Commun. Optim. Theory 2026 (2026) 40 https://doi.org/10.23952/cot.2026.40

MATHRES

Communications in Optimization Theory

Available online at http://cot.mathres.org

HIGHER ORDER NECESSARY CONDITIONS FOR OPTIMAL CONTROLS NOT RANGING IN THE INTERIOR

FRANCESCA ANGRISANI[1], FRANCO RAMPAZZO[2,*]

[1]Sorbonne Université, Jacques Louis Lions Laboratory (LJLL), Paris, France
[2]Department of Mathematics T. Levi-Civita, University of Padova, Italy

Abstract. Goh and Legendre-Clebsch necessary conditions for optimal controls of affine-control systems are usually established under the hypothesis that the minimizing control lies in the interior of the control set U. In this paper, we investigate the possibility of establishing Goh and Legendre-Clebsch necessary conditions without this assumption, so that even control sets with empty interiors or optimal controls touching the boundary of U can be taken into consideration.

Keywords. Control systems; Higher order necessary conditions; Optimal control; Singularity.

2020 Mathematics Subject Classification. 49J52, 49N25.

1. Introduction

Besides the classical, first order, Maximum Principle, higher order necessary conditions are classically considered for optimal controls of ordinary differential equations. In particular, when the differential system is control-affine and the minimum problem is of the form

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

where the vector fields f, g_i and the cost Ψ are sufficiently regular, the target $\mathfrak{T}$ is a subset of the state space, and the controls take values in a subset $U \subset \mathbb{R}^m$, $m \geq 1$— Goh conditions and Legendre-Clebsch conditions for an optimal control-trajectory pair $(\bar{u}, \bar{x})$ read

$$\begin{aligned}
&(a) \quad p(t) \cdot [g_i, g_j](\bar{x}(t)) = 0 \quad \forall 1 \leq i < j \leq m \qquad \text{and} \\
&(b) \quad p(t) \cdot [f, g_i](\bar{x}(t)) = 0 \quad \forall 1 \leq i \leq m \qquad \forall t \in [0,T],
\end{aligned} \tag{1.1}$$

*Corresponding author.

E-mail address: francesca.angrisani@sorbonne-universite.fr (F. Angrisani), rampazzo@math.unipd.it (F. Rampazzo).

Received May 11, 2025; Accepted August 27, 2025.

respectively, where $p(\cdot)$ is the adjoint map whose existence is postulated in the first order Maximum Principle. Furthermore, in the special case when $m = 1$ (and $g := g_1$), they are complemented by the third order Legendre-Clebsch condition

$$0 \geq p(t) \cdot [g,[f,g]](\overline{x}(t)) \qquad \forall t \in [0,T]. \tag{1.2}$$

Let us remark that the crucial hypothesis (S) below is made on the optimal control $\overline{u}$ in order to obtain the higher order conditions (1.1) and (1.2):

$$\text{(S)} \qquad \overline{u}(t) \in int(U) \textit{ for almost every } t \in [0,T],$$

$int(U)$ denoting the interior of U. An optimal control verifying hypothesis (S) is often called *singular*, and the reason why one assumes the singularity hypothesis (S) consists in the following fact: on a left neighbourhood of almost any time t, the primitive of a singular control can be perturbed by means of suitable infinitesimal continuous paths having the effect of producing infinitesimal variations of the corresponding trajectory $\overline{x}$ in the directions of the Lie brackets involved in (1.1) and (1.2).

The present paper is an attempt to obtain higher order optimality conditions of type (1.1) and (1.2) in situations where the singularity assumption (S) is *not* verified. For this purpose, for any given control $u : [0,T] \to U$, any time $t \in]0,T[$, and any $i \in \{1,\dots,m\}$, we introduce the notion of *"u is i-balanced at t"*: this means that there must exist two positive numbers α^i, β^i such that $\{u(s) + \alpha^i \mathbf{e}_i, u(s) - \beta^i \mathbf{e}_i,\} \subset U$, for any s in a left neighbourhood of t. In the case of Goh conditions (1.1)-(a), for a given pair (i, j), $0 \leq i < j \leq m$, we replace (S) with weaker hypothesis stating that the control $\overline{u}$ has to be both i-balanced and j-balanced at almost every $t \in [0,1]$: namely, there must exist four positive numbers $\alpha^i, \alpha^j, \beta^i, \beta^j$ such that $\{u(s) + \alpha^i \mathbf{e}_i, u(s) + \alpha^j \mathbf{e}_j, u(s) - \beta^i \mathbf{e}_i, u(s) - \beta^j \mathbf{e}_j\} \subset U$ for any s in a left neighbourhood of t. Similarly, in the case of Legendre-Clebsch conditions (1.1)-(b) and (1.2), for a given i we assume that $\overline{u}$ has to be *i-balanced*. Notice that under our weakened hypotheses Goh conditions (1.1) [resp. Legendre-Clebsch conditions (1.2)] might be valid only for certain pairs (i, j) [resp. for certain i], while failing for the other pair of indexes [resp. the other indexes].

The paper is organized as follows. In Section 2, we state the problem and present the main result (Theorem 2.4). Moreover, we give some instances of controls that are (i, j)-balanced or (i)-balanced but are not in the interior of U. Additionally, in a toy example (see Example (2.1)) our higher order conditions allow to establish, in a case where the control set has empty interior, that a certain control map verifying the first order, Pontryagin Maximum Principle is in fact not optimal. In Section 3 we construct variations of a process $(\overline{u},\overline{x})$ at a time t in the x-direction of specific Lie brackets in the case when (at t) the control $\overline{u}$ is either i-balanced or both i-balanced and j-balanced, for some $i, j \in \{1,\dots,m\}$. Section 4, which deals with approximation of trajectories through products of exponential maps corresponding to Lie brackets, is not at all original and is put there with the only purpose of adding self-contained character to the paper. Section 5 is devoted to the proof of Theorem 3.7, which concerns infinitesimal variations. In Section 6, we exploit set-separation issues to deduce the proof of Theorem 2.4. A short remark concerning a possible generalization of the presented result to the case of less smooth control systems concludes the paper.

1.1. **Remarks on the notation. Duality.** Sometimes we use the standard identification between $\mathbb{R}^n$ and its dual $(\mathbb{R}^n)^*$.

Lie brackets. The Lie bracket $[X,Y]$ (on a differential manifold) of two vector fields X,Y at a point x, is defined (in any local system of coordinates) as $x \mapsto [X,Y](x) := DY(x)\cdot X(x) - DX(x)\cdot Y(x)$.

Differential equations. If we have uniqueness of the solution $x(\cdot)$ on some interval $[0,T]$ to a Cauchy problem $\dot{x}(t) = F(t,x(t)), x(0) = \tilde{x}$, we will use the exponential notation $e^{\int_0^t F}(\tilde{x}) := x(t)$. In the particular case of an autonomous vector field $F = F(x)$, we will use the notation $e^{tF}(\tilde{x}) := x(t)$.

Areas. In order to recall the geometrical meaning in some coefficients appearing in exponential maps of Lie brackets (see Sect. 3), for any plane, closed, curve $(C^1,C^2) : [a,b] \to \mathbb{R}^2$ of class $W^{1,2}$, we define —according to Green's Theorem— the signed area $Area(C^1,C^2)$ of the region *encircled* by (C^1,C^2) as the quantity

$$Area(C^1,C^2) := \left\langle C^1, \frac{dC_2}{ds} \right\rangle_{L^1} = \int_a^b C^1(s)\frac{dC_2}{ds}(s)ds = \left(-\int_a^b C^2(s)\frac{dC_1}{ds}(s)ds \right). \tag{1.3}$$
[1]

2. The Main Results

We will be concerned with the optimal control problem

$$(P) \qquad \begin{cases} \min \Psi(x(T)), \\ \begin{cases} \dfrac{dx}{dt} = f(x(t)) + \displaystyle\sum_{i=1}^m g_i(x(t))u^i(t), \quad a.e.\ t \in [0,T], \\ x(0) = \hat{x}, \qquad x(T) \in \mathfrak{T}, \end{cases} \end{cases}$$

where, for some integers n,m,

- the state x takes values in $\mathbb{R}^n$;
- the vector fields $f, g_1, \dots, g_m$ as well as the *cost* Ψ are assumed to be of class C^1;
- the control u takes values in a subset $U \subseteq \mathbb{R}^m$ (possibly with empty interior);
- $\mathfrak{T}$ is a subset of $\mathbb{R}^n$, called the *target*;
- by *process* we mean a pair (u,x) such that $u \in L^1([0,T],U)$ and $x \in W^{1,1}([0,T],\mathbb{R}^n)$ is the corresponding (Carathéodory) solution of the above Cauchy problem;
- a process (u,x) is called *feasible* as soon as $x(T) \in \mathfrak{T}$;
- the minimization is performed over the set of *feasible processes* (u,x).

Definition 2.1. A feasible process $(\hat{u},\hat{x})$ is called a local weak minimizer of problem (P) if there exists an L^1 neighbourhood $\mathscr{U}$ of $\hat{u}$ such that $\Psi(\hat{x}(T)) \le \Psi(x(T))$ for all $x(\cdot) \in W^{1,1}([0,T],\mathbb{R}^n)$ such that (x,u) is a feasible process with $u \in \mathscr{U}$. A feasible process $(\hat{u},\hat{x})$ is called a local strong minimizer of problem (P) if there exists a C^0 neighbourhood $\mathscr{V}_{\mathscr{C}}$ of $\hat{x}$ such that $\Psi(\hat{x}(T)) \le \Psi(x(T))$ for all feasible processes (x,u) such that $x \in \mathscr{V}_{\mathscr{C}}$.

[1] The *Area* operator is invariant for positive reparameterization of the curve. Namely, if $\phi : [\alpha,\beta] \to [a,b]$ is a $W^{1,1}$ increasing map such that $\phi(\alpha) = a, \phi(\beta) = b$, and we set $(\hat{C}_1,\hat{C}_2) := (C_1,C_2)\circ\phi$, then $((\hat{C}_1,\hat{C}_2)$ is a closed curve, with $(\hat{C}_1,\hat{C}_2)([a,b]) = (C_1,C_2)([\alpha,\beta])$, and) one has $Area(\hat{C}_1,\hat{C}_2) = Area(C_1,C_2)$.

Clearly any local strong minimizer is also a local weak minimizer.

In order to state Lie brackets-including higher order necessary conditions for minima, we need to recall the concept of Boltyanski approximating cone and to introduce the notion of i-balanced control u at a time t. Let $\mathbf{e}_0,\dots,\mathbf{e}_m$ be the canonical basis of $\mathbb{R}\times\mathbb{R}^m$.

Definition 2.2. ([3, 10]) Let Z be a subset of $\mathbb{R}^N$ for some integer $N\geq 1$, and fix $z\in Z$. We say that a convex cone[2] $K\subseteq\mathbb{R}^N$ is a *Boltyanski approximating cone* for Z at z if there exist a convex cone $C\subset\mathbb{R}^M$ 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

- $F(0)=z$,
- there exists a linear map $L:\mathbb{R}^M\to\mathbb{R}^N$ verifying
$$F(v)=F(0)+Lv+o(|v|)\quad \text{for all } v\in V\cap C,$$
- $LC=K$.

Definition 2.3. If $i\in\{1,\dots m\}$, $t\in]0,T[$, a control $u:[0,T]\to U$ is called ***i*-balanced at** t if there exists $\delta\in]0,\min\{t,T-t\}[$ and $\alpha^i,\beta^i>0$ such that $\{u(s)+\alpha^i\mathbf{e}_i,u(s)-\beta^i\mathbf{e}_i\}\subset U$ for a.e. $s\in]t-\delta,t[$. Furthermore, a control $u:[0,T]\to U$ is called ***i*-balanced a.e.** if there exists a full-measure subset $\Lambda_u^i\subseteq]0,T[$ such u is i-balanced at all $t\in\Lambda_u^i$.

We are now ready to state the main result of the paper, providing, in particular, higher order conditions for a minimum without the usual *singularity* assumption.

Theorem 2.4 (A higher order maximum principle)**.** *Let $(\overline{u},\overline{x})$ be a local weak minimizer of problem* (P)*, and set $H(x,p,u):=p\cdot\Big(f(x)+\sum_{i=1}^m g_i(x)u^i\Big)$. Furthermore, let C be the Boltyanski approximating cone to the target $\mathfrak{T}$ at $\overline{x}(T)$. Then there exist multipliers $(p,\lambda)\in AC\Big([0,T],(\mathbb{R}^n)^*\Big)\times\mathbb{R}^*$, with $\lambda\geq 0$[3] such that the following properties are satisfied:*

i) (NON-TRIVIALITY) $(p,\lambda)\neq(0,0)$;

ii) (ADJOINT EQUATION)
$$\frac{dp}{dt}=-\frac{\partial H}{\partial x}(\overline{x}(t),p(t),\overline{u}(t))\qquad \textit{for a.e.}\quad t\in[0,T]\,;$$

iii) (NON-TRANSVERSALITY)
$$p(T)\in-\lambda\frac{\partial\Psi}{\partial x}(\overline{x}(T))-C^{\perp}\,;$$

iv) (MAXIMUM PRINCIPLE)
$$\max_{u\in U}H(x(t),p(t),u)=H(x(t),p(t),\overline{u}(t))\qquad \textit{for a.e.}\quad t\in[0,T],\ \forall u\in U\,.$$

Moreover, (p,λ) can be chosen so that the following further three properties are verified:

v) (i,j)-(GOH CONDITION) *If $i,j\in\{1,\dots m\},i<j$, and the control $\overline{u}$ is both i-balanced and j-balanced a.e., then*
$$0=p(t)\cdot[g_i,g_j](\overline{x}(t))\qquad\forall t\in[0,T]\,;\tag{2.1}$$

[2]A subset D of a real linear space is called a cone if $\alpha v\in C,\forall\alpha\in[0,+\infty[$ and $\forall v\in D$.

[3]The relation $\lambda\geq 0$ is meaningful thanks to the already mentioned identification between $\mathbb{R}$ and its dual $\mathbb{R}^*$.

vi) i-(LEGENDRE–CLEBSCH CONDITION OF STEP 2) *If $i \in \{1,\dots m\}$ and the control $\overline{u}$ is i-balanced a.e., then*

$$0 = p(t)\cdot[f,g_i](\overline{x}(t)) \qquad \forall t \in [0,T]\,;^4 \tag{2.2}$$

vii) (LEGENDRE–CLEBSCH CONDITION OF STEP 3 WITH $m = 1$) *If $m = 1$, $f, g := g_1$ are of class C^2 around $\overline{x}([0,T])$, and the control $\overline{u}$ is* 1*-balanced a.e., then*

$$0 \geq p(t)\cdot[g,[f,g]](\overline{x}(t)) \qquad \forall t \in [0,T]\,. \tag{2.3}$$

Remark 2.5. If a control u is *singular*, i.e., it takes values in the interior of U, then it is i-balanced a.e. for all $i \in \{1,\dots m\}$. Because of this, Goh and Legendre-Clebsch conditions in their classical form are a particular case of Theorem 2.4.

