On the validity of the transversality condition for different concepts of tangent cone to a set

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*— In the nonsmooth versions of the Pontryagin Maximum Principle, the transversality condition involves a normal cone to the terminal set. General versions of the principle for highly non-smooth systems have been proved by separation methods for cases that include, for example, a reference vector field which is classically differentiable along the reference trajectory but not Lipschitz. In these versions, the notion of normal cone used is that of the polar of a Boltyanskii approximating cone. Using a recent result of A. Bressan, we prove that these versions can fail to be true if the Clarke normal cone (and, a fortiori, any smaller normal cone, such as the Mordukhovich cone) is used instead. The key fact is A. Bressan's recent example of two closed sets that intersect at a point p and are such that (a) one of the sets has a Boltyanskii approximating cone C_1 at p, (b) the other set has a Clarke tangent cone C_2 at p, and (c) the cones C_1 and C_2 are strongly transversal, but (d) the sets only intersect at p.**

I. Introduction

In a series of papers (e.g., [2], [3], [4], [5], [6], [7], [8], [9], [10], [11], [12], [13], [14], [15], [16], [17], [18], [19]) we have proposed versions of the Pontryagin Maximum Principle for highly non-smooth systems, based on generalized differentials and flows, and proved by "primal" methods, using packets of needle variations. All these proofs are based on separation theorems for sets, which assert that a necessary condition for two sets S_1, S_2 containing a point p to be separated at p (in the sense that $S_1 \cap S_2 = \{p\}$) is that C_1 and C_2 not be strongly transversal (cf. Definition 3.2 below), where C_1 and C_2 are "tangent cones" to S_1 and S_2 at p. These separation theorems are true if "tangent cone" is interpreted to mean "Boltyanskii approximating cone" and also if it is taken to mean "Clarke tangent cone."

It had been an open question for several years whether there exists a more general theory that contains both separation results, i.e., whether there exists a concept of (non-unique) "tangent cone to a set" such that both Boltyanskii cones and Clarke cones are "tangent" in this more general sense, and the separation theorem is still true.

Recently, A. Bressan has shown (cf. [1]) that such a concept cannot exist. Clearly, if it did exist, then it would follow in particular that, if two sets S_1, S_2 have tangent cones C_1, C_2 at p, one of them in the Boltyanskii sense and the other one in the Clarke one, then it is necessary for S_1 and S_2 to be separated at p that C_1 and C_2 not be strongly transversal. Bressan's paper [1] gives a very nice construction that provides a counterexample to this assertion.

Using Bressan's counterexample, we will construct an optimal control problem and an optimal trajectory for which the usual conclusions of the Maximum Princple are not true, if the Clarke tangent cone to the terminal set is used in the transversality condition instead of a Boltyanskii approximating cone.

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

II. Boltyanskii and Clarke cones

We begin by reviewing the concepts of "Boltyanskii approximating cone," "Bouligand tangent cone," and "Clarke tangent cone" to a set at a point.

In this paper, a ***cone*** in a real linear space X is a nonempty subset C of X such that $rc \in C$ whenever $r \in \mathbb{R}$, $r \geq 0$, and $c \in C$. (In particular, if C is a cone then necessarily $0 \in C$.)

Let S be a subset of $\mathbb{R}^n$ and let $p \in S$.

Definition 2.1: A ***Boltyanskii approximating cone*** to S at p is a convex cone C in $\mathbb{R}^n$ having the property that there exist

(i) a nonnegative integer m,
(ii) a closed convex cone D in $\mathbb{R}^m$,
(iii) a neighborhood U of 0 in $\mathbb{R}^m$,
(iv) a continuous map $F : U \cap D \mapsto S$,
(v) a linear map $L : \mathbb{R}^m \mapsto \mathbb{R}^n$,

such that

$$F(x) = p + Lx + o(\|x\|) \quad \text{as} \quad x \to 0\,,\ \ x \in D\,,$$

and $LD = C$. ■

Definition 2.2: The ***Bouligand tangent cone*** to S at p is the set of all vectors $v \in \mathbb{R}^n$ such that there exist

(i) a sequence $\{p_j\}_{j\in\mathbb{N}}$ of points of S converging to p,
(ii) a sequence $\{h_j\}_{j\in\mathbb{N}}$ of positive real numbers converging to 0,

such that

$$v = \lim_{j\to\infty} \frac{p_j - p}{h_j}\,. \qquad \blacksquare$$

We use $T_p^B S$ to denote the Bouligand tangent cone to S at p. It is then clear that $T_p^B S$ is a closed cone.

