Necessary Conditions without Differentiability Assumptions in Unilateral Control Problems*

J. WARGA

Northeastern University, Boston, Massachusetts 02115

Received May 2, 1974

We derive two theorems combining existence with necessary conditions for the relaxed unilateral problem of the optimal control of ordinary differential equations in which the functions that define the problem are Lipschitz-continuous in the state variables. These theorems generalize the results presented in a previous paper [8] by the addition of unilateral constraints on the state and control functions. As in that paper, the new necessary conditions have a canonical form obtained by replacing, in the "customary" conditions, the partial derivatives with respect to the state variables by finite difference quotients at neighboring arguments, and then applying limiting processes and convexification. More general necessary conditions are also obtained in terms of the representations of the Lipschitz-continuous functions as compositions.

1. INTRODUCTION

In this paper we continue our investigation of necessary conditions without differentiability assumptions, and generalize the results of [8] to unilateral control problems defined by ordinary differential equations. Specifically, we consider the relaxed optimal control problem defined, in its original (unrelaxed) version, by the cost functional $h^0(y(t))$ and the relations

$$\dot{y}(t) = f(t, y(t), u(t)) \quad \text{a.e. in } T = [t_0, t_1],$$

$$y(t_0) \in A_0, \quad h^1(y(t_1)) \in A_1,$$

$$u(t) \in U^*(t) \quad \text{a.e. in } T,$$

$$a^i(t, y(t)) \leq 0 \quad (t \in T, i = 1, 2, \ldots, m),$$

where the functions $f(t, \cdot, u), h^0, h^1,$ and $a^i(t, \cdot)$ are assumed Lipschitz-continuous over bounded sets but not necessarily differentiable or with any particular convexity properties. This problem differs from the one investi-
gated in [8] by the addition of the control restriction (1.3) and the unilateral
restriction (1.4); it differs from the problems previously investigated by the
author [4-7] and by Rockafellar [3] by the absence of differentiability or
convexity assumptions.

Our present approach is similar to that of [8] and is based on the use of
convolutions with mollifiers to approximate the functions \( h^0, h^1, a^i(t, \cdot) \) and
\( f(t, \cdot, u) \) by \( C^1 \) functions. This yields a sequence of "approximating" uni-
lateral problems of a type for which necessary conditions were previously
derived [6, 7]. Our final results are obtained by investigating the behavior of
these necessary conditions for ever finer approximations. The only phase of
this research using techniques other than those encountered in [8] is the
study, based on Lemmas 3.2 and 3.3, of the convergence of the "dual"
functions \( \kappa \) for the approximating unilateral problems.

Our basic results are presented in Theorems 2.2 and 2.3. The proofs are
carried out in Section 3.

2. Existence and Necessary Conditions

2.1. Definitions and Assumptions. We shall use the terms "a.e.",
"a.a." (almost all) and "measurable" in the sense of the Borel–Lebesgue
measure on \( T \) which we denote by \( \mu \). Our optimal control problem is relaxed
by replacing relations (1.1) and (1.3) by

\[
y(t) \in \text{co } f(t, y(t), \overline{U}^*(t)) \quad \text{a.e. in } T,
\]

where \( \overline{U}^*(t) \) is the closure of \( U^*(t) \) and \text{co} denotes the convex hull. We refer
to an absolutely continuous function \( y \) that satisfies relations (1.2), (1.4),
and (2.1.1) as an admissible relaxed solution. If \( \bar{y} \) minimizes \( h^0(\gamma(t)) \) among
admissible relaxed solutions, then it is a minimizing relaxed solution. For
\( x = (x^1, \ldots, x^d) \in \mathbb{R}^d \), we define the norm of \( x \) by \( |x| = \text{Max}_i |x^i| \). We make
the following assumptions:

(a) \( V \) is an open subset of \( \mathbb{R}^n \);
(b) \( A_0 \) and \( A_1 \) are closed convex subsets of \( \mathbb{R}^n \) and \( \mathbb{R}^m \), respectively;
(c) \( U \) is a compact metric space, \( U^*(t) \in U \), the function \( t \to U^*(t) \) is
measurable (that is, for every open subset \( G \) of \( U \), the set

\[
\{ t \in T \mid U^*(t) \cap G \neq \emptyset \}
\]

is measurable), and either \( U^*(t) \) is closed for all \( t \in T \) or, for all \( t \in T \), \( U^*(t) \)
is contained in the closure of its interior and for every \( \epsilon > 0 \) there exists
$T_e \subset T$ such that $\mu(T - T_e) \leq \epsilon$ and the set $\{(t, u) \in T_e \times U | u \in \text{interior } U^*(t)\}$ is open relative to $T_e \times U$;

(d) there exist a number $c > 0$ and a compact set $D \subset V$, such that the functions $h^0: V \to \mathbb{R}$, $h^1: V \to \mathbb{R}^m$, $a = (a^1, ..., a^m): T \times V \to \mathbb{R}^m$, and $f: T \times V \times U \to \mathbb{R}^n$ satisfy the following conditions for all $(t, v, u) \in T \times V \times U$:

(d1) the functions $f(t, \cdot, u)$, $h^0$, $h^1$ and $a(t, \cdot)$ admit $c$ as a common bound (for the norms of their values) and a common Lipschitz-constant;

(d2) $f(\cdot, v, u)$ is measurable and both $f(t, \cdot, \cdot)$ and $a(\cdot, \cdot)$ are continuous;

(d3) $y([t_0, \tau]) \subset D$ for every $\tau \in T$ and every absolutely continuous $y: [t_0, \tau] \to V$ satisfying (2.1.1) a.e. in $[t_0, \tau]$ and with $y(t_0) \in A_0$.

Remark. As it is easy to see, we may weaken these assumptions as far as $f$ is concerned by replacing $c$ with an integrable function $\psi: T \to (0, \infty)$, and then choosing as the new independent variable $\theta = \int_{t_0}^t \psi(\tau) \, d\tau$. This transformation will cause the function $\hat{f}(\theta, \cdot, u)$ of the transformed problem to admit 1 as a bound and a Lipschitz-constant and will not affect the other assumptions.

