This paper will be presented at the 35th IEEE Conference on Decision and Control, to be held in Kobe, Japan, December 1996, and will appear in the proceedings of the conference.

A strong maximum principle for systems of differential inclusions*

Héctor J. Sussmann
Department of Mathematics, Rutgers University, New Brunswick, NJ 08903
E-mail: sussmann@hamilton.rutgers.edu

Abstract

We present a version of the finite-dimensional Maximum Principle for systems of differential inclusions, that contains as particular cases previous versions for control systems as well as for a single differential inclusion. The result incorporates high-order point variations and is valid under minimal technical assumptions, weaker than those of most classical and nonsmooth versions. The proof follows the classical idea of using needle variations and applying an appropriate open mapping theorem to a multiparameter variation, whose effect is computed in terms of those of the needle variations by means of the chain rule. However, this has to be carried out in a new setting, namely, the class of "semidifferentiable maps," that contains all maps arising in the optimal control problem and has a concept of generalized differential with all the right properties. We depart from the most common approaches to the problem by reducing the differential inclusions case to the vector field system case, in spite of the fact that set-valued functions —even if they are continuous or Lipschitz or smooth— in general fail to have continuous selections. The key technical tool enabling us to carry out the reduction anyhow is a generalization of a selection theorem due to A. Bressan, which yields the existence, for almost lower semicontinuous set-valued maps, of selections that, while not necessarily continuous, are time-varying vector fields with good properties, such as local existence of trajectories and upper semicontinuity of the (set-valued) flow.

1. Introduction

The Maximum Principle (MP) of optimal control theory deals with a trajectory $[a,b] \ni t \mapsto \xi_*(t) \in \Omega$ of a control system Σ with state space Ω, and gives a necessary condition for ξ_* to be a "boundary trajectory" (i.e. such that $\xi_*(b)$ lies on the boundary of the set of points reachable from $\xi_*(a)$) or a minimizer of some cost functional of the Lagrange, Mayer or Bolza type, subject to various constraints. We will only be concerned here with the finite-dimensional versions of the MP, in which Ω is an open subset of $\mathbb{R}^n$ or a smooth manifold. In most such versions, the trajectories are either supposed to be generated by a dynamical equation $\dot{x} = f(x,t,u)$, together with the specification of a class $\mathcal{U}$ of "admissible open-loop controls" $t \mapsto u(t)$, or assumed to arise as the solutions of a differential inclusion (DI) $\dot{x} \in F(x,t)$. The former approach can be slightly generalized by considering parametrized families $f = \{f_\eta : \eta \in \mathcal{U}\}$ where each f_η is a time-varying vector field (TVVF), and $\mathcal{U}$ is just an abstract set. (The special case corresponding to $\dot{x} = f(x,t,u)$ gives rise to such a family, by taking $\mathcal{U}$ to be the class of all admissible open-loop controls and defining $f_\eta(x,t) = f(x,t,\eta(t))$ for $\eta \in \mathcal{U}$.) The work in this direction —henceforth referred to as "the TVVF approach"— follows the ideas of B. Kaskosz and S. Lojasiewicz Jr. on "generalized control systems," cf. [6], and those of Lojasiewicz in [7]. A version of the MP for TVVF systems was presented in [9] and [10].

The conventional wisdom on the subject holds that the DI approach is strictly more general than the TVVF method, because (a) one can always associate a DI $\dot{x} \in F(x,t)$ to a TVVF system $f = \{f_\eta : \eta \in \mathcal{U}\}$, by defining $F(x,t) = \{f_\eta(x,t) : \eta \in \mathcal{U}\}$, while on the other hand (b) set-valued maps —even very smooth ones— in general do not have continuous selections.

In this note we choose to follow a direction that contradicts this conventional wisdom, by pursuing the TVVF approach in spite of these alleged drawbacks. We work with *parametrized families* $F = \{F_\eta : \eta \in \mathcal{U}\}$, where each F_η is a set-valued map $(x,t) \mapsto F_\eta(x,t)$. This includes as a special case the usual MP for DI's as well as that for TVVF systems, but is in fact more general than either of those results. Using a generalization of a selection theorem due to A. Bressan [2, 3, 4], one can show that every "sufficiently reasonable" set-valued map $(x,t) \mapsto F(x,t)$ with closed values admits "sufficiently many" selections $(x,t) \mapsto f(x,t)$ that are "good enough for all our purposes," even though such selections will in general fail to be continuous in x for fixed t. Here, "sufficiently reasonable" will mean "almost lower semicontinuous" (ALSC), as defined below.† The selections f are "good enough" in the sense that the equation $\dot{x} = f(x,t)$ has local existence of solutions, and the set-valued map $(x,s,t) \mapsto \Phi^f(x,s,t)$ determined by the flow of f is upper semicontinuous. Finally, the words "sufficiently many" will refer to the fact that for every (x_0,t_0,y_0) in the graph of F there exists a selection f in our class such that $f(x_0,t_0) = y_0$, with the property that $\Phi^f(x,s,t)$ is "differentiable at (x_0,t_0,t_0), with differential $dx + (ds - dt) \cdot f(x_0,t_0)$," in the sense that

$$\begin{aligned}&\sup\{||z-x-(s-t)f(x_0,t_0)|| : z \in \Phi^f(x,s,t)\}\\ &\quad = o(||x-x_0||+|t-t_0|+|s-t_0|) \text{ as } x \to x_0,\ (t,s) \to (t_0,t_0).\end{aligned}$$

Bressan's selection result makes it possible to reduce the general DI situation to that of TVVF systems, as long as we are willing to work with discontinuous vector fields. As in all other versions of the MP, we must assume that the reference trajectory arises from a TVVF or DI that satisfies a suitable Lipschitz condi-

*This research was supported in part by NSF Grant DMS95-00798 and AFOSR Grant 0923.

