ON THE SMOOTHNESS OF THE VALUE FUNCTION FOR AFFINE OPTIMAL CONTROL PROBLEMS

DAVIDE BARILARI AND FRANCESCO BOAROTTO

Abstract. We prove that the value function associated with an affine optimal control problem with quadratic cost plus a potential always admits points of smoothness. The result is obtained by a careful analysis of points of continuity of the value function, without assuming any condition on singular minimizers.

Contents

1. Introduction 1
2. Preliminaries 3
3. Tame points 8
4. Smooth points 17
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 nonsmooth analysis. It is indeed well-known that the value function associated with an optimal control problem fails to be everywhere differentiable, in general. This is typically the case at those points where the uniqueness of minimizers is not guaranteed. Actually, the value function associated with affine optimal control problem 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. Inspired by the results and the arguments of [1, 15], we prove that in such cases the value function always admits smoothness points. Indeed, we show that the value function is smooth on an open dense subset of the interior of the attainable set, without assuming any condition on singular minimizers. Let us introduce briefly the setting and the main results.

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:

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

Date: October 23, 2016.

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* $\Omega^T_{x_0}$ 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 $u \in \Omega^T_{x_0}$ we say that $x_u(\cdot)$ is an *admissible trajectory*. By classical results of ODE theory, the set $\Omega^T_{x_0}$ is open.

(ii) the *attainable set* $A^T_{x_0}$ (in time $T > 0$, from the point x_0) as the set of points of M that can be reached from x_0 by admissible trajectories in time T, i.e.,

$$A^T_{x_0} = \{x_u(T) \mid u \in \Omega^T_{x_0}\}.$$

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

$$C_T : \Omega^T_{x_0} \to \mathbb{R}, \qquad C_T(u) = \frac{1}{2}\int_0^T \left(\sum_{i=1}^k u_i(t)^2 - Q(x_u(t))\right) dt. \tag{1.2}$$

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

$$S^T_{x_0}(x) = \inf\left\{C_T(u) \mid u \in \Omega^T_{x_0},\, x_u(T) = x\right\}; \tag{1.3}$$

with the understanding that $S^T_{x_0}(x) = +\infty$ 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 will assume the following hypotheses:

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

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

where $(\mathrm{ad}\, X)Y = [X, Y]$ and $\mathrm{Lie}_x\mathcal{F} \subset T_xM$ denotes the values at the point x of vector fields that belong to the Lie algebra generated by a family of vector fields $\mathcal{F}$.

(H2) For every bounded family $\mathcal{U} \subset \Omega^T_{x_0}$ of admissible controls, there exists a compact subset $K_T \subset M$ (depending on $\mathcal{U}$) 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 nonempty interior, i.e., $\mathrm{int}\left(A^T_{x_0}\right) \neq \emptyset$ (cf. [11, 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.

The main result of this paper reads as follows.

Theorem 1. *Fix $x_0 \in M$ and let $S^T_{x_0}$ be the value function associated with an optimal control problem of the form* (1.1)-(1.2) *satysfying assumptions* (H1)-(H3)*. Then $S^T_{x_0}$ is smooth on a non-empty open and dense subset of* $\mathrm{int}\left(A^T_{x_0}\right)$.

In the paper [1] it is proved the analogous result for the value functions associated with sub-Riemannian optimal control problems, i.e., driftless systems with zero potential. 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 $X_1, \dots, X_d$.

A key point in our proof of Theorem 1 is the characterization of points where the value function is continuous. It turns out that the continuity of the value function at a point s is strictly related with the openness of the end-point map on the optimal control steering to x. Here by end-point map we mean the map that with every control u associates the final point of the corresponding trajectory (cf. Section 2 for precise definitions).

In the sub-Riemannian case (more precisely, when $X_0 = 0$) the end-point map is open at every point (even if it is not a submersion in presence of singular minimizers) and this implies that the sub-Riemannian distance is continuous everywhere. This is no more true in presence of a drift field and the characterization of the set of points where the end-point is open is more delicate.

In the affine case, in [17, 19] the authors assumes that there are no abnormal optimal controls, which gives openness of the end-point at 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 abnormals controls exists. For more details on Goh abnormals we refer the reader to [5, Chapter 20] (see also [3, 16]). Let us mention that in [8] the authors proved 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 [13] the author proves the continuity (and other Hölder properties) of the L^1 cost under a strong bracket generating assumption.

In our setting, we admit the presence of abnormal optimal minimizers of any kind, without any further assumption on their nature. We introduce the set $\Sigma_t \subset \operatorname{int}\left(A^T_{x_0}\right)$ of points x such that the end-point map is open *and* a submersion on every optimal control steering to x (called *tame points* in what follows). The core of the proof of Theorem 1, which leads to a non trivial modification of the argument in the driftless case, is that we can find a large set of tame points.

Theorem 2. *The set Σ_t of tame points is open and dense in* $\operatorname{int}\left(A^T_{x_0}\right)$. *Moreover $S^T_{x_0}$ is continuous on Σ_t.*

Let us mention that, even in the sub-Riemannian case, it is an open question whether the set of smoothness points of the value function has full measure in $\operatorname{int}\left(A^T_{x_0}\right)$ or not.

For different approaches investigating the regularity of the value function through techniques of nonsmooth analysis, one can see for instance the monographs [6, 9, 7, 10].

1.2. **Structure of the paper.** In Section 2 we recall some properties of the end-point map, we prove the existence of minimizers in our setting and recall their characterization in terms of an Hamiltonian equation. Section 3 is devoted to the study of tame points and the proof of Theorem 2. In the last Section 4 we complete the proof of Theorem 1.