Definition 2.3: If S is closed, then the ***Clarke tangent cone*** to S at p is the set of all vectors $v \in \mathbb{R}^n$ such that, whenever $\{p_j\}_{j\in\mathbb{N}}$ is a sequence of points of S converging to p, it follows that there exist Bouligand tangent vectors $v_j \in T_{p_j}^B S$ such that $\lim_{j\to\infty} v_j = v$. $\blacksquare$

We use $T_p^C S$ to denote the Clarke tangent cone to S at p. Then $T_p^C S$ is a closed convex cone.

III. Transversality

The ***polar*** of a cone C in $\mathbb{R}^n$ is the set $C^\perp$ of all $w \in \mathbb{R}^n$ such that $\langle w, c\rangle \le 0$ for all $c \in C$. It is clear that $C^\perp$ is always a closed convex cone, and $C^{\perp\perp}$ is the smallest closed convex cone containing C, so in particular $C^{\perp\perp} = C$ if and only if C is closed and convex.

Definition 3.1: Two convex cones C_1, C_2 in $\mathbb{R}^n$ are ***transversal*** if $C_1 - C_2 = \mathbb{R}^n$, i.e., if for every $x \in \mathbb{R}^n$ there exist $c_1 \in C_1$, $c_2 \in C_2$, such that $x = c_1 - c_2$. $\blacksquare$

Definition 3.2: Two convex cones C_1, C_2 in $\mathbb{R}^n$ are ***strongly transversal*** if they are transversal and in addition $C_1 \cap C_2 \neq \{0\}$. $\blacksquare$

The following trivial observation says that transversality and strong transversality are almost equivalent, and that the only gap between the two conditions occurs when the cones C_1 and C_2 are linear subspaces such that $C_1 \oplus C_2 = \mathbb{R}^n$, in which case C_1 and C_2 are transversal but not strongly transversal.

Lemma 3.3: If C_1, C_2 are convex cones in $\mathbb{R}^n$, then C_1 and C_2 are transversal if and only if either

(i) C_1 and C_2 are strongly transversal,

or

(ii) C_1 and C_2 are linear subspaces and $C_1 \oplus C_2 = \mathbb{R}^n$.

Proof: It suffices to assume that C_1 and C_2 are transversal but not strongly transversal and show that (ii) holds. Let us prove that C_1 is a linear subspace. Pick $c \in C_1$. Using the transversality of C_1 and C_2 write $-c = c_1 - c_2$, $c_i \in C_i$. Then $c_1 + c = c_2$. But $c_1 + c \in C_1$ and $c_2 \in C_2$. So $c_1 + c \in C_1 \cap C_2$, and then $c_1 + c = 0$, since C_1 and C_2 are not strongly transversal. Therefore $-c = c_1$, so $-c \in C_1$. This shows that $c \in C_1 \Rightarrow -c \in C_1$. So C_1 is a linear subspace. A similar argument shows that C_2 is a linear subspace. Then the transversality of C_1 and C_2 implies that $C_1 + C_2 = \mathbb{R}^n$, and the fact that they are not strongly transversal implies that $C_1 \cap C_2 = \{0\}$. Hence $C_1 \oplus C_2 = \mathbb{R}^n$. $\blacksquare$

The following fact tells us that nontransversality of C_1 and C_2 is exactly equivalent to the existence of a separating linear functional, i.e. of a $\lambda \in \mathbb{R}^n$ such that $\langle \lambda, c\rangle \le 0$ for all $c \in C_1$ and $\langle \lambda, c\rangle \ge 0$ for all $c \in C_2$. This point is particularly important in conjunction with Lemma 3.3: in the proof of the maximum principle, one constructs a subset of a reachable set which turns out to be separated from another set, and then concludes that the tangent cones to the two sets are not strongly transversal. Lemma 3.3 then tells us (provided that we manage to eliminate the special case when $C_1 \oplus C_2 = \mathbb{R}^n$) that the cones are not transversal, so the separating covector exists.

Lemma 3.4: If C_1, C_2 are convex cones in $\mathbb{R}^n$, then C_1 and C_2 are transversal if and only if

$$C_1^\perp \cap (-C_2^\perp) = \{0\}\,.$$

Proof: Observe that $C_1 - C_2$ is a convex cone, so $C_1 - C_2 = \mathbb{R}^n$ if and only if $\mathrm{Clos}(C_1 - C_2) = \mathbb{R}^n$, and if $\mathrm{Clos}(C_1 - C_2) \neq \mathbb{R}^n$ then the Hahn-Banach theorem implies that $C_1^\perp \cap (-C_2^\perp) \neq \{0\}$. $\blacksquare$

IV. Set separation

Two subsets S_1, S_2 of $\mathbb{R}^n$ containing a point $p \in \mathbb{R}^n$ are ***separated*** at p if $S_1 \cap S_2 = \{p\}$. We say that S_1 and S_2 are ***locally separated*** at p if there exists a neighborhood V of p such that $S_1 \cap S_2 \cap V = \{p\}$.