†The ALSC condition on F is in fact more general than the conditions usually imposed for DI versions of the MP.

tion. We do this by requiring that ξ_* be a solution of $\dot{x} \in F_{\eta_*}(x,t)$, where the "reference DI" F_{η_*} is *integrably pseudolipschitz along* ξ_*. (The precise definition is given in Section 8 below. The pseudolipschitz condition was introduced by J.P. Aubin in [1].)

It then turns out that F_{η_*} admits "integrably Lipschitz" (IL) selections $(x,t) \mapsto f(x,t)$ (cf. Theorem 1). So ξ_* can actually be thought of as generated by an ordinary IL TVVF, and this makes it possible to write the usual "adjoint inclusion" $-\dot{\zeta}(t) \in \zeta(t).\partial_x f(\xi_*(t),t)$, where $\partial_x f(x,t)$ is the Clarke generalized Jacobian at x of the map $x' \mapsto f(x',t)$.

Remark 1 The version of the MP presented here can be improved in a number of ways. For example, one can allow "jump maps," in which case the adjoint equation must be modified to allow the adjoint vector to jump as well. Or, instead of assuming that the reference trajectory ξ_* arises from a DI which is integrably pseudolipschitz along ξ_*, one can assume that ξ_* is generated by a flow $\{\Phi_{s,t}\}_{a\leq t\leq s\leq b}$, such that each map $\Phi_{s,t}$ is semidifferentiable at $\xi_*(t)$. This yields a version of the MP that covers many cases where ξ_* is generated by a *discontinuous* vector field, such as, for example, in the derivation of Snell's law of refraction from Fermat's minimum time principle, or in the "reflected brachystochrone problem" described in our companion history paper [11].

It is possible to formulate a single version of the maximum principle that contains all these cases. This requires extending the theory of semidifferentiable maps to the larger class of "multidifferentiable maps." Full details of this will be given in [12]. □

2. Statement of the main theorem

We use $\mathbb{R}$, $\mathbb{R}^{n\times m}$, $\mathbb{R}^n$, $\mathbb{R}_n$, to denote, respectively, the real line, the set of n by m real matrices, and the spaces $\mathbb{R}^{n\times 1}$, $\mathbb{R}^{1\times n}$. "Interval" means "nonempty connected subset of $\mathbb{R}$." If μ is a map, then $\mathrm{Dom}(\mu)$ is the domain of μ.

If I is an interval, then a *DI control system with time set* I is a triple $\Sigma = (\Omega, \mathcal{U}, F)$ such that Ω (the *state space* of Σ) is an open subset of $\mathbb{R}^n$ for some n, $\mathcal{U}$ is a set, and $F = \{F_\eta : \eta \in \mathcal{U}\}$ is a parametrized family of set-valued maps from $\Omega \times I$ to $\mathbb{R}^n$. A *trajectory-control pair* of a DI control system $\Sigma = (\Omega, \mathcal{U}, F)$ is a pair $\gamma = (\xi, \eta)$ such that $\eta \in \mathcal{U}$ and $\xi : J \mapsto \Omega$ is a locally absolutely continuous map, defined on a subinterval J of I, such that $\dot{\xi}(t) \in F_\eta(\xi(t),t)$ for almost all $t \in J$. We use $TCP(\Sigma)$ to denote the set of all trajectory-control pairs of Σ.

Let $\Sigma = (\Omega, \mathcal{U}, F)$ be a DI control system with time set I. If $p, x \in M$, $a \in I$, $b \in I$, $a \leq b$, we say that x is Σ*-reachable from* p *over* $[a,b]$ if there exists a $\gamma = (\xi, \eta)$ in $TCP(\Sigma)$ such that $\xi(a) = p$ and $\xi(b) = x$. We define $\mathcal{R}^\Sigma_{[a,b]}(S)$, the Σ*-reachable set from a subset* S *of* Ω *over* $[a,b]$, to be the set of all $x \in \Omega$ that are Σ-reachable from p over $[a,b]$ for some $p \in S$. We write $\mathcal{R}^\Sigma(S) \stackrel{\mathrm{def}}{=} \cup_{a,b} \mathcal{R}^\Sigma_{[a,b]}(S)$.

If F and G are two set-valued maps on $\Omega \times I$ and E is a subset of I, we use $F\#_E G$ to denote the set-valued map H such that $H(x,t) = F(x,t)$ when $x \in \Omega$, $t \notin E$, and $H(x,t) = G(x,t)$ when $x \in \Omega$, $t \in E$.

If F is a set-valued map whose values are subsets of $\mathbb{R}^n$, we use $\mathrm{co}(F)$ to denote the *convexification* of F, i.e. the map whose value for every x is the closed convex hull of $F(x)$.

If $\varepsilon > 0$, then the ε*-tube about* a curve $\xi : [a,b] \mapsto \mathbb{R}^n$ is the set $\mathcal{T}(\xi,\varepsilon) = \{(x,t) \in \mathbb{R}^n \times [a,b] : ||x - \xi_*(t)|| \leq \varepsilon\}$. We use $\mathcal{T}_x(\xi,\varepsilon)$, $\mathcal{T}^t(\xi,\varepsilon)$ to denote, respectively, the sets $\{t : (x,t) \in \mathcal{T}(\xi,\varepsilon)\}$, and $\{x : (x,t) \in \mathcal{T}(\xi,\varepsilon)\}$.

If $a, b \in \mathbb{R}$, $a \leq b$, f is a single-valued map with values in $\mathbb{R}^m$, and $\xi : [a,b] \mapsto \mathbb{R}^n$ is a curve, we say that f is *integrably Lipschitz* (IL) *near* ξ if there exists an $\varepsilon > 0$ such that (a) $\mathcal{T}(\xi,\varepsilon) \subseteq \mathrm{Dom}(f)$, (b) for every x, $f(x,t)$ is measurable with respect to t on $\mathcal{T}_x t(\xi,\varepsilon)$, (c) for almost every $t \in [a,b]$, $f(x,t)$ is Lipschitz with respect to x on $\mathcal{T}^t(\xi,\varepsilon)$, with a Lipschitz constant $k(t)$ which is integrable on $[a,b]$ as a function of t.

