Combining high-order necessary conditions for optimality with nonsmoothness

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 version of the Pontryagin Maximum Principle valid for systems of flows rather than for systems governed by ordinary differential equations. The flow maps are required to be differentiable in a generalized sense (using the theory of "generalized differential quotients") which is much weaker than ordinary differentiability and allows the "differentials" to be sets of linear maps rather than single linear maps. The resulting conditions apply to control dynamics with a right-hand side that needs not be smooth, or even Lipschitz, and could even be discontinuous. This is so because the usual adjoint equation, in which there occur derivatives of the reference vector field with respect to the state, is replaced by an integrated form. This form only involves differentials of the reference flow maps, and therefore makes sense as long as these flow maps are differentiable, which can happen even when the reference vector field itself fails to be Lipschitz or even continuous. The resulting "integrated adjoint equation" gives rise to "adjoint vectors" that need not be absolutely continuous, and could be discontinuous and unbounded. Furthermore, this integrated adjoint equation relates the values of the adjoint vector on intervals that could be disjoint and contain singularities in between. This makes it possible to establish necessary conditions for an optimum that yield a global adjoint vector that satisfies various nonsmooth conditions everywhere and at the same time satisfies extra "high-order" requirements, such as the Goh condition, on intervals where the dynamics is sufficiently smooth.**

I. Preliminary definitions

As a preliminary to the statement of our main theorem, we give a long series of definitions and background results. The definitions themselves are all rather trivial and natural. The background results are also rather simple, except for the chain rule for GDQs and the separation theorem 2.6, for both of which proofs are available in the literature (cf. [3]). Theorem 2.6 is the crux of the matter, and involves a refinement of the "topological argument" used in classical proofs of the Maximum Principle such as the one due to Boltyanskii (cf. Pontryagin *et al.* [1]).

The abbreviations "LS" and "FDLS" will stand for "linear space" and "finite-dimensional linear space," respectively. (All LSs in this paper are over the field $\mathbb{R}$ of real numbers.) If A, B are LSs, we use $LM(A,B)$, $A^\dagger$ to denote, respectively, the space of all linear maps from A to B and the (algebraic) dual of a LS A, so $A^\dagger = LM(A,\mathbb{R})$.

The symbols $\mathbb{Z}$, $\mathbb{Z}_+$, $\mathbb{N}$ will denote, respectively, the set of all integers and the sets $\{n\in\mathbb{Z}: n\geq 0\}$, $\{n\in\mathbb{Z}: n>0\}$. If $n,m\in\mathbb{Z}_+$, then $\mathbb{R}^{m\times n}$ is the set of all real matrices with m rows and n columns. We write $\mathbb{R}^n$, $\mathbb{R}_n$ for $\mathbb{R}^{n\times 1}$, $\mathbb{R}^{1\times n}$. We identify $LM(\mathbb{R}^n,\mathbb{R}^m)$ with $\mathbb{R}^{m\times n}$ in the usual way. The norms on $\mathbb{R}^n$ and $\mathbb{R}_n$ are the Euclidean norms, and for $\mathbb{R}^{m\times n}$ we use the operator norm $\|\cdot\|_{op}$, defined by $\|L\|_{op} = \sup\{\|L\cdot x\| : x\in\mathbb{R}^n, \|x\|\leq 1\}$. If $\Lambda \subseteq \mathbb{R}^{m\times n}$ and $\delta \geq 0$, we write Λ^δ to denote the set $\{L\in\mathbb{R}^{m\times n} : \mathrm{dist}(L,\Lambda)\leq\delta\}$, where $\mathrm{dist}(L,\Lambda)\stackrel{\mathrm{def}}{=}\inf\{\|L-L'\|_{op} : L'\in\Lambda\}$.

Ordinary, set-valued, and augmented maps. A ***set-valued map*** (abbr. SVM) is a triple $F=(A,B,G)$ such that A and B are sets and G is a subset of $A\times B$. The sets A, B, G are, respectively, the ***source***, ***target***, and ***graph*** of the SVM F, and we write $A=\mathrm{So}(F)$, $B=\mathrm{Ta}(F)$, $G=\mathrm{Gr}(F)$. If x is any object, we write $F(x)=\{y : (x,y)\in\mathrm{Gr}(F)\}$. (Hence $F(x)=\emptyset$ unless $x\in\mathrm{So}(F)$.) The sets $\mathrm{Do}(F)=\{x\in\mathrm{So}(F) : F(x)\neq\emptyset\}$, $\mathrm{Im}(F)=\bigcup_{x\in\mathrm{So}(F)}F(x)$, are, respectively, the ***domain*** and ***image*** of F. If $F=(A,B,G)$ is an SVM, we say that F is an ***SVM from*** A ***to*** B, and write $F: A\mapsto\!\!\!\!\rightarrow B$. We say that F is ***single-valued*** if $\mathrm{card}(F(x))\leq 1$ for all x. We use $SVM(A,B)$ to denote the set of all SVMs from A to B. The expression "ppd map" stands for "possibly partially defined (that is, not necessarily everywhere defined) ordinary (that is, single-valued) map," and we write $f: A\hookrightarrow B$ to indicate that f is a ppd map from A to B.

If X is a set, then $\mathbb{I}_X$ denotes the ***identity map*** of X, i.e., the triple (X,X,Δ_X), where $\Delta_X=\{(x,x) : x\in X\}$. If F is an SVM, and S is a set, then the ***restriction*** of F to S is the SVM $F\lceil S : \mathrm{So}(F)\cap S \mapsto\!\!\!\!\rightarrow \mathrm{Ta}(F)$ whose graph is $\mathrm{Gr}(F)\cap(S\times\mathrm{Ta}(F))$. If F_1,F_2 are SVMs, then the ***composite*** $F_2\circ F_1$ is defined if and only if $\mathrm{So}(F_2)=\mathrm{Ta}(F_1)$, and in that case, by definition, $\mathrm{So}(F_2\circ F_1)=\mathrm{So}(F_1)$, $\mathrm{Ta}(F_2\circ F_1)=\mathrm{Ta}(F_2)$, and $\mathrm{Gr}(F_2\circ F_1)=\{(x,z) : \exists y, (y\in F_1(x)\wedge z\in F_2(y))\}$.

If A is an LS and X is a set, the A***-augmentation of*** X

Research supported in part by NSF Grant DMS01-03901

is the set X^A defined by $X^A \stackrel{\mathrm{def}}{=} X \times A$. If X, Y are sets, an A-***augmented set-valued map*** (abbr. A-SVM) from X to Y is a set-valued map $F : X^A \mapsto\!\!\!\rightarrow Y^A$ such that

(*) $F(x,a) = F(x,0) + \{(0,a)\}$ *(that is,* $F(x,a) = \{(y,a'+a) : (y,a') \in F(x,0)\}$*) whenever* $x \in X$, $a \in A$.

We use $SVM^A(X,Y)$ to denote the set of all A-SVMs from X to Y. If $F \in SVM^A(X,Y)$ we define an SVM $\check{F} : X \mapsto\!\!\!\rightarrow Y \times A$ by letting $\check{F}(x) \stackrel{\mathrm{def}}{=} F(x,0)$ for $x \in X$. It is then clear that F can be recovered from $\check{F}$, because $F(x,a) = \{(x', a'+a) : (x',a') \in \check{F}(x)\}$.

