A HIGHER-ORDER MAXIMUM PRINCIPLE
FOR IMPULSIVE OPTIMAL CONTROL PROBLEMS

M. SOLEDAD ARONNA, MONICA MOTTA, AND FRANCO RAMPAZZO

Abstract. We consider a nonlinear control system, affine with respect to an unbounded control $u$ taking values in a closed cone of $\mathbb{R}^m$, and with drift depending on a second, ordinary control $a$, ranging on a bounded set. We provide first and higher order necessary optimality conditions for a Bolza problem associated to this system. The lack of coercivity assumptions gives an impulsive character to the problem: minimizing sequences of trajectories happen to converge toward impulsive, namely discontinuous, trajectories. As is known, a distributional approach does not make sense in such a nonlinear setting, where instead, a suitable embedding in the graph space is needed. We will illustrate how the chance of using impulse perturbations, makes it possible to derive a Maximum Principle which includes both the usual maximum condition and higher order conditions involving the Lie brackets of the vector fields involved in the dynamics.

1. Introduction

In this paper we investigate necessary optimality conditions for the following optimal control problem:

(1) minimize $\Psi(T, x(T)) + \int_0^T \ell(x(t), u(t), a(t))dt$

over the set of processes $(T, u, a, x)$ verifying the control system

(2) \[
\begin{aligned}
\frac{dx}{dt}(t) &= f(x(t), a(t)) + \sum_{i=1}^{m} g_i(x(t))u^i(t), \quad t \in [0, T], \\
x(0) &= \hat{x}
\end{aligned}
\]

with the integral constraint

(3) $\int_0^T |u(t)|dt \leq K, \quad (T, x(T)) \in \mathcal{T},$

for some $K \leq +\infty$. For simplicity, the state variable $x$ ranges over $\mathbb{R}^n$, but treated issues and the presented results can be easily extended to the case of a differential manifold. The set $\mathcal{T} \subseteq \mathbb{R}^{1+n}$ is regarded as a target. The main feature of the control system (2) consists in the fact that the control $u$ is unbounded: more precisely, $u$ takes values in a closed, convex, cone $C \subseteq \mathbb{R}^m$, while $a$ is a standard control ranging in a compact set $A \subset \mathbb{R}^q$. The Lagrangian $\ell$ is assumed to have the form

\[
\ell(x, u, a) = \ell_0(x, a) + \ell_1(x, u),
\]

where $\ell_0, \ell_1$ are non-negative functions and $\ell_1$ is positively homogeneous in $u$ of degree $\gamma \in (0, 1]$. Since the dynamics is affine in $u$, this sublinearity of $\ell_1$ with respect to $u$ is crucial for it causes the minimum problem (1)-(3) to have an impulsive character. Indeed, in an unbounded control problem it is exactly the superlinearity of the Lagrangian in the variable $u$ that makes, so to speak, very large speeds not convenient. On the contrary, the lack of such growth condition makes jumps of optimal trajectories a likely occurrence. A trivial example is represented by the time optimal problem with the target intersecting the orbit of the vector field $g_1$ issuing from the initial point $\hat{x}$: the optimal time is clearly equal
to zero and an optimal trajectory is represented by running across the mentioned orbit with infinite speed. Let us recall that, because of the nonlinearity of the dynamics, a measure-theoretical approach, with $u$ to be interpreted as a Radon measure, is out of question (see e.g. [13]). Instead, following a nowadays standard approach [31, 24, 5, 17, 16, 26], we embed the original system into the space-time system (6) below, namely the control system for the graphs, possibly reparametrized in order to include instantaneous evolutions. Preliminarily, let us rewrite the integral constraint in (3) as the final constraint

$$v(T) \leq K$$

on an additional state variable $v$ verifying

$$\begin{align*}
\frac{dv}{dt}(t) &= |u(t)|, \\
v(0) &= 0.
\end{align*}$$

Considering a new time parameter $s$, related to $t$ through the relation $t = y^0(s)$, for a nondecreasing function $y^0$, we obtain the system

$$\begin{align*}
\frac{dy^0}{ds}(s) &= w^0(s), \\
\frac{dy}{ds}(s) &= f(y(s), \alpha(s))w^0(s) + \sum_{i=1}^{m} g_i(y(s))w^i(s), \\
\frac{d\beta}{ds}(s) &= |w(s)|, \\
(y^0, y, \beta)(0) &= (0, \bar{x}, 0), \\
(y^0(S), y(S), \beta(S)) &= \in \Sigma \times [0, K],
\end{align*}$$

which has bounded controls $(w^0, w, \alpha)$ with values in a subset of $[0, +\infty) \times C \times A$, and where the reparameterized version of the variable $v$ is denoted by $\beta$. Observe that when the function $t = y^0(s)$ is strictly increasing, that is $w^0(s) > 0$ for almost every $s$, the space-time system is a mere rescaling of time from $t$ to $s$ of the original control system (2)-(5). Instead, by allowing subsets $I \subseteq [0, S]$ of positive measure with $w^0(s) = 0$ a.e. $s \in I$, we are actually extending our control system into a new larger system where space-time trajectories may contain instantaneous arcs: when $w^0 = 0$ the time $t$ is stopped, while the state evolves (in the pseudo-time $s$) obeying to the non-drift dynamics $\frac{dy}{ds} = \sum_{i=1}^{m} g_i(y)w^i$. It is then natural that necessary conditions concern, more or less explicitly (see e.g. [27, 21, 16]), an optimal space-time process $(S, w^0, w, \alpha, y^0, y, \beta)$ for the system (6) and the extended cost

$$\Psi(y^0(S), y(S)) + \int_{0}^{S} \ell^{e}(y(s), w^0(s), w(s), \alpha(s)) ds,$$

where $\ell^{e} : \mathbb{R} \times [0, +\infty) \times C \times A \to [0, +\infty)$ is given by

$$\ell^{e}(x, w^0, w, a) := \lim_{r \to w^0} \ell \left( x, \frac{w}{r}, a \right) r = \ell_0(x, a)w^0 + \ell_1(x, w)(w^0)^{1-\gamma}.$$

To begin with, we exploit the rate-independence of system (6) to establish a First Order Maximum Principle, from which the identical vanishing of some Hamiltonians $H_{g_i}(x, p) := p \cdot g_i$ follows.

The main novelty of the paper consists in a Higher Order Maximum Principle, namely Theorem 4.1. The result contains new, higher-order constants of motion, the latter involving iterated Lie brackets of the vector fields \{f, g_1, \ldots, g_m\}.

Optimal control problems of the form (11)-(13) are motivated by several important applications [6, 12, 7, 15]. For instance, in Classical Mechanics, when moving D’Alambert
constraints are regarded as controls, the $u^i$’s turn out to be the velocities of these moving constraints [21].

As far as we know, higher-order necessary optimality conditions – up to second order – are present in the literature on optimal impulsive control only for the commutative case, namely when all brackets $[g_i, g_j]$ vanish identically (see e.g. [10] [2, 3]). This is the so-called Frobenius property, under which the vector fields $g_1, \ldots, g_m$ are constant in a suitable coordinate chart. Instead, our necessary conditions are established for the general, non-commutative case. Furthermore, our results are obtained under very low regularity assumptions on the involved vector fields. Moreover, for some nonnegative integer $m_1$, the control components $u_{m_1+1}, \ldots, u_m$ are allowed to range only on a convex cone which is not a subspace. Even more importantly, unlike [8], we are not assuming any constancy hypothesis on the dimension of the Lie algebra generated by $g_1, \ldots, g_m$.

The article is organized as follows. In the next subsection 1.1 we introduce some notation and a set of definitions and technical results involving Lie brackets. The optimal control problem to be dealt with is given in Section 2 together with its space-time extension. In Section 3 we state a First Order Maximum Principle (see Theorem 3.1). We establish a Higher-Order Maximum Principle, that is our main result, in Section 4 (see Theorem 4.1). Section 5 is devoted to the proof of the main result. Some generalizations are discussed along Section 6. In particular, while Theorems 3.1, 4.1 are stated and proved on $\mathbb{R}^n$, in Subsection 6.1 we give the details for a version of the results on a differential manifold. Moreover, in Subsection 6.2 we give some hints for the extension to the case of non-smooth vector fields and in Subsection 6.3 discuss the connection of our main theorem with a higher order maximum principle, recently proved in [8], for non-impulsive but still unbounded control systems.

1.1. Notations and preliminaries.

1.1.1. Notation. Let $N$ be a natural number. For every $i \in \{1, \ldots, N\}$, we write $e_i$ for the $i$th element of the canonical basis of $\mathbb{R}^N$. Given $\bar{x} \in \mathbb{R}^N$, $B_N(\bar{x})$ is the closed unit ball $\{x \in \mathbb{R}^N : |x - \bar{x}| \leq 1\}$, just $B_N$ when $\bar{x} = 0$, and $\partial B_N := \{x \in \mathbb{R}^N : |x| = 1\}$ is its boundary. A subset $K \subseteq \mathbb{R}^N$ is called cone if $\alpha x \in K$ whenever $\alpha > 0$, $x \in K$. Given a real interval $I$ and $X \subseteq B_N$, we write $AC(I, X)$ for the space of absolutely continuous functions, $C^0(I, X)$ for the space of continuous functions, $L^1(I, X)$ for the Lebesgue space of $L^1$-functions, and $L^\infty(I, X)$ for the Lebesgue space of measurable essentially bounded functions, respectively, defined on $I$ and assuming values in $X$. As customary, we shall use $\|\cdot\|_{C^0(I, X)}$, $\|\cdot\|_{L^\infty(I, X)}$, and $\|\cdot\|_{L^1(I, X)}$ to denote the $C^0$, the essential supremum norm and the $L^1$-norm, respectively. Sometimes, when no confusion may arise, we will simply write $\|\cdot\|_{\infty}$, $\|\cdot\|_1$. Given an integer $k \geq 1$ and an open subset $\Theta \subseteq \mathbb{R}^N$, we say that a function $F : \Theta \to \mathbb{R}^N$ is of class $C^k$ if it possesses continuous partial derivatives up to order $k$ in $\Theta$. Given a real valued function $F : [a, b] \to \mathbb{R}$, we define, as usual, the essential infimum of $F$ as $\text{ess inf}_{[a, b]} F := \sup \{r \in \mathbb{R} : m\{x \in [a, b] : F(x) < r\} = 0\}$, where $m$ denotes the Lebesgue measure.

1.1.2. Boltyanskii approximating cones. Let us recall the notion of Boltyanskii approximating cone. Let $N$ be a natural number.

Definition 1.1. Let $Z \subseteq \mathbb{R}^N$, $z$ a point in $Z$ and $K$ a convex cone in $\mathbb{R}^N$. We say that $K$ is a Boltyanskii approximating cone to $Z$ at $z$ if there exist a convex cone $C \subseteq \mathbb{R}^M$ (for some integer $M \geq 0$), a neighborhood $V$ of $0$ in $\mathbb{R}^M$, and a continuous map $F : V \cap C \to Z$ such that $F(0) = z$, $F$ is differentiable at 0 in the direction of $C$ with differential $L$ at 0, namely, for some neighborhood of 0, there exists a linear map $L : \mathbb{R}^M \to \mathbb{R}^N$ such that $F(v) = F(0) + Lv + o(|v|)$, $\forall v \in V \cap C$,
and $LC = K$.

**Definition 1.2.** Let us consider two subsets $A_1, A_2$ of a topological space $X$. If $y \in A_1 \cap A_2$, we say that $A_1$ and $A_2$ are locally separated at $y$ provided there exists a neighborhood $C$ of $y$ such that

$$A_1 \cap A_2 \cap C = \{y\}$$

A basic relation between local separation and linear separation of their approximating cones is established by the following result.

**Theorem 1.1.** Let $Z_1$ and $Z_2$ be subsets of $\mathbb{R}^N$, $z \in Z_1 \cap Z_2$ and let $K_1, K_2 \subseteq \mathbb{R}^N$ be approximating cones to $Z_1$ and $Z_2$, respectively, at $z$. If one of the two cones $K_1$ or $K_2$ is not a subspace and $Z_1, Z_2$ are locally separated at $z$, then there exists a covector $\lambda \neq 0$ such that

$$\lambda \cdot v_1 \leq 0, \quad \lambda \cdot v_2 \geq 0, \quad \text{for all } v_1 \in K_1, v_2 \in K_2.$$

1.1.3. **Lie brackets of vector fields and formal brackets.** As is well-known, the Lie bracket of two $C^1$ vector fields $F_1, F_2$ is the vector field $[F_1, F_2]$ defined, in any system of coordinates, by

$$DF_2 \cdot F_1 - DF_1 \cdot F_2$$

where $D$ denotes differentiation. By iterating the bracketing procedure, for vector fields with the appropriate regularity, we obtain iterated vector fields. The Lie bracket is antisymmetric and verifies the Jacobi identity

$$[F_1, F_2] = -[F_2, F_1], \quad [[F_1, F_2], F_3] + [[F_3, F_1], F_2] + [[F_2, F_3], F_1] = 0.$$  

Let us define a **formal bracket** as the object that one obtains by formally replacing vector fields with a set of adjacent abstract variables $(X_1, X_2, \ldots)$ ordered increasingly. For instance,

$$b_1(X_4, X_5, X_6) := [X_4, [X_5, X_6]], \quad b_2(X_2, X_3, X_4, X_5, X_6) := [[X_2, [X_3, X_4]], [X_5, X_6]],$$

are formal brackets. The length of a formal bracket $b$ is the number of variables in $b$. For instance, the lengths of the above $b_1$ and $b_2$ are 3 and 5, respectively.

Clearly, for every formal bracket $b$ there exists a unique pair $(b_1, b_2)$ of formal brackets, called the factorization of $b$, such that $b = [b_1, b_2]$.

We now introduce recursively a notion of regularity for $q$-tuples $F = (F_1, \ldots, F_q)$ of vector fields, in connection with a given formal bracket $b$.

**Definition 1.3** (Classes $C^{b+k}$). Let us fix integers $k, \mu \geq 0$, $p, q \geq 1$, $\mu + p < q$, and let $b = b(X_{\mu+1}, \ldots, X_{\mu+p})$ be a formal bracket. We say that a $q$-tuple $F = (F_1, \ldots, F_q)$ of vector fields is of class $C^{b+k}$ if the following conditions hold true:

i) whenever the length of $b$ is 1, i.e. $p = 1$ and $b = X_{\mu+1}$, the vector field $F_{\mu+1}$ is of class $C^k$;

ii) when $p > 1$ and $(b_1, b_2)$ is the factorization of $b$, $F$ is of class $C^{B_1+k+1}$ and of class $C^{B_2+b+k+1}$.

Roughly speaking, $(F_1, \ldots, F_q)$ is of class $C^{b+k}$ if there exists non-negative integers $k_1, \ldots, k_q$ such that every vector field $F_j$ is of a class $C^{k_j}$, and this information is enough to conclude that the Lie bracket $b(F_1, \ldots, F_q)$ is well defined and is of class $C^k$. For instance, if $b = [[X_3, X_4], [X_5, X_6], X_7]$ and $F = (F_1, \ldots, F_8)$ is a 8-tuple of vector fields, (so $m = 5$, $q = 8$, $\mu = 2$), then $F$ is of class $C^{b+5}$ if and only if $F_3, F_4, F_7 \in C^5$ and $F_5, F_6 \in C^6$.  

2. The optimization problems

In this section we introduce rigorously the optimization problem over \( L^1 \)-controls and its embedding in an impulsive problem.

Throughout the paper we shall assume the following hypotheses:

\[ \text{(Hp)} \]

1. The target \( \mathcal{K} \) is a closed subset of \([0, +\infty) \times \mathbb{R}^n\); the control set \( A \subset \mathbb{R}^q \) is compact; the unbounded control set \( \mathcal{C} \subseteq \mathbb{R}^m \) is a convex cone of the form

\[
\mathcal{C} = \mathbb{R}^{m_1} \times C_2
\]

where \( m_1 + m_2 = m \) and \( C_2 \subset \mathbb{R}^{m_2} \) is a convex cone which does not contain straight lines;

2. the drift dynamics \( f : \mathbb{R}^n \times A \to \mathbb{R}^n \) is continuous and continuously differentiable in \( x \), for every fixed \( a \in A \), and has continuous partial derivatives \( \frac{\partial f}{\partial x}, \ldots, \frac{\partial f}{\partial x^n} \);

3. the vector fields \( g_1, \ldots, g_m : \mathbb{R}^n \to \mathbb{R}^n \) are continuously differentiable in \( \mathbb{R}^n \);

4. the running cost \( \ell : \mathbb{R}^n \times \mathcal{C} \times A \to [0, +\infty) \) is continuous and continuously differentiable in \( x \) for every fixed \( (u, a) \in \mathcal{C} \times A \) and has continuous partial derivatives \( \frac{\partial \ell}{\partial x}, \ldots, \frac{\partial \ell}{\partial x^n} \). Moreover,

\[
\ell(x, u, a) = \ell_0(x, a) + \ell_1(x, u)
\]

for all \( (x, u, a) \in \mathcal{C} \times A \) and there exists \( \gamma \in (0, 1) \) such that

\[
\ell_1(x, pu) = \rho^n \ell_1(x, u), \quad \text{for all } (x, u, \rho) \in \mathbb{R}^n \times \mathcal{C} \times (0, +\infty).
\]

5. the final cost \( \Psi : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R} \) is continuously differentiable around the final point of a given reference trajectory.