In the following statement of our version of the MP, we use a number of concepts —"generalized approximating cone" (abbr. GAC), "semidifferentiable map," "semidifferential," "variational vectors," and "compatibility"— that were introduced in [10], as well as those of ALSC, LILB and integrably pseudolipschitz maps, to be defined below.

The statement of the MP is then as follows:

Theorem 1 *Let I be an interval, and let $\Sigma = (\Omega, \mathcal{U}, F)$ be a DI control system with time set I. Assume that each set-valued map F_η has closed values and is ALSC. Let $\gamma_* = (\xi_*, \eta_*) \in TCP(\Sigma)$ (the "reference trajectory-control pair") be such that ξ_* has domain $[a,b]$, where $a, b \in I$ and $a \leq b$. Assume that F_{η_*} is integrably pseudolipschitz along ξ_*. Assume also that for every $\eta \in \mathcal{U}$ one of the following two substitution conditions hold: (s1) F_η is LILB, and for every compact subinterval E of I there is an $\eta' \in \mathcal{U}$ such that $F_{\eta_*}\#_E F_\eta = F_{\eta'}$, (s2) for every measurable subset E of I there is an $\eta' \in \mathcal{U}$ such that $F_{\eta_*}\#_E F_\eta = F_{\eta'}$. Let $p_* = \xi_*(a)$, $q_* = \xi_*(b)$. Let g be a (possibly set-valued) map, with values in $\mathbb{R}^m$, defined on a neighborhood U of q_* in $\mathbb{R}^n$. Let S_0, S be subsets of of $\mathbb{R}^n$, $\mathbb{R}^m$, respectively, such that $p_* \in S_0$ and $g(q_*) \in S$. Assume that S_0, S have GAC's C_0, C at p_* and $g(q_*)$, respectively. Assume also that either (i) C is not a linear subspace of $\mathbb{R}^m$ or (ii) $C = \{0\}$ and $g(q_*)$ belongs to $\mathrm{Clos}(S - \{g(q_*)\})$. Assume that g is semidifferentiable at q_*, and let $G \subseteq \mathbb{R}^{m\times n}$ be a semidifferential of g at q_*. Let $\mathbf{v}$ be a compatible set of point variational vectors for γ_*. Assume that $g(\mathcal{R}^\Sigma_{[a,b]}(S_0)) \cap S \cap W = \{g(q_*)\}$ for some neighborhood W of $g(q_*)$. Then*

(1) There exists a selection f of $\mathrm{co}(F_{\eta_*})$ which is IL near ξ_* and such that $\dot{\xi}_*(t) = f(\xi_*(t),t)$ for a.e. t.

(2) For every f with the properties of (1) there exist: (I) a measurable matrix function $A : [a,b] \mapsto \mathbb{R}^{n\times n}$ such that $A(t) \in \partial_x f(\xi_*(t),t)$ for a.e. $t \in [a,b]$, (II) a $B \in G$, (III) a $c^* \in \mathbb{R}_m - \{0\}$, such that, if $\zeta : [a,b] \mapsto \mathbb{R}_n$ is the unique absolutely continuous

solution of $\dot\zeta(t) = -\zeta(t)A(t)$ with terminal condition $\zeta(b) = c^*B$, then (i) for every $\eta \in \mathcal{U}$ the inequality $\zeta(t).\dot\xi_*(t) \le \inf\{\zeta(t).y : y \in F_\eta(\xi_*(t),t)\}$ holds for almost all t in $[a,b]$, (ii) $c^*c \le 0$ for all $c \in C$, (iii) $\zeta(a)c_0 \ge 0$ for all $c_0 \in C_0$, and (iv) $\zeta(t)v \ge 0$ for all $(v,t) \in \mathbf{v}$. □

3. Quasicontinuous time-varying maps

Assume that X, Y are finite-dimensional real normed spaces, $\Omega \subseteq X$ is Borel measurable and nonempty, and I is a nontrivial interval.

A *modulus of local integrability* (MLI) on $\Omega \times I$ is a function $\psi : \Omega \times I \mapsto \bar{\mathrm{I\!R}}_+$ such that

(i) $\psi(x,t)$ is measurable as a function of t for all $x \in \Omega$, and continuous as a function of x for almost all $t \in I$,

(ii) the function $t \mapsto \varphi_K(t) \overset{\text{def}}{=} \sup_{x\in K} \psi(x,t)$ is locally integrable for every compact subset K of Ω. □

If ψ, $\tilde\psi$ are MLI's on $\Omega \times I$, we say that $\tilde\psi$ *dominates* ψ, and write $\tilde\psi \succeq \psi$, if for every $(x_0,t_0) \in \Omega \times I$ there exist neighborhoods V_{x_0}, V_{t_0} of x_0, t_0 in Ω, I, such that $\psi(x,t) < \tilde\psi(x',t)$ for all $t \in V_{t_0}$, $x, x' \in V_{x_0}$.

We use $C^0(I,X)$ to denote the space of continuous functions $\xi : I \mapsto X$, endowed with the topology of uniform convergence on compact subintervals of I, so $C^0(I,X)$ is a Fréchet space, and a Banach space if I is compact. We use $C^0(I,\Omega)$ to denote the set of all continuous curves in Ω with domain I, topologized as a subspace of $C^0(I,X)$, and write $\mathcal{BL}(\Omega,I)$ to denote the σ-algebra of Borel$\times$Lebesgue-measurable subsets of $\Omega \times I$.

A map $f : \Omega \times I \mapsto Y$ is *essentially $\mathcal{BL}(\Omega,I)$-measurable* if there exists a subset J of I of zero Lebesgue measure such that $f\lceil\Big(\Omega\times(I\backslash J)\Big)$ is $\mathcal{BL}(\Omega,I)$-measurable. We use $\mathcal{M}_e(\mathcal{BL}(\Omega,I),Y)$ to denote the set of all essentially $\mathcal{BL}(\Omega,I)$-measurable maps $\Omega\times I \mapsto Y$.

