Set separation and the Lipschitz maximum principle

Héctor J. Sussmann
Department of Mathematics
Rutgers, the State University of New Jersey
Hill Center—Busch Campus
110 Frelinghuysen Road
Piscataway, NJ 08854-8019, USA
sussmann@math.rutgers.edu

***Abstract*— We present a general necessary condition for separation of the reachable set of a Lipschitz control system from another given set, expressed in terms of an "approximating multicone" to the set in a sense that contains as special cases the Clarke and Mordukhovich cones. We then show how this separation result implies a stronger form of the usual necessary condition for optimality.**

I. Introduction

This paper announces results, proved in [24], pursuing the investigation, initiated in [22], of the possibility of deriving general versions of the finite-dimensional Pontryagin maximum principle (PMP) by means of a unified method, based on set separation.

Proofs of various versions of the PMP using set separation together with a topological argument based on the Brouwer fixed point theorem have appeared in Pontryagin *et al.* [14], Berkovitz [1], as well as in many papers that incorporate high-order conditions (cf. Bianchini [2], Knobloch [10], Krener [11], Stefani [15], Sussmann [16]). Another set of proofs use a different approach, based on a limiting argument, in which a sequence $\boldsymbol{\pi} = \{\pi_j\}_{j\in\mathbb{N}}$ of approximate terminal adjoint vectors π_j—normalized so that $\|\pi_j\| = 1$—is constructed, and then an exact adjoint vector is obtained by taking the limit of some convergent subsequence of $\boldsymbol{\pi}$. These proofs (cf. Clarke [5], [6], Clarke *et al.* [7], Ioffe [8], Ioffe-Rockafellar [9], Mordukhovich [13], Vinter [25]) have successfuly dealt with nonsmooth Lipschitz dynamical laws, which appeared inaccessible to the topological method.

In 1993, S. Łojasiewicz Jr. ([12]) discovered a powerful new technique that made it possible to deal with nonsmoothness by means of the topological set separation method. Subsequently, in a series of papers (cf. [18], [19], [20], [21], [23]), we pursued this idea and developed topological methods for the non-smooth PMP, based on generalized differentials, flows, and general variations. These methods, however, always led to results where the transversality conditions involved a tangent cone to the target set that was a Boltyanskii approximating cone or some generalization thereof, and resisted all attempts to deal with transversality conditions involving the Clarke tangent cone or the Mordukhovich normal cone. Recently, A. Bressan (cf. [4]) found an explanation for this fact, by proving, by means of a counterexample, that the usual necessary conditions for set separation that can be derived for a pair of sets and corresponding Boltyanskii approximating cones, as well as for a pair of sets and corresponding Clarke or Mordukhovich normal cones, can fail to be true if a Boltyanskii aproximating cone is specified for one of the sets and the Clarke or Mordukhovich normal cone is used for the other one. This shows that versions of the PMP with "mixed" technical conditions—some suitable for the topological approach and others for the limiting method—are likely to be false in general, and that there probably does not exist a single unified version of the PMP that contains both types of results.

The second-best alternative is that, even though a single commmon generalization of both approaches does not exist, it may at least be possible to deal with both kinds of results by means of set-separation techniques, using different but parallel separation theorems for the two kinds of results. As a first step in this direction, we proposed in [22] a notion of "approximating multicone" to a set at a point that extends the concepts of Clarke and Mordukhovich cones and has the property that "strong transversality of the approximating cones implies nontrivial intersection of the sets." (In our setting, "convex multicones" have polars that are ordinary cones but can fail to be convex. In particular, the usual Mordukhovich normal cone is the polar of a convex multicone, that we call the "Mordukhovich tangent multicone.")

In this note we apply the set-separation approach to Lipschitz control problems. We first present, in Theorem 3.5, a necessary and sufficient condition for two multicones not to be strongly transversal, expressed in terms of polar covectors. We then use this result to obtain, in Theorem 4.1, a general necessary condition for separation of the reachable set $\mathcal{R}$ of a control system from another given set S, expressed in terms of an approximating multicone to S in the sense of our theory, and observe that, if the necessary condition is not satisfied, then there exists a nonconstant Lipschitz curve $[0,1] \ni s \mapsto \zeta(s) \in \mathcal{R} \cap S$ such that $\zeta(0) = \xi_*(b)$, where ξ_* is the reference trajectory and b is the terminal time. We then show how this condition can be used to derive a strengthened form of the usual necessary conditions for optimal control.

Research supported in part by NSF Grant DMS-05-09930

II. Preliminaries and background

Some abbreviations and basic notations. We use "FDRLS" for "finite-dimensional real linear space".

If X, Y are FDRLSs, then $Lin(X,Y)$ will denote the space of all linear maps from X to Y. If X is a FDRLS, then $\dim X$, $X^\dagger$ denote, respectively, the dimension and the dual of X (so that $X^\dagger = Lin(X,\mathbb{R})$). We identify the double dual $X^{\dagger\dagger}$ with X in the usual way.

We write $\mathbb{R}^n$, $\mathbb{R}_n$ to denote, respectively, the spaces of all real n-dimensional column vectors $x = (x^1, \dots, x^n)^\dagger$ and of all real n-dimensional row vectors $p = (p_1, \dots, p_n)$. We identify $Lin(\mathbb{R}^n, \mathbb{R}^m)$ with the space $\mathbb{R}^{m\times n}$ of real $m \times n$ matrices in the usual way, by assigning to each $M \in \mathbb{R}^{m\times n}$ the linear map $\mathbb{R}^n \ni x \mapsto M \cdot x \in \mathbb{R}^m$. Also, we identify $\mathbb{R}_n$ with $(\mathbb{R}^n)^\dagger$, by assigning to a $y \in \mathbb{R}_n$ the linear functional $\mathbb{R}^n \ni x \mapsto y \cdot x \in \mathbb{R}$. If X, Y are FDRLSs, and $L \in Lin(X,Y)$, then the ***adjoint*** (or ***transpose***) of L is the map $L^\dagger : Y^\dagger \mapsto X^\dagger$ such that $L^\dagger(y) = y \circ L$ for $y \in Y^\dagger$. In the special case when $X = \mathbb{R}^n$ and $Y = \mathbb{R}^m$, so $L \in \mathbb{R}^{m\times n}$, the map $L^\dagger$ goes from $\mathbb{R}_m$ to $\mathbb{R}_n$, and is given by $L^\dagger(y) = y \cdot L$ for $y \in \mathbb{R}_m$.

