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.

Key words. regularity, control-affine systems, value function, smoothness

AMS subject classifications. 93B27, 49J15, 49J52, 93B03

DOI. 10.1137/17M1123948

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 nonsmooth 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, 34]).

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 the 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 to 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.

[†]Institut de Mathematique de Jussieu, Université Paris Diderot - Paris 7, UMR CNRS 7586 - UFR de Mathématiques, 75013 Paris, France (davide.barilari@imj-prg.fr).

[‡]CMAP, École Polytechnique and Inria, Team GeCo, Palaiseau, France (francesco.boarotto@polytechnique.edu).

Adapting then the arguments of [1, 31], 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 driftless 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 the presence of deep singular minimizers. The absence of such a property for affine control systems makes the study of the continuity of the value function more delicate in our context.

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

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

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

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

Given $x_0 \in M$ we define the following:

(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 $Q : M \to \mathbb{R}$, we are interested in those trajectories minimizing the *cost* given by

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

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

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

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

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

where $(\mathrm{ad}\, X) Y = [X, Y]$, and $\mathrm{Lie}_x \mathcal{F} \subset T_x M$ 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 nonempty interior, i.e., $\text{int}\left(A_{x_0}^T\right) \neq \emptyset$ (cf. [33] or [26, 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 $X_0, \ldots, X_d$. We refer the reader 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 $\Sigma_t \subset \text{int}\left(A_{x_0}^T\right)$ 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 Σ_t of tame points is open and dense in* $\text{int}\left(A_{x_0}^T\right)$ *and $S_{x_0}^T$ is continuous on Σ_t.*

In the driftless 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 [10]. This is no longer 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 [2, 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 driftless 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, 31] let us derive the following result.

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

In [1], the author proves the analogue of Theorem 2 for the value function associated with sub-Riemannian optimal control problems, i.e., driftless 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 $X_1, \ldots, X_d$.

Let us further mention that, even in the sub-Riemannian situation, it still re-

mains an open question to establish whether the set of smoothness points of the value function has full measure in int $\left(A_{x_0}^T\right)$ or not.

1.2. Comparison with recent literature. Regularity of the value function for these kinds of control systems with techniques of geometric control has been also studied in [15, 34], 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. In fact, as one can see following the proof of Lemma 20 in the appendix, openness of the end-point map at every optimal control is what is needed to ensure the continuity of the value function.

For more details on Goh abnormals we refer the reader to [5, Chapter 20] (see also [2, 30]). Let us mention that in [17] the authors prove that the system (1.1) admits no Goh optimal trajectories for the generic choice of the $(d+1)$-tuple $X_0, \ldots, X_d$ (in the Whitney topology).

Finally, in [29] the author proves the Hölder continuity of the value function under a strong bracket generating assumption, when one considers the L^1 cost.

In the investigation of the regularity of the value function, techniques of nonsmooth analysis have been also extensively used. A complete overview of the vast literature on this approach being not possible in this short discussion, we refer the interested reader to the monographs [6, 18, 16, 23, 19, 36] for a general introduction, and we discuss here some results that are more closely related to those investigated in this paper.

In the paper [12] the authors consider an optimal control problem of Bolza type. Their main results are analogous to that of our Theorem 2, stating that as soon as the proximal subdifferential (cf. Definition 12) of the value function of the Bolza problem is nonempty at a point, then the value function turns out to be of class C^2 in a neighborhood of that point.

A similar result concerning a Hamilton–Jacobi equation related to the Bolza problem of the calculus of variations was already obtained in [13], assuming coercivity of the Hamiltonian under consideration. Combined with the fact that points where the proximal subdifferential is nonempty are dense in the domain of the value function (cf. Proposition 13), these results are used to derive deep regularity properties on the optimal synthesis.

Analogous techniques have also been employed to treat an optimal control problem of Mayer type [14], in which the dynamic is given by a differential inclusion. In this case the Hamiltonian is no longer coercive, but still the local C^2 (and $C^{2,m}$, for $0 < m \leq 1$) smoothness of the value function is proven at points where the proximal subdifferential is nonempty.

For a discussion on the relation between optimal control problems, Mayer-type problems for differential inclusions, and the Bolza problem of the calculus of variations, we refer the reader to [19, Chapter 1].

1.3. Further comments. As we have seen in the aforementioned results, the regularity of the value function has important consequences for the structure of optimal control problems. In fact, this is also the paradigm of this paper. The main issue is that, in general, the value function does not exhibit sufficient regularity in principle (cf. Proposition 7) and has to be analyzed using the so-called sensitivity relations. It should be mentioned that, since the seminal papers [20, 37], sensitivity analysis has proved itself to be successful for a wide range of optimal control problems [8, 9, 24, 25].

The idea behind sensitivity relations is to give necessary optimality conditions (e.g., as in Proposition 8) in the form of inclusions into suitable generalized differentials of the value function.

In this paper we exploit the nonemptiness of the proximal differential (cf. Proposition 14), only to prove that the corresponding point is reached by a normal trajectory, i.e., a trajectory that can be reconstructed from the exponential map (cf. subsection 2.3 and Definition 9). Finer regularity properties of the corresponding trajectory (such as absence of singular minimizer), and as a consequence of the value function, are deduced through direct investigation.