If $f : \Omega \times I \mapsto Y$ is a map, and $\xi \in C^0(I,\Omega)$, then $f\circ\xi$ denotes the map $I \ni t \mapsto f(\xi(t),t) \in Y$.

In the following definitions, $\mathcal{S}$ is a subset of $C^0(I,\Omega)$, endowed with the subspace topology, and $f : \Omega\times I \mapsto Y$ is a mapping.

Definition 1 We call f *locally integrably bounded* (LIB) if there exists a MLI ψ on $\Omega \times I$ such that

$$||f(x,t)|| \le \psi(x,t) \ \text{ for all } \ x \in \Omega\,,\ t \in I\,. \qquad (1)$$

A set $\mathcal{F} \subseteq \mathcal{M}_e(\mathcal{BL}(\Omega,I),Y)$ is *uniformly LIB* if every $f \in \mathcal{F}$ is LIB and, in addition, the MLI ψ such that (1) holds can be chosen independently of f. □

Definition 2 We say that f is *quasiintegrable on $\mathcal{S}$* if $f\circ\xi$ is locally integrable for every $\xi \in \mathcal{S}$. □

If f is quasiintegrable on $\mathcal{S}$, then we can define a map $\mathcal{T}_f^{\mathcal{S}} : I \times I \times \mathcal{S} \mapsto Y$ by

$$\mathcal{T}_f^{\mathcal{S}}(a,t,\xi) = \int_a^t f(\xi(s),s)\,ds\,. \qquad (2)$$

Definition 3 We say that f is *quasicontinuous on $\mathcal{S}$* if

(QC.1) $f \in \mathcal{M}_e(\mathcal{BL}(\Omega,I),Y)$,

(QC.2) f is quasiintegrable on $\mathcal{S}$,

(QC.3) the map $\mathcal{T}_f^{\mathcal{S}} : I\times I\times\mathcal{S} \mapsto Y$ is continuous. □

If $\psi : \Omega \times I \mapsto \mathrm{I\!R}_+ \cup \{+\infty\}$ is a function, we use $C^0_\psi(I,X)$ to denote the set of all $\xi \in C^0(I,X)$ such that

$$\limsup_{h\downarrow 0} \frac{||\xi(t+h)-\xi(t)||}{h} \le \psi(\xi(t),t) \text{ for a.e. } t \in I\,. \quad (3)$$

Naturally, $C^0_\psi(I,X)$ is a convex subset of $C^0(I,X)$. We write

$$C^0_\psi(I,\Omega) \overset{\text{def}}{=} C^0_\psi(I,X) \cap C^0(I,\Omega)\,. \qquad (4)$$

If $C > 0$, and $(x,t) \in X \times \mathrm{I\!R}$, we write

$$\Gamma_C(x,t) \overset{\text{def}}{=} \{(x',s) \in X\times\mathrm{I\!R} : s \ge t,\ ||x'-x|| \le C(s-t)\},$$

and we let $\Gamma^2_C(x,t)$ be the set of all (x',x'',s) such that $(x',s)\in\Gamma_C(x,t)$ and $(x'',s)\in\Gamma_C(x,t)\}$.

Definition 4 Let $C > 0$. We say that f is *forward Γ_C-continuous* at $(x,t)\in\Omega\times I$ if $\lim f(x',s) = f(x,t)$ as $(x',s)\to(x,t)$ via values in $\Gamma_C(x,t)$.

We say that f is *weakly forward Γ_C-continuous* at (x,t) if $\lim\Big(f(x',s) - f(x'',s)\Big) = 0$ as $(x',x'',s)\to(x,x,t)$ via values in $\Gamma^2_C(x,t)$.

If $(x,t)\mapsto C(x,t)\in[0,\infty]$ is a function on $\Omega\times I$, we say that f is *forward Γ_C-continuous* (resp. *weakly forward Γ_C-continuous*) if it is forward (resp. weakly forward) $\Gamma_{C(x,t)}$-continuous at every point $(x,t)\in\Omega\times I$ such that $C(x,t) < \infty$. □

Theorem 2 *Let C, ψ be functions on $\Omega\times I$ such that, for almost all $t\in I$, the inequality $\psi(x,t) < C(x,t) < \infty$ holds for all $x\in\Omega$. Assume that $f\in\mathcal{M}_e(\mathcal{BL}(\Omega,I),Y)$ is weakly forward Γ_C-continuous on $\Omega\times I$ and LIB. Then f is quasicontinuous on $C^0_\psi(I,\Omega)$.* □

Quasicontinuous maps have the following *measurable substitution property*:

Theorem 3 *Let $\{f_j\}$ be a uniformly LIB sequence of maps $\Omega\times I\mapsto Y$. Let $\mathcal{S} \subseteq C^0(I,\Omega)$ be such that the f_j are quasicontinuous on $\mathcal{S}$. Let $\{E_j\}$ be a sequence of pairwise disjoint Lebesgue measurable subsets of I such that $I = \cup_j E_j$. Define $f : \Omega \times I \mapsto Y$ via*

$$f(x,t) = f_j(x,t) \ \text{ for } \ x \in \Omega\,,\ t \in E_j\,. \qquad (5)$$

Then f is quasicontinuous on $\mathcal{S}$. □

4. Discontinuous vector fields and their flows

We assume that X is a finite-dimensional real normed space, $\Omega \subseteq X$ is open and nonempty, and I is a nontrivial interval.

A *time-varying vector field* (abbreviated "TVVF") on Ω, with *time domain* I, is an essentially $\mathcal{BL}(\Omega,I)$-measurable map $f:\Omega\times I \mapsto X$. We use $TVVF(\Omega,I)$ to denote the set of all TVVF's on Ω with time domain I.