Observe that the hypothesis (i) on \( \mathcal{C} \) is by no means restrictive, since it can be recovered by replacing the single vector fields \( g_i \) with suitable linear combinations of \( \{g_1, \ldots, g_m\} \).

2.1. The original optimal control problem. We define the set of strict-sense controls as

\[
\mathcal{U} := \bigcup_{T > 0} \{ T \} \times \{(u, a) \in L^1([0, T], \mathcal{C}) \times L^1([0, T], A) \}.
\]

**Definition 2.1.** For any \((T, u, a, x, v) \in \mathcal{U}\), we say that \((T, u, a, x, v)\) is a strict-sense process if \((x, v)\) is the unique Carathéodory solution to

\[
\begin{align*}
\frac{dx}{dt}(t) &= f(x(t), a(t)) + \sum_{i=1}^{m} g_i(x(t)) u^i(t) \\
\frac{dv}{dt}(t) &= |u(t)| \\
(x, v)(0) &= (\hat{x}, 0)
\end{align*}
\]

on \([0, T] \). Furthermore, we say that a strict-sense process \((T, u, a, x, v)\) is feasible if \((T, x(T), v(T)) \in \mathcal{K} \times [0, K] \).

Let

\[ \text{(P)} \]

\[
\text{minimize} \left\{ \Psi(T, x(T)) + \int_0^T \ell(x, u, a) dt : (T, u, a, x, v) \text{ is a feasible strict-sense process} \right\}
\]

denote the optimization problem over the set of strict-sense processes. In order to define local minimizers of \((P)\), we introduce a metric over the set of trajectories. To this aim,

\[ \text{Under our assumptions on the control system, for any strict-sense control } (T, u, v), \text{ there exists a unique solution of } (8) \text{ which is defined in general on a maximal interval of definition } [0, \tau) \subseteq [0, T]. \]
for any strict-sense process \((T, u, a, x, v)\), we consider the extension of \((x, v)\) to \([0, +\infty)\) obtained by setting \((x(t), v(t)) := (x(T), v(T))\) for \(t > T\). For any pair of strict-sense processes \((T, u, a, x, v), (\bar{T}, \bar{u}, \bar{a}, \bar{x}, \bar{v})\), we define the distance:

\[
\begin{align*}
\text{d}\left((T, x, v), (\bar{T}, \bar{x}, \bar{v})\right) & := |T - \bar{T}| + \| (x, v) - (\bar{x}, \bar{v}) \|_{L^\infty([0, +\infty), \mathbb{R}^{n+1})}.
\end{align*}
\]

**Definition 2.2 (Local and global strict-sense minimizers).** We say that a feasible strict-sense process \((\bar{T}, \bar{u}, \bar{a}, \bar{x}, \bar{v})\) is a local strict-sense minimizer of \((P)\) if there exists \(\delta > 0\) such that

\[
\Psi(\bar{T}, \bar{x}(\bar{T})) + \int_0^T \ell(\bar{x}(t), \bar{u}(t), \bar{a}(t))dt \leq \Psi(T, x(T)) + \int_0^T \ell(x(t), u(t), a(t))dt
\]

for every feasible strict-sense process \((T, u, a, x, v)\) verifying

\[
\text{d}\left((T, x, v), (\bar{T}, \bar{x}, \bar{v})\right) < \delta.
\]

If relation \((P)\) is satisfied for all admissible strict-sense processes, we say that \((\bar{T}, \bar{u}, \bar{a}, \bar{x}, \bar{v})\) is a global strict-sense minimizer.

**Remark 2.1.** By adding the trivial equations \(\frac{dx^0}{dt}(t) = 1\), \(\frac{d\hat{x}}{dt}(t) = u(t)\), where \(\hat{x} = (x^{n+1}, \ldots, x^{n+m})\), we can allow \(\ell, f, g_j, j = 1, \ldots, m\) in \((P)\) to depend on \(t\) and on the integral of \(u\), and \(\Psi\) in \((P)\) to depend on the integral of \(u\) as well.

### 2.2. The space-time optimal control problem.

Define the set of space-time controls

\[
\mathcal{W} := \bigcup_{S > 0} \{ S \} \times \left\{ (w^0, w, \alpha) \in L^\infty([0, S], [0, +\infty) \times \mathcal{C} \times A) : \text{ess inf}_{[0, S]} (w^0 + |w|) > 0 \right\}.
\]

Since every \(L^1\)-equivalence class contains Borel measurable representatives, here and in the sequel we tacitly assume that all \(L^1\)-maps are Borel measurable on compact intervals, when necessary.

**Definition 2.3.** For any space-time control \((W, w^0, w, \alpha) \in \mathcal{W}\), we say that \((S, w^0, w, \alpha, y^0, \beta)\) is a space-time process if \((y^0, y, \beta)\) is the unique Carathéodory solution of

\[
\begin{align*}
\frac{dy^0}{ds}(s) & = w^0(s), \\
\frac{dy}{ds}(s) & = f(y(s), \alpha(s))w^0(s) + \sum_{i=1}^m g_i(y(s))w_i(s), \\
\frac{d\beta}{ds}(s) & = |w(s)|, \\
(y^0, y, \beta)(0) & = (0, \bar{x}, 0),
\end{align*}
\]

on the interval \([0, S]\). Furthermore, we say that a space-time process \((S, w^0, w, \alpha, y^0, y, \beta)\) is feasible if \((y^0(S), y(S), \beta(S))\) belongs to \(\Sigma \times [0, K]\).

Let us write

\[
(P^{st}) \quad \text{minimize} \left\{ \Psi(y^0(S), y(S)) + \int_0^S \ell^\varepsilon(y, w^0, w, \alpha)ds : (S, w^0, w, \alpha, y^0, y, \beta) \text{ is a feasible space-time process} \right\},
\]

where we remind that

\[
\ell^\varepsilon(x, w^0, w, a) = \ell_0(x, a)w^0 + \ell_1(x, w)(w^0)^{1-\gamma} \quad \forall (x, w^0, w, a) \in \mathbb{R}^n \times [0, +\infty) \times \mathcal{C} \times A,
\]
for some $\gamma \in (0, 1]$. Given a space-time process $(S, w^0, w, \alpha, y^0, y, \beta)$, we extend $(y^0, y, \beta)$ to $[0, +\infty)$ by setting $(y^0, y, \beta)(s) := (y^0, y, \beta)(S)$ for every $s > S$. For any pair of space-time processes $(S, w^0, w, \alpha, y^0, y, \beta), (\bar{S}, \bar{w}, \bar{\alpha}, \bar{y^0}, \bar{y}, \bar{\beta})$, we define the distance

\begin{equation}
\begin{aligned}
d_e\left((S, y^0, y, \beta), (\bar{S}, \bar{y^0}, \bar{y}, \bar{\beta})\right) := |S - \bar{S}| + \|(y^0, y, \beta) - (\bar{y}^0, \bar{y}, \bar{\beta})\|_{L^\infty([0, +\infty), \mathbb{R}^{n+2})}.
\end{aligned}
\end{equation}

**Definition 2.4.** A feasible space-time process $(S, \bar{w}, \bar{\alpha}, \bar{y}^0, \bar{y}, \bar{\beta})$ is said to be a local minimizer for the space-time problem $\mathcal{P}_S^\alpha$ if there exists $\delta > 0$ such that

\begin{equation}
\begin{aligned}
\Psi((\bar{y}^0, \bar{y})(\bar{S})) + \int_0^{\bar{S}} \ell^e(\bar{y}, \bar{w}, \bar{\alpha})ds \leq \Psi((y^0, y)(S)) + \int_0^S \ell^e(y, w^0, w, \alpha)ds
\end{aligned}
\end{equation}

for all feasible space-time processes $(S, w^0, w, \alpha, y^0, y, \beta)$ satisfying

\begin{equation}
\begin{aligned}
d_e\left((S, y^0, y, \beta), (\bar{S}, \bar{y}^0, \bar{y}, \bar{\beta})\right) < \delta.
\end{aligned}
\end{equation}

If (14) is satisfied for all feasible space-time processes, we say that $(S, \bar{w}, \bar{\alpha}, \bar{y}^0, \bar{y}, \bar{\beta})$ is a global space-time minimizer.

The space-time system (12) enjoys a key property: it is rate-independent. Precisely, given a strictly increasing, surjective and bi-Lipschitz continuous function $\sigma : [0, S] \to [0, \bar{S}], (S, \bar{w}, \bar{\alpha}, \bar{y}^0, \bar{y}, \bar{\beta})$ is a space-time process of (12) if and only if $(S, w^0, w, \alpha, y^0, y, \beta)$ verifying

\begin{equation}
\begin{aligned}
(w^0, w) = (\bar{w}, \bar{w}) \circ \frac{d\sigma}{ds}, \quad (\alpha, y^0, y, \beta) = (\bar{\alpha}, \bar{y}^0, \bar{y}, \bar{\beta}) \circ \sigma
\end{aligned}
\end{equation}

is a space-time process of (12) (see Sect. 3). In this case, $(S, \bar{w}, \bar{\alpha}, \bar{y}^0, \bar{y}, \bar{\beta}$) is feasible if and only if $(S, w^0, w, \alpha, y^0, y, \beta)$ is feasible, and in this case it holds

\begin{equation}
\begin{aligned}
\Psi((\bar{y}^0, \bar{y})(\bar{S})) + \int_0^{\bar{S}} \ell^e(\bar{y}, \bar{w}, \bar{\alpha})ds = \Psi((y^0, y)(S)) + \int_0^S \ell^e(y, w^0, w, \alpha)ds.
\end{aligned}
\end{equation}

This allows us to introduce an equivalence relation: we say that two space-time processes $(S, \bar{w}, \bar{\alpha}, \bar{y}^0, \bar{y}, \bar{\beta}), (S, w^0, w, \alpha, y^0, y, \beta)$ are equivalent if there exists a function $\sigma$ with the above properties. It is not difficult to prove the following result.

**Lemma 2.1.** A feasible space-time process $(S, \bar{w}, \bar{\alpha}, \bar{y}^0, \bar{y}, \bar{\beta})$ is a local minimizer (resp., a global minimizer) for the space-time problem $\mathcal{P}_S^\alpha$ if and only if every equivalent space-time process is a local minimizer (resp., a global minimizer), and in this case the costs coincide.

As a consequence, one can replace without loss of generality a minimizer with its canonical parameterization:

**Definition 2.5.** For any space-time process $(S, w^0, w, \alpha, y^0, y, \beta)$, let us set

\begin{equation}
\begin{aligned}
\sigma(s) := \int_0^s (w^0(r) + |w(r)|) dr \quad \text{for } s \in [0, S], \quad S_c := \sigma(S) = S + \beta(S).
\end{aligned}
\end{equation}

Observing that $\sigma : [0, S] \to [0, S_c]$ is strictly increasing, surjective and bi-Lipschitz continuous, we call canonical parameterization of $(S, w^0, w, \alpha, y^0, y, \beta)$ the equivalent process $(S_c, w^0_c, w_c, \alpha_c, y^0_c, y_c, \beta_c)$, verifying

\begin{equation}
\begin{aligned}
(w^0_c, w_c) := (w^0, w) \circ \sigma^{-1} \frac{d\sigma^{-1}}{ds}, \quad (\alpha_c, y^0_c, y_c, \beta_c) := (\alpha, y^0, y, \beta) \circ \sigma^{-1}.
\end{aligned}
\end{equation}

Note that

\begin{equation}
\begin{aligned}
w^0_c(s) + |w_c(s)| = 1 \quad \text{a.e. } s \in [0, S_c].
\end{aligned}
\end{equation}

Therefore, we set

\begin{equation}
\begin{aligned}
\mathcal{W}_c := \{(S, w^0, w, \alpha) \in \mathcal{W} : w^0(s) + |w(s)| = 1 \quad \text{a.e. } s \in [0, S]\}.
\end{aligned}
\end{equation}
and call \textit{canonical} the space-time controls in $\mathcal{W}_c$ and the corresponding space-time processes of \cite{12}. Of course, a canonical space-time process coincides with its canonical parametrization.

2.3. \textbf{The space-time embedding}. The original control system \cite{8} can be embedded into the space-time system \cite{12}. Precisely, as an easy consequence of the chain-rule, given a strict-sense process $(T, u, a, x, v)$, by setting

$$\sigma(t) := \int_0^t (1 + |u(\tau)|) \, d\tau, \quad y^0 := \sigma^{-1} : [0, S] \rightarrow [0, T],$$

one obtains that

$$\begin{align*}
(S, w^0, w, \alpha, y^0, y, \beta) := & \left( S, \frac{dy^0}{ds}, (u \circ y^0) \cdot \frac{dy^0}{ds}, a \circ y^0, y^0, x \circ y^0, \nu \circ y^0 \right) \\
\mathcal{W}_+ := & \left\{(S, w^0, w, \alpha) \in \mathcal{W} : w^0(s) > 0 \text{ a.e. } s \in [0, S]\right\}.
\end{align*}$$

(19) is a (canonic) space-time process with $(S, w^0, w, \alpha) \in \mathcal{W}_+$, where

$$\begin{align*}
\mathcal{W}_+ := & \left\{(S, w^0, w, \alpha) \in \mathcal{W} : w^0(s) > 0 \text{ a.e. } s \in [0, S]\right\}.
\end{align*}$$

Conversely, if $(S, w^0, w, \alpha, y^0, y, \beta)$ is a space-time process with $(S, w^0, w, \alpha) \in \mathcal{W}_+$, then the function $y^0 : [0, S] \rightarrow [0, T]$, has an absolutely continuous inverse

$$\sigma : [0, T] \rightarrow [0, S],$$

and

$$\begin{align*}
(T, u, a, x, v) := & \left( T, (w \circ \sigma) \cdot \frac{d\sigma}{dt}, \alpha \circ \sigma, y \circ \sigma, \beta \circ \sigma \right)
\end{align*}$$

is a strict-sense process. Hence, the family of strict-sense processes can be identified with the subfamily of space-time processes $(S, w^0, w, \alpha, y^0, y, \beta)$ for which $(S, w^0, w, \alpha)$ belongs to $\mathcal{W}_+$.

The impulsive, space-time extension of the original optimal control problem consists just in allowing the control $w^0$ to vanish (on a set of positive measure): observe that the $s$-intervals where $w^0$ vanishes represent the ‘impulses’, namely the arcs of instantaneous evolution of both the control and the state (see e.g. \cite{5, 17, 11, 19}).

The notion of local minimizer for \cite{Pst} is consistent with the definition of local minimizer for the original problem \cite{P}:

\textbf{Lemma 2.2.} \textit{A feasible strict-sense process $(T, u, a, x, v)$ is a strict-sense local minimizer for problem \cite{P} if and only if $(S, w^0, w, \alpha, y^0, y, \beta)$ defined as in \cite{19} is a space-time local minimizer for \cite{Pst} among feasible space-time processes with controls in $\mathcal{W}_+$. Moreover,}

$$\Psi((y^0, \bar{y})(\bar{S})) + \int_0^S \ell^c(\bar{y}, \bar{w}, \bar{\alpha}) \, ds = \Psi(T, \bar{x}(T)) + \int_0^T \ell(x, u, a) \, dt.$$

We omit the proof of this result, which is an easy consequence of Lemma \cite{21} above.

3. \textbf{A First Order Maximum Principle}

In this section we prove a First Order Maximum Principle and deduce some constants of motion. Due to the rate-independence of the space-time control system discussed in Subsection 2.2, we can always assume that a local minimizer $(\bar{S}, \bar{w}, \bar{\alpha}, \bar{y}, \bar{y}, \bar{\beta})$ for the space-time problem \cite{Pst} is canonical.

