GEOMETRIC APPROACH TO PONTRYAGIN'S MAXIMUM PRINCIPLE

M. Barbero-Liñán, M. C. Muñoz-Lecanda *

Departamento de Matemática Aplicada IV

Edificio C-3, Campus Norte UPC. C/ Jordi Girona 1. E-08034 Barcelona. Spain

Phone numbers: +34934015992, +34934015959. *Fax:* +34934015981.

May 6, 2008

Abstract

Pontryagin's Maximum Principle has been widely discussed and used in Optimal Control Theory since the second half of the 20th century. Here, we focus on the proof and on the understanding of this Principle, using as much as possible geometric ideas and tools. This approach provides a better and clearer understanding of the Principle and, in particular, of the role of the abnormal extremals. In order to give a detailed exposition of the proof, the paper is mostly self-contained.

Key words: *Pontryagin's Maximum Principle, perturbation vectors, tangent perturbation cones, optimal control problems.*

AMS s. c. (2000): 34A12, 49J15, 49J30, 49K05, 49K15, 93C15.

* **e**-*mail and phones*: mbarbero@ma4.upc.edu, matmcml@ma4.upc.edu

Contents

1 Introduction

The importance of Pontryagin's Maximum Principle as a method to find solutions to optimal control problems is the main justification for this work. The use and the comprehension of this Principle does not always gather together. The understanding of this Maximum Principle never finishes as shows the continuous wide number of references in this topic [1, 2, 6, 8, 9, 12, 17, 19, 20, 21, 28, 31, 32, 33] and references therein. We try to contribute to this process through a differential geometric approach.

In 1958 the International Congress of Mathematicians was held in Edinburgh, Scotland, where for the first time L. S. Pontryagin talked publicly about the Maximum Principle. This Principle was developed by a research group on automatic control created by Pontryagin in the fifties. He was engaged in applied mathematics by his friend A. Andronov and because scientists in the Steklov Mathematical Institute were asked to carry out applied research, especially in the field of aircraft dynamics.

At the same time, in the regular seminars on automatic control in the Institute of Automatics and Telemechanics, A. Feldbaum introduced Pontryagin and his collegues to the time-optimization problem. This allowed them to study how to find the best way of piloting an aircraft in order to defeat a zenith fire point in the shortest time as a time-optimization problem.

Since the equations for modelling the aircraft's problem are nonlinear and the control of the rear end of the aircraft runs over a bounded subset, it was necessary to reformulate the calculus of variations known at that time. Taking into account ideas suggested by E. J. McShane in [24], Pontryagin and his collaborators managed to state and prove the Maximum Principle, which was published in Russian in 1961 and translated into English [28] the following year. See [7] for more historical remarks.

Pontryagin's Maximum Principle is considered as an outstanding achievement of the Optimal Control Theory. It has been used in a wide range of applications, such as medicine, traffic flow, robotics, economy, etc. Nevertheless, it is worth remarking that the Maximum Principle does not give sufficient conditions to compute an optimal trajectory; it only provides necessary conditions. Thus only candidates to be optimal trajectories are found. To determine if they are optimal or not, other results related to the existence of solutions for these problems are needed. See [1, 2, 15, 21] for more details.

In this report we present a full detailed proof of the Maximum Principle, trying to be self-contained as much as possible. Hence some of the necessary results from other areas of mathematics (differential equations, differential geometry, convexity, separating hyperplanes) are included as appendices, although we assume some knowledge in differential geometry, such as the core chapters of [22], differential equations [13, 14, 16], and convexity [5, 29].

One of the main elements in this report is a vector field along a projection defined in §2 together with its properties. In the heart of the report there are two big parts corresponding with two different statements of Pontryagin's Maximum Principle. If the optimal control problem has the time interval given and also the endpoints, then we have to look at §3, §4. If the final time is not given and the endpoints are submanifolds, then the corresponding paragraphs are §5, §6. These two areas have been written in an analogous way. First of all, two different but equivalent statements of the optimal control problems are given. The so-called extended system is the useful one in §4, §6 because the functional to be minimized is included as a new coordinate of the system. One part of the proof of Pontryagin's Maximum Principle consists of perturbing the given optimal curve, therefore we introduce how this curve can be perturbed depending on the known data. The last subsection in §3, §5 explains a hamiltonian problem that leads to the statements of Maximum Principle. In this way, the proof is just in the following paragraph. Our purpose is to give an intrinsic proof of the Maximum Principle, but at some point it is necessary the use of local results and coordinate expressions.

The appendices contain essential results for the core of the report and also some explanation to make clear some well-known ideas related with the reachable set and the tangent perturbation cone, Appendix C. The way to think about it is that the tangent perturbation cone contains, in some sense, the vectors tangent to perturbation curves.

The solutions to optimal control problems are of different kinds (abnormal and normal), mainly determined by the role of the cost function. An extremal is a curve candidate to be a solution to the problem. The abnormal ones are extremals with the particularity that the cost function is not used to define them. Nevertheless, the cost function is important to guarantee the optimality of the abnormal extremals, as is pointed out in §3.4. Any chance we have along the report to make a comment about abnormality will be made.

All the effort to elaborate this work is being used to enlighten the research, from a geometric point of view, on abnormal and strict abnormal extremals in optimal control problems in general [3], and for mechanical systems [4]. Those extremals are of the great interest since their optimality was proved in specific problems in subRiemannian geometry [23, 25, 26].

The origin of this report was a series of seminars and talks with Professor Andrew D. Lewis during his stay in our Department on sabbatical during the first term of 2005. We tried to understand the details of the proof as a way to work on some aspects of controllability and accesibility of control systems with a cost function, [12, 17], and where abnormal solutions are in the accesibility sets.

In the sequel, unless otherwise stated, all the manifolds are real, second countable and $\mathcal{C}^\infty$ and the maps are assumed to be $\mathcal{C}^\infty$. Sum over repeated indices is understood.

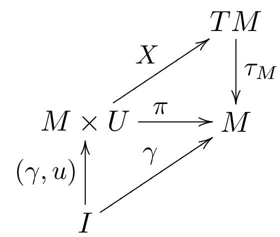
2 General setting

Control theory is studied from a differential geometric point of view as long as vector fields depending on parameters are introduced. In §3.3, §5.2, we concentrate on how a vector field depending on parameters evolves when the parameters change. We refer the reader to [13, 14, 16, 18] for more details.

Let M be a differentiable manifold and $U \subset \mathbb{R}^m$ an open set. Consider the trivial bundle $\pi\colon M \times U \to M$.

Let X be a vector field on M along the projection π; that is, if $\tau_M\colon TM \to M$ is the natural projection, then $X\colon M \times U \to TM$ and $\tau_M \circ X = \pi$. If (x^i) are local coordinates on M, its local expression is $X = f^i \partial/\partial x^i$ where f^i are functions defined on an open set of $M \times U$.

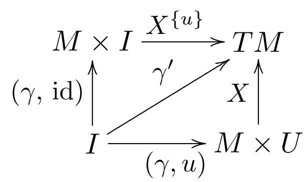
Let $I \subset \mathbb{R}$ be a closed interval and $(\gamma, u)\colon I \to M \times U$ a curve. All these elements come together in the following diagram:



Usually the parameters are called *controls* and are assumed to be a mapping $u\colon I \to U$. A given map u defines a time-dependent vector field on M,

$$\begin{array}{rcl} X^{\{u\}}\colon M \times I & \longrightarrow & TM \\ (x,t) & \longmapsto & X^{\{u\}}(x,t) = X(x,u(t)). \end{array}$$

If γ is an integral curve of $X^{\{u\}}$, the following diagram commutes:



(2.1)

That is, $\gamma' = X^{\{u\}} \circ (\gamma, id)$.

A differentiable time-dependent vector field X has associated the *time dependent flow* or *evolution operator of* X defined as

$$\begin{array}{rl} \Phi^X\colon \quad I \times I \times M & \longrightarrow M \\ (t,s,x) & \longmapsto \Phi^X(t,s,x) = \Phi^X_{(s,x)}(t) \end{array}$$

where $\Phi^X_{(s,x)}$ is the integral curve of X with initial condition x at time s. See Appendix B.1 for more details. Moreover, the evolution operator defines a diffeomorphism on M that is used in the following section $\Phi^X_{(t,s)}\colon M \to M$, $x \mapsto \Phi^X_{(t,s)}(x) = \Phi^X_{(s,x)}(t)$.

From now on, we assume the following conditions.

1. $I = [a,b]$ is a fixed interval.
2. The vector field X is continuous on $M \times \overline{U}$, where $\overline{U}$ is the closure of U, and continuously differentiable on M for every $u \in \overline{U}$.
3. The controls $u\colon I \to U$ are measurable and bounded. Hence the vector fields $X^{\{u\}}$ are measurable on t, and for a fixed t, they are differentiable on M. We need to impose that the curves $\gamma\colon I \to M$ to be absolutely continuous. So they are generalized integral curves of the vector field $X^{\{u\}}$, that is, they only satisfy condition $\dot{x}^i = f^i(x,u)$ at points where γ is derivable, which happens almost everywhere.

   We will suppose that the integral curves of vector fields in this paper are integral curves in this generalized sense. For more details, see Appendix A and [13, 14].

3 Pontryagin's Maximum Principle for fixed time and fixed end-points

In the general setting a manifold M has been used to introduce the concepts. In optimal control theory the manifold is denoted by Q.

3.1 Statement of optimal control problem and notation

Let Q be a differentiable manifold of dimension n and $U \subset \mathbb{R}^m$ a subset. Let us consider the trivial fiber bundle $\pi\colon Q \times U \to Q$.

Let X be a vector field along the projection $\pi\colon Q \times U \to Q$. If (x^i) are local coordinates on Q, the local expression of the vector field is $X = f^i \partial/\partial x^i$ where f^i are functions defined on an open set of $Q \times U$.

Let $I \subset \mathbb{R}$ be an interval and $(\gamma, u)\colon I \to Q \times U$ a curve. Given $F\colon Q \times U \to \mathbb{R}$, let us consider the functional

$$\mathcal{S}[\gamma, u] = \int_I F(\gamma, u)\, dt$$

defined on curves (γ, u) with a compact interval as domain. The function $F\colon Q \times \overline{U} \to \mathbb{R}$ is continuous on $Q \times \overline{U}$ and continuously differentiable with respect to Q on $Q \times \overline{U}$.

Statement 3.1. ***(Optimal Control Problem,** OCP)* *Given the elements Q, U, X, F, $I = [a, b]$ and the endpoint conditions x_a, $x_b \in Q$, consider the following problem.*

Find (γ^*, u^*) such that

(1) endpoint conditions: $\gamma^*(a) = x_a$, $\gamma^*(b) = x_b$,

(2) γ^* is an integral curve of $X^{\{u^*\}}$: $\dot{\gamma}^*(t) = X(\gamma^*(t), u^*(t))$, $t \in I$, and

(3) minimal condition: $\mathcal{S}[\gamma^*, u^*]$ is minimum over all curves (γ, u) satisfying (1) *and* (2)*.*

The tuple $(Q, U, X, F, I, x_a, x_b)$ denotes the *optimal control problem*. The function F is called the *cost function* of the problem. The mappings $u\colon I \to U$ are called *controls*.

Comments: Remember from §2:

1. Given (γ, u), the function $u\colon I \to U$ allows us to construct a time-dependent vector field on Q, $X^{\{u\}}\colon Q \times I \to TQ$, defined by $X^{\{u\}}(x, t) = X(x, u(t))$. Condition (2) shows that γ^* is an integral curve of $X^{\{u^*\}}$.
2. The curve (γ^*, u^*) locally satisfies the differential equation $\dot{x}^i = f^i$, the endpoint conditions $\gamma(a) = x_a$, $\gamma(b) = x_b$ and minimizes the functional $\mathcal{S}[\gamma, u]$.

3.2 The extended problem

Taking into account the elements defining the optimal control problem and their properties, we state an equivalent problem.

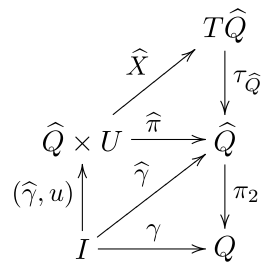
Given the *OCP* $(Q, U, X, F, I, x_a, x_b)$, let us consider $\widehat{Q} = \mathbb{R} \times Q$ and the trivial fiber bundle $\widehat{\pi}\colon \widehat{Q} \times U \to \widehat{Q}$.

Let $\widehat{X}$ be the following vector field along the projection $\widehat{\pi}\colon \widehat{Q} \times U \to \widehat{Q}$:

$$\widehat{X}(x^0, x, u) = F(x, u)\partial/\partial x^0|_{(x^0, x, u)} + X(x, u),$$

where x^0 is the natural coordinate on $\mathbb{R}$.

Given a curve $(\widehat{\gamma}, u) = ((x^0 \circ \widehat{\gamma}, \gamma), u) \colon I \to \widehat{Q} \times U$. The previous elements come together in the following diagram:



where π_2 is the projection of $\widehat{Q}$ onto Q.

Statement 3.2. ***(Extended Optimal Control Problem,*** $\widehat{\mathbf{OCP}}$***)*** *Given the OCP* $(Q, U, X, F, I, x_a, x_b)$, $\widehat{Q}$ *and* $\widehat{X}$ *as defined above, consider the following problem.*

Find $(\widehat{\gamma}^*, u^*)$ *such that*

(1) endpoint conditions: $\widehat{\gamma}^*(a) = (0, x_a)$, $\gamma^*(b) = x_b$,

(2) $\widehat{\gamma}^*$ *is an integral curve of* $\widehat{X}^{\{u^*\}}$: $\dot{\widehat{\gamma}}^*(t) = \widehat{X}(\widehat{\gamma}^*(t), u^*(t))$, $t \in I$, *and*

(3) minimal condition: ${\gamma^*}^0(b)$ *is minimum over all curves* $(\widehat{\gamma}, u)$ *satisfying* (1) *and* (2).

The tuple $(\widehat{Q}, U, \widehat{X}, I, x_a, x_b)$ denotes the *extended optimal control problem.*

Comments: The extended problem fulfills analogous properties to those satisfied by the optimal control problem.

1. The functional ${\gamma^*}^0(b)$ to be minimized in the $\widehat{OCP}$ is equal to the functional defined in the OCP. That is to say, we have

$$\widehat{\mathcal{S}}[\widehat{\gamma}, u] = \gamma^0(b) = \int_a^b F(\gamma, u) dt = \mathcal{S}[\gamma, u]$$

   for curves $(\widehat{\gamma}, u)$.

2. The curve $(\widehat{\gamma}^*, u^*)$ locally satisfies the differential equations $\dot{x}^0 = F$, $\dot{x}^i = f^i$, the conditions $\widehat{\gamma}(a) = (0, x_a)$, $\gamma(b) = x_b$, and optimizes the functional $\widehat{\mathcal{S}}[\widehat{\gamma}, u]$.

The elements defining the problem $(\widehat{Q}, U, \widehat{X}, I, x_a, x_b)$ satisfy the same properties as the elements of the problem $(Q, U, X, F, I, x_a, x_b)$, see §2, §3.1.

3.3 Perturbation and associated cones

The following constructions can be defined for any vector field depending on parameters §2, in particular, for those vector fields defining a control system. In order to point out this generality of these constructions we use the manifold M instead of Q. In this way M can be Q, $\widehat{Q}$ or any other convenient manifold along this report.

3.3.1 Elementary perturbation vectors: class I

Now we study how integral curves of the time-dependent vector field $X^{\{u\}} \colon M \times I \to TM$, introduced in §2, change when the control u is perturbed in a small interval.

In the sequel, a control $u\colon I \to U$ and an integral curve $\gamma\colon I \to M$ of $X^{\{u\}}$ are given. Let $\pi_1 = \{t_1, l_1, u_1\}$, where t_1 is a Lebesgue time in (a,b) always for the $X \circ (\gamma, u)$ (i.e. it satisfies Equation (A.15)), $l_1 \in \mathbb{R}^+$, $u_1 \in U$. From now on to simplify t_1 is called just a Lebesgue time. For every $s \in \mathbb{R}^+$ small enough such that $a < t_1 - l_1 s$, consider $u[\pi_1^s]\colon I \to U$ defined by

$$u[\pi_1^s](t) = \begin{cases} u_1, & t \in [t_1 - l_1 s, t_1], \\ u(t), & \text{elsewhere.} \end{cases}$$

Definition 3.3. *The function $u[\pi_1^s]$ is called an **elementary perturbation of** u **specified by the data** $\pi_1 = \{t_1, l_1, u_1\}$. It is also called a **needle-like variation**.*

Associated to $u[\pi_1^s]$, consider the mapping $\gamma[\pi_1^s]\colon I \to M$, the generalized integral curve of $X^{\{u[\pi_1^s]\}}$ with initial condition $(a, \gamma(a))$.

Given $\epsilon > 0$, define the map

$$\begin{array}{rcl} \varphi_{\pi_1}\colon I \times [0,\epsilon] & \longrightarrow & M \\ (\, t \,,\, s \,) & \longmapsto & \varphi_{\pi_1}(t,s) = \gamma[\pi_1^s](t) \end{array}$$

For every $t \in I$, $\varphi_{\pi_1}^t\colon [0,\epsilon] \to M$ is given by $\varphi_{\pi_1}^t(s) = \varphi_{\pi_1}(t,s)$.

As the controls are measurable and bounded, it makes sense to define the distance between two controls $u, \overline{u}\colon I \to U$ as follows

$$d(u, \overline{u}) = \int_I \|u(t) - \overline{u}(t)\| \; dt$$

where $\|\cdot\|$ is the usual norm in $\mathbb{R}^m$. The control $u[\pi_1^s]$ depends continuously on the parameters s and $\pi_1 = \{t_1, l_1, u_1\}$, that is, given $\epsilon > 0$ there exists $\delta > 0$ such that if $|t_1 - t_2| < \delta$, $|l_1 - l_2| < \delta$, $\|u_1 - u_2\| < \delta$, $|s_1 - s_2| < \delta$, then $d(u[\pi_1^{s_1}], u[\pi_2^{s_2}]) < \epsilon$.

Hence the curve $\varphi_{\pi_1}^t$ depends continuously on s and $\pi_1 = \{t_1, l_1, u_1\}$, then it converges uniformly to γ as s tends to 0. See [13, 14] for more details of the differential equations depending continuously on parameters.

Let us prove that the curve $\varphi_{\pi_1}^{t_1}$ has a tangent vector at $s = 0$. Let $u[\pi_1^s]$ be an elementary perturbation of u specified by $\pi_1 = \{t_1, l_1, u_1\}$ and consider the curve $\varphi_{\pi_1}^{t_1}\colon [0,\epsilon] \to M$, $\varphi_{\pi_1}^{t_1}(s) = \gamma[\pi_1^s](t_1)$.

Proposition 3.4. *If t_1 is a Lebesgue time, the curve $\varphi_{\pi_1}^{t_1}\colon [0,\epsilon] \to M$ is differentiable at $s = 0$. Its tangent vector is $[X(\gamma(t_1), u_1) - X(\gamma(t_1), u(t_1))]\; l_1$.*

Proof. It is enough to prove that for every differentiable function $g\colon M \to \mathbb{R}$, there exists

$$A = \lim_{s \to 0} \frac{g(\varphi_{\pi_1}^{t_1}(s)) - g(\varphi_{\pi_1}^{t_1}(0))}{s}$$

As this is a derivation on the functions defined on a neighbourhood of $\gamma(t_1)$, it is enough to prove the proposition for the coordinate functions x^i of a local chart at $\gamma(t_1)$. Thus take $g = x^i$,

$$A = \lim_{s \to 0} \frac{(x^i \circ \varphi_{\pi_1}^{t_1})(s) - (x^i \circ \varphi_{\pi_1}^{t_1})(0)}{s} = \lim_{s \to 0} \frac{(x^i \circ \gamma[\pi_1^s])(t_1) - (x^i \circ \gamma)(t_1)}{s} = \lim_{s \to 0} \frac{\gamma^i[\pi_1^s](t_1) - \gamma^i(t_1)}{s}$$

As γ is an absolutely continuous integral curve of $X^{\{u\}}$, $\dot{\gamma}(t) = X(\gamma(t), u(t))$ at every Lebesgue time. Then integrating

$$\gamma^i(t_1) - \gamma^i(a) = \int_a^{t_1} f^i(\gamma(t), u(t))dt$$

and similarly for $\gamma[\pi_1^s]$ and $u[\pi_1^s]$. Observe that $\gamma[\pi_1^s](t) = \gamma(t)$ and $u[\pi_1^s](t) = u(t)$ for $t \in [a, t_1 - l_1 s)$. Then,

$$A = \lim_{s\to 0} \frac{\int_a^{t_1} f^i(\gamma[\pi_1^s](t), u[\pi_1^s](t))dt - \int_a^{t_1} f^i(\gamma(t), u(t))dt}{s} =$$

$$= \lim_{s\to 0} \frac{\int_{t_1-l_1 s}^{t_1} f^i(\gamma[\pi_1^s](t), u_1)dt - \int_{t_1-l_1 s}^{t_1} f^i(\gamma(t), u(t))dt}{s}.$$

As t_1 is a Lebesgue time, we use Equation (A.15): $\int_{t-h}^{t} X(\gamma(s), u(s))ds = hX(\gamma(t), u(t)) + o(h)$.

$$A = \lim_{s\to 0} \frac{f^i(\gamma[\pi_1^s](t_1), u_1)l_1 s - f^i(\gamma(t_1), u(t_1))l_1 s + o(s)}{s} = \lim_{s\to 0}[f^i(\gamma[\pi_1^s](t_1), u_1) - f^i(\gamma(t_1), u(t_1))]\, l_1$$

As f^i is continuous on M, we have

$$A = \lim_{s\to 0}[f^i(\gamma[\pi_1^s](t_1), u_1) - f^i(\gamma(t_1), u(t_1))]\, l_1 = [f^i(\lim_{s\to 0} \gamma[\pi_1^s](t_1), u_1) - f^i(\gamma(t_1), u(t_1))]\, l_1 =$$

$$= [f^i(\gamma(t_1), u_1) - f^i(\gamma(t_1), u(t_1))]\, l_1 = [(X(\gamma(t_1), u_1) - X(\gamma(t_1), u(t_1)))\ \ l_1]\,(x^i).$$

□

Definition 3.5. *The tangent vector $v[\pi_1] = (X(\gamma(t_1), u_1) - X(\gamma(t_1), u(t_1)))\ l_1 \in T_{\gamma(t_1)}M$ is the* ***elementary perturbation vector associated to the perturbation data*** *$\pi_1 = \{t_1, l_1, u_1\}$. It is also called a* ***perturbation vector of class I****.*

Comments:

(a) The previous proof shows the importance of defining perturbations only at Lebesgue times, otherwise the elementary perturbation vectors do not exist.

(b) Observe that if we change $\pi_1 = \{t_1, l_1, u_1\}$ for $\pi_2 = \{t_1, l_2, u_1\}$, then $v[\pi_1] = (l_1/l_2)\ v[\pi_2]$. If $v[\pi_1]$ is a perturbation vector of class I and $\lambda \in \mathbb{R}^+$, then $\lambda\ v[\pi_1]$ is also a perturbation vector of class I with perturbation data $\{t_1, \lambda\ l_1, u_1\}$.

(c) We write $L(w)g$ for the derivative of the function g in the direction given by the vector $w \in T_xM$. Due to Proposition 3.4, for every $g\colon M \to \mathbb{R}$ differentiable function we have

$$\frac{g(\varphi_{\pi_1}^{t_1}(s)) - g(\gamma(t_1)) - s\ L(v[\pi_1])g}{s} \underset{s\to 0}{\longrightarrow} 0.$$

Hence

$$g\left(\varphi_{\pi_1}^{t_1}(s)\right) = g\left(\gamma(t_1)\right) + s\, L(v[\pi_1])g + o(s).$$

If (x^i) are local coordinates of a chart at $\gamma(t_1)$,

$$x^i\left(\varphi_{\pi_1}^{t_1}(s)\right) = x^i\left(\gamma(t_1)\right) + s\, v[\pi_1]^i + o(s),$$

that is,

$$\left(\varphi_{\pi_1}^{t_1}\right)^i(s) = \gamma^i(t_1) + s\, v[\pi_1]^i + o(s).$$

Now, if we identify the open set of the local chart and the tangent space to M at $\gamma(t_1)$ with the same space $\mathbb{R}^n$, we write the following linear approximation

$$\varphi_{\pi_1}^{t_1}(s) = \gamma(t_1) + s\, v[\pi_1] + o(s). \tag{3.2}$$

Let $V[\pi_1]\colon [t_1, b] \to TM$ be the integral curve of the complete lift $\left(X^T\right)^{\{u\}}$ of $X^{\{u\}}$ with initial condition $(t_1, (\gamma(t_1), v[\pi_1]))$. See Appendix B for more details.

Note that $\varphi_{\pi_1}^{t}(s) = \Phi_{(t,t_1)}^{X^{\{u\}}}\left(\varphi_{\pi_1}^{t_1}(s)\right)$ for $t \geq t_1$ because of the definition of φ_{π_1} and $u[\pi_1^s]$.

Proposition 3.6. *For every Lebesgue time $t \in (t_1, b]$, $V[\pi_1](t)$ is the tangent vector to the curve $\varphi_{\pi_1}^t\colon [0, \epsilon] \to M$ at $s = 0$.*

3.3.2 Perturbation vectors of class II

The control can be perturbed twice instead of only once, in fact it may be modified a finite number of times. If t_2 is a Lebesgue time greater than t_1, we perturb the control with $\pi_1 = \{t_1, l_1, u_1\}$ and $\pi_2 = \{t_2, l_2, u_2\}$, and we obtain the perturbation data $\pi_{12} = \{(t_1, t_2), (l_1, l_2), (u_1, u_2)\}$ given by

$$u[\pi_{12}^s](t) = \begin{cases} u_1, & t \in [t_1 - l_1 s, t_1], \\ u_2, & t \in [t_2 - l_2 s, t_2], \\ u(t), & \text{elsewhere} \end{cases}$$

for every $s \in \mathbb{R}^+$ small enough such that $[t_1 - l_1 s, t_1] \cap [t_2 - l_2 s, t_2] = \emptyset$. Then $\gamma[\pi_{12}^s] \colon I \longrightarrow M$ is the generalized integral curve of $X^{\{u[\pi_{12}^s]\}}$ with initial condition $(a, \gamma(a))$. Observe that $\gamma[\pi_{12}^0](t) = \gamma(t)$. Consider the curve $\varphi_{\pi_{12}}^{t_2} \colon [0, \epsilon] \to M$ given by $\varphi_{\pi_{12}}^{t_2}(s) = \gamma[\pi_{12}^s](t_2)$.