If $f \in TVVF(\Omega,I)$, a *trajectory* of f is a locally absolutely continuous map $\xi : J \mapsto \Omega$, defined on a nonempty subinterval J of I, such that $\dot\xi(t) = f(\xi(t),t)$

for almost every $t \in J$. We use $\mathrm{Traj}(f)$ to denote the set of all trajectories of f, and $\mathrm{Traj}_c(f)$ to denote the set of all $\xi \in \mathrm{Traj}(f)$ whose domain is a compact interval. If $S \subseteq \Omega \times I$, then $\mathrm{Traj}_c(f, S)$ denotes the set of all $\xi \in \mathrm{Traj}_c(f)$ such that $(\xi(t), t) \in S$ for all $t \in \mathrm{Dom}(\xi)$.

A *maximal trajectory* of f is a trajectory $\xi : J \mapsto \Omega$ that cannot be extended to a trajectory of f defined on a strictly larger interval. For $(x, s, t) \in \Omega \times I \times I$, we define

$$\Phi^f(x, s, t) = \{\xi(s) : \xi \in \mathrm{Traj}(f)\, ,\, \xi(t) = x\}\, . \tag{6}$$

The set-valued map $\Phi^f : \Omega \times I \times I \mapsto 2^\Omega$ is the *flow* of f.

The sets $\mathrm{Traj}_c(f)$ are subsets of the space

$$CRV(\Omega) \stackrel{\mathrm{def}}{=} \bigcup_{-\infty < a \le b < \infty} C^0([a, b], \Omega) \tag{7}$$

of curves in Ω. For each $\xi \in CRV(\Omega)$, we let $a(\xi)$, $b(\xi)$ be the unique a, b such that $\xi \in C^0([a, b], \Omega)$. We regard $CRV(\Omega)$ as a metric space, endowed with the metric $d_{CRV} : CRV(\Omega) \times CRV(\Omega) \mapsto \mathbb{R}$ given by

$$d_{CRV}(\xi, \eta) = d^1_{CRV}(\xi, \eta) + \sup\{||\tilde{\xi}(t) - \tilde{\eta}(t)|| : t \in \mathbb{R}\}, \tag{8}$$

where $d^1_{CRV}(\xi, \eta) = |a(\xi) - a(\eta)| + |b(\xi) - b(\eta)|$ and, for $\xi \in CRV(\Omega)$, we use $\tilde{\xi}$ to denote the continuous curve $\tilde{\xi} : \mathbb{R} \mapsto \Omega$ obtained by extending ξ to the whole real line by just letting $\tilde{\xi}$ be constant on $]-\infty, a(\xi)]$ and on $[b(\xi), +\infty[$. Then a sequence $\{\xi^j\}$ of curves, possibly defined on different time intervals $[a^j, b^j]$, converges to a curve $\xi : [a, b] \mapsto \Omega$ iff (i) $a^j \to a$, (ii) $b^j \to b$, and (iii) $\xi^j \to \xi$ uniformly, i.e. $\xi^j(t^j) \to \xi(t)$ whenever $t^j \in [a^j, b^j]$, $t^j \to t$.

Zorn's Lemma implies that every trajectory can be extended to a maximal trajectory. With our definitions, if $(x, t) \in \Omega \times I$, $J = \{t\}$, and $\xi : J \mapsto \Omega$ is given by $\xi(t) = x$, then ξ is a trajectory of f. Therefore, if $f \in TVVF(\Omega, I)$, then for every $(x, t) \in \Omega \times I$ there is a maximal trajectory ξ of f such that $\xi(t) = x$, although it may turn out that the domain $\mathrm{Dom}(\xi)$ of such a trajectory is just $\{t\}$.

If ψ is a MLI on $\Omega \times I$, we define $\mathcal{S}(\psi)$ to be the space of all locally absolutely continuous curves $\xi : I \mapsto \Omega$ such that $||\dot{\xi}(t)|| \le \psi(\xi(t), t)$ for almost all $t \in I$. Notice that $\mathcal{S}(\psi) \subseteq C^0_\psi(I, \Omega)$.

We will be interested in time-varying vector fields $f \in TVVF(\Omega, I)$ that satisfy the following property:

(AD) *There exist moduli of local integrability ψ, $\tilde{\psi}$ on $\Omega \times I$ such that $\tilde{\psi} \succeq \psi$, for which*

- *(AD1)* $||f(x, t)|| \le \psi(x, t)$ *for all* $(x, t) \in \Omega \times I$,
- *(AD2)* f *is quasicontinuous on* $\mathcal{S}(\tilde{\psi})$,

Definition 5 An $f \in TVVF(\Omega, I)$ for which (AD) holds is called *admissible.* We use $TVVF_{AD}(\Omega, I)$ to denote the class of all admissible $f \in TVVF(\Omega, I)$. □

If $f \in TVVF(\Omega, I)$, we say that f has the *existence property* (EP) if the domain of every maximal trajectory ξ of f is relatively open in I, and we call f *trajectory-compact* if the set $\mathrm{Traj}_c(f, K)$ is compact for every compact subset K of $\Omega \times I$. We say that f is *flow-continuous* at $(x, t) \in \Omega \times I$ if there exists a $\delta > 0$ such that, on the set $\{(x', s', s) : ||x' - x|| \le \delta, |s' - t| \le \delta, |s - t| \le \delta, s' \in I, s \in I\}$, the set-valued map Φ^f is upper semicontinuous and has nonempty compact values. We say that f is *flow-continuous* if it is flow-continuous at every point of $\Omega \times I$.