If $F_i \in SVM^A(X_{i-1}, X_i)$ for $i = 1,2$, then the composite $F = F_2 \circ F_1$ belongs to $SVM^A(X_0, X_2)$, and $\check{F} = \check{F}_2 \circ^A \check{F}_1$, where the A-***augmented composition*** $\nu \circ^A \mu$ of two SVMs $\mu : X \mapsto\!\!\!\rightarrow Y \times A$, $\nu : Y \mapsto\!\!\!\rightarrow Z \times A$, is the set-valued map $\nu \circ^A \mu : X \mapsto\!\!\!\rightarrow Z \times A$ such that, for $x \in X$, $(\nu \circ^A \mu)(x)$ is the set

$$\{(x'',a'') \in Z \times A : (\exists (x',a') \in \mu(x)) : (x'', a''-a') \in \nu(x')\}.$$

Augmented linear maps. If A, X, Y are LSs, an A-***augmented linear map*** from X to Y is a linear map $L : X^A \mapsto Y^A$ such that L is an A-augmented SVM from X to Y. Clearly, an $L \in LM(X^A, Y^A)$ is A-augmented if and only if $L(x,a) = L(x,0) + (0,a)$ for every $(x,a) \in X^A$. We use $LM^A(X,Y)$ to denote the space of augmented linear maps from X to Y. A member L of $LM^A(X,Y)$ is determined (via the formula $L(x,a) = (\ell(x), a + \ell^0(x))$) by specifying two linear maps $\ell : X \mapsto Y$ and $\ell^0 : X \mapsto A$, known, respectively, as the ***state space component*** and ***cost component*** of L. So $LM^A(X,Y)$ is canonically identified with the product $LM(X,Y) \times LM(X,A)$. An A-***augmented linear functional*** on a LS X is a linear functional on X^A. An A-augmented linear functional $\Lambda : X^A \mapsto \mathbb{R}$ is determined (via the formula $\Lambda(x,a) = \lambda(x) + \lambda^0(a)$) by its ***state space component*** $\lambda \in A^\dagger$ and its ***augmented component*** $\lambda^0 \in A^\dagger$. The augmented component λ^0 is then given by $\lambda^0(a) = \Lambda(0,a)$ for $a \in A$.

If $L_i \in LM^A(X_{i-1}, X_i)$ for $i = 1,2$, then $L = L_2 \circ L_1 : X_0^A \mapsto X_2^A$ is also an A-augmented linear map. Indeed, if $x \in X_0$ and $a \in \mathbb{R}$, then

$$\begin{aligned} L(x,a) &= L_2(L(_1(x,a)) = L_2\Big(L_1(x,0)) + (0,a)\Big) \\ &= L_2(L_1(x,0)) + L_2(0,a) \\ &= L_2(L_1(x,0)) + (0,a) = L(x,0) + (0,a)\,, \end{aligned}$$

because (a) $L_1(x,a) = L_1(x,0) + (0,a)$ and in addition (b) $L_2(0,a) = L_2(0,0) + (0,a) = (0,a)$, since the fact that L_2 is linear implies that $L_2(0,0) = (0,0)$. If we identify L_1, L_2, L, with pairs (ℓ_1, ℓ_1^0), (ℓ_2, ℓ_2^0), (ℓ, ℓ^0), as explained above, then a simple calculation shows that $\ell = \ell_2 \circ \ell_1$ and $\ell^0 = \ell_1^0 + \ell_2^0 \circ \ell_1$.

If $\Lambda \in (Y^A)^\dagger$ and $L \in LM^A(X,Y)$, then the composite (or "pullback") map $L^*(\Lambda) \stackrel{\mathrm{def}}{=} \Lambda \circ L$ is of course a linear functional on X^A, i.e., an A-augmented linear functional on X. Furthermore, if $\tilde{\Lambda} = \Lambda \circ L$, λ^0, $\tilde{\lambda}^0$ are the augmented components of Λ, $\tilde{\Lambda}$, and $a \in A$, then $\tilde{\lambda}^0(a) = \tilde{\Lambda}(0,a) = \Lambda(L(0,a)) = \Lambda(0,a) = \lambda^0(a)$ (since $L(0,a) = L(0,0) + (0,a) = (0,a)$), so ***the augmented component of the pullback*** $\Lambda \circ L$ ***of*** Λ ***by*** L ***is equal to the augmented component of*** Λ.

Manifolds. If $k \in \mathbb{N}$, X is a manifold of class C^k, and $x \in X$, we use T_xX, $T_x^\dagger X$, TX to denote, respectively, the tangent and cotangent spaces of X at x, and the tangent bundle of X. (Clearly, then, TX is a manifold of class C^{k-1}.)

If X, Y are manifolds of class C^1, $f : X \hookrightarrow Y$, and $x \in X$ is such that f is defined and of class C^1 on a neighborhood of x, then $Df(x)$ denotes the differential of f at X, so $Df(x) \in LM(T_xX, T_{f(x)}Y)$.

We use the following precise definition of "chart": a ***cubic coordinate chart*** on a μ-dimensional manifold X of class C^k is a ppd map $\kappa : X \hookrightarrow \mathbb{R}^\mu$ such that (a) $\mathrm{Do}(\kappa)$ is a nonempty open subset of X, (b) $\mathrm{Im}(\kappa)$ is an open cube $]-\alpha, \alpha[^\mu$ for some positive α, (c) $\kappa \lceil \mathrm{Do}(\kappa)$ and $\kappa^{-1} \lceil \mathrm{Im}(\kappa)$ are injective maps of class C^k. A cubic chart κ is ***centered*** at a point x of X if $\kappa(x) = 0$.

Approximating cones. A ***cone*** in a LS A is a nonempty subset C of A such that $r \cdot c \in C$ whenever $c \in C$ and $r \geq 0$. If X is a manifold of class C^1, $x \in X$, and $v \in T_xX$, we use ∇_v to denote directional differentiation in the direction of v. That is, $\nabla_v \varphi$ is equal, if $\varphi : X \hookrightarrow \mathbb{R}$ is of class C^1 near x, to the derivative $\frac{d}{dt}\varphi(\xi(t))|_{t=0}$, if $\xi : [-\varepsilon, \varepsilon] \mapsto X$ is any curve of class C^1 such that $\xi(0) = x$ and $\dot{\xi}(0) = v$. If $s \in S \subseteq X$, a ***Boltyanski approximating cone*** to S at s is a convex cone $C \subseteq T_sX$ such that there exist a neighborhood U of 0 in T_sX and a continuous map $f : U \cap C \mapsto S$ for which (a) $f(0) = s$ and (b) $\varphi(f(v)) - \varphi(s) - \nabla_v \varphi = o(\|v\|)$ as $v \to 0$ via values in C for every function $\varphi : X \hookrightarrow \mathbb{R}$ which is of class C^1 near s. A ***limiting approximating cone*** to S at s is a closed convex cone $C \subseteq T_sX$ which is the closure of an increasing union $\bigcup_{j=1}^\infty C_j$ of Boltyanski approximating cones to S at s.

