Second-Order Necessary Conditions for Differential Inclusion Problems

H. Zheng

Department of Mathematics, University of British Columbia, Vancouver, British Columbia, Canada V6T 1Y4

Communicated by F. H. Clarke

Abstract. We study second-order necessary conditions for optimality in the unbounded differential inclusion control problem and recover the accessory problem in optimal control theory.

Key Words. Optimal control, Unbounded differential inclusion, Second-order necessary condition, Second variation.

AMS Classification. 49K24.

1. Introduction

Consider the following differential inclusion control problem:

\[
(P_{D}) \quad \text{minimize } g(x(T)) \quad \text{subject to } x'(t) \in F(t, x(t)), \quad x(0) \in C_0, \quad x(T) \in C_1.
\]

First-order necessary conditions for an arc \( \bar{x} \) to be optimal in \( P_D \) are well known (see [4] and [7]). Second-order necessary conditions may be developed in different ways such as variations of extremals for the classical Lagrangian problem, the implicit function theorem for the optimal control problem [11], etc. Since the second-order theory in finite-dimensional optimization has been well developed by Rockafellar and others (see [2] and [10]), another possible way to get second-order conditions in \( P_D \) is to reduce it to a finite-dimensional minimization problem and then apply the known optimality conditions, for example, those in [2] and [10]. We follow this approach here.
Problem \((P_D)\) is equivalent to minimizing the extended real-valued function \(G(x)\) over all \(x\) in \(\mathbb{R}^n\), where \(G\) is defined by
\[
G(x) := g(x) + \Psi_{R(T)}(x) + \Psi_C(x).
\]
Here \(R(T)\) is the reachable set at time \(T\) to the differential inclusion (see [1])
\[
\begin{aligned}
(D) \quad \begin{cases}
    x'(t) \in F(t, x(t)) & \text{a.e. } t \in [0, T], \\
    x(0) \in C_0,
\end{cases}
\end{aligned}
\]
and \(\Psi_S(x)\) denotes the indicator of a subset \(S\) in \(\mathbb{R}^n\), defined by
\[
\Psi_S(x) = 0 \quad \text{when } x \in S \quad \text{and} \quad \Psi_S(x) = +\infty \quad \text{when } x \notin S.
\]
If an arc \(\bar{x}\) solves problem \((P_D)\), then its endpoint \(\bar{x}(T)\) minimizes \(G\). Hence the optimality conditions in [2] tell us that some generalized directional derivatives of \(G\) at \(\bar{x}(T)\) must be nonnegative. To express these derivatives in terms of those of \(g\), \(\Psi_{R(T)}\), and \(\Psi_C\), we must check a constraint qualification, so we are naturally led to study the Clarke normal cone to the reachable set \(R(T)\) at the optimal endpoint. It seems there is no easy way to do this.

In this paper we first study the Clarke normal cone to the reachable set via proximal analysis and get necessary conditions for optimality in \((P_D)\) with bounded differential inclusions. Then we discuss unbounded differential inclusions with the reduction method in [8]. Finally we apply the results to the optimal control problem and recover the accessory problem in [12].

2. Background Material

Let \(S\) be a closed subset of \(\mathbb{R}^n\) and let \(\Gamma\) be a set-valued function from \(\mathbb{R}^n\) to \(\mathbb{R}^n\). Given \(s \in S\) and \((x, u) \in \text{gph } \Gamma\), recall the following definitions:

- The proximal normal cone to \(S\) at \(s\) given by
  \[
P_{\partial S}(s) := \{\xi \in \mathbb{R}^n | \exists M > 0, \text{ s.t. } \langle \xi, \tilde{s} - s \rangle \leq M|\tilde{s} - s|^2, \forall \tilde{s} \in S\}.
  \]

- The limiting proximal normal cone to \(S\) at \(s\) is given by
  \[
  L_{\partial S}(s) = \left\{ \xi = \lim_{k \to \infty} \xi_k | \xi_k \in P_{\partial S}(s_k), s_k \xrightarrow{\ast} s \right\}.
  \]

- The Clarke normal cone to \(S\) at \(s\) is given by [4, Theorem 2.5.6]
  \[
  N_{\partial S}(s) = \partial S \setminus L_{\partial S}(s).
  \]

- The derivative cone to \(S\) at \(s\) is given by
  \[
  D_{\partial S}(s) := \{\xi \in \mathbb{R}^n | \forall t_k \downarrow 0, \exists \varepsilon_k \to \xi, \text{ s.t. } s + t_k \varepsilon_k \in S\}.
  \]
The second-order derivative set to $S$ at $s$ relative to $\xi \in D_2(s)$ is

$$D_2^2(s, \xi) := \{\omega \in \mathbb{R}^n | \forall t_k \downarrow 0, \exists \omega_k \to \omega, \text{s.t. } s + t_k \xi + t_k^2 \omega_k \in S\}.$$  

The first-order derivative of $\Gamma$ at $(x, u)$ is a set-valued function $d\Gamma(x, u)$, defined by

$$\text{gph } d\Gamma(x, u) = D_{\text{gph}} \Gamma(x, u).$$

The second-order derivative of $\Gamma$ at $(x, u)$ relative to $(y, v) \in D_{\text{gph}} \Gamma(x, u)$ is a set-valued function $d^2\Gamma(x, u; y, v)$, defined by

$$\text{gph } d^2\Gamma(x, u; y, v) = D_{\text{gph}}^2 \Gamma(x, u; y, v).$$

The first three concepts are described in [4] and [5]; the last four are discussed in [2] under the name of adjacent cone and adjacent derivative. The term "derivative cone" was first used by Rockafellar [9]. The main results in this paper are based on the following necessary conditions for general optimization problems in finite-dimensional spaces. In the statement we use the notation

$$D_1 g(x, y) = \lim \inf_{h \to 0, y' \to y} \frac{g(x + hy') - g(x)}{h},$$

$$D_2^2 g(x, y, z) = \lim \inf_{h \to 0, z' \to z} \frac{g(x + hy + h^2z') - g(x) - hD_1 g(x, y) - h^2D_2 g(x, y, z)}{h^2}.$$ 