Example 2.6. If $U := \mathbb{N}^3$, then any control map $u = (u^1,u^2,u^3) : [0,T] \to U$ verifing $u^i(t) > 0$ for every $i = 1,2,3$ and almost every $t \in [0,T]$ is i-balanced for every $i = 1,2,3$. Therefore, if the control u is optimal it satisfies (the usual maximum principle and) the i-Legendre-Clebsch condition for all $i = 1,2,3$ as well as the $(1,2)$-,the $(1,3)$- and the $(2,3)$-Goh conditions.

Instead, an optimal control $\hat{u} = (\hat{u}^1,\hat{u}^2,\hat{u}^3) : [0,T] \to U$ such that $\hat{u}^1(t) \equiv 0$ while $\hat{u}^2(t) > 0$ and $\hat{u}^3(t) > 0$ for almost every $t \in [0,T]$, satisfies i-Legendre-Clebsch condition for $i = 2,3$, and the $(2,3)$-Goh condition.

2.1. **An worked out example.** Let us consider the optimal control problem

$$\begin{cases} \qquad\qquad \min \Psi(x(1)), \\ \dfrac{dx}{dt} = g_1(x(t))u^1(t) + g_2(x(t))u^2(t), \quad a.e.\ t \in [0,1] \\ x(0) = \begin{pmatrix} 2 \\ 0 \\ 0 \end{pmatrix} \qquad\qquad x(1) \in \mathfrak{T} \end{cases}$$

where, for every $x \in \mathbb{R}^3$,

$$\Psi(x) := x^3 \qquad g_1(x) := \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \quad g_2(x) := \begin{pmatrix} 0 \\ 1 \\ x^2 \end{pmatrix}, \qquad \text{and } \mathfrak{T} := \{0\}\times\{0\}\times\mathbb{R}.$$

The control set $U \subset \mathbb{Z}^2$ is defined as

$$U := \Big\{-4,-2,0,3\Big\} \times \Big\{-6,0,4,7\Big\}.$$

We wish to establish whether the constant control map

$$\hat{u}(t) = \begin{pmatrix} \hat{u}^1(t) \\ \\ \hat{u}^2(t) \end{pmatrix} := \begin{pmatrix} -2 \\ \\ 0 \end{pmatrix} \qquad \forall t \in [0,1]$$

[4] Because of the antisimmetry of the Lie bracket, (2.1) and (2.2) are equivalent to $0 = p(t)\cdot[g_j,g_i](\overline{x}(t))$ and $0 = p(t)\cdot[g_i,f](\overline{x}(t))$, respectively.

is (feasible and) allowed to be optimal. The trajectory $\hat{x} : [0,1] \to \mathbb{R}^3$ corresponding to the control $\hat{u}$ is given by

$$\hat{x}(t) = \begin{pmatrix} \hat{x}^1(t) \\ \hat{x}^2(t) \\ \hat{x}^3(t) \end{pmatrix} = \begin{pmatrix} 2-2t \\ 0 \\ -2t \end{pmatrix} \qquad t \in [0,1],$$

which, in particular, says that $\hat{u}$ is feasible, in that $\hat{x}(1) = \begin{pmatrix} 0 \\ 0 \\ -2 \end{pmatrix} \in \mathfrak{T}$. Let us check that the process $(\hat{u},\hat{x})$ verifies the first order Pontryagin Maximum Principle, namely there exists a pair $(p,\lambda) \in AC\Big([0,T],(\mathbb{R}^n)^*\Big) \times \mathbb{R}^*$ verifying **i**) $-$ **iv**) in Theorem 2.4. Indeed, since the vector field g_1 is constant and $\hat{u}^2 \equiv 0$, the adjoint equation reduces to

$$\frac{dp}{dt}(t) = -p(t)\cdot\frac{\partial}{\partial x}\Big(g_1(x)\hat{u}^1(t)+g_2(x)\hat{u}^2(t)\Big)_{x=\hat{x}(t)} = (0,0,0) \quad \forall t \in [0,1]. \tag{2.4}$$

On the other hand, the non-transversality condition reads

$$p(1) \in \big(\mathbb{R}^2 \times \{0\}\big) - \lambda\,(\{(0,0)\}\times\mathbb{R}) \qquad \text{for some } \lambda \geq 0.$$

Therefore, there exist constants $\alpha,\beta \in \mathbb{R}$ and $\lambda \geq 0$ such that $p(t) = (\alpha,\beta,-\lambda)$ for all $t \in [0,1]$. Let us see that, if the maximization **iv)** in Theorem 2.4 holds true, then λ cannot vanish. Indeed, this maximization implies that, for all $\begin{pmatrix} u^1 \\ u^2 \end{pmatrix} \in U$, one has

$$\big(\alpha,\beta,-\lambda\big)\cdot\Big(g_1(\hat{x}(t))(-2-u^1)+g_2(\hat{x}(t))(0-u^2)\Big) \geq 0 \qquad \forall t \in [0,1]. \tag{2.5}$$

Taking $(u^1,u^2) = (-2,-6)$ and $(u^1,u^2) = (-2,4)$ in (2.5), we obtain, for every $t \in [0,1]$, $6\big(\alpha,\beta,-\lambda\big)\cdot g_2(\hat{x}(t)) \geq 0$ and $-4\big(\alpha,\beta,-\lambda\big)\cdot g_2(\hat{x}(t)) \geq 0$, respectively, which implies

$$0 = \big(\alpha,\beta,-\lambda\big)\cdot g_2(\hat{x}(t)) = \big(\alpha,\beta,-\lambda\big)\cdot\begin{pmatrix} 0 \\ 1 \\ x_2(t) \end{pmatrix} = \beta - \lambda x_2(t),$$

for all $t \in [0,1]$. Since $x_2(t) = 0$ for all $t \in [0,1]$, it follows that $\beta = 0$.

On the other hand, taking $(u^1,u^2) = (-4,0)$ and $(u^1,u^2) = (0,0)$ in (2.5) we get, for every $t \in [0,1]$, $2\big(p_1(t),p_2(t),p_3(t)\big)\cdot g_1(\hat{x}(t)) \geq 0$ and $-2\big(p_1(t),p_2(t),p_3(t)\big)\cdot g_1(\hat{x}(t)) \geq 0$, respectively, which in turn implies

$$0 = \big(\alpha,\beta,-\lambda\big)\cdot g_1(\hat{x}(t)) = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} = \alpha - \lambda,$$

for all $t \in [0,1]$. Hence one deduces that $\alpha = \lambda$. Therefore $p(t) = (\lambda,\ 0,\ -\lambda)$ for every $t \in [0,1]$. Hence the non-triviality condition **i)** yields $p(t) = (\lambda,0,-\lambda)$ with $\lambda > 0$. This implies that the higher order principle stated in Theorem 2.4 *is not satisfied*, in that the control $\hat{u}$ is both 1-balanced and 2-balanced a.e. while

$$p(t)\cdot[g_1,g_2](\hat{x}(t)) = (\lambda,\beta,-\lambda)\cdot\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = -\lambda \neq 0,$$

namely the $(1,2)$-Goh condition **v)** is violated. Therefore $\hat{u}(\cdot)$ is not optimal, a conclusion that would not be deducible from the standard (first order) Maximum Principle.

3. Variations

The proof of the main result (Theorem 2.4) will be based on set-separation arguments (see Section 6), which, in particular, require the generation of adequate *variations* of the optimal process. As explained in the Introduction, this is the crucial issue of the present paper, for we are going to produce Lie bracket-based variations even when the classical assumption that the optimal control's values are in the interior of the control set U (which might be even empty) fails to be verified.

3.1. **Variation builders.**

Let us introduce a class of indexes, that we shall call the set of *variation signals*.

Definition 3.1 (Variation signals)**.** Let us define the set $\mathfrak{V}$ of *variation signals* as

$$\mathfrak{V} := \mathfrak{V}_{ndl} \bigcup \mathfrak{V}_{Goh} \bigcup \mathfrak{V}_{LC2} \bigcup \mathfrak{V}_{LC3},$$

where $\mathfrak{V}_{ndl} := U$, $\mathfrak{V}_{Goh} := \{(i,j)\ i<j\ \ j,i=1,\dots,m\}$, $\mathfrak{V}_{LC2} := \{(0,i)\ ,i=1,\dots,m\}$, and $\mathfrak{V}_{LC3} := \{(1,0,1)\}$.

Every variation signal $\mathbf{c} \in \mathfrak{V}$ will represent an associated *variation builder*, in a way we now describe.

Definition 3.2 (Variation builders)**.**

- **i)** If $\mathbf{c} \in \mathfrak{V}_{ndl}$, namely $\mathbf{c} = \bar{u} \in U$ the associated variation builder will be the standard needle variation corresponding to $\bar{u}$.
- **ii)** If $\mathbf{c} \in \mathfrak{V}_{Goh}$, that is $\mathbf{c} = (i,j)$, for some $0 < i < j \le m$, let us choose four positive real numbers α^i, β^i, α^j, β^j and let us define the times $\tau_0, \tau_1, \dots, \tau_4 (= 1/2)$ by setting

$$r := (\alpha^i)^{-1} + (\alpha^j)^{-1} + (\beta^i)^{-1} + (\beta^j)^{-1}$$

$$\tau_0 = 0, \quad \tau_1 := \tau_0 + \frac{(\alpha^i)^{-1}}{2r} \quad \tau_2 := \tau_1 + \frac{(\alpha^j)^{-1}}{2r} \quad \tau_3 := \tau_2 + \frac{(\beta^i)^{-1}}{2r} \quad \tau_4 := \tau_3 + \frac{(\beta^j)^{-1}}{2r} = \frac{1}{2}.$$

The variation builder associated to $\mathbf{c} = (i,j)$ is the (piece-wise constant) map $\gamma_{(i,j)}$ on $[0,1]$

$$\gamma_{(i,j)}(s) := \begin{cases} \hat{\gamma}_{(i,j)}(s) & \forall s \in [0,1/2], \\ -\hat{\gamma}_{(i,j)}(s-1/2) & \forall s \in [1/2,1], \end{cases} \tag{3.1}$$

where $\hat{\gamma}_{(i,j)}$ is defined (on $[0,1/2]$) as follows:

$$\hat{\gamma}_{(i,j)}(s) := \alpha^i \mathbf{1}_{[0,\tau_1]}(s)\mathbf{e}_i + \alpha^j \mathbf{1}_{[\tau_1,\tau_2]}(s)\mathbf{e}_j - \beta^i \mathbf{1}_{[\tau_2,\tau_3]}(s)\mathbf{e}_i - \beta^j \mathbf{1}_{[\tau_3,\frac{1}{2}]}(s)\mathbf{e}_j \quad \forall s \in [0,1/2].$$

More explicitly, if we consider the times

$$\tau_5 := \tau_4 + \frac{(\beta^j)^{-1}}{2r} \quad \tau_6 := \tau_5 + \frac{(\beta^i)^{-1}}{2r} \quad \tau_7 := \tau_6 + \frac{(\alpha^j)^{-1}}{2r} \quad \tau_8 := \tau_7 + \frac{(\alpha^i)^{-1}}{2r} = 1,$$

we get, for every $s\in[0,1]$,

$$\begin{aligned}\gamma_{(i,j)}(s) &:= \alpha^i\mathbf{1}_{[0,\tau_1]}(s)\mathbf{e}_i+\alpha^j\mathbf{1}_{[\tau_1,\tau_2]}(s)\mathbf{e}_j-\beta^i\mathbf{1}_{[\tau_2,\tau_3]}(s)\mathbf{e}_i-\beta^j\mathbf{1}_{[\tau_3,\tau_4]}(s)\mathbf{e}_j\\ &-\beta^i\mathbf{1}_{[\tau_4,\tau_5]}(s)\mathbf{e}_i-\beta^j\mathbf{1}_{[\tau_5,\tau_6]}(s)\mathbf{e}_j+\alpha^i\mathbf{1}_{[\tau_6,\tau_7]}(s)\mathbf{e}_i+\alpha^j\mathbf{1}_{[\tau_7,1]}(s)\mathbf{e}_j.\end{aligned} \tag{3.2}$$

iii) If we consider a variation signal $\mathbf{c}\in\mathfrak{V}_{LC2}$, namely $\mathbf{c}=(0,i)$ for some $i\in\{1,\dots m\}$, after choosing two positive numbers α^i,β^i and setting $\bar\tau:=\dfrac{\alpha^i}{\alpha^i+\beta^i}$, let us associate the variation builder

$$\gamma_{(0,i)}(s):=\gamma^i_{(0,i)}(s)\mathbf{e}_i:=\Big(-\beta^i\mathbf{1}_{[0,\bar\tau]}(s)+\alpha^i\mathbf{1}_{[\bar\tau,1]}(s)\Big)\mathbf{e}_i\qquad\forall s\in[0,1]. \tag{3.3}$$

iv) If $\mathbf{c}\in\mathfrak{V}_{LC3}$, namely $m=1$ and $\mathbf{c}=(1,0,1)$, after choosing two positive numbers α^1,β^1 and setting $\hat\tau_1:=\dfrac{\alpha^1}{2(\alpha^1+\beta^1)}$ $\quad\hat\tau_2:=1-\hat\tau_1=\dfrac{\alpha^1+2\beta^1}{2(\alpha^1+\beta^1)}$, as variation builder let us consider the map

$$\gamma_{(1,0,1)}(s):=\gamma^1_{(1,0,1)}(s)\mathbf{e}_1:=\Big(-\beta^1\mathbf{1}_{[0,\hat\tau_1]}(s)+\alpha^1\mathbf{1}_{[\hat\tau_1,\hat\tau_2]}-\beta^1\mathbf{1}_{[\hat\tau_2,1]}\Big)\mathbf{e}_1\qquad\forall s\in[0,1].$$

3.1.1. *Primitives of the variation generators* $\gamma_{i,j},\gamma_{0,i}$ $\gamma_{1,1,0}$. To prove Theorem 3.7 below, we will consider peculiar properties of the (continuous, piecewise linear) primitives

$$\Gamma_{(i,j)}(s):=\int_0^s\gamma_{(i,j)}(\sigma)d\sigma\qquad\Gamma_{(0,i)}(s):=\int_0^s\gamma_{(0,i)}(\sigma)d\sigma,\quad s\in[0,1],\ 1\le i<j\le m,$$

$$\Gamma_{(1,0,1)}(s):=\int_0^s\gamma_{(1,0,1)}(\sigma)d\sigma,\qquad\forall s\in[0,1].$$