2. Preliminaries

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

$$P^u_{0,t} : U_{x_0} \subset M \to M, \qquad t \in [0, T],$$

defined on some neighborhood U_{x_0} of x_0 by $P^u_{0,t}(y) = x_{u,y}(t)$, 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 Lipschitz with respect to t. Similarly, given $u \in \Omega^T_{x_0}$ it is possible to define the family of diffeomorphism flow $P^u_{s,t} : U_{x_0} \to M$ by solving (1.1) with initial condition $x_{u,y}(s) = y$; notice then that $P^u_{t,t} = \mathrm{Id}$, and that the following composition formulas hold true (at those points where

all terms are defined):

$$P^u_{s,t} \circ P^u_{r,s} = P^u_{r,t} \qquad \text{and} \qquad \left(P^u_{s,t}\right)^{-1} = P^u_{t,s}.$$

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

$$E^T_{x_0} : \Omega^T_{x_0} \to M, \quad E^T_{x_0}(u) = x_u(T),$$

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

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

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

$$d_u E^T_{x_0}(v) = \int_0^T \sum_{i=1}^d v_i(s) \left(P^u_{s,T}\right)_* X_i(x_u(s)) ds. \tag{2.1}$$

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 $x_u(\cdot) = \lim_{n\to\infty} x_{u_n}(\cdot)$ (in the uniform topology) is its associated trajectory (see for example [20]).

This proves that the end-point map $E^T_{x_0}$ is *weakly continuous*. Indeed one can prove that the same holds true for its differential $d_u E^T_{x_0}$. 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 that

$$\lim_{n\to\infty} E^T_{x_0}(u_n) = E^T_{x_0}(u) \quad \text{and} \quad \lim_{n\to\infty} d_{u_n} E^T_{x_0} = d_u E^T_{x_0},$$

where 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 $A^T_{x_0}$ 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 $A^T_{x_0} \subset M$ can be proper, that is the end-point map $E^T_{x_0}$ may not be surjective on M; nevertheless, the weak Hörmander condition (1.4) implies that for every initial point x_0 one has $\operatorname{int}\left(A^T_{x_0}\right) \neq \emptyset$ [11, Ch. 3, Thm. 3].

2.2. **Value function and optimal trajectories.** If we introduce the Tonelli Lagrangian

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

then the *cost* $C_T : \Omega^T_{x_0} \to \mathbb{R}$ 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

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

that is obtained by writing $x_u(t) = E^t_{x_0}(u)$ 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 $S^T_{x_0} : M \to \mathbb{R}$, we write in fact

$$S^T_{x_0}(x) = \inf\left\{C_T(u) \mid u \in \Omega^T_{x_0} \cap \left(E^T_{x_0}\right)^{-1}(x)\right\}; \tag{2.2}$$

moreover, we set $S^T_{x_0}(x) = +\infty$ if the preimage $\left(E^T_{x_0}\right)^{-1}(x)$ 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 aways concentrate on the case that the final point x of an admissible trajectory belongs to the interior of the attainable set $A^T_{x_0}$. Indeed, it is a general fact that $\operatorname{int}\left(A^T_{x_0}\right)$ is densely contained in $A^T_{x_0}$ [5, 11], and the weak Hörmander condition ensures that $\operatorname{int}\left(A^T_{x_0}\right)$ is nonempty; moreover, for every point $x \in \operatorname{int}\left(A^T_{x_0}\right)$, we trivially have that $S^T_{x_0}(x) < +\infty$, 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) is classical. For the reader's convenience we give a proof below.

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

$$E^T_{x_0}(u) = x, \quad \textit{and} \quad C_T(u) = S^T_{x_0}(x).$$

Proof. Let $c > 0$ be any real constant satisfying $Q \le c$ (the existence of c follows from (H3)). Then, for every admissible control $v \in \Omega^T_{x_0}$ such that $E^T_{x_0}(v) = x$, we have:

$$C_T(v) = \frac{1}{2}\int_0^T \left(\sum_{i=1}^d v_i(t)^2 - Q(x_v(t))\right) dt \ge -\frac{cT}{2};$$

therefore, taking the infimum on both sides, we deduce the lower bound $S^T_{x_0}(x) \ge -cT/2 > -\infty$.

Consider a minimizing sequence $\{u_n\}_{n\in\mathbb{N}}$ such that $\liminf_{n\to\infty} C_T(u_n) = S^T_{x_0}(x)$; for n large enough, the relation $C_T(u_n) \le S^T_{x_0}(x) + 1$ holds, and yields that the sequence of L^2-norms $\{\|u_n\|_{L^2}\}_{n\in\mathbb{N}}$ is uniformly bounded:

$$\|u_n\|^2_{L^2} = \int_0^T \sum_{i=1}^d u_{i,n}(t)^2 dt \le 2(S^T_{x_0}(x) + 1) + cT.$$

Since balls are weakly compact for the L^2-topology, there exist a subsequence $\{u_{n_k}\}_{k\in\mathbb{N}} \subset \{u_n\}_{n\in\mathbb{N}}$ and $u \in \Omega^T_{x_0}$ such that $u_{n_k} \rightharpoonup u$, and then we call $x_u(\cdot)$ the admissible trajectory associated with u.

The weak continuity of the end-point map $E^T_{x_0}$ ensures that $x = \lim_{k\to\infty} E^T_{x_0}(u_{n_k}) = E^T_{x_0}(u)$ and moreover, since $x_{u_{n_k}}(\cdot) \to x_u(\cdot)$ uniformly as trajectories (cf. Section 2.2), the integral terms strongly converge:

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

Finally, since the L^2-norm is weakly lower semicontinuous, we have the inequalities:

$$\begin{aligned} C_T(u) &\leq \liminf_{k\to\infty} \frac{1}{2}\int_0^T \left(\sum_{i=1}^d u_{i,n_k}(t)^2 - Q(x_{u_{n_k}}(t))\right) dt \\ &= \liminf_{k\to\infty} C_T(u_{n_k}) = S^T_{x_0}(x) \leq C_T(u), \end{aligned}$$

which imply that $S^T_{x_0}(x) = C_T(u)$, as claimed. □

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.

