A geometrically based criterion to avoid infimum gaps in Optimal Control

M. Palladino

Gran Sasso Science Institute - GSSI
via Crispi 7, L'Aquila, 67100, Italy

F. Rampazzo

University of Padova, Department of Mathematics
via Trieste 63, Padova, 35121, Italy

Abstract

In optimal control theory, *infimum gap* means a non-zero difference between the infimum values of a given minimum problem and an extended problem obtained by embedding the original family $\mathcal{V}$ of controls in a larger family $\mathcal{W}$. For some embeddings –like standard convex relaxations or impulsive extensions – the normality of an extended minimizer has been shown to be sufficient for the avoidance of infimum gaps. A natural issue is then the search of a general hypothesis under which the criterium "normality implies no gap" holds true. We prove that this criterium is actually valid as soon as $\mathcal{V}$ is *abundant* in $\mathcal{W}$, without any convexity assumption on the extended dynamics. *Abundance*, which was introduced by J. Warga in a convex context and was later generalized by B. Kaskosz, strengthens density, the latter being not sufficient for the mentioned criterium to hold true.

Keywords: Optimal Control, Infimum Gap, Necessary Conditions, Set Separation.

1. Introduction

One of the main reason for enlarging the domain of a minimum problem relies on the aim of establishing the existence of at least one solution. Actually, domain extension is a quite common and variously motivated practice, in particular in the Calculus of Variations and in Optimal Control. Of course, a crucial requisite of such a domain enlargement consists in the *density* of the original problem in the new one: the extended minimum should be approximable by processes of the original problem. However, because of the presence of a final target and of dynamic constraints, even a dense extension of the domain may result in the occurrence of an *infimum gap*: namely, it can happen that the infimum value of the original problem is strictly greater than the infimum value of the extended problem. This might be undesirable in many respects, for instance in the convergence of numerical schemes as well as in the identification of the value function via Hamilton-Jacobi equations. This raises a natural question: how can one avoid this gap phenomenon? A sufficient condition for gap avoidance seems to emerge from investigations by J. Warga [47, 48, 49, 50] and from some other more recent papers [2, 33, 36, 37, 38, 39], dealing with some particular cases: this criterion is the so-called *normality* of minimizers. Therefore, the mentioned question can be turned into the following one:

(Q) *Under which hypotheses on a general optimal control problem normality is sufficient for gap-avoidance?*

In order to be more precise, let us briefly sketch the abstract setting of our optimal control problem. The *state* variable y will range on a Riemannian manifold $\mathcal{M}$, while the control maps v will belong to an *original* family $\mathcal{V} \subset \mathcal{W} := L^1([0,S],\mathfrak{W})$ (where $\mathfrak{W}$ is a subset of a metric space) or to a larger set $\mathcal{W}$, which will be called the *extended* family of controls. Given an initial state $\bar{y} \in \mathcal{M}$ and a time interval $[0,S]$, we will consider the control system

$$(E) \qquad \begin{cases} \dfrac{dy}{ds}(s) = f(s, y(s), w(s)) \\ y(0) = \bar{y}, \end{cases}$$

and, for every $w \in \mathcal{W}$, we will use $y[w] : [0,S] \to \mathcal{M}$ to denote the corresponding (supposedly unique) solution. The *original optimal control problem* is defined as

$$(P)_{\mathcal{V}} \qquad \textit{Minimize}\Big\{\ h\big(y[v](S)\big) \quad | \quad v \in \mathcal{V}, \quad y[v](S) \in \mathfrak{T}\Big\},$$

where the *cost function* $h : \mathcal{M} \to \mathbb{R}$ is continuous, and $\mathfrak{T} \subset \mathcal{M}$ is a closed set called *target.*

Replacing the family of controls $\mathcal{V}$ with the larger set $\mathcal{W}$, one obtains the *extended optimal control problem:*

$$(P)_{\mathcal{W}} \qquad \textit{Minimize}\Big\{\ h\big(y[w](S)\big) \quad | \quad w \in \mathcal{W}, \quad y[w](S) \in \mathfrak{T}\Big\}.$$

We will assume the existence of a local minimum for the extended problem, namely a control $\hat{w} \in \mathcal{W}$ such that, for some C^0 neighbourhood $\mathcal{O}$ of $y[\hat{w}]$, $h(y[\hat{w}](S)) \leq h(y[w](S))$ for all $w \in \mathcal{W}$ such that $y[w](S) \in \mathfrak{T}$ and $y[w] \in \mathcal{O}$. The non-occurrence of infimum gaps means that the original infimum value is unaffected by the introduction of the extended controls, namely

$$h(y[\hat{w}](S)) = \inf\Big\{h(y[v](S)) \quad | \quad v \in \mathcal{V}, \quad y[v](S) \in \mathfrak{T},\ y[v] \in \mathcal{O}\Big\}$$

for all sufficiently small neighbourhoods $\mathcal{O}$ of $y[\hat{w}]$.

If, on the contrary, there exists a neighbourhood $\mathcal{O}$ such that

$$h(y[\hat{w}](S)) < \inf\Big\{h(y[v](S)) \quad | \quad v \in \mathcal{V}, \quad y[v](S) \in \mathfrak{T},\ y[v] \in \mathcal{O}\Big\},$$

one says that the optimal control problem *satisfies the infimum gap condition* (see Def. 3.2). (Obviously, via the usual reductions, one can formulate a notion of infimum gap for a general Bolza problem as well).

For problems defined on Euclidean spaces and such that the extended dynamics is convex, an insightful investigation of the gap question and its relation with normality was carried out by J.Warga (see e.g. [48]). More recently, two specific classes of domain extensions have been studied in [33, 37, 38, 39]. As mentioned above, these investigations share the fact that the following necessary condition turns out to be valid:

(A) *There is an infimum-gap only if the minimum of the extended problem is an abnormal extremal.* [1]

Since 'extremal' means 'satisfying the thesis of the Maximum Principle', in order for **(A)** to have a precise meaning, one has to specify which kind of *approximating cones* we are going to utilize for both the *reachable set* and the target $\mathfrak{T}$. For this purpose, we shall introduce a generalized differential called *Quasi Differential Quotient* (QDQ) (Def. 2.3) [2] and the associated notion of QDQ *approximating cone* (Def. 2.5). While it is impossible at this stage to give an exhaustive description of what QDQ approximating cones are, let us point out that, on the one hand, they are *sufficiently small* for a certain open mapping theorem to hold true and, on the other hand, they are *large enough* to allow the utilization of the notion of *abundance* (of $\mathcal{V}$ in $\mathcal{W}$), which, as we shall see, is crucial to prove that normality implies the absence of gaps.

This said, let us give the precise notions of *normal and abnormal extremal.* For simplicity, we consider here only the case when the state ranges on a Euclidean space. Moreover, if $C \subset \mathbb{R}^n$ is a cone, we use $C^\perp$ to denote the *polar cone of* C, namely the set of linear forms $\lambda \in (\mathbb{R}^n)^*$ such that $\lambda \cdot c \leq 0$ for all $c \in C$.

Definition 1.1 (Extremal)**.** *Consider a control $\hat{w} \in \mathcal{W}$ and the corresponding trajectory $\hat{y} := y[\hat{w}]$. Assume that $\hat{y}(S) \in \mathfrak{T}$, and let C be a QDQ approximating cone of the target $\mathfrak{T}$ at $\hat{y}(S)$. We say that the process $(\hat{y}, \hat{w})$ is an* extremal *(with respect to h and C) if there exist an absolutely continuous (adjoint) path $\lambda \in W^{1,1}([0, S]; (\mathbb{R}^n)^*)$ and a 'cost multiplier' $\lambda_c \in \{0, 1\}$ such that the following conditions are verified:*

(i) $(\lambda, \lambda_c) \neq 0$;

(ii) $$\frac{d\lambda}{ds} = -\lambda \cdot \frac{\partial f}{\partial y}(s, \hat{y}(s), \hat{w}(s))$$

(iii) $$\max_{\mathfrak{w} \in \mathfrak{W}} \lambda(s) \cdot f(s, \hat{y}(s), \mathfrak{w}) = \lambda(s) \cdot f(s, \hat{y}(s), \hat{w}(s)) \quad a.e.\ s \in [0, S];$$

(iv) $\lambda(S) \in -\lambda_c \nabla h(\hat{y}(S)) - C^\perp$.

Furthermore, we say that an extremal $(\hat{y}, \hat{w})$ is normal *if for every choice of the pair (λ, λ_c) one has $\lambda_c = 1$. We say that an extremal $(\hat{y}, \hat{w})$ is* abnormal *if it is not normal, namely, if there exists a choice of (λ, λ_c) with $\lambda_c = 0$.*

---

[1] Equivalently: if the minimum is normal (=not abnormal) there is no gap.

[2] A QDQ is a special case of Sussmann's Approximate Generalized Differential Quotient [45].

The validity of criterion **(A)** was proven for two specific cases:

- when (the dynamics is bounded and) the original set of controls $\mathcal{V}$ is embedded in the set $\mathcal{W}$ of *relaxed* controls ([37, 38, 39]);
- when the system is control-affine and the original set $\mathcal{V}$ comprises *unbounded* controls ranging in a convex cone ([33]). In this case, a space-time, impulsive, extension is considered, namely the larger set of trajectories corresponding to $\mathcal{W}$ comprises space-time paths which are allowed to evolve along fixed time directions.[3]

It is worth noticing that, in all previously investigated cases, the original set of trajectories is *dense* in the set of extended trajectories, when the latter is endowed with C^0 topology. So, one might conjecture that criterion **(A)** is true as soon as the trajectories corresponding to $\mathcal{V}$ are dense in the set of trajectories corresponding to $\mathcal{W}$. In fact, *this is not the case*, as shown by the simple example in Section 9.

Hence, a condition stronger than density is needed. For this goal we recall Kaskosz' formulation of Warga's notion of $\mathcal{V}$*-abundance in* $\mathcal{W}$ (Def. 4.1). This condition strengthens density by requiring the trajectories of the extended system's convexification to be uniformly approachable by trajectories of the original system. We further generalize the notion of $\mathcal{V}$ being abundant in $\mathcal{W}$ to control systems defined on Riemannian manifolds and to fairly general classes of controls (which are merely required to belong to a metric space). Then, aiming to express normality of extended trajectories in geometric terms, we invoke local set separation of the target from the original reachable set.

A crucial result for the achievement of the main theorem consists in showing that, under the abundance hypothesis, *every needle-variational cone* $\mathbf{C}$ *at* $\hat{y}$ *corresponding to the enlarged domain* $\mathcal{W}$ *is also a QDQ approximating cone to the original reachable set* (Theorem 4.1).

The next step consists in showing that the local set separation of the target from the original reachable set implies the linear separability between a QDQ approximating cone to the target and the above mentioned needle-variational cone $\mathbf{C}$ (Theorem 5.1). This is exactly the point where the choice of QDQ approximating cones –rather than other more classical cones, e.g. Boltiansky cones– plays essential. By expressing this linear separation in terms of adjoint paths, one finally gets the main result of the paper (Corollary 5.2), where, under the abundance hypothesis, statement **(A)** is turned into an actual theorem. Finally, since normality cannot be verified *a priori*, in Theorem 6.1 we provide a sufficient condition on the data guaranteeing that a given extremal is normal.

In Section 8 we provide an application of the main theorem to nonlinear systems whose dynamics are neither bounded nor convex.

[3] It is well-known that, under commutativity hypotheses, the extended dynamics could be regarded as a measure, while such a measure-theoretical approach is unfit for general, non-commutative problems, see e.g. [10],[29].

1.1. Basic notions and notations

1.1.1. Linear spaces, manifolds

Let E be a real linear space, and let us use E^* to denote the algebraic dual of E. If $\langle\cdot,\cdot\rangle$ is a given scalar product on E, we will use $|\cdot|$ to denote the norm associated with $\langle\cdot,\cdot\rangle$, namely, for every $e \in E$ we set $|e| = \sqrt{\langle e, e\rangle}$. For every $e \in E$ and every real number $r \geq 0$ let us use $e + B_r$ to denote the closed ball of center e and radius r, namely $e + B_r = \{e + f \mid |f| \leq r\}$. When $e = 0$ we will write B_r instead of $0 + B_r$

If E_1, E_2 are real linear spaces, $e_1 \in E_1$ we shall use $Lin\{E_1, E_2\}$ to denote the set of linear maps from E_1 to E_2. If $L \in Lin\{E_1, E_2\}$, we shall use $L \cdot e_1$ to denote the image of e_1. We will use the symbol $\cdot$ also to mean duality. Furthermore, if $\lambda \in E_2^*$ and and $L \in Lin\{E_1, E_2\}$, sometimes we will use the notation $\lambda \cdot L$ to mean *the element of E_1^* coinciding with the image of λ through the dual map of L.* While this doesn't generate any confusion, it makes the writing $\lambda \cdot L \cdot e$ unambiguous, for one has $(\lambda \cdot L) \cdot e = \lambda \cdot (L \cdot e)$ for all $(e, \lambda) \in E_1 \times E_2^*$.

For any integer $r \geq 0$, by saying that $\Big(\mathcal{M}, \langle\cdot,\cdot\rangle\Big)$ is a Riemannian differentiable manifold of class C^{r+1} we will mean that $\mathcal{M}$ is a C^{r+1} differential manifold and $\langle\cdot,\cdot\rangle$ is a C^r Riemannian metric. For every $x \in \mathcal{M}$ and $e, f \in T_x\mathcal{M}$, $\langle e, f\rangle_x$ will denote the corresponding scalar product of e, f, and $|e|_x := \sqrt{\langle e, e\rangle_x}$ will be called the *norm* of e. We will often omit the subscript and we will write $\langle e, f\rangle$ and $|e|$ instead of $\langle e, f\rangle_x$ and $|e|_x$.

We will use d to denote the distance induced on $\mathcal{M}$ by $\langle\cdot,\cdot\rangle$. We recall that, if $x_1, x_2 \in \mathcal{M}$, the *distance* $d(x_1, x_2)$ is defined as the minimum among the $\langle\cdot,\cdot\rangle$-lengths of the absolutely continuous curves having x_1, x_2 as end-points.

1.1.2. Cones

Let E be a real linear space. A subset $K \subset E$ is a *cone* if $\alpha k \in K$ for all $(\alpha, k) \in [0, +\infty[\times K$. If $A \subset E$ is any subset, we use $span^+ A$ to denote the smallest convex cone containing A. Let us introduce a notion of *transversality* for cones (see [45]).

Definition 1.2. *Let E be a linear space and let $K_1, K_2 \subseteq E$ be convex cones. We say that*

1. *K_1 and K_2 are* transverse, *if $K_1 - K_2 := \Big\{k_1 - k_2,\ (k_1, k_2) \in K_1 \times K_2\Big\} = E$;*

2. *K_1 and K_2 are* strongly transverse, *if they are transverse and $K_1 \cap K_2 \supsetneq \{0\}$.*

Transversality differs from strong transversality only when K_1 and K_2 are complementary subspaces. Indeed:

Proposition 1.1. *Let E be a linear space, and let $K_1, K_2 \subseteq E$ be convex cones. Then K_1, K_2 are transverse if and only if either K_1, K_2 are strongly transverse or K_1, K_2 are complementary linear subspaces, namely $K_1 + K_2 = E$ and $K_1 \cap K_2 = \{0\}$.*

Definition 1.3. *Let E be a finite-dimensional linear space, and let E^* be its dual space. For any subset $A \subset E$, the (convex) cone $A^\perp \subset E^*$ defined as*

$$A^\perp \doteq \{p \in E^* : \quad p \cdot w \leq 0 \quad \forall\, w \in A\}$$

will be called the polar cone *of A.*

The transversality of two cones is equivalent to their linearly separability. More precisely:

Proposition 1.2. *Two convex cones K_1 and K_2 are not transverse if and only if $(-K_1)^\perp \cap K_2^\perp \backslash \{0\} \neq \emptyset$, namely there exists a linear form $\lambda \neq 0$ such that*

$$\lambda \cdot k_1 \geq 0 \;\forall k_1 \in K_1 \qquad \text{and} \quad \lambda \cdot k_2 \leq 0 \;\forall k_2 \in K_2.$$

In this case one also says that K_1 and K_2 are linearly separable.

1.1.3. Scorza-Dragoni points

Definition 1.4 (Scorza-Dragoni point)**.** *Given a compact set $X \subset \mathcal{M}$ and an interval $[a, b] \subseteq \mathbb{R}$, $a < b$, let us consider a function $\varphi : [a, b] \times X \to \mathbb{R}^n$ verifying*

i) $[a, b] \ni s \mapsto \varphi(s, y) \in \mathbb{R}^n$ *is measurable for each* $y \in X$;

ii) $X \ni y \mapsto \varphi(s, y)$ *is continuous for each* $s \in [a, b]$,

A function φ satisfying conditions i)-ii) is said to be a Carathéodory function. We say that $\bar{s} \in [a, b]$ is a Scorza-Dragoni point for φ if, for all $y \in X$,

$$\lim_{r \to 0} \lim_{\delta \to 0} \frac{1}{\delta} \int_{\bar{s}}^{\bar{s}+\delta} \Lambda_r(s, y)\, ds = 0 \tag{1.1}$$

where

$$\Lambda_r(s, y) := \sup_{x \in X,\, d(x,y) \leq r} |\varphi(s, x) - \varphi(\bar{s}, y)| \tag{1.2}$$

We shall use SD $\{\varphi\}$ *to denote the* set of all the Scorza-Dragoni points for the function φ.

Notice in particular that, if $s \in \mathrm{SD}\{\varphi\}$, one has $\lim\limits_{\substack{x \to y \\ \delta \searrow 0}} \varphi(s + \delta, x) = \varphi(s, y)$, for any $y \in X$. Let us also mention that, when φ is independent of y and $s \mapsto \varphi(s)$ is integrable on $[\bar{s}, \bar{s} + \delta]$, the definition of Scorza-Dragoni point reduces to the definition of *Lebesgue point.* Namely, relations (1.1)-(1.2) become

$$\lim_{\delta \to 0} \frac{1}{\delta} \int_{\bar{s}}^{\bar{s}+\delta} |\varphi(s) - \varphi(\bar{s})|\, ds = 0. \tag{1.3}$$

We shall use $\mathcal{L}\{\varphi\}$ to denote the set of Lebesgue points of φ.

The importance of Scorza-Dragoni points relies on the fact that they form a full measure set [41]:

Theorem 1.1 (Scorza-Dragoni)**.** *The set of all the Scorza-Dragoni points of a Carathéodory function $\varphi : [a, b] \times X \to \mathbb{R}^n$ has measure equal to $b - a$.*

2. Set separation and open mappings

2.1. Quasi Differential Quotients

In the statement of the set-separation theorem (Th. 2.3), we will make use of the notions of Quasi Differential Quotient (QDQ) and of the corresponding approximating cone to a set. A QDQ is a particular case of Sussmann's Approximate Generalized Differential Quotient (AGDQ) [45]. The related set-separation theorem (Theorem 2.3) is based on an open mapping result provided by Theorem 2.2 below.

