ON THE REGULARITY OF ABNORMAL MINIMIZERS FOR RANK 2 SUB-RIEMANNIAN STRUCTURES

D. BARILARI, Y. CHITOUR, F. JEAN, D. PRANDI, AND M. SIGALOTTI

Abstract. We prove the C^1 regularity for a class of abnormal length-minimizers in rank 2 sub-Riemannian structures. As a consequence of our result, all length-minimizers for rank 2 sub-Riemannian structures of step up to 4 are of class C^1.

Contents

1. Introduction 1
2. Notations and preliminary notions 3
3. Desingularisation and nilpotentization 4
4. Dynamics of abnormal extremals: preliminary results 7
5. Dynamics of abnormal extremals in a special case: proof of Proposition 3 8
6. Dynamics of abnormal extremals: the general case 10
7. Proof of Theorem 1 15
Appendix A. An elementary lemma 18
References 18

1. Introduction

The question of regularity of length-minimizers is one of the main open problems in sub-Riemannian geometry, cf. for instance [Mon02, Problem 10.1] or [Agr14, Problem II] and the survey [Mon14].

Length-minimizers are solutions to a variational problem with constraints and satisfy a first-order necessary condition resulting from the Pontryagin Maximum Principle. With every length-minimizer $\gamma : [0, T] \to M$ we can associate a lift $\lambda : [0, T] \to T^*M$ in the cotangent space, satisfying a Hamiltonian equation. This lift can be either *normal* or *abnormal*, although a length-minimizer γ can actually admit several lifts, each of them being either normal or abnormal.

If a length-minimizer admits a normal lift, then it is smooth, i.e., C^∞, since normal lifts are solutions of smooth autonomous Hamiltonian systems in T^*M. Note that we assume length-minimizers to be arclength parameterized and their regularity is meant with respect to this time parameterization. The question of regularity is then reduced to length-minimizers that are strictly abnormal, i.e., those which do not admit normal lifts. For such length-minimizers, from the first order necessary condition (and actually from the second order one as well) it is a priori not possible to deduce any regularity other than Lipschitz continuity.

In this paper we investigate the following.

Open Problem. *Are all length-minimizers in a sub-Riemannian manifold of class C^1?*

If the sub-Riemannian structure has step 2, there are no strictly abnormal length-minimizers, see e.g. [AS95, ABB17], thus every length-minimizer admits a normal lift, and

Date: April 3, 2018.

is hence smooth. For step 3 structures, the situation is already more complicated and a positive answer to the above problem is known only for Carnot groups (where, actually, length-minimizers are proved to be C^∞), see [LDLMV13, TY13].

To state our main result, we introduce some notations. We refer the reader to Section 2 for precise definitions. Recall that a sub-Riemannian structure (D, g) on M is defined by a bracket generating distribution D endowed with a metric g. Hence D defines a flag of subspaces at every point $x \in M$

$$D_x = D_x^1 \subset D_x^2 \subset D_x^3 \subset \cdots \subset D_x^r = T_x M,$$

where D_x^i is the subspace of the tangent space spanned by Lie brackets of length at most i between horizontal vector fields. This induces a dual decreasing sequence of subspaces of $T_x^* M$

$$0 = (D_x^r)^\perp \subset \cdots \subset (D_x^4)^\perp \subset (D_x^3)^\perp \subset (D_x^2)^\perp \subset (D_x^1)^\perp \subset T_x^* M,$$

where perpendicularity is considered with respect to the duality product. By construction, any abnormal lift satisfies $\lambda(t) \in (D^1)^\perp$ for every t. If the lift is strictly abnormal, then by Goh conditions $\lambda(t) \in (D^2)^\perp$ for every t.

When the distribution has rank 2, it is known that if $\lambda(t)$ does not cross $(D^3)^\perp$, then the length-minimizer is C^∞ [LS95, Sect. 6.2, Cor. 4]. Our main result pushes this analysis further and establishes that the answer to the Open Problem is positive for length-minimizers whose abnormal lift does not enter in $(D^4)^\perp$.

Theorem 1. *Let (D, g) be a rank 2 sub-Riemannian structure on M. Assume that $\gamma : [0, T] \to M$ is an arclength parameterized abnormal minimizer. If γ admits a lift satisfying $\lambda(t) \notin (D^4)^\perp$ for every $t \in [0, T]$, then γ is of class C^1.*

If the sub-Riemannian manifold has rank 2 and step at most 4, the assumption in Theorem 1 is trivially satisfied by every abnormal minimizer γ and we immediately obtain the following corollary.

Corollary 2. *Assume that the sub-Riemannian structure has rank 2 and step at most 4. Then all length-minimizers are of class C^1.*

It is legitimate to ask whether the C^1 regularity in the Open Problem can be further improved. Indeed, the argument behind our proof permits to obtain C^∞ regularity of length-minimizers under an additional nilpotency condition on the Lie algebra generated by horizontal vector fields.

Proposition 3. *Assume that D is generated by two vector fields X_1, X_2 such that $\mathrm{Lie}\{X_1, X_2\}$ is a nilpotent Lie algebra of step at most 4. Then for every sub-Riemannian structure (D, g) on M, the corresponding length-minimizers are of class C^∞.*

Notice that, from a geometric viewpoint, the assumption of Proposition 3 implies that M is a homogeneous space. Proposition 3 applies in particular to Carnot groups of rank 2 and step at most 4. In this case we recover the results obtained in [LM08, Example 4.6]. Moreover, under the hypotheses of the above proposition, we also deduce a Sard property for minimizers (Corollary 13).

The strategy of proof of Theorem 1 is to show that, at points where they are not of class C^1, length-minimizers can admit only corner-like singularities. This is done by a careful asymptotic analysis of the differential equations satisfied by the abnormal lift, which exploits their Hamiltonian structure. We can then conclude thanks to the following result.

Theorem 4 ([HL16])**.** *Let M be a sub-Riemannian manifold. Let $T > 0$ and let $\gamma : [-T, T] \to M$ be a horizontal curve such that, in local coordinates, there exist*

$$\dot\gamma^+(0) := \lim_{t\downarrow 0} \frac{\gamma(t) - \gamma(0)}{t}, \qquad \dot\gamma^-(0) := \lim_{t\to 0} \frac{\gamma(t) - \gamma(0)}{t}.$$

If $\dot\gamma^+(0) \neq \dot\gamma^-(0)$, then γ is not a length-minimizer.

We oberve that the proof contained in [HL16] requires a previous result stated in [LM08]. A complete argument for the latter, addressing some issues raised in [Rif17, p. 1113-15], is provided in [MPV17]. For sub-Riemannian structures of rank 2 and step at most 4 (and indeed also for higher step, under an additional condition on the Lie algebra generated by horizontal vector fields), the fact that corners are not length-minimizers is already contained in [LM08].

1.1. **Structure of the paper.** In Section 2 we recall some notations and preliminary notions. Section 3 is devoted to a desingularization and nilpotentization argument. Section 4 contains a preliminary analysis on the dynamics of abnormal extremals. To illustrate our approach in a simpler case, we discuss in Section 5 the proof of the main result for a nilpotent structure of rank up to 4. Then in Sections 6 and 7 we complete our analysis to prove the general result. Appendix A contains a technical lemma.

Acknowledgments. This work was supported by the Grant ANR-15-CE40-0018 SRGI "Sub-Riemannian geometry and interactions" and by a public grant as part of the Investissement d'avenir project, reference ANR-11-LABX-0056-LMH, LabEx LMH, in a joint call with Programme Gaspard Monge en Optimisation et Recherche Opérationnelle.

2. Notations and preliminary notions

Let M be a smooth n-dimensional manifold. A *sub-Riemannian structure* of rank m on M is a triplet (E, g^E, f) where E is a vector bundle of rank m over M, g^E is an Euclidean metric on E, and $f : E \to TM$ is a morphism of vector bundles such that $f(E_x) \subseteq T_xM$ for every $x \in M$. Fix such a structure and define a family of subspaces of the tangent spaces by

$$D_x = \{X(x) \mid X \in D\} \subseteq T_xM, \qquad \forall\, x \in M,$$

where $D = \{f \circ Y \mid Y \text{ smooth section of } E\}$ is a submodule of the set of vector fields on M. We assume that the structure is *bracket generating*, i.e., the tangent space T_xM is spanned by the vector fields in D and their iterated Lie brackets evaluated at x.

The sub-Riemannian structure induces a quadratic form g_x on D_x by

$$g_x(v, v) = \inf\{g^E_x(u, u) \mid f(u) = v, u \in E_x\}, \quad v \in D_x.$$

In analogy with the classic sub-Riemannian case and to simplify notations, in the sequel we will refer to the sub-Riemannian structure as the pair (D, g) rather than (E, g^E, f). This is justified since all the constructions and definitions below rely only on D and g. The triplet (M, D, g) is called a *sub-Riemannian manifold.*

Remark 5. Usually, a sub-Riemannian manifold denotes a triplet (M, D, g), where M is a smooth manifold, D is a subbundle of TM, and g is a Riemannian metric on D (see, e.g., [Bel96]). This corresponds to the case where $f(E_x)$ is of constant rank. The definition given above follows, for instance, [ABB17].

