Here one takes account of interaction be-
 tween the state trajectory and the measure by choosing the integrand \( \tilde{g}(\tau, x(\tau^-); \mu(\{t\})) \)
of the impulsive term on the right to depend, at time \( \tau \), on \( \mu(\{t\}) \). \( \tilde{g}(\tau, x(\tau^-); \mu(\{t\})) \)
is determined by motion along integral curves of the impulsive dynamics, from the
initial state \( x(\tau^-) \). The magnitude of \( \mu(\{t\}) \) governs how far we move along an inte-
gral curve. If \( \tau \) is not an atom of \( \mu \), then \( \mu(\{t\}) = 0 \) and there is no motion. In this
case \( \tilde{g}(\tau, x(\tau^-); \mu(\{t\})) = g(\tau, x(\tau)) \).

The central result in this paper is a necessary condition, in the form of a maximum
principle, governing minimizers for (P) over control policies and corresponding state
trajectories, when the latter are interpreted as robust solutions of the dynamic equa-
tions. Our derivation of the condition covers problems involving a general, time-de-
dependent control constraint set \( U_t \) associated with the conventional control and
having dynamics which are nonsmooth in the state variable and measurably time dependent.

There is a substantial literature on necessary conditions of optimality for problem \( P \), relating to robust solutions of the dynamical equations. (See [11] and the survey article, with extensive bibliography, provided by Miller [12].) Earlier necessary conditions have been proven under the assumptions that the control constraint set is time independent and that the dynamics are at least Lipschitz continuous in the time variable. Indeed these conditions do not even make sense when the data are merely measurably dependent on the time variable, because they involve time derivatives of the data. Earlier approaches have been to reparameterize the independent variable in such a manner that the measure-driven dynamic equation reduces to a conventional equation, and \( P \) is transformed into a standard optimal control problem. Necessary conditions for \( P \) follow directly from the maximum principle (as conventionally understood) applied to the transformed problem. The transformation renders the original time variable a state variable; the extra restrictions arise then, because the conventional maximum principle does not allow for a state dependent control constraint set or data measurable in the state variable.

The proof given here follows a different route, in which we approximate \( P \) by a conventional problem with the help of Ekeland's theorem and pass to the limit. No reparameterization of the independent variable is involved in this approximation itself (this would appear to be crucial for problems with data measurably dependent on the time variable and with a time-dependent control set), although a delicate consideration of the dynamical equations in both their original and their reparameterized forms is involved in the convergence analysis. A similar, simpler approximation procedure was followed in [21], which does not allow state dependence of \( g \).

The optimality conditions assert the existence of a costate function \( p \) which satisfies, among other things, a measure-driven Hamiltonian system involving state derivatives of the data. Because the data are nonsmooth, and the derivatives are set valued, the Hamiltonian system is a measure-driven inclusion. The formulation and derivation of the optimality conditions make use of the concept of “robust solutions” to measure driven differential inclusions, and associated closure properties, provided in a recent paper [18].

The impulse controls considered here are scalar valued. For problems involving vector-valued controls the picture is complicated by the possibility that different sequences of approximating measures (absolutely continuous with respect to Lebesgue measure) of the same measure \( \mu \) can give different “state trajectories” (for a fixed initial state and conventional control); see, e.g., [2], [3], [4], [19]. The optimal control problem posed over processes \((x, u, \mu)\) still makes sense even though a unique “state trajectory” \( x \) no longer is associated with a fixed control policy and initial state. In the case that the data are smooth with respect to the time variable and the conventional control constraint set is constant, necessary conditions can once again be derived by applying the conventional maximum principle to a standard optimal control problem obtained by transforming the time variable (see Motta and Rampazzo [15] and Miller [12]). It would appear that necessary conditions for problems involving vector-valued impulse controls, when the data are assumed merely measurable in time, can be derived by reformulating the problem as one with scalar impulsive controls but one whose dynamics involve a measure differential inclusion, and adapting the techniques of this paper to allow for these more general dynamics.

We list notation and conventions adhered to below.
B denotes the open unit ball in Euclidean space. $C([0, 1]; \mathbb{R}^n)$ denotes the vector space of continuous $\mathbb{R}^n$-valued functions on $[0, 1]$ with supremum norm, and $C^*(0, 1); \mathbb{R}^n)$ its topological dual.

$C^+([0, 1]; \mathbb{R}^n) \subset C^*([0, 1]; \mathbb{R}^n)$ is the cone of functionals taking nonnegative values on nonnegative functions.

$AC([0, 1]; \mathbb{R}^n)$ is the space of absolutely continuous $\mathbb{R}^n$-valued functions on $[0, 1]$.

$BV^+([0, 1]; \mathbb{R}^n)$ denotes the vector space of $\mathbb{R}^n$-valued functions on $[0, 1]$, of bounded variation, which are continuous from the right on $(0, 1)$. The Borel measure associated with some $x \in BV^+([0, 1]; \mathbb{R}^n)$ is denoted $dx$.

For brevity we often do not distinguish between elements in $C^*([0, 1]; \mathbb{R}^n)$ and the Borel measures which represent them.

The weak* topology on $BV^+([0, 1]; \mathbb{R}^n)$ refers to the weak* topology on $(\mathbb{R}^n \times C([0, 1]; \mathbb{R}^n))^*$ under the isomorphism

$$x \to (x(0), dx).$$

Thus “$x_i \to x$ (weakly*)” indicates that $x_i(0) \to x(0)$ and $dx_i \to dx$ (weakly* in $C^*([0, 1]; \mathbb{R}^n)$). For simplicity we write $C(0, 1)$ in place of $C([0, 1]; \mathbb{R}^1)$, $C^*(0, 1)$ in place of $C^*([0, 1]; \mathbb{R}^1)$, and so on.

$\mathcal{L}$ denotes the Lebesgue subsets of $[0, 1]$, $\mathcal{B}$ the Borel sets in $\mathbb{R}^k$, and $\mathcal{L} \times \mathcal{B}$ the product $\sigma$-field.

The following concepts from nonsmooth analysis are required. Consider a closed set $A \in \mathbb{R}^k$ and points $x \in A$, $p \in \mathbb{R}^k$. We say that $p$ is a limiting normal to $A$ at $x$ if and only if there exist $p_i \to p$ and $x_i \overset{A}{\to} x$ such that for each $i$ we have

$$p_i : (z - x_i) \leq o(|z - x_i|) \quad \forall \; z \in A$$

(i.e., limiting normals are limits of vectors which support $A$ at points near $x$, to first order). The limiting normal cone to $A$ at $x$, written $N_A(x)$, comprises the limiting normals to $A$ at $x$.

Given a lower semicontinuous function $f : \mathbb{R}^k \to \mathbb{R} \cup \{+\infty\}$ and a point $x \in \mathbb{R}^k$ such that $f(x) < \infty$, we define the limiting subdifferential of $f$ at $x$, written $\partial f(x)$, to be

$$\partial f(x) := \{\xi : (-1, \xi) \in N_{\text{epi}(f)}(f(x), x)\},$$

in which $\text{epi}(f)$ denotes the epigraph set $\{(\eta, x) : \eta \geq f(x)\}$. In the event $f$ is Lipschitz continuous on a neighborhood of $x$, $\text{co}H\{f(x)\}$ coincides with the (Clarke) generalized gradient of $f$ at $x$, which may be defined directly [5].