Let us mention that, in the geometric language, the exact characterization of the regularity of the value function at the so-called cut locus (the set of points where trajectories satisfying first-order necessary conditions lose optimality) is not trivial. This phenomenon is well understood in the Riemannian setting and is characterized by the loss of semiconvexity of $S_{x_0}^T$ [21]. An analogous property has been recently proved in the sub-Riemannian setting in [7], but only in the absence of singular minimizers (cf. also the discussion in [22]).

1.4. Structure of the paper. In section 2 we recall some properties of the end-point map and the existence of minimizers in our setting, and we 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 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 $u \in \Omega_{x_0}^T$, the family of diffeomorphisms

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

defined by $P_{0,t}^u(y) = x_{u,y}(t)$, is well defined on some neighborhood U_{x_0} of x_0 [11]. Here, by $x_{u,y}(t)$ we denote the solution to 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 $P_{s,t}^u : U_{x_0} \to M$ by solving (1.1) with initial condition $x_{u,y}(s) = y$. Notice then that $P_{t,t}^u = \mathrm{Id}$ and that the composition formulas

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

hold true (at those points where all terms are defined). Finally, the notation $\left(P_{s,t}^u\right)_*$ refers 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 $\left(P_{s,t}^u\right)_* X$ is defined by

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

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

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

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

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 [2, 30, 34].

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

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

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 $x_u(\cdot) = \lim_{n \to \infty} x_{u_n}(\cdot)$ (in the uniform topology) is its associated trajectory (see, for example, [35]).

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

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

and the last convergence is in the (strong) operator norm (see [34]).

Remark 1. 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 [15], one could ask for a sublinear growth condition on the vector fields $X_0, \ldots, X_d$. 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 set $\mathrm{int}\left(A_{x_0}^T\right)$ is densely contained in $A_{x_0}^T$ [5, 26, 33], and the weak Hörmander condition (1.4) implies that for every initial point x_0 one has $\mathrm{int}\left(A_{x_0}^T\right) \neq \emptyset$ [33] (see also [26, Ch. 3, Thm. 3]).

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

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

then the *cost* $C_T : \Omega_{x_0}^T \to \mathbb{R}$ 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

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

which 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 $S^T_{x_0} : M \to \mathbb{R}$ via the position

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

with the agreement that $S^T_{x_0}(x) = +\infty$ if the preimage $\left(E^T_{x_0} \right)^{-1}(x)$ is empty; then the *optimal control problem* consists in 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.

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

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*

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

Remark 2. The assumptions (H2)–(H3) play a crucial role for the existence of optimal controls. An equivalent approach could be to work directly inside a given compact set (see [3]) 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. subsection 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 $S^T_{x_0} : A^T_{x_0} \to \mathbb{R}$ is lower semicontinuous.*

Proofs of Proposition 6 and Proposition 7 are classical and follow from standard arguments in the literature (see, e.g., [27, 34]); hence their proofs are 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, 2, 28]).

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 holds:*
(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 to be *normal* in case (a) and *abnormal* in case (b). Notice that (b)

implies that $d_u E_{x_0}^T$ is not surjective in the abnormal case. Moreover, any covector λ_T, either normal or abnormal, is defined only modulo the subspace $\ker(d_u E_{x_0}^T)^*$. 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 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 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 $\overrightarrow{h}$ the Hamiltonian vector field associated with a smooth function $h : T^*M \to \mathbb{R}$, defined by the identity

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

The Pontryagin Maximum Principle [28, 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 $\mathcal{H} : \mathbb{R}^d \times (-\infty, 0] \times T^*M \to \mathbb{R}$, 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 nonpositive real constant ν 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) $$u_i(t) = \langle p(t), X_i(x(t)) \rangle \qquad \text{for } i = 1, \ldots, d.$$

Normal extremals are therefore solutions to the differential system

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

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* Exp *with base point* x_0 *is defined as*

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

where $\mathcal{D}_{x_0}^T$ *is the open subset of covectors in* $T_{x_0}^* M$ *such that the solution to* (2.4) *is defined up to time* T. *When the first argument is fixed, we employ the notation* $\mathrm{Exp}_{x_0}^s : \mathcal{D}_{x_0}^T \to M$ *to denote the exponential map with base point* x_0 *at time* s, *that is to say, we set* $\mathrm{Exp}_{x_0}^s(\lambda) := \mathrm{Exp}_{x_0}(s, \lambda)$.

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

DEFINITION 10. *We say that a point* $x = \mathrm{Exp}_{x_0}(s, \lambda)$, $s \in [0, T]$, *is conjugate to* x_0 *along the normal extremal* $t \mapsto \mathrm{Exp}_{x_0}(t, \lambda)$, $t \in [0, T]$, *if* (s, λ) *is a critical point of* Exp_{x_0}, *i.e., if the differential* $d_{(s,\lambda)}\mathrm{Exp}_{x_0}$ *is not surjective.*

3. On the continuity. In this section we study fine properties of the value function on $A_{x_0}^T$. Eventually, we investigate differentiability properties of $S_{x_0}^T$. It is thus natural to restrict the analysis on the nonempty open subset $\mathrm{int}\left(A_{x_0}^T\right)$ (cf. also the discussion at the end of subsection 2.1).

