Bianchini-Stefani variations and generalized differentials

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 a series of nonsmooth versions of the Pontryagin Maximum Principle, we used generalized differentials of set-valued maps, flows, and abstract variations. Bianchini and Stefani have introduced a notion of possibly high-order variational vector that has the summability property. We consider a slightly more general class of variational vectors than that defined by Bianchini and Stefani, and prove that the convex combinations of these vectors arise as "differentials" of variations that are differentiable in the sense of one of our generalized differentiation theories, namely, that of "approximate generalized differential quotients" (AGDQs).**

A NOTE FOR THE REVIEWERS (which, naturally, will be omitted in the final version): In the final version I will make the paper shorter, by removing some technical details, in order to conform to the 6-page limit.

I. Introduction

In a series of papers, we showed how to derive general, nonsmooth versions of the Pontryagin Maximum Principle using generalized differentials of set-valued maps, flows, and abstract point variations. The use of general variations rather than the nedle variations used to prove the ordinary maximum principle makes it possible to obtain high-order versions of the maximum principle. The main technical difficulty with these general abstract variations is that they need not have the summability property, which is absolutely essential in order to derive the necessary conditions for optimality.

R. M. Bianchini and G. Stefani (cf. [1], [2], [3], [4], [5]) proposed a concept of high-order variation that has good summability properties. The goal of this note is to relate this concept to a theory generalized differentials, by describing a slightly more general version of the Bianchini-Stefani variations, and showing that they are differentiable in the precise sense of the theory of "Approximate Generalized Differential Quotients" (AGDQs). This makes is possible to use these variations in order to get additional necessary conditions for an optimum in situations such as the very general one described in [12], where the differentials involved are generalized differential quotients, and *a fortiori* AGDQs.

1) Preliminary remarks on notation: We will use the notations and abbreviations of [12]. In particular, "FDRLS" stands for "finite-dimensional real linear space," "FDNRLS" for "normed FDRLS," "SVM" for "set-valued map," and "CCA" for "Cellina continuously approximable." Also, $\mathbb{I}_S$ is the identity map of the set S, $\mathrm{So}(f)$, $\mathrm{Ta}(f)$ and $\mathrm{Gr}(f)$ are, respectively, the source, target, and graph of a SVM f, $f : A \mapsto\!\!\!\rightarrow B$ means "f is a set-valued map from A to B," and $f : A \hookrightarrow B$ means "f is a single-valued possibly partially defined map from A to B."

We use $\boldsymbol{\Theta}$ to denote the class of all functions $\theta : [0, +\infty\,[\mapsto [0, +\infty]$ such that

- θ is monotonically nondecreasing (that is, $\theta(s) \le \theta(t)$ whenever s, t are such that $0 \le s \le t < +\infty$);
- $\theta(0) = 0$ and $\lim_{s\downarrow 0} \theta(s) = 0$.

If X is a FDNRLS, $x_* \in X$, $r > 0$, then $\mathbb{B}_X(x_*, r)$, $\bar{\mathbb{B}}_X(x_*, r)$ are, respectively, the open ball $\{x \in X : \|x - x_*\| < r\}$ and the closed ball $\{x \in X : \|x - x_*\| \le r\}$. If X, Y are FDNRLSs, then $Lin(X, Y)$, $Aff(X, Y)$ will denote, respectively, the set of all linear maps and the set of all affine maps from X to Y. Hence the members of $Aff(X, Y)$ are the maps $X \ni x \mapsto A(x) = L \cdot x + h$, $L \in Lin(X, Y)$, $h \in Y$. (For a map A of this form, the linear map $L \in Lin(X, Y)$ and the vector $h \in Y$ are the ***linear part*** and the ***constant part*** of A.) We identify $Aff(X, Y)$ with $Lin(X, Y) \times Y$ by identifying each $A \in Aff(X, Y)$ with the pair $(L, h) \in Lin(X, Y) \times Y$, where L, h are, respectively, the linear part and the constant part of A.

We endow $Lin(X, Y)$ with the operator norm $\|\cdot\|_{op}$. If $\Lambda \subseteq Lin(X, Y)$ and $\delta > 0$, we define

$$\Lambda^\delta = \{L \in Lin(X, Y) : \mathrm{dist}(L, \Lambda) \le \delta\}\,,$$

where $\mathrm{dist}(L, \Lambda) = \inf\{\|L - L'\|_{op} : L' \in \Lambda\}$. Notice that if $L \in Lin(X, Y)$, then $\mathrm{dist}(L, \emptyset) = +\infty$. In particular, if $\Lambda = \emptyset$ then $\Lambda^\delta = \emptyset$. Notice also that Λ^δ is compact if Λ is compact and Λ^δ is convex if Λ is convex.

II. Approximate Generalized Differential Quotients

Definition 2.1: Assume that X, Y are FDNRLSs, $F : X \mapsto\!\!\!\rightarrow Y$ is a set-valued map, Λ is a compact subset of $Lin(X, Y)$, $\bar{x}_* \in X$, $\bar{y}_* \in Y$, and $S \subseteq X$. We say that Λ is an ***approximate generalized differential quotient of*** F ***at*** $(\bar{x}_*, \bar{y}_*)$ ***in the direction of*** S—and write $\Lambda \in AGDQ(F, \bar{x}_*, \bar{y}_*, S)$—if there exists a function

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

$\theta \in \boldsymbol{\Theta}$—called an ***AGDQ modulus for*** $(\Lambda, F, \bar{x}_*, \bar{y}_*, S)$—having the property that

(**) *for every* $\varepsilon \in]0, +\infty[$ *such that* $\theta(\varepsilon) < \infty$ *there exists a set-valued map*

$$A^\varepsilon \in CCA(\bar{\mathbb{B}}_X(\bar{x}_*, \vec{\varepsilon}) \cap S, Aff(X, Y))$$

such that, whenever $x \in \bar{\mathbb{B}}_X(\bar{x}_*, \varepsilon) \cap S$ *and* $(L, h) \in A^\varepsilon(x)$, *it follows that* $L \in \Lambda^{\theta(\varepsilon)}$, $\|h\| \leq \theta(\varepsilon)\varepsilon$, *and* $\bar{y}_* + L \cdot (x - \bar{x}_*) + h \in F(x)$. ■

1) Properties of AGDQs: The following statement is the ***chain rule*** for AGDQs.

Theorem 2.2: For $i = 1, 2, 3$, let X_i be a FDNRLS, and let $\bar{x}_{*,i}$ be a point of X_i. Assume that, for $i = 1, 2$, (i) $F_i : X_i \mapsto\!\!\!\to X_{i+1}$ is a set-valued map, (ii) S_i is a subset of X_i, and (iii) $\Lambda_i \in AGDQ(F_i, \bar{x}_{*,i}, \bar{x}_{*,i+1}, S_i)$. Assume, in addition, that (iv) $F_1(S_1) \subseteq S_2$, and either (v) S_2 is a local quasiretract (cf. [13]) of X_2 at $\bar{x}_{*,2}$ or (v') there exists a neighborhood U of $\bar{x}_{*,1}$ in X_1 such that the restriction $F_1 \lceil (U \cap S_1)$ of F_1 to $U \cap S_1$ is single-valued. Then $\Lambda_2 \circ \Lambda_1 \in AGDQ(F_2 \circ F_1, \bar{x}_{*,1}, \bar{x}_{*,3}, S_1)$. ■

If M and N are manifolds of class C^1, $\bar{x}_* \in M$, $\bar{y}_* \in N$, $S \subseteq M$, and $F : M \mapsto\!\!\!\to N$, then it is possible to define a set $AGDQ(F, \bar{x}_*, \bar{y}_*, S)$ of compact subsets of the space $Lin(T_{\bar{x}_*}M, T_{\bar{y}_*}N)$ of linear maps from $T_{\bar{x}_*}M$ to $T_{\bar{y}_*}N$ as follows. We let $m = \dim M$, $n = \dim N$, and pick coordinate charts $M \ni x \hookrightarrow \xi(x) \in \mathbb{R}^m$, $N \ni y \hookrightarrow \eta(y) \in \mathbb{R}^n$, defined near $\bar{x}_*$, $\bar{y}_*$ and such that $\xi(\bar{x}_*) = 0$ and $\eta(\bar{y}_*) = 0$, and declare that a subset Λ of $Lin(T_{\bar{x}_*}M, T_{\bar{y}_*}N)$ belongs to $AGDQ(F, \bar{x}_*, \bar{y}_*, S)$ if the composite map $D\eta(\bar{y}_*) \circ \Lambda \circ D\xi(\bar{x}_*)^{-1}$ is in $AGDQ(\eta \circ F \circ \xi^{-1}, 0, 0, \xi(S)))$. It then follows easily from the chain rule that, with this definition, ***the set*** $AGDQ(F, \bar{x}_*, \bar{y}_*, S)$ ***does not depend on the choice of the charts*** ξ, η. In other words, ***the notions of an AGDQ is invariant under*** C^1 ***diffeomorphisms and therefore makes sense intrinsically on manifolds of class*** C^1.

