

Set separation, tangent cones, and the Pontryagin Maximum Principle

H. J. Sussmann

Department of Mathematics
Rutgers University
Piscataway, NJ 08854, USA
`sussmann@math.rutgers.edu`

CONFERENCE IN MEMORY OF W. P. DAYAWANSA
TEXAS TEX UNIVERSITY
April 20, 2007

These slides are available at <http://www.math.rutgers.edu/~sussmann/slides.html>

This work is partially supported by grant NSF-DMS 05-09930. The author is grateful to NSF for the support.

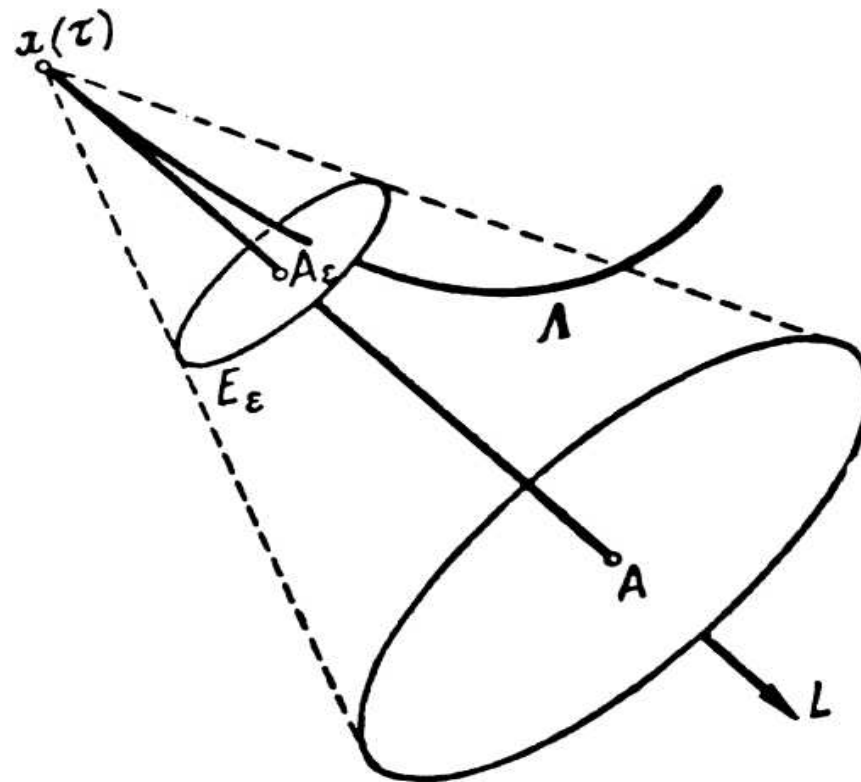


Figure 31

The figure on page 96

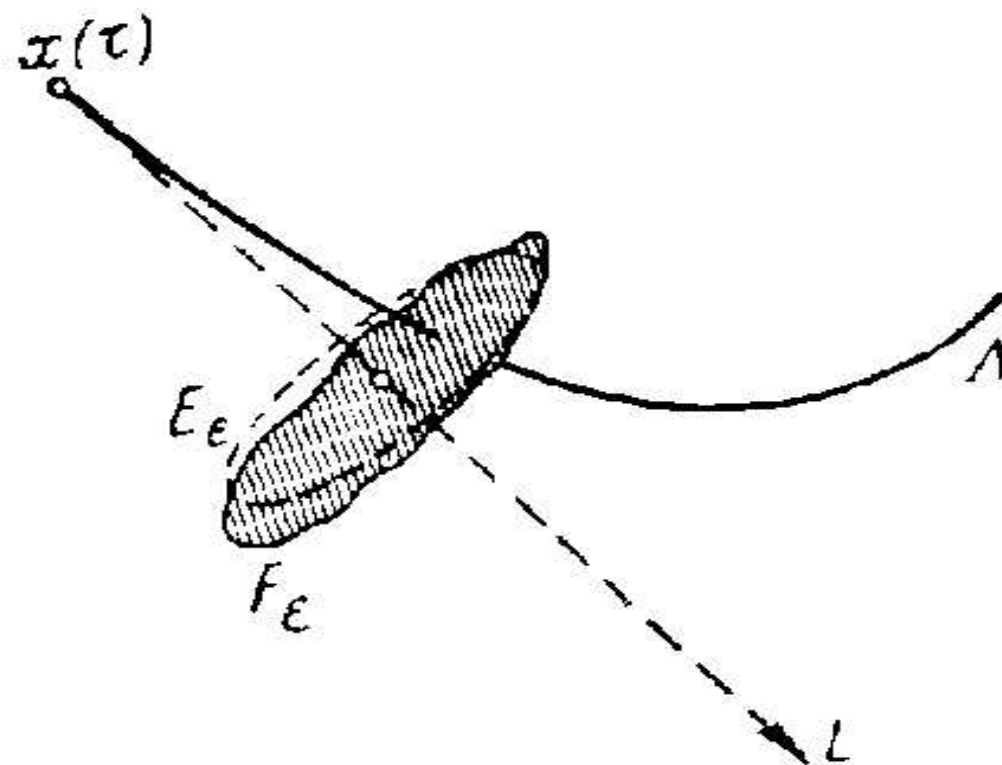


Figure 32

The figure on page 98

† A set \tilde{M} in a vector space is called a convex cone with vertex at the point o if (1) it is a *cone*, i.e., together with every point a distinct from o , it also contains the entire ray \overrightarrow{oa} ; (2) it is *convex*, i.e., together with each pair of points, it contains the entire segment joining them. Note that if the convex cone M does not fill out the entire vector space X to which it belongs, there exists a hyperplane in X , passing through the vertex o of M , such that M lies in one of the (closed) half-spaces defined by this hyperplane. The point $x \in X$ is called an *interior point* of $M \subset X$ if some neighborhood of x (in X) is entirely contained in M . The set of all interior points of $M \subset X$ is called the *interior* of M .

Furthermore, let M_1 and M_2 be two convex cones in X with a common vertex o . We shall say that M_1 and M_2 are *separated* in X if there exists a hyperplane which separates them; i.e., a hyperplane such that M_1 is entirely contained in one (closed) half-space defined by this hyperplane, and M_2 is entirely contained in the other. In order that M_1 and M_2 be separated, it is necessary and sufficient that one of the following two conditions be satisfied: (1) there exists a hyperplane containing both M_1 and M_2 ; (2) there is no point which is a relative interior point of both M_1 and M_2 (where relative refers to the relative topology in the respective carrier plane defined as the