A *horizontal curve* $\gamma : [0, T] \to M$ is an absolutely continuous path such that $\dot\gamma(t) \in D_{\gamma(t)}$ for almost every (a.e. for short) $t \in [0, T]$. The *length* of a horizontal curve is defined by

$$\ell(\gamma) = \int_0^T \sqrt{g_{\gamma(t)}(\dot\gamma(t), \dot\gamma(t))}dt.$$

The *sub-Riemannian distance* between two arbitrary points x, y in M is then

$$d(x,y) = \inf\{\ell(\gamma) \mid \gamma(0) = x,\, \gamma(T) = y,\, \gamma \text{ horizontal}\}.$$

A *length-minimizer* is a horizontal curve γ which realizes the distance between its extremities, that is, $\ell(\gamma) = d(\gamma(0), \gamma(T))$. Note that any time-reparameterization of a length-minimizer is a length-minimizer as well.

A *generating frame* of the sub-Riemannian structure is a family of vector fields $X_1, \dots, X_k$ such that D is generated by $X_1, \dots, X_k$ as a module and

$$g_x(v,v) = \inf \left\{ \sum_{i=1}^k u_i^2 \mid \sum_{i=1}^k u_i X_i(x) = v \right\}, \quad x \in U,\ v \in D_x.$$

There always exists a global generating frame (see [ABB17, Corollary 3.26]), with, in general, a number k of elements greater than the rank m of the structure. However, every point $x \in M$ admits a neighborhood on which there exists a (local) generating frame with exactly $k = m$ elements, e.g., by taking the image via f of a local orthonormal frame of (E, g^E).

Fix now a (local or global) generating frame $X_1, \dots, X_k$ of (D, g). For any horizontal curve γ of finite length, there exists $u \in L^\infty([0,T], \mathbb{R}^k)$ satisfying

$$\dot\gamma(t) = \sum_{i=1}^k u_i(t) X_i(\gamma(t)), \qquad \text{for a.e. } t \in [0,T]. \tag{1}$$

The curve is said to be *arclength parameterized* if $g_{\gamma(t)}(\dot\gamma(t), \dot\gamma(t)) = 1$ for a.e. $t \in [0,T]$, i.e., if there exists $u \in L^\infty([0,T], \mathbb{S}^{k-1})$ satisfying (1). In that case $\ell(\gamma) = T$.

To state the first order necessary conditions, let us first introduce some notations. For $\lambda \in T^*M$ and $x = \pi(\lambda)$, where $\pi : T^*M \to M$ is the canonical projection, we set $h_i(\lambda) = \langle \lambda, X_i(x) \rangle$, for $i = 1, \dots, k$ (here $\langle \lambda, \cdot \rangle$ denotes the dual action of covectors on vectors). Recall also that, for a function $H : T^*M \to \mathbb{R}$, the corresponding Hamiltonian vector field $\vec H$ is the unique vector field such that $\sigma(\cdot, \vec H) = dH$, where σ is the canonical symplectic form on the cotangent bundle.

Applying the Pontryagin Maximum Principle to the sub-Riemannian length minimization problem yields the following theorem.

Theorem 6. *Let (M, D, g) be a sub-Riemannian manifold with generating frame $X_1, \dots, X_k$ and $\gamma : [0,T] \to M$ be a length-minimizer. Then there exists a nontrivial absolutely continuous curve $t \mapsto \lambda(t) \in T^*_{\gamma(t)}M$ such that one of the following conditions is satisfied:*

- *(N) $\dot\lambda(t) = \vec H(\lambda(t))$ for all $t \in [0,T]$, where $H(\lambda) = \frac{1}{2}\sum_{i=1}^k h_i^2$,*
- *(A) $\dot\lambda(t) = \sum_{i=1}^k u_i(t) \vec h_i(\lambda(t))$ for almost every $t \in [0,T]$, with $u_1, \dots, u_k \in L^1([0,T])$. Moreover, $\lambda(t) \in (D_{\gamma(t)})^\perp$ for all t, i.e., $h_i(\lambda(t)) \equiv 0$ for $i = 1, \dots, k$.*

In case (N) (respectively, case (A)), λ is called a *normal* (respectively, *abnormal*) *extremal*. Normal extremals are integral curves of $\vec H$. As such, they are smooth. A length-minimizer is normal (respectively, abnormal) if it admits a normal (respectively, abnormal) extremal lift. We stress that both conditions can be satisfied for the same curve γ, with different lifts λ_1 and λ_2.

3. Desingularisation and nilpotentization

3.1. **Desingularisation.** Let (M, D, g) be a sub-Riemannian manifold. We define recursively the following sequence of submodules of the set of vector fields,

$$D^1 = D, \quad D^{i+1} = D^i + [D, D^i].$$

At every point $x \in M$, the evaluation at x of these modules induces a flag of subspaces of the tangent space,

$$D^1_x \subset D^2_x \subset \cdots \subset D^r_x = T_xM.$$

The smallest integer $r = r(x)$ satisfying $D^r_x = T_xM$ is called the *step of D at x*. A point is said to be *regular* if the dimensions of the subspaces of the flag are locally constant in an open neighborhood of the point. When every point in M is regular, the sub-Riemannian manifold is said to be *equiregular*.

In general a sub-Riemannian manifold may admit non-regular points. However, for our purposes, we can restrict ourselves with no loss of generality to equiregular manifolds thanks to a desingularisation procedure.

Lemma 7. *Fix an integer $m \geq 2$. Assume that for every rank m equiregular sub-Riemannian structure the following property holds: every arclength parameterized abnormal minimizer admitting a lift $\lambda(t) \notin (D^4)^\perp$ is of class C^1. Then the same property holds true for every rank m sub-Riemannian structure.*

Proof. Let (M, D, g) be a non-equiregular sub-Riemannian manifold of rank m and γ be an abnormal length-minimizer of (M, D, g) which admits an abnormal extremal lift such that $\lambda(t) \notin (D^4)^\perp$ for every $t \in [0, T]$. Assume moreover that γ is arclength parameterized. We have to prove that γ is of class C^1.

Fix $t_0 \in [0, T]$ and a generating frame $X_1, \dots, X_m$ on a neighborhood of $\gamma(t_0)$. By [Jea14, Lemma 2.5], there exists an equiregular sub-Riemannian manifold $(\widetilde{M}, \widetilde{D}, \widetilde{g})$ of rank m with a generating frame $\xi_1, \dots, \xi_m$ and a map $\varpi : \widetilde{M} \to M$ onto a neighborhood $U \subset M$ of $\gamma(t_0)$ such that $\varpi_* \xi_i = X_i$. Up to reducing the interval $[0, T]$ we assume that $\gamma(t) \in U$ for all $t \in [0, T]$. Let $u \in L^\infty([0, T], \mathbb{S}^{m-1})$ be such that

$$\dot\gamma(t) = \sum_{i=1}^m u_i(t) X_i(\gamma(t)), \qquad \text{a.e. } t.$$

By construction, since γ is a length-minimizer, there exists a length-minimizer $\widetilde{\gamma}$ in $\widetilde{M}$ with $\varpi(\widetilde{\gamma}) = \gamma$ associated with the same u, that is,

$$\dot{\widetilde{\gamma}}(t) = \sum_{i=1}^m u_i(t) \xi_i(\widetilde{\gamma}(t)), \qquad \text{a.e. } t,$$

which is arclength parameterized as well. Hence the trajectory γ has at least the same regularity as $\widetilde{\gamma}$.

Moreover, if λ is an abnormal lift of γ in T^*M, then $\widetilde{\gamma}$ admits an abnormal lift $\widetilde{\lambda}$ in $T^*\widetilde{M}$ such that $\widetilde{\lambda}(t) = \varpi^* \lambda(t)$ for every t. Since $\varpi^*(\widetilde{D}^k)^\perp = (D^k)^\perp$ for any positive integer k, the property $\lambda(t) \notin (D^4)^\perp$ implies $\widetilde{\lambda}(t) \notin (\widetilde{D}^4)^\perp$.

It results from the hypothesis that $\widetilde{\gamma}$ is C^1, so γ is of class C^1 in an open neighborhood of $t_0 \in [0, T]$, which ends the proof. □

As a consequence of Lemma 7, we can assume in the rest of the paper that the sub-Riemannian manifold is equiregular.

3.2. **Nilpotentization.** Let us recall the construction of the nilpotent approximation (see for instance [Bel96] for details).

Let (M, D, g) be an equiregular sub-Riemannian manifold. We fix a point $x \in M$ and a local generating frame $X_1, \dots, X_m$ in a neighborhood of x.