Then the chain rule also holds on manifolds.

Proposition 2.3: Assume that (I) for $i = 1, 2, 3$, M_i is a manifold of class C^1 and $\bar{x}_{*,i} \in M_i$, and (II) for $i = 1, 2$, (II.1) $S_i \subseteq M_i$, (II.2) $F_i : M_i \mapsto\!\!\!\to M_{i+1}$, and (II.3) $\Lambda_i \in AGDQ(F_i, \bar{x}_{*,i}, \bar{x}_{*,i+1}, S_i)$. Assume, in addition, that either S_2 is a local quasiretract of M_2 or F_1 is single-valued on $U \cap S_1$ for some neighborhood U of $\bar{x}_{*,1}$. Then the composite $\Lambda_2 \circ \Lambda_1$ belongs to $AGDQ(F_2 \circ F_1, \bar{x}_{*,1}, \bar{x}_{*,3}, S_1)$. ■

2) Partial AGDQs: Suppose that (a) for $i = 1, 2$, X_i is a manifold of class C^1, $\bar{x}_i \in X_i$, and S_i is a subset of 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) $F : X \mapsto\!\!\!\to Y$, and (g) $\Lambda \in AGDQ(F; \bar{x}, \bar{y}; S)$. 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 AGDQs" Λ_{X_1}, Λ_{X_2} (where Λ_{X_1} is the set of all maps $T_{\bar{x}_1}X_1 \ni v \mapsto L(v, 0) \in T_{\bar{y}}Y$ for all $L \in \Lambda$, Λ_{X_2} is the set of all maps $T_{\bar{x}_2}X_2 \ni v \mapsto L(0, v) \in T_{\bar{y}}Y$ for all $L \in \Lambda$, and we are using the canonical identification of $T_{\bar{x}}X$ with $T_{\bar{x}_1}X_1 \times T_{\bar{x}_2}X_2$) are AGDQs 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.

3) Approximating multicones.: A ***multicone*** is a set of cones. A ***convex multicone*** is a set of convex ones.

Assume that M is a manifold of class C^1, S is a subset of M and $\bar{x}_* \in S$.

Definition 2.4: An ***AGDQ approximating multicone to*** S ***at*** $\bar{x}_*$ is a convex multicone $\mathcal{C}$ in $T_{\bar{x}_*}M$ such that there exist an $m \in \mathbb{Z}_+$, a set-valued map $F : \mathbb{R}^m \mapsto\!\!\!\to M$, a convex cone D in $\mathbb{R}^m$, and a $\Lambda \in AGDQ(F, 0, \bar{x}_*, D)$, such that $F(D) \subseteq S$ and $\mathcal{C} = \{LD : L \in \Lambda\}$. ■

4) Transversality of cones and multicones.: If S_1, S_2 are subsets of a linear space X, we define the ***difference*** $S_1 - S_2$ by letting $S_1 - S_2 = \{s_1 - s_2 : s_1 \in S_1, s_2 \in S_2\}$.

Definition 2.5: Let X be a FDRLS, and let C^1, C^2 be two convex cones in X. We say that C^1 and C^2 are ***transversal***, and write $C^1 \overline{\pitchfork} C^2$, if $C^1 - C^2 = X$. We say that C^1 and C^2 are ***strongly transversal***, and write $C^1 \overline{\overline{\pitchfork}} C^2$, if $C^1 \overline{\pitchfork} C^2$ and in addition $C^1 \cap C^2 \neq \{0\}$. ■

Definition 2.6: Let X be a FDRLS. We say that two convex multicones $\mathcal{C}^1$ and $\mathcal{C}^2$ in X are ***transversal***, and write $\mathcal{C}^1 \overline{\pitchfork} \mathcal{C}^2$, if $C^1 \overline{\pitchfork} C^2$ for all $C^1 \in \mathcal{C}^1$, $C^2 \in \mathcal{C}^2$. ■

Definition 2.7: Let X be a finite-dimensional real linear space. Let $\mathcal{C}^1$, $\mathcal{C}^2$ be convex multicones in X. We say that $\mathcal{C}^1$ and $\mathcal{C}^2$ are ***strongly transversal***, and write $\mathcal{C}^1 \overline{\overline{\pitchfork}} \mathcal{C}^2$, if (i) $\mathcal{C}^1 \overline{\pitchfork} \mathcal{C}^2$, and (ii) there exists a nontrivial linear functional $\lambda \in X^\dagger$ such that $C^1 \cap C^2 \cap \{x \in X : \lambda(x) > 0\} \neq \emptyset$ for every $(C^1, C^2) \in \mathcal{C}^1 \times \mathcal{C}^2$. ■

5) The nonseparation theorem.: The crucial fact about AGDQs that leads to the maximum principle is the transversal intersection property, that we now state (cf. [13] for the proof).

Theorem 2.8: Let M be a manifold of class C^1, let S_1, S_2 be subsets of M, and let $\bar{s}_* \in S_1 \cap S_2$. Let $\mathcal{C}_1$, $\mathcal{C}_2$ be AGDQ-approximating multicones to S_1, S_2 at $\bar{s}_*$ such that $\mathcal{C}_1 \overline{\overline{\pitchfork}} \mathcal{C}_2$. Then S_1 and S_2 are not locally separated at $\bar{s}_*$. (That is, the set $S_1 \cap S_2$ contains a sequence of points s_j converging to $\bar{s}_*$ but not equal to $\bar{s}_*$.) ■

III. Flows and trajectories.

1) State space bundles: If I is a nonempty real interval, we define $I^{2,\geq} = \{(t, s) \in I \times I : t \geq s\}$, and $I^{3,\geq} = \{(t, s, r) \in I \times I \times I : t \geq s \geq r\}$. A ***state-space bundle*** (abbr. SSB) ***over*** I is an indexed family $\mathbf{X} = \{X_t\}_{t \in I}$ of sets. A ***state-space bundle*** is a pair $\mathcal{X} = (\mathbf{X}, I)$ such that I is a nonempty real interval and $\mathbf{X}$ is an SSB over I. If $\mathcal{C}$ is a category of sets with some additional structure (for example, topological spaces, metric spaces, manifolds of class C^k, linear spaces, FDRLSs), then an SSB $(\mathbf{X}, I)$ is a ***bundle of*** $\mathcal{C}$***-objects*** if each X_t is a member of $\mathcal{C}$. In particular, if $k \in \mathbb{Z}_+$, a C^k ***SSB*** is an SSB of manifolds of class C^k. Also, an ***FDRLS SSB*** is an SSB of finite-dimensional real linear spaces.

2) Sections: Assume that

(A1) $\mathcal{X} = (\mathbf{X}, I) = (\{X_t\}_{t\in I}, I)$ is an SSB.

Definition 3.1: A ***section of*** $\mathcal{X}$ is a single-valued everywhere defined map ξ on I such that $\xi(t) \in X_t$ for every $t \in I$. We use $Sec(\mathcal{X})$ to denote the set of all sections of $\mathcal{X}$. ■

3) The tangent bundle of a section: Suppose that (A1) holds, $\mathcal{X}$ is a C^1 SSB, and $\xi \in Sec(\mathcal{X})$. The family $\mathbf{T}_\xi\mathcal{X} = \{T_{\xi(t)}X_t\}_{t\in I}$ is the ***tangent bundle*** of $\mathcal{X}$ along ξ. Clearly, *the tangent bundle* $\mathbf{T}_\xi\mathcal{X}$ *of a* C^1 *SSB* $\mathcal{X}$ *along a section* $\xi \in Sec(\mathcal{X})$ *is an FDRLS SSB.*

4) Flows: Assume that (A1) holds.

Definition 3.2: A ***flow*** on $\mathcal{X}$ is an indexed family $\mathbf{f} = \{f_{t,s}\}_{(t,s)\in I^{2,\geq}}$ such that

1) $f_{t,s}$ is a set-valued map from X_s to X_t whenever $(t,s) \in I^{2,\geq}$;
2) $f_{t,t}$ is the identity map of X_t whenever $t \in I$;
3) $f_{t,s} \circ f_{s,r} = f_{t,r}$ whenever $(t,s,r) \in I^{3,\geq}$.

A ***flow*** is a pair $\mathcal{F} = (\mathcal{X}, \mathbf{f})$ such that $\mathcal{X}$ is a state space bundle and $\mathbf{f}$ is a flow on $\mathcal{X}$. ■

5) Comparison of flows: If, for $i = 1, 2$, $\mathcal{F}^i = (\mathcal{X}, \mathbf{f}^i)$ are flows on the same SSB $\mathcal{X}$, and $\mathbf{f}^i = \{f^i_{t,s}\}_{(t,s)\in I^{2,\geq}}$, we say that $\mathcal{F}^1$ is a ***subflow*** of $\mathcal{F}^2$, or $\mathcal{F}^2$ is a ***superflow*** of $\mathcal{F}^1$, and write $\mathcal{F}^1 \preceq \mathcal{F}^2$, if

$$f^1_{t,s}(x) \subseteq f^2_{t,s}(x) \quad \text{for all} \quad (t,s) \in I^{2,\geq}\,.$$

6) Trajectories: Assume that (A1) holds, and

(A2) $\mathcal{F} = (\mathcal{X}, \mathbf{f})$ *is a flow, and* $\mathbf{f} = \{f_{t,s}\}_{(t,s)\in I^{2,ge}}$.