Let us recall the notion of *Cellina continuously approximable (CCA) set-valued function*:

Definition 2.1 (CCA). *Let $F : \mathbb{R}^N \rightsquigarrow \mathbb{R}^n$ be a set-valued map. We say that F is a* Cellina continuously approximable (CCA) set-valued map *if, for any compact set $K \subset \mathbb{R}^N$:*

- *the restriction of F on K has compact graph, that is, the set $\mathrm{Gr}(F_{|_K}) := \{(x, y) \in K \times \mathbb{R}^n : \ y \in F(x)\}$ is compact, and*
- *there exists a sequence of single-valued, continuous maps $f_k : K \to \mathbb{R}^n$, $k \in \mathbb{N}$, such that the following condition holds: for every open set $\Omega \subseteq \mathbb{R}^N \times \mathbb{R}^n$ satisfying $\mathrm{Gr}(F_{|_K}) \subset \Omega$, there exists k_Ω such that $\mathrm{Gr}\,(f_k) := \{(x, y) \in K \times \mathbb{R}^n : \ \ y \in f_k(x)\} \subset \Omega$ for every $k \geq k_\Omega$.*

We will say that a function $\rho : [0, +\infty[\to [0, +\infty]$ is a *a pseudo-modulus* if it is monotonically nondecreasing and $\lim_{s \to 0^+} \rho(s) = \rho(0) = 0$. We call *modulus* a pseudo-modulus taking values in $[0, +\infty[$.

Definition 2.2 (AGDQ). *Assume that $F : \mathbb{R}^N \rightsquigarrow \mathbb{R}^n$ is a set-valued map, $(\bar{\gamma}, \bar{y}) \in \mathbb{R}^N \times \mathbb{R}^n$, $\Lambda \subset Lin\{\mathbb{R}^N, \mathbb{R}^n\}$ is a compact set, and $\Gamma \subset \mathbb{R}^N$ is any subset. We say that Λ is an* Approximate Generalized Differential Quotient (AGDQ) of F at $(\bar{\gamma}, \bar{y})$ *in the direction of* Γ *if there exists a pseudo-modulus ρ having the property that*

(*) for every $\delta > 0$ such that $\rho(\delta) < +\infty$, there exists a CCA set-valued map $A^\delta : (\bar{\gamma} + B_\delta) \cap \Gamma \rightsquigarrow Lin\{\mathbb{R}^N, \mathbb{R}^n\} \times \mathbb{R}^n$ such that

$$\inf_{L' \in \Lambda} |L - L'| \leq \rho(\delta), \quad |h| \leq \delta\rho(\delta), \quad \textit{and } \bar{y} + L \cdot (\gamma - \bar{\gamma}) + h \in F(\gamma) \quad {}^{4}$$

whenever $\gamma \in (\bar{\gamma} + B_\delta) \cap \Gamma$ and $(L, h) \in A^\delta(\gamma)$.

We now introduce a subclass of AGDQs, which we call *Quasi Differential Quotients*. Their main property consists in the validity of an actual, *not punctured*, open mapping theorem (see Theorem 2.2 below).

[4] Here $|\cdot|$ denotes the operator norm, namely $|M| = \sup\limits_{|v|=1} |M \cdot v|$, for every linear operator $M \in Lin(\mathbb{R}^N, \mathbb{R}^n)$.

Definition 2.3 (QDQ). *Assume that* $F : \mathbb{R}^N \rightsquigarrow \mathbb{R}^n$ *is a set-valued map,* $(\bar{\gamma}, \bar{y}) \in \mathbb{R}^N \times \mathbb{R}^n$, $\Lambda \subset Lin\{\mathbb{R}^N, \mathbb{R}^n\}$ *is a compact set, and* $\Gamma \subset \mathbb{R}^N$ *is any subset. We say that* Λ *is a* Quasi Differential Quotient (QDQ) of F at $(\bar{\gamma}, \bar{y})$ *in the direction of* Γ *if there exists a modulus* $\rho : [0, +\infty[\to [0, +\infty[$ *having the property that*

(*) for every $\delta > 0$ *there is a continuous map* $(L_\delta, h_\delta) : (\bar{\gamma} + B_\delta) \cap \Gamma \to Lin\{\mathbb{R}^N, \mathbb{R}^n\} \times \mathbb{R}^n$ *such that*

$$\min_{L' \in \Lambda} |L_\delta(\gamma) - L'| \le \rho(\delta), \quad |h_\delta(\gamma)| \le \delta\rho(\delta), \quad \textit{and } \bar{y} + L_\delta(\gamma) \cdot (\gamma - \bar{\gamma}) + h_\delta(\gamma) \in F(\gamma),$$

whenever $\gamma \in (\bar{\gamma} + B_\delta) \cap \Gamma$ *.*

Clearly a (QDQ) is a (AGDQ) as well.

Definition 2.4 (AGDQ and QDQ on manifolds). *Let* $\mathcal{N}, \mathcal{M}$ *be* C^1 *Riemannian manifolds. Assume that* $\tilde{F} : \mathcal{N} \rightsquigarrow \mathcal{M}$ *is a set-valued map,* $(\bar{\gamma}, \bar{y}) \in \mathcal{N} \times \mathcal{M}$, $\tilde{\Lambda} \subset Lin\{T_\gamma\mathcal{N}, T_y\mathcal{M}\}$ *is a compact set, and* $\Gamma \subset \mathcal{N}$ *is any subset. Moreover, let* $\phi : U \to \mathbb{R}^N$ *and* $\psi : V \to \mathbb{R}^n$ *be charts defined on neighbourhoods* U *and* V *of* $\bar{\gamma}$ *and* $\bar{y}$*, respectively, and assume that* $\phi(\bar{\gamma}) = 0$, $\psi(\bar{y}) = 0$*. Consider the map* $\psi \circ \tilde{F} \circ \phi^{-1} : \phi(U) \to \mathbb{R}^n$ *and extend it arbitrarily to a map* $F : \mathbb{R}^N \to \mathbb{R}^n$*. We say that* $\tilde{\Lambda}$ *is an* Approximate Generalized Differential Quotient (AGDQ) [resp. a Quasi Differential Quotient (QDQ)] of $\tilde{F}$ at $(\bar{\gamma}, \bar{y})$ *in the direction of* $\tilde{\Gamma}$ *if* $\Lambda := D\psi(\bar{y}) \circ \tilde{\Lambda} \circ D\phi^{-1}(0)$ *is an Approximate Generalized Differential Quotient [resp. a Quasi Differential Quotient] of* F *at* $(0, 0)$ *in the direction of* $\Gamma := \phi(\tilde{\Gamma} \cap U)$*.*

As pointed out in [45], this definition is intrinsic, that is, it is independent of the choice of the charts ϕ and ψ.

2.2. Open Mapping results

Let us recall a directional open mapping result for AGDQ's ([45]).

Theorem 2.1 (Directional Open Mapping). *Let* N, n *be positive integers, and let* Γ *be a convex cone in* $\mathbb{R}^N$*. Let* $F : \mathbb{R}^N \rightsquigarrow \mathbb{R}^n$ *be a set-valued map, and let* Λ *be a AGDQ of* F *at* $(\bar{\gamma}, \bar{y})$ *in the direction of* Γ*. Let us assume that there is an element* $\bar{w} \in \mathbb{R}^n$ *such that* $\bar{w} \in Int(L \cdot \Gamma)$ *for every* $L \in \Lambda$*. Then there exist a closed convex cone* $D \subseteq \mathbb{R}^n$ *and positive constants* α, β *verifying* $\bar{w} \in Int(D)$ *and*

$$\bar{y} + \big(B_a \backslash \{0\} \cap D\big) \subset F(\bar{\gamma} + (B_{a\beta} \cap \Gamma)) \quad \textit{for all} \quad a \in]0, \alpha]. \tag{2.4}$$

If one takes $\bar{w} = 0$ in the statement of Theorem 2.1, the cone D necessarily coincides with the whole $\mathbb{R}^n$. As a consequence, one obtains the following 'punctured' Open Mapping Theorem.

Corollary 2.1 ('Punctured' Open Mapping). [5] *Let* N, n *be positive integers, and let* Γ *be a convex cone in* $\mathbb{R}^N$*. Let* $F : \mathbb{R}^N \rightsquigarrow \mathbb{R}^n$ *be a set-valued map, and let* Λ *be an QDQ of* F *at* $(\bar{\gamma}, \bar{y})$

[5] The adjective *punctured* here refers to the fact that $\bar{y}$ does not belong to $F(\bar{\gamma} + (B_{a\beta} \cap \Gamma))$.

in the direction of Γ. *Let us assume that* Λ *is* surjective, *by which we mean that* $L\cdot\Gamma=\mathbb{R}^n$ *for every* $L\in\Lambda$. *Then there are positive constants* α,β *verifying*

$$\bar{y}+\big(B_a\backslash\{0\}\big)\subset F(\bar{\gamma}+(B_{a\beta}\cap\Gamma))\quad \textit{for all}\quad a\in]0,\alpha]. \tag{2.5}$$

By further assuming that Λ is a QDQ (rather then a mere AGDQ), we get an actual, non-punctured, open mapping result:

Theorem 2.2 (Open Mapping)**.** *Let* N,n *be positive integers, and let* Γ *be a convex cone in* $\mathbb{R}^N$. *Let* $F:\mathbb{R}^N\rightsquigarrow\mathbb{R}^n$ *be a set-valued map, and let* Λ *be a QDQ of* F *at* $(\bar{\gamma},\bar{y})$ *in the direction of* Γ. *As in Corollary 2.1, let us assume that* Λ *is* surjective, *by which we mean that* $L\cdot\Gamma=\mathbb{R}^n$ *for every* $L\in\Lambda$. *Then the following statements hold true:*

(*i*) *there are positive constants* α,β *having the property that*

$$\bar{y}+\big(B_a\backslash\{0\}\big)\subset F(\bar{\gamma}+(B_{a\beta}\cap\Gamma))\quad \textit{for all}\quad a\in]0,\alpha]; \tag{2.6}$$

(*ii*) *there exists* $\check{\delta}>0$ *such that, for every* $\delta\leq\check{\delta}$ *and every* (L_δ,h_δ) *as in Definition 2.3, there exists* $\gamma_\delta\in\bar{\gamma}+(\Gamma\cap B_\delta)$ *such that*

$$\bar{y}=\bar{y}+L_\delta(\gamma_\delta)\cdot(\gamma_\delta-\bar{\gamma})+h_\delta(\gamma_\delta)\quad\Big[\in F(\gamma_\delta)\Big]. \tag{2.7}$$

In particular, by possibly reducing the size of α, *one gets the open-mapping inclusions*

$$\bar{y}+B_a\subset F(\bar{\gamma}+(B_{a\beta}\cap\Gamma))\quad \textit{for all}\quad a\in]0,\alpha].$$

Proof. Without loss of generality, we can assume $(\bar{\gamma},\bar{y})=(0,0)$. Furthermore, since a QDQ is an AGDQ, in view of Theorem 2.1, it is sufficient to prove only statement (*ii*). Namely, for every $\delta>0$ sufficiently small, we have to establish the existence of a $\gamma_\delta\in B_\delta\cap\Gamma$ such that

$$0=L_\delta(\gamma_\delta)\cdot\gamma_\delta+h_\delta(\gamma_\delta). \tag{2.8}$$

For every $\delta>0$, let us define the set-valued map $L_\delta^{-1_r}:B_\delta\cap\Gamma\rightsquigarrow Lin(\mathbb{R}^n,\mathbb{R}^N)$ by setting, for every $\gamma\in B_\delta\cap\Gamma$,

$$L_\delta^{-1_r}(\gamma):=\Big\{M\in Lin(\mathbb{R}^n,\mathbb{R}^N),\ L_\delta(\gamma)\circ M=Id_{\mathbb{R}^n}\Big\}.$$

Namely, $L_\delta^{-1_r}(\gamma)$ is the set of *right inverse of* $L_\delta(\gamma)$. Let us first observe that, for every $\gamma\in B_\delta\cap\Gamma$, and δ sufficiently small, $L_\delta^{-1_r}(\gamma)$ is non-empty. Indeed, since $L\in\Lambda$ and Λ is surjective, $L_\delta^{-1_r}(\gamma)$ contains the *Moore-Penrose pseudo-inverse*

$$M_\delta^\sharp(\gamma):=L_\delta^{tr}(\gamma)\circ\big(L_\delta(\gamma)\circ L_\delta^{tr}(\gamma)\big)^{-1}, \tag{2.9}$$

where ^{tr} denotes transposition. Furthermore, it is trivial to verify that *the set-valued map* $L_\delta^{-1_r}$ *is convex-valued.* Finally, by possibly reducing the size of $\check{\delta}$, for every $\delta\in[0,\check{\delta}]$, *the* *set-valued map $L_\delta^{-1_r}$ has compact graph.* Indeed, there exist a constant $K > 0$ such that $\Lambda^{\rho(\delta)}$ is a compact subset made of linear operators whose right inverse are bounded (in the operator norm) by K. Moreover, let us consider a sequence $(\gamma_m)_{m\in\mathbb{N}} \subset B_\delta \cap \Gamma$ converging to $\tilde{\gamma} \in B_\delta \cap \Gamma$, and, for every $m \in \mathbb{N}$, let us choose $M_m \in L_\delta^{-1_r}(\gamma_m)$. Hence, one has that $L_\delta(\gamma_m) \circ M_m = Id_{\mathbb{R}^n}$ and, since the sequence (M_m) ranges in a compact set, there exists a subsequence $(M_{m_k})_k$ converging to a linear operator $\tilde{M}$. In particular,

$$L_\delta(\tilde{\gamma}) \circ \tilde{M} = \lim_{k\to\infty} \Big(L_\delta(\gamma_{m_k}) \circ M_{m_k}\Big) = Id_{\mathbb{R}^n},$$

so that $\tilde{M} \in L_\delta^{-1_r}(\tilde{\gamma})$. This proves that the set-valued map $\gamma \mapsto L_\delta^{-1_r}(\gamma)$ has compact graph.

Now consider the set-valued map $\Psi_\delta : B_\delta \cap \Gamma \rightsquigarrow \mathbb{R}^N$ defined by setting

$$\Psi_\delta(\gamma) := \Big\{ - M \cdot h_\delta(\gamma) \ | \quad M \in L_\delta^{-1_r}(\gamma)\Big\} \bigcap \Gamma, \quad \gamma \in B_\delta \cap \Gamma.$$

To prove that this map has non-empty values for every $\gamma \in B_\delta \cap \Gamma$, it is sufficient to determine a linear mapping $M^\flat : \mathbb{R}^n \to \mathbb{R}^N$ and an element $v \in \Gamma$ such that

$$(L_\delta(\gamma) \circ M^\flat) \cdot w = w \quad \forall w \in \mathbb{R}^n \quad \Big(\iff M^\flat \in L_\delta^{-1_r}(\gamma)\Big), \qquad -M^\flat \cdot h_\delta(\gamma) = v \tag{2.10}$$

Fix $\gamma \in B_\delta \cap \Gamma$ and choose $v \in \Gamma$ verifying $L_\delta(\gamma) \cdot v = -h_\delta(\gamma)$. Such a v exists, since $L_\delta(\gamma)$ is surjective. Now, a geometrical intuition suggests that $M^\flat$ might be obtained by adding a suitable linear operator to an element of $L_\delta^{-1_r}(\gamma)$, for instance the Moore-Penrose inverse $M_\delta^\sharp$ defined in (2.9). Actually, following [45], if $\langle\cdot,\cdot\rangle$ is any scalar product on $\mathbb{R}^n$, we define the linear map $M^\flat : \mathbb{R}^n \to \mathbb{R}^N$ by setting, for every $w \in \mathbb{R}^n$,

$$M^\flat \cdot w := M_\delta^\sharp \cdot w - \frac{\langle w, h_\delta(\gamma)\rangle}{\langle h_\delta(\gamma), h_\delta(\gamma)\rangle} \Big(v + M_\delta^\sharp \cdot h_\delta(\gamma)\Big).$$

It is straightforward to verify that $M^\flat$ verifies conditions (2.10), so that $\Psi_\delta(\gamma)$ is not empty.

Since for every δ the map h_δ is continuous and $|h_\delta(\gamma)| \le \delta\rho(\delta)$ for all $\gamma \in B_\delta \cap \Gamma$, by possibly reducing further the size of $\check{\delta}$ we conclude that, for every $\delta \in [0, \check{\delta}]$, the set-valued map Ψ_δ verifies $\Psi_\delta(B_\delta \cap \Gamma) \subset \bar{B}_\delta \cap \Gamma$ and has non-empty, convex values, and a closed graph. Since the domain of Ψ_δ is compact and convex, the set-valued map Ψ_δ verifies the hypotheses of the Kakutani fixed point theorem, so that there exists $\gamma_\delta \in B_\delta \cap \Gamma$ such that $\gamma_\delta \in \Psi_\delta(\gamma_\delta)$. It follows that there is a matrix $M \in L_\delta^{-1_r}(\gamma_\delta)$ such that $0 = \gamma_\delta + M \cdot h_\delta(\gamma_\delta)$. Therefore, one gets

$$0 = L_\delta(\gamma_\delta) \cdot \Big(\gamma_\delta + M \cdot h_\delta(\gamma_\delta)\Big) = L_\delta(\gamma_\delta) \cdot \gamma_\delta + h_\delta(\gamma_\delta),$$

which concludes the proof.

$\square$

2.3. QDQ approximating cones and set separation

Assume that $\mathcal{M}$ is a C^1 differentiable manifold, $\mathcal{E} \subset \mathcal{M}$, and $z \in \mathcal{E}$. If X is a linear space, let us call *convex multicone* in X any family of convex cones of X.