If $m \in \mathbb{Z}_+$, $x \in \mathbb{R}^m$, $r \in \mathbb{R}$, and $r > 0$, then $\bar{\mathbb{B}}^m(x,r)$, $\mathbb{B}^m(x,r)$ denote, respectively, the closed and open balls in $\mathbb{R}^m$ with center x and radius r. We write $\bar{\mathbb{B}}^m(r)$, $\mathbb{B}^m(r)$ for $\bar{\mathbb{B}}^m(0,r)$, $\mathbb{B}^m(0,r)$, and $\bar{\mathbb{B}}^m$, $\mathbb{B}^m$ for $\bar{\mathbb{B}}^m(1)$, $\mathbb{B}^m(1)$.

Generalized Jacobians and Warga derivate containers. If Ω is open in $\mathbb{R}^n$, and $\varphi : \Omega \mapsto \mathbb{R}^m$ is locally Lipschitz, we use $diff(\varphi)$ to denote the set of points of differentiability of φ. It follows from the well known Rademacher theorem that $\Omega \backslash diff(\varphi)$ is a null subset of Ω. If $x \in \Omega$, we let $\tilde{\partial}\varphi(x)$ be the set of all limits as $k \to \infty$ of convergent sequences $\{D\varphi(x_k)\}_{k=1}^\infty$ such that $x_k \in diff(\varphi)$ for all k and $\lim_{k\to\infty} x_k = x$. Then $\tilde{\partial}\varphi(x)$ is a nonempty compact subset of $Lin(\mathbb{R}^n, \mathbb{R}^m)$. We use $\partial\varphi(x)$ to denote the convex hull of $\tilde{\partial}\varphi(x)$, and refer to it as the ***Clarke generalized Jacobian*** of φ at x.

"Warga derivate containers" are defined as follows.

Definition 2.1: Assume that Ω is open in $\mathbb{R}^n$, and $\Omega \ni x \mapsto F(x) \subseteq \mathbb{R}^m$ is a multivalued map from Ω to $\mathbb{R}^m$. Let $(x,y) \in \Omega \times \mathbb{R}^m$, and let Λ be a compact subset of $Lin(\mathbb{R}^n, \mathbb{R}^m)$. We say that Λ is a ***Warga derivate container of*** F ***at*** (x,y), and write "$\Lambda \in WDC(F;x,y)$," if for every open subset $\mathcal{O}$ of $Lin(\mathbb{R}^n, \mathbb{R}^m)$ such that $\Lambda \subseteq \mathcal{O}$ there exist (a) an open subset U of Ω such that $x \in U$, (b) a sequence $\boldsymbol{\varphi} = \{\varphi_k\}_{k=1}^\infty$ of maps of class C^1 from U to $\mathbb{R}^m$, such that (i) $D\varphi_k(x') \in \mathcal{O}$ whenever $x' \in U$ and $k \in \mathbb{N}$, and (ii) $\boldsymbol{\varphi} \to \varphi$ uniformly for some locally Lipschitz function $\varphi : U \mapsto \mathbb{R}^m$ such that $\varphi(x) = y$ and $\varphi(x') \in F(x')$ for all $x' \in U$. ∎

The above definition reduces, when F is single-valued, to the standard one. If F is signle-valued and $F(x) = y$, we will write "$\Lambda \in WDC(F;x)$" rather than "$\Lambda \in WDC(F;x,y)$."

III. Cones, multicones, transversality, and set separation

Cones, multicones, polars. A ***cone*** in a FDRLS X is a nonempty subset C of X such that $r \cdot c \in C$ whenever $c \in C, r \in \mathbb{R}$ and $r \geq 0$. If C is a cone in X, the ***polar*** of C is the convex cone $C^\perp = \{\lambda \in X^\dagger : \lambda(c) \leq 0 \text{ for all } c \in C\}$. Then $C^\perp$ is a closed convex cone in $X^\dagger$, and $C^{\perp\perp}$ is the smallest closed convex cone containing C. In particular, $C^{\perp\perp} = C$ if and only if C is closed and convex.

A ***multicone*** in X is a nonempty set of convex cones in X. A multicone $\mathcal{C}$ is ***convex*** if every member C of $\mathcal{C}$ is convex and ***closed*** if every $C \in \mathcal{C}$ is closed. The ***polar*** of $\mathcal{C}$ is the set $\mathcal{C}^\perp = \mathrm{Clos}\Big(\bigcup\{C^\perp : C \in \mathcal{C}\}\Big)$, so $\mathcal{C}^\perp$ is a (not necessarily convex) closed cone in $X^\dagger$.

Transversality of cones. We say that two convex cones C_1, C_2 in a FDRLS X are ***transversal***, and write $C_1 \overline{\pitchfork} C_2$, if $C_1 - C_2 = X$, i.e., if for every $x \in X$ there exist $c_1 \in C_1$, $c_2 \in C_2$, such that $x = c_1 - c_2$. We say that C_1 and C_2 are ***strongly transversal***, and write $C_1 \overline{\overline{\pitchfork}} C_2$, if $C_1 \overline{\pitchfork} C_2$ and in addition $C_1 \cap C_2 \neq \{0\}$. Then "$\sim C_1 \overline{\pitchfork} C_2$", "$\sim C_1 \overline{\overline{\pitchfork}} C_2$" will stand for "$C_1$ and C_2 are not transversal," and "C_1 and C_2 are not strongly transversal," respectively.

The following is easily proved.

Lemma 3.1: Assume that X is a FDRLS and C_1, C_2 are convex cones in X. Then $C_1 \overline{\pitchfork} C_2 \Leftrightarrow \overline{C_1} \overline{\pitchfork} \overline{C_2} \Leftrightarrow C_1^\perp \cap (-C_2^\perp) = \{0\}$. Furthermore, $C_1 \overline{\pitchfork} C_2$ if and only if either (i) $C_1 \overline{\overline{\pitchfork}} C_2$, or (ii) C_1 and C_2 are both linear subspaces and $C_1 \oplus C_2 = X$. ∎

If C is a convex cone in a FDRLS X, then $span(C)$ will denote the linear span of C. It is then clear that $span(C) = span(\bar{C})$. It is then easy to see that

Lemma 3.2: Assume that X is a FDRLS and C is a convex cone in X. Then $\overline{\mathrm{Int}_{span(C)}(C)} = \bar{C}$ and $\mathrm{Int}_{span(C)}(\bar{C}) = \mathrm{Int}_{span(C)}(C) \neq \emptyset$. ∎

Lemma 3.3: Assume that X is a FDRLS, C_1, C_2 are convex cones in X, and $C_1 \overline{\pitchfork} C_2$. Then $\overline{C_1 \cap C_2} = \overline{C_1} \cap \overline{C_2}$ and $(C_1 \cap C_2)^\perp = C_1^\perp + C_2^\perp$. ∎