The properties of limiting normal cones, limiting subdifferentials, and generalized gradients are developed in [9], [14] and [5].

2. Change of variables. We outline a change of variables technique, previously used in Rishel [16], Warga [22], Dal Maso and Rampazzo [6], and elsewhere, whose role will be to reduce measure-driven differential equations and inclusions to ordinary differential equations and inclusions.

Fix a measure $\mu \in C^+(0, 1)$. Let $F$ be its distribution function

$$F(t) := \begin{cases} \int_{[0, t]} \mu(ds), & t \in (0, 1], \\ 0 & \text{if} \; t = 0. \end{cases}$$

Define the reparameterization function $\eta$ corresponding to $\mu$ to be

$$\eta(t) := \begin{cases} (t + \int_{[0, t]} \mu(d\tau))/(1 + \mu([0, 1])), & t \in (0, 1], \\ 0 & \text{if} \; t = 0. \end{cases}$$
Evidently, \( \eta \) is a \( BV^+(0,1) \) function which is strictly increasing. Define also the continuous, nondecreasing function \( \theta : [0,1] \to [0,1] \) to be

\[
\theta(s) := \sup \{ t \in [0,1] : \eta(t) \leq s \} \quad \forall \ s \in [0,1].
\]

Let \( \{t_i\} \) be an enumeration of the atoms of \( \mu \), and let \( S_i(= [\sigma_i, \sigma_i']) \) be the subintervals

\[
S_i := \theta^{-1}(\{t_i\}) \text{ for } i = 1, 2, \ldots. \]

Now define the function \( \gamma : [0,1] \to \mathbb{R}^+ \) to be

\[
\gamma(s) := \begin{cases} 
F(\theta(s)) & \text{if } s \in [0,1] \setminus \bigcup_{i=1}^{\infty} S_i, \\
F(t_i^-) + \frac{(s-\sigma_i)}{\sigma_i'-\sigma_i}(F(t_i) - F(t_i^-)) & \text{for } s \in S_i, \ i = 1, 2, \ldots.
\end{cases}
\]

(In this formula \( F(t_i^-) \) and \( F(t_i) \) are interpreted as \( F(0) \) and \( F(0^+) \) if \( t_i = 0 \).)

The function \( (\theta, \gamma) : [0,1] \to (\mathbb{R}^+)^2 \) is called the graph completion of the measure \( \mu \). It is so called because it corresponds to filling in with straight line segments the graph of \( F \) and reparameterizing the resulting curve in \( \mathbb{R}^2 \).

Basic properties of the graph completion are as follows.

**Proposition 2.1.** Let \( (\theta, \gamma) \) be the graph completion of \( \mu \in C^+(0,1) \). Then

(i) \( \theta \) and \( \gamma \) are Lipschitz continuous, nonnegative, nondecreasing functions and

\[
\dot{\theta}(s) + \dot{\gamma}(s) = 1 + \mu([0,1]) \quad \mathcal{L} \text{ a.e.}
\]

(ii) For any Borel measurable function \( h \) which is \( \mu \) integrable and any Borel set \( T \subset [0,1] \) we have

\[
\int_{\theta^{-1}(T)} h(\theta(s)) \gamma(s)ds = \int_{T} h(\tau)\mu(d\tau).
\]

(iii) For any \( \mathcal{L} \)-integrable function \( g \) and Borel set \( S \subset [0,1] \), \( \theta(S) \) is also a Borel set and

\[
\int_{\theta(S)} g(\theta(s))\dot{\theta}(s)ds = \int_{\theta(S)} g(\tau)d\tau.
\]

(iv) Let \( \{\mu_i\} \) be a sequence of elements in \( C^+(0,1) \), and let \( \{(\theta_i, \gamma_i)\} \) be the corresponding graph completions. Suppose that \( \mu_i \rightharpoonup \mu \) (weakly*). Then

\( (\theta_i, \gamma_i) \to (\theta, \gamma) \) uniformly and \( (\dot{\theta}_i, \dot{\gamma}_i) \to (\dot{\theta}, \dot{\gamma}) \) weakly in \( L^1 \).

Parts (i) and (iv) are proven by Dal Maso and Rampazzo [6]. We comment briefly on the other assertions. (ii) will be recognized as an example of the “change of variables” lemma [7, Theorem 6.9], since \( \mu \) can be interpreted as the measure induced by the measure \( \dot{\gamma}(s)ds \) under the mapping \( \theta \), i.e.,

\[
\mu(A) = \int_{\theta^{-1}(A)} \dot{\gamma}(s)ds \quad \forall \ A \in \mathcal{B}.
\]

As for (iii), \( \theta(S) \) is a Borel set since \( \theta \) is monotone. The identity is another consequence of the change-of-variables lemma in view of the fact that

\[
\int_{\theta^{-1}\theta(S)} h\dot{\theta}ds = \int_{S} h\dot{\theta}ds.
\]

This last relationship comes about because \( \dot{\theta}(s) = 0 \) almost everywhere on the set where \( \theta \) is not one-to-one.
3. Measure-driven differential inclusions. In this section we give precise meaning to robust solutions of measure-driven differential inclusions (MDIs) of the form

\begin{equation}
\begin{aligned}
    dx(t) & \in F_1(t, x(t))dt + F_2(t, x(t))\mu(dt), \quad t \in [0, 1], \\
x(0) & = x_0
\end{aligned}
\end{equation}

for which the data are multifunctions \(F_1 : [0, 1] \times \mathbb{R}^n \longrightarrow \mathbb{R}^n\) and \(F_2 : [0, 1] \times \mathbb{R}^n \longrightarrow \mathbb{R}^n\). In the equation, \(\mu \in C^+(0, 1)\) is a given measure, and \(x_0 \in \mathbb{R}^n\) a given initial state.

The definition of robust solution involves the multifunction \(\tilde{F}_2 : [0, 1] \times \mathbb{R}^n \times [0, \infty) \longrightarrow \mathbb{R}^n:\)

\[\tilde{F}_2(t, v; \alpha) := \{\alpha^{-1}[\xi(1) - \xi(0)] : \xi \in AC([0, 1]; \mathbb{R}^n), \ \dot{\xi}(\sigma) \in \alpha F_2(t, \xi(\sigma)) \text{ a.e.,}
\text{ and } \xi(0) = v\}\]

if \(\alpha > 0\) and \(\tilde{F}_2(t, v; 0) := F_2(t, v)\).

**Definition 3.1.** We say that a function \(x \in BV^+([0, 1]; \mathbb{R}^n)\) is a robust solution to (3.1) if there exist an \(\mathcal{L}\)-integrable function \(\phi_1\) and \(\mu\)-integrable function \(\phi_2\) such that

\[\phi_1(t) \in F_1(t, x(t)), \quad \mathcal{L}\text{-a.e.,}
\phi_2(t) \in \tilde{F}_2(t, x(t^{-}); \mu(\{t\})), \quad \mu\text{-a.e.,}\]

and

\[x(t) = x(0) + \int_0^t \phi_1(\tau)d\tau + \int_{[0, t]} \phi_2(\tau)\mu(d\tau) \quad \forall \ t \in (0, 1].\]

Reparameterization by means of the graph completion of \(\mu\) results in a (conventional) differential inclusion as described in the following proposition. Here \(\eta\) is the reparameterization function of \(\mu\), and \((\theta(\cdot), \gamma(\cdot))\) is the graph completion of this measure (see section 2).