Proposition 3.7. *Let $t_1 < t_2$. The vector tangent to $\varphi_{\pi_{12}}^{t_2} \colon [0, \epsilon] \to M$ at $s = 0$ is $v[\pi_2] + V[\pi_1](t_2)$, where $V[\pi_1] \colon [t_1, b] \to TM$ is the generalized integral curve of $\left(X^T\right)^{\{u\}}$ with initial condition $(t_1, (\gamma(t_1), v[\pi_1]))$.*

Proof. Here we perturb the control first with π_1 along γ and we obtain $u[\pi_1^s]$. Then we perturb this last control with the other perturbation data, π_2, along $\gamma[\pi_1^s]$. Then the superindeces of the tangent vectors denotes the curve along which the perturbation is made. As in the proof of Proposition 3.4,

$$A = \lim_{s \to 0} \frac{(x^i \circ \varphi_{\pi_{12}}^{t_2})(s) - (x^i \circ \varphi_{\pi_{12}}^{t_2})(0)}{s} = \lim_{s \to 0} \frac{(x^i \circ \gamma[\pi_{12}^s])(t_2) - (x^i \circ \gamma)(t_2)}{s} =$$

$$= \lim_{s \to 0} \frac{\gamma^i[\pi_{12}^s](t_2) - \gamma^i(t_2)}{s} = \lim_{s \to 0} \left\{ \frac{\gamma^i[\pi_{12}^s](t_2) - \gamma^i[\pi_1^s](t_2)}{s} + \frac{\gamma^i[\pi_1^s](t_2) - \gamma^i(t_2)}{s} \right\}$$

We understand $\gamma[\pi_{12}^s]$ as the result of perturbing $\gamma[\pi_1^s]$ with π_2, and use the linear approximation in Equation (3.2) for $\gamma[\pi_{12}^s](t_2)$ and $\gamma[\pi_1^s](t_2)$ according to Proposition 3.4.

$$\varphi_{\pi_{12}}^{t_2}(s) = \gamma[\pi_{12}^s](t_2) = \gamma[\pi_1^s](t_2) + s\, v[\pi_2]^{\gamma[\pi_1^s]} + o(s),$$
$$\gamma[\pi_1^s](t_2) = \gamma(t_2) + s\, V[\pi_1]^\gamma(t_2) + o(s),$$

then

$$A = \lim_{s \to 0} \left\{ \frac{s(v[\pi_2]^{\gamma[\pi_1^s]})^i}{s} + \frac{s(V[\pi_1]^\gamma)^i(t_2)}{s} \right\} = \lim_{s \to 0} \left\{ (v[\pi_2]^{\gamma[\pi_1^s]})^i + (V[\pi_1]^\gamma)^i(t_2) \right\}.$$

As $\gamma[\pi_1^s]$ depends on s, satisfying $\lim_{s \to 0} \gamma[\pi_1^s](t) = \gamma(t)$. Thus $A = \mathrm{L}\left(v[\pi_2]^\gamma + V[\pi_1]^\gamma(t_2)\right) x^i$. □

Considering identifications similar to the ones used to write Equation (3.2), we have

$$\varphi_{\pi_{12}}^{t_2}(s) = \gamma(t_2) + s v[\pi_2] + s V[\pi_1](t_2) + o(s).$$

Now we define how the control changes when it is perturbed twice at the same time. If t_1 is a Lebesgue time, $\pi_1' = \{t_1, l_1', u_1'\}$ and $\pi_1'' = \{t_1, l_1'', u_1''\}$ are perturbation data, then $\pi_{11} = \{(t_1, t_1), (l_1', l_1''), (u_1', u_1'')\}$ is a perturbation data given by

$$u[\pi_{11}^s](t) = \begin{cases} u_1', & t \in [t_1 - (l_1' + l_1'')s, t_1 - l_1'' s], \\ u_1'', & t \in [t_1 - l_1'' s, t_1], \\ u(t), & \text{elsewhere.} \end{cases}$$

for every $s \in \mathbb{R}^+$ small enough such that $a < t_1 - (l_1' + l_1'')s$. Then $\gamma[\pi_{11}^s] \colon I \longrightarrow M$ is the generalized integral curve of $X^{\{u[\pi_{11}^s]\}}$ with initial condition $(a, \gamma(a))$. Observe that $\gamma[\pi_{11}^0](t) = \gamma(t)$. Consider the curve $\varphi_{\pi_{11}}^{t_1} \colon [0, \epsilon] \to M$, defined by $\varphi_{\pi_{11}}^{t_1}(s) = \gamma[\pi_{11}^s](t_1)$.

Proposition 3.8. *The vector tangent to* $\varphi^{t_1}_{\pi_{11}}\colon [0,\epsilon] \to M$ *at* $s = 0$ *is* $v[\pi'_1] + v[\pi''_1]$*, where* $v[\pi'_1]$*,* $v[\pi''_1]$ *are the perturbation vectors of class I associated to* π'_1*,* π''_1 *respectively.*

Proof. As in the proof of Proposition 3.4

$$A = \lim_{s\to 0} \frac{(x^i \circ \varphi^{t_1}_{\pi_{11}})(s) - (x^i \circ \varphi^{t_1}_{\pi_{11}})(0)}{s} = \lim_{s\to 0} \frac{\gamma^i[\pi^s_{11}](t_1) - \gamma^i(t_1)}{s}$$

As γ is an integral curve of $X^{\{u\}}$ absolutely continuous, $\dot\gamma(t) = X(\gamma(t), u(t))$ at every Lebesgue time. Then integrating $\gamma^i(t_1) - \gamma^i(a) = \int_a^{t_1} f^i(\gamma(t), u(t))dt$ and similarly for $\gamma[\pi^s_{11}]$ and $u[\pi^s_{11}]$. Observe that $\gamma[\pi^s_{11}](t) = \gamma(t)$ and $u[\pi^s_{11}](t) = u(t)$ for $t \in [a, t_1 - (l'_1 + l''_1)s)$.

$$A = \lim_{s\to 0} \frac{\int_a^{t_1} f^i(\gamma[\pi^s_{11}](t), u[\pi^s_{11}](t))dt - \int_a^{t_1} f^i(\gamma(t), u(t))dt}{s} =$$

$$= \lim_{s\to 0} \frac{\int_{t_1-(l'_1+l''_1)s}^{t_1} f^i(\gamma[\pi^s_{11}](t), u[\pi^s_{11}](t))dt - \int_{t_1-(l'_1+l''_1)s}^{t_1} f^i(\gamma(t), u(t))dt}{s} =$$

$$= \lim_{s\to 0} \left\{ \frac{\int_{t_1-(l'_1+l''_1)s}^{t_1-l''_1 s} \left[f^i(\gamma[\pi'^s_1](t), u'_1) - f^i(\gamma(t), u(t))\right] dt}{s} + \right.$$

$$\left. + \frac{\int_{t_1-l''_1 s}^{t_1} \left[f^i(\gamma[\pi^s_{11}](t), u''_1) - f^i(\gamma(t), u(t))\right] dt}{s} \right\}$$

As t_1 and $t_1 - l''_1 s$ are Lebesgue times (s is chosen conveniently), Equation (A.15) is used.

$$A = \lim_{s\to 0} \left\{ \frac{f^i(\gamma[\pi'^s_1](t_1 - l''_1 s), u'_1) l'_1 s - f^i(\gamma(t_1 - l''_1 s), u(t_1 - l''_1 s)) l'_1 s}{s} + \right.$$

$$\left. + \frac{f^i(\gamma[\pi^s_{11}](t_1), u''_1) l''_1 s - f^i(\gamma(t_1), u(t_1)) l''_1 s}{s} \right\} =$$

$$= \lim_{s\to 0} \left\{ [f^i(\gamma[\pi'^s_1](t_1 - l''_1 s), u'_1) - f^i(\gamma(t_1 - l''_1 s), u(t_1 - l''_1 s))]\, l'_1 + \right.$$

$$\left. + [f^i(\gamma[\pi^s_{11}](t_1), u''_1) - f^i(\gamma(t_1), u(t_1))]\, l''_1 \right\}$$

As f^i is continuous on $M \times U$, we have

$$A = \left[f^i\left(\lim_{s\to 0} \gamma[\pi'^s_1](t_1 - l''_1 s), u'_1\right) - f^i\left(\lim_{s\to 0} \gamma(t_1 - l''_1 s), \lim_{s\to 0} u(t_1 - l''_1 s)\right)\right] l'_1 +$$

$$+ \left[f^i\left(\lim_{s\to 0} \gamma[\pi^s_{11}](t_1), u''_1\right) - f^i(\gamma(t_1), u(t_1))\right] l''_1 = [f^i(\gamma(t_1), u'_1) - f^i(\gamma(t_1), u(t_1))]\, l'_1 +$$

$$= [f^i(\gamma(t_1), u''_1) - f^i(\gamma(t_1), u(t_1))]\, l''_1 = \mathrm{L}\left(\mathrm{v}[\pi'_1] + \mathrm{v}[\pi''_1]\right)(\mathrm{x^i}).$$

□

Analogous to the linear approximation (3.2), we have

$$\varphi^{t_1}_{\pi_{11}}(s) = \gamma(t_1) + sv[\pi'_1] + sv[\pi''_1] + o(s). \tag{3.3}$$

If we perturb the control n times, $\pi = \{\pi_1, \ldots, \pi_r\}$, with $a < t_1 \leq \ldots \leq t_r < b$, then $\gamma[\pi^s](t)$ is the generalized integral curve of $X^{\{u[\pi^s]\}}$ with initial condition $(a, \gamma(a))$. Consider the curve $\varphi^t_\pi\colon [0,\epsilon] \to M$ for $t \in [t_r, b]$ given by $\varphi^t_\pi(s) = \gamma[\pi^s](t)$.

Corollary 3.9. *The vector tangent to the curve* $\varphi_\pi^t\colon [0,\epsilon] \to M$ *at* $s=0$ *is* $V[\pi_1](t) + \ldots + V[\pi_r](t)$*, where* $V[\pi_i]\colon [t_i, b] \to TM$ *is the generalized integral curve of* $\left(X^T\right)^{\{u\}}$ *with initial condition* $(t_i, (\gamma(t_i), v[\pi_i]))$*.*

It may be easily proved by induction using Propositions 3.4, 3.7, 3.8, where all the possibilities of combination of perturbation data have been studied. If w is the vector tangent to φ_π^t at $s=0$, the perturbation data will be denoted by π_w. Bearing in mind Definition D.2, we have:

Definition 3.10. *The conic non-negative combinations of perturbation vectors of class I and displacements by the flow of* $X^{\{u\}}$ *of perturbation vectors of class I are called* ***perturbation vectors of class II****.*

3.3.3 Perturbation cones

Considering all the elementary perturbation vectors, we define a closed convex cone at every time containing at least all displacements of these vectors. To transport all the elementary perturbation vectors, the pushforward of the flow of the vector field $X^{\{u\}}$ is used. See Appendix B. Observe that the first comment after Definition 3.5 guarantees the fact that the set of elementary perturbation vectors is a cone.

Definition 3.11. *For* $t \in (a,b]$*, the* ***tangent perturbation cone*** K_t *is the smallest closed convex cone in the tangent space to the manifold* M *at the point* $\gamma(t)$ *that contains all the displacements by the flow of* $X^{\{u\}}$ *of all the elementary perturbations vectors from all Lebesgue times* τ *smaller than* t*:*

$$K_t = \overline{\operatorname{conv}\left(\bigcup_{a<\tau<t} (\Phi_{t,\tau}^{X^{\{u\}}})_*(\mathcal{V}_\tau)\right)}, \tag{3.4}$$

where $\mathcal{V}_\tau$ *denotes the set of elementary perturbation vectors at* τ *and* $\operatorname{conv}(A)$ *means the convex hull of the set* A*.*

To prove the following statement, we use results in Appendices D and E; precisely Proposition D.4, D.5 and Corollary E.2.

Proposition 3.12. *Let* $t \in (a,b]$*. If* v *is a nonzero vector in the interior of* K_t*, then there exists* $\epsilon > 0$ *such that for every* $s \in (0,\epsilon)$ *there are* $s' > 0$ *and a perturbation of the control* $u[\pi^s]$ *such that* $\gamma[\pi^s](t) = \gamma(t) + s'v$*.*

Proof. As v is interior to K_t, by Proposition D.5, item d), v is in the interior of the cone $\mathcal{C} = \operatorname{conv}\bigcup_{a<\tau\leq t}\left(\Phi_{(t,\tau)}^{X^{\{u\}}}\right)_* \mathcal{V}_\tau$, where $\mathcal{V}_\tau$ is the cone of elementary perturbation vectors of class I at time τ. Hence, v can be expressed as a convex finite combination of perturbation vectors of class I by Proposition D.4.

Let $(W,(x^i))$ be a local chart of M at $\gamma(t)$. We suppose that the image of the local chart and W are identified locally with $\mathbb{R}^n$. Through the local chart we also identify $T_{\gamma(t)}M$ with the same $\mathbb{R}^n$. We consider the affine hyperplane Π orthogonal to v at the endpoint of the vector v and identify Π with $\mathbb{R}^{n-1}$.

We can choose a "closed" convex cone $\widetilde{\mathcal{C}}$, that is, a closed cone without the vertex, contained in the interior of $\mathcal{C}$ such that v lies in the interior of $\widetilde{\mathcal{C}}$ and $\langle w, v\rangle > 0$ for every $w \in \widetilde{\mathcal{C}}$. For example, we can consider a circular cone with axis v satisfying the two previous conditions, as we will suppose from now on. Hence

$$\Pi \cap \widetilde{\mathcal{C}} = v + \overline{B(0,R)}$$

where $\overline{B(0,R)}$ is the closure of an open ball in the subspace orthogonal to v, denoted by $v^{\perp}$ and identified with $\mathbb{R}^{n-1}$. For $r \in v^{\perp} \simeq \mathbb{R}^{n-1}$, we will write r instead of $0v + r$ as a vector in $\mathbb{R}^n$.

Let us construct a diffeomorphism from the cone $\widetilde{\mathcal{C}}$ to a cylinder of $\mathbb{R}^n$. If $w \in \widetilde{\mathcal{C}}$, the orthogonal decomposition of w induced by v and $v^{\perp}$ is

$$w = \frac{\langle w, v\rangle}{\|v\|}\frac{v}{\|v\|} + \left(w - \frac{\langle w, v\rangle}{\langle v, v\rangle}v\right) = \frac{\langle w, v\rangle}{\langle v, v\rangle}\left[v + \left(\frac{\langle v, v\rangle}{\langle w, v\rangle}w - v\right)\right].$$

Observe that $\frac{\langle v,v\rangle}{\langle w,v\rangle}w - v$ is a vector in $\overline{B(0,R)} \subset v^{\perp}$. Considering the "closed" cone $\widetilde{\mathcal{C}}$ without the vertex, we have the map

$$\begin{array}{rccl} g\colon \widetilde{\mathcal{C}} & \longrightarrow & \mathbb{R}^+ \times \overline{B(0,R)} \\ w & \longmapsto & \left(\frac{\langle w,v\rangle}{\langle v,v\rangle}, \frac{\langle v,v\rangle}{\langle w,v\rangle}w - v\right) = (s,r), \end{array}$$

that is a $\mathcal{C}^{\infty}$ diffeomorphism with inverse given by

$$\begin{array}{rccl} g^{-1}\colon & \mathbb{R}^+ \times \overline{B(0,R)} & \longrightarrow & \widetilde{\mathcal{C}} \\ & (s,r) & \longmapsto & s(v+r) = w. \end{array}$$

Note that g and g^{-1} can be extended to an open cone, without the vertex, containing $\widetilde{\mathcal{C}}$, so the condition of diffeomorphism is clear.

If we truncate $\widetilde{\mathcal{C}}$ by the affine hyperplane Π, we obtain a bounded convex set $\widetilde{\mathcal{C}}_v$. The restriction of g to $\widetilde{\mathcal{C}}_v$ is $g_v\colon \widetilde{\mathcal{C}}_v \to (0,1] \times \overline{B(0,R)}$, that is also a $\mathcal{C}^{\infty}$ diffeomorphism with inverse $g_v^{-1}\colon (0,1] \times \overline{B(0,R)} \to \widetilde{\mathcal{C}}_v$.

If $r \in \overline{B(0,R)}$, then $w_0 = v + r$ is interior to $\mathcal{C}$. Hence, associated to w_0 we have a perturbation π_{w_0} of the control u. Let $\gamma[\pi^s_{w_0}]\colon I \to M$ be the generalized integral curve of $X^{\{u[\pi^s_{w_0}]\}}$ with initial condition $(a, \gamma(a))$ and consider the map

$$\begin{array}{rccl} \Gamma\colon & [0,1] \times \overline{B(0,R)} & \longrightarrow & M \\ & (s,r) & \longmapsto & \Gamma(s,r) = \gamma[\pi^s_{w_0}](t) \\ & (0,r) & \longmapsto & \Gamma(0,r) = \gamma(t). \end{array}$$

which is continuous because $\gamma[\pi^s_{w_0}](t)$ depends continuously on s and $\pi^s_{w_0}$ and $\lim_{(s,r)\to(0,r_0)} \Gamma(s,r) = \gamma(t) = \Gamma(0,r_0)$. Hence, for every $\epsilon > 0$, there exist $\delta_1, \delta_2 > 0$ such that if $|s| < \delta_1$ and $\|r\| < \delta_2$, then $\|\Gamma(s,r) - \Gamma(0,0)\| = \|\gamma[\pi^s_{w_0}](t) - \gamma(t)\| < \epsilon$.

Taking ϵ such that $B(\gamma(t), \epsilon)$ is contained in W, there exist $\delta_1, \delta_2 > 0$ such that if $|s| < \delta_1$ and $\|r\| < \delta_2$, then $\gamma[\pi^s_{w_0}](t) \in W$.

We consider now the map

$$\begin{array}{rccl} \Delta\colon & [0,\delta_1] \times \overline{B(0,\delta_2)} & \longrightarrow & T_{\gamma(t)}M \simeq \mathbb{R}^n \\ & (s,r) & \longmapsto & \Delta(s,r) = \gamma[\pi^s_{w_0}](t) - \gamma(t) \\ & (0,r) & \longmapsto & \Delta(0,r) = 0 \end{array}$$

that is continuous because $\lim_{(s,r)\to(0,r_0)} \Delta(s,r) = 0 = \Delta(0,r_0)$. Remember that we have identified W with $\mathbb{R}^n$ through the local chart. With this in mind and Equation 3.2, we can write

$$\gamma[\pi^s_{w_0}](t) - \gamma(t) = s(v+r) + o_r(s),$$

where $o_r(s) \in \mathbb{R}^n$.

We are going to show that, taking (s,r) in an adequate subset, $\Delta(s,r)$ lies in the interior of the cone $\widetilde{\mathcal{C}}$.

Take a section of the cone through a plane containing v and w, and compute the distance from the endpoint of w to the boundary of the cone $\widetilde{\mathcal{C}}$. It is given by

$$\frac{s\ (R - \|r\|)}{\sqrt{1 + \left(\frac{R}{\|v\|}\right)^2}}.$$

This is the maximum value for the radius of an open ball centered at the endpoint of $s(v+r)$ to be contained in $\widetilde{\mathcal{C}}$.

Define the function

$$\begin{array}{rccl} \Theta\colon & [0,\delta_1] \times \overline{B(0,\delta_2)} & \longrightarrow & \mathbb{R}^n \\ & (s,r) & \longmapsto & \left(\gamma[\pi^s_{w_0}](t) - \gamma(t) - s(v+r)\right)/s \ = \ o_r(s)/s \\ & (0,r) & \longmapsto & 0. \end{array}$$

which is continuous because $\lim_{(s,r)\to(0,r_0)} \Theta(s,r) = 0 = \Theta(0,r_0)$. Take

$$\epsilon = \frac{R - \delta_2}{\sqrt{1 + \left(\frac{R}{\|v\|}\right)^2}},$$

then there exist $\overline{\delta}_1$, $\overline{\delta}_2 > 0$ such that if $|s| < \overline{\delta}_1$ and $\|r\| < \overline{\delta}_2$, then $\|\Theta(s,r)\| = \|o_r(s)/s\| < \epsilon$.

If $(s,r) \in (0,\overline{\delta}_1) \times \overline{B(0,\overline{\delta}_2)}$, then let us show that $\Delta(s,r)$ lies in the interior of the cone $\widetilde{\mathcal{C}}$ without the vertex,

$$\|\Delta(s,r) - s(v+r)\| = \|s(v+r) + o_r(s) - s(v+r)\| = \|o_r(s)\| \le s\epsilon < s\frac{R - \|r\|}{\sqrt{1 + \left(\frac{R}{\|v\|}\right)^2}}$$

since $\|r\| \le \overline{\delta}_2 < \delta_2 < R$.

Thus we conclude that $s(v+r) + o_r(s)$ is in the interior of the cone $\widetilde{\mathcal{C}}$ for every $(s,r) \in (0,\overline{\delta}_1) \times \overline{B(0,\overline{\delta}_2)}$.

Now, for $s \in (0,\overline{\delta}_1)$, we define the continuous mapping

$$\begin{array}{rccl} G_s\colon & \overline{B(0,\overline{\delta}_2)} & \longrightarrow & \overline{B(0,R)} \subset \mathbb{R}^{n-1} \\ & r & \longmapsto & G_s(r) = (\pi_2 \circ g \circ \Delta)\,(s,r), \end{array}$$

where $\pi_2\colon \mathbb{R}^+ \times \overline{B(0,R)} \to \overline{B(0,R)}$, $\pi_2(s,r) = r$. Observe that for $r_0 \in \overline{B(0,\overline{\delta}_2)}$ we have

$$\lim_{(s,r)\to(0,r_0)} G_s(r) = \lim_{(s,r)\to(0,r_0)} \left[\frac{\langle v,v\rangle}{s\langle v,v\rangle + \langle o(s),v\rangle}(s(v+r)+o(s)) - v\right] = r$$

and

$$(g\circ\Delta)(s,r) = g(\gamma[\pi^s_{w_0}](t) - \gamma(t)) = g(s(v+r) + o_r(s)) = (s',r').$$

Suppose that there exists $r \in \overline{B(0,R)}$ such that $G_s(r) = 0$, then applying g^{-1} to the above equation we have

$$\Delta(s,r) = \gamma[\pi^s_{w_0}](t) - \gamma(t) = g^{-1}(s',0) = s'v.$$

Hence, to conclude the proof we need to show that there exists r with $G_s(r) = 0$ for s small enough. To apply Corollary E.2, there must exist $r' \in B(0,\overline{\delta}_2)$ such that $\|G_s(r) - r\| < \|r - r'\|$ for every $r \in \partial\left(\overline{B(0,\overline{\delta}_2)}\right)$. We will show that the condition is fulfilled for $r' = 0$.

Consider the mapping

$$\begin{array}{rccc} \mathcal{G}\colon & [0,\bar{\delta}_1]\times\overline{B(0,\bar{\delta}_2)} & \longrightarrow & \overline{B(0,R)}\subset\mathbb{R}^{n-1} \\ & (s,r) & \longmapsto & \mathcal{G}(s,r)=G_s(r)-r \\ & (0,r) & \longmapsto & \mathcal{G}(0,r)=0. \end{array}$$

For $r_0 \in \overline{B(0,\bar{\delta}_2)}$, we have $\lim_{(s,r)\to(0,r_0)} \mathcal{G}(s,r) = \lim_{(s,r)\to(0,r_0)} G_s(r) - r = 0$. Thus $\mathcal{G}$ is continuous.

Given $r_0 \in \partial\left(\overline{B(0,\bar{\delta}_2)}\right)$, take $\epsilon = \bar{\delta}_2/2$, then there exist $\delta_0(0,r_0)$, $\delta_1(0,r_0) > 0$ such that if $|s| < \delta_0(0,r_0)$ and $\|r-r_0\| < \delta_1(0,r_0)$, then $\|\mathcal{G}(s,r)-\mathcal{G}(0,r_0)\| < \bar{\delta}_2/2$. Hence $\{B(r_0,\delta_1(0,r_0))\,|\,r_0\in\partial\left(\overline{B(0,\bar{\delta}_2)}\right)\}$ is an open covering of $\partial\left(\overline{B(0,\bar{\delta}_2)}\right)$. As this is a compact set, there exists a finite subcovering,

$$\{B(r_1,\delta_1(0,r_1)),\dots,B(r_k,\delta(0,r_k))\}.$$

Take δ as the minimum of $\{\delta_0(0,r_1),\dots,\delta_0(0,r_k)\}$. Let us see that for every $(s,r)\in[0,\delta]\times\partial\left(\overline{B(0,\bar{\delta}_2)}\right)$, $\|G_s(r)-r\|<\|r\|$. As r is in an open set of the finite subcovering

$$\|\mathcal{G}(s,r)\| = \|G_s(r)-r\| < \frac{\bar{\delta}_2}{2} < \bar{\delta}_2 = \|r\|.$$

Hence using Corollary E.2, for every $s\in(0,\delta)$, $G_s(\overline{B(0,\bar{\delta}_2)})$ covers the origin: there exists $r\in\overline{B(0,\bar{\delta}_2)}$ such that

$$G_s(r) = (\pi_2\circ g\circ\Delta)(s,r) = 0,$$

so there exists $s'\in\mathbb{R}^+$ such that

$$\gamma[\pi^s_{w_0}](t) = \gamma(t) + s'v,$$

and we have found a trajectory coming from a perturbation of the control that meets the ray generated by v, as we wanted. □

3.4 Pontryagin's Maximum Principle in the symplectic formalism for the optimal control problem

In this section, the *OCP* is transformed into a hamiltonian problem that will allow us to state Pontryagin's Maximum Principle.