Let us set

$$\begin{align*}
\mathcal{W}(C) := & \{(w^0, w) \in [0, \infty) \times C : w^0 + |w| = 1\}.
\end{align*}$$

We will consider the \textit{unmaximized Hamiltonian}

$$H(x, p_0, p, \pi, \lambda, w^0, w, a) := p_0 w^0 + p \cdot \left(f(x, a)w^0 + \sum_{i=1}^m g_i(x)w^i\right) + \pi|w| + \lambda \varepsilon(x, w^0, w, a),$$

\begin{align*}
\Psi((y^0, \bar{y})(\bar{S})) + \int_0^S \ell^c(\bar{y}, \bar{w}, \bar{\alpha}) \, ds = & \Psi(T, \bar{x}(T)) + \int_0^T \ell(x, u, a) \, dt.
\end{align*}
and the Hamiltonian $H : \mathbb{R}^n \times \mathbb{R} \times \mathbb{R}^n \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ defined by

$$H(x, p_0, p, \pi, \lambda) := \max_{(w^0, w, a) \in \mathcal{W}(C) \times A} H(x, p_0, p, \pi, \lambda, w^0, w, a)$$

**Theorem 3.1 (First Order Maximum Principle).** Let $(\bar{S}, \bar{w}, \bar{v}, \bar{y}, \bar{\beta})$ be a canonical local minimizer for the space-time problem $(P_{\text{st}})$. Then, for every Boltyanskiy approximating cone $\Gamma$ of the target $\Sigma$ at $(\bar{y}, \bar{y})(\bar{S})$, there exists a multiplier $(p_0, p, \pi, \lambda) \in \mathbb{R} \times AC\left([0, \bar{S}], \mathbb{R}^n\right) \times (-\infty, 0] \times [0, +\infty)$ such that the following conditions hold true:

1. **(Non-triviality)**
   
   $$(p_0, p, \lambda) \neq (0, 0, 0).$$
   
   Particularly, if the trajectory $\bar{y}$ is not instantaneous, namely, if $\bar{y}(\bar{S}) > 0$, then (22) can be strengthened to
   
   $$(p, \lambda) \neq (0, 0).$$

2. **(Non-transversality)**
   
   $$(p_0, p(\bar{S}), \pi) \in \left[-\lambda \left(\frac{\partial \Psi}{\partial t}(\bar{y}(\bar{S})), \frac{\partial \Psi}{\partial x}(\bar{y}(\bar{S}))\right) - \Gamma^\perp\right] \times J,$$
   
   with $J = \{0\}$ if $\bar{\beta}(T) < K$, and $J = (0, +\infty)$ if $\bar{\beta}(T) = K$. In particular,
   
   $\pi = 0$ provided $\bar{\beta}(S) < K$.

3. **(Adjoint equation)** The path $p$ solves the adjoint equation
   
   $$\frac{dp}{ds}(s) = -\frac{\partial H}{\partial x}(\bar{y}(s), p(s), \pi, \lambda, w^0(s), \bar{w}(s), \bar{\alpha}(s)) \quad a.e. \ s \in [0, \bar{S}].$$

4. **(First order maximization)** For almost every $s \in [0, \bar{S}]$, one has
   
   $$H(\bar{y}(s), p_0, p(s), \pi, \lambda, w^0(s), \bar{w}(s), \bar{\alpha}(s)) = H(\bar{y}(s), p_0, p(s), \pi, \lambda).$$

5. **(Vanishing of the Hamiltonian)** For every $s \in [0, \bar{S}]$, one has
   
   $$H(\bar{y}(s), p_0, p(s), \pi, \lambda) = 0.$$

**Proof.** In a slightly different version (utilizing the limiting normal cone instead of the polar of a Boltyanskiy approximating cone), this result has been already proved in [13] as an easy consequence of the standard Maximum Principle. For the sake of self-consistency, let us give the precise proof of the version stated here, which, again, is an almost straightforward corollary of the classical Pontryagin Maximum Principle. Indeed, if we apply the classical Pontryagin Maximum Principle – in the version one can find e.g. in [20] and [25] – we get the existence of a multiplier $(p_0, p, \pi, \lambda) \in \mathbb{R} \times AC\left([0, \bar{S}], \mathbb{R}^n\right) \times (-\infty, 0] \times [0, +\infty)$ verifying the non-transversality condition (24), the adjoint equation (26), the maximum relation (27), the conservation (28), and the non-triviality condition (29)

$$(p_0, p, \pi, \lambda) \neq 0.$$ 

To get the strengthened non-triviality condition (22), just observe that if it were false, then from (29) one would get $\pi < 0$, which in turn would imply (by (24)) $\bar{\beta}(\bar{S}) = K(> 0)$. Yet, this would contrast with the fact that, by the maximum relation (27) and the conservation (28), $\bar{w}(s) = 0$ for almost all $s \in [0, \bar{S}]$: indeed, this implies $0 = \int_0^{\bar{S}} |\bar{w}| ds = \bar{\beta}(\bar{S})$. To prove (23), assume by contradiction that $\bar{y}(\bar{S}) > 0$ (i.e., that the trajectory does not reduce to a unique instantaneous jump) and $(p, \pi) = 0$. Then, by (22) one obtains $p_0 \neq 0$, and

---

2It is tacitly meant that, as an approximating cone to the $(T, x, v)$-target $\Sigma \times [0, K]$ at $(\bar{y}, \bar{y}, \bar{\beta})(\bar{S})$, one chooses $\Gamma \times \mathbb{R}$ if $\bar{\beta}(\bar{S}) < K$ and $\Gamma \times (-\infty, 0]$ if $\bar{\beta}(\bar{S}) = K$. In particular, $(\Gamma \times \mathbb{R})^\perp = \Gamma^\perp \times \{0\}$ if $\bar{\beta}(\bar{S}) < K$ and $(\Gamma \times (-\infty, 0))^\perp = \Gamma^\perp \times [0, +\infty)$ when $\bar{\beta}(\bar{S}) = K$. 

**Nonlinear Impulsive Systems** 9
the maximum relation implies \( p_0 < 0 \) and \( \bar{w}^0(s) = 0 \) for almost every \( s \in [0, \bar{S}] \). This contradicts the assumption \( 0 < \gamma^0(\bar{S}) = \int_{0}^{\bar{S}} \bar{w}^0(s)ds = 0 \).

**Definition 3.1.** A process \((\bar{S}, \bar{w}^0, \bar{w}, \bar{\alpha}, \bar{y}^0, \bar{y}, \bar{\beta})\) is called an extremal if it verifies the conditions in Theorem \( \text{[P3]} \) for some multiplier \((p_0, p, \pi, \lambda)\). If there is a choice of the multiplier with \( \lambda = 0 \), then the extremal \((\bar{S}, \gamma^0, \bar{y}, \bar{\beta}, \bar{w}^0, \bar{\alpha})\) is called abnormal, otherwise it is normal. Finally, the extremal is said to be strictly abnormal if every choice of the multiplier \((p_0, p, \pi, \lambda)\) verifies \( \lambda = 0 \).

When \( \ell_1(x, \cdot) \) is positive 1-homogeneous, so that

\[
\ell^c(x, w^0, a) = \ell_0(x, a)w^0 + \ell_1(x, w), \quad \forall w \neq 0,
\]

we can introduce the following drift Hamiltonian:

\[
H^{(dr)}(x, p, p, \lambda) := \max_{a \in A} \left\{ p_0 + p \cdot f(x, a) + \lambda \ell_0(x, a) \right\},
\]

and the impulse Hamiltonian:

\[
H^{(imp)}(x, p, \pi, \lambda) := \max_{w \in C, |w| = 1} \left\{ p \cdot \sum_{i=1}^{m} g_i(x)w^i + \pi + \lambda \ell_1(x, w) \right\}.
\]

**Corollary 3.1.** Assume that \( \ell_1(x, \cdot) \) is positive 1-homogeneous, and let \((\bar{S}, \bar{w}^0, \bar{w}, \bar{\alpha}, \bar{y}^0, \bar{y}, \bar{\beta})\) be a canonical extremal for the space-time problem \([P3]\). Furthermore, let \((p_0, p, \pi, \lambda)\) be a multiplier (whose existence is established Theorem \( \text{[P3]} \)). Then there exists a zero-measure subset \( N \subset [0, \bar{S}] \) such that, for every \( s \in [0, \bar{S}] \setminus N \), one has

\[
H(\bar{y}(s), p_0, p(s), \pi, \lambda, \bar{w}^0(s), \bar{w}(s), \bar{\alpha}(s)) = \max \left\{ H^{(dr)}(\bar{y}(s), p_0, p(s), \lambda), H^{(imp)}(\bar{y}(s), p(s), \pi, \lambda) \right\} = 0,
\]

and

\[
w^0(s) \left[ p_0 + p(s) \cdot f(\bar{y}(s), \bar{\alpha}(s)) + \lambda \ell_0(\bar{y}(s), \bar{\alpha}(s)) \right] = 0,
\]

\[
p(s) \cdot \sum_{i=1}^{m} g_i(\bar{y}(s))\bar{w}^i(s) + \pi |\bar{w}(s)| + \lambda \ell_1(\bar{y}(s), \bar{w}(s)) = 0.
\]

In particular, if for some \( s \in [0, \bar{S}] \setminus N \) one has

1) \( H^{(dr)}(\bar{y}(s), p_0, p(s), \lambda) < 0 \), then \( \bar{w}^0(s) = 0 \) and

\[
p(s) \cdot \sum_{i=1}^{m} g_i(\bar{y}(s))\bar{w}^i(s) + \pi + \lambda \ell_1(\bar{y}(s), \bar{w}(s)) = H^{(imp)}(\bar{y}(s), p_0, p(s), \pi, \lambda) = 0;
\]

2) \( H^{(imp)}(\bar{y}(s), p(s), \pi, \lambda) < 0 \), then \( \bar{w}(s) = 0 \) and

\[
p_0 + p(s) \cdot f(\bar{y}(s), \bar{\alpha}(s)) + \lambda \ell_0(\bar{y}(s), \bar{\alpha}(s)) = H^{(dr)}(\bar{y}(s), p_0, p(s), \lambda) = 0.
\]

**Proof.** By \( \text{[P28]} \) it follows that for every \( s \in [0, \bar{S}] \), one has

\[
p_0w^0 + p(s) \left( f(\bar{y}(s), a)w^0 + \sum_{i=1}^{m} g_i(\bar{y}(s))w^i \right) + \pi |w| + \lambda \left( \ell_0(\bar{y}(s), a)w^0 + \ell_1(\bar{y}(s), w) \right) \leq 0
\]

for all \((w^0, w, a) \in W(C) \times A\). Now, by choosing \( w = 0 \) one gets that \( w^0 = 1 \) and

\[
p_0 + p(s) \cdot f(\bar{y}(s), a) + \lambda \ell_0(\bar{y}(s), a) \leq 0, \quad \text{for all } a \in A,
\]
while taking \( w^0 = 0 \) one obtains
\[
p(s) \cdot \sum_{i=1}^{m} g_i(\bar{y}(s))w^i + \pi + \lambda \ell_1(\bar{y}(s), w) \leq 0, \quad \text{for all } (w, a) \in \mathcal{C} \times A, \quad |w| = 1.
\]

Therefore, \( H^{(\text{dr})}(\bar{y}(s), p(s), \pi, \lambda) \leq 0 \) and \( H^{(\text{imp})}(\bar{y}(s), p(s), \pi, \lambda) \leq 0 \). In fact, it must be that \( \max \{ H^{(\text{dr})}(\bar{y}(s), p_0, p(s), \lambda), H^{(\text{imp})}(\bar{y}(s), p(s), \pi, \lambda) \} = 0 \), since otherwise both the Hamiltonians would be negative, which easily contradicts (28). Let \( N \subset [0, S] \) be the zero-measure subset such that the first order maximization (27) is verified in \([0, S] \setminus N\). Let \( s \in [0, S] \setminus N \). Then by (27), (28) one has that
\[
\begin{align*}
w^0(s) \left[ p_0 + p(s) \cdot f(\bar{y}(s), \bar{\alpha}(s)) + \lambda \ell_0(\bar{y}(s), \bar{\alpha}(s)) \right] \\
+ \left[ p(s) \cdot \sum_{i=1}^{m} g_i(\bar{y}(s))w^i(s) + \pi |w(s)| + \lambda \ell_1(\bar{y}(s), w(s)) \right] = 0.
\end{align*}
\]

Since the above argument implies that both terms in this equality are nonnegative, they necessarily vanish, namely (30) and (31) are verified.

Suppose now that \( H^{(\text{dr})}(\bar{y}(s), p(s), \pi, \lambda) < 0 \). Then (30) implies \( \bar{w}(s) = 0 \), so that \( |\bar{w}(s)| = 1 \) and the thesis (32) follows by (31). If instead \( H^{(\text{imp})}(\bar{y}(s), p(s), \pi, \lambda) < 0 \), then \( \bar{w}(s) = 0 \) due to (31) and in view of the positive 1-homogeneity of the Hamiltonian \( H \) w.r.t. \((w, 0)\). Hence \( \bar{w}(s) = 1 \) and (30) yields (33).

**Remark 3.1.** Notice that under the same hypotheses of Corollary 3.1 on any non-impulsive subinterval \([s_1, s_2] \subset [0, S] \), namely where \( \bar{w}(s) > 0 \) a.e., one has that the drift Hamiltonian \( H^{(\text{dr})}(\bar{y}(s), p_0, p(s), \lambda) \) vanishes. In particular, if \((\bar{S}, \bar{w}, \bar{\alpha}, \bar{y}, \bar{\beta})\) can be identified with a strict-sense process, i.e. if \( \bar{w}(0) > 0 \) for a.e. \( s \in [0, S] \), then
\[
H^{(\text{dr})}(\bar{y}(s), p_0, p(s), \lambda) = 0, \quad \text{for all } s \in [0, \bar{S}].
\]

For any continuous vector field \( F : \mathbb{R}^n \rightarrow \mathbb{R}^n \), let us introduce the classical \( F \)-Hamiltonian
\[
H_F(x, p) := p \cdot F(x), \quad \text{for } (x, p) \in \mathbb{R}^n \times \mathbb{R}^n.
\]

**Corollary 3.2.** Let \((\bar{S}, \bar{w}, \bar{\alpha}, \bar{y}, \bar{\beta})\) be a canonical extremal for the space-time problem (28) and let \((p_0, p, \pi, \lambda)\) be a corresponding multiplier. If
\[
\pi = 0 \quad \text{and} \quad \lambda \ell^c(\bar{y}(s), 0, e_i) = 0, \quad \text{for all } s \in [0, \bar{S}], \quad i = 1, \ldots, m_1,
\]
then one has
\[
\begin{align*}
(H_{g_i}(\bar{y}(s), p(s)) = p(s) \cdot g_i(\bar{y}(s)) = 0, \quad \text{for all } s \in [0, \bar{S}], \quad i = 1, \ldots, m_1.
\end{align*}
\]

**Proof.** By (28) it follows that for every \( s \in [0, \bar{S}] \) one has
\[
p_0 w^0 + p(s) \cdot \left( f(\bar{y}(s), a)w^0 + \sum_{i=1}^{m} g_i(\bar{y}(s))w^i \right) + \lambda \ell^c(\bar{y}(s), w^0, w, a) \leq 0
\]
for all \((w^0, w, a) \in W(\mathcal{C}) \times A\). Therefore, choosing \( w^0 = 0 \) and \( w = \pm e_i \) for any \( i = 1, \ldots, m_1 \), one gets (34).

**Remark 3.2.** Let us recall from Theorem 3.1 that the hypothesis \( \pi = 0 \) is verified as soon as \( \bar{\beta}(\bar{S}) < K \). Moreover, the hypothesis \( \lambda \ell^c(\bar{y}(s), 0, e_i) = 0 \) is obviously satisfied when the extremal \((\bar{S}, \bar{w}, \bar{\alpha}, \bar{y}, \bar{\beta})\) is abnormal and one chooses \( \lambda = 0 \), or when \( \ell_1 \equiv 0 \), or even when \( u_1 \mapsto \ell_1(x, (u_1, 0)) \) is positive homogeneous of degree \( \gamma < 1 \) for all \((x, u) \in \mathbb{R}^n \times \mathbb{R}^{m_1} \).
4. A Higher Order Maximum Principle

This section is devoted to establish a Higher Order Maximum Principle.

Let us begin with a regularity notion for Lie brackets of vector fields in \( \{g_1, \ldots, g_m\} \).