For $i = 1, \dots, n$, let w_i be the smallest integer j such that $\dim D^j_x \geq i$. We define the dilations $\delta_\nu : \mathbb{R}^n \to \mathbb{R}^n$ for $\nu \in \mathbb{R}$ as $\delta_\nu(z) = (\nu^{w_1} z_1, \dots, \nu^{w_n} z_n)$. Let z^x be a system of privileged coordinates at x and set $\delta^x_\nu = \delta_\nu \circ z^x$. Then, for $i = 1, \dots, m$, the vector field $\varepsilon \big(\delta^x_{1/\varepsilon}\big)_* X_i$ converges locally uniformly as $\varepsilon \to 0$ to a vector field $\widehat{X}^x_i$ on $\mathbb{R}^n$. The space $\mathbb{R}^n$ endowed with the sub-Riemannian structure having $\widehat{X}^x_1, \dots, \widehat{X}^x_m$ as generating frame is called the *nilpotent approximation of* (M,D,g) *at* x and is denoted by $\widehat{M}_x$. This nilpotent approximation $\widehat{M}_x$ is a Carnot group equipped with a left-invariant sub-Riemannian structure.

Since (M,D,g) is equiregular, we can locally choose systems of privileged coordinates z^x depending continuously on x [Jea14, Sect. 2.2.2]. Note that the w_i's and δ_ν are independent of x. Thus an easy adaptation of the proof of [AGM15, Prop. 3.4] (see also [ABB17, Sect. 10.4.1]) shows that, for $i=1,\dots,m$, the vector field $\varepsilon(\delta^x_{1/\varepsilon})_* X_i$ converges locally uniformly to $\widehat{X}_i^{x_0}$ as $\varepsilon\to 0$ and $x\to x_0$.

Lemma 8. *Let* $(a_n)_{n\in\mathbb{N}},(b_n)_{n\in\mathbb{N}}\subset[0,T]$, $\bar a\in[0,T]$, *be such that* $a_n,b_n\to\bar a$ *and* $a_n<b_n$ *for any* $n\in\mathbb{N}$. *Given* $u\in L^\infty([0,T],\mathbb{S}^{m-1})$ *and* $n\in\mathbb{N}$, *define* $u_n\in L^\infty([0,1],\mathbb{S}^{m-1})$ *by*

$$u_n(\tau)=u(a_n+\tau(b_n-a_n)).$$

Assume that $(u_n)_{n\in\mathbb{N}}$ *converges to* $u_\star\in L^\infty([0,1],\mathbb{R}^m)$ *for the weak-*$\star$ *topology of* $L^\infty([0,1],\mathbb{R}^m)$ *and, moreover, that the trajectory* $\gamma:[0,T]\to M$ *associated with* u *is a length-minimizer. If* $x=\gamma(\bar a)$, *then the trajectory* $\gamma_\star:[0,1]\to\widehat{M}_x$ *satisfying*

$$\dot\gamma_\star(s)=\sum_{i=1}^m u_{\star,i}(s)\widehat{X}_i^x(\gamma_\star(s)),\qquad \gamma_\star(0)=0,$$

is also a length-minimizer. In particular, $u_\star(t)\in\mathbb{S}^{m-1}$ *for almost every* $t\in[0,1]$.

Proof. We consider a continuously varying family of privileged coordinates $z^{\gamma(t)}$, $t\in[0,T]$, and the corresponding 1-parameter family of dilations $\delta^t_\nu:=\delta^{\gamma(t)}_\nu$. It is not restrictive to assume that $\delta^{a_n}_{\frac{1}{b_n-a_n}}\gamma(t)$ is well-defined for every $n\in\mathbb{N}$ and $t\in[a_n,b_n]$.

Let γ_n be defined by $\gamma_n(\tau)=\delta^{a_n}_{\frac{1}{b_n-a_n}}(\gamma(a_n+\tau(b_n-a_n)))$. Then, γ_n is a length-minimizing curve for the sub-Riemannian structure on $\mathbb{R}^n$ with orthonormal frame

$$\varepsilon_n\Big(\delta^{a_n}_{\frac{1}{b_n-a_n}}\Big)_* X_1,\dots,\varepsilon_n\Big(\delta^{a_n}_{\frac{1}{b_n-a_n}}\Big)_* X_m.$$

The corresponding control is u_n.

Since the sequence $\Big((b_n-a_n)\Big(\delta^{a_n}_{\frac{1}{b_n-a_n}}\Big)_* X_i\Big)_{n\in\mathbb{N}}$ converges locally uniformly to $\widehat{X}^x_i$, it follows by standard ODE theory that $(\gamma_n)_{n\in\mathbb{N}}$ converges uniformly to $\gamma_\star$.

We claim that $\widehat{d}(\gamma_\star(0),\gamma_\star(1))=1$. Indeed, $\ell(\gamma_\star)\ge\widehat{d}(\gamma_\star(0),\gamma_\star(1))$ and, by [Bel96, Theorem 7.32], we have

$$\begin{aligned}\widehat{d}(\gamma_\star(0),\gamma_\star(1))&=\lim_{n\to\infty}\frac{1}{b_n-a_n}d(\gamma(a_n),\gamma(b_n))\\&=\lim_{n\to\infty}\int_0^1|u(a_n+\tau(b_n-a_n))|\,d\tau=1,\end{aligned}$$

where $|\cdot|$ denotes the norm in $\mathbb{R}^m$. On the other hand, by weak-$\star$ convergence we have

$$\ell(\gamma_\star)=\|u_\star\|_{L^1([0,1],\mathbb{R}^m)}\le\liminf_{n\to\infty}\|u_n\|_{L^1([0,1],\mathbb{R}^m)}=1,$$

proving the claim.

To conclude the proof, it suffices now to observe that the above implies that $\gamma_\star$ is minimizing. In particular, since $|u_\star(t)|\le 1$ a.e. on $[0,1]$ by the properties of weak-$\star$ convergence, this shows that $|u_\star(t)|=1$ a.e. on $[0,1]$. □

Corollary 9. *Let* γ, u, $\bar a$, $(a_n)_{n\in\mathbb{N}}$, $(b_n)_{n\in\mathbb{N}}$, *and* $u_\star$ *be as in Lemma 8. Assume that there exist* $u_+,u_-\in\mathbb{S}^{m-1}$ *such that* $u_\star=u_-$ *almost everywhere on* $[0,1/2]$ *and* $u_\star=u_+$ *almost everywhere on* $[1/2,1]$. *Then* $u_-=u_+$.

Proof. If $u_-\neq u_+$, then $\gamma_\star$ is not length-minimizing by Theorem 4, which contradicts Lemma 8. □

4. Dynamics of abnormal extremals: preliminary results

In this section we present the dynamical system associated with the abnormal extremal, whose analysis is the basis for the proof of Theorem 1, and we derive a first result on its structure.

4.1. **Introduction to the dynamical system.** Let (M, D, g) be an equiregular sub-Riemannian manifold of rank 2. Since the arguments are local, in what follows we fix a local generating frame $\{X_1, X_2\}$ of (D, g).

Consider an abnormal length-minimizer $\gamma : [0, T] \to M$ parameterized by arclength. Then $T = d(\gamma(0), \gamma(T))$ and there exists $u \in L^\infty([0, T], \mathbb{S}^1)$ such that

$$\dot\gamma(t) = u_1(t)X_1(\gamma(t)) + u_2(t)X_2(\gamma(t)), \quad \text{a.e. } t \in [0, T].$$

Moreover from Theorem 6, γ admits a lift $\lambda : [0, T] \to T^*M$ which satisfies

$$\dot\lambda(t) = u_1\vec{h}_1(\lambda(t)) + u_2\vec{h}_2(\lambda(t)) \quad \text{and} \quad h_1(\lambda(t)) \equiv h_2(\lambda(t)) \equiv 0.$$

By a slight abuse of notation, set $h_i(t) = \langle\lambda(t), X_i(\gamma(t))\rangle$, $i = 1, 2$, and for every $i_1, \dots, i_m \in \{1, 2\}$,

$$h_{i_1\cdots i_m}(t) = \langle\lambda(t), [X_{i_1}, \dots, [X_{i_m-1}, X_{i_m}]](\gamma(t))\rangle.$$

Such a function $h_{i_1\cdots i_m}$ is absolutely continuous and satisfies

$$\dot h_{i_1\cdots i_m}(t) = u_1(t)h_{1i_1\cdots i_m}(t) + u_2(t)h_{2i_1\cdots i_m}(t) \quad \text{for a.e. } t \in [0, T]. \tag{2}$$

Differentiating the equalities $h_1 \equiv h_2 \equiv 0$ and using (2) we obtain $h_{12} \equiv 0$. Differentiating again we get

$$0 = \dot h_{12} = u_1h_{112} + u_2h_{212} \quad \text{a.e. on } [0, T]. \tag{3}$$

Remark 10. The identities $h_1(t) = h_2(t) = h_{12}(t) = 0$ imply that $\lambda(t) \in (D^2)^\perp$ for every t. The latter is known as *Goh condition* and is in general (i.e., for sub-Riemannian structures of any rank) a necessary condition for the associated curve to be length-minimizing [AS99]. It is known that a generic sub-Riemannian structure of rank larger than 2 does not have non-constant abnormal extremals satisfying the Goh condition [CJT06].