**Proposition 3.1.** Suppose that the data for MDI (3.1) satisfies the following:

- \(F_1\) has values-closed sets and is \(\mathcal{L} \times \mathcal{B}\) measurable
- \(F_2\) has values-closed sets and is Borel measurable.

Fix a measure \(\mu \in C^+(0, 1)\) and an initial state \(x_0\). We have the following:

(i) Suppose that \(x(\cdot) \in BV^+([0, 1]; \mathbb{R}^n)\) is a robust solution to MDI (3.1). Then there exists a solution \(y(\cdot) \in AC([0, 1]; \mathbb{R}^n)\) to

\begin{equation}
\begin{aligned}
    \dot{y}(s) & = F_1(\theta(s), y(s))\dot{\theta}(s) + F_2(\theta(s), y(s))\gamma(s), \quad s \in [0, 1], \\
y(0) & = x_0
\end{aligned}
\end{equation}

for which

\begin{equation}
    x(t) = y(\eta(t)) \quad \forall \ t \in [0, 1].
\end{equation}

Conversely,

(ii) suppose that \(y(\cdot) \in AC([0, 1]; \mathbb{R}^n)\) is a solution to (3.2). Then there exists a robust solution \(x(\cdot) \in BV^+([0, 1]; \mathbb{R}^n)\) to MDI (3.1) for which (3.3) is satisfied.

**Proof.** See [18, Theorem 4.1].
Solutions to the MDI (3.1) (as defined above) are “robust” in the sense that the set of solutions has desirable “closure” properties with respect to perturbations of the driving measure $\mu$ and the initial state. A result of this kind, suitable for future applications, is conveniently described in terms of a sequence of MDIs approximating (3.1), namely,

$$
\begin{align*}
& dx_i(t) \in F_1^{(i)}(t, x_i(t)) dt + F_2(t, x_i(t))\mu_i(dt) \quad \text{on } [0, 1], \\
& x_i(0) = x_0^i,
\end{align*}
$$

$i = 1, 2, \ldots$

Here $F_1^{(i)} : [0, 1] \times \mathbb{R}^n \Rightarrow \mathbb{R}^n$, $i = 1, 2, \ldots$; $F_1 : [0, 1] \times \mathbb{R}^n \Rightarrow \mathbb{R}^n$; and $F_2 : [0, 1] \times \mathbb{R}^n \mapsto \mathbb{R}^n$ are given multifunctions. $\mu_i$, $i = 1, 2, \ldots$, and $\mu$ are elements in $C^+(0, 1)$, and $x_0^i, i = 1, 2, \ldots$, and $x_0$ are $n$-vectors.

The associated reparameterized equations are

$$
\begin{align*}
& \dot{\theta}_i(s) \in F_1^{(i)}(\theta_i(s), y_i(s))\dot{\gamma}_i(s) + F_2(\theta_i(s), y_i(s))\gamma_i(s), \quad s \in [0, 1], \\
& y_i(0) = x_0^i,
\end{align*}
$$

in which $(\theta_i, \gamma_i)$ is the graph completion of $\mu_i$, $i = 1, 2, \ldots$, and $(\theta, \gamma)$ is the graph completion of $\mu$. Denote by $\eta : [0, 1] \rightarrow [0, 1]$ the reparameterization function for $\mu$.

We refer also to the reparameterization function $\eta : [0, 1] \rightarrow [0, 1]$ of $\mu_i$, $i = 1, 2, \ldots$. Reparameterization of the nominal MDI (3.1) results in the differential inclusion

$$
\begin{align*}
& \dot{\theta}(s) \in F_1(\theta(s), y(s))\dot{\gamma}(s) + F_2(\theta(s), y(s))\gamma(s), \quad s \in [0, 1], \\
& y(0) = x_0.
\end{align*}
$$

We have the following proposition.

**PROPOSITION 3.2.** Consider multifunctions $F_1$, $F_1^{(i)}$, $i = 1, 2, \ldots$, and $F_2$ with domain $[0, 1] \times \mathbb{R}^n$ and taking values-compact subsets of $\mathbb{R}^n$. Assume that

- $F_1^{(i)}(t, \cdot), i = 1, 2, \ldots$, and $F_1(t, \cdot)$ have closed graph and $F_1^{(i)}(\cdot, \cdot), i = 1, 2, \ldots$, and $F_1(\cdot, \cdot)$ are $\mathcal{L} \times \mathcal{B}$ measurable.
- $F_1(t, x)$ is convex for all $(t, x)$.
- $F_2(\cdot, \cdot)$ has closed graph and takes values convex sets.

Assume further that

- $\mathcal{L}$-measure $\{t : F_1^{(i)}(t, x) = F_1(t, x) \quad \forall x \in \mathbb{R}^n\} \rightarrow 1$ as $t \rightarrow \infty$.

Take a sequence $\{x_0^i\}$ in $\mathbb{R}^n$, a sequence $\{\mu_i\}$ in $C^+(0, 1)$, and elements $x_0 \in \mathbb{R}^n$ and $\mu \in C^+(0, 1)$. Take also a sequence $\{x_i\} \in BV^+(0, 1; \mathbb{R}^n)$ such that $x_i$ is a robust solution to (3.4) for each $i$ and

$$
x_0^i \rightarrow x_0 \quad \text{and} \quad \mu_i \rightarrow \mu \quad \text{(weakly*)} \quad \text{as} \quad i \rightarrow \infty.
$$

Assume that there exists $\beta(t) \in L^1$ and $c > 0$ such that $F_1^{(i)}(t, x_i(t)) \subset \beta(t)B$ a.e. and $F_2(t, x_i(t)) \subset cB$ for all $t$.

Then there exists a sequence $\{y_i\} \subset AC([0, 1]; \mathbb{R}^n)$ such that $y_i$ is a solution to (3.5) for each $i$, a solution $y$ to (3.6), and a robust solution $x$ to (3.1) such that

$$
x_i(t) = y_i(\eta_i(t)) \quad \forall \quad t \in [0, 1]
$$

and

$$
x(t) = y(\eta(t)) \quad \forall \quad t \in [0, 1].
$$
Along a subsequence we have
\[ x_i \longrightarrow x \quad \text{(weakly*)}, \]
\[ x_i(t) \longrightarrow x(t) \quad \forall \ t \in \left( [0,1] \setminus \mathcal{M}_\mu \right) \cup \{0,1\} \]
(where \( \mathcal{M}_\mu \) denotes the set of atoms of \( \mu \)), and
\[ y_i \longrightarrow y \quad \text{strongly in } C([0,1]; \mathbb{R}^n). \]

Proof. See [18, Theorem 5.1]. \qed

Consider now the dynamic equation of the problem (P) (the optimal control problem introduced in section 1)
\[
\begin{cases}
  dx(t) = f(t, x(t), u(t))dt + g(t, x(t))\mu(dt), \\
  x(0) = x_0.
\end{cases}
\]
(3.7)

For a fixed control policy \((u, \mu)\) this is just an example of MDI (3.1) (set \( F_1(t, x) = \{ f(t, x, u(t)) \} \) and \( f_2 = \{ g \} \)). We may therefore speak of robust solutions of the dynamic equation (3.4) (corresponding to a control policy \((u, \mu)\)). Henceforth, a process is taken to be a triple of elements \((x, u, \mu)\), in which \((u, \mu)\) is a control policy (see section 1) and \(x\) is a robust solution of the dynamical equation (3.7). Notice that, under the hypotheses imposed below there will be a unique robust solution to (3.7) for given \((u, \mu)\). This follows from the characterization of robust solutions provided by Proposition (3.1) since, for the MDI associated with (3.7), the differential inclusion (3.2) reduces to a differential equation which is known to have a unique solution.