Now let $e_k$ denote the $k$th column of the unit matrix $I$ of appropriate dimension, $S^a(x, r)$ the closed ball with center $x$ and radius $r$, and $L^p(\mathbb{R}^n, \mathbb{R}^a)$ the space of real $b \times a$ matrices. If $\phi = (\phi^1, ..., \phi^b): G \to \mathbb{R}^b$ has a Lipschitz-constant $\hat{c}$, $G$ is an open subset of $\mathbb{R}^a$, $x \in G$, $i \in \{1, 2, ..., b\}$, and $k \in \{1, 2, ..., a\}$, we set $A_{i,k}(x) = [-c, c]$ if $S^a(x, 2e) \subset G$ and $A_{i,k}(x) = \{(2a)^{-1}[\phi^i(\xi + \alpha e_k) - \phi^i(\xi - \alpha e_k)] | \xi - x | \leq e, 0 < \alpha \leq e\}$ if $S^a(x, 2e) \cap G$, and then define $\Delta^e\phi(x)$ as the set of all $b \times a$ matrices $(M_{i,k})$ such that $M_{i,k} \in [\inf A_{i,k}(x), \sup A_{i,k}(x)]$ for all $i$ and $k$. We write $\Delta^e\phi$ for the “partial” $\Delta^e$ operator with respect to the argument in $V$; thus $\Delta^e f(t, v, u)$, $\Delta^e a^i(t, v)$, etc., denote $\Delta^e \phi(v)$, where $\phi$ represents $f(t, \cdot, u)$, $a^i(t, \cdot)$, etc.

Finally, we denote transposition by $T$, treat each element of $\mathbb{R}^a$ as a column vector, and write $M \mathcal{A}$ for $\{MA | A \in \mathcal{A}\}$ if $\mathcal{A}$ is a collection of matrices.

We can now state our first theorem combining existence with necessary conditions.

Theorem 2.2. Assume that the set of admissible relaxed solutions is non-empty. Then there exist a minimizing relaxed solution $\bar{y}$ and a sequence $((y_j, u_j))_{j=1}^\infty$ such that each $y_j: T \to V$ is Lipschitz-continuous, each $u_j$ is a measurable selection of $U^*$, each $(y_j, u_j)$ satisfies the differential equation (1.1), and $\lim_j y_j = \bar{y}$ uniformly. Furthermore, there exist

$l_0 \geq 0$, \quad l_1 \in \mathbb{R}^m$, \quad h^0 \in L^p(\mathbb{R}^n, \mathbb{R})$, \quad h^1 \in L^p(\mathbb{R}^n, \mathbb{R}^m)$,
Borel-measurable \( \bar{a}^i : T \rightarrow \mathcal{L}(\mathbb{R}^n, \mathbb{R}) \) and positive Radon measures \( \omega^i \) on \( T(i = 1, 2, \ldots, m_o) \), Lipschitz-continuous \( Z : T \rightarrow \mathcal{L}(\mathbb{R}^n, \mathbb{R}^n) \), and \( k : T \rightarrow \mathbb{R}^n \) such that

\[
l_0 + |l_1| + \sum_{i=1}^{m_o} \omega^i(T) > 0, \quad Z(t_1) = I; \tag{1}
\]

\( \omega^i \) (\( i = 1, 2, \ldots, m_o \)) is supported on the set \( \{ t \in T \mid a^i(t, \bar{y}(t)) = 0 \} \); \( \bar{h}^i \in \bigcap_{\varepsilon > 0} \Delta \bar{h}^i(\bar{y}(t_i)) \) (\( l = 0, 1 \)); \( \bar{a}^i(t) \in \bigcap_{\varepsilon > 0} \overline{\text{co}} \left( \bigcup_{|\tau-1| \leq \epsilon} \Delta a^i(t, \bar{y}(t)) \omega^i \right) \) a.e. \( (i = 1, 2, \ldots, m_o) \); \( h(t) = \left[ l_0 \bar{h}^0 + l_1 \bar{h}^1 + \sum_{i=1}^{m_o} \int_{[t, t_1]} \bar{a}^i(t) Z(t) Z(t) \right] Z(t) \) (\( t \in T \)); \( \text{(the maximum principle)} \)

\[
(d/dt)(\bar{y}(t), Z(t)) \in \text{co} \bigcap_{\varepsilon > 0} K_\varepsilon(t) \quad \text{a.e. in } T, \tag{6}
\]

where

\[
K_\varepsilon(t) = \text{closure } \{ (f(t, \bar{y}(t), \bar{u}), -W) \mid W \in Z(t) \Delta \varepsilon f(t, \bar{y}(t), \bar{u}), \quad \bar{u} \in U, \}
\]

\[
h(t) \cdot f(t, \bar{y}(t), \bar{u}) = \min_{u \in \mathcal{G}^*(t)} h(t) \cdot f(t, y(t), u));
\]

and

\[
k(t_0) \cdot \bar{y}(t_0) = \min_{\bar{a}_0 \in A_0} k(t_0) \cdot a_0, \quad l_1 \cdot h^1(\bar{y}(t_1)) = \max_{\bar{a}_1 \in A_1} l_1 \cdot a_1. \tag{7}
\]

Our second theorem, generalizing Theorem 2.2, provides a different set of necessary conditions for each representation of \( f(t, \cdot, u), h^0, h^1, \) and \( a^i(t, \cdot) \) as compositions. (See [8, 2.5, p. 471] for a discussion of this matter.) In order to state this theorem, we must use the definition of derivat \( (\mathcal{D}^\varepsilon cont) \) introduced in [8].