**Definition 4.1.** For every non-negative integer \( k \), we say that a vector field \( B \) is a \( C^k \)-admissible Lie bracket if \( B = b(F_1, \ldots, F_q) \), where \( b \) is a formal bracket and \( (F_1, \ldots, F_q) \) is a \( q \)-tuple of class \( C^{k+1} \) of vector fields belonging to \( \{g_1, \ldots, g_m\} \) (see Definition 1.3). We will use \( \mathfrak{B}^k \) to denote the set of \( C^k \)-admissible Lie brackets.

Roughly speaking, a Lie bracket \( B = b(F_1, \ldots, F_q) \) is a \( C^k \)-admissible Lie bracket if one can deduce its \( C^k \) regularity from the regularity of the vector fields \( F_1, \ldots, F_q \).

**Theorem 4.1 (Higher Order Maximum Principle).** Assume hypothesis (Hp) with \( \gamma < 1 \), and let \( (\bar{S}, \bar{w}, \bar{w}, \bar{\alpha}, \bar{\gamma}, \bar{\beta}) \) be a canonical local minimizer for the space-time problem \( \left[ \mathcal{P}^\mathbb{M} \right] \). Then, for every Boltyanskii approximating cone \( \Gamma \) of the target \( \Sigma \) at \( (\bar{y}, \bar{\gamma})(\bar{S}) \), there exists a multiplier \( (p_0, p, \lambda) \in \mathbb{R} \times AC \left( [0, S], \mathbb{R}^n \right) \times [0, +\infty) \), such that the following conditions hold true:

(i) (Non-triviality)

\[
(p_0, p, \lambda) \neq (0, 0, 0);
\]

(ii) (Non-transversality)

\[
(p_0, p(S)) \in -\lambda \left( \frac{\partial \Psi}{\partial t}(\bar{y}(S), \bar{\gamma}(S)), \frac{\partial \Psi}{\partial x}(\bar{y}(S), \bar{\gamma}(S)) \right) - \Gamma^\perp;
\]

(iii) (Adjoint equation)

\[
\frac{dp}{ds}(s) = -\frac{\partial H}{\partial x}(\bar{y}(s), p(s), 0, \lambda, \bar{w}(s), \bar{a}(s)) \quad a.e. \ s \in [0, \bar{S}];
\]

(iv) (First order maximization)

\[
H(\bar{y}(s), p_0, p(s), 0, \lambda, \bar{w}(s), \bar{a}(s)) = H(\bar{y}(s), p_0, p(s), 0, \lambda), \quad a.e. \ on \ [0, \bar{S}];
\]

(v) (Vanishing of Hamiltonians)

\[
H(\bar{y}(s), p_0, p(s), 0, \lambda) = 0, \quad \forall s \in [0, \bar{S}];
\]

(vi) (Vanishing of higher order Hamiltonians)

\[
H_B(\bar{y}(s), p(s)) = 0, \quad \forall s \in [0, \bar{S}], \forall B \in \mathfrak{B}^0.
\]

Furthermore, if the trajectory \( \bar{y} \) is not instantaneous, namely, if \( \bar{y}(0) > 0 \), then (55) can be strengthened to

\[
(p, \lambda) \neq (0, 0).
\]

The proof of this theorem is postponed to Section 5.

To make the meaning of Lie brackets involving the vector field \( f \), in the sequel we will use the notation \( f_a(\cdot) := f(\cdot, a) \). In particular, if \( h \) is a vector field of class \( C^1 \), \( [h, f_a] \) denotes the Lie bracket of the vector fields \( x \mapsto h(x) \) and \( x \mapsto f_a(x) \).

**Corollary 4.2.** Assume the general hypothesis (Hp) with \( \gamma < 1 \), and let \( (\bar{S}, \bar{w}, \bar{w}, \bar{\alpha}, \bar{\gamma}, \bar{\beta}) \) be a canonical local minimizer for the space-time problem \( \left[ \mathcal{P}^\mathbb{M} \right] \) such that \( \beta(\bar{S}) < K \). Given
Consider the so-called linear control system (Kalman Condition) Remark 4.3

\[ (46) \quad \text{for a.e. } s \]

Also when (44) is verified, by (45) one is left with the only constants of motion

\[ (47) \quad p(s) \cdot \left( f_{\bar{\alpha}(s)}(s) \right) = \sum_{j=m+1}^{m} \left[ g_j(s) \right] \bar{w}^j(s) = \lambda \frac{\partial f^e}{\partial x}(\bar{y}(s), \bar{w}^0(s), \bar{w}(s), \bar{\alpha}(s)) \cdot B(\bar{y}(s)), \]

for a.e. \( s \in [0, S] \). In particular, if

\[ (48) \quad m_1 = m \quad \text{and} \quad \lambda \frac{\partial f^e}{\partial x}(\bar{y}(s), \bar{w}^0(s), \bar{w}(s), \bar{\alpha}(s)) \cdot B(\bar{y}(s)) = 0, \quad \text{for a.e. } s \in [0, S], \]

one obtains

\[ (49) \quad p(s) \cdot \left( f_{\bar{\alpha}(s)}(s) \right) \bar{w}^0(s) = 0, \quad \text{for a.e. } s \in [0, S]. \]

**Remark 4.1.** Let us point out that hypothesis (47) covers at least two important situations, namely the abnormal case, i.e. \( \lambda = 0 \), and the minimum time problem, i.e. \( t \equiv 1 \).

**Proof of Corollary 4.2.** Condition (43) can be obtained by differentiating (41) and remembering that the derivative of \( p \) verifies the adjoint equation (26).

**Remark 4.2.** (The case with commuting vector fields). Under the same assumptions of Corollary 4.2, if the vector fields \( g_1, \ldots, g_{m_1} \) commute, namely \( [g_i, g_j] \equiv 0 \) for all \( i, j = 1, \ldots, m_1 \), the higher order relations in (41) are trivial identities, so one only has that

\[ p(s) \cdot g_i(\bar{y}(s)) = 0, \quad \text{for all } s \in [0, S], i = 1, \ldots, m_1. \]

Also when (44) is verified, by (45) one is left with the only constants of motion

\[ (46) \quad p(s) \cdot [f_{\bar{\alpha}(s)}, g_i](\bar{y}(s)) = 0, \quad \text{for all } i = 1, \ldots, m, \]

for a.e. \( s \) in any interval \([s_1, s_2] \subseteq [0, S]\) where \( \bar{w}^0 > 0 \) a.e..

**Remark 4.3.** (Kalman Condition). As a special case with commuting vector fields, let us consider the so-called linear control system

\[ (47) \quad \frac{dx}{dt}(t) = Cx(t) + Eu(t), \]

where \( C, E \) are an \( n \times n \) and an \( n \times m \) real matrix, respectively. When (44) holds and \( \bar{w}^0(s) > 0 \) for a.e. \( s \) in some interval \([s_1, s_2] \subseteq [0, S]\), setting

\[ f(x) := Cx, \quad g_i := (E_{i1}, \ldots, E_{im})^1, \]

simply by differentiation, from condition (46) one easily obtains the constants of motion:

\[ p(s) \cdot CE = 0, \quad p(s) \cdot C^2E = 0, \ldots, \quad p(s) \cdot C^{n-1}E = 0. \]

for all \( s \in [s_1, s_2] \). In particular, since one also has \( p(s) \cdot E = 0 \), one gets \( p \equiv 0 \) on these intervals provided the linear system (47) verifies the Kalman Controllability Condition (see e.g. [28]).

---

\(^3\text{i.e., } B \text{ is a } C^1 \text{-admissible Lie bracket, see Definition 4.1.}\)
5. Proof of Theorem 4.1

This section is entirely devoted to prove Theorem 4.1. We assume that a canonical local minimizer \((\bar{S}, \bar{w}, \bar{\alpha}, \bar{v}, \bar{y}, \bar{\beta})\) of \(\mathcal{P}_{\text{st}}\) verifying \(\beta(\bar{S}) < K\) is given, and we call it reference process. Let us recall that throughout this section the homogeneity degree \(\gamma\) of \(\ell_1\) is strictly less than 1, as required in the statement of Theorem 4.1.

We begin by transforming our problem with variable terminal time \(S\) into a problem on the fixed interval \([0, \bar{S}]\). This is achieved by a standard procedure, which consists in the following steps:

(i) one introduces a rescaling scalar control \(\zeta : [0, \bar{S}] \to [-\frac{1}{2}, \frac{1}{2}]\) and a reference time variable \(\sigma : [0, \bar{S}] \to [0, +\infty)\) related by

\[
\sigma(s) = \int_0^s (1 + \zeta(\eta))d\eta;
\]

(ii) one considers the rescaled space-time control system on \([0, \bar{S}]\)

\[
\begin{align*}
\frac{dy^0}{ds} &= w^0(1 + \zeta), \\
\frac{dy}{ds} &= \left(f(y, \alpha)w^0 + \sum_{i=1}^{m_1} g_i(y)w^i\right)(1 + \zeta), \\
\frac{dy^\ell}{ds} &= \ell^\ell(y, w^0, w, \alpha)(1 + \zeta), \\
\frac{d\beta}{ds} &= |w|(1 + \zeta), \\
(y^0, y^\ell, \beta)(0) &= (0, \bar{x}, 0, 0); \\
\end{align*}
\]

(iii) one shows that a process \((\tilde{S}, \tilde{w}, \tilde{\alpha}, \tilde{v}, \tilde{y}, \tilde{y}, \tilde{\beta})\) is a local minimizer for the variable end-time problem if and only if the process \((\bar{S}, \bar{w}, \bar{\alpha}, \bar{v}, \bar{y}, \bar{y}, \bar{\beta})\) is a local minimizer for the fixed end-time problem where the local minimum is achieved among the processes \((\tilde{S}, \tilde{w}, \tilde{\alpha}, \tilde{v}, \tilde{y}, \tilde{y}, \tilde{\beta})\) of \(\mathcal{P}\) (that agree with the terminal constraint \((y^0, y, y^\ell, \beta)\) of \(\mathcal{P}\)) (that agree with the terminal constraint \((y^0, y, y^\ell, \beta)\) of \(\mathcal{P}\)).

The proof’s main point consists in defining the variations that produce at the same time Lie brackets of various lengths. Once this is done, the proof proceeds by making use of some arguments that are akin to the ones utilized in one of the standard approaches to the Pontryagin Maximum Principle, namely local set-separation.

Let us observe that we can consider an optimal control problem related to \(\mathcal{P}_{\text{st}}\), in which \(f\) and the \(g_i\) ’s are modified outside a ball in \(\mathbb{R} \times \mathbb{R}^n\) containing the graph of \((\bar{y}^0, \bar{y})\) in its interior, by truncation and mollification. Since the analysis involves consideration of space-time trajectories with graphs arbitrarily close to the graph of \((\bar{y}^0, \bar{y})\) in the Hausdorff sense and the relations appearing in the statement of the theorem concern properties of the data ‘near’ the graph of \((\bar{y}^0, \bar{y})\), it suffices to prove the assertions for only the modified problem. Precisely, we may assume, without loss of generality, that \((\text{Hp})\) is replaced by the stronger hypothesis:

\((\text{Hp})^\ast\) all the assumptions in \((\text{Hp})\) are verified and, moreover, \(\ell, f,\) the \(g_i\)’s, their derivatives \(\frac{\partial f}{\partial y^i}, \frac{\partial g_i}{\partial y^i}, \frac{\partial g_i}{\partial \beta}\) and all the iterated brackets \(\mathcal{B} \in \mathcal{B}^0\) of \(g_1, \ldots, g_{m_1}\) are bounded.

Hypothesis \((\text{Hp})^\ast\) guarantees that for any control \((w^0, w, \alpha, \zeta) \in L^\infty([0, \bar{S}], W(C) \times A \times [-\frac{1}{2}, \frac{1}{2}])\) there exists a unique solution \((y^0, y, y^\ell, \beta)\) to \(\mathcal{P}\), which is defined on the
whole interval \([0, \bar{S}]\). Moreover, the input-output map
\[
\Phi : L^\infty \left( [0, \bar{S}], W(C) \times A \times \left[ -\frac{1}{2}, \frac{1}{2} \right] \right) \to C^0([0, \bar{S}], \mathbb{R} \times \mathbb{R}^n \times \mathbb{R} \times \mathbb{R})
\]
which associates to any control the corresponding solution to (18) turns out to be Lipschitz continuous, when one considers the usual sup-norm over the set of trajectories, while the set of controls is equipped with the following distance \(\tilde{d}\):
\[
(49) \quad \tilde{d}(\{w^0, w, \alpha, \zeta\}, \{\tilde{w}^0, \tilde{w}, \tilde{\alpha}, \tilde{\zeta}\}) := m \{ (w^0, w, \alpha, \zeta)(s) \neq (\tilde{w}^0, \tilde{w}, \tilde{\alpha}, \tilde{\zeta})(s) : s \in [0, \bar{S}] \}
\]
for every pair \((w^0, w, \alpha, \zeta), (\tilde{w}^0, \tilde{w}, \tilde{\alpha}, \tilde{\zeta})\) of controls.

5.1. Needle and bracket-like approximations.

Definition 5.1 (Variation generator). Let us set
\[
\mathcal{W} := \left( W(C) \times A \times \left[ -\frac{1}{2}, \frac{1}{2} \right] \right) \cup \mathcal{W}^0,
\]
where \(W(C)\) is as in (21) and \(\mathcal{W}^0\) as in Definition 4.7. Namely, the elements of \(\mathcal{W}\) are either four-tuples \((w^0, w, a, \zeta) \in \mathbb{R}^n \times C \times \mathbb{R} \times \mathbb{R}\) such that \(w^0 + \|w\| = 1\) or iterated Lie brackets of the vector fields \(g_1, \ldots, g_m\). The elements of \(\mathcal{W}\) will be called variation generators. More specifically, any \(c \in W(C) \times A \times \left[ -\frac{1}{2}, \frac{1}{2} \right]\) will be called a needle variation generator, or a variation generator of length 1, while, for every integer \(h \geq 2\), any \(c = B \in \mathcal{W}^0\) of length \(h\) will be called a bracket-like variation generator of length \(h\).

To every variation generator and to each instant \(\bar{s}\), we now associate an infinitesimal space-time variation of the reference trajectory \((\bar{y}^0, \bar{y}, \bar{y}^T, \beta)\), whose \(y\)-component coincides with either a standard needle variation or a Lie bracket.

As usual, the needle variations will be considered at points \(\bar{s}\) where the augmented optimal dynamics enjoys some continuity-like property, e.g. Lebesgue points.

Definition 5.2. Given an \(L^1\) function \(F : [a, b] \to \mathbb{R}^N\), a point \(s \in (a, b)\) is called a Lebesgue point if \(\lim_{\delta \to 0} \frac{1}{\delta} \int_{s-\delta}^{s+\delta} |F(\sigma) - F(s)|d\sigma = 0\). By Lebesgue Differentiation Theorem, the set of Lebesgue points has measure \(b - a\).

We will use the notation
\[
(0, \bar{S})_{Leb} \subseteq (0, \bar{S})
\]
to denote the full measure subset of Lebesgue points of the function
\[
s \mapsto \begin{pmatrix}
    w^0(s) \\
    f(\bar{y}(s), \bar{\alpha}(s))w^0(s) + \sum_{i=1}^{m} g_i(\bar{y}(s))w^i(s) \\
    \ell^e(\bar{y}(s), w^0(s), \bar{w}(s), \bar{\alpha}(s)) \\
    \bar{w}(s)
\end{pmatrix}, \quad s \in [0, \bar{S}].
\]

Definition 5.3 (Variations).

(NEEDLE) For every \(\bar{s} \in (0, \bar{S})_{Leb}\) and every needle variation generator \(c = (w^0, w, a, \zeta)\), consider the vector
\[
(51) \quad \begin{pmatrix}
    v^0_{c, \bar{s}} \\
    v_{c, \bar{s}} \\
    \ell^e_{c, \bar{s}} \\
    \ell^e_{c, \bar{s}}
\end{pmatrix} := \begin{pmatrix}
    w^0(1 + \zeta) - w^0(\bar{s}) \\
    f(\bar{y}(\bar{s}), a)w^0(1 + \zeta) - f(\bar{y}(\bar{s}), \bar{\alpha}(\bar{s}))w^0(\bar{s}) + \sum_{i=1}^{m} g_i(\bar{y}(\bar{s}))(w^i(1 + \zeta) - w^i(\bar{s})) \\
    \ell^e(\bar{y}(\bar{s}), w^0, w, a)(1 + \zeta) - \ell^e(\bar{y}, w^0, \bar{w}, \bar{\alpha})(\bar{s}) \\
    |w|(1 + \zeta) - |\bar{w}(\bar{s})|
\end{pmatrix}.
\]
For every \( \bar{s} \in (0, \bar{S}) \) and every bracket-like variation generator \( c = B \in \mathfrak{B}^0 \) of length \( h \geq 2 \), one sets

\[
\begin{pmatrix}
  v^{0}_{c, \bar{s}} \\
  v_{c, \bar{s}} \\
  v^{\ell}_{c, \bar{s}}
\end{pmatrix} := \begin{pmatrix}
  0 \\
  \frac{B(y(\bar{s}))}{r_{B}} \\
  0
\end{pmatrix},
\]

where \( r_{B} \) is recursively defined by

(i) \( r_{B} := 1 \) if \( B \) has length 1 (i.e. \( B = g_{j} \), for some \( j = 1, \ldots, m_{1} \)), and

(ii) \( r_{[B_{1}, B_{2}]} := 2(r_{B_{1}} + r_{B_{2}}) \).