Notice that, for every $s\in[0,\tau_4]=[0,1/2]$,

$$\Gamma_{(i,j)}(s)=\hat\Gamma_{(i,j)}(s):=\int_0^s\hat\gamma_{(i,j)}(\sigma)(\sigma)=\begin{cases}s\alpha^i\mathbf{e}_i & \forall s\in[0,\tau_1]\\ (2r)^{-1}\mathbf{e}_i+(s-\tau_1)\alpha^j\mathbf{e}_j & \forall s\in[\tau_1,\tau_2]\\ (2r)^{-1}\mathbf{e}_i+(2r)^{-1}\mathbf{e}_j-(s-\tau_2)\beta^1\mathbf{e}_i & \forall s\in[\tau_2,\tau_3]\\ (2r)^{-1}\mathbf{e}_j-(s-\tau_3)\beta^j\mathbf{e}_j & \forall s\in[\tau_3,1/2].\end{cases}$$

Therefore, for every $s\in[0,\tau_8]=[0,1]$,

$$\Gamma_{(i,j)}(s)=\begin{cases}\hat\Gamma_{(i,j)}(s) & \forall s\in[0,1/2],\\ \hat\Gamma_{(i,j)}(1/2)-\hat\Gamma_{(i,j)}(s-1/2) & \forall s\in[1/2,1].\end{cases} \tag{3.4}$$

Indeed,

$$\Gamma_{(i,j)}(s)=\int_0^s\left(\gamma_{(i,j)}(\sigma)\mathbf{1}_{[0,1/2]}(\sigma)+\gamma_{(i,j)}(\sigma)\mathbf{1}_{[1/2,1]}(\sigma)\right)d\sigma$$

$$=\int_0^s\Big(\gamma_{(i,j)}(\sigma)\mathbf{1}_{[0,1/2]}(\sigma)-\hat\gamma_{(i,j)}(\sigma-1/2)\mathbf{1}_{[1/2,1]}(\sigma)\Big)d\sigma$$

$$=\begin{cases}\hat\Gamma_{(i,j)}(s) & \forall s\in[0,1/2]\\ \hat\Gamma_{(i,j)}(1/2)-\displaystyle\int_{1/2}^s\hat\gamma_{(i,j)}(\sigma-1/2)d\sigma & \forall s\in[1/2,1]\end{cases}=$$

$$= \begin{cases} \hat{\Gamma}_{(i,j)}(s) & \forall s \in [0,1/2] \\ \hat{\Gamma}_{(i,j)}(1/2) - \displaystyle\int_0^{s-1/2} \hat{\gamma}_{(i,j)}(\xi)d\xi & \forall s \in [1/2,1] \end{cases}$$
$$= \begin{cases} \hat{\Gamma}_{(i,j)}(s) & \forall s \in [0,1/2] \\ \hat{\Gamma}_{(i,j)}(1/2) - \hat{\Gamma}_{(i,j)}(s-1/2) & \forall s \in [1/2,1]\,. \end{cases}$$

More explicitly, one has

$$\Gamma_{(i,j)}(s) := \int_0^s \gamma_{(i,j)}(\sigma)d\sigma = \begin{cases} s\alpha^i \mathbf{e}_i & \forall s \in [0,\tau_1] \\ (2r)^{-1}\mathbf{e}_i + (s-\tau_1)\alpha^j \mathbf{e}_j & \forall s \in [\tau_1,\tau_2] \\ (2r)^{-1}\mathbf{e}_i + (2r)^{-1}\mathbf{e}_j - (s-\tau_2)\beta^i \mathbf{e}_i & \forall s \in [\tau_2,\tau_3] \\ (2r)^{-1}\mathbf{e}_j - (s-\tau_3)\beta^j \mathbf{e}_j & \forall s \in [\tau_3,\tau_4] \\ -(s-\tau_4)\alpha^i \mathbf{e}_i & \forall s \in [\tau_4,\tau_5] \\ -(2r)^{-1}\mathbf{e}_i - (s-\tau_5)\alpha^j \mathbf{e}_j & \forall s \in [\tau_5,\tau_6] \\ -(2r)^{-1}\mathbf{e}_i - (2r)^{-1}\mathbf{e}_j + (s-\tau_6)\beta^i \mathbf{e}_i & \forall s \in [\tau_6,\tau_7] \\ -(2r)^{-1}\mathbf{e}_j + (s-\tau_7)\beta^j \mathbf{e}_j & \forall s \in [\tau_7,1] \end{cases}$$

(observe that $\Gamma_{(i,j)}([0,1]) \subset \mathbf{e}_i\mathbb{R} \times \mathbf{e}_j\mathbb{R}$). Furthermore, for every $i = 1\dots,m$, one gets

$$\Gamma_{(0,i)}(s) := \Gamma^i_{(0,i)}(s)\mathbf{e}_i = \int_0^s \gamma_{(0,i)}(\sigma)d\sigma = \begin{cases} -s\beta^i \mathbf{e}_i & \forall s \in [0,\bar{\tau}] \\ \big(-\bar{\tau}\beta^1 + (s-\bar{\tau})\alpha^i\big)\mathbf{e}_i & \forall s \in [\bar{\tau},1] \end{cases}$$

(observe that $\Gamma_{(0,i)}([0,1]) \subset \mathbf{e}_i\mathbb{R}$). Finally,

$$\Gamma_{(1,0,1)}(s) := \Gamma^1_{(1,0,1)}(s)\mathbf{e}_1 = \int_0^s \gamma_{(1,0,1)}(\sigma)d\sigma$$
$$= \begin{cases} -s\beta^1 \mathbf{e}_1, & \forall s \in [0,\hat{\tau}_1] \\ \Big(-\hat{\tau}_1\beta^1 + (s-\hat{\tau}_1)\alpha^1\Big)\mathbf{e}_1 & \forall s \in [\hat{\tau}_1,\hat{\tau}_2] \\ \Big(-\hat{\tau}_1\beta^1 + (\hat{\tau}_2-\hat{\tau}_1)\alpha^1 - (s-\hat{\tau}_2)\beta^1\Big)\mathbf{e}_1, & \forall s \in [\hat{\tau}_2,1] \end{cases}$$

(observe that $\Gamma_{(1,0,1)}([0,1]) \subset \mathbf{e}_1\mathbb{R}$). For every $h = 0,\dots,m$ let us use $\Gamma^h_{(i,j)}$, $\Gamma^h_{(0,i)}$, and $\Gamma^h_{(1,0,1)}$ to denote the $h-th$ components of $\Gamma_{(i,j)}$ $\Gamma_{(0,i)}$, and $\Gamma_{(1,0,1)}$, respectively. Notice that, in particular, $\Gamma^0_{(i,j)}(s) = \Gamma^0_{(0,i)}(s) = \Gamma^0_{(1,0,1)}(s) = 0$, $\forall s \in [0,1]$.

The following two lemmas will be crucial in the construction of control variations.

Lemma 3.3. *For every $i,j \in \{1,\dots,m\}$, $i<j$, one has*

$$\begin{aligned} &0 = \Gamma_{(i,j)}(0) = \Gamma_{(i,j)}(1) = \Gamma_{(0,i)}(0) = \Gamma_{(0,i)}(1) \\ &0 = \Gamma^1_{(1,0,1)}(0) = \Gamma^1_{(1,0,1)}(1) \qquad \int_0^1 \Gamma_{(1,0,1)}(s)ds = 0, \end{aligned} \tag{3.5}$$

$$\int_0^1 \Gamma_{(i,j)}(s)ds = 0, \tag{3.6}$$

and

$$\int_0^1 \Gamma^h_{(0,i)}(s)ds = \begin{cases} -\dfrac{\alpha^i\beta^i}{2(\alpha^i+\beta^i)} & \text{if } h=i \\ \\ 0 & \text{if } h\in\{1,\cdots\}\backslash\{i\}. \end{cases} \tag{3.7}$$

Proof. Equalities (3.5) are trivial. As for (3.6), by (3.4), one gets

$$\begin{aligned}\int_0^1 \Gamma_{(i,j)}(s)ds &= \int_0^{1/2} \hat\Gamma_{(i,j)}(s)ds - \int_{1/2}^1 \hat\Gamma_{(i,j)}(s-1/2)ds \\ &= \int_0^{1/2} \hat\Gamma_{(i,j)}(s)ds - \int_0^{1/2} \hat\Gamma_{(i,j)}(s)ds = 0.\end{aligned}$$

Equality (3.7) is just the computation of the integral of a piecewise linear continuous map on $[0,1]$. □

Other important quantities in the proof of the main theorem are computed in the following Lemma:

Lemma 3.4. *For every pair (i,j), $1\le i<j\le m$ one has*

$$\int_0^1 \Gamma^i_{(i,j)}(t)\dot\Gamma^j_{(i,j)}(t)dt = Area\Big(\Gamma^i_{(i,j)},\Gamma^j_{(i,j)}\Big) = \frac{1}{2r^2} = -Area\Big(\Gamma^j_{(i,j)},\Gamma^i_{(i,j)}\Big). \tag{3.8}$$

Moreover, for all $h\in\{0,\dots m\}\backslash\{i,j\}$,

$$\int_0^1 \Gamma^i_{(i,j)}(t)\dot\Gamma^h_{(i,j)}(t)dt = Area\Big(\Gamma^i_{(i,j)},\Gamma^h_{(i,j)}\Big) = 0 = Area\Big(\Gamma^j_{(i,j)},\Gamma^h_{(i,j)}\Big) = \int_0^1 \Gamma^j_{(i,j)}(t)\dot\Gamma^h_{(i,j)}(t)dt.^5 \tag{3.9}$$

Proof. Equality (3.8) can obviously be obtained by direct computation. However it is geometrically deductible by the fact that the image is the the union of two equal squares –run in counterclockwise sense– whose edges' lenghts are $1/2r$. So $Area\Big(\Gamma^i_{(i,j)},\Gamma^j_{(i,j)}\Big) = 2\cdot(1/2r)^2 = \dfrac{1}{2r^2}$. Similarly, equalities (3.9) can be trivially deduced by computation, though they also might be deduced by inspection of their running senses. □

3.2. **Control variations.**

Definition 3.5 (Families of control variations)**.** Let us fix a map $u\in L^\infty([0,T],\mathbb{R}^m)$, a Lebesgue point $\bar t\in]0,T[$ for u[6], and a parameter $\varepsilon>0$ such that $\sqrt[3]{\varepsilon}\le\bar t$.

[5]The introduction of the functions $\tilde\Gamma^h_{(0,i)}$, $\tilde\Gamma^0_{(1,0,1)}$, $\tilde\Gamma^1_{(1,0,1)}$ is motivated by the fact that, in order to consider the notion of *Area*, we need *closed* curves.

[6]Let us recall that, if $\varphi:[a,b]\mapsto\mathbb{R}^k$ is a L^1 map, a $t\in]a,b[$ is called a Lebesgue point for φ if

$$\lim_{r\to0+}\frac{1}{2r}\int_{t-r}^{t+r}|\varphi(s)-\varphi(t)|ds = 0.$$

- If $\mathbf{c} = \hat{u} \in \mathfrak{V}_{ndl}(= U)$ we define the family of ε-dependent controls $\left\{u_{\varepsilon,\mathbf{c},\bar{t}},\ \varepsilon \in [0,\bar{t})\right\}$ by setting

$$u_{\varepsilon,\mathbf{c},\bar{t}}(t) := \begin{cases} u(t) & \text{if } t \in [0,\bar{t}-\varepsilon) \cup (\bar{t},T] \\ \hat{u} & \text{if } t \in [\bar{t}-\varepsilon,\bar{t}]. \end{cases} \tag{3.10}$$

This family is usually referred to as a needle variation of u at $\bar{t}$.

- If $\mathbf{c} = (i,j) \in \mathfrak{V}_{Goh}$, $1 \le i < j \le m$, and the control u is both i-balanced and j-balanced at $\bar{t}$, we define the ε-parameterized family $\left\{u_{\varepsilon,\mathbf{c},\bar{t}},\ 0 < \sqrt{\varepsilon} \le \bar{t}\right\}$ of controls by setting

$$u_{\varepsilon,\mathbf{c},\bar{t}}(t) := \begin{cases} u(t) & \text{if } t \notin [\bar{t}-\sqrt{\varepsilon},\bar{t}] \\ u(t) + \gamma_{(i,j)}\left(\dfrac{t-(\bar{t}-\sqrt{\varepsilon})}{\sqrt{\varepsilon}}\right) & \text{if } t \in [\bar{t}-\sqrt{\varepsilon},\bar{t}]. \end{cases} \tag{3.11}$$

- If $\mathbf{c} = (0,i) \in \mathfrak{V}_{LC2}$, $0 < i \le m$, and the control u is i-balanced at $\bar{t}$, we define the ε-parameterized family $\left\{u_{\varepsilon,\mathbf{c},\bar{t}}(t),\ 0 < \sqrt{\varepsilon} \le \bar{t},\right\}$ of controls by setting

$$u_{\varepsilon,\mathbf{c},\bar{t}}(t) := \begin{cases} u(t) & \text{if } t \notin [\bar{t}-\sqrt{\varepsilon},\bar{t}] \\ u(t) + \gamma_{(0,i)}\left(\dfrac{t-(\bar{t}-\sqrt{\varepsilon})}{\sqrt{\varepsilon}}\right) & \text{if } t \in [\bar{t}-\sqrt{\varepsilon},\bar{t}]. \end{cases} \tag{3.12}$$

- If $m=1$, $g := g_1$, $\{\mathbf{c} = (1,0,1)\} = \mathfrak{V}_{LC3}$, and u is 1-balanced at $\bar{t}$, we define the family $\left\{u_{\varepsilon,\mathbf{c},\bar{t}}(t),\ 0 < \sqrt[3]{\varepsilon} \le \bar{t}\right\}$ of controls by setting

$$u_{\varepsilon,\mathbf{c},\bar{t}}(t) := \begin{cases} u(t) & \text{if } t \notin [\bar{t}-\sqrt[3]{\varepsilon},\bar{t}] \\ u(t) + \gamma_{(1,0,1)}\left(\dfrac{t-(\bar{t}-\sqrt[3]{\varepsilon})}{\sqrt[3]{\varepsilon}}\right) & \text{if } t \in [\bar{t}-\sqrt[3]{\varepsilon},\bar{t}] \end{cases} \tag{3.13}$$

Remark 3.6. Let us observe that, for every control u and every variation signal $\mathbf{c} \in \mathfrak{V}$ as in Definition 3.5, the corresponding perturbed control $u_{\varepsilon,\mathbf{c},\bar{t}}$ is admissible for every $\varepsilon > 0$, i.e. it takes values in U.

Parts **2)-4)** of following theorem establish how the above control variations produce variations of the trajectories, in terms of Lie brackets (while part **1)** deals with the standard needle variations.).