Definition 3.3: A ***trajectory*** of $\mathcal{F}$ is a section ξ of $\mathcal{X}$ such that

$$\xi(t) \in f_{t,s}(\xi(s)) \quad \text{whenever} \quad (t,s) \in I^{2,\geq}\,. \tag{1}$$

We use $Traj(\mathcal{F})$ to denote the set of all trajectories of the flow $\mathcal{F}$. ■

7) Generalized differentials of flows along trajectories: Assume that (A1-2) hold, an in addition

(A3) $\mathcal{X}$ *is a* C^1 *SSB;*

(A4) $\xi \in Traj(\mathcal{F})$.

Definition 3.4: An ***AGDQ*** of $\mathcal{F}$ along ξ is a family $\mathbf{g} = \{g_{t,s}\}_{(t,s)\in I^{2,\geq}}$ such that

1) if $(t,s) \in I^{2,\geq}$, then $g_{t,s}$ is a nonempty compact set of linear maps from $T_{\xi(s)}X_s$ to $T_{\xi(t)}X_t$ such that $g_{t,s} \in AGDQ(f_{t,s}; \xi(s), \xi(t); X_s)$;
2) $g_{t,t} = \{\mathbb{I}_{T_{\xi(t)}X_t}\}$ whenever $t \in I$;
3) $g_{t,s} \circ g_{s,r} = g_{t,r}$ whenever $(r,s,t) \in I^{3,\geq}$. ■

8) Compatible selections: Assume that (A1-2-3-4), hold, and in addition

(A5) $\mathbf{g} = \{g_{t,s}\}_{(t,s)\in \mathbf{T}^{\geq}}$ *is an AGDQ of* $\mathcal{F}$ *along* ξ.

Definition 3.5: A ***compatible selection*** of $\mathbf{g}$ is a family $\gamma = \{\gamma_{t,s}\}_{(t,s)\in I^{2,\geq}}$ such that

1) $\gamma_{t,s} \in g_{t,s}$ whenever $(t,s) \in I^{2,\geq}$;
2) $\gamma_{t,t} = \mathbb{I}_{T_{\xi(t)}X}$ for each $t \in I$;
3) $\gamma_{t,s}\gamma_{s,r} = \gamma_{t,r}$ whenever $(t,s,r) \in I^{3,\geq}$.

We write $CSel(\mathbf{g})$ to denote the set of all compatible selections of $\mathbf{g}$. ■

IV. VARIATIONS

1) Variations of set-valued maps: If f, f' are SVMs, we write $f \preceq f'$ if $\mathrm{So}(f) = \mathrm{So}(f')$, $\mathrm{Ta}(f) = \mathrm{Ta}(f')$, and $\mathrm{Gr}(f) \subseteq \mathrm{Gr}(f')$.

Definition 4.1: Assume that f, f' are SVMs such that $f \preceq f'$. Then

1) A ***variation of*** f ***in*** f' is a family $v = \{v_\varepsilon\}_{\varepsilon\in C}$ such that
   a) C is a closed convex cone with nonempty interior in some "ambient" $FDRLS$ $\mathcal{A}(C)$;
   b) each v_ε is a SVM such that $\mathrm{So}(v_\varepsilon) = \mathrm{So}(f)$ and $\mathrm{Ta}(v_\varepsilon) = \mathrm{Ta}(f)$;
   c) $v_0 = f$;
   d) $v_\varepsilon \preceq f'$ whenever $\varepsilon \in C$.
2) A ***variation of*** f is a variation of f in the "maximal" SVM $(\mathrm{So}(f), \mathrm{Ta}(f), \mathrm{So}(f) \times \mathrm{Ta}(f))$. ■

If v is a variation of f, and $v = \{v_\varepsilon\}_{\varepsilon\in C}$, then the cone C and the linear space $\mathcal{A}(C)$ are, respectively, the ***parameter cone*** and the ***parameter space*** of v. The dimension of C (or of $\mathcal{A}(C)$) is the ***number of parameters*** of v. We will use $\tilde{v}$ to denote the SVM with source $\mathcal{A}(C) \times \mathrm{So}(v_0)$ and target $\mathrm{Ta}(v_0)$ such that $\tilde{v}(\varepsilon, x) = v_\varepsilon(x)$ for all $\varepsilon \in \mathcal{A}(C)$, $x \in \mathrm{So}(v_0)$. (In particular, $\tilde{v}(\varepsilon, x) = \emptyset$ if $\varepsilon \in \mathcal{A}(C)\backslash C$.)

2) Infinitesimal impulse variations: Assume that (A1-2-3-4) hold.

Definition 4.2: An ***infinitesimal impulse variation*** (abbr, IIV) for $(\mathcal{F}, \xi)$ is a pair (v, t) such that $t \in I$ and $v \in T_{\xi(t)}X_t$. ■

3) Summability: Assume that (A1-2-3-4-5) hold. Assume in addition that

(A6) $\mathcal{F}' = (\mathcal{X}, \mathbf{f}') = (\mathcal{X}, \{f'_{t,s}\}_{(t,s)\in I^{2,\geq}})$ *is a flow on* $\mathcal{X}$ *which is a superflow of* $\mathcal{F}$.

If $\mathbf{V}$ is a finite set of IIVs for $(\mathcal{F}, \xi)$, we let $\mathbb{R}^{\mathbf{V}}$, $\mathbb{R}^{\mathbf{V}}_+$ denote, respectively, the set of all families $\vec{\varepsilon} = \{\varepsilon^V\}_{V\in\mathbf{V}}$ of real numbers, and the set of all $\vec{\varepsilon} = \{\varepsilon^V\}_{V\in\mathbf{V}} \in \mathbb{R}^{\mathbf{V}}$ such that $\varepsilon^V \geq 0$ for all $V \in \mathbf{V}$. (Hence, if m is the cardinality of $\mathbf{V}$, and $\mathbf{V} = \{(v^1, t^1), \ldots, (v^m, t^m)\}$, the space $\mathbb{R}^{\mathbf{V}}$ can be identified with $\mathbb{R}^m$, by identifying each family $\vec{\varepsilon} = \{\varepsilon^V\}_{\varepsilon\in\mathbf{V}}$ with the m-tuple $(\tilde{\varepsilon}^1, \ldots, \tilde{\varepsilon}^m)$, where $\tilde{\varepsilon}^j = \varepsilon^{(v^j,t^j)}$ for $j = 1, \ldots, m$.) We define

$$t_-(\mathbf{V}) \;=\; \min\{t : (\exists v)(v,t) \in \mathbf{V}\}\,,$$
$$t_+(\mathbf{V}) \;=\; \max\{t : (\exists v)(v,t) \in \mathbf{V}\}\,,$$

and let $I^{2,\geq}(\mathbf{V})$ be the set of all pairs $(b,a) \in I^{2,\geq}$ such that $a < t_-(\mathbf{V})$ and $b > t_-(\mathbf{V})$. If $\gamma = \{\gamma_{t,s}\}_{(t,s)\in I^{2,\geq}} \in CSel(\mathbf{g})$ and $(b,a) \in I^{2,\geq}(\mathbf{V})$, we define a linear map $L^{\mathbf{V},\gamma,a,b} : \mathbb{R}^{\mathbf{V}} \times T_{\xi(a)}X_a \mapsto T_{\xi(b)}X_b$ by letting

$$L^{\mathbf{V},\gamma,a,b}(\vec{\varepsilon}) = \gamma_{b,a}(x) + \sum_{(v,t)\in\mathbf{V}} \varepsilon^{(v,t)}\gamma_{b,t}(v)\,.$$

We let $\Lambda^{\mathbf{V},\mathbf{g},a,b}$ be the set of all the maps $L^{\mathbf{V},\gamma,a,b}$, for all $\gamma \in CSel(\mathbf{g})$.

Definition 4.3: Let $\mathcal{V}$ be a set of IIVs for $(\mathcal{F}, \xi)$. We say that $\mathcal{V}$ is $\mathbf{g}$-***summable within*** $\mathcal{F}'$ if the following is true:

- *for every finite subset* $\mathbf{V}$ *of* $\mathcal{V}$, *and every pair* $(b,a) \in I^{2,\geq}(\mathbf{V})$, *there exists a variation* $w = \{w_{\vec{\varepsilon}}\}_{\vec{\varepsilon}\in\mathbb{R}_+^{\mathbf{V}}}$ *of* $f_{b,a}$ *in* $f'_{b,a}$ *such that the set* $\Lambda^{\mathbf{V},\mathbf{g}}$ *is an AGDQ of the map* $\tilde{w}$ *at* $((0,\xi(a)),\xi(b))$ *along* $\mathbb{R}_+^{\mathbf{V}} \times X_a$. ■

V. Generalized Bianchini-Stefani variations

Assume that $\mathcal{F}$, $\mathcal{X}$, $\mathbf{f}$, $\mathbf{X}$, I, ξ, $\mathbf{g}$, $\mathcal{F}'$, $\mathbf{f}'$, are such that (A1-2-3-4-5-6) hold.