The following results are well known.

Proposition 4.1: Let S_1, S_2 be subsets of $\mathbb{R}^n$ that are locally separated at a point $p \in \mathbb{R}^n$, and let C_1, C_2 be Boltyanskii approximating cones to S_1, S_2 at p. Then C_1 and C_2 are not strongly transversal. $\blacksquare$

Proposition 4.2: Let S_1, S_2 be closed subsets of $\mathbb{R}^n$ that are locally separated at a point $p \in \mathbb{R}^n$, and let C_1, C_2 be the Clarke tangent cones to S_1, S_2 at p. Then C_1 and C_2 are not strongly transversal. $\blacksquare$

V. Bressan's counterexample

The following result, proved recently by Bressan in [1], answers negatively the natural question whether Propositions 4.1 and 4.2 can be combined into a single more general separation theorem.

Proposition 5.1: There exist two closed subsets S_1, S_2 of $\mathbb{R}^4$ that are separated at 0, and closed convex cones C_1, C_2 in $\mathbb{R}^4$ such that

(i) C_1 is a Boltyanskii approximating cone to S_1 at 0,

(ii) C_2 is the Clarke tangent cone to S_2 at 0,

(iii) C_1 and C_2 are strongly transversal. $\blacksquare$

It will be important for us to know explicitly the cone C_1 of Bressan's example. It turns out that

$$C_1 = \{(\tau, x, y, z) \in \mathbb{R}^4 : \tau \ge 0,\ |z| \le \tau,\ x = y = 0\}\,. \qquad (1)$$

In particular, the cone C_1 is two-dimensional, and is the image of the sector

$$\Sigma = \{(\tau, z) \in \mathbb{R}^2 : \tau \geq 0\,,\ \ |z| \leq \tau\}$$

in the (τ, z) plane under the embedding map $\mathbb{R}^2 \ni (\tau, z) \mapsto (\tau, 0, 0, z) \in \mathbb{R}^4$.

Furthermore, in Bressan's counterexample the number m, the cone D, and the map L of Definition 2.1 can be taken to be 4, C_1 and the identity map, respectively. In other words:

(iv) There exist a neighborhood U of 0 in $\mathbb{R}^4$, and a continuous map $\Gamma : U \cap C_1 \mapsto S_1$, such that

$$\Gamma(v) = v + o(\|v\|) \qquad \text{as} \qquad v \to 0\,,\ v \in C_1\,.$$

VI. The Optimal Control Counterexample

We will first quote a version of the Maximum Principle which is known to be true—because it is a special case of the general results of [5], [8], [9], [10], [11], [12], [13], [14], [15], [16], [18]—in which the transversality condition involves a Boltyanskii cone. This will be done in order to contrast the result with the nearly identical statement where the Clarke tangent cone is used instead, for which we will give a counterexample.

Both statements deal with an optimal control problem in which we are given a *state space*, which will be taken to be $\mathbb{R}^n$, a *control space* U, a *class of admissible controls* $\mathcal{U}$, a *time interval* $[a, b] \subseteq \mathbb{R}$, an *initial state* $\bar{x}$, a *terminal set* S, a *dynamical law*

$$\dot{x} = f(x, u, t)\,,$$

where f is a map from $\mathbb{R}^n \times U \times [a, b]$ to $\mathbb{R}^n$, and a *Lagrangian* L, which is a map from $\mathbb{R}^n \times U \times [a, b]$ to $\mathbb{R}$. It is desired to minimize the integral

$$J(\xi, \eta) = \int_a^b L(\xi(t), \eta(t), t)\, dt$$

in the class of all pairs (ξ, η) that satisfy:

(a1) η is a map from $[a, b]$ to U,

(a2) $\eta \in \mathcal{U}$,

(a3) ξ is an f-trajectory corresponding to η (that is, $\xi : [a, b] \mapsto \mathbb{R}^n$ is absolutely continuous and satisfies $\dot{\xi}(t) = f(\xi(t), \eta(t), t)$ for almost every $t \in [a, b]$),

(a4) $\xi(a) = \bar{x}$ and $\xi(b) \in S$.