Theorem 3.7. *Let (u,x) be a process for the control system in* (1)*, and let $\bar{t} \in]0,T]$ be a Lebesgue point for u. Then, for any $\mathbf{c} \in \mathfrak{V}$ and any $\varepsilon > 0$ sufficiently small, using x_ε to denote the trajectory corresponding to the perturbed control $u_{\varepsilon,\mathbf{c},\bar{t}}$ (see Definition 3.5), the following four statements hold true.*

1) *If $\bar{u} \in \mathfrak{V}_{ndl}$, then*

$$x_\varepsilon(\bar{t}) - x(\bar{t}) = \varepsilon \sum_{\ell=1}^{m} g_\ell(x(\bar{t}))\big(\bar{u}^\ell - u^\ell(\bar{t})\big) + o(\varepsilon)\,. \tag{3.14}$$

2) *If $\mathbf{c} = (i,j) \in \mathfrak{V}_{Goh}$, , for some $1 \le i < j \le m$, and the control u is both i-balanced and j-balanced at $\bar{t}$, then there exists $M > 0$ such that*

$$x_\varepsilon(\bar{t}) - x(\bar{t}) = M\varepsilon[g_i,g_j](x(\bar{t})) + o(\varepsilon). \tag{3.15}$$

3) *If* $\mathbf{c} = (0,i) \in \mathfrak{V}_{LC2}$*, for some* $i = 1,\dots,m$*, and the control* u *is* i-balanced at $\bar{t}$*, then there exists* $\tilde{M} > 0$ *such that*

$$x_\varepsilon(\bar{t}) - x(\bar{t}) = \tilde{M}\varepsilon[f,g_i](x(\bar{t})) + o(\varepsilon)\,. \tag{3.16}$$

4) *If* $m = 1$*,* $\mathbf{c} = \mathfrak{V}_{LC3}$*, the control* u *is* 1*-balanced, and* $f, g(:= g_1)$ *are of class* C^2 *around* $x(\bar{t})$*, then there exists* $\check{M} > 0$ *such that*

$$x_\varepsilon(\bar{t}) - x(\bar{t}) = \check{M}\varepsilon[g,[f,g]](x(\bar{t})) + o(\varepsilon)\,. \tag{3.17}$$

The proof of this theorem is given in Section 5.

The following simple remark is crucial in the proof of Theorem 2.4:

Remark 3.8. If in the previous construction we replace the maps $\Gamma_{(i,j)}$ and $\Gamma_{(0,i)}$, $1 \le i < j \le m$ with the maps $\check{\Gamma}_{(i,j)}$ and $\check{\Gamma}_{(0,i)}$ defined as $\check{\Gamma}_{(i,j)}(s) := \Gamma_{(i,j)}(1-s)$ and $\check{\Gamma}_{(0,i)} := \Gamma_{(0,i)}(1-s)$ for all $s \in [0,1]$, respectively, we do obtain $x_\varepsilon(\bar{t}) - x(\bar{t}) = -M\varepsilon[g_i,g_j](x(\bar{t})) + o(\varepsilon)$ in place of (3.15), and $x_\varepsilon(\bar{t}) - x(\bar{t}) = -\tilde{M}\varepsilon[f,g_i](x(\bar{t})) + o(\varepsilon)$ in place of (3.16).

4. Approximation of Solutions

In this section, we recall some classical facts (see, e.g., [6, 8, 9, 11]) about local approximations of trajectories of control affine systems through product of exponential maps of Lie brackets up to a certain degree. Together with the use of variations builders presented in the previous sections, these kinds of approximation will allow us to prove Theorem 3.7.

4.1. **The case with sub-linear controls.**

Most of the results are of local nature, so we limit our exposition to Euclidean spaces $\mathbb{R}^n$, $n \ge 1$. Let us begin with a trivial remark: if X is a given Lipschitz continuous vector field defined on an open subset $\Omega \subseteq \mathbb{R}^n$ and $b \in L^1([0,T],\mathbb{R})$, if, for some $T > 0$, a (necessarily unique) solution $[0,T] \ni t \mapsto x(t)$ of the Cauchy problem

$$\dot{x}(t) = b(t)X(x(t)) \;\; x(0) = \hat{x} \quad t \in [0,T]$$

exists, then $x(t) = e^{B(t)X}(\hat{x})$ for all $t \in [0,T]$, where we have set $B(t) = \int_0^t b(s)ds$, $\forall t \in [0,T]$. In other words, the value $x(t)$ of the solution at time t coincides with the value $\tilde{x}(\tau)$ of the solution of the autonomous problem $\dot{\tilde{x}}(\tau) = X(\tilde{x}(\tau))$, $\tilde{x}(0) = \hat{x}$ at time $\tau = B(t)$.

An estimate of the error caused by non-commutativity is given by the following elementary result:

Lemma 4.1. *Let* m *be any natural number, let* $\chi > 0$ *be a positive constant and let* $b_i(s) : [0,T] \to \mathbb{R}, \quad i = 0,\dots,m$ *be* $m+1$ *arbitrary* L^1 *functions such that* $|b_i(s)| \le \chi s$*, for almost every* $s \in [0,T]$*. Let us consider* $m+1$ *vector fields* $X_i \in C^1(\Omega;\mathbb{R}^n)$ *defined on an open subset* $\Omega \subseteq \mathbb{R}^n, \quad i = 0,\dots,m$*, and let* $x(\cdot)$ *be a solution on an interval* $[0,T]$ *of* $\dot{x}(t) = \sum_{i=0}^m b^i(t)X_i(x(t))$ *Then, setting, for every* $t \in (0,T)$ *and every* $i = 0,\dots,m$*,* $B_i(t) := \int_0^t b^i(s)$*,, we have* $x(t) = e^{B^0(t)X_0} \circ \dots \circ e^{B^m(t)X_m}(x(0)) + o(t^3)$*,* $t \in [0,T]$*.*

For a proof, we refer the reader to the vast literature on this subject (see, e.g., [2, 5, 6]). We just recall that, in order to prove the above result, it is essential to make use of the following basic fact:

Lemma 4.2. *If* $\Omega \subseteq \mathbb{R}^q$ *(for some* $q \ge 1$*) is an open subset,* $X,Y \in C^n(\Omega;\mathbb{R}^q)$ *for some* $n \ge 1$*, and* $x \in \Omega$*, then* $\frac{d^n}{d\tau^n}\Big(D(e^{-\tau X}) \cdot Y(e^{\tau X}x)\Big) = D(e^{-\tau X}) \cdot ad^n X(Y)(e^{\tau X}x)$*, for all* $n \ge 1$*, for any* *sufficiently small τ, where we have used the "ad" operator defined recursively as $ad^0X(Y) = Y$ $\dots ad^nX(Y) = [X, ad^{n-1}X(Y)]$ for all $n \in \mathbb{N}$. In particular, we get the Taylor expansion $D(e^{-\tau X}) \cdot \big(Y(e^{\tau X}x)\big) = \sum_{i=0}^{n} \frac{\tau^i}{i!} ad^iX(Y)(x) + o(\tau^n)$.*

In Proposition 4.3 below, the hypothesis of sublinearity of the controls is dismissed, and an estimate is given for the flow of $\dot{x} = \sum_{i=0}^{m} a^i(s) \cdot g_i(x(s))$ in terms of the original vector fields $g_0, \dots, g_m$ and their Lie brackets.

Proposition 4.3. *For any $i = 0, \dots, m$, consider a function $a^i(s) \in L^\infty([0,T];\mathbb{R})$ and let g_i be a vector field belonging to $C^2(\Omega;\mathbb{R}^n)$. Then, if we set, for all $i = 0, \dots, m$ and $i < j \le m$,*

$$A^i(s) := \int_0^s a^i(\sigma)d\sigma,\ A^{i,j}(s) := \int_0^s A^i(\sigma)\frac{dA^j}{d\sigma}(\sigma)d\sigma = \int_0^s A^i(\sigma)a^j(\sigma)d\sigma, \qquad s \in [0,T],$$

the solution $x(\cdot)$ (on some interval $[0,\tau]$) to $\dot{x}(t) = \sum_{i=0}^{m} a^i(t)\, g_i(x(t))$ verifies

$$x(t) = e^{A^0(t)g_0} \circ \dots \circ e^{A^m(t)g_m} \circ e^{A^{0,1}(t)[g_0,g_1]} \circ \dots \circ e^{A^{m-1,m}(t)[g_{m-1},g_m]})(x(0)) + o(t^2)$$

(where it is meant that the order in the product with two indexes is the lexicographic one).

As a straightforward consequence of Proposition 4.3, one gets the following result.

Corollary 4.4. *Let us fix $t > 0$ (sufficiently small), let us choose $i^*, j^* \in \{0, \dots, m\}$, $j^* > i^*$ and let $g_{i^*}, g_{j^*} \in C^2(\Omega,\mathbb{R}^n)$. If, for every $i = 0, 1, \dots, m$, the L^∞ controls $a^i : [0,t] \to U$ in Proposition 4.3 are chosen such that $A^i(t) = 0$ and $A^{j^*,i^*}(t) = ct^2 + o(t^2)$ with $c > 0$ is the only non-vanishing term in the family $\{A^{ji}(t)\}_{j>i,i=0,\dots,m}$, then $x(t) - x(0) = ct^2[g_{i^*}, g_{j^*}](x(0)) + o(t^2)$.*

Finally, we provide an estimate for the flow of our differential equation $\dot{x}(t) = \sum_{i=0}^{m} a^i(t)\, g_i(x(t))$ in terms of the Lie brackets up to the length 3:

Proposition 4.5. *Let $a_i(s)$, $\quad i = 0, \dots, m$ be $m+1$ arbitrary functions in $L^\infty([0,T];\mathbb{R})$. Let $t \in (0,T)$ and let g_i, $i = 0, \dots, m$, be $m+1$ vector fields belonging to $C^3(\Omega,\mathbb{R}^n)$. Then, if we set, for all k, i, j with $i = 0, \dots, m$, $k, i < j \le m$,*

$$A^{kij}(s) := \int_0^s A^k(\sigma)\frac{dA^{i,j}}{d\sigma}(\sigma)d\sigma = \int_0^s A^k(\sigma)A^i(\sigma)a^j(\sigma)d\sigma \qquad s \in [0,T],$$

the solution $x(\cdot)$(on some interval $[0,\tau]$) to $\dot{x}(s) = \sum_{i=0}^{m} a^i(s)\, g_i(x(s))$ verifies

$$\begin{aligned} x(t) &= e^{A^0(t)g_0} \circ \dots \circ e^{A^m(t)g_m} \circ \\ &\quad e^{A^{0,1}(t)[g_0,g_1]} \circ \dots \circ e^{A^{m-1,m}(t)[g_{m-1},g_m]} \circ \\ &\quad e^{A^{0,0,1}(t)[g_0,[g_0,g_1]]} \circ e^{A^{0,0,1}(t)[g_0,[g_0,g_2]]} \circ \dots \circ e^{A^{m-1,m-1,m}(t)[g_{m-1},[g_{m-1},g_m]]} \circ \\ &\quad e^{A^{1,0,1}(t)[g_1,[g_0,g_1]]} \circ e^{A^{1,0,2}(t)[g_1,[g_0,g_2]]} \circ \dots \circ e^{A^{m,m-1,m}(t)[g_m,[g_{m-1},g_m]]}\big(x(0)\big) + o(t^3) \end{aligned}$$

(still a lexicographic order has been adopted).

Remark 4.6. We will use this third order approximation only when $m=1$, so that, if $f:=g_0$, $g:=g_1$, it reduces to $x(t)=e^{A^0(t)f}\circ e^{A^1(t)g}\circ e^{A^{0,1}(t)[f,g]}\circ e^{A^{0,0,1}(t)[f,[f,g]]}\circ e^{A^{1,0,1}(t)[g,[f,g]]}(x(0))+o(t^3)$.

As a straightforward consequence of Proposition 4.5, one gets

Corollary 4.7. *Let us fix $t>0$. Assume that $m=1$ and that the vector fields $f:=g_0, g=g_1$ belong to $C^3(\Omega,\mathbb{R}^n)$. If one chooses the L^∞ control $a^0,a^1:[0,t]\to U$ in Remark 4.6 such that $A^0(t)=A^1(t)=A^{0,1}(t)=A^{0,0,1}(t)=0$ and $A^{(1,0,1)}(t)=kt^3+o(t^3)$ with $k>0$, then*

$$x(t)-x(0)=kt^3[g,[f,g]](x(0))+o(t^3).$$

5. Proof of Theorem 3.7

The proof of 1) in Theorem 3.7 is a standard issue in the proof of the classical Pontryagin maximum Principle, so we skip it.

5.0.1. Proof of 2) in Theorem 3.7. Let us recall that i,j, with $0<i<j$, are fixed. Observe that

$$x_\varepsilon(\bar t)=\hat x(2\sqrt{\varepsilon}), \tag{5.1}$$

where $\hat x$ is the solution of the Cauchy problem

$$\begin{cases} \dot{\hat x}(s)=\hat a^0(s)\cdot f(\hat x(s))+\sum_{r=1}^m \hat a^r(s)(s)\cdot g_r(\hat x(s)) \\ \hat x(0)=x(\bar t), \end{cases} \qquad s\in[0,2\sqrt{\varepsilon}] \tag{5.2}$$

where the controls $\hat a^h$, $h=0,\dots,m$, are defined as follows:

$$(\hat a^0,\hat a^1,\dots\hat a^m)(s):=\begin{cases} -(1,u(\bar t-s)) & \text{if } s\in[0,\sqrt{\varepsilon}] \\ \big(1,u(\bar t-2\sqrt{\varepsilon}+s)\big)+\gamma_{(i,j)}\left(\dfrac{s-\sqrt{\varepsilon}}{\sqrt{\varepsilon}}\right) & \text{if } s\in[\sqrt{\varepsilon},2\sqrt{\varepsilon}]. \end{cases}$$

Indeed, notice that one has $\hat x(\sqrt{\varepsilon})=x(\bar t-\sqrt{\varepsilon})$, while in $[\bar t-\sqrt{\varepsilon},\bar t]$ we are implementing the corresponding control $\mathbf{u}_{\varepsilon,\mathbf{c},\bar t}$ in (3.11), so obtaining (5.2). As in Section 4, let us set

$$\hat A^h(s):=\int_0^s \hat a^h(\sigma)d\sigma,\quad \hat A^{h,k}(s):=\int_0^s \hat A^h(\sigma)\hat a^k(\sigma)d\sigma \quad \forall h,k=0,\dots m.$$