Notice that when $g$ is of class $C^2$,

$$D_1 g(x, y) = g'(x)^T y,$$

$$D_2^2 g(x, y, z) = g'(x)^T z + \frac{1}{2} z' g''(x) y.$$ 

**Theorem 1.** Let $g: \mathbb{R}^n \to \mathbb{R}$ be Lipschitzian in some open set containing $\hat{x}$, and let $S_1, S_2$ be nonempty closed subsets of $\mathbb{R}^n$ containing $\hat{x}$. If $\hat{x}$ solves the following minimization problem,

$$(P) \quad \text{minimize } g(x) \quad \text{over all } x \in S_1 \cap S_2,$$

and also satisfies the constraint qualification

$$N_{S_1}(\hat{x}) \cap (-N_{S_2}(\hat{x})) = \{0\},$$

then we have the first-order necessary condition:

$$D_1 g(\hat{x}, y) \geq 0, \quad \forall y \in D_{S_1}(\hat{x}) \cap D_{S_2}(\hat{x}).$$

Furthermore, if equality holds for some $\hat{y}$, then we have the second-order necessary condition:

$$D_2^2 g(\hat{x}, \hat{y}, z) \geq 0, \quad \forall z \in D_{S_1}^2(\hat{x}, \hat{y}) \cap D_{S_2}^2(\hat{x}, \hat{y}).$$
Proof. We apply some results in [2]. Let

\[ G(x) := g(x) + \Psi_{s_{1}}(x) + \Psi_{s_{2}}(x). \]

Then \( \hat{x} \) is a minimum point of \( G \), so by the Fermat rule [2, Theorem 6.1.9] we have \( D_{1} G(\hat{x}, y) \geq 0 \) for all \( y \) in \( R^{n} \). Since \( g \) is Lipschitzian around \( \hat{x} \) and \( \hat{x} \) satisfies the constraint qualification, by Theorem 6.3.1, Proposition 6.2.4, and Corollary 4.3.5 of [2], we have

\[ D_{1} G(\hat{x}, y) = D_{1} g(\hat{x}, y) + \Psi_{D_{s_{1}}(\hat{x})}^{}(y) + \Psi_{D_{s_{2}}(\hat{x})}^{}(y). \]

The first-order necessary condition follows.

The second-order condition is obtained by combining Proposition 6.6.3 and Theorems 6.6.2 and 4.7.4 from [2]. \( \square \)

3. Hypotheses

A tube about an arc \( x \) is a relatively open subset \( \Omega \) of \([0, T] \times R^{n}\) containing the graph of \( x \). An arc \( x \) is admissible for \((P_{D})\) if it is a trajectory of \( F \) and satisfies both end constraints. An admissible arc \( x \) is a local solution to \((P_{D})\) if there is some tube \( \Omega \) about \( x \) such that any other admissible arc \( y \) lying in \( \Omega \) satisfies \( g(y(T)) \geq g(x(T)) \).

From now on let \( \bar{x} \) be a given local solution to \((P_{D})\) relative to a tube \( \Omega \). The following hypotheses about \((P_{D})\) are assumed throughout the paper:

(H1) On the tube \( \Omega \) the multifunction \( F \) has nonempty, compact, convex images. A nonnegative integrable function \( k \) exists such that

\[ F(t, x) \subseteq k(t) \text{ cl } B, \quad \forall (t, x) \in \Omega; \]

\[ F(t, y) \subseteq F(t, x) + k(t)\|y - x\| \text{ cl } B, \quad \forall (t, x), (t, y) \in \Omega, \]

where \( B \) is the open unit ball in \( R^{n} \).

(H2) The multifunction \( F \) is measurable with respect to the \( \sigma \)-field \( \mathcal{L} \times \mathcal{B} \) generated by products of Lebesgue subsets of \([0, T]\) with Borel subsets of \( R^{n} \).

(H3) The integrable function \( k \) above can be chosen so that the following additional condition is satisfied: for every arc \( y \) satisfying

\[ (y(t), y'(t)) \in D_{\text{gph} F(t, \cdot)}(\bar{x}(t), \bar{x}'(t)), \]

a constant \( \alpha_{0} > 0 \) exists such that

\[ d(\bar{x}'(t) + \alpha y'(t), F(t, \bar{x}(t) + \alpha y(t))) \leq \alpha^{2} k(t) \]

for all \( 0 \leq \alpha \leq \alpha_{0} \).

(H4) The subsets \( C_{0}, C_{1} \) of \( R^{n} \) are nonempty and closed.

(H5) The function \( g: R^{n} \to R \) is Lipschitzian near \( \bar{x}(T) \).
Here the assumption \((H3)\) is used in deriving a second-order approximation to the reachable set. In some sense it means "the derivative of \(F\)" is Lipschitz on the tube \(\Omega\), as we will see later.