(A pair (ξ, η) such that (a1), (a2), and (a3) hold is called a ***trajectory-control pair***, abbr. "TCP." A TCP for which (a4) holds is an ***admissible trajectory-control pair***, abbr. "ATCP.")

We define

$$f^L(x, u, t) = (f(x, u, t), L(x, u, t))$$

for $x \in \mathbb{R}^n$, $u \in U$, $t \in [a, b]$, so f^L is a map from $\mathbb{R}^n \times U \times [a, b]$ to $\mathbb{R}^{n+1}$.

We assume that

(a) U is a separable metric space,

(b) The class $\mathcal{U}$ is a set of maps from $[a, b]$ to U, which is a "variational neighborhood" of η_*, in the following sense: for every positive integer N and every N-tuple $(u_1, \ldots, u_N)$ of members of U there exists a positive number ε such that, whenever $(I_1, \ldots, I_N)$ is an N-tuple of pairwise disjoint subintervals of $[a, b]$ for which $\sum_{j=1}^N \operatorname{meas}(I_j) \leq \varepsilon$, and η is the function obtained from η_* by substituting the constant value u_j for $\eta_*(t)$ for $t \in I_j$, $j = 1, \ldots, N$, it follows that $\eta \in \mathcal{U}$.

(c) $f : \mathbb{R}^n \times U \times [a, b] \mapsto \mathbb{R}^n$ and $L : \mathbb{R}^n \times U \times [a, b] \mapsto \mathbb{R}$ are maps such that

  (c.1) the maps $\mathbb{R}^n \times U \ni (x, u) \mapsto f^L(x, u, t) \in \mathbb{R}^{n+1}$ is continuous for every $t \in \mathbb{R}$,

  (c.2) the map $[a, b] \ni t \mapsto f^L(x, \eta(t), t) \in \mathbb{R}^{n+1}$ is measurable for every $\eta \in \mathcal{U}$, $x \in \mathbb{R}^n$,

  (c.3) for every compact subset K of $\mathbb{R}^n$ and every $\eta \in \mathcal{U}$ there exists an integrable function $\varphi : [a, b] \mapsto [0, +\infty]$ having the property that $\|f^L(x, \eta(t), t)\| \leq \varphi(t)$ whenever $x \in K$ and $t \in [a, b]$,

(d) S is a subset of $\mathbb{R}^n$ and $\bar{x} \in \mathbb{R}^n$.

In addition, we are given an ATCP (ξ_*, η_*) called the "reference ATCP." For this ATCP, we assume that

(i) The map

$$\mathbb{R}^n \ni h \mapsto f^L(\xi_*(t) + h, \eta_*(t), t) \in \mathbb{R}^{n+1}$$

is differentiable at $h = 0$ for almost every $t \in [a, b]$.

(ii) There exist

  (1) an integrable function $k : [a, b] \mapsto [0, +\infty]$

  (2) a positive number δ

such that

$$\|f^L(\xi_*(t) + h, \eta_*(t), t) - f^L(\xi_*(t), \eta_*(t), t)\| \leq k(t)\|h\|$$

whenever $t \in [a, b]$, $h \in \mathbb{R}^n$, and $\|h\| \leq \delta$.

Theorem 6.1: Assume that the above conditions hold, and the reference ATCP (ξ_*, η_*) is optimal, in the sense that $J(\xi_*, \eta_*) \leq J(\xi, \eta)$ for all ATCPs (ξ, η). Assume, moreover, that C is a convex cone in $\mathbb{R}^n$ which is a Boltyanskii approximating cone to S at $\xi_*(b)$. Then there exist an absolutely continuous map $\pi : [a, b] \mapsto \mathbb{R}^n$ and a nonnegative constant π_0 such that

  (I) the *adjoint equation*

$$\dot{\pi}(t) = -\frac{\partial H_{\pi_0}}{\partial x}(\xi_*(t), \pi(t), \eta_*(t, t)$$

  and the *Hamiltonian maximization condition*

$$\begin{aligned} H_{\pi_0}(\xi_*(t), \pi(t), \eta_*(t), t) \\ = \ \max\{H_{\pi_0}(\xi_*(t), \pi(t), u, t) : u \in U\} \end{aligned}$$

hold for almost all $t \in [a,b]$, where, for each $p_0 \in \mathbb{R}$, the Hamiltonian H_{p_0} is the function from $\mathbb{R}^n \times \mathbb{R}^n \times U \times [a,b]$ to $\mathbb{R}$ defined by the formula

$$H_{p_0}(x,p,u,t) \stackrel{\text{def}}{=} \langle p, f(x,u,t)\rangle - p_0 L(x,u,t)\,,$$

(II) $(\pi(b),\pi_0) \neq (0,0)$ (the *nontriviality condition*),

(III) $-\pi(b) \in C^\dagger$ (the *transversality condition*). ∎

The main observation of this paper is then the following.