Given the *OCP* (Q,U,X,F,I,x_a,x_b) and the $\widehat{OCP}$ $(\widehat{Q},U,\widehat{X},I,x_a,x_b)$ let us consider the cotangent bundle $T^*\widehat{Q}$ with its natural symplectic structure that will be denoted by ω. If $(\widehat{x},\widehat{p}) = (x^0,x,p_0,p) = (x^0,x^1,\dots,x^n,p_0,p_1,\dots,p_n)$ are local natural coordinates on $T^*\widehat{Q}$, the form ω has as its local expression $\omega = dx^0\wedge dp_0 + dx^i\wedge dp_i$.

For each $u\in U$, $H^u\colon T^*\widehat{Q}\to\mathbb{R}$ is the hamiltonian function defined by

$$H^u(\widehat{p}) = H(\widehat{p},u) = \langle\widehat{p},\widehat{X}(\widehat{x},u)\rangle = p_0F(x,u) + \sum_{i=1}^n p_if^i(x,u),$$

where $\widehat{p}\in T^*_{\widehat{x}}\widehat{Q}$. The tuple $(T^*\widehat{Q},\omega,H^u)$ is a hamiltonian system. The associated hamiltonian vector field, $Y^{\{u\}}$, satisfies the equation

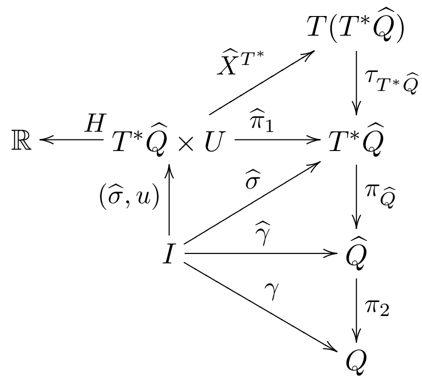
$$i(Y^{\{u\}})\omega = dH^u.$$

Thus we get a family of hamiltonian systems parameterized by u, $H\colon T^*\widehat{Q}\times U\to\mathbb{R}$, and the associated hamiltonian vector field $Y\colon T^*\widehat{Q}\times U\to T(T^*\widehat{Q})$ which is a vector field along the projection $\widehat{\pi}_1\colon T^*\widehat{Q}\times U\to T^*\widehat{Q}$. Its local expression is

$$Y(\widehat{p},u)=\left(F(x,u)\frac{\partial}{\partial x^0}+f^i(x,u)\frac{\partial}{\partial x^i}+0\frac{\partial}{\partial p_0}+\left(-p_0\frac{\partial F}{\partial x^i}(x,u)-p_j\frac{\partial f^j}{\partial x^i}(x,u)\right)\frac{\partial}{\partial p_i}\right)_{(\widehat{x},\widehat{p},u)}$$

for $i,j=1,\ldots,n$. It should be noted that $Y=\widehat{X}^{T^*}$ is the cotangent lift of $\widehat{X}$. See Appendix B.3 for definition and properties of the cotangent lift.

Given a curve $(\widehat{\sigma},u)\colon I\to T^*\widehat{Q}\times U$ with $\widehat{\gamma}=\pi_{\widehat{Q}}\circ\widehat{\sigma}$, if $\pi_{\widehat{Q}}\colon T^*\widehat{Q}\to\widehat{Q}$ is the natural projection. The previous elements come together in the following diagram:



Statement 3.13. ***(Hamiltonian Problem,*** HP***)*** *Given the OCP* (Q,U,X,F,I,x_a,x_b)*, and the equivalent* $\widehat{OCP}$ $(\widehat{Q},U,\widehat{X},I,x_a,x_b)$*, consider the following problem.*

Find $(\widehat{\sigma}^*,u^*)$ *such that*

(1) $\widehat{\gamma}^*(a)=(0,x_a)$ *and* $\gamma^*(b)=x_b$*, if* $\widehat{\gamma}^*=\pi_{\widehat{Q}}\circ\widehat{\sigma}^*$*,* $\gamma^*=\pi_2\circ\widehat{\gamma}^*$*.*

(2) $\dot{\widehat{\sigma}}^*(t)=\widehat{X}^{T^*}(\widehat{\sigma}^*(t),u^*(t))$*,* $t\in I$*.*

The tuple $(T^*\widehat{Q},U,\widehat{X}^{T^*},I,x_a,x_b)$ denotes the *hamiltonian problem* as it has just been defined and the elements satisfy the same properties as in §2.

Comments: The hamiltonian problem satisfies analogous conditions to those satisfied by the OCP and the $\widehat{OCP}$ defined in §3.1 and §3.2 respectively.

1. Given $(\widehat{\sigma},u)$, the function $u\colon I\to U$ allows us to construct a time-dependent vector field on $T^*\widehat{Q}$, $(\widehat{X}^{T^*})^{\{u\}}\colon T^*\widehat{Q}\times I\to T(T^*\widehat{Q})$, defined by

   $$(\widehat{X}^{T^*})^{\{u\}}(\widehat{x},\widehat{p},t)=\widehat{X}^{T^*}(\widehat{x},\widehat{p},u(t)).$$

   Condition (2) shows that $\widehat{\sigma}^*$ is an integral curve of $(\widehat{X}^{T^*})^{\{u^*\}}$.

2. Condition (2) is equivalent to the commutativity of Diagram 2.1 with $M=T^*\widehat{Q}$ and the vector field $(\widehat{X}^{T^*})^{\{u^*\}}$ for $(\widehat{\sigma}^*,u^*)$.

3. The vector field $(\widehat{X}^{T^*})^{\{u\}}$ is $\pi_{\widehat{Q}}$−projectable and projects onto $\widehat{X}^{\{u\}}$. Thus if $\widehat{\sigma}$ is an integral curve of $(\widehat{X}^{T^*})^{\{u\}}$, $\widehat{\gamma}=\pi_{\widehat{Q}}\circ\widehat{\sigma}$ is an integral curve of $\widehat{X}^{\{u\}}$.

4. Locally the curve $(\widehat{\sigma}^*, u^*)$ satisfies the Hamilton equations of the system $(T^*\widehat{Q}, \omega, H^u)$,

$$\begin{aligned}
\dot{x}^0 &= \frac{\partial H^u}{\partial p_0} = F \\
\dot{x}^i &= \frac{\partial H^u}{\partial p_i} = f^i \\
\dot{p}_0 &= -\frac{\partial H^u}{\partial x^0} = 0 \Rightarrow p_0 = \mathrm{ct} \qquad (3.5)\\
\dot{p}_i &= -\frac{\partial H^u}{\partial x^i} = -p_0\frac{\partial F}{\partial x^i} - p_j\frac{\partial f^j}{\partial x^i}, \qquad (3.6)
\end{aligned}$$

   and satisfies the conditions $\widehat{\gamma}(a) = (0, x_a)$, $\gamma(b) = x_b$.

   In the literature of optimal control, the system of differential equations given by Equations (3.5), (3.6) is called the *adjoint system*. In differential geometry, the adjoint system is the differential equations satisfied by the fiber coordinates of an integral curve of the cotangent lift of a vector field on Q. See Appendix B.3 for more details.

   Note that there is no initial condition for $\widehat{p} = (p_0, p_1, \ldots, p_n)$, hence HP is not a Cauchy problem.

Comment: So far we have considered a fixed control $u \in U$. Therefore we have been working with a family of hamiltonian systems on the manifold $(T^*\widehat{Q}, \omega)$ given by the Hamiltonians $\{H^u |\, u \in U\}$.

Given $u\colon I \to U$, then we consider the Hamiltonian $H^{u(t)}$. The equation of the hamiltonian vector field for the hamiltonian system $(T^*\widehat{Q}, \omega, H^{u(t)})$ is

$$i(Y^{\{u(t)\}})\omega = d_{\widehat{Q}} H^{u(t)},$$

where $d_{\widehat{Q}}$ is the exterior differential on the manifold $T^*\widehat{Q}$. Observe that we have studied the system defined by $(T^*\widehat{Q}, \omega, H^{u(t)})$ as an autonomous system by fixing the time t. The hamiltonian vector field obtained, $Y^{\{u(t)\}}$, is a time-dependent vector field whose integral curves satisfy the equation

$$\dot{\widehat{\sigma}}(t) = Y^{\{u(t)\}}(\widehat{\sigma}(t)), \quad t \in I. \qquad (3.7)$$

Observe that $Y^{\{u(t)\}} = (\widehat{X}^{T^*})^{\{u(t)\}}$.

Now we are ready to state Pontryagin's Maximum Principle that provides the necessary conditions, which are in general not sufficient, to find solutions of the optimal control problem.

Theorem 3.14. ***(Pontryagin's Maximum Principle, PMP)***
Let $(\widehat{\gamma}^*, u^*)\colon I \to \widehat{Q} \times U$ be a solution of the extended optimal control problem, Statement 3.2. Then there exists $(\widehat{\sigma}^*, u^*)\colon I \to T^*\widehat{Q} \times U$ such that:

1. *it is a solution of the hamiltonian problem, that is, it satisfies Equation (3.7) and the initial conditions $\widehat{\gamma}^*(a) = (0, x_a)$ and $\gamma^*(b) = x_b$, if $\gamma^* = \pi_2 \circ \widehat{\gamma}^*$;*

2. *$\widehat{\gamma}^* = \pi_{\widehat{Q}} \circ \widehat{\sigma}^*$, with fiber $\widehat{\alpha}^*(t) = (\alpha_0^*(t), \alpha^*(t)) \in T^*_{\widehat{\gamma}^*(t)}\widehat{Q}$;*

3. (a) *$H(\widehat{\sigma}^*(t), u^*(t)) = \max_{u \in \overline{U}} H(\widehat{\sigma}^*(t), u)$ almost everywhere;*
   (b) *$\max_{u \in \overline{U}} H(\widehat{\sigma}^*(t), u)$ is constant everywhere;*
   (c) *$\widehat{\alpha}^*(t) \neq 0 \in T^*_{\widehat{\gamma}^*(t)}\widehat{Q}$ for each $t \in [a, b]$;*
   (d) *$\alpha_0^*(t)$ is constant and $\alpha_0^*(t) \leq 0$.*

Comments:

1. Condition (2) is immediately satisfied because $\widehat{\sigma}^*$ is a covector along $\widehat{\gamma}^*$.

2. Conditions ($3a$) and ($3b$) imply that the hamiltonian function is constant almost everywhere for $t \in [a, b]$.

3. In item ($3a$), if U is not a closed set, then either we will write the maximum over the controls in the closure of U or the supremum over the controls in U. But in condition ($3b$) we can consider the maximum because item ($3a$) guarantees that the supremum of the Hamiltonian is reached in the optimal curve.

4. From the Hamilton equations of the system $(T^*\widehat{Q}, \omega, H^{u(t)})$, it is concluded that p_0 is constant along the integral curves of $(\widehat{X}^{T^*})^{\{u(t)\}}$, since $\dot{p_0} = 0$. Hence the first result in (d) is immediate for every integral curve of $(\widehat{X}^{T^*})^{\{u(t)\}}$.

5. Condition ($3c$) implies that $\alpha_0^*(t) \neq 0$ or $\alpha^*(t) \neq 0 \in T^*_{\gamma^*(t)}Q$ for each $t \in [a, b]$. Locally the condition (c) states that for each $t \in [a, b]$ there exists a coordinate of $\widehat{\alpha}^*(t)$ nonzero, $(p_i \circ \widehat{\alpha}^*)(t) = \alpha_i^*(t) \neq 0$.

6. Pontryagin's Maximum Principle only guarantees that given a solution of $\widehat{OCP}$ there exists a solution of HP. Hence, in principle, both problems are not equivalent.

Observe that Maximum Principle guarantees the existence of a covector along the optimal curve, but it does not say anything about the uniqueness of the covector. Indeed, this covector may not be unique. Depending on the covector we associate with the optimal curves, different kind of curves can be defined.

Definition 3.15. *A curve* $(\widehat{\gamma}, u)\colon [a, b] \to \widehat{Q} \times U$ *for* $\widehat{OCP}$ *is*

1. *an* ***extremal*** *if there exist* $\widehat{\sigma}\colon [a, b] \to T^*\widehat{Q}$ *such that* $\widehat{\gamma} = \pi_{T\widehat{Q}} \circ \widehat{\sigma}$ *and* $(\widehat{\sigma}, u)$ *satisfies the necessary conditions of PMP;*

2. *a* ***normal extremal*** *if it is an extremal with* $\alpha_0 < 0$*;*

3. *an* ***abnormal extremal*** *if it is an extremal with* $\alpha_0 = 0$*;*

4. *a* ***strictly abnormal extremal*** *if it is not a normal extremal, but it is abnormal;*

5. *a* ***strictly normal extremal*** *if it is not a abnormal extremal, but it is normal.*

In [1, 34] there are some examples of optimal control problems whose solutions are searched using Pontryagin's Maximum Principle.

Before proceeding, let us point out the new contribution of Maximum Principle: the initial condition for the fibers in $T^*\widehat{Q}$ to solve the Hamiltonian Problem must be chosen conveniently in order to get the maximization of the Hamiltonian, as shows the proof of Theorem 3.14, see §4.

Observe that if $\widehat{\gamma}\colon I \to \widehat{Q}$ is an integral curve of a vector field, there always exists a lift of $\widehat{\gamma}$ to a curve $\widehat{\sigma}\colon I \to T^*\widehat{Q}$, given an initial condition for the cofibers, which is an integral curve of the cotangent lift of the given vector field on $\widehat{Q}$. Analogously, if the system is given by a vector field along the projection $\widehat{\pi}\colon \widehat{Q} \times U \to \widehat{Q}$.

For abnormal extremals, $p_0 = 0$, the Hamiltonian is $H(\widehat{p}, u) = \langle p, X(x, u)\rangle$, for $\widehat{p} \in T^*_{\widehat{x}}\widehat{Q}$. Then σ is an integral curve of $(X^{T^*})^{\{u\}}$, the cotangent lift of $X^{\{u\}}$, once initial conditions for the fibers are given. Thus in the abnormality case the hamiltonian problem can be restricted to Q instead of $\widehat{Q}$, that is, it is not necessary to consider the extended system to associate a lift to the abnormal extremal.

For normal extremals, the Hamiltonian is $H(\widehat{p}, u) = \langle p, X(x, u)\rangle + p_0 F(x, u)$. Then $\widehat{\sigma}$ is an integral curve of $(\widehat{X}^{T^*})^{\{u\}}$, the cotangent lift of $\widehat{X}$, once initial conditions for the fibers are given. In contrast to the lift of abnormal extremals, now we must use the extended system.

Therefore, the items 1 and 2 in Theorem 3.14 do not say anything new except for the fact that a final condition must be also satisfied. The accomplishment of this depends on the accessibility of the problem, see [11, 27].

The real contribution of PMP is the third item related mostly with the maximization of the Hamiltonian, that will be only satisfied if the initial conditions for the fibers are chosen suitably. This is the key point of the proof of Pontryagin's Maximum Principle. In other words, we can always find a cotangent lift of an integral curve, but it is not guaranteed the fulfilment of conditions in item 3 in Theorem 3.14.

Previously, we have written the hamiltonian function for the abnormal and the normal case. The difference is that the cost function does not play any role in the Hamiltonian for abnormal extremals. That is why it is said the abnormal extremals only depend on the geometry of the control system. But to determine the optimality of the abnormal extremals the cost function is essential. In fact, for the same control system different optimal control problems can be stated depending on the cost function, in such a way that the abnormal extremals are minimizers only in some of them.

To conclude, the strict abnormality characterizes the abnormal extremals that are not normal. An extremal is not normal when there does not exist any covector that satisfies Hamilton's equations for normality, thus it is necessary to the cost function.

4 Proof of Pontryagin's Maximum Principle for fixed time and fixed endpoints

Previous comments: To prove Pontryagin's Maximum Principle it is necessary to use analytic results about absolute continuity and lower semicontinuity for real functions, and properties of convex cones. For the details see Appendix A and Appendix D and references therein, respectively.

The reader is referred to §3.3 and §5.2 for results on perturbations of a trajectory in a dynamical system with controls and constructions obtained from them.

In the literature of optimal control, the proof of the Maximum Principle has been discussed taking into account different hypotheses, [1, 2, 8, 17, 31, 32, 33]. Most authors believe and justify that the origin of this Principle is the calculus of variations, see [37] for instance.

Proof. (Theorem 3.14: Pontryagin's Maximum Principle, PMP)
1. As $(\widehat{\gamma}^*, u^*)$ a solution of $\widehat{OCP}$, if τ is in $[a, b]$, for every initial condition $\widehat{\alpha}_\tau$ in $T^*_{\widehat{\gamma}^*(\tau)}\widehat{Q}$, we have a solution of HP, $(\widehat{\gamma}^*, \widehat{\alpha})\colon [a, b] \to T^*\widehat{Q}$, satisfying that initial condition. The covector $\widehat{\alpha}_\tau$ must be chosen conveniently so that the remaining conditions of the PMP are satisfied.

According to §3.3, we construct a tangent perturbation cone $\widehat{K}_b$ in $T_{\widehat{\gamma}^*(b)}\widehat{Q}$ that contains all tangent vectors associated with perturbations of the trajectory $\widehat{\gamma}^*$ corresponding to variations of u^*; see Definition 3.11.

Let us consider the vector $(-1, \mathbf{0})_{\widehat{\gamma}^*(b)} \in T_{\widehat{\gamma}^*(b)}\widehat{Q}$, satisfying the following properties:

1. The variation of $x^0(t) = \int_a^t F(\gamma^*(s), u^*(s))ds$ along $(-1, \mathbf{0})$ is negative.

2. It is not interior to $\widehat{K}_b$.

   Let us prove this second item. Take a local chart at $\widehat{\gamma}^*(b)$ and work on the image of the local chart, in $\mathbb{R}^{n+1}$, without changing the notation.

   If $(-1, \mathbf{0})_{\widehat{\gamma}^*(b)}$ was interior to $\widehat{K}_b$, by Proposition 3.12 there would exist a positive number ϵ such that for every $s \in (0, \epsilon)$ there would exist a positive number s', close to s, and a perturbation of the control $u[\pi^s]$ such that

   $$\widehat{\gamma}[\pi^s](b) = (\gamma^0[\pi^s](b), \gamma[\pi^s](b)) = \widehat{\gamma}^*(b) + s'(-1, \mathbf{0}).$$

   For this perturbed trajectory we have

   $$\gamma^0[\pi^s](b) < {\gamma^*}^0(b) \quad \text{and} \quad \gamma[\pi^s](b) = \gamma^*(b).$$

   Hence there would be a trajectory, $\widehat{\gamma}[\pi^s]$, from $\gamma^*(a)$ to $\gamma^*(b)$ with less cost than $\widehat{\gamma}^*$. Hence $\widehat{\gamma}^*$ would not be optimal. In other words, $(-1, \mathbf{0})_{\widehat{\gamma}^*(b)}$ is the direction of decreasing of the functional to be minimized in the extended optimal control problem.

The second property implies that $\widehat{K}_b$ cannot be equal to $T_{\widehat{\gamma}^*(b)}\widehat{Q}$. As $(-1, \mathbf{0})_{\widehat{\gamma}^*(b)}$ is not interior to $\widehat{K}_b$, there exist separating hyperplanes of $\widehat{K}_b$ and $(-1, \mathbf{0})_{\widehat{\gamma}^*(b)}$ by Proposition D.15; that is, there exists a nonzero covector determining a separating hyperplane. Let $\widehat{\alpha}_b \in T^*_{\widehat{\gamma}^*(b)}\widehat{Q}$ be nonzero such that $\operatorname{Ker}\widehat{\alpha}_b$ is one of those separating hyperplanes satisfying

$$\langle \widehat{\alpha}_b, (-1, \mathbf{0}) \rangle \geq 0,$$
$$\langle \widehat{\alpha}_b, \widehat{v}_b \rangle \leq 0 \quad \forall \widehat{v}_b \in \widehat{K}_b.$$

Observe that if $\widehat{\alpha}_b = 0 \in T^*_{\widehat{\gamma}^*(b)}\widehat{Q}$, $\operatorname{Ker}\widehat{\alpha}_b$ does not determine a hyperplane, but the whole space $T^*_{\widehat{\gamma}^*(b)}\widehat{Q}$.

Given the initial condition $\widehat{\alpha}_b \in T^*_{\widehat{\gamma}^*(b)}\widehat{Q}$, there exists only one integral curve $\widehat{\sigma}^* = (\widehat{\gamma}^*, \widehat{\alpha}^*)$ of $(\widehat{X}^{T^*})^{\{u^*\}}$ such that $\widehat{\sigma}^*(b) = (\widehat{\gamma}^*(b), \widehat{\alpha}_b)$. Hence $(\widehat{\sigma}^*, u^*)$ is a solution of HP.

2. Obviously, by construction, $\widehat{\gamma}^* = \pi_{\widehat{Q}} \circ \widehat{\sigma}^*$.

Now we prove that $\widehat{\sigma}^*$, the solution of HP, satisfies the remaining conditions of the PMP.

(3.a) $H(\widehat{\sigma}^*(t), u^*(t)) = \max_{u \in \overline{U}} H(\widehat{\sigma}^*(t), u)$ almost everywhere.

We are going to prove the statement for every Lebesgue time, hence it will be true almost everywhere. Suppose that there exists a control $\widetilde{u} \colon I \to \overline{U}$ and a Lebesgue time t_1 such that u^* does not maximize the Hamiltonian at t_1; that is,

$$H(\widehat{\sigma}^*(t_1), \widetilde{u}(t_1)) > H(\widehat{\sigma}^*(t_1), u^*(t_1)).$$

As $H(\widehat{p}, u) = \langle \widehat{p}, \widehat{X}(\widehat{x}, u) \rangle$,

$$\langle \widehat{\alpha}^*(t_1), \widehat{X}(\widehat{\gamma}^*(t_1), \widetilde{u}(t_1)) - \widehat{X}(\widehat{\gamma}^*(t_1), u^*(t_1)) \rangle > 0,$$

that is, $\langle \widehat{\alpha}^*(t_1), \widehat{v}[\pi_1] \rangle > 0$ where $\widehat{v}[\pi_1] = \widehat{X}(\widehat{\gamma}^*(t_1), \widetilde{u}(t_1)) - \widehat{X}(\widehat{\gamma}^*(t_1), u^*(t_1)) \in \widehat{K}_{t_1} \subset T_{\widehat{\gamma}^*(t_1)}\widehat{Q}$ is the elementary perturbation vector associated with the perturbation data $\pi_1 = \{t_1, 1, \widetilde{u}(t_1)\}$ by Proposition 3.4.

Let $\widehat{V}[\pi_1]\colon [t_1,b] \to T\widehat{Q}$ be the integral curve of $(\widehat{X}^T)^{\{u^*\}}$ with $(t_1, \widehat{\gamma}^*(t_1), \widehat{v}[\pi_1])$ as initial condition. For $\widehat{\sigma}^*$, solution of HP, the continuous function $\langle \widehat{\alpha}^*, \widehat{V}[\pi_1]\rangle\colon [t_1,b] \to \mathbb{R}$ is constant everywhere by Proposition B.3. Hence $\langle \widehat{\alpha}^*(t_1), \widehat{v}[\pi_1]\rangle > 0$ implies that $\langle \widehat{\alpha}_b, \widehat{V}[\pi_1](b)\rangle > 0$, which is a contradiction with $\langle \widehat{\alpha}_b, \widehat{v}_b\rangle \leq 0$ for every $\widehat{v}_b \in \widehat{K}_b$, since $\widehat{V}[\pi_1](b) \in \widehat{K}_b$.

Therefore

$$H(\widehat{\sigma}^*(t), u^*(t)) = \max_{u\in\overline{U}} H(\widehat{\sigma}^*(t), u)$$

at every Lebesgue time on $[a,b]$, so almost everywhere.

(3.b) $\max_{u\in\overline{U}} H(\widehat{\sigma}^*(t), u)$ is constant everywhere.

To simplify the notation we define the function

$$\begin{array}{rcl} \mathcal{M}\circ\widehat{\sigma}^*\colon I & \longrightarrow & \mathbb{R} \\ t & \longmapsto & \mathcal{M}(\widehat{\sigma}^*(t)) = \max_{u\in\overline{U}} H(\widehat{\sigma}^*(t), u). \end{array}$$

In order to prove $(3.b)$, it is enough to see that $\mathcal{M}(\widehat{\sigma}^*(t))$ is constant everywhere.

First let us see that $\mathcal{M}\circ\widehat{\sigma}^*$ is lower semicontinuous on I. See Appendix A for details of this property. As $\mathcal{M}(\widehat{\sigma}^*(t))$ is the maximum of the hamiltonian function over the controls, for every $\epsilon > 0$, there exists a control $u_{\mathcal{M}}\colon I \to \overline{U}$ such that

$$H(\widehat{\sigma}^*(t), u_{\mathcal{M}}(t)) \geq \mathcal{M}(\widehat{\sigma}^*(t)) - \frac{\epsilon}{2} \tag{4.8}$$

everywhere.

For each constant control $\widetilde{u} \in U$, $H^{\widetilde{u}} \circ \widehat{\sigma}^* = H(\widehat{\sigma}^*, \widetilde{u})\colon I \to \mathbb{R}$ is continuous on I. Hence for every $t_0 \in I$ and $\epsilon > 0$, there exists $\delta > 0$ such that $\mid t - t_0 \mid < \delta$, we have

$$\mid H^{\widetilde{u}}(\widehat{\sigma}^*(t)) - H^{\widetilde{u}}(\widehat{\sigma}^*(t_0)) \mid < \frac{\epsilon}{2}.$$

If $\widetilde{u} = u_{\mathcal{M}}(t_0)$, then using the continuity of $H^{\widetilde{u}} \circ \widehat{\sigma}^*$

$$\mathcal{M}(\widehat{\sigma}^*(t)) = \max_{u\in\overline{U}} H(\widehat{\sigma}^*(t), u) \geq H(\widehat{\sigma}^*(t), u_{\mathcal{M}}(t_0)) \geq$$

$$\geq H(\widehat{\sigma}^*(t_0), u_{\mathcal{M}}(t_0)) - \frac{\epsilon}{2} \geq \mathcal{M}(\widehat{\sigma}^*(t_0)) - \epsilon.$$

The last inequality is true evaluating Equation (4.8) at t_0. Hence $\mathcal{M}\circ\widehat{\sigma}^*$ is lower semicontinuous at every $t_0 \in I$, that is, $\mathcal{M}\circ\widehat{\sigma}^*$ is lower semicontinuous on I.

