

# ON THE SET OF POINTS OF SMOOTHNESS FOR THE VALUE FUNCTION OF AFFINE OPTIMAL CONTROL PROBLEMS

DAVIDE BARILARI AND FRANCESCO BOAROTTO

ABSTRACT. We study the regularity properties of the value function associated with an affine optimal control problem with quadratic cost plus a potential, for a fixed final time and initial point. Without assuming any condition on singular minimizers, we prove that the value function is continuous on an open and dense subset of the interior of the attainable set. As a byproduct we obtain that it is actually smooth on a possibly smaller set, still open and dense.

## CONTENTS

1. Introduction	1
2. Preliminaries	4
3. On the continuity	8
4. Density of tame points	10
5. On the smoothness	16
Appendix A. A few technical results	18
Acknowledgments	19
References	19

## 1. INTRODUCTION

The regularity of the value function associated with an optimal control problem is a classical topic of investigation in control theory and has been deeply studied in the last decades, extensively using tools from geometric control theory and non-smooth analysis. It is well-known that the value function associated with an optimal control problem fails to be everywhere differentiable and this is typically the case at those points where the uniqueness of minimizers is not guaranteed. Actually, it is not even continuous, in general, as soon as singular minimizers are allowed (see for instance [4, 19]).

In this paper we investigate the regularity of the value function associated with affine optimal control problems, whose cost is written as a quadratic term plus a potential

The key starting point of our work is the characterization of points where the value function is continuous. As we said, in presence of singular minimizers for the control problem one could not expect the value function to be continuous. Indeed, for a fixed final time  $T > 0$  and initial point  $x_0$ , the continuity of the value function  $S_{x_0}^T$  at a point  $x$  is strictly related with the openness of the end-point map on the optimal controls steering the initial fixed point  $x_0$  to  $x$  in time  $T > 0$ . Here by end-point map, we mean the map that to every control  $u$  associates the final point of the corresponding trajectory (cf. Section 2 for precise definitions).

Without assuming any condition on singular minimizers, we focus on the set of points, that we call *tame points*, in the interior of the attainable set such that the end-point map is open and a submersion at every optimal control. The main result of this paper is that we can find a large set of tame points. Since tame points are points of continuity for the value function, we deduce that  $S_{x_0}^T$  is continuous on an open and dense set of the interior of the attainable set.

Adapting then the arguments of [1, 15], we prove that the value function is actually smooth on a (possibly smaller) open dense subset of the interior of the attainable set.

The main novelty with respect to the known results, valid in the drift-less case and with zero potential, is that in the latter case the value function is everywhere continuous as a consequence of the openness of the end-point map, even in presence of deep singular minimizers. The absence of such a property for affine control system makes the study of the continuity of the value function more delicate in our context.

Let us introduce briefly the notation and present the main results more in details.

**1.1. Setting and main results.** Let  $M$  be a smooth, connected  $m$ -dimensional manifold and let  $T > 0$  be a given *fixed* final time. A smooth *affine control system* is a dynamical system which can be written in the form:

$$(1.1) \quad \dot{x}(t) = X_0(x(t)) + \sum_{i=1}^d u_i(t) X_i(x(t)),$$

where  $X_0, X_1, \dots, X_d$  are smooth vector fields on  $M$ , and the map  $t \mapsto u(t) = (u_1(t), \dots, u_d(t))$  belongs to the Hilbert space  $L^2([0, T], \mathbb{R}^d)$ .

Given  $x_0 \in M$  we define:

- (i) the set of *admissible controls*  $\Omega_{x_0}^T$  as the subset of  $u \in L^2([0, T], \mathbb{R}^d)$ , such that the solution  $x_u(\cdot)$  to (1.1) satisfying  $x_u(0) = x_0$  is defined on the interval  $[0, T]$ . If  $u \in \Omega_{x_0}^T$  we say that  $x_u(\cdot)$  is an *admissible trajectory*. By classical results of ODE theory, the set  $\Omega_{x_0}^T$  is open.
- (ii) the *attainable set*  $A_{x_0}^T$  (from the point  $x_0$ , in time  $T > 0$ ), as the set of points of  $M$  that can be reached from  $x_0$  by admissible trajectories in time  $T$ , i.e.,

$$A_{x_0}^T = \{x_u(T) \mid u \in \Omega_{x_0}^T\}.$$

For a given smooth function  $Q : M \rightarrow \mathbb{R}$ , we are interested in those trajectories minimizing the *cost* given by:

$$(1.2) \quad C_T : \Omega_{x_0}^T \rightarrow \mathbb{R}, \quad C_T(u) = \frac{1}{2} \int_0^T \left( \sum_{i=1}^d u_i(t)^2 - Q(x_u(t)) \right) dt.$$

More precisely, given  $x_0 \in M$  and  $T > 0$ , we are interested in the regularity properties of the *value function*  $S_{x_0}^T : M \rightarrow \mathbb{R}$  defined as follows:

$$(1.3) \quad S_{x_0}^T(x) = \inf \{C_T(u) \mid u \in \Omega_{x_0}^T, x_u(T) = x\};$$

with the understanding that  $S_{x_0}^T(x) = +\infty$  if  $x$  cannot be attained by admissible curves in time  $T$ . We call *optimal control* any control  $u$  which solves the optimal control problem (1.3).

**Main assumptions.** For the rest of the paper we make the following assumptions:

- (H1) The *weak Hörmander condition* holds on  $M$ . Namely, we require for every point  $x \in M$  the equality

$$(1.4) \quad \text{Lie}_x \left\{ (\text{ad } X_0)^j X_i \mid j \geq 0, i = 1, \dots, d \right\} = T_x M.$$

where  $(\text{ad } X)Y = [X, Y]$ , and  $\text{Lie}_x \mathcal{F} \subset T_x M$  denotes the evaluation at the point  $x$  of the Lie algebra generated by a family  $\mathcal{F}$  of vector fields.

- (H2) For every bounded family  $\mathcal{U}$  of admissible controls, there exists a compact subset  $K_T \subset M$  such that  $x_u(t) \in K_T$ , for every  $u \in \mathcal{U}$  and  $t \in [0, T]$ .
- (H3) The potential  $Q$  is a smooth function bounded from above.

The assumption (H1) is needed to guarantee that the attainable set has at least non-empty interior, i.e.,  $\text{int}(A_{x_0}^T) \neq \emptyset$  (cf. [12, Ch. 3, Thm. 3]). The second assumption (H2) is a completeness/compactness assumption on the dynamical system that, together with (H3), is needed to guarantee the existence of optimal controls. We stress that (H2) and (H3) are automatically satisfied when  $M$  is compact. When  $M$  is not compact, (H2) holds true under a sublinear growth condition on the vector fields  $X_0, \dots, X_d$ . We refer to Section 2 for more details on the role of these assumptions.

As already anticipated, the key starting point of our work is the characterization of points where the value function is continuous through the study of the set of *tame points*. This is the set  $\Sigma_t \subset \text{int}(A_{x_0}^T)$  of all points  $x$  such that the end-point map is open *and* a submersion at every optimal control steering  $x_0$  to  $x$ . The main result of this paper, whose proof comprises its technical core, is that we can find a large set of tame points.

**Theorem 1.** *Fix  $x_0 \in M$  and let  $S_{x_0}^T$  be the value function associated with an optimal control problem of the form (1.1)-(1.2) satisfying assumptions (H1)-(H3). Then the set  $\Sigma_t$  of tame points is open and dense in  $\text{int}(A_{x_0}^T)$  and  $S_{x_0}^T$  is continuous on  $\Sigma_t$ .*

In the drift-less case (more precisely, when  $X_0 = 0$  and  $Q = 0$ ) the end-point map is open at every point, even if it is not a submersion in the presence of singular minimizers. This, however, suffices for the sub-Riemannian distance to be continuous everywhere. Moreover this remains true for any  $L^p$ -topology on the space of controls, for  $p < +\infty$ , see [7]. This is no more true if we introduce a drift field, and the characterization of the set of points where the end-point is open and the choice of the topology in the space of controls is more delicate.

The proof of Theorem 1 is inspired by the arguments, dealing with the sub-Riemannian case, presented among others by the first author in [3, Chapter 11], and starts by characterizing the set of points reached by a unique minimizer trajectory that is not strictly singular (called *fair points*). The classical argument proves that this set is dense in the attainable set but, while in the drift-less case each of these points is also a continuity point for the value function, in this setting in principle it could likely be that the set of fair points and the set of continuity points, both dense, may have empty intersection. Completing this gap requires ad hoc new arguments developed in Section 4.

Once Theorem 1 is proved, an adaptation of arguments from [1, 15] let us derive the following result.

**Theorem 2.** *Under the assumptions of Theorem 1,  $S_{x_0}^T$  is smooth on a non-empty open and dense subset of  $\text{int}(A_{x_0}^T)$ .*

In [1], the author proves the analogue of Theorem 2 for the value function associated with sub-Riemannian optimal control problems, i.e., drift-less systems with zero potential. Notice that in this case (H1) reduces to the classical Hörmander condition, and the value function (at time  $T$ ) coincides with one half of the square of the sub-Riemannian distance (divided by  $T$ ) associated with the family of vector fields  $X_1, \dots, X_d$ .

Let us further mention that, even in the sub-Riemannian situation, it still remains an open question to establish whether the set of smoothness points of the value function has full measure in  $\text{int}(A_{x_0}^T)$  or not.

**1.2. Further comments.** Regularity of the value function for these kinds of control system with techniques of geometric control has been also studied in [17, 19], where the authors assume that there are no abnormal optimal controls, a condition which yields the openness of the end-point already at the first order, while in [4] the authors obtain the openness of the end-point map on optimal controls with second-order techniques, assuming no optimal Goh abnormal controls exist.

For more details on Goh abnormality we refer the reader to [5, Chapter 20] (see also [3, 16]). Let us mention that in [9] the authors prove that the system (1.1) admits no Goh optimal trajectories for the generic choice of the  $(d+1)$ -tuple  $X_0, \dots, X_d$  (in the Whitney topology). Finally in [14] the author proves the Hölder continuity of the value function under a strong bracket generating assumption, when one considers the  $L^1$  cost.

For different approaches investigating the regularity of the value function through techniques of non-smooth analysis, one can see for instance the monographs [6, 10, 8, 11].

**1.3. Structure of the paper.** In Section 2 we recall some properties of the end-point map, the existence of minimizers in our setting and recall their characterization in terms of the Hamiltonian equation. Section 3 introduces different sets of points that are relevant in our analysis. Section 4 is devoted to the study of tame points and the proof of Theorem 1. In the last Section 5 we complete the proof of Theorem 2. Finally, in Appendix A we present for readers' convenience the proof of a few technical facts, adapted with minor modifications to our setting.