**Remark.** Without loss of generality, we may suppose that the domain of \(F\) is \(\Omega\). Otherwise we could define a new multifunction \(\tilde{F}\) by restricting \(F\) to \(\Omega\). The arc \(\tilde{x}\) would remain a local solution to the associated problem \((\tilde{P}_D)\), for which we are about to provide necessary conditions. Since these necessary conditions involve only the local behavior of \(\tilde{F}\) along \(\tilde{x}\), their conclusions for \(\tilde{F}\) are identical to the desired results for \(F\). Similarly, we may suppose that the sets \(C_0\) and \(C_1\) are compact.

4. The Clarke Normal Cone to the Reachable Set

Under the above assumptions, we know the reachable set \(R(T)\) is a closed set [5]. In order to apply Theorem 1 to the function \(G\) described in the introduction, we must first verify the constraint qualification. This requires that we describe the Clarke normal cone to the reachable set \(R(T)\) at \(\tilde{x}(T)\).

**Definition.** A trajectory \(x\) for \((D)\) is said to be proper if, for any sequence \(x_k\) of trajectories for \((D)\) having the property that \(x_k(T) \rightarrow x(T)\), there is a subsequence along which \(\|x_k - x\|_{\infty} \rightarrow 0\).

Let us denote by \(H\) the Hamiltonian associated with the multifunction \(F\), that is, \(H(t, x, p) = \max\{\langle v, p \rangle | v \in F(t, x)\}\). We know \(H(t, \cdot, \cdot)\) is locally Lipschitz \([4, \text{Proposition 3.2.4}]\), and denote by \(\partial H(t, x, p)\) the Clarke subgradient set of \(H(t, \cdot, \cdot)\) with respect to \((x, p)\).

**Theorem 2.** Let \(\tilde{x}\) be a local solution for \((P_D)\); suppose \(\tilde{x}\) is proper. Then

\[
N_{R(T)}(\tilde{x}(T)) \subseteq \overline{\partial} Q(\tilde{x}),
\]

where

\[
Q(x) = \{p(T) | (-p', x') \in \partial H(t, x, p), p(0) \in LN_{C_0}(x(0))\}. \tag{1}
\]

**Proof.** Let \(x\) be a trajectory of \((D)\), and \(\xi \in PN_{R(T)}(x(T))\). Then a constant \(M > 0\) exists, such that \(x(T)\) minimizes the smooth function \(\langle -\xi, y(T) \rangle + M \|y(T) - x(T)\|^2\) over all \(y(T) \in R(T)\). According to Theorem 1.2 of [8] (see also Theorem 4.1, footnote, of [5]), this implies that an arc \(p\) satisfying the Hamiltonian inclusion and initial condition defining \(Q(x)\) and \(-p(T) = -\xi\) exists, so \(\xi = p(T) \in Q(x)\). Since \(\xi\) is arbitrary, this shows that \(PN_{R(T)}(x(T)) \subseteq Q(x)\).

It follows that, for any \(\xi\) in \(LN_{R(T)}(\tilde{x}(T))\), there must be a sequence of trajectories \(x_k\) and a sequence of adjoint arcs \(p_k\) such that

\[
x_k(T) \rightarrow \tilde{x}(T), \quad p_k(T) \rightarrow \xi
\]
with
\[ (-p'_k, x'_k) \in \partial H(t, x_k, p_k), \quad p_k(0) \in LN_{C_0}(x_k(0)). \]
Since $\bar{x}$ is proper, there is a subsequence of $\{x_k\}$ along which $x_k$ uniformly converges to $\bar{x}$. Now $H$ is locally Lipschitzian, so $|p'_k(t)| \leq k_1(t)$, $|x'_k(t)| \leq k_1(t)$ almost everywhere for some integrable function $k_1$. By Theorem 3.1.7 of [4], applied to $\Gamma = \partial H$, a further subsequence $p_k$ uniformly converging to an adjoint arc $p$ exists. Since the graph of $LN_{C_0}$ is closed, we also have $p(0) \in LN_{C_0}(\bar{x}(0))$, that is, $p(T) \in Q(\bar{x})$. So $LN_{R(T)}(\bar{x}(T)) \subseteq Q(\bar{x})$, as required.

5. Approximation to the Reachable Set

In order to apply Theorem 1 to problem $(P_D)$, we also need to know the derivative cone and the second-order derivative set to the reachable set $R(T)$ at $\bar{x}(T)$. The first-order approximation has been discussed thoroughly in [6], where the following result is given.

Theorem 3. Let $r_1(T)$ be the reachable set of the following system:
\[
\begin{align*}
\dot{y}(t) &\in dF(t, \bar{x}(t), \bar{x}'(t))y(t) & \text{a.e. } t \in [0, T], \\
y(0) &\in D_{C_0}(\bar{x}(0)).
\end{align*}
\]
Then we have
\[ r_1(T) \subseteq D_{R(T)}(\bar{x}(T)). \]

Now we turn to the second-order derivative set to $R(T)$ at $\bar{x}(T)$ relative to $\bar{y}(T) \in r_1(T)$.