smallest dimensional plane containing the respective cone). Thus, if the cones M_1 and M_2 (with common vertex o) are not separated in X , the linear span of their carrier planes coincides with the entire space X , and in addition, there exists a point a which is a relative interior point of both M_1 and M_2 . In this case, it is possible to pass a plane C through a (which plane is orthogonal to the line oa and intersects M_2 only at a) such that all the points of C which are sufficiently close to a belong to M_1 , and in addition, such that the linear span of C and the carrier plane of M_2 coincides with X . In other words, the intersection of C with a sphere of small radius centered at a is a "complimentary area" to the carrier plane of M_2 . This area is orthogonal to the line oa , and is entirely contained in M_1 . The dimension of this "complimentary area" is equal to the difference of the dimensions of X and M_2 .

I. What the footnote says

Two convex cones C_1, C_2 are **separated** iff there exists a hyperplane H such that C_1 is contained in one of the closed halfspaces determined by H and C_2 is contained in the other one.

Equivalently, C_1, C_2 are separated if there exists a nontrivial linear functional λ such that $\lambda(c_1) \leq 0$ and $\lambda(c_2) \geq 0$ whenever $c_1 \in C_1$ and $c_2 \in C_2$.

From now on, we use **linearly separated** to mean “separated in the sense of the footnote.”

THEOREM: C_1 and C_2 are not linearly separated if and only if the following two conditions hold:

- $C_1 \cup C_2$ linearly spans the whole space.
- The relative interiors of C_1 and C_2 intersect.

FIXED TIME INTERVAL OPTIMAL CONTROL PROBLEM:

minimize $\int_{a_*}^{b_*} f_0(\xi(t), \eta(t), t) dt$

subject to $\dot{\xi}(t) = f(\xi(t), \eta(t), t)$ for a.e. $t \in [a_*, b_*]$
($\xi(\cdot)$ absolutely continuous map from $[a_*, b_*]$ to \mathbb{R}^n)
 $\xi(a_*) = \bar{x}$
 $\xi(b_*) \in S$.

TCP: trajectory-control pair

The **Pontryagin Maximum Principle** (PMP) is a necessary condition for a reference TCP (ξ_*, η_*) to be a solution.