Proposition 7. *The map $S^T_{x_0} : \mathrm{int}\left(A^T_{x_0}\right) \to \mathbb{R}$ is lower semicontinuous.*

Proof. Let $x \in \mathrm{int}\left(A^T_{x_0}\right)$. To prove the claim it is sufficient to show that for every sequence $x_n \to x$ there holds:

$$S^T_{x_0}(x) \leq \liminf_{n\to\infty} S^T_{x_0}(x_n).$$

If $\liminf_{n\to\infty} S^T_{x_0}(x_n) = +\infty$ the statement is trivially true, and therefore we can assume that $\liminf_{n\to\infty} S^T_{x_0}(x_n) = L < +\infty$; in particular we can extract a subsequence $\{x_{n_k}\}_{k\in\mathbb{N}} \subset \{x_n\}_{n\in\mathbb{N}}$ such that $\lim_{k\to\infty} S^T_{x_0}(x_{n_k}) = L$.

Let $c > 0$ be an upper bound for the map Q. If we denote by u_{n_k} an optimal control associated with x_{n_k}, then for k large enough $S^T_{x_0}(x_{n_k}) \leq L + 1$ and, in turn, this yields

$$\|u_{n_k}\|^2_{L^2} \leq 2(L+1) + \int_0^T Q(x_{u_{n_k}}(t))dt \leq 2(L+1) + cT.$$

It is therefore not restrictive to assume, up to subsequences, that $u_{n_k} \rightharpoonup u$ weakly in L^2. As in Proposition 6, the weak convergence $u_{n_k} \rightharpoonup u$ implies both that $x = \lim_{k\to\infty} E^T_{x_0}(u_{n_k}) = E^T_{x_0}(u)$ and also that $x_{u_{n_k}}(\cdot) \to x_u(\cdot)$ uniformly in the space of trajectories. Since the L^2-norm is weakly sequentially lower semicontinuous, we can write:

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

which completes the proof. □

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, 12]).

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

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

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

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 $d_u E^T_{x_0}$ 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 $\pi : T^*M \to M$ 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 $v \in T_x M$. In canonical coordinates (p,x) on the cotangent space, we can express the Liouville form $s = \sum_{i=1}^m p_i dx_i$ and the standard symplectic form $\sigma = ds = \sum_{i=1}^m dp_i \wedge dx_i$. We denote by $\overrightarrow{h}$ the Hamiltonian vector field associated with a smooth function $h : T^*M \to \mathbb{R}$, defined by the identity:

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

Consider the (control-dependent) Hamiltonian $\mathcal{H} : \mathbb{R}^d \times T^*M \to \mathbb{R}$:

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

associated with the system (1.1). The Pontryagin Maximum Principle [12] tells us that normal optimal trajectories are integral curves of the *maximized* Hamiltonian

$$H(p,x) = \max_{u\in\mathbb{R}^d} \mathcal{H}(u,p,x);$$

in particular, since $\mathcal{H}$ is smooth, they satisfy the relation $\partial_u \mathcal{H}(u,p,x) = 0$ so that the optimal control u can be recovered using the equality

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

Notice that if $p' \in T^*_x M$ can be written as $p' = p + \xi$, where ξ satisfies $\langle \xi, X_i(x)\rangle = 0$ for $i = 1,\dots,d$, then we have $u_i = \langle p, X_i(x)\rangle = \langle p', X_i(x)\rangle$ for all the indexes i. The maximized Hamiltonian takes the expression:

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

and we see that, along normal trajectories, the relation $u_i(t) = \langle p(t), X_i(x(t))\rangle$ holds. Since the pair $\dot{x}(t) = \frac{\partial H}{\partial p}(p(t),x(t))$ and $\dot{p}(t) = -\frac{\partial H}{\partial x}(p(t),x(t))$ is smooth, being the solution to a smooth autonomous system of differential equations, we deduce that the normal control $u(t) = \langle p(t), X_i(x(t))\rangle$ is smooth as well.

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

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

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

The exponential map parametrizes normal trajectories. 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]). We end this section by defining the notion of *conjugate points* along a normal trajectory.

Definition 10. We say that a point $x = \mathcal{E}_{x_0}(\lambda, s)$ is conjugate to x_0 along the normal trajectory $t \mapsto \mathcal{E}_{x_0}(\lambda, t)$ if (s, λ) is a critical point of $\mathcal{E}_{x_0}$.

3. Tame points

In this section we study fine properties of the value function on different subsets of $\operatorname{int}\left(A^T_{x_0}\right)$.

3.1. Fair points.

We start by introducing the set of fair points.

Definition 11. A point $x \in \operatorname{int}\left(A^T_{x_0}\right)$ 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 Σ_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 $S^T_{x_0}$ permits to find a great abundance of fair points; their existence is related to the notion of proximal subdifferential (see for instance [9, 15] for more details).

Definition 12. Let $F : \operatorname{int}\left(A^T_{x_0}\right) \to \mathbb{R}$ be a lower semicontinuous function. For every $x \in \operatorname{int}\left(A^T_{x_0}\right)$ we call the *proximal subdifferential* at x the subset of T^*_xM defined by:

$$\partial_P F(x) = \left\{\lambda = d_x\phi \in T^*_xM \mid \phi \in C^\infty(\operatorname{int}\left(A^T_{x_0}\right)) \text{ and } F - \phi \text{ attains a local minimum at } x\right\}.$$

The proximal subdifferential is a convex subset of T^*_xM which is nonempty for a lower semicontinuous function [9, Theorem 3.1].