We now define a subfamily of Sussmann's AGDQ approximating multicones [45], which we call *QDQ approximating multicones.*

Definition 2.5. *An* AGDQ approximating multicone [*resp. a* QDQ approximating multicone] *to $\mathcal{E}$ at z is a convex multicone $\mathcal{C} \subseteq T_z\mathcal{M}$ such that there exist a non-negative integer N, a set-valued map $F : \mathbb{R}^N \rightsquigarrow \mathcal{M}$, a convex cone $\Gamma \subset \mathbb{R}^N$, and an AGDQ* [*resp. a* QDQ] *Λ of F at $(0, z)$ in the direction of Γ such that $F(\Gamma) \subset \mathcal{E}$ and $\mathcal{C} = \{L \cdot \Gamma \ : L \in \Lambda\}$.*

In the particular case when an AGDQ approximating multicone [resp. a QDQ approximating multicone] *is a singleton, namely $\Lambda = \{L\}$ for some $L \in Lin(\mathbb{R}^N, \mathbb{R}^n)$, we say that $C := L \cdot \Gamma$ is an* AGDQ approximating cone [resp. a QDQ approximating cone] *to $\mathcal{E}$ at z.*

Let us introduce the notion of local set-separation:

Definition 2.6. *Let $\mathcal{X}$ be a topological space, and let us consider two subsets $\mathcal{A}_1, \mathcal{A}_2 \subset \mathcal{X}$ and a point $z \in \mathcal{A}_1 \cap \mathcal{A}_2$. We say that $\mathcal{A}_1$ and $\mathcal{A}_2$ are* locally separated *at z provided there exists a neighborhood V of z such that*

$$\mathcal{A}_1 \cap \mathcal{A}_2 \cap V = \{z\}.$$

We are now ready to state our set-separation result, which connects set separation with the linear separability of QDQ approximating cones. Furthermore, the result includes a crucial special approximation property (see *ii*) in Theorem 2.3) for the case in which the approximating cones are complementary linear subspaces.[6]

Theorem 2.3 (Set separation)**.** *Let $\mathcal{E}_1, \mathcal{E}_2$ be subsets of $\mathcal{M}$, and let $z \in \mathcal{E}_1 \cap \mathcal{E}_2$ Assume that C_1, C_2 are AGDQ approximating cones of $\mathcal{E}_1$ and $\mathcal{E}_2$, respectively, at z.*

i) If C_1 and C_2 are strongly transverse, then the sets $\mathcal{E}_1$ and $\mathcal{E}_2$ are not locally separated.

ii) If, moreover,

1. *C_1, C_2 are QDQ cones,*
2. *C_1 and C_2 are complementary linear subspaces, i.e. $C_1 + C_2 = T_z\mathcal{M}$ and $C_1 \cap C_2 = \{0\}$,*
3. *for each $i = 1, 2$, N_i is a non-negative integer, $\Gamma_i \subset \mathbb{R}^{N_i}$ is a convex cone, $F_i : \mathbb{R}^{N_i} \rightsquigarrow \mathcal{M}$ is a set-valued map, and $\Lambda_i = \{L^i\} \in Lin(\mathbb{R}^{N_i}, T_z\mathcal{M})$ is a QDQ of F_i at $(0, z)$ in the direction of Γ_i, $F_i(\Gamma_i) \subseteq \mathcal{E}_i$ and $C_i = L^i \cdot \Gamma_i$,*

then there exists a sequence $(\gamma_{1_k}, \gamma_{2_k}) \in \Gamma_1 \times \Gamma_2$ such that $z_k \in F_1(\gamma_{1_k}) \cap F_2(\gamma_{2_k})$ and $z_k \to z$.

[6]We recall that this is the only case when non-transversality differs from non-strong-transversality.

Remark 2.1. Property ii), whose proof is based on the Open Mapping result stated in Theorem 2.2, is not true if we replace QDQ approximating cones with AGDQ approximating cones. Of course, this is connected with the non validity of a non-punctured open mapping result for AGDQs.

Proof of Theorem 2.3. Statement i) of Theorem 2.3 is direct consequence of [45], Theorem 4.37, where an analogous result concerning the non-separation of multicones is provided.

Let us prove statement ii). Because of the local character of the statement, there is not loss of generality in considering only the Euclidean case when $\mathcal{M} = \mathbb{R}^n$. For every $i = 1, 2$, let $n_i \geq 0$ be the dimensions of the subspace C_i, so that $n_1 + n_2 = n$. By hypothesis, for every $i = 1, 2$ there exists a modulus $\rho_i : [0, +\infty[\to [0, +\infty[$ having the property that, for every $\delta > 0$, there exists a continuous map $(L^i_\delta, h^i_\delta) : B_\delta \cap \Gamma_i \to Lin\{\mathbb{R}^{N_i}, \mathbb{R}^{n_i}\} \times \mathbb{R}^{n_i}$, such that

$$|L^i_\delta(\gamma_i) - L^i| \leq \rho_i(\delta), \quad |h^i_\delta| \leq \delta \cdot \rho_i(\delta), \quad \text{and } z + L^i_\delta(\gamma_i) \cdot \gamma_i + h^i_\delta(\gamma_i) \in F_i(\gamma_i)$$

whenever $\gamma_i \in B_\delta \cap \Gamma_i$. Let us consider the cone $\Gamma := \Gamma_1 \times \Gamma_2 \subset \mathbb{R}^{N_1+N_2}$ and the set-valued map $F : \Gamma \rightsquigarrow \mathbb{R}^n$ defined by setting

$$F(\gamma_1, \gamma_2) := F_2(\gamma_2) - F_1(\gamma_1) = \Big\{ z_2 - z_1 \mid (z_1, z_2) \in F_1(\gamma_1) \times F_2(\gamma_2) \Big\} \quad \forall (\gamma_1, \gamma_2) \in \Gamma_1 \times \Gamma_2,$$

and observe that

$$\textit{if } (\bar\gamma_1, \bar\gamma_2) \textit{ is such that } 0 \in F(\bar\gamma_1, \bar\gamma_2) \textit{ then } \emptyset \neq F_2(\bar\gamma_2) \cap F_1(\bar\gamma_1) \subseteq \mathcal{E}_2 \cap \mathcal{E}_1.$$

Furthermore, let us set $\rho(\delta) := \rho_1(\delta) + \rho_2(\delta)$ and let us define the continuous map

$$(L_\delta, h_\delta)(\gamma_1, \gamma_2) := \Big(\big(-L^1_\delta(\gamma_1), L^2_\delta(\gamma_2)\big) \ , \ h^2_\delta(\gamma_2) - h^1_\delta(\gamma_1) \Big), \qquad (\gamma_1, \gamma_2) \in B_\delta \cap \Gamma.$$

Defining the linear map $L \in Lin\{\mathbb{R}^{N_1+N_2}, \mathbb{R}^n\}$ by setting $L(v_1, v_2) := L^2 \cdot v_2 - L^1 \cdot v_1$, one obtains $|L_\delta(\gamma_1, \gamma_2) - L| \leq \rho(\delta)$, $|h_\delta(\gamma_1, \gamma_2)| \leq \delta \cdot \rho(\delta)$, and

$$L_\delta(\gamma_1, \gamma_2) \cdot (\gamma_1, \gamma_2) + h_\delta(\gamma_1, \gamma_2) \in F(\gamma_1, \gamma_2).$$

whenever $(\gamma_1, \gamma_2) \in B_\delta \cap \Gamma$. Hence, $\Lambda := \{L\}$ is a QDQ of F at $(0, 0)$. Moreover, one has $L \cdot \Gamma = C_1 + C_2 = \mathbb{R}^n$, so that, by the Open Mapping result stated in Theorem 2.2 for $k \in \mathbb{N}$ sufficiently large, we get the existence of $(\gamma_{1_k}, \gamma_{2_k}) \in \Gamma \cap B_{\frac{1}{k}} \subset \Gamma_1 \times \Gamma_2$ such that

$$z_k := z + L^1_{\frac{1}{k}}(\gamma_{1_k}) \cdot \gamma_{1_k} + h^1_{\frac{1}{k}}(\gamma_{1_k}) = z + L^2_{\frac{1}{k}}(\gamma_{2_k}) \cdot \gamma_{2_k} + h^2_{\frac{1}{k}}(\gamma_{2_k}) \in F_1(\gamma_{1_k}) \cap F_2(\gamma_{2_k}).$$

Notice that, by $h^1_{\frac{1}{k}}(\gamma_{1_k}) \leq \rho_1(\frac{1}{k}) \cdot \frac{1}{k}$, and $|\gamma_{1_k}| \leq \frac{1}{k}$ one has $\lim_{k\to\infty} z_k = z$, which concludes the proof. $\square$

3. Gaps and set-separation

3.1. Original and extended controls

Let $(\mathcal{M}, \langle\cdot,\cdot\rangle)$ be a Riemannian differentiable manifold of class C^2, let $[0,S]$ be a *time*-interval and let $\mathfrak{W}$ be a metric space which we call the set of *control values.* For every $(s,\mathfrak{w}) \in [0,S]\times\mathfrak{W}$, let $\mathcal{M} \ni y \mapsto (y, f(s,y,\mathfrak{w})) \in T\mathcal{M}$ be a vector field. We will consider *two families of controls* $\mathcal{V}$, $\mathcal{W} := L^1([0,S],\mathfrak{W})$, with $\mathcal{V} \subset \mathcal{W}$. We will call $\mathcal{V}$ and $\mathcal{W}$ the *original family of controls* and the *extended family of controls*, respectively.

Let us choose an *initial point* $\bar{y} \in \mathcal{M}$, and, for any *control* map $w \in \mathcal{W}$, let us consider the Cauchy problem

$$(E) \qquad \begin{cases} \dfrac{dy}{ds}(s) = f(s,y(s),w(s)) \quad \text{a.e. } s \in [0,S] \\[2mm] y(0) = \bar{y} \end{cases} .$$

We shall assume the following regularity hypothesis:

Hypothesis (SH) :

(i) for each $(s,\mathfrak{w}) \in [0,S]\times\mathfrak{W}$, the vector field $y \mapsto f(s,y,\mathfrak{w})$ is of class C^1 on $\mathcal{M}$;

(ii) there exists an integrable function $c \in L^1([0,S];\mathbb{R})$ such that, for a.e. $s \in [0,S]$,

$$|f(s,y,\mathfrak{w})| \le c(s), \qquad \left|\frac{\partial f}{\partial y}(s,y,\mathfrak{w})\right| \le c(s) \tag{3.11}$$

for every $(y,\mathfrak{w}) \in \mathcal{M}\times\mathfrak{W}$.

(iii) for every $(y,\mathfrak{w}) \in \mathcal{M}\times\mathfrak{W}$, the map $s \mapsto f(s,y,\mathfrak{w})$ is measurable;

(iv) for every $s \in [0,S]$, the map $(y,\mathfrak{w}) \mapsto f(s,y,\mathfrak{w})$ is continuous.

In particular, for every $w \in \mathcal{W}$ there exists a unique trajectory $y[w]$ of (E).

Let us fix a closed set $\mathfrak{T} \subseteq \mathcal{M}$, which we will refer to as *target.*

Remark 3.1. Of course, through standard cut-off arguments, in many situations one can replace (ii) in hypothesis (SH) with a weaker assumption concerning a neighbourhood of the reference trajectory $s \mapsto \hat{y}(s)$ *instead of the whole state-space* $\mathcal{M}$*.*

Definition 3.1. *For any control* $v \in \mathcal{V}$ *[resp.* $w \in \mathcal{W}$*], the pair* $(y,v) := (y[v],v)$ *[resp.* $(y,w) := (y[w],w)$*] will be called* original process *[resp.* extended processes*]. An extended process –in particular, an original process–* (y,w) *is called* feasible *if* $y(S) \in \mathfrak{T}$*.*

3.2. Infimum gaps

Let us endow the set of controls $\mathcal{W}$ with the pseudo-distance d_f defined by setting

$$d_f(w_1, w_2) := d_\infty(y[w_1], y[w_2]) \left(:= \max_{s\in[0,S]} d(y[w_1](s), y[w_2](s))\right), \tag{3.12}$$

for all controls $w_1, w_2 \in \mathcal{W}$.

The set

$$\mathcal{R}_\mathcal{V} := \Big\{ y[v](S) : \quad v \in \mathcal{V} \Big\} \subset \mathcal{M} \tag{3.13}$$

will be called the *original reachable set*, and the set

$$\mathcal{R}_\mathcal{W} := \Big\{ y[w](S) : \quad w \in \mathcal{W} \Big\} \subset \mathcal{M} \tag{3.14}$$

will be called the *extended reachable set.*

We will also consider local versions of the above reachable sets. Precisely, for a given extended process $(\hat{y}, \hat{w})$ and $r \geq 0$, we set

$$\begin{aligned} \mathcal{R}_\mathcal{V}^{\hat{w},r} &:= \Big\{ y[v](S) : \qquad v \in \mathcal{V}, \quad d_f(\hat{w}, v) < r \Big\} \\ \mathcal{R}_\mathcal{W}^{\hat{w},r} &:= \Big\{ y[w](S) : \qquad w \in \mathcal{W}, \quad d_f(\hat{w}, w) < r \Big\}. \end{aligned}$$

Clearly $\mathcal{R}_\mathcal{W} \supseteq \mathcal{R}_\mathcal{V}$ and $\mathcal{R}_\mathcal{W}^{\hat{w},r} \supseteq \mathcal{R}_\mathcal{V}^{\hat{w},r}$, for all $r \geq 0$.

The occurrence of a local infimum gap is captured by the following definition:

Definition 3.2. *Let $(\hat{y}, \hat{w})$ be a feasible extended process such that $\hat{y}(S) \in \mathcal{R}_\mathcal{W} \backslash \mathcal{R}_\mathcal{V}$. We say that $(\hat{y}, \hat{w})$* satisfies the infimum gap condition *if, for any continuous cost function $h : \mathcal{M} \to \mathbb{R}$, there exists $r > 0$ such that one has*

$$h\big(\hat{y}(S)\big) < \inf \Big\{ h(y) : \quad y \in \mathcal{R}_\mathcal{V}^{\hat{w},r} \cap \mathfrak{T} \Big\}. \tag{3.15}$$

Despite the name, the infimum gap condition (3.15) is clearly a fully dynamical property. Actually, it can be as well rephrased in terms of 'supremum gap' or even independently of any optimization procedure, as it results from Lemma 3.1 below.

Definition 3.3. *Let $(\hat{y}, \hat{w})$ be a feasible extended process such that $\hat{y}(S) \in \mathcal{R}_\mathcal{W} \backslash \mathcal{R}_\mathcal{V}$. We say that $(\hat{y}, \hat{w})$* is isolated from $\mathcal{V}$ *if, for some $r > 0$ the sets $\Big(\mathcal{R}_\mathcal{V}^{\hat{w},r} \cup \{\hat{y}(S)\}\Big)$ and $\mathfrak{T}$ are locally separated at $\hat{y}(S)$, namely, there exists a neighborhood $\mathcal{N} \subset \mathcal{M}$ of $\hat{y}(S)$ such that $\Big(\mathcal{R}_\mathcal{V}^{\hat{w},r} \cup \{\hat{y}(S)\}\Big) \cap \mathfrak{T} \cap \mathcal{N} = \{\hat{y}(S)\}$.*

Lemma 3.1. *Let $(\hat{y}, \hat{w})$ be an extended feasible process such that $\hat{y}(S) \in \mathcal{R}_\mathcal{W} \backslash \mathcal{R}_\mathcal{V}$. Then the following conditions are equivalent:*

i) $(\hat{y}, \hat{w})$ satisfies (3.15) *for a given continuous cost function h and $\hat{r} > 0$;*

ii) the process $(\hat{y},\hat{w})$ is isolated from $\mathcal{V}$;

iii) the process $(\hat{y},\hat{w})$ satisfies the infimum gap condition. Furthermore the right hand-side of (3.15) *is equal to* $+\infty$.

Proof. We give a proof just for the sake of completeness, all arguments being trivial.

Let us start proving that $i)$ implies $ii)$. This means that one has to show that there exists $\hat{r}>0$ such that

$$\mathcal{R}_{\mathcal{V}}^{\hat{w},r}\cap\mathfrak{T}=\emptyset \quad \forall r<\hat{r}. \tag{3.16}$$

Assume that (3.16) is false, which means that there exists a sequence $r_n\downarrow 0$ such that $\mathcal{R}_{\mathcal{V}}^{\hat{w},r_n}\cap\mathfrak{T}\neq\emptyset$ for all natural n. This implies that there exists a sequence $(y_k)_{k\in\mathbb{N}}$ verifying $y_k\in\Big(\mathcal{R}_{\mathcal{V}}^{\hat{w},r_k}\cap\mathfrak{T}\Big)$ for every $k\in\mathbb{N}$, so that $y_n\to\hat{y}(S)$, which, in view of the continuity of h, contradicts $i)$. Hence, (3.16) holds true, from which we get $ii)$.

Let us now prove that $ii)\Rightarrow iii)$. By hypothesis, there exists a neighborhood $\mathcal{N}$ of $\hat{y}(S)$ such that $\Big(\mathcal{R}_{\mathcal{V}}^{\hat{w},r}\cup\{\hat{y}(S)\}\Big)\cap\mathfrak{T}\cap\mathcal{N}=\{\hat{y}(S)\}$. Since $\hat{y}(S)\in\mathcal{R}_{\mathcal{W}}\backslash\mathcal{R}_{\mathcal{V}}$, by possibly reducing the size of $r>0$ one obtains that $\mathcal{R}_{\mathcal{V}}^{\hat{w},r}\cap\mathfrak{T}=\emptyset$, which obviously implies $iii)$, with the right hand-side of (3.15) equal to $+\infty$. Finally, the relation $iii)\Rightarrow i)$ is trivial.

□

4. Abundance

Our main results –namely Theorems 5.1, 5.2, and 5.3– strongly rely on a property introduced by J. Warga and called "abundance". It consists in a particular pervasiveness of $\mathcal{V}$ in $\mathcal{W}$, which happens to be stronger than density. In fact, because of the presence of a closed final constraint, the mere density of $\mathcal{R}_{\mathcal{V}}$ into $\mathcal{R}_{\mathcal{W}}$ is not enough in order to normality to be a sufficient condition for gaps' avoidance (see Section 9). We will make use of a generalization of abundance introduced by B. Kaskosz in [24] and we will extend it to manifolds.

For every positive integer N, let Γ_N be the convex hull of the union of the origin $0\in\mathbb{R}^N$ with the N-simplex, namely

$$\Gamma_N:=\Big\{\gamma=(\gamma^1,...,\gamma^N)\in\mathbb{R}^N:\sum_{j=1}^N\gamma^j\leq 1,\ \gamma^j\geq 0,\ j=1,...,N\Big\}.$$

For any $\gamma\in\Gamma_N$, let us consider the control system on $\mathcal{M}$

$$\begin{cases}\dfrac{dy}{ds}(s)=f_\gamma\Big(s,y(s),w(s),w_1(s),...,w_N(s)\Big)\\[2ex] y(0)=\bar{y},\end{cases} \tag{4.17}$$

where: i) the *control* $(w, w_1, ..., w_N)$ belongs to $\mathcal{W}^{1+N}$, and ii) the vector field f_γ is defined by setting, for every $(s, y) \in [0, S] \times \mathcal{M}$ and $(\mathfrak{w}, \mathfrak{w}_1, ..., \mathfrak{w}_N) \in \mathfrak{W}^{1+N}$,

$$f_\gamma\big(s, y, \mathfrak{w}, (\mathfrak{w}_1, ..., \mathfrak{w}_N)\big) := f(s, y, \mathfrak{w}) + \sum_{i=1}^{N} \gamma^i \Big(f(s, y, \mathfrak{w}_i) - f(s, y, \mathfrak{w})\Big).$$

For every value of the parameter $\gamma \in \Gamma_N$ and every control $\big(w, w_1, ..., w_N\big) \in \mathcal{W}^{1+N}$, let us use $y_\gamma\Big[w, w_1, ..., w_N\Big]$ to denote the corresponding solution of (4.17).[7] Notice, in particular, that $y[w] = y_\gamma\Big[w, w, ..., w\Big]$ for all $w \in \mathcal{W}$ and for all $\gamma \in \Gamma_N$.