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

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

Equivalently, a fair point is reached by a unique optimal trajectory, and this trajectory is not strictly abnormal.

The lower semicontinuity of $S_{x_0}^T$ permits us to find a great abundance of fair points. Their existence is related to the notion of the proximal subdifferential (see, for instance, [18, 31] for more details).

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

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

The proximal subdifferential is a convex subset of $T_x^* M$ which is often nonempty in the case of a lower semicontinuous function [18, Theorem 3.1].

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

We showed in Proposition 7 that the value function $S_{x_0}^T : \mathrm{int}\left(A_{x_0}^T\right) \to \mathbb{R}$ is lower semicontinuous. By classical arguments, the proximal subdifferential machinery yields the following result (cf. also [31, 1]).

PROPOSITION 14. *Let $x \in \mathrm{int}\left(A_{x_0}^T\right)$ be such that $\partial_P S_{x_0}^T(x) \neq \emptyset$. Then there exists a unique optimal trajectory $x_u(\cdot) : [0,T] \to M$ steering x_0 to x, and this trajectory 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 λ as final covector.

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

$$\psi : \mathrm{int}\left(A_{x_0}^T\right) \to \mathbb{R}, \quad \psi(y) = S_{x_0}^T(y) - \phi(y)$$

has a local minimum at x; i.e., there exists an open neighborhood $O \subset \mathrm{int}\left(A_{x_0}^T\right)$ 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

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

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:

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

where in the second inequality we used the fact that ψ has a minimum at $x = E_{x_0}^T(u)$. 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 $\lambda(t) = e^{(t-T)\vec{H}}(\lambda)$ is the desired normal lift of the trajectory $x_u(\cdot)$.

Observe that the extremal normal lift $\lambda(t)$ is uniquely reconstructed from λ, and in fact its projection onto the manifold M does not depend on u. Then there is only one optimal trajectory between x_0 and x, admitting a normal lift, which precisely means that $x \in \Sigma_f$ is a fair point.

As a final remark, we stress that nothing prevents the optimal trajectory from admitting also an abnormal lift. In particular, when $\partial_P S_{x_0}^T(x) \neq \emptyset$, 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. $\square$

COROLLARY 15 (density of fair points). *The set Σ_f of fair points is dense in $\mathrm{int}\left(A_{x_0}^T\right)$.*

Along the same lines of Proposition 14, we show that all differentiability points of $S_{x_0}^T$ are fair points.

PROPOSITION 16. *Suppose that $S_{x_0}^T$ is differentiable at some point $x \in \mathrm{int}\left(A_{x_0}^T\right)$. 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 nonnegative 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

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

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 Σ_c of the *points of continuity* for the value function. It is a fact from general topology that a lower semicontinuity function has plenty of continuity points.

LEMMA 17. *The set Σ_c is a residual subset of* $\mathrm{int}\left(A_{x_0}^T\right)$.

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 is shown in the next lemma.

LEMMA 18. *Let $x \in \mathrm{int}\left(A_{x_0}^T\right)$ be a continuity point of $S_{x_0}^T$. Let $\{x_n\}_{n\in\mathbb{N}} \subset \mathrm{int}\left(A_{x_0}^T\right)$ 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 $\{x_n\}_{n\in\mathbb{N}} \subset \mathrm{int}\left(A_{x_0}^T\right)$ 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

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

Then we have

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

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

3.3. Tame points. We have introduced so far two subsets of $\mathrm{int}\left(A_{x_0}^T\right)$, namely the sets Σ_c of the continuity points of $S_{x_0}^T$, and the set Σ_f of fair points, which are essentially points that are well parametrized by the exponential map. While both these sets are dense in $\mathrm{int}\left(A_{x_0}^T\right)$, their intersection can still be empty. Here we have 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 Σ_c and Σ_f, we introduce the following set.

DEFINITION 19 (tame points). *Let $x \in \mathrm{int}\left(A_{x_0}^T\right)$. We say that x is a* tame point *if for every optimal control u steering x_0 to x there holds*

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

We call Σ_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 [34, Theorem 4.6]. A complete proof is contained in Appendix A.

LEMMA 20. *Let* $x \in \text{int}\left(A_{x_0}^T\right)$ *be a tame point. Then the following hold:*
 (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* $S_{x_0}^T\big|_{O_x}$ *is a continuous map.*

The previous lemma can be restated as follows.

COROLLARY 21. *The set* Σ_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*

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

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 $S_{x_0}^T(x) = C_T(u_n)$ 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

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

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

We introduce now the notion of the *class* of a point. Heuristically, the class of a point $x \in \text{int}\left(A_{x_0}^T\right)$ 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 $x \in \text{int}\left(A_{x_0}^T\right)$ 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:*

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

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

It turns out that the function class is locally constant around points of continuity in the interior of the attainable set.