Theorem 4. Let the arc $\bar{y}$ satisfy (2), and let $r_2(T)$ denote the reachable set of the following system:
\[
\begin{align*}
\dot{z}(t) &\in d^2F(t, \bar{x}(t), \bar{x}'(t); \bar{y}(t), \bar{y}'(t))z(t) & \text{a.e. } t \in [0, T], \\
z(0) &\in D_{C_0}^2(\bar{x}(0); \bar{y}(0)).
\end{align*}
\]
Then we have
\[ r_2(T) \subseteq D_{R(T)}^2(\bar{x}(T), \bar{y}(T)). \]

Proof. Take any $z(T) \in r_2(T)$, and any sequence $t_k > 0$ converging to 0. Since $z(0) \in D_{C_0}^2(\bar{x}(0), \bar{y}(0))$, by the definition of the second-order derivative set there must be a sequence $s_k \to z(0)$, such that $\bar{x}(0) + t_k \bar{y}(0) + t_k^2 s_k \in C_0$. Let $z_k(t) := \bar{x}(t) + ...$
$t_k \bar{y}(t) + t_k^2 z(t) + t_k^3 (s_k - z(0))$. Then $z_k$ is an arc on $[0, T]$ and $z_k(0) \in C_0$. For $k$ sufficiently large, we have the estimate

$$\frac{d(z_k'(t), F(t, z_k(t)))}{t_k^2} \leq \left| z'(t) \right| + \frac{d(\bar{x}'(t) + t_k \bar{y}'(t), F(t, \bar{x}(t) + t_k \bar{y}(t)))}{t_k^2} + k(t)(|z(t)| + |s_k - z(0)|)$$

$$\leq \left| z'(t) \right| + k(t) + k(t)(|z(t)| + 1). \quad (3)$$

(The last inequality relies on assumption (H3).) Fix $t \in [0, T]$ outside the measure zero set implicitly specified below. Since

$$z'(t) \in d^2 F(t, \bar{x}(t), \bar{y}(t); y'(t))z(t)$$

for the same $t_k \downarrow 0$, a sequence $\{(\xi_k, \bar{z}_k)\}$ which converges to $(z(t), z'(t))$ exists such that

$$\bar{x}'(t) + t_k \bar{y}'(t) + t_k^2 \bar{z}_k \in F(t, \bar{x}(t) + t_k \bar{y}(t) + t_k^2 \bar{z}_k).$$

Therefore

$$\frac{d(z_k'(t), F(t, z_k(t)))}{t_k^2} \leq \left| \xi_k - z'(t) \right| + k(t)(|\xi_k - z(t)| + |s_k - z(0)|) \rightarrow 0 \quad \text{as} \quad k \rightarrow \infty.$$ 

Upon writing $\rho_k := \int_0^T d(z_k'(t), F(t, z_k(t))) \, dt$, we have, by (3) and the dominated convergence theorem, that

$$\lim_{k \to \infty} \frac{\rho_k}{t_k^2} = 0.$$ 

By Filippov’s theorem [4, Theorem 3.1.6], arcs $x_k$ on $[0, T]$ and a constant $c \geq 0$ exist such that the $x_k$ are trajectories of $F$ with $x_k(0) = z_k(0) \in C_0$ and $\|x_k - z_k\|_{\infty} \leq c \rho_k$. Each endpoint $x_k(T)$ lies in $R(T)$ and we have

$$\frac{x_k(T) - \bar{x}(T) - t_k \bar{y}(T)}{t_k^2} = \frac{x_k(T) - z_k(T)}{t_k^2} + z(T) + s_k - z(0) \rightarrow z(T) \quad \text{as} \quad k \rightarrow \infty.$$ 

So we must have $z(T) \in D^2_{R(T)}(\bar{x}(T), \bar{y}(T))$, as required. \qed
6. Second-Order Necessary Condition for (P_D)

Now we are ready to state the necessary conditions for problem (P_D).

**Theorem 5.** Let \( \bar{x} \) be a local solution to (P_D) and satisfy the constraint qualification

\[
(CQ) \quad \overline{Q}(\bar{x}) \cap (-N_{C_1}(\bar{x}(T))) = \{0\}.
\]

Then we have the first-order necessary condition

\[
D_1 g(\bar{x}(T), y(T)) \geq 0, \quad \forall y(T) \in r_1(T) \cap D_{C_1}(\bar{x}(T)).
\]

Furthermore, if equality holds for some \( \bar{y}(T) \), then we have the second-order necessary condition

\[
D_2^2 g(\bar{x}(T), \bar{y}(T), z(T)) \geq 0, \quad \forall z(T) \in r_2(T) \cap D_{C_1}^2(\bar{x}(T), \bar{y}(T)).
\]

**Proof.** If \( \bar{x} \) is proper, the result follows directly from Theorem 1 and the results above. Theorem 4 assures that assumption (CQ) implies the constraint qualification of Theorem 1, while the estimates of certain derivative sets in terms of \( r_1(T) \) and \( r_2(T) \) are described in Theorems 3 and 4.

If \( \bar{x} \) is not proper, instead of problem (P_D) we may study the auxiliary problem (\( \tilde{P}_D \)) with state variable \( \tilde{x} = (x_0, x) \in \mathbb{R}^{n+1} \) and data

\[
\tilde{g}(\tilde{x}) = x_0^4 + g(x), \quad \tilde{C}_0 = \{0\} \times C_0, \quad \tilde{C}_1 = [-1, 1] \times C_1, \quad \tilde{F}(t, \tilde{x}) = \{|x - \tilde{x}(t)|^4\} \times F(t, x).
\]

It is easy to see that \( \tilde{x} = (0, \bar{x}) \) is a local solution for problem (\( \tilde{P}_D \)). We show it is also proper: Let \( x_k \) be a sequence of trajectories for the differential inclusion (\( \tilde{D} \)) having the property \( x_k(T) \to \bar{x}(T) \), that is, \( \int_0^T |x_k(t) - \tilde{x}(t)|^4 \, dt \to 0 \), and \( x_k(T) \to \bar{x}(T) \). The sequence \( x_k \) is uniformly bounded and equicontinuous, so it has a subsequence converging uniformly to some limit \( \bar{x} \), but the integral condition above forces \( \bar{x} = \tilde{x} \), so \( \tilde{x} \) is proper. Therefore \( N_{\tilde{G}}(\tilde{x}(T)) \subseteq \overline{Q}(\tilde{x}) \). Here \( \tilde{Q} \) is expressed by (1) and the extended Hamiltonian \( \tilde{H} \) is given by

\[
\tilde{H}(t, \tilde{x}, \tilde{p}) = p_0 |x - \tilde{x}(t)|^4 + H(t, x, p).
\]