The control u^* is bounded, so $\operatorname{Im} u^*$ is contained in a compact set $D \subset \mathbb{R}^m$. Let us define a new function

$$\begin{array}{rcl} m\colon T^*\widehat{Q} & \longrightarrow & \mathbb{R} \\ \beta & \longmapsto & m(\beta) = \max_{\widetilde{u}\in D\cap\overline{U}} H(\beta, \widetilde{u}). \end{array}$$

As $H(\beta, \cdot)\colon D\cap\overline{U} \to \mathbb{R}$, $\widetilde{u} \mapsto H(\beta, \widetilde{u})$ is continuous by hypothesis and $D\cap\overline{U}$ is compact, for every $\beta \in T^*\widehat{Q}$ there exists a control $\widetilde{w}_\beta$ that gives us the maximum of $H(\beta, \widetilde{u})$

$$m(\beta) = \max_{\widetilde{u}\in D\cap\overline{U}} H(\beta, \widetilde{u}) = H(\beta, \widetilde{w}_\beta). \tag{4.9}$$

Hence m is well-defined on $T^*\widehat{Q}$. The following sketch explains in a compact way the necessary steps to prove that $\mathcal{M}\circ\sigma^*$ is constant everywhere. In this sketch, the figures refer to statements which are going to be proved in the next paragraphs and a.c. stands for absolutely continuous and a.e. for almost everywhere.

$$
\left.
\begin{array}{l}
\qquad\qquad\qquad\qquad\qquad\qquad \mathcal{M}\circ\widehat{\sigma}^* \text{ is lower semicontinuous on } I \\
\begin{array}{l}
H^{\widetilde{u}} \in \mathcal{C}^1(T^*\widehat{Q}) \\
\qquad \Downarrow_1 \\
H^{\widetilde{u}} \text{ is locally Lipschitz } \forall\, \widetilde{u}\in D \\
\qquad \Downarrow_2
\end{array} \\
\left.\begin{array}{l}
m \text{ is locally Lipschitz on } \mathrm{Im}\widehat{\sigma}^* \\
\\
\widehat{\sigma}^* \text{ is a.c.}
\end{array}\right\} \Rightarrow^3 m\circ\widehat{\sigma}^* \text{ is a.c.} \Rightarrow m\circ\widehat{\sigma}^* \text{ is continuous} \\
\qquad\qquad\qquad\qquad\qquad\qquad {}^4\, m(\widehat{\sigma}^*(t)) \le \mathcal{M}(\widehat{\sigma}^*(t)),\ \forall\, t\in[a,b] \\
\qquad\qquad\qquad\qquad\qquad\qquad {}^5\, m(\widehat{\sigma}^*(t)) = \mathcal{M}(\widehat{\sigma}^*(t)) \text{ a.e.}
\end{array}
\right\} \Rightarrow^6
$$

$$
\left.
\begin{array}{l}
m\circ\widehat{\sigma}^* \text{ is a.c.} \\
{}^7\, m\circ\widehat{\sigma}^* \text{ has zero derivative}
\end{array}
\right\} \Rightarrow^8
$$

$$
\left.
\begin{array}{l}
\Rightarrow^6\ \text{(A. 15)}\, m(\widehat{\sigma}^*(t)) = \mathcal{M}(\widehat{\sigma}^*(t))\ \forall\, t\in[a,b] \\
\Rightarrow^8\ \text{(A. 16)}\, m(\widehat{\sigma}^*(t)) \text{ is constant } \forall\, t\in[a,b]
\end{array}
\right\} \Rightarrow^9 \mathcal{M}(\widehat{\sigma}^*(t)) \text{ is constant } \forall\, t\in[a,b]
$$

1. $H^{\widetilde{u}} \in \mathcal{C}^1(T^*\widehat{Q}) \Rightarrow H^{\widetilde{u}}$ is locally Lipschitz $\forall\, \widetilde{u}\in D$.

The Lipschitzian property applies to functions defined on a metric space. As the property we want to prove is local, we define the distance on a local chart as is explained in Appendix A. For every $\beta\in T^*\widehat{Q}$, let (V_β,ϕ) be a local chart centered at β such that $\phi(\beta)=0$ and $\phi(V_\beta)=B$, where B is an open ball centered at $0\in\mathbb{R}^{2n+2}$. If β_1 and β_2 are in V_β, define $d_\phi(\beta_1,\beta_2)=d(\phi(\beta_1),\phi(\beta_2))$ where d is the euclidean distance in $\mathbb{R}^{2n+2}$.

For every β in $T^*\widehat{Q}$, we get an open neighbourhood V_β using the local chart (V_β,ϕ). As $H^{\widetilde{u}}$ is $\mathcal{C}^1(T^*\widehat{Q})$ and $\widetilde{u}$ lies in the compact set D, by the Mean Value Theorem for every β in $T^*\widehat{Q}$ there exists an open neighbourhood V_β such that $|H^{\widetilde{u}}(\beta_1)-H^{\widetilde{u}}(\beta_2)| < K_\beta d_\phi(\beta_1,\beta_2)$ where K_β does not depend on the control $\widetilde{u}$. Thus $H^{\widetilde{u}}$ is locally Lipschitz on $T^*\widehat{Q}$. Moreover, the Lipschitz constant and the open neighbourhood V_β do not depend on the control since $\widetilde{u}$ is in a compact set.

2. $H^{\widetilde{u}}$ is locally Lipschitz $\forall\,\widetilde{u}\in D \Rightarrow m$ is locally Lipschitz on $\mathrm{Im}\,\widehat{\sigma}^*$.

Let β be in $\mathrm{Im}\,\widehat{\sigma}^*$, there exists an open convex neighbourhood V_β such that $|H^{\widetilde{u}}(\beta_1) - H^{\widetilde{u}}(\beta_2)| < K_\beta d(\beta_1,\beta_2)$ for every $\widetilde{u}$ in D and β_1, β_2 in V_β. If $\widetilde{w}_1$, $\widetilde{w}_2$ are the controls in $D\cap\overline{U}$ maximizing $H(\beta_1,\widetilde{u})$ and $H(\beta_2,\widetilde{u})$ respectively, then

$$H(\beta_1,\widetilde{w}_2)\le H(\beta_1,\widetilde{w}_1),$$

$$H(\beta_2,\widetilde{w}_1)\le H(\beta_2,\widetilde{w}_2).$$

Moreover, $H^{\widetilde{w}_1}$ and $H^{\widetilde{w}_2}$ are Lipschitz on V_β since the Lipschitz constant and the neighbourhood is independent of the control. Then using the last inequalities

$$-K_\beta d(\beta_1,\beta_2)\le H^{\widetilde{w}_2}(\beta_1)-H^{\widetilde{w}_2}(\beta_2)\le H^{\widetilde{w}_1}(\beta_1)-H^{\widetilde{w}_2}(\beta_2)\le H^{\widetilde{w}_1}(\beta_1)-H^{\widetilde{w}_1}(\beta_2)\le K_\beta d(\beta_1,\beta_2).$$

Observe that by Equation (4.9), $H^{\widetilde{w}_1}(\beta_1)-H^{\widetilde{w}_2}(\beta_2)=m(\beta_1)-m(\beta_2)$. Hence

$$|m(\beta_1)-m(\beta_2)|\le K_\beta d(\beta_1,\beta_2),\quad \forall\,\beta_1,\beta_2\in V_\beta, \tag{4.10}$$

that is, m is locally Lipschitz on $\mathrm{Im}\,\widehat{\sigma}^*$. As $\widehat{\sigma}^*$ is absolutely continuous, $\mathrm{Im}\,\widehat{\sigma}^*$ is compact. Thus we may choose a Lipschitz constant independent of the point β. Hence

$$|m(\beta_1)-m(\beta_2)|\le K d(\beta_1,\beta_2),\quad \forall\,\beta_1,\beta_2\in V_\beta.$$

3. m is locally Lipschitz on $\operatorname{Im}\widehat{\sigma}^*$ and $\widehat{\sigma}^*$ is absolutely continuous $\Rightarrow m\circ\widehat{\sigma}^*\colon I\to\mathbb{R}$ is absolutely continuous $\Rightarrow m\circ\widehat{\sigma}^*\colon I\to\mathbb{R}$ is continuous.

For every $t\in I$, let us consider the neighbourhood $V_{\widehat{\sigma}^*(t)}$ where Equation (4.10) is satisfied. As $\operatorname{Im}\widehat{\sigma}^*$ is a compact set,

- there exists a finite open subcovering $V_{\widehat{\sigma}^*(t_1)},\dots,V_{\widehat{\sigma}^*(t_r)}$ of $\{V_{\widehat{\sigma}^*(t)}\ ;\ t\in I\}$, and
- there exists a Lebesgue number l of the subcovering, that is, for every two points in an open ball of diameter l there exists an open set of the finite subcovering containing both points.

For the Lebesgue number l, by the uniform continuity of $\widehat{\sigma}^*$, there exists a $\delta_l>0$ such that for each t_1, t_2 in I with $|t_2-t_1|<\delta_l$, then $d(\widehat{\sigma}^*(t_2),\widehat{\sigma}^*(t_1))<l$. Thus there exists an open set of the finite subcovering containing $\widehat{\sigma}^*(t_1)$ and $\widehat{\sigma}^*(t_2)$.

On the other hand, taken $\epsilon>0$ the absolutely continuity of $\widehat{\sigma}^*$ determines a $\delta_\epsilon>0$.

To prove the absolute continuity of $m\circ\widehat{\sigma}^*$, take $\delta=\min\{\delta_l,\delta_\epsilon\}$, then for every finite number of nonoverlapping subintervals (t_{i_1},t_{i_2}) of I, with $\sum_{i=1}^n|t_{i_2}-t_{i_1}|<\delta$,

$$\sum_{i=1}^k|m(\widehat{\sigma}^*(t_{i_2}))-m(\widehat{\sigma}^*(t_{i_1}))|\leq\sum_{i=1}^k Kd(\widehat{\sigma}^*(t_{i_2}),\widehat{\sigma}^*(t_{i_1}))\leq K\epsilon.$$

In the first step we use that $\delta<\delta_l$ to guarantee that $\widehat{\sigma}^*(t_{i_2})$ and $\widehat{\sigma}^*(t_{i_1})$ are contained in the same open set of the finite subcovering of $\operatorname{Im}\widehat{\sigma}^*$. That allows us to use the property of being locally Lipschitzian. Secondly, we use that $\delta<\delta_\epsilon$ to apply the absolute continuity of $\widehat{\sigma}^*$.

As $m\circ\widehat{\sigma}^*$ is absolutely continuous on I, $m\circ\widehat{\sigma}^*$ is continuous on I.

4. $m(\widehat{\sigma}^*(t))\leq\mathcal{M}(\widehat{\sigma}^*(t))$ everywhere.

Observe that

$$m(\widehat{\sigma}^*(t))=\max_{u\in D\cap\overline{U}}H(\widehat{\sigma}^*(t),u)\leq\max_{u\in\overline{U}}H(\widehat{\sigma}^*(t),u)=\mathcal{M}(\widehat{\sigma}^*(t)),$$

for each $t\in I$.

5. $\mathcal{M}(\widehat{\sigma}^*(t))=m(\widehat{\sigma}^*(t))$ almost everywhere.

For each $t\in I$ there exists a control $w(t)$ maximizing $H(\widehat{\sigma}^*(t),u)$ over the controls in $D\cap\overline{U}$,

$$m(\widehat{\sigma}^*(t))=\max_{u\in D\cap\overline{U}}H(\widehat{\sigma}^*(t),u)=H(\widehat{\sigma}^*(t),w(t)).$$

As $u^*(t)\in D\cap\overline{U}$ for each $t\in I$,

$$\max_{u\in D\cap\overline{U}}H(\widehat{\sigma}^*(t),u)=\max_{u\in\overline{U}}H(\widehat{\sigma}^*(t),u)=\mathcal{M}(\widehat{\sigma}^*(t))=H(\widehat{\sigma}^*(t),u^*(t))$$

almost everywhere by (3.a). Thus $\mathcal{M}(\widehat{\sigma}^*(t))=m(\widehat{\sigma}^*(t))$ a.e..

6. Applying Proposition A.8, we have $m(\widehat{\sigma}^*(t))=\mathcal{M}(\widehat{\sigma}^*(t))$ everywhere on I, because $m\circ\widehat{\sigma}^*$ is continuous on I, $\mathcal{M}\circ\widehat{\sigma}^*$ is lower semicontinuous, $m(\widehat{\sigma}^*(t))\leq\mathcal{M}(\widehat{\sigma}^*(t))$ everywhere and $m(\widehat{\sigma}^*(t))=\mathcal{M}(\widehat{\sigma}^*(t))$ almost everywhere.

7. $m\circ\widehat{\sigma}^*$ has zero derivative.

As $m\circ\widehat{\sigma}^*$ is absolutely continuous on I, by Corollary A.4 it has a derivative almost everywhere. As the intersection of two sets of full measure is not empty, see Appendix A, there exists a $t_0 \in I$ such that $m \circ \widehat{\sigma}^*$ is derivable at t_0 and $m(\widehat{\sigma}^*(t_0)) = H(\widehat{\sigma}^*(t_0), u^*(t_0))$. For each $t \neq t_0$, by the definition of m, we have

$$m(\widehat{\sigma}^*(t)) = \max_{u \in D \cap \overline{U}} H(\widehat{\sigma}^*(t), u) \geq H(\widehat{\sigma}^*(t), u^*(t_0))$$

because $u^*(t_0) \in D \cap \overline{U}$. Thus $m(\widehat{\sigma}^*(t)) - m(\widehat{\sigma}^*(t_0)) \geq H(\widehat{\sigma}^*(t), u^*(t_0)) - H(\widehat{\sigma}^*(t_0), u^*(t_0))$.

If $t - t_0 > 0$,

$$\frac{m(\widehat{\sigma}^*(t)) - m(\widehat{\sigma}^*(t_0))}{t - t_0} \geq \frac{H(\widehat{\sigma}^*(t), u^*(t_0)) - H(\widehat{\sigma}^*(t_0), u^*(t_0))}{t - t_0}.$$

Let us compute the right derivative of $m \circ \widehat{\sigma}^*$ at t_0

$$\left.\frac{d(m \circ \widehat{\sigma}^*)}{dt}\right|_{t=t_0^+} = \lim_{t \to t_0^+} \frac{m(\widehat{\sigma}^*(t)) - m(\widehat{\sigma}^*(t_0))}{t - t_0} \geq \lim_{t \to t_0^+} \frac{H^{u^*(t_0)}(\widehat{\sigma}^*(t)) - H^{u^*(t_0)}(\widehat{\sigma}^*(t_0))}{t - t_0} =$$

$$= \mathrm{L}_{\widehat{X}^{T^*\{u^*(t_0)\}}_{\widehat{\sigma}^*(t_0)}} H^{u^*(t_0)} = 0$$

since $i\left(\widehat{X}^{T^*\{u^*(t_0)\}}_{\widehat{\sigma}^*(t_0)}\right)\omega = \left(dH^{u^*(t_0)}\right)_{\widehat{\sigma}^*(t_0)}$.

Similarly, if $t - t_0 < 0$,

$$\left.\frac{d(m \circ \widehat{\sigma}^*)}{dt}\right|_{t=t_0^-} \leq 0.$$

Hence the derivative of $m \circ \widehat{\sigma}^*$ is zero almost everywhere.

8. Applying Theorem A.5, $m \circ \widehat{\sigma}^*$ is constant everywhere, because $m \circ \widehat{\sigma}^*$ is absolutely continuous.

9. As $m(\widehat{\sigma}^*(t))$ and $\mathcal{M}(\widehat{\sigma}^*(t))$ coincide everywhere, $\mathcal{M} \circ \widehat{\sigma}^*$ is constant everywhere on I.

(3.c) $\widehat{\alpha}^*(t) \neq 0 \in T^*_{\widehat{\gamma}^*(t)}\widehat{Q}$ for each $t \in [a, b]$.

Let us suppose that there exists $\tau \in [a, b]$ such that $\widehat{\alpha}^*(\tau) = 0 \in T^*_{\widehat{\gamma}^*(\tau)}\widehat{Q}$. As $\widehat{\sigma}^*$ is a generalized integral curve of $(\widehat{X}^{T^*})^{\{u^*\}}$, a linear vector field over $\widehat{X}$, then $\widehat{\alpha}^*(t) = 0$ for each $t \in [a, b]$. As there exists at least a time such that $\widehat{\alpha}^*(\tau) \neq 0$, we arrive at a contradiction. Hence $\widehat{\alpha}^*(t) \neq 0$ for each $t \in [a, b]$.

(3.d) $\alpha_0^*(t)$ is constant, $\alpha_0^*(t) \leq 0$.

From the equations satisfied by the generalized integral curves of $(\widehat{X}^{T^*})^{\{u^*\}}$, we have p_0 is constant. It was seen that $\langle \widehat{\alpha}_b, (-1, \mathbf{0}) \rangle \geq 0$ is equivalent to $(p_0 \circ \widehat{\alpha}^*)(b) = \alpha_0(b) \leq 0$. Hence $\alpha_0 \leq 0$ for each $t \in [a, b]$. □

Comment: As $\widehat{\alpha}_b$ is determined up to multiply by a positive real number, we may assume that $\alpha_0 \in \{-1, 0\}$.

The way how perturbations have been used in this proof give some clues of the fact that the tangent perturbation cone is understood as an approximation of the reachable set. The reachable set is the set of points swept by the integral curves of the control system taking all the admissible controls. A precise meaning of this approximation is explained in Appendix C.

The covector in the proof has been chosen such that

$$\langle \widehat{\alpha}_b, (-1, \mathbf{0}) \rangle \geq 0,$$
$$\langle \widehat{\alpha}_b, \widehat{v}_b \rangle \leq 0 \quad \forall \widehat{v}_b \in \widehat{K}_b.$$

In the abnormal case $\alpha_0 = 0$ and the first inequality is satisfied with equality. Thus the covector is contained in the separating hyperplane. It would be interesting to determine geometrically what else must happen in order to have abnormal minimizers.

5 Pontryagin's Maximum Principle for nonfixed time and nonfixed endpoints

Once Pontryagin's Maximum Principle has been proved for time and endpoints fixed, let us state the different problems related to Pontryagin's Maximum Principle with nonfixed time and nonfixed endpoints.

5.1 Statement of the optimal control problem with time and endpoints nonfixed

We consider the elements Q, U, X, F, $\mathcal{S}$ and π_2 with the same properties as in §2, §3.1. Let S_a and S_f be submanifolds of Q.

Statement 5.1. ***(Free Optimal Control Problem,*** *FOCP)* *Given the elements Q, U, X, F, and the disjoint submanifolds of Q, S_a and S_f, consider the following problem.*

Find $b \in \mathbb{R}$ and $(\gamma^*, u^*)\colon [a,b] \to Q \times U$ such that

(1) endpoint conditions: $\gamma^*(a) \in S_a$, $\gamma^*(b) \in S_f$,

(2) γ^* is an integral curve of $X^{\{u^*\}}$: $\dot{gamma}^* = X^{\{u^*\}} \circ (\gamma^*, \mathrm{id})$, and

(3) minimal condition: $\mathcal{S}[\gamma^*, u^*] = \int_a^b F(\gamma^*(t), u^*(t))dt$ is minimum over all curves (γ, u) satisfying (1) and (2).

The tuple (Q, U, X, F, S_a, S_f) denotes the *free optimal control problem.*

Statement 5.2. ***(Extended Free Optimal Control Problem,*** $\widehat{\textbf{FOCP}}$*)* *Given the FOCP, (Q, U, X, F, S_a, S_f), and the elements $\widehat{Q}$ and $\widehat{X}$ defined in §3.2, consider the following problem.*

Find $b \in \mathbb{R}$ and $(\widehat{\gamma}^*, u^*)\colon [a,b] \to \widehat{Q} \times U$, with $\gamma^* = \pi_2 \circ \widehat{\gamma}^*$, such that

(1) endpoint conditions: $\widehat{\gamma}^*(a) \in \{0\} \times S_a$, $\gamma^*(b) \in S_f$,

(2) $\widehat{\gamma}^*$ is an integral curve of $\widehat{X}^{\{u^*\}}$: $\dot{\widehat{\gamma}}^* = \widehat{X}^{\{u^*\}} \circ (\widehat{\gamma}^*, \mathrm{id})$, and

(3) minimal condition: ${\gamma^*}^0(b)$ is minimum over all curves $(\widehat{\gamma}, u)$ satisfying (1) and (2).

The tuple $(\widehat{Q}, U, \widehat{X}, S_a, S_f)$ denotes the *extended free optimal control problem.*

5.2 Perturbation of the time and the endpoints

In this case of nonfixed time and nonfixed endpoint optimal control problems, we do not only just modify the control as explained in §3.3, but also modify the final time and the endpoint conditions. As was mentioned in 3.3, the following constructions obtained from perturbing the final time and the endpoint conditions are also general for any vector field depending on parameters.

5.2.1 Time perturbation vectors and associated cones

We study how to perturb the interval of definition of the control taking advantage of the fact that the final time is another unknown for the free optimal control problems.

Let X be a vector field on M along the projection $\pi\colon M \times U \to M$, $I \subset \mathbb{R}$ be a closed interval and $(\gamma, u)\colon I = [a,b] \to M \times U$ a curve such that γ is an integral curve of $X^{\{u\}}$.

Let $\pi_\pm = \{\tau, l_\tau, \delta\tau, u_\tau\}$, where τ is a Lebesgue time in (a,b) for $X \circ (\gamma, u)$, $l_\tau \in \mathbb{R}^+ \cup \{0\}$, $\delta\tau \in \mathbb{R}$, $u_\tau \in U$. For every $s \in \mathbb{R}^+$ small enough such that $a < \tau - (l_\tau - \delta\tau)s$, consider $u[\pi^s_\pm]\colon [a, b+\delta\tau s] \to U$ defined by

$$u[\pi^s_\pm](t) = \begin{cases} u(t), & t \in [a, \tau - (l_\tau - \delta\tau)s], \\ u_\tau, & t \in (\, \tau - (l_\tau - \delta\tau)s, \tau + \delta\tau s\,], \\ u(t), & t \in (\, \tau + \delta\tau s, b + \delta\tau s\,], \end{cases}$$

if $\delta\tau < 0$, and by

$$u[\pi^s_\pm](t) = \begin{cases} u(t), & t \in [a, \tau - (l_\tau - \delta\tau)s], \\ u_\tau, & t \in (\, \tau - (l_\tau - \delta\tau)s, \tau + \delta\tau s\,], \\ u(t - \delta\tau s), & t \in (\, \tau + \delta\tau s, b + \delta\tau s\,], \end{cases}$$

if $\delta\tau \geq 0$.

Definition 5.3. *The function $u[\pi^s_\pm]$ is called a* ***perturbation of*** u ***specified by the data*** $\pi_\pm = \{\tau, l_\tau, \delta\tau, u_\tau\}$.

Associated to $u[\pi^s_\pm]$ we consider the mapping $\gamma[\pi^s_\pm]\colon [a, b+\delta\tau s] \to M$, the generalized integral curve of $X^{\{u[\pi^s_\pm]\}}$ with initial condition $(a, \gamma(a))$.

Given $\epsilon > 0$, define

$$\begin{array}{rcl} \varphi_{\pi_\pm}\colon [\tau, b] \times [0, \epsilon] & \longrightarrow & M \\ (t,s) & \longmapsto & \varphi_{\pi_\pm}(t,s) = \gamma[\pi^s_\pm](t + \delta\tau s) \end{array}$$

For every $t \in [\tau, b]$, $\varphi^t_{\pi_\pm}\colon [0,\epsilon] \to M$ is given by $\varphi^t_{\pi_\pm}(s) = \varphi_{\pi_\pm}(t,s)$.

As explained in §3.3, the control $u[\pi^s_\pm]$ depends continuously on the parameters s and $\pi_\pm = \{\tau, l_\tau, \delta\tau, u_\tau\}$. Hence the curve $\varphi^t_{\pi_\pm}$ depends continuously on s and $\pi_\pm = \{\tau, l_\tau, \delta\tau, u_\tau\}$, then it converges uniformly to γ as s tends to 0. See [13, 14] for more details of the differential equations depending continuously on parameters.

Let us prove that the curve $\varphi^\tau_{\pi_\pm}$ has a tangent vector at $s = 0$, compare with Proposition 3.4.

Proposition 5.4. *The curve $\varphi^\tau_{\pi_\pm}\colon [0,\epsilon] \to M$ is differentiable at $s = 0$. Its tangent vector is $X(\gamma(\tau), u(\tau))\,\delta\tau + [X(\gamma(\tau), u_\tau) - X(\gamma(\tau), u(\tau))]\; l_\tau$.*

Proof. As in the proof of Proposition 3.4, we compute the limit

$$A = \lim_{s\to 0} \frac{(x^i \circ \varphi^\tau_{\pi_\pm})(s) - (x^i \circ \varphi^\tau_{\pi_\pm})(0)}{s} = \lim_{s\to 0} \frac{\gamma^i[\pi^s_\pm](\tau + \delta\tau s) - \gamma^i(\tau)}{s}$$

As γ is an absolutely continuous integral curve of $X^{\{u\}}$, $\dot\gamma(t) = X(\gamma(t), u(t))$ at every Lebesgue time. Then by integration

$$\gamma^i(\tau) - \gamma^i(a) = \int_a^\tau f^i(\gamma(t), u(t))dt$$

and similarly for $\gamma[\pi^s_\pm]$ and $u[\pi^s_\pm]$. Observe that $\gamma[\pi^s_\pm](t) = \gamma(t)$ and $u[\pi^s_\pm](t) = u(t)$ for $t \in [a, \tau - (l_\tau - \delta\tau)s]$.