Let \( \phi : V \rightarrow \mathbb{R}^b \) be Lipschitz-continuous. We say that the sets \( \mathcal{A}^\varepsilon(\mathcal{V}) \) (\( \varepsilon > 0, \mathcal{V} \in V \)) determine a deriviative container for \( \phi \) if there exist positive integers \( l, k_0, \ldots, k_l \), open sets \( V_i \subset \mathbb{R}^{k_i} (i = 0, 1, \ldots, l), \gamma > 0, \) and Lipschitz-continuous \( \phi_i : V_i \rightarrow V_{i-1} (i = 1, \ldots, l) \), such that \( V_l = V, \ V_0 = \mathbb{R}^b, \ \phi = \phi_1 \circ \phi_2 \circ \cdots \circ \phi_l, \ S^\varepsilon(\phi(V_i), \gamma) \subset V_{i-1} \), and

\[
\mathcal{A}^\varepsilon(\mathcal{V}) = \{ M_1 M_2 \cdots M_l \mid w_i = \mathcal{V}, \ w_i = \phi_{i+1} \circ \cdots \circ \phi_1(\mathcal{V}), \ M_i \in \mathcal{A}^\varepsilon(\phi_i(w_i)) \}.
\]

We say that \( \mathcal{A}^\varepsilon(t, v, u) \) (\( \varepsilon > 0, \ t \in T, \ v \in V, \ u \in U \)) determine a derivat
container for $f$ if there exist positive integers $l, k_0, \ldots, k_l$, open $V_i \subset \mathbb{R}^{k_i}$, $\gamma > 0$, and $f_i : T \times V_i \times U \rightarrow V_{i-1}$ ($1 = 0, \ldots, l$), such that $V_1 = V$, $V_0 = \mathbb{R}^n$, $S^F(f_i(t, V_i, u), \gamma) \subset V_{i-1}$, 

$$f(t, \cdot, u) = f_i(t, \cdot, u) \circ \cdots \circ f_1(t, \cdot, u) \quad (t \in T, u \in U),$$

$A_f(t, v, u)$ is defined as $A^F(v)$ with $\phi_i = f_i(t, \cdot, u)$ and, furthermore, each $f_i(t, \cdot, u)$ is measurable, each $f_i(t, \cdot, \cdot)$ continuous, each $f_i(t, \cdot, u)$ Lipschitz-continuous with a Lipschitz-constant independent of $i, t,$ and $u$, and each $f_i$ bounded. We similarly define a derivate container $A_i^a(t, v)$ for $a^i$ which must be based on a composition

$$a^i(t, \cdot) = \tilde{a}^i(t, \cdot) \circ \cdots \circ \tilde{a}_1(t, \cdot)$$

with each function $\tilde{a}_j(\cdot, \cdot)$ continuous and each $\tilde{a}_j(t, \cdot)$ admitting the same Lipschitz-constant (independent of $i, j,$ and $t$). (We might add, for the sake of clarity, that $l, k_i, V_i$, etc., may be different for each of the functions $f, h^0, h^1, a^i$).

**Theorem 2.3.** Let the conditions of Theorem 2.2 be satisfied and $A_f(t, v, u), A_0^a(v), A_1^a(v), A_{a, i}(t, v)(i = 1, \ldots, m_a)$ determine derivate containers for $f, h^0, h^1, a^i,$ respectively. Then the conclusions of Theorem 2.2 remain valid with $A_{ef}, A^{h^0}, A^{h^1}, A^{a^i}$ replaced by $A_f, A_0^a, A_1^a, A_{a, i}$, respectively.

3. **Proofs**

3.1. **Notation.** If $X$ is a compact metric space, we denote by $\text{frm}(X)$ the real vector space of Radon measures on $X$ with the weak star topology of $C(X)^*$, and by $\text{frm}^+(X)$, respectively, $\text{rpm}(X)$, the subset of $\text{frm}(X)$ whose elements are positive, respectively, probability measures. (We use the conventional term "positive measure" to mean "nonnegative measure.") We write $\mathcal{S}$ for the collection of all measurable functions $\sigma : T \rightarrow \text{rpm}(U)$, set

$$\mathcal{S}^\# = \{\sigma \in \mathcal{S} \mid \sigma(\delta^\sigma(t)) = 1 \text{ a.e. in } T\},$$

and identify each measurable $\rho : T \rightarrow U$ with the element $\sigma \in \mathcal{S}$ such that $\sigma(t)$ is the Dirac measure at $\rho(t)$ for all $t \in T$. We identify each $\sigma \in \mathcal{S}$ with the element

$$\phi \rightarrow \int_{t_0}^{t_1} dt \int \phi(t, u) \sigma(t)(du)$$

of $L^1(\mu, C(U))^*$ and endow $\mathcal{S}$ with the corresponding weak star topology.
We write

$$f(t, y(t), \sigma(t)) = \int f(t, y(t), u) \sigma(t)(du),$$

$\mathcal{D}$ for the Frechet derivative, and $\mathcal{D}_2$ for the partial Frechet derivative with respect to the second argument.

We refer to a function $\rho: \mathbb{R}^i \to \mathbb{R}$ as a mollifier if $\rho(x_1, x_2, ..., x^i) = \prod_{i=1}^{i} \pi(x^i)$ and $\pi$ is a $C^\infty$ function that vanishes outside a finite interval, is nonnegative, symmetric, nonincreasing for nonnegative arguments, and with $\int_{-\infty}^{\infty} \pi(\alpha) \, d\alpha = 1$. The mollifier $\rho$ has a radius $\varepsilon$ if $\pi$ vanishes outside $[-\varepsilon, \varepsilon]$, and a sequence $(\rho_1, \rho_2, ..., \rho_j)$ of mollifiers is a $\delta$-sequence in $\mathbb{R}^l$ if $\rho_j$ has a radius $\varepsilon_j$ and $(\varepsilon_j)$ decreases to 0. We write $\phi * \rho$ for the convolution $x \mapsto \int_S \rho(\alpha) \phi(x - \alpha) \, d\alpha \cdots d\alpha$, where $S$ is the support of $\rho$, $\alpha = (\alpha_1, ..., \alpha^l)$ and $x \in \mathbb{R}^l$.

We denote by $d(x, y)$ ($d(x, A)$) the distance between two points (a point and a set) and by $S^\delta(A, r)$ the set $\{b \mid d(b, A) \leq r\}$.