LEMMA 25. *Let $O \subset \mathrm{int}\left(A_{x_0}^T\right)$ be an open set, and let*

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

Then there exists a neighborhood $O' \subset O$, such that $\mathrm{class}\,(y) = k_O$, 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 $\mathrm{class}\,(x) = k_O$. Assume by contradiction that we can find a sequence $\{x_n\}_{n \in \mathbb{N}}$ converging to x and satisfying $\mathrm{class}\,(x_n) \le k_O - 1$ for every $n \in \mathbb{N}$. Accordingly, let $u_n \in \mathcal{U}_{x_n}^{\min}$ be an associated sequence of optimal controls. In particular, for every $n \in \mathbb{N}$, we have by definition that $\mathrm{class}\,(x_n) = \mathrm{rank}\, d_{u_n} E_{x_0}^T$.

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

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

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

We can now state the main result of this section.

THEOREM 26. *The set Σ_t of tame points is dense in $\mathrm{int}\left(A_{x_0}^T\right)$.*

We postpone the proof of Theorem 26 until the end of the section, since we need first a series of preliminary results. We begin attaching to each $u \in \mathcal{U}_x^{\min}$ the set of all normal covectors λ satisfying condition (a) in Proposition 8.

DEFINITION 27. *Pick x in $\mathrm{int}\left(A_{x_0}^T\right)$ and let $u \in \mathcal{U}_x^{\min}$. If u is not strictly abnormal, then we choose any normal covector $\lambda_{T,x} \in T_x^* M$ associated to u and satisfying item* (a) *in Proposition 8, and we define*

$$\widehat{\Xi}_x^u = \left\{\lambda \in T_x^* M \mid \lambda d_u E_{x_0}^T = \lambda_{T,x} d_u E_{x_0}^T\right\} = \lambda_{T,x} + \ker\left(d_u E_{x_0}^T\right)^* \subset T_x^* M.$$

If instead u is strictly abnormal, we simply set $\widehat{\Xi}_x^u = \ker\left(d_u E_{x_0}^T\right)^* \subset T_x^* M$. 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. The dimension of these subspaces equals $m - \mathrm{class}\,(x) \ge 0$.

Fix any Riemannian metric $\mathbf{g}_{x_0}$ on $T_{x_0}^* M$. If u is an admissible control associated to a trajectory steering x_0 to x in time T, i.e., $x = P_{0,T}^u(x_0)$, unless otherwise stated we will always consider on $T_x^* M$ the Riemannian metric $\mathbf{g}_x^u$ defined as the pull-back of $\mathbf{g}_{x_0}$ by $P_{0,T}^u$; i.e., we set

(4.1) $$\mathbf{g}_x^u(\xi, \eta) := \mathbf{g}_{x_0}((P_{0,T}^u)^* \xi, (P_{0,T}^u)^* \eta) \qquad \forall \xi, \eta \in T_x^* M.$$

Observe that this metric depends continuously on the control.

We call $\widehat{Z}_u \subset T_x^* M$ the orthogonal (with respect to $\mathbf{g}_x^u$) subspace to $\ker\left(d_u E_{x_0}^T\right)^*$, of dimension equal to $\mathrm{class}\,(x)$, so that

(4.2) $$T_x^* M = \ker\left(d_u E_{x_0}^T\right)^* \oplus \widehat{Z}_u.$$

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

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

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

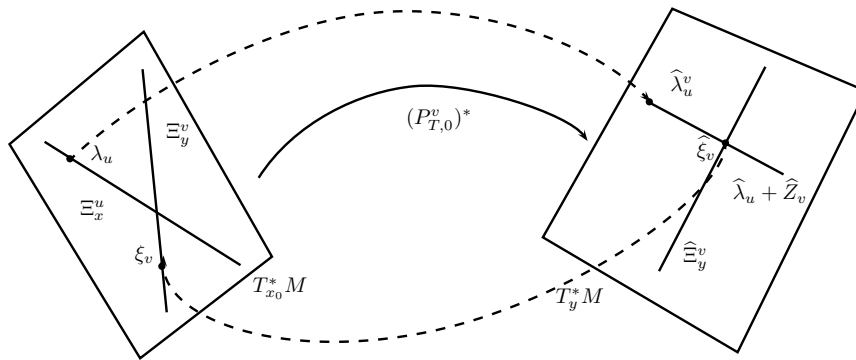$$\Xi_x^u := \left(P_{0,T}^u\right)^* \widehat{\Xi}_x^u \subset T_{x_0}^* M.$$

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

The following estimate will be crucial in what follows.