Let $h = (-h_{212}, h_{112})$ and $(t_0, t_1) \subset (0, T)$ be a maximal (i.e., non-extendable) open interval on which $h \neq 0$. Equation (3) then implies that $u = \pm h/|h|$ almost everywhere on (t_0, t_1).

Moreover, by length-minimality of γ we can assume without loss of generality that $u = h/|h|$ on (t_0, t_1) (see Lemma 21 in the appendix). Thus γ may be non-differentiable only at a time t such that $h(t) = 0$. In particular, if the step of the sub-Riemannian structure is not greater than 3, then γ is differentiable everywhere. We assume from now on that the step is at least 4.

Observe that from (2) and using $u = h/|h|$ one obtains

$$\dot h = A\frac{h}{|h|}, \qquad A = \begin{pmatrix} -h_{2112} & -h_{2212} \\ h_{1112} & h_{2112} \end{pmatrix}, \qquad \text{on } (t_0, t_1). \tag{4}$$

Here, we used the relation $h_{1212} = h_{2112}$, which follows from the Jacobi identity. Observe that the matrix A has zero trace and is absolutely continuous on the whole interval $[0, T]$.

Lemma 11. *Assume that $\lambda(t) \notin (D^4_{\gamma(t)})^\perp$ for every $t \in [0, T]$. If $h(t_0) = 0$ for some $t_0 \in [0, T]$, then $A(t_0) \neq 0$.*

Proof. The fact that γ is abnormal implies that the non-zero covector $\lambda(t)$ annihilates $D_{\gamma(t)}$ for every $t \in [0, T]$. The Goh condition $h_{12} \equiv 0$ guarantees that it also annihilates $D^2_{\gamma(t)}$. The fact that $h(t_0) = 0$ says, moreover, that $\lambda(t_0)$ annihilates $D^3_{\gamma(t_0)}$. If $A(t_0)$ is equal to zero, then $\lambda(t_0)$ annihilates $D^4_{\gamma(t_0)}$, which contradicts the assumption. □

4.2. **The sign of** $\det A$ **is non-negative where** h **vanishes.** A key step in the proof of Theorem 1 is the following result.

Proposition 12. *Let* (t_0, t_1) *be a maximal open interval of* $[0, T]$ *on which* $h \neq 0$ *and assume that* $t_1 < T$. *Then* $\det A(t_1) \leq 0$.

Proof. Assume by contradiction that $\det A(t_1) > 0$. Since $\operatorname{trace} A(t_1) = 0$, there exists $P \in \mathrm{GL}(2, \mathbb{R})$ such that

$$P^{-1}A(t_1)P = \begin{pmatrix} 0 & -a \\ a & 0 \end{pmatrix}, \qquad a > 0. \tag{5}$$

Define the scalar functions α, β and ζ through the relation

$$P^{-1}A(t)P = \begin{pmatrix} -\alpha(t) & \beta(t) \\ \zeta(t) & \alpha(t) \end{pmatrix},$$

and notice that α, β, ζ are absolutely continuous with bounded derivatives on (t_0, t_1), since they are linear combinations of $h_{2112}, h_{2212}, h_{1112}$, according to (2). Clearly, (5) implies that $\alpha(t) \to 0$, $\beta(t) \to -a$, and $\zeta(t) \to a$ as $t \to t_1$.

Consider a time rescaling and a polar coordinates representation so that $P^{-1}h(t) = \rho(s(t))e^{i\vartheta(s(t))}$, where

$$s(t) := \int_{t_0}^{t} \frac{|P^{-1}h(\tau)|}{|h(\tau)|} d\tau.$$

It is useful to introduce $\mu := (\zeta + \beta)/2$ and $\eta := (\zeta - \beta)/2$. Then, denoting by ρ' and ϑ' the derivatives of ρ and ϑ with respect to the parameter s, (4) can be rewritten as

$$\begin{cases} \rho' = (-\alpha \cos 2\vartheta + \mu \sin 2\vartheta), \\ \vartheta' = \frac{1}{\rho}(\alpha \sin 2\vartheta + \mu \cos 2\vartheta + \eta). \end{cases}$$

Let $w = \alpha \sin 2\vartheta + \mu \cos 2\vartheta + \eta$ and notice that $2a > w > a/2$ in a left-neighborhood of $s(t_1)$. Therefore,

$$\begin{aligned} (\rho^2 w)' =& 2\rho(-\alpha \cos 2\vartheta + \mu \sin 2\vartheta)w + \rho^2(\alpha' \sin 2\vartheta + \mu' \cos 2\vartheta + \dot\eta) \\ &+ \rho^2(\alpha \cos 2\vartheta - \mu \sin 2\vartheta)2\vartheta' \\ =& \rho^2 w \frac{\alpha' \sin 2\vartheta + \mu' \cos 2\vartheta + \eta'}{w} \geq -M\rho^2 w, \end{aligned}$$

for some constant $M > 0$. This implies at once that $t \mapsto e^{Mt}\rho^2(t)w(t)$ is increasing, and hence that it is impossible for $\rho^2 w$ to tend to zero as $s \to s(t_1)$. This contradicts the assumption that $\rho(t) \to 0$ as $t \to t_1$, completing the proof of the statement. □

5. Dynamics of abnormal extremals in a special case: proof of Proposition 3

In this section we prove Proposition 3. We present it here to illustrate in a simpler context the general procedure used later to complete the proof of Theorem 1.

Assume that D is generated by two vector fields X_1, X_2 such that the Lie algebra $\mathrm{Lie}\{X_1, X_2\}$ is nilpotent of step at most 4. This means that all Lie brackets of X_1, X_2 of length 5 vanish. Notice that this implies that M is a homogeneous space (see [Bel96], [Jea14, Sect. 2.3.1] or [ABB17, Sect. 10.5]). In particular $\dim M \leq 8$.

Proof of Proposition 3. Without loss of generality, we assume that the step is equal to 4. Recall that for an abnormal minimizer on an interval I we have

$$h_1 \equiv h_2 \equiv h_{12} \equiv 0, \qquad 0 = \dot h_{12} = u_1 h_{112} + u_2 h_{212} \quad \text{a.e. on } I.$$

The vector $h = (-h_{212}, h_{112})$ satisfies the differential equation

$$\dot{h} = Au, \qquad A = \begin{pmatrix} -h_{2112} & -h_{2212} \\ h_{1112} & h_{2112} \end{pmatrix}, \quad \text{a.e. on } I. \tag{6}$$

Notice that A is a constant matrix (with zero trace), as follows from (2) and the nilpotency assumption.

As we have already seen in the general case, on every interval where $h(t) \neq 0$ we have that u is smooth and equal to either $\frac{h(t)}{|h(t)|}$ or $-\frac{h(t)}{|h(t)|}$.

We are then reduced to the case where h vanishes at some point $\bar{t} \in I$. In this case the matrix A cannot be zero, as it follows from Lemma 11.

We consider separately the following cases:

(a) $h(t) = 0$ for all $t \in I$;
(b) there exist $\bar{t}, t_* \in I$ such that $h(\bar{t}) = 0$ and $h(t_*) \neq 0$.

Case (a). From (6) it follows that $u(t)$ is in the kernel of A for a.e. $t \in I$. Since u is nonzero for a.e. $t \in I$, then necessarily A has one-dimensional kernel $\ker A = \operatorname{span}\{\bar{u}\}$, where $\bar{u}$ has norm one. Then $u(t) = \sigma(t)\bar{u}$ for a.e. $t \in I$, with $\sigma(t) \in \{-1, 1\}$ and

$$\dot{\gamma}(t) = \sigma(t) X_{\bar{u}}(\gamma(t)), \quad \text{a.e. } t \in I,$$

with $X_{\bar{u}}$ a constant vector field. Since γ is a length-minimizer then σ is constant, and u is smooth, thanks to Lemma 21 in the appendix.

Case (b). Consider the maximal neighborhood $J = (t_0, t_1)$ of t_* on which h is non-vanishing. In particular $h(t_0) = 0$ or $h(t_1) = 0$. Consider the case $h(t_1) = 0$.

The trajectories of (4) are time reparameterizations of those of the linear system $\dot{z} = Az$. Hence h stays in the stable or in the unstable manifold of A. Recall that $\det A \leq 0$ by Proposition 12 and notice that if $\det A = 0$ then no nontrivial stable nor unstable manifold exists. We deduce that $\det A < 0$.