The function \( \tilde{H} \) does not depend on \( x_0 \), so \( p_0 \) is a constant, and we have \( \tilde{Q}(\tilde{x}) \subseteq \mathbb{R} \times Q(\tilde{x}) \). Our hypothesis \( \overline{Q}(\tilde{x}) \cap (-N_{C_1}(\tilde{x}(T))) = \{0\} \) implies that \( \overline{Q}(\tilde{x}) \cap (-N_{C_1}(\tilde{x})) = \{0\} \). By Theorem 1, we have

\[
D_1 \tilde{g}(\tilde{x}(T), \tilde{y}(T)) \geq 0, \quad \forall \tilde{y}(T) \in \tilde{r}_1(T) \cap D_{C_1}(\tilde{x}(T)).
\]

Furthermore, if equality holds for some \( \tilde{y}(T) \), then we have

\[
D_2^2 \tilde{g}(\tilde{x}(T), \tilde{y}(T), \tilde{z}(T)) \geq 0, \quad \forall \tilde{z}(T) \in \tilde{r}_2(T) \cap D_{C_1}^2(\tilde{x}(T), \tilde{y}(T)).
\]

It is easy to verify that if \( y(T) \in r_1(T) \) (or \( z(T) \in r_2(T) \)), then \( (0, y(T)) \in \tilde{r}_1(T) \) (or \( (0, z(T)) \in \tilde{r}_2(T) \)). Substituting these into the above expressions yields the required result. \( \square \)

The first-order necessary conditions implicit in Theorem 5 are not directly comparable with those in the literature \cite{4}, \cite{7} because they are in primal rather
than dual form. In other words, they assert the nonexistence of admissible descent directions for the objective function directly, instead of passing to a dual description involving an adjoint arc $p(t)$. In Section 8 below we make explicit the relationship between the necessary conditions of Theorem 5 and Pontryagin's maximum principle for a problem in which $F(t, x) = f(t, x, U)$ arises from a model in optimal control.

7. Unbounded Differential Inclusions

To get optimality conditions for unbounded differential inclusions, we first quote a definition in [8]:

**Definition.** Let $F: \Omega \to \mathbb{R}^n$ be a multifunction with closed values, and suppose $F$ is $\mathcal{L} \times \mathcal{B}$ measurable on $\Omega$. Consider a point $(\hat{t}, \hat{x})$ in $\Omega$.

(a) The multifunction $F$ is called **sub-Lipschitzian at** $(\hat{t}, \hat{x})$ if, for every constant $\rho \geq 0$, constants $\varepsilon > 0$ and $\alpha \geq 0$ exist such that

$$\Gamma(t, y) \cap \rho \text{ cl } B \subseteq \Gamma(t, x) + \alpha|y - x| \text{ cl } B$$

for all $t \in (\hat{t} - \varepsilon, \hat{t} + \varepsilon) \cap [0, T]$ and all $x, y$ in $\hat{x} + \varepsilon B$.

(b) The multifunction $F$ is called **integrably sub-Lipschitzian in the large at** $(\hat{t}, \hat{x})$ if constants $\varepsilon > 0$ and $\beta \geq 0$, together with a nonnegative function $\alpha(t)$ integrable on $(\hat{t} - \varepsilon, \hat{t} + \varepsilon)$, exist such that

$$\Gamma(t, y) \cap \rho \text{ cl } B \subseteq \Gamma(t, x) + (\alpha(t) + \beta \rho)|y - x| \text{ cl } B$$

for all $t \in (\hat{t} - \varepsilon, \hat{t} + \varepsilon) \cap [0, T]$, all $x, y$ in $\hat{x} + \varepsilon B$, and all $\rho \geq 0$.

Now for the differential inclusion control problem $(P_D)$ we retain hypotheses (H2)–(H5) and replace (H1) with the following hypothesis:

(H6) The multifunction $F$ has nonempty, closed, convex images on $\Omega$, and one of the following two conditions holds:

(a) The arc $\bar{x}$ is Lipschitzian, and the multifunction $F$ is sub-Lipschitz at every point $(t, \bar{x}(t))$ in $\text{gph} \bar{x}$.

(b) The multifunction $F$ is integrably sub-Lipschitzian in the large at every point $(t, \bar{x}(t))$ in $\text{gph} \bar{x}$.

We have the following optimality conditions for problem $(P_D)$:

**Theorem 6.** Under hypotheses (H2)–(H6), the conclusions of Theorem 5 still hold.

**Proof.** By Proposition 2.4 of [8] integrable functions $R$ and $\varphi$ on $[0, T]$ and a relatively open subset $\bar{\Omega}$ of $\Omega$ containing the graph of $\bar{x}$ on which the truncated multifunction $\bar{F}$ defined by

$$\bar{F}(t, x) := F(t, x) \cap (\bar{x}'(t) + R(t) \text{ cl } B)$$
satisfies the hypothesis (H1) on $\bar{\Omega}$ exist. The arc $\bar{x}$ is also a local solution for the new problem $(\bar{P}_D)$, so the conclusions of Theorem 5 can be applied to $(\bar{P}_D)$ with the corresponding constraint qualification and $\bar{r}_1(T)$ and $\bar{r}_2(T)$.