From now on, "GBS" will stand for "Generalized Bianchini-Stefani".

Definition 5.1: An ***$\mathcal{F}'$-compatible right GBS IIV for*** $(\mathcal{F},\xi)$ is an IIV (v,t) for $(\mathcal{F},\xi)$ such that there exist

- a nonempty set Σ,
- positive real numbers $\bar{c}$, ν,
- a pair (τ_-,τ_+) of functions from Σ to I,
- a function $\alpha : \Sigma \mapsto]\,0,+\infty\,[$,
- a pair of families

$$\boldsymbol{\varphi} = \{\varphi_{\sigma,c}\}_{(\sigma,c)\in\Sigma\times[0,\bar{c}]}\,, \quad \boldsymbol{\psi} = \{\psi_{\sigma,c}\}_{(\sigma,c)\in\Sigma\times[0,\bar{c}]}\,,$$

- a neighborhood $\mathcal{N}$ of $\xi(t)$ in X_t,

such that

1) $t < \tau_-(\sigma) < \tau_+(\sigma)$ for every $\sigma\in\Sigma$;
2) for every $\sigma\in\Sigma$ and every real number ρ such that $0<\rho\leq 1$, there exists a $\sigma'\in\Sigma$ such that the following three identities are satisfied:
$$\begin{aligned}\tau_-(\sigma')-t &= \rho(\tau_-(\sigma)-t)\,,\\ \tau_+(\sigma')-t &= \rho(\tau_+(\sigma)-t)\,,\\ \alpha(\sigma') &= \rho^\nu\alpha(\sigma)\,;\end{aligned}$$
3) for every $(\sigma,c)\in\Sigma\times[0,\bar{c}]$, $\varphi_{\sigma,c}$ belongs to $SVM(X_{\tau_+(\sigma)},X_{\tau_-(\sigma)})$ and $\psi_{\sigma,c}\in SVM(X_t,X_t)$;
4) $\varphi_{\sigma,0} = f_{\tau_+(\sigma),\tau_-(\sigma)}$ and $\psi_{\sigma,0} = \mathbb{I}_{X_t}$ whenever $\sigma\in\Sigma$;
5) $\varphi_{\sigma,c}(x) \subseteq f'_{\tau_+(\sigma),\tau_-(\sigma)}(x)$ whenever $\sigma\in\Sigma$, $c\in[0,\bar{c}]$, $x\in X_{\tau_-(\sigma)}$,
6) for every $\sigma\in\Sigma$, $c\in[0,\bar{c}]$, and $x\in X_t$, the inclusion
$$(f_{\tau_+(\sigma),t}\circ\psi_{\sigma,c})(x) \subseteq (\varphi_{\sigma,c}\circ f_{\tau_-(\sigma),t})(x) \qquad (2)$$
is satisfied;
7) for every $\sigma\in\Sigma$, the map
$$[0,\bar{c}]\times X_t \ni (c,x) \mapsto\!\!\!\rightarrow \psi_{\sigma,c}(x) \stackrel{\text{def}}{=} \Psi_\sigma(c,x) \subseteq X_t$$
is Cellina continuously approximable on $[0,\bar{c}]\times\mathcal{N}$;
8) the continuity formula
$$\lim_{x\to\xi(t),\alpha(\sigma)\to 0}\psi_{\sigma,c}(x) = \xi(t) \qquad (3)$$
holds in the following precise sense:
   - if N is any neighborhood of $\xi(t)$ in X_t then there exist a neighborhood N' of $\xi(t)$ in X_t and a positive number β such that $\psi_{\sigma,c}(x)\subseteq N$ whenever $x\in N'$, $c\in[0,\bar{c}]$, and $\alpha(\sigma)<\beta$;
9) the asymptotic formula
$$\psi_{\sigma,c}(x) = x + \alpha(\sigma)cv + o\Big(\alpha(\sigma) + \|x-\xi(t)\|\Big) \quad (4)$$
holds as $(\alpha(\sigma), x-\xi(t)) \to (0,0)$, in the following precise sense:
   - if $\kappa:\Omega\mapsto\mathbb{R}^n$ is any coordinate chart of X_t defined on an open subset Ω of X_t such that $\xi(t)\in\Omega$, then
   $$\lim_{\beta\downarrow 0}\sup\left\{\frac{\|\kappa(y)-\kappa(x)-\alpha(\sigma)c\kappa_*v\|}{\alpha(\sigma)+\|\kappa(x)-\kappa(\xi(t))\|}\right\} = 0$$
   where (i) the supremum is taken over all 4-tuples $(\sigma,c,x,y)\in\Sigma\times[0,\bar{c}]\times\Omega\times\Omega$ such that $\|\kappa(x)-\kappa(\xi(t))\|\leq\beta$, $\alpha(\sigma)\leq\beta$, $y\in\psi_{\sigma,c}(x)$, and (ii) $\kappa_*v = D\kappa(\xi(t))\cdot v$. ■

An 9-tuple $\mathcal{D} = (\Sigma,\bar{c},\nu,\tau_-,\tau_+,\alpha,\boldsymbol{\varphi},\boldsymbol{\psi},\mathcal{N})$ for which the conditions of the above definition hold will be called a ***GBS data 9-tuple*** for $(\mathcal{F},\xi,\mathcal{F}',v,t)$. Given a number $\bar{\nu}$, an IIV (v,t) is ***of order*** $\bar{\nu}$ (as an $\mathcal{F}'$-compatible right GBS IIV for $(\mathcal{F},\xi)$) if it admits a GBS data 9-tuple $\mathcal{D} = (\Sigma,\bar{c},\nu,\tau_-,\tau_+,\alpha,\boldsymbol{\varphi},\boldsymbol{\psi},\mathcal{N})$ for $(\mathcal{F},\xi,\mathcal{F}',v,t)$ for which $\nu=\bar{\nu}$.

The following is the main theorem of this paper:

Theorem 5.2: Assume that (A1-2-3-4-5-6) hold. Let $\mathcal{V}$ be the set of all $\mathcal{F}'$-compatible right GBS IIVs for $(\mathcal{F},\xi)$. Then $\mathcal{V}$ is $\mathbf{g}$-summable within $\mathcal{F}'$. ■

Proof: It clearly suffices to prove that every finite subset $\mathbf{V}$ of $\mathcal{V}$ is $\mathbf{g}$-summable within $\mathcal{F}'$. Furthermore, it is well known that there is no problem with the summability of variations at different times, so it suffices to establish the summability of a nonempty finite set $\mathbf{V}$ of $\mathcal{F}'$-compatible right GBS IIVs for $(\mathcal{F},\xi)$ all of which have the same insertion time t.

Fix such a set $\mathbf{V}$, and let m be its cardinality, so $m\geq 1$. Choose for each $(v,t)\in\mathbf{V}$ a GBS data 9-tuple $\mathcal{D}^v = (\Sigma^v,\bar{c}^v,\nu^v,\tau_-^v,\tau_+^v,\alpha^v,\boldsymbol{\varphi}^v,\boldsymbol{\psi}^v,\mathcal{N}^v)$ for $(\mathcal{F},\xi,\mathcal{F}',v,t)$. Order the members of $\mathbf{V}$ as a sequence $\tilde{\mathbf{V}} = ((v^1,t),(v^2,t),\ldots,(v^m,t))$, in such a way that $\nu^{v^1}\leq\nu^{v^2}\leq\cdots\leq\nu^{v^m}$.

From now on, we will write Σ^j, $\bar{c}^j$, $\nu(j)$, τ_-^j, τ_+^j, α^j, $\boldsymbol{\varphi}^j$, $\boldsymbol{\psi}^j$, $\mathcal{N}^j$ instead of Σ^{v^j}, $\bar{c}^{v^j}$, ν^{v^j}, $\tau_-^{v^j}$, $\tau_+^{v^j}$, α^{v^j}, $\boldsymbol{\varphi}^{v^j}$, $\boldsymbol{\psi}^{v^j}$, $\mathcal{N}^{v^j}$. Also, we fix a $b\in I$ such that $b>t$.

Define $\Sigma=\Sigma^1\times\cdots\times\Sigma^m$, and let $\hat{\Sigma}$ be the set of all m-tuples $\vec{\sigma}=(\sigma^1,\ldots,\sigma^m)\in\Sigma$ such that $\tau_+(\sigma^m)\leq b$ and $\tau_+^j(\sigma^j)\leq\tau_-^{j+1}(\sigma^{j+1})$ for $j=1,2,\ldots,m-1$. Let $C=[0,\bar{c}^1]\times\cdots\times[0,\bar{c}^m]$. For each $\vec{c}=(c^1,\ldots,c^m)\in C$, and each $\vec{\sigma}=(\sigma^1,\ldots,\sigma^m)\in\hat{\Sigma}$, write $\alpha(\vec{\sigma})\cdot\vec{c}$ to denote the m-tuple $(\alpha^1(\sigma^1)c^1,\ldots,\alpha^m(\sigma^m)c^m)$.