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

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

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[1]

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

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 27. Therefore, we can assume from the beginning that all such subspaces Ξ_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 $(\widehat{\lambda}_u^v + \widehat{Z}_v) \cap \widehat{\widehat{\Xi}}_y^v$ (cf. with (4.2) 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

---

[1] We omit the explicit dependence on the base point when it is clear from the context on which fiber of T^*M we are evaluating the norm.

computation of the norm of $|\pi_{\widehat{Z}_v}(\widehat{\xi}_v)|$. In fact, since $\ker(d_v E_{x_0}^T)^* = (\operatorname{Im} d_v E_{x_0}^T)^\perp$, this amounts to evaluating

$$(4.3) \qquad |\pi_{\widehat{Z}_v}(\widehat{\xi}_v)| = \sup_{f \in \operatorname{Im} d_v E_{x_0}^T} \frac{|\langle \widehat{\xi}_v, f \rangle|}{|f|}.$$

We deduce immediately from (4.3) 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)

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

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 (4.3) as follows (W_v denotes the k_O-dimensional subspace of $L^2([0,T], \mathbb{R}^d)$ on which the restriction $d_v E_{x_0}^T|_{W_v}$ is invertible):

$$(4.4) \qquad \begin{aligned} |\pi_{\widehat{Z}_v}(\widehat{\xi}_v)| &= \sup_{w \in W_v} \frac{|\langle \widehat{\xi}_v, d_v E_{x_0}^T(w) \rangle|}{|d_v E_{x_0}^T(w)|} \\ &\leq \sup_{w \in W_v} \frac{|\langle \widehat{\xi}_v, d_v E_{x_0}^T(w) \rangle|}{\|w\|_{L^2}} \|(d_v E_{x_0}^T|_{W_v})^{-1}\| \\ &= \sup_{w \in W_v} \frac{|\langle v, w \rangle_{L^2}|}{\|w\|_{L^2}} \|(d_v E_{x_0}^T|_{W_v})^{-1}\| \\ &\leq \|v\|_{L^2} \|(d_v E_{x_0}^T|_{W_v})^{-1}\|. \end{aligned}$$

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 (4.4) we conclude that for some constant $C > 0$, the estimate $|\pi_{\widehat{Z}_v}(\widehat{\xi}_v)| \leq C$ holds true, which implies as well, by the triangular inequality, that

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

Thus, setting $\xi_v = (P_{0,T}^v)^* \widehat{\xi}_v \in T_{x_0}^* M$ (cf. Figure 1) we compute thanks to (4.1)

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

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

Remark 3. 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 (4.2). Then if we call $Z_v := (P_{0,T}^v)^* \widehat{Z}_v \subset T_{x_0}^* M$, the map

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

is continuous.[2] Moreover, Z_v is by construction transversal to Ξ_y^v, and $\xi_v \in (\lambda_u + Z_v) \cap \Xi_y^v$.

Having in mind Remark 3, we deduce the following.

COROLLARY 29. *Let $O \subset \mathrm{int}\left(A_{x_0}^T\right)$ be an open set, and assume that*

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

Let $x \in O$, $u \in \mathcal{U}_x^{\min}$, and consider $\mathcal{V}_u \subset \Omega_{x_0}^T$ as in Proposition 28. Then, for every $\lambda_u \in \Xi_x^u$, there exists a k_O-dimensional compact ball A_u, centered at λ_u and transversal to Ξ_x^u, such that

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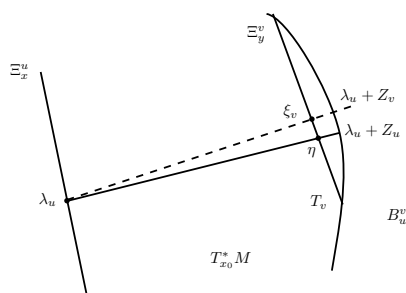

FIG. 2. *On the fiber $T_{x_0}^* M$, the point η denotes the intersection between T_v and the affine space $\lambda_u + Z_u$.*

Proof. Let $\lambda_u \in \Xi_x^u$ 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 28), and centered at λ_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 28, implies that the intersection of B_u^v with Ξ_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 $\mathrm{int}\,(T_v)$ is not empty.

Let us consider as before (cf. Remark 3) the k_O-dimensional affine subspace $\lambda_u + Z_u$, which is transversal to Ξ_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 subcover $\mathcal{W}_{v_1}, \ldots, \mathcal{W}_{v_l}$ of $\mathcal{V}_u$ and choose C_0 as the maximum between $C_0^{v_1}, \ldots, C_0^{v_l}$.

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

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

---

[2]The continuity of the map $v \mapsto Z_v$ is to be intended in the Grassmannian $\mathrm{Gr}(k_O, T_{x_0}^* M)$, of k_O-dimensional subspaces in $T_{x_0}^* M$. Also, continuity is not affected by the common translation $Z_v \mapsto Z_v + \lambda_u$.

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

Collecting all the results, we can now prove Theorem 26.

Proof of Theorem 26. Let O be an open set in int $\left(A_{x_0}^T\right)$, 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 int $\left(A_{x_0}^T\right)$ by Lemma 17. Then we may suppose that k_O is strictly less than m, for otherwise there would be nothing to prove. Moreover, by Lemma 25 it is not restrictive to assume that class $(y) = k_O$ for every $y \in O$.