Definition 4.1. *[24] We say that a subclass of controls $\mathcal{V} \subset \mathcal{W}$ is* abundant in $\mathcal{W}$ *if, for every integer N, every $(1+N)$-tuple of controls $(w, w_1, ..., w_N) \in \mathcal{W}^{1+N}$, and every $\delta > 0$, there exists a continuous mapping $\theta^\delta_{w,w_1,...,w_N} : \Gamma_N \to \mathcal{V}$ such that*

$$d\Big(y_\gamma\Big[w, w_1, ..., w_N\Big](S), y\big[\theta^\delta_{w,w_1,...,w_N}(\gamma)\big](S)\Big) < \delta, \qquad \forall \gamma \in \Gamma_N. \tag{4.18}$$

A sufficient condition for abundance, based on *concatenation*, is given in Proposition 4.1 below.

Definition 4.2. *We say that a set of controls $\mathcal{V}$ satisfies the* concatenation property *if, for every $\bar{s} \in]0, S[$ and for any $v_1, v_2 \in \mathcal{V}$, one has $v_1\chi_{[0,\bar{s}[} + v_2\chi_{[\bar{s},S]} \in \mathcal{V}$,[8] where we have used χ_E to denote the indicator function of a subset $E \subseteq [0, S]$*

Proposition 4.1. *([24], Theorem IV.3.9) Assume that the subfamily $\mathcal{V} \subset \mathcal{W}$ satisfies the concatenation property and is dense in $\mathcal{W}$ with respect to the pseudo-metric d_f. Then $\mathcal{V}$ is an abundant subset of $\mathcal{W}$.*

The proof of this result for the special case when $\mathcal{M} = \mathbb{R}^n$ was given in ([24], Theorem IV.3.9) by developing some arguments in [21]. The required, obvious, changes to prove the result on a Riemannian manifold consists in a reformulation of estimate (4.18) in local coordinates, so we omit them.

4.1. Approximating the original reachable set by extended cones

Let us fix a a feasible extended process $(\hat{y}, \hat{w})$, and, for any $s, \check{s} \in [0, S]$, $s > \check{s}$, let $M(s, \check{s}) : T_{\hat{y}(\check{s})}\mathcal{M} \to T_{\hat{y}(s)}\mathcal{M}$ denote the differential of the diffeomorphism established by the differential equation $\dot{y} = f(s, y, \hat{w})$ from a neighborhood of $\hat{y}(\check{s})$ to a neighborhood of $\hat{y}(s)$. As it is known, $s \to M(s, \check{s})$ is the solution of the *variational* Cauchy problem having the following coordinate representation:

$$\frac{dM}{ds}(s) = \frac{\partial f}{\partial y}(s, \hat{y}(s), \hat{w}(s)) \circ M(s), \qquad M(\check{s}, \check{s}) = id_{T_{\hat{y}(\check{s})}\mathcal{M}}. \tag{4.19}$$

[7] Under hypothesis (SH) such a solution exists and is unique.

[8] Concatenation is weaker than *decomposability* of a set $\mathcal{S}$ of paths on an interval $[a, b]$ [19, 28, 35], which prescribes that for any pair of paths $v_1, v_2 \in \mathcal{S}$ and any measurable set $E \subset [0, S]$, one has $v_1\chi_E + v_2\chi_{([0,S]\setminus E)} \in \mathcal{S}$.

Definition 4.3. *Let N be a positive integer, and consider N control values $\mathfrak{w}_1, ..., \mathfrak{w}_N \in \mathfrak{W}$ and N instants $s_1, ..., s_N \in \mathrm{SD}\{f(\cdot,\cdot,\hat{w}(\cdot))\} \cap \mathrm{SD}\{f(\cdot,\cdot,\mathfrak{w}_1)\} \cap \ldots \cap \mathrm{SD}\{f(\cdot,\cdot,\mathfrak{w}_N)\} \cap \mathcal{L}\{c(\cdot)\}$,*[9] *$0 < s_1 < \ldots, < s_N \leq S$. The convex cone*

$$\mathbf{C}^{s_1,\ldots,s_N}_{\mathfrak{w}_1,\ldots,\mathfrak{w}_N} = span^+ \Big\{ M(S, s_i) \cdot \Big(f(s_i, \hat{y}(s_i), \mathfrak{w}_i) - f(s_i, \hat{y}(s_i), \hat{w}(s_i)) \Big) : \ i = 1, \ldots, N \Big\} \subset T_{\hat{y}(S)}\mathcal{M}$$

will be called extended variational cone *corresponding to the feasible extended process $(\hat{y}, \hat{w})$.*

The following result can be regarded as claiming a sort of *infinitesimal thickness* of $\mathcal{V}$ in $\mathcal{W}$.

Theorem 4.1. *Let the original family of controls $\mathcal{V} \subset \mathcal{W}$ be abundant in $\mathcal{W}$, and let a feasible extended process $(\hat{y}, \hat{w})$ be given. Let N be a positive integer, and consider N control values $\mathfrak{w}_1, ..., \mathfrak{w}_N \in \mathfrak{W}$ and N instants $s_1, ..., s_N \in \mathrm{SD}\{f(\cdot,\cdot,\hat{w}(\cdot))\} \cap \mathrm{SD}\{f(\cdot,\cdot,\mathfrak{w}_1)\} \cap \ldots \cap \mathrm{SD}\{f(\cdot,\cdot,\mathfrak{w}_N)\} \cap \mathcal{L}\{c(\cdot)\}$, $0 < s_1 < \ldots, < s_N \leq S$. Then, for any $r > 0$, the extended variational cone $\mathbf{C}^{s_1,\ldots,s_N}_{\mathfrak{w}_1,\ldots,\mathfrak{w}_N}$ is a QDQ approximating cone to $\mathcal{R}^{\hat{w},r}_{\mathcal{V}} \cup \{\hat{y}(S)\}$ at $\hat{y}(S)$.*

Remark 4.1. While the fact that $\mathbf{C}^{s_1,\ldots,s_N}_{\mathfrak{w}_1,\ldots,\mathfrak{w}_N}$ is a QDQ approximating cone to the extended reachable set $\mathcal{R}^{\hat{w},r}_{\mathcal{W}}$ at $\hat{y}(S)$ (for any $r > 0$) is a classical argument, utilized in the proof of the Maximum Principle,[10] the fact that $\mathbf{C}^{s_1,\ldots,s_N}_{\mathfrak{w}_1,\ldots,\mathfrak{w}_N}$ is a first order approximation for the original reachable set $\mathcal{R}^{\hat{w},r}_{\mathcal{V}}$ is anything but obvious: it means, in a sense, that this cone is *not too large.*

Proof of Theorem 4.1. We will prove this theorem assuming that $\mathcal{M}$ is an open subset of $\mathbb{R}^n$, so that we can identify $T_{\hat{y}(S)}\mathcal{M}$ with $\mathbb{R}^n$. Clearly, this is not restrictive because of the local character of the result.

Let us set $s_0 = 0$ and, for every $i = 1, ..., N$, consider a number $\delta^i \leq s_i - s_{i-1}$ and the control

$$w_i^{\delta^i}(s) := \begin{cases} \hat{w}(s), & \forall s \in [0, S] \backslash [s_i - \delta^i, s_i] \\ \mathfrak{w}_i & \forall s \in [s_i - \delta^i, s_i]. \end{cases} \tag{4.20}$$

Let us set $\bar{\delta} := \frac{N}{2} \min\{s_i - s_{i-1}, \ i = 1, \ldots, N\}$. Let us define the set-valued map $F : \mathbb{R}^N \rightsquigarrow \mathbb{R}^n$ as

$$F(\epsilon) = \left\{ y\left[\theta^{\delta^2}_{\hat{w}, w_1^{\delta/N}, \ldots, w_N^{\delta/N}}\left(\pi\left(\frac{\epsilon}{\delta}\right)\right)\right](S) : \qquad 0 < \delta \leq \bar{\delta} \right\} \quad \forall \epsilon \in \mathbb{R}^N, \tag{4.21}$$

where $\pi : \mathbb{R}^N \to \Gamma_N$ denotes the orthogonal projection on Γ_N (which, because of the convexity of Γ_N, is a continuous, single-valued, map). Notice that, by construction $F(\epsilon) \subseteq \mathcal{R}_{\mathcal{V}}$, for every $\epsilon \in \mathbb{R}^N$.

For each $\delta \in]0, \bar{\delta}]$ and $\epsilon \in \Gamma_N \cap B_\delta$, let us choose $\gamma = \left(\gamma^1, \ldots, \gamma^N\right) := \left(\frac{\epsilon^1}{\delta}, \ldots, \frac{\epsilon^N}{\delta}\right) \in \Gamma_N$.

[9] We recall (see Subsection 1.1) that $SD(\phi)$ denotes the (full measure) set of Scorza-Dragoni points of a function $\phi = \phi(s, y)$, while $\mathcal{L}\{c\}$ denotes the set of Lebesgue points of the integrable function c.

[10] Actually the same is true for other, more classical, cones, e.g. the Boltyanski cone and the regular tangent cone.

From (4.25) in Lemma 4.1 below it follows that

$$
\begin{aligned}
&y_{\epsilon/\delta}[\hat{w}, w_1^{\delta/N}, \dots, w_N^{\delta/N}](S) = \\
&= \hat{y}(S) + \frac{1}{N}\sum_{i=1}^{N} \epsilon^i \, M(S, s_i) \cdot \Big(f(s_i, \hat{y}(s_i), \mathfrak{w}_i) - f(s_i, \hat{y}(s_i), \hat{w}(s_i)) \Big) + \phi(\epsilon, \delta),
\end{aligned}
\tag{4.22}
$$

for every $\epsilon \in \Gamma_N \cap B_\delta$, where $\phi : \bigcup_{0<\delta\leq\bar{\delta}} ((\Gamma_N \cap B_\delta) \times \{\delta\}) \to \mathbb{R}^n$ is a continuous function which verifies $\max\big\{|\phi(\epsilon,\delta)|,\ \epsilon \in \Gamma_N \cap B_\delta\big\} = o(\delta)$. In view of the abundance property, for each $\delta \in]0, \bar{\delta}]$ and $\epsilon \in \Gamma_N \cap B_\delta$, there exists $\tilde{\phi} : \bigcup_{0<\delta\leq\bar{\delta}} ((\Gamma_N \cap B_\delta) \times \{\delta\}) \to \mathbb{R}^n$ such that

$$
y\Big[\theta^{\delta^2}_{\hat{w}, w_1^{\delta/N}, \dots, w_N^{\delta/N}}\Big(\frac{\epsilon}{\delta}\Big)\Big](S) - y_{\epsilon/\delta}[\hat{w}, w_1^{\delta/N}, \dots, w_N^{\delta/N}](S) = \tilde{\phi}(\epsilon, \delta)
$$

for all $\epsilon \in \Gamma_N \cap B_\delta$, with $\max\big\{|\tilde{\phi}(\epsilon,\delta)|,\ \epsilon \in \Gamma_N \cap B_\delta\big\} \leq \delta^2$. Therefore

$$
\begin{aligned}
&y\Big[\theta^{\delta^2}_{\hat{w}, w_1^{\delta/N}, \dots, w_N^{\delta/N}}\Big(\frac{\epsilon}{\delta}\Big)\Big](S) = \\
&= \hat{y}(S) + \frac{1}{N}\sum_{i=1}^{N} \epsilon^i \, M(S, s_i) \cdot \Big(f(s_i, \hat{y}(s_i), \mathfrak{w}_i) - f(s_i, \hat{y}(s_i), \hat{w}(s_i)) \Big) + h^\delta(\epsilon),
\end{aligned}
\tag{4.23}
$$

where $h^\delta(\epsilon) := \phi(\epsilon, \delta) + \tilde{\phi}(\epsilon, \delta)$. Observe that

$$
|h^\delta(\epsilon)| \leq \delta\rho(\delta), \quad \forall \epsilon \in \Gamma_N \cap B_\delta, \tag{4.24}
$$

where we have set $\rho(\delta) := \dfrac{\max\big\{|\phi(\epsilon,\delta)|,\ \epsilon \in \Gamma_N \cap B_\delta\big\}}{\delta} + \delta.$

For every $\delta \in]0, \bar{\delta}]$, let us define the map

$$
\begin{aligned}
A^\delta : \Gamma_N \cap B_\delta &\to Lin(\mathbb{R}^N, \mathbb{R}^n) \times \mathbb{R}^n \\
\epsilon &\mapsto A^\delta(\epsilon) := (L, h^\delta(\epsilon)),
\end{aligned}
$$

where L is the linear map defined as

$$
L \cdot b = \frac{1}{N}\sum_{i=1}^{N} b^i M(S, s_i) \cdot \Big(f(s_i, \hat{y}(s_i), \mathfrak{w}_i) - f(s_i, \hat{y}(s_i), \hat{w}(s_i)) \Big), \quad \forall b \in \mathbb{R}^N.
$$

Notice that, because of the continuity w.r.t. ϵ of the left-hand side of (4.23), for every $\delta > 0$, the map $\epsilon \mapsto A^\delta(\epsilon)$ is continuous. By rewriting relation (4.23) as

$$
y\Big[\theta^{\delta^2}_{\hat{w}, w_1^{\delta/N}, \dots, w_N^{\delta/N}}\Big(\frac{\epsilon}{\delta}\Big)\Big](S) = \hat{y}(S) + L \cdot \epsilon + h^\delta(\epsilon),
$$

we get

$$
\hat{y}(S) + L \cdot \epsilon + h^\delta(\epsilon) \in F(\epsilon),
$$

which means that L is a QDQ of F at $(0, \hat{y}(S))$ in the direction of the set Γ_N. Since $\mathbf{C}^{s_1,\dots,s_N}_{\mathfrak{w}_1,\dots,\mathfrak{w}_N} = L \cdot \Gamma$, one concludes that $\mathbf{C}^{s_1,\dots,s_N}_{\mathfrak{w}_1,\dots,\mathfrak{w}_N}$ is a QDQ approximating cone to $\mathcal{R}_{\mathcal{V}} \cup \{\hat{y}(S)\}$ at $\hat{y}(S)$. □

Lemma 4.1. *Fix $\gamma \in \Gamma_N$ and consider N control values $\mathfrak{w}_1, ..., \mathfrak{w}_N \in \mathfrak{W}$ and N instants $s_1, ..., s_N \in \mathrm{SD}\{f(\cdot,\cdot,\hat{w}(\cdot))\} \cap \mathrm{SD}\{f(\cdot,\cdot,\mathfrak{w}_1)\} \cap \ldots \cap \mathrm{SD}\{f(\cdot,\cdot,\mathfrak{w}_N)\} \cap \mathcal{L}\{c(\cdot)\}$, $0 < s_1 < \ldots, < s_N \leq S$. Then, the map $\epsilon \to y_\gamma\Big[\hat{w}, w_1^{\epsilon^1}, ..., w_N^{\epsilon^N}\Big](S)$ verifies*

$$
\begin{aligned}
& y_\gamma\left[\hat{w}, w_1^{\epsilon^1}, ..., w_N^{\epsilon^N}\right](S) - \hat{y}(S) \\
& \quad = \sum_{i=1}^{N} \gamma^i \epsilon^i M(S, s_i) \cdot \Big(f(s_i, \hat{y}(s_i), \mathfrak{w}_i) - f(s_i, \hat{y}(s_i), \hat{w}(s_i)) \Big) + \phi(\epsilon, \gamma),
\end{aligned} \tag{4.25}
$$

where ϕ is a continuous function verifying $\max\{\phi(\epsilon,\gamma) : \ \gamma \in \Gamma_N\} = o(\epsilon)$.