We let
$$\alpha_+(\vec{\sigma}) = \max\Big\{\alpha^j(\sigma^j) : j\in\{1,\ldots,m\}\Big\}\,,$$
and, if $\beta>0$, define $\hat{\Sigma}(\beta) = \{\vec{\sigma}\in\hat{\Sigma} : \alpha_+(\vec{\sigma})\leq\beta\}$.

Define set-valued maps $\Phi^j_{\vec{\sigma},\vec{c}} : X_t \mapsto\!\!\!\rightarrow X_{\tau_+(\sigma^j)}$ inductively, by letting $\Phi^1_{\vec{\sigma},\vec{c}} = \varphi^1_{\sigma^1,c^1}\circ f_{\tau_-^1(\sigma^1),t}$ and, for $j=2,\ldots,m$,
$$\Phi^j_{\vec{\sigma},\vec{c}} = \varphi^j_{\sigma^j,c^j}\circ f_{\tau_-^j(\sigma^j),\tau_+^{j-1}(\sigma^{j-1})}\circ\Phi^{j-1}_{\vec{\sigma},\vec{c}}\,.$$
Then let $\Phi_{\vec{\sigma},\vec{c}} = f_{b,\tau_+^m(\sigma^m)}\circ\Phi^m_{\vec{\sigma},\vec{c}}$, so $\Phi_{\vec{\sigma},\vec{c}} : X_t \mapsto\!\!\!\rightarrow X_b$.

It is then clear that $\Phi_{\vec\sigma,\vec c}(x) \subseteq f'_{b,t}(x)$ for all triples $(\sigma,c,x) \in \hat\Sigma \times C \times X_t$.

Introduce nonnegative real parameters ε^j, for $j=1,\ldots,m$, and define, for each $\vec\varepsilon = (\varepsilon^1,\ldots,\varepsilon^m) \in \mathbb{R}^m_+$,

$$\Phi^{\#}_{\vec\varepsilon}(x) = \bigcup \left\{ \Phi_{\vec\sigma,\vec c}(x) : (\vec\sigma,\vec c) \in \hat\Sigma \times C,\ \alpha(\vec\sigma)\cdot\vec c = \vec\varepsilon \right\}.$$

Then $\mathbf{\Phi}^{\#} = \{\Phi^{\#}_{\vec\varepsilon}\}_{\vec\varepsilon\in\mathbb{R}^m_+}$ is a variation of $f_{b,t}$ in $f'_{b,t}$.

Now, for each $\gamma \in g_{b,t}$, we let L^γ be the linear map from $\mathbb{R}^m \times T_{\xi(t)}X_t$ to $T_{\xi(b)}X_b$ given by

$$L^\gamma(\varepsilon^1,\ldots,\varepsilon^m,w) = \gamma(w + \varepsilon^1 v^1 + \cdots + \varepsilon^m v^m)$$

for $(\varepsilon^1,\ldots,\varepsilon^m) \in \mathbb{R}^m$, $w \in T_{\xi(t)}X_t$.

Let Λ be the set of all maps L^γ, for all $\gamma \in g_{b,t}$. We will show that

$$\Lambda \in AGDQ(\tilde{\mathbf{\Phi}}^{\#};(0,\xi(t)),\xi(b);\mathbb{R}^m_+ \times X_t)\,. \qquad (5)$$

(Recall that $\tilde{\mathbf{\Phi}}^{\#}$ is the set-valued map $\mathbb{R}^m \times X_t \ni (\vec\varepsilon,x) \mapsto \mathbf{\Phi}^{\#}_{\vec\varepsilon}(x) \subseteq X_b$.)

To prove (5), we first define

$$\Psi_{\vec\sigma,\vec c} = \psi^m_{\sigma^m,c^m} \circ \cdots \circ \psi^2_{\sigma^2,c^2} \circ \psi^1_{\sigma^1,c^1}\,,$$

and let $\Upsilon_{\vec\sigma,\vec c} = f_{b,t} \circ \Psi_{\vec\sigma,\vec c}$. We then observe that

$$\Upsilon_{\vec\sigma,\vec c}(x) \subseteq \Phi_{\vec\sigma,\vec c}(x) \qquad (6)$$

whenever $\vec\sigma \in \hat\Sigma$, $\vec c \in C$, and $x \in X$. (The proof of this is straightforward: we have

$$(f_{\tau^1_+(\sigma^1),t} \circ \psi^1_{\sigma^1,c^1})(x) \subseteq \varphi^1_{\sigma^1,c^1} \circ f_{\tau^1_-(\sigma^1),t}(x) = \Phi^1_{\vec\sigma,\vec c}(x)$$

for all x. If we compose both sides with $f_{\tau^2_-(\sigma^2),\tau^1_+(\sigma^1)}$, we find

$$(f_{\tau^2_-(\sigma^2),t} \circ \psi^1_{\sigma^1,c^1})(x) \subseteq (f_{\tau^2_-(\sigma^2),\tau^1_+(\sigma^1)} \circ \Phi^1_{\vec\sigma,\vec c})(x)$$

and then

$$\begin{aligned}&(\varphi^2_{\sigma^2,c^2} \circ f_{\tau^2_-(\sigma^2),t} \circ \psi^1_{\sigma^1,c^1})(x)\\ &\qquad\subseteq (\varphi^2_{\sigma^2,c^2} \circ f_{\tau^2_-(\sigma^2),\tau^1_+(\sigma^1)} \circ \Phi^1_{\vec\sigma,\vec c})(x)\,,\end{aligned}$$

so that

$$(\varphi^2_{\sigma^2,c^2} \circ f_{\tau^2_-(\sigma^2),t} \circ \psi^1_{\sigma^1,c^1})(x) \subseteq \Phi^2_{\vec\sigma,\vec c}(x)\,. \qquad (7)$$

Next, we have

$$(f_{\tau^2_+(\sigma^2),t} \circ \psi^2_{\sigma^2,c^2})(y) \subseteq (\varphi^2_{\sigma^2,c^2} \circ f_{\tau^2_-(\sigma^2),t})(y)$$

for all y. By applying this to the members y of $\psi^1_{\sigma^1,c^1}(x)$, we find

$$\begin{aligned}&(f_{\tau^2_+(\sigma^2),t} \circ \psi^2_{\sigma^2,c^2} \circ \psi^1_{\sigma^1,c^1})(x)\\ &\qquad\subseteq (\varphi^2_{\sigma^2,c^2} \circ f_{\tau^2_-(\sigma^2),t} \circ \psi^1_{\sigma^1,c^1})(x)\,,\end{aligned}$$

and this, together with (7), implies that

$$(f_{\tau^2_+(\sigma^2),t} \circ \psi^2_{\sigma^2,c^2} \circ \psi^1_{\sigma^1,c^1})(x) \subseteq \Phi^2_{\vec\sigma,\vec c}(x)\,.$$

Continuing in this way, we show that

$$(f_{\tau^j_+(\sigma^j),t} \circ \psi^j_{\sigma^j,c^j} \circ \cdots \circ \psi^1_{\sigma^1,c^1})(x) \subseteq \Phi^j_{\vec\sigma,\vec c}(x)$$

for $j=1,\ldots,m$. In particular,

$$(f_{\tau^m_+(\sigma^m),t} \circ \Psi_{\vec\sigma,\vec c})(x) \subseteq \Phi^j_{\vec\sigma,\vec c}(x)$$

If we then compose both sides with $f_{b,\tau^m_+(\sigma^m)}$, we conclude that (6) holds.)