From the proof of Proposition 2.4 of [8] we know a constant $R_0 > 0$ exists such that $R(t) \geq R_0$ for all $t \in [0, T]$. Fixing $t \in [0, T]$ outside of a measure zero set and taking a small neighborhood of $(\bar{x}(t), \bar{x}'(t))$, we see immediately that

$$\eta \in F(t, \zeta) \iff \eta \in \bar{F}(t, \zeta)$$

whenever $\eta \in \bar{x}'(t) + R_0 B$. So the graph of $F$ and that of $\bar{F}$ are the same near the point $(\bar{x}(t), \bar{x}'(t))$. Since the derivative of $F$ at $(\bar{x}(t), \bar{x}'(t))$ is only related to the local behavior of the graph of $F$ at that point, we have

$$dF(t, \bar{x}(t), \bar{x}'(t)) = d\bar{F}(t, \bar{x}(t), \bar{x}'(t)).$$

So $\bar{r}_1(T) = r_1(T)$. Similarly we have $\bar{r}_2(T) = r_2(T)$.

Finally, note that any $(-p', \bar{x}') \in \partial \bar{H}(t, \bar{x}, p)$ implies that $\langle p, \bar{x}' \rangle \geq \langle p, v \rangle$ for all $v \in \bar{F}(t, \bar{x})$. Formula (3.4) in [8] tells us that

$$\partial \bar{H}(t, \bar{x}, p) \subseteq \partial H(t, \bar{x}, p).$$

So we have $\overline{\partial \bar{H}(\bar{x})} \subseteq \overline{\partial H(\bar{x})}$. We get the desired result by putting all these together.

**Remark.** If the multifunction $F$ satisfies assumption (H1), it will satisfy assumption (H6) just by taking $\alpha \equiv k$ and $\beta = 0$. Thus Theorem 6 is a true extension of Theorem 5.

### 8. Application to Optimal Control Problems

Theorem 6 is the second-order necessary condition for a general multifunction $F$. It can be used to study many problems, such as closed-loop control systems, implicit dynamical systems, and so on (see [6]). Here we just discuss its application to the optimal control problem below:

$$(P_D) \quad \text{minimize} \quad g(x(T))$$

subject to $x'(t) = f(t, x(t), u(t)), \quad x(0) \in C_0, \quad x(T) \in C_1,$

where $u(t)$ is a measurable selection of $U$. Denote the reachable set at time $T$ of the control system $(P_C)$ by $\bar{R}(T)$, that is,

$$\bar{R}(T) := \{x(T)|x'(t) = f(t, x(t), u(t)), x(0) \in C_0, u(t) \in U\}.$$  

Let $\Omega$ be a tube about an arc $\bar{x}$. Assume that $U$ is nonempty and closed, and the function $f: \Omega \times U \to \mathbb{R}^n$ satisfies the following conditions:
The function \( f(\cdot, x, u) \) is measurable on \([0, T]\).

The function \( f(t, \cdot, \cdot) \) is twice continuously differentiable, and an integrable function \( k \) exists such that, at any point \((t, x, u) \in \Omega \times U\), we have \(|f_x|, |f_u| \leq k(t)\).

Consider the multifunction \( F(t, x) := f(t, x, U) \). Clearly, \( F \) satisfies assumptions (H2) and (H6)(b), and \( F \) has nonempty images on \( \Omega \). Here we explicitly suppose that \( F(t, x) \) is convex for all \((t, x) \in \Omega\) (this is a natural requirement for the existence of optimal solutions, see [5]), that \( F(t, x) \) is closed (if \( U \) is compact, this automatically holds), and that \( F \) satisfies (H3) (see Theorem 1 for a condition under which (H3) holds).

**Theorem 7.** Let \( \hat{r}_1(T) \) be the reachable set of the following system:

\[
\begin{cases}
y'(t) = \bar{f}_x(t)y(t) + \bar{f}_u(t)v(t) & \text{a.e. } t, \\
y(0) \in D_C(\bar{x}(0)), \\
v \in L^\infty[0, T] \text{ and } v(t) \in D_U(\bar{u}(t)).
\end{cases}
\]

Then we have

\[
\hat{r}_1(T) \subseteq D_{R(T)}(\bar{x}(T)).
\]

Here \( \bar{f}_x(t) \) denotes \( f'_x(t, \bar{x}(t), \bar{u}(t)) \), and \( \bar{f}_u(t) \) is defined similarly.

**Proof.** According to Theorem 3, it suffices to show that, for fixed \( t \in [0, T] \) outside some measure zero set, any \( \alpha \in \mathbb{R}^n, v \in D_U(\bar{u}(t)) \), we have

\[\bar{f}_x(t)\alpha + \bar{f}_u(t)v \in dF(t, \bar{x}(t), \bar{x}'(t))\alpha.\]

This is equivalent to, \( \forall t_k \downarrow 0, \exists \xi_{k} \rightarrow \alpha, \zeta_{k} \rightarrow \bar{f}_x(t)\alpha + \bar{f}_u(t)v, u_k \in U \), such that

\[f(t, \bar{x}(t), \bar{u}(t)) + t_k \zeta_{k} = f(t, \bar{x}(t) + t_k \xi_{k}, u_k).\]