Here, we should consider three different possibilities

- if $0 \leq \delta\tau \leq l_\tau$, then $\tau - (l_\tau - \delta\tau)s < \tau < \tau + \delta\tau s$,
- if $\delta\tau < 0$, then $\tau - (l_\tau - \delta\tau)s < \tau + \delta\tau s < \tau$,
- if $0 < l_\tau < \delta\tau$, then $\tau < \tau - (l_\tau - \delta\tau)s < \tau + \delta\tau s$.

We prove the proposition for the first case and the other cases follow analogously.

$$A = \lim_{s\to 0} \frac{\int_a^{\tau+\delta\tau s} f^i(\gamma[\pi^s_\pm](t), u[\pi^s_\pm](t))dt - \int_a^\tau f^i(\gamma(t), u(t))dt}{s} =$$

$$= \lim_{s\to 0} \frac{\int_{\tau-(l_\tau-\delta\tau)s}^{\tau+\delta\tau s} f^i(\gamma[\pi^s_\pm](t), u_\tau)dt - \int_{\tau-(l_\tau-\delta\tau)s}^{\tau} f^i(\gamma(t), u(t))dt}{s}$$

As $\tau + \delta\tau s$ is a Lebesgue time, we use Equation (A.15).

$$A = \lim_{s\to 0} \frac{f^i(\gamma[\pi^s_\pm](\tau + \delta\tau s), u_\tau)l_\tau s - f^i(\gamma(\tau), u(\tau))(l_\tau - \delta\tau)s + o(s)}{s} =$$

$$= \lim_{s\to 0} f^i(\gamma[\pi^s_\pm](\tau + \delta\tau s), u_\tau)l_\tau - f^i(\gamma(\tau), u(\tau))(l_\tau - \delta\tau)$$

As f^i is continuous on M, we have

$$A = [f^i(\gamma(\tau), u_\tau) - f^i(\gamma(\tau), u(\tau))]\, l_\tau + f^i(\gamma(\tau), u(\tau))\ \delta\tau =$$

$$= \mathrm{L}([X(\gamma(\tau), u(\tau))\ \delta\tau + (X(\gamma(\tau), u_\tau) - X(\gamma(\tau), u(\tau)))\ \ l_\tau])(x^i).$$

□

Definition 5.5. *The tangent vector* $v[\pi_\pm] = X(\gamma(\tau), u(\tau))\ \delta\tau + [X(\gamma(\tau), u_\tau) - X(\gamma(\tau), u(\tau))]\ l_\tau$ *is the* ***perturbation vector associated to the perturbation data*** $\pi_\pm = \{\tau, l_\tau, \delta\tau, u_\tau\}$.

If we disturb the control r times at r different Lebesgue times as in §3.3.1 and also the domain of the curve (γ, u) as just described, that is, $\pi = \{\pi_1, \ldots, \pi_r, \pi_\pm\}$, with $a < t_1 \leq \ldots \leq t_r \leq \tau < b$, then $\gamma[\pi^s]$ is the generalized integral curve of $X^{\{u[\pi^s]\}}$ with initial condition $(a, \gamma(a))$. Consider the curve $\varphi^t_\pi \colon [0, \epsilon] \to M$ for $t \in [\tau, b]$ given by $\varphi^t_\pi(s) = \gamma[\pi^s](t + \delta\tau s)$.

Corollary 5.6. *The vector tangent to the curve* $\varphi^t_{\pi_\pm} \colon [0, \epsilon] \to M$ *at* $s = 0$ *is* $X(\gamma(t), u(t))\ \delta\tau + V[\pi_1](t) + \ldots + V[\pi_n](t)$*, where* $V[\pi_i] \colon [t_i, b] \to TM$ *is the generalized integral curve of* $\left(X^T\right)^{\{u\}}$ *with initial condition* $(t_i, (\gamma(t_i), v[\pi_i]))$.

This corollary may be proved taking into account Proposition 3.6, Corollary 3.9 and Appendix B.

Now, at a Lebesgue time $t \in (a, b)$, take the union of the tangent perturbation cone K_t, see Definition 3.11, with $\pm X(\gamma(t), u(t))$ and close it convexly and topologically.

Definition 5.7. *The* ***time perturbation cone*** $K^\pm_t$ *at every Lebesgue time* t *is the smallest closed cone in the tangent space at* $\gamma(t)$ *containing* K_t *and* $\pm X(\gamma(t), u(t))$.

$$K^\pm_t = \overline{\mathrm{conv}\left(\{\pm\lambda X(\gamma(t), u(t)) \mid \lambda \in \mathbb{R}\} \bigcup \left(\bigcup_{a<\tau\leq t} \left(\Phi^{X\{u\}}_{(t,\tau)}\right)_* \mathcal{V}_\tau\right)\right)}$$

where $\mathcal{V}_\tau$ *denotes the set of elementary perturbation vectors at* τ*, see Definition 3.11.*

Enlarging the cone K_τ to $K_\tau^\pm$ allows us to introduce time variations, increasing or decreasing the final time.

Proposition 5.8. *If t_2 is a Lebesgue time greater than t_1, then $\left(\Phi^{\widehat{X}\{u^*\}}_{(t_2,t_1)}\right)_* \widehat{K}^\pm_{t_1} \subset \widehat{K}^\pm_{t_2}$.*

Proof. We have

$$\widehat{K}^\pm_{t_1} = \overline{\operatorname{conv}\left(\{\pm\widehat{X}(\widehat{\gamma}^*(t_1), u^*(t_1)) \mid \lambda \in \mathbb{R}\} \bigcup \left(\bigcup_{a<\tau\leq t_1} \left(\Phi^{X\{u\}}_{(t_1,\tau)}\right)_* \mathcal{V}_\tau\right)\right)}.$$

Just for simplicity we use $\mathcal{C}^\pm_{t_1}$ to denote

$$\operatorname{conv}\left(\{\pm\widehat{X}(\widehat{\gamma}^*(t_1), u^*(t_1)) \mid \lambda \in \mathbb{R}\} \bigcup_{a<\tau<t_1} \left(\Phi^{X\{u\}}_{(t_1,\tau)}\right)_* \mathcal{V}_\tau\right).$$

1. The set $\mathcal{C}^\pm_{t_1}$ being convex, if $\widehat{v}$ is interior to $\widehat{K}^\pm_{t_1}$, then $\widehat{v}$ is interior to $\mathcal{C}^\pm_{t_1}$ by Proposition D.5, item (d). Hence by Proposition D.4

   $$\widehat{v} = \delta t_1 \widehat{X}(\widehat{\gamma}^*(t_1), u^*(t_1)) + \sum l_i \widehat{V}[\pi_i](t_1),$$

   where every $\widehat{V}[\pi_i](t_1)$ is the transported of the elementary perturbation vector $\widehat{v}[\pi_i]$ of class I from t_i to t_1 by the flow of $\widehat{X}^{\{u^*\}}$. By definition of the cone and the linearity of the flow, $\left(\Phi^{\widehat{X}\{u^*\}}_{(t_2,t_1)}\right)_* \widehat{v}$ is in $\widehat{K}^\pm_{t_2}$, since $\left(\Phi^{\widehat{X}\{u^*\}}_{(t_2,t_1)}\right)_* \left(\widehat{X}(\widehat{\gamma}^*(t_1), u^*(t_1))\right) = \widehat{X}(\widehat{\gamma}^*(t_2), u^*(t_2))$, because both sides of the equality are the unique solutions of the variational equation along $\widehat{\gamma}^*$ associated with $\widehat{X}^{\{u^*\}}$ with initial condition $(t_1, \widehat{X}(\widehat{\gamma}^*(t_1), u^*(t_1)))$. See Appendix B.2 for more details.

2. If $\widehat{v}$ is in the boundary of $\widehat{K}^\pm_{t_1}$, then there exists a sequence of vectors $(v_j)_{j\in\mathbb{N}}$ in the interior of $\widehat{K}^\pm_{t_1}$ such that

   $$\lim_{j\to\infty} \widehat{v}_j = \widehat{v}.$$

   Due to the continuity of the flow

   $$\lim_{j\to\infty} \left(\Phi^{\widehat{X}\{u^*\}}_{(t_2,t_1)}\right)_* \widehat{v}_j = \left(\Phi^{\widehat{X}\{u^*\}}_{(t_2,t_1)}\right)_* \widehat{v}.$$

   All the elements of the convergent sequence are in the closed cone $\widehat{K}^\pm_{t_2}$, hence the limit $\left(\Phi^{\widehat{X}\{u^*\}}_{(t_2,t_1)}\right)_* \widehat{v}$ is also in $\widehat{K}^\pm_{t_2}$. □

We can state and prove other properties similar to Propositions 3.4, 3.6, 3.7 and 3.8, already proved for the tangent perturbation cone, but for the time perturbation cone $K^\pm_\tau$. These results will not be proved, but they are assumed as true from now on.

Proposition 5.9. *Let $t \in (a, b)$ be a Lebesgue time. If v is a nonzero vector interior to $K^\pm_t$, then there exists $\epsilon > 0$ such that for every $s \in (0, \epsilon)$ there are $s' > 0$ and a perturbation of the control $u[\pi^s_\pm]$ such that $\gamma[\pi^s_\pm](t + s\delta t) = \gamma(t) + s'v$.*

Proof. The proof follows the same line as the proof of Proposition 3.12, but now the tangent space to M at $\gamma(t + \delta ts)$ is also identified with $\mathbb{R}^n$ through the local chart of M at $\gamma(t)$.

We use the same functions as in the proof of Proposition 3.12, but changing $\Gamma(s, r) = \gamma[\pi^s_{w_0}](t)$ by $\Gamma(s, r) = \gamma[\widetilde{\pi}^s_{w_0}](t + s\delta t)$. Observe that this change modifies the expression of the mappings Δ, Θ, G_s and $\mathcal{G}$. □

5.2.2 Perturbing the endpoint conditions

Now we consider that the endpoint conditions for the integral curves of $X^{\{u\}}$ varies on submanifolds of M. Let S_a be a submanifold of M and $\gamma(a)$ in S_a; consider the integral curve $\gamma\colon I \to M$ of $X^{\{u\}}$ with initial condition $(a, \gamma(a))$.

We consider the curve $\gamma[\pi^s_\pm]$ obtained from a time perturbation of the control u associated with a vector in the time perturbation cone. The initial condition is disturbed along a curve $\delta\colon [0,\epsilon] \to S_a$ with initial tangent vector v_a in $T_{\gamma(a)}S_a$ and $\delta(0) = \gamma(a)$. Taking into account Appendix B.2.1, §3.3.1 and considering that $T_{\gamma(a)}S_a$ and an open set at $\delta(a)$ are identified with $\mathbb{R}^n$, the integral curve $\gamma_{\delta(s)}[\pi^s_\pm]\colon I \to M$ of $X^{\{u[\pi^s_\pm]\}}$ with initial condition $(a, \delta(s))$ can be written as

$$\gamma_{\delta(s)}[\pi^s_\pm](t) = \gamma(t) + s\left(\Phi^{X^{\{u\}}}_{t,a}\right)_* v_a + sv[\pi_\pm](t) + o(s).$$

We define a cone that includes the time perturbation vectors, the elementary perturbation vectors and the vectors coming from changing the initial condition on S_a along different curves $\delta\colon [0,\epsilon] \to S_a$ through $\gamma(a)$ and contained in S_a.

Definition 5.10. *Let t be a Lebesgue time. The cone $\mathcal{K}_t$ is the smallest closed and convex cone containing the time perturbation cone at time t and the transported of the tangent space to S_a from a to t through the flow of $X^{\{u\}}$.*

$$\mathcal{K}_t = \overline{\operatorname{conv}(K^\pm_t \bigcup (\Phi^{X^{\{u\}}}_{t,a})_*(T_{\gamma(a)}S_a))}$$

Proposition 5.11. *Let t be a Lebesgue time in (a,b) and $S \subset M$ be a submanifold with boundary. Suppose that $\gamma(t)$ is on the boundary of S. Let T be the half-plane tangent to S at $\gamma(t)$. If $\mathcal{K}_t$ and T are not separated, then there exists a perturbation of the control $u[\pi^s_\pm]$ and $x_a \in S_a$ such that the integral curve $\gamma_{x_a}[\pi^s_\pm]$ of $X^{\{u[\pi^s_\pm]\}}$ with initial condition (a, x_a) meets S at a point in the relative interior of S.*

Proof. As $\mathcal{K}_t$ and T are not separated, by Proposition D.15 there no exists any hyperplane containing both and there is a vector v in the relative interior of both $\mathcal{K}_t$ and T. By Corollary D.16, if $\mathcal{K}_t$ and T are not separated,

$$T_{\gamma(t)}M = \mathcal{K}_t - T.$$

See Appendix D for the notation and properties. If V is an open set of a local chart at $\gamma(t)$, we identify V with $\mathbb{R}^n$ and also the tangent space at $\gamma(t)$, $T_{\gamma(t)}M$, in the same sense defined for Equation (3.2). Let us consider an orthonormal basis in $T_{\gamma(t)}M$, $\{e_1, \ldots, e_n\}$. If we take $e_0 = -(e_1 + \ldots + e_n)$, the vector $0 \in T_{\gamma(t)}M$ is expressed as an affine combination of $e_0, e_1, \ldots, e_n$:

$$0 = \frac{1}{n+1}e_0 + \ldots + \frac{1}{n+1}e_n.$$

Each w in $T_{\gamma(t)}M$ is written uniquely as

$$w = a^1 e_1 + \ldots + a^n e_n$$

and as an affine combination of $e_0, e_1, \ldots, e_n$:

$$w = \sum_{i=0}^{n} b^i(w) e_i = r e_0 + \sum_{i=0}^{n} (r + a^i) e_i \quad \text{with} \quad r = \frac{1 - \sum_{i=1}^{n} a^i}{n+1}.$$

Hence, we define the continuous mapping

$$\begin{array}{rccc} \mathcal{G}\colon & T_{\gamma(t)}M & \longrightarrow & \mathbb{R}^{n+1} \\ & w & \longmapsto & (b^0(w), b^1(w), \ldots, b^n(w)) \end{array}$$

As $b^i(0) > 0$ for every $i = 0, \ldots, n$, there exists an open ball $B(0, r)$ centered at 0 with radius r such that for every $w \in B(0, r)$, $b^i(w) > 0$ for $i = 0, \ldots, n$. Now we consider the restriction of $\mathcal{G}$ to the closed ball $\overline{B(0, r)}$, $\mathcal{G}_{\overline{B(0,r)}} \colon \overline{B(0, r)} \to [0, 1]^{n+1}$. Choose vectors $e_i^{\mathcal{K}} \in \mathcal{K}_t$ and $e_i^T \in T$ such that

$$e^i = e_i^{\mathcal{K}} - e_i^T.$$

As v lies in the relative interior of both convex sets, $e^i = (e_i^{\mathcal{K}} + v) - (e_i^T + v) = e_i^{\mathcal{K}} - e_i^T$. Then $e_i^{\mathcal{K}}$ and e_i^T are in the relative interior of $\mathcal{K}$ and T respectively. For any $w \in \overline{B(0, r)}$,

$$w = \sum_{i=0}^{n} b^i(w) e_i = \sum_{i=0}^{n} b^i(w) \left(e_i^{\mathcal{K}} - e_i^T\right).$$

Then we can define

$$\begin{array}{rccc} F_1 \colon & \overline{B(0,r)} & \longrightarrow & \mathcal{K}_t \\ & w & \longmapsto & F_1(w) = \sum_{i=0}^{n} b^i(w) e_i^{\mathcal{K}}, \end{array}$$

$$\begin{array}{rccc} F_2 \colon & \overline{B(0,r)} & \longrightarrow & T \\ & w & \longmapsto & F_2(w) = \sum_{i=0}^{n} b^i(w) e_i^T \end{array}$$

and let us consider the mapping

$$\begin{array}{rccc} G \colon & \mathbb{R} \times \overline{B(0,r)} & \longrightarrow & \mathbb{R}^n \\ & (s, w) & \longmapsto & (\gamma[\pi^s_{F_1(w)}](t) - \gamma[sF_2(w)](t))/s \end{array}$$

where $\gamma[\pi^s_{F_1(w)}]$ is the perturbation curve associated to $\pi^s_{F_1(w)}$ and $\gamma[sF_2(w)](t) = \gamma(t) + sF_2(w)$ is the straight line through $\gamma(t)$ with tangent vector $F_2(w)$. In local coordinates $G(s, w) = F_1(w) - F_2(w) + o(1)$, hence

$$\lim_{s \to 0} G(s, w) = F_1(w) - F_2(w) = w.$$

Hence, for any positive number ϵ, there exists $s_0 > 0$ such that if $s < s_0$ then $\|G(s, w) - w\| < \epsilon$. Take $\epsilon < r$, then

$$\|G(s, w) - w\| < \epsilon < r = \|w\|$$

for every w in the boundary of $\overline{B(0, r)}$. Thus the map $G_s \colon \overline{B(0, r)} \to \mathbb{R}^n$, $G_s(w) = G(s, w)$, satisfies the hypotheses of Corollary E.2 for the point 0 in $B(0, r)$. Hence, the point 0 is in the image of $\overline{B(0, r)}$ through G_s and there exists w such that $G_s(w) = 0$, that is,

$$\gamma[\pi^s_{F_1(w)}](t) = \gamma[sF_2(w)](t).$$

Therefore, there exists a perturbation of the control such that the associated trajectory meets S in an interior point since $F_2(w)$ lies in the relative interior of T. $\square$

5.3 Pontryagin's Maximum Principle with time and endpoints nonfixed

Bearing in mind the symplectic formalism introduced in §3.4, we define the corresponding Hamiltonian Problem when the time and the endpoints are nonfixed.

Statement 5.12. ***(Free Hamiltonian Problem,*** *FHP) Given the FOCP,* (Q, U, X, F, S_a, S_f), *and the equivalent* $\widehat{FOCP}$, $(\widehat{Q}, U, \widehat{X}, S_a, S_f)$, *consider the following problem.*

Find $b \in \mathbb{R}$ *and* $(\widehat{\sigma}^*, u^*) \colon [a, b] \to T^*\widehat{Q} \times U$, *with* $\widehat{\gamma}^* = \pi_{\widehat{Q}} \circ \widehat{\sigma}^*$ *and* $\gamma^* = \pi_2 \circ \widehat{\gamma}^*$, *such that*

(1) $\widehat{\gamma}^*(a) \in \{0\} \times S_a$, $\gamma^*(b) \in S_f$, *and*

(2) $\dot{\widehat{\sigma}}^* = (\widehat{X}^{T^*})^{\{u^*\}} \circ (\widehat{\sigma}^*, \mathrm{id})$.

The tuple $(T^*\widehat{Q}, U, \widehat{X}^{T^*}, S_a, S_f)$ denotes the *free hamiltonian problem.*

Comments:

1. The minimum of the interval of definition of the curves is a, but the maximum is not fixed.

2. In the degenerate case that the submanifolds S_a and S_f are only a point, then Pontryagin's Maximum Principle for fixed endpoints with free time is applied. This statement does not appear in this paper, but it may be deduced from Theorems 3.14, 5.13.

3. The curves γ, $\widehat{\gamma}$ and $\widehat{\sigma}$ are assumed to be absolutely continuous. So they are generalized integral curves of $X^{\{u\}}$, $\widehat{X}^{\{u\}}$ and $(\widehat{X}^{T^*})^{\{u\}}$, respectively, in the sense defined in §2.

Now, we are ready to state the Free Pontryagin's Maximum Principle that provides the necessary conditions, but in general not sufficient, for finding solutions of the free optimal control problem.

Theorem 5.13. ***(Free Pontryagin's Maximum Principle, FPMP)***
Let $(\widehat{\gamma}^*, u^*) \colon [a, b] \to \widehat{Q} \times U$ be a solution of the extended free optimal control problem, Statement 5.2. Then there exists $(\widehat{\sigma}^*, u^*) \colon [a, b] \to T^*\widehat{Q} \times U$ such that:

1. *it is a solution of the associated free hamiltonian problem;*

2. *$\widehat{\gamma}^* = \pi_{\widehat{Q}} \circ \widehat{\sigma}^*$, with fiber $\widehat{\alpha}^*(t) = (\alpha_0^*(t), \alpha^*(t)) \in T^*_{\widehat{\gamma}^*(t)}\widehat{Q}$;;*

3. *(a) $H(\widehat{\sigma}^*(t), u^*(t)) = \max_{u\in\overline{U}} H(\widehat{\sigma}^*(t), u)$ almost everywhere;*
   (b) $\max_{u\in\overline{U}} H(\widehat{\sigma}^*(t), u) = 0$ everywhere;
   (c) $\widehat{\alpha}^*(t) \neq 0 \in T^*_{\widehat{\gamma}^*(t)}\widehat{Q}$ for each $t \in [a, b]$;
   (d) $\alpha_0^*(t)$ is constant, $\alpha_0^*(t) \leq 0$;
   (e) transversality conditions:$\alpha^*(a) \in \operatorname{ann} T_{\gamma^*(a)}S_a$ and $\alpha^*(b) \in \operatorname{ann} T_{\gamma^*(b)}S_f$.

Observe that if we have a solution of the $\widehat{FOCP}$, the final time and the endpoints of an optimal solution are known. Thus, the final time is known by hypothesis. Nevertheless in the proof of Theorem 5.13 the freedom to chose the final time in Problem is used to consider variations of the optimal curve that are slightly different from the variations used in the case of fixed time, see §4 and §5 to compare them.

The main difference between FPMP and PMP, putting aside the transversality conditions, is the fact that the domain of the curves in the optimal control problems is unknown. That introduces a new necessary condition: the maximum of the Hamiltonian must be zero, not just constant. Then, from (a) and (b) it may be concluded that the Hamiltonian is zero almost everywhere. For instance, in the time optimal problems the Hamiltonian along extremals must be zero.

The transversality conditions show up when the endpoints are not fixed. The annihilators of the tangent space to the initial and final submanifolds, $\operatorname{ann} T_{\gamma^*(a)}S_f a$ and $\operatorname{ann} T_{\gamma^*(b)}S_f$ are subspaces of the cotangent space.

There are different statements of Pontryagin's Maximum Principle. In §3.4 we have considered the statement of PMP for a fixed-time problem without transversality conditions to simplify the proof, although it may be stated the PMP for the fixed-time problem with variable endpoints where the transversality conditions appear.

6 Proof of Pontryagin's Maximum Principle for nonfixed time and nonfixed endpoints

In the proof of Theorem 5.13 we use notions about perturbations of the trajectories of a system introduced in §5.2, but they are slightly different from the perturbations in §3.3 used to prove Theorem 3.14.

Proof. (Theorem 5.13: Free Pontryagin's Maximum Principle, FPMP)

Given a solution of the $\widehat{FOCP}$, we only need an appropriate initial condition in the fibers of $\pi_{\widehat{Q}}\colon T^*\widehat{Q} \to \widehat{Q}$ to find a solution of the FHP, because this initial condition is not given in the hypotheses of the Free Pontryagin's Maximum Principle. It is not possible to use Theorem 3.14 directly because the perturbation cones are not the same. Indeed, we need to consider changes in the interval of definition of the curves. These changes imply the inclusion of $\pm\widehat{X}(\widehat{\gamma}^*(t_1), u^*(t_1))$ in the perturbation cone at time t_1. All the times considered in this proof are Lebesgue times for the vector field giving the optimal curve.

By Proposition 5.8, for $t_2 > t_1$,

$$\left(\Phi^{\widehat{X}\{u^*\}}_{(t_2,t_1)}\right)_* \widehat{K}^{\pm}_{t_1} \subset \widehat{K}^{\pm}_{t_2}.$$

Let us consider the limit cone as follows

$$\widehat{K}^{\pm}_b = \overline{\bigcup_{a<\tau<b} \left(\Phi^{\widehat{X}\{u\}}_{(b,\tau)}\right)_* \widehat{K}^{\pm}_{\tau}}.$$

Observe that it is a closed cone and it is convex because it is the union of an increasing family of convex cones. Let us show that $(-1,\mathbf{0})_{\widehat{\gamma}^*(b)}$ is not interior to $\widehat{K}^{\pm}_b$. Indeed, suppose that $(-1,\mathbf{0})_{\widehat{\gamma}^*(b)}$ is interior to the limit cone, then it will be interior to $\bigcup_{a<\tau<b}(\Phi^{\widehat{X}\{u\}}_{(b,\tau)})_*\widehat{K}^{\pm}_{\tau}$ by Proposition D.5, item (d). As we have an increasing family of cones, there exists a time τ such that $(-1,\mathbf{0})_{\widehat{\gamma}^*(b)}$ is interior to $(\Phi^{\widehat{X}\{u\}}_{(b,\tau)})_*\widehat{K}^{\pm}_{\tau}$. Let us see that this is not possible.

If $(-1,\mathbf{0})_{\widehat{\gamma}^*(b)}$ is interior to $(\Phi^{\widehat{X}\{u\}}_{(b,\tau)})_*\widehat{K}^{\pm}_{\tau}$, then, by Proposition 5.9, there exists $\epsilon > 0$ such that for every $s \in (0,\epsilon)$ there exist $s' > 0$ and a perturbation of the control $u[\pi^s_{\pm}]$ such that

$$\widehat{\gamma}[\pi^s_{w_0}](b + s\delta\tau) = \widehat{\gamma}^*(b) + s'(-1,\mathbf{0}).$$

Hence

$$\gamma^0[\pi^s_{w_0}](b + s\delta\tau) < {\gamma^*}^0(b) \quad \text{and} \quad \gamma[\pi^s_{w_0}](b + s\delta\tau) = \gamma^*(b).$$

That is, the trajectory $\gamma[\pi^s_{w_0}]$ arrives at the same endpoint as γ^* but with less cost. Then $\widehat{\gamma}^*$ cannot be optimal as assumed. Thus $(-1,\mathbf{0})_{\widehat{\gamma}^*(b)}$ is not interior to $\widehat{K}^{\pm}_b$.