Lemma 3.4: Assume that X is a FDRLS and C_1, C_2 are convex cones in X. Then $C_1 \overline{\overline{\pitchfork}} C_2$ if and only if $\overline{C_1} \overline{\overline{\pitchfork}} \overline{C_2}$. ∎

Transversality of multicones. Two convex multicones $\mathcal{C}_1$, $\mathcal{C}_2$ in a fdrls X are ***transversal*** if $C_1 \overline{\pitchfork} C_2$ whenever $C_1 \in \mathcal{C}_1$, $C_2 \in \mathcal{C}_2$. A linear functional $\mu \in X^\dagger$ is ***intersection positive*** on $(\mathcal{C}_1, \mathcal{C}_2)$ if

$$(\forall\, C_1 \in \mathcal{C}_1,\, C_2 \in \mathcal{C}_2)(\exists x \in C_1 \cap C_2)(\mu(x) > 0)\,. \qquad (1)$$

The convex multicones $\mathcal{C}_1$, $\mathcal{C}_2$ are ***strongly transversal*** if they are transversal and in addition there exists a $\mu \in X^\dagger$ which is intersection positive on $(\mathcal{C}_1, \mathcal{C}_2)$.

We use for transversality of multicones the same notations as for cones: "$\mathcal{C}_1 \overline{\pitchfork} \mathcal{C}_2$" (resp. "$\sim \mathcal{C}_1 \overline{\pitchfork} \mathcal{C}_2$") means that $\mathcal{C}_1$ and $\mathcal{C}_2$ are (resp. are not) transversal, and "$\mathcal{C}_1 \overline{\overline{\pitchfork}} \mathcal{C}_2$"

(resp. "$\sim \mathcal{C}_1 \pitchfork \mathcal{C}_2$") means that $\mathcal{C}_1$ and $\mathcal{C}_2$ are (resp. are not) strongly transversal. If $\mathcal{C}$ is a multicone in a FDRLS X, we write $\overline{\mathcal{C}}$ to denote the multicone $\{\bar{C} : C \in \mathcal{C}\}$.

The following theorem is proved in [24].

Theorem 3.5: Assume that X is a FDRLS and $\mathcal{C}_1$, $\mathcal{C}_2$ are convex multicones in X. Then the following conditions are equivalent:

(i) $\mathcal{C}_1$ and $\mathcal{C}_2$ are not strongly transversal;
(ii) $\overline{\mathcal{C}_1}$ and $\overline{\mathcal{C}_2}$ are not strongly transversal;
(iii) for every $\nu \in X^\dagger \backslash \{0\}$ there exist $C_1 \in \mathcal{C}_1$, $C_2 \in \mathcal{C}_2$, $\omega_1 \in C_1^\perp$, $\omega_2 \in C_2^\perp$, $\omega_0 \in [0, +\infty[$, such that $(\omega_1, \omega_2, \omega_0) \neq (0, 0, 0)$ and $\omega_1 + \omega_2 = \omega_0 \nu$. ■

Mordukhovich tangent multicones. Let S be a subset of $\mathbb{R}^n$, and let $\bar{s} \in S$. The ***Bouligand tangent cone*** to S at $\bar{s}$ is the set of all vectors $v \in \mathbb{R}^n$ such that there exist a sequence $\{s_j\}_{j \in \mathbb{N}}$ of points of S converging to $\bar{s}$, and a sequence $\{h_j\}_{j \in \mathbb{N}}$ of positive real numbers converging to 0, such that $v = \lim_{j \to \infty} \frac{s_j - \bar{s}}{h_j}$. We use $T_{\bar{s}}^B S$ to denote the Bouligand tangent cone to S at $\bar{s}$. It is clear, and well known, that $T_{\bar{s}}^B S$ is a closed cone. The ***Bouligand normal cone*** of S at $\bar{s}$ is the polar cone $(T_{\bar{s}}^B S)^\perp$ of $T_{\bar{s}}^B S$, that is, the set of all covectors $p \in \mathbb{R}_n$ such that $p \cdot v \leq 0$ for all $v \in T_{\bar{s}}^B S$. The ***limiting normal cone***, or ***Mordukhovich normal cone*** of S at $\bar{s}$ is the set of all covectors $p \in \mathbb{R}_n$ such that $p = \lim_{j \to \infty} p_j$ for some sequence $\{s_j\}_{j \in \mathbb{N}}$ of members of S that converges to $\bar{s}$ and some sequence $\{p_j\}_{j \in \mathbb{N}}$ of members of $\mathbb{R}_n$ such that $p_j \in (T_{s_j}^B S)^\perp$ for each j.

We use $N_{\bar{s}}^{Mo} S$ to denote the Mordukhovich normal cone of S at $\bar{s}$. For each $p \in \mathbb{R}_n$, we let $p^\perp = \{v \in \mathbb{R}^n : p \cdot v \leq 0\}$, so $p^\perp$ is a half space if $p \neq 0$, and $p^\perp$ is the whole space $\mathbb{R}^n$ if $p = 0$. The ***Mordukhovich tangent multicone to*** S ***at*** $\bar{s}$ is the set $T_{\bar{s}}^{Mo} S \stackrel{\text{def}}{=} \{p^\perp : p \in N_{\bar{s}}^{Mo} S\}$, so $T_{\bar{s}}^{Mo} S$ is a set all whose members are closed half-spaces in $T_{\bar{s}} M$, except for one "trivial member," namely, the whole space $T_{\bar{s}} M$.

The Clarke tangent and normal cones. If S is a closed subset of $\mathbb{R}^n$, and $\bar{s} \in S$, then the ***Clarke tangent cone*** to S at $\bar{s}$ is the set of all vectors $v \in \mathbb{R}^n$ such that, whenever $\{s_j\}_{j \in \mathbb{N}}$ is a sequence of points of S converging to $\bar{s}$, it follows that there exist Bouligand tangent vectors $v_j \in T_{s_j}^B S$ such that $\lim_{j \to \infty} v_j = v$. We use $T_{\bar{s}}^{Cl} S$ to denote the Clarke tangent cone to S at $\bar{s}$. It is well known that $T_{\bar{s}}^{Cl} S$ is a closed convex cone. Also, it is well-known that $T_{\bar{s}}^{Cl} S$ is the polar of the Mordukhovich cone $N_{\bar{s}}^{Mo} S$. Therefore $T_{\bar{s}}^{Cl} = \bigcap \{C : C \in T_{\bar{s}}^{Mo} S\}$. The ***Clarke normal cone*** $N_{\bar{s}}^{Cl} S$ of S at $\bar{s}$ is the polar $(T_{\bar{s}}^{Cl} S)^\perp$ of the Clarke tangent cone, so $(T_{\bar{s}}^{Cl} S)^\perp$ is the smallest closed convex cone in $\mathbb{R}_n$ containing $N_{\bar{s}}^{Mo} S$.