Actually, the PMP is a necessary condition for \mathcal{R} and S to be separated, in the sense that $\mathcal{R} \cap S = \{\xi_*(b_*)\}$, where

- M is the state space,
- $\xi_* : [a_*, b_*] \mapsto M$ is the reference trajectory,
- \mathcal{R} is the reachable set from $\bar{x} = \xi_*(a_*)$ over $[a_*, b_*]$,
- S is a given set, such that $\xi_*(b_*) \in S$.

(The reduction of the optimal control problem to this separation problem is done in the book by Pontryagin et al. Basically, we add the cost as a new state variable, and considering the “augmented reachable set” $\mathcal{R}^a = \{(c, x) : x \text{ is reachable from } \bar{x} \text{ over } [a_*, b_*] \text{ with cost } c\}$. The “forbidden zone” is $F = \{(c, x) : x \in S, c < c_*\}$, where $c_* = \text{cost of the reference trajectory}$. Let $S^a = F \cup \{q_*\}$, where $q_* = (c_*, \xi_*(b_*))$. Then **If the reference trajectory is optimal then $\mathcal{R}^a \cap S^a = \{q_*\}$.**)

The proof of the PMP roughly works as follows: one constructs a “tangent cone” C_1 to the reachable set \mathcal{R} and selects a “tangent cone” C_2 to the set S of forbidden terminal states.

If the sets \mathcal{R} and S are “separated”, in the sense that $\mathcal{R} \cap S = \{\xi_*(b_*)\}$, then their “tangent cones” C_1 and C_2 should be linearly separated.

This would give the separating linear functional λ , and then by transporting λ backwards via the differential of the reference flow, one gets the “adjoint vector” of the PMP.

The way one proves that the “tangent cones” C_1 and C_2 must be linearly separated is, roughly, as follows: assume they are not linearly separated. Then there is a ray (i.e. half-line starting at 0) L contained in the intersection. And then the intersection of the sets themselves will contain points other than $\xi_*(b_*)$ arbitrarily close to $\xi_*(b_*)$.

But

1. this argument is not quite rigorous (or even clear),
2. if valid, the argument appears to yield something stronger, namely, that the intersection $\mathcal{R} \cap S = \{\xi_*(b_*)\}$ contains a nontrivial curve,
3. cone “separation” in the sense of the footnote does not exactly correspond to set separation.

II. CONES

A. Definition

A *cone* in a real linear space X is a subset C of X which is nonempty, and closed under multiplication by nonnegative scalars. (In particular, if C is a cone then necessarily $0 \in C$.)

B. Definition

The *polar* of a cone C in a real linear normed space X is the set C^\perp of all $w \in X^\dagger$ such that $\langle w, c \rangle \leq 0$ for all $c \in C$. Clearly, C^\perp is always a closed convex cone. If X is finite-dimensional (so $X \sim X^{\dagger\dagger}$ canonically), then $C^{\perp\perp}$ is the smallest closed convex cone containing C , from which it follows in particular that $C^{\perp\perp} = C$ if and only if C is closed and convex.

REMARK: X^\dagger is the dual of X .

C. Definition

Assume that $S \subseteq \mathbb{R}^n$ and $p \in S$. 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

$$v = \lim_{j \rightarrow \infty} \frac{p_j - p}{h_j}.$$

D. Notation

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 always a closed cone.)

E. Definition

Assume that $S \subseteq \mathbb{R}^n$ and $p \in S$. 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

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

and $LD = C$.

F. Definition

Assume that $S \subseteq \mathbb{R}^n$, S is closed, and $p \in S$. 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 $\lim_{j \rightarrow \infty} v_j = v$.

G. Notation

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

A. Definition

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 $x = c_1 - c_2$.

B. Remark

This is a very natural generalization to cones of the ordinary notion of transversality of linear subspaces. For subspaces S_1, S_2 , it is customary to require that $S_1 + S_2 = \mathbb{R}^n$, but it would make no difference if we required $S_1 - S_2 = \mathbb{R}^n$ instead.

C. Intuition

The basic idea of transversality is that, if two objects O_1, O_2 have first-order approximations A_1, A_2 near a point p , and A_1 and A_2 are transversal, then $O_1 \cap O_2$ looks, near p , like $A_1 \cap A_2$.