Proposition 13. *Let $F : \operatorname{int}\left(A^T_{x_0}\right) \to \mathbb{R}$ be a lower semicontinuous function. Then the proximal subdifferential $\partial_P F(x)$ is not empty for a dense set of points $x \in \operatorname{int}\left(A^T_{x_0}\right)$.*

We proved in Proposition 7 that the value function $S^T_{x_0} : \operatorname{int}\left(A^T_{x_0}\right) \to \mathbb{R}$ is lower semicontinuous. Then the proximal subdifferential machinery yields the following result (cf. also [15, 1]):

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

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

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

$$\psi : \operatorname{int}\left(A^T_{x_0}\right) \to \mathbb{R}, \quad \psi(y) = S^T_{x_0}(y) - \phi(y)$$

has a local minimum at x, i.e. there exists an open neighborhood $O \subset \operatorname{int}\left(A^T_{x_0}\right)$ 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:

$$\Phi : \Omega^T_{x_0} \to \mathbb{R}, \quad \Phi(v) = C_T(v) - \phi(E^T_{x_0}(v)).$$

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

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

Then

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

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

In particular, since any two normal extremal lifts having λ 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. □

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

Corollary 15 (Density of fair points)**.** *The set Σ_f of fair points is* dense *in* $\operatorname{int}\left(A^T_{x_0}\right)$.

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

Corollary 16. *Suppose that $S^T_{x_0}$ is differentiable at some point $x \in \operatorname{int}\left(A^T_{x_0}\right)$. Then x is a fair point, and its normal covector is $\lambda = d_x S^T_{x_0} \in T^*_x M$.*

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

$$v \mapsto C_T(v) - S^T_{x_0}(E^T_{x_0}(v)),$$

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

$$0 = d_u C_T - \left(d_x S^T_{x_0}\right) d_u E^T_{x_0},$$

and the uniqueness of u (hence the claim) follows as in the previous proof. □

3.2. **Points of continuity.** We are also interested in the subset Σ_c of the *points of continuity* for the value function. The first result we give concerns their existence: again the lower semicontinuity of $S^T_{x_0}$ permits to find plenty of them.

Proposition 17. *The set Σ_c is a residual subset of* $\operatorname{int}\left(A^T_{x_0}\right)$.

Proof. This is a general fact; however, since the proof for real-valued functions is easy, we report it here for the sake of completeness.

We will show that the complement of Σ_c is a meager set, i.e. it can be written as a countable union of closed, nowhere dense subsets of $\operatorname{int}\left(A^T_{x_0}\right)$. Then the claim will follow from the classical Baire category theorem.

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

$$S^T_{x_0}(x) + \varepsilon \leq S^T_{x_0}(x_n).$$

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

$$K_q = \left\{x \in \operatorname{int}\left(A^T_{x_0}\right) \mid S^T_{x_0}(x) \leq q\right\};$$

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

$$\operatorname{int}\left(A^T_{x_0}\right) \setminus \Sigma_c \subset \bigcup_{r\in\mathbb{Q}} \left(K_r \setminus \operatorname{int}(K_r)\right),$$

as desired. □

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 $x \in \operatorname{int}\left(A^T_{x_0}\right)$ be a continuity point of $S^T_{x_0}$. Let $\{x_n\}_{n\in\mathbb{N}} \subset \operatorname{int}\left(A^T_{x_0}\right)$ 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 $\{x_n\}_{n\in\mathbb{N}} \subset \operatorname{int}\left(A^T_{x_0}\right)$ be a sequence converging to x and let $\{u_n\}_{n\in\mathbb{N}}$ 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

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

Then we have

$$\begin{aligned}
\frac{1}{2}\|u\|^2_{L^2} - \frac{1}{2}\int_0^T Q(x_u(t))dt &\le \liminf_{k\to\infty} \frac{1}{2}\|u_{n_k}\|^2_{L^2} - \frac{1}{2}\int_0^T Q(x_{u_{n_k}}(t))dt \\
&\le \limsup_{k\to\infty} \frac{1}{2}\|u_{n_k}\|^2_{L^2} - \frac{1}{2}\int_0^T Q(x_{u_{n_k}}(t))dt \\
&= \lim_{k\to\infty} S^T_{x_0}(E^T_{x_0}(u_{n_k})) = \lim_{k\to\infty} S^T_{x_0}(x_{n_k}) \\
&= S^T_{x_0}(x) = S^T_{x_0}(E^T_{x_0}(u)) \\
&\le \frac{1}{2}\|u\|^2_{L^2} - \frac{1}{2}\int_0^T Q(x_u(t))dt,
\end{aligned}$$

which readily means both that $\lim_{k\to\infty} \|u_{n_k}\|_{L^2} = \|u\|_{L^2}$ (from which the convergence in L^2 follows), and that $C_T(u) = S^T_{x_0}(E^T_{x_0}(u)) = S^T_{x_0}(x)$. □

3.3. **From regularity to continuity.** We have introduced so far two subsets of $\operatorname{int}\left(A^T_{x_0}\right)$, namely the sets Σ_c of the continuity points of $S^T_{x_0}$, and the set Σ_f of fair points, which are well-parametrized by the exponential map. Both are dense in $\operatorname{int}\left(A^T_{x_0}\right)$, still their intersection can be empty. To see that this is not actually the case, we need to introduce one more class of points.

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

$$\operatorname{rank} d_u E^T_{x_0} = \dim M = m.$$

We call Σ_t the set of tame points.