WDC approximating multicones. If $\mathcal{C}$, $\mathcal{D}$ are convex multicones, then we say that $\mathcal{C}$ ***is a full submulticone of*** $\mathcal{D}$, and write $\mathcal{C} \preceq \mathcal{D}$, if for every $D \in \mathcal{D}$ there exists a $C \in \mathcal{C}$ such that $C \subseteq D$.

If X, Y are FDRLSs, $\mathcal{C}$ is a multicone in X, and $\Lambda \subseteq Lin(X, Y)$, then we define $\Lambda \cdot \mathcal{C} \stackrel{\text{def}}{=} \{L \cdot C : L \in \Lambda, C \in \mathcal{C}\}$.

Definition 3.6: If $\bar{s} \in S \subseteq \mathbb{R}^n$, and $\mathcal{C}$ is a convex multicone in $\mathbb{R}^n$, we say that $\mathcal{C}$ is a ***WDC approximating multicone of*** S ***at*** $\bar{s}$ if there exist (i) a nonnegative integer d, (ii) a compact subset K of $\mathbb{R}^d$ such that $0 \in K$, (iii) an open neighborhood U of K in $\mathbb{R}^d$, (iv) a set-valued map $U \ni u \mapsto F(u) \subseteq \mathbb{R}^n$, (v) a compact subset Λ of $Lin(\mathbb{R}^d, \mathbb{R}^n)$, and (vi) a convex multicone $\mathcal{D}$ in $\mathbb{R}^d$, such that (I) $F(K) \subseteq S$, (II) $\Lambda \in WDC(F; 0, \bar{s})$, (III) $\mathcal{D} \preceq T_{\bar{s}}^{Mo} K$, and, finally, (IV) $\mathcal{C} = \Lambda \cdot \mathcal{D}$. ■

We will use $WDCAM(S, \bar{s})$ to denote the set of all WDC approximating multicones of S at $\bar{s}$, so "$\mathcal{C} \in WDCAM(S, \bar{s})$" is an alternative way of saying that "$\mathcal{C}$ is a WDC approximating multicone of S at $\bar{s}$."

The transversal intersection property. If X is a topological space, and S_1, S_2 are subsets of X, we say that S_1 and S_2 are ***locally separated*** at a point $p \in X$ if there exists a neighborhood U of p in X such that $S_1 \cap S_2 \cap U \subseteq \{p\}$.

The following result is proved in [24]..

Theorem 3.7: Let S_1, S_2 be subsets of $\mathbb{R}^n$, and let $\bar{x} \in S_1 \cap S_2$. Let $\mathcal{C}_1$, $\mathcal{C}_2$, be WDC approximating multicones of S_1, S_2 at $\bar{x}$. Assume that $\mathcal{C}_1$ and $\mathcal{C}_2$ are strongly transversal. Then S_1 and S_2 are not locally separated at $\bar{x}$. (That is, there exists a sequence $\{x_j\}_{j \in \mathbb{N}}$ of points of $(S_1 \cap S_2) \backslash \{\bar{x}\}$ such that $\lim_{j \to \infty} x_j = \bar{x}$.) Furthermore, there exists a Lipschitz arc $\gamma : [0, 1] \mapsto \mathbb{R}^n$ such that $\gamma(0) = \bar{x}$, $\gamma(t)$ does not identically equal $\bar{x}$, and $\gamma(t) \in S_1 \cap S_2$ for all $t \in [0, 1]$.

IV. The Maximum Principle

If U is a set, a U***-control*** is a U-valued function η whose domain $\mathrm{Do}\,\eta$ is a nonempty compact subinterval of $\mathbb{R}$.

We will consider ***Lipschitz control systems with a fixed time interval***,

$$\begin{aligned} \dot{\xi}(t) &= f(\xi(t), \eta(t), t) \quad \text{for a.e. } t \in [a, b], \\ \eta(t) &\in U \quad \text{for all } t \in [a, b], \\ \xi(\cdot) &\in \mathcal{W}^{1,1}(M), \quad \eta(\cdot) \in \mathcal{U}, \quad \text{and } \mathrm{Do}\,\xi = \mathrm{Do}\,\eta, \end{aligned}$$

specified by giving a ***system data*** 7***-tuple*** $\mathcal{D} = (n, \Omega, a, b, f, U, \mathcal{U})$ such that

(H1) *n is a nonnegative integer and Ω (the state space) is an open subset of $\mathbb{R}^n$;*
(H2) *U (the control space) is a set;*
(H3) *$[a, b]$ (the time interval) is a nonempty compact subinterval of $\mathbb{R}$;*
(H3) *f (the dynamical law) is a family $\{f_u\}_{u \in U}$ of maps $f_u : \Omega \times [a, b] \mapsto \mathbb{R}^n$;*
(H4) *$\mathcal{U}$ (the class of admissible controllers) is a set of U-controls defined on $[a, b]$.*

Given such a $\mathcal{D}$,

- We use $f(x, u, t)$ as an alternative notation for $f_u(x, t)$.
- If η is a U-control defined on $[a, b]$, then
  - the expression f_η denotes the map $\Omega \times [a, b] \ni (x, t) \mapsto f(x, \eta(t), t) \in \mathbb{R}^n$;

- if $t \in [a,b]$, then $f_{\eta,t}$ denotes the vector field $\Omega \ni x \mapsto f(x,\eta(t),t) \in \mathbb{R}^n$;
- if $\xi : [\alpha,\beta] \mapsto \Omega$ and $[\alpha,\beta] \subseteq [a,b]$, , then $f_{\eta,\xi}(t) \stackrel{\text{def}}{=} f(\xi(t),\eta(t),t)$ for $t \in [\alpha,\beta]$;
- a ***trajectory*** for η is an absolutely continuous curve $\xi : [a,b] \mapsto \Omega$ such that $\dot{\xi}(t) = f(\xi(t),\eta(t),t)$ for almost all $t \in [a,b]$;

- A ***trajectory-control pair*** (abbr. TCP) is a pair (ξ,η) such that η is a U-control and ξ is a trajectory for η.
- If $\gamma = (\xi,\eta)$ is a TCP, then the ***domain*** $\mathrm{Do}\,\gamma$ is the set $\mathrm{Do}\,\eta$, which, by definition, is the same as $\mathrm{Do}\,\xi$.
- An ***admissible control*** is a member of $\mathcal{U}$.
- A TCP (ξ,η) is ***admissible*** if $\eta \in \mathcal{U}$.
- We write $TCP(\mathcal{D})$, $TCP_{adm}(\mathcal{D})$, to denote, respectively, the set of all TCPs of $\mathcal{D}$ and the set of all admissible TCPs of $\mathcal{D}$.