Let us check that $\hat A^0(2\sqrt{\varepsilon})=0$ and that the choice of the map $\gamma_{(i,j)}$ yields $\hat A^h(2\sqrt{\varepsilon})=0$ $\forall h\in\{1,\dots,m\}$, $\hat A^{h,k}(2\sqrt{\varepsilon})=0$ $\forall (h,k)\in\{1,\dots,m\}^2\backslash\{(i,j)\}$, and $\hat A^{i,j}(2\sqrt{\varepsilon})\neq 0$.[7] For all $s\in[0,2\sqrt{\varepsilon}]$, one has

$$\hat A^0(s)=-s\mathbf{1}_{[0,\sqrt{\varepsilon}]}+(s-2\sqrt{\varepsilon})\mathbf{1}_{[\sqrt{\varepsilon},2\sqrt{\varepsilon}]},$$

and, for all $h\in\{1,\dots,m\}$,

$$\hat A^h(s)=\int_0^s\left[-\mathbf{1}_{[0,1\sqrt{\varepsilon}]}(\sigma)\cdot u^h(\bar t-\sigma)+\mathbf{1}_{[\sqrt{\varepsilon},2\sqrt{\varepsilon}]}(\sigma)\cdot\left(u^h(\bar t+\sigma-2\sqrt{\varepsilon})+\gamma_{(j,i)}\left(\tfrac{\sigma-\sqrt{\varepsilon}}{\sqrt{\varepsilon}}\right)\right)\right]d\sigma.$$

Hence,

$$\hat A^0(2\sqrt{\varepsilon})=0, \tag{5.3}$$

[7]This will allow us to apply Corollary 4.4 and get the desired conclusion.

and, by $\Gamma_{(i,j)}(1)-\Gamma_{(i,j)}(0)=0$, one also has

$$\begin{aligned}\hat{A}^h(2\sqrt{\varepsilon}) &= \int_0^{2\sqrt{\varepsilon}} \left[-\mathbf{1}_{[0,1\sqrt{\varepsilon}]}(\sigma)\cdot u^h(\bar{t}-\sigma)+\mathbf{1}_{[\sqrt{\varepsilon},2\sqrt{\varepsilon}]}(\sigma)\cdot\Big(u^h(\bar{t}+\sigma-2\sqrt{\varepsilon})\Big)\right]d\sigma\\ &\quad+\sqrt{\varepsilon}\Big(\Gamma^h_{(i,j)}(1)-\Gamma^h_{(i,j)}(0)\Big)\\ &= -\int_0^{\sqrt{\varepsilon}} u^h(\bar{t}-\sigma)d\sigma+\int_0^{\sqrt{\varepsilon}} u^h(\bar{t}-\sigma)d\sigma=0.\end{aligned}$$

In other words , for any $h,k=1,\dots,m$ the time-space curves $(\hat{A}^0,\hat{A}^h)$ and the space curve $(\hat{A}^h,\hat{A}^k)$ are closed, in that $(\hat{A}^0,\hat{A}^k)(0)=(\hat{A}^0,\hat{A}^k)(2\sqrt{\varepsilon})=0$ and $(\hat{A}^h,\hat{A}^k)(0)=(\hat{A}^h,\hat{A}^k)(2\sqrt{\varepsilon})=0$. In order to compute the coefficients $\hat{A}^{h,k}(2\sqrt{\varepsilon})$ when $h,k=1,\dots,m$, let us set, for every $h=1,\dots,m$ and every $s\in[0,2\sqrt{\varepsilon}]$,

$$\check{a}^h_u(s):=-u^h(\bar{t}-s)\cdot\mathbf{1}_{[0,\sqrt{\varepsilon}]}+u^h(\bar{t}-2\sqrt{\varepsilon}+s)\cdot\mathbf{1}_{[\sqrt{\varepsilon},2\sqrt{\varepsilon}]},$$

$$\check{A}^h_u(s):=\int_0^s\check{a}^h_u(\sigma)d\sigma=\int_0^s\Big(-u^h(\bar{t}-\sigma)\cdot\mathbf{1}_{[0,\sqrt{\varepsilon}]}+u^h(\bar{t}-2\sqrt{\varepsilon}+\sigma)\cdot\mathbf{1}_{[\sqrt{\varepsilon},2\sqrt{\varepsilon}]}\Big)d\sigma$$

and

$$\check{a}^h(s):=\gamma^h_{(i,j)}\left(\frac{s-\sqrt{\varepsilon}}{\sqrt{\varepsilon}}\right)\cdot\mathbf{1}_{[\sqrt{\varepsilon},2\sqrt{\varepsilon}]}$$

$$\check{A}^h(s):=\int_0^s\check{a}^h(\sigma)d\sigma=\begin{cases}0 & \forall s\in[0,\sqrt{\varepsilon}]\\ \sqrt{\varepsilon}\left(\Gamma^h_{(i,j)}\left(\frac{s-\sqrt{\varepsilon}}{\sqrt{\varepsilon}}\right)-\Gamma^h_{(i,j)}(0)\right) & \forall s\in[\sqrt{\varepsilon},2\sqrt{\varepsilon}].\end{cases}$$

Hence

$$\hat{a}^h(s)=\check{a}^h_u(s)+\check{a}^h(s)\qquad\hat{A}^h(s)=\check{A}^h_u(s)+\check{A}^h(s)\qquad\forall s\in[0,2\sqrt{\varepsilon}].$$

Observing that $\int_0^{2\sqrt{\varepsilon}}\check{A}^h_u(s)\check{a}^k_u(s)ds=0$, we get

$$\hat{A}^{h,k}(2\sqrt{\varepsilon})=\int_0^{2\sqrt{\varepsilon}}\check{A}^h(s)\check{a}^k_u(s)ds+\int_0^{2\sqrt{\varepsilon}}\check{A}^h_u(s)\check{a}^k(s)ds+\int_0^{2\sqrt{\varepsilon}}\check{A}^h(s)\check{a}^k(s)ds.\tag{5.4}$$

Since the curves $(\hat{A}^0,\hat{A}^k)$, $(\hat{A}^h,\hat{A}^k)$, and $(\check{A}^h_u,\check{A}^k_u)$ are closed, we can give the following interpretation to some of above coefficients:

$$\hat{A}^{0,k}(2\sqrt{\varepsilon})=Area(\hat{A}^0,\hat{A}^k),\quad\hat{A}^{h,k}(2\sqrt{\varepsilon})=Area(\hat{A}^h,\hat{A}^k),$$

and

$$\check{A}^{h,k}_u(2\sqrt{\varepsilon}):=\int_0^{2\sqrt{\varepsilon}}\check{A}^h_u(s)\check{a}^k_u(s)ds=Area(\check{A}^h_u,\check{A}^k_u)=0.$$

Let us compute the three terms on the right-hand side of (5.4). Since $2\sqrt{\varepsilon}$ is a Lebesgue point for the map

$$[\sqrt{\varepsilon},2\sqrt{\varepsilon}]\ni s\to\check{A}^h(s)\check{a}^k_u(s)=u^k\left(\bar{t}-2\sqrt{\varepsilon}+s\right)\cdot\left(\sqrt{\varepsilon}\int_0^{\frac{s-\sqrt{\varepsilon}}{\sqrt{\varepsilon}}}\gamma^h_{(i,j)}(\sigma)\,d\sigma\right),$$

one gets

$$\begin{aligned}\int_0^{2\sqrt{\varepsilon}} \breve{A}^h(s)\breve{a}^k_u(s)ds &= \int_{\sqrt{\varepsilon}}^{2\sqrt{\varepsilon}} \breve{A}^h(s)\breve{a}^k_u(s)ds \\ &= \sqrt{\varepsilon}\sqrt{\varepsilon}\left(u^k(\bar{t})\cdot\left(\int_0^1 \frac{d\Gamma^h_{(j,i)}}{d\sigma}(\sigma)\,d\sigma\right)\right)ds + o(\varepsilon) \\ &= \varepsilon u^k(\bar{t})\left(\Gamma^h_{(j,i)}(1)-\Gamma^h_{(j,i)}(0)\right)ds + o(\varepsilon) = o(\varepsilon) \qquad \forall h,k\in\{1,\dots,m\}.\end{aligned}$$

Moreover, since $\breve{A}^k(2\sqrt{\varepsilon}) = \breve{A}^k(\sqrt{\varepsilon}) = 0$ for all $k\in\{1,\dots,m\}$, the above estimate implies

$$\begin{aligned}\int_0^{2\sqrt{\varepsilon}} \breve{A}^h_u(s)\cdot\breve{a}^k(s)ds &= \int_{\sqrt{\varepsilon}}^{2\sqrt{\varepsilon}} \breve{A}^h_u(s)\cdot\breve{a}^k(s)ds \\ &= \breve{A}^h_u(s)\cdot\breve{A}^k(s)\Big|_{\sqrt{\varepsilon}}^{2\sqrt{\varepsilon}} - \int_{\sqrt{\varepsilon}}^{2\sqrt{\varepsilon}} \breve{a}^h_u(s)\cdot\breve{A}^k(s)ds \\ &= \left(\breve{A}^h_u(2\sqrt{\varepsilon})\cdot\breve{A}^k(2\sqrt{\varepsilon}) - \breve{A}^h_u(\sqrt{\varepsilon})\cdot\breve{A}^k(\sqrt{\varepsilon})\right) + o(\varepsilon) = o(\varepsilon).\end{aligned}$$

Finally, for any $h,k = 1,\dots,m$ (see (3.8) and (3.9)) and

$$\hat{A}^{h,k}(2\sqrt{\varepsilon}) = \int_0^{2\sqrt{\varepsilon}} \breve{A}^h(s)\breve{a}^k(s)ds = \int_{\sqrt{\varepsilon}}^{2\sqrt{\varepsilon}}\left(\int_{\sqrt{\varepsilon}}^s \frac{d\Gamma^h_{(i,j)}}{dt}\left(\frac{\sigma-\sqrt{\varepsilon}}{\sqrt{\varepsilon}}\right)d\sigma\right)\frac{d\Gamma^k_{(i,j)}}{dt}\left(\frac{s-\sqrt{\varepsilon}}{\sqrt{\varepsilon}}\right)ds$$

$$= \varepsilon\int_0^1\left(\int_0^s \frac{d\Gamma^h_{(i,j)}}{d\sigma}(\sigma)\,d\sigma\right)\frac{d\Gamma^k_{(i,j)}}{ds}(s)\,ds = \varepsilon\int_0^1\left(\Gamma^h_{(i,j)}(s)-\Gamma^h_{(i,j)}(0)\right)\gamma^k_{(i,j)}(s)ds =$$

$$\varepsilon\int_0^1 \Gamma^h_{(i,j)}(s)\gamma^k_{(i,j)}(s)ds = \begin{cases} \varepsilon Area\left(\Gamma^h_{(i,j)},\Gamma^k_{(i,j)}\right) = 0 & \text{if } (h,k)\neq(i,j) \text{ or } (h,k)\neq(j,i) \\ \varepsilon Area\left(\Gamma^i_{(i,j)},\Gamma^j_{(i,j)}\right) = \varepsilon/2r^2 & \text{if } (h,k)=(i,j) \\ -\varepsilon Area\left(\Gamma^i_{(i,j)},\Gamma^j_{(i,j)}\right) = -\varepsilon/2r^2 & \text{if } (h,k)=(j,i). \end{cases}{}^{8} \tag{5.5}$$

Moreover, if $k>0$, one has

$$\hat{A}^{0,k}(2\sqrt{\varepsilon}) =$$

$$\int_0^{\sqrt{\varepsilon}} \sigma\cdot u^i(\bar{t}-\sigma)\,d\sigma + \int_{\sqrt{\varepsilon}}^{2\sqrt{\varepsilon}}(\sigma-2\sqrt{\varepsilon})\cdot u^i(\bar{t}-2\sqrt{\varepsilon}+\sigma)\,d\sigma + \int_{\sqrt{\varepsilon}}^{2\sqrt{\varepsilon}} \gamma^k_{(i,j)}\left(\frac{\sigma-\sqrt{\varepsilon}}{\sqrt{\varepsilon}}\right)\cdot(\sigma-2\sqrt{\varepsilon})\,d\sigma$$

$$= \int_{\sqrt{\varepsilon}}^{2\sqrt{\varepsilon}} \gamma^k_{(i,j)}\left(\frac{\sigma-\sqrt{\varepsilon}}{\sqrt{\varepsilon}}\right)\cdot(\sigma-2\sqrt{\varepsilon})\,d\sigma = \tag{5.6}$$

$$= \varepsilon\int_0^1 \frac{d\Gamma^k_{(i,j)}}{ds}(s)\cdot(s-1)\,ds = \Gamma^k_{(i,j)}(s)\cdot(s-1)\Big|_0^1 - \varepsilon\int_0^1 \Gamma^k_{(i,j)}(s)ds = 0.$$

[8] We recall from Definition 3.1 that $r := (\alpha^i)^{-1} + (\alpha^j)^{-1} + (\beta^i)^{-1} + (\beta^j)^{-1}$.

By applying Corollary 4.4, we then get $x_\varepsilon(\bar t)-x(\bar t)=\hat x(2\sqrt{\varepsilon})-x(\bar t)=\varepsilon M\,[g_i,g_j](x(\bar t))+o(\varepsilon)$, where $M:=\dfrac{1}{2r^2}$, so (3.15) is proved.