II. GENERALIZED DIFFERENTIAL QUOTIENTS (GDQS)

If X, Y are metric spaces, then $SVM_{comp}(X,Y)$ will denote the subset of $SVM(X,Y)$ whose members are the set-valued maps from X to Y that have a compact graph. We say that a sequence $\{F_j\}_{j \in \mathbb{N}}$ of members of $SVM_{comp}(X,Y)$ ***inward graph-converges*** to an $F \in SVM_{comp}(X,Y)$—and write $F_j \xrightarrow{\mathrm{igr}} F$—if for every open subset Ω of $X \times Y$ such that $\mathrm{Gr}(F) \subseteq \Omega$ there exists a $j_\Omega \in \mathbb{N}$ such that $\mathrm{Gr}(F_j) \subseteq \Omega$ whenever $j \geq j_\Omega$.

Definition 2.1: Let X, Y be metric spaces. A set-valued map $F : X \mapsto\!\!\!\rightarrow Y$ is ***Cellina continuously approximable***[2] (abbr. "CCA") if

- *for every compact subset* K *of* X*, the restriction* $F \lceil K$ *of* F *to* K *belongs to* $SVM_{comp}(K,Y)$ *and is a limit—in the sense of inward graph-convergence—of a sequence of continuous single-valued maps from* K *to* Y.

We use $\mathrm{CCA}(X;Y)$ to denote the set of all CCA set-valued maps from X to Y. ■

It is easy to see that if $F : X \hookrightarrow Y$ is a single-valued ppd map, then F belongs to $\mathrm{CCA}(X;Y)$ iff it is everywhere defined and continuous. It is not hard to prove the following.

Theorem 2.2: If X, Y, Z are metric spaces, and F, G are in $\mathrm{CCA}(X;Y)$, $\mathrm{CCA}(Y;Z)$, then the composite SVM $G \circ F$ belongs to $\mathrm{CCA}(X;Z)$. ■

Definition 2.3: Assume that $m, n \in \mathbb{Z}_+$, $F : \mathbb{R}^n \mapsto\!\!\!\!\rightarrow \mathbb{R}^m$, $\Lambda \subseteq \mathbb{R}^{m\times n}$, $S \subseteq \mathbb{R}^m$, $0 \in S$. We say that Λ is a ***generalized differential quotient*** (abbreviated "GDQ") ***of*** F ***at*** $(0,0)$ ***in the direction of*** S (or ***"along*** S***"***), and write $\Lambda \in GDQ(F;0,0;S)$, if (a) Λ is compact and nonempty, and (b) for every positive $\delta \in \mathbb{R}$ there exist a neighborhood U of 0 in $\mathbb{R}^n$ and a $G \in \mathrm{CCA}(U \cap S; \Lambda^\delta)$ such that $G(x)\cdot x \subseteq F(x)$ for every $x \in U \cap S$. ■

The chain rule. If X_1, X_2, X_3 are real linear spaces, and Λ_1, Λ_2 are subsets of $LM(X_1, X_2)$, $LM(X_2, X_3)$, then the ***composite*** $\Lambda_2 \circ \Lambda_1$ is the subset of $LM(X_1, X_3)$ defined by

$$\Lambda_2 \circ \Lambda_1 \stackrel{\mathrm{def}}{=} \{L_2 \circ L_1 : L_2 \in \Lambda_2, L_1 \in \Lambda_1\}\,.$$

A subset S of a topological subspace T is a *local retract* at a point $\bar{s} \in S$ if for every neighborhood U of $\bar{s}$ there exist a neighborhood V of $\bar{s}$ asuch that $V \subseteq U$ and a continuous map $\rho : V \mapsto V \cap S$ such that $\rho(s) = s$ whenever $s \in V \cap S$.

The following is the well-known ***chain rule*** for GDQs.

Theorem 2.4: Assume that n_1, n_2, n_3, F_1, F_2, S_1, S_2, Λ_1, Λ_2 are such that $n_1, n_2, n_3 \in \mathbb{Z}_+$ and, for $i = 1, 2$,

1. $0 \in S_i$ and $F_i : \mathbb{R}^{n_i} \mapsto\!\!\!\!\rightarrow \mathbb{R}^{n_{i+1}}$;
2. $\Lambda_i \in GDQ(F_i;0,0;S_i)$.

Assume, moreover, that $F_1(S_1) \subseteq S_2$, and either S_2 is a local retract at 0 or F_1 is single-valued. Then $\Lambda_2 \circ \Lambda_1$ belongs to $GDQ(F_2 \circ F_1;0,0;C_1)$. ■

GDQs on manifolds. If X, Y are manifolds of class C^1, $m = \dim X$, $n = \dim Y$, $\bar{x} \in S \subseteq X$, $\bar{y} \in Y$, and $F : X \mapsto\!\!\!\!\rightarrow Y$, then we can define a set $GDQ(F;\bar{x},\bar{y};S)$ of subsets of $LM(T_{\bar{x}}X, T_{\bar{y}}Y)$ by picking cubic coordinate charts $X \ni x \mapsto \xi(x) \in \mathbb{R}^m$, $Y \ni y \mapsto \eta(y) \in \mathbb{R}^n$ centered at $\bar{x}$, $\bar{y}$, and declaring a subset Λ of $LM(T_{\bar{x}}X, T_{\bar{y}}Y)$ to belong to $GDQ(F;\bar{x},\bar{y};S)$ if the set $D\eta(\bar{y}) \circ \Lambda \circ D\xi(\bar{x})^{-1}$ (which, by definition, is equal to $\{\, D\eta(\bar{y}) \circ L \circ D\xi(\bar{x})^{-1} : L \in \Lambda \,\}$) belongs to $GDQ(\eta \circ F \circ \xi^{-1};0,0;\xi(S))$. With these definitions, $\mathrm{So}(\xi^{-1}) = \mathbb{R}^m$, $\mathrm{Ta}(\xi^{-1}) = \mathrm{So}(F) = X$, $\mathrm{Ta}(F) = \mathrm{So}(\eta) = Y$, and $\mathrm{Ta}(\eta) = \mathbb{R}^n$, so $\eta \circ F \circ \xi^{-1}$ is a well defined member of $SVM(\mathbb{R}^m, \mathbb{R}^n)$. The chain rule then implies that, with this definition, ***the set*** $GDQ(F;\bar{x},\bar{y};S)$ ***does not depend on the choice of the charts*** ξ***,*** η***.*** Moreover, the following two results can be proved.