4. A maximum principle. Our aim is to obtain optimality conditions for problem (P) in the form of a maximum principle. These will follow from conditions on processes (as defined is section 3) which generate boundary points of some “reachable set” of the control system with dynamic equations:
\[
\begin{align*}
  dx(t) &= f(t, x(t), u(t))dt + g(t, x(t))\mu(dt), \\
  x(0) &= x_0.
\end{align*}
\]
(3.7)

Take a locally Lipschitz continuous function \( \psi : \mathbb{R}^n \to \mathbb{R}^k \) and a closed set \( D \subset \mathbb{R}^n \). We define the \((\psi, D)\)-reachable set \( \mathcal{R}_{\psi, D} \) to be:
\[
\mathcal{R}_{\psi, D} := \{ \psi(x(1)) : (x, u, \mu) \text{ is a process and } x(0) \in D \}.
\]

The following hypotheses will be invoked:

(H1) There exists a constant \( K_f(\cdot) \in L^1 \) such that
\[
|f(t, x, u) - f(t, y, u)| \leq K_f(t) |x - y| \quad \text{for } (x, u) \in \mathbb{R}^n \times \mathbb{R}^m \text{ and } t \in [0,1].
\]

(H2) \( f(\cdot, x, \cdot) \) is \( \mathcal{L} \times \mathcal{B} \)-measurable.

(H3) \( g(\cdot, \cdot) \) is continuous and there exists a constant \( K_g \) such that
\[
|g(t, x) - g(t, y)| \leq K_g |x - y| \quad \forall \ x, y \in \mathbb{R}^n, \ t \in [0,1].
\]

(H4) \( U \in \mathbb{R}^{1+m} \) is a Borel set.

Theorem 4.1. Let \( \{ \tilde{x}(\cdot), u(\cdot), \tilde{\mu}(\cdot) \} \) be a process for which \( x(0) \in D \) and \( \psi(\tilde{x}(1)) \) is a boundary point of \( \mathcal{R}_{\psi, D} \). Assume that hypotheses (H1)–(H4) are satisfied. Then
there exist a function \( p \in BV^+([0,1];\mathbb{R}^n) \) and a unit vector \( d \in \mathbb{R}^k \) such that 
(\( \bar{x}(\cdot), p(\cdot) \)) is a robust solution of the MDI
\[
(4.1) \quad d \quad \bar{x}(t)
\[
\begin{bmatrix}
\frac{d}{dt} \bar{x}(t) \\
p(t)
\end{bmatrix}
\in \left[ \begin{bmatrix}
f(t, \bar{x}(t), \bar{u}(t)) \\
-p(t) \cdot \co\partial_x f(t, \bar{x}(t), \bar{u}(t))
\end{bmatrix}
+ \left[ -p(t) \cdot \partial_x g(t, \bar{x}(t)) \right] \bar{\mu}(dt) \right],
\]

Furthermore,
\[
(4.2) \quad -p(1) \in d \cdot \partial \psi(\bar{x}(1)),
\]
\[
(4.3) \quad p(0) \in N_D(\bar{x}(0)),
\]
\[
(4.4) \quad p(t) \cdot f(t, \bar{x}(t), \bar{u}(t)) = \max_{u \in U_t} \{ p(t) \cdot f(t, \bar{x}(t), u) \} \quad \text{a.e.} \quad t \in [0, 1],
\]
\[
(4.5) \quad p(t) \cdot g(t, \bar{x}(t)) \leq 0 \quad \forall \quad t \in (0, 1),
\]
\[
p(t) \cdot g(t, \bar{x}(t)) \geq 0, \quad \bar{\mu} - \text{a.e.} \quad t \in [0, 1],
\]
and corresponding to every atom \( t \) of \( \bar{\mu} \), there exists a solution \((\xi_t, \alpha_t)\) to
\[
\left[ \begin{array}{c}
\xi_t(s) \\
\alpha_t(s)
\end{array} \right] \in \bar{\mu}(\{t\}) \left[ \begin{array}{c}
g(t, \xi_t(s)) \\
-\alpha_t(s) \cdot \partial_x g(t, \xi_t(s))
\end{array} \right], \quad t \in [0, 1],
\]
(\(4.6\))

\[
(\xi_t(0), \alpha_t(0)) = (\bar{x}(t^-), p(t^-)), \quad (\xi_t(1), \alpha_t(1)) = (\bar{x}(t), p(t))
\]

that satisfies
\[
(4.7) \quad \alpha_t(s) \cdot g(t, \xi_t(s)) \geq 0 \quad \forall \quad s \in [0, 1].
\]

Here \(4.6\) is interpreted as
\[
(\xi_t(0), \alpha_t(0)) := (\bar{x}(0), p(0)), \quad (\xi_t(1), \alpha_t(1)) := (\bar{x}(0^+), p(0^+))
\]
if \( t = 0 \). Also \( \partial_x g(t, x) \) denotes the set
\[
\partial_x g(t, x) := \left\{ \lim_{i} a_i : a_i \in \co\partial_x g(t, x_i) \text{ for some } t_i \to t, \; x_i \to x \right\}.
\]

A proof is given in section 5.

The transition from Theorem 4.1 to a maximum principle for problem (P) follows the standard lines. Indeed, suppose that \( h : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R} \) is a locally Lipschitz continuous function and \( C \) is a closed subset of \( \mathbb{R}^n \times \mathbb{R}^n \), and let \((\bar{x}, \bar{u}, \bar{\mu})\) be a minimizing process for problem (P).

Consider the control system in which state trajectories are triples \((x, y, z)\):
\[
d \quad (x(t), y(t), z(t)) = \left( f(t, x(t), u(t)), 0, 0 \right) dt + \left( g(t, x(t)), 0, 0 \right) \mu(dt),
\]
\[
(x(0), y(0), z(0)) \in D.
\]

Here
\[
D := \{(x, y, z) : (x, y) \in C \text{ and } z \geq h(x, y)\}.
\]

It is easy to deduce from the optimality of \((\bar{x}, \bar{u}, \bar{\mu})\) that \((0, h(\bar{x}(0), \bar{x}(1))\) is a boundary point of \( \mathcal{R}_{\psi,\mu}, \psi(x, y, z) := (y - x, z) \).

Applying the earlier theorem to this boundary process we arrive at the following maximum principle for (P).
THEOREM 4.2. Let \((\bar{x}(\cdot), \bar{u}(\cdot), \bar{\mu}(\cdot))\) be a minimizing process for \((P)\). Assume that 
\(h\) is locally Lipschitz continuous, that \(C\) is a closed subset, and that hypotheses \((H1)\)–
\((H4)\) are satisfied.