**Definition 5.4** (Needle approximation). Let \( c = (w^{0}, w, a, \zeta) \) be a needle variation generator and let \( \bar{s} \in (0, \bar{S}) \). For any control \((\bar{w}^{0}, \bar{w}, \bar{\alpha}, \bar{\zeta}) \in L^{\infty}([0, \bar{S}], W(\mathcal{C}) \times A \times [-\frac{1}{2}, \frac{1}{2}])\)

we say that the family of controls \( \{(\bar{w}^{0}, \bar{w}, \bar{\alpha}, \bar{\zeta})_{c, \bar{s}} : \varepsilon \in (0, \bar{s}]\} \), defined by

\[
(\bar{w}^{0}, \bar{w}, \bar{\alpha}, \bar{\zeta})_{c, \bar{s}}(s) := \begin{cases}
  (w^{0}, w, a, \zeta), & \text{if } s \in [\bar{s} - \varepsilon, \bar{s}] \\
  (\bar{w}^{0}, \bar{w}, \bar{\alpha}, \bar{\zeta})(s), & \text{if } s \in [\bar{s}, \bar{s} - \varepsilon) \cup (\bar{s}, \bar{S}],
\end{cases}
\]

is a needle control approximation of \((\bar{w}^{0}, \bar{w}, \bar{\alpha}, \bar{\zeta})\) at \( \bar{s} \) associated to \( c \).

Lemma 5.1 below, concerning the asymptotics of needle variations, is a standard result (see e.g. [3]). In order to state it, let use \( M(\cdot, \cdot) \) to denote the \( n \times n \) fundamental matrix corresponding to the variational equation

\[
\frac{dV}{ds}(s) = (\frac{\partial f}{\partial x}(\bar{y}(s), \bar{\alpha}(s)) \bar{w}^{0}(s) + \sum_{i=1}^{m} \frac{\partial g_{i}}{\partial x}(\bar{y}(s)) \bar{w}^{i}(s)) \cdot V(s) \text{ a.e. } s \in [0, \bar{S}].
\]

More precisely, for each vector \( v \in \mathbb{R}^{n} \) and each \( s_{1}, s_{2} \in [0, \bar{S}] \), \( s_{1} < s_{2} \), \( M(s_{2}, s_{1}) \cdot v = V(s_{2}) \), where \( V \) is the solution of \([51] \) such that \( V(s_{1}) = v \).

**Lemma 5.1** (Asymptotics of needle variations). Assume that \( \bar{s} \in (0, \bar{S})_{\text{Leb}} \). For every needle variation generator \( c = (w^{0}, w, a, \zeta) \in W(\mathcal{C}) \times A \times [-\frac{1}{2}, \frac{1}{2}] \) and for every \( s \in (\bar{s}, \bar{S}] \), one has

\[
\begin{pmatrix}
  y^{0}(s) - y^{0}(\bar{s}) \\
  y^{\ell}(s) - y^{\ell}(\bar{s}) \\
  y^{\ell}(s) - y^{\ell}(\bar{s}) \\
  \beta^{\ell}(s) - \beta^{\ell}(\bar{s})
\end{pmatrix} = \varepsilon \begin{pmatrix}
  v^{0}_{c, \bar{s}} \\
  v_{c, \bar{s}} \\
  v^{\ell}_{c, \bar{s}}
\end{pmatrix} + o(\varepsilon),
\]

where \((y^{0}, y^{\ell}, y^{\ell}, \beta^{\ell})\) denotes the solution of the rescaled system \([48] \) corresponding to the needle control approximation \((\bar{w}^{0}, \bar{w}, 0)_{c, \bar{s}} \) of \((\bar{w}^{0}, \bar{w}, 0)\) at \( \bar{s} \) associated to \( c \).

Bracket-like approximations, which can be performed in various ways (see e.g. [3]), are here based on the following result:

**Lemma 5.2.** Assume \((H_{p})^{*}\) with \( \gamma < 1 \). Let us choose a point \((\bar{y}^{0}, \bar{y}, \bar{y}^{\ell}, \bar{\beta}) \in \mathbb{R} \times \mathbb{R}^{n} \times \mathbb{R} \times \mathbb{R} \) and an element \( a \in A \). For every Lie bracket \( B \in \mathfrak{B}^0 \) of length \( h \), with \( h \geq 2 \), there exists \( \bar{\varepsilon} > 0 \) such that, for any \( s \in (0, \bar{\varepsilon}^{1/h}] \), there exists a piecewise constant, looping \(4\) control \((w^{0}_{c, \bar{s}}, w_{c, \bar{s}})\), with

\[
\begin{align*}
  w^{0}_{c, \bar{s}}(\sigma) &= 0 \quad \text{for } \sigma \in [0, s], \\
  w_{c, \bar{s}} : [0, s] &\rightarrow \left\{ \pm e_{1}, \ldots, \pm e_{m_{1}} \right\} \times \{0_{m_{2}}\},
\end{align*}
\]

verifying

\[
(\bar{y}^{0}, \bar{y}^{\ell})(\sigma) = (\bar{y}^{0}, \bar{y}^{\ell}) \quad \text{for } \sigma \in [0, s], \quad \beta^{\ell}(s) = \bar{\beta} + s,
\]

[4] "Looping" here refers to the fact that \((w^{0}_{c, \bar{s}}, w_{c, \bar{s}})(s) = (w^{0}_{c, \bar{s}}, w_{c, \bar{s}})(0)\)
For any control

\[ y(s) = \tilde{y} + \left( \frac{s}{r_B} \right)^h B(\tilde{y}) + o(s^h), \tag{57} \]

where we have used \((y^0, y^\ell, \beta)\) to denote the solution to the space-time control system in \((45)\) corresponding to the control \((u_{c,0}^0, w_{c,s}, 0, 0)\) \(^5\) and the initial condition \((y^0, y^\ell, \beta)(0) = (\tilde{y}, \tilde{y}, y^\ell, \tilde{\beta})\), while \(r_B\) is as in Definition \((5.3)\).

**Proof.** While the first relation in \((56)\) is trivial, in that \(u_{c,s}^0 \equiv 0\) and

\[ \ell^c(y, a, w_{c,s}^0, w_{c,s}) = \ell_0(y, a)w_{c,s}^0 + \ell_1(y, w_{c,s})(w_{c,s})^{1-\gamma} = 0, \]

a proof of \((57)\) can be found in \([11]\). Finally, the second relation in \((56)\) is trivial as well, for \(|w_{c,s}| \equiv 1\) on \([0, s]\).

**Definition 5.5** (Bracket-like approximation). Fix an instant \(\bar{s} \in (0, \bar{S})\) and let \(c = B \in \mathfrak{B}^0\) be a bracket-like variation generator of length \(h\), with \(h \geq 2\). For any real number \(\varepsilon > 0\) such that \(\varepsilon < \bar{\varepsilon}\) and \(2\varepsilon^{-1/h} < \bar{s}\), where \(\bar{\varepsilon}\) is as in Lemma \((5.2)\) consider the dilation \(\gamma^\varepsilon: [\bar{s} - 2\varepsilon^{-1/h}, \bar{s} - \varepsilon^{-1/h}] \to [\bar{s} - 2\varepsilon^{-1/h}, \bar{s}]\) given by

\[ \gamma^\varepsilon(\sigma) := (\bar{s} - 2\varepsilon^{-1/h}) + 2(\sigma - (\bar{s} - 2\varepsilon^{-1/h})), \quad \text{for } \sigma \in [\bar{s} - 2\varepsilon^{-1/h}, \bar{s} - \varepsilon^{-1/h}]. \tag{58} \]

For any control \((\bar{w}^0, \bar{w}, \bar{\alpha}, 0) \in L^\infty([0, \bar{S}], W(C) \times A \times [-1/2, 1/2])\) and any \(\varepsilon\) as above, let us define the control \((\bar{w}^0, \bar{w}, \bar{\alpha}, 0)^{\varepsilon}_{c,s}\) by setting

\[ (\bar{w}^0, \bar{w}, \bar{\alpha}, 0)^{\varepsilon}_{c,s}(s) := \left\{ \begin{array}{ll} (2\bar{w}^0, 2\bar{w}, \bar{\alpha}, 0) \circ \gamma^\varepsilon(s), & \text{if } s \in [\bar{s} - 2\varepsilon^{-1/h}, \bar{s} - \varepsilon^{-1/h}], \\ (0, w_{c,\varepsilon^{-1/h}}(s - (\bar{s} - \varepsilon^{-1/h})), a), & \text{if } s \in [\bar{s} - \varepsilon^{-1/h}, \bar{s}], \\ (\bar{w}^0, \bar{w}, \bar{\alpha}, 0)(s), & \text{if } s \in [0, \bar{s} - 2\varepsilon^{-1/h}] \cup (\bar{s}, \bar{S}], \end{array} \right. \tag{59} \]

where \(a \in A\) is arbitrary and \(w_{c,\varepsilon^{-1/h}}\) is as in Lemma \((5.2)\). We refer to the family of controls \(\{(\bar{w}^0, \bar{w}, \bar{\alpha}, 0)^{\varepsilon}_{c,s}: \varepsilon \in (0, \bar{\varepsilon}), 2\varepsilon^{-1/h} < \bar{s}\}\) as a bracket-like control approximation of \((\bar{w}^0, \bar{w}, \bar{\alpha}, 0)\) at \(\bar{s}\) associated to \(c = B\).

**Lemma 5.3** (Asymptotics of bracket-like variations). Let us consider a bracket-like variation generator \(c = B \in \mathfrak{B}^0\), with \(B\) of length \(h \geq 2\). For every point \(\bar{s} \in (0, \bar{S})\) and for each \(\varepsilon > 0\) such that \(\varepsilon < \bar{\varepsilon}\) and \(2\varepsilon^{-1/h} < \bar{s}\), let \((\bar{w}^0, \bar{w}, \bar{\alpha}, 0)^{\varepsilon}_{c,s}\) be a bracket-like control approximation of \((\bar{w}^0, \bar{w}, \bar{\alpha}, 0)\) at \(\bar{s}\) associated to \(c = B\) as in Definition \((5.3)\) and let \((y^0\varepsilon, y^\ell\varepsilon, y^\ell_\beta, \bar{\beta}^\varepsilon)\) be the corresponding solution of the rescaled system \((48)\). Then, for every \(s \in (\bar{s}, \bar{S})\) one has

\[ y^0\varepsilon(s) - \bar{y}^0(s) = 0, \quad y^\ell(s) - \bar{y}(s) = \varepsilon M(s, \bar{s}) \frac{B(\bar{y}(\bar{s}))}{(r_B)^h} + o(\varepsilon), \quad y^\ell_\beta(s) - \bar{y}^\ell(s) = 0 \]

while

\[ \beta^\varepsilon(s) - \bar{\beta}(s) = \varepsilon^{1/h}. \tag{61} \]

**Proof.** The relations concerning the variables \(y^0\) and \(y^\ell\) are trivial. Moreover, by the rate-independence of the space-time control system \((48)\), \(y^\ell = \bar{y} \circ \gamma^\varepsilon\) on \([\bar{s} - 2\varepsilon^{-1/h}, \bar{s} - \varepsilon^{-1/h}]\), and thus one has

\[ y^\ell(\bar{s} - \varepsilon^{-1/h}) = \bar{y}(\bar{s}). \tag{62} \]

\(^5\)Note that the choice of the element \(a\) is irrelevant and purely formal.
Hence
\[
y^\varepsilon(s) - \bar{y}(s) = \int_{s-\varepsilon/h}^{s} \left[ \sum_{i=1}^{m} g_i(y^\varepsilon(s)) w_{c,\varepsilon,1/h}(s - (\bar{s} - \varepsilon/h)) \right] \, ds
\]
(63)
\[
= \int_{0}^{\varepsilon/h} \left[ \sum_{i=1}^{m} g_i(y^\varepsilon(s + (\bar{s} - \varepsilon/h))) w_{c,\varepsilon,1/h}(s) \right] \, ds
\]
where \( w_{c,\varepsilon,1/h} \) is the control associated to the bracket \( B \) as in Lemma 5.2. It follows that \( y^\varepsilon(s) = y^\varepsilon(s + (\bar{s} - \varepsilon/h)) \big|_{s=\varepsilon/h} = y^\varepsilon(\varepsilon/h) \), where we have used \( y^\varepsilon \) to denote the solution to the Cauchy problem
\[
dy{dy}{ds} = \sum_{i=1}^{m} g_i(y(s)), \quad w_{c,\varepsilon,1/h}(s) \quad y(0) = \bar{y}(s),
\]
so that, by Lemma 5.2, we get
\[
y^\varepsilon(s) - \bar{y}(s) - \left( \frac{\varepsilon/h}{r_B} \right)^h B(\bar{y}(s)) = y^\varepsilon(\varepsilon/h) - \bar{y}(s) - \left( \frac{\varepsilon/h}{r_B} \right)^h B(\bar{y}(s)) = o(\varepsilon).
\]
Therefore,
\[
y^\varepsilon(s) - \bar{y}(s) = \varepsilon B(\bar{y}(s)) + o(\varepsilon).
\]
Since for every \( s \in (\bar{s}, \bar{S}] \), the fundamental matrix \( M(s, \bar{s}) \) is the differential of the flow map from \( \bar{s} \) to \( s \), one obtains
\[
y^\varepsilon(s) - \bar{y}(s) = \varepsilon M(s, \bar{s}) \cdot \frac{B(\bar{y}(s))}{(r_B)^h} + o(\varepsilon).
\]
Finally, by the second relation in (56), one has
\[
\beta^\varepsilon(s) - \bar{\beta}(s) = \beta^\varepsilon(\bar{s}) - \bar{\beta}(\bar{s}) = \varepsilon h.
\]
This concludes the proof of the lemma. \( \square \)

5.2. Composition of variations.

Let \( c \in \mathfrak{G} \) be a variation generator of length \( h \), with \( h \geq 1 \), and let \( \bar{s} \in (0, \bar{S}) \). For any \( \varepsilon > 0 \) small enough, let us introduce the map
\[
A_{c,\bar{s}}^\varepsilon : L^\infty \left( [0, \bar{S}], [0, +\infty) \times C \times A \times \left[ -\frac{1}{2}, \frac{1}{2} \right] \right) \rightarrow L^\infty \left( [0, \bar{S}], [0, +\infty) \times C \times A \times \left[ -\frac{1}{2}, \frac{1}{2} \right] \right)
\]
(65)
\[
A_{c,\bar{s}}^\varepsilon(w^0, w, \alpha, \zeta) := (w^0, w, \alpha, \zeta)^{\varepsilon}_{c,\bar{s}},
\]

**Lemma 5.4** (Multiple variations at different times). Let \( N > 0 \) be an integer and let \( \bar{c} := (c_1, \ldots, c_N) \in \mathfrak{G}^N \) be an \( N \)-uple of variations of lengths \( \bar{h} := (h_1, \ldots, h_N) \in \mathbb{N}^N \). Fix \( \bar{s} := (\bar{s}_1, \ldots, \bar{s}_N) \in (0, \bar{S})^N \), with \( 0 =: \bar{s}_0 < \bar{s}_1 < \cdots < \bar{s}_N < \bar{S} \) and \( \bar{s}_j \in (0, \bar{S})_{\text{Leb}} \) as soon as \( h_j = 1 \). For each \( \varepsilon := (\varepsilon_1, \ldots, \varepsilon_N) \in (0, +\infty)^N \) small enough, let us consider the composition varied control
\[
(w^{0\varepsilon}, w^\varepsilon, \alpha^\varepsilon, \zeta^\varepsilon) := A_{c,\bar{s}}^{\varepsilon_1} \circ \cdots \circ A_{c,\bar{s}_N}^{\varepsilon_N} \circ \cdots \circ A_{c,\bar{s}_1}^{\varepsilon_1} (w^0, \bar{w}, \bar{\alpha}, 0),
\]
and let \( (\bar{s}, w^{0\varepsilon}, w^\varepsilon, \alpha^\varepsilon, \zeta^\varepsilon, y^{0\varepsilon}, y^\varepsilon, \beta^\varepsilon) \) denote the corresponding process of (48).