As $(-1,\mathbf{0})_{\widehat{\gamma}^*(b)}$ is not in the interior of $\widehat{K}^{\pm}_b$, by Proposition D.15 there exists a covector $\widehat{\alpha}_b \in T^*_{\widehat{\gamma}^*(b)}\widehat{Q}$ such that

$$\begin{aligned} \langle \widehat{\alpha}_b, (-1,\mathbf{0})\rangle &\geq 0, \\ \langle \widehat{\alpha}_b, \widehat{v}_b\rangle &\leq 0 \quad \forall\, \widehat{v}_b \in \widehat{K}^{\pm}_b. \end{aligned}$$

The initial condition for the covector not only satisfies the previous inequalities, but also the transversality conditions. In order to prove this, it is necessary to have the separability of two new cones.

(3.e) Hence, the initial condition in the fibers of $T^*\widehat{Q}$ may be chosen satisfying the **transversality conditions**. We consider the manifold with boundary given by $M_f = \{(x^0, x) \mid x \in S_f,\ x^0 \le {\gamma^*}^0(b)\}$. The set of tangent vectors to M_f at $\widehat{\gamma}^*(b)$ is the convex set whose generators are $(-1, \mathbf{0})_{\widehat{\gamma}^*(b)}$ and $T_f = \{0\} \times T_{\gamma^*(b)}S_f$.

Given $\tau \in [a, b]$, consider the following closed convex sets

$$\mathcal{K}_\tau = \overline{\mathrm{conv}(\widehat{K}_\tau^\pm \bigcup (\Phi^{\widehat{X}\{u^*\}}_{(\tau,a)})_*(T_a))} \qquad \text{where} \quad T_a = \{0\} \times T_{\gamma^*(a)}S_a,$$

$$\mathcal{J}_\tau = \overline{\mathrm{conv}((-1, \mathbf{0})_{\widehat{\gamma}^*(\tau)} \bigcup (\Phi^{\widehat{X}\{u^*\}}_{(b,\tau)})_*^{-1}(T_f))} \qquad \text{where} \quad T_f = \{0\} \times T_{\gamma^*(b)}S_f,$$

and the manifold M_τ obtained transporting M_f from b to τ using the flow of $\widehat{X}^{\{u^*\}}$. Observe that $\mathcal{J}_\tau$ is the closure of the set of tangent vectors to M_τ at the point $\widehat{\gamma}^*(\tau)$. We are going to show that the cones $\mathcal{K}_b$ and $\mathcal{J}_b$ are separated, using Proposition 5.11.

Observe that $\mathcal{J}_b$ is a half-plane tangent to M_f and $\widehat{\gamma}^*(b)$ is on the boundary of M_f by construction. Hence, if $\mathcal{K}_b$ and $\mathcal{J}_b$ were not separated, by Proposition 5.11 there would exist a perturbation of the control $u[\pi^s_\pm]$ and $x_a \in S_a$ such that the integral curve $\gamma_{x_a}[\pi^s_\pm]$ with initial condition (a, x_a) meets M_f at a point in the relative interior of M_f. Hence we have found a trajectory with less cost than the optimal one because of the definition of M_f and this is not possible because of the optimality of $\widehat{\gamma}^*$. Thus $\mathcal{K}_b$ and $\mathcal{J}_b$ are separated. So, by Proposition D.15, there exists a covector $\widehat{\alpha}_b \in T^*_{\widehat{\gamma}^*(b)}\widehat{Q}$ such that

$$\langle \widehat{\alpha}_b, \widehat{v}_b \rangle \le 0 \quad \forall\, \widehat{v}_b \in \mathcal{K}_b, \tag{6.11}$$

$$\langle \widehat{\alpha}_b, \widehat{w}_b \rangle \ge 0 \quad \forall\, \widehat{w}_b \in \mathcal{J}_b. \tag{6.12}$$

This covector separates the vector $(-1, \mathbf{0})_{\widehat{\gamma}^*(b)} \in \mathcal{J}_b$ and the cone $\widehat{K}^\pm_b \subset \mathcal{K}^\pm_b$. Let $\widehat{\sigma}^* = (\widehat{\gamma}^*, \widehat{\alpha}^*)$ be the integral curve of $(\widehat{X}^{T^*})^{\{u^*\}}$ with initial condition $(\widehat{\gamma}^*(b), \widehat{\alpha}_b)$.

As T_f is contained in $\mathcal{J}_b$, we have $\langle \widehat{\alpha}_b, \widehat{v} \rangle \ge 0$ for every $\widehat{v} \in T_f$. As T_f is a vector space, if $\widehat{v} \in T_f$, then $-\widehat{v} \in T_f$. Hence, we have

$$\langle \widehat{\alpha}_b, \widehat{v} \rangle = 0 \quad \text{for every} \quad \widehat{v} \in T_f,$$

that is,

$$\langle \widehat{\alpha}_b, (0, v) \rangle = 0 \quad \text{for every} \quad v \in T_{\gamma^*(b)}S_f.$$

This is equivalent to $\langle \alpha_b, v \rangle = 0$ for every $v \in T_{\gamma^*(b)}S_f$, that is, $\alpha_b = \alpha^*(b)$ is in the annihilator of $T_{\gamma^*(b)}S_f$ as wanted.

For the initial transversality condition at a, for every $\widehat{w}_b \in \mathcal{J}_b$, if $\widehat{W} \colon I \to TQ$ is the integral curve of $(\widehat{X}^T)^{\{u^*\}}$ with initial condition $(b, \widehat{\gamma}(b), \widehat{w}_b)$, then by Proposition B.3 the pairing continuous function $\langle \widehat{\alpha}^*, \widehat{W} \rangle \colon I \to \mathbb{R}$ is constant everywhere and $\langle \widehat{\alpha}^*(a), \widehat{W}(a) \rangle \ge 0$ by Equation (6.12). As $(\Phi^{\widehat{X}\{u^*\}}_{(b,a)})_*^{-1}(\mathcal{J}_b) = \mathcal{J}_a$ by the continuity and the linearity of the flow, the transversality condition at a is proved analogously as the transversality condition at b proved above.

Being $(\widehat{\gamma}^*, u^*)$ a solution of the $\widehat{FOCP}$, it is also a solution of $\widehat{OCP}$ with time and endpoints fixed and given by the curve. Hence, we can apply Pontryagin's Maximum Principle for time and endpoints fixed, Theorem 3.14. If the curve $(\widehat{\gamma}^*, u^*)$ is a solution of $\widehat{OCP}$ with $I = [a, b]$ and endpoints $\widehat{\gamma}^*(a)$ and $\widehat{\gamma}^*(b)$, $(\widehat{\sigma}^*, u^*) \colon [a, b] \to T^*\widehat{Q} \times U$ is a solution of the HP, such that $\widehat{\gamma}^* = \pi_{\widehat{Q}} \circ \widehat{\sigma}^*$, and moreover $\widehat{\sigma}^*$ satisfies that

(3.a) $H(\widehat{\sigma}^*(t), u^*(t)) = \max_{u \in \overline{U}} H(\widehat{\sigma}^*(t), u)$ almost everywhere.

(3.b) $\max_{u\in\overline{U}} H(\widehat{\sigma}^*(t),u)$ is constant everywhere.

(3.c) $\widehat{\alpha}^*(t) \neq 0 \in T^*_{\widehat{\gamma}^*(t)}\widehat{Q}$ for every $t\in[a,b]$.

(3.d) $\alpha_0^*(t)$ is constant, $\alpha_0^*(t)\leq 0$.

Observe that it only remains to prove **(3.b)** of the Free Pontryagin's Maximum Principle, since **(3.a)**, **(3.c)** and **(3.d)** are the same in both Theorems 3.14, 5.13.

(3.b) We already know that the maximum of the Hamiltonian is constant everywhere along $(\widehat{\sigma}^*, u^*)$. Now, let us prove that it is zero everywhere.

Take $\widehat{v}_b = \pm\widehat{X}(\widehat{\gamma}^*(b), u^*(b)) \in \widehat{K}_b^{\pm}$, let $\widehat{V}\colon I\to T\widehat{Q}$ be the integral curve of $(\widehat{X}^T)^{\{u^*\}}$ with initial condition $(b, \widehat{\gamma}(b), \widehat{v}_b)$, then the continuous function $\langle\widehat{\alpha}^*, \widehat{V}\rangle\colon I\to\mathbb{R}$ is constant everywhere by Proposition B.3. Thus,

$$\langle\widehat{\alpha}^*(t), \widehat{V}(t)\rangle = \langle\widehat{\alpha}^*(t), \pm\widehat{X}(\widehat{\gamma}^*(t), u^*(t))\rangle \leq 0 \quad \text{for every } t\in I$$

by Equation (6.11), and this implies that

$$\langle\widehat{\alpha}^*(t), \widehat{X}(\widehat{\gamma}^*(t), u^*(t))\rangle = 0.$$

As $\langle\widehat{\alpha}^*(t), \widehat{X}(\widehat{\gamma}^*(t), u^*(t))\rangle = H(\widehat{\sigma}^*(t), u^*(t))$, the hamiltonian function is zero everywhere and the maximum of the hamiltonian function is zero everywhere by Theorem 3.14. □

Observe that the initial condition for the covector in this proof has been chosen such that the tangent spaces to the initial and final submanifolds are contained in the separating hyperplane defined. In this statement of the Maximum Principle the initial condition for the covector must satisfy more conditions than in Theorem 3.14. Those are the transversality conditions.

Appendices

This last part of the report is mainly devoted to state and prove some of the results used in the proof of the Maximum Principle and to give more understanding to some key points.

A Results on real functions

In this section, we focus on the analytic results used in the core part of the paper. For more details, see [30, 35].

Definition A.1. *Let (X, d_X) and (Y, d_Y) be metric spaces. A function $f\colon X\to Y$ is **Lipschitz** if there exists $K\in\mathbb{R}$ such that $d_Y(f(x_1), f(x_2)) \leq K\, d_X(x_1,x_2)$ for all $x_1, x_2\in X$.*

*A function $f\colon X\to Y$ is **locally Lipschitz** if, for every $x\in X$ there exists an open neighbourhood V of x such that $d_Y(f(x_1), f(x_2)) \leq K_x\, d_X(x_1,x_2)$ for all x_1 and x_2 in V, where K_x depends on the neighbourhood.*

If M is a differentiable manifold, let g be a Riemannian metric on M and $d_g\colon M\times M\to\mathbb{R}$ be the induced distance, then (M, d_g) is a metric space and the notion of Lipschitz on M does not depend on g. A $\mathcal{C}^1$ function $F\colon M\to\mathbb{R}$ is locally Lipschitz if, for every $p\in M$ we take the local chart (V,ϕ) such that $\phi(p)=0$, $\phi(V)=B(0,r)$, the open ball centered at the origin with radius $r > 0$ in the standard euclidean space, and $F \circ \phi^{-1} \colon B(0,r) \to \mathbb{R}$ is Lipschitz, that is, there exists $K \in \mathbb{R}$ with

$$|F(p_1) - F(p_2)| = |(F \circ \phi^{-1})(\phi(p_1)) - (F \circ \phi^{-1})(\phi(p_2))| \leq K d(\phi(p_1), \phi(p_2)), \quad \forall p_1, p_2 \in V.$$

Hence, given the local chart (V, ϕ) we define a distance on V as follows $d_\phi \colon V \times V \to \mathbb{R}$, $d_\phi(p_1, p_2) = d(\phi(p_1), \phi(p_2))$. Consequently, (V, ϕ) is a metric space with the topology induced by the open set V in M. This distance is equivalent to the distance induced by the riemannian metric on V, where $\partial/\partial x^i$ is an orthonormal basis induced by the local chart (V, x^i).

The definition of locally Lipschitz for functions on manifolds depends on the local chart, but it determines the same notion of locally Lipschitz.

Definition A.2. *A function $f \colon [a,b] \to \mathbb{R}$ is* ***uniformly continuous on*** *$[a,b]$ if, for every $\epsilon > 0$, there exists $\delta > 0$ such that for any $x, y \in [a,b]$ with $|x - y| < \delta$ we have $|f(x) - f(y)| < \epsilon$.*

Definition A.3. *A function $f \colon [a,b] \to \mathbb{R}$ is* ***absolutely continuous on*** *$[a,b]$ if, for every $\epsilon > 0$, there exists $\delta > 0$ such that for every finite number of nonoverlapping subintervals (a_i, b_i) of $[a,b]$, with $\sum_{i=1}^{n} |b_i - a_i| < \delta$, then $\sum_{i=1}^{n} |f(b_i) - f(a_i)| < \epsilon$.*

We consider an interval $I = [a,b]$ in $\mathbb{R}$, with the usual Lebesgue measure. A statement is said to be satisfied almost everywhere, if it is fulfilled in I but a zero measure set. A measurable subset $A \subset I$ is said of full measure if $I - A$ has measure zero. Recall that if $A, B \subset I$ and $I - A$, $I - B$ have measure zero, then $A \cap B$ is not empty.

Results in [30], pp. 96, 100, 105 allow to proof the following result.

Proposition A.4. *If f is absolutely continuous, then f has a derivative almost everywhere.*

Theorem A.5. *[[30], pp.105 and [35], pp.836] If f is absolutely continuous and $f'(x) = 0$ almost everywhere on $[a,b]$, then f is a constant function.*

Definition A.6. *A real-valued function f on a metric space (X, d) is called* ***lower semicontinuous at*** *$t_0 \in X$ if, for every $\epsilon > 0$, there exists a $\delta(\epsilon, t_0) > 0$ such that $f(t) \geq f(t_0) - \epsilon$ whenever $d(t, t_0) \leq \delta(\epsilon, t_0)$.*

Definition A.7. *If f is lower semicontinuous at every point of (X, d), it is said to be* ***lower semicontinuous on*** *(X, d).*

The following result is stated by Pontryagin et al. in [28], page 102, but it is neither proved nor stated as a proposition. We believe it is appropriate to write it with more detail.

Proposition A.8. *Let f and g be real functions, $f, g \colon [a,b] \to \mathbb{R}$, if f is continuous, g is lower semicontinuous, $f \leq g$ and $f = g$ almost everywhere then $f = g$ everywhere.*

Proof. Let $t_0 \in [a,b]$, as g is a lower semicontinuous on $[a,b]$, for every $\epsilon > 0$ there exists $\delta(\epsilon, t_0) = \delta > 0$ such that

$$g(t) \geq g(t_0) - \epsilon$$

whenever $|t - t_0| < \delta(\epsilon, t_0)$.

Since f and g coincide almost everywhere on $[a,b]$, there exists $t_1 \in (t_0 - \delta, t_0 + \delta)$ such that $f(t_1) = g(t_1)$. Moreover, $f \leq g$, so

$$f(t_0) \leq g(t_0) \leq g(t_1) + \epsilon = f(t_1) + \epsilon. \tag{A.13}$$

The continuity of f guarantees that for every $\epsilon' > 0$, there exists $\delta' > 0$ such that if $|t_1 - t_0| < \delta'$, then $f(t_1) - \epsilon' < f(t_0) < f(t_1) + \epsilon'$. Hence Equation (A.13) is rewritten as follows

$$f(t_0) \leq g(t_0) \leq f(t_0) + \epsilon' + \epsilon.$$

As this inequality is valid for every $\epsilon, \epsilon' > 0$, then $g(t_0) = f(t_0)$ for every $t_0 \in [a, b]$. Thus $f = g$ everywhere. □

A.1 Lebesgue points for a real function

After introducing the concept of measurable function and some properties, we will state Lebesgue's differentiation theorem, which enables us to distinguish certain points for a measurable function. In the whole paper we consider the Lebesgue measure in $\mathbb{R}$. See the book by Zaanen [36] for more details.

Definition A.9. *A function $f\colon [a, b] \to \mathbb{R}$ is **measurable** if the set $\{x \in [a, b] : f(x) > \alpha\}$ is measurable for every $\alpha \in \mathbb{R}$.*

Definition A.10. *A function $f\colon [a, b] \to \mathbb{R}$ is Lebesgue summable over each Lebesgue measurable set of finite measure if $\nu(x) = \int_a^x f d\mu$ is well defined for every $x \in [a, b]$.*

Theorem A.11. ***(Lebesgue's differentiation Theorem [36])*** *Let μ be the Lebesgue measure, if $f\colon [a, b] \to \mathbb{R}$ is a Lebesgue summable function over every Lebesgue measurable set of finite measure, then*

$$D\nu(x_+) = D\nu(x_-) = f(x)$$

holds for μ-almost every $x \in [a, b]$, where $D\nu(x_+)$, $D\nu(x_-)$ are the right and left derivatives of ν respectively.

The equality $D\nu(x_-) = f(x)$ almost everywhere may be rewritten as follows for $h > 0$

$$\lim_{h\to 0} \frac{\nu(x-h) - \nu(x)}{-h} = f(x) \quad \text{a.e.} \Leftrightarrow \lim_{h\to 0} \frac{\int_a^{x-h} f(t)dt - \int_a^x f(t)dt}{-h} = f(x) \quad \text{a.e.} \Leftrightarrow$$

$$\Leftrightarrow \lim_{h\to 0} \frac{\int_{x-h}^x f(t)dt}{h} = f(x) \quad \text{a.e.} \Leftrightarrow \int_{x-h}^x f(t)dt = hf(x) + o(h) \quad \text{a.e.} \tag{A.14}$$

Definition A.12. *If $f\colon [a, b] \to \mathbb{R}$ is a measurable function, $x \in (a, b)$ is a **Lebesgue point for* f** *if,*

$$\lim_{h\to 0} \int_{x-h}^x \frac{f(t) - f(x)}{h} dt = 0.$$

Remark A.13. As Theorem A.11 is true almost everywhere, the set of Lebesgue points for a measurable function has full measure.

Remark A.14. Observe that if $u\colon I \to U$ is measurable, then the set of Lebesgue points for u has full measure. If $f\colon U \to \mathbb{R}$ is continuous, then $f \circ u\colon I \to \mathbb{R}$ is measurable. The intersection of Lebesgue points for u and $f \circ u$ has full measure.

Note: Assume we have a manifold Q, an open set $U \subset \mathbb{R}^m$ and a continuous vector field X along the projection $\pi\colon Q \times U \to Q$. If $(\gamma, u)\colon I = [a, b] \to Q \times U$, where γ is absolutely continuous and u is measurable and bounded, then $X \circ (\gamma, u)\colon I \to TQ$ is a measurable vector field along (γ, u), in the sense that in any coordinate system its coordinate functions are measurable. For $t \in (a, b)$, it is a *Lebesgue point for u* if

$$\int_{t-h}^t X(\gamma(s), u(s))ds = hX(\gamma(t), u(t)) + o(h). \tag{A.15}$$

The Lebesgue points for a vector field will be useful in the following appendix to guarantee the differentiability of some curves, that is, the existence of its tangent vector. See [13, 14] for more details about differential equations and measurability.

B Time-dependent variational equations

The variational equations give us an approach to how the integral curves of vector fields vary when the initial condition varies along a curve. These equations have a formulation on the tangent and the cotangent bundle. Here we are interested in studying the variational equations associated to time-dependent vector fields, and in proving some relationship between the solutions of variational equations on the tangent bundle and the ones on the cotangent bundle. See [18] for definitions.

B.1 Time-dependent vector fields

For $I \subset \mathbb{R}$, let $X\colon I \times M \to TM$ be a differentiable time-dependent vector field. For every $(s,x) \in I \times M$, *the integral curve of X with initial condition* (s,x) is denoted by $\Phi^X_{(s,x)}\colon J_{(s,x)} \subset I \to M$ and satisfies

1. $\Phi^X_{(s,x)}(s) = x$.
2. $\frac{d}{dt}\big|_t \Phi^X_{(s,x)} = X(t, \Phi^X_{(s,x)}(t))$, $t \in J$.

Observe that the domain of the integral curves depends on the initial condition, that is why the domain of $\Phi^X_{(s,x)}$ is denoted by $J_{(s,x)} \subset I$.

The *time dependent flow* or *evolution operator of* X is the mapping

$$\begin{aligned} \Phi^X\colon \quad I \times I \times M &\longrightarrow M \\ (t,s,x) &\longmapsto \Phi^X(t,s,x) = \Phi^X_{(s,x)}(t) \end{aligned}$$

defined in a maximal open neighborhood $V \times M$ of $\Delta_I \times M$, where Δ_I is the diagonal of $I \times I$, and that satisfies

1. $\Phi^X(s,s,x) = x$.
2. $\frac{d}{dt}\big|_t \left(\Phi^X(t,s,x)\right) = X(t, \Phi^X(t,s,x))$.

To compute the vector field through the evolution operator, we must evaluate the last expression at $s = t$,

$$\left[\frac{d}{dt}\bigg|_t \left(\Phi^X(t,s,x)\right)\right]\bigg|_{s=t} = X(t,x).$$

There is a vector field on the manifold $I \times M$ associated to X and given by $\widetilde{X}(t,x) = \partial/\partial t_{|(t,x)} + X(t,x)$. For $(t,s,x) \in V \times M$, the flow of $\widetilde{X}$ is $\Phi^{\widetilde{X}}(t,(s,x)) = (s+t, \Phi^X(s+t,(s,x)))$. The existence and uniqueness of the flow of $\widetilde{X}$ is given by the existence and uniqueness theorems of ordinary differential equations and the smooth dependence of the initial conditions. Hence this guarantees the existence and uniqueness of the maximally defined evolution operator Φ^X.

For a fixed $(t,s) \in V \subset I \times I$,

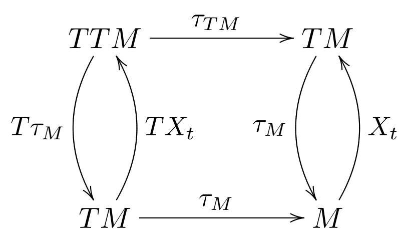
$$\begin{aligned} \Phi^X_{(t,s)}\colon M &\longrightarrow M \\ x &\longmapsto \Phi^X_{(t,s)}(x) = \Phi^X_{(s,x)}(t) \end{aligned}$$

is a diffeomorphism satisfying $\Phi^X_{(t,s)} = \Phi^X_{(t,r)} \circ \Phi^X_{(r,s)}$ for $r \in I$, whenever one side makes sense. For more details see I. Kolář et. [18].

B.2 Complete lift

Once the evolution operator of time-dependent vector fields is defined, we determine the evolution operator of its tangent or complete lift.

Let $X_t\colon M \to TM$ be a vector field on M such that $X_t(x) = X(t,x)$ for every $t \in I$. Consider the following commutative diagram:



If $(x,v) \in TM$, $TX_t(x,v) = (x, X_t(x), T_xX_t(v)) \in T_{(x,X_t(x))}(TM)$.

The *complete lift of* X_t *to* TM is the time-dependent vector field X_t^T on TM satisfying

$$X_t^T = \kappa_M \circ TX_t,$$

where κ_M is the canonical involution of TTM, see [18] for the definition.

Let (x^i) be local coordinates in M and $X_t = X_t^i\,\partial/\partial x^i$. Let (x^i, v^i) be the induced local coordinates in TM. Locally, the complete lift X^T of X is given by

$$X^T(t,x,v) = X^i(t,x)\,\frac{\partial}{\partial x^i}\bigg|_{(x,v)} + \frac{\partial X^i}{\partial x^j}(t,x)\,v^j\,\frac{\partial}{\partial v^i}\bigg|_{(x,v)}.$$

The equations satisfied by the integral curves of X^T are called *variational equations.*

Proposition B.1. *If* Φ^X *is the evolution operator of* X*, then* $\Psi\colon I \times I \times TM \to TM$ *such that*

$$\Psi(t,s,(x,v)) = \left(\Phi^X(t,s,x), T_x\Phi^X_{(t,s)}(v)\right)$$

is the evolution operator of X^T*.*

Proof. We have to prove that

$$\begin{cases} \Psi(s,s,(x,v)) = (x,v) \\ \\ \frac{d}{dt}\Big|_t \left(\Psi(t,s,(x,v))\right) = X^T(t,\Psi(t,s,(x,v))). \end{cases}$$

The first item is proved easily,

$$\Psi(s,s,(x,v)) = (\Phi^X(s,s,x), T_x\left(\Phi^X_{(s,s)}\right)(v)) = (x, Id\ (v)) = (x,v).$$

As $\Phi_s^X\colon I \times M \to M$ is $\mathcal{C}^\infty$, in local coordinates we have

$$T_t\left(T_x\Phi^X_{(t,s)}(v)\right)1 = \frac{d}{dt}\bigg|_t \left(T_x\Phi^X_{(t,s)}(v)\right)1 = \left(T_x\left(\frac{d}{dt}\bigg|_t\left(\Phi^X_{(t,s)}\right)1\right)\right)(v) = T_x\left(T_t\Phi^X_{(s,x)}1\right)(v).$$

Let us prove the second item using the last equality.

$$\frac{d}{dt}\Big|_t (\Psi(t,s,x,v)) = \left(\frac{d}{dt}\Big|_t \left(\Phi^X(t,s,x)\right), \frac{d}{dt}\Big|_t \left(T_x\Phi^X_{(t,s)}(v)\right)\right)$$

$$= \left(X(t,\Phi^X(t,s,x)), \left(T_x\left(\frac{d}{dt}\Big|_t \left(\Phi^X_{(t,s)}\right)\right)\right)(v)\right) = \left(X(t,\Phi^X(t,s,x)), \left(T_x\left(X_t(\Phi^X(t,s,x))\right)\right)(v)\right)$$

$$= \left(X(t,\Phi^X(t,s,x)), \left(T_{\Phi^X(t,s,x)}(X_t) \circ T_x(\Phi^X(t,s,x))\right)(v)\right)$$

$$= \left(X(t,\Phi^X(t,s,x)), T_{\Phi^X(t,s,x)}(X_t)\left(T_x(\Phi^X_{(t,s)})(v)\right)\right) = X^T(t,\Psi(t,s,x,v)).$$

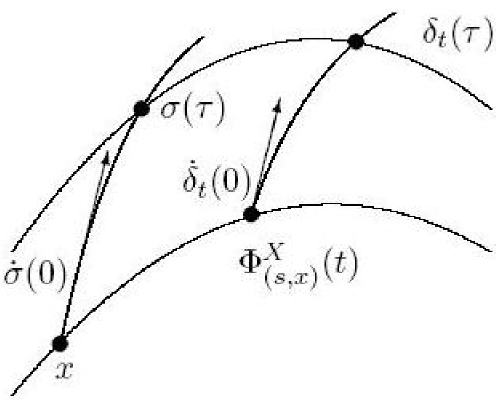
□

Hence, the evolution operator of X^T is the complete lift of the evolution operator of X. The integral curves of X^T are vector fields along the integral curves of X.