## 2. PRELIMINARIES

For a fixed admissible control  $u \in \Omega_{x_0}^T$ , it is well-defined the family of diffeomorphisms

$$P_{0,t}^u : U_{x_0} \subset M \rightarrow M, \quad t \in [0, T],$$

defined on some neighborhood  $U_{x_0}$  of  $x_0$  by  $P_{0,t}^u(y) = x_{u,y}(t)$ , where  $x_{u,y}(t)$  is the solution of the equation (1.1) with initial condition  $x_{u,y}(0) = y$ . It is a classical fact that this family is absolutely continuous with respect to  $t$ . Similarly, given  $u \in \Omega_{x_0}^T$  it is possible to define the family of flow diffeomorphisms  $P_{s,t}^u : U_{x_0} \rightarrow M$  by solving (1.1) with initial condition  $x_{u,y}(s) = y$ ; notice then that  $P_{t,t}^u = \text{Id}$ , and that the following composition formulas hold true (at those points where all terms are defined):

$$P_{s,t}^u \circ P_{r,s}^u = P_{r,t}^u \quad \text{and} \quad (P_{s,t}^u)^{-1} = P_{t,s}^u.$$

Finally, we employ the notation  $(P_{s,T}^u)_*$  to refer to the push-forward map defined from  $T_{x_u(s)}M$  to  $T_{x_u(t)}M$ : in particular if  $X$  is any vector field on  $M$ , then the push-forward  $(P_{s,t}^u)_* X$  is defined by:

$$(P_{s,t}^u)_*(X(y)) = \left( (P_{s,t}^u)_* X \right) (P_{s,t}^u(y)).$$

**2.1. The end-point map.** In what follows we fix a point  $x_0 \in M$  and a time  $T > 0$ .

**Definition 3** (end-point map). The *end-point map* at time  $T$  is the map

$$E_{x_0}^T : \Omega_{x_0}^T \rightarrow M, \quad E_{x_0}^T(u) = x_u(T),$$

where  $x_u(\cdot)$  is the admissible trajectory driven by the control  $u$ .

The end-point map is smooth on  $\Omega_{x_0}^T \subset L^2([0, T], \mathbb{R}^d)$ . The computation of its Fréchet differential is classical and can be found for example in [3, 16, 19]:

**Proposition 4.** *The differential  $d_u E_{x_0}^T : L^2([0, T], \mathbb{R}^d) \rightarrow T_{x_u(T)}M$  of the end-point map at  $u \in \Omega_{x_0}^T$  is given by the formula:*

$$(2.1) \quad d_u E_{x_0}^T(v) = \int_0^T \sum_{i=1}^d v_i(s) (P_{s,T}^u)_* X_i(x_u(s)) ds.$$

Let us consider a sequence of admissible controls  $\{u_n\}_{n \in \mathbb{N}}$ , which weakly converges to some element  $u \in L^2([0, T], \mathbb{R}^d)$ . Then the sequence  $\{u_n\}_{n \in \mathbb{N}}$  is bounded in  $L^2$  and, thanks to our assumption (H2), there exists a compact set  $K_T$  such that  $x_{u_n}(t) \in K_T$  for all  $n \in \mathbb{N}$  and  $t \in [0, T]$ .

This yields that the family of trajectories  $\{x_{u_n}(\cdot)\}_{n \in \mathbb{N}}$  is uniformly bounded, and from here it is a classical fact to deduce that the weak limit  $u$  is an admissible control, and that  $x_u(\cdot) = \lim_{n \rightarrow \infty} x_{u_n}(\cdot)$  (in the uniform topology) is its associated trajectory (see for example [20]).

This proves that the end-point map  $E_{x_0}^T$  is *weakly continuous*. Indeed, one can prove that the same holds true for its differential  $d_u E_{x_0}^T$ . More precisely: if  $\{u_n\}_{n \in \mathbb{N}}$  is a sequence of admissible controls which weakly converges in  $L^2([0, T], \mathbb{R}^d)$  to  $u$  (which is admissible by the previous discussion), we have both that

$$\lim_{n \rightarrow \infty} E_{x_0}^T(u_n) = E_{x_0}^T(u) \quad \text{and} \quad \lim_{n \rightarrow \infty} d_{u_n} E_{x_0}^T = d_u E_{x_0}^T,$$

and the last convergence is in the (strong) operator norm (see [19]).

*Remark.* There are other possible assumptions to ensure that the weak limit of a sequence of admissible controls is again an admissible control; for example, as suggested in [17], one could ask a sublinear growth condition on the vector fields  $X_0, \dots, X_d$ . In this case the uniform bound on the trajectories (equivalent to (H2)) follows as a consequence of the Gronwall inequality, and the observation that a weakly convergent sequence in  $L^2$  is necessarily bounded.

**Definition 5** (Attainable set). For a fixed final time  $T > 0$ , we denote by  $A_{x_0}^T$  the image of the end-point map at time  $T$ , and we call it the *attainable set* (from the point  $x_0$ ).

In general the inclusion  $A_{x_0}^T \subset M$  can be proper, that is the end-point map  $E_{x_0}^T$  may not be surjective on  $M$ ; nevertheless, the weak Hörmander condition (1.4) implies that for every initial point  $x_0$  one has  $\text{int}(A_{x_0}^T) \neq \emptyset$  [12, Ch. 3, Thm. 3].

**2.2. Value function and optimal trajectories.** Let  $Q : M \rightarrow \mathbb{R}$  be a smooth function, which plays in what follows the role of a potential; if we introduce the Tonelli Lagrangian

$$L : M \times \mathbb{R}^d \rightarrow \mathbb{R}, \quad L(x, u) = \frac{1}{2} \left( \sum_{i=1}^d u_i^2 - Q(x) \right),$$

then the *cost*  $C_T : \Omega_{x_0}^T \rightarrow \mathbb{R}$  is written as:

$$C_T(u) = \int_0^T L(x_u(t), u(t)) dt = \frac{1}{2} \int_0^T \left( \sum_{i=1}^d u_i(t)^2 - Q(x_u(t)) \right) dt.$$

The differential  $d_u C_T$  of the cost can be recovered similarly as for the differential of the end-point map, and is given, for every  $v \in L^2([0, T], \mathbb{R}^d)$ , by the formula

$$d_u C_T(v) = \int_0^T \langle u(t), v(t) \rangle dt - \frac{1}{2} \int_0^T Q'(x_u(t)) \left( \int_0^t \sum_{i=1}^d v_i(s) (P_{s,t}^u)_* X_i(x_u(s)) ds \right) dt,$$

that is obtained by writing  $x_u(t) = E_{x_0}^t(u)$  and applying (2.1).

Fix two points  $x_0$  and  $x$  in  $M$ . The problem of describing optimal trajectories steering  $x_0$  to  $x$  in time  $T$  can be naturally reformulated in the following way: introducing the value function  $S_{x_0}^T : M \rightarrow \mathbb{R}$  via the position

$$(2.2) \quad S_{x_0}^T(x) := \inf \left\{ C_T(u) \mid u \in \Omega_{x_0}^T \cap (E_{x_0}^T)^{-1}(x) \right\},$$

with the agreement that  $S_{x_0}^T(x) = +\infty$  if the preimage  $(E_{x_0}^T)^{-1}(x)$  is empty, then, for any fixed  $x \in M$ , the *optimal control problem* consists into looking for elements  $u \in L^2([0, T], \mathbb{R}^d)$  realizing the infimum in (2.2). Accordingly, from now on we will call *optimal control* any admissible control  $u$  which solves the optimal control problem.

In this paper we will always concentrate on the case that the final point  $x$  of an admissible trajectory belongs to the interior of the attainable set  $A_{x_0}^T$ . Indeed, it is a general fact that  $\text{int}(A_{x_0}^T)$  is densely contained in  $A_{x_0}^T$  [5, 12], and the weak Hörmander condition ensures that  $\text{int}(A_{x_0}^T)$  is non-empty; moreover, for every point  $x \in \text{int}(A_{x_0}^T)$ , we trivially have that  $S_{x_0}^T(x) < +\infty$ , since by definition there exists at least one admissible control  $v$  steering  $x_0$  to  $x$ .

Existence of minimizers under our main assumptions (H1)-(H3) follows from classical arguments.

**Proposition 6** (Existence of minimizers). *Let  $x \in A_{x_0}^T$ . Then there exists an optimal control  $u \in \Omega_{x_0}^T$  satisfying:*

$$E_{x_0}^T(u) = x, \quad \text{and} \quad C_T(u) = S_{x_0}^T(x).$$

*Remark.* The assumptions (H2)-(H3) play a crucial role for the existence of optimal control. An equivalent approach could be to work directly inside a given compact set (see [2]) or with  $M$  itself a compact manifold. For some specific cases, as in the classical case of the harmonic oscillator, one is able to integrate directly Hamilton's equations (cf. Section 2.4), and the existence of optimal trajectories could be proved with ad hoc arguments.

As already pointed out in the introduction, one could not expect global continuity for the value function. Nevertheless, it is well-known that under our assumptions, we have the following.

**Proposition 7.** *The map  $S_{x_0}^T : A_{x_0}^T \rightarrow \mathbb{R}$  is lower semicontinuous.*

Proofs of Propositions 6 and 7 are classical and follows from standard arguments in the literature, hence their proof is omitted and left to the reader.

**2.3. Lagrange multipliers rule.** In this section we briefly recall the classical necessary condition satisfied by optimal controls  $u$  realizing the infimum in (2.2). It is indeed a restatement of the classical Lagrange multipliers' rule (see [5, 3, 13]).

**Proposition 8.** *Let  $u \in L^2([0, T], \mathbb{R}^d)$  be an optimal control with  $x = E_{x_0}^T(u)$ . Then at least one of the following statements is true:*

- a)  $\exists \lambda_T \in T_x^*M$  such that  $\lambda_T d_u E_{x_0}^T = d_u C_T$ ,
- b)  $\exists \lambda_T \in T_x^*M$ , with  $\lambda_T \neq 0$ , such that  $\lambda_T d_u E_{x_0}^T = 0$ .

Here  $\lambda_T d_u E_{x_0}^T : L^2([0, T]) \rightarrow \mathbb{R}$  denotes the composition of the linear maps  $d_u E_{x_0}^T : L^2([0, T]) \rightarrow T_x M$  and  $\lambda_T : T_x M \rightarrow \mathbb{R}$ .