We now define $\Psi^{\#}$ and $\Upsilon^{\#}$ exactly as we defined $\Phi^{\#}$, but using the $\Psi_{\vec\sigma,\vec c}$ and $\Upsilon_{\vec\sigma,\vec c}$ instead of the $\Phi_{\vec\sigma,\vec c}$. That is, we define

$$\begin{aligned}\Psi^{\#}_{\vec\varepsilon}(x) &= \bigcup \left\{ \Psi_{\vec\sigma,\vec c}(x) : (\vec\sigma,\vec c) \in \hat\Sigma \times C,\ \alpha(\vec\sigma)\cdot\vec c = \vec\varepsilon \right\},\\ \Upsilon^{\#}_{\vec\varepsilon}(x) &= \bigcup \left\{ \Upsilon_{\vec\sigma,\vec c}(x) : (\vec\sigma,\vec c) \in \hat\Sigma \times C,\ \alpha(\vec\sigma)\cdot\vec c = \vec\varepsilon \right\},\end{aligned}$$

for each $\vec\varepsilon = (\varepsilon^1,\ldots,\varepsilon^m) \in \mathbb{R}^m_+$. Then the family $\mathbf{\Upsilon}^{\#} = \{\Upsilon^{\#}_{\vec\varepsilon}\}_{\vec\varepsilon\in\mathbb{R}^m_+}$ is a variation of $f_{b,t}$ in $f'_{b,t}$, such that

$$\Upsilon^{\#}_{\vec\varepsilon}(x) \subseteq \Phi^{\#}_{\vec\varepsilon}(x) \quad \text{whenever} \quad \vec\varepsilon \in \mathbb{R}^m_+\,,\ x \in X_t\,. \qquad (8)$$

Clearly, (5) is a consequence of

$$\Lambda \in AGDQ(\tilde{\mathbf{\Upsilon}}^{\#};(0,\xi(t)),\xi(b);\mathbb{R}^m_+ \times X_t)\,, \qquad (9)$$

and this in turn will follow—thanks to the chain rule—if we prove that

$$\{L\} \in AGDQ(\tilde{\mathbf{\Psi}}^{\#};(0,\xi(t)),\xi(t);\mathbb{R}^m_+ \times X_t)\,. \qquad (10)$$

where L is the linear map from $\mathbb{R}^m \times T_{\xi(t)}X_t$ to $T_{\xi(t)}X_t$ given by

$$L(\varepsilon^1,\ldots,\varepsilon^m,w) = w + \varepsilon^1 v^1 + \cdots + \varepsilon^m v^m \qquad (11)$$

for $(\varepsilon^1,\ldots,\varepsilon^m) \in \mathbb{R}^m$, $w \in T_{\xi(t)}X_t$.

To prove (10), we construct an AGDQ modulus for $(\{L\},\tilde{\mathbf{\Psi}}^{\#};(0,\xi(t)),\xi(t);\mathbb{R}^m_+ \times X_t)$.

Our first step will be to introduce coordinates on X_t near $\xi(t)$. We let $n = \dim X_t$, and write $\mathbb{B}^n = \{y \in \mathbb{R}^n : \|y\| < 1\}$, $\bar{\mathbb{B}}^n = \{y \in \mathbb{R}^n : \|y\| \le 1\}$, $r\mathbb{B}^n = \{y \in \mathbb{R}^n : \|y\| < r\}$, $r\bar{\mathbb{B}}^n = \{y \in \mathbb{R}^n : \|y\| \le r\}$. We fix a coordinate chart $\kappa : \Omega \mapsto \mathbb{R}^n$ of class C^1, mapping an open subset Ω of X_t onto $\mathbb{B}^n$, and such that $\kappa(\xi(t)) = 0$. We then choose $\bar r$ such that $0 < \bar r < 1$, having the property that $\kappa^{-1}(\bar r\bar{\mathbb{B}}^n) \subseteq \bigcap_{j=1}^m \mathcal{N}^j$, so all the maps $[0,\bar c] \times \kappa^{-1}(\bar r\bar{\mathbb{B}}^n) \ni (c,x) \mapsto\!\!\!\to \psi_{\sigma,c}(x) \subseteq X_t$ are CCA. We then use the continuity property of the ψ^j to find radii $r_0,r_1,\ldots,r_m$ such that $0 < r_0 < r_1 < r_2 < \cdots < r_m \le \bar r$, and a positive number β_*, having the property that,

(*) *whenever* $j=1,\ldots,m$, $c \in [0,\bar c^j]$, $\sigma \in \Sigma^j$, *and* $\alpha^j(\sigma) \le \beta_*$, *the inclusion* $\psi^j_{\sigma,c}(x) \subseteq \kappa^{-1}(r_j\mathbb{B}_n)$ *holds for all* $x \in \kappa^{-1}(r_{j-1}\bar{\mathbb{B}}_n)$.

We then identify Ω with $\mathbb{B}^n$ by means of κ, and restrict the maps $x \mapsto\!\!\!\to \psi^j_{\sigma,c}(x)$, for $\alpha^j(\sigma) \le \beta_*$, to the ball $r_{j-1}\bar{\mathbb{B}}_n$. Then $\xi(t) = 0$, and

$$\psi^j_{\sigma,c}(x) \subseteq r_j\mathbb{B}_n \quad \text{if} \quad x \in r_{j-1}\bar{\mathbb{B}}_n\,,\ \alpha^j(\sigma) \le \beta_*\,.$$

Furthermore, since $r_{j-1}\bar{\mathbb{B}}_n \subseteq \bar r\bar{\mathbb{B}}_n \subseteq \mathcal{N}^j$, it is clear that each map $[0,\bar c^j] \times r_{j-1}\bar{\mathbb{B}}_n \ni (c,x) \mapsto\!\!\!\to \psi^j_{\sigma,c}(x) \subseteq r_j\mathbb{B}_n$ is CCA.

Next, we pick in an arbitrary fashion an m-tuple $\vec{\sigma}_* = (\sigma_*^1, \ldots, \sigma_*^m) \in \Sigma$. (Clearly, σ_* exists because the Σ^j are nonempty.) Then, even though it could happen that one or both inequalities $\tau_+^m(\sigma_*^m) > b$, $\alpha^m(\sigma_*^m) > \beta_*$ hold, we can always pick a number ρ such that $0 < \rho \leq 1$, $\rho(\tau_+^m(\sigma_*^m) - t) \leq b - t$, and $\rho^{\nu(m)}\alpha^m(\sigma_*^m) \leq \beta_*$. We can then find a $\tilde{\sigma}_*^m$ such that $\tau_+^m(\tilde{\sigma}_*^m) - t = \rho(\tau_+^m(\sigma_*^m) - t)$ and $\alpha^m(\tilde{\sigma}_*^m) = \rho^{\nu(m)}\alpha^m(\sigma_*^m)$. If we then relabel $\tilde{\sigma}_*^m$ to be σ_*^m, we now have both inequalities $\tau_+^m(\sigma_*^m) \leq b$ and $\alpha^m(\sigma_*^m) \leq \beta_*$.

Similarly, it could happen that $\tau_+^{m-1}(\sigma_*^{m-1}) > \tau_-^m(\sigma_*^m)$ and $\alpha^{m-1}(\sigma_*^{m-1}) > \beta_*$, but we can pick a ρ such that $0 < \rho \leq 1$, $\rho(\tau_+^{m-1}(\sigma_*^{m-1}) - t) \leq \tau_-^m(\sigma_*^m) - t$, and $\rho^{\nu(m-1)}\alpha^{m-1}(\sigma_*^{m-1}) < \beta_*$. We can then find a $\tilde{\sigma}_*^{m-1}$ such that $\tau_+^{m-1}(\tilde{\sigma}_*^{m-1}) - t = \rho(\tau_+^{m-1}(\sigma_*^{m-1}) - t)$ and $\alpha^{m-1}(\tilde{\sigma}_*^{m-1}) = \rho^{\nu(m-1)}\alpha^{m-1}(\sigma_*^{m-1})$. Then, if we take $\tilde{\sigma}_*^{m-1}$ to be our new σ_*^{m-1}, we get the inequalities $\tau_+^{m-1}(\sigma_*^{m-1}) \leq \tau_-^m(\sigma_*^m)$, and $\alpha^{m-1}(\sigma_*^{m-1}) \leq \beta_*$.

Continuing in this way we obtain, after appropriate relabelings, a $\vec{\sigma}_* = (\sigma_*^1, \ldots, \sigma_*^m) \in \Sigma$ such that

$$\begin{aligned} \tau_+^m(\sigma_*^m) &\leq b\,, && (12) \\ \tau_+^{j-1}(\sigma_*^{j-1}) &\leq \tau_-^j(\sigma_*^j) \quad \text{for} \quad j \in \{2, \ldots, m\}, && (13) \\ \alpha^j(\sigma_*^j) &\leq \beta_* \qquad \text{for} \quad j \in \{1, \ldots, m\}. && (14) \end{aligned}$$

It follows, in particular, that $\vec{\sigma}_*$ belongs to $\hat{\Sigma}$.

Having selected $\vec{\sigma}_*$, we now modify β_* by taking $\beta_* = \alpha_+(\vec{\sigma}_*)$. It is clear that this new choice makes β_* smaller, so it does not interfere with the condition (*) that was used to make our first choice of β_*.

Now, for $j = 1, \ldots, m$, $0 < \beta \leq \beta_*$, $0 < r \leq r_{j-1}$, we define

$$\omega^j(\beta, r) = \sup\left\{ \frac{\|y - x - \alpha^j(\sigma)cv^j\|}{\alpha^j(\sigma) + \|x\|} : \|x\| \leq r_{j-1}\,, \right.$$
$$\left. y \in \psi^j_{\sigma,c}(x)\,,\ \sigma \in \Sigma^j,\ \alpha^j(\sigma) \leq \beta,\ c \in [0, \bar{c}^j] \right\}.$$

Then $\lim_{\beta\downarrow 0, r\downarrow 0} \omega^j(\beta, r) = 0$ for each j.