The importance of tame points resides in that they locate open sets on which the value function $S^T_{x_0}$ is continuous. This is crucial in the affine case where, as opposed to driftless one, one does not have the Chow-Rashevskii [14, 3] theorem. The precise statement is contained in the following (cf. also the arguments of [19, Theorem 4.6]):

Lemma 20. *Let $x \in \operatorname{int}\left(A^T_{x_0}\right)$ be a tame point. Then*

(i) *x is a point of continuity of $S^T_{x_0}$;*
(ii) *there exists a neighborhood O_x of x such that every $y \in O_x$ is a tame point. In particular, the restriction $S^T_{x_0}\big|_{O_x}$ 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 $\lim_{n\to+\infty} S^T_{x_0}(x_n) = S^T_{x_0}(x)$; in particular we will prove the latter equality by showing that $S^T_{x_0}(x)$ is the unique cluster point for all such sequences $\{S^T_{x_0}(x_n)\}_{n\in\mathbb{N}}$.

Let u be any optimal control steering x_0 to x; by hypothesis $d_u E^T_{x_0}$ is surjective, and therefore $E^T_{x_0}$ is locally open at u, which means that there exists a neighborhood $\mathcal{V}_u \subset \Omega^T_{x_0}$ of u such that the image $E^T_{x_0}(\mathcal{V}_u)$ covers a full neighborhood of x in $\operatorname{int}\left(A^T_{x_0}\right)$. 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 $\{S^T_{x_0}(x_n)\}_{n\in\mathbb{N}}$. Then, it is not restrictive to assume that $\lim_{n\to\infty} S^T_{x_0}(x_n) = a$. 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

$$S^T_{x_0}(x) \le C_T(u) \le \liminf_{k\to\infty} C_T(u_{n_k}) = \liminf_{k\to\infty} S^T_{x_0}(x_{n_k}) = a.$$

Let us assume by contradiction that $S^T_{x_0}(x) = b < a$, 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 end-point map $E^T_{x_0}$ is open in a (strong) neighborhood $\mathcal{V}_v \subset \Omega^T_{x_0}$ 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 $C_T(w) \le b + \varepsilon < a$. But this gives a contradiction since $S^T_{x_0}(x_{n_k})$ 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 λ_n such that:

$$\lambda_n d_{u_n} E^T_{x_0} = 0. \tag{3.1}$$

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 $\overline{\lambda} = \lim_{k\to\infty} \lambda_{n_k}$ exists. Thus, passing to the limit as k tend to infinity in (3.1), 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 $S^T_{x_0}\big|_{O_x}$ is indeed a continuous map. □

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

3.4. **Density of tame points.** We start with the somewhat trivial observation that the set of optimal controls reaching a fixed point x is compact in the L^2-topology.

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

$$\mathcal{U}_x = \left\{u \in \Omega^T_{x_0} \mid u \text{ is an optimal control steering } x_0 \text{ to } x\right\}$$

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 $S^T_{x_0}(x) = C_T(u_n)$ for every $n \in \mathbb{N}$, and consequently there exists $C > 0$ such that $\|u_n\|_{L^2} \le C$ 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

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

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

We introduce now the notion of *class* of a point. Heuristically, the class of a point $x \in \operatorname{int}\left(A^T_{x_0}\right)$ measures how much that point "fails" to be tame (see Definition 19).

Definition 23. Let $x \in A^T_{x_0}$. We define

$$\operatorname{class}(x) = \min_{u\in\mathcal{U}_x} \operatorname{rank} d_u E^T_{x_0}.$$

Any point $x \in \operatorname{int}\left(A^T_{x_0}\right)$ satisfying $\operatorname{class}(x) = m$ is necessarily a tame point.

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

$$\mathcal{U}^{\min}_x = \left\{u \in \mathcal{U}_x \mid \operatorname{rank} d_u E^T_{x_0} = \operatorname{class}(x)\right\}.$$

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

We can now state and prove the main result of this section.

Theorem 25. *The set* Σ_t *of tame points is dense in* $\operatorname{int}\left(A^T_{x_0}\right)$.

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 $\operatorname{int}\left(A^T_{x_0}\right)$ and let $u \in \mathcal{U}^{\min}_x$. If u is not strictly abnormal, then we choose a normal covector $\eta_x \in T^*_x M$ associated to u and we define

$$\widehat{\Xi}^u_x = \left\{\xi \in T^*_x M \mid \xi d_u E^T_{x_0} = \eta_x d_u E^T_{x_0}\right\} = \eta_x + \ker\left(d_u E^T_{x_0}\right)^* \subset T^*_x M.$$

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

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

$$T^*_x M = \ker\left(d_u E^T_{x_0}\right)^* \oplus \widehat{Z}_u; \tag{3.2}$$

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

$$\ker(\pi_{\widehat{Z}_u}) = \ker\left(d_u E^T_{x_0}\right)^*.$$

Finally, by means of the adjoint map $\left(P^u_{0,T}\right)^*$, we can pull the spaces $\widehat{\Xi}^u_x$ "back" to $T_{x_0}M$, and set