A control  $u$ , satisfying the necessary conditions for optimality stated in Proposition 8, is said *normal* in case (a) and *abnormal* in case (b); moreover, directly from the definition we see that  $d_u E_{x_0}^T$  is not surjective in the abnormal case. We stress again that the two possibilities are not mutually exclusive, and we define accordingly a control  $u$  to be *strictly normal* (resp. *strictly*

*abnormal*) if it is normal but not abnormal (resp. abnormal but not normal). Slightly abusing of the notation, we extend this language even to the associated optimal trajectories  $t \mapsto x_u(t)$ .

**2.4. Normal extremals and exponential map.** Let us denote by  $\pi : T^*M \rightarrow M$  the canonical projection of the cotangent bundle, and by  $\langle \lambda, v \rangle$  the duality pairing between a covector  $\lambda \in T_x^*M$  and a vector  $v \in T_xM$ . In canonical coordinates  $(p, x)$  on the cotangent space, we can express the Liouville form as  $s = \sum_{i=1}^m p_i dx_i$  and the standard symplectic form becomes  $\sigma = ds = \sum_{i=1}^m dp_i \wedge dx_i$ . We denote by  $\vec{h}$  the Hamiltonian vector field associated with a smooth function  $h : T^*M \rightarrow \mathbb{R}$ , defined by the identity:

$$\vec{h} = \sum_{i=1}^m \frac{\partial h}{\partial p_i} \frac{\partial}{\partial x_i} - \frac{\partial h}{\partial x_i} \frac{\partial}{\partial p_i}.$$

The Pontryagin Maximum Principle [13, 5] tells us that candidate optimal trajectories are projections of extremals, which are integral curves of the constrained Hamiltonian system:

$$\dot{x}(t) = \frac{\partial \mathcal{H}}{\partial p}(u(t), \nu, p(t), x(t)), \quad \dot{p}(t) = -\frac{\partial \mathcal{H}}{\partial x}(u(t), \nu, p(t), x(t)), \quad 0 = \frac{\partial \mathcal{H}}{\partial u}(u(t), \nu, p(t), x(t)),$$

where the (control-dependent) Hamiltonian  $\mathcal{H} : \mathbb{R}^d \times (-\infty, 0] \times T^*M \rightarrow \mathbb{R}$ , associated with the system (1.1), is defined by:

$$\mathcal{H}^\nu(u, \nu, p, x) = \langle p, X_0(x) \rangle + \sum_{i=1}^d u_i \langle p, X_i(x) \rangle + \frac{\nu}{2} \sum_{i=1}^d u_i^2 - \frac{\nu}{2} Q(x).$$

In particular, the non-positive real constant  $\nu$  remains constant along extremals; recalling the result of Proposition 8, there holds either the identity  $(p(T), \nu) = (\lambda_T, 0)$  in the case of abnormal extremals, or  $(p(T), \nu) = (\lambda_T, -1)$  for the normal ones. Moreover, we see that under the previous normalizations, the optimal control  $u(t)$  along normal extremals can be recovered using the equality:

$$(2.3) \quad u_i(t) = \langle p(t), X_i(x(t)) \rangle, \quad \text{for } i = 1, \dots, d.$$

Normal extremals are therefore solutions to the differential system:

$$\dot{x}(t) = \frac{\partial H}{\partial p}(p(t), x(t)), \quad \dot{p}(t) = -\frac{\partial H}{\partial x}(p(t), x(t)),$$

where the Hamiltonian  $H$  has the expression:

$$H(p, x) = \langle p, X_0(x) \rangle + \frac{1}{2} \sum_{i=1}^d \langle p, X_i(x) \rangle^2 + \frac{1}{2} Q(x).$$

In particular, being the solution to a smooth autonomous system of differential equations, the pair  $(x(t), p(t))$  is smooth as well, which eventually implies that the control  $u_i(t) = \langle p(t), X_i(x(t)) \rangle$  associated to normal trajectories is itself smooth by (2.3). It is well known that, under our assumptions, small pieces of normal trajectories are optimal among all the admissible curves that connect their end-points (see for instance [5]), that is, if  $x_1 = x_u(t_1)$  and  $x_2 = x_u(t_2)$  are sufficiently close points on the normal trajectory  $x_u(\cdot)$ , then the cost-minimizing admissible trajectory between  $x_1$  and  $x_2$  that solves (2.2) is precisely  $x_u(\cdot)$ .

**Definition 9** (Exponential map). The exponential map  $\mathcal{E}$  with base point  $x_0$  is defined as

$$\mathcal{E}_{x_0}(\cdot, \cdot) : [0, T] \times T_{x_0}^*M \rightarrow M, \quad \mathcal{E}_{x_0}(s, \lambda) = \pi(e^{s\vec{H}}(\lambda)).$$

When the first argument is fixed, we employ the notation  $\mathcal{E}_{x_0}^s : T_{x_0}^* M \rightarrow M$  to denote the exponential map with base point  $x_0$  at time  $s$ ; that is to say, we set  $\mathcal{E}_{x_0}^s(\lambda) := \mathcal{E}_{x_0}(s, \lambda)$ .

Then we see that the exponential map parametrizes normal extremals; moreover, mimicking the classical notion in the Riemannian setting, it permits to define *conjugate points* along these trajectories.

**Definition 10.** We say that a point  $x = \mathcal{E}_{x_0}(s, \lambda)$  is conjugate to  $x_0$  along the normal extremal  $t \mapsto \mathcal{E}_{x_0}(t, \lambda)$  if  $(s, \lambda)$  is a critical point of  $\mathcal{E}_{x_0}$ , i.e. if the differential  $d_{(s, \lambda)} \mathcal{E}_{x_0}$  is not surjective.

### 3. ON THE CONTINUITY

In this section we study fine properties of the value function on different subsets of  $\text{int}(A_{x_0}^T)$ .

**3.1. Fair points.** We start by introducing the set of fair points.

**Definition 11.** A point  $x \in \text{int}(A_{x_0}^T)$  is said to be a *fair point* if there exists a unique optimal trajectory steering  $x_0$  to  $x$ , which admits a normal lift. We call  $\Sigma_f$  the set of all fair points contained in the attainable set.

We stress that only the uniqueness of the optimal trajectory matters in the definition of a fair point; abnormal lifts are as well admitted for the moment.

The lower semicontinuity of  $S_{x_0}^T$  permits to find a great abundance of fair points; their existence is related to the notion of proximal subdifferential (see for instance [10, 15] for more details).

**Definition 12.** Let  $F : \text{int}(A_{x_0}^T) \rightarrow \mathbb{R}$  be a lower semicontinuous function. For every  $x \in \text{int}(A_{x_0}^T)$  we call the *proximal subdifferential* at  $x$  the subset of  $T_x^* M$  defined by:

$$\partial_P F(x) = \{ \lambda = d_x \phi \in T_x^* M \mid \phi \in C^\infty(\text{int}(A_{x_0}^T)) \text{ and } F - \phi \text{ attains a local minimum at } x \}.$$

The proximal subdifferential is a convex subset of  $T_x^* M$  which is often non-empty in the case of a lower semicontinuous function [10, Theorem 3.1].

**Proposition 13.** *Let  $F : \text{int}(A_{x_0}^T) \rightarrow \mathbb{R}$  be a lower semicontinuous function. Then the proximal subdifferential  $\partial_P F(x)$  is not empty for a dense set of points  $x \in \text{int}(A_{x_0}^T)$ .*

We showed in Proposition 7 that the value function  $S_{x_0}^T : \text{int}(A_{x_0}^T) \rightarrow \mathbb{R}$  is lower semicontinuous. By classical arguments, the proximal subdifferential machinery yields the following result (cf. also [15, 1]).

**Proposition 14.** *Let  $x \in \text{int}(A_{x_0}^T)$  be such that  $\partial_P S_{x_0}^T(x) \neq \emptyset$ . Then there exists a unique optimal trajectory  $x_u(\cdot) : [0, T] \rightarrow M$  steering  $x_0$  to  $x$ , which admits a normal lift. In particular  $x$  is a fair point.*

*Proof.* Fix any  $\lambda \in \partial_P S_{x_0}^T(x)$ . Let us prove that every optimal trajectory steering  $x_0$  to  $x$  admits a normal lift having  $\lambda$  as final covector.

Indeed, if  $\phi$  is a smooth function such that  $\lambda = d_x \phi \in \partial_P S_{x_0}^T(x)$ , by definition the map

$$\psi : \text{int}(A_{x_0}^T) \rightarrow \mathbb{R}, \quad \psi(y) = S_{x_0}^T(y) - \phi(y)$$

has a local minimum at  $x$ , i.e. there exists an open neighborhood  $O \subset \text{int}(A_{x_0}^T)$  of  $x$  such that  $\psi(y) \geq \psi(x)$  for every  $y \in O$ . Then, let  $t \mapsto x_u(t)$ ,  $t \in [0, T]$  be an optimal trajectory from  $x_0$  to  $x$ , let  $u$  be the associated optimal control, and define the smooth map:

$$\Phi : \Omega_{x_0}^T \rightarrow \mathbb{R}, \quad \Phi(v) = C_T(v) - \phi(E_{x_0}^T(v)).$$

There exists a neighborhood  $\mathcal{V} \subset \Omega_{x_0}^T$  of  $u$  such that  $E_{x_0}^T(\mathcal{V}) \subset O$ , and since  $C_T(v) \geq S_{x_0}^T(E_{x_0}^T(v))$  we have the following chain of inequalities:

$$\begin{aligned} \Phi(v) &= C_T(v) - \phi(E_{x_0}^T(v)) \geq S_{x_0}^T(E_{x_0}^T(v)) - \phi(E_{x_0}^T(v)) \\ &\geq S_{x_0}^T(E_{x_0}^T(u)) - \phi(E_{x_0}^T(u)) = C_T(u) - \phi(E_{x_0}^T(u)) = \Phi(u), \quad \forall v \in \mathcal{V}. \end{aligned}$$

Then:

$$0 = d_u \Phi = d_u C_T - (d_x \phi) d_u E_{x_0}^T,$$

and therefore we see that the curve  $\lambda(t) = e^{(t-T)\bar{H}}(\lambda)$  is the desired normal lift of the trajectory  $x_u(\cdot)$ .

In particular, since any two normal extremal lifts having  $\lambda$  as common final point have to coincide, we see that there can only be one optimal trajectory between  $x_0$  and  $x$ , which precisely means that  $x \in \Sigma_f$  is a fair point.  $\square$

*Remark.* Notice that from the previous proof it follows that, when  $\partial_P S_{x_0}^T(x) \neq \emptyset$ , then the unique normal trajectory steering  $x_0$  to  $x$  is strictly normal if and only if  $\partial_P S_{x_0}^T(x)$  is a singleton.