IV. NON-TRANSVERSALITY = LINEAR SEPARATION

Suppose C_1, C_2 are convex cones in \mathbb{R}^n . Then the following conditions are equivalent:

- C_1 and C_2 are not transversal,
- $C_1^\perp \cap (-C_2)^\perp \neq \{0\}$,
- there exists a nonzero linear functional $\bar{p} : \mathbb{R}^n \mapsto \mathbb{R}$ such that

$$\langle \bar{p}, c_1 \rangle \leq 0 \quad \text{for all } c_1 \in C_1,$$

and

$$\langle \bar{p}, c_2 \rangle \geq 0 \quad \text{for all } c_2 \in C_2.$$

V. STRONG TRANSVERSALITY

A. Definition

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

B. Intuition:

If two sets S_1, S_2 have first-order approximations C_1, C_2 near a point p , and the cones C_1, C_2 are strongly transversal, it should follow that $S_1 \cap S_2$ contains points p_j converging to p and $\neq p$.

Reason:

Near p , $S_1 \cap S_2$ should look like $C_1 \cap C_2$, because C_1 and C_2 are transversal.

Since $C_1 \cap C_2$ contains a full half-line through 0, $S_1 \cap S_2$ should also contain a nontrivial curve through p .

C. An important caveat:

The above intuition is, of course, not a proof, and when one does things carefully, it turns out that, for very reasonable notions of “first-order approximation,” all one can prove is that $S_1 \cap S_2$ must contain a nontrivial connected set through p , but this set could fail to be path-connected. And for other reasonable notions one can prove even less. (For example, that $S_1 \cap S_2$ contains a sequence of points $p_j \neq p$ that converges to p .)

The following lemma says that transversality and strong transversality are almost equivalent.

More precisely, 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.

D. Lemma

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. (Recall that (ii) says: “ C_1 and C_2 are linear subspaces and $C_1 \oplus C_2 = \mathbb{R}^n$.”)

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

$$-c = c_1 - c_2, \quad c_1 \in C_1, \quad c_2 \in C_2.$$

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 $C_1 \oplus C_2 = \mathbb{R}^n$. **END OF PROOF.**

VI. Set separation

Two subsets S_1, S_2 of a Hausdorff topological space T are **locally separated** at a point $p \in T$ if there exists a neighborhood U of p in T such that

$$S_1 \cap S_2 \cap U \subseteq \{p\}.$$

VII. The Transversal Intersection Property

If two subsets S_1, S_2 of \mathbb{R}^n have tangent cones C_1, C_2 at a point p , and the cones C_1, C_2 are strongly transversal, then S_1 and S_2 are not locally separated at p .

The statement that “ S_1 and S_2 are not locally separated at p ” means the following:

$S_1 \cap S_2$ contains a sequence of points p_j converging to p and $\neq p$.

A. Remark. This is exactly the “intuition” discussed earlier.

B. Question. For what notions of “tangent cone to a set at a point” is the TIP (Transversal Intersection Property) true?

THEOREM: The TIP is true if “tangent cone” is taken to mean “Boltyanskii approximating cone.” (The proof of this is Type T.)

Furthermore, when the cones are strongly transversal, not only do the sets have a nontrivial intersection, but their intersection contains a nontrivial connected set containing p .

THEOREM: The TIP is true if “tangent cone” is interpreted to mean “Clarke tangent cone.” (The proof of this is Type L.)

Furthermore, when the cones are strongly transversal, not only do the sets have a nontrivial intersection, but their intersection contains a nontrivial Lipschitz curve going through p .

Both theorems extend to **tangent multicones**:

A **multicone** is a nonempty set of cones.

A **convex multicone** is a nonempty set of convex cones.

Two multicones $\mathcal{C}_1, \mathcal{C}_2$, are **transversal** if every $C_1 \in \mathcal{C}_1$ is transversal to every $C_2 \in \mathcal{C}_2$.

Two multicones $\mathcal{C}_1, \mathcal{C}_2$, are **strongly transversal** if they are transversal and, in addition, there is a nonzero linear functional μ such that $C_1 \cap C_2 \cap \{v : \mu(v) > 0\} \neq \emptyset$ for every $C_2 \in \mathcal{C}_2$.

VIII. The Transversal Intersection Property for Multicones

If two subsets S_1, S_2 of \mathbb{R}^n have tangent multicones $\mathcal{C}_1, \mathcal{C}_2$ at a point p , and the multicones $\mathcal{C}_1, \mathcal{C}_2$ are strongly transversal, then S_1 and S_2 are not locally separated at p .

A. Question. For what notions of “tangent multicone to a set at a point” is the TIP true?

Two multicones are **weakly linearly separated** if they are not strongly transversal. (Recall that two multicones are **linearly separated** if they are not transversal.)