Theorem 6.2: The statement of Theorem 6.1 is not true if S is assumed to be closed and the words "a Boltyanskii approximating cone to S at $\xi_*(b)$" are replaced by "the Clarke tangent cone to S at $\xi_*(b)$."

VII. Proof of Theorem 6.2

We construct an optimal control problem and an optimal reference ATCP for which the pair (π,π_0) of Theorem 6.1 does not exist.

For this purpose, we use the sets S_1, S_2 and the cones C_1, C_2 of Bressan's counterexample (cf. §V). Then the following conditions hold:

- (c1) $S_1 \subseteq \mathbb{R}^4$ and $S_2 \subseteq \mathbb{R}^4$,
- (c2) $S_1 \cap S_2 = \{0\}$,
- (c3) C_1 and C_2 are closed convex cones in $\mathbb{R}^4$,
- (c4) C_1 and C_2 are strongly transversal,
- (c5) C_1 is a Boltyanskii approximating cone to S_1 at 0, and satisfies Condition (iv) of §V,
- (c6) C_2 is the Clarke tangent cone to S_2 at 0.
- (c7) C_1 is given by (1).

In view of (c5), there exist ρ, Γ such that

- (c8) $\rho \in \mathbb{R}$ and $\rho > 0$,
- (c9) Γ is a continuous map from $\rho\bar{\mathbb{B}}^4 \cap C_1$ into S_1 (where $\rho\bar{\mathbb{B}}^4$ is the closed ball in $\mathbb{R}^4$ with center 0 and radius ρ),
- (c10) $\Gamma(v) = v + o(\|v\|)$ as $v \to 0$ via values in C_1.

It follows from (c4) that

- (c11) $C_1 - C_2 = \mathbb{R}^4$, i.e., there does not exist a nonzero vector $v \in \mathbb{R}^4$ having the property that $\langle v,c\rangle \geq 0$ whenever $c \in C_1$ and $\langle v,c\rangle \leq 0$ whenever $c \in C_2$.

Using (c10), we can assume without loss of generality, by choosing ρ small enough, that

$$\|\Gamma(v) - v\| \leq \frac{1}{2}\|v\| \quad \text{whenever} \;\; v \in \rho\bar{\mathbb{B}}^4 \cap C_1\,.$$

This implies, in particular, that

$$\text{if} \;\; v \in \rho\bar{\mathbb{B}}^4 \cap C_1 \;\; \text{and} \;\; \Gamma(v) = 0 \;\; \text{then} \;\; v = 0\,. \tag{2}$$

We now let

$$U = \rho\bar{\mathbb{B}}^4 \cap C_1\,,$$

and consider the optimal control problem in $\mathbb{R}^8 \sim \mathbb{R}^4 \times \mathbb{R}^4$ with dynamical law

$$\begin{aligned} \dot{x} &= \begin{cases} 0 & \text{if} \quad t < 1 \\ \Gamma(y) & \text{if} \quad t \geq 1\,, \end{cases} \\ \dot{y} &= \begin{cases} u & \text{if} \quad t < 1 \\ 0 & \text{if} \quad t \geq 1 \end{cases} \end{aligned}$$

(where x and y belong to $\mathbb{R}^4$), and control constraint $u \in U$.

We seek to maximize the integral

$$\mathcal{I} = \int_0^2 \|u(t)\|^{1/2} dt$$

in the class of all triples $(x(\cdot),y(\cdot),\eta(\cdot))$ such that $\eta(\cdot) \in \mathcal{U}$, where $\mathcal{U}$ is the set of all measurable functions on $[0,2]$ with values in U, and $x(\cdot)$, $y(\cdot)$ are the components of a trajectory of $u(\cdot)$.

We impose the initial constraint

$$x(0) = 0 \qquad \text{and} \qquad y(0) = 0\,,$$

as well as the terminal constraint

$$x(2) \in S_2 \tag{3}$$

(i.e., $(x(2),y(2)) \in S_2 \times \mathbb{R}^4$).

We let

$$\xi_*(t) \equiv 0\,, \;\; \eta_*(t) \equiv 0\,.$$

Then the TCP (ξ_*,η_*) satisfies the initial and terminal constraints, and the integrand $\mathcal{I}$ to be maximized has the value 0.