**Corollary 15** (Density of fair points). *The set  $\Sigma_f$  of fair points is dense in  $\text{int}(A_{x_0}^T)$ .*

In particular we have that all differentiability points of  $S_{x_0}^T$  are fair points.

**Corollary 16.** *Suppose that  $S_{x_0}^T$  is differentiable at some point  $x \in \text{int}(A_{x_0}^T)$ . Then  $x$  is a fair point, and its normal covector is  $\lambda = d_x S_{x_0}^T \in T_x^* M$ .*

*Proof.* Indeed, let  $u$  be any optimal control steering  $x_0$  to  $x$ ; then it is sufficient to consider the non-negative map

$$v \mapsto C_T(v) - S_{x_0}^T(E_{x_0}^T(v)),$$

which has by definition a local minimum at  $u$  (equal to zero). Then

$$0 = d_u C_T - (d_x S_{x_0}^T) d_u E_{x_0}^T,$$

and the uniqueness of  $u$  (hence the claim) follows as in the previous proof.  $\square$

**3.2. Continuity points.** We are also interested in the subset  $\Sigma_c$  of the *points of continuity* for the value function. It is a fact from general topology that a lower semicontinuity functions has plenty of continuity points.

**Lemma 17.** *The set  $\Sigma_c$  is a residual subset of  $\text{int}(A_{x_0}^T)$ .*

Recall that a residual subset of a topological space  $X$  is the complement of a union of countably many nowhere dense subsets of  $X$ . This fact is well-known but the proof is often presented for functions defined on complete metric spaces. For the sake of completeness, we give a proof in the Appendix.

The existence of points of continuity is tightly related to the compactness of optimal controls, as it is shown in the next lemma.

**Lemma 18.** *Let  $x \in \text{int}(A_{x_0}^T)$  be a continuity point of  $S_{x_0}^T$ . Let  $\{x_n\}_{n \in \mathbb{N}} \subset \text{int}(A_{x_0}^T)$  be a sequence converging to  $x$  and let  $u_n$  be an optimal control steering  $x_0$  to  $x_n$ . Then there exists a subsequence  $\{x_{n_k}\}_{k \in \mathbb{N}} \subset \{x_n\}_{n \in \mathbb{N}}$ , whose associated sequence of optimal controls  $\{u_{n_k}\}_{k \in \mathbb{N}}$ , strongly converges in  $L^2([0, T], \mathbb{R}^d)$  to some optimal control  $u$  which steers  $x_0$  to  $x$ .*

*Proof.* Let  $\{x_n\}_{n \in \mathbb{N}} \subset \text{int}(A_{x_0}^T)$  be a sequence converging to  $x$  and let  $\{u_n\}_{n \in \mathbb{N}}$  be the corresponding sequence of optimal controls. Since  $x$  is a continuity point for the value function, it is not restrictive to assume that the sequence of norms  $\{\|u_n\|_{L^2}\}_{n \in \mathbb{N}}$  remains uniformly bounded, and thus we can suppose to extract a subsequence  $\{u_{n_k}\}_{k \in \mathbb{N}} \subset \{u_n\}_{n \in \mathbb{N}}$  such that  $u_{n_k} \rightharpoonup u$  weakly in  $L^2([0, T], \mathbb{R}^d)$ , which in turn implies

$$\lim_{k \rightarrow \infty} \int_0^T Q(x_{u_{n_k}}(t)) dt = \int_0^T Q(x_u(t)) dt.$$

Then we have

$$\begin{aligned} \frac{1}{2} \|u\|_{L^2}^2 - \frac{1}{2} \int_0^T Q(x_u(t)) dt &\leq \liminf_{k \rightarrow \infty} \left( \frac{1}{2} \|u_{n_k}\|_{L^2}^2 - \frac{1}{2} \int_0^T Q(x_{u_{n_k}}(t)) dt \right) \\ &= \lim_{k \rightarrow \infty} S_{x_0}^T(E_{x_0}^T(u_{n_k})) = \lim_{k \rightarrow \infty} S_{x_0}^T(x_{n_k}) \\ &= S_{x_0}^T(x) = S_{x_0}^T(E_{x_0}^T(u)) \\ &\leq \frac{1}{2} \|u\|_{L^2}^2 - \frac{1}{2} \int_0^T Q(x_u(t)) dt, \end{aligned}$$

which readily means both that  $\lim_{k \rightarrow \infty} \|u_{n_k}\|_{L^2} = \|u\|_{L^2}$  (from which the convergence in  $L^2$  follows), and that  $C_T(u) = S_{x_0}^T(E_{x_0}^T(u)) = S_{x_0}^T(x)$ .  $\square$

**3.3. Tame points.** We have introduced so far two subsets of  $\text{int}(A_{x_0}^T)$ , namely the sets  $\Sigma_c$  of the continuity points of  $S_{x_0}^T$ , and the set  $\Sigma_f$  of fair points, which are essentially points that are well-parametrized by the exponential map; both these sets are dense in  $\text{int}(A_{x_0}^T)$ , still their intersection can be empty. Here is the main differences with respect to the arguments of [1]: indeed in that context every fair point is a point of continuity. In our setting, to relate  $\Sigma_c$  and  $\Sigma_f$ , we introduce the following set.

**Definition 19** (Tame Points). Let  $x \in \text{int}(A_{x_0}^T)$ . We say that  $x$  is a *tame point* if for every optimal control  $u$  steering  $x_0$  to  $x$  there holds

$$\text{rank } d_u E_{x_0}^T = \dim M = m.$$

We call  $\Sigma_t$  the set of tame points.

Tame points locate open sets on which the value function  $S_{x_0}^T$  is continuous. The precise statement is contained in the following lemma, whose first part of the proof is an adaptation of the arguments of [19, Theorem 4.6]. A complete proof is contained in Appendix A.

**Lemma 20.** *Let  $x \in \text{int}(A_{x_0}^T)$  be a tame point. Then*

- (i)  *$x$  is a point of continuity of  $S_{x_0}^T$ ;*
- (ii) *there exists a neighborhood  $O_x$  of  $x$  such that every  $y \in O_x$  is a tame point. In particular, the restriction  $S_{x_0}^T|_{O_x}$  is a continuous map.*

The previous lemma can be restated as follows.

**Corollary 21.** *The set  $\Sigma_t$  of tame points is open. Moreover  $\Sigma_t \subset \Sigma_c$ .*

#### 4. DENSITY OF TAME POINTS

This section is devoted to the proof that the set of tame point is open and dense in the interior of the attainable set. We start with the observation that the set of optimal controls reaching a fixed point  $x$  is compact in the  $L^2$ -topology.

**Lemma 22.** For every  $x \in A_{x_0}^T$ , the set

$$\mathcal{U}_x = \{u \in \Omega_{x_0}^T \mid u \text{ is an optimal control steering } x_0 \text{ to } x\}$$

is strongly compact in  $L^2([0, T], \mathbb{R}^d)$ .

*Proof.* Let  $\{u_n\}_{n \in \mathbb{N}} \subset \mathcal{U}_x$ . Then we have  $S_{x_0}^T(x) = C_T(u_n)$  for every  $n \in \mathbb{N}$ , and consequently there exists  $C > 0$  such that  $\|u_n\|_{L^2} \leq C$  for every  $n \in \mathbb{N}$ . Thus we may assume that there exists a subsequence  $\{u_{n_k}\}_{k \in \mathbb{N}} \subset \{u_n\}_{n \in \mathbb{N}}$ , and a control  $u$  steering  $x_0$  to  $x$ , such that  $u_{n_k} \rightharpoonup u$  weakly in  $L^2([0, T], \mathbb{R}^d)$ . This, on the other hand, implies that

$$\begin{aligned} \frac{1}{2}\|u\|_{L^2}^2 - \frac{1}{2}\int_0^T Q(x_u(t))dt &\leq \liminf_{k \rightarrow \infty} \frac{1}{2}\|u_{n_k}\|_{L^2}^2 - \frac{1}{2}\int_0^T Q(x_{u_{n_k}}(t))dt \\ &= \liminf_{k \rightarrow \infty} C_T(u_{n_k}) = S_{x_0}^T(x) \\ &= C_T(u) = \frac{1}{2}\|u\|_{L^2}^2 - \frac{1}{2}\int_0^T Q(x_u(t))dt, \end{aligned}$$

therefore  $\|u\|_{L^2} = \lim_{k \rightarrow \infty} \|u_{n_k}\|_{L^2}$ , and the claim is proved.  $\square$

We introduce now the notion of the *class* of a point. Heuristically, the class of a point  $x \in \text{int}(A_{x_0}^T)$  measures how much that point “fails” to be tame (see Definition 19).

**Definition 23.** Let  $x \in A_{x_0}^T$ . We define

$$\text{class}(x) = \min_{u \in \mathcal{U}_x} \text{rank } d_u E_{x_0}^T.$$

Any point  $x \in \text{int}(A_{x_0}^T)$  satisfying  $\text{class}(x) = m$  is necessarily a tame point.

**Definition 24.** We also define the subset  $\mathcal{U}_x^{\min} \subset \mathcal{U}_x$  as follows:

$$\mathcal{U}_x^{\min} = \{u \in \mathcal{U}_x \mid \text{rank } d_u E_{x_0}^T = \text{class}(x)\}.$$

By the lower semicontinuity of the rank function, the set  $\mathcal{U}_x^{\min}$  is closed in  $\mathcal{U}_x$ , hence (strongly) compact in  $L^2([0, T], \mathbb{R}^d)$ .

We can now state the main result of this section.

**Theorem 25.** The set  $\Sigma_t$  of tame points is dense in  $\text{int}(A_{x_0}^T)$ .

We postpone the proof of Theorem 25 at the end of the section, since we need first a series of preliminary results.

**Definition 26.** Pick  $x$  in  $\text{int}(A_{x_0}^T)$  and let  $u \in \mathcal{U}_x^{\min}$ . If  $u$  is not strictly abnormal, then we choose a normal covector  $\eta_x \in T_x^* M$  associated to  $u$  and we define

$$\widehat{\Xi}_x^u = \{\xi \in T_x^* M \mid \xi d_u E_{x_0}^T = \eta_x d_u E_{x_0}^T\} = \eta_x + \ker(d_u E_{x_0}^T)^* \subset T_x^* M.$$

If instead  $u$  is strictly abnormal, we simply set  $\widehat{\Xi}_x^u = \ker(d_u E_{x_0}^T)^* \subset T_x^* M$ .