Theorem 4 *Assume that Ω is open in X, I is an interval, and $f \in TVVF_{AD}(\Omega, I)$. Then f has the EP and is trajectory-compact and flow-continuous.* □

Theorem 4 contains as a particular case the usual Carathéodory existence theorem: *if Ω is open in X, $f : \Omega \times I \mapsto X$ is LIB and such that the map $x \mapsto f(x, t)$ is continuous for almost every $t \in I$, and the map $t \mapsto f(x, t)$ is measurable for all $x \in \Omega$, then f is quasicontinuous on the space $C^0(I, \Omega)$. In particular, $f \in TVVF_{AD}(\Omega, I)$, and f has the existence property and is trajectory-compact and flow-continuous.*

The following is a trivial consequence of flow-continuity:

Corollary 1 *Let $f \in TVVF_{AD}(\Omega, I)$, $(x, t) \in \Omega \times I$. Then*

$$\limsup \left\{ ||y - x|| : y \in \Phi^f(x', s', s) \right\} = 0 \tag{9}$$

as $(x', s', s) \to (x, t, t)$, $x' \in \Omega$, $s', s \in I$. □

5. Differentiability of the flow

We assume in this section that X is a finite-dimensional real linear space, Ω is an open subset of X, and I is an interval.

Definition 6 Let $f : \Omega \times I \mapsto X$ be a single-valued map. We say that f is *almost continuous* (abbr. AC) at a point $(x, t) \in \Omega \times I$ if there exists a measurable subset E of I such that

- *(i)* t is a point of density of E,
- *(ii)* $\lim_{(x', s) \to (x, t), s \in E} f(x', s) = f(x, t)$. □

Definition 7 Let $f \in TVVF_{AD}(\Omega, I)$ and let (x, t) belong to $\Omega \times I$. We say that (x, t) is a *regular point* of f if the following two conditions hold:

- *(i)* f is AC at (x, t),
- *(ii)* there exists a compact subset K of Ω, containing x in its interior, such that there is a locally integrable function $\varphi : I \mapsto \mathbb{R} \cup \{+\infty\}$ for which
  - *(a)* $||f(x', s)|| \le \varphi(s)$ for all $x' \in K$, $s \in I$,
  - *(b)* t is a Lebesgue point of φ. □

(Recall that a *Lebesgue point* of φ is a point $s \in I$ such that $\lim_{h \to 0+} h^{-1} \int_{[s-h, s+h] \cap I} \left| \varphi(\sigma) - \varphi(s) \right| d\sigma = 0$. It is well known that if φ is locally integrable on I then almost every point of I is a Lebesgue point.)

The differentiation theorem says, roughly, that if (x, t) is a regular point of f, then

$$\Phi^f(x', s', s) = x' + (s' - s) f(x, t) + o(||x' - x|| + |s' - t| + |s - t|) \tag{10}$$

as $x' \to x$ and $(s', s) \to (t, t)$.

Since Φ^f is set-valued, (10) requires interpretation. The precise meaning of (10) is that the linear map $D\Phi^f(x,t,t) : X \times \mathbb{R} \times \mathbb{R} \mapsto X$ given by

$$D\Phi^f(x,t,t)(v,h,k) = v + (h-k)f(x,t) \qquad (11)$$

is a first-order approximation to $\Phi^f(x+v, t+h, t+k) - x$ near $(0,0,0)$, i.e.

$$\frac{\sup\Big\{||y - x' - (s' - s)f(x,t)|| : y \in \Phi^f(x', s', s)\Big\}}{||x' - x|| + |s' - t| + |s - t|} \to 0 \quad (12)$$

as $(x', s', s) \to (x,t,t)$ via $(\Omega \times I \times I) \backslash \{(x,t,t)\}$.

When this happens, we say that Φ^f is *differentiable at* (x,t,t), and *the differential of* Φ^f *at* (x,t,t) is the map $D\Phi^f(x,t,t)$.

Theorem 5 *If $f \in TVVF_{AD}(\Omega, I)$ and (x,t) is a regular point, then (12) holds, that is, the flow Φ^f is differentiable at (x,t,t), and the differential of Φ^f at (x,t,t) is the map $D\Phi^f(x,t,t)$ defined by (11).* □

Theorem 6 *Let X,Ω, I be, respectively, a finite-dimensional real linear space, an open subset of X, and an interval. Let $f \in TVVF_{AD}(\Omega, I)$. Let (x,t) be a regular point of f. Then the flow Φ^f is semidifferentiable at (x,t,t), and a semidifferential of Φ^f at (x,t,t) is the set $\Big\{D\Phi^f(x,t,t)\Big\}$, where $D\Phi^f(x,t,t)$ is the map defined by (11).* □

6. LSC and ALSC set-valued maps

Suppose X and Y are metric spaces, with distance functions d_X, d_Y, and $F : X \mapsto 2^Y$ is a set-valued map. As usual, $\mathrm{Dom}(F)$ is the set $\{x \in X : F(x) \neq \emptyset\}$.

Definition 8 If $x \in X$, then F is *lower semicontinuous* (LSC) at x if for every $y \in F(x)$ and every $\varepsilon > 0$ there exists a $\delta > 0$ such that, whenever $x' \in X$ and $0 < d_X(x', x) \leq \delta$, it follows that there exists a $y' \in F(x')$ such that $d_Y(y', y) \leq \varepsilon$. We call F *lower semicontinuous on a subset S of X* if F is LSC at every point of S. We call F *lower semicontinuous* if it is LSC on X. □

Assume now that F is a "time-varying set-valued map from X to Y," i.e. that S is a subset of $X \times \mathbb{R}$, and $F : S \mapsto 2^Y$ is a set-valued map. If $J \subseteq \mathbb{R}$, we write $S_J = S \cap (X \times J)$, and we let F_J denote the restriction of F to S_J, so F_J is a set-valued map from S_J to Y, and $F_J(x,t) = F(x,t)$ whenever $(x,t) \in S_J$.