5.0.2. PROOF OF 3) IN THEOREM 3.7. Let us fix $i\in 1,\dots,m$ and, similarly to the previous step, let us consider again the solution $\tilde x$ of Cauchy problem

$$\begin{cases} \tilde x(s)=\tilde a^0(s)\cdot f(\tilde x(s))+\sum_{r=1}^m \tilde a^r(s)(s)\cdot g_r(\tilde x(s)) \\ \tilde x(0)=x(\bar t)\end{cases} \qquad s\in[0,2\sqrt{\varepsilon}], \tag{5.7}$$

the controls $\tilde a^h$ being now defined as

$$(\tilde a^0,\tilde a^1,\dots\tilde a^m)(s):=\begin{cases}-\,(1,u(\bar t-s)) & \text{if } s\in[0,\sqrt{\varepsilon}]\\ \big(1,u(\bar t-2\sqrt{\varepsilon}+s)\big)+\gamma_{(0,i)}\left(\dfrac{s-\sqrt{\varepsilon}}{\sqrt{\varepsilon}}\right) & \text{if } s\in[\sqrt{\varepsilon},2\sqrt{\varepsilon}].\end{cases}$$

Once again, we will finally exploit equality (5.1), namely $x_\varepsilon(\bar t)=\tilde x(2\sqrt{\varepsilon})$. Let us set

$$\tilde A^h(s):=\int_0^s \tilde a^h(\sigma)d\sigma,\quad \tilde A^{h,k}(s):=\int_0^s \tilde A^h(\sigma)\tilde a^k(\sigma)d\sigma \quad \forall h,k=0,\dots m.$$

As in the previous case we have $\tilde A^h(2\sqrt{\varepsilon})=0$ for all $h=0,\dots m$. Moreover, from

$$\int_0^1 \Gamma^i_{(0,i)}(s)ds=-\frac{\alpha^i\beta^i}{2(\alpha^i+\beta^i)}$$

(see (3.7)), we get

$$\begin{aligned}\tilde A^{0,i}(2\sqrt{\varepsilon})&=\int_0^{\sqrt{\varepsilon}}\sigma\cdot u^i(\bar t-\sigma)\,d\sigma+\int_{\sqrt{\varepsilon}}^{2\sqrt{\varepsilon}}(\sigma-2\sqrt{\varepsilon})\cdot u^i(\bar t-2\sqrt{\varepsilon}+\sigma)\,d\sigma\\ &\quad+\int_{\sqrt{\varepsilon}}^{2\sqrt{\varepsilon}}\gamma^i_{(0,i)}\left(\frac{\sigma-\sqrt{\varepsilon}}{\sqrt{\varepsilon}}\right)\cdot(\sigma-2\sqrt{\varepsilon})\,d\sigma\\ &=\int_{\sqrt{\varepsilon}}^{2\sqrt{\varepsilon}}\gamma^i_{(0,i)}\left(\frac{\sigma-\sqrt{\varepsilon}}{\sqrt{\varepsilon}}\right)\cdot(\sigma-2\sqrt{\varepsilon})\,d\sigma\\ &=\varepsilon\int_0^1\frac{d\Gamma^i_{(0,i)}}{ds}(s)\cdot(s-1)\,ds\\ &=-\varepsilon\int_0^1\Gamma^i_{(0,i)}(s)ds=\frac{\alpha^i\beta^i}{2(\alpha^i+\beta^i)}\varepsilon,\end{aligned}$$

while, for every $k\neq i$, one has

$$\tilde A^{0,k}(2\sqrt{\varepsilon})=\int_0^{\sqrt{\varepsilon}}\sigma\cdot u^i(\bar t-\sigma)\,d\sigma+\int_{\sqrt{\varepsilon}}^{2\sqrt{\varepsilon}}(\sigma-2\sqrt{\varepsilon})\cdot u^i(\bar t-2\sqrt{\varepsilon}+\sigma)\,d\sigma=0.$$

By applying Corollary 4.4, we now get $x_\varepsilon(\bar t)-x(\bar t)=\tilde x(2\sqrt{\varepsilon})-x(\bar t)=\varepsilon\tilde M\,[f,g_i](x(\bar t))$, where $\tilde M:=\dfrac{\alpha^i\beta^i}{2(\alpha^i+\beta^i)}$, so (3.16) is proved.

5.0.3. PROOF OF 4) IN THEOREM 3.7. Once again, we will exploit the equality $x_\varepsilon(\bar t)=\bar x(2\sqrt{\varepsilon})$, where $\bar x$ is the solution on $[0,2\sqrt{\varepsilon}]$ of the Cauchy problem

$$\begin{cases} \bar x(s)=\overline{a}^0(s)\cdot f(\bar x(s))+\overline{a}^1(s)\cdot g_1(\bar x(s)),\\ \bar x(0)=x(\bar t), \end{cases} \tag{5.8}$$

the controls $(\overline{a}^0,\overline{a}^1)$ being now defined as

$$(\overline{a}^0,\overline{a}^1)(s):=\begin{cases} -(1,u(\bar t-s)) & \text{if } s\in[0,\sqrt{\varepsilon}]\\ \big(1,u(\bar t-2\sqrt{\varepsilon}+s)\big)+\gamma_{(1,0,1)}\left(\dfrac{s-\sqrt{\varepsilon}}{\sqrt{\varepsilon}}\right) & \text{if } s\in[\sqrt{\varepsilon},2\sqrt{\varepsilon}]. \end{cases}$$

Let us set

$$\overline{A}^h(s):=\int_0^s \overline{a}^h(\sigma)d\sigma,\qquad \forall h=0,1\qquad \overline{A}^{0,1}(s):=\int_0^s \overline{A}^0(\sigma)\overline{a}^1(\sigma)d\sigma,$$

$$\overline{A}^{1,(0,1)}(s):=\int_0^s \overline{A}^1(\sigma)\frac{d\overline{A}^{0,1}}{d\sigma}(\sigma)d\sigma=\int_0^s \overline{A}^1(\sigma)\overline{A}^0(\sigma)a^1(\sigma)d\sigma,$$

$$\overline{A}^{0,(0,1)}(s):=\int_0^s \overline{A}^0(\sigma)\frac{d\overline{A}^{0,1}}{d\sigma}(\sigma)d\sigma=\int_0^s \overline{A}^0(\sigma)\overline{A}^0(\sigma)a^1(\sigma)d\sigma.$$

As in the previous case, one has $\overline{A}^0(2\sqrt{\varepsilon})=0$,

$$\begin{aligned}\overline{A}^1(2\sqrt[3]{\varepsilon}) &:=\int_0^{2\sqrt[3]{\varepsilon}}\overline{a}^1(s)ds\\ &=-\int_0^{\sqrt[3]{\varepsilon}}u(\bar t-s)ds+\int_{\sqrt[3]{\varepsilon}}^{2\sqrt[3]{\varepsilon}}u(\bar t-2\sqrt[3]{\varepsilon}+s)ds+\int_{\sqrt[3]{\varepsilon}}^{2\sqrt[3]{\varepsilon}}\gamma^1_{(1,0,1)}\left(\frac{s-\sqrt[3]{\varepsilon}}{\sqrt[3]{\varepsilon}}\right)ds\\ &=\sqrt[3]{\varepsilon}\int_0^1\gamma^1_{(1,0,1)}(\sigma)d\sigma=\sqrt[3]{\varepsilon}\Big(\Gamma^1_{(1,0,1)}(1)-\Gamma^1_{(1,0,1)}(0)\Big)\\ &=0.\end{aligned}$$

Moreover, $\overline{A}^{0,1}(2\sqrt[3]{\varepsilon})=\overline{A}^{0,0,1}(2\sqrt[3]{\varepsilon})=0$. Indeed,

$$\begin{aligned}&\overline{A}^{0,1}(2\sqrt[3]{\varepsilon}):=\int_0^{2\sqrt[3]{\varepsilon}}\overline{A}^0(s)\overline{a}^1(s)ds=\\ &=\int_0^{\sqrt[3]{\varepsilon}}su(\bar t-s)ds+\int_{\sqrt[3]{\varepsilon}}^{2\sqrt[3]{\varepsilon}}(s-\sqrt[3]{\varepsilon})u(\bar t-2\sqrt[3]{\varepsilon}+s)ds+\int_{\sqrt[3]{\varepsilon}}^{2\sqrt[3]{\varepsilon}}(s-\sqrt[3]{\varepsilon})\gamma^1_{(1,0,1)}\left(\frac{s-\sqrt[3]{\varepsilon}}{\sqrt[3]{\varepsilon}}\right)ds\\ &=\sqrt[3]{\varepsilon^2}\int_0^1\sigma\gamma^1_{(1,0,1)}(\sigma)ds=\sqrt[3]{\varepsilon^2}\left(1\cdot\Gamma^1_{(1,0,1)}(1)-0\cdot\Gamma^1_{(1,0,1)}(0)-\int_0^1\Gamma^1_{(1,0,1)}(s)ds\right)=0\end{aligned}$$

and

$$\begin{aligned}&\overline{A}^{0,0,1}(2\sqrt[3]{\varepsilon}):=\int_0^{2\sqrt[3]{\varepsilon}}\left(\overline{A}^0(s)\right)^2\overline{a}^1(s)ds=\\ &-\int_0^{\sqrt[3]{\varepsilon}}s^2u(\bar t-s)ds+\int_{\sqrt[3]{\varepsilon}}^{2\sqrt[3]{\varepsilon}}(s-\sqrt[3]{\varepsilon})^2u(\bar t-2\sqrt[3]{\varepsilon}+s)ds+\int_{\sqrt[3]{\varepsilon}}^{2\sqrt[3]{\varepsilon}}(s-\sqrt[3]{\varepsilon})^2\gamma^1_{(1,0,1)}\left(\frac{s-\sqrt[3]{\varepsilon}}{\sqrt[3]{\varepsilon}}\right)ds\\ &=\sqrt[3]{\varepsilon^2}\int_0^1\sigma^2\gamma^1_{(1,0,1)}(\sigma)ds=\sqrt[3]{\varepsilon^2}\left(1\cdot\Gamma^1_{(1,0,1)}(1)-0\cdot\Gamma^1_{(1,0,1)}(0)-\int_0^1\Gamma^1_{(1,0,1)}(s)ds\right)=0.\end{aligned}$$

Finally,

$$
\begin{aligned}
\overline{A}^{1,0,1}(2\sqrt[3]{\varepsilon}) &= \int_0^{2\sqrt[3]{\varepsilon}} (\overline{A}^1)^2 \overline{a}^0 d\sigma = -\int_0^{\sqrt[3]{\varepsilon}} \Big(\int_0^{\eta} u(t-\sigma)d\sigma\Big)^2 d\eta \\
&+ \int_{\sqrt[3]{\varepsilon}}^{2\sqrt[3]{\varepsilon}} \left(-\int_{\eta}^{2\sqrt[3]{\varepsilon}} u(t-\sigma-2\sqrt[3]{\varepsilon})d\sigma + \int_{\sqrt[3]{\varepsilon}}^{\eta} \gamma^1_{(1,0,1)}\Big(\frac{\sigma-\sqrt[3]{\varepsilon}}{\sqrt[3]{\varepsilon}}\Big)d\sigma\right)^2 d\eta \\
&= -\int_0^{\sqrt[3]{\varepsilon}} \Big(\int_0^{\eta} u(t-\sigma)d\sigma\Big)^2 d\eta + \int_{\sqrt[3]{\varepsilon}}^{2\sqrt[3]{\varepsilon}} \left(\int_{\eta}^{2\sqrt[3]{\varepsilon}} u(t-\sigma-2\sqrt[3]{\varepsilon})\right)^2 d\sigma \\
&+ \int_{\sqrt[3]{\varepsilon}}^{2\sqrt[3]{\varepsilon}} \left(-2\int_{\eta}^{2\sqrt[3]{\varepsilon}} u(t-\sigma-2\sqrt[3]{\varepsilon})d\sigma \int_{\sqrt[3]{\varepsilon}}^{\eta} \gamma^1_{(1,0,1)}\left(\frac{\sigma-\sqrt[3]{\varepsilon}}{\sqrt[3]{\varepsilon}}\right)d\sigma \right. \\
&+ \left.\Big(\int_{\sqrt[3]{\varepsilon}}^{\eta} \gamma^1_{(1,0,1)}\left(\frac{\sigma-\sqrt[3]{\varepsilon}}{\sqrt[3]{\varepsilon}}\right)d\sigma\Big)^2\right) d\eta \\
&= \int_{\sqrt[3]{\varepsilon}}^{2\sqrt[3]{\varepsilon}} \left(-2\int_{\eta}^{2\sqrt[3]{\varepsilon}} u(t-\sigma-2\sqrt[3]{\varepsilon})d\sigma \int_{\sqrt[3]{\varepsilon}}^{\eta} \gamma^1_{(1,0,1)}\Big(\frac{\sigma-\sqrt[3]{\varepsilon}}{\sqrt[3]{\varepsilon}}\Big)d\sigma \right. \\
&+ \left.\Big(\int_{\sqrt[3]{\varepsilon}}^{\eta} \gamma^1_{(1,0,1)}\Big(\frac{\sigma-\sqrt[3]{\varepsilon}}{\sqrt[3]{\varepsilon}}\Big)d\sigma\Big)^2\right) d\eta^9 \\
&= -4\sqrt[3]{\varepsilon^2}\int_{\sqrt[3]{\varepsilon}}^{2\sqrt[3]{\varepsilon}} \left[\left(⨍_{\eta}^{2\sqrt[3]{\varepsilon}} u(t-\sigma-2\sqrt[3]{\varepsilon})d\sigma\right)\cdot \Gamma^1_{(1,0,1)}\left(\frac{\eta-\sqrt[3]{\varepsilon}}{\sqrt[3]{\varepsilon}}\right)\right] d\eta \\
&- 2\sqrt[3]{\varepsilon}\int_{\sqrt[3]{\varepsilon}}^{2\sqrt[3]{\varepsilon}} \left[\eta\left(-⨍_{\eta}^{2\sqrt[3]{\varepsilon}} u(t-\sigma-2\sqrt[3]{\varepsilon})d\sigma\right)\cdot \Gamma^1_{(1,0,1)}\left(\frac{\eta-\sqrt[3]{\varepsilon}}{\sqrt[3]{\varepsilon}}\right)\right] d\eta \\
&+ \varepsilon\int_0^1 \left(\Gamma^1_{(1,0,1)}(s)\right)^2 ds \\
&= -4\varepsilon\int_0^1 \left[\left(⨍_{(s+1)\sqrt[3]{\varepsilon}}^{2\sqrt[3]{\varepsilon}} u(t-\sigma-2\sqrt[3]{\varepsilon})d\sigma\right)\cdot \Gamma^1_{(1,0,1)}\Big(s\Big)\right] ds \\
&+ 2\varepsilon\int_0^1 (s+1)\left[\left(⨍_{(s+1)\sqrt[3]{\varepsilon}}^{2\sqrt[3]{\varepsilon}} u(t-\sigma-2\sqrt[3]{\varepsilon})d\sigma\right)\cdot \Gamma^1_{(1,0,1)}\Big(s\Big)\right] ds + \varepsilon\int_0^1 \left(\Gamma^1_{(1,0,1)}(s)\right)^2 ds
\end{aligned}
$$

[9]The symbol $⨍$ denotes the integral average, namely $⨍_a^b \varphi(t)dt := \frac{\int_a^b \varphi(t)dt}{b-a}$.

$$= -2\varepsilon \int_0^1 \left[\left(\fint_{(s+1)\sqrt[3]{\varepsilon}}^{2\sqrt[3]{\varepsilon}} u(t-\sigma-2\sqrt[3]{\varepsilon})d\sigma\right)\cdot \Gamma^1_{(1,0,1)}\Big(s\Big)\right] ds$$
$$+2\varepsilon \int_0^1 \left[\left(\fint_{(s+1)\sqrt[3]{\varepsilon}}^{2\sqrt[3]{\varepsilon}} u(t-\sigma-2\sqrt[3]{\varepsilon})d\sigma\right)\cdot s\cdot \Gamma^1_{(1,0,1)}\Big(s\Big)\right] ds + \varepsilon\int_0^1 \Big(\Gamma^1_{(1,0,1)}(s)\Big)^2 ds$$
(using integration by parts and the equality $\Gamma^1_{(1,0,1)}(1)=\Gamma^1_{(1,0,1)}(0)=0$)
$$= -4\varepsilon \int_0^1 \left[\left(\fint_{(s+1)\sqrt[3]{\varepsilon}}^{2\sqrt[3]{\varepsilon}} u(t-\sigma-2\sqrt[3]{\varepsilon})d\sigma\right)\cdot \Gamma^1_{(1,0,1)}\Big(s\Big)\right] ds + \varepsilon\int_0^1 \Big(\Gamma^1_{(1,0,1)}(s)\Big)^2 ds$$

Notice that
$$\lim_{\varepsilon\to 0}\frac{1}{\sqrt[3]{\varepsilon}-s\sqrt[3]{\varepsilon}}\int_{(s+1)\sqrt[3]{\varepsilon}}^{2\sqrt[3]{\varepsilon}}\Big(u(t-\sigma-2\sqrt[3]{\varepsilon})-u(t)\Big)d\sigma = 0$$
$$\implies \quad \fint_{(s+1)\sqrt[3]{\varepsilon}}^{2\sqrt[3]{\varepsilon}} u(t-\sigma-2\sqrt[3]{\varepsilon})d\sigma\Big) = u(t)+o(1).$$

Therefore,
$$\overline{A}^{1,0,1}(2\sqrt[3]{\varepsilon}) = -4\varepsilon\Big(u(t)+o(1)\Big)\int_0^1 \Gamma^1_{(1,0,1)}\Big(s\Big)ds + \varepsilon\int_0^1\Big(\Gamma^1_{(1,0,1)}(s)\Big)^2 ds$$
$$= \varepsilon\int_0^1\Big(\Gamma^1_{(1,0,1)}(s)\Big)^2 ds.$$

Thanks to Lemma 3.3, Lemma 3.4, and Corollary 4.7, we then deduce that
$$x_\varepsilon(t)-x(t) = k\varepsilon[g,[f,g]](x(t))+o(\varepsilon)$$
with
$$k := \int_0^1\Big(\Gamma^1_{(1,0,1)}(s)\Big)^2 ds = \frac{2}{3}\left((\beta^1)^2\left(\frac{\alpha^1}{2(\alpha^1+\beta^1)}\right)^3 + (\alpha^1)^2\left(\frac{\alpha^1+2\beta^1}{2(\alpha^1+\beta^1)}\right)^3\right).$$

Hence, equality (3.17) is proved, with $\check{M} = k$.