Theorem 2.5: If X, Y are manifolds of class C^1, $x \in X$, and $F : X \hookrightarrow Y$ is such that F is defined and continuous on a neighborhood of x and F is classically differentiable at x, then $\{DF(x)\} \in GDQ(F;x,F(x);X)$. ■

Theorem 2.6: If X, Y are manifolds of class C^1, $F : X \mapsto\!\!\!\!\rightarrow Y$, $\bar{x} \in S \subseteq X$, $\bar{y} \in R \subseteq Y$, $\Lambda \in GDQ(F;\bar{x},\bar{y};S)$, S, R have limiting approximating cones C_S, C_R at $\bar{x}$, $\bar{y}$, and C_R is not a linear subspace, then a necessary condition for the sets $F(S)$ and R to be separated at $\bar{y}$ (in the sense that $F(S) \cap R = \{\bar{y}\}$) is that there exist a $\pi \in T^\dagger_{\bar{y}}Y \backslash \{0\}$ and an $L \in \Lambda$ such that $\pi(v) \geq 0$ for every $v \in C_R$ and $\pi(L(w)) \leq 0$ for every $w \in C_S$. ■

Partial GDQs. Suppose that (a) for $i = 1, 2$, X_i is a manifold of class C^1 and $\bar{x}_i \in S_i \subseteq X_i$, (b) Y is a manifold of class C^1 and $\bar{y} \in Y$, (c) $X = X_1 \times X_2$, (d) $S = S_1 \times S_2$, (e) $\bar{x} = (\bar{x}_1, \bar{x}_2)$, (f) $F : X \mapsto\!\!\!\!\rightarrow Y$, and (g) $\Lambda \in GDQ(F;\bar{x},\bar{y};S)$. Then, if we let ι_1, ι_2 be the partial maps $X_1 \ni x \mapsto (x, \bar{x}_2) \in X$ and $X_2 \ni x \mapsto (\bar{x}_1, x) \in X$, the chain rule implies that the ***partial GDQs*** Λ_{X_1}, Λ_{X_2} (where $\Lambda_{X_j} = \{L_{X_j} : L \in \Lambda\}$ and, for $L \in \Lambda$, L_{X_1}, L_{X_2} are the maps $T_{\bar{x}_1}X_1 \ni v \mapsto L(v, 0) \in T_{\bar{y}}Y$, and $T_{\bar{x}_2}X_2 \ni v \mapsto L(0, v) \in T_{\bar{y}}Y$, and we canonically identify $T_{\bar{x}}X$ with $T_{\bar{x}_1}X_1 \times T_{\bar{x}_2}X_2$) are GDQs of $F \circ \iota_1$, $F \circ \iota_2$, respectively, at $(\bar{x}_1, \bar{y})$ and $(\bar{x}_2, \bar{y})$, along S_1, S_2.

III. Flows and trajectories.

A ***time set*** is a nonempty subset of $\mathbb{R}$. If $\mathbf{T}$ is a time set, and $m \in \mathbb{N}$, we use $\mathbf{T}^{m,\geq}$ to denote the set of all ordered m-tuples $(t_1, t_2, \cdots, t_m)$ of members of $\mathbf{T}$ such that $t_1 \geq t_2 \geq \cdots \geq t_m$. A ***state-space bundle over a time set*** $\mathbf{T}$ is an indexed family $X = \{X_t\}_{t\in\mathbf{T}}$ of sets. A ***state-space bundle*** (abbr. SSB) is a pair $\mathcal{X} = (\mathbf{T}, X)$ such that $\mathbf{T}$ is a time set and X is a state-space bundle over $\mathbf{T}$. An SSB $(\mathbf{T}, X)$ is a ***bundle of topological spaces*** (resp. ***of metric spaces***, ***of manifolds of class*** C^k, LSs, FDLSs, etc.) if each X_t is a topological space (resp. a metric space, a manifold of class C^k, an LS, an FDLS, etc.). We will use the abbreviations C^k-SSB, FDLS-SSB, for "bundle of manifolds of class C^k" and "bundle of FDLSs," respectively.

A ***section*** of an SSB $\mathcal{X} = (\mathbf{T}, X)$ is a single-valued everywhere defined map ξ on $\mathbf{T}$ such that $\xi(t) \in X_t$ for every $t \in \mathbf{T}$. We use $Sec(\mathcal{X})$ to denote the set of all sections of $\mathcal{X}$. If $\mathcal{X} = (\mathbf{T}, X)$ is a C^k-SSB, and $\xi \in Sec(\mathcal{X})$, the familiy $T_\xi\mathcal{X} = \{T_{\xi(t)}X_t\}_{t\in\mathbf{T}}$ is the ***tangent bundle*** of $\mathcal{X}$ along ξ. Clearly, $T_\xi\mathcal{X}$ is an FDLS-SSB.

If $(\mathbf{T}, X)$ is an SSB, a ***flow*** on $(\mathbf{T}, X)$ is a family $f = \{f_{t,s}\}_{(t,s)\in\mathbf{T}^{2,\geq}}$ such that (1) $f_{t,s} : X_s \mapsto\!\!\!\!\rightarrow X_t$ whenever $(t,s) \in \mathbf{T}^{2,\geq}$; (2) $f_{t,t}$ is the identity map of X_t whenever $t \in \mathbf{T}$; and (3) $f_{t,s} \circ f_{s,r} = f_{t,r}$ whenever $(t,s,r) \in \mathbf{T}^{3,\geq}$. A ***flow*** is a triple $\mathcal{F} = (\mathbf{T}, X, f)$ such that $(\mathbf{T}, X)$ is an SSB and f is a flow on $(\mathbf{T}, X)$. If, for $i = 1, 2$, $\mathcal{F}^i = (\mathbf{T}, X, f^i)$ are flows with the same SSB $\mathcal{X} = (\mathbf{T}, X)$, we say that $\mathcal{F}^1$ is a ***subflow*** of $\mathcal{F}^2$ if $\mathrm{Gr}(f^1_{t,s}) \subseteq \mathrm{Gr}(f^2_{t,s})$ whenever $(t,s) \in \mathbf{T}^{2,\geq}$. If $\mathcal{F} = (\mathbf{T}, X, f)$ is a flow, then a ***trajectory*** of $\mathcal{F}$ is a section $\xi \in Sec(\mathcal{X})$ such that

[2] In our previous papers on the subject, CCA maps were called "regular maps." We have now adopted the name "Cellina continuously approximable" because these maps were actually introduced by A. Cellina in his work of the 1960s.

$\xi(t) \in f_{t,s}(\xi(s))$ whenever $(t,s) \in \mathbf{T}^{2,\geq}$. We use $Traj(\mathcal{F})$ to denote the set of all trajectories of $\mathcal{F}$.

Augmented flows. If $\mathcal{X} = (\mathbf{T}, X)$ is an SSB and A is an LS, we use X^A to denote the family $\{X_t^A\}_{t\in\mathbf{T}}$ (recall that $X_t^A \stackrel{\text{def}}{=} X_t \times A$), and $\mathcal{X}^A$ to denote the SSB $(\mathbf{T}, X^A)$. An A-***augmented flow*** on $\mathcal{X}$ is a flow $(\mathbf{T}, X^A, F)$ on $\mathcal{X}^A$ such that $F_{t,s} \in SVM^A(X_s, X_t)$ whenever $(t,s) \in \mathbf{T}^{2,\geq}$. We recall that the SVMs $F_{t,s}$ can be recovered from the SVMs $\check{F}_{t,s} : X_s \mapsto\!\!\!\!\rightarrow X_t^A$ defined by $\check{F}_{t,s}(x) = F_{t,s}(x,0)$, and that the flow composition law for F amounts to saying, in terms of the $\check{F}_{t,s}$, that $\check{F}_{t,r} = \check{F}_{t,s} \circ^A \check{F}_{s,r}$ whenever $(t,s,r) \in \mathbf{T}^{3,\geq}$, where $\circ^A$ stands for A-augmented composition.