In addition, we specify x_*, S such that

(H5) $x_* \in \Omega$ *and* $S \subseteq \Omega$;

as well as a ***reference trajectory-control pair*** (ξ_*,η_*), that is, a pair (ξ_*,η_*) such that

(H6.a) $(\xi_*,\eta_*) \in TCP_{adm}(\mathcal{D})$,

(H6.b) $\xi_*(a) = x_*$, and $\xi_*(b_*) \in S$.

In order to state precisely the technical hypotheses on the map f, we first let $\mathcal{U}^c_{[a,b]}$ denote the set of all constant U-controls defined on $[a,b]$, and define $\mathcal{U}^{c,*}_{[a,b]} = \mathcal{U}^c_{[a,b]} \cup \{\eta_*\}$, so $\mathcal{U}^{c,*}_{[a,b]}$ consists of the reference control η_* and all the constant controls whose domain is $[a,b]$.

The key technical hypothesis on our control dynamical law is then

(H7) *For each* $\eta \in \mathcal{U}^{c,*}_{[a,b]}$, *the vector field* f_η *is integrably Lipschitz near* ξ_*.

(This means that: (H7.i) the map $[a,b] \ni t \mapsto f(x,\eta(t),t)$ is measurable for each $x \in \Omega$, and (H.7ii) there exist an integrable function $k_\eta : [a,b] \mapsto [0,+\infty]$ and a positive number δ_η such that $\|f(x,\eta(t),t)\| \le k_\eta(t)$ and $\|f(x,\eta(t),t) - f(x',\eta(t),t)\| \le k_\eta(t)\|x-x'\|$ whenever $\max(\|x-\xi_*(t)\|,\|x'-\xi_*(t)\|) \le \delta_\eta$ and $t \in [a,b]$.)

In addition, we also specify $\mathcal{C}$ such that

(H8) $\mathcal{C}$ *is a WDC approximating multicone of* S *at* $\xi_*(b)$,

Our last hypothesis will require the concept of an *equal-time measurable-variational neighborhood* (abbr. ETMVN) of a controller η. We say that a set $\mathcal{V}$ of controllers is an ETMVN of a controller η if

- *for every* $N \in \mathbb{Z}_+$ *and every* N*-tuple* $\mathbf{u} = (u_1,\ldots,u_N)$ *of members of* U, *there exists a positive number* $\varepsilon = \varepsilon(N,\mathbf{u})$ *such that whenever* $\eta : [a,b] \mapsto U$ *is a map obtained from* η_* *by first selecting an* N*-tuple* $\mathbf{M} = (M_1,\ldots,M_M)$ *of pairwise disjoint measurable subsets of* $[a,b]$ *with the property that* $\sum_{j=1}^M \mathrm{meas}(M_j) \le \varepsilon$, *and then substituting the constant value* u_j *for the value* $\eta_*(t)$ *for every* $j = 1,\ldots,N$ *and every* $t \in M_j$, *it follows that* $\eta \in \mathcal{U}$.

We will then assume

(H9) *The class* $\mathcal{U}$ *is an ETMVN of* η_*.

A. *The maximum principle for set separation*

For the set separation problem, we specify a data 12-tuple

$$\mathcal{D}^{sep} = (n,\Omega,a,b,f,U,\mathcal{U},x_*,S,\xi_*,\eta_*,\mathcal{C})\,. \qquad (2)$$

We let $\mathcal{D} = (n,\Omega,a,b,f,U,\mathcal{U})$, and define the ***$\mathcal{D}$-reachable set from x_* over*** $[a,b]$ to be the set $\mathcal{R}_{\mathcal{D};[a,b]}(x_*) \stackrel{\text{def}}{=} \{\xi(b) : (\xi,\eta) \in TCP_{adm}(\mathcal{D})\,,\ \xi(a_*) = x_*\}$.

The local separation condition is then

(H^{sep}) *there exists a neighborhood* V *of* $\xi_*(b)$ *in* Ω *such that* $\mathcal{R}_{\mathcal{D};[a,b]}(x_*) \cap S \cap V = \{\xi(b)\}$.

It will also be convenient to single out the following strong form of the negation of H^{sep}, that we will call the ***Lispchitz arc intersection property***:

($\mathrm{H}^{Lip,in}$) *There exists a nonconstant Lipschitz arc* $\gamma : [0,1] \mapsto \mathcal{R}_{\mathcal{D};[a,b]}(x_*) \cap S$ *having the property that* $\gamma(0) = \xi_*(b)$ *and* $\gamma(1) \ne \gamma(0)$.

We define the ***Hamiltonian of*** f by first writing $Z = \Omega \times \mathbb{R}_n \times U \times [a,b]$, and then letting $H^f : Z \mapsto \mathbb{R}$ be the function given by the formula $H^f(z) = p \cdot f(x,u,t)$, for $z = (x,p,u,t) \in Z$.

The following is then our version of the Lipschitz maximum principle for set separation. The proof is given in [24]..

Theorem 4.1: Assume that the data $\mathcal{D}^{sep}$ satisfy Hypotheses (H1) to (H9). Let $\mathcal{L}$ be the set of all pairs (u,τ) such that $u \in U$, $\tau \in]a,b[$, and τ is a Lebesgue point of both functions $t \mapsto f(\xi_*(t),u,t)$ and $t \mapsto f(\xi_*(t),\eta_*(t),t)$. Then either ($\mathrm{H}^{Lip,in}$) holds, or

(*) for every covector $\mu \in \mathbb{R}_n \backslash \{0\}$ there exists a triple $(\pi_0,\pi^\#,L)$ having the property that (i) L is a measurable selection of the set-valued map $[a,b] \ni t \mapsto \partial f_{\eta_*,t}(\xi_*(t)) \subseteq \mathbb{R}^{n\times n}$, (ii) $\pi_0 \in [0,+\infty[$, and (iii) $\pi^\# \in \mathbb{R}_n$, such that, if $\pi : [a,b] \mapsto \mathbb{R}_n$ is the unique absolutely continuous solution of the adjoint Cauchy problem $\dot{\pi}(t) = -\pi(t)\cdot L(t)$, $\pi(b) = \pi^\#$, then the following three conditions are satisfied:

I. ***Hamiltonian maximization:*** The inequality $H^f(\xi_*(\tau),\pi(\tau),\eta_*(\tau),\tau) \ge H^f(\xi_*(\tau),\pi(\tau),u,\tau)$ holds whenever $(u,\tau) \in \mathcal{L}$.