**Lemma 3.2.** Let $X$ be a compact subset of $\mathbb{R}^n$ and, for $j = 1, 2, ..., \omega_j \in \text{frm}^+(T)$, $\rho_j = (\rho_{j1}, ..., \rho_{jn}): T \to X$, and $\Gamma_\varepsilon$ a mapping from $T$ to the collection of nonempty, closed, and convex subsets of $X$. Assume that

$$\omega_j(T) \leq 1, \quad \lim_j \omega_j = \omega \text{ weakly},$$

$$\Gamma_\varepsilon(t) \subset \Gamma_\varepsilon(t) \quad (t \in T, 0 < \varepsilon < \varepsilon'),$$

$$G(\Gamma_\varepsilon) = \{(t, x) \mid x \in \Gamma_\varepsilon(t)\} \text{ is closed},$$

and for every $\varepsilon > 0$ there exists $j_0(\varepsilon)$ such that $\rho_j$ is a Borel measurable selection of $\Gamma_\varepsilon$ for $j \geq j_0(\varepsilon)$. Then there exist a Borel measurable $\rho: T \to X$ and a sequence $j(1, 2, ...)$ such that

$$\rho(t) \in \bigcap_{\varepsilon > 0} \Gamma_\varepsilon(\tau) \quad \text{for } \omega\text{-a.a. } \tau \in T$$

and

$$\lim_{t \to t_0} \int_{[t, t_1]} \rho_\tau(t) \omega_\tau(d\tau) = \int_{[t, t_1]} \rho(t) \omega(d\tau) \quad \text{for } t = t_0 \text{ and a.a. } t \in T.$$

**Proof.** Step 1. Let $\xi_j \in \text{frm}(T \times X)$ be defined by

$$\int \phi(\tau, \alpha) \xi_j(\alpha) \, d\alpha = \int \phi(\tau, \rho_j(\alpha)) \omega_j(\alpha) \, d\alpha \quad (\phi \in C(T \times X)). \quad (1)$$

We have $\xi_j(T \times X) = \omega_j(T) \leq 1$ and $\xi_j$ is clearly a positive measure. Thus there exist $j(1, 2, ...) \in \text{frm}^+(T \times X)$ such that $\lim_{j \to j} \xi_j \rightarrow \xi$ weakly.
Now let $n_\omega$ be the norm of $L^1(\omega, C(X))$ and $| \cdot |_{\sup}$ the sup norm. We observe that, by (1), for each $\phi \in C(T \times X)$ we have

$$\int |\phi(\tau, x) \xi(d(\tau, x))| \leq \int |\phi(\tau, \cdot)|_{\sup} \xi(d(\tau, x))$$

$$= \lim_{j \to \infty} \int |\phi(\tau, \cdot)|_{\sup} \xi_j(d(\tau, x)) = \lim_{j \to \infty} \int |\phi(\tau, \cdot)|_{\sup} \omega_j(d\tau)$$

$$= \int |\phi(\tau, \cdot)|_{\sup} \omega(d\tau) = n_\omega(\phi).$$

Thus $\zeta$ is a continuous linear functional on the normed vector space $(C(T \times X), n_\omega)$ and, by the Hahn–Banach theorem, $\xi$ can be extended as a continuous linear functional to $L^1(\omega, C(X))$. It follows, by a variant of the Dunford–Pettis theorem [7, IV.1.8, p. 268] (whose proof remains valid when $\mu$ is replaced by any positive Radon measure on $T$), that there exists an $\omega$-measurable $\lambda: T \to \text{frm}(X)$ such that $\text{ess sup}_{t \in T} |\lambda(t)| (X) < \infty$ and

$$\int f(\tau, x) \xi(d(\tau, x)) = \int \omega(d\tau) \int f(\tau, x) \lambda(d\tau)(dx) \quad (f \in L^1(\omega, C(X))). \quad (2)$$

It is clear that $\lambda(\tau)$ is a positive measure for $\omega$-a.a. $\tau \in T$.

Since $\omega(T) \leq 1$, the set $S_\omega$ of all the atoms of $\omega$ is at most denumerable. If $t = t_0$ or $t \in T - S_\omega$ and if $i \in \{1, 2, \ldots, n\}$, then the function

$$(\tau, x) \mapsto \phi(\tau, x_1, \ldots, x_n) = x_i \quad \text{if} \quad \tau \geq t,$$

$$= 0 \quad \text{if} \quad \tau < t,$$

is continuous $\xi$-a.e. and, by [1, 4.5.1(c), p. 196] and (2), $\lim_{j \to \infty} \int \phi(t, x) \xi_j(d(t, x)) = \int \phi(t, x) \xi(d(t, x)) = \int \omega(d\tau) \int \phi(\tau, x) \lambda(t)(dx)$; hence,

$$\lim_{j \to \infty} \int_{[t, t_1]} \rho_j(\tau) \omega_j(d\tau) = \int_{[t, t_1]} \omega(d\tau) \int x^i \lambda(\tau)(dx). \quad (3)$$

**Step 2.** We shall next show that $\lambda(\tau)(G(\tau_\epsilon)) = 1$ $\omega$-a.e. for every $\epsilon > 0$. Let $t \in \{1, 2, \ldots\}$ and $c_\epsilon : T \times X \to [0, 1]$ be continuous and such that $c_\epsilon(\tau, x) = 1$ for $(\tau, x) \in G(\tau_\epsilon)$ and $c_\epsilon(\tau, x) = 0$ if $d((\tau, x), G(\tau_\epsilon)) > 1/i$. Then, for each $\phi \in C(T)$,

$$\int \phi(\tau) \omega(d\tau) = \lim_{j \to \infty} \int \phi(\tau) c_\epsilon(\tau, x) \xi_j(d(\tau, x)) = \int \phi(\tau) c_\epsilon(\tau, x) \xi(d(\tau, x))$$

$$= \int \psi(\tau) \omega(d\tau) \int c_\epsilon(\tau, x) \lambda(\tau)(dx).$$
Since $c_i$ converges pointwise to the characteristic function $\chi_\varepsilon$ of the closed set $G(\Gamma_\varepsilon)$, it follows that
\[
\int \phi(\tau) \omega(d\tau) = \int \phi(\tau) \omega(d\tau) \int \chi_\varepsilon(\tau, x) \lambda(\tau)(dx);
\]
hence,
\[
\int \chi_\varepsilon(\tau, x) \lambda(\tau)(dx) = \lambda(\tau)(\Gamma_\varepsilon(\tau)) = 1 \quad \omega\text{-a.e.}
\]