Definition 9 We call F *almost lower semicontinuous* (ALSC) if for every $\varepsilon > 0$ there exists a closed subset J of $\mathbb{R}$ such that $\mathrm{meas}(\mathbb{R} \backslash J) \leq \varepsilon$ and F_J is LSC. □

An equivalent form of the above condition is

(ALSC') *there exists a sequence $J_0, J_1, J_2, \ldots$ of pairwise disjoint subsets of $\mathbb{R}$ such that J_0 has measure zero, the J_i for $i > 0$ are compact, and $\mathbb{R} = \cup_{j=0}^{\infty} J_j$, with the property that the set-valued maps F_{J_j} are LSC for $j > 0$.*

If I is an interval, $F : X \times I \mapsto 2^Y$ is a set-valued map, and $g : X \times I \mapsto Y$ is a single-valued map, we let

$$\rho_{g,F}(x,t) = d_Y(g(x,t), F(x,t)) \ \text{ for } \ (x,t) \in X \times I. \ (13)$$

Definition 10 If I is an interval and $F : X \times I \mapsto 2^Y$ is a set-valued map, then F is *locally integrably lower-bounded* (abbr. LILB) if for some —and hence every— $\bar{y} \in Y$ there exists a MLI ψ on $X \times I$ with the property that $\rho_{\bar{y},F}(x,t) < \psi(x,t)$ for all $(x,t) \in X \times I$. □

7. Selections of ALSC DI's

We now present a generalization of Bressan's selection theorem for ALSC DI's. We assume that X and Y are finite-dimensional real linear spaces, Ω is an open subset of X, I is an interval, and $F : \Omega \times I \mapsto 2^Y$ is an ALSC set-valued map.

If $f : \Omega \times I \mapsto Y$ is a single-valued map, and $J \subseteq I$, we define $f_J : \Omega \times I \mapsto Y$ by letting $f_J(x,t) = \chi_J(t) f(x,t)$, where χ_J is the indicator function of J.

Definition 11 Let $f : \Omega \times I \mapsto Y$ be a single-valued map. Let $C : \Omega \times I \mapsto]0, \infty[$ be a function. We say that f is *almost weakly forward Γ_C-continuous* if there exists a family $\{J_j\}_{j=0}^{\infty}$ of pairwise disjoint subsets of I such that $I = \cup_j J_j$, the J_j are compact for $j > 0$, $\mathrm{meas}(J_0) = 0$, and the maps f_{J_j} are weakly forward Γ_C-continuous for $j > 0$. □

If Z is any finite-dimensional real linear space, then the well-known Scorza-Dragoni theorem says that a single-valued map $g : \Omega \times I \mapsto Z$ is ALSC as a set-valued map iff $g(x,t)$ is measurable with respect to t for each x and continuous with respect to x for each t.

Theorem 7 *Let $F : \Omega \times I \mapsto 2^Y$ be an ALSC multifunction with nonempty closed values. Let $g : \Omega \times I \mapsto Y$ be an ALSC single-valued map and let $\beta : \Omega \times I \mapsto \mathbb{R}$ be ALSC and such that*

$$\beta(x,t) > \rho_{g,F}(x,t) \ \text{ for all } \ (x,t) \in \Omega \times I \qquad (14)$$

for a.e. $t \in \mathbb{R}$. Let $C : \Omega \times I \mapsto]0, \infty[$ be ALSC. Then the following is true for almost all $t \in I$:

(#) For every $x \in \Omega$ and every $y \in F(x,t)$ with the property that $||g(x,t) - y|| < \beta(x,t)$ there exists a selection $f : \Omega \times I \mapsto Y$ of F such that

- *(i) $f(x,t) = y$ and f is AC at (x,t),*
- *(ii) f is almost weakly forward Γ_C-continuous,*
- *(iii) for a.e. $s \in I$, $||f(x', s) - g(x', s)|| < \beta(x', s)$ for all $x' \in \Omega$,*
- *(iv) f is essentially $\mathcal{BL}(\Omega, I)$-measurable.*

An immediate corollary of Theorems 3 and 7 is the following result, which involves two moduli of local integrability for f. In the applications, it will be convenient to take ψ as small as possible and make $\tilde{\psi}$ very large.

Theorem 8 *Let $F : \Omega \times I \mapsto 2^X$ be a LILB and ALSC multifunction. Let ψ, $\tilde{\psi}$ be moduli of local integrability on S such that $\tilde{\psi} \succeq \psi \succeq \rho_{0,F}$. Then the following is true for almost all $t \in I$:*

(#') For every $x \in \Omega$ and every $y \in F(x,t)$ such that $||y|| < \psi(x,t)$ there exists a selection $f : \Omega \times I \mapsto X$ of F such that

(i) $f \in TVVF_{AD}(\Omega, I)$,

(ii) $f(x,t) = y$ *and* (x,t) *is a regular point of* f,

(iii) f *is quasicontinuous on* $C^0_{\tilde{\psi}}(I, \Omega)$,

(iv) for almost every $s \in I$, *the inequality* $||f(x', s)|| < \psi(x', s)$ *holds for all* $x' \in \Omega$. □

8. Pseudolipschitz maps

Suppose X, Y are finite-dimensional normed real linear spaces, $F : X \mapsto 2^Y$ is a multifunction, and $y_0 \in F(x_0)$. We say that F is *pseudolipschitz* at (x_0, y_0) if there exist constants $C > 0$, $\bar{\varepsilon} > 0$, $\bar{\delta} > 0$ such that $F(x) \cap B_Y(y_0, \bar{\varepsilon})$ is closed whenever $x \in B_X(x_0, \bar{\delta})$, and $F(x) \cap B_Y(y_0, \bar{\varepsilon}) \subseteq F(x') + B_Y(0, C||x - x'||)$ whenever $x, x' \in B_X(x_0, \bar{\delta})$, where we are writing

$$B_Z(z, r) = \{z' \in Z : ||z' - z|| \leq r\}, \qquad (15)$$

if Z is a normed space with norm $||\ldots||$, and $r > 0$.

For $\varepsilon > 0$, define $F^{\varepsilon, y_0}(x) = F(x) \cap B_Y(y_0, \varepsilon)$.