6. Proof of Theorem 2.3

6.1. **Local set separation.** [4, 7]

Definition 6.1 (Local separation of sets)**.** Two subsets E_1 and E_2 are locally separated at x if there exists a neighbourhood W of x such that $E_1\cap E_2\cap W = \{x\}$.

Definition 6.2. Let V be a finite-dimensional real vector space. For any given subset $E\subseteq V$, the set $E^\perp := \{p\in V^*,\, p\cdot c\le 0\,\forall c\in C\}\subseteq V^*$[10] is a closed cone, called the polar cone of E. We say that two convex cones C_1, C_2 are linearly separable if $C_1^\perp\cap -C_2^\perp \supsetneq \{0\}$, that is, if there exists a linear form $\mu\in V^*\backslash\{0\}$ such that $\mu\cdot c_1\ge 0, \mu\cdot c_2\le 0$ for all $(c_1,c_2)\in C_1\times C_2$.

Definition 6.3. Two convex cones C_1, C_2 of a vector space V are said to be transversal if
$$C_1 - C_2 := \Big\{c_1 - c_2, \quad (c_1,c_2)\in C_1\times C_2\Big\} = V.$$
Two transversal cones C_1, C_2 are called strongly transversal if $C_1\cap C_2 \supsetneq \{0\}$.

[10] If V is a finite-dimensional real vector space we use V^* to denote its dual space.

One easily checks the following equivalences for a pair of convex cones C_1 and C_2 (see e.g. [10]):

- *C_1 and C_2 are linearly separable if and only if they are not transversal.*
- *C_1 and C_2 are strongly transversal if and only if they are transversal and there exist a non-zero linear form μ and an element $c \in C_1 \cap C_2$ such that $\mu(c) > 0$.*

The following open-mapping-based result characterizes set-separation in terms of linear separation of approximating cones (see, e.g., [10]).

Theorem 6.4 (Set separation of approximating cones)**.** *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 Boltyanski approximating cones (see Def. 2.2) for Z_1 and Z_2, respectively, at z. If K_1 or K_2 is not a subspace and Z_1, Z_2 are locally separated at z, then K_1 and K_2 are linearly separated, namely there exists a covector $\lambda \in R^n$ such that $0 \neq \lambda \in K_1^\perp \cap K_2^\perp$.*

6.2. **Finitely many variations.** Let us use $(0,T)_{\rm Leb} \subset [0,T]$ to denote the full-measure subset of Lebesgue points of the L^1 map $[0,T] \ni t \mapsto F(t) := f(\bar{x}(t)) + \sum_{i=1}^m g_i(\bar{x}(t))\bar{u}^i(t)$. By Lusin's Theorem there exists a sequence of subsets $E_q \subset [0,\bar{T}]$, $q \geq 0$, such that i) E_0 has null measure, ii) for every $q > 0$ E_q is a compact set such that the restriction of F to E_q is continuous, and iii) $(0,T)_{\rm Leb} = \bigcup_{q=0}^{+\infty} E_q$. For every $q > 0$ let us use $D_q \subseteq E_q$ to denote the set of all density points of E_q[11], which, by Lebesgue's Theorem has the same Lebesgue measure as E_q. In particular, the subset $D := \bigcup_{q=1}^{+\infty} D_q \subset [0,T]$ is full-measure, i.e., it has measure equal to T.

For any $(\bar{t}, \mathbf{c}) \in]0,T] \times \mathfrak{V}$ and any ε sufficiently small, consider the operator

$$\mathscr{A}_{\varepsilon,\mathbf{c},\tau} : L^\infty([0,T],\mathbb{R}^m) \to L^\infty([0,T],\mathbb{R}^m) \qquad \mathscr{A}_{\varepsilon,\mathbf{c},\tau}(u) := u_{\varepsilon,\mathbf{c},\tau}.$$

Clearly, *the control $\mathscr{A}_{\varepsilon,\mathbf{c},\tau}(\bar{u}) = u_{\varepsilon,\mathbf{c},\tau}$ might be not admissible* , namely it can happen that $\mathscr{A}_{\varepsilon,\mathbf{c},\tau}(\bar{u})(\mathscr{T}) \not\subset U$ for some subset $\mathscr{T} \subset [0,T]$ having positive measure. To avoid this drawback, let us define the following subsets of $\mathfrak{V}$:

$$\begin{aligned}
\mathfrak{V}^{\bar{u}}_{Goh} &:= \Big\{ \mathbf{c} = (i,j) \in \mathfrak{V}_{Goh},\ \bar{u} \text{ is } i\text{-balanced and } j\text{-balanced a.e. },\ i,j = 1,\dots,m \Big\} \\
\mathfrak{V}^{\bar{u}}_{LC2} &:= \Big\{ \mathbf{c} = (0,i) \in \mathfrak{V}_{LC2},\ \bar{u} \text{ is } i\text{-balanced a.e. } i = 1,\dots,m \Big\} \\
\mathfrak{V}^{\bar{u}}_{LC3} &:= \Big\{ \mathbf{c} = (1,0,1),\ m = 1\ \bar{u} \text{ is 1-balanced a.e.} \Big\}
\end{aligned} \tag{6.1}$$

and

$$\mathfrak{V}^{\bar{u}} := \mathfrak{V}_{ndl} \cup \mathfrak{V}^{\bar{u}}_{Goh} \cup \mathfrak{V}^{\bar{u}}_{LC2} \cup \mathfrak{V}^{\bar{u}}_{LC3}\,. \tag{6.2}$$ [12]

Let us consider the full-measure subset $\Lambda := \bigcap_r \Lambda_r \subset]0,T[$, where $\Lambda_r \subseteq]0,T[$ is the (full-measure) subset in Definition 2.3, and the intersection is extended to all r such that $r = i$ or $r \in \{i,j\}$, for some of the indexes i,j appearing in hypotheses **v**) $-$ **vii**) of Theorem 2.4.[13]

Let N be a natural number and let us consider N variation signals $\mathbf{c}_1, \dots, \mathbf{c}_N \in \mathfrak{V}^{\bar{u}}$ and N instants $0 < t_1 < \dots t_N < T$, with $t_h \in D \cap \Lambda$. For some $\tilde{\varepsilon} > 0$, let us define the *multiple variation*

[11] A point x is called a *density point* of a Lebesgue–measurable set E if $\lim_{\rho\to 0} \frac{|B_\rho(x)\cap E|}{|B_\rho(x)|} = 1$.

[12] Depending on the hypotheses **v)-vii)** some or even all subsets $\mathfrak{V}^{\bar{u}}_{Goh}$, $\mathfrak{V}^{\bar{u}}_{LC2}$, $\mathfrak{V}^{\bar{u}}_{LC3}$ might be empty. However $\mathfrak{V}^{\bar{u}}$ is never empty, in that $\mathfrak{V}^{\bar{u}} \supseteq \mathfrak{V}_{ndl} (= U)$.

[13] In particular, $\bigcap_r \Lambda_r$ is a *finite* intersection, so that Λ is full-measure, namely $meas(\Lambda) = meas([0,T]) = T$.

of $\overline{u}$

$$[0,\tilde{\varepsilon}]^N \ni \varepsilon \mapsto \overline{u}_\varepsilon := \mathscr{A}_{\varepsilon_N,\mathbf{c}_N,t_N} \circ \ldots \circ \mathscr{A}_{\varepsilon_1,\mathbf{c}_1,t_1}(\overline{u}).$$

Notice that, in view of the hypotheses of the theorem, $\overline{u}_\varepsilon$ turns out to be an admissible control (i.e., $\overline{u}_\varepsilon(t) \in U$ for almost every $t \in [0,T]$) as soon as $\tilde{\varepsilon}$ is sufficiently small. Let $(\overline{u}_\varepsilon, x_\varepsilon)$ be the process corresponding to the control $\overline{u}_\varepsilon$.[14]

The effect of multiple perturbations consists in the sum of single perturbations, as stated in the following elementary result.

Lemma 6.5. *The map $\varepsilon \mapsto x_\varepsilon(T)$ from $[0,\tilde{\varepsilon}]^N$ into $\mathbb{R}^n$ satisfies*

$$x_\varepsilon(T) - x_0(T) = \sum_{i=1}^N \Big(x_{\varepsilon_i \mathbf{e}_i}(T) - x_0(T) \Big) + o(|\varepsilon|), \qquad \forall \varepsilon = (\varepsilon^1, \ldots, \varepsilon^N) \in [0,\tilde{\varepsilon}]^N. \tag{6.3}$$

In relation with Theorem 3.7, let us adopt the notation

$$v_{\mathbf{c},t} := \begin{cases} \displaystyle\sum_{r=1}^m g_r(x(t))\big(\bar{u}^r - u^r(t)\big) & \text{if } \mathbf{c} = \bar{u} \in \mathfrak{V}_{ndl} \\ [g_i, g_j](x(t)) & \text{if } \mathbf{c} = (i,j) \in \mathfrak{V}_{Goh} \\ [f, g_i](x(t)) & \text{if } \mathbf{c} = (0,i) \in \mathfrak{V}_{LC2} \\ [g,[f,g]](x(\bar{t})) & \text{if } m = 1 \text{ and } \{\mathbf{c}\} = \{(1,0,1)\} = \mathfrak{V}_{LC3}\,. \end{cases} \tag{6.4}$$

Using the foundamental matrix $M(\cdot,\cdot)$ of the the variational equation[15]

$$\dot{v}(t) = \frac{\partial}{\partial x}\left(f(x) + \sum_{r=1}^m g_i(x)\overline{u}^r(t) \right)_{x=\bar{x}(t)} \cdot v(t) \tag{6.5}$$

associated to the state equation, for every $r = 1,\ldots,N$ we deduce the first order approximation

$$x_{\varepsilon_r \mathbf{e}_i}(T) - x_0(T) = \varepsilon_r M(T,t_r) \cdot v_{\mathbf{c_r},t_r} + o(|\varepsilon|)$$

Hence, by Lemma 6.5 we get the following fact:

Corollary 6.6. *The map*

$$\varepsilon \mapsto x_\varepsilon(T)$$

from $[0,\tilde{\varepsilon}]^N$ into $\mathbb{R}^n$ satisfies

$$x_\varepsilon(T) - x_0(T) = \sum_{r=1}^N \varepsilon_r v_{\mathbf{c_r},t_r} + o(|\varepsilon|), \qquad \forall \varepsilon = (\varepsilon^1, \ldots, \varepsilon^N) \in [0,\tilde{\varepsilon}]^N. \tag{6.6}$$

---

[14] Of course, $(\overline{u}_\varepsilon, x_\varepsilon)$ depends on the parameters $\mathbf{c}_k$ and t_k as well.

[15] Namely, $t \mapsto M(t,T)$ is the (matrix) solution of the variational Cauchy problem

$$\frac{dM}{dt}(t,T) = \frac{\partial}{\partial x}\left(f(x) + \sum_{r=1}^m g_i(x)\overline{u}^r(t) \right)_{x=\bar{x}(t)} \cdot \frac{dM}{dt}(t,T) \qquad M(T,T) = Id_{\mathbb{R}^n}.$$

This allows us to build a Boltyanski approximation cone at $\begin{pmatrix} x(T) \\ \Psi(x(T)) \end{pmatrix}$ to the $\delta-$ *reachable set* (for some $\delta > 0$)

$\mathscr{R}_\delta =: \left\{ \begin{pmatrix} x(T) \\ \Psi(x(T)) \end{pmatrix}, \exists \text{ an admissible control } u \text{ s.t. } (u,x) \text{is a process s.t.} \|x-\bar{x}\|_{C^0} + \|u-\bar{u}\|_1 < \delta \right\} \subset \mathbb{R}^{n+1}.$

Indeed, Lemma 6.6 can be rephrased by stating that

Lemma 6.7. *Choose $\delta > 0$. The cone $B := L \cdot [0,+\infty)^N \subset \mathbb{R}^{n+1}$, where the homomorphism $L \in Lin(\mathbb{R}^N, \mathbb{R}^{n+1})$ is defined by setting*

$$L(\varepsilon_1,\dots,\varepsilon_N) := \sum_{r=1}^{N} \begin{pmatrix} M(t_r,T)\cdot v_{r,t_r}\varepsilon_r \\ \\ \dfrac{\partial\Psi}{\partial x}(\bar{x}(T))\cdot M(t_r,T)\cdot v_{r,t_r}\varepsilon_r \end{pmatrix} \qquad \forall(\varepsilon_1,\dots,\varepsilon_N)\in\mathbb{R}^N,$$

is a Boltyanski approximating cone at $\begin{pmatrix} \bar{x}(T) \\ \Psi(x(T)) \end{pmatrix}$ of the $\delta-$ reachable set *$\mathscr{R}_\delta$ in the direction of* $[0,+\infty[^N$.