Now we prove the latter statement. By assumption \( v \in F_U(\bar{u}(t)) \), so a sequence \( v_k \rightarrow v \) exists such that \( u_k := \bar{u}(t) + t_k v_k \in U \). Let \( \xi_k = \alpha \) and \( \zeta_k = (f(t, \bar{x}(t) + t_k \alpha, u_k) - f(t, \bar{x}(t), \bar{u}(t)))/t_k.\)

Using a first-order Taylor's expansion, we see immediately that \( \zeta_k \rightarrow \bar{f}_x(t)\alpha + \bar{f}_u(t)v.\)

**Theorem 8.** Let \( \bar{y}(T) \in \hat{r}_1(T) \) be as described in Theorem 7 with control \( \bar{u} \), and define

\[H(t) = \frac{1}{2}\bar{f}_x(x(t))(\bar{y}(t), \bar{y}(t)) + \frac{1}{2}\bar{f}_u(u(t))(\bar{y}(t), \bar{u}(t)) + \frac{1}{2}\bar{f}_{uu}(t)(\bar{y}(t), \bar{u}(t)).\]

Let \( \hat{r}_2(T) \) be the reachable set of the following control system:

\[
\begin{cases}
z'(t) = \bar{f}_x(t)z(t) + \bar{f}_u(t)w(t) + H(t) & \text{a.e. } t \in [0, T], \\
z(0) \in D_{C(\bar{x}(0), \bar{y}(0))}, \\
w \in L^\infty[0, T] \text{ and } w(t) \in D_{U}(\bar{u}(t), \bar{u}(t)).
\end{cases}
\]

Then we have
\[ \hat{\mathbf{r}}_2(T) \subseteq D_{R(T)}(\hat{x}(T), \hat{y}(T)). \]

**Proof.** Similar to Theorem 7.

We can now state a second-order necessary condition for the optimal control problem \((P_C)\). This result follows from Theorems 1, 2, 7, and 8.

**Theorem 9.** Let \( \bar{x} \) be a local solution for \((P_C)\) with the corresponding control \(\bar{u}\). If \( \bar{x} \) satisfies the constraint qualification
\[ Q(\bar{x}) \cap (-N_{C_1}(\bar{x}(T))) = \{0\}, \]
then we have
\[ D_1g(\bar{x}(T), y(T)) \geq 0, \quad \forall (T) \in \hat{r}_1(T) \cap D_{C_1}(\bar{x}(T)). \]

Furthermore, if equality holds for some \( \bar{y}(T) \), then we have
\[ D_2^2g(\bar{x}(T), \bar{y}(T), z(T)) \geq 0, \quad \forall (T) \in \hat{r}_2(T) \cap D_{C_2}(\bar{x}(T), \bar{y}(T)). \]

Recall that the set \( Q(\bar{x}) \) is defined in (1). In particular, the constraint qualification in Theorem 9 is expressed in terms of a Hamiltonian inclusion.

Recall [12] that the admissible pair \((\bar{x}, \bar{u})\) is said to be normal if the only solution of the following system (*) is \( p = 0 \):
\[
\begin{cases}
- p'(t) = \bar{f}_x(t)p(t), \\
p(0) \in N_{C_0}(\bar{x}(0)), \\
- p(T) \in N_{C_1}(\bar{x}(T)), \\
\max_{u \in U} \langle p(t), f(t, \bar{x}(t), u) \rangle = \langle p(t), f(t, \bar{x}(t), \bar{u}(t)) \rangle.
\end{cases}
\]

The final result in this paper shows how our constraint qualification (CQ) can be replaced by this normality condition. It shows that Theorem 9 accurately recovers the admissible directions and the accessory problem in Theorem 2.2 of [12].

**Theorem 10.** Suppose the optimal pair \((\bar{x}, \bar{u})\) is normal and the control set \(U\) is convex. Then the necessary conditions of Theorem 9 hold regardless of (CQ). In particular, if the function \( g \) is twice continuously differentiable near the point \( \bar{x}(T) \), and the arc \( y \) is a trajectory of the following system for some function \( v \in L^\infty[0, T] \),
\[
\begin{cases}
y'(t) = \bar{f}_x(t)y(t) + \bar{f}_u(t)v(t), \\
y(0) \in D_{C_0}(\bar{x}(0)), \\
y(T) \in D_{C_1}(\bar{x}(T)), \\
g'(\bar{x}(T))y(T) = 0, \\
v(t) \in U - \bar{u}(t),
\end{cases}
\]

\( \dagger \)
then we have
\[ g'(\bar{x}(T))^T z(T) + \frac{1}{2} y(T) g''(\bar{x}(T)) y(T) \geq 0 \]
for any arc \( z \) satisfying the following equation:
\[
\begin{aligned}
z'(t) &= \tilde{f}_{\bar{x}}(t) z(t) + H(t), \\
z(0) &\in D^\infty_c(\bar{x}(0), y(0)), \\
z(T) &\in D^\infty_c(\bar{x}(T), y(T)).
\end{aligned}
\]

Proof. Since \( U \) is a convex set, for all \( 0 < \alpha < 1 \), we have \( \tilde{u}(t) + \alpha v(t) \in U \)
and
\[
d(\tilde{x}'(t) + \alpha x'(t), F(t, \tilde{x}(t) + \alpha y(t))) \\
\leq |f(t, \tilde{x}(t) + \alpha y(t), \tilde{u}(t) + \alpha v(t)) - \tilde{x}'(t) - \alpha x'(t)| \\
\leq \alpha^2 k(t).
\]

So assumption (H3) holds in this case.