Fix then a point $x \in \Sigma_c \cap O$. Since the hypotheses of Proposition 28 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 λ_u and transversal to Ξ_x^u, such that (Corollary 29)

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

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

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

The union $A_{u_1} \cup \cdots \cup A_{u_l}$ is again of positive codimension. Now we claim that there exists an open neighborhood $O_x \subset O$ of x such that, for every $y \in O_x$ and for every $v \in \mathcal{U}_y^{\min}$, there exists $i \in \{1, \ldots, l\}$ such that $A_{u_i} \cap \Xi_y^v \neq 0$. Indeed, assume by contradiction that this is not true. Then there exists a sequence of points $\{x_n\}_{n \in \mathbb{N}}$ converging to x, and a sequence of optimal controls $v_n \in \mathcal{U}_{x_n}^{\min}$ such that $v_n \notin \bigcup_{i=1}^l \mathcal{V}_{u_i}$. This, however, is in contradiction with Lemma 18 and we get the absurd.

On the other hand, for any fair point $z \in \Sigma_f \cap O_x$, its (unique) optimal control v admits a normal lift, and we have the equality

$$\text{Exp}_{x_0}^T(\Xi_z^v) = z,$$

where $\text{Exp}_{x_0}^T$ is the exponential map with base point x_0 at time T of Definition 9. Eventually, we deduce the inclusion

(4.5) $$\Sigma_f \cap O_x \subset \text{Exp}_{x_0}^T\left(A_{u_1} \cup \cdots \cup A_{u_l}\right).$$

The set on the right-hand side is closed, being the continuous image of a compact set. Moreover, it is of measure zero by the classical Sard lemma [32], as it is the image of a set of positive codimension by construction. Since the set $\Sigma_f \cap O_x$ is dense in O_x by Corollary 15, passing to the closures in (4.5) we conclude that meas$(O_x) = 0$, which is impossible. $\qquad \square$

Combining now Lemma 20 and Theorem 26, we obtain the following (cf. Theorem 1).

COROLLARY 30. *The set Σ_t of tame points is open and dense in* int $\left(A_{x_0}^T\right)$.

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*

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

is strongly compact in the L^2-topology.

The first result of this section, which is an adaptation of an argument of [31, 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 $O_x \subset \text{int}\left(A_{x_0}^T\right)$ of x such that

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

is surjective, and there exists a smooth right inverse $\Phi : O_x \to \mathcal{V}_u$ such that $E_{x_0}^T(\Phi(y)) = y$ 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 Φ is smooth, there exists $R > 0$ and $C_0 > 0$ such that

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

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

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

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 [34, 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 $|x - y| \leq C_0 R$. $\qquad\square$

DEFINITION 33. *We define the set* $\Sigma \subset \mathrm{int}\left(A_{x_0}^T\right)$ *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, and*

(b) x *is not conjugate to* x_0 *along* $x_u(\cdot)$ *(cf. Definition* 10*).*

Item (a) in Definition 33 is equivalent to requiring 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 $\mathrm{int}\left(A_{x_0}^T\right)$.

The following result finally proves Theorem 2.

THEOREM 34 (density of smooth points). Σ *is open and dense in* $\mathrm{int}\left(A_{x_0}^T\right)$. *Moreover,* $S_{x_0}^T$ *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 Σ_t of tame points is open and dense in $\mathrm{int}\left(A_{x_0}^T\right)$, 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 Proposition 16, $\mathrm{meas}(\Sigma_f \cap O') = \mathrm{meas}(O')$. Moreover, any point in $\Sigma_f \cap O'$ is also contained in the image of the exponential map $\mathrm{Exp}_{x_0}^T$, and the 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 $\mathrm{meas}(\Sigma \cap O') = \mathrm{meas}(\Sigma_f \cap O') = \mathrm{meas}(O')$, 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 $\mathrm{int}\left(A_{x_0}^T\right)$. 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 $d_{u_n}E_{x_0}^T$ 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 λ_n and ξ_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 $\lim_{n\to\infty} \lambda_n = \lim_{n\to\infty} \xi_n = \lambda$, 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 $\mathrm{Exp}_{x_0}^T$. Then there exist full neighborhoods $V \subset T_{x_0}^* M$ of λ^0 and $O_x \subset \mathrm{int}\left(A_{x_0}^T\right)$ of x such that the exponential map $\mathrm{Exp}_{x_0}^T\big|_V : V \to O_x$ 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 $\mathrm{Exp}_{x_0}^T$, and from the equality $E_{x_0}^T(u) = \mathrm{Exp}_{x_0}^T(\lambda_y)$, 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 the previous point, there are neighborhoods $V_\lambda \subset T_{x_0}^* M$ of λ and $O_x \subset \text{int}\left(A_{x_0}^T\right)$ of x such that $\text{Exp}_{x_0}^T\big|_{V_\lambda} : V_\lambda \to O_x$ 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. $\square$

Appendix A. A few technical results. We give here the proof of Lemma 17 and Lemma 20.