Now, we exploit the fact that, by the definition of local weak minimizer,

- *the sets $\mathscr{R}_\delta$ and the* profitable set $\mathscr{P} := \Big(\mathfrak{T} \times \,]-\infty, \Psi\big(\bar{x}(T)\big)\big[\Big) \cup \left\{ \begin{pmatrix} \bar{x}(T) \\ \Psi\big(\bar{x}(T)\big) \end{pmatrix} \right\} \subset \mathbb{R}^{n+1}$ *are locally separated at* $\begin{pmatrix} \bar{x}(T) \\ \Psi\big(\bar{x}(T)\big) \end{pmatrix}$.

Let us now consider a Boltyanski approximating cone C to the target $\mathfrak{T}$ at $\bar{x}(T)$. Therefore, since $C\times]-\infty,0]$ is a Boltyanski approximating cone of $\mathscr{P}$ at $\begin{pmatrix} \bar{x}(T) \\ \Psi(x(T)) \end{pmatrix}$, in view of Theorem 6.4, the cones B and $C\times]-\infty,0]$ are not strongly transversal, i.e. they are linearly separable.[16]
In other words, there exists an adjoint vector $(\xi,\xi_c) \in -(C\times]-\infty,0])^\perp = -C^\perp\times]-\infty,0]$ verifying

$$\begin{aligned} (\xi,\xi_c)\cdot L(\varepsilon_1,\dots,\varepsilon_N) &= \sum_{r=1}^{N}(\xi,\xi_c)\cdot \begin{pmatrix} M(t_r,T)\cdot v_{r,t_r}\varepsilon_r \\ \\ \dfrac{\partial\Psi}{\partial x}(\bar{x}(T))\cdot M(t_r,T)\cdot v_{r,t_r}\varepsilon_r \end{pmatrix} \\ &= \sum_{r=1}^{N}\left(\xi\cdot M(t_r,T)\cdot v_{r,t_r} + \xi_c\frac{\partial\Psi}{\partial x}(\bar{x}(T))\cdot M(T,t_r)v_{r,t_r}\right)\varepsilon_r \le 0 \end{aligned}$$

for every $(\varepsilon_1,\dots,\varepsilon_N)\in[0,+\infty[^N$, which is equivalent to say that

$$\left(\left(\xi - \lambda\frac{\partial\Psi}{\partial x}(\bar{x}(t))\right)\cdot M(T,t_r)\right)\cdot v_{r,t_r} \le 0 \qquad \forall k=1,\dots,N, \tag{6.7}$$

where we have set $\lambda := -\xi_c(\ge 0)$.

[16] We use the name *profitable set* because this set is made of points which at the same time are admissible and have a cost which is less than or equal to the optimal cost.

Now we shall utilize the invariance of the product of a solution of the adjoint system with a solution of the variational system. Let us use $p:[0,T]\to(\mathbb{R}^n)^*$ to denote the solution of the adjoint Cauchy problem

$$\begin{cases} \dot{p}(t) = -p(t)\dfrac{\partial}{\partial x}\left(f(x)+\displaystyle\sum_1^m g_i(x)\bar{u}(t)\right)_{x=\bar{x}(t)} \\ p(T) = \xi - \lambda\dfrac{\partial \Psi}{\partial x}(\bar{x}(T)). \end{cases} \tag{6.8}$$

As it is well-known,, one has $p(t) = \left(\xi - \lambda\dfrac{\partial\Psi}{\partial x}(\bar{x}(t))\right)\cdot M(T,t)$, for all $t\in[0,T]$. Thus, in particular, the pair $(p(\cdot),\lambda)$ verifies **i**), **ii**) and **iii**) of Theorem 2.4.[17] Furthermore, by the invariance of the scalar product $p(t)\cdot v(t)$ on $[0,T]$, by (6.7) we obtain

$$p(t_k)v_{\mathbf{c}_k,t_k}\le 0 \qquad \forall k=1,\dots,N. \tag{6.9}$$

Now, specializing (6.9) to the various cases described in (6.4), we obtain

$$\begin{aligned}
&p(t_k)\cdot\left(\sum_{\ell=1}^m g_\ell(\bar{x}(t_k))\big(u^\ell-\bar{u}^\ell(t_k)\big)\right)\le 0 && \text{if } \mathbf{c}_k = u\in\mathfrak{V}_{ndl}\\
&p(t_k)\cdot[g_i,g_j](\bar{x}(t_k))\le 0 && \text{if } \mathbf{c_k}=(i,j)\in\mathfrak{V}_{Goh}\\
&p(t_k)\cdot[f,g_i](\bar{x}(t_k))\le 0 && \text{if } \mathbf{c}_k=(0,i)\in\mathfrak{V}_{LC2}\\
&p(t_k)\cdot[g,[f,g]](\bar{x}(\bar{t_k}))\le 0 && \text{if } m=1 \text{ and } \{\mathbf{c_k}\}=\{(1,0,1)\}=\mathfrak{V}_{LC3}.
\end{aligned} \tag{6.10}$$

Furthermore, in view of Remark 3.8, we also get

$$\begin{aligned}
&p(t_k)\cdot[g_i,g_j](\bar{x}(t_k))\ge 0 && \text{if } \mathbf{c_k}=(i,j)\in\mathfrak{V}_{Goh}\\
&p(t_k)\cdot[f,g_i](\bar{x}(t_k))\ge 0 && \text{if } \mathbf{c_k}=(0,i)\in\mathfrak{V}_{LC2}\,,
\end{aligned} \tag{6.11}$$

so that we can improve (6.11) up to obtain

$$\begin{aligned}
&p(t_k)\cdot\left(\sum_{\ell=1}^m g_\ell(\bar{x}(t_k))\big(u^\ell-\bar{u}^\ell(t_k)\big)\right)\le 0 && \text{if } \mathbf{c}_k = u\in\mathfrak{V}_{ndl}\\
&p(t_k)\cdot[g_i,g_j](\bar{x}(t_k))= 0 && \text{if } \mathbf{c_k}=(i,j)\in\mathfrak{V}_{Goh}\\
&p(t_k)\cdot[f,g_i](\bar{x}(t_k))= 0 && \text{if } \mathbf{c}_k=(0,i)\in\mathfrak{V}_{LC2}\\
&p(t_k)\cdot[g,[f,g]](\bar{x}(\bar{t_k}))\le 0 && \text{if } m=1 \text{ and } \{\mathbf{c_k}\}=\{(1,0,1)\}=\mathfrak{V}_{LC3}.
\end{aligned} \tag{6.12}$$

[17]Indeed $(p(\cdot),\lambda)\neq 0$, which coincides with **i**) of Theorem 2.4. Moreover the equation in (6.8) coincides with the adjoint equation $\frac{dp}{dt} = -\frac{\partial H}{\partial x}(\bar{x}(t),p(t),\bar{u}(t))$ in **ii**), while the initial condition in (6.8) is exactly the non-trasversality condition **iii**).

Hence, the restriction to the instants $t_1,\dots,t_N$ of **iv**), **v**, **vi**) and **vii**) in Theorem 2.4 have been proved.

6.3. **Infinitely many variations.** To complete the proof of Theorem 2.4, we need to extend the validity of (6.12) from a finite set of distict instants $0<t_1,\dots,t_k<T$ to a full-measure subset of $[0,T]$. Even though this is a standard procedure —based on Cantor's non–empty intersection theorem— for the sake of self-consistency we will indicate the main steps of this final part of the proof.

Definition 6.8. Let $X\subseteq (D\cap\Lambda)\times\mathfrak{V}^{\overline{u}}$ be any subset of pairs $(t,\mathbf{c})$. We will say that multipliers $(p,\lambda)\in AC\big([0,T],(\mathbb{R}^n)^*\big)\times[0,+\infty[$ satisfies property $\mathscr{P}_X$ if the following conditions **(1)-(6)** are verified:

(1) p is a solution on $[0,T]$ of the differential inclusion

$$\frac{dp}{dt}=-\frac{\partial H}{\partial x}(\overline{x}(t),p(t),\overline{u}(t))\qquad a.e.\quad t\in[0,T].$$

(2) One has

$$p(T)\in-\lambda\frac{\partial^C\Psi}{\partial x}\big(\overline{x}(T)\big)-C^{\perp}.$$

(3) For every $(t,\mathbf{c})\in X$, if $\mathbf{c}=u\in\mathfrak{V}_{ndl}$, then

$$p(t)\left(f(\overline{x}(t))+\sum_{i=1}^m g_i(\overline{x}(t))u^i\right)\le p(t)\left(f(\overline{x}(t))+\sum_{i=1}^m g_i(\overline{x}(t))\overline{u}^i(t)\right).$$

(4) For every $(t,\mathbf{c})\in X$ and if $\mathbf{c}=(i,j)$, with $1\le i<j\le m$, then $0=p(t)\cdot[g_i,g_j](\overline{x}(t))$.
(5) For every $(t,\mathbf{c})\in X$ and if $\mathbf{c}=(0,i)$, with $1\le i\le m$, then $0=p(t)\cdot[f,g_i](\overline{x}(t))$.
(6) For every $(t,\mathbf{c})\in X$, if $m=1$, $g:=g_1$, $\mathbf{c}=(1,0,1)$, and f,g are of class C^2 near $\overline{x}(t)$, then $0\ge p(t)\cdot[g,[f,g]](\overline{x}(t))$.

Finally, let us define the subset $\Theta(X)\subset AC\big([0,T],(\mathbb{R}^n)^*\big)\times[0,+\infty[$ as

$$\Theta(X):=\left\{\begin{array}{l}(p,\lambda)\in AC\big([0,T],(\mathbb{R}^n)^*\big)\times[0,+\infty[:\ |(p(T),\lambda)|=1,\\ (p,\lambda)\text{ verifies the property }\mathscr{P}_X\end{array}\right\}.^{18}$$

Clearly Theorem 2.4 is proved as soon as we are able to show that $\Theta\Big((D\cap\Lambda)\times\mathfrak{V}^{\overline{u}}\Big)\neq\emptyset$. By (6.12) we already know that $\Theta(X)\neq\emptyset$ whenever X comprises N couples $(t_k,\mathbf{c}_k)\in(D\cap\Lambda)\times\mathfrak{V}^{\overline{u}}$ such that $0<t_1<\ldots<t_N<0$. It can be shown that if we allow X to have the general form $X=\Big\{(t_k,\mathbf{c}_k)\in(D\cap\Lambda)\times\mathfrak{V}^{\overline{u}},\ 0<t_1\le t_2\ldots\le t_N<0\Big\}$, then X is still not empty. This is clearly true for the continuity of vector fields involved in the problem and their Lie Brackets (see e.g. [1] for details)[19] . To conclude the proof of Theorem 2.4, notice that

$$\Theta(X_1\cup X_2)=\Theta(X_1)\cap\Theta(X_2),\quad\forall X_1,X_2\subseteq D\times\mathfrak{V},$$

[18]The norm inside the parentheses is the operator norm of $(\mathbb{R}^n\times\mathbb{R})^*$.

[19]For every t_k one can find a sequence $(t_{n,k})_{n\in\ltimes\ltimes}$ such that $t_{n,k}<t_{n+1,k}$ for all n and $t_{n,k}\to t_k+$ and argue taking the limit of the points $t_{n,k}$ and the corresponding multipliers $p(t_k)$.

so that

$$\Theta\Big((D\cap\Lambda)\times\mathfrak{V}^{\overline{u}}\Big) = \bigcap_{\substack{X\subseteq (D\cap\Lambda)\times\mathfrak{V}^{\overline{u}}\\ X \text{ finite}}} \Theta(X). \tag{6.13}$$

So we can deduce that $\Theta\Big((D\cap\Lambda)\times\mathfrak{V}^{\overline{u}}\Big) \neq \emptyset$ by invoking Cantor's theorem, which says that the intersection of an infinite family of sets is non-empty provided every finite intersection of sets of the family is non-empty. The proof of Theorem 2.4 is concluded.

Remark 6.9 (Possible generalizations to non-smooth systems)**.** In [1], the authors have investigated the problem of establishing higher order necessary conditions for minima under weak hypotheses of regularity on the vector fields of the dynamics. In particular, Goh and Legendre-Clebsch conditions have been obtained by making use of *set-valued Lie brackets* of Lipschitz continous vector fields. We conjecture that these kinds of results can be extended to the setting considered in the present paper.

Acknowledgements
The first author is supported by replace with MathGreaterInParis project by Fondation Sciences mathématiques de Paris (FSMP), funding from the European Union's Horizon 2020 research and innovation programme, under the Marie Skłodowska-Curie grant agreement No 101034255. She is also supported by Sorbonne Universite, being affiliated at Laboratoire Jacques Louis Lions (LJLL). Both the author during the writing of this paper were members of the Gruppo Nazionale per l'Analisi Matematica, la Probabilità e le loro Applicazioni (GNAMPA) of the Istituto Nazionale di Alta Matematica (INdAM). In particular they were members of "INdAM -GNAMPA Project 2024", codice CUP E53C23001670001, "Non-smooth optimal control problems". The second author is still GNAMPA member. He is also member of PRIN 2022, Progr-2022238YY5-PE1, "Optimal control problems: analysis, approximations and applications".

REFERENCES

[1] F. Angrisani, F. Rampazzo, Goh and Legendre-Clebsch conditions for nonsmooth control systems, SIAM J. Control Optim. 63 (2025) 86-111.
[2] F. Angrisani, F. Rampazzo, Goh and Legendre-Clebsch conditions for nonsmooth control systems, arXiv:2308.06867.
[3] J. Aubin, H. Frankowska, Set-Valued Analysis, Birkhauser, Boston, Basel, Berlin, 1990.
[4] A.D. Ioffe, V.M. Tikhomirov, Theory of Extremal Problems, Studies in Mathematics and its Applications, Vol. 6, North-Holland, Amsterdam, 1979.
[5] M. Kawski, H. Sussmann, Noncommutative power series and formal Lie-algebraic techniques in nonlinear control theory, In: Operators, Systems and Linear Algebra: Three Decades of Algebraic Systems Theory, U. Helmke, D. Praetzel-Wolters, and E. Zerz, (eds.), B.G. Teubner, Stuttgart, pp. 111–129, 1997.
[6] U. Ledzewicz, H. Schättler, Geometric Optimal Control, Springer, 2016.
[7] F. Rampazzo, Set separation and directional openness with applications to optimal control, J. Convex Anal. 17 (2010), 931-955.
[8] H.J. Sussmann, A product expansion of the Chen series, Theory and Applications of Nonlinear Control Systems, C.I. Byrnes and A. Lindquist (ed.) Elsevier, 1986.
[9] H.J. Sussmann, Generalized Differentials, Variational Generators, and the Maximum Principle with State Constraints, Lecture Notes in Mathematics, vol.1932, pp.221-287, Springer, 2008.
[10] H.J. Sussmann, Geometry and Optimal Control, Mathematical Control Theory, pp. 140–198, Springer, 1999.

[11] H.J. Sussmann, High-order point variations and generalized differentials, Geometric control and nonsmooth analysis, Ser. Adv. Math. Appl. Sci., 76, pp. 327-357, World Sci. Publ., Hackensack, NJ, 2008.