Then there exist \(\lambda \geq 0\) and \(p \in BV^+ ([0, 1]; \mathbb{R}^n)\) such that 
\(\|p(\cdot)\|_{L^\infty} + \lambda > 0\) and 
\((\bar{x}(\cdot), p(\cdot))\) is a robust solution of the MDI

\[ 
dx \begin{bmatrix} x(t) \\ p(t) \end{bmatrix} \in \begin{bmatrix} -\frac{f(t, x(t), u(t))}{p(t)} g(t, x(t)) \end{bmatrix} dt + \begin{bmatrix} \frac{g(t, x(t))}{p(t)} \end{bmatrix} \bar{\mu}(dt). \]

Furthermore,

\( (p(0), -p(1)) \in NC (\bar{x}(0), \bar{x}(1)) + \lambda \partial h \bar{x}(0), \bar{x}(1) \),

\( p(t) \cdot f(t, \bar{x}(t), \bar{u}(t)) = \max_{u \in U_t} \{ p(t) \cdot f(t, \bar{x}(t), u) \} \quad \text{a.e. } t \in [0, 1], \)

\( p(t) \cdot g(t, \bar{x}(t)) \leq 0 \quad \forall \ t \in [0, 1], \)

\( p(t) \cdot g(t, \bar{x}(t)) = 0, \quad \bar{\mu} - \text{a.e. on } [0, 1]. \)

Corresponding to every atom \(t\) of \(\bar{\mu}\), there exists a solution \(\begin{bmatrix} \xi_t(\cdot) \\ \alpha_t(\cdot) \end{bmatrix}\) to

\[ 
\begin{bmatrix} \xi_t(s) \\ \alpha_t(s) \end{bmatrix} \in \bar{\mu}(\{t\}) \begin{bmatrix} g(t, \xi_t(s)) \\ -\alpha_t(s) \cdot \partial g(t, \xi_t(s)) \end{bmatrix} \quad \text{on } [0, 1]
\]

which satisfies

\( (\xi_t(0), \alpha_t(0)) = (\bar{x}(t^-), p(t^-)), \quad (\xi_t(1), \alpha_t(1)) = (\bar{x}(t), p(t)), \)

\( \alpha_t(s) \cdot g(t, \xi_t(s)) \geq 0 \quad \forall \ s \in [0, 1]. \)

5. Proof of Theorem 4.1. Take a process \((\bar{x}, \bar{u}, \bar{\mu})\) with the stated “boundary”
property. The assertions of the theorem are proven first in the special case when an
interim hypothesis,

(\(\bar{H}\)) There exists \(\alpha \in L^1\) such that \(\sup_{u \in U(t)} |f(t, \bar{x}(t), u)| \leq \alpha(t) \quad \text{a.e.}\)

is added to \((H1)\)–\((H4)\). We show how to remove \((\bar{H})\) in the final stage of the proof.
(\(\bar{H}\)) has the role of ensuring suitable “linear growth properties” of the data, namely,
the following lemma.

LEMMA 5.1. There exist \(\alpha_1, \alpha_2 \in L^1\) and \(\beta_1 \geq 0, \beta_2 \geq 0\) such that

\[ |f(t, x, u)| \leq \alpha_1(t)|x| + \alpha_2(t) \quad \forall \ x \in \mathbb{R}^n \text{ and } u \in U_t, \mathcal{L}-\text{a.e.} \]

and

\[ |g(t, x)| \leq \beta_1|x| + \beta_2 \quad \forall \ x \in \mathbb{R}^n \text{ and } t \in [0, 1]. \]

Proof. An appeal to \((H1), (H3),\) and \((\bar{H})\) and repeated applications of the triangle
inequality validate the inequalities with

\( \alpha_1(t) := K_f(t), \quad \alpha_2(t) := \alpha(t) + K_f(t) \|\bar{x}(\cdot)\|_{L^\infty}, \)

\( \beta_1 := K_g, \quad \beta_2 := \sup_{s \in [0, 1]} |g(s, \bar{x}(s))| + K_g \|\bar{x}(\cdot)\|_{L^\infty}. \)

Next, we have a lemma describing the continuity properties of state trajectories
with respect to control policies and initial states.
Lemma 5.2. Given any control policy \((u, \mu)\) and initial state \(x_0\) there is a unique robust solution \(x\) of the dynamical equation

\[
\begin{aligned}
&dx(t) = f(t, x(t), u(t))\, dt + g(t, x(t))\, \mu(dt), \quad t \in [0, 1], \\
x(0) = x_0
\end{aligned}
\]

and a unique solution \(y\) of the reparameterized equation

\[
\begin{aligned}
&\dot{y}(s) = f(\theta(s), y(s), u(\theta(s)))\, \dot{\theta}(s) + g(\theta(s), y(s))\, \dot{\gamma}(s), \quad s \in [0, 1], \\
y(0) = x_0.
\end{aligned}
\]

(Here \((\theta(s), \gamma(s))\) is the graph completion of the measure \(\mu\).) If \(\{(u_i(\cdot), \mu_i)\}\) and \(\{x_i^0\}\) are sequences of control policies and initial states, respectively, such that

\[
\mathcal{L}\text{-meas}\{t : u_i(t) \neq u(t)\} \to 0, \\
\mu_i \rightharpoonup \mu \quad \text{weakly*},
\]

then \(x_i(t) \to x(t) \forall t \in \mathcal{C}_\mu \cup \{0, 1\}, \) \(dx_i \to dx\) (weakly*), and \(y_i(t) \to y(t)\) uniformly, where \(\{x_i(\cdot)\}\) and \(\{y_i(\cdot)\}\) are the corresponding sequences of state trajectories and reparameterized state trajectories and \(\mathcal{C}_\mu\) is the set of continuity points of \(\mu\).

Proof. Notice that the right-hand side

\[
\phi(s, y) := f(\theta(s), y, u(\theta(s)))\, \dot{\theta}(s) + g(\theta(s), y)\, \dot{\gamma}(s)
\]

of the reparameterized equation (5.2) satisfies the growth condition,

\[
|\phi(s, y)| \leq \alpha_1'(s)|y| + \alpha_2'(s) \quad \text{a.e.} \quad s \in [0, 1]
\]

\[\forall y \in \mathbb{R}^n\], where \(\alpha_1'\) and \(\alpha_2'\) are the integrable functions

\[
\begin{array}{l}
\alpha_1'(s) := \alpha_1(\theta(s))\, \dot{\theta}(s) + \beta_1(1 + \mu([0, 1])), \\
\alpha_2'(s) := \alpha_2(\theta(s))\, \dot{\theta}(s) + \beta_2(1 + \mu([0, 1])),
\end{array}
\]

and also the Lipschitz condition,

\[
|\phi(s, y) - \phi(s, z)| \leq k'(s)|y - z|
\]

for all \(y, z \in \mathbb{R}^n\), where \(k'(s)\) is the integrable function

\[
k'(s) := \int f(\theta(s))\, \dot{\theta}(s) + K_g(1 + \mu([0, 1])).
\]

These are known conditions under which the reparameterized equation (5.2) has a unique solution \(y \in AC([0, 1]; \mathbb{R}^n)\). But then by Proposition 3.1 there is a unique robust solution to (5.1).