II. ***Transversality:*** $\pi_0\mu - \pi^\# \in \mathcal{C}^\perp$.

III. ***Nontriviality:*** $(\pi_0,\pi^\#) \ne (0,0)$.

In particular, if the local separation condition H^{sep} is satisfied, then (*) holds. ∎

B. *The maximum principle for optimal control*

We now show how the maximum principle for Lipschitz optimal control (Clarke [5]) follows from Theorem 4.1.

We consider a ***fixed time-interval Lagrangian optimal control problem***

$$\begin{array}{ll}\text{minimize} & \int_a^b f_0(\xi(t),\eta(t),t)\,dt \\ \text{subject to} & \left\{\begin{array}{l}\xi(\cdot)\in W^{1,1}([a,b],\mathbb{R}^n)\,,\\ \dot\xi(t)=f(\xi(t),\eta(t),t)\ \text{ a.e.}\,,\\ \xi(a)=x_*\ \text{ and }\ \xi(b)\in S\,,\\ \eta(t)\in U\ \text{ for all } t\in[a,b]\,,\ \text{ and }\ \eta(\cdot)\in\mathcal{U}\,.\end{array}\right.\end{array}$$

We assume, as before, that we are given a *reference trajectory-control pair* (ξ_*,η_*). And, finally, we assume that we are given a multicone $\mathcal{C}$. So we are specifying a data 12-tuple $\mathcal{D}^{sep}$ as in (2), and we will assume that all the conditions (H1) to (H9) hold.

In addition to $\mathcal{D}^{sep}$, we now need to specify a cost functional. For this purpose, we give f_0, such that

(H10) *f_0 is a real-valued function on $\Omega\times U\times[a,b]$.*

Then, if $\mathcal{D}^{opt}=(n,\Omega,a,b,f,U,\mathcal{U},x_*,S,\xi_*,\eta_*,\mathcal{C},f_0)$ is our data 13-tuple,

- A TCP (ξ,η) with domain $[a,b]$ is ***endpoint-cost-admissible*** if it satisfies: (i) (ξ,η) is admissible, (ii) $\xi(a)=x_*$, (iii) $\xi(b)\in S$, and, finally (iv) the function $[a,b]\ni t\mapsto f_0(\xi(t),\eta(t),t)\in\mathbb{R}$ is a. e. defined, measurable, and such that $\int_\alpha^\beta \min\Big(0,f_0(\xi(t),\eta(t),t)\Big)\,dt>-\infty$.
- We write $TCP_{adm,ec}(\mathcal{D}^{opt})$ to denote the set of all TCPs of $\mathcal{D}^{opt}$ that are endpoint-cost-admissible.

It follows that if (ξ,η) belongs to $TCP_{adm,ec}(\mathcal{D}^{opt})$ then the number $J(\xi,\eta)=\int_a^b f_0(\xi(t),\eta(t),t)\,dt$ —called the ***cost*** of (ξ,η)—is well defined and belongs to $]-\infty,+\infty]$.

To the data $\mathcal{D}^{opt}$, we associate the map $\mathbf{f}:\Omega\times U\times[a,b]\mapsto\mathbb{R}\times\mathbb{R}^n$, defined by letting $\mathbf{f}(z)=(f_0(z),f(z))$ for $z=(x,u,t)\in\Omega\times U\times[a,b]$. We refer to $\mathbf{f}$ as the ***augmented dynamics***.

If η is a U-control, we write $f_{0,\eta}(x,t)=f_0(x,\eta(t),t)$, $\mathbf{f}_\eta(x,t)=\mathbf{f}(x,\eta(t),t)$, so $f_{0,\eta}$ is a function from $\Omega\times[a,b]$ to $\mathbb{R}$, and $\mathbf{f}_\eta$ is a map from $\Omega\times[a,b]$ to $\mathbb{R}^n$. If in addition $\xi:[\alpha,\beta]\mapsto\Omega$ is an arc in Ω, we write $f_{0,\eta,\xi}(t)=f_0(\xi(t),\eta(t),t)$, $\mathbf{f}_{\eta,\xi}(t)=\mathbf{f}(\xi(t),\eta(t),t)$,

The precise technical hypothesis on f_0 is

(H11) *for each control $\eta\in\mathcal{U}^{c,*}_{[a,b]}$, the time-varying function $f_{0,\eta}$ is integrably Lipschitz near ξ_*.*

(The definition of the "integrably Lipschitz" property for real-valued functions is identical to that for vector fields, with the obvious trivial modifications.)

We write $\xi_{0,*}(t)=\int_a^t f_0(\xi_*(s),\eta_*(s),s)\,ds$, so the function $\xi_{0,*}$ is the ***running Lagrangian cost*** along (ξ_*,η_*), initialized so that $\xi_{0,*}(a)=0$. We then let $\Xi_*(t)=(\xi_{0,*}(t),\xi_*(t))$, so $\Xi_*:[a,b]\mapsto\mathbb{R}\times M$ is the ***cost-augmented reference trajectory***. Clearly, Ξ_* is an integral curve of $\mathbf{f}_{\eta_*}$, if we regard $\mathbf{f}_{\eta_*}$ as a time-varying vector field on $\mathbb{R}\times\Omega$, as explained above, and our assumptions imply that $\mathbf{f}_{\eta_*}$ is integrably Lipschitz near Ξ_*. This makes it possible to talk about the Clarke generalized Jacobian $\partial\mathbf{f}_{\eta_*,t}(\Xi_*(t))$, for which we will also use the notation $\partial\mathbf{f}_{\eta_*,t}(\xi_*(t))$, since $\mathbf{f}_{\eta_*,t}$ does not depend on the first component. Then $\partial\mathbf{f}_{\eta_*,t}(\xi_*(t))$ is a compact convex subset of the space $\mathbb{R}_n\times\mathbb{R}^{n\times n}$. So every member of $\partial\mathbf{f}_{\eta_*,t}(\xi_*(t))$ can be regarded as a pair (L_0,L), where $L_0\in\mathbb{R}_n$ and $L\in\mathbb{R}^{n\times n}$. It is then clear that

$$\partial\mathbf{f}_{\eta_*,t}(\xi_*(t))\subseteq\partial f_{0,\eta_*,t}(\xi_*(t))\times\partial f_{\eta_*,t}(\xi_*(t))\,. \tag{3}$$