Then, for every \( s \in (\bar{s}_N, \bar{S}] \), one has
\[
y^{0\varepsilon}(s) - \bar{y}^\varepsilon(s) = \sum_{j=1}^{N} \varepsilon_j \left( M(s, \bar{s}_j) v_{c,\bar{s}_j}^{\varepsilon} + o(|\varepsilon|) \right),
\]
(67)

\(^{6}\)Precisely, if \( h \geq 2 \), we require \( 0 < \varepsilon < \bar{\varepsilon}, 2\varepsilon/h < \bar{s} \), as in Definition 5.5 while, in case \( h = 1, \varepsilon < \bar{s} \).
and

$$\beta^*(s) - \beta(s) = \sum_{j \in I_1} \varepsilon_j \left(|w_j|(1 + \zeta_j) - |\tilde{w}(\tilde{s}_j)|\right) + o(|\varepsilon|) + \sum_{j \in I_1^c} (\varepsilon_j)^{\frac{1}{N}},$$

where $I_1 := \{j = 1, \ldots, N : h_j = 1\}$ and $I_1^c := \{1, \ldots, N\} \setminus I_1$. In particular, if all $c_j$ are needle variations, i.e. $c_j := (w_j^0, w_j, a_j, \zeta_j)$ for every $j = 1, \ldots, N$, one gets

$$\beta^*(s) - \beta(s) = \sum_{j=1}^{N} \varepsilon_j \left(|w_j|(1 + \zeta_j) - |\tilde{w}(\tilde{s}_j)|\right) + o(|\varepsilon|).$$

**Proof.** Let us prove the theorem by induction on $N$, the number of composed variations. For $N = 1$, the result is proved in Lemmas 5.1 and 5.3.

If $N \geq 2$, let us assume that the result holds true for $N - 1$ and let us show that it holds for $N$ as well. Let us use $(y^0, y, \beta)^N$ and $(y^0, y, \beta)^{N-1}$ to denote the trajectories associated to the $N$ variations and to the first $N - 1$ variations, respectively (we omit to indicate the dependence on $\varepsilon$ for brevity). We have

$$\begin{pmatrix} y^0(N)(\tilde{s}_N) - \tilde{y}(\tilde{s}_N) \\ y^N(\tilde{s}_N) - \tilde{y}(\tilde{s}_N) \\ y^t(N-1)(\tilde{s}_N) - \tilde{y}^t(\tilde{s}_N) \\ \beta^N(\tilde{s}_N) - \beta(\tilde{s}_N) \end{pmatrix} = \begin{pmatrix} y^0(N-1)(\tilde{s}_N) - \tilde{y}(\tilde{s}_N) \\ y^N(\tilde{s}_N) - y^{N-1}(\tilde{s}_N) \\ y^t(N-1)(\tilde{s}_N) - \tilde{y}^t(\tilde{s}_N) \\ \beta^N(\tilde{s}_N) - \beta^{N-1}(\tilde{s}_N) \end{pmatrix}.$$
Case I: the \( N \)th variation is a needle variation, namely \( c_N = (w_N^0, w_N, a_N, \zeta_N) \). One has
\[
y^N(\bar{s}_N) - y^{N-1}(\bar{s}_N)
= \int_{\bar{s}_N - \epsilon_N}^{\bar{s}_N} \left[ f(y^N(s), a_N)w_N^0(1 + \zeta_N) + \sum g_i(y^N(s))w_N^i(1 + \zeta_N)
- f(y^{N-1}(s), \bar{a}(s))\bar{w}^0(s) - \sum g_i(y^{N-1}(s))\bar{w}^i(s) \right] \, ds
\]
where
\[
r_1(s) := f(y^N(s), a_N)w_N^0(1 + \zeta_N) + \sum g_i(y^N(s))w_N^i(1 + \zeta_N)
- f(\bar{y}(\bar{s}_N), a_N)w_N^0(1 + \zeta_N) - \sum g_i(\bar{y}(\bar{s}_N))w_N^i(1 + \zeta_N),
\]
\[
r_2(s) := f(\bar{y}(\bar{s}_N), \bar{a}(\bar{s}_N))\bar{w}^0(\bar{s}_N) - \sum g_i(\bar{y}(\bar{s}_N))\bar{w}^i(\bar{s}_N),
\]
\[
r_3(s) := f(y^{N-1}(s), \bar{a}(\bar{s}_N))\bar{w}^0(\bar{s}_N) + \sum g_i(y^{N-1}(s))\bar{w}^i(\bar{s}_N)
- f(y^{N-1}(s), \bar{a}(s))\bar{w}^0(\bar{s}_N) - \sum g_i(y^{N-1}(s))\bar{w}^i(s).
\]
Let us start by estimating \( r_1 \). Observe that, for \( s \in [\bar{s}_N - \epsilon_N, \bar{s}_N] \),
\[
|y^N(s) - \bar{y}(\bar{s}_N)| \leq |y^N(s) - y^{N-1}(s)| + |y^{N-1}(s) - \bar{y}(s)| + |\bar{y}(s) - \bar{y}(\bar{s}_N)|,
\]
where there is some \( K > 0 \) such that
\[
\|y^N - y^{N-1}\|_{C^0([\bar{s}_N - \epsilon_N, \bar{s}_N], \mathbb{R}^n)} \leq K \epsilon_N
\]
in view of the Lipschitz continuity of the input-output map for the control set equipped with the distance \( \tilde{d} \), introduced in (19). Moreover, by the inductive hypothesis, one has
\[
\|y^N - \bar{y}\|_{C^0([\bar{s}_N - \epsilon_N, \bar{s}_N], \mathbb{R}^n)} \leq K(\epsilon_1 + \cdots + \epsilon_{N-1}),
\]
while, by the Lipschitz continuity of the reference trajectory (w.r.t. time), one deduces that
\[
\|\bar{y}(s) - \bar{y}(\bar{s}_N)\|_{C^0([\bar{s}_N - \epsilon_N, \bar{s}_N], \mathbb{R}^n)} \leq K \epsilon_N.
\]
Hence, one obtains the estimate
\[
\|y^N(s) - \bar{y}(\bar{s}_N)\|_{C^0([\bar{s}_N - \epsilon_N, \bar{s}_N], \mathbb{R}^n)} \leq 3K|\bar{\epsilon}|,
\]
so that
\[
\left| \int_{\bar{s}_N - \epsilon_N}^{\bar{s}_N} r_1(s) \, ds \right| \leq \int_{\bar{s}_N - \epsilon_N}^{\bar{s}_N} L(w_N^0 + |w_N|)(1 + \zeta_N)|y^N(s) - \bar{y}(\bar{s}_N)| \, ds
\]
\[
\leq \int_{\bar{s}_N - \epsilon_N}^{\bar{s}_N} \frac{9}{2} LK|\bar{\epsilon}| \, ds \leq \epsilon_N O(|\bar{\epsilon}|),
\]
where \( L \) is the maximum of the Lipschitz constants of \( f \) and the \( g_i \)'s. By estimate (79), and recalling that \( \bar{s}_N \) is a Lebesgue point as in Definition 5.2, we derive that,
\[
\left| \int_{\bar{s}_N - \epsilon_N}^{\bar{s}_N} r_3(s) \, ds \right|
\leq \int_{\bar{s}_N - \epsilon_N}^{\bar{s}_N} \sup_{|y - \bar{y}(\bar{s}_N)| \leq K|\bar{\epsilon}|} \left| f(\bar{y}(\bar{s}_N), \bar{a}(\bar{s}_N))\bar{w}^0(\bar{s}_N)
+ \sum g_i(\bar{y}(\bar{s}_N))\bar{w}^i(\bar{s}_N) - f(y, \bar{a}(s))\bar{w}^0(s) - \sum g_i(y)\bar{w}^i(s) \right| \, ds \leq O(\epsilon_N).
\]
Notice that \( \frac{\mathcal{O}(\varepsilon^N)}{\varepsilon^N} \to 0 \) when \( |\bar{z}| \to 0 \). Therefore,

\begin{equation}
(83) \quad y^N(\bar{s}_N) - y^{N-1}(\bar{s}_N) = \varepsilon_N r_2 + \varepsilon_N \mathcal{O}(|\bar{z}|),
\end{equation}

and the relation in (73) concerning the state variables is proven.

**Case II:** the \( N \)th variation has length bigger than 2, i.e. \( c_N = B_N \in \mathcal{B}^0 \), with \( h_N \geq 2 \).

We know that

\begin{equation}
(84) \quad y^N(\bar{s}_N - \varepsilon_N^{1/h_N}) = y^{N-1}(\bar{s}_N).
\end{equation}

Hence we have

\[
y^N(\bar{s}_N) - y^{N-1}(\bar{s}_N) = \int_{\bar{s}_N - \varepsilon_N^{1/h_N}}^{\bar{s}_N} \sum g_i(y^N(s)) w^i \left( \frac{1}{h_N} \right) ds,
\]

where \( w = (\varepsilon_N^{1/h_N}) \) is the control associated to the bracket \( B_N \) as in Lemma 5.2. In view of Lemma 5.2, we get the estimate

\begin{equation}
(85) \quad \left| \int_{\bar{s}_N - \varepsilon_N^{1/h_N}}^{\bar{s}_N} \sum g_i(y^N(s)) w^i \left( \frac{1}{h_N} \right) ds - \frac{\varepsilon_N B_N(\bar{y}(\bar{s}_N))}{(r_N(B_N))^{h_N}} \right|
\end{equation}

\[
\leq \left| y[y^{N-1}(\bar{s}_N), w] \left( \frac{1}{h_N} \right) \right| - y^{N-1}(\bar{s}_N) - \frac{\varepsilon_N B_N(\bar{y}(\bar{s}_N))}{(r_N(B_N))^{h_N}}
\]

\[
+ \varepsilon N \left| B_N(y^N(\bar{s}_N)) - B_N(\bar{y}(\bar{s}_N)) \right|
\]

\[
\leq \mathcal{O}(\varepsilon_N) + \varepsilon N \mathcal{O}(|\bar{z}|) = o(\varepsilon_N),
\]

where \( \mathcal{O}(\varepsilon_N) \to 0 \) when \( \bar{z} \to 0 \). This yields (73).

At this point, by (71)–(74) the proof of (67), (68) can be concluded, recalling that, for every \( s \in (\bar{s}_N, \bar{s}) \), the fundamental matrix \( M(s, \bar{s}_N) \) is the differential of the flow map from \( \bar{s}_N \) to \( s \), so that

\[
y^N(s) - \bar{y}(s) = M(s, \bar{s}_N) \cdot (y^N(\bar{s}_N) - \bar{y}(\bar{s}_N)).
\]

\[\square\]

### 5.3. Set separation

Given a control \((w^0, w, \alpha, \zeta) \in L^\infty \left( [0, \bar{S}], [0, +\infty) \times C \times A \times [-\frac{1}{T}, \frac{1}{T}] \right)\) and the corresponding solution \((y^0, y, y^f, \beta)\) of the rescaled space-time control system (43), we set

\[
y^f(s) := \Psi(y^0(s), y(s)) + y^f(s), \quad \text{for } s \in [0, \bar{S}],
\]

and we call

\[
(\bar{S}, w^0, w, \alpha, \zeta, y^0, y, y^f, \beta)
\]

a cost-valued space-time process\(^7\). We will use \((\bar{S}, w^0, w, \alpha, 0, \bar{y}, \bar{y}, \bar{y}, \bar{y}, \bar{y}, \beta)\) to denote the cost-valued space-time process corresponding to the reference process.

For every \( \delta > 0 \), let us define the augmented \( \delta \)-reachable set \( \mathcal{R}_\delta \) and its projection \( \mathcal{R}_\delta' \):

\(^7\)As is well known, the presence of the relation \( y^f = \Psi(y_0, y) \) doesn’t change the fact that (43) is a genuine Cauchy problem, for it can be replaced by the differential relation \( \frac{d\bar{y}^f}{ds} = \frac{\partial\Psi}{\partial y} w^0 + \frac{\partial\Psi}{\partial \alpha} f(y, \alpha) w^0 + \sum_{i=1}^m g_i w^i + \frac{\partial\Psi}{\partial \zeta} \) with the initial condition \( y^f(0) = \Psi(0, \bar{y}, 0) \).
5.4. By (67) we get

\[ E = \text{span} \]  

Lemma 5.5.

When all \( c_j = (w_j^0, w_j, a_j, \zeta_j) \) are needle variations, \( j=1, \ldots N \), let us consider the set

\[ E := \left\{ \left( \frac{\partial \psi}{\partial t}(\bar{y}^0, \bar{y})((\bar{S}), \mathbf{v})_{\bar{S}}, \bar{S} \right) + \frac{\partial \psi}{\partial x}((\bar{y}^0, \bar{y}))(\bar{S}) \cdot M((\bar{S}, \bar{\zeta}), \mathbf{v}(\bar{S}, \bar{\zeta})) \cdot \mathbf{v}(\bar{S}, \bar{\zeta}) + \mathbf{v}^\ell_{\bar{S}, \bar{\zeta}} : j = 1, \ldots N \right\}. \]

Furthermore, in the general case (i.e. when some \( c_j \) are allowed to be bracket-like variations) but under the hypothesis that \( \gamma < 1 \), let us consider the set

\[ E' := \left\{ \left( \frac{\partial \psi}{\partial t}(\bar{y}^0, \bar{y})((\bar{S}), \mathbf{v})_{\bar{S}}, \bar{S} \right) + \frac{\partial \psi}{\partial x}((\bar{y}^0, \bar{y}))(\bar{S}) \cdot M((\bar{S}, \bar{\zeta}), \mathbf{v}(\bar{S}, \bar{\zeta})) \cdot \mathbf{v}(\bar{S}, \bar{\zeta}) + \mathbf{v}^\ell_{\bar{S}, \bar{\zeta}} : j = 1, \ldots N \right\}. \]

Finally, let us define the convex cones

\[ R := \text{span}_+(E) \subset \mathbb{R}^{1+n+2}, \quad R' := \text{span}_+(E') \subset \mathbb{R}^{1+n+1}, \]

where, for a given subset \( \Theta \) of a vector space we have used \( \text{span}_+(\Theta) \) to denote its positive span.

Lemma 5.5.

(i) \( R' \) is a Boltyanskii approximating cone of the set \( \mathcal{R}'_\delta \) at the point \((\bar{y}^0, \bar{y}, \bar{y}')(\bar{S})\).

(ii) Furthermore, when all \( c_j = (w_j^0, w_j, a_j, \zeta_j) \) are needle variations, \( j=1, \ldots N \), \( R \) is a Boltyanskii approximating cone of the set \( \mathcal{R} \) at \((\bar{y}^0, \bar{y}, \bar{y}')(\bar{S})\).

Proof. Let us set \( y^{\bar{c}}(s) := y^{\bar{c}}(s) + \psi((y^{\bar{c}}'', y^{\bar{c}}')(s)) \), where \( y^{\bar{c}}, y^{\bar{c}}', \) and \( y^{\bar{c}}'' \) are as in Lemma 5.3. By (87) we get

\[ y^{\bar{c}}(S) - y^{\bar{c}}(S) = \sum_{j=1}^N \epsilon_j \left( \frac{\partial \psi}{\partial t}(\bar{y}^0, \bar{y})(\bar{S}) + \frac{\partial \psi}{\partial x}((\bar{y}^0, \bar{y}))(\bar{S}) \cdot M((\bar{S}, \bar{\zeta}), \mathbf{v}(\bar{S}, \bar{\zeta})) \cdot \mathbf{v}(\bar{S}, \bar{\zeta}) + \mathbf{v}^\ell_{\bar{S}, \bar{\zeta}} + o(|\epsilon|). \]

Therefore, part (ii) of the statement is the standard result by which one defines the variational cone in the standard proof of the Maximum Principle.