**Step 3.** By (3), we have for $t = t_0$ or $t \in T - S_\omega$,
\[
\lim_{j \to \infty} \int_{[t, t_1]} p_j(\tau) \omega_j(d\tau) = \int_{[t, t_1]} \omega(d\tau) \int x\lambda(\tau)(dx).
\]
For each $\varepsilon > 0$ and $\omega$-a.a. $\tau \in T$, $\Gamma_\varepsilon(\tau)$ is closed and convex and $\lambda(\tau)$ is a probability measure with support in $\Gamma_\varepsilon(\tau)$. It follows, setting
\[
\tilde{p}(\tau) = \int x\lambda(\tau)(dx) \quad (\tau \in T),
\]
and selecting a Borel-measurable $\bar{p}$ in the $\omega$-equivalence class of $\tilde{p}$, that $\bar{p}(\tau) \in \Gamma_\varepsilon(\tau)$ $\omega$-a.e. and
\[
\lim_{j \to \infty} \int_{[t, t_1]} p_j(\tau) \omega_j(d\tau) = \int_{[t, t_1]} \omega(d\tau) \int x\lambda(\tau)(dx) \quad \text{for } t = t_0 \text{ and } \text{a.a. } t \in T.
\]
Since $\bigcap_{\varepsilon > 0} \Gamma_\varepsilon(\tau) = \bigcap_{j=1}^{\infty} \Gamma_1/\varepsilon(\tau)$ for all $\tau \in T$, we also have
\[
\bar{p}(\tau) \in \bigcap_{\varepsilon > 0} \Gamma_\varepsilon(\tau) \quad \text{for } \omega\text{-a.a. } \tau \in T. \quad \text{Q.E.D.}
\]

**Lemma 3.3.** Let $\omega^i_j \in \text{frm}^+(T)(i = 1, \ldots, m_2; j = 1, 2, \ldots)$, $\sum_{i=1}^{m_2} \omega^i_j(T) \leq 1$, and $\lim_{j} \omega^i_j = \omega^i$ weakly. Let $Y_j : T \to \mathcal{L}(\mathbb{R}^n, \mathbb{R})$ be continuous, $\lim_{j} Y_j = Y$ uniformly, and $Y(t)$ be invertible for each $t \in T$. Finally, let $p_{i,j} : T \to [-c, c]^n$ be continuous and assume that for every $\varepsilon > 0$ there exists $j_0(\varepsilon)$ such that
\[
p_{i,j}(t) \in A^i_{\varepsilon,t}(t, \bar{y}(t)) \quad (t \in T; i = 1, \ldots, m_2; j \geq j_0(\varepsilon)).
\]
Then there exist Borel-measurable $p_i : T \to \mathcal{L}(\mathbb{R}^n, \mathbb{R})$ and $J \subset (1, 2, \ldots)$ such that
\[
p_i(t) \in \bigcap_{\varepsilon > 0} \bigcup_{|\tau-t| < \varepsilon} A^i_{\varepsilon,t}(t, \bar{y}(t)) \quad \text{for } \omega^i\text{-a.a. } t \in T, \quad i = 1, 2, \ldots, m_2, \quad (1)
\]
and
\[
\lim_{j \to \infty} \sum_{i=1}^{m_2} \int_{[t, t_1]} p_{i,j}(\tau) Y_j(\tau) \omega^i_j(d\tau) = \sum_{i=1}^{m_2} \int_{[t, t_1]} p_i(\tau) Y(\tau) \omega^i(d\tau) \quad (2)
\]
for $t \to t_0$ and $t \in T$. 

Proof. Let \( X = [-\epsilon, \epsilon]^n, i \in \{1, 2, \ldots, m_2\} \), and
\[
\Omega^i_\epsilon(t) = \overline{\text{co}}(wY(t)| w \in \mathcal{A}^*_\epsilon(i, \bar{y}(t))), \quad |\tau - t| \leq \epsilon \quad (t \in T, \epsilon > 0).
\]
We also set
\[
G(\Omega^i_\epsilon) = \{(t, x)| x \in \Omega^i_\epsilon(t)\}
\]
and
\[
\Gamma^i_\epsilon(t) = \text{co}\{x | (t, x) \in G(\Omega^i_\epsilon)\}.
\]
If we identify \( \mathcal{L}(\mathbb{R}^n, \mathbb{R}) \) with \( \mathbb{R}^n \), then we can verify that \( \Gamma^i_\epsilon \) satisfy the assumptions (about \( \Gamma^i_\epsilon \)) of Lemma 3.2 and, for all \( \epsilon > 0 \) and \( t \in T \), we have
\[
\Omega^i_\epsilon(t') \subseteq \Omega^i_{2\epsilon}(t) \quad \text{provided } |t - t'| \leq \epsilon \quad \text{and} \quad |\bar{y}(t) - \bar{y}(t')| \leq \epsilon.
\]
Thus
\[
\Omega^i_\epsilon(t) \subseteq \Gamma^i_\epsilon(t) \subseteq \Omega^i_{2\epsilon}(t);
\]
hence
\[
\bigcap_{\epsilon > 0} \Omega^i_\epsilon(t) = \bigcap_{\epsilon > 0} \Gamma^i_\epsilon(t).
\]
It follows now from Lemma 3.2 that there exist an \( \omega^i \)-measurable \( q_i : T \rightarrow \mathcal{L}(\mathbb{R}^n, \mathbb{R}) \) and \( J_i \subseteq (1, 2, \ldots) \) such that
\[
q_i(\tau) \in \bigcap_{\epsilon > 0} \Gamma^i_\epsilon(\tau) - \bigcap_{\epsilon > 0} \Omega^i_\epsilon(\tau) \quad \omega^i-\text{a.e.} \quad (3)
\]
and
\[
\lim_{j \in J_i, t_1 \in [t, t_1]} p_i, j(\tau) Y(\tau) \omega^i_1(d\tau) = \int_{[t, t_1]} q_i(\tau) \omega^i(d\tau)
\]
for \( t = t_0 \) and a.a. \( t \in T \).