THEOREM: Two convex multicones $\mathcal{C}_1, \mathcal{C}_2$ in \mathbb{R}^n are weakly linearly separated if and only if

(*) for every nonzero linear functional μ on \mathbb{R}^n there exist $\pi_0, \pi_1, \pi_2, C_1, C_2$, such that $\pi_0 \in \mathbb{R}, \pi_0 \geq 0, C_1 \in \mathcal{C}_1, C_2 \in \mathcal{C}_2, \pi_1 \in C_1^\perp, \pi_2 \in C_2^\perp$, and $\pi_0\mu = \pi_1 + \pi_2$.

REMARK. $\mathcal{C}_1, \mathcal{C}_2$ in \mathbb{R}^n are linearly separated if and only if (*) can be satisfied with $\pi_0 = 0$.

There are two natural notions of “tangent multicone to a set at a point”, generalizing, respectively, the notion of Boltyanskii approximating cones and that of Clarke cone.

A **GDQ approximating multicone** to a set S at a point p is a set \mathcal{C} of convex cones which is the image $\Lambda \cdot D$ of a Boltyanskii approximating cone D to a set A at 0 under a GDQ Λ at $(0, p)$ of a set-valued map F such that $F(A) \subseteq S$ and $p \in F(0)$.

A **WDC approximating multicone** to a set S at a point p is a set \mathcal{C} of convex cones which is the image $\Lambda \cdot D$ of the Clarke tangent cone D to a closed set A at 0 under a WDC Λ at 0 of a Lipschitz map F such that $F(A) \subseteq S$ and $F(0) = p$.

GDQ: Generalized Differential Quotient

WDC: Warga Derivate Container

GDQs and WDCs are “generalized differentials, such that that a GDQ or WDC of a map F at a point p is a compact set of linear maps.

THEOREM: The TIP is true if “tangent multicone” is taken to mean “GDQ approximating multicone.” (The proof of this is Type T.)

Furthermore, when the multicones are strongly transversal, not only do the sets have a nontrivial intersection, but their intersection contains a nontrivial connected set containing p .

THEOREM: The TIP is true if “tangent multicone” is interpreted to mean “WDC approximating multicone” (The proof of this is Type L.)

Furthermore, when the multicones are strongly transversal, not only do the sets have a nontrivial intersection, but their intersection contains a nontrivial Lipschitz curve going through p .

GDQs are studied in detail in my Cetraro lecture notes, to appear in a Springer book, available in my Web page.

WDC approximating multicones are studied in detail in my 2005 CDC paper, and in a much more detailed JDE paper to appear, and to be made available in my Web page shortly.

The definition of “WDC approximating multicone” considered in the JDE paper is more general.

IX. The Directional Open Mapping Property (DOMP)

If a set S has an approximating multicone \mathcal{C} at a point p , and a vector v belongs to the interior of every $C \in \mathcal{C}$, then the set S contains a “conic sector” at p in the direction of v , that is, a set of the form $\{p + rw : 0 \leq r \leq \alpha, \|w - v\| \leq \beta\}$, for some positive α, β .

THEOREM: For any reasonable concept of “approximating multicone”, the DOMP is equivalent to the TIP.

REASON: Suppose the DOMP holds. Let S_1, S_2 be subsets of \mathbb{R}^n having approximating multicones $\mathcal{C}_1, \mathcal{C}_2$, at 0. Suppose \mathcal{C}_1 and \mathcal{C}_2 are strongly transversal. Let μ be a nontrivial linear functional such that $C_1 \cap C_2 \cap \{v : \mu(v) > 0\} \neq \emptyset$ for every $C_1 \in \mathcal{C}_1, C_2 \in \mathcal{C}_2$.

Let $S = S_1 \times S_2$. Let F be the map $(x_1, x_2) \mapsto (x_1 - x_2, \mu(x_1))$. Let $\mathcal{C} = \mathcal{C}_1 \times \mathcal{C}_2$. Then $DF(0) \cdot \mathcal{C}$ is an approximating multicone of $F(S)$ at $(0, 0)$, and the strong transversality hypothesis amounts to the assertion that the vector $(0, 1)$ belongs to the interior of every member of $DF(0) \cdot \mathcal{C}$. So $F(S)$ contains an arc $\{(0, r) : 0 \leq r \leq \alpha\}$. This means that there are points $x(r) \in S_1$ such that $x(r) \in S_2$ and $\mu(x(r)) = r$, so $x(r) \neq 0$.