To prove part (i), for some \( \bar{\epsilon} > 0 \) small enough let us define the map \( F : (0, +\infty)^N \cap \bar{\epsilon} \mathbb{B}_N \to \mathbb{R}^{1+n+2} \) by setting

\[ F(\bar{\epsilon}) = \left( y^{\bar{c}}(S), y^{\bar{c}}(S), y^{\bar{c}}(S) \right). \]

It is straightforward to prove that

\[ F(\bar{\epsilon}) = \left( y^0(S), y(S), y^{\bar{c}}(S) \right) + L \cdot \bar{\epsilon} + o(|\bar{\epsilon}|) \]

where the linear operator \( L \in \text{Hom}(\mathbb{R}_N, \mathbb{R}^{1+n+1}) \) is defined by

\[ L \cdot \bar{\epsilon} = \sum_{j=1}^N \epsilon_j \left( \frac{\partial \psi}{\partial t}(\bar{y}^0, \bar{y})(\bar{S}) + \frac{\partial \psi}{\partial x}((\bar{y}^0, \bar{y}))(\bar{S}) \cdot M((\bar{S}, \bar{\zeta}), \mathbf{v}(\bar{S}, \bar{\zeta})) \cdot \mathbf{v}(\bar{S}, \bar{\zeta}) + \mathbf{v}^\ell_{\bar{S}, \bar{\zeta}} + \mathbf{v}^\ell_{\bar{S}, \bar{\zeta}} \right), \]

\footnote{Let us remind that \( \mathbf{v}^\ell_{\bar{S}, \bar{\zeta}} = 0 \) as soon as the variation generator \( c_j \) has length > 1.}
and \( \mathbf{v}_{c_i,s_j}^0, \mathbf{v}_{c_j,s_j}, \mathbf{v}_{c_j,s_j}^\ell \) are as in Definition 5.3. Hence (i) is proved, for \( R' = L(0, +\infty)^N \).

Let us consider the profitable set

\[
P := \left( \Sigma \times (-\infty, \bar{y}^r(\bar{S})) \times [0, K] \right) \bigcup \left\{ (\bar{y}^0, \bar{y}, \bar{y}^c, \bar{\beta})(\bar{S}) \right\},
\]

and let \( \Gamma \) be a Boltyanskii approximating cone for the target \( \Sigma \) at \( (\bar{y}^0, \bar{y})(\bar{S}) \). Set: i) \( J = \mathbb{R} \) if \( \bar{\beta}(\bar{S}) \in (0, K) \), ii) \( J = (-\infty, 0] \) if \( \bar{\beta}(\bar{S}) = K \), and iii) \( J = [0, +\infty) \) if \( \bar{\beta}(\bar{S}) = 0 \). It is trivial to check that

\[
P := \Gamma \times (-\infty, 0] \times J,
\]
is a Boltyanskii approximating cone for \( P \) at \( (\bar{y}^0, \bar{y}, \bar{y}^c, \bar{\beta})(\bar{S}) \).

Let us also consider the projected profitable set

\[
P' := \left( \Sigma \times (-\infty, \bar{y}^r(\bar{S})) \right) \bigcup \{(\bar{y}^0, \bar{y}, \bar{y}^c)(\bar{S})\}.
\]
It also is trivial to check that

\[
P' := \Gamma \times (-\infty, 0],
\]
is a Boltyanskii approximating cone for \( P' \) at \( (\bar{y}^0, \bar{y}, \bar{y}^c)(\bar{S}) \).

We will need the following elementary result:

**Lemma 5.6.** Let \( (\bar{S}, \bar{v}^0, \bar{w}, \bar{\alpha}, \bar{y}^0, \bar{y}, \bar{\beta}) \) be a local minimizer such that \( \bar{\beta}(\bar{S}) < K \). Then there exists a \( \delta > 0 \) such that the sets \( P' \) and \( R'_\delta \) are locally separated at \( (\bar{y}^0, \bar{y}, \bar{y}^c)(\bar{S}) \).

**Proof.** Suppose that for every \( \delta > 0 \) the sets \( P' \) and \( R'_\delta \) are not locally separated at \( (\bar{y}^0, \bar{y}, \bar{y}^c)(\bar{S}) \). Fix \( \delta > 0 \) such that \( \delta < K - \bar{\beta}(\bar{S}) \). Then there would exist a sequence of controls \( (\bar{S}, \bar{w}_n, \bar{w}_n, \bar{\alpha}_n, \bar{\zeta}_n) \) such that the corresponding trajectories of (45) verify

\[
(y_n^0, y_n, y_n^c, \beta_n)(\bar{S}) \in R'_\delta \cap P' \quad \forall n \in \mathbb{N}
\]
and

\[
(y_n^0, y_n, y_n^c)(\bar{S}) \to (\bar{y}^0, \bar{y}, \bar{y}^c)(\bar{S}).
\]
Notice that by (55) one has \( \delta - \bar{\beta}(\bar{S}) \leq \beta_n(S_n) \leq \delta + \bar{\beta}(\bar{S}) < K \), so that \( (y_n^0, y_n, y_n^c, \beta_n(S_n)) \in P \cap R'_\delta. \) In particular, for every \( \delta \), the sets \( P' \) and \( R'_\delta \) are not locally separated, which is in contradiction with the fact that \( (\bar{s}, \bar{w}, \bar{\alpha}, \bar{y}^0, \bar{y}, \bar{\beta}) \) is a local minimizer. Theorem 5.6.}

Since \( (\bar{S}, \bar{w}^0, \bar{w}, \bar{\alpha}, 0, \bar{y}^0, \bar{y}, \bar{\beta}) \) is local minimizer and \( \bar{\beta}(\bar{S}) < K \), by Lemma 5.6 the projected reachable set \( R'_\delta \) must be locally separated from the projected profitable set \( P' \) at \( (\bar{y}^0, \bar{y}, \bar{y}^c)(\bar{S}) \), for some \( \delta > 0 \). Therefore, since \( R' \) and \( P' \) are approximating cone to \( R' \) and \( P' \), respectively, and, furthermore, \( P' \) is not a subspace, there exists \( (\xi_0, \xi, \xi_c) \in (\mathbb{R}^{1+n+1} \cup (-P') \) verifying

\[
0 \neq (\xi_0, \xi, \xi_c) \in R'_\perp \cap (-P'_\perp).
\]
Since \( P'_\perp = \Gamma \times [0, +\infty) \) one gets

\[
(\xi_0, \xi) \in -\Gamma \perp, \quad \xi_c = -\lambda \leq 0,
\]
and

\[
\xi_0 \mathbf{v}^0 + \xi \cdot \mathbf{v} + \xi_c \mathbf{v}^c \leq 0 \quad \forall (\mathbf{v}^0, \mathbf{v}, \mathbf{v}^c) \in R'.
\]
Recalling the form of \( R' \), (59) is verified if and only if

\[
\left( \xi_0 - \lambda \frac{\partial \Psi}{\partial x}(\bar{y}^0, \bar{y})(\bar{S}) \right) \mathbf{v}_{c_i,s_j}^0 + \left( \xi - \lambda \frac{\partial \Psi}{\partial x}(\bar{y}^0, \bar{y})(\bar{S}) \right) \cdot M(\bar{S}, s_j) \cdot \mathbf{v}_{c_j,s_j} - \lambda \mathbf{v}_{c_j,s_j}^\ell \leq 0,
\]
for all \( j = 1, \ldots, N \). Setting

\[
(p_0, p)(s) := (\xi_0, \xi) \cdot \frac{\partial \Psi}{\partial t}(\bar{y}^0, \bar{y})(\bar{S}), \quad \frac{\partial \Psi}{\partial x}(\bar{y}^0, \bar{y})(\bar{S}) \cdot M(\bar{S}, s)
\]
we obtain that the multipliers \((p_0, p, \lambda) \in \mathbb{R} \times AC \left([0, \bar{S}], \mathbb{R}^n\right) \times [0, +\infty)\) verify
\[
(90) \quad p_0 \mathbf{v}^0_{\bar{c}_j, \bar{s}_j} + p(\bar{s}_j) \cdot \mathbf{v}_{\bar{c}_j, \bar{s}_j} - \lambda \mathbf{v}^\ell_{\bar{c}_j, \bar{s}_j} \leq 0
\]
for every \(j = 1, \ldots, N\), the non-triviality condition \(\text{[(35)]}\) i.e. \((p_0, p, \lambda) \neq 0\), and (by \(M(\bar{S}, \bar{S}) = Id\)) the non-transversality condition \(\text{[(35)]}\), i.e.
\[
(p_0, p(\bar{S})) \in -\lambda \left( \frac{\partial \Psi}{\partial t}((\bar{y}^0, \bar{y})(\bar{S})), \frac{\partial \Psi}{\partial x}((\bar{y}^0, \bar{y})(\bar{S})) \right) - \Gamma^\perp.
\]

Moreover, by the definition of \(\mapsto M(\bar{S}, \cdot)\), the path \(p\) solves the adjoint equation \(\text{[(37)]}\).

Finally, for a needle variation generator \(\mathbf{c}_j = (w^0_j, w_j, a_j, \zeta_j)\), by \(\text{[(91)]}\) we get
\[
(91) \quad H\left(\bar{y}(\bar{s}_j), p_0, p(\bar{s}_j), 0, \lambda, w^0_j(1+\zeta_j), w_j(1+\zeta_j), a_j\right) - H\left(\bar{y}(\bar{s}_j), p_0, p(\bar{s}_j), 0, \lambda, \bar{w}^0(\bar{s}_j), \bar{w}(\bar{s}_j), \bar{\alpha}(\bar{s}_j)\right) \leq 0,
\]
while, for a bracket-like variation generator \(\mathbf{c}_j = B\), we obtain
\[
(92) \quad p(\bar{s}_j) \cdot B(\bar{y}(\bar{s}_j)) \leq 0.
\]

5.4. Conclusion of the proof. To conclude the proof we need to extend the previous inequalities to almost all \(s \in [0, \bar{S}]\) and to all variations generators \(\mathbf{c} \in \mathfrak{M}\). This will obtained via density arguments coupled with infinite intersection criteria. Though this is a quite standard procedure, we give the details for the sake of completeness.

By Lusin’s Theorem, one has that \(E = \bigcup_{k=0}^{+\infty} E_k\) where \(E_0\) has null measure and, for every \(k \in \mathbb{N}\), the set \(E_k\) is compact and the restriction to \(E_k\) of the measurable map considered in \(\text{[(50)]}\) is continuous (see Definition 5.2). For every \(k\), let \(D_k \subseteq E_k\) be the set of density points of \(E_k\). Since \(D_k\) and \(E_k\) have the same measure, by Lebesgue density Theorem, \(D = \bigcup_{k=0}^{+\infty} D_k \subset [0, \bar{S}]\) has full measure.

**Definition 5.6.** Let us consider an arbitrary subset \(F\) of \(D \times \mathfrak{M}\). One says that a triple \((\bar{p}_0, \bar{p}, \lambda) \in \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}\) verifies (P) \(\text{\(F\)}\) if, setting \(p(\cdot) = \bar{p} \cdot M(\bar{S}, \cdot)\), one has

(i) \((\bar{p}_0, p(\bar{S})) + \lambda \left( \frac{\partial \Psi}{\partial t}((\bar{y}^0, \bar{y})(\bar{S})), \frac{\partial \Psi}{\partial x}((\bar{y}^0, \bar{y})(\bar{S})) \right) \in -\Gamma^\perp;\)

(ii) inequality

\[\text{[(93)]} \quad H\left(\bar{y}(s), p_0, p(s), 0, \lambda, w^0(1+\zeta), w(1+\zeta), a\right) - H\left(\bar{y}(s), p_0, p(s), 0, \lambda, \bar{w}^0(s), \bar{w}(s), \bar{\alpha}(s)\right) \leq 0,\]

holds true for every \((s, \mathbf{c}) \in F\) with \(\mathbf{c} = (w^0, w, a, \zeta)\), while for every \((s, \mathbf{c}) \in F\) such that \(\mathbf{c}_j = B \in \mathfrak{M}^0\), one has

\[\text{[(94)]} \quad p(s) \cdot B(\bar{y}(s)) \leq 0;\]

(iii) \(\lambda \geq 0.\)

For any given subset \(F \subset D \times \mathfrak{M}\), let us set
\[
\Lambda(F) := \left\{ (\bar{p}_0, \bar{p}, \lambda) \in \mathbb{R}^x \mathbb{R}^n \times \mathbb{R} : \left| (\bar{p}_0, \bar{p}, \lambda) \right| = 1, (\bar{p}_0, \bar{p}, \lambda) \text{ verifies (P)\(\text{\(F\)}\)} \right\},
\]

Notice that, in principle, this set \(\Lambda(F)\) might be empty. Instead, our goal consists in showing that \(\Lambda(F) \neq \emptyset\) for some \(F\) comprising pairs \((s, \mathbf{c})\) such that the union of all times \(s\) is a full measure subset of \([0, \bar{S}]\) and \(\mathbf{c}\) can range all over \(\mathfrak{M}\). Clearly, for arbitrary subsets \(F_1, F_2\) of \(D \times \mathfrak{M}\) the sets \(\Lambda(F_1), \Lambda(F_2)\) are compact and \(\Lambda(F_1) \cap \Lambda(F_2) = \Lambda(F_1) \cap \Lambda(F_2)\).
We have already proved in the previous step that $\Lambda(F) \neq \emptyset$ as soon as $F$ is finite and of the form
\begin{equation}
\left\{(\bar{s}_1, c_1), \ldots, (\bar{s}_N, c_N)\right\}, \text{ with } 0 =: \bar{s}_0 < \bar{s}_1 < \cdots < \bar{s}_N < \bar{S}.
\end{equation}
In order to prove that $\Lambda(F) \neq \emptyset$ for an arbitrary finite set $F \subset D \times \mathcal{W}$, we have to show that it is non-empty even in case
\begin{equation}
F = \left\{(\bar{s}_1, c_1), \ldots, (\bar{s}_N, c_N)\right\}, \text{ with } 0 =: \bar{s}_0 \leq \bar{s}_1 \leq \cdots \leq \bar{s}_N \leq \bar{S},
\end{equation}
where one allows that $\bar{s}_j = \bar{s}_{j+1}$ for some $j = 1, \ldots, N - 1$. To this end, observe that every $\bar{s}_j$ belongs to some set of density points $D_k$, that we denote $D_{k(j)}$. Hence there exists a sequence $(\bar{s}_{j,i})_{i \in \mathbb{N}}$ such that
\begin{equation*}
\bar{s}_{j,i} \in D_{k(j)} \quad \text{and} \quad \bar{s}_{1,i} < \cdots < \bar{s}_{N,i}, \quad \forall i, \quad \text{and} \quad \lim_{i \to +\infty} \bar{s}_{j,i} = \bar{s}_j,
\end{equation*}
For each $i$, set $F_i := \left\{(\bar{s}_{1,i}, c_1), \ldots, (\bar{s}_{N,i}, c_N)\right\}$, so that $F_i$ has the form (95) and hence $\Lambda(F_i) \neq \emptyset$. For each $i \in \mathbb{N}$, let us select $(\bar{p}_0, \bar{p}_i, \lambda_i) \in \Lambda(F_i)$. Since $|(\bar{p}_0, \bar{p}_i, \lambda_i)| = 1$, by possibly considering a subsequence, we can assume that $(\bar{p}_0, \bar{p}_i, \lambda_i)$ converges to a point $(\bar{p}_0, \lambda)$ with $|(\bar{p}_0, \lambda)| = 1$. By the definition of $D_{k(j)}(\subseteq E_{k(j)})$, passing to the limit as $j \to +\infty$ one obtains that $(\bar{p}_0, \lambda) \in \Lambda(F)$. Hence we have proved that $\Lambda(F) \neq \emptyset$ as soon as $\operatorname{card}(F) < +\infty$. In particular, if we take a finite family of subsets $F_1, \ldots, F_M \subset D \times \mathcal{W}$ with $\operatorname{card}(F_i) < +\infty$ for all $i = 1, \ldots, M$, we get
\begin{equation*}
\Lambda(F_1) \cap \cdots \cap \Lambda(F_M) = \Lambda\left(\bigcup_{i=1}^M F_i\right) \neq \emptyset.
\end{equation*}
Hence
\begin{equation*}
\left\{\Lambda(F) : F \subset D \times \mathcal{W}, \operatorname{card}(F) < +\infty\right\}
\end{equation*}
is a family of compact subsets such that the intersection of each finite subfamily is non-empty. This implies that also the (finite) intersection of all $\Lambda(F)$ over finite sets $F$ is non-empty. Therefore
\begin{equation*}
\Lambda(D \times \mathcal{W}) = \Lambda\left(\bigcup_{\operatorname{card}(F) < +\infty} F\right) = \bigcap_{\operatorname{card}(F) < +\infty} \Lambda(F) \neq \emptyset.
\end{equation*}
This means that there exists some co-vector $(\bar{p}_0, \bar{p}, \lambda) \neq 0$ such that, setting $p(s) := \bar{p} \cdot M(\bar{S}, s)$, one gets
\begin{equation}
\begin{aligned}
H(\bar{y}(s), p_0, p(s), 0, \lambda) &= H\left(\bar{y}(s), p_0, p(s), 0, \lambda, \bar{w}(s), \bar{\alpha}(s)\right) \\
&= \max_{(\omega, \omega, a, \zeta) \in W(\mathcal{C}) \times Ax[-\frac{1}{2}, \frac{1}{2}]} H\left(\bar{y}(s), p_0, p(s), 0, \lambda, \bar{w}(1 + \zeta), w(1 + \zeta), a\right) \\
&= \max_{\zeta \in [-\frac{1}{2}, \frac{1}{2}]} (1 + \zeta) H(\bar{y}(s), p_0, p(s), 0, \lambda),
\end{aligned}
\end{equation}
and
\begin{equation}
p(s) \cdot B(\bar{y}(s)) \leq 0, \quad \text{for all } B \in \mathfrak{W},
\end{equation}
for all instants $s$ in the full-measure set $D$.