We may clearly apply Lemma 3.2 consecutively for \( i = 1, 2, \ldots, m_2 \), each time choosing \( J_i \) as a subsequence of \( J_{i-1} \) [with \( J_0 = (1, 2, \ldots) \)]. If we set \( J = J_{m_2} \) and \( p_i(t) = q_i(t) Y(t)^{-1} \quad (t \in T; i = 1, 2, \ldots, m_2) \), and recall that the \( p_{ij} \) and \( \omega^i \) are all uniformly bounded, then relations (1) and (2) follow from (3) and (4), respectively.

We may assume that each \( p_i \) is Borel-measurable by replacing it with an appropriate element of its \( \omega^i \)-equivalence class. Q.E.D.

3.4 Proof of Theorems 2.2 and 2.3. Step 1. Since Theorem 2.2 is a special case of Theorem 2.3, we shall only deal with the latter, defining statements (1)-(7) of Theorem 2.3 to be the corresponding ones of Theorem 2.2 with \( \Delta^c \) replaced by the appropriate \( \mathcal{A}^c \).
We first observe that, by [7, VI.3.2, p. 370], an absolutely continuous function \( y \) satisfies
\[
\dot{y}(t) \in \text{co} \{ f(t, y(t), \bar{U}^*(t)) \} \quad \text{a.e. in } T
\]
if and only if there exists \( \sigma \in \mathcal{P}^\# \) such that
\[
\dot{y}(t) = f(t, y(t), \sigma(t)) \quad \text{a.e. in } T.
\]
Thus every admissible relaxed solution \( y \) has an associated relaxed control \( \sigma \in \mathcal{P}^\# \), and we shall also refer to such a couple \( (y, \sigma) \) as an admissible relaxed solution. It follows easily from our assumptions that the equation
\[
y(t) = a + \int_0^t f(\tau, y(\tau), \sigma(\tau)) \, d\tau \quad (t \in T)
\]
has a unique solution \( y(\sigma, a) \) for all \( (\sigma, a) \in \mathcal{P}^\# \times A_0 \), and that \( y(\sigma, a)(T) \subset D \).

Step 2. Now let \( l, k_0, \ldots, k_l, V_0, \ldots, V_l, \gamma \), and
\[
f(t, \cdot, u) = f_i(t, \cdot, u) \circ \cdots \circ f_1(t, \cdot, u) \quad (t \in T, u \in U)
\]
provide the representation of \( f \) that yields \( \Lambda f(t, v, u) \). For each \( i = 1, 2, \ldots, l \), we select a \( \delta \)-sequence \( (p_i^j)_{j=1}^\infty \) in \( \mathbb{R}^{k_i} \) such that each \( p_i^j \) has a radius \( \epsilon_j < \gamma/2 \). We then set, for all \( (t, u) \in T \times U \),
\[
f_{i,0}(t, \cdot, u) = f_i(t, \cdot, u) \circ p_i\circ \cdots \circ f_1(t, \cdot, u),
\]
We also apply a similar procedure to \( h_0, h_1, \) and \( a^i(t, \cdot) \), using for each the representation that gives rise to \( \Lambda h^0, \Lambda a^0, \Lambda a^i, \) respectively, and choosing the same radii \( \epsilon_i \) and the same number \( \gamma \) for each. We thus obtain sequences \( h_{i,0}, h_{i,1}, a^i(t, \cdot) (j = 1, 2, \ldots) \). It is easy to verify (see, e.g., [8, 3.2, p. 51]) that
\[
(a) \text{ all the functions } f_i(t, \cdot, u), h_{i,0}, h_{i,1}, a^i(t, \cdot) \text{ have a common bound and a common Lipschitz constant which we shall continue to denote by } c,
\]
(b) \( f_i(\cdot, v, u) \) and \( \mathcal{D} f_i(\cdot, v, u) \) are measurable,
(c) \( f_i(t, \cdot, \cdot), \mathcal{D} f_i(t, \cdot, \cdot), a^i(\cdot, \cdot), \) and \( \mathcal{D} a^i(\cdot, \cdot) \) are continuous, and
(d) for each \( \epsilon > 0 \), there exists \( j_0(\epsilon) \) such that, for all
\[
(t, v, u) \in T \times S^T(D, \gamma/2) \times U \quad \text{and} \quad j \geq j_0(\epsilon),
\]
\[
\mathcal{D} f_j(t, v, u) \in \Lambda f(t, v, u),
\]
\[
\mathcal{D} h^k_j(v) \in \Lambda h^k_j(v) \quad (k = 0, 1),
\]
\[
\mathcal{D} a^i_j(t, v) \in \Lambda a^i_j(t, v) \quad (i = 1, 2, \ldots, m_2),
\]
\begin{equation}
\tag{1}
\end{equation}
and \( f(t, v, u), h^0(v), h^1(v), a_j(t, v) \) are within a distance at most \( \epsilon \) from \( f(t, v, u), h^0(v), h^1(v), a(t, v) \), respectively.

**Step 3.** We now consider the differential equation

\[
y_j(t) = a + \int_{t_0}^{t} f_j(\tau, y_j(\tau), \sigma(\tau)) \, d\tau \quad (t \in T).
\]

Our assumptions and the properties of \( f_j \) listed above ensure that there exist an integer \( j_0 \) and a constant \( c_1 \) (depending only on \( c \) and \( t_1 - t_0 \)) such that this equation has a unique solution \( y_j(\sigma, a) \) for all \((\sigma, a) \in \mathcal{P}^n \times A_0 \) and \( j \geq j_0 \), and

\[
| y_j(\sigma, a)(t) - y(\sigma, a)(t) | \leq c_1 \epsilon_j \quad (t \in T).
\]

Thus there exists a sequence \((\epsilon_j^1)\) decreasing to 0 such that \( c_1 \epsilon_j^1 \leq \gamma/2 \) and \( h^1_j( y_j(\sigma, a)(t_1)) \in S^p(A_1, \epsilon_j^1), \quad a_j^i( y_j(\sigma, a)(t)) - \epsilon_j^1 \leq 0 \quad (i = 1, 2, \ldots, m_a) \)

for all \( j \geq j_0 \) and \( t \in T \) provided that \((y(\sigma, a), \sigma)\) is an admissible relaxed solution.