IV. GDQs of flows along trajectories.

If $\mathcal{F} = (X, \mathbf{T}, \mathbf{f})$ is a flow, $(\mathbf{T}, X)$ is a C^1-SSB, and ξ is a trajectory of $\mathcal{F}$, then a ***GDQ*** of $\mathcal{F}$ along ξ is a family $g = \{g_{t,s}\}_{(t,s)\in\mathbf{T}^{2,\geq}}$ such that

(1) $g_{t,s} \in GDQ(f_{t,s}; \xi(s), \xi(t); X_s)$ whenever (t,s) belongs to $\mathbf{T}^{2,\geq}$;

(2) $g_{t,t} = \{\mathbb{I}_{T_{\xi(t)}X_t}\}$ whenever $t \in \mathbf{T}$;

(3) $g_{t,s} \circ g_{s,r} = g_{t,r}$ whenever $(t,s,r) \in \mathbf{T}^{3,\geq}$.

Adjoint vectors. If $\mathcal{F} = (\mathbf{T}, X, f)$ is a flow on a bundle $(\mathbf{T}, X)$ of manifolds of class C^1, $\xi \in Traj(\mathcal{F})$, and g is a GDQ of $\mathcal{F}$ along ξ, then a g-***adjoint vector*** (or g-***momentum***) for $\mathcal{F}$ along ξ is a map $\mathbf{T} \ni t \mapsto \pi(t) \in T^\dagger_{\xi(t)}X_t$ that satisfies the following ***integrated adjoint differential inclusion*** for g:

$$\pi(s) \in \pi(t) \circ g_{t,s} \quad \text{whenever } (t,s) \in \mathbf{T}^{2,\geq}\,. \tag{1}$$

A family $\ell = \{\ell_{t,s}\}_{(t,s)\in\mathbf{T}^{2,\geq}}$ such that $\ell_{t,s} \in g_{t,s}$ for each t,s and $\ell_{t,s} \circ \ell_{s,r} = \ell_{t,r}$ whenever $(t,s,r) \in \mathbf{T}^{3,\geq}$ is called a ***compatible selection*** of g.

Proposition 4.1: A map $\mathbf{T} \ni t \mapsto \pi(t) \in T^\dagger_{\xi(t)}X_t$ is a g-momentum if and only if there exists a compatible selection ℓ of g such that $\pi(s) = \pi(t) \circ \ell_{t,s}$ whenever $(t,s) \in \mathbf{T}^{2,\geq}$.

Proof: The "if" implication is trivial. To prove the other one, we assume that π is a g-momentum and find ℓ.

The product space $\Gamma = \Pi_{(t,s)\in\mathbf{T}^{2,\geq}} g_{t,s}$, endowed with the product topology, is compact by Tikhonov's theorem. The members of Γ are all the families $\ell = \{\ell_{t,s}\}_{(t,s)\in\mathbf{T}^{2,\geq}}$ such that $\ell_{t,s} \in g_{t,s}$ for each t,s. If S is a subset of $\mathbf{T}$, let $\Gamma^{c,S}$ denote the set of all $\ell \in \Gamma$ that are compatible over S, in the sense that $\ell_{t,s} \circ \ell_{s,r} = \ell_{t,r}$ whenever $r \leq s \leq t$ and $r,s,t \in S$. Then $\Gamma^{c,S}$ is the intersection of the sets $\Gamma^{c,\{r,s,t\}}$ over all triples $(t,s,r) \in S \times S \times S$ such that $r \leq s \leq t$. Each $\Gamma^{c,\{r,s,t\}}$ is closed, because the map $\ell \mapsto \ell_{t,s} \circ \ell_{s,r} - \ell_{t,r}$—from Γ to $LM(T_{\xi(r)}X_r, T_{\xi(t)}X_t)$—is continuous. Therefore $\Gamma^{c,S}$ is closed. Let $\Gamma^{\pi,S}$ be the set of those $\ell \in \Gamma$ that are π-compatible over S, in the sense that $\pi(s) = \pi(t) \circ \ell_{t,s}$ whenever $s,t \in S$ and $s \leq t$. Again, $\Gamma^{\pi,S}$ is the intersection of the sets $\Gamma^{\pi,\{s,t\}}$ over all pairs $(s,t) \in S \times S$ such that $s \leq t$, and the sets $\Gamma^{\pi,\{s,t\}}$ are closed, because the map $\ell \mapsto \pi(s) - \pi(t) \circ \ell_{t,s}$ (from Γ to $T_{\xi(s)}X_s$) is continuous, so $\Gamma^{\pi,S}$ is closed. Hence all the $\Gamma^{c,S}$ and $\Gamma^{\pi,S}$ are compact subsets of Γ. Let $\Gamma^S = \Gamma^{c,S} \cap \Gamma^{\pi,S}$. We want to prove that $\Gamma^{\mathbf{T}} \neq \emptyset$. Clearly, $\Gamma^{\mathbf{T}} = \bigcap \Big\{ \Gamma^S : S \subseteq \mathbf{T}\,,\ S \text{ finite} \Big\}$. Hence our conclusion will follow if we show that the family $\{\Gamma^S\}_{S\subseteq\mathbf{T}\,,\,S \text{ finite}}$ has the finite intersection property, i.e., that every finite intersection $\Gamma^{S_1} \cap \Gamma^{S_2} \cap \cdots \cap \Gamma^{S_m}$ is nonempty. But $\Gamma^{S_1} \cap \Gamma^{S_2} \cap \cdots \cap \Gamma^{S_m} \supseteq \Gamma^{S_1\cup S_2\cup\cdots\cup S_m}$. So it suffices to prove that $\Gamma^S \neq \emptyset$ whenever S is a finite subset of $\mathbf{T}$. Write $S = \{s_1, s_2, \ldots, s_m\}$ with $s_1 < s_2 < \cdots < s_m$. For $j = 1, \ldots, m-1$, pick $\ell^j \in g_{s_j,s_{j-1}}$ such that $\pi(s_{j-1}) = \pi(s_j) \circ \ell^j$. If $s,t \in S$ and $s \leq t$, define

$$\ell_{t,s} = \ell^{j-1} \circ \ell^{j-2} \circ \cdots \circ \ell^i \tag{2}$$

if $s = s_i$, $t = s_j$. (It is clear that $j \geq i$; the right-hand side of (2) is $\mathbb{I}_{T_{\xi(s)}}X_s$ if $s = t$.) Extend the family $\{\ell_{t,s}\}_{s\leq t, s\in S, t\in S}$ by picking $\ell_{t,s}$ to be an arbitrary member of $g_{t,s}$ if $s \leq t$ but $(s,t) \notin S \times S$. (This, of course, uses the Axiom of Choice.) Then the resulting family ℓ belongs to Γ^S, proving that $\Gamma^S \neq \emptyset$ as desired. ∎

Augmented GDQs of augmented flows. According to our general definitions, if $\mathcal{F} = (\mathbf{T}, X^A, F)$ is an A-augmented flow on an SSB $\mathcal{X} = (\mathbf{T}, X)$, a trajectory of $\mathcal{F}$ is a map $\mathbf{T} \ni t \mapsto \Xi(t) \in X_t \times A$ such that $\Xi(t) \in F_{t,s}(\Xi(s))$ whenever $(t,s) \in \mathbf{T}^{2,\geq}$. In terms of the maps $\check{F}_{t,s}$, we can then regard a trajectory Ξ as being a pair (ξ, ξ^0) of maps $\mathbf{T} \ni t \mapsto \xi(t) \in X_t$ and $\mathbf{T} \ni t \mapsto \xi^0(t) \in A$ such that $(\xi(t), \xi^0(t) - \xi^0(s)) \in \check{F}_{t,s}(\xi(s))$ whenever $(t,s) \in \mathbf{T}^{2,\geq}$.