Proof. Let us begin by proving the lemma in the case when $N = 1$ and $0 \leq \gamma \leq 1$. One has

$$
\begin{aligned}
& y_\gamma\left[\hat{w}, w_1^{\epsilon^1}\right](s_1) - \hat{y}(s_1) = y_\gamma\left[\hat{w}, w_1^{\epsilon^1}\right](s_1 - \epsilon^1) - \hat{y}(s_1 - \epsilon^1) \\
& \quad + \int_{s_1-\epsilon^1}^{s_1} \Big(f_\gamma\Big(s, y_\gamma[\hat{w}, w_1^{\epsilon^1}](s), \hat{w}(s), \mathfrak{w}_1\Big) - f(s, \hat{y}(s), \hat{w}(s)) \Big)\, ds \\
& \quad = \int_{s_1-\epsilon^1}^{s_1} \Big(f_\gamma\Big(s, y_\gamma[\hat{w}, w_1^{\epsilon^1}](s), \hat{w}(s), \mathfrak{w}_1\Big) - f(s, \hat{y}(s), \hat{w}(s)) \Big)\, ds \\
& = \epsilon^1 \Big(f_\gamma(s_1, \hat{y}(s_1), \hat{w}(s_1), \mathfrak{w}_1) - f(s_1, \hat{y}(s_1), \hat{w}(s_1)) \Big) + \Phi_1(\epsilon^1, \gamma) + \Phi_2(\epsilon^1) \\
& = \gamma\, \epsilon^1 \cdot \Big(f(s_1, \hat{y}(s_1), \mathfrak{w}_1) - f(s_1, \hat{y}(s_1), \hat{w}(s_1)) \Big) + \Phi_1(\epsilon^1, \gamma) + \Phi_2(\epsilon^1),
\end{aligned} \tag{4.26}
$$

where

$$
\Phi_1(\epsilon^1, \gamma) := \int_{s_1-\epsilon^1}^{s_1} \Big(f_\gamma\Big(s, y_\gamma[\hat{w}, w_1^{\epsilon^1}](s), \hat{w}(s), \mathfrak{w}_1\Big) - f_\gamma(s_1, \hat{y}(s_1), \hat{w}(s_1), \mathfrak{w}_1) \Big)\, ds,
$$

$$
\Phi_2(\epsilon^1) := \int_{s_1-\epsilon^1}^{s_1} \Big(f(s_1, \hat{y}(s_1), \hat{w}(s_1)) - f(s, \hat{y}(s), \hat{w}(s)) \Big) ds.
$$

To simplify the notation, in what follows we will write $y_\gamma(s)$ in place of $y_\gamma[\hat{w}, w_1^{\epsilon^1}](s)$. Using hypothesis **(SH)**-(ii), for every $\sigma \in [s_1 - \epsilon^1, s_1]$, one obtains the following estimate:

$$
\begin{aligned}
& |y_\gamma(\sigma) - \hat{y}(\sigma)| \leq \int_{s_1-\epsilon^1}^{\sigma} |f(s, y_\gamma(s), \hat{w}(s)) - f(s, \hat{y}(s), \hat{w}(s))|\, ds + \\
& \gamma \int_{s_1-\epsilon^1}^{\sigma} |f(s, y_\gamma(s), \mathfrak{w}_1) - f(s, y_\gamma(s), \hat{w}(s)) + f(s, \hat{y}(s), \hat{w}(s)) - f(s, \hat{y}(s), \mathfrak{w}_1)|\, ds + \\
& \gamma \int_{s_1-\epsilon^1}^{\sigma} |f(s, \hat{y}(s), \mathfrak{w}_1) - f(s, \hat{y}(s), \hat{w}(s))|\, ds \leq \\
& (1 + 2\gamma) \int_{s_1-\epsilon^1}^{\sigma} c(s)\, |y_\gamma(s) - \hat{y}(s)|\, ds + \gamma \int_{s_1-\epsilon^1}^{\sigma} |f(s, \hat{y}(s), \mathfrak{w}_1) - f(s, \hat{y}(s), \hat{w}(s))|\, ds.
\end{aligned} \tag{4.27}
$$

Setting

$$
\alpha(\sigma) = \gamma \int_{s_1-\epsilon^1}^{\sigma} |f(s, \hat{y}(s), \mathfrak{w}_1) - f(s, \hat{y}(s), \hat{w}(s))|\, ds, \tag{4.28}
$$

from the Grönwall's Lemma, we obtain that

$$
\begin{aligned}
&|y_\gamma(s_1)-\hat y(s_1)| \le \alpha(s_1) + \int_{s_1-\epsilon^1}^{s_1} (1+2\gamma)c(s)\exp\left\{(1+2\gamma)\int_s^{s_1} c(\sigma)d\sigma\right\}\alpha(s)ds\\
&\le \gamma\int_{s_1-\epsilon^1}^{s_1} |f(s,\hat y(s),\mathfrak{w}_1)-f(s,\hat y(s),\hat w(s))|\,ds + (1+2\gamma)\alpha(s_1)\int_{s_1-\epsilon^1}^{s_1} c(s)\exp\left\{(1+2\gamma)\int_s^{s_1} c(\sigma)d\sigma\right\}ds\\
&= \gamma\int_{s_1-\epsilon^1}^{s_1} |f(s,\hat y(s),\mathfrak{w}_1)-f(s,\hat y(s),\hat w(s))|\,ds + o(\epsilon^1)\to 0,
\end{aligned}
\tag{4.29}
$$

as soon as $\epsilon^1\to 0$. Therefore,

$$
\begin{aligned}
&|y_\gamma(s)-\hat y(s_1)| \le |y_\gamma(s)-y_\gamma(s_1)| + |y_\gamma(s_1)-\hat y(s_1)| \le (1+4\gamma)\int_{s_1-\epsilon^1}^{s_1} c(s)ds + o(\epsilon^1)\\
&\le (1+4\gamma)\int_{s_1-\epsilon^1}^{s_1} |c(s)-c(s_1)|\,ds + (1+4\gamma)\epsilon^1 c(s_1) + o(\epsilon^1) = (1+4\gamma)\epsilon^1 c(s_1) + o(\epsilon^1),
\end{aligned}
\tag{4.30}
$$

where the last equality follows from the fact that $s_1\in\mathcal{L}\{c(\cdot)\}$. Since

- s_1 is a Scorza-Dragoni point of $f(\cdot,\cdot,\hat w(\cdot))$ and $f(\cdot,\cdot,\mathfrak{w}_1)$, and
- the maps $y\mapsto f(s,y,\hat w(s))\}$, $y\mapsto f(s,y,\mathfrak{w}_1)$ are Lipschitz continuous in a neighbourhood of $\hat y([0,S])$,

in view of (4.29), (4.30), one gets

$$
\max\left\{\Phi_1(\epsilon^1,\gamma):\ 0\le\gamma\le 1\right\} = o(\epsilon^1), \qquad \Phi_2(\epsilon^1) = o(\epsilon^1). \tag{4.31}
$$

If we set $\phi(\epsilon^1,\gamma) := \Phi_1(\epsilon^1,\gamma)+\Phi_2(\epsilon^1)$, ϕ is a continuous map and, by estimates (4.26) and (4.31), it follows that

$$
\begin{aligned}
&y_\gamma\left[\hat w, w_1^{\epsilon^1}\right](s_1) - \hat y(s_1)\\
&= \gamma\,\epsilon^1\Big(f\left(s_1,\hat y(s_1),\hat{\mathfrak{w}}_1\right) - f\left(s_1,\hat y(s_1),\hat w(s_1)\right)\Big) + \phi(\epsilon^1,\gamma).
\end{aligned}
$$

Hence, one has

$$
\frac{d}{d\epsilon^1} y_\gamma\left[\hat w, w_1^{\epsilon^1}\right](s_1)_{|\epsilon^1=0} = \gamma\Big(f\left(s_1,\hat y(s_1),\hat{\mathfrak{w}}_1\right) - f\left(s_1,\hat y(s_1),\hat w(s_1)\right)\Big),
$$

which, by the basic theory of linear ODE's, implies that

$$
\begin{aligned}
&\frac{d}{d\epsilon^1} y_\gamma\left[\hat w, w_1^{\epsilon^1}\right](S)_{|\epsilon^1=0} = M(S,s_1)\cdot\frac{d}{d\epsilon^1} y_\gamma\left[\hat w, w_1^{\epsilon^1}\right](s_1)_{|\epsilon^1=0} =\\
&M(S,s_1)\cdot\gamma\Big(f\left(s_1,\hat y(s_1),\hat{\mathfrak{w}}_1\right) - f\left(s_1,\hat y(s_1),\hat w(s_1)\right)\Big).
\end{aligned}
$$

Therefore, the lemma is proved for $N=1$ and for any $0\le\gamma\le 1$. The general case $N\ge 2$ is easily obtained by a finite induction argument. The latter doesn't display any new difficulty with respect to the proof of the case $N=1$. Actually, the argument is almost verbatim the one utilized in the proof of the Pontryagin Maximum Principle when passing from single to multiple, finitely many needle variations (see e.g. [40], Theorem 4.2.1). Hence we omit it. $\square$

The reason why we have adopted QDQ approximating cones as tangential objects relies on the validity of the following result. [11]

Theorem 4.2. *Let the original family of controls $\mathcal{V} \subset \mathcal{W}$ be abundant in $\mathcal{W}$, and let a feasible extended process $(\hat{y}, \hat{w})$ be given. Let N be a positive integer, and consider N control values $\mathfrak{w}_1, ..., \mathfrak{w}_N \in \mathfrak{W}$ and N instants $s_1, ..., s_N \in \mathrm{SD}\{f(\cdot, \cdot, \hat{w}(\cdot))\} \cap \mathrm{SD}\{f(\cdot, \cdot, \mathfrak{w}_1)\} \cap \ldots \cap \mathrm{SD}\{f(\cdot, \cdot, \mathfrak{w}_N)\} \cap \mathcal{L}\{c(\cdot)\}$, $0 < s_1 < \ldots, < s_N \leq S$. Moreover, let C be a QDQ approximating cone to the target $\mathfrak{T}$ at $\hat{y}(S)$. If $\mathbf{C}^{s_1,\ldots,s_N}_{\mathfrak{w}_1,\ldots,\mathfrak{w}_N}$ and C, are complementary subspaces, then there exists a sequence $(z_k)_{k\in\mathbb{N}} \subset \mathcal{R}_{\mathcal{V}} \cap \mathfrak{T}$ such that*

$$\lim_{k\to\infty} z_k = \hat{y}(S)$$

.

Proof. In view of Theorem 4.1, $\mathbf{C}^{s_1,\ldots,s_N}_{\mathfrak{w}_1,\ldots,\mathfrak{w}_N}$ is a QDQ approximating cone to $\mathcal{R}_{\mathcal{V}} \cup \{\hat{y}(S)\}$ at $\hat{y}(S)$. Furthermore, since C is a QDQ approximating cone to the target $\mathfrak{T}$ at $\hat{y}(S)$, there exist a positive integer M, a set-valued map $G : \mathbb{R}^M \rightsquigarrow \mathcal{M}$, a convex cone $\Gamma \subset \mathbb{R}^M$, and a Quasi Differential Quotient L of G at $(0, \hat{y}(S))$ in the direction of Γ such that $G(\Gamma) \subseteq \mathfrak{T}$ and $C = L \cdot \Gamma$. In order to conclude the proof, it is enough to apply Theorem 2.3, *ii*), with $C_1 = \mathbf{C}^{s_1,\ldots,s_N}_{\mathfrak{w}_1,\ldots,\mathfrak{w}_N}$, $C_2 = C$, $N_1 = N$, $N_2 = M$, $\Gamma_1 = [0, \infty)^N$, $\Gamma_2 = \Gamma$, F_1 defined as in (4.21), and $F_2 = G$. This concludes the proof. □

5. The main results

Theorem 5.1 (A GEOMETRIC PRINCIPLE FOR GAPS). *Let us assume that the family of controls $\mathcal{V}$ is abundant in $\mathcal{W}$. Let $(\hat{y}, \hat{w})$ be a feasible extended process satisfying the infimum gap condition. Then any QDQ approximating cone C to $\mathfrak{T}$ at $\hat{y}(S)$ is linearly separable from any extended variational cone $\mathbf{C}^{s_1,\ldots,s_N}_{\mathfrak{w}_1,\ldots,\mathfrak{w}_N}$, i.e. there exists a non-zero linear form $\xi \in T^*_{\hat{y}(S)}\mathcal{M}$ such that*

$$\xi \cdot c_1 \leq 0 \leq \xi \cdot c_2, \qquad \forall (c_1, c_2) \in \mathbf{C}^{s_1,\ldots,s_N}_{\mathfrak{w}_1,\ldots,\mathfrak{w}_N} \times C.$$

Let us give the definition of *abnormal extremal, normal h-extremal, and h-abnormal extremal.*

Definition 5.1 (Abnormal extremal). *Let $(\hat{y}, \hat{w})$ be a feasible extended process, and let C be a QDQ approximating cone to the target $\mathfrak{T}$ at $\hat{y}(S)$. We say that the process $(\hat{y}, \hat{w})$ is an* abnormal extremal *(with respect to C) if there exists a lift $(\hat{y}, \lambda) \in W^{1,1}([0, S]; T^*\mathcal{M})$ of $\hat{y}$ verifying the following conditions:*

$$(i) \quad \frac{d\lambda}{ds} = -\lambda \cdot \frac{\partial f}{\partial y}(s, \hat{y}(s), \hat{w}(s));$$

$$(ii) \quad \max_{\mathfrak{w}\in\mathfrak{W}} \lambda(s) \cdot f(s, \hat{y}(s), \mathfrak{w}) = \lambda(s) \cdot f(s, \hat{y}(s), \hat{w}(s)) \quad a.e.\ s \in [0, S];$$

[11] For the validity of this result, it is crucial to utilize QDQ approximating cones instead of the more general AGDQ approximating cones.

(iii) $\lambda(S) \in -C^{\perp}$;

(iv) $\lambda \neq 0$.

Definition 5.2 (h-extremal)**.** *Let $(\hat{y}, \hat{w})$ be a feasible extended process, let a (cost) function $h : \mathcal{M} \to \mathbb{R}$ be differentiable at $\hat{y}(S)$, and let C be a QDQ approximating cone of the target $\mathfrak{T}$ at $\hat{y}(S)$. We say that the process $(\hat{y}, \hat{w})$ is an* h-extremal *(with respect to h and C) if there exist a lift $(\hat{y}, \lambda) \in W^{1,1}([0, S]; T^*\mathcal{M})$ of $\hat{y}(\cdot)$ and a* cost multiplier $\lambda_c \in \{0, 1\}$ *such that:*

(i) $\dfrac{d\lambda}{ds} = -\lambda \cdot \dfrac{\partial f}{\partial y}(s, \hat{y}(s), \hat{w}(s))$;

(ii) $\max_{\mathfrak{w} \in \mathfrak{W}} \lambda(s) \cdot f(s, \hat{y}(s), \mathfrak{w}) = \lambda(s) \cdot f(s, \hat{y}(s), \hat{w}(s)) \quad$ *a.e.* $s \in [0, S]$;

(iii) $\lambda(S) \in -\lambda_c \nabla h(\hat{y}(S)) - C^{\perp}$;

(iv) $(\lambda, \lambda_c) \neq 0$.

Furthermore, we say that an h-extremal $(\hat{y}, \hat{w})$ is normal *if for every choice of the pair (λ, λ_c), one has $\lambda_c = 1$. We say that an* h-extremal $(\hat{y}, \hat{w})$ is abnormal *if it is not normal, namely if exists a choice of (λ, λ_c) with $\lambda_c = 0$.*

Remark 5.1. Though these definitions have intrinsic meanings, we have chosen to adopt a notation reminiscent of coordinates. Of course, the adjoint equation (i) might be expressed –when coupled with the dynamics– as the Hamiltonian system

$$\frac{d}{dt}(y, \lambda) = X_H(s, x, \lambda) := J \cdot DH(s, x, \lambda),$$

where $H : T^*\mathcal{M} \to \mathbb{R}$ is the maximized Hamiltonian defined by setting

$$H(s, x, \lambda) := \max_{\mathfrak{w} \in \mathfrak{W}} \lambda(s) \cdot f(s, x, \mathfrak{w}) \qquad \forall (s, x, \lambda) \in [0, S] \times T^*\mathcal{M}$$

and X_H is the Hamiltonian vector field, namely $X_H := J \cdot DH$, J being the symplectic matrix and D the differential operator with respect to x and λ.

Observe that every abnormal extremal is an abnormal h-extremal for any cost h differentiable at $\hat{y}(S)$, while every abnormal h-extremal is an abnormal extremal. We are now ready to state our main result on infimum gaps.

Theorem 5.2 (Normality No-Gap Criterion)**.** *Let us assume that the family of controls $\mathcal{V}$ is abundant in $\mathcal{W}$. If a feasible extended process $(\hat{y}, \hat{w})$ satisfies the infimum gap condition, then, for every QDQ approximating cone C to $\mathfrak{T}$ at $\hat{y}(S)$, $(\hat{y}, \hat{w})$ is an abnormal extremal with respect to C.*

When referred to a specific cost h, the contrapositive version of this theorem provides a sufficient condition for the absence of local infimum gaps. Precisely:

Theorem 5.3 (A sufficient condition for avoiding infimum gaps)**.** *Let us assume that the family of controls $\mathcal{V}$ is abundant in $\mathcal{W}$, and let $(\hat{y},\hat{w})$ be a feasible extended process. Let $h:\mathcal{M}\to\mathbb{R}$ be a continuous cost function, differentiable at $\hat{y}(S)$, and let $(\hat{y},\hat{w})$ be a normal h-extremal for some QDQ approximating cone C to $\mathfrak{T}$ at $\hat{y}(S)$. Then $(\hat{y},\hat{w})$ does not satisfy the infimum gap condition.*

As we have mentioned in the Introduction, the relation between gap phenomena and abnormality has been quite investigated in two cases of embeddings: the embedding of bounded optimal control problems into their *convex relaxation* [37, 38, 39] and the embedding of unbounded (convex) control systems into their *impulsive, space-time closure* [33]. Since (by Proposition 4.1) the original control families in such embeddings turn out to be abundant in their extensions, these kinds of results can be also obtained by Theorem 5.2.[12] In Section 8, we are going to present a new application to a dynamics which is *neither convex nor bounded.*

Remark 5.2. Let us mention that the *Lavrentiev phenomenon* (see e.g. [15, 26, 27]) in the Calculus of Variations is strictly connected with the notion of infimum gap. In our opinion this deserves future investigation. Indeed, since there are no dynamical constraints, one has full local controllability, which, in turn, implies normality. On the one hand, in several examples of Lavrentiev phenomenon, even the Calculus of Variations reduction of the Maximum Principle, namely the Euler-Lagrange equations, often fails to hold true. Hence, it is not at all clear what notion of *normality* one should take into account. On the other hand, the Lavrentiev phenomenon is likely due more to the lack of some kind of *abundance* of the set of original velocities in the set of the extended velocities than to abnormality.

6. A verifiable sufficient condition for normality

In practical situations, it may be difficult or even impossible to directly verify the normality of an extremal, which, in view of Theorem 5.3, would guarantee the absence of gaps. This motivates Proposition 6.1 below, which provides a sufficient condition in order for a process not to be an abnormal extremal.

In the following definition we assume that a control system (E) as above is given, with an initial condition $y(0)=\bar{y}$, and we still use $\mathcal{R}_{\mathcal{W}}$ to denote the reachable set from $\bar{y}$.

Definition 6.1. *Consider a feasible process $(\tilde{y},\tilde{w}):[0,S]\to\mathcal{M}\times\mathfrak{W}$ of (E). Let C be a QDQ approximating cone to $\mathfrak{T}$ at $\tilde{y}(S)$. We say that the process $(\tilde{y},\tilde{w})$ is C-*needle-controllable *at S if, for every $\boldsymbol{\xi}\in C^{\perp}\backslash\{0\}$, there exist $\delta_1>0$ and $\delta_2\in(0,S]$ such that*

$$\sup_{\mathfrak{w}\in\mathfrak{W}}\boldsymbol{\xi}\cdot\left(f(s,\tilde{y}(s),\mathfrak{w})-f(s,\tilde{y}(s),\tilde{w}(s))\right)\geq\delta_1\qquad a.e.\ s\in[S-\delta_2,S].\tag{6.32}$$

Proposition 6.1. *Consider a feasible process $(\hat{y},\hat{w}):[0,S]\to\mathcal{M}\times\mathfrak{W}$ of (E). Let C be a QDQ approximating cone to $\mathfrak{T}$ at $\hat{y}(S)$, and let the process $(\hat{y},\hat{w})$ be C-needle-controllable at S. Then the process $(\hat{y},\hat{w})$ is not an abnormal extremal, so, in particular, it cannot satisfy the infimum-gap condition.*

[12] Although the use of different types of cones describing the non-transversality condition makes Theorem 5.2 and the results in [33, 37, 38, 39] distinct (see [9], [36] for the details).