Notice that whenever  $u$  is strictly abnormal, then  $\widehat{\Xi}_x^u$  is a linear subspace, while if  $u$  admits at least one normal lift,  $\widehat{\Xi}_x^u$  is affine; also, the dimension of these subspaces equals  $m - \text{class}(x) \geq 0$ . We call  $\widehat{Z}_u \subset T_x^* M$  the orthogonal subspace to  $\ker(d_u E_{x_0}^T)^*$ , of dimension equal to  $\text{class}(x)$ , for which:

$$(4.1) \quad T_x^* M = \ker(d_u E_{x_0}^T)^* \oplus \widehat{Z}_u;$$

moreover we let  $\pi_{\widehat{Z}_u} : T_x^*M \rightarrow \widehat{Z}_u$  to be the orthogonal projection subordinated to this splitting, that is satisfying:

$$\ker(\pi_{\widehat{Z}_u}) = \ker(d_u E_{x_0}^T)^*.$$

Finally, by means of the adjoint map  $(P_{0,T}^u)^*$ , we can pull the spaces  $\widehat{\Xi}_x^u$  “back” to  $T_{x_0}M$ , and set

$$\Xi_x^u := (P_{0,T}^u)^* \widehat{\Xi}_x^u \subset T_{x_0}^*M.$$

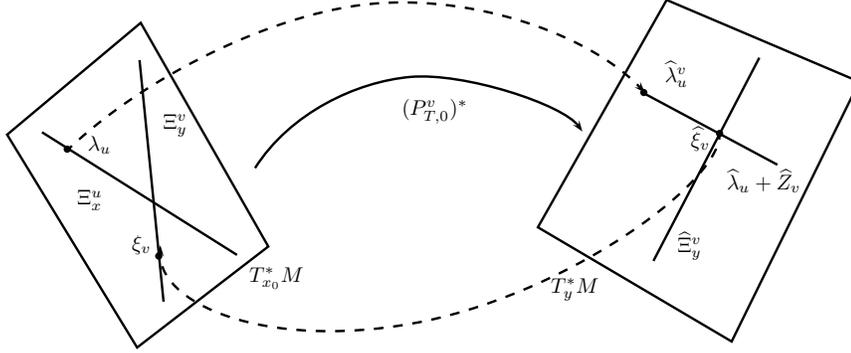


FIGURE 1. We set  $y = E_{x_0}^T(v)$ . The subspace  $\widehat{\Xi}_y^v$  is linear if  $v$  is strictly abnormal, and affine otherwise;  $\widehat{Z}_v$  and  $\ker(d_v E_{x_0}^T)^*$  are orthogonal. The point  $\widehat{\xi}_v$  belong to  $T_y^*M$ , and is then pulled back on  $T_{x_0}^*M$ .

The following estimate will be crucial in what follows.

**Proposition 27.** *Let  $O \subset \text{int}(A_{x_0}^T)$  be an open set, and assume that:*

$$\text{class}(z) \equiv k_O < m, \quad \text{for every } z \in O.$$

*Let  $x \in O$  and  $u \in \mathcal{U}_x^{\min}$ . Then there exists a neighborhood  $\mathcal{V}_u \subset \Omega_{x_0}^T$  of  $u$  such that, for every  $\lambda_u \in \Xi_x^u \subset T_{x_0}^*M$ , there exists a constant  $K = K(\lambda_u) > 1$  such that, for every  $v \in \mathcal{V}_u \cap \mathcal{U}_{E_{x_0}^T(v)}^{\min}$ , there is  $\xi_v \in \Xi_{E_{x_0}^T(v)}^v \subset T_{x_0}^*M$  satisfying:*

$$|\lambda_u - \xi_v| \leq K.$$

*Proof.* Let us choose a neighborhood  $\mathcal{V}_u \subset \Omega_{x_0}^T$  of  $u$ , such that all the endpoints of admissible trajectories driven by controls in  $\mathcal{V}_u$  belong to  $O$ .

Then, if  $y = E_{x_0}^T(v)$  for some  $v \in \mathcal{V}_u$ , it follows that  $y \in O$ ; moreover, if also  $v \in \mathcal{U}_y^{\min}$ , we can define the  $(m - k_O)$ -dimensional subspace  $\Xi_y^v \subset T_{x_0}^*M$  as in Definition 26. Therefore we can assume from the beginning that all such subspaces  $\Xi_y^v$  have dimension constantly equal to  $m - k_O > 0$ .

Fix  $\lambda_u \in \Xi_x^u$ , and set

$$\widehat{\lambda}_u^v = (P_{T,0}^v)^* \lambda_u \in T_y^*M, \quad v \in \mathcal{V}_u \cap \mathcal{U}_y^{\min}, \quad y = E_{x_0}^T(v).$$

The intersection  $(\widehat{\lambda}_u^v + \widehat{Z}_v) \cap \widehat{\Xi}_y^v$  (cf. with (4.1) and Figure 1) consists of the single point  $\widehat{\xi}_v$ ; since both  $\widehat{\lambda}_u^v$  and  $\widehat{\xi}_v$  belong to the affine subspace  $\widehat{\lambda}_u^v + \widehat{Z}_v$ , in order to estimate the norm  $|\widehat{\lambda}_u^v - \widehat{\xi}_v|$

it is sufficient to evaluate the norm  $|\pi_{\widehat{Z}_v}(\widehat{\lambda}_u^v) - \pi_{\widehat{Z}_v}(\widehat{\xi}_v)|$  of the projections onto the linear space  $\widehat{Z}_v = (\ker(d_v E_{x_0}^T)^*)^\perp$ . The key point is the computation of the norm of  $|\pi_{\widehat{Z}_v}(\widehat{\xi}_v)|$ : in fact, since  $\ker(d_v E_{x_0}^T)^* = (\text{Im } d_v E_{x_0}^T)^\perp$ , this amounts to evaluate

$$(4.2) \quad |\pi_{\widehat{Z}_v}(\widehat{\xi}_v)| = \sup_{f \in \text{Im } d_v E_{x_0}^T} \frac{|\langle \widehat{\xi}_v, f \rangle|}{|f|}.$$

We deduce immediately from (4.2) that, whenever  $v$  is strictly abnormal, then  $\pi_{\widehat{Z}_v}(\widehat{\xi}_v) = 0$ , while from the expression for the normal control (2.3)

$$v_i(t) = \langle \widehat{\xi}_v(t), X_i(x_v(t)) \rangle = \langle \widehat{\xi}_v, (P_{T,t}^v)^* X_i(x_v(t)) \rangle,$$

we see that  $\langle v, w \rangle_{L^2} = \langle \widehat{\xi}_v, d_v E_{x_0}^T(w) \rangle$ , and we can continue from (4.2) as follows ( $W_v$  denotes the  $k_O$ -dimensional subspace of  $L^2([0, T], \mathbb{R}^d)$  on which the restriction  $d_v E_{x_0}^T|_{W_v}$  is invertible):

$$(4.3) \quad \begin{aligned} |\pi_{\widehat{Z}_v}(\widehat{\xi}_v)| &= \sup_{w \in W_v} \frac{|\langle \widehat{\xi}_v, d_v E_{x_0}^T(w) \rangle|}{|d_v E_{x_0}^T(w)|} \\ &\leq \sup_{w \in W_v} \frac{|\langle \widehat{\xi}_v, d_v E_{x_0}^T(w) \rangle|}{\|w\|_{L^2}} \|(d_v E_{x_0}^T|_{W_v})^{-1}\| \\ &= \sup_{w \in W_v} \frac{|\langle v, w \rangle|}{\|w\|_{L^2}} \|(d_v E_{x_0}^T|_{W_v})^{-1}\| \\ &\leq \|v\|_{L^2} \|(d_v E_{x_0}^T|_{W_v})^{-1}\|. \end{aligned}$$

It is not restrictive to assume that the  $L^2$ -norm of any element  $v \in \mathcal{V}_u \cap \mathcal{U}_y^{\min}$  remains bounded; moreover, since all the subspaces have the same dimension, the map  $v \mapsto W_v$  is continuous, which implies that so is the map  $v \mapsto (d_v E_{x_0}^T|_{W_v})^{-1}$ . This, on the other hand, guarantees that the operator norm  $\|(d_v E_{x_0}^T|_{W_v})^{-1}\|$  remains bounded for all  $v \in \mathcal{V}_u \cap \mathcal{U}_y^{\min}$ , and then from (4.3) we conclude that for some  $C > 1$ , the estimate  $|\pi_{\widehat{Z}_v}(\widehat{\xi}_v)| \leq C$  holds true, which implies as well, by the triangular inequality, that:

$$|\widehat{\lambda}_u^v - \widehat{\xi}_v| \leq |\widehat{\lambda}_u^v| + C.$$

Finally, the continuity of both the map  $v \mapsto P_{0,T}^v$  and its inverse, implies that for another real constant  $C > 1$  we have:

$$\sup_{v \in \mathcal{V}_u} \{ \|(P_{0,T}^v)^*\|, \|(P_{T,0}^v)^*\| \} \leq C.$$

Thus, setting  $\xi_v = (P_{0,T}^v)^* \widehat{\xi}_v \in T_{x_0}^* M$  (cf. Figure 1) we can compute (here  $C$  denotes a constant that can change from line to line):

$$\begin{aligned} |\lambda_u - \xi_v| &\leq C |\widehat{\lambda}_u^v - \widehat{\xi}_v| \\ &\leq C |\widehat{\lambda}_u^v| + C^2 \\ &\leq C^2 (|\lambda_u| + 1) \\ &\leq 2C^2 \max\{|\lambda_u|, 1\}. \end{aligned}$$

Setting  $K(\lambda_u) := 2C^2 \max\{|\lambda_u|, 1\}$  the claim is proved.  $\square$

*Remark.* Let us fix  $\lambda_u \in \Xi_x^u \subset T_{x_0}^*M$  and consider the  $k_O$ -dimensional affine subspace

$$(P_{0,T}^v)^*(\widehat{\lambda}_u^v + \widehat{Z}_v) = \lambda_u + (P_{0,T}^v)^*\widehat{Z}_v,$$

with  $\widehat{Z}_v$  defined as in (4.1). Then if we call  $Z_v := (P_{0,T}^v)^*\widehat{Z}_v \subset T_{x_0}^*M$ , the map

$$v \mapsto \lambda_u + Z_v, \quad v \in \mathcal{V}_u \cap \mathcal{U}_y^{\min}, \quad y = E_{x_0}^T(v)$$

is continuous; moreover,  $Z_v$  is by construction transversal to  $\Xi_y^v$ , and  $\xi_v \in (\lambda_u + Z_v) \cap \Xi_y^v$ .