We now define maps $\zeta_{\vec{\sigma}} : C \times r_0\bar{\mathbb{B}}^n \mapsto\!\!\!\!\rightarrow r_m\mathbb{B}^n$, for $\vec{\sigma} \in \hat{\Sigma}(\beta_*)$, by letting $\zeta_{\vec{\sigma}}(\vec{c}, x) = \Psi_{\vec{\sigma},\vec{c}}(x)$ for $\vec{c} \in C$, $x \in r_0\bar{\mathbb{B}}^n$. Then

$$\zeta_{\vec{\sigma}} \in CCA(C \times r_0\bar{\mathbb{B}}^n, r_m\bar{\mathbb{B}}^n) \ \text{ if } \ \sigma \in \hat{\Sigma}(\beta_*)$$

because, if $\vec{c} = (c^1, \ldots, c^m)$, then

$$\zeta_{\vec{\sigma}}(c^1, \ldots, c^m, x) = (\psi^m_{\sigma^m,c^m} \circ \cdots \circ \psi^1_{\sigma^1,c^1})(x)\,,$$

and each map $[0, \bar{c}^j] \times \bar{r}^{j-1}\bar{\mathbb{B}}^n \ni (c, x) \mapsto \psi^j_{\sigma^j,c}(x) \subseteq r^j\mathbb{B}^n$ is CCA.

We then define, for $\vec{\sigma} \in \hat{\Sigma}(\beta_*)$, $\vec{c} \in C$, and $x \in r_0\bar{\mathbb{B}}^n$,

$$h_{\vec{\sigma}}(\vec{c}, x) = \zeta_{\vec{\sigma}}(\vec{c}, x) - x - \sum_{j=1}^m \alpha^j(\sigma^j)c^jv^j\,.$$

(The precise meaning of the above inequality is that $h_{\vec{\sigma}}(\vec{c}, x) = \{y - \sum_{j=1}^m \alpha^j(\sigma^j)c^jv^j : y \in \zeta_{\vec{\sigma}}(\vec{c}, x)\}$.) Then

$$h_{\vec{\sigma}} \in CCA(C \times r_0\bar{\mathbb{B}}^n; \mathbb{R}^n) \ \text{ if } \ \sigma \in \hat{\Sigma}(\beta_*)\,.$$

We now get an estimate for $h_{\vec{\sigma}}$. Suppose that $\sigma \in \hat{\Sigma}(\beta_*)$, $x \in r_0\bar{\mathbb{B}}^n$, $\vec{c} \in C$, and $z \in h_{\vec{\sigma}}(\vec{c}, x)$. Write $\vec{\sigma} = (\sigma^1, \ldots, \sigma^m)$, $\vec{c} = (c^1, \ldots, c^m)$. Then there exist $y_0, y_1, \ldots, y_m$ such that $y_0 = x$, $y_j \in \psi_{\sigma^j,c^j}(y_{j-1})$ for $j = 1, \ldots, m$, and $z = y_m - y_0 - \sum_{j=1}^m \alpha^j(\sigma^j)c^jv^j$. Since $y_{j-1} \in r_{j-1}\bar{\mathbb{B}}^n$, $y_j \in \psi_{\sigma^j,c^j}(y_{j-1})$, and $\alpha^j(\sigma^j) \leq \beta_*$, we can conclude that

$$\frac{\|y_j - y_{j-1} - \alpha^j(\sigma^j)c^jv^j\|}{\alpha^j(\sigma^j) + \|y_{j-1}\|} \leq \omega^j(\alpha_+(\vec{\sigma}), \|y_{j-1}\|)\,,$$

so that

$$\|y_j - y_{j-1} - \alpha^j(\sigma^j)c^jv^j\| \leq \Big(\alpha^j(\sigma^j) + \|y_{j-1}\|\Big)\omega^j(\alpha_+(\vec{\sigma}), \|y_{j-1}\|)\,.$$

It follows, in particular, that

$$\begin{aligned} \|y_j\| \ \leq\ & \|y_{j-1}\| + \lambda\alpha_+(\vec{\sigma}) \\ & + (\alpha_+(\vec{\sigma}) + \|y_{j-1}\|)\omega^j(\alpha_+(\vec{\sigma}), \|y_{j-1}\|)\,. \end{aligned}$$

where $\lambda = \max\{\bar{c}^j\|v^j\| : j = 1, \ldots, m\}$. Therefore

$$\|y_j\| \leq \Big(\|y_{j-1}\| + \alpha_+(\vec{\sigma})\Big)\Big(\hat{\lambda} + \omega^j(\alpha_+(\vec{\sigma}), \|y_{j-1}\|)\Big)\,,$$

where $\hat{\lambda} = \max(1, \lambda)$.

Given a positive number δ, let $\nu(\delta)$ be such that $0 < \nu(\delta) \leq \beta_*$, $\nu(\delta) < r_0$, and $\omega^j\Big(\nu(\delta)\,,\ (m{+}1)(\hat{\lambda}{+}\nu(\delta))^m\nu(\delta)\Big) \leq \min(1, \delta)$ whenever $j = 1, \ldots, m$. Suppose that $\sigma \in \hat{\Sigma}(\nu(\delta))$ and $\|x\| \leq \nu(\delta)$. Then

$$\|y_1\| \leq (\|x\| + \alpha_+(\vec{\sigma}))\Big(\hat{\lambda} + \omega^1(\nu(\delta), \nu(\delta))\Big) \leq 2\nu(\delta)(\hat{\lambda}+1)\,.$$

Since $\|y_1\| \leq 2\nu(\delta)(\hat{\lambda} + 1)$, we can conclude that $\omega^2(\alpha_+(\vec{\sigma}), \|y_1\|) \leq \omega^2(\nu(\delta), 2\nu(\delta)(\hat{\lambda} + 1)) \leq 1$, so

$$\|y_2\| \leq \Big(\|y_1\| + \alpha_+(\vec{\sigma})\Big)(\hat{\lambda}+1) \leq 2\nu(\delta)(\hat{\lambda}+1)^2 + \nu(\delta)(\hat{\lambda}+1)\,,$$

and then $\|y_2\| \leq 3\nu(\delta)(\hat{\lambda}+1)^2$. Continuing in this way, we show that $\|y_j\| \leq (j{+}1)\nu(\delta)(\hat{\lambda}{+}1)^j$ for all $j \in \{1, \ldots, m\}$. It then follows that

$$\omega^j\Big(\alpha_+(\vec{\sigma}), \|y_{j-1}\|\Big) \leq \omega^j\Big(\nu(\delta), (m{+}1)(\hat{\lambda}{+}\nu(\delta))^m\nu(\delta)\Big) \leq \delta$$

for all j. Hence

$$\begin{aligned} \|y_j - y_{j-1} - \alpha^j(\sigma^j)c^jv^j\| \ &\leq\ \delta\Big(\alpha^j(\sigma^j) + \|y_{j-1}\|\Big) \\ &\leq \delta\Big(\nu(\delta) + (m+1)(\hat{\lambda}+1)^m\Big)\nu(\delta)\,. \end{aligned}$$

The identity $z = y_m - y_0 - \sum_{j=1}^m \alpha^j(\sigma^j)c^jv^j$ implies that $z = \sum_{j=1}^m (y_j - y_{j-1} - \alpha^j(\sigma^j)c^jv^j)$. Since $\nu(\delta) \leq \beta_*$, it follows that

$$\|z\| \leq K\delta\,\nu(\delta)\,,$$

where $K = m\Big(\beta_* + (m+1)(\hat{\lambda}+1)^m\Big)$.

Summarizing, we have proved that

(#) *If $\delta > 0$, and $\vec{\sigma}$ belongs to $\hat{\Sigma}(\nu(\delta))$, then the map $h_{\vec{\sigma}}$ is in $CCA(C \times r_0\bar{\mathbb{B}}^n, \mathbb{R}^n)$, and the bound $\|z\| \leq K\delta\,\nu(\delta)$ holds whenever $\|x\| \leq \nu(\delta)$ $\vec{c} \in C$, and $z \in h_{\vec{\sigma}}(\vec{c}, x)$.*

Using (#), we will now conclude our proof. Recall that we have chosen a member $\vec{\sigma}_* = (\sigma^1_*, \ldots, \sigma^m_*)$ of $\hat{\Sigma}$ such that $\alpha_+(\vec{\sigma}_*) = \beta_*$, and the identities (12), (13), (14) hold.