Indeed, if the process $(\hat{y}, \hat{w})$ were an abnormal extremal, there would exist an absolutely continuous lift $(\hat{y}, \lambda) : [0, S] \to T^*\mathcal{M}$ of $\hat{y}$ such that $\lambda \neq 0$, $\lambda(S) \in -C^\perp$ and, furhermore, the inequality

$$\lambda(s) \cdot (f(s, \hat{y}(s), \mathfrak{w}) - f(s, \hat{y}(s), \hat{w}(s))) \leq 0$$

would hold for almost every $s \in [0, S] \backslash I_0$ and every $\mathfrak{w} \in \mathfrak{W}$, I_0 having zero Lebesgue measure. This contradicts (6.32) as soon as one we set $\boldsymbol{\xi} := \lambda(S)$.

7. Proofs of the main results

7.1. Proof of the Geometric Principle (Theorem 5.1)

By a basic result on control system (see e.g. [40], [11]), $\mathbf{C}^{s_1,\ldots,s_N}_{\mathfrak{w}_1,\ldots,\mathfrak{w}_N}$ turns out to be a QDQ approximating cone to the (local) extended reachable set $\mathcal{R}^{\hat{w},r}_{\mathcal{W}}$ at $\hat{y}(S)$.[13] More importantly, Theorem 4.1 states that $\mathbf{C}^{s_1,\ldots,s_N}_{\mathfrak{w}_1,\ldots,\mathfrak{w}_N}$ is also a QDQ approximating cone to $\mathcal{R}^{\hat{w},r}_{\mathcal{V}} \cup \{\hat{y}(S)\}$ at $\hat{y}(S)$. Therefore, by Lemma 3.1, the sets $\left(\mathcal{R}^{\hat{w},r}_{\mathcal{V}} \cup \{\hat{y}(S)\}\right)$ and $\mathfrak{T}$ are locally separated at $\hat{y}(S)$, which by Theorem 2.3, *i*), implies that the cones $\mathbf{C}^{s_1,\ldots,s_N}_{\mathfrak{w}_1,\ldots,\mathfrak{w}_N}$ and C are not strongly transverse. Since linear separability is equivalent to non-transversality (Proposition 1.2) we have to prove that *these cones are not transverse as well.* Actually, in view of Proposition 1.1, the only case in which they might happen to be transverse (and not strongly transverse) is the one in which they are complementary subspaces of $T_{\hat{y}(S)}\mathcal{M}$. However, such an instance happen happens to be excluded by Theorem 4.2 and the occurrence of an infimum gap: indeed, if $\mathbf{C}^{s_1,\ldots,s_N}_{\mathfrak{w}_1,\ldots,\mathfrak{w}_N} + C = T_{\hat{y}(S)}\mathcal{M}$ and $\mathbf{C}^{s_1,\ldots,s_N}_{\mathfrak{w}_1,\ldots,\mathfrak{w}_N} \cap C = \{0\}$, then Theorem 4.2 provides the existence of a sequence $(y_k)_{k\in\mathbb{N}} \subset \mathcal{R}_{\mathcal{V}} \cap \mathfrak{T}$ such that $y_k \to \hat{y}(S)$, which contradicts the fact that $(\hat{y}, \hat{w})$ verifies the infimum gap condition. This concludes the proof. □

7.2. Proof of Theorem 5.2

By Theorem 5.1, the cones $\mathbf{C}^{s_1,\ldots,s_N}_{\mathfrak{w}_1,\ldots,\mathfrak{w}_N}$ and C are linearly separable. This means that there exists $\xi \in (T_{\hat{y}(S)}\mathcal{M})^* \backslash \{0\}$ such that $\xi \in -C^\perp \cap (\mathbf{C}^{s_1,\ldots,s_N}_{\mathfrak{w}_1,\ldots,\mathfrak{w}_N})^\perp$. Now let us set $\lambda(s) := \xi \cdot M(S, s)$, where $M(S, s)$ is the fundamental matrix defined in (4.19), so that

$$\lambda \neq 0, \qquad \lambda(S) \in -C^\perp, \quad \frac{d\lambda}{ds}(s) = -\lambda(s) \cdot \frac{\partial f}{\partial y}(s, \hat{y}(s), \hat{w}(s)), \quad \text{for a.e. } s \in [0, S].$$

By $\xi \in (\mathbf{C}^{s_1,\ldots,s_N}_{\mathfrak{w}_1,\ldots,\mathfrak{w}_N})^\perp$, it follows that, for every $i = 1, \ldots, N$,

$$\begin{aligned} 0 &\geq \xi \cdot \Big(M(S, s_i) \cdot \Big(f(s_i, \hat{y}(s_i), \mathfrak{w}_i) - f(s_i, \hat{y}(s_i), \hat{w}(s_i))\Big)\Big) \\ &= \Big(\xi \cdot M(S, s_i)\Big) \cdot \Big(f(s_i, \hat{y}(s_i), \mathfrak{w}_i) - f(s_i, \hat{y}(s_i), \hat{w}(s_i))\Big) \\ &= \lambda(s_i) \cdot \Big(f(s_i, \hat{y}(s_i), \mathfrak{w}_i) - f(s_i, \hat{y}(s_i), \hat{w}(s_i))\Big). \end{aligned} \tag{7.33}$$

[13]For instance: it is well-known that $\mathbf{C}^{s_1,\ldots,s_N}_{\mathfrak{w}_1,\ldots,\mathfrak{w}_N}$ is a Boltyanski approximating cone to $\mathcal{R}^{\hat{w},r}_{\mathcal{W}}$ at $\hat{y}(S)$ (see e.g. [44]). Furthermore, a Boltyanski approximating cone is clearly a QDQ approximating cone.

Therefore the lift $(\hat{y}, \lambda)$ verifies (i)-(iv) of Definition 5.1, except that (iii) is verified only for every finite set of pairs $(s_i, \mathfrak{w}_i) \in [0,S] \times \mathfrak{W}$, $i = 1, \dots, N$, such that $s_1, ..., s_N \in \mathrm{SD}\{f(\cdot,\cdot,\hat{w}(\cdot))\} \cap \mathrm{SD}\{f(\cdot,\cdot,\mathfrak{w}_1)\} \cap \ldots \cap \mathrm{SD}\{f(\cdot,\cdot,\mathfrak{w}_N)\} \cap \mathcal{L}\{c(\cdot)\}$, $0 < s_1 < \ldots, < s_N \leq S$. To conclude the proof we have to show the validity of (iii) in the whole control value set $\mathfrak{W}$ and almost all times. This is achieved through standard non-empty intersection arguments borrowed from those utilized, for instance, in [44] to prove the Maximum Principle.

7.2.1. *The case of a finite subset of controls*

Let us consider a *finite subset* of control values $\bar{\mathfrak{W}} \subseteq \mathfrak{W}$ and let us set

$$E(\bar{\mathfrak{W}}) := \bigcap_{w \in \bar{\mathfrak{W}}} \mathrm{SD}\{f(\cdot, \hat{y}(\cdot), \mathfrak{w})\} \bigcap \mathrm{SD}\{f(\cdot, \hat{y}(\cdot), \hat{w}(\cdot))\} \bigcap \mathcal{L}\{c(\cdot)\} \quad \Big(\subset [0,S] \Big).$$

Since $\bar{\mathfrak{W}}$ is finite, $E(\bar{\mathfrak{W}})$ has measure equal to S. Therefore, by Lusin's theorem we can write

$$E(\bar{\mathfrak{W}}) = \bigcup_{j=0}^{\infty} E_j, \tag{7.34}$$

where E_0 has zero measure and, for every j, the set E_j is compact, and, for every $w \in \bar{\mathfrak{W}}$, the restrictions to E_j of the map

$$s \mapsto r^{\mathfrak{w}}(s) := f(s, \hat{y}(s), \mathfrak{w}) - f(s, \hat{y}(s), \hat{w}(s)),$$

is continuous.

For every integer j, let D_j be the set of *density points* [14] of E_j. Since, for every natural number j, E_j and D_j have the same measure, one obtains that $meas(E(\bar{\mathfrak{W}})) = meas(D)$,[15] where we have set $D := \cup_{j=0}^{\infty} D_j$.

Now let F be an arbitrary, non-empty, subset of $D \times \bar{\mathfrak{W}}$, and let us define the subset $\Lambda(F, \bar{\mathfrak{W}}) \subseteq (T_{\hat{y}(S)}\mathcal{M})^*$ by setting

$$\Lambda(F, \bar{\mathfrak{W}}) := \left\{ \bar{\lambda} \in (T_{\hat{y}(S)}\mathcal{M})^*, \quad |\bar{\lambda}| = 1, \quad \bar{\lambda} \text{ verifies } \ \mathbf{(P)}_F \right\},$$

where property $\mathbf{(P)}_F$ is as follows:

Property $\mathbf{(P)}_F$: The pair $(\hat{y}, \lambda) \in W^{1,1}([0,S]; T^*\mathcal{M})$ is a lift of $\hat{y}(\cdot)$ such that:

(1) $\lambda(S) = \bar{\lambda} \in -C^{\perp}$;

(2) $\dfrac{d\lambda}{ds} = -\lambda \cdot \dfrac{\partial f}{\partial y}(s, \hat{y}(s), \hat{w}(s)), \qquad$ a.e $s \in [0,S]$;

(3) $\lambda(s) \cdot f(s, \hat{y}(s), \hat{w}(s)) \geq \lambda(s) \cdot f(s, \hat{y}(s), \mathfrak{w}), \qquad$ for every $(s, \mathfrak{w}) \in F$.

---

[14] We recall that an element $t \in B \subset \mathbb{R}$ is a *density point for B* if $\lim_{\delta \to 0+} \frac{meas([t-\delta, t+\delta])}{2\delta} = 1$.

[15] For every measurable subset $A \subseteq [0,S]$, we use $meas(A)$ to denote the Lebesgue measure of A.

Notice that, for every subset $F \in D \times \bar{\mathfrak{W}}$, $\Lambda(F, \bar{\mathfrak{W}})$ is compact and, moreover,

$$\Lambda(F_1 \cup F_2, \bar{\mathfrak{W}}) = \Lambda(F_1, \bar{\mathfrak{W}}) \cap \Lambda(F_2, \bar{\mathfrak{W}})$$

for all $F_1, F_2 \in D \times \bar{\mathfrak{W}}$.

We have already proved that $\Lambda(F, \bar{\mathfrak{W}}) \neq \emptyset$ in the case of a *finite* F having the form

$$F = \{(s_1, \mathfrak{w}_1), ..., (s_m, \mathfrak{w}_m)\}, \qquad 0 \leq s_1 < ... < s_i < ... < s_m < S, \quad \mathfrak{w}_i \in \bar{\mathfrak{W}}.$$

Claim: *One has* $\Lambda(F, \bar{\mathfrak{W}}) \neq \emptyset$ *even when* F *is an* arbitrary finite *subset of* $D \times \bar{\mathfrak{W}}$*, namely* F *can be written as*

$$F = \{(s_1, \mathfrak{w}_1), ..., (s_m, \mathfrak{w}_m)\} \qquad 0 \leq s_1 \leq ... \leq s_i \leq ... \leq s_m, \ \mathfrak{w}_i \in \bar{\mathfrak{W}}.$$

Indeed, every s_i belongs to a suitable D_h, which can be labelled as $D_{h(i)}$. Since $D_{h(i)}$ is made of density points, there exist sequences $(s_{i,j})$ such that

$$s_{i,j} \in D_{h(i)} \quad \forall j, \qquad s_i = \lim_{j \to \infty} s_{i,j},$$

and

$$s_{1,j} < ... < s_{m,j} \quad \forall j \in \mathbb{N}.$$

Set $F_j = \{(s_{i,j}, \mathfrak{w}_1), ..., (s_{m,j}, \mathfrak{w}_m)\}$ –so that $\Lambda(F_j, \bar{\mathfrak{W}}) \neq \emptyset$– and choose $\bar{\lambda}_j \in \Lambda(F_j, \bar{\mathfrak{W}})$. Since $|\bar{\lambda}_j| = 1$ for all j, by possibly taking a subsequence we can assume that $(\bar{\lambda}_j)_{j \in \mathbb{N}}$ converges to some $\bar{\lambda}$. For every $s \in [s_1, S]$, define the lifts $(\hat{y}, \bar{\lambda}), (\hat{y}, \bar{\lambda}_j) \in W^{1,1}([s_1, S]; T\mathcal{M})$ of $\hat{y}$ such that $\bar{\lambda}(S) = \bar{\lambda}$, $\bar{\lambda}_j(S) = \bar{\lambda}_j$ and both satisfying the equation

$$\frac{d\lambda}{ds} = -\lambda \cdot \frac{\partial f}{\partial y}(s, \hat{y}(s), \hat{w}(s)), \qquad \text{a.e } s \in [s_1, S].$$

The mapping $s \mapsto \bar{\lambda}_j(s)$ satisfies the inequality

$$\bar{\lambda}_j(s_{i,j}) \cdot f(s_{i,j}, \hat{y}(s_{i,j}), \hat{w}(s_{i,j})) \geq \bar{\lambda}_j(s_{i,j}) \cdot f(s_{i,j}, \hat{y}(s_{i,j}), \mathfrak{w})$$

for all $j \in \mathbb{N}$, every $i = 1, \ldots, m$ and $\mathfrak{w} \in \bar{\mathfrak{W}}$. Since, for every $i = 1, \ldots m$, the map $s \mapsto r^{\mathfrak{w}_i}(s) := f(s, \hat{y}(s), \mathfrak{w}_i) - f(s, \hat{y}(s), \hat{w}(s))$ is continuous on $D_{h(i)}$, the function $s \mapsto \bar{\lambda}_j(s) \cdot r^{\mathfrak{w}_i}(s)$ is also continuous on $D_{h(i)}$, so passing to the limit we can conclude that

$$\bar{\lambda}(s_i) \cdot f(s_i, \hat{y}(s_i), \hat{w}(s_i)) \geq \bar{\lambda}(s_i) \cdot f(s_i, \hat{y}(s_i), \mathfrak{w})$$

for every $i = 1, ..., m$ and $\mathfrak{w} \in \bar{\mathfrak{W}}$. Since one also has $0 \neq \bar{\lambda} = \bar{\lambda}(S) \in -C^{\perp}$, the claim is proved.

7.2.2. The general case of an infinite control set

Up to now we have shown that, *if* $\bar{\mathfrak{W}}$ *is finite*, and $F \subset D \times \bar{\mathfrak{W}}$ is *finite* —and we write $card(F) < \infty$—, then $\Lambda(F, \bar{\mathfrak{W}})$ is a nonempty compact set. We now conclude the proof through a standard non-empty intersection argument (see e.g. [44]). If we take a finite family $F^1, \dots, F^r \subset D \times \bar{\mathfrak{W}}$ such that $card(F_i) < \infty$ for every $i = 1, \dots, r$, one has

$$\Lambda(F^1, \bar{\mathfrak{W}}) \cap \cdots \cap \Lambda(F^r, \bar{\mathfrak{W}}) = \Lambda(F^1 \cup \cdots \cup F^r, \bar{\mathfrak{W}}) \neq \emptyset,$$

(for $card(F^1 \cup \cdots \cup F^r) < \infty$). Hence,

$$\{\Lambda(F, \bar{\mathfrak{W}}) \quad | \quad F \subset D \times \bar{\mathfrak{W}}, \quad \mathrm{card} F < \infty\}$$

is *a family of compact subsets such that each finite intersection is nonempty.* This implies that the (infinite) intersection of *all* $\Lambda(F, \bar{\mathfrak{W}})$ such that $cardF < \infty$ is nonempty. Therefore

$$\Lambda(D \times \bar{\mathfrak{W}}, \mathfrak{W}) = \Lambda\left(\bigcup_{card(F)<\infty} F, \mathfrak{W}\right) = \bigcap_{card(F)<\infty} \Lambda(F, \bar{\mathfrak{W}}) \neq \emptyset.$$

To end the proof in the general case when $card(\mathfrak{W})$ is infinite, for any arbitrary subset $\hat{\mathfrak{W}} \subseteq \mathfrak{W}$ define

$$\Lambda(\hat{\mathfrak{W}}) := \left\{\bar{\lambda} \in (T_{\hat{y}(S)}\mathcal{M})^*, \quad |\bar{\lambda}| = 1, \quad \bar{\lambda} \text{ verifies } \mathbf{(PP)}_{\hat{F}}\right\},$$

where property $\mathbf{(PP)}_{\hat{F}}$ is as follows:

Property $\mathbf{(PP)}_{\hat{F}}$: The pair $(\hat{y}, \lambda) \in W^{1,1}([0, S]; T^*\mathcal{M})$ is a lift of $\hat{y}(\cdot)$ such that:

(1) $\lambda(S) = \bar{\lambda} \in -C^{\perp}$;

(2) $\dfrac{d\lambda}{ds} = -\lambda \cdot \dfrac{\partial f}{\partial y}(s, \hat{y}(s), \hat{w}(s)), \qquad$ a.e $s \in [0, S]$;

(3) For each $\mathfrak{w} \in \hat{\mathfrak{W}}$, there exists a subset of full measure $I_{\mathfrak{w}} \subseteq [0, S]$ such that

$$\lambda(s) \cdot f(s, \hat{y}(s), \hat{w}(s)) \geq \lambda(s) \cdot f(s, \hat{y}(s), \mathfrak{w})$$

for every $s \in I_{\mathfrak{w}}$, $\mathfrak{w} \in \hat{\mathfrak{W}}$.

So, proving Theorem 5.2 is equivalent to showing that

$$\Lambda(\mathfrak{W}) \neq \emptyset. \tag{7.35}$$

Since

$$\Lambda(\mathfrak{W}) = \bigcap_{card(\hat{\mathfrak{W}})<\infty} \Lambda(\hat{\mathfrak{W}}), \tag{7.36}$$

once again we have to show that the (possibly infinite) family

$$\Big\{\Lambda(\hat{\mathfrak{W}}), \quad card(\hat{\mathfrak{W}}) < \infty\Big\}$$

has non-empty intersection. This can easily achieved by the same arguments as above. Indeed, $\Lambda(\hat{\mathfrak{W}})$ is not empty and compact as soon as $\hat{\mathfrak{W}}$ is finite. Furthermore, for every $\mathfrak{W}_1, \mathfrak{W}_2 \subseteq \mathfrak{W}$ one has

$$\Lambda(\mathfrak{W}_1 \cup \mathfrak{W}_2) = \Lambda(\mathfrak{W}_1) \cap \Lambda(\mathfrak{W}_2).$$

In particular, the family $\Big\{\Lambda(\hat{\mathfrak{W}}) : \ card(\hat{\mathfrak{W}}) < \infty\Big\}$ is made of compact subsets and satisfies the finite intersection property, that is, the intersection of any finite finite subfamily $\Big\{\Lambda(\hat{\mathfrak{W}}) : \ card(\hat{\mathfrak{W}}) < \infty\Big\}$ is not empty. Therefore, it has non-empty intersection, namely

$$\Lambda(\mathfrak{W}) = \bigcap_{card(\hat{\mathfrak{W}})<\infty} \Lambda(\hat{\mathfrak{W}}) \neq \emptyset \ .$$

This concludes the proof of Theorem 5.2.