Having in mind this remark, we deduce the following:

**Corollary 28.** *Let  $O \subset \text{int}(A_{x_0}^T)$  be an open set, and assume that*

$$\text{class}(z) \equiv k_O < m, \quad \text{for every } z \in O.$$

*Let  $x \in O$ ,  $u \in \mathcal{U}_x^{\min}$ , and consider  $\mathcal{V}_u \subset \Omega_{x_0}^T$  as in Proposition 27. Then, for every  $\lambda_u \in \Xi_x^u$ , there exists a  $k_O$ -dimensional compact ball  $A_u$ , centered at  $\lambda_u$  and transversal to  $\Xi_x^u$ , such that:*

$$A_u \cap \Xi_y^v \neq \emptyset \quad \text{for every } v \in \mathcal{V}_u \cap \mathcal{U}_y^{\min}, \quad \text{where } y = E_{x_0}^T(v).$$

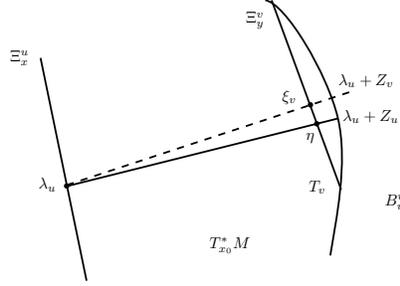


FIGURE 2. On the fiber  $T_{x_0}^*M$ , the point  $\eta$  denotes the intersection between  $T_v$  and the affine space  $\lambda_u + Z_u$ .

*Proof.* Let  $\lambda_u \in \Xi_x^u$  be chosen, and assume without loss of generality that  $\mathcal{V}_u$  is relatively compact. For every  $v \in \mathcal{V}_u$ , we can construct an  $m$ -dimensional ball  $B_u^v$ , of radius  $C_0^v$  strictly greater than  $K = K(\lambda_u)$  (given by Proposition 27), and centered at  $\lambda_u$ .

Then, the existence of an element  $\xi_v \in (\lambda_u + Z_v) \cap \Xi_y^v$  satisfying  $|\lambda_u - \xi_v| \leq K$ , proved in Proposition 27, implies that the intersection of  $B_u^v$  with  $\Xi_y^v$  is a compact submanifold  $T_v$  (with boundary); moreover, since the radius of  $B_u^v$  is strictly greater than  $|\lambda_u - \xi_v|$ , it is also true that the intersection of  $\lambda_u + Z_v$  with  $\text{int}(T_v)$  is not empty.

Let us consider as before (cf. previous remark) the  $k_O$ -dimensional affine subspace  $\lambda_u + Z_u$ , which is transversal to  $\Xi_x^u$ : possibly increasing the radius  $C_0^v$ , the continuity of the map  $w \mapsto \lambda_u + Z_w$  ensures that  $\lambda_u + Z_u$  remains transversal to  $T_v$ , and in particular that the intersection  $T_v \cap (\lambda_u + Z_u)$  is not empty (see Figure 2). Moreover, it is clear that this conclusion is local, that is with the same choice of  $C_0^v$  it can be drawn on some full neighborhood  $\mathcal{W}_v$  of  $v$ . Then, to find a ball  $B_u$  and a radius  $C_0$  uniformly for the whole set  $\mathcal{V}_u$ , it is sufficient to extract a finite sub-cover  $\mathcal{W}_{v_1}, \dots, \mathcal{W}_{v_l}$  of  $\mathcal{V}_u$ , and choose  $C_0$  as the maximum between  $C_0^{v_1}, \dots, C_0^{v_l}$ .

We conclude the proof setting  $A_u = B_u \cap (\lambda_u + Z_u)$ ; indeed  $A_u$  is a compact  $k_O$ -dimensional ball by construction, and moreover if we call  $\eta_v$  any element in the intersection  $T_v \cap (\lambda_u + Z_u)$ , for  $v \in \mathcal{V}_u$ , then it follows that:

$$\eta_v \in \Xi_y^v \cap B_u \cap (\lambda_u + Z_u) = \Xi_y^v \cap A_u,$$

that is, the intersection  $\Xi_y^v \cap A_u$  is not empty for every  $v \in \mathcal{V}_u \cap \mathcal{U}_y^{\min}$ .  $\square$

**Lemma 29.** *Let  $O \subset \text{int}(A_{x_0}^T)$  be an open set, and let*

$$k_O = \max_{x \in \Sigma_c \cap O} \text{class}(x).$$

*Then there exists a neighborhood  $O' \subset O$ , such that  $\text{class}(y) = k_O$ , for every  $y \in O'$ .*

*Proof.* Let  $x \in \Sigma_c \cap O$  be a point of continuity for the value function  $S_{x_0}^T$ , having the property that  $\text{class}(x) = k_O$ . Assume by contradiction that we can find a sequence  $\{x_n\}_{n \in \mathbb{N}}$  converging to  $x$  and satisfying  $\text{class}(x_n) \leq k_O - 1$  for every  $n \in \mathbb{N}$ . Accordingly, let  $u_n \in \mathcal{U}_{x_n}^{\min}$  an associated sequence of optimal controls; in particular, for every  $n \in \mathbb{N}$ , we have by definition that  $\text{class}(x_n) = \text{rank } d_{u_n} E_{x_0}^T$ .

By Lemma 18, we can extract a subsequence  $\{u_{n_k}\}_{k \in \mathbb{N}} \subset \{u_n\}_{n \in \mathbb{N}}$  which converges to some optimal control  $u$  steering  $x_0$  to  $x$ , strongly in the  $L^2$ -topology, and write:

$$\text{class}(x) \leq \text{rank } d_u E_{x_0}^T \leq \liminf_{k \rightarrow \infty} \text{rank } d_{u_{n_k}} E_{x_0}^T = \liminf_{k \rightarrow \infty} \text{class}(x_{n_k}) \leq k_O - 1,$$

which is absurd by construction, and the claim follows.  $\square$

Collecting all the results we can now prove Theorem 25.

*Proof of Theorem 25.* Let  $O$  be an open set in  $\text{int}(A_{x_0}^T)$  and define

$$k_O = \max_{x \in \Sigma_c \cap O} \text{class}(x);$$

notice that this definition makes sense, since points of continuity are dense in  $\text{int}(A_{x_0}^T)$  by Proposition 17. Then we may suppose that  $k_O$  is strictly less than  $m$ , for otherwise there would be nothing to prove. Moreover, by Lemma 29 it is not restrictive to assume that  $\text{class}(y) = k_O$  for every  $y \in O$ .

Fix then a point  $x \in \Sigma_c \cap O$ ; since the hypotheses of Proposition 27 are satisfied, for every  $u \in \mathcal{U}_x^{\min}$  we can find a neighborhood  $\mathcal{V}_u \subset \Omega_{x_0}^T$  of  $u$ , fix  $\lambda_u \in \Xi_x^u$ , and construct accordingly a compact  $k_O$ -dimensional ball  $A_u$ , centered at  $\lambda_u$  and transversal to  $\Xi_x^u$ , such that (Corollary 28)

$$A_u \cap \Xi_y^v \neq \emptyset \quad \text{for every } v \in \mathcal{V}_u, \text{ and with } y = E_{x_0}^T(v).$$

Since  $\mathcal{U}_x^{\min}$  is compact (Definition 24), we can choose finitely many elements  $u_1, \dots, u_l$  in  $\mathcal{U}_x^{\min}$  such that

$$\mathcal{U}_x^{\min} \subset \bigcup_{i=1}^l \mathcal{V}_{u_i}.$$

The union  $A_{u_1} \cup \dots \cup A_{u_l}$  is again of positive codimension. Moreover, for every sequence  $x_n$  of fair points converging to  $x$ , and whose associated sequence of optimal controls (by uniqueness of the optimal control, necessarily  $u_n \in \mathcal{U}_{x_n}^{\min}$ ) the sequence  $u_n$  converges to some  $v \in \mathcal{V}_{u_i} \subset \mathcal{U}_x^{\min}$ , we have that  $A_{u_i}$  is also transversal to  $\Xi_{x_n}^{u_n}$ . In particular, possibly enlarging the ball  $A_{u_i}$ , we can assume that

$$A_{u_i} \cap \Xi_{x_n}^{u_n} \neq \emptyset, \quad \text{for every } n \in \mathbb{N}.$$

For any fair point  $z \in \Sigma_f \cap O$ , the optimal control admits a normal lift, and we have the equality

$$\mathcal{E}_{x_0}^T(\Xi_z^u) = z,$$

where  $\mathcal{E}_{x_0}^T$  is the exponential map with base point  $x_0$  at time  $T$  of Definition 9, so that we eventually deduce the inclusion:

$$(4.4) \quad \Sigma_f \cap O \subset \mathcal{E}_{x_0}^T(A_{u_1} \cup \dots \cup A_{u_l}).$$

The set on the right-hand side is closed, being the image of a compact set; moreover, it is of measure zero by the classical Sard Lemma [18], as it is the image of a set of positive codimension by construction. Since the set  $\Sigma_f \cap O$  is dense in  $O$  by Corollary 15, passing to the closures in (4.4) we conclude that  $\text{meas}(O) = 0$ , which is impossible.  $\square$

Combining now Lemma 20 and Theorem 25 we obtain the following (cf. Theorem 1).

**Corollary 30.** *The set  $\Sigma_t$  of tame points is open and dense in  $\text{int}(A_{x_0}^T)$ .*

## 5. ON THE SMOOTHNESS

In this section we deduce smoothness of the value function  $S_{x_0}^T$  in the presence of tame points. Since tame points are in particular points of continuity for  $S_{x_0}^T$ , the arguments of Lemma 18, with minor changes, prove the following result.

**Lemma 31.** *Let  $K \subset \Sigma_t$  be a compact subset of tame points. Then the set of optimal controls reaching points of  $K$*

$$\mathcal{M}_K = \{u \in \Omega_{x_0}^T \mid E_{x_0}^T(u) \in K \text{ and } C_T(u) = S_{x_0}^T(E_{x_0}^T(u))\}$$

*is strongly compact in the  $L^2$ -topology.*

The first result of this section, which is an adaptation of an argument of [15, 1], is as follows:

**Proposition 32.** *Let  $K \subset \Sigma_t$  be a compact subset of tame points. Then  $S_{x_0}^T$  is Lipschitz continuous on  $K$ .*

*Proof.* By compactness, it is sufficient to show that  $S_{x_0}^T$  is locally Lipschitz continuous on  $K$ .