B.2.1 About the geometric meaning of the complete lift

In this paragraph we give a rough geometric interpretation of this last idea. We are going to observe that the integral curves of X^T are the linear approximation of the integral curves of X when the initial condition varies along a curve in M.

Let $\Phi^X_{(s,x)}\colon I \to M$ be the integral curve of X with initial condition (s,x), and $\sigma\colon(-\epsilon,\epsilon)\to M$ be a curve $\mathcal{C}^\infty$ such that $\sigma(0)=x$.



The integral curve of X with initial condition $(s,\sigma(\tau))$ is given by $\Phi^X_{(s,\sigma(\tau))} : I \to M$. For $t\in I$, let $\delta_t\colon(-\epsilon,\epsilon)\to M$ be a curve such that

- $\delta_t(\tau) = \Phi^X_{(s,\sigma(\tau))}(t)$,
- $\delta_s(\tau) = \sigma(\tau)$,
- $\delta_t(0) = \Phi^X_{(s,x)}(t)$,

Let us prove that $\dot{\delta}_t(0) = T_x\Phi^X_{(t,s)}(\dot{\sigma}(0))$.

$$\dot{\delta}_t(0) = (T_0\delta_t(\tau))\,\frac{d}{d\tau}\Big|_0 = \left(T_0\left(\Phi^X_{(s,\sigma(\tau))}(t)\right)\right)\frac{d}{d\tau}\Big|_0 = \left(T_0\left(\Phi^X_{(t,s)}(\sigma(\tau))\right)\right)\frac{d}{d\tau}\Big|_0$$

$$= T_{\sigma(0)}\Phi^X_{(t,s)}\left(T_0(\sigma(\tau))\,\frac{d}{d\tau}\Big|_0\right) = T_{\sigma(0)}\Phi^X_{(t,s)}\left(\dot{\sigma}(0)\right) = T_x\Phi^X_{(t,s)}\left(\dot{\sigma}(0)\right).$$

If $\dot{\sigma}(0) = v$, then $(\delta_\cdot(0), \dot{\delta}_\cdot(0))\colon I \to TM$, $t \mapsto (\delta_t(0), \dot{\delta}_t(0))$ is the integral curve of X^T with initial condition $(s, (x, v)) = (s, (\sigma(0), \dot{\sigma}(0)))$.

B.3 Cotangent lift

Given $(t, s) \in I \times I$, the evolution operator $\Phi^{X^T}_{(t,s)}$ is a diffeomorphism on TM and a linear isomorphism on the fibers on TM, so it makes sense to consider its transpose and inverse, $({}^{\tau}\Phi^{X^T}_{(t,s)})^{-1} = \Lambda_{(t,s)}$. It is a linear isomorphism on the fibers on T^*M and satisfies $\Lambda_{(t,s)} = \Lambda_{(t,r)} \circ \Lambda_{(r,s)}$ for $r \in I$. Hence $\Lambda\colon I \times I \times T^*M \to T^*M$ is the evolution operator of a time-dependent vector field on T^*M, called the cotangent lift X^{T^*} of X to T^*M. In other words, since the complete lift X^T is a linear vector field over X, the cotangent lift $X^{T^*}\colon I \times T^*M \to TT^*M$ is the dual of X^T. See pages 379-381 in the book by I. Kolář *et al.* [18] for more details.

Another vector field on T^*M may be associated to X as follows. For every $t \in I$, the hamiltonian vector field Z_t of the hamiltonian system $(M, \omega, \widehat{X_t})$ is a vector field on T^*M associated to X. If (x^i, p_i) are the induced local coordinates in T^*M, the function $\widehat{X_t}\colon T^*M \to \mathbb{R}$ is given by $\widehat{X_t}(x,p) = X^i(t,x)p_i$. Let ω be the natural symplectic structure on T^*M. The Hamilton equations are given by $i_{Z_t}\omega = d\widehat{X_t}$.

Locally, $Z\colon I \times T^*M \to TT^*M$ is given by

$$Z(t,x,p) = X^i(t,x)\, \frac{\partial}{\partial x^i}\bigg|_{(x,p)} - \frac{\partial X^j}{\partial x^i}(t,x)\, p_j\, \frac{\partial}{\partial p_i}\bigg|_{(x,p)}.$$

The equations satisfied by the integral curves of Z in the fibers are the *adjoint variational equations on the cotangent bundle.* In the literature, they are sometimes called *adjoint equations.*

Let us prove that both vector fields Z and X^{T^*} associated to X are the same.

Proposition B.2. *If Φ^X is the evolution operator of X, then $\Lambda\colon I \times I \times T^*M \to T^*M$ such that*

$$\Lambda(t,s,(x,p)) = (\Phi^X(t,s,x), \left({}^{\tau}T_x\Phi^X_{(t,s)}\right)^{-1}(p))$$

is the evolution operator of Z. Thus $Z = X^{T^*}$.

Proof. We have to prove that

$$\begin{cases} \Lambda(s,s,(x,p)) = (x,p) \\ \frac{d}{dt}\big|_t (\Lambda(t,s,(x,p))) = Z(t, \Lambda(t,s,(x,p))). \end{cases}$$

The first item is proved easily,

$$\Lambda(s,s,(x,p)) = \left(\Phi^X(s,s,x), \left({}^{\tau}T_x\Phi^X_{(s,s)}\right)^{-1}(p)\right) = (x, Id\ (p)) = (x,p).$$

As $\Phi^X_s\colon I \times M \to M$ is $\mathcal{C}^\infty$, in local coordinates we have

$$\frac{d}{dt}\bigg|_t \left({}^{\tau}T_x\Phi^X_{(t,s)}\right) = {}^{\tau}T_x\left(\frac{d}{dt}\bigg|_t \Phi^X_{(t,s)}\right),$$

considering that

$$\frac{d}{dt}\bigg|_t \left({}^{\tau}T_x\Phi^X_{(t,s)}\right)\colon T^*_xM \to T^*_{\Phi^X(t,(s,x))}M, \quad \text{and} \quad {}^{\tau}T_x\left(\frac{d}{dt}\bigg|_t \Phi^X_{(t,s)}\right)\colon T^*_xM \to T^*_{\Phi^X(t,(s,x))}M.$$

Let us prove the second item using the last equality:

$$\begin{aligned}
&\left.\tfrac{d}{dt}\right|_t (\Lambda(t,s,x,p)) = \left(\left.\tfrac{d}{dt}\right|_t \left(\Phi^X(t,s,x)\right), \left.\tfrac{d}{dt}\right|_t \left(\left({}^\tau T_x \Phi^X_{(t,s)} \right)^{-1} (p) \right) \right) \\
&= \left(X(t,\Phi^X(t,s,x)), \left(- \left({}^\tau T_x \Phi^X_{(t,s)} \right)^{-1} \circ \left(\left.\tfrac{d}{dt}\right|_t \left({}^\tau T_x \Phi^X_{(t,s)} \right) \right) \circ \left({}^\tau T_x \Phi^X_{(t,s)} \right)^{-1} \right) (p) \right) \\
&= \left(X(t,\Phi^X(t,s,x)), \left(- \left({}^\tau T_x \Phi^X_{(t,s)} \right)^{-1} \circ \left({}^\tau T_x \left(\left.\tfrac{d}{dt}\right|_t \Phi^X_{(t,s)} \right) \right) \right) \left(\left({}^\tau T_x \Phi^X_{(t,s)} \right)^{-1} (p) \right) \right) \\
&= \left(X(t,\Phi^X(t,s,x)), \left(- \left({}^\tau T_{\Phi^X_{(t,s)}(x)} \left(\Phi^X_{(t,s)} \right)^{-1} \right) \circ \left({}^\tau T_x \left(X_t \circ \Phi^X_{(t,s)} \right) \right) \right) \left(\left({}^\tau T_x \Phi^X_{(t,s)} \right)^{-1} (p) \right) \right) \\
&= \left(X(t,\Phi^X(t,s,x)), - \left({}^\tau T_{\Phi^X_{(t,s)}(x)} \left(X_t \circ \Phi^X_{(t,s)} \circ \left(\Phi^X_{(t,s)} \right)^{-1} \right) \right) \left(\left({}^\tau T_x \Phi^X_{(t,s)} \right)^{-1} (p) \right) \right) \\
&= \left(X(t,\Phi^X(t,s,x)), - \left({}^\tau T_{\Phi^X_{(t,s)}(x)} (X_t) \right) \left(\left({}^\tau T_x \Phi^X_{(t,s)} \right)^{-1} (p) \right) \right) = Z(t,\Lambda(t,s,x,p)).
\end{aligned}$$

Hence, the evolution operator of Z is the cotangent lift of the evolution operator of X. Thus $Z = X^{T^*}$. □

B.4 A property for the complete and cotangent lift

The previous propositions allow to determine an invariant function along integral curves of X.

Proposition B.3. *Let $X\colon I\times M \to TM$ be a time-dependent vector field, $X^T\colon I\times TM \to TTM$ and $X^{T^*}\colon I \times T^*M \to TT^*M$ be the complete lift and cotangent lift of X, respectively. If $\gamma\colon I \to M$ is an integral curve of X with initial condition (s,x), $(\gamma,u)\colon I \to TM$ is the integral curve of X^T with initial condition $(s,(x,v))$, $v \in T_{\gamma(s)}M$, and $(\gamma,\alpha)\colon I \to T^*M$ is the integral curve of X^{T^*} with initial condition $(s,(x,p))$, $p \in T^*_{\gamma(s)}M$, then*

$$\begin{array}{rcl} \langle \alpha, u\rangle\colon I & \to & \mathbb{R} \\ t & \mapsto & \langle \alpha(t), u(t)\rangle \end{array}$$

is constant.

Proof. If Φ^X is the evolution operator of X, the evolution operators of X^T and X^{T^*} are

$$\Phi^{X^T}(t,s,(x,v)) = (\Phi^X(t,s,x), T_x\Phi^X_{(t,s)}(v)),$$

$$\Phi^{X^{T^*}}(t,s,(x,p)) = (\Phi^X(t,s,x), \left({}^\tau T_x \Phi^X_{(t,s)} \right)^{-1} (p)),$$

respectively. Hence

$$\begin{aligned}
\langle \alpha(t), u(t)\rangle &= \left\langle \left({}^\tau T_x \Phi^X_{(t,s)} \right)^{-1} (p), T_x\Phi^X_{(t,s)}(v) \right\rangle = \left\langle {}^\tau \left(\left(T_x \Phi^X_{(t,s)} \right)^{-1} \right) (p), T_x\Phi^X_{(t,s)}(v) \right\rangle = \\
&= \left\langle p, \left(\left(T_x \Phi^X_{(t,s)} \right)^{-1} \circ \left(T_x \Phi^X_{(t,s)} \right) \right) (v) \right\rangle = \langle p, v\rangle = \text{constant.}
\end{aligned}$$

□

C The tangent perturbation cone as an approximation of the reachable set

In dynamical systems, the reachable sets are useful to determine the controllability of the systems. In more applied areas, as optimal control, the reachable set has great interest for studying optimal curves through Pontryagin's Maximum Principle. A key point of the proof of the

Maximum Principle is to understand the so-called tangent perturbation cone along a reference trajectory, that is a solution of the dynamical system, as an approximation of the reachable set in a neighborhood of a point in the reference trajectory. This interpretation will become slightly clearer at the end of this appendix. Hence, the cone gives locally how to arrive at reachable points from another. In [1] this approximation is studied.

In the sequel, we explain why this interpretation of the tangent perturbation cone is feasible. First of all, we need a general result about the flow of time-dependent vector fields. Then we concentrate on clarifying the meaning of the title of this appendix.

Let M be a manifold and I be an interval of $\mathbb{R}$. Let X be a time-dependent vector field on M. The flow of X is a mapping $\Phi^X\colon I\times I\times M\to M$, $(t,s,x)\mapsto\Phi^X(t,s,x)$ as defined in Appendix B.1.

For a fixed initial time s, according to Appendix B.1, we define the mapping $\Phi^X_s\colon I\times M\to M$, $(t,x)\mapsto\Phi^X(t,s,x)$, satisfying $\Phi^X_s(s,x)=x$, $\Phi^X_{(s,x)}\colon I\to M$, $\Phi^X_{(t,s)}\colon M\to M$.

The vector field X can be interpreted as a time-independent vector field $\overline{X}\in\mathfrak{X}(I\times M)$ such that $\overline{X}(t,x)=\left.\dfrac{\partial}{\partial t}\right|_{(t,x)}+X(t,x)$. The flow $\mathrm{Fl}^{\overline{X}}\colon I\times I\times M\to I\times M$ of $\overline{X}$ is the mapping $\mathrm{Fl}^{\overline{X}}$ defined by $(r,s,x)\mapsto\mathrm{Fl}^{\overline{X}}(r,s,x)$, such that

(i) the initial condition is (s,x) at time 0, that is, $\mathrm{Fl}^{\overline{X}}(0,s,x)=(s,x)$;

(ii) $\mathrm{Fl}^{\overline{X}}_{(s,x)}\colon I\to I\times M$, $\mathrm{Fl}^{\overline{X}}_{(s,x)}(r)=\mathrm{Fl}^{\overline{X}}(r,s,x)$, is the integral curve of $\overline{X}$ with initial condition (s,x) at time 0, that is,

$$\left.\frac{\mathrm{d}}{\mathrm{d}r}\right|_r \mathrm{Fl}^{\overline{X}}_{(s,x)}(r)=\overline{X}(\mathrm{Fl}^{\overline{X}}_{(s,x)}(r))\,,\quad \mathrm{Fl}^{\overline{X}}_{(s,x)}(0)=(s,x)\,;$$

(iii) $\mathrm{Fl}^{\overline{X}}_s\colon I\times M\to I\times M$, $\mathrm{Fl}^{\overline{X}}_s(r,x)=\mathrm{Fl}^{\overline{X}}(r,s,x)$, is a diffeomorphism on $I\times M$.

The flows of X and $\overline{X}$ are related as follows

$$\mathrm{Fl}^{\overline{X}}(t-s,s,x)=(t,\Phi^X(t,s,x))\,,\quad \mathrm{Fl}^{\overline{X}}(r,s,x)=(r+s,\Phi^X(r+s,s,x))\,.$$

Let X, Y be time-dependent vector fields, we are interested in computing a time-dependent vector field Z that allows to write

$$\mathrm{Fl}^{\overline{X}+\overline{Y}}_s(t,x)=(\mathrm{Fl}^{\overline{X}}_s\circ\mathrm{Fl}^{\overline{Z}}_s)(t,x)\;,$$

due to the relationships of the flows this is equivalent to write

$$(t+s,\Phi^{X+Y}(t+s,s,x))=(\mathrm{Fl}^{\overline{X}}_s)(t+s,\Phi^Z(t+s,s,x))=(t+2s,\Phi^X(t+2s,s,\Phi^Z(t+s,s,x)))\,.$$

It makes sense if and only if $s=0$, then

$$(t,\Phi^{X+Y}(t,0,x))=(t,\Phi^X(t,0,\Phi^Z(t,0,x)))\,. \tag{C.16}$$

If we define the mapping $\tilde{\Phi}^X\colon I\times M\to I\times M$, $(t,x)\to(t,\Phi^X(t,0,x))=(t,\Phi^X_0(t,x))$ that satisfies $\tilde{\Phi}^X(0,x)=(0,x)$, then Equation (C.16) can be rewritten as follows

$$\tilde{\Phi}^{X+Y}(t,x)=(\tilde{\Phi}^X\circ\tilde{\Phi}^Z)(t,x)\,.$$

This expression has been assumed as true in [1, 10], but it has not been carefully proved.

Observe that an integral curve of a given vector field with initial condition at time s will be an integral curve of another vector field with initial condition at time 0. That is, if $\Phi^X_{(s,y)}(t)$ is the integral curve of the time-dependent vector field X with initial condition y at $t = s$, then $\Phi^X_{(s,y)}(t-s)$ is the integral curve of $Y\colon I \times M \to TM$, given by $Y(t,x) = X(t-s,x)$, with initial condition y at $t = 0$. In other words, $\Phi^Y(t,0,y) = \Phi^X(t-s,s,y)$. Thus, the initial time could be assumed to be equal to 0.

But, for the sake of generality we do not assume $s = 0$. For a fixed initial condition s, we also define the mapping $\tilde{\Phi}^X_s\colon I \times M \to I \times M$, $(t,x) \to (t, \Phi^X_s(t,x))$ that satisfies $\tilde{\Phi}^X_s(s,x) = (s,x)$. Then, Equation (C.16) can be written as

$$\tilde{\Phi}^{X+Y}_s(t,x) = (\tilde{\Phi}^X_s \circ \tilde{\Phi}^Z_s)(t,x)\ . \tag{C.17}$$

On the left-hand side of Equation (C.17) we have

$$\tilde{\Phi}^{X+Y}_s(t,x) = (t, \Phi^{X+Y}_s(t,x))$$

and the right-hand side is

$$(\tilde{\Phi}^X_s \circ \tilde{\Phi}^Z_s)(t,x) = \tilde{\Phi}^X_s(t, \Phi^Z_s(t,x)) = (t, \Phi^X_s(t, \Phi^Z_s(t,x)))\ .$$

Thus Equation (C.17) is satisfied if and only if

$$\Phi^{X+Y}_s(t,x) = \Phi^X_s(t, \Phi^Z_s(t,x)) = (\Phi^X_s \circ \tilde{\Phi}^Z_s)(t,x)\ , \tag{C.18}$$

$$\Phi^{X+Y}_{(t,s)} = \Phi^X_{(t,s)} \circ \Phi^Z_{(t,s)}\ . \tag{C.19}$$

Let us differentiate with respect to t the left-hand side of Equation (C.18),

$$\frac{\mathrm{d}}{\mathrm{d}t}\Phi^{X+Y}_{(s,x)}(t) = (X+Y)(t, \Phi^{X+Y}_{(s,x)}(t)) = (X+Y)(t, \Phi^X_s(t, \Phi^Z_s(t,x)))\ . \tag{C.20}$$

The differentiation with respect to time of the right-hand side of Equation (C.18), for f in $\mathcal{C}^\infty(M)$, is

$$\frac{\mathrm{d}}{\mathrm{d}t}(\Phi^X_s(t, \Phi^Z_s(t,x)))f = \lim_{h\to 0} \frac{f((\Phi^X_s(t+h, \Phi^Z_s(t+h,x)))) - f((\Phi^X_s(t, \Phi^Z_s(t,x))))}{h}$$

$$= \lim_{h\to 0}\left\{ \frac{(f\circ\Phi^X_{(t+h,s)})(\Phi^Z_{(t+h,s)}(x)) - (f\circ\Phi^X_{(t+h,s)})(\Phi^Z_{(t,s)}(x))}{h}\right.$$

$$\left. + \frac{(f\circ\Phi^X_s)(t+h, \Phi^Z_s(t,x)) - (f\circ\Phi^X_s)(t, \Phi^Z_s(t,x))}{h}\right\}$$

$$= Z(t, \Phi^Z_s(t,x))(f\circ\Phi^X_{(t,s)}) + X(t, \Phi^X_s(t, \Phi^Z_s(t,x)))f$$

$$= \mathrm{T}_{\Phi^Z_s(t,x)}\Phi^X_{(t,s)} Z(t, \Phi^Z_s(t,x))f + X(t, \Phi^X_s(t, \Phi^Z_s(t,x)))f\ .$$

Hence

$$\frac{\mathrm{d}}{\mathrm{d}t}(\Phi^X_s(t, \Phi^Z_s(t,x))) = \mathrm{T}_{\Phi^Z_s(t,x)}\Phi^X_{(t,s)} Z(t, \Phi^Z_s(t,x)) + X(t, \Phi^X_s(t, \Phi^Z_s(t,x)))\ .$$

From Equation (C.20) we have

$$X(t, \Phi^X_s(t, \Phi^Z_s(t,x))) + Y(t, \Phi^X_s(t, \Phi^Z_s(t,x))) = \mathrm{T}_{\Phi^Z_s(t,x)}\Phi^X_{(t,s)} Z(t, \Phi^Z_s(t,x)) + X(t, \Phi^X_s(t, \Phi^Z_s(t,x)))\ ,$$

that is,

$$Y((\tilde{\Phi}_s^X \circ \tilde{\Phi}_s^Z)(t,x)) = \mathrm{T}_{\Phi_s^Z(t,x)} \Phi_{(t,s)}^X Z(t, \Phi_s^Z(t,x)) \; .$$

The push-forward of a time-dependent vector field Z is another time-dependent vector field given by

$$(\Phi_{(t,s)*}^X Z)(t,x) = \mathrm{T}_{(\Phi_{(t,s)}^X)^{-1}(x)} \Phi_{(t,s)}^X (Z(t, (\Phi_{(t,s)}^X)^{-1}(x))) \; .$$

Then

$$(Y \circ \tilde{\Phi}_s^X \circ \tilde{\Phi}_s^Z)(t,x) = (\Phi_{(t,s)*}^X Z)(t, \Phi_{(t,s)}^X(\Phi_s^Z(t,x)))$$

$$= (\Phi_{(t,s)*}^X Z)(\tilde{\Phi}_s^X(t, \Phi_s^Z(t,x))) = (\Phi_{(t,s)*}^X Z)(\tilde{\Phi}_s^X \circ \tilde{\Phi}_s^Z)(t,x) \; ,$$

or equivalently,

$$Y \circ \tilde{\Phi}_s^X \circ \tilde{\Phi}_s^Z = (\Phi_{(t,s)*}^X Z) \circ \tilde{\Phi}_s^X \circ \tilde{\Phi}_s^Z \; ,$$

that is, $Y = \Phi_{(t,s)*}^X Z$.

Hence $Z = (\Phi_{(t,s)*}^X)^{-1} Y = (\Phi_{(t,s)}^X)^* Y$. Now, going back to Equation (C.19) we have

$$\Phi_{(t,s)}^{X+Y}(x) = (\Phi_{(t,s)}^X \circ \Phi_{(t,s)}^{(\Phi_{(t,s)}^X)^* Y})(x) \; . \tag{C.21}$$

Once these initial computations have been made, we consider a dynamical system given by a vector field depending on controls, that is, we look for integral curves of a vector field $X \colon M \times U \to TM$ where U is a subset of $\mathbb{R}^m$.

Let q_0 be a point in M determining an initial condition at time t_0, we consider the whole set of reachable points from q_0 at time t_1 that is obtained as follows

(1) take an admissible control $u \colon \mathbb{R} \to U$;

(2) consider the time-dependent vector field $X^u \colon I \times M \to TM$, $X^u(t,q) = X(q, u(t))$;

(3) solve the Cauchy problem given by X^u with initial condition q_0 at t_0;

(4) evaluate the integral curve at time t_1;

and repeat this process for all admissible controls u.

Let u be a control, consider the integral curve of X^u with initial condition q_0 at t_0 denoted by γ, the reference trajectory. Then $\gamma(t_1)$ is a reachable point from q_0 at time t_1. Let us consider another control $\tilde{u} \colon I \to U$ and the integral curve of $X^{\tilde{u}}$ with initial condition q_0 at t_0 denoted by $\tilde{\gamma}$. Then $\tilde{\gamma}(t_1)$ is another reachable point from q_0 at time t_1.

Let us see how to reach the point $\tilde{\gamma}(t_1)$ using Equation (C.21),

$$\begin{aligned}
\tilde{\gamma}(t_1) &= \Phi_{(t_1,t_0)}^{X^{\tilde{u}}}(q_0) = \Phi_{(t_1,t_0)}^{X^u + (X^{\tilde{u}} - X^u)}(q_0) = (\Phi_{(t_1,t_0)}^{X^u} \circ \Phi_{(t_1,t_0)}^{(\Phi_{(t_1,t_0)}^{X^u})^*(X^{\tilde{u}} - X^u)})(q_0) \\
&= (\Phi_{(t_1,t_0)}^{X^u} \circ \Phi_{(t_1,t_0)}^{(\Phi_{(t_1,t_0)}^{X^u})^*(X^{\tilde{u}} - X^u)} \circ (\Phi_{(t_1,t_0)}^{X^u})^{-1} \circ \Phi_{(t_1,t_0)}^{X^u})(q_0) \\
&= (\Phi_{(t_1,t_0)}^{X^u} \circ \Phi_{(t_1,t_0)}^{(\Phi_{(t_1,t_0)}^{X^u})^*(X^{\tilde{u}} - X^u)} \circ (\Phi_{(t_1,t_0)}^{X^u})^{-1})(\gamma(t_1)) \; .
\end{aligned} \tag{C.22}$$

Hence, from $\gamma(t_1)$ we can get every reachable point from q_0 at time t_1 through Equation (C.22), that is, composing integral curves of the vector fields X^u and $(\Phi_{(t_1,t_0)}^{X^u})^*(X^{\tilde{u}} - X^u) \colon I \times M \to TM$, this latter with initial condition q_0 at t_0.