If $\mathcal{X} = (\mathbf{T}, X)$ is C^1-SSB, A is a FDLS, $\mathcal{F} = (\mathbf{T}, X^A, F)$ is an A-augmented flow on $\mathcal{X}$, and $\Xi = (\xi, \xi^0) \in Traj(\mathcal{F})$, an ***augmented GDQ*** of $\mathcal{F}$ along Ξ is a GDQ $G = \{G_{t,s}\}_{(t,s)\in\mathbf{T}^{2,\geq}}$ of $\mathcal{F}$ along Ξ such that every set $G_{t,s}$ is a subset of $LM^A(T_{\xi(s)}X_s, T_{\xi(t)}X_t)$, i.e., a set of A-augmented linear maps from $T_{\xi(s)}X_s$ to $T_{\xi(t)}X_t$. Therefore an A-augmented GDQ G of $\mathcal{F}$ along Ξ is such that each $G_{t,s}$ is a nonempty compact set of pairs $L = (\ell, \ell^0)$ such that $\ell \in LM(T_{\xi(s)}X_s, T_{\xi(t)}X_t)$ and $\ell^0 \in LM(T_{\xi(s)}X_s, A)$. A G-momentum is then a map $\mathbf{T} \ni t \mapsto \Pi(t) \in (T_{\xi(t)}X_t \times A)^\dagger$ such that $\Pi(s) \in \Pi(t) \circ G_{t,s}$ whenever $(t,s) \in \mathbf{T}^{2,\geq}$. Hence a G-momentum is a pair (π, π^0) of functions $\mathbf{T} \ni t \mapsto \pi(t) \in T^\dagger_{\xi(t)}X_t$, $\mathbf{T} \ni t \mapsto \pi^0(t) \in A^\dagger$ that satisfy

$$\pi^0(s) = \pi^0(t) \tag{3}$$

$$(\exists (\ell, \ell^0) \in G_{t,s})\ \pi(s) = \pi(t) \circ \ell + \pi^0(t)\ell^0 \tag{4}$$

for $s,t \in \mathbf{T}^{2,\geq}$. Hence *the momentum-augmenting function* $t \mapsto \pi^0(t) \in A^\dagger$ *for a G-momentum for an A-augmented GDQ G is constant*. Condition (4) is the ***integrated*** A-***augmented adjoint differential inclusion*** for G.

V. Variations

Variations of SVMs. A *pointed finite-dimensional convex subset* (abbr. "PFDCS") is a set Q endowed with a structure consisting of a FDLS $\mathcal{A}_Q$ (the "ambient space" of Q) such that (1) Q is a convex subset of $\mathcal{A}_Q$ with nonempty interior, and (2) the origin 0_Q of $\mathcal{A}_Q$ belongs to Q. If

X, Y are sets and $f, f' \in SVM(X,Y)$, a ***variation of*** f ***in*** f' is a set-valued map $v : Q(v) \times X \mapsto\!\!\!\!\rightarrow Y$ such that (a) $Q(v)$ is a compact PFDCS, (b) $v(0_{Q(v)}, x) = f(x)$ for every $x \in X$, and (c) $v(q,x) \subseteq f'(x)$ whenever $q \in Q(v)$, $x \in X$. The PFDCS $Q(v)$ is the ***parameter domain*** of the variation v. A ***variation of*** f is a variation of f in the "maximal" set-valued map $(X, Y, X \times Y)$. If X, Y are manifolds of class C^1, $\bar{x} \in X$, and $\bar{y} \in Y$, a ***variational GDQ*** (abbr. VGDQ) ***at*** $(\bar{x}, \bar{y})$ of a variation $v : Q(v) \times X \mapsto\!\!\!\!\rightarrow Y$ of an SVM $f : X \mapsto\!\!\!\!\rightarrow Y$ is a GDQ of $v^{\mathcal{A}}$ at $((0_{Q(v)}, \bar{x}), \bar{y})$ along $Q(v) \times X$, where $v^{\mathcal{A}} : \mathcal{A}_{Q(v)} \times X \mapsto\!\!\!\!\rightarrow Y$ is the SVM such that $v^{\mathcal{A}}(q,x) = v(q,x)$ if $q \in Q(v)$, $x \in X$, and $v^{\mathcal{A}}(q,x) = \emptyset$ if $q \in \mathcal{A}_{Q(v)} \backslash Q(v)$, $x \in X$. We use $VGDQ(v; \bar{x}, \bar{y})$ to denote the set of all VGDQs of v at $(\bar{x}, \bar{y})$, so $VGDQ(v; \bar{x}, \bar{y}) = GDQ(v^{\mathcal{A}}; (0_{Q(v)}, \bar{x}), \bar{y}; Q(v) \times X)$. If $\gamma \in VGDQ(v; \bar{x}, \bar{y})$, then the partial GDQs γ_X, $\gamma_{\mathcal{A}_{Q(v)}}$ are, respectively, ***the state part*** and the ***the parameter part*** of γ, and we write $s(\gamma) \stackrel{\text{def}}{=} \gamma_X$ and $p(\gamma) \stackrel{\text{def}}{=} \gamma_{\mathcal{A}_{Q(v)}}$. We know that $s(\gamma) \in GDQ(f; \bar{x}, \bar{y}; X)$. Given a $g \in GDQ(f; \bar{x}, \bar{y}; X)$, and a variation v of f, a $\gamma \in VGDQ(v; \bar{x}, \bar{y})$ such that $s(\gamma) = g$ will be said to be an ***extension*** of g. We write $VGDQ_g(v; \bar{x}, \bar{y})$ to denote the set of all $\gamma \in VGDQ(v; \bar{x}, \bar{y})$ that are extensions of g.

Infinitesimal variations. If $\mathcal{F} = (\mathbf{T}, X, f)$ is a flow on a C^1-SSB $\mathcal{X} = (\mathbf{T}, X)$, $\xi \in Traj(\mathcal{F})$, and g is a GDQ of $\mathcal{F}$ along ξ, an ***infinitesimal variation of*** g ***along*** ξ is a triple $\Gamma = (|\Gamma|, Q, \gamma)$ such that (a) $|\Gamma|$ (the "carrier" of Γ) is a compact interval $[\alpha, \beta]$ such that $(\beta, \alpha) \in \mathbf{T}^{2,\geq}$, (b) Q is a PFDCS, and (c) γ is a nonempty compact subset of $LM(\mathcal{A}_Q \times T_{\xi(\alpha)} X_\alpha, T_{\xi(\beta)} X_\beta)$ such that $s(\gamma) = g_{\beta,\alpha}$, where $s(\gamma)$ is the set of all linear maps $T_{\xi(\alpha)} X_\alpha \ni x \mapsto L(0_Q, x) \in T_{\xi(\beta)} X_\beta$ for all $L \in \gamma$. If $\Gamma = ([\alpha, \beta], Q, \gamma)$ is an infinitesimal variation of g along ξ, and $(\tilde{\beta}, \beta, \alpha, \tilde{\alpha}) \in \mathbf{T}^{4,\geq}$, then we can ***expand*** Γ to an infinitesimal variation $\Gamma^{[\tilde{\alpha}, \tilde{\beta}]}$ of g along ξ by letting $\Gamma^{[\tilde{\alpha}, \tilde{\beta}]} = ([\tilde{\alpha}, \tilde{\beta}], Q, \hat{\gamma})$, where $\hat{\gamma} = g_{\tilde{\beta}, \beta} \circ \gamma \circ (\mathbb{I}_{\mathcal{A}_Q} \times g_{\alpha, \tilde{\alpha}})$.