The first relation in (96) coincides with (S3), while the second relation immediately implies (39), that is,
\begin{equation*}
H(\bar{y}(s), p_0, p(s), 0, \lambda) = 0 \quad \text{for all } s \in [0, \bar{S}].
\end{equation*}
Finally, observe that $B \in \mathfrak{W}$ if and only if $-B \in \mathfrak{W}$, so that (97) yields that, for all $s \in [0, \bar{S}]$,
\begin{equation*}
p(s) \cdot B(\bar{y}(s)) = 0, \quad \text{for all } B \in \mathfrak{W},
\end{equation*}
which coincides with the conservation relation stated in (41). This concludes the proof, since, in case $\bar{y}(\bar{S}) > 0$, the strengthened non-triviality condition (42), i.e. $(p, \lambda) \neq (0, 0)$, can be obtained as in proof of the First Order Maximum Principle.

6. Generalizations and final remarks

6.1. The result on manifolds. One of the main motivation for studying impulsive systems is Classical Mechanics, where it often happens that the state ranges over a differential manifold (typically, the cotangent bundle of a state-space manifold). This, as well as other applications [12, 7, 15], motivate the interest for a generalization of the previous results to the case of a differential manifold. Though this extension does not present any special difficulty, we devote this subsection to make the corresponding notions and statements precise. This is also the occasion to point out that all definitions and results in this paper have a chart-independent character. Let us begin by generalizing the notion of Boltyanskii cone.

Definition 6.1 (Boltyanskii cones on manifolds). Let $\Omega$ be a $C^1$ Riemannian manifold with a metric $d$, and let $Z \subseteq \Omega$ be any subset. Consider a point $z \in Z$ and let $K \subset T_z\Omega$ be a convex cone, where $T_z\Omega$ denotes the tangent space at $z$. We say that $K$ is a Boltyanskii approximating cone to $Z$ at $z$ if there exist a convex cone $C \subset \mathbb{R}^\beta$ (for some integer $\beta \geq 0$), a neighborhood $V \subseteq \mathbb{R}^\beta$ of 0, and a continuous map $F : V \cap C \to Z$ such that, for one (hence any) coordinate chart $(\xi, U)$ on a neighborhood $U$ of $z$, there exists an $n \times \beta$ matrix $L$ verifying

$$\xi(F(v)) = \xi(z) + Lv + o(|v|) \quad \text{for all } v \in V \cap C \quad \text{and} \quad LC = K,$$

where $o(|v|)$ denotes a function $g : F : V \cap C \to \xi(U)$ such that $\lim_{v \to 0} \frac{g(v)}{|v|} = 0$.

By passing to coordinates it is trivial to check that Theorem 4.1 connecting transversality and set separation remains unchanged if we take $\Omega$ to be a $C^1$ manifold.

Aiming to the definition of Hamiltonians on the cotangent bundle $T^*\Omega$, we use $\Pi : T^*\Omega \to \Omega$ to denote the fiber-bundle projection of $T^*\Omega$ on $\Omega$, namely $\Pi(P) = x$ as soon as $P \in T^*_x\Omega$.

Definition 6.2 (Hamiltonian). Let $\Omega$ be a $C^1$ manifold. We will call unmaximized Hamiltonian the function $H : T^*\Omega \times \mathbb{R}^2 \times [0, +\infty) \times C \times A \to \mathbb{R}$ defined by setting, for every $P \in T^*\Omega$, $(p_0, \pi) \in \mathbb{R} \times \mathbb{R}$ and $(w^0, w, a) \in [0, +\infty) \times C \times A$,

$$H(P, p_0, \pi, w^0, w, a) := p_0 w^0 + \langle P, f(\Pi(P), a) w^0 + \sum_{i=1}^{m} g_i(\Pi(P)) w^i \rangle + \pi |w|,$$

where $\langle \cdot, \cdot \rangle$ denotes the duality product.

Furthermore, we set

$$H(P, p_0, \pi) := \sup_{(w^0, w, a) \in W(C) \times A} H(P, p_0, \pi, w^0, w, a).$$

and call the so defined function $H : T^*\Omega \times \mathbb{R}^2 \to \mathbb{R}$ the (maximized) Hamiltonian.

The optimal control problem, both in its original version and in the extended form, makes perfectly sense when the state ranges on a Riemannian manifold. Of course, in the notion of local minimizer one has to use the Riemannian metric instead of the Euclidean norm.

Finally, the Maximum Principle on differential manifolds reads the same as in the Euclidean case (Theorem 5.1), except for the part concerning the adjoint equation, which is no longer meaningful and is now replaced by Hamilton equation on the cotangent bundle:
Theorem 6.1 (Maximum Principle on manifolds). Let \((\bar{S}, \bar{y}^0, \bar{\beta}, \bar{w}^0, \bar{\omega}, \bar{\alpha})\) be a canonical local minimizer for the extended problem \(\mathcal{P}^{\text{st}}\). For every Boltyanskii approximating cone \(\Gamma\) of the target \(\mathcal{T}\) at \((\bar{y}^0, \bar{\gamma})(\bar{S})\), there exist a path \(P \in AC([0, S]; T^*\Omega)\) and multipliers \((p_0, \pi, \lambda) \in \mathbb{R} \times (-\infty, 0] \times [0, +\infty)\) such that, mutatis mutandis and according to the above definition, \(\bar{y} = \Pi \circ P\) and conditions (i),(iii), and (iv) of Theorem 5.1 hold true. Furthermore, one has:

(ii)\(^\dagger\) (HAMILTON EQUATION) The path \(P\) solves Hamiltonian system

\[
\frac{d(P(s))}{ds} = X_H(P(s), p_0, \pi)
\]

at almost every \(s \in [0, S]\), where \(X_H\) denotes the Hamiltonian vector field\(^\dagger\) corresponding to \(H\), the latter being regarded as a function (parameterized by \((p_0, \pi)\)) on \(T^*\Omega\).

6.2. The case of brackets of nonsmooth vector fields. There are several aspects of our result that suggest for a possible generalization to non-smooth data.

The first and more obvious one concerns even the first order Maximum Principle, namely the part of Theorems 3.1 and 6.1 that do not involve Lie brackets, and consists in

\[\text{Maximum Principle (applied to the reparameterized problem , see e.g. [9, 30, 29]) surely holds true for an impulsive trajectory. Indeed, such a generalization reduces to the application of a non-smooth maximum principle to the reparameterized problem, which happens to be a standard optimal control problem. In particular the adjoint equation should be replaced by a differential inclusion.}

However, another kind of non-smoothness might affect a problem where the vector fields are not sufficiently smooth for a given bracket to be continuous. For instance, if the vector fields \(g_1, \ldots, g_m\) are just locally Lipschitz continuous the bracket \([g_1, g_2]\) is a bounded measurable function defined almost everywhere.

Besides the afore-mentioned issue of using an adjoint differential inclusion in place of an adjoint equation, one might consider the chance of using the multi-valued Lie brackets of locally Lipschitz vector fields defined in [22, 23] and [11]. For instance, we wonder whether a relation like \(p(s) \cdot [g_1, g_2](\dot{y}(s)) = 0\) keeps on holding even in the Lipschitz case in some form, perhaps in the form

\[p(s) \cdot b = 0 \quad \text{for all } b \in [g_1, g_2]_{\text{set}}(\dot{y}(s)),\]

where \([g_1, g_2]_{\text{set}}\) is the set-valued bracket of the Lipschitz vector fields \(g_1, g_2\), defined by

\([g_1, g_2]_{\text{set}}(y) = \co \left\{ v = \lim_{y_n \to y} Dg_2(y_n) \cdot g_1(y_n) - Dg_1(y_n) \cdot g_2(y_n) \right\},\]

where we have used \(\co E\) to denote the convex hull of \(E\). We do not have an answer at the moment.

6.3. Unbounded control problems. Under some additional hypotheses –like e.g. the normality of the examined extremal, the \(C^\infty\)-smoothness of the vector fields, and the identity \(C = \mathbb{R}^m\) – in [8] the authors investigate higher-order sufficient conditions for the optimal time problem associated to the original unbounded system [2]. Part of the paper is devoted to higher-order necessary conditions, and it is noticeable that a certain number of the latter formally coincide with our conditions. This deserves some interest, for the time-optimal unbounded control problem is clearly different from its impulsive extension: the minimum must be a strict sense trajectory, whose performance, in turn, has to be compared only with strict sense trajectories. In particular, it is not obvious that a necessary condition for the impulsive problem may be applied to a non-impulsive minimum. On the contrary, the coincidence between some of our conditions and those proved in [8] seems to suggest that, for some reason a (possibly higher-order) Maximum Principle valid for the

\[^\dagger\] As is well known, the Hamiltonian vector field \(X_H\) is defined on \(T^*\Omega\) by the relation \(dH(Y) = \omega(X_H, Y)\), where \(Y\) is any vector field on \(T^*\Omega\) and \(\omega\) is the simplectic form.
impulsive system might be a necessary condition for the strict sense unbounded system as well. The advantage of such a criterium would prove useful in easing the search of new necessary conditions, for instance the ones proved here under weakened hypotheses. Our guess is that such automatic applicability of the conditions for the impulsive problem to the original problem holds actually true. The reason why we formulate this conjecture lies on the observation that the set of strict sense controls turns out to be abundant in the set of extended, impulsive controls, according to the definition of abundantness introduced by Jack Warga in the 70s [31]. Referring the reader to [31] and [14] for the definition of abundant set, we just point out that abundantness is stronger than density and seems suitable for the construction of variational cones which approximate both the original reachable set and the extended one.

ACKNOWLEDGMENTS

This research is partially supported by the INdAM-GNAMPA Project 2017 “Optimal impulsive control: higher-order necessary conditions and gap phenomena”; by the Padova University grant PRAT 2015 “Control of dynamics with reactive constraints”; by the European Union under the 7th Framework Programme FP7-PEOPLE-2010-ITN Grant agreement number 264735-SADCO; by FAPERJ (Brazil) through the “Jovem Cientista do Nosso Estado” Program; by CNPq and CAPES (Brazil) and by the Alexander von Humboldt Foundation (Germany).

REFERENCES

1. M.S. Aronna and F. Rampazzo, \(L^1\) limit solutions for control systems, J. Differential Equations 258 (2015), no. 3, 954–979.
2. A.V. Arutyunov, V.A. Dykhta, and F.L. Pereira, Necessary conditions for impulsive nonlinear optimal control problems without a priori normality assumptions, J. Optim. Theory Appl. 124 (2005), no. 1, 55–77.
3. A.V. Arutyunov, D.Yu. Karamzin, F.L. Pereira, and N.Yu. Chernikova, Second-order necessary optimality conditions in optimal impulsive control problems, Differential Equations 54 (2018), no. 8, 1083–1101.
4. A. Bressan and B. Piccoli, Introduction to the mathematical theory of control, AIMS Series on Applied Mathematics, vol. 2, American Institute of Mathematical Sciences (AIMS), Springfield, MO, 2007.
5. A. Bressan and F. Rampazzo, On differential systems with vector-valued impulsive controls, Boll. Un. Mat. Ital. B (7) 2 (1988), no. 3, 641–656.
6. A. Bressan and F. Rampazzo, Moving constraints as stabilizing controls in classical mechanics, Archive for rational mechanics and analysis 196 (2010), no. 1, 97–141.
7. A.J. Caltiá, D.G. Schaeffer, T.P. Witelski, E.E. Monson, and A.L. Lin, On spiking models for synaptic activity and impulsive differential equations, SIAM Rev. 50 (2008), no. 3, 553–569.
8. F. Chittaro and G. Stefani, Minimum-time strong optimality of a singular arc: the multi-input non involutive case, ESAIM Control Optim. Calc. Var. 22 (2016), no. 3, 786–810.
9. F. Clarke, Functional analysis, calculus of variations and optimal control, vol. 264, Springer, 2013.
10. V.A. Dykhta, The variational maximum principle and quadratic conditions for the optimality of impulse and singular processes, Sibirsik. Mat. Zh. 35 (1994), no. 1, 70–82, ii.
11. E. Feleqi and F. Rampazzo, Iterated lie brackets for nonsmooth vector fields, Nonlinear Differential Equations and Applications NoDEA 24 (2017), no. 61, 1–43.
12. P. Gajardo, H. Ramirez C., and A. Rapaport, Minimal time sequential batch reactors with bounded and impulsive controls for one or more species, SIAM J. Control Optim. 47 (2008), no. 6, 2827–2856.
13. O. Hájek, Book review, Bull. Amer. Math. Soc. 12 (1985), no. 2, 272–279.
14. B. Kaskosz, Extremality, controllability, and abundant subsets of generalized control systems, Journal of Optimization Theory and Applications 101 (1999), no. 1, 73–108.
15. C. Marle, Géométrie des systèmes mécaniques liaisons actives, Symplectic Geometry and Mathematical Physics (1991), 260–287.
16. B.M. Miller and E.Y. Rubinovich, Impulsive control in continuous and discrete-continuous systems, Kluwer Academic/Plenum Publishers, New York, 2003.
17. M. Motta and F. Rampazzo, Space-time trajectories of nonlinear systems driven by ordinary and impulsive controls, Differential Integral Equations 8 (1995), no. 2, 269–288.
18. M. Motta, F. Rampazzo, and R. Vinter, Normality and gap phenomena in optimal unbounded control, ESAIM: COCV (2019).
19. M. Motta and C. Sartori, On $L^1$ limit solutions in impulsive control, Discrete Contin. Dyn. Syst. Ser. S 11 (2018), no. 6, 1201–1218.
20. F.L. Pereira and G.N. Silva, Necessary conditions of optimality for vector-valued impulsive control problems, Systems & Control Letters 40 (2000).
21. F. Rampazzo, On the riemannian structure of a lagrangian system and the problem of adding time-dependent constraints as controls, Eur. J. Mech. A Solids 10 (1991).
22. F. Rampazzo and H.J. Sussmann, Set-valued differentials and a nonsmooth version of chow’s theorem, Decision and Control, 2001. Proceedings of the 40th IEEE Conference on, vol. 3, IEEE, 2001, pp. 2613–2618.
23. , Commutators of flow maps of nonsmooth vector fields, J. Differential Equations 232 (2007), no. 1, 134–175.
24. R.W. Rishel, An extended Pontryagin principle for control systems whose control laws contain measures, J. Soc. Indust. Appl. Math. Ser. A Control 3 (1965), 191–205.
25. H. Schättler and U. Ledzewicz, Geometric optimal control: theory, methods and examples, vol. 38, Springer Science & Business Media, 2012.
26. G.N. Silva and R. Vinter, Measure driven differential inclusions, J. Math. Anal. Appl. 202 (1996), no. 3, 727–746.
27. , Necessary conditions for optimal impulsive control problems, SIAM J. Control Optim. 35 (1997), no. 6, 1829–1846.
28. E.D. Sontag, Mathematical control theory: deterministic finite dimensional systems, vol. 6, Springer Science & Business Media, 1998.
29. H.J. Sussmann, Geometry and optimal control, Mathematical control theory, Springer, 1999, pp. 140–198.
30. R. Vinter, Optimal control, Systems & Control: Foundations & Applications, Birkhäuser Boston, Inc., Boston, MA, 2000.
31. J. Warga, Optimal control of differential and functional equations, Academic Press, New York-London, 1972. MR 0372708

M.S. Aronna, Escola de Matemática Aplicada, Fundação Getúlio Vargas, Rio de Janeiro 22250-900, Brazil
E-mail address: soledad.aronna@fgv.br

M. Motta, Dipartimento di Matematica, Università di Padova, Padova 35121, Italy
E-mail address: motta@math.unipd.it

F. Rampazzo, Dipartimento di Matematica, Università di Padova, Padova 35121, Italy
E-mail address: rampazzo@math.unipd.it