8. An application to non-convex, unbounded, problems

Impulsive optimal control problems –where the dynamics is *unbounded*– have been extensively studied together with their applications [4, 5, 6, 10, 12, 13, 16, 17, 22, 23, 25, 29, 30, 42, 51]. The space-time representation (see (8.38) below) can be regarded as an *extension* of unbounded control systems. An important case is the one of a minimum problem with a control-affine dynamics:

$$(P)\left\{\begin{array}{l} \qquad \text{Minimize}\, h(t_2, x(t_2), \eta(t_2)) \\ \\ \text{over } \ t_2 \in \mathbb{R}, \ t_2 > t_1, \ (x, \eta, u)(\cdot) \in AC([t_1, t_2], \mathcal{M} \times \mathbb{R}) \times L^1([t_1, t_2], U) \\ \quad \text{such that} \\ \left\{\begin{array}{l} \dfrac{dx}{dt}(t) = f(t, x(t)) + \displaystyle\sum_{j=1}^{m} g_j(t, x(t))u^j(t) \quad \text{a.e. t} \in [\mathrm{t}_1, \mathrm{t}_2] \\ \dfrac{d\eta}{dt}(t) = |u(t)| \quad \text{a.e. t} \in [\mathrm{t}_1, \mathrm{t}_2] \end{array}\right. \\ (x(t_1), \eta(t_1)) = (\bar{x}, 0), \quad (t_2, x(t_2), \eta(t_2)) \in \bar{\mathfrak{T}} \times [0, K] \end{array}\right. ,$$

Here the set U where the controls u take values is *unbounded*. Furthermore, the state x range over a n-dimensional Riemannian manifold $\mathcal{M}$ of class C^2, and the time-dependent vector fields $f, g_1, \ldots, g_m$ are of class C^1 in x, continuous in t, and uniformly bounded by a L^1 map. Moreover, the cost $h : \mathbb{R} \times \mathcal{M} \times \mathbb{R} \to \mathbb{R}$ is a continuous function, $(t_1, \bar{x}) \in \mathbb{R} \times \mathcal{M}$ is a fixed initial condition, K is a non negative fixed constant, possibly equal to $+\infty$, and the end-point constraint $\bar{\mathfrak{T}} \subseteq \mathbb{R} \times \mathbb{R}^n$ is a closed subset. Notice incidentally that the function $\eta(t)$ coincides with the L^1-norm of the control function $u := (u^1, u^2, ..., u^m)$ on the interval $[t_1, t]$.

The gap-abnormality criterion for this kind of systems (where one considers the space-time extension (8.38) below) has been already investigated in the case when the set of controls U is a convex cone [33]. Actually, thanks to Theorem 5.2 (see also [36]), the main result in [33] can be extended to the case in which the state ranges on a Riemannian manifold. However the generalization made possible by Theorem 5.2 allows one to go much further. Indeed, in what follows we are able to deduce from Corollary 5.2 that the gap-abnormality criterium holds true also in the situation when the control set U is unbounded but *neither is convex nor is a cone.*

More precisely, we will consider the following two cases:

CASE (I) **(Space-time convex extension)** *The controls take values on a (necessarily unbounded) subset $U \subseteq \mathbb{R}^m$ such that $\mathbf{co}\, U$ is a (convex) cone, where we have used $\mathbf{co}\, E$ to denote the convex hull of a subset $E \subseteq \mathbb{R}^m$;*

For instance, one could consider the set $U = \mathbb{N}^m$, so that $\mathbf{co}(U) = [0, +\infty[^m$.

CASE (II) **(Space-time non-convex extension)** *The controls take values on a (necessarily unbounded) subset $U \subseteq \mathbb{R}^m$ such that*

$$(r, u) \in [0, +\infty[\times U \Longrightarrow \exists \rho > r \text{ s.t. } \rho u \in U \tag{8.37}$$

Notice that, if for a given set E we consider the $\mathbf{conic}(E) := \{re \mid (r, e) \in [0, +\infty[\times E\}$ —a cone which we call the *conic envelope* of E—, one has that

$$\inf_{u \in U} d(u, \mathbf{conic}(U)) = 0.$$

For instance, one could consider the set $U = \{(n^2, 0), (0, -m^3) \mid m, n \in \mathbb{N}\}$, so that $\mathbf{conic}\,(U) = [0, +\infty[\times \{0\} \bigcup \{0\} \times] - \infty, 0]$.

Remark 8.1. We will treat CASE (I) in detail, describing the extension to the convex space-time system obtained by both convexification of the dynamics and the closure of suitably reparameterized processes. Instead, we will only suggest the needed changes to deal with CASE (II), where the only extension comes from reparameterization. However, CASE (II) is somehow more significative, in that it marks the most important improvement with respect to the former literature initiated by Warga's work. Indeed, in this case not only the original dynamics but *also the extended dynamics is non-convex.* This can be of interest in those application where the convexification of the dynamics is not needed (for instance because one gets existence of minima without invoking convexification).

8.1. CASE (I) *(Space-time convex extension)*

In order to formulate this problem by means of the terminology adopted in Theorem 5.2, we need to embed our system into a suitably extended one. To this aim we need to perform both a 'compactification' (to manage unboundedness) and a 'convexification'. Let us begin by setting

$$\begin{array}{ll}
A := \{a = (a^1, \dots a^{n+3}) \in [0,1]^{n+3},\ \sum_{i=1}^{n+3} a^i = 1\} & \mathcal{A} := L^1([0,S], A) \\
\mathbf{W} := \{(w^0, w) \in [0, \infty) \times \mathbf{co}\, U : \quad w^0 + |w| = 1\} & \hat{\mathcal{W}} := L^1([0,S], \mathbf{W}) \\
\mathbf{V} := \{(v^0, v) \in (0, \infty) \times U : \quad v^0 + |v| = 1\} & \hat{\mathcal{V}} := L^1([0,S], \mathbf{V}) \\
D := [-0.5, 0.5] & \mathcal{D} := L^1\big([0,S], D\big)
\end{array}$$

$$\mathcal{W} := \mathcal{A} \times (\hat{\mathcal{W}})^{n+3} \times \mathcal{D}, \qquad \mathcal{V} := \{(1, 0, \ldots, 0)\} \times (\hat{\mathcal{V}})^{n+3} \times \mathcal{D},$$

and let us consider the optimal control problem

$$(P)^h_{\mathcal{W}} \begin{cases} \qquad \text{Minimize}\, h\big(z^0(\hat{S}), z(\hat{S}), \nu(\hat{S})\big) \\ \text{over } \; (z^0, z, \nu, a, (w_1^0, w_1), \ldots, (w_{n+3}^0, w_{n+3}), d)(\cdot) \in AC([0, \hat{S}], \mathbb{R} \times \mathcal{M} \times \mathbb{R}) \times \mathcal{W} \times \mathcal{D} \\ \text{s. t., } \; \text{for a.e. } \; \mathrm{s} \in [0, \hat{\mathrm{S}}], \\ \begin{cases} \dfrac{dz^0}{ds}(s) = (1 + d(s)) \displaystyle\sum_{i=1}^{n+3} a^i(s) w_i^0(s) \\ \dfrac{dz}{ds}(s) = (1 + d(s)) \displaystyle\sum_{i=1}^{n+3} a^i(s) \Big(f(z^0(s), z(s)) w_i^0(s) + \sum_{j=1}^{m} g_j(z^0(s), z(s)) w_i^j(s) \Big) \\ \dfrac{d\nu}{ds}(s) = (1 + d(s)) \displaystyle\sum_{i=1}^{n+3} a^i(s) |w_i(s)| \end{cases} \\ (z^0(0), z(0), \nu(0)) = (0, \bar{x}, 0), \qquad (z^0(\hat{S}), z(\hat{S}), \nu(\hat{S})) \in \bar{\mathfrak{T}} \times [0, K] \end{cases} . \tag{8.38}$$

Accordingly, a pair

$$\Big(\big(z^0, z, \nu\big) \;,\; \big(a, (w_1^0, w_1), \ldots, (w_{n+3}^0, w_{n+3}), d\big) \Big)$$

such that $\big(z^0, z, \nu\big)$ is the solution of the above control system corresponding to the *control* $\big(a, (w_1^0, w_1), \ldots, (w_{n+3}^0, w_{n+3}), d\big)$ is called *a process of* $(P)^h_{\mathcal{W}}$. The embedding of the problem (P) into $(P)^h_{\mathcal{W}}$ is as follows: fix $\hat{S} > 0$, and, for every control $u : [t_1, t_2^u] \to U$, consider the function $\sigma_u : [t_1, t_2^u] \to [0, \hat{S}]$ defined by

$$\sigma_u(t) := \frac{\hat{S}}{t_2^u + \|u\|_1} \int_{t_1}^{t} \Big(1 + |u(\tau)| \Big)\, d\tau = \frac{\hat{S}}{t_2^u + \|u\|_1} (t + \eta(t)). \tag{8.39}$$

Then define $\mathcal{I} : \mathbb{R} \times AC([t_1, t_2], \mathcal{M} \times \mathbb{R} \times \mathbb{R}^m) \to \mathbb{R} \times AC([0, S], \mathbb{R} \times \mathcal{M} \times \mathbb{R} \times \mathbb{R} \times \mathbb{R}^m)$ by setting

$$\mathcal{I}(x, \eta, u) := \Big(\big((z^0, z, \nu\big) \;,\; \big(a, (w_1^0, w_1), \ldots, (w_{n+3}^0, w_{n+3}), d\big) \Big)$$

where, forall $s \in [0, \hat{S}]$ and all $i = 1, \ldots, m$,

$$(z^0, z, \nu)(s) := (id, x, \eta) \circ \sigma_u^{-1}(s), \qquad \forall\, s \in [0, \hat{S}],$$

$$a := (1, 0, \ldots, 0), \quad d := \frac{t_2^u + \|u\|_1}{\hat{S}} - 1, \quad (w_i^0, w_i) := \left(\frac{1}{1 + |u|} (1, u) \right) \circ \sigma_u^{-1}(s).$$

By a trivial use of the chain rule one gets the following result (see e.g. [4] for a similar embedding):

Lemma 8.1. *The embedding $\mathcal{I}$ is injective* [16]. *Moreover, the image space of the embedding $\mathcal{I}$ coincides with the set of all processes $\Big(\big((z^0,z,\nu\big)\ ,\ \big(a,(w_1^0,w_1),\dots,(w_{n+3}^0,w_{n+3}),d\big)\Big)$ such that*

$$\big(a,(w_1^0,w_1),\dots,(w_{n+3}^0,w_{n+3}),d\big)\in\mathcal{V}.$$

Thanks to Lemma 8.1 we can identify the original problem (P) with the problem

$$(P)^h_{\mathcal{V}}\left\{\begin{array}{l} \qquad\text{Minimize}\, h\big(z^0(\hat{S}),z(\hat{S}),\nu(\hat{S})\big)\\ \text{over }\ (z^0,z,\nu,a,(w_1^0,w_1),\dots,(w_{n+3}^0,w_{n+3}),d)(\cdot)\in AC([0,\hat{S}],\mathbb{R}\times\mathcal{M}\times\mathbb{R})\times\mathcal{V}\times\mathcal{D}\\ \text{s. t.,}\ \ \text{for a.e.}\ \ \mathrm{s}\in[0,\hat{\mathrm{S}}],\\ \left\{\begin{array}{l} \dfrac{dz^0}{ds}(s)=(1+d(s))\displaystyle\sum_{i=1}^{n+3}a^i(s)w_i^0(s)\\ \dfrac{dz}{ds}(s)=(1+d(s))\displaystyle\sum_{i=1}^{n+3}a^i(s)\Big(f(z^0(s),z(s))w_i^0(s)+\sum_{j=1}^{m}g_j(z^0(s),z(s))w_i^j(s)\Big)\\ \dfrac{d\nu}{ds}(s)=(1+d(s))\displaystyle\sum_{i=1}^{n+3}a^i(s)|w_i(s)| \end{array}\right.\\ (z^0(0),z(0),\nu(0))=(0,\bar{x},0),\qquad (z^0(\hat{S}),z(\hat{S}),\nu(\hat{S}))\in\bar{\mathfrak{T}}\times[0,K] \end{array}\right. .$$

We can now apply the theory developed in the previous sections. In view of the sufficient condition provided by Proposition 4.1, it is quite easy [17] to verify that *the family of controls $\mathcal{V}$ is abundant in $\mathcal{W}$ w.r.t. the dynamics of problem $(P)^h_{\mathcal{W}}$*. Therefore, by Theorem 5.2, one obtains the following infimum-gap result:

Theorem 8.1. *Consider a feasible extended process*

$$\Big(\big(y^0,y,\nu\big)\ ,\ \big(\hat{a},(\hat{w}_1^0,\hat{w}_1),\dots,(\hat{w}_{n+3}^0,\hat{w}_{n+3}),\hat{d}\big)\Big),\quad \hat{d}\equiv 0,$$

and assume that it satisfies the infimum gap condition. Then, for all approximating cones C to $\mathfrak{T}:=\bar{\mathfrak{T}}\times[0,K]$ at $(\hat{y}^0,\hat{y},\hat{\nu})(\hat{S})$, there exist a number $\beta\le 0$, an absolutely continuous path $(\lambda^0,\lambda,\lambda^\nu)\in W^{1,1}([0,\hat{S}];\mathbb{R}^{(1+n+1)})$ and a zero-measure subset I_0 such that the following conditions hold true:

(*i*) $(\lambda^0,\lambda,\lambda^\nu)\neq 0$;

[16] Notice that the injectivity is a consequence of the fact that $w_i^0(s)+|w_i(s)|=1$ for a.e. $s\in[0,S]$ and for $i=1,\dots,m$.

[17] In particular the density is a consequence of standard relaxation results for ode's (see, e.g., [46], Theorem 2.7.2). The concatenation property of $\mathcal{V}$ is instead easy to verify. So the abundance of $\mathcal{V}$ into $\mathcal{W}$ follows from Proposition 4.1.

$$(ii.1)\quad \frac{d\lambda^0}{ds}(s) = -\lambda(s)\cdot\left[\sum_{i=1}^{n+3}\hat a^i(s)\left(\frac{\partial f}{\partial y^0}(\hat y^0(s),\hat y(s))\hat w_i^0(s) + \sum_{j=1}^m \frac{\partial g_j}{\partial y^0}(\hat y^0(s),\hat y(s))\hat w_i^j(s)\right)\right],$$

$$(ii.2)\quad \frac{d\lambda}{ds}(s) = -\lambda(s)\cdot\left[\sum_{i=1}^{n+3}\hat a^i(s)\left(\frac{\partial f}{\partial y}(\hat y^0(s),\hat y(s))\hat w_i^0(s) + \sum_{j=1}^m \frac{\partial g_j}{\partial y}(\hat y^0(s),\hat y(s))\hat w_i^j(s)\right)\right],$$

a.e. $s\in[0,\hat S]$;

$$(iii)\quad (1+d)\sum_{i=1}^{n+3} a^i\Big[\lambda^0(s)w_i^0 + \lambda(s)\cdot\Big(\hat f(s)w_i^0 + \sum_{j=1}^m \hat g_j(s)w_i^j\Big) + \beta|w_i|\Big]$$
$$\le \sum_{i=1}^{n+3}\hat a^i(s)\Big[\lambda^0(s)\hat w_i^0(s) + \lambda(s)\cdot\Big(\hat f(s)\hat w_i^0(s) + \sum_{j=1}^m \hat g_j(s)\hat w_i^j(s)\Big) + \beta|\hat w_i(s)|\Big],$$

for every $(w^0,w,d)\in\mathbf{W}\times[-0.5,0.5]$ *and* $s\in[0,\hat S]\backslash I_0$; [18]

$$(iv)\quad (\lambda^0(\hat S),\lambda(\hat S),\beta)\in -C^\perp.$$

8.2. Case (ii) *(Space-time non-convex extension)*

Let us recall that we are assuming (8.37), namely

$$(r,u)\in[0,+\infty[\times U \Longrightarrow \exists\rho>r \text{ s.t. } \rho u\in U.$$

Unlike the previous case, we are not going to convexify the dynamics, while we will consider just the impulsive extension. Without repeating all the steps, we just observe that the sought extension is obtained by neglecting the sets A and $\mathcal{A}$, and by replacing $\mathbf{W}$ with the (generally non-convex) set $\mathbf{W}^{nc} := \{(w^0,w)\in[0,\infty)\times U:\quad w^0+|w|=1\}$, respectively. In turn, problem $(P)^h_{\mathcal{W}}$ simplifies into the following non-convex problem $(P^{nc})^h_{\mathcal{W}}$:

$$(P^{nc})^h_{\mathcal{W}}\begin{cases} \text{Minimize } h\big(z^0(\hat S), z(\hat S), \nu(\hat S)\big) \\ \text{over } (z^0,z,\nu,(w^0,w),d)\in AC([0,\hat S],\mathbb{R}\times\mathcal{M}\times\mathbb{R})\times\mathcal{W}\times\mathcal{D} \\ \text{s. t., for a.e. } \mathrm{s}\in[0,\hat{\mathrm{S}}], \\ \begin{cases} \dfrac{dz^0}{ds}(s) = (1+d(s))w^0(s) \\ \dfrac{dz}{ds}(s) = (1+d(s))\Big(f(z^0(s),z(s))w^0(s) + \displaystyle\sum_{j=1}^m g_j(z^0(s),z(s))w^j(s)\Big) \\ \dfrac{d\nu}{ds}(s) = (1+d(s))|w(s)| \end{cases} \\ (z^0(0),z(0),\nu(0)) = (0,\bar x,0), \qquad (z^0(\hat S),z(\hat S),\nu(\hat S))\in\bar{\mathfrak{T}}\times[0,K] \end{cases}. \tag{8.40}$$

[18] We have set $\hat f(s) := f(\hat y^0(s),\hat y(s))$, $\hat g_j(s) := g_j(\hat y^0(s),\hat y(s))$, for all $s\in[0,\hat S]$ and $i=1,\dots,m$

The other objects simplify accordingly, and, still because of the concatenation property (and of the d_f-density of $\mathcal{V}$ in $\mathcal{W}$), *the subfamily* $\mathcal{V}$ *of controls is abundant in the family* $\mathcal{W}$. Therefore, by applying the infimum-gap result stated in Theorem 5.2 we get:

Theorem 8.2. *Consider a feasible extended process*

$$\Big((y^0, y, \nu)\ ,\ ((\hat{w}^0, \hat{w}), \hat{d})\Big), \quad \hat{d} \equiv 0,$$

and assume that it satisfies the infimum gap condition. Then, for all QDQ approximating cones C *to* $\mathfrak{T} := \bar{\mathfrak{T}} \times [0, K]$ *at* $(\hat{y}^0, \hat{y}, \hat{\nu})(\hat{S})$*, there exist a number* $\beta \leq 0$*, an absolutely continuous path* $(\lambda^0, \lambda, \lambda^\nu) \in W^{1,1}([0, \hat{S}]; \mathbb{R}^{(1+n+1)})$ *and a zero-measure subset* I_0 *such that the following conditions hold true:*

(i) $(\lambda^0, \lambda, \lambda^\nu) \neq 0$;

(ii) $$\frac{d\lambda^0}{ds}(s) = -\lambda(s) \cdot \left(\frac{\partial f}{\partial y^0}(\hat{y}^0(s), \hat{y}(s))\hat{w}^0(s) + \sum_{j=1}^{m} \frac{\partial g_j}{\partial y^0}(\hat{y}^0(s), \hat{y}(s))\hat{w}^j(s)\right),$$
$$\frac{d\lambda}{ds}(s) = -\lambda(s) \cdot \left(\frac{\partial f}{\partial y}(\hat{y}^0(s), \hat{y}(s))\hat{w}_i^0(s) + \sum_{j=1}^{m} \frac{\partial g_j}{\partial y}(\hat{y}^0(s), \hat{y}(s))\hat{w}_i^j(s)\right),$$

for a.e. $s \in [0, \hat{S}]$*;*

(iii) $$(1+d)\left[\lambda^0(s)w^0 + \lambda(s) \cdot \Big(\hat{f}(s)w^0 + \sum_{j=1}^{m} \hat{g}_j(s)w^j\Big) + \beta|w|\right]$$
$$\leq \left[\lambda^0(s)\hat{w}^0(s) + \lambda(s) \cdot \Big(\hat{f}(s)\hat{w}^0(s) + \sum_{j=1}^{m} \hat{g}_j(s)\hat{w}^j(s)\Big) + \beta|\hat{w}(s)|\right],$$

for every $(w^0, w, d) \in \mathbf{W}^{nc} \times [-0.5, 0.5]$ *and* $s \in [0, \hat{S}] \backslash I_0$*;*

(iv) $(\lambda^0(\hat{S}), \lambda(\hat{S}), \beta) \in -C^\perp$.