Theorem 9 *If* F *is pseudolipschitz at* (x_0, y_0), *then for every number* $\varepsilon > 0$ *there exists a Lipschitz multifunction* $V^\varepsilon : N(\varepsilon) \mapsto 2^Y$, *defined on a neighborhood* $N(\varepsilon)$ *of* x_0 *and having compact convex values, such that* $V^\varepsilon(x_0) = \{y_0\}$ *and* $V^\varepsilon(x) \subseteq \mathrm{co}\Big(F^{\varepsilon, y_0}(x)\Big)$ *for all* $x \in N(\varepsilon)$. □

Theorem 9 implies in particular the existence of a Lipschitz selection $f^\varepsilon : N(\varepsilon) \mapsto Y$ of $\mathrm{co}(F^{\varepsilon, y_0})$ such that $f^\varepsilon(x_0) = y_0$, since one can choose $f^\varepsilon(x)$ to be the Steiner point of $V^\varepsilon(x)$.

The proof of Theorem 9 is by an explicit construction. It is easy to follow through the steps of this construction and show that, if the original F depends measurably on a parameter t, then V^ε will also be measurable with respect to t. This is important for the main application that motivates our result, namely, the case when we are given a trajectory $[a, b] \ni t \mapsto \xi_*(t) \in X$ of a differential inclusion $\dot{x} \in F(x, t)$, and we assume that $F : X \times [a, b] \mapsto 2^X$ is "integrably pseudo-Lipschitz along ξ_*," in the sense that there exist an integrable function $[a, b] \ni t \mapsto C(t) > 0$, and constants $\bar{\varepsilon} > 0$, $\bar{\delta} > 0$ such that: (i) for every $x \in X$, the multifunction $t \mapsto F(x, t) \cap B_Y(\dot{\xi}_*(t), \bar{\varepsilon})$ is measurable on the set $\{t : ||x - \xi_*(t)|| \leq \bar{\delta}\}$, (ii) for almost every $t \in [a, b]$, $x \in B_X(\xi_*(t), \bar{\delta})$ implies that $F(x, t) \cap B_Y(\dot{\xi}_*(t), \bar{\varepsilon})$ is closed, and (iii) $x, x' \in B_X(\xi_*(t), \bar{\delta})$ implies that $F(x, t) \cap B_Y(\dot{\xi}_*(t), \bar{\varepsilon}) \subseteq F(x', t) + B_Y(0, C(t)||x - x'||)$.

In that case, our result yields the existence, for each $\varepsilon > 0$, of an *integrably Lipschitz selection* f^ε of $\mathrm{co}(F^\varepsilon)$ along ξ_* such that $f^\varepsilon(\xi_*(t), t) = \dot{\xi}_*(t)$ for a.e. t.

Precisely, we have

Theorem 10 *Let* $F : X \times [a, b] \mapsto 2^X$ *be a set-valued map, and let* $\xi_* : [a, b] \mapsto X$ *be a solution of the DI* $\dot{x} \in F(x, t)$. *Assume that* F *is integrably pseudolipschitz along* ξ_*. *Then for every* $\varepsilon > 0$ *there is a* $\delta > 0$ *such that there exists a time-varying vector field* $(x, t) \mapsto f(x, t)$, *defined on the tube* $T(\xi_*, \delta)$, *such that*

(IL) f *is integrably Lipschitz near* ξ_*,

(SEL) $f(x, t) \in \mathrm{co}\Big(F(x, t) \cap B_Y(\dot{\xi}_*(t), \varepsilon)\Big)$ *for each* $(x, t) \in T(\xi_*, \delta)$,

(C) $f(\xi_*(t), t) = \dot{\xi}_*(t)$ *for a.e.* $t \in [a, b]$. □

9. Outline of the proof of the MP

The proof uses the same techniques as that of the MP for control systems, sketched in [10]. Theorem 10 gives the existence of the selection f of F_{η_*}. Using Theorems 6 and 8 we construct, for each $\eta \in \mathcal{U}$, for almost all $t \in I$, and for each $y_0 \in F_\eta(\xi_*(t), t)$, a point variation with variational vector y_0. These variations are then combined exactly as in the proof of the theorem of [10], and the existence of ζ then follows by a compactness argument.

References

[1] Aubin, J.-P.: "Lipschitz behavior of solutions to convex minimization problems," *Math. Oper. Res.* **9** (1984), pp. 87-111.

[2] Bressan, A., "On differential relations with lower continuous right-hand side. An existence theorem." *J. Diff. Equations* **37** (1980), pp. 89-97.

[3] Bressan, A., "Solutions of lower semicontinuous differential inclusions on closed sets," *Rend. Sem. Mat. Univ. Padova* **69** (1983), pp. 99-107.

[4] Bressan, A., "Directionally continuous selections and differential inclusions," *Funk. Ekvac.* **31** (1988), pp. 459-470.

[5] Clarke, F.H., "The Maximum Principle under minimal hypotheses," *S.I.A.M. J. Control Optim.* **14** (1976), 1078-1091.

[6] Kaskosz, B., and S. Lojasiewicz Jr., "A Maximum Principle for generalized control systems," *Nonlinear Anal. TMA*, **9** (1985), pp. 109-130.

[7] Lojasiewicz Jr., S., "Local controllability of parametrized differential equations," in preparation.

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

[9] Sussmann, H.J., "A strong version of the Lojasiewicz Maximum Principle," in *Optimal Control of Differential Equations,* Nicolai H. Pavel Ed., M. Dekker Inc., New York (1994).

[10] Sussmann, H.J., "A strong version of the Maximum Principle under weak hypotheses," *Proc. 33rd IEEE CDC*, Orlando, FL, Dec. 1994, pp. 1950-1956.

[11] Sussmann, H.J., "From the Brachystochrone to the Maximum Principle and back," in these proceedings (history session).

[12] Sussmann, H. J., "The maximum principle of optimal control theory," in preparation.