An important observation is that, if \(c > 0\) is a constant such that \(\mu([0, 1]) < c\), then the coefficients in the linear growth inequality (5.3) have \(L^1\) norm bounded above by a number which depends only on \(c\). (This is clear from (5.4) and the fact that \(\int \alpha_1'(\theta(s))\, \dot{\theta}(s)\, ds = \int \alpha_1'(t)\, dt\).) Since \(\{\mu_i\}\) is a weak* convergent sequence, the \(\mu_i\)'s and \(\mu\) are uniformly bounded in total variation, and the \(x_i^0\)'s too are uniformly
bounded, it follows from an application of Gronwall’s lemma that the solutions $y_i$ to the reparameterized equations, resulting from replacing $(u, \mu)$ by $(u_i, \mu_i)$ and $x_0$ by $x_0^{(i)}$, $i = 1, 2, \ldots$, are uniformly bounded in the supremum norm. Since $\|x_i\|_{L^\infty} \leq \|y_i\|_{L^\infty}$, $i = 1, 2, \ldots$, the $x_i$’s are likewise bounded. The $y_i$’s are, in addition, equicontinuous. To show this, we choose constants $r > 0$ such that $\|y_i\|_{L^\infty} \leq r$, $i = 1, 2, \ldots$, and $\rho > 0$ such that $|g(t, y)| \leq \rho$ for all $(t, y) \in [0, 1] \times rB$. Fix $\epsilon > 0$. Choose $\delta > 0$ and let $[\sigma_1, \sigma_2] \subset [0, 1]$ be an arbitrary interval such that $|\sigma_2 - \sigma_1| < \delta$ and $i$ an arbitrary index value. We have

$$|y_i(\sigma_2) - y_i(\sigma_1)| \leq \left| \int_{\sigma_1}^{\sigma_2} f(\theta_i(s), y_i(s), u_i(\tilde{\theta}_i(s)))\gamma(s)ds \right| + \left| \int_{\sigma_1}^{\sigma_2} g(\theta_i(s), y_i(s))\gamma(s)ds \right|$$

$$\leq \left| \int_{t_1}^{t_2} f(t, x_i(t), u_i(t))dt \right| + \rho(1 + \mu([0, 1]))|\sigma_2 - \sigma_1|$$

$$\leq r \left| \int_{t_1}^{t_2} \alpha_1(t)dt + \int_{t_1}^{t_2} \alpha_2(t)dt + \rho(1 + \mu([0, 1]))|\sigma_2 - \sigma_1| \right|. $$

Here $t_1 = \theta_i(\sigma_1)$ and $t_2 = \theta_i(\sigma_2)$. Since $|t_2 - t_1| \leq (1 + \mu([0, 1]))|\sigma_2 - \sigma_1|$ and the $|\mu_k([0, 1])|$’s are uniformly bounded, this last inequality implies that, by choosing $\delta > 0$ sufficiently small, we can arrange that $|y_i(\sigma_2) - y_i(\sigma_1)| < \epsilon$ independently of our choice of interval $[\sigma_1, \sigma_2]$ and index value $i$. This confirms the equicontinuity of the $y_i$’s.

The hypotheses are satisfied under which Proposition 3.2 may be applied when $F_i(t, x) := \{f(t, x, u_i(t))\}$, $i = 1, 2, \ldots$, and $F_2(t, x) := \{g(t, x)\}$. For arbitrary subsequences, further subsequences may be chosen such that $\{y_i\}$ and $\{x_i\}$ converge as described. Since, however, the limits are unique, the original sequences must converge.

Since $\psi(\bar{x}(1)) \in \partial R_{\psi, D}$ there exists a sequence of vectors $\{\xi_j\}$ such that $\xi_j \notin R_{\psi, D}$ for $j = 1, 2, \ldots$ and

$$\xi_j \rightarrow \psi(\bar{x}(1)) \quad \text{as} \quad j \rightarrow \infty. $$

We may construct, using standard procedures (see, e.g., [20]), a sequence of nonnegative-valued $L^\infty$ functions $\{\bar{m}_j(\cdot)\}$ such that

$$\bar{m}_j(t)dt \rightarrow \bar{\mu}(dt) \quad \text{weakly*}. $$

Let $\bar{x}_j(\cdot)$ be the state trajectory corresponding to the conventional and measure controls, $\bar{u}(\cdot)$ and $m_j(t)dt$, and initial state $\bar{x}_0$. According to Lemma 5.2,

$$\bar{x}_j \rightarrow \bar{x}(t) \quad \forall \ t \in C_{\overline{\mu}} \cup \{0, 1\}$$

and

$$d\bar{x}_j(t) \rightarrow d\bar{x}(t) \quad \text{weakly*}. $$

Now define $\epsilon_j := |\xi_j - \psi(\bar{x}(t))|^{1/2}$. Note that, by the properties of $\{\xi_j\}$, $\epsilon_j \rightarrow 0$ as $i \rightarrow \infty$. 
For each \( j \), set \( r_j := j + \| \tilde{m}_j(\cdot) \|_{L^{\infty}} \) and consider the optimal control problem

\[
(P_j) \quad \left\{ \begin{array}{l}
\text{Minimize} \quad |\xi_j - \phi(\tilde{x}(t))| \\
\text{subject to} \quad \dot{x}(t) = f(t, x(t), u(t)) + g(t, x(t))m(t), \quad t \in [0, 1], \\
(u(t), m(t)) \in U_t \times [0, r_j] \quad \mathcal{L}\text{-a.e.,} \\
x_0 \in D.
\end{array} \right.
\]

Since the dynamic constraint defines a unique state trajectory \( x_{a,u,m}(\cdot) \) corresponding to a given control function \( (u(\cdot), m(\cdot)) \) and initial state \( a \), we may regard this as a minimization problem over triples \( (a, u(\cdot), m(\cdot)) \) such that \( a \in D; u(\cdot) \) and \( m(\cdot) \) are measurable functions; and \( (u(t), m(t)) \in U_t \times [0, r_j] \) a.e. Denote this collection of triples by \( W_j \). We provide \( W_j \) with the metric

\[
\rho((a, u, m); (a', u', m')) := |a - a'| + \mathcal{L}\text{-meas}\{t \in [0, 1] : u(t) \neq u'(t)\} + \int_0^1 |m(t) - m'(t)|dt.
\]

The newly reformulated problem can be expressed as

\[
\text{minimize}\{\Psi_j(a, u, m) : (a, u, m) \in W_j\},
\]

in which

\[
\Psi_j(a, u, m) := |\xi_j - \phi(x_{a,u,m}(1))|.
\]