We now consider for each \( j \geq j_0 \) a new optimal control problem \( P_j \) which differs from our basic problem \( P \) in that \( f, h^0, h^1, a, A_1 \) are replaced by \( f_j, h_j^0, h_j^1, a_j^i - \epsilon_j^1, S^p(A_1, \epsilon_j^1) \), respectively. Since \( P \) is assumed to have an admissible relaxed solution \((\tilde{y}_j, \tilde{\sigma})\), it follows that \( P_j \) has an admissible relaxed solution \((y_j(\sigma, a)(t_1)), \sigma(\tau)\). We may now apply a known existence theorem and necessary conditions \([7, V.1.1.1 \text{ and } V.1.2.3, \text{ pp. 348, 357–358}] \) and conclude that \( P_j \) admits a minimizing relaxed solution \((y_j(\sigma, a), \sigma)\) and there exist \( l_0^i \geq 0, l_i^1 \in \mathbb{R}^m, \omega_1^i, \ldots, \omega_m^i \in \operatorname{frm}^+(T), Z_j : T \to \mathcal{L}(\mathbb{R}^n, \mathbb{R}^n) \) and \( k_j : T \to \mathbb{R}^n \) such that

\[
l_0^i \cdot |l_1^i| - \sum_{i=1}^{m_2} \omega_i^i(T) = 1, \quad \omega_i^i \text{ is supported on } \{ t \in T \mid a_j^i(t, \tilde{y}(t)) = \epsilon_j^1 \},
\]

\[
Z_j(t) = I + \int_{\tau}^{t_1} Z_j(\tau) \mathcal{D}_2 f_j(\tau, \tilde{y}_j(\tau), \sigma_j(\tau)) \, d\tau \quad (t \in T),
\]

\[
k_j(t)^T = \left[ \int_{\tau}^{t_1} \mathcal{D}_2 h_j^1(\tilde{y}_j(t_1)) + \sum_{i=1}^{m_2} \int_{\tau}^{t_1} \mathcal{D}_2 a_j^i(\tau, \tilde{y}_j(\tau)) Z_j(\tau)^{-1} \omega_j^i(\tau) \, d\tau \right] \cdot Z(t) \quad (t \in T),
\]

\[
k_j(t) \cdot f_j(t, \tilde{y}_j(t), \sigma_j(t)) = \min_{u \in \mathcal{D}(\sigma_j(t))} k_j(t) \cdot f_j(t, \tilde{y}_j(t), u) \quad \text{a.e. in } T,
\]

\[
k_j(t_0) \cdot y_j(t_0) = \min_{a_j \in A_0} k_j(t_0) \cdot a_j^0, \quad l_1^i \cdot h^1(\tilde{y}_j(t_1)) = \max_{a_j \in A_0} l_1^i a_j.
\]
We observe that the functions $\bar{y}_j$ and $Z_j$ ($j \geq j_0$) have a common Lipschitz-constant and a common bound, and recall [7, IV.3.11, p. 287] that $\mathcal{F}$ is sequentially compact. Thus, in view of (2), we may determine an increasing sequence $J$ of positive integers, $l_0 \geq 0$, $l_1 \in \mathbb{Z}^n$, $\bar{d} \in \mathcal{F}$, Lipschitz-continuous $\bar{y} : T \rightarrow D$ and $Z : T \rightarrow \mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)$, and $\omega^i \in \text{frm}^+(T)$ ($i = 1, 2, \ldots, m_a$) such that

$$
\lim_{j \in J} l^i_k = l^i_k \quad (k = 0, 1), \quad \lim_{j \in J} \sigma^k = \bar{d}, \quad \lim_{j \in J} \omega^i = \omega^i \text{ weakly},
$$

and

$$
\lim_{j \in J} \bar{y}_j = \bar{y} \quad \text{and} \quad \lim_{j \in J} Z_j = Z \text{ uniformly on } T.
$$

Furthermore, $Z(t)^{-1}$ exists for all $t \in T$ because, as a consequence of (3), $Z(t)^{-1}$ are uniformly bounded. We also verify that each $\omega^i$ is supported on the set $\{t \in T \mid a^i(t, \bar{y}(t)) = 0\}$. Thus relations (1) and (2) of Theorem 2.3 are satisfied.

**Step 4.** We shall assume that $J = (1, 2, \ldots)$ by appropriately relabeling the indices if necessary. We have

$$
\bar{y}(t) = \lim_j \bar{y}_j(t) = \lim_j \left[ \bar{y}_j(t_0) + \int_{t_0}^t f_j(\tau, \bar{y}_j(\tau), \sigma_j(\tau)) \, d\tau \right]
$$

and therefore, by [7, IV.2.9., p. 278],

$$
\bar{y}(t) = \bar{y}(t_0) + \int_{t_0}^t f(\tau, \bar{y}(\tau), \bar{d}(\tau)) \, d\tau \quad (t \in T).
$$

Furthermore, we clearly have $\bar{y}(t_0) = \lim_j \bar{y}_j(t_0) \in A_0$, and similarly,

$$
h^i(\bar{y}(t_0)) \in A_1, \quad a^i(t, \bar{y}(t)) \leq 0 \quad (t \in T, \quad i = 1, \ldots, m_a).
$$

Thus ($\bar{y}, \bar{d}$) is an admissible relaxed solution of $P$. If ($\bar{y}, \bar{d}$) is another such solution, then, as we have seen before, ($y_j(\bar{d}, \bar{y}(t_0)), \bar{d}$) is an admissible relaxed solution of $P^j$, and therefore,

$$
h^0(\bar{y}(t)) = \lim_j h^0(\bar{y}_j(t)) \leq \lim_j h^0(y_j(\bar{d}, \bar{y}(t_0))(t)) = h^0(\bar{y}(t_0)).
$$

Thus ($\bar{y}, \bar{d}$) is a minimizing relaxed solution of $P$. By [7, VI.1.3, p. 350], there exists a sequence $((y_j, u_j))_{i=1}^\infty$ as described in the theorem.