Fix a point  $x \in K$  and let  $u$  be associated with an optimal trajectory joining  $x_0$  and  $x$ . By assumption,  $d_u E_{x_0}^T$  is surjective, so that there are neighborhoods  $\mathcal{V}_u \subset \Omega_{x_0}^T$  of  $u$  and  $O_x \subset \text{int}(A_{x_0}^T)$  of  $x$  such that

$$E_{x_0}^T|_{\mathcal{V}_u} : \mathcal{V}_u \rightarrow O_x$$

is surjective, and there exists a smooth right inverse  $\Phi : O_x \rightarrow \mathcal{V}_u$  such that  $E_{x_0}^T(\Phi(y)) = y$  for every  $y \in O_x$ .

Fix local coordinates around  $x$ , and let  $B_x(r) \subset M$  and  $\mathcal{B}_u(r) \subset \Omega_{x_0}^T$  denote some balls of radius  $r > 0$  centered at  $x$  and  $u$  respectively. As  $\Phi$  is smooth, there exists  $R > 0$  and  $C_0 > 0$  such that:

$$(5.1) \quad B_x(C_0 r) \subset E_{x_0}^T(\mathcal{B}_u(r)), \quad \text{for every } 0 \leq r \leq R.$$

Observe that there also exists  $C_1 > 0$  such that, for every  $v, w \in \mathcal{B}_u(R)$  we have

$$(5.2) \quad |C_T(v) - C_T(w)| \leq C_1 \|v - w\|_{L^2}.$$

Indeed our main assumption implies that the subset  $\{x_v(t) \mid t \in [0, T], v \in \mathcal{B}_u(R)\}$  is contained in a compact set  $K$  of  $M$ , on which the smooth function  $Q$ , together with its differential

$Q'$ , attains both a maximum and a minimum. Then, using the mean value theorem and [19, Proposition 3.5], we deduce that

$$\int_0^T |Q(x_v(t)) - Q(x_w(t))| dt \leq \sup_{y \in K} |Q'(y)| \int_0^T |x_v(t) - x_w(t)| dt \leq C \|v - w\|_{L^2},$$

and by means of the triangular inequality, (5.2) is proved.

Pick any point  $y \in K$  such that  $|y - x| = C_0 r$ , with  $0 \leq r \leq R$ . Then by (5.1) there exists  $v \in \mathcal{B}_u(r)$  satisfying  $\|u - v\|_{L^2} \leq r$  and such that  $E_{x_0}^T(v) = y$ ; since  $C_T(u) = S_{x_0}^T(x)$  and  $S_{x_0}^T(y) \leq C_T(v)$ , we have

$$S_{x_0}^T(y) - S_{x_0}^T(x) \leq C_T(v) - C_T(u) \leq C_1 \|v - u\|_{L^2} \leq \frac{C_1}{C_0} |y - x|.$$

Using the compactness of both  $K$  and  $\mathcal{M}_K$  (cf. Lemma 31), all the constants can be made uniform, and the role of  $x$  and  $y$  can be exchanged, so that we have indeed

$$|S_{x_0}^T(x) - S_{x_0}^T(y)| \leq \frac{C_1}{C_0} |x - y|,$$

for every pair of points  $x$  and  $y$  such that  $|x - y| \leq C_0 R$ .  $\square$

**Definition 33.** We define the set  $\Sigma \subset \text{int}(A_{x_0}^T)$  of the *smooth points* as the set of points  $x$  such that

- (a) there exists a unique optimal trajectory  $t \mapsto x_u(t)$  steering  $x_0$  to  $x$  in time  $T$ , which is strictly normal,
- (b)  $x$  is not conjugate to  $x_0$  along  $x_u(\cdot)$  (cf. Definition 9).

Item (a) in the Definition 33 is equivalent to require that  $x$  is in fact a point that is at the same time fair and tame. Notice that as a consequence of the results of Section 3, and in particular of Corollary 30, the set  $\Sigma_f \cap \Sigma_t$  is dense in  $\text{int}(A_{x_0}^T)$ .

The following result finally proves Theorem 2.

**Theorem 34** (Density of smooth points).  *$\Sigma$  is open and dense in  $\text{int}(A_{x_0}^T)$ . Moreover  $S_{x_0}^T$  is smooth on  $\Sigma$ .*

*Proof.* (i.a) Let us show that  $\Sigma$  is dense. First we prove that, for any open set  $O$ , we have  $\Sigma \cap O \neq \emptyset$ . Since the set  $\Sigma_t$  of tame points is open and dense in  $\text{int}(A_{x_0}^T)$ , we can choose a subset  $O' \subset O \cap \Sigma_t$  relatively compact, and assume by Proposition 32 that  $S_{x_0}^T$  is Lipschitz on  $O'$ . Thanks to the classical Rademacher theorem we know that  $S_{x_0}^T$  is differentiable almost everywhere on  $O'$ , and therefore, since any point of differentiability is a fair point by Corollary 16,  $\text{meas}(\Sigma_f \cap O') = \text{meas}(O')$ . Moreover, any point in  $\Sigma_f \cap O'$  is also contained in the image of the exponential map  $\mathcal{E}_{x_0}^T$ , and Sard Lemma implies that the set of regular points is of full measure in  $\Sigma_f \cap O'$ . By definition any such point is in  $\Sigma$ , that is we have  $\text{meas}(\Sigma \cap O') = \text{meas}(\Sigma_f \cap O') = \text{meas}(O')$ , which implies that  $\Sigma \cap O' \neq \emptyset$ , and this concludes the proof.

(i.b) Let us prove that  $\Sigma$  is open. Fix as before an open set  $O$  having compact closure in  $\text{int}(A_{x_0}^T)$ . Assume by contradiction that there exists a sequence of points  $x_n \in O$  converging to  $x \in \Sigma$  and such that there are (at least) two optimal trajectories connecting them with  $x_0$ . Call  $\{u_n\}_{n \in \mathbb{N}}$  and  $\{v_n\}_{n \in \mathbb{N}}$  the corresponding sequences of optimal controls associated with such trajectories. Lemma 18 then guarantees that, up to considering subsequences, it is not restrictive to assume the existence of both  $u = \lim_{n \rightarrow \infty} u_n$  and  $v = \lim_{n \rightarrow \infty} v_n$  in  $L^2([0, T], \mathbb{R}^d)$ ; however, the uniqueness of the minimizer steering  $x_0$  to  $x$  implies that  $u = v$ .

Then both  $d_{u_n} E_{x_0}^T$  and  $d_{v_n} E_{x_0}^T$  have maximal rank for  $n$  large enough ( $u$  is strictly normal because  $x$  is a smooth point), and we can define the families of covectors  $\lambda_n$  and  $\xi_n$ , as elements of  $T_{x_n}^* M$ , satisfying the identities:

$$\lambda_n d_{u_n} E_{x_0}^T = d_{u_n} C_T, \quad \xi_n d_{v_n} E_{x_0}^T = d_{v_n} C_T.$$

Taking the limit on these two equations we see that  $\lim_{n \rightarrow \infty} \lambda_n = \lim_{n \rightarrow \infty} \xi_n = \lambda$ , where  $\lambda$  is the covector associated with the unique optimal control  $u$  steering  $x_0$  to  $x$ . If, for any  $s \in [0, T]$ , we let  $\lambda_n^s = (P_{s,T}^{u_n})^* \lambda_n$  and  $\xi_n^s = (P_{s,T}^{v_n})^* \xi_n$ , then we see that even the ‘‘initial covectors’’  $\lambda_n^0$  and  $\xi_n^0$  converge to the same element  $\lambda^0$ .

On the other hand, since by the point (b) of Definition 33  $x$  is not conjugate to  $x_0$  along the unique optimal trajectory  $x_u(\cdot)$ , we have that  $\lambda^0$  is a regular point for the exponential map  $\mathcal{E}_{x_0}^T$ . Then there exist full neighborhoods  $V \subset T_{x_0}^* M$  of  $\lambda^0$  and  $O_x \subset \text{int}(A_{x_0}^T)$  of  $x$  such that the exponential map  $\mathcal{E}_{x_0}^T|_V : V \rightarrow O_x$  is a diffeomorphism. In particular, if we pick some point  $y \in O_x$ , there is a unique optimal trajectory  $x_u(\cdot)$  steering  $x_0$  to  $y$ ; moreover the covector  $\lambda_y$  associated with  $x_u(\cdot)$  is a regular point for  $\mathcal{E}_{x_0}^T$ , and from the equality  $E_{x_0}^T(u) = \mathcal{E}_{x_0}^T(\lambda_y)$ , we see that  $u$  has to be strictly normal. This shows that  $O_x \subset \Sigma$ , which in the end is an open set.

(ii). Next we prove the smoothness of  $S_{x_0}^T$  on  $\Sigma$ . Let us consider a covector  $\lambda \in T_{x_0}^* M$  associated with the unique optimal trajectory connecting  $x_0$  and  $x$ . By the arguments of the previous point, there are neighborhoods  $V_\lambda \subset T_{x_0}^* M$  of  $\lambda$  and  $O_x \subset \text{int}(A_{x_0}^T)$  of  $x$  such that  $\mathcal{E}_{x_0}^T|_{V_\lambda} : V_\lambda \rightarrow O_x$  is a diffeomorphism.

It is then possible to define a smooth inverse  $\Phi : O_x \rightarrow V_\lambda$  sending  $y$  to the corresponding ‘‘initial’’ covector  $\lambda_y$ . Along (strictly normal) trajectories associated with covectors  $\lambda_y$  in  $V_\lambda$  we have therefore (compare with (2.3)):

$$u_i^y(t) = \langle \Phi(y), X_i(x_u^y(t)) \rangle,$$

which means that the control  $u^y \in \Omega_{x_0}^T$  and, in turn, the cost  $C_T(u)$  itself, are smooth on  $O_x$ .  $\square$

#### APPENDIX A. A FEW TECHNICAL RESULTS

We give here the proof of Lemmas 17 and 20.

**Lemma.** *The set  $\Sigma_c$  is a residual subset of  $\text{int}(A_{x_0}^T)$ .*

*Proof.* We will show that the complement of  $\Sigma_c$  is a meager set, i.e. it can be included into a countable union of closed, nowhere dense subsets of  $\text{int}(A_{x_0}^T)$ . Then the claim will follow from the classical Baire category theorem, that holds on smooth manifolds.