By Lemma 5.2, \( W_j \) is a complete space and \( \Psi_j \) is continuous with respect to the above metric. According to Ekeland’s theorem [5] then there exists a triple \( (a_j, u_j, m_j) \in W_j \) such that

\[
(5.5) \quad \Psi_j(a_j, u_j, m_j) \leq \Psi_j(a, u, m) + \epsilon_j \rho((a, u, m), (a_j, u_j, m_j))
\]

for all \( (a, u, m) \in W_j \) and

\[
(5.6) \quad \rho(\tilde{x}(0), \tilde{u}, \tilde{m}_j), (a_j, u_j, m_j)) \leq \epsilon_j
\]

for each \( j \). Let \( x_j \) be the trajectory corresponding to \( (a_j, u_j, m_j) \). Since \( \epsilon_j \to 0 \) we deduce from (5.6) that

\[
x_j(0) \to \bar{x}(0),
\]

\[
\mathcal{L} - \text{meas}\{t \in [0, 1] : u_j(t) \neq \bar{u}(t)\} \to 0,
\]

and

\[
m_j(t)dt \to \bar{m}(dt) \quad \text{weakly}^*.
\]

For each \( j \), (5.5) implies that \( (x_j(\cdot), u_j(\cdot), m_j(\cdot)) \) is a minimizer for the problem

\[
\text{Minimize} \quad |\xi_j - \phi(x(1))| + \epsilon_j \left\{ |x(0) - x_j(0)| + \int_0^1 |m_j(t) - m(t)|dt + \int_0^1 \chi_j(t, u(t))dt \right\}
\]

subject to \( \dot{x}(t) = f(t, x(t), u(t)) + g(t, x(t))m(t) \quad \mathcal{L}\text{-a.e.,} \)

\[
(u(t), m(t)) \in U_t \times [0, r_j] \quad \mathcal{L}\text{-a.e.,}
\]
in which

$$\chi_j(t, u) := \begin{cases} 0 & \text{if } u = u_j(t), \\ 1 & \text{if } u \neq u_j(t). \end{cases}$$

Now apply the maximum principle to this last problem (see, e.g., [17]). Since $\xi_j \notin \Psi(x_j(1))$, this tells us that there exists an absolutely continuous function $p_j$ and a vector $d_j \in \mathbb{R}^k$ of unit length such that

$$- \dot{p}_j(t) \in p_j \cdot \nabla \partial_x f(t, x_j(t), u_j(t)) + p_j \cdot \nabla \partial_x g(t, x_j(t))m_j, \quad t \in [0,1],$$

(5.7)

$$p_j(0) \in N_D(x_j(0)) + \epsilon_j \partial_x |x - x_0|_{x=x_j(0)},$$

(5.8)

$$- p_j(1) \in d_j \cdot \partial \psi(x_j(1)),$$

and

(5.9) $$H_j(t, u_j(t), m_j(t)) = \max\{H_j(t, u, m) : u \in U_t, m \in [0, r_j]\} \text{ a.e.}$$

Here

$$H_j(t, u, m) := p_j(t) \cdot f(t, x_j(t), u) + p_j(t) \cdot g(t, x_j(t))m - \epsilon_j(|m - m_j(t)| + \chi_j(t, u)).$$

Notice that the maximization of the Hamiltonian condition (5.9) implies that $m_j(t) = r_j$ a.e. on the set $\{t : p_j(t) \cdot g(t, x_j(t)) > \epsilon_j\}$. Since $m_j(t)dt \rightarrow \mu(dt)$ weakly$^*$, the sequence $\{m_j\}$ is bounded in $L^1$ norm. Observing that $r_j \rightarrow \infty$, we deduce

(5.10) $$\mathcal{L}\text{-meas}\{t : p_j(t) \cdot g(t, x_j(t)) \leq \epsilon_j\} \rightarrow 1 \quad \text{as } j \rightarrow \infty.$$ 

It follows also from (5.9) that $p_j(t) \cdot g(t, x_j(t)) \geq -\epsilon_j$ a.e. (with respect to Lebesgue measure) on $\{t : m_j(t) > 0\}$. This property can be expressed as

(5.11) $$\int_0^1 \max\{-p_j(t) \cdot g(t, x_j(t)) - \epsilon_j, 0\} m_j(t)dt = 0, \quad j = 1, 2, \ldots.$$ 

We can regard $x_j$ and $p_j$ as solutions to the following MDI, corresponding to controls $(u, \mu) = (u_j, m_j(t)dt)$ and initial values $(x_j(0), p_j(0))$:

(5.12) $$d(x(t), p(t)) \in F_1(t, x(t), p(t), u(t))dt + F_2(t, x(t), p(t))\mu(dt),$$ 

in which

$$F_1(t, x, p, u) := \{f(t, x, u)\} \times \{-p \cdot \nabla \partial_x f(t, x, u)\}$$

and

$$F_2(t, x, p) := \{g(t, x)\} \times \{-p \cdot \partial_x g(t, x)\}.$$ 

Incorporation of the hybrid gradient $\partial_x g$ in these relationships ensures that $F_2$ has the requisite upper semicontinuity properties for application of the convergence results of section 5.

Bearing in mind the sequences $\{x_j\}$ and $\{p_j\}$ are uniformly bounded (in the case of $p_j$, this follows from the uniform bound on the right endpoints, and an application of Gronwall’s lemma), we deduce from Proposition 4.2 that (following an extraction
of subsequences) there exist \( \pi \in BV^+(0,1;\mathbb{R}^n) \), \( p \in BV^+(0,1;\mathbb{R}^n) \), and absolutely continuous functions \( y(\cdot) \) and \( q(\cdot) \) with the following properties:

\[
(5.13) \quad x_j(t) \to \pi(t) \quad \text{and} \quad p_j(t) \to p(t) \quad \forall \ t \in \mathcal{C}_\mu \cup \{0,1\},
\]

\[
x_j(\theta_j(s)) \to y(s) \quad \text{and} \quad p_j(\theta_j(s)) \to q(t) \quad \text{uniformly,}
\]

and

\[
(5.14) \quad \ddot{x}(t) = y(\eta(t)) \quad \text{and} \quad \dot{p}(t) = q(\eta(t)) \quad \forall \ t \in [0,1].
\]

Here \( \eta \) is the reparameterization function of \( \mu \), \( (\theta, \gamma) \) is the graph completion of \( \mu \), and \( (\theta_j, \gamma_j) \) is the graph completion of \( \mu_j \) for \( j = 1,2,\ldots \). Furthermore \( (\pi(\cdot), \dot{p}(\cdot)) \) is a robust solution of the MDI (5.12), and \( y \) and \( q \) satisfy the differential inclusion

\[
(\dot{y}(s), \dot{q}(s)) \in \{f(\theta(s), y(s), \pi(\theta(s)) \} \times \{-q(s) \cdot \cos \theta f(\theta(s), y(s), \pi(\theta(s)))\} \theta(s)
\]

\[
+ \{g(\theta(s), y(s)) \} \times \{-q(s) \cdot \varphi g(\theta(s), y(s))\} \gamma(s), \quad s \in [0,1].
\]

Notice that we are justified in taking the cluster point of the sequence \( \{x_j(\cdot)\} \) to be the “boundary” state trajectory \( \pi(\cdot) \) because the original sequence is known to have converged to \( \pi(\cdot) \) on \( \mathcal{C}_\mu \cup \{0,1\} \), and two functions in \( BV^+(0,1;\mathbb{R}^n) \) coincide if they have the same values on a dense set including \( \{0,1\} \).

We now arrange, by further subsequence extraction, that the sequence of vectors \( \{d_j\} \) in (5.8) has a limit:

\[
d_j \to d \quad \text{as} \quad j \to \infty
\]

for some vector \( d \) of unit length.