Denote by $\lambda_\pm$ the eigenvalues of A and by $v_\pm$ the corresponding unit eigenvectors. Since h belongs to the stable (respectively, unstable) manifold of A then $\frac{h(t)}{|h(t)|}$ is constantly equal to v_- or $-v_-$ on J (respectively, v_+ or $-v_+$). Then we can integrate (6) and get

$$h(t) = h(t_*) \pm (t - t_*)\lambda_- v_-, \qquad t \in J,$$

or

$$h(t) = h(t_*) \pm (t - t_*)\lambda_+ v_+, \qquad t \in J.$$

In particular $\lim_{t \downarrow t_0} h(t) \neq 0$ and $J = I \cap (-\infty, t_1)$. Similarly, if $h(t_0) = 0$ then one has $\lim_{t \uparrow t_1} h(t) \neq 0$ and $J = I \cap (t_0, +\infty)$.

If there exist two distinct maximal intervals $J_1, J_2 \subset I$ where h does not vanish, then

$$I \setminus (J_1 \cup J_2) = [\tau_0, \tau_1],$$

for some $\tau_0 \leq \tau_1$ in I. If $\tau_0 < \tau_1$, we can apply case (a) on the interval (τ_0, τ_1), which leads to a contradiction since A should have nontrivial kernel. We are thus left to consider the case where $\tau_0 = \tau_1 = \bar{t}$, that is, where $h(t) \neq 0$ for $t \in I \setminus \{\bar{t}\}$. In this case u is piecewise constant on $I \setminus \{\bar{t}\}$ and satisfies

$$\lim_{t \downarrow \bar{t}} u(t) \in \{v_-, -v_-, v_+, -v_+\}, \quad \lim_{t \uparrow \bar{t}} u(t) \in \{v_-, -v_-, v_+, -v_+\}.$$

Theorem 4 and the length-minimizing assumption on γ imply that the two limits must be equal. Hence, u is constant on I, and in particular it is smooth. □

As a byproduct of the previous proof we get that the Sard conjecture for minimizers holds [Agr14], which is known in the free case [LDMO+16]. More precisely we have the following.

Corollary 13. *Let (D, g) be a rank 2 sub-Riemannian structure on M. Assume that D is generated by two vector fields X_1, X_2 such that* $\mathrm{Lie}\{X_1, X_2\}$ *is a nilpotent Lie algebra of step at most* 4*. Then the set of points reached by abnormal minimizers has codimension at least 2.*

Proof. We proved that given any initial covector in $(D^2)^\perp$, there exists at most four length-minimizing curves, whose extremal lifts start with this covector (one if $h \neq 0$, four if $h = 0$). Hence the set of final points of abnormal minimizers has codimension 2, that is, the codimension of $(D^2)^\perp$ (which is equal to three) minus one, taking into account the time parameterization. □

For recent results on the Sard conjecture for rank 2 structures in 3-dimensional manifolds, see [BdSRar], which extends the analysis in [ZZ95].

6. Dynamics of abnormal extremals: the general case

The goal of this section is to prove the following result.

Proposition 14. *Let (t_0, t_1) be a maximal interval on which $h \neq 0$. Assume that $t_1 < T$ and $A(t_1) \neq 0$. Then $u(t)$ has a limit as $t \uparrow t_1$.*

We split the analysis in two steps. The first one, which is a rather straightforward adaptation of the proof of Proposition 3, corresponds to the case where $\det A(t_1) < 0$. We will then turn to the case where $\det A(t_1) = 0$ (recall that, according to Proposition 12, $\det A(t_1)$ cannot be positive).

For this purpose, we start by proving a preliminary result.

6.1. A time-rescaling lemma.

The result below highlights the fact that equation (4) is "almost invariant" with respect to similarity of A.

Lemma 15. *For $P \in \mathrm{GL}(2, \mathbb{R})$ and $t_* \in (t_0, t_1)$, we consider the time reparameterization given by*

$$\varphi : [t_*, t_1) \ni t \mapsto s := \int_{t_*}^{t} \frac{d\tau}{|h(\tau)|}.$$

Let $\mathfrak{h} = P^{-1} h \circ \varphi^{-1}$ and $\mathfrak{A} = P^{-1}(A \circ \varphi^{-1})P$. Then,

(i) *$\varphi(t) \to +\infty$ as $t \to t_1$;*
(ii) *for any $p \in [1, +\infty]$ we have $\mathfrak{h} \in L^p((0, +\infty), \mathbb{R}^2)$;*
(iii) *for every $s \in (0, +\infty)$ we have*

$$\mathfrak{h}'(s) = \mathfrak{A}(s)\mathfrak{h}(s). \tag{7}$$

Proof. We start by proving point (iii). Observe that $\dot\varphi = 1/|h|$. Then, simple computations yield

$$\mathfrak{h}' = \frac{P^{-1}\dot h \circ \varphi^{-1}}{\dot\varphi \circ \varphi^{-1}} = \mathfrak{A}\,\mathfrak{h}.$$

Assume now that $\lim_{t \to t_1} \varphi(t) = s_* < +\infty$. Then, since $\mathfrak{h}(s_*) = h(t_1) = 0$, we have that $\mathfrak{h}$ is the solution to the (backward) Cauchy problem

$$\begin{cases} \mathfrak{h}' = \mathfrak{A}\mathfrak{h} & \text{on } (0, s_*), \\ \mathfrak{h}(s_*) = 0. \end{cases}$$

This implies that $\mathfrak{h} \equiv 0$ on $(0, s_*)$ and thus $h \equiv 0$ on (t_*, t_1), which contradicts the definition of the interval (t_0, t_1).

To complete the proof of the statement, observe that $t \mapsto h(t)$ is bounded on $[t_*, t_1]$ and thus belongs to $L^\infty((t_*, t_1), \mathbb{R}^2)$. Then, for every $p \geq 1$,

$$\int_0^{+\infty} |\mathfrak{h}(s)|^p \, ds = \int_{t_*}^{t_1} |P^{-1}h|^p |h|^{-1} \, dt \leq \|P^{-1}\|^p \int_{t_*}^{t_1} |h|^{p-1} \, dt$$
$$\leq \|P^{-1}\|^p \|h\|_{L^\infty}^{p-1} (t_1 - t_*) < +\infty. \qquad \square$$

6.2. **Proof of Proposition 14 in the case** $\det A(t_1) < 0$**.** Since $\operatorname{trace}(A) = 0$ and $\det A(t_1) < 0$, there exists $P \in \mathrm{GL}(2, \mathbb{R})$ such that

$$PA(t_1)P^{-1} = \begin{pmatrix} -a & 0 \\ 0 & a \end{pmatrix}, \qquad a > 0.$$

Up to applying the change of coordinates associated with P and defining the time-rescaled curves $\mathfrak{h}$ and $\mathfrak{A}$ as in Lemma 15, we have

$$\mathfrak{A}(s) = \begin{pmatrix} -\alpha(s) & \beta(s) \\ \zeta(s) & \alpha(s) \end{pmatrix}, \tag{8}$$

where

$$\lim_{s\to\infty} \alpha(s) = a, \qquad \lim_{s\to\infty} \zeta(s) = \lim_{s\to\infty} \beta(s) = 0. \tag{9}$$

Let $\mathfrak{h} = \rho e^{i\vartheta}$ for $\rho > 0$ and $\vartheta \in [0, 2\pi)$. To prove the statement it is enough to show that ϑ has a limit as $s \to \infty$. Let us show that $\lim_{s\to\infty} \tan 2\vartheta = 0$.

Observe that, letting $\mathfrak{h} = (x_1, x_2)$ with $x_1, x_2 \in \mathbb{R}$, we have

$$\frac{1}{2} \tan 2\vartheta = \frac{\sin\vartheta \cos\vartheta}{\cos^2\vartheta - \sin^2\vartheta} = \frac{x_1 x_2}{x_2^2 - x_1^2}. \tag{10}$$

By (7) and simple computations we obtain

$$(x_1x_2)' = \zeta x_1^2 + \beta x_2^2, \qquad \frac{(x_1^2 - x_2^2)'}{2} = -\alpha(x_1^2 + x_2^2) + (\beta - \zeta) x_1 x_2. \tag{11}$$

Upon integration and exploiting (9), we get

$$x_1 x_2 = o(R), \quad x_1^2 - x_2^2 = 2aR(1 + o(1)), \quad \text{where} \quad R(s) := \int_s^{+\infty} |\mathfrak{h}|^2 \, ds.$$

Observe that, by Lemma 15, $\mathfrak{h} \in L^2((0, +\infty), \mathbb{R}^2)$ and, in particular, $R \to 0$ as $s \to +\infty$. Finally, substituting the above in (10) shows that $\tan 2\vartheta \to 0$, completing the proof of Proposition 14 in the case $\det A(t_1) < 0$.