$$\Xi^u_x := \left(P^u_{0,T}\right)^* \widehat{\Xi}^u_x \subset T^*_{x_0}M.$$

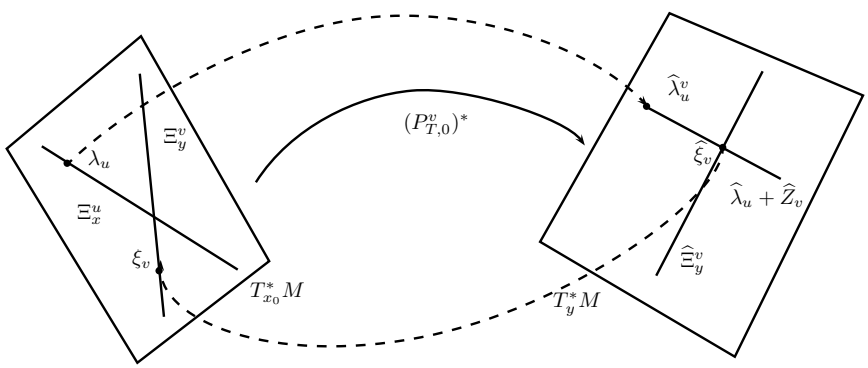

FIGURE 1. We set $y = E^T_{x_0}(v)$. The subspace $\widehat{\Xi}^v_y$ is linear if v is strictly abnormal, and affine otherwise; $\widehat{Z}_v$ and $\ker\left(d_vE^T_{x_0}\right)^*$ are orthogonal. The point $\widehat{\xi}_v$ belong to T^*_yM, and is then pulled back on $T^*_{x_0}M$.

The following estimate will be crucial in what follows.

Proposition 27. *Let $O \subset \operatorname{int}\left(A^T_{x_0}\right)$ be an open set, and assume that:*

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

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

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

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

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

Fix $\lambda_u \in \Xi^u_x$, and set

$$\widehat{\lambda}^v_u = (P^v_{T,0})^*\lambda_u \in T^*_yM, \quad v \in \mathcal{V}_u \cap \mathcal{U}^{\min}_y, \quad y = E^T_{x_0}(v).$$

The intersection $(\widehat{\lambda}^v_u + \widehat{Z}_v) \cap \widehat{\Xi}^v_y$ (cf. with (3.2) and Figure 1) consists of the single point $\widehat{\xi}_v$; since both $\widehat{\lambda}^v_u$ and $\widehat{\xi}_v$ belong to the affine subspace $\widehat{\lambda}^v_u + \widehat{Z}_v$, in order to estimate the norm $|\widehat{\lambda}^v_u - \widehat{\xi}_v|$ it is sufficient to evaluate the norm $|\pi_{\widehat{Z}_v}(\widehat{\lambda}^v_u) - \pi_{\widehat{Z}_v}(\widehat{\xi}_v)|$ of the projections onto the linear space $\widehat{Z}_v = (\ker(d_vE^T_{x_0})^*)^\perp$. The key point is the computation of the norm of $|\pi_{\widehat{Z}_v}(\widehat{\xi}_v)|$: in fact, since

$\ker(d_v E^T_{x_0})^* = (\operatorname{Im} d_v E^T_{x_0})^\perp$, this amounts to evaluate

$$|\pi_{\widehat{Z}_v}(\widehat{\xi}_v)| = \sup_{f \in \operatorname{Im} d_v E^T_{x_0}} \frac{|\langle \widehat{\xi}_v, f\rangle|}{|f|}. \tag{3.3}$$

We deduce immediately from (3.3) 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)

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

we see that $\langle v, w\rangle_{L^2} = \langle \widehat{\xi}_v, d_v E^T_{x_0}(w)\rangle$, and we can continue from (3.3) as follows (W_v denotes the k_O-dimensional subspace of $L^2([0,T],\mathbb{R}^d)$ on which the restriction $d_v E^T_{x_0}\big|_{W_v}$ is invertible):

$$\begin{aligned} |\pi_{\widehat{Z}_v}(\widehat{\xi}_v)| &= \sup_{w \in W_v} \frac{|\langle \widehat{\xi}_v, d_v E^T_{x_0}(w)\rangle|}{|d_v E^T_{x_0}(w)|} \\ &\le \sup_{w \in W_v} \frac{|\langle \widehat{\xi}_v, d_v E^T_{x_0}(w)\rangle|}{\|w\|_{L^2}} \|(d_v E^T_{x_0}\big|_{W_v})^{-1}\| \\ &= \sup_{w \in W_v} \frac{|\langle v, w\rangle|}{\|w\|_{L^2}} \|(d_v E^T_{x_0}\big|_{W_v})^{-1}\| \\ &\le \|v\|_{L^2} \|(d_v E^T_{x_0}\big|_{W_v})^{-1}\|. \end{aligned} \tag{3.4}$$

It is not restrictive to assume that the L^2-norm of any element $v \in \mathcal{V}_u \cap \mathcal{U}^{\min}_y$ 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 $v \mapsto (d_v E^T_{x_0}\big|_{W_v})^{-1}$. This, on the other hand, guarantees that the operator norm $\|(d_v E^T_{x_0}\big|_{W_v})^{-1}\|$ remains bounded for all $v \in \mathcal{V}_u \cap \mathcal{U}^{\min}_y$, and then from (3.4) we conclude that for some $C > 1$, the estimate $|\pi_{\widehat{Z}_v}(\widehat{\xi}_v)| \le C$ holds true, which implies as well, by the triangular inequality, that:

$$|\widehat{\lambda}^v_u - \widehat{\xi}_v| \le |\widehat{\lambda}^v_u| + C.$$

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

$$\sup_{v \in \mathcal{V}_u} \left\{ \|(P^v_{0,T})^*\|, \|(P^v_{T,0})^*\| \right\} \le C.$$

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

$$\begin{aligned} |\lambda_u - \xi_v| &\le C|\widehat{\lambda}^v_u - \widehat{\xi}_v| \\ &\le C|\widehat{\lambda}^v_u| + C^2 \\ &\le C^2 \left(|\lambda_u| + 1\right) \\ &\le 2C^2 \max\{|\lambda_u|, 1\}. \end{aligned}$$