Now, if ρ is any real number such that $0 < \rho \le 1$, we can pick for each j a $\sigma^j_\rho \in \Sigma^j$ such that

$$\begin{aligned}\tau^j_-(\sigma^j_\rho) - t &= \rho^{1/\nu(j)}(\tau^j_-(\sigma^j_*) - t)\,,\\ \tau^j_+(\sigma^j_\rho) - t &= \rho^{1/\nu(j)}(\tau^j_+(\sigma^j_*) - t)\,,\\ \alpha^j(\sigma^j_\rho) &= \rho\alpha^j(\sigma^j_*)\,.\end{aligned}$$

Let $\vec{\sigma}_\rho = (\sigma^1_\rho, \ldots, \sigma^m_\rho)$. It is then clear that $\alpha_+(\vec{\sigma}_\rho) = \rho\alpha_+(\vec{\sigma}_*) = \rho\beta_*$. Furthermore, $\vec{\sigma}_\rho$ belongs to $\hat{\Sigma}$. (Indeed, if $j = 2, \ldots, m$, we have

$$\begin{aligned}\tau^{j-1}_+(\sigma^{j-1}_\rho) - t &= \rho^{1/\nu(j-1)}(\tau^{j-1}_+(\sigma^{j-1}_*) - t)\\ &\le \rho^{1/\nu(j)}(\tau^{j-1}_+(\sigma^{j-1}_*) - t)\\ &\le \rho^{1/\nu(j)}(\tau^j_-(\sigma^j_*) - t)\\ &\le \tau^j_-(\sigma^j_\rho) - t\,,\end{aligned}$$

since $\nu(j-1) \le \nu(j)$ and $0 < \rho \le 1$. Hence $\tau^{j-1}_+(\sigma^{j-1}_\rho) \le \tau^j_-(\sigma^j_\rho)$ for all j. Also,

$$\begin{aligned}\tau^m_+(\sigma^m_\rho) - t &= \rho^{1/\nu(m)}(\tau^m_+(\sigma^m_*) - t)\\ &\le \tau^m_+(\sigma^m_*) - t)\\ &\le b - t\,.\end{aligned}$$

Hence $\tau^m_+(\sigma^m_\rho) \le b$.)

Given a positive δ, let $\rho(\delta) = \nu(\delta)/\beta_*$, so $0 < \rho(\delta) \le 1$, because $0 < \nu(\delta) \le \beta_*$. Let $Q(\delta)$ be the set of all $(\vec{\varepsilon}, x) = (\varepsilon_1, \ldots, \varepsilon_m, x) \in \mathbb{R}^m \times \mathbb{R}^n$ such that $\|x\| \le \nu(\delta)$ and $0 \le \varepsilon_j \le \alpha^j(\sigma^j_{\rho(\delta)})\bar{c}^j$ for $j = 1, \ldots, m$.

Then, if $(\vec{\varepsilon}, x) \in Q(\delta)$, it follows, if we let $\vec{c}(\vec{\varepsilon}) = (\varepsilon_1/\alpha^1(\sigma^1_{\rho(\delta)}), \ldots, \varepsilon_m/\alpha^m(\sigma^m_{\rho(\delta)}))$, that $\vec{c}(\vec{\varepsilon}) \in C$. Therefore the map $H_\delta : Q(\delta) \mapsto\!\!\!\!\to \mathbb{R}^n$ given by $H_\delta(\vec{\varepsilon}, x) = h_{\vec{\sigma}_\rho}(\vec{c}(\vec{\varepsilon}), x)$ belongs to $CCA(Q(\delta), \mathbb{R}^n$. Furthermore, this map satisfies the bound $\|z\| \le K\delta\nu(\delta)$ whenever $z \in H_\delta(\vec{\varepsilon}, x)$, $(\vec{\varepsilon}, x) \in Q(\delta)$.

Let $\varepsilon(\delta)$ be the minimum of $\nu(\delta)$ and the numbers $\alpha^j(\sigma^j_{\rho(\delta)})\bar{c}^j$, for $j = 1 \ldots, m$. Let $\mathcal{B}(\delta)$ be the closed Euclidean ball in $\mathbb{R}^m \times \mathbb{R}^n$ having radius $\varepsilon(\delta)$ and center 0. Then $\mathcal{B}(\delta)$ is the set of all $m{+}n$-tuples $(\varepsilon_1, \ldots, \varepsilon_m, x_1, \ldots, x_n) \in \mathbb{R}^m \times \mathbb{R}^n$ that satisfy $\varepsilon_1^2 + \cdots + \varepsilon_m^2 + x_1^2 + \cdots + x_n^2 \le \varepsilon(\delta)$. Let $\mathcal{B}_+(\delta) = \mathcal{B}(\delta) \cap (\mathbb{R}^m_+ \times \mathbb{R}^n)$. Then $\mathcal{B}_+(\delta)$ is clearly a subset of $Q(\delta)$, so H_δ, restricted to $\mathcal{B}_+(\delta)$ is a CCA map from $\mathcal{B}_+(\delta)$ to $\mathbb{R}^n$.

For $(\vec{\varepsilon}, x) \in \mathcal{B}_+(\delta)$, define a set $A^\delta(\vec{\varepsilon}, x)$ of affine maps from $\mathbb{R}^m \times \mathbb{R}^n$ by letting $A^\delta(\vec{\varepsilon}, x)$ be the set of all maps of the form $(\vec{\gamma}, w) \to L(\vec{\gamma}, w) + u$, for $u \in H_\delta(\vec{\varepsilon}, x)$. (Recall that the map L was defined in (11).) Then A^δ is a CCA map from $\mathcal{B}_+(\delta)$ to $Aff(\mathbb{R}^m \times \mathbb{R}^n; \mathbb{R}^n)$.

Given an $(\vec{\varepsilon}, x) \in \mathcal{B}_+(\delta)$, and any map $M \in A^\delta(\vec{\varepsilon}, x)$, we have, if $u \in H_\delta(\vec{\varepsilon}, x)$ is such that $M(\vec{\gamma}, w) \equiv L(\vec{\gamma}, w) + u$, and $\vec{c}(\vec{\varepsilon}) = (c^1, \ldots, c^m)$

$$M(\vec{\varepsilon}, x) = L(\vec{\varepsilon}, x) + u = x + \sum_{j=1}^m \varepsilon_j v^j = x + \sum_{j=1}^m \alpha^j(\sigma^j_\rho)c^j v^j\,.$$

Since $u \in H_\delta(\vec{\varepsilon}, x) = h_{\vec{\sigma}_\rho}(\vec{c}(\vec{\varepsilon}), x)$, it follows that $M(\vec{\varepsilon}, x) \in \zeta_{\vec{\sigma}_\rho}(\vec{c}, x) = \Psi_{\vec{\sigma}_\rho, \vec{c}}(x)$. Since $\alpha(\vec{\sigma}_\rho) \cdot \vec{c} = \vec{\varepsilon}$, we conclude that $M(\vec{\varepsilon}, x) \in \tilde{\Psi}^\#(\vec{\varepsilon}, x)$.

Therefore the map A^δ belongs to $CCA(\mathcal{B}_+(\delta); Aff(\mathbb{R}^m \times \mathbb{R}^n; \mathbb{R}^n)$ and is such that $L(\vec{\varepsilon}, x) + u \in \Psi^\#(\vec{\varepsilon}, x)$ whenever $(L, u) \in A^\delta(\vec{\varepsilon}, x)$.

This almost proves (9). To complete the proof, we must derive, for the u's such that $(L, u) \in A^\delta(\vec{\varepsilon}, x)$, a bound $\|u\| \le \theta(\varepsilon)\varepsilon$, where ε is the radius of $\mathcal{B}(\delta)$, i. e., $\varepsilon = \varepsilon(\delta)$. We already have the bound $\|u\| \le K\delta\nu(\delta)$, so it suffices to show that $\nu(\delta) \le q\varepsilon(\delta)$ for some constant q. But

$$\varepsilon(\delta) = \min\Big(\nu(\delta), \min\{\alpha^j(\sigma^j_{\rho(\delta)})\bar{c}^j : j = 1 \ldots, m\}\Big)\,,$$

and $\alpha^j(\sigma^j_{\rho(\delta)}) = \rho(\delta)\alpha^j(\sigma^j_*) = \nu(\delta)/\beta_*$. Hence $\alpha^j(\sigma^j_{\rho(\delta)})\bar{c}^j \ge \frac{\nu(\delta)\hat{c}}{\beta_*}$, where $\hat{c} = \min\{\bar{c}^j : j = 1 \ldots, m\}$. Therefore, if we let $q = \max(1, \beta_*/\hat{c})$, we find that $\nu(\delta) \le q\varepsilon(\delta)$, concluding our proof. ∎

References

[1] Bianchini, R. M., and G. Stefani, "Graded approximations and controllability along a trajectory." *SIAM J. Control and Optim.* **28**, 1990, pp. 903–924.

[2] Bianchini, R. M., and G. Stefani, "Controllability along a trajectory: a variational approach." *SIAM J. Control and Optim.* **31**, 1993, pp. 900–927.

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

[4] G. Stefani, "On the local controllability of a scalar-input system." In *Theory and Applications of Nonlinear Control Systems*, North Holland, 1986, pp. 167–179.

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

[6] 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.

[7] 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.

[8] Sussmann, H.J., "Résultats récents sur les courbes optimales." In *15 e Journée Annuelle de la Société Mathématique de France (SMF)*, Publications de la SMF, Paris, 2000, pp. 1-52.

[9] 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.

[10] 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.

[11] 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.*

[12] 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.

[13] 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.