LEMMA. *The set Σ_c is a residual subset of $\text{int}\left(A_{x_0}^T\right)$.*

Proof. We will show that the complement of Σ_c is a meager set; i.e., it can be included into a countable union of closed, nowhere dense subsets of $\text{int}\left(A_{x_0}^T\right)$. Then the claim will follow from the classical Baire category theorem, which 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 exist $\varepsilon > 0$ and a sequence $x_n \to x$ 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

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

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

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

LEMMA. *Let $x \in \text{int}\left(A_{x_0}^T\right)$ be a tame point. Then the following hold:*
(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 $S_{x_0}^T\big|_{O_x}$ 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 $\lim_{n \to +\infty} S_{x_0}^T(x_n) = S_{x_0}^T(x)$. 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 $\text{int}\left(A_{x_0}^T\right)$. 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 $\{S_{x_0}^T(x_n)\}_{n \in \mathbb{N}}$. Then, it is not restrictive to assume that $\lim_{n \to \infty} S_{x_0}^T(x_n) = a$. 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 converges in $L^2([0,T],\mathbb{R}^d)$ to some admissible control $\widehat{u}$ steering x_0 to x, which in turn yields the inequality

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

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 end-point 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 $C_T(w) \le b + \varepsilon < a$. 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 λ_n such that

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

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 $\overline{\lambda} = \lim_{k \to \infty} \lambda_{n_k}$ exists. Thus, passing to the limit as k tends 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 $S_{x_0}^T\big|_{O_x}$ is indeed a continuous map. $\qquad\square$

REFERENCES

[1] A. A. AGRACHEV, *Any sub-Riemannian metric has points of smoothness*, Dokl. Akad. Nauk, 424 (2009), pp. 295–298, https://doi.org/10.1134/S106456240901013X.

[2] A. A. AGRACHEV, D. BARILARI, AND U. BOSCAIN, *Introduction to Riemannian and Sub-Riemannian Geometry*, Preprint SISSA 09/2012/M, 2012, http://people.sissa.it/agrachev/agrachev_files/notes.html.

[3] A. A. AGRACHEV, D. BARILARI, AND L. RIZZI, *Curvature: A variational approach*, Mem. Amer. Math. Soc., in press; preprint, https://arxiv.org/pdf/1306.5318.pdf, 2013.

[4] A. A. AGRACHEV AND P. LEE, *Optimal transportation under nonholonomic constraints*, Trans. Amer. Math. Soc., 361 (2009), pp. 6019–6047, https://doi.org/10.1090/S0002-9947-09-04813-2.

[5] A. A. AGRACHEV AND Y. L. SACHKOV, *Control Theory from the Geometric Viewpoint*, Encyclopaedia of Mathematical Sciences 87, Springer-Verlag, Berlin, 2004, https://doi.org/10.1007/978-3-662-06404-7.

[6] M. BARDI AND I. CAPUZZO-DOLCETTA, *Optimal Control and Viscosity Solutions of Hamilton-Jacobi-Bellman Equations*, Systems Control Found. Appl., Birkhäuser Boston, Boston, MA, 1997, https://doi.org/10.1007/978-0-8176-4755-1.

[7] D. BARILARI AND L. RIZZI, *Sub-Riemannian Interpolation Inequalities: Ideal Structures*, preprint, https://arxiv.org/abs/1705.05380, 2017.

[8] P. BETTIOL, H. FRANKOWSKA, AND R. B. VINTER, *Improved sensitivity relations in state constrained optimal control*, Appl. Math. Optim., 71 (2015), pp. 353–377, https://doi.org/10.1007/s00245-014-9260-6.

[9] P. Bettiol and R. B. Vinter, *Sensitivity interpretations of the costate variable for optimal control problems with state constraints*, SIAM J. Control Optim., 48 (2010), pp. 3297–3317, https://doi.org/10.1137/080732614.

[10] F. Boarotto and A. Lerario, *Homotopy properties of horizontal path spaces and a theorem of Serre in sub-Riemannian geometry*, Comm. Anal. Geom., 25 (2017), pp. 269–301, https://doi.org/10.4310/CAG.2017.v25.n2.a1.

[11] A. Bressan and B. Piccoli, *Introduction to the Mathematical Theory of Control*, AIMS Ser. Appl. Math. 2, American Institute of Mathematical Sciences (AIMS), Springfield, MO, 2007.

[12] P. Cannarsa and H. Frankowska, *Local regularity of the value function in optimal control*, Systems Control Lett., 62 (2013), pp. 791–794.

[13] P. Cannarsa and H. Frankowska, *From pointwise to local regularity for solutions of Hamilton-Jacobi equations*, Calc. Var. Partial Differential Equations, 49 (2014), pp. 1061–1074, https://doi.org/10.1007/s00526-013-0611-y.

[14] P. Cannarsa, H. Frankowska, and T. Scarinci, *Second-order sensitivity relations and regularity of the value function for Mayer's problem in optimal control*, SIAM J. Control Optim., 53 (2015), pp. 3642–3672, https://doi.org/10.1137/14098346X.

[15] P. Cannarsa and L. Rifford, *Semiconcavity results for optimal control problems admitting no singular minimizing controls*, Ann. Inst. H. Poincaré Anal. Non Linéaire, 25 (2008), pp. 773–802, https://doi.org/10.1016/j.anihpc.2007.07.005.