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

Remark. Let us fix $\lambda_u \in \Xi^u_x \subset T^*_{x_0} M$ and consider the k_O-dimensional affine subspace

$$(P^v_{0,T})^*(\widehat{\lambda}^v_u + \widehat{Z}_v) = \lambda_u + (P^v_{0,T})^* \widehat{Z}_v,$$

with $\widehat{Z}_v$ defined as in (3.2). Then if we call $Z_v := (P^v_{0,T})^* \widehat{Z}_v \subset T^*_{x_0}M$, the map

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

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

Having in mind this remark, we deduce the following:

Corollary 28. *Let $O \subset \operatorname{int}\left(A^T_{x_0}\right)$ be an open set, and assume that*

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

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

$$A_u \cap \Xi^v_y \neq \emptyset \quad \textit{for every } v \in \mathcal{V}_u \cap \mathcal{U}^{\min}_y, \quad \textit{where } y = E^T_{x_0}(v).$$

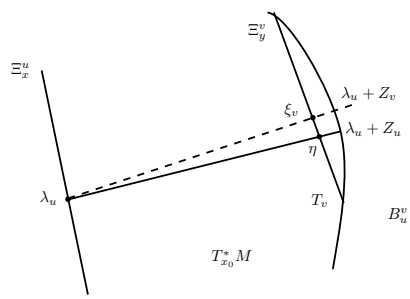

FIGURE 2. On the fiber $T^*_{x_0}M$, the point η denotes the intersection between T_v and the affine space $\lambda_u + Z_u$.

Proof. Let $\lambda_u \in \Xi^u_x$ 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^v_u, of radius C^v_0 strictly greater than $K = K(\lambda_u)$ (given by Proposition 27), and centered at λ_u.

Then, the existence of an element $\xi_v \in (\lambda_u + Z_v) \cap \Xi^v_y$ satisfying $|\lambda_u - \xi_v| \leq K$, proved in Proposition 27, implies that the intersection of B^v_u with Ξ^v_y is a compact submanifold T_v (with boundary); moreover, since the radius of B^v_u is strictly greater than $|\lambda_u - \xi_v|$, it is also true that the intersection of $\lambda_u + Z_v$ with $\operatorname{int}(T_v)$ 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 Ξ^u_x: possibly increasing the radius C^v_0, 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^v_0 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 $C^{v_1}_0, \dots, C^{v_l}_0$.

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 η_v any element in the intersection $T_v \cap (\lambda_u + Z_u)$, for $v \in \mathcal{V}_u$, then it follows that:

$$\eta \in \Xi^v_y \cap B_u \cap (\lambda_u + Z_u) = \Xi^v_y \cap A_u,$$

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

Lemma 29. *Let $O \subset \operatorname{int}\left(A^T_{x_0}\right)$ be an open set, and let*

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

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

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

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:

$$\operatorname{class}(x) \leq \operatorname{rank} d_u E^T_{x_0} \leq \liminf_{k\to\infty} \operatorname{rank} d_{u_{n_k}} E^T_{x_0} = \liminf_{k\to\infty} \operatorname{class}(x_{n_k}) \leq k_O - 1,$$

which is absurd by construction, and the claim follows. □

Collecting all the results we can now prove Theorem 25.

Proof of Theorem 25. Let O be an open set in $\operatorname{int}\left(A^T_{x_0}\right)$ and define

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

notice that this definition makes sense, since points of continuity are dense in $\operatorname{int}\left(A^T_{x_0}\right)$ 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 $\operatorname{class}(y) = k_O$ 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 $u \in \mathcal{U}^{\min}_x$ we can find a neighborhood $\mathcal{V}_u \subset \Omega^T_{x_0}$ of u, fix $\lambda_u \in \Xi^u_x$, and construct accordingly a compact k_O-dimensional ball A_u, centered at λ_u and transversal to Ξ^u_x, such that (Corollary 28)

$$A_u \cap \Xi^v_y \neq \emptyset \quad \text{for every } v \in \mathcal{V}_u, \text{ and with } y = E^T_{x_0}(v).$$

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

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

The union $A_{u_1} \cup \ldots \cup A_{u_l}$ 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 $u_n \in \mathcal{U}^{\min}_{x_n}$) the sequence u_n converges to some $v \in \mathcal{V}_{u_i} \subset \mathcal{U}^{\min}_x$, we have that A_{u_i} is also transversal to $\Xi^{u_n}_{x_n}$. In particular, possibly enlarging the ball A_{u_i}, we can assume that

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

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

$$\mathcal{E}^T_{x_0}(\Xi^u_z) = z,$$

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

$$\Sigma_f \cap O \subset \mathcal{E}^T_{x_0}\left(A_{u_1} \cup \ldots \cup A_{u_l}\right). \tag{3.5}$$

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 (3.5) we conclude that $\mathrm{meas}(O) = 0$, which is impossible. □

Combining now Lemma 20 and Theorem 25 we obtain the following (cf. Theorem 2).

Corollary 30. *The set Σ_t of tame points is open and dense in* $\mathrm{int}\left(A^T_{x_0}\right)$.

4. Smooth points

In this section we investigate the regularity properties of the value function $S^T_{x_0}$ in the presence of tame points. Since tame points are in particular points of continuity for $S^T_{x_0}$, 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*

$$\mathcal{M}_K = \left\{u \in \Omega^T_{x_0} \mid E^T_{x_0}(u) \in K \text{ and } C_T(u) = S^T_{x_0}(E^T_{x_0}(u))\right\}$$

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 $S^T_{x_0}$ is Lipschitz continuous on K.*

Proof. By compactness, it is sufficient to show that $S^T_{x_0}$ 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, $d_u E^T_{x_0}$ is surjective, so that there are neighborhoods $\mathcal{V}_u \subset \Omega^T_{x_0}$ of u and $O_x \subset \mathrm{int}\left(A^T_{x_0}\right)$ of x such that

$$E^T_{x_0}\big|_{\mathcal{V}_u} : \mathcal{V}_u \to O_x$$

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

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

$$B_x(C_0 r) \subset E^T_{x_0}(\mathcal{B}_u(r)), \quad \text{for every } 0 \le r \le R. \tag{4.1}$$

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

$$|C_T(v) - C_T(w)| \le C_1 \|v - w\|_{L^2}. \tag{4.2}$$

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

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

and by means of the triangular inequality, (4.2) is proved.