It then follows that every measurable selection of the set-valued map $[a,b]\ni t\mapsto\partial\mathbf{f}_{\eta_*,t}(\xi_*(t))$ can be regarded as a pair (L_0,L), where

(i) $L_0:[a,b]\mapsto\mathbb{R}_n$ is a measurable selection of $[a,b]\ni t\mapsto\partial f_{0,\eta_*,t}(\xi_*(t))$,
(ii) $L:[a,b]\mapsto\mathbb{R}^{n\times n}$ is a measurable selection of the map $[a,b]\ni t\mapsto\partial f_{\eta_*,t}(\xi_*(t))$,

Then we can write the "inhomogeneous adjoint equation" $\dot\pi=-\pi\cdot L+\pi_0L_0$, for any real number π_0. The corresponding Cauchy problem, with terminal condition $\pi(b)=\pi^\#$, clearly has a unique solution π for any given $(L_0,L,\pi_0,\pi^\#)$. A field π of covectors (or a pair (π_0,π)) arising in this way is known as an ***adjoint covector***, or ***adjoint vector***.

The hypothesis on the reference TCP (ξ_*,η_*) is that it is a local cost-minimizer in $TCP_{adm,ec}(\mathcal{D}^{opt})$. In other words,

(H^{opt}) *$(\xi_*,\eta_*)\in TCP_{adm,ec}(\mathcal{D}^{opt})$, and there exists a neighborhood V of $\xi_*(b)$ in Ω having the property that $J(\xi_*,\eta_*)\le J(\xi,\eta)$ for all pairs $(\xi,\eta)\in TCP_{adm,ec}(\mathcal{D}^{opt})$ such that $\xi(b)\in V$.*

We also single out the following very strong form of the negation of H^{opt}, that we will call the ***Lispchitz arc nonoptimality property***:

($\mathrm{H}^{Lip,no}$) *There exist a nonconstant Lipschitz arc $[0,1]\ni s\mapsto\gamma(s)=(\gamma_0(s),\vec\gamma(s))\in\mathbb{R}\times\Omega$ and TCPs $(\xi_s,\eta_s)\in TCP_{adm,ec}(\mathcal{D}^{opt})$ for $s\in[0,1]$, such that (i) $(\xi_0,\eta_0)=(\xi_*,\eta_*)$, and (ii) $\gamma(s)=(J(\xi_s,\eta_s),\xi_s(b))$ and $J(\xi_s,\eta_s)<J(\xi_*,\eta_*)$ for $0<s\le 1$.*

We define the ***Hamiltonian of*** $\mathbf{f}$ to be the parametrized family of functions $H^{\mathbf{f}}_\alpha:\Omega\times\mathbb{R}_n\times U\times[a,b]\mapsto\mathbb{R}$, (depending on the real parameter α), given by $H^{\mathbf{f}}_\alpha(x,p,u,t)=p\cdot f(x,u,t)-\alpha f_0(x,u,t)\,$,

The following is our stronger form of the maximum principle for Lipschitz optimal control problems, extending the result proved by Clarke in [5]. The proof is given in [24].

Theorem 4.2: Assume that the data 13-tuple $\mathcal{D}^{opt}$ satisfies Hypotheses (H1) to (H11). Let $\mathcal{L}$ denote the set of all pairs (u,τ) such that $u\in U$, $\tau\in]a_*,b_*[$, and τ is a Lebesgue time of both maps $t\mapsto\mathbf{f}(\xi_*(t),u,t)$ and $t\mapsto\mathbf{f}(\xi_*(t),\eta_*(t),t)$. Then either ($\mathrm{H}^{Lip,no}$) holds, or

(*) there exist
1. a pair $(\pi_0,\pi^\#)\in[0,+\infty[\times\mathbb{R}_n$,
2. a measurable selection

$$[a,b]\ni t\mapsto(L_0(t),L(t))\in\partial\mathbf{f}_{\eta_*,t}(\xi_*(t))$$

of the set-valued map $t \mapsto \partial \mathbf{f}_{\eta_*,t}(\xi_*(t))$,

having the property that, if $\pi : [a,b] \mapsto \mathbb{R}_n$ is the unique absolutely continuous solution of the "inhomogeneous adjoint Cauchy problem" $\dot{\pi}(t) = -\pi(t) \cdot L(t) + \pi_0 L_0(t)$, $\pi(b) = \pi^{\#}$, then the following three conditions are satisfied:

I. ***Hamiltonian maximization:*** the inequality $H^f_{\pi_0}(\xi_*(\tau), \eta_*(\tau), \tau) \geq H^f_{\pi_0}(\xi_*(\tau), u, \tau)$ is satisfied

whenever $(u, \tau) \in \mathcal{L}$,

II ***Transversality:*** $-\pi^{\#} \in \mathcal{C}^{\perp}$,

III. ***Nontriviality:*** $(\pi_0, \pi^{\#}) \neq (0, 0)$.

In particular, if the optimality condition (H^{opt}) is satisfied then (*) holds. ■

An example. We now show, by means of an example, how our result is stronger than the usual Lipschitz maximum principle.

Consider the optimal control problem in $\mathbb{R}^2$ in which it is desired to maximize the integral $\int_0^1 u(t)\,dt$—i.e., to minimize $\int_0^1 (-u(t))\,dt$—subject to $\dot{x}(t) = u(t)$, $\dot{y}(t) = v(t)$, $u \geq 0$, $v \in \mathbb{R}$, $(x(0), y(0)) = (0, 0)$, and $(x(1), y(1)) \in S$, where the target set S is given by $S = \{(x, x \sin 1/x) : x > 0\} \cup \{(0, 0)\}$. The constant trajectory $x(t) \equiv 0$, $y(t) \equiv 0$, corresponding to the control $u(t) \equiv 0$, $v(t) \equiv 0$, is obviously not optimal. The Mordukhovich normal cone N to S at $(0, 0)$ is $\mathbb{R}_2$, which means that the necessary conditions for optimality of the usual Lipschitz maximum principle—including the transversality condition $-\pi(1) \in N$—are satisfied, because N also the Mordukhovich normal cone to the set $\hat{S} = \{(0, 0)\}$, and if the target set was $\hat{S}$ rather than S then our reference TCP would be optimal. On the other hand, if we consider the optimal control problem with target $\tilde{S}$, where $\tilde{S} = \{(x, y) : |y| \leq x\}$, then a simple calculation shows that our reference TCP does not satisfy (*). (The adjoint equation obviously says that the momentum $\pi(t) = (\pi_1(t), \pi_2(t))$ must be constant. The Hamiltonian is $\pi_1 u + \pi_2 v + \pi_0 u$, so Hamiltonian maximimization implies $\pi_2 = 0$ and $\pi_1 + \pi_0 \leq 0$. Then $\pi_1 \leq 0$, while $-\pi(1)$ cannot be polar to $\tilde{S}$ at $(0, 0)$ unless $\pi_1 \geq 0$. So $\pi_1 = 0$, and then $\pi_0 = 0$, contradicting the nontriviality condition.) This implies that there exists a Lipschitz arc $\gamma : [0, 1] \mapsto \mathbb{R}^3$ such that $\gamma(0) = (0, 0, 0)$ and, if we write $\gamma(s) = (\gamma_0(s), \vec{\gamma}(s))$, then $\gamma_0(s) = J(\xi_s, \eta_s)$ and $\vec{\gamma}(s) = \xi_s(1)$ for some cost-admissible TCP (ξ_s, η_s) such that, if $s > 0$, then $\xi_s(0) = (0, 0)$, $\xi_s(1) \in \tilde{S}$, and $\gamma_0(s) < \gamma_0(0)$. It follows immediately that $\xi_s(1) \in S$ for some s, so the reference TCP is not optimal. ■