In the previous theorem, “reasonable” means “such that the proof given above is valid. All that is needed is the Cartesian product property and the “image under a smooth map” property.

X. A ROUGH CLASIFICATION OF VERSIONS OF THE FDPMP (FINITE-DIMENSIONAL PONTRYAGIN MAXIMUM PRINCIPLE)

Every known version of the FDPMP is of one of the following two types:

- Type T. (The “T” stands for “topological.”)
- Type L. (The “L” stands for “limiting.”)

In the transversality condition:

- Type T versions involve some kind of Boltyanskii tangent cone to the terminal set.
- Type L versions involve the Clarke tangent cone to the terminal set, or the Mordukhovich normal cone

The proofs of Type T versions typically use a topological separation argument, based on the Brouwer fixed point theorem or some variant thereof.

All versions of the finite-dimensional Pontryagin maximum principle with high-order conditions (Knobloch, Krener, Bianchini-Stefani, Agrachev, Sarychev, Gamkrelidze, and many others) appear to be Type T.

The finite dimensionality comes in where the Brouwer fixed-point theorem is used, since that theorem depends essentially on being in a finite-dimensional space.

The proofs of Type L versions usually produce a sequence $\{\bar{p}_k\}_{k \in \mathbb{N}}$ of “approximate terminal adjoint covectors” (using, for example, the Ekeland variational principle) and then extract a convergent (or weakly convergent) subsequence whose limit \bar{p}_∞ is the terminal value of the adjoint covector.

The finite dimensionality comes in when one tries to establish that $\bar{p}_\infty \neq 0$. The \bar{p}_k can be normalized so that $\|\bar{p}_k\| = 1$, and the existence of a weak*-convergent subsequence (if, say, we are working on a Hilbert space) follows from the weak*-compactness of the closed unit ball, but in infinite dimensions one cannot prove in general that $\bar{p}_\infty \neq 0$, since the unit sphere is not weak*-compact.

The first TIP result leads to a number of versions of the FDPMP with a Boltyanskii or Boltyanskii-like tangent cones in the transversality condition. In these versions, high-order conditions can easily be included. (Classical work by Pontryagin et al., work by Knobloch, Krener, Agrachev, Sarychev, Gamkrelidze, Bianchini, Stefani, HJS, and lots of others.) These results are all proved using the TIP for Boltyanskii cones or for some generalization of them, such as the “approximating multicones” used by HJS.

The second TIP result leads to a number of versions of the FDPMP with a Clarke or Mordukhovich normal cone in the transversality condition. (Work by Clarke, Vinter, Rockafellar, Ioffe, Mordukhovich, Loewen, da Pinho, Franskowska, and lots of others.) In these versions, it does not seem that high-order conditions can be incorporated. Most of these results are not proved by explicitly using the TIP for Clarke cones or for some generalization thereof, but work is now in progress by HJS which, it is hoped, will show that they can be proved that way.

It may seem natural to expect that a more general TIP might be true, containing both results. I conjectured (and even briefly believed I had proved) about 10 years ago that such a result was true.

The problem was solved in January, 2006, by [Alberto Bressan](#), who proved the following:

XI. Bressan's Theorem

There exist two closed subsets S_1, S_2 of \mathbb{R}^4 , and two closed convex cones C_1, C_2 in \mathbb{R}^4 , such that

- C_1 is a Boltyanskii approximating cone to S_1 at 0;
- C_2 is the Clarke tangent cone to S_2 at 0;
- C_1, C_2 are strongly transversal;
- $S_1 \cap S_2 = \{0\}$.

Using Bressan's example, one can construct an example of a Lagrange optimal control problem in \mathbb{R}^8 with a terminal state constraint, and an optimal trajectory-control pair (ξ_*, η_*) , defined on an interval $[a_*, b_*]$, such that

- the dynamics and Lagrangian satisfy conditions that lend themselves to Type T arguments,
- the terminal set S has a Clarke tangent cone C at the terminal point of $\xi_*(b)$,
- there does not exist a nontrivial multiplier $(\pi(\cdot), \pi_0)$ (consisting of an adjoint covector $\pi(\cdot)$ and “abnormal multiplier” π_0) that satisfies the adjoint equation, the Hamiltonian maximization condition, and the transversality condition $-\pi(b_*) \in C^\perp$.

The actual construction is done in complete detail in my 2006 CDC paper, and it's sort of technical.

Remark: In this particular example, the usual nonsmooth “adjoint differential inclusion” is actually a true “adjoint differential equation.”