We claim that (ξ_*,η_*) is optimal. To see this, we pick an arbitrary TCP (ξ,η) that satisfies the initial and terminal constraints, and prove that $\xi \equiv \xi_*$ and $\eta \equiv \eta_*$.

Write

$$\xi(t) = (x(t),y(t))\,.$$

Then

$$y(1) = \int_0^1 \eta(t)\,dt\,,$$

so $y(1) \in U$, because η is U-valued and U is compact and convex.

Furthemore, it is clear that

$$y(2) = y(1)\,.$$

On the other hand, $x(1) = 0$, and the function $[1,2] \ni t \mapsto y(t) \in \mathbb{R}^4$ is constant and has the value $y(1)$. It follows that $x(2) = \Gamma(y(1))$. Since Γ takes values in S_1, it follows that

$$x(2) \in S_1\,.$$

Then the terminal condition (3) implies that

$$x(2) \in S_1 \cap S_2\,,$$

and it follows from (c2) that $x(2) = 0$.

Therefore

$$\Gamma(y(1)) = 0\,,$$

and then (2) implies that $y(1) = 0$.

If we write

$$\eta(t) = (\tau(t), \hat{x}(t), \hat{y}(t), \hat{z}(t)),$$

then $\tau(t) \geq 0$ for all $t \in [0,1]$ and $\int_0^1 \tau(t)\,dt = 0$, so $\tau(t) = 0$ for almost all $t \in [0,1]$. Then $\hat{z}(t) = 0$ almost everywhere as well, because $|\hat{z}(t)| \leq \tau(t)$.

It follows that $\eta \equiv 0 \equiv \eta_*$, so $\xi \equiv \xi_*$. Hence (ξ_*, η_*) is the *only* TCP that satisfies the initial and terminal constraints, so *a fortiori* (ξ_*, η_*) is optimal.

We now show that the map $\pi : [0,2] \mapsto \mathbb{R}^8$ and the constant π_0 of the Maximum Principle do not exist. Assume they do, and write

$$\pi(t) = (\mu(t), \nu(t)).$$

Then the Hamiltonian is given by

$$H_{\pi_0}(x,y,p,q,u,t) = \begin{cases} q \cdot u + \pi_0\|u\|^{1/2} & \text{if} \quad t < 1 \\ p \cdot \hat{\Gamma}(y) & \text{if} \quad t \geq 1, \end{cases}$$

where $p \in \mathbb{R}^4$, $q \in \mathbb{R}^4$ are, respectively, the momentum variables conjugate to x and y. (In particular, $H_{\pi_0}(x,y,p,q,u,t)$ does not depend on t.)

For the Hamiltonian maximization condition (HMC) to be satisfied, we need $\pi_0 = 0$, because if $\pi_0 > 0$ then, if p is arbitrary, the function

$$U \ni u \mapsto p \cdot u + \pi_0\|u\|^{1/2}$$

does not have a maximum at $u = 0$, since

$$p \cdot u + \pi_0\|u\|^{1/2} = 0 \quad \text{when} \quad u = 0,$$

while, on the other hand,

$$\begin{aligned} p \cdot u + \pi_0\|u\|^{1/2} &= \pi_0\|u\|^{1/2}\Big(1 + \frac{p \cdot u}{\pi_0\|u\|^{1/2}}\Big) \\ &= \pi_0\|u\|^{1/2}\Big(1 + o(1)\Big), \end{aligned}$$

so that $p \cdot u + \pi_0\|u\|^{1/2} > 0$ if $u \neq 0$ and $\|u\|$ is small.

If $t < 1$, then the HMC implies that

$$\nu(t) \cdot u \leq 0 \qquad \text{for all} \quad u \in U,$$

from which it follows easily that

$$\nu(t) \cdot u \leq 0 \qquad \text{for all} \quad u \in C_1.$$

Therefore

$$\nu(t) \in C_1^\dagger \qquad \text{for all} \quad t \in [0,1]. \tag{4}$$

The adjoint equation says that

$$\mu(t) = \text{constant} \qquad \text{for} \quad t \in [0,2],$$

$$\nu(t) = \text{constant} \qquad \text{for} \quad t \in [0,1],$$

and

$$\dot{\nu}(t) = -\mu(t) \qquad \text{for} \quad t \in [1,2],$$

(because the map $y \mapsto \Gamma(y)$ is differentiable at $y = 0$ and its differential at 0 is the identity map).