Let then  $x$  be a discontinuity point of  $S_{x_0}^T$ . This implies that  $S_{x_0}^T$  is not upper semicontinuous at  $x$ , i.e. there exists  $\varepsilon > 0$  and a sequence  $x_n \rightarrow x$  such that for all  $n$

$$S_{x_0}^T(x) + \varepsilon \leq S_{x_0}^T(x_n).$$

For any  $q \in \mathbb{Q}$  define the set

$$K_q = \{x \in \text{int}(A_{x_0}^T) \mid S_{x_0}^T(x) \leq q\};$$

the lower semicontinuity of  $S_{x_0}^T$  implies that  $K_q$  is closed. Moreover, let us choose  $r \in \mathbb{Q}$  such that  $S_{x_0}^T(x) < r < S_{x_0}^T(x) + \varepsilon$ ; then by construction  $x \in K_r \setminus \text{int}(K_r)$ , which means that

$$\text{int}(A_{x_0}^T) \setminus \Sigma_c \subset \bigcup_{r \in \mathbb{Q}} (K_r \setminus \text{int}(K_r)).$$

$\square$

**Lemma.** *Let  $x \in \text{int}(A_{x_0}^T)$  be a tame point. Then*

- (i)  $x$  is a point of continuity of  $S_{x_0}^T$ ;
- (ii) there exists a neighborhood  $O_x$  of  $x$  such that every  $y \in O_x$  is a tame point. In particular, the restriction  $S_{x_0}^T|_{O_x}$  is a continuous map.

*Proof.* To prove (i) we will show that, for every sequence  $\{x_n\}_{n \in \mathbb{N}}$  converging to  $x$ , there holds  $\lim_{n \rightarrow +\infty} S_{x_0}^T(x_n) = S_{x_0}^T(x)$ ; in particular we will prove the latter equality by showing that  $S_{x_0}^T(x)$  is the unique cluster point for all such sequences  $\{S_{x_0}^T(x_n)\}_{n \in \mathbb{N}}$ .

Let  $u$  be any optimal control steering  $x_0$  to  $x$ ; by hypothesis  $d_u E_{x_0}^T$  is surjective, and therefore  $E_{x_0}^T$  is locally open at  $u$ , which means that there exists a neighborhood  $\mathcal{V}_u \subset \Omega_{x_0}^T$  of  $u$  such that the image  $E_{x_0}^T(\mathcal{V}_u)$  covers a full neighborhood of  $x$  in  $\text{int}(A_{x_0}^T)$ . This implies that, for  $n$  large enough, the  $L^2$ -norms  $\{\|u_n\|_{L^2}\}_{n \in \mathbb{N}}$  of optimal controls steering  $x_0$  to  $x_n$  remain uniformly bounded by some positive constant  $C$ .

Let now  $a$  be a cluster point for the sequence  $\{S_{x_0}^T(x_n)\}_{n \in \mathbb{N}}$ . Then, it is not restrictive to assume that  $\lim_{n \rightarrow \infty} S_{x_0}^T(x_n) = a$ . Moreover, our previous point implies that we can find a subsequence  $\{x_{n_k}\}_{k \in \mathbb{N}}$ , whose associated sequence of optimal controls  $\{u_{n_k}\}_{k \in \mathbb{N}}$  weakly converge in  $L^2([0, T], \mathbb{R}^d)$  to some admissible control  $u$  steering  $x_0$  to  $x$ , which in turn yields the inequality

$$S_{x_0}^T(x) \leq C_T(u) \leq \liminf_{k \rightarrow \infty} C_T(u_{n_k}) = \liminf_{k \rightarrow \infty} S_{x_0}^T(x_{n_k}) = a.$$

Let us assume by contradiction that  $S_{x_0}^T(x) = b < a$ , and let  $\varepsilon > 0$  be such that  $b + \varepsilon < a$ ; moreover, let  $v$  be an optimal control attaining that cost. By the tameness assumption, the endpoint map  $E_{x_0}^T$  is open in a (strong) neighborhood  $\mathcal{V}_v \subset \Omega_{x_0}^T$  of  $v$ , which means that all points  $y$  sufficiently close to  $x$  can be reached by admissible (but not necessarily optimal) trajectories, driven by controls  $w \in \mathcal{V}_v$  satisfying  $C_T(w) \leq b + \varepsilon < a$ . But this gives a contradiction since  $S_{x_0}^T(x_{n_k})$  must become arbitrarily close to  $a$ , as  $k$  goes to infinity.

To prove (ii), assume by contradiction that such a neighborhood  $O_x$  does not exist; then we can find a sequence  $\{x_n\}_{n \in \mathbb{N}}$  convergent to  $x$ , and such that for every  $n \in \mathbb{N}$  there exists a choice of an abnormal optimal control  $u_n$  steering  $x_0$  to  $x_n$ , that is for every  $n \in \mathbb{N}$  there exists a norm-one covector  $\lambda_n$  such that:

$$(A.1) \quad \lambda_n d_{u_n} E_{x_0}^T = 0.$$

By Lemma 18, there exists a subsequence  $u_{n_k}$  which converges strongly in  $L^2([0, T], \mathbb{R}^d)$  to some optimal control  $u$  reaching  $x$ ; moreover, since we assumed  $|\lambda_n| = 1$  for all  $n \in \mathbb{N}$ , it is not restrictive to suppose that  $\bar{\lambda} = \lim_{k \rightarrow \infty} \lambda_{n_k}$  exists. Thus, passing to the limit as  $k$  tend to infinity in (A.1), we see that  $u$  is forced to be abnormal, and thus we have a contradiction, as  $x$  is tame. It follows then from point (i) that  $S_{x_0}^T|_{O_x}$  is indeed a continuous map. □

#### ACKNOWLEDGMENTS

The authors wish to thank A.A. Agrachev for bringing the problem to our attention, and for stimulating discussions. This research has been supported by the European Research Council, ERC StG 2009 ‘‘GeCoMethods’’, contract number 239748 and by the ANR project SRGI ‘‘Sub-Riemannian Geometry and Interactions’’, contract number ANR-15-CE40-0018.

#### REFERENCES

- [1] Andrei A. Agrachev. Any sub-Riemannian metric has points of smoothness. *Dokl. Akad. Nauk*, 424(3):295–298, 2009.
- [2] Andrei A. Agrachev, D. Barilari, and L. Rizzi. Curvature: a variational approach, <https://arxiv.org/pdf/1306.5318.pdf>. *Memoirs of the AMS*, in press.

- [3] Andrei A. Agrachev, Davide Barilari, and Ugo Boscin. Introduction to Riemannian and sub-Riemannian geometry, [http://people.sissa.it/agrachev/agrachev\\_files/notes.html](http://people.sissa.it/agrachev/agrachev_files/notes.html). *Preprint SISSA 09/2012/M*, 2012.
- [4] Andrei A. Agrachev and Paul Lee. Optimal transportation under nonholonomic constraints. *Trans. Amer. Math. Soc.*, 361(11):6019–6047, 2009.
- [5] Andrei A. Agrachev and Yuri L. Sachkov. *Control theory from the geometric viewpoint*, volume 87 of *Encyclopaedia of Mathematical Sciences*. Springer-Verlag, Berlin, 2004. Control Theory and Optimization, II.
- [6] Martino Bardi and Italo Capuzzo-Dolcetta. *Optimal control and viscosity solutions of Hamilton-Jacobi-Bellman equations*. Systems & Control: Foundations & Applications. Birkhäuser Boston, Inc., Boston, MA, 1997. With appendices by Maurizio Falcone and Pierpaolo Soravia.
- [7] Francesco Boarotto and Antonio Lerario. Homotopy Properties of Horizontal Path Spaces and a Theorem of Serre in sub-Riemannian Geometry. *Accepted on CAG*.
- [8] Piermarco Cannarsa and Carlo Sinestrari. *Semiconcave functions, Hamilton-Jacobi equations, and optimal control*. Progress in Nonlinear Differential Equations and their Applications, 58. Birkhäuser Boston, Inc., Boston, MA, 2004.
- [9] Y. Chitour, F. Jean, and E. Trélat. Singular trajectories of control-affine systems. *SIAM J. Control Optim.*, 48(2):1078–1095, 2008.
- [10] Francis Clarke, Yuri Ledyev, Ronald Jay Stern, and Peter Wolenski. *Nonsmooth Analysis and Control Theory*, volume 178. Springer-Verlag, New York, 1998.
- [11] Hélène Frankowska. Value function in optimal control. In *Mathematical control theory, Part 1, 2 (Trieste, 2001)*, ICTP Lect. Notes, VIII, pages 516–653 (electronic). Abdus Salam Int. Cent. Theoret. Phys., Trieste, 2002.
- [12] Velimir Jurdjevic. *Geometric control theory*, volume 52 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 1997.
- [13] L. S. Pontryagin, V. G. Boltyanskiĭ, R. V. Gamkrelidze, and E. F. Mishchenko. *Selected works. Vol. 4*. Classics of Soviet Mathematics. Gordon & Breach Science Publishers, New York, 1986. The mathematical theory of optimal processes, Edited and with a preface by R. V. Gamkrelidze, Translated from the Russian by K. N. Trirogoff, Translation edited by L. W. Neustadt, With a preface by L. W. Neustadt and K. N. Trirogoff, Reprint of the 1962 English translation.
- [14] Dario Prandi. Hölder equivalence of the value function for control-affine systems. *ESAIM Control Optim. Calc. Var.*, 20(4):1224–1248, 2014.
- [15] L. Rifford and E. Trélat. Morse-Sard type results in sub-Riemannian geometry. *Math. Ann.*, 332(1):145–159, 2005.
- [16] Ludovic Rifford. *Sub-Riemannian geometry and optimal transport*. Springer Briefs in Mathematics. Springer, Cham, 2014.
- [17] Ludovic Rifford and Piermarco Cannarsa. Semiconcavity results for optimal control problems admitting no singular minimizing controls. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, 25(4):773–802, 2008.
- [18] Shlomo Sternberg. *Lectures on Differential Geometry*. Englewood Cliffs, NJ: Prentice-Hall, 1964.
- [19] Emmanuel Trélat. Some Properties of the Value Function and Its Level Sets for Affine Control Systems with Quadratic Cost. *Journal of Dynamical and Control Systems*, 61(4):511–541, 2000.
- [20] Emmanuel Trélat. *Contrôle optimal*. Mathématiques Concrètes. [Concrete Mathematics]. Vuibert, Paris, 2005. Théorie & applications. [Theory and applications].

UNIVERSITÉ PARIS DIDEROT - PARIS 7, INSTITUT DE MATHÉMATIQUE DE JUSSIEU, UMR CNRS 7586 - UFR DE MATHÉMATIQUES.

*E-mail address:* `davide.barilari@imj-prg.fr`

CMAP, ÉCOLE POLYTECHNIQUE AND INRIA, TEAM GECO, PALAISEAU, FRANCE.

*E-mail address:* `francesco.boarotto@polytechnique.edu`