We have seen that \( (\pi, p) \) can be interpreted as a robust solution of the combined state and costate equations (5.12). Our goal now is to show that the \( p(\cdot) \) and \( d \) we have constructed satisfy the remaining conditions in the theorem statement.

For each \( j \) let \( S_j \) be the set of points \( t \in [0,1] \) such that

\[
p_k(t) \cdot g(t, x_k(t)) \leq \epsilon_k \quad \forall \ k \geq j,
\]

\[
u_k(t) = \pi(t) \quad \forall \ k \geq j,
\]

\[
H_k(t, u_k(t), m_k(t)) = \max\{H_k(t, u, m) : u \in U_t, m \in [0, r_k]\} \quad \forall \ k \geq j,
\]

and

\[
x_k \to \pi(t), \quad p_k(t) \to p(t).
\]

In view of (5.6), (5.9), (5.10), and (5.13), we can arrange by a further subsequence extraction that \( \mathcal{L} \)-meas \( \{S_j\} \to 1 \).

Take any \( t \in \bigcup_j S_j \) and any \( u \in U_t \). The above relationships imply that, in the limit,

\[
p(t) \cdot f(t, \pi(t), \pi(t)) = \max\{p(t) \cdot f(t, \pi(t), u) : u \in U_t\}
\]

and

\[
p(t) \cdot g(t, \pi(t)) \leq 0.
\]
We have shown (4.5) on $\cup_j S_j$. Then (4.5) holds on $(0, 1)$ because $\cup_j S_j$ is dense and $t \to p(t) \cdot g(t, \mathbf{x}(t))$ is right continuous on $(0, 1)$. The boundary conditions on the costate function $p(\cdot)$ are verified by passage to the limit in (5.7) and (5.8).

Next we examine the consequences of (5.11). Since $\{m_j(\cdot)\}$ is bounded in $L^1$ norm, we know that

$$\lim_{j \to \infty} \int_0^1 \max\{-p_j(t) \cdot g(t, x_j(t)), \ 0\} m_j(t) dt = 0.$$ 

Applying the change-of-variables lemma and using the facts that

$$\int_0^1 \max\{-p_j(\theta_j(s)) \cdot g(\theta_j(s), x_j(\theta_j(s))), \ 0\} \gamma_j(s) ds = 0$$ 

(see Proposition 2.1), we conclude that

$$\int_0^1 \max\{-q(s) \cdot g(\theta(s), y(s)), \ 0\} \dot{\gamma}(s) ds = \lim_{j \to \infty} \int_0^1 \max\{-p_j(\theta_j(s)) \cdot g(\theta_j(s, x_j(\theta_j(s))), \ 0\} \dot{\gamma}_j(s) ds = 0.$$ 

Since the integrand of the expression on the left is nonnegative, it follows that

$$(5.16) \quad \int_{[0,1]\setminus(\cup_i I_i)} \max\{-p(\theta(s)) \cdot g(\theta(s), \mathbf{x}(\theta(s)), \ 0\} \dot{\gamma}(s) ds = 0$$

(we have noted that $\mathbf{x}(\theta(s)) = y(s), p(\theta(s)) = q(s)$ on $[0, 1] \setminus (\cup_i I_i)$) and, for each $i$,

$$(5.17) \quad \int_{[s'_i, s''_i]} \max\{-q(s) \cdot g(t, y(s)), \ 0\} ds = 0$$

(in view of the fact that $\dot{\gamma} \equiv 1 + \mu([0, 1])$ on $[s'_i, s''_i]$). Here $\{t_i\}$ is an enumeration of the atoms of $\mu$ and

$$[s'_i, s''_i] = I_i = \theta^{-1}(\{t_i\}), \ i = 1, 2, \ldots.$$ 

The change-of-variables lemma applied to (5.16) gives

$$(5.18) \quad \int_{[0,1]\setminus(\cup t_i)} \max\{-p(t) \cdot g(t, \mathbf{x}(t)), \ 0\} \mathbf{\Pi}(dt) = 0.$$ 

Since the integrand in (5.17) is continuous, we conclude from this equation that

$$(5.19) \quad q(s) \cdot g(t_i, y(s)) \geq 0 \quad \forall \ s \in [s'_i, s''_i], \ i = 1, 2, \ldots.$$ 

We note, however, that, by (5.14), the continuity of $y$ and $q$, and the continuity of $\mathbf{x}$ from the right,

$$(5.20) \quad \mathbf{x}(t_i) = y(s''_i) \quad \text{and} \quad p(t_i) = q(s''_i).$$
and
\[ x(t_i^-) = y(s_i) \quad \text{and} \quad p(t_i^-) = q(s_i). \]

Equations (5.19) and (5.20) now give
\[ p(t_i) \cdot g(t, \pi(t)) = q(s_i^\mu) \cdot g(t, y(s_i)) \geq 0 \]
for all \( i \). Combining this relationship with (5.18) we conclude that
\[ p(t) \cdot g(t, \pi(t)) \geq 0 \quad \mu\text{-a.e.} \]

Next, for each \( i \), define \( \xi_i : [0, 1] \rightarrow \mathbb{R}^n \), \( \pi_i : [0, 1] \rightarrow \mathbb{R}^n \) to be
\[ \xi_i(\sigma) = y(s_i + \sigma(s_i^{\mu} - s_i)), \quad \pi_i(\sigma) = q(s_i + \sigma(s_i^{\mu} - s_i)) \]

By (5.20) and (5.21)
\[ (\xi_i(0), \pi_i(0)) = (x(t_i^-), p(t_i^-)), \quad (\xi_i(1), \pi_i(1)) = (x(t_i), p(t_i)) \]
for each \( i \). Furthermore, since \( \hat{\theta}(s) \equiv 0 \) and \( \hat{\gamma}(s) \cdot |s_i^{\mu} - s_i| \equiv \mu(\{t_i\}) \) on \([s_i, s_i^{\mu}]\), we deduce from (5.2) that
\[ (\xi_i(s), \hat{\pi}_i(s)) \in g(t_i, \xi_i(s)) \times (-\xi_i(s) \cdot \overline{\pi}g(t_i, \xi_i(s))\mu(\{t_i\}) \quad \text{a.e.} \]

This concludes the proof of the theorem in the case that \( H \) is added to the hypotheses. It remains to deal with the case when \( H \) is not valid. For each index value \( k \) replace \( U(t) \) with \( \tilde{U}_k(t) \):
\[ \tilde{U}_k(t) = U(t) \cap \{ u \in \mathbb{R}^m : |f(t, \pi(t), u)| \leq |f(t, \pi(t), \pi(t))| + k \}. \]

We see that \( (\pi, \pi, \pi) \) remains a minimizer. Furthermore, \( H \) is now satisfied. By our earlier analysis there exist \( (p_k(\cdot), d_k) \) with the properties described in the above theorem (except that now \( \tilde{U}_k(t) \) replaces \( U(t) \)). Passage to the limit as \( k \rightarrow \infty \) (the convergence analysis is along the lines of the first half of the proof) yields the assertions of Theorem 4.1 in this case also. \( \square \)

REFERENCES

1846  G. N. SILVA AND R. B. VINTER