Hence, if we let

$$\bar{\mu} = \mu(2), \qquad \bar{\nu} = \nu(2),$$

we see that

$$\mu(t) = \bar{\mu} \qquad \text{for} \quad t \in [0,2],$$

$$\nu(t) = (2-t)\bar{\mu} + \bar{\nu} \qquad \text{for} \quad t \in [1,2], \tag{5}$$

and

$$\nu(t) = \bar{\mu} + \bar{\nu} \qquad \text{for} \quad t \in [0,1]. \tag{6}$$

The terminal condition $(x(2), y(2)) \in S_2 \times \mathbb{R}^4$ tells us that

$$-\bar{\mu} \in C_2^\dagger \qquad \text{and} \qquad \bar{\nu} = 0.$$

It then follows from (5) (or from (6)) that

$$\nu(1) = \bar{\mu}.$$

So (4) implies that

$$\bar{\mu} \in C_1^\dagger.$$

We have therefore established that

$$-\bar{\mu} \in C_1^\dagger \cap (-C_2)^\dagger.$$

On the other hand, C_1 and C_2 are transversal, Hence $\bar{\mu} = 0$. Since we have already shown that $\bar{\nu} = 0$, we have in fact proved that

$$\mu(2) = \nu(2) = 0 \qquad \text{and} \qquad \pi_0 = 0.$$

This contradicts the nontriviality condition.

Hence we have completed our proof that a set of multipliers satisfying all the conditions of the conclusion of the Maximum Principle does does not exist, proving our theorem.

VIII. Concluding remarks

The result of Theorem 6.2 shows that the separation methods used to prove theorems such as those of [16] and [18] will definitely not work if the terminal condition involves a set with a given Clarke tangent cone C and the transversality condition is expected to require that the terminal adjoint covector belong to the negative polar $-C^\dagger$ of C. Naturally, then, these methods will not work either if one tries to use a "normal cone" N smaller that $C^\dagger$, such as for example the Mordukhovich normal cone.

On the other hand, these methods are all based on extensions of Proposition 4.1, whose proof depends on a topological argument involving the Brouwer fixed point theorem or some variation thereof.

It is conceivable that there might exist another family of separation theorems, generalizing Proposition 4.2. In [17] we have proposed a concept of "tangent multicone" to a set at a point (called a "Mordukhovich-Warga approximating multicone", henceforth abbreviated as MWAMC), which contains as a particular case the Clarke tangent cone and, more generally, the Mordukhovich tangent multicone. (In the ordinary theory, there exists no such thing as a “Mordukhovich tangent cone,” because the Mordukhovich normal cone is not convex, and therefore it is not the polar of an ordinary cone. In our theory, in which “multicones” are sets of cones, the polar of a convex multicone is a cone which need not be convex, and the Mordukhovich normal cone is the polar of a cone that we call the “Mordukhovich tangent multicone.”)

This suggests the possibility that there might exist another family of versions of the Maximum Principle related to our MWAMCs, or perhaps to some other notion of tangent cone or multicone, or of normal cone, or of normal object of some kind.

The first step towards developing such a theory is the separation theorem for MWAMCs, which was proved in [17], and constitutes a natural extension of Proposition 4.2. The next step would be to derive versions of the Maximum Principle in this new setting. It is clear that the methods would have to be different from those of the theory based on topological arguments. In particular, the proofs of the separation theorems would involve analytical arguments. (Typically, these proofs use the finite-dimensionality in a completely different way from that of the topological arguments. The standard technique is to somehow construct a sequence $\{p_j\}_{j\in\mathbb{N}}$ of “approximate adjoint covectors”, normalized so that $\|p_j\| = 1$, and then pass to the limit and extract a subsequence that convergence to a true adjoint covector $\bar{p}$. The technical problem is then to guarantee that $\bar{p} \neq 0$, which is possible in finite dimensions using ordinary convergence and the compactness of the unit sphere, while in infinite dimensions, even when one can extract a subsequence that converges in some weak sense, it may very well happen that $\bar{p} = 0$.)

The hope is that there will turn out to be two (and no more than two) general versions of the finite-dimensional maximum principle, one involving generalizations of Boltyanksii cones (such as our “GDQ approximating multicones”), the other one involving generalizations of Clarke cones (perhaps the MWAMCs).

The unifying principle of both theories would be the separation theorems. Unfortunately, the counterexample discovered by Bressan shows that the idea of developing one single theory that contains both is almost certainly doomed to failure.

REFERENCES

[1] Bressan, A., “On the intersection of a Clarke cone with a Boltyanskii cone.” January 2006, *SIAM J. Control Optim.*, to appear.