6.3. **Proof of Proposition 14 in the case** $\det A(t_1) = 0$**.** Assume that $\det A(t_1) = 0$ and recall that $\operatorname{trace} A(t_1) = 0$. Since, moreover, $A(t_1) \neq 0$, there exists $P \in \mathrm{GL}(2, \mathbb{R})$ such that

$$PA(t_1)P^{-1} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}. \tag{12}$$

As before, using the change of variables of Lemma 15, we let

$$\mathfrak{A}(s) = \begin{pmatrix} -\alpha(s) & \beta(s) \\ \zeta(s) & \alpha(s) \end{pmatrix},$$

where α, β, ζ are linear combinations of $h_{2112} \circ \varphi^{-1}, h_{2212} \circ \varphi^{-1}$, and $h_{1112} \circ \varphi^{-1}$, and hence absolutely continuous with bounded derivatives on $(0, +\infty)$, according to (2). Equality (12) implies that $\alpha \to 0$, $\beta \to 1$, and $\zeta \to 0$ as $s \to +\infty$.

We also introduce $\mu := \zeta + \beta$ and we notice that $\mu \to 1$ as $s \to +\infty$. (Beware that the same letters are used for different parameters in the proof of Proposition 12.) Then, (7) reads

$$\frac{\rho'}{\rho} = \mu \sin\vartheta \cos\vartheta - \alpha \cos 2\vartheta, \qquad \vartheta' = -\mu \sin^2\vartheta + \alpha \sin 2\vartheta + \zeta,$$

and can be written as

$$\frac{\rho'}{\rho} = \sin\vartheta\cos\vartheta + f, \qquad \vartheta' = -\sin^2\vartheta + g, \tag{13}$$

where the functions

$$f = -\alpha\cos 2\vartheta + (1-\mu)\sin\vartheta\cos\vartheta, \qquad g = \alpha\sin 2\vartheta + \zeta + (1-\mu)\sin^2\vartheta,$$

tend to zero as $s \to +\infty$.

Lemma 16. *We have the following dichotomy:*

(i) $\sin\vartheta \to 0$ *as* $s \to +\infty$. *(That is,* ϑ *admits a limit as* $s \to +\infty$*, which is equal to* 0 *modulus* π*.)*

(ii) $\vartheta \to -\infty$ *as* $s \to +\infty$. *Moreover, in this case, for any* $0 < \varepsilon < \pi/2$ *there exist two sequences of positive real numbers* $(s_n)_{n\in\mathbb{N}}$ *and* $(\bar{s}_n)_{n\in\mathbb{N}}$ *tending to infinity and such that*

$$\bar{s}_{2n-1} < s_{2n} < s_{2n+1} < \bar{s}_{2n} < \bar{s}_{2n+1},$$
$$\vartheta(s_{2n}) = \pi - \varepsilon \mod 2\pi, \qquad \vartheta(s_{2n+1}) = \varepsilon \mod 2\pi,$$
$$\vartheta(\bar{s}_{2n}) = -\varepsilon \mod 2\pi, \qquad \vartheta(\bar{s}_{2n+1}) = \varepsilon - \pi \mod 2\pi,$$
$$\vartheta'(s) < 0 \quad \forall s \in [s_{2n}, s_{2n+1}] \cup [\bar{s}_{2n}, \bar{s}_{2n+1}].$$

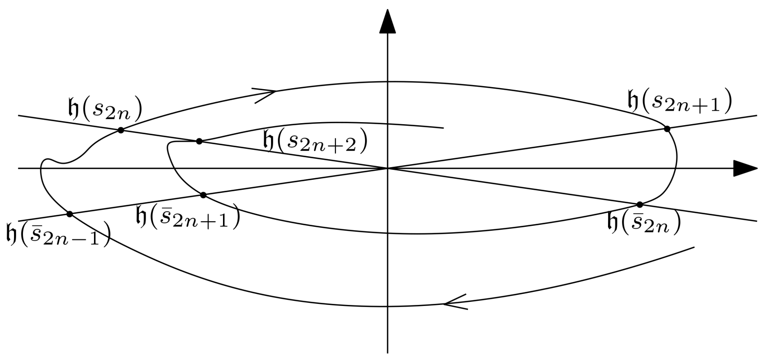

FIGURE 1. The sequences $(s_n)_{n\in\mathbb{N}}$ and $(\bar{s}_n)_{n\in\mathbb{N}}$ in Lemma 16

Proof. Notice that the dynamics of ϑ is a perturbation via g of

$$\vartheta' = -\sin^2\vartheta.$$

The phase portrait of the latter on $\mathbb{S}^1$ is made of two equilibria in 0 and π joined by two clock-wise oriented heteroclinic trajectories.

Assume that (i) does not hold. Therefore, there exists $c > 0$ such that

$$\limsup_{s\to+\infty} |\sin\vartheta(s)| > c. \tag{14}$$

Let $\varepsilon > 0$ be such that $\sin\varepsilon \in (0, c)$ and $s^* > 0$ be such that, for $s > s^*$, $\vartheta'(s) < -\varepsilon^2/2$ as soon as $|\sin\vartheta(s)| > \varepsilon$.

Pick $q_1 > s^*$ such that $|\sin\vartheta(q_1)| > c > \sin\varepsilon$. Since ϑ' is bounded from zero as long as $|\sin\vartheta|$ stays larger than $\sin\varepsilon$, there exists $r_1 > q_1$ such that $|\sin\vartheta(r_1)| = \sin\varepsilon$. By definition of c, there exists $q_2 > r_1$ such that $|\sin\vartheta(q_2)| > c$. Moreover, q_1 and q_2 can be taken so that $\vartheta(q_2) = \vartheta(q_1) - \pi$ and (14) holds with c arbitrarily close to 1. By iterating the procedure leading from q_1 to q_2, we prove that $\vartheta \to -\infty$. The construction also shows how to construct the sequences $(s_n)_{n\in\mathbb{N}}$ and $(\bar{s}_n)_{n\in\mathbb{N}}$ as in (ii). □

The rest of the argument consists in showing that case (ii) in Lemma 16 cannot hold true. For that purpose, we argue by contradiction.

Lemma 17. *Assume that property (ii) in Lemma 16 holds true. Then there exists* $0 < \varepsilon_0 < \pi/2$ *such that for any* $0 < \varepsilon < \varepsilon_0$ *there exists* N_ε *for which, given any* $n \geq N_\varepsilon$*,*

$$\frac{2}{\varepsilon}\left(1-\varepsilon^2\right) \leq s_{2n+1} - s_{2n} \leq \frac{2}{\varepsilon}\left(1+\varepsilon^2\right), \tag{15}$$

and

$$(1-\varepsilon)\varepsilon\rho(s_{2n}) \leq \rho(s)\sin\vartheta(s) \leq (1+\varepsilon)\varepsilon\rho(s_{2n}), \text{ for } s \in [s_{2n}, s_{2n+1}]. \tag{16}$$

As a consequence, for every $n \geq N_\varepsilon$*, one has the following estimates*

$$2(1-2\varepsilon)\rho(s_{2n}) \leq \int_{s_{2n}}^{s_{2n+1}} \sin\vartheta(s)\rho(s)\,ds \leq 2(1+2\varepsilon)\rho(s_{2n}), \tag{17}$$

$$(1-4\varepsilon)\frac{2\rho(s_{2n})}{\varepsilon} \leq \int_{s_{2n}}^{s_{2n+1}} \rho(s)\,ds \leq (1+4\varepsilon)\frac{2\rho(s_{2n})}{\varepsilon}, \tag{18}$$

$$\left|\int_{s_{2n}}^{s_{2n+1}} \cos\vartheta(s)\rho(s)\,ds\right| \leq \rho(s_{2n}). \tag{19}$$

Proof. Set $M_f(s) = \sup_{\tau\geq s}|f(\tau)|$ and $M_g(s) = \sup_{\tau\geq s}|g(\tau)|$. Observe that these two functions tend to zero as s tends to infinity.

By Lemma 16, for n large enough and $s \in [s_{2n}, s_{2n+1}]$, equation (13) becomes

$$(\cot\vartheta)' = 1 - \frac{g}{\sin^2\vartheta}. \tag{20}$$

For n large enough, for every $s \in [s_{2n}, s_{2n+1}]$ we have

$$\left|\frac{g(s)}{\sin^2\vartheta(s)}\right| \leq \frac{M_g(s_{2n})}{\sin^2\varepsilon} \leq \frac{\varepsilon^2}{2}.$$

Equation (15) follows by integrating (20) on the interval $[s_{2n}, s_{2n+1}]$.