**Step 5.** Let

$$
Y_j(t) = Z_j(t)^{-1}, \quad p_{i,j}(t) = \mathcal{A}_j(t, \bar{y}_j(t)) \quad (t \in T; \quad i = 1, 2, \ldots, m_a; \quad j = 1, 2, \ldots).
$$
Then it follows from (1), (7), and Lemma 3.3 that there exist $J_1 \subset \{1, 2, \ldots\}$ and Borel-measurable $\tilde{a}^i: T \to \mathcal{P}((\mathbb{R}^n, \mathbb{R})$ that satisfy relation (4) of Theorem 2.3, and such that

$$\lim_{j \to 1} \sum_{i=1}^{m_1} \int_{[t, t_1]} \mathcal{P} \tilde{a}^i(\tau, \bar{y}(\tau)) Z(\tau)^{-1} \omega_j^i(d\tau) = \sum_{i=1}^{m_1} \int_{[t, t_1]} \tilde{a}^i(\tau) Z(\tau)^{-1} \omega^i(d\tau) \tag{9}$$

for $t = t_0$ and a.a. $t \in T$.

We may choose $J_1$ in such a manner that the bounded sequences $(\mathcal{P} \tilde{a}^i(\bar{y}(t_1)))_{0, 1}$ for $l = 0, 1$ converge to limits $\tilde{k}^0$ and $\tilde{k}^1$, respectively, and, by (1), these limits satisfy relation (3) of Theorem 2.3. Furthermore, by (1), (4), (6), and (9), there exists $k: T \to \mathbb{R}^n$ such that

$$\lim_{j \to 1} k_j(t) = k(t) \quad \text{for } t = t_0 \text{ and a.a. } t \in T, \tag{10}$$

and relations (5) and (7) of Theorem 2.3 are satisfied.

**Step 6.** It remains to prove the "maximum principle" (6) of Theorem 2.3. It follows from (5) that for every $\sigma \in \mathcal{S}^e$, we have

$$\int_{t_0}^t k_j(\tau) \cdot f_j(\tau, \bar{y}(\tau), \sigma(\tau)) d\tau \leq \int_{t_0}^t k(\tau) \cdot f(\tau, \bar{y}(\tau), \sigma(\tau)) d\tau \quad (t \in T, j = 1, 2, \ldots).$$

It follows now from (10) and [7, IV.2.9, p. 278] that

$$\int_{t_0}^t k(\tau) \cdot f(\tau, \bar{y}(\tau), \sigma(\tau)) d\tau \leq \int_{t_0}^t k(\tau) \cdot f(\tau, \bar{y}(\tau), \sigma(\tau)) d\tau \quad (t \in T). \tag{11}$$

By [7, IV.3.2, p. 381], there exists an at most denumerable set $\mathcal{Y}_\infty = \{\rho_1, \rho_2, \ldots\}$ of measurable selections of $U^e$ such that $\{\rho_1(t), \rho_2(t), \ldots\}$ is dense in $U^e(t)$ for a.a. $t \in T$. If we set in (11)

$$\sigma(t) = \rho_i(t) \quad (t \in E), \quad \sigma(t) = \tilde{\sigma}(t) \quad (t \notin E)$$

for various choices of $i \in \{1, 2, \ldots\}$ and measurable sets $E \subset T$, then we can deduce easily that

$$k(t) \cdot f(t, \bar{y}(t), \sigma(t)) = \min_{u \in U^e(t)} k(t) \cdot f(t, \bar{y}(t), u) \quad \text{a.e. in } T. \tag{12}$$

Since $\sigma_j(t)$ is a probability measure and

$$\dot{y}_j(t) = f(t, y_j(t), \sigma_j(t)) \quad \text{a.e. in } T,$$
it follows from relation (3) that

\[(y_j(t), Z_j(t)) \in \overline{co}\{(f_j(t, \bar{y}_j(t), u), -Z_j(t) \delta f_j(t, \bar{y}_j(t), u)) | u \in U\}\]

a.e. in \(T\), say for \(t \in T'\). Now let \(0 < \epsilon < \gamma/4\). Then, by (1) and (7),

\[(\bar{y}_j(t), Z_j(t)) \in \overline{co}\{(f_j(t, \bar{y}_j(t), u), -W) | u \in U, W \in Z_j(t) \cdot A^\text{p}_j(t, \bar{y}_j(t), u)\}\]

\((t \in T')\),

for all sufficiently large \(j\). This implies, in turn, that there exists \(j_1\) such that

\[(\bar{y}_j(t), Z_j(t)) \in S^\text{c}(A^\epsilon_j(t), \epsilon) \quad (t \in T', j \geq j_1),\]

(13)

where

\[A^\epsilon_j(t) = \overline{co}\{(f(t, \bar{y}(t), u), -W) | u \in U, W \in Z(t) \cdot A^\text{p}_j(t, \bar{y}(t), u)\}.

We can easily show (using arguments similar to those of [8, 3.6, p. 55] with the dense denumerable subset of \(U\) replaced by the set \(\{\rho(t) | \rho \in \mathcal{U}_\infty\}\), where \(\mathcal{U}_\infty\) is defined as above) that the mapping \(t \rightarrow A^\epsilon_j(t)\) is measurable. It follows then, by a theorem of Castaing [2, 9.2, p. 88] (or as in [8, 3.7, Step 2, p. 58]) that

\[(\bar{y}(t), Z(t)) \in S^\text{c}(A^\epsilon(t), \epsilon) \quad \text{a.e. in } T.\]

(14)

Relation (6) of Theorem 2.3 now follows from (8), (12), (14), and [8, 3.5, p. 54].

Q.E.D.

REFERENCES