[16] P. Cannarsa and C. Sinestrari, *Semiconcave Functions, Hamilton-Jacobi Equations, and Optimal Control*, Progr. Nonlinear Differential Equations Appl. 58, Birkhäuser Boston, Boston, MA, 2004.

[17] Y. Chitour, F. Jean, and E. Trélat, *Singular trajectories of control-affine systems*, SIAM J. Control Optim., 47 (2008), pp. 1078–1095, https://doi.org/10.1137/060663003.

[18] F. Clarke, Y. Ledyaev, R. J. Stern, and P. Wolenski, *Nonsmooth Analysis and Control Theory*, Grad. Texts in Math. 178, Springer-Verlag, New York, 1998.

[19] F. H. Clarke, *Optimization and Nonsmooth Analysis*, Canadian Mathematical Society Series of Monographs and Advanced Texts, A Wiley-Interscience Publication, John Wiley & Sons, New York, 1983.

[20] F. H. Clarke and R. B. Vinter, *The relationship between the maximum principle and dynamic programming*, SIAM J. Control Optim., 25 (1987), pp. 1291–1311, https://doi.org/10.1137/0325071.

[21] D. Cordero-Erausquin, R. J. McCann, and M. Schmuckenschläger, *Prékopa-Leindler type inequalities on Riemannian manifolds, Jacobi fields, and optimal transport*, Ann. Fac. Sci. Toulouse Math. (6), 15 (2006), pp. 613–635, http://afst.cedram.org/item?id=AFST_2006_6_15_4_613_0.

[22] A. Figalli and L. Rifford, *Mass transportation on sub-Riemannian manifolds*, Geom. Funct. Anal., 20 (2010), pp. 124–159, https://doi.org/10.1007/s00039-010-0053-z.

[23] H. Frankowska, *Value function in optimal control*, in Mathematical Control Theory, Part 1, 2 (Trieste, 2001), ICTP Lect. Notes VIII, Abdus Salam Int. Cent. Theoret. Phys., Trieste, 2002, pp. 516–653.

[24] H. Frankowska and M. Mazzola, *On relations of the adjoint state to the value function for optimal control problems with state constraints*, NoDEA Nonlinear Differential Equations Appl., 20 (2013), pp. 361–383, https://doi.org/10.1007/s00030-012-0183-0.

[25] R. Goebel, *Regularity of the optimal feedback and the value function in convex problems of optimal control*, Set-Valued Anal., 12 (2004), pp. 127–145, https://doi.org/10.1023/B:SVAN.0000023404.46708.60.

[26] V. Jurdjevic, *Geometric Control Theory*, Cambridge Stud. Adv. Math. 52, Cambridge University Press, Cambridge, UK, 1997.

[27] E. B. Lee and L. Markus, *Foundations of Optimal Control Theory*, 2nd ed., Robert E. Krieger Publishing, Melbourne, FL, 1986.

[28] L. S. Pontryagin, V. G. Boltyanskiĭ, R. V. Gamkrelidze, and E. F. Mishchenko, *Selected Works. Vol. 4. The Mathematical Theory of Optimal Processes*, Classics of Soviet Mathematics, Gordon & Breach Science Publishers, New York, 1986, 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.

[29] D. Prandi, *Hölder equivalence of the value function for control-affine systems*, ESAIM Control Optim. Calc. Var., 20 (2014), pp. 1224–1248, https://doi.org/10.1051/cocv/2014014.

[30] L. Rifford, *Sub-Riemannian Geometry and Optimal Transport*, Springer Briefs in Mathematics, Springer, Cham, 2014, https://doi.org/10.1007/978-3-319-04804-8.

[31] L. RIFFORD AND E. TRÉLAT, *Morse-Sard type results in sub-Riemannian geometry*, Math. Ann., 332 (2005), pp. 145–159, https://doi.org/10.1007/s00208-004-0622-2.

[32] S. STERNBERG, *Lectures on Differential Geometry*, Prentice-Hall, Englewood Cliffs, NJ, 1964.

[33] H. J. SUSSMANN AND V. JURDJEVIC, *Controllability of nonlinear systems*, J. Differential Equations, 12 (1972), pp. 95–116, https://doi.org/10.1016/0022-0396(72)90007-1.

[34] E. TRÉLAT, *Some properties of the value function and its level sets for affine control systems with quadratic cost*, J. Dynam. Control Syst., 61 (2000), pp. 511–541, https://doi.org/10.1023/A:1009552511132.

[35] E. TRÉLAT, *Contrôle optimal. Théorie & applications*, Mathématiques Concrètes, Vuibert, Paris, 2005.

[36] R. VINTER, *Optimal Control*, reprint of 2000 ed., Modern Birkhäuser Classics, Birkhäuser Boston, Boston, MA, 2010, https://doi.org/10.1007/978-0-8176-8086-2.

[37] R. B. VINTER, *New results on the relationship between dynamic programming and the maximum principle*, Math. Control Signals Syst., 1 (1988), pp. 97–105, https://doi.org/10.1007/BF02551239.