On the interval $[s_{2n}, s_{2n+1}]$, one has

$$\frac{\rho'(s)}{\rho(s)} + \frac{\vartheta'(s)\cos\vartheta(s)}{\sin\vartheta(s)} = -f(s) - \frac{\cos\vartheta(s)g(s)}{\sin\vartheta(s)}.$$

For n large enough, for every $s \in [s_{2n}, s_{2n+1}]$ we have

$$\left|f(s) + \frac{\cos\vartheta(s)g(s)}{\sin\vartheta(s)}\right| \leq M_f(s_{2n}) + \frac{M_g(s_{2n})}{\sin\varepsilon} \leq \frac{\varepsilon^2}{4}. \tag{21}$$

By integrating between s_{2n} and any $s \in [s_{2n}, s_{2n+1}]$, one gets

$$\left|\ln\left(\frac{\rho(s)\sin\vartheta(s)}{\rho(s_{2n})\sin\varepsilon}\right)\right| = \left|\int_{s_{2n}}^{s}\left(f(s) + \frac{\cos\vartheta(s)g(s)}{\sin\vartheta(s)}\right)ds\right| \leq \frac{(s_{2n+1}-s_{2n})\varepsilon^2}{4} \leq \frac{\varepsilon}{2},$$

yielding (16) for ε small enough.

We now turn to the proof of the three estimates (17)–(19). The first one simply follows by integrating (16) on $[s_{2n}, s_{2n+1}]$ and using (15). Estimate (18) is obtained by first dividing (16) by $\sin\vartheta(s)$ and then integrating the resulting inequalities on $[s_{2n}, s_{2n+1}]$. One gets that

$$(1-\varepsilon)\varepsilon\rho(s_{2n})\int_{s_{2n}}^{s_{2n+1}}\frac{ds}{\sin\vartheta(s)} \leq \int_{s_{2n}}^{s_{2n+1}}\rho(s)\,ds \leq (1+\varepsilon)\varepsilon\rho(s_{2n})\int_{s_{2n}}^{s_{2n+1}}\frac{ds}{\sin\vartheta(s)}.$$

On the other hand, the following holds true,

$$\int_{s_{2n}}^{s_{2n+1}}\frac{ds}{\sin\vartheta(s)} = \int_{s_{2n}}^{s_{2n+1}}\frac{\vartheta'(s)}{-\sin^3\vartheta(s)(1-\frac{g(s)}{\sin^2\vartheta(s)})}\,ds,$$

which implies that

$$(1-\varepsilon^2)\int_{\varepsilon}^{\pi-\varepsilon}\frac{d\vartheta}{\sin^3\vartheta}\leq\int_{s_{2n}}^{s_{2n+1}}\frac{ds}{\sin\vartheta(s)}\leq(1+\varepsilon^2)\int_{\varepsilon}^{\pi-\varepsilon}\frac{d\vartheta}{\sin^3\vartheta}.$$

A direct computation shows that $\int_{\varepsilon}^{\pi-\varepsilon}\frac{d\vartheta}{\sin^3\vartheta}=\frac{1}{\varepsilon^2(1+o(\varepsilon))}$ as ε tends to zero. One finally deduces estimate (18).

To derive estimate (19), one notices that

$$\begin{aligned}\int_{s_{2n}}^{s_{2n+1}}\cos\vartheta(s)\rho(s)\,ds &= \int_{s_{2n}}^{s_{2n+1}}\frac{\sin\vartheta(s)\cos\vartheta(s)\rho(s)}{\sin\vartheta(s)}\,ds\\ &= \int_{s_{2n}}^{s_{2n+1}}\frac{\rho'(s)-f(s)\rho(s)}{\sin\vartheta(s)}\,ds\\ &= -\int_{s_{2n}}^{s_{2n+1}}\frac{f(s)\rho(s)}{\sin\vartheta(s)}\,ds+\frac{\rho(s_{2n+1})-\rho(s_{2n})}{\sin\varepsilon}\\ &\quad+\int_{s_{2n}}^{s_{2n+1}}\frac{\rho(s)\cos\vartheta(s)\vartheta'(s)}{\sin\vartheta(s)}\,ds.\end{aligned}$$

By using the expression of ϑ' in the last integral, one deduces that

$$2\int_{s_{2n}}^{s_{2n+1}}\cos\vartheta(s)\rho(s)\,ds=\frac{\rho(s_{2n+1})-\rho(s_{2n})}{\sin\varepsilon}-\int_{s_{2n}}^{s_{2n+1}}\rho(s)\frac{f(s)+\frac{\cos\vartheta(s)g(s)}{\sin\vartheta(s)}}{\sin\vartheta(s)}\,ds.$$

By using (16) for $s=s_{2n}$ and $s=s_{2n+1}$ and then (21), one deduces (19). □

Fix a sequence $(\varepsilon_k)_{k\in\mathbb{N}}$, strictly decreasing to 0. For each $k\in\mathbb{N}$, we use $(s_{k,n})_{n\in\mathbb{N}}$ to denote the sequence $(s_n)_{n\in\mathbb{N}}$ given by Lemma 16 and corresponding to $\varepsilon=\varepsilon_k$. For all $k\in\mathbb{N}$ let $n_k\geq N_{\varepsilon_k}$ be an integer to be fixed later, where N_{ε_k} is as in Lemma 17. We use $(\xi_\ell)_{\ell\in\mathbb{N}}$ to denote the sequence defined by

$$\xi_{2k}=s_{k,2n_k},\quad \xi_{2k+1}=s_{k,2n_k+1},\qquad \forall k\in\mathbb{N}.$$

We choose $k\mapsto n_k$ so that the sequence $(\xi_\ell)_{\ell\in\mathbb{N}}$ is strictly increasing and tends to infinity as $\ell\to+\infty$.

Let $t_\ell=\varphi^{-1}(\xi_\ell)$, where φ is the change of variables introduced in Lemma 15. For every $\ell\geq 0$ consider the function $u_\ell\in L^\infty([0,1],\mathbb{S}^1)$ defined by $u_\ell(\tau)=u(t_{2\ell}+\tau(t_{2\ell+1}-t_{2\ell}))$. By the weak-⋆ compactness of all bounded subsets of $L^\infty([0,1],\mathbb{R}^2)$, we can assume without loss of generality that $u_\ell\rightharpoonup u_\star$ in the weak-⋆ topology. Applying Lemma 8 with $a_\ell=t_{2\ell}$ and $b_\ell=t_{2\ell+1}$, we deduce that $u_\star$ is minimizing and $|u_\star|\equiv 1$ almost everywhere in $[0,1]$.

For every subinterval $[a,b]$ of $[0,1]$, by the properties of weak-⋆ convergence, we have that

$$\int_a^b v^T u_\ell(\tau)\,d\tau\to\int_a^b v^T u_\star(\tau)\,d\tau,\qquad \forall\, v\in\mathbb{R}^2.$$

Moreover, one has

$$\begin{aligned}\int_a^b v^T u_\ell(\tau)\,d\tau &= \frac{1}{t_{\ell+1}-t_\ell}\int_{(1-a)t_{2\ell}+at_{2\ell+1}}^{(1-b)t_{2\ell}+bt_{2\ell+1}}\frac{v^T h}{|h|}\,dt\\ &= \frac{1}{t_{\ell+1}-t_\ell}\int_{\varphi((1-a)t_{2\ell}+at_{2\ell+1})}^{\varphi((1-b)t_{2\ell}+bt_{2\ell+1})}v^T P\mathfrak{h}(s)\,ds,\end{aligned}\tag{22}$$

where P has been introduced in (12).

In addition

$$t_{\ell+1}-t_\ell=\int_{\xi_{2\ell}}^{\xi_{2\ell+1}}|P\mathfrak{h}(s)|\,ds.\tag{23}$$

Lemma 18. *Under the above assumptions, there exists a unit vector $v_\star\in\mathbb{R}^2$ such that $u_\star(t)=v_\star$ for a.e. $t\in[0,1]$. Moreover, $v_\star$ is parallel to $P(1,0)$.*

Proof. Let $v_\star, w_\star \in \mathbb{R}^2$ be two orthogonal unit vectors such that $v_\star$ is parallel to $P(1,0)$. Notice that $P^T w_\star$ is orthogonal to $(1,0)$, that is, it is parallel to $(0,1)$. We start by showing that $w_\star^T u_\star(t) = 0$ for a.e. $t \in [0,1]$. This amounts to show that for all $0 \le a < b \le 1$ it holds

$$\frac{1}{t_{\ell+1} - t_\ell} \int_{\varphi((1-a)t_{2\ell}+at_{2\ell+1})}^{\varphi((1-b)t_{2\ell}+bt_{2\ell+1})} \mathfrak{h}_2(s)\, ds \to 0 \qquad \text{as } \ell \to +\infty.$$

Since $\mathfrak{h}_2 = \rho \sin\vartheta$ is positive on $[\xi_{2\ell}, \xi_{2\ell+1}]$ by construction, using (23) it is enough to show that

$$\frac{\int_{\xi_{2\ell}}^{\xi_{2\ell+1}} \rho(s) \sin\vartheta(s)\, ds}{\int_{\xi_{2\ell}}^{\xi_{2\ell+1}} |P\mathfrak{h}(s)|\, ds} \longrightarrow 0 \qquad \text{as } \ell \to +\infty.$$

Since $|P\mathfrak{h}(s)| \ge \|P^{-1}\|^{-1}\rho(s)$ for all s, the latter limit holds true according to (17) and (18) in Lemma 17, applied to $\varepsilon = \varepsilon_\ell$ for $\ell \ge 0$.