If $\Gamma_1, \ldots, \Gamma_m$ are infinitesimal variations of g along ξ, we define the ***combined infinitesimal variation*** $\Gamma = \Gamma_1 \Box \Gamma_2 \Box \ldots \Box \Gamma_m$ as follows. Let $\Gamma_j = ([\alpha_j, \beta_j], Q_j, \gamma_j)$ for $j = 1, \ldots, m$. Let $\alpha = \min\{\alpha_j\}_{j=1,\ldots,m}$, $\beta = \max\{\beta_j\}_{j=1,\ldots,m}$. Let σ be an $m+1$-tuple $(\ell, \mu_1, \ldots, \mu_m)$ consisting of a compatible selection $\ell = \{\ell_{t,s}\}_{(t,s) \in \mathbf{T}^{2,\geq}}$ of g and linear maps $\mu_j \in \gamma_j$ such that $s(\mu_j) = \ell_{\beta_j, \alpha_j}$. Define $\theta(\sigma)$ to be the linear map from $\mathcal{A}_{Q_1} \times \cdots \mathcal{A}_{Q_m} \times T_{\xi(\alpha)} X_\alpha$ to $T_{\xi(\beta)} X_\beta$ that sends a point $(q_1, q_2, \ldots, q_m, h)$ belonging to $\mathcal{A}_{Q_1} \times \cdots \mathcal{A}_{Q_m} \times T_{\xi(\alpha)} X_\alpha$ to the vector $\ell_{\beta,\alpha}(h) + \sum_{j=1}^m \ell_{\beta, \beta_j} \nu_j(q_j)$, where $\nu_j = p(\mu_j)$. We then let $\Gamma = ([\alpha, \beta], Q, \gamma)$, where $Q = Q_1 \times Q_2 \cdots \times Q_m$ and γ is the set of all maps $\theta(\sigma)$, for all σ.

Compatibility of sets of infinitesimal variations. Given $\mathbf{T}$, X, f, $\mathcal{F}$, $\mathcal{X}$, ξ, g, as in the previous subsection, as well as points a, b of $\mathbf{T}$ such that $a \leq b$, and an $R \in SVM(X_a, X_b)$, a set $\mathcal{V}$ of infinitesimal variations of g along ξ is said to be $\mathcal{F}$-R***-compatible over*** $[a,b]$ if (a) all the carrier intervals of the members V of $\mathcal{V}$ are subintervals of $[a,b]$, and (b) whenever $\mathcal{V}_0$ is a finite subset of $\mathcal{V}$, the members of $\mathcal{V}_0$ can be arranged in a sequence $\Gamma_1, \ldots, \Gamma_m$ such that the expansion $\Gamma^{[a,b]} = ([a,b], Q, \gamma)$ of the combined infinitesimal variation $\Gamma = \Gamma_1 \Box \Gamma_2 \Box \ldots \Box \Gamma_m$ satisfies: *there exists a variation* v *of* $f_{b,a}$ *in* R *such that* $\gamma \in VGDQ(v; \xi(a), \xi(b))$.

We are now ready to state our version of the maximum principle. The proof is straighforward, based on applying the definitions and the separation theorem 2.6.

VI. A NONSMOOTH MAXIMUM PRINCIPLE.

The reader should think of ξ, $\mathcal{F}$, and R as the "reference trajectory," "reference flow," and "reachability relation over $[a,b]$" (so that "$y \in R(x)$" means "y is reachable at time b by means of a trajectory that starts at x at time a") of a control system. We will first state the result as a necessary condition for the graph of R to be separated from some other subset S of $X_a \times X_b$, and then outline how the necessary conditions for optimal control are obtained by applying the theorem to $\mathbb{R}$-augmented flows.

Theorem 6.1: Assume that $\mathcal{F} = (\mathbf{T}, X, f)$ is a flow on a bundle $\mathcal{X} = (\mathbf{T}, X)$ of C^1 manifolds, $a, b \in \mathbf{T}$, $a \leq b$, $R \in SVM(X_a, X_b)$, ξ is a trajectory of $\mathcal{F}$, and g is a GDQ of $\mathcal{F}$ along ξ. Let S be a subset of $X_a \times X_b$ such that $(\xi(a), \xi(b)) \in S$, and let $C \subseteq T_{\xi(a)} X_a \times T_{\xi(b)} X_b$ be a limiting approximating cone of S at $(\xi(a), \xi(b))$ which is not a linear subspace. Let $\mathcal{V}$ be a set of infinitesimal variations of $\mathcal{F}$ along ξ which is $\mathcal{F} - R$-compatible over $[a,b]$. Then, a necessary condition for the separation condition $\mathrm{Gr}(R) \cap S = \{(\xi(a), \xi(b))\}$ is that there exist (a) a compatible selection $\{\ell_{t,s}\}_{(s,t) \in \mathbf{T}^{2,\geq}}$ of g, (b) a family $\{\lambda_V\}_{V \in \mathcal{V}}$ such that, for each $V = ([\alpha, \beta], Q, v) \in \mathcal{V}$, λ_V is a member of v such that $s(\lambda_V) = \ell_{\beta,\alpha}$, and (c) a covector $\bar{\pi} \in T^\dagger_{\xi(b)} X_b \backslash \{0\}$, such that, if we let $\pi(t) = \bar{\pi} \circ \ell_{b,t}$ for every $t \in \mathbf{T}$, then (a) $-\pi(a)h + \pi(b)(\tilde{h}) \geq 0$ for every $(h, \tilde{h}) \in C$, and (b) $(\pi(\beta)(\lambda_V(q, 0)) \leq 0$ whenever $V = ([\alpha, \beta], Q, v) \in \mathcal{V}$ and $q \in Q$. ■

If we take $\mathrm{Gr}(R) = \{\xi(a)\} \times R_0$, $S = \{\xi(a)\} \times S_0$, where R_0 and S_0 are subsets of X_b, then the theorem gives a necessary condition for a control trajectory ξ to be such that the reachable set from $\xi(a)$ over the interval $[a,b]$ is separated at $\xi(b)$ from another set S_0, provided that S_0 has a limiting approximating cone C_0 which is not a subspace.