Pick any point $y \in K$ such that $|y-x| = C_0 r$, with $0 \le r \le R$. Then by (4.1) there exists $v \in \mathcal{B}_u(r)$ satisfying $\|u-v\|_{L^2} \le r$ and such that $E^T_{x_0}(v) = y$; since $C_T(u) = S^T_{x_0}(x)$ and $S^T_{x_0}(y) \le C_T(v)$, we have

$$S^T_{x_0}(y) - S^T_{x_0}(x) \le C_T(v) - C_T(u) \le C_1\|v-u\|_{L^2} \le \frac{C_1}{C_0}|y-x|.$$

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

$$|S^T_{x_0}(x) - S^T_{x_0}(y)| \le \frac{C_1}{C_0}|x-y|,$$

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

Definition 33. We define the set of $\Sigma \subset \operatorname{int}\left(A^T_{x_0}\right)$ 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 $\operatorname{int}\left(A^T_{x_0}\right)$.

The following result finally proves Theorem 1.

Theorem 34 (Density of smooth points)**.** *Σ is open and dense in* $\operatorname{int}\left(A^T_{x_0}\right)$*. Moreover $S^T_{x_0}$ is smooth on Σ.*

Proof. (i.a) Let us show that Σ is dense. First we prove that, for any open set O, we have $\Sigma \cap O \neq \emptyset$. Since the set points satisfying (a) (i.e. the set $\Sigma_f \cap \Sigma_t$) is dense in $\operatorname{int}\left(A^T_{x_0}\right)$, by Proposition 32 we can choose a point $x \in \Sigma_f \cap \Sigma_t \cap O$ and $O' \subset O \cap \Sigma_t$ relatively compact, such that $S^T_{x_0}$ is Lipschitz on O'. Thanks to the classical Rademacher theorem we know that $S^T_{x_0}$ is differentiable almost everywhere on O', and therefore, since any point of differentiability is a fair point by Corollary 16, $\operatorname{meas}(\Sigma_f \cap O') = \operatorname{meas}(O')$. Moreover, any point in $\Sigma_f \cap O'$ is also contained in the image of the exponential map $\mathcal{E}^T_{x_0}$, 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 Σ, that is we have $\operatorname{meas}(\Sigma \cap O') = \operatorname{meas}(\Sigma_f \cap O') = \operatorname{meas}(O')$, which implies that $\Sigma \cap O' \neq \emptyset$, and this concludes the proof.

(i.b) Let us prove that Σ is open. Fix as before an open set O having compact closure in $\operatorname{int}\left(A^T_{x_0}\right)$. 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 $u = \lim_{n\to\infty} u_n$ and $v = \lim_{n\to\infty} v_n$ 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 $d_{u_n}E^T_{x_0}$ and $d_{v_n}E^T_{x_0}$ 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 λ_n and ξ_n, as elements of $T^*_{x_n}M$, satisfying the identities

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

Taking the limit on these two equations we see that $\lim_{n\to\infty}\lambda_n = \lim_{n\to\infty}\xi_n = \lambda$, where λ 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" λ_n^0 and ξ_n^0 converge to the same element λ^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 λ^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 λ^0 and $O_x \subset \operatorname{int}\left(A_{x_0}^T\right)$ of x such that the exponential map $\mathcal{E}_{x_0}^T\big|_V : V \to O_x$ 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 λ_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 Σ. 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 Theorem 34, there are neighborhoods $V_\lambda \subset T_{x_0}^* M$ of λ and $O_x \subset \operatorname{int}\left(A_{x_0}^T\right)$ of x such that $\mathcal{E}_{x_0}^T\big|_{V_\lambda} : V_\lambda \to O_x$ is a diffeomorphism.

It is then possible to define a smooth inverse $\Phi : O_x \to V_\lambda$ sending y to the corresponding "initial" covector λ_y. Along (strictly normal) trajectories associated with covectors λ_y in V_λ 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. □

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 Boscain. Introduction to Riemannian and sub-Riemannian geometry, 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] 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.

[8] Y. Chitour, F. Jean, and E. Trélat. Singular trajectories of control-affine systems. *SIAM J. Control Optim.*, 48(2):1078 – 1095, 2008.

[9] Francis Clarke, Yuri Ledyaev, Ronald Jay Stern, and Peter Wolenski. *Nonsmooth Analysis and Control Theory*, volume 178. Springer-Verlag, New York, 1998.

[10] 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.

[11] Velimir Jurdjevic. *Geometric control theory*, volume 52 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 1997.

[12] 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.

[13] Dario Prandi. Hölder equivalence of the value function for control-affine systems. *ESAIM Control Optim. Calc. Var.*, 20(4):1224–1248, 2014.

[14] P. K. Rashevski. Any Two Points of a Totally Nonholonomic Space May Be Connected By an Admissible Line . *Uch. Zap. Ped Inst. im. Liebknechta*, 21:83–94, 1938.

[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 Mathematique de Jussieu, UMR CNRS 7586 - UFR de Mathématiques.
E-mail address: `davide.barilari@imj-prg.fr`

CMAP, École Polytechnique, Palaiseau, France.
E-mail address: `francesco.boarotto@polytechnique.edu`