In fact this is true for any time τ in $[t_0, t_1]$, that is,

$$\tilde{\gamma}(\tau) = (\Phi^{X^u}_{(\tau,t_0)} \circ \Phi^{(\Phi^{X^u}_{(\tau,t_0)})^*(X^{\tilde{u}}-X^u)}_{(\tau,t_0)} \circ (\Phi^{X^u}_{(\tau,t_0)})^{-1})(\gamma(\tau)) \,.$$

If we compose with the flow of X^u, we get a reachable point from q_0 at time t_1 because it is a concatenation of integral curves of the dynamical system,

$$\begin{aligned} \Phi^{X^u}_{(t_1,\tau)}(\tilde{\gamma}(\tau)) &= (\Phi^{X^u}_{(t_1,\tau)} \circ \Phi^{X^u}_{(\tau,t_0)} \circ \Phi^{(\Phi^{X^u}_{(\tau,t_0)})^*(X^{\tilde{u}}-X^u)}_{(\tau,t_0)} \circ (\Phi^{X^u}_{(\tau,t_0)})^{-1})(\gamma(\tau)) \\ &= (\Phi^{X^u}_{(t_1,t_0)} \circ \Phi^{(\Phi^{X^u}_{(\tau,t_0)})^*(X^{\tilde{u}}-X^u)}_{(\tau,t_0)} \circ (\Phi^{X^u}_{(t_1,t_0)})^{-1})(\gamma(t_1)) \,. \end{aligned} \tag{C.23}$$

Hence, from $\gamma(t_1)$ we can also get reachable points from q_0 at time t_1 through composition of integral curves of the vector fields X^u and $(\Phi^{X^u}_{(\tau,t_0)})^*(X^{\tilde{u}} - X^u)$, the latter with initial condition $\gamma(t_0)$ at time t_0.

On the other hand, the tangent perturbation cone at $\gamma(t_1)$ is given by the closure of the convex hull of all the tangent vectors $(\Phi^{X^u}_{(t_1,\tau)})_*(X^{\tilde{u}}(\tau,\gamma(\tau)) - X^u(\tau,\gamma(\tau)))$ for every Lebesgue time τ in $[t_0, t_1]$. These vectors are related with the vector fields X^u through Equations (C.22) and (C.23).

In this sense, we say that the tangent perturbation cone at $\gamma(t_1)$ is an approximation of the reachable set in a neighborhood of $\gamma(t_1)$.

D Convex sets, cones and hyperplanes

We study some properties satisfied by convex sets and cones, see [5, 29] for details. Unless otherwise is stated, we suppose that all the sets are in a n-dimensional vector space E. We need to define the different kinds of cones and linear combinations used in this report.

Definition D.1. *A **cone** C with vertex at $0 \in E$ satisfies that if $v \in C$, then $\lambda\, v \in C$ for every $\lambda \geq 0$.*

Definition D.2. *Given a family of vectors $V \subset E$.*

1. *A **conic non-negative combination** of elements in V is a vector of the form $\lambda_1 v_1 + \ldots, \lambda_r v_r$, with $\lambda_i \geq 0$ and $v_i \in V$ for all $i \in \{1, \ldots, r\}$.*
2. *The **convex cone** generated by V is the set of all conic non-negative combinations of vectors in V.*
3. *An **affine combination** of elements in V is a vector of the form $\lambda_1 v_1 + \ldots, \lambda_r v_r$, with $v_i \in V$, $\lambda_i \in \mathbb{R}$ for alli $\in \{1, \ldots, r\}$ and $\sum_{i=1}^r \lambda_i = 1$.*
4. *A **convex combination** of elements in V is a vector of the form $\lambda_1 v_1 + \ldots, \lambda_r v_r$, with $v_i \in V$, $0 \leq \lambda_i \leq 1$ for all $i \in \{1, \ldots, r\}$ and $\sum_{i=1}^r \lambda_i = 1$.*

Remember that a set $A \subset E$ is *convex* if given two different elements in A, then all their convex combinations are contained in A. Thus, all the convex combination of elements in A are in A.

Definition D.3. *The **convex hull** of a set $A \subset E$, $\mathrm{conv}(A)$, is the smallest convex subset containing A.*

Let us prove another characterization of the convex hull that will be useful.

Proposition D.4. *The convex hull of a set A is the set of the convex combinations of elements in A.*

Proof. Let us denote by C the set of all convex combinations of elements in A. First, we prove that C is a convex set. If x, y are in C, then they are convex combinations of elements in A: $x = \sum_{i=1}^{k} \lambda_i v_i$, $y = \sum_{i=1}^{r} \mu_i w_i$, with $\sum_{i=1}^{k} \lambda_i = 1$, $\sum_{i=1}^{r} \mu_i = 1$. For $s \in (0,1)$, consider

$$sx + (1-s)y = s\left(\sum_{i=1}^{k} \lambda_i v_i\right) + (1-s)\left(\sum_{i=1}^{r} \mu_i w_i\right),$$

that will be in C if the sum of the coefficients is equal to 1 and each of the coefficients lies in $[0,1]$. Observe that $s\sum_{i=1}^{k} \lambda_i + (1-s)\sum_{i=1}^{r} \mu_i = s + (1-s) = 1$ and the other condition is satisfied trivially. As C is convex and contains A, the convex hull of A is a subset of C.

Second, we prove that $C \subset \operatorname{conv}(A)$ by induction of the number of vectors in the convex combinations of elements in A. Trivially, when the convex combination is given by an element in A, it lies in the convex hull of A.

Now, suppose that a convex combination of $k-1$ elements of A is in $\operatorname{conv}(A)$, and we prove that a convex combination of k elements of A is in $\operatorname{conv}(A)$: $x = \sum_{i=1}^{k} \mu_i v_i = \sum_{i=1}^{k-1} \mu_i v_i + \mu_k v_k$. If $\sum_{i=1}^{k-1} \mu_i = 0$, then $\mu_k = 1$. By the first step of induction, x is in $\operatorname{conv}(A)$. If $\sum_{i=1}^{k-1} \mu_i \in (0,1]$, then $\mu_k \in [0,1)$ and we can rewrite x as

$$x = (1-\mu_k)\left(\sum_{i=1}^{k-1} \mu_i (1-\mu_k)^{-1} v_i\right) + \mu_k v_k.$$

Observe that $\sum_{i=1}^{k-1} \mu_i (1-\mu_k)^{-1} = (1-\mu_k)(1-\mu_k)^{-1} = 1$, then $\sum_{i=1}^{k-1} \mu_i (1-\mu_k)^{-1} v_i$ is in $\operatorname{conv}(A)$. By the first step of induction, v_k is in $\operatorname{conv}(A)$. As $(1-\mu_k) + \mu_k = 1$, x is in $\operatorname{conv}(A)$ because of the convexity of $\operatorname{conv}(A)$. Thus $C \subset \operatorname{conv}(A)$ and $C = \operatorname{conv}(A)$. □

Proposition D.5. *Let C be a convex set. If $\overline{C}$ and $\operatorname{int} C$ are the topological closure and the interior of C, respectively, we have:*

(a) For every $x \in \operatorname{int} C$, if $y \in \overline{C}$, then $(1-\lambda)x + \lambda y \in \operatorname{int} C$ for all $\lambda \in [0,1)$.

(b) $\overline{C} = \overline{\operatorname{int} C}$.

(c) The interior of C is empty if and only if the interior of $\overline{C}$ is empty.

(d) $\operatorname{int} C = \operatorname{int} \overline{C}$.

Proof. **(a)** If $x \in \operatorname{int}\ C$, then there exists $\epsilon_x > 0$ such that $B(x, \epsilon_x) \subset C$, where $B(x, \epsilon_x)$ denotes the open ball centered at x of radius ϵ_x.

If $y \in \overline{C}$, for all $\epsilon > 0$, $y \in C + \epsilon B(0,1)$.

For every $\lambda \in [0,1)$, we consider $x_\lambda = (1-\lambda)x + \lambda y$. Let us compute the value of ϵ_λ such that $x_\lambda + \epsilon_\lambda B(0,1) \subset C$.

$$x_\lambda + \epsilon_\lambda B(0,1) = (1-\lambda)x + \lambda y + \epsilon_\lambda B(0,1) \subset$$

$$\subset (1-\lambda)x + \lambda C + \lambda \epsilon B(0,1) + \epsilon_\lambda B(0,1) = (1-\lambda)x + (\lambda\epsilon + \epsilon_\lambda)B(0,1) + \lambda C.$$

We need that

$$(1-\lambda)x + (\lambda\epsilon + \epsilon_\lambda)B(0,1) \subset (1-\lambda)C \Leftrightarrow \lambda\epsilon + \epsilon_\lambda = (1-\lambda)\epsilon_x \Leftrightarrow \epsilon_\lambda = (1-\lambda)\epsilon_x - \lambda\epsilon$$

so that $x_\lambda + \epsilon_\lambda B(0,1) \subset C$. As ϵ_λ is positive, ϵ is chosen conveniently. Here, we use the sum operation of convex sets, which is well-defined if the coefficients are positive (if C_1 and C_2 are convex sets, $\mu_1 C_1 + \mu_2 C_2$ is a convex set for all $\mu_1, \mu_2 \geq 0$).

(b) As $\text{int } C \subset C$, $\overline{\text{int } C} \subset \overline{C}$.

On the other hand, each point in the closure of C can be approached along a line segment by points in the interior of C by (a), then $\overline{C} \subset \overline{\text{int } C}$.

(c) As $\text{int } C \subset \text{int } \overline{C}$, if $\text{int } \overline{C}$ is empty, then $\text{int } C$ is empty.

Conversely, $\text{int } C$ is empty, then by (b) $\overline{C}$ is empty. So C is empty and $\text{int } C$ is also empty.

(d) Trivially $\text{int } C \subset \text{int } \overline{C}$.

As the equality of the sets is true when they are empty because of (c), let us suppose that $\text{int } C$ is not empty. If $z \in \text{int } \overline{C}$ and take $x \in \text{int } C$, then there exists a small enough positive number δ such that $y = z + \delta(z - x) \in \text{int } \overline{C} \subset \overline{C}$.

Hence,

$$z = \frac{1}{1+\delta}y + \frac{\delta}{1+\delta}x.$$

Note that

$$0 < \frac{1}{1+\delta} < 1, \quad 0 < \frac{\delta}{1+\delta} < 1, \quad \frac{1}{1+\delta} + \frac{\delta}{1+\delta} = 1.$$

As $y \in \overline{C}$, $x \in \text{int } C$ and $1/(1+\delta)$ lies in $(0,1)$, by (a) $z \in \text{int } C$. □

Remark D.6. Consequently, if C is convex and dense, then C is the whole space.

The following paragraphs introduce elements playing an important role in the proof of Pontryagin's Maximum Principle.

Definition D.7. *Let C be a cone with vertex at $0 \in E$, a* ***supporting hyperplane to C at*** 0 *is a hyperplane such that C is contained in one of the half-spaces defined by the hyperplane.*

Remark D.8. In a geometric framework, we will define a hyperplane in E as the kernel of a 1-form, that is, if $\alpha \neq 0$, $\alpha \in E^*$, the dual space of E, then the hyperplane P_α associated to α is $\text{Ker}\,\alpha$. Hence the supporting hyperplane to C at 0 is a hyperplane P_α such that $\alpha(v) \leq 0$ for all $v \in C$. A supporting hyperplane to C at 0 is not necessarily unique.

From now on, we consider that all the cones have vertex at 0.

Definition D.9. *Let C be a cone, the* ***polar of*** *C is $C^* = \{\alpha \in E^* \mid \alpha(v) \leq 0 \ , \quad \forall\, v \in C\}$.*

Note that the polar of a cone is a closed and convex cone in E^*.

Definition D.10. *Let C be a cone, the set $C^{**} = \{w \in E \mid \alpha(w) \leq 0 \ , \quad \forall\, \alpha \in C^*\}$ is called the* ***polar of the polar of*** *C.*

Observe that $C \subset C^{**}$. The following lemma is used in the proof of the existence of a supporting hyperplane to a cone with vertex at 0.

Lemma D.11. *The cone C is closed and convex if and only if $C^{**} = C$.*

Proof. Observe that

$$C^{**} = \{w \in E \mid \alpha(w) \leq 0 \ , \quad \forall \alpha \in C^*\} = \cap_{\alpha \in C^*} \{w \in E \mid \alpha(w) \leq 0\}.$$

Then $C^{**} = \overline{\text{conv}(C)}$, because of Theorem 6.20 in Rockafellar [29]: the closure of the convex hull of a set is the intersection of all the closed half-spaces containing the set. Now, the result is immediate. □

The following proposition guarantees the existence of a supporting hyperplane to a cone with vertex at 0. This result is used throughout the proof of Pontryagin's Maximum Principle.

Proposition D.12. *If C is a convex and closed cone that is not the whole space, then there exists a supporting hyperplane to C at 0.*

Proof. If there is no supporting hyperplane containing the cone in one of the two half-spaces, then for all $\alpha \in E^*$ there exist $v_1, v_2 \in C$ with $\alpha(v_1) \leq 0$ and $\alpha(v_2) \geq 0$. Thus $C^* = \{0\}$ and $C^{**} = E$, then by Lemma D.11, $C = C^{**} = E$ in contradiction with the hypothesis on C. □

Corollary D.13. *If C is a convex cone that is not the whole space, then there exists a supporting hyperplane to C at 0.*

Proof. If $C \neq E$, then $\overline{C} \neq E$ by Proposition D.5 at 0. Hence, by Proposition D.12, there exists a supporting hyperplane to $\overline{C}$, which is also a supporting hyperplane to C. □

Definition D.14. *Let C_1 and C_2 be cones with common vertex 0. They are separated if there exists a hyperplane P such that each cone lies in a different closed half-space defined by P. P is called a* ***separating hyperplane of $\mathbf{C_1}$ and $\mathbf{C_2}$****, and* **$\mathbf{C_1}$ *and* $\mathbf{C_2}$ *are separated*** *if there exists a separating hyperplane.*

A useful characterization of separated cones is the following:

Proposition D.15. *The convex cones C_1 and C_2, with common vertex 0, are separated if and only if one of the two following conditions are satisfied:*

(1) there exists a hyperplane containing both C_1 and C_2,

(2) there is no point that is a relative interior point of both C_1 and C_2.

Proof. ⇒ If C_1 and C_2 are separated then there exists a separating hyperplane P_α such that

$$\alpha(v_1) \leq 0 \quad \forall\, v_1 \in C_1, \ \alpha(v_2) \geq 0 \quad \forall\, v_2 \in C_2.$$

If $\alpha(v_i) = 0$ for all $v_i \in C_i$ and $i = 1, 2$, then we are in the first case.

If some $v_i \in C_i$ satisfies the strict inequality, then both sets do not lie in the hyperplane. They lie in a different closed half-space. If the convex cones intersect, the intersection lies in the boundary of the cones and in the hyperplane. Hence, there is no point that is a relative interior point of both C_1 and C_2.

⇐ First, we are going to prove that if (1) is true, then C_1 and C_2 are separated. As there exists a hyperplane determined by α such that $\alpha(v_i) = 0$ for all $v_i \in C_i$, α determines a separating hyperplane of C_1 and C_2.

Now, we are going to prove that if (2) is true, then C_1 and C_2 are separated. As C_1 and C_2 are convex cones, $C_1 - C_2 = \{u \in E \mid u = v_1 - v_2,\ v_1 \in C_1,\ v_2 \in C_2\}$ is a convex cone. Since there is no relative interior point of both C_1 and C_2, 0 does not lie in $C_1 - C_2$. By Corollary D.13 there exists a supporting hyperplane P_α to $C_1 - C_2$ such that $\alpha(v_1 - v_2) \leq 0$, that is, $\alpha(v_1) \leq \alpha(v_2)$, for all $v_1 \in C_1$, $v_2 \in C_2$.

Observe that a supporting hyperplane to $C_1 - C_2$ is a supporting hyperplane to C_1, because taking $v_2 = 0$, $\alpha(v_1) \leq \alpha(v_2) = 0$ for all $v_1 \in C_1$.

As $\partial(C_1 - C_2) \cap C_1 \subset \partial C_1$, we consider a supporting hyperplane P_α to $C_1 - C_2$ such that $\alpha(v_1) = 0$ for some $v_1 \in \partial C_1$. Hence $\alpha(v_2) \geq \alpha(v_1) = 0$ for all $v_2 \in C_2$. As $\alpha(v_1) \leq 0$ for all $v_1 \in C_1$, α determines a separating hyperplane of C_1 and C_2. □

This proposition gives us necessary and sufficient conditions for the existence of a separating hyperplane of two convex cones with common vertex. Observe that a separating hyperplane of two cones with common vertex is also a supporting hyperplane to each cone at the vertex.

Corollary D.16. *If the convex cones C_1 and C_2 with common vertex 0 are not separated, then $E = C_1 - C_2$.*

Proof. If the cones are not separated, by Proposition D.15 there no exists any hyperplane containing both and the intersection of their relative interior is not empty.

Let us suppose that the convex cone $C_1 - C_2 \neq E$, then by Corollary D.13 there exists a supporting hyperplane determined by λ at the vertex such that $\lambda(v) \geq 0$ for every v in $C_1 - C_2$.

Because of the definition of cones, if $v_1 \in C_1$, then $v_1 \in C_1 - C_2$ and $\lambda(v_1) \geq 0$. Analogously, if $v_2 \in C_2$, then $-v_2 \in C_1 - C_2$ and $\lambda(-v_2) \geq 0$, that is, $\lambda(v_2) \leq 0$. □

E One corollary of Brouwer Fixed-Point Theorem

From the statement of Brouwer Fixed-point Theorem, it is possible to prove a useful corollary for this report [21].

Theorem E.1. ***(Brouwer Fixed-point Theorem)*** *Let B_1^n be the closed unit ball in $\mathbb{R}^n$. Any continuous function $G\colon B_1^n \to B_1^n$ has a fixed point.*

Corollary E.2. *Let $g\colon B_1^n \to \mathbb{R}^n$ be a continuous map. Let P be an interior point of B_1^n. If $\|g(x) - x\| < \|x - P\|$ for every x in the boundary ∂B_1^n, then the image $g(B_1^n)$ covers P.*

Proof. We assume that P is the origin of $\mathbb{R}^n$. Consider the mapping g as a continuous vector field on the unit ball B_1^n.

As $\|g(x) - x\| < \|x\|$, we are going to show that $g(x)$ makes an acute angle with the outward ray from the origin through x for every $x \in \partial B^n$. Let us consider the equality

$$\|y - z\|^2 + \|z - x\|^2 = \|y - x\|^2 + 2\langle y - z, x - z\rangle,$$

and take $y = g(x)$ and $z = 0$, then

$$2\langle g(x), x\rangle = \|g(x)\|^2 + \|x\|^2 - \|g(x) - x\|^2 > \|g(x)\|^2 + \|x\|^2 - \|x\|^2 = \|g(x)\|^2 \geq 0.$$

Thus $g(x)$ makes an acute angle with x. So $g(x)$ has an outward radial component at every point $x \in \partial B_1^n$. The vector $-g(x)$ has a negative radial component. For a sufficiently small positive number α the function $x \to x - \alpha\, g(x)$ goes from B_1^n to B_1^n. By Theorem E.1 there exists a fixed point x_0 such that $x_0 = x_0 - \alpha\, g(x_0)$, then $\alpha\, g(x_0) = 0$ and $g(x_0) = 0$ since $\alpha \in \mathbb{R}^+$. As g is continuous and $g(x_0) = 0$, the image of a neighbourhood of x_0 covers the origin. □

Acknowledgements

We acknowledge the financial support of *Ministerio de Educación y Ciencia*, Project MTM2005-04947 and the Network Project MTM2006-27467-E/. MBL also acknowledges the financial support of the FPU grant AP20040096.

We are grateful to Professor Andrew D. Lewis for encouraging us to work on a deep understanding of this subject. We thank Marina Delgado for collaborating with us on this report during her postdoctoral stay in our Department in 2005.

References

[1] A. Agrachev, Y. Sachkov, *Control Theory from the Geometric Viewpoint*, Springer-Verlag, Berlin-Heidelberg 2004.

[2] M. Athans, P. L. Falb, *Optimal Control: An Introduction to the Theory and Its Applications*, McGraw-Hill Book Company, New York 1966.

[3] M. Barbero-Liñán, M.C. Muñoz-Lecanda, Constraint Algorithm for Extremals in Optimal Control Problems, arXiv:0708.0102 math.OC, preprint (February 2008).

[4] M. Barbero-Liñán, M.C. Muñoz-Lecanda, Optimal control problems for affine connection control systems: characterization of extremals, Proceedings of the XVI International Workshop on Geometry and Physics, Lisbon, Portugal, September 5-8, 2007 (to be published).

[5] D. P. Bertsekas, A. Geary-Nedic, A. Koksal, *Lectures Notes on Convexity, Duality and Lagrange Multipliers*, Chapter 1, Massachussets Institute of Technology 2001. http://citeseer.ist.psu.edu/499868.html .

[6] A. M. Bloch et al, *Nonholonomic Mechanics and Control*, volume 24 of Interdisciplinary Applied Mathematics, Springer-Verlag, New York-Heidelberg-Berlin, 2004.

[7] V. Boltyanski, H. Martini, V. Soltan, *Geometric Methods and Optimization Problems*, Kluwer Academic Publishers, Dordrecht 1999.

[8] B. Bonnard, J. B. Caillau, *Introduction to Nonlinear Optimal Control*, Lecture notes in FAP (Formation en Automatique Paris) course 2005. To appear at Springer-Verlag.

[9] B. Bonnard, M. Chyba, *Singular Trajectories and their Role in Control Theory*, Mathematics & Applications 40, Springer-Verlag, New York-Heidelberg-Berlin, 2003.

[10] F. Bullo, A. D. Lewis, *On the homogeneity of the affine connection model for mechanical control systems*, Proceedings of the 39th IEEE Conference on Decision and Control, 2000, 1260-1265, vol.2.

[11] F. Bullo, A. D. Lewis, *Geometric Control of Mechanical Systems. Modeling, analysis and design for simple mechanical control*, Springer-Verlag, New York-Heidelberg-Berlin 2004.

[12] F. Bullo, A. D. Lewis, Supplementary Chapters of *Geometric Control of Mechanical Systems. Modeling, analysis and design for simple mechanical control*, Texts in Applied Mathematics 49, Springer-Verlag, New York-Heidelberg-Berlin 2004.

[13] J. A. Cañizo Rincón, *Ecuaciones diferenciales ordinarias en el sentido de Carathéodory*, July 2004, http://www.ceremade.dauphine.fr/∼canizo/tex/ordinarias-debil.pdf

[14] E. A. Coddington, N. Levison, *Theory of Ordinary Differential Equations*, McGraw-Hill Book Company, New York 1955.

[15] A. F. Filippov, On certain questions in the theory of optimal control, *J. SIAM Control Ser. A*, **1** (1962), 76-84.

[16] E. Hairer, *Analyse II, Partie B*, Université de Genève 1998/99. http://www.unige.ch/∼hairer/polycop.html .

[17] V. Jurdjevic, *Geometric Control Theory*, Cambridge Studies in Advanced Mathematics 51, Cambridge University Press, New York 1997.

[18] I. Kolář, P. W. Michor, J. Slovák, *Natural Operations in Differential Geometry*, Springer, Berlin 1993.

[19] B. Langerock, Generalised Connections and Applications to Control Theory, PhD dissertation, 2003, https://archive.ugent.be/retrieve/957/801001362505.pdf

[20] B. Langerock, Geometric Aspects of the Maximum Principle and Lifts over a Bundle Map, *Acta Appl. Math.* **77**, 71-104, 2003.

[21] E. B. Lee, L. Markus, *Foundations of Optimal Control Theory*, John Wiley and Sons, New York 1967.

[22] J. M. Lee, *Introduction to Smooth Manifolds*, Springer-Verlag, New York-Heidelberg-Berlin 2003.

[23] W. Liu, H. J. Sussmann, Shortest paths for sub-Riemannian metrics on rank-two distribution, *Mem. Amer. Math. Soc.* 564, Jan. 1996.

[24] E. J. McShane, On Multipliers for Lagrange Problems, *Am. J. Math.* **61** (1939), 809-19.

[25] R. Montgomery, Abnormal Minimizers, *SIAM J. Control Optim.*, **32**(6), 1605-1620, November 1994.

[26] R. Montgomery, *A Tour of Subriemannian Geometries, Their Geodesics and Applications*, number 91 in Mathematical Surveys and Monographs, American Mathematical Society, Providence, RI, 2002.

[27] H. Nijmeijer, A. J. Van der Schaft, *Nonlinear Dynamical Control Systems*, Springer-Verlag, New York 1990.

[28] L. S. Pontryagin, V. G. Boltyanski, R. V. Gamkrelidze, E. F. Mischenko, *The Mathematical Theory of Optimal Processes*, Interscience Publishers, Inc., New York 1962.

[29] R. T. Rockafellar, R. J-B. Wets, *Variational Analysis*, Comprehensive Studies in Mathematics 317, Springer-Verlag, Berlin-Heidelberg-New York 1998.

[30] H. L. Royden, *Real Anlaysis*, Macmillan, New York 1968.

[31] H. J. Sussmann, An introduction to the coordinate-free maximum principle, *Geometry of feedback and optimal control*, 463–557, Monogr. Textbooks Pure Appl. Math. 207, Marcel Dekker, New York, 1998.

[32] H. J. Sussmann, *Course on Optimal Control of Finite-Dimensional Deterministic Systems* at the Weizmann Institute, 2000.
http://www.math.rutgers.edu/∼sussmann/weizmann-course-2000.html

[33] H. J. Sussmann, A very non-smooth maximum principle with state constraints, Proceedings of the 44th IEEE CDC, Sevilla, Spain, December 12-15, 2005.

[34] J. L. Troutman, *Variational Calculus and Optimal Control: Optimization with Elementary Convexity*, Undergraduate Texts in Mathematics, Springer-Verlag, Berlin-Heidelberg-New York 1996.

[35] Dale E. Varberg, On absolutely continuous functions, *Amer. Math. Monthly*, **72** (1965), 831-841.

[36] A. C. Zaanen, *Continuity, Integration and Fourier Theory*, Springer-Verlag, Berlin 1989.

[37] E. Zeidler, *Nonlinear Functional Analysis and its Applications III, Variational Methods and Optimization*, Springer-Verlag, New York 1985.