REFERENCES

[1] Berkovitz, L. D., *Optimal Control Theory.* Springer-Verlag, New York, 1974.

[2] Bianchini, R. M., "Variation of a control process at the initial point." J. Opt.Theory and Appl. **81** (1994), pp. 249–258.

[3] Bressan, A., "A high-order test for optimality of bang-bang controls." SIAM J. Control Optimiz. **23** (1985), pp. 38–48.

[4] Bressan, A., "On the intersection of a Clarke cone with a Boltyanskii cone." SIAM J. Control Optim., to appear.

[5] Clarke, F. H., *The Maximum Principle under minimal hypotheses.* SIAM J. Control Optim. **14**, 1976, pp. 1078–1091.

[6] Clarke, F. H., *Optimization and Nonsmooth Analysis.* Wiley Interscience, New York, 1983.

[7] Clarke, F. H., Yu. S. Ledyaev, R. J. Stern, and P. R. Wolenski, *Nonsmooth Analysis and Control Theory*. Springer Verlag, Graduate Texts in Mathematics No. 178; New York, 1998.

[8] Ioffe. A., "Euler-Lagrange and Hamiltonian formalisms in dynamic optimization." Trans. Amer. Math. Soc. **349** (1997), pp. 2871–2900

[9] Ioffe, A., and R. T. Rockafellar, "The Euler and Weierstrass conditions for nonsmooth variational problems." Calc. Var. Partial Differential Equations **4** (1996), pp. 59–87.

[10] Knobloch, H. W., "High Order Necessary Conditions in Optimal Control." Springer-Verlag, Berlin, 1975.

[11] Krener, A. J. , "The High Order Maximal Principle and Its Application to Singular Extremals." SIAM J. Contrl. Optim. **15** (1977), pp. 256-293.

[12] Łojasiewicz Jr., S, personal communication, unpublished.

[13] Mordukhovich, B., *Variational Analysis and Generalized Differentiation, I: Basic Theory; II: Applications.* Grundlehren Series (Fundamental Principles of Mathematical Sciences), Vol. 330 and 331, Springer Verlag, Berlin, 2006.

[14] Pontryagin, L. S., V. G. Boltyanskii, R. V. Gamkrelidze and E. F. Mischenko, *The Mathematical Theory of Optimal Processes.* Wiley, New York, 1962.

[15] Stefani, G., "On maximum principles." In *Analysis of Controlled Dynamical Systems*, Bonnard, Bride, Gauthier, Kupka, eds., Progress in Systems and Control Theory, vol. 8, Birkhäuser, Boston 1991.

[16] Sussmann, H. J., "A general theorem on local controllability." SIAM J. Control Otim. **25** (1987), pp. 158–194.

[17] Sussmann, H. J., "A strong version of the Łojasiewicz Maximum Principle." In *Optimal Control of Differential Equations, (Athens, OH, 1993)*, Nicolai H. Pavel Ed., Lect. Notes in Pure and Applied Math. **160**, Marcel Dekker Inc., New York, 1994, pp. 293-309.

[18] Sussmann, H. J., "A nonsmooth hybrid maximum principle." In *Stability and Stabilization of Nonlinear Systems*, D. Aeyels, F. Lamnabhi-Lagarrigue and A. J. van der Schaft Eds. Lect. Notes Control and Information Sciences **246**, Springer-Verlag, 1999, pp. 325–354.

[19] Sussmann, H.J., "New theories of set-valued differentials and new versions of the maximum principle of optimal control theory." In *Nonlinear Control in the year 2000*, A. Isidori, F. Lamnabhi-Lagarrigue and W. Respondek Eds., Springer-Verlag, London, 2000, pp. 487-526.

[20] Sussmann, H.J., "Optimal control of nonsmooth systems with classically differentiable flow maps." In *Proc. 6th IFAC Symposium on Nonlinear Control Systems (NOLCOS 2004), Stuttgart, Germany, September 1-3, 2004*, Vol. 2, pp. 609-704.

[21] Sussmann, H.J., "Combining high-order necessary conditions for optimality with nonsmoothness." In *Proc. 43rd IEEE 2004 Conference on Decision and Control (Paradise Island, the Bahamas, December 14-17, 2004)*, IEEE Publications, New York, 2004.

[22] Sussmann, H.J., "Set transversality, approximating multicones, Warga derivate containers and Mordukhovich cones." In the *Proceedings of the 44th IEEE 2005 Conference on Decision and Control, held in Sevilla, Spain, December 12-15, 2005)*, IEEE Publications, New York, 2005.

[23] Sussmann, H.J., "Generalized differentials, variational generators, and the maximum principle with state constraints." To appear in *Nonlinear and Optimal Control Theory; Cetraro, Italy, June 2004, proceedings of a summer school of the Fondazione C.I.M.E.* (Book coauthored by Andrei Agrachev, Stephen Morse, Eduardo Sontag, Héctor Sussmann, and Vadim Utkin.) Gianna Stefani and Paolo Nistri Eds. Springer-Verlag Lecture Notes in Mathematics.

[24] Sussmann, H.J., "Set separation, approximating multicones, and the Lipschitz maximum principle." *J. Diff. Equations*, to appear.

[25] Vinter, R. B., *Optimal Control*. Birkhäuser, Boston, 2000.

[26] Warga, J., "Optimization and controllability without differentiability assumptions." *SIAM J. Control and Optimization* **21**, 1983, pp. 837–855.

[27] Warga, J., "Homeomorphisms and local C^1 approximations." *Nonlinear Anal. TMA* **12**, 1988, pp. 593–597.