[2] Sussmann, H. J., “A strong version of the Łojasiewicz maximum principle.” In *Optimal Control of Differential Equations*, N. H. Pavel Ed., Marcel Dekker, New York, 1994.

[3] Sussmann, H. J., “A strong version of the maximum principle under weak hypotheses.” In *Proc. 33rd IEEE Conf. Decision and Control, Orlando, FL, 1994*. IEEE Publications, New York, 1994, pp. 1950-1956.

[4] Sussmann, H. J., “A strong maximum principle for systems of differential inclusions.” In *Proc. 35th IEEE Conf. Decision and Control, Kobe, Japan, Dec. 1996*. IEEE Publications, New York, 1996, pp. 1809-1814.

[5] Sussmann, H. J., “Some recent results on the maximum principle of optimal control theory.” In *Systems and Control in the Twenty-First Century*, C. I. Byrnes, B. N. Datta, D. S. Gilliam and C. F. Martin Eds., Birkhäuser, Boston, 1997, pp. 351–372.

[6] Sussmann, H. J., “Transversality conditions and a strong maximum principle systems of differential inclusions. In *Proceedings 37th IEEE Conf. Decision and Control, Tampa, FL, Dec. 1998). IEEE publications, New York, 1998, pp 1-6.*

[7] Sussmann, H. J., “Multidifferential calculus: chain rule, open mapping and transversal intersection theorems.” In *Optimal Control: Theory, Algorithms, and Applications; Proceedings of the IFIP conference held at the University of Florida, Gainesville, on February 27-28 and March 1, 1997*, W. W. Hager and P. M. Pardalos, Editors, Kluwer Academic Publishers, 1998, pp. 436-487.

[8] 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 in Control and Information Sciences, no. 246, Springer-Verlag, pp. 325-354.

[9] Sussmann, H.J., “A maximum principle for hybrid optimal control problems.” In *Proc. 38th IEEE Conf. Decision and Control, Phoenix, AZ, Dec. 1999*. IEEE publications, New York, 1999, pp. 425–430.

[10] Sussmann, H.J., “Résultats récents sur les courbes optimales.” In *15 ^{e} Journée Annuelle de la Société Mathémathique de France (SMF)*, Publications de la SMF, Paris, 2000, pp. 1-52.

[11] 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.

[12] Sussmann, H.J., “Set-valued differentials and the hybrid maximum principle.” In *Proc. 39th IEEE Conf. Decision and Control, Sydney, Australia, December 12-15, 2000*. IEEE publications, New York, 2000.

[13] Sussmann, H.J., “Needle variations and almost lower semicontinuous differential inclusions.” *Set Valued Analysis*, June-September 2002, Vol. 10, Issue 2-3, pp. 233-285..

[14] Sussmann, H.J., “Optimal control of nonsmooth systems with classically differentiable flow maps.” In *Proc. of the 6th IFAC Symposium on Nonlinear Control Systems (NOLCOS 2004), Stuttgart, Germany, September 1-3, 2004.*

[15] Sussmann, H.J., “Combining high-order necessary conditions for optimality with nonsmoothness.” In *Proc. 43rd IEEE Conf. Decision and Control, The Atlantis, Paradise Island, the Bahamas, December 14-17, 2004.*

[16] Sussmann, H.J., “A very non-smooth maximum principle with state constraints.” In *Proc. 44th IEEE CDC, Sevilla, Spain, Dec. 12-15, 2005*. IEEE publications, New York, 2005.

[17] Sussmann, H.J., “Set transversality, approximating multicones, Warga derivate containers and Mordukhovich cones.” In *Proc. 44th IEEE CDC, Sevilla, Spain, Dec. 12-15, 2005*. IEEE publications, New York, 2005.

[18] Sussmann, H.J., “Generalized differentials, variational generators, and the maximum principle with state constraints.” To appear in *Andrei Agrachev, Stephen Morse, Eduardo Sontag, Hector Sussmann, Vadim Utkin, Nonlinear and Optimal Control Theory. Cetraro, Italy, June 2004* (Fondazione C.I.M.E.), Gianna Stefani and Paolo Nistri Eds., Lect. Notes in Mathematics, Springer-Verlag; 64-page preprint.

[19] Sussmann, H.J., “Chattering variations, finitely additive measures, and the nonsmooth maximum principle with state space constraints.” *Proc. MTNS 2006 (17th International Symposium on the Mathematical Theory of Networks and Systems)*, held in Kyoto, Japan, July 24-28, 2006.