Following the trick in [3], we want to show that if \( (\tilde{x}, \tilde{u}) \) is normal, then our constraint qualification (CQ) holds for the new differential inclusion control problem \( \bar{P}_D \) with state variable \( \tilde{x} = (x_0, x) \) in \( R^{n+1} \) and data \( \tilde{g}(\tilde{x}) = g(x) \), \( \tilde{f}(t, \tilde{x}) = \partial \tilde{g}\left(\{u - \tilde{u}(t)\}^4, f(t, x, u)\right) : u \in U \), \( \tilde{C}_0 = \{0\} \times C_0 \), and \( \tilde{C}_1 = [-1, 1] \times C_1 \). If \( \tilde{x} = (x_0, x) \) is admissible for \( \bar{P}_D \), by the closedness and convexity of \( f(t, x, U) \) and the measurable selection theorem, a control \( u \) exists such that \( (x, u) \) is admissible for \( \bar{P}_C \). Since \( (\bar{x}, \bar{u}) \) is a local solution for \( \bar{P}_C \), we have \( \bar{g}(\bar{x}(T)) \geq g(\bar{x}(T)) \). So \( \bar{x} = (0, \bar{x}) \) is a local solution for \( \bar{P}_D \). The extended Hamiltonian is
\[
\bar{H}(t, \bar{x}, \bar{p}) = \max_{u \in U} \{p_0||u - \tilde{u}(t)||^4 + \langle p, f(t, x, u) \rangle\}.
\]

By Theorem 2.8.6 of [4] its generalized gradient \( \partial \bar{H}(t, \bar{x}, \bar{p}) \) is the convex hull of the vectors \((0, f_x(t, \bar{x}, \bar{u})^T p, ||u - \tilde{u}(t)||^4, f(t, \bar{x}, \bar{u})) \) over all \( \bar{u} \) such that
\[
\max_{u \in U} \{p_0||u - \tilde{u}(t)||^4 + \langle p, f(t, \bar{x}, u) \rangle\} = p_0 ||\bar{u} - \tilde{u}(t)||^4 + \langle p, f(t, \bar{x}, \bar{u}) \rangle.
\]

Now consider a Hamiltonian adjoint arc \((p_0, p)\) associated with \( \bar{x} \). \( \bar{H} \) does not depend on \( x_0 \), so \( p_0 \) is a constant. Since \((0, -p', 0, \bar{x}') \) is in \( \partial \bar{H} \), the only \( \bar{u} \) would be \( \bar{u}(t) \). So \( \bar{Q}(\bar{x}) \) is the set of all \((p_0(T), p(T))\) such that \( p_0(t) \) is a constant, \( p \) satisfies \(-p'(t) = f_\bar{x}(t)^T p(t), p(0) \in N_{C_0}(\bar{x}(0))\), and
\[
\max_{u \in U} \{p_0||u - \tilde{u}(t)||^4 + \langle p, f(t, \bar{x}(t), u) \rangle\} = \langle p(t), f(t, \bar{x}(t), \bar{u}(t)) \rangle.
\]

Now any \( \bar{p}(T) \in \bar{Q}(\bar{x}) \cap (-N_{C_1}(\bar{x}(T))) \) must have \( p_0 = 0 \), so \( p \) satisfies the system (\( \star \)). By the normality of \( (\bar{x}, \bar{u}) \) we have \( p \equiv 0 \). So the constraint qualification for \( \bar{P}_D \) holds. It is easy to show that if \( y(T) \in \bar{p}_1(T) \) (or \( z(T) \in \bar{p}_2(T) \)), then we have \((0, y(T)) \in \bar{p}_1(T) \) (or \((0, z(T)) \in \bar{p}_2(T)) \); also note that \( 0 \in D^\infty_c(\bar{u}(t), v(t)) \). By applying Theorem 5 we get the required conclusion. \( \square \)
To complete the comparison between Theorem 10 and Theorem 2.2 of [12], let \( C_0 = \{ x_0 \} \) and \( C_1 = \{ x \in \mathbb{R}^n | \Phi(x) = 0 \} \). Then by Propositions 4.3.7 and 4.7.5 of [2] the endpoint constraints become \( y(0) = 0, \Phi'((\tilde{x}(T))z(T) + \frac{1}{2} \Phi''((\tilde{x}(T))(y(T), y(T))) = 0 \) in (\( \dagger \)) and \( \tilde{z}(0) = 0, \Phi'((\tilde{x}(T)))z(T) + \frac{1}{2} \Phi''((\tilde{x}(T))(y(T), y(T))) = 0 \) in (\( \ddagger \)). So (\( \dagger \)) is equivalent to (2.3), (2.4), and (2.5) in [12]. If \( p \) is an adjoint arc [12, Theorem 2.1] with \( \lambda_0 = 1 \), that is, \( -p'(t) = \tilde{f}_z(t)p(t), p(T) = g'((\tilde{x}(T)) + \Phi'((\tilde{x}(T)))z(T) \), then \((p(t))^Tz(t)' = p(t)^TH(t)\).

These observations allow us to rewrite the central inequality of Theorem 10 as

\[
0 \leq \frac{1}{2}y(T)^Tg''((\tilde{x}(T))y(T) + p(T)^Tz(T) - \gamma^T\Phi'((\tilde{x}(T)))z(T)
= \frac{1}{2}y(T)^T(g + \gamma^T\Phi')((\tilde{x}(T)))y(T) + \int_0^T p(t)^TH(t) \, dt.
\]

Thus Theorem 10 agrees with Theorem 2.2 of [12].

Acknowledgment

I would like to thank Professor Loewen for his advice and encouragement during the preparation of this paper.

References


Accepted 15 February 1993