If we take $A = \mathbb{R}$, we can apply the theorem to an A-augmented flow $\mathcal{F}$, A-augmented trajectory $\Xi = (\xi, \xi^0)$, and A-augmented GDQ G of $\mathcal{F}$ along Ξ, using in the role of S a subset of the product $X_a^{\mathbb{R}} \times X_b^{\mathbb{R}}$ (i.e., of $(X_a \times \mathbb{R}) \times (X_b \times \mathbb{R})$) of the form $\{(x,r), (x',r') : (x,x') \in S_0 \wedge r = \xi^0(a) \wedge r' - r \leq \xi^0(b) - \xi^0(a) - d((\xi(a), \xi(b)), (x,x'))^2\}$, for some subset S_0 of $X_a \times X_b$, where d is the distance induced by some Riemannian metric on $X_a \times X_b$. We then get a necessary condition for minimization: if Ξ is such that

$\xi^0(b) - \xi^0(a) \le r' - r$ whenever $((x, r), (x', r')) \in \mathrm{Gr}(R)$ and $(x, x') \in S_0$, then $\mathrm{Gr}(R) \cap S = \{(\Xi(a), \Xi(b))\}$. If C_0 is a limiting approximating cone to S_0 at $(\xi(a), \xi(b))$, then $C = \{((c, 0), (c', \rho)) : (c, c') \in C \wedge \rho \le 0\}$ is a limiting approximating cone to S at $(\Xi(a), \Xi(b))$. (Notice that in this case there is no need to require that C_0 not be a linear subspace, because C is never a linear subspace, even if C_0 is.) The theorem then gives a momentum $\Pi = (\pi, -\pi^0)$, where the momentum-augmenting function π^0 is then just a constant real number. The transversality condition (a) of the theorem then says that $\pi(b)\tilde{h} - \pi(a)h - \pi^0\rho \ge 0$ for all $(h, \tilde{h}) \in C_0$ and all $\rho \in \mathbb{R}$ such that $\rho \le 0$, which is of course equivalent to the usual pair of separate conditions (i) $\pi(b)\tilde{h} - \pi(a)h \ge 0$ for all $(h, \tilde{h}) \in C_0$ and (ii) $\pi^0 \ge 0$.

Finally, the theorem can be applied with higher-dimensional augmentation spaces, to obtain, for example, necessary conditions for Pareto optimality.

VII. Classical and high-order conditions

Theorem 6.1 can be applied to systems that are "classical" on several nonoverlapping subintervals $[c_j, d_j]$ of $[a, b]$, and makes it possible to obtain "adjoint vectors" that satisfy all the necessary conditions given by these theorems as well as the necessary conditions of the classical nonsmooth and high-order versions of the maximum principle in the intervals $[c_j, d_j]$. We now discuss two examples of this kind of result. We just consider the case of augmented systems corresponding to minimization problems.

First of all, suppose that our augmented control system is given, on each of a finite collection $I_1, I_2, \ldots, I_m$ of subintervals $I_j = [c_j, d_j]$ of $[a, b]$ such that $c_j < d_j$ for all j and $d_{j-1} \le c_j$ for $j = 2, \ldots, m$, by classical control equations $\dot{\xi}(t) = f_j(\xi(t), \eta(t), t)$, $\dot{\xi}^0(t) = f_j^0(\xi(t), \eta(t), t)$, where f_j and f_j^0 are, respectively, vector-valued and scalar-valued functions of $x \in X_j$, $u \in U_j$, $t \in [c_j, d_j]$, the X_js are manifolds of class C^1, and the U_j are sets. Suppose that the reference trajectory ξ_* arises, on I_j, from a control $\eta_{*,j}$ which is such that $f_j(x, \eta_{*,j}(t), t)$ and $f_j^0(x, \eta_{*,j}(t), t)$ are Lipschitz-continuous as functions of x for (x, t) in some tube containing the graph of the restriction $\xi_* \lceil I_j$, measurable with respect to t for each x, and bounded by an integrable function of t. Assume, moreover, that for each fixed $u \in U_j$ both $f_j(x, u, t)$ and $f_j^0(x, u, t)$ are measurable functions of t for each x, and continuous functions of x for each t. Finally, assume that the class of admissible controls contains all constant controls and the reference control, and is closed under intertwining on mesaurable sets. The well-known Łojasiewicz Maximum Principle (cf., e.g., [2]) then applies on each I_j, and yields an adjoint vector π_j and abnormal multiplier π_j^0 such that the usual adjoint differential inclusion (i.e., $-\dot{\pi}_j(t) \in \partial_x H_j(x, \pi(t), \pi^0, \eta_{*,j}(t), t)|_{x=\xi_*(t)}$, where $H_j(x, p, p^0, u, t) \stackrel{\mathrm{def}}{=} p \cdot f_j(x, u, t) - p^0 f_j^0(x, u, t)$, and ∂_x stands for "Clarke generalized gradient with respect to x keeping all other variables fixed") holds, as well as the Hamiltonian maximization condition (i.e., the statement that for every $u \in U$ the inequality $H_j(\xi_*(t), \pi_j(t), \pi^0, u, t) \le H_j(\xi_*(t), \pi_j(t), \pi^0, \eta_{*,j}, t)$ is satisfied for almost every $t \in I_j$). Theorem 6.1 produces a *single* multiplier (π, π^0), which satisfies the adjoint inclusion and the Hamiltonian maximization condition on each I_j as well as all the other conditions arising from variations on intervals that do not overlap with the I_j.

For a second example, we specialize further and look at the minimum-time problem, in the case when our system is autonomous and, on some of the intervals I_j (more precisely, for all j in some subset J of $\{1, \ldots, m\}$), the control equations are of the form $\dot{\xi}(t) = f_{0j}(x) + \sum_{i=1}^{k_j} u_i f_{ij}(\xi(t))$, and the set U_j of control values is the cube $[-1, 1]^{k_j}$. (Naturally, minimum-time problems do not have a fixed time interval, but autonomous problems with a variable time interval can be reduced to problems with a fixed time interval in a well-known way. Assume in addition that the vector fields f_{ij} are of class C^2. Then it is well known that to the usually necessary conditions for a minimum one can add, on each I_j such that $j \in J$, the "Goh condition," i.e., the requirement that $\pi_j(t) \cdot [f_{ij}, f_{i'j}](\xi_*(t)) = 0$ for every triple (t, i, i') such that $t \in I_j$, t is a Lebesgue point of the reference control $\eta_{*,j}$, and the inequalities $-1 < (\eta_{*,j})_i(t) < 1$ and $-1 < (\eta_{*,j})_{i'}(t) < 1$ hold. Our result yields the stronger conclusion that there exists a single multiplier on the whole interval $[a, b]$ that satisfies all these high-order conditions on all the I_j for $j \in J$.

VIII. Concluding remarks

It would be desirable to strengthen our results so that, for example, the regularity requirement of the second example of the previous section is weakened to allow the vector fields f_{ij} to be just of class C^1, or perhaps even Lipschitz, but at the moment it is unclear how this could be done. An even more delicate question is that of extending the "Goh condition" of the previous section to the case when the contol sets U_j are balls rather than cubes. This seems to lead to a whole new set of questions that appear difficult to handle with our methods. Specifically, the Goh condition for this case gives the equalities $\pi_j(t) \cdot [f_{ij}, f_{i'j}](\xi_*(t)) = 0$ *provided that the reference trajectory is such that all the abnormal multipliers, for all Hamiltonian-maximizing adjoint vectors, vanish.* This is a global condition, and we do not know how to split it into conditions that would apply on separate intervals.

References

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

[2] Sussmann, H. J., *A strong version of the Łojasiewicz maximum principle.* In *Optimal Control of Differential Equations,* N. H. Pavel Ed., Marcel Dekker Inc., New York, 1994, pp. 293–309.

[3] 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.