9. Why abundance is crucial: an example

The following example, which is due to H.J. Sussmann,[19] shows how the abundance hypothesis plays crucial for the validity of Theorem 5.3.

Consider the families of controls $\mathcal{V} \subset \mathcal{W}$ defined as

$$\mathcal{W} := L^1([0,1], [0,5]), \qquad \mathcal{V} := \left\{v \in \mathcal{W} : \int_0^1 v(s)ds \neq 1\right\},$$

[19] Personal communication.

and the optimal control problems

$$(P)_{\mathcal{V}} \begin{cases} \text{Minimize } y(1) \\ \text{over processes } (y,v)(\cdot) \in W^{1,1}([0,1],\mathbb{R}) \times \mathcal{V} \\ \dfrac{dy}{ds}(s) = v(s), \qquad \text{a.e. s} \in [0,1] \\ y(0)=0, \qquad y(1)=1, \end{cases},$$

$$(P)_{\mathcal{W}} \begin{cases} \text{Minimize } y(1) \\ \text{over processes } (y,w)(\cdot) \in W^{1,1}([0,1],\mathbb{R}) \times \mathcal{W} \\ \dfrac{dy}{ds}(s) = w(s), \qquad \text{a.e. s} \in [0,1] \\ y(0)=0, \qquad y(1)=1, \end{cases}.$$

The process $(\hat{y},\hat{w})(s) := (s,1)$ is a minimizer of the extended problem $(P)_{\mathcal{W}}$, with cost equal to 1. If we restrict the controls to the *original* family of controls $\mathcal{V}$, the cost of the problem raises to $+\infty$, since every solution $y[v]$ with $v \in \mathcal{V}$ fails to be feasible. In other words the process $(\hat{y},\hat{w})$ satisfies the infimum gap condition.

By applying the Pontryagin's Maximum Principle to the minimizer $(\hat{y},\hat{w})$ of $(P)_{\mathcal{W}}$, we get that there exist multipliers $(\lambda(\cdot),\lambda_c) \neq (0,0)$ such that

$$\frac{d\lambda}{ds}(s) \equiv 0, \qquad \lambda(s)\mathfrak{w} \leq \lambda(s) \qquad \forall \mathfrak{w} \in [0,5], \quad s \in [0,1].$$

In particular this implies $\lambda(s) \equiv 0$ and $\lambda_c > 0$. Therefore, if we set $h(y) := y$ for every $y \in \mathbb{R}$, the process $(\hat{y},\hat{w})$ turns out to be a *normal h-extremal.* Therefore, in view of Theorem 5.2 the set $\mathcal{V}$, though being dense in $\mathcal{W}$, cannot be abundant in $\mathcal{W}$. As a matter of fact, one can easily find a positive integer N, $\delta > 0$ and $N+1$ controls $w, w_1, \dots, w_N$ for which $\theta^{\delta}_{w,w_1,\dots,w_N} : \Gamma_N \to \mathcal{V}$ verifying the properties of Definition 4.1 does not exist. Indeed, consider $\Gamma_1 (= [0,1])$, $w(s) := 0, w_1(s) := 2$, $\forall s \in [0,1]$, $\delta > 0$, and take any mapping $\theta^{\delta} : [0,1] \to \mathcal{V}$. In view of Definition 4.1, one has

$$\lim_{\delta \to 0} \int_0^1 \theta^{\delta}(\gamma)(s)ds = \int_0^1 w(s) + \gamma\big(w_1(s) - w(s)\big) = 2\gamma \qquad \forall \gamma \in \Gamma_1.$$

Then, for every δ sufficiently small, either there exists a $\gamma_{\delta} \in [0,1]$ such that

$$\int_0^1 \theta^{\delta}(\gamma_{\delta})(s)ds = 1,$$

or the map $\gamma \mapsto \int_0^1 \theta^{\delta}(\gamma)(s)ds$ is not continuous. Since the former case is ruled out by the fact that the map $\theta^{\delta}(\cdot)(s)$ has to take values in $\mathcal{V}$, the map $\gamma \mapsto \int_0^1 \theta^{\delta}(\gamma)(s)ds$ is not continuous, so providing a contradiction.

Acknowledgement

The authors would like to thank Hector Sussmann for many useful discussions and for suggesting the use of the notion of abundance. The authors would like also to thank Monica Motta for the careful reading of the manuscript.

Grants

This research was partially supported by the INdAM, G.N.A.M.P.A. with the project "Problemi di controllo estesi: gap, condizioni di ordine superiore e funzioni di Lyapunov".

References

[1] M. S. Aronna, J. F. Bonnans, P. Martinon, *A well-posed shooting algorithm for optimal control problems with singular arcs*, J. Optim. Theory Applic. vol. 158, no. 2, 2013, pp. 419-459. https://doi.org/10.1007/s10957-012-0254-8

[2] M.S. Aronna, A. Boccia, C.M. Campos, M. Mazzola, M. Palladino, T. Scarinci, F.J. Silva, D. Tonon, L. Van Nguyen, *Optimality conditions (In Pontryagin form)*, Lecture Notes in Mathematics, vol. 2180, 2017, pp. 1-125. https://doi.org/10.1007/978-3-319-60771-9_1

[3] M.S. Aronna, M. Motta, F. Rampazzo, *Infimum gaps for limit solutions*, Set-Valued Var. Anal. 23, no. 1, 2015, pp. 3–22. https://doi.org/10.1007/s11228-014-0296-1

[4] M.S. Aronna, F. Rampazzo, *L^1 limit solutions for control systems.* J. Differential Equations, vol. 258, no. 3, 2015, pp. 954–979. https://doi.org/10.1016/j.jde.2014.10.013

[5] A. Arutyunov, D. Karamzin and F. L. Pereira, *Pontryagin's maximum principle for constrained impulsive control problems*, Nonlinear Anal., vol. 75, no. 3, 2012, pp. 1045-1057. https://doi.org/10.1016/j.na.2011.04.047

[6] A. Arutyunov, D. Karamzin and F. L. Pereira, *A nondegenerate maximum principle for the impulse control problem with state constraints.* SIAM J. Control Optim. 43, no. 5, 2005, pp. 1812-1843. https://doi.org/10.1137/S0363012903430068

[7] D. Azimov, R. Bishop, *New trends in astrodynamics and applications: optimal trajectories for space guidance.* Ann. New York Acad. Sci., vol 1065, no. 1, 2005, pp. 189–209. https://doi.org/10.1196/annals.1370.002

[8] J. F. Bonnans, A. Shapiro, *Perturbation Analysis of Optimization Problems*, Springer, New York, 2000. https://doi.org/10.1007/978-1-4612-1394-9

[9] A. Bressan, *On the Intersection of a Clarke Cone with a Boltyanskii Cone*, SIAM J. Control Optim., vol. 45, no. 6, 2007, pp. 2054-2064. https://doi.org/10.1137/060654852

[10] A. Bressan, F. Rampazzo, *On differential systems with vector-valued impulsive controls.* Boll. Un. Mat. Ital. VII Ser., B2, no. 3, 1988, pp. 641-656.

[11] A. Bressan, B. Piccoli, *Introduction to the mathematical theory of control.* AIMS Series on Applied Mathematics, 2. American Institute of Mathematical Sciences (AIMS), Springfield, MO, 2007.

[12] Aldo Bressan, *Hyper-impulsive motions and controllizable coordinates for Lagrangean systems,* Atti Accad. Naz. Lincei, Memorie, Serie VIII, Vol. XIX, 1991, pp. 197–246.

[13] Aldo Bressan, *On some control problems concerning the ski or swing,* Atti Accad. Naz. Lincei, Memorie, Serie IX, Vol. I, 1991, pp. 147–196.

[14] A. Catllá, D. Schaeffer, T. Witelski, E. Monson, A. Lin, *On spiking models for synaptic activity and impulsive differential equations,* SIAM Rev., vol 50, no 3, 2008, pp. 553-569. https://doi.org/10.1137/060667980

[15] A. Cellina, A. Ferriero and E.M. Marchini, *Reparametrizations and approximate values of integrals of the calculus of variations,* Journal of Differential Equations, vol.193, no. 2, 2003, pp. 374-384. https://doi.org/10.1016/S0022-0396(02)00176-6

[16] V. A. Dykhta, *The variational maximum principle and quadratic conditions for the optimality of impulse and singular processes.* (Russian) Sibirsk. Mat. Zh. 35 (1994), no. 1, 70–82, ii; translation in Siberian Math. J. 35, no. 1, 1994, pp. 65–76.

[17] V. A. Dykhta *Second order necessary optimality conditions for impulse control problem and multiprocesses.* Singular solutions and perturbations in control systems (Pereslavl-Zalessky, 1997), 97–101, IFAC Proc. Ser., IFAC, Laxenburg, 1997.

[18] A.L. Dontchev, T. Zolezzi, *Well-posed optimization*, Springer-Verlag, New York, 1993. DOI: 10.1007/BFb0084195

[19] A. Fryszowsky, *Continuous selections for a class of non-convex multivalued maps*, Studia Math., vol. 76, 1983, pp. 163 - 174. DOI: 10.4064/sm-76-2-163-174

[20] P. Gajardo, H. Ramirez C., A. Rapaport, *Minimal time sequential batch reactors with bounded and impulse controls for one or more species,* SIAM J. Control Optim. 47 (6), 2008, pp. 2827-2856. https://doi.org/10.1137/070695204

[21] R. V. Gamkrelidze, *On Some Extremal Problems in the Theory of Differential Equations with Applications to Optimal Control*, SIAM J. Control Optim., vol. 3, 1965, pp. 106-128. https://doi.org/10.1137/0303010

[22] M. Guerra, A. Sarychev, *Fréchet generalized trajectories and minimizers for variational problems of low coercivity,* J. Dyn. Control Syst., vol. 21, no. 3, 2015, pp. 351–377. https://doi.org/10.1007/s10883-014-9231-x

[23] D.Y. Karamzin; V.A. de Oliveira, F.L. Pereira, G.N. Silva, *On the properness of an impulsive control extension of dynamic optimization problems,* ESAIM Control Optim. Calc. Var. 21, no. 3, 2015, pp. 857–875. https://doi.org/10.1051/cocv/2014053

[24] B. Kaskosz, *Extremality, controllability and abundant subsets of generalized control systems*, Journal of Optimization Theory and Applications, vol. 101, 1999, pp. 73-108. https://doi.org/10.1023/A:1021719027140

[25] A. C. Lai, M. Motta and F. Rampazzo, *Minimum restraint functions for unbounded dynamics: general and control-polynomial systems*, Pure and Applied Functional Analysis, vol. 1, no. 4, 2016, pp. 583-612.

[26] M. Lavrentiev, *Sur quelques problémes du calcul des variations*, Ann. Mat. Pura Appl. 4, 1927, pp. 9-28.

[27] B. Manià, *Sopra un esempio di Lavrentieff*, Boll. Un. Matem. Ital. 13, 1934, 147-153.

[28] C. Mariconda, *Contractibility and Fixed Point Property: the case of Decomposable Sets,* Nonlinear Analysis, T.M.A., vol. 17, no. 7, 1992, pp. 689-695. https://doi.org/10.1016/0362-546X(92)90007-2

[29] M. Miller, E. Y. Rubinovich, *Impulsive control in continuous and discrete-continuous systems.* Kluwer Academic/Plenum Publishers, New York, 2003. DOI: 10.1007/978-1-4615-0095-7

[30] M. Motta, F. Rampazzo, *Space-time trajectories of nonlinear systems driven by ordinary and impulsive controls*, Differential Integral Equations, vol. 8, no. 2, 1995, pp. 269-288.

[31] M. Motta, F. Rampazzo, *Dynamic programming for nonlinear systems driven by ordinary and impulsive controls*, SIAM J. Control Optim. 34, no. 1, 1996, pp. 199–225. https://doi.org/10.1137/S036301299325493X

[32] M. Motta, F. Rampazzo, *State-constrained control problems with neither coercivity nor L^1 bounds on the controls*, Ann. Mat. Pura Appl., vol. 177,, no. 4, 1999, pp. 117-142. https://doi.org/10.1007/BF02505906

[33] M. Motta, F. Rampazzo, R. B. Vinter, *Normality and Gap Phenomena in Optimal Unbounded Control*, ESAIM Control Optim. Calc. Var., vol. 24, no. 4, 2018, pp. 1645 - 1673. https://doi.org/10.1051/cocv/2018069

[34] M. Motta, C. Sartori, *On asymptotic exit-time control problems lacking coercivity*, ESAIM Control Optim. Calc. Var., vol. 20, no. 4, 2014, pp. 957-982. https://doi.org/10.1051/cocv/2014003

[35] C. Olech, *Decomposability as a substitute of convexity*, Lecture Notes in Mathematics "Multifunctions and Integrands", vol. 1091, 1984, pp. 193-205. https://doi.org/10.1007/BFb0098812

[36] M. Palladino, F. Rampazzo, *A No Infimum-Gap Criterion*, IEEE 58th Annual Conference on Decision and Control (CDC), 11-13 Dec. 2019, pp. 535-540. DOI: 10.1109/CDC40024.2019.9029870

[37] M. Palladino, R. B. Vinter, *Minimizers that are not also Relaxed Minimizers*, SIAM J. Control and Optim., vol. 52, no. 4, 2014, pp. 2164 - 2179. https://doi.org/10.1137/130909627

[38] M. Palladino, R. B. Vinter, *When are Minimizing Controls also Minimizing Relaxed Controls?*, Discrete Contin. Dyn. Syst. Series A, vol. 35, no. 9, 2015, pp. 4573 - 4592. DOI: 10.3934/dcds.2015.35.4573

[39] M. Palladino and R. B. Vinter, *When does relaxation reduce the minimum cost of an optimal control problem?*, IEEE 52nd Annual Conference on Decision and Control (CDC), 10-13 Dec. 2013, pp. 526-531. DOI: 10.1109/CDC.2013.6759935

[40] H. Schättler, U. Ledzewicz, *Geometric Optimal Control: Theory, Methods and Examples*, Springer, 2012. DOI: 10.1007/978-1-4614-3834-2

[41] G. Scorza Dragoni, *Un teorema sulle funzioni continue rispetto ad una e misurabili rispetto ad un'altra variabile*, (Italian) Rend. Sem. Mat. Univ. Padova 17, 1948, pp. 102-106.

[42] G. Silva, R. B. Vinter, *Measure driven differential inclusions,* J. Math. Anal. Appl., vol. 202 no. 3, 1996, pp. 727-746. https://doi.org/10.1006/jmaa.1996.0344

[43] G. Silva, R. B. Vinter, *Necessary conditions for optimal impulsive control problems*, SIAM J. Control Optim., vol. 35, no. 6, 1997, pp. 1829-1846. https://doi.org/10.1137/S0363012995281857

[44] H. J. Sussmann, *Geometry and Optimal Control*, Mathematical Control Theory, J. Baillieul and J. C. Willems, Eds., Springer-Verlag, New York, 1998, pp. 140-198. https://doi.org/10.1007/978-1-4612-1416-8_5

[45] H. J. Sussmann, *Generalized differentials, variational generators, and the maximum principle with state constraints*, Mathematical Control Theory, J. Baillieul and J. C. Willems, Eds., Springer-Verlag, New York, 1998, pp. 140-198. https://doi.org/10.1007/978-3-540-77653-6_4

[46] R. B. Vinter, *Optimal Control*, Birkhäuser, Boston, 2000. DOI: 10.1007/978-0-8176-8086-2

[47] J. Warga, *Normal Control Problems have no Minimizing Strictly Original Solutions*, Bulletin of the Amer. Math. Soc., 77, 4, 1971, pp. 625-628.

[48] J. Warga, *Optimal Control of Differential and Functional Equations*, Academic Press, New York, 1972. https://doi.org/10.1016/C2013-0-11669-8

[49] J. Warga, *Controllability, extremality, and abnormality in nonsmooth optimal control*, J. Optim. Theory and Applic., vol. 41, no. 1, 1983, pp. 239-260. https://doi.org/10.1007/BF00934445

[50] J. Warga, *Optimization and Controllability Without Differentiability Assumptions*, SIAM J. Control and Optim., vol. 21, no. 6, 1983, pp. 837-855. https://doi.org/10.1137/0321051

[51] P. Wolenski, S. Zabić, *A differential solution concept for impulsive systems,* Dyn. Contin. Discrete Impuls. Syst. Ser. A Math. Anal., 13B, 2006, pp. 199-210.