Recall that the control $u_\star$ is minimizing and $|u_\star(t)| = 1$ for a.e. $t \in [0,1]$. From what precedes, one deduces that $u_\star$ is almost everywhere perpendicular to $w_\star$, hence equal to $v_\star$ or $-v_\star$. It then follows from Lemma 21 in the appendix that, up to replacing $v_\star$ by $-v_\star$, the equality $u_\star(t) = v_\star$ holds for a.e. $t \in [0,1]$. □

Let $\bar v \in \mathbb{R}^2$ be such that $P^T \bar v = (1,0)$. We have, according to Lemma 18,

$$\lim_{\ell\to\infty} \int_0^1 \bar v^T u_\ell(\tau)\, d\tau = \int_0^1 \bar v^T u_\star(\tau)\, d\tau = \bar v^T v_\star \neq 0.$$

We conclude the proof by contradiction by showing that the limit in the left-hand side is zero. Indeed, according to (22), we have

$$\left|\int_0^1 \bar v^T u_\ell(\tau)\, d\tau\right| = \frac{\left|\int_{\xi_{2\ell}}^{\xi_{2\ell+1}} \rho(s)\cos\vartheta(s)\, ds\right|}{\int_{\xi_{2\ell}}^{\xi_{2\ell+1}} |P\mathfrak{h}(s)|\, ds} \le \|P^{-1}\| \frac{\left|\int_{\xi_{2\ell}}^{\xi_{2\ell+1}} \rho(\tau)\cos\vartheta(\tau)\, d\tau\right|}{\int_{\xi_{2\ell}}^{\xi_{2\ell+1}} \rho(\tau)\, ds}.$$

The right-hand side of the above equation tends to zero thanks to (18) and (19) in Lemma 17 applied to $\varepsilon = \varepsilon_\ell$ for $\ell \ge 0$.

We have therefore proved that (ii) in Lemma 16 cannot hold true, which completes the proof of Proposition 14.

7. Proof of Theorem 1

Let M be as in the statement of Theorem 1. Denote, as in the previous sections, by $\gamma : [0,T] \to M$ a length-minimizing trajectory parameterized by arclength and by $\lambda : [0,T] \to T^*M$ an abnormal extremal lift of γ.

Proposition 14, together with Theorem 4, proves the C^1 regularity of γ provided that h vanishes only at isolated points.

We consider in this section the case where $t_0 \in (0,T)$ is a density point of $\{t \in [0,T] \mid h(t) = 0\}$. We want to prove that $u(t)$ (up to modification on a set of measure zero) has a limit as $t \uparrow t_0$ and as $t \downarrow t_0$. By symmetry, we restrict our attention to the existence of the limit of $u(t)$ as $t \uparrow t_0$.

We are going to consider separately the situations where $h \equiv 0$ on a left neighborhood of t_0 and where there exists a sequence of maximal open intervals (t_0^n, t_1^n) with $h|_{(t_0^n,t_1^n)} \neq 0$ and such that $t_1^n \to t_0$.

Assume for now on that $h \equiv 0$ on a left neighborhood $(t_0 - \eta, t_0]$ of t_0. Then, since $\dot h = Au$ almost everywhere on $(t_0 - \eta, t_0]$, we have that $u(t)$ belongs to $\ker A(t)$ for almost every t in $(t_0 - \eta, t_0]$. By Lemma 11, moreover, $\ker A(t)$ is one-dimensional for every $t \in (t_0 - \eta, t_0]$.

Fix an open neighborhood V_0 of $\lambda(t_0)$ in T^*M such that there exists a smooth map $V_0 \ni \lambda \mapsto v(\lambda) \in \mathbb{S}^1$ such that $v(\lambda(t)) \in \ker A(t)$ if $\lambda(t) \in V_0$ and $t \in (t_0 - \eta, t_0]$. Up to

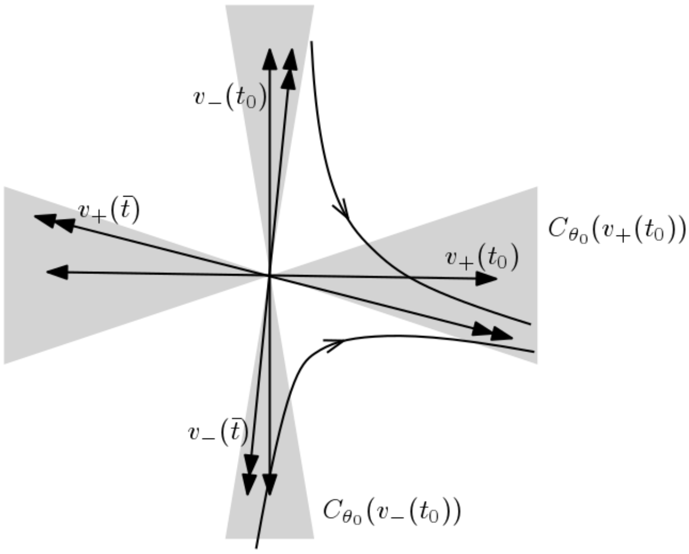

FIGURE 2. Phase portrait of $\dot x(t) = A(\bar t)x(t)$ for $\bar t \in [t_0 - \eta, t_0]$

reducing η, we assume that $\lambda(t) \in V_0$ for every $t \in (t_0 - \eta, t_0]$. Notice that $\lambda|_{(t_0-\eta,t_0]}$ is a solution of the time-varying system

$$\dot\lambda = \sigma(t)\vec X_{v(\lambda)}(\lambda),$$

where $\sigma : (t_0 - \eta, t_0] \to \{-1, 1\}$ is measurable. Hence, by length-minimality of γ and by Lemma 21 in the appendix, either $u = v$ almost everywhere on $(t_0 - \eta, t_0]$ or $u = -v$ almost everywhere on $(t_0 - \eta, t_0]$. We conclude that u is continuous on $(t_0 - \eta, t_0]$ and the proof in this case in concluded.

We are left to consider the case where every left neighborhood of t_0 contains a maximal interval (τ_0, τ_1) such that $h \neq 0$ on (τ_0, τ_1).

Notice that, by Proposition 12 and by continuity of $t \mapsto A(t)$, we have that $\det A(t_0) \leq 0$.

The case $\det A(t_0) < 0$ can be ruled out thanks to the following lemma.

Lemma 19. *Let* $\det A(t_0) < 0$. *Then there exists* $\eta \in (0, t_0)$ *such that, if* $0 \leq \tau_0 < \tau_1 \leq t_1$ *and* $x : (\tau_0, \tau_1) \to \mathbb{R}^2 \setminus \{0\}$ *is a maximal solution to* $\dot x = A\frac{x}{|x|}$, *then* $\tau_0 < t_0 - \eta$.

Proof. For every $v \in \mathbb{R}^2 \setminus \{0\}$ and every $\vartheta > 0$ denote by $C_\vartheta(v)$ the cone of all vectors in $\mathbb{R}^2 \setminus \{0\}$ making an (unoriented) angle smaller than ϑ with v or $-v$.

Let $\eta_0 \in (0, t_0)$ be such that $\det(A(t)) < 0$ for every $t \in [t_0 - \eta_0, t_0]$. For $t \in [t_0 - \eta_0, t_0]$, denote by $v_-(t)$ and $v_+(t)$ two eigenvectors of $A(t)$, the first corresponding to a negative and the second to a positive eigenvalue.

Let $\eta \in (0, \eta_0)$ and $\vartheta_0 > 0$ be such that $C_{\vartheta_0}(v_+(t_0)) \cap C_{\vartheta_0}(v_-(t_0)) = \emptyset$ and $v_\pm(t) \in C_{\vartheta_0}(v_\pm(t_0))$ for every $t \in [t_0 - \eta, t_0]$.

Notice that $C_{\vartheta_0}(v_+(t_0))$ is positively invariant for the dynamics of $\dot x(t) = A(\bar t)x(t)$ for every fixed $\bar t \in [t_0 - \eta, t_0]$ (see Figure 2).

Assume by contradiction that $x : (\tau_0, \tau_1) \to \mathbb{R}^2 \setminus \{0\}$ is a maximal solution of $\dot x = A\frac{x}{|x|}$ with $(\tau_0, \tau_1) \subset (t_0 - \eta, t_0)$. It follows from Proposition 14 that

$$\lim_{\tau \downarrow \tau_0} \frac{x(\tau)}{|x(\tau)|} \to \pm v_+(\tau_1), \quad \lim_{\tau \to \tau_1} \frac{x(\tau)}{|x(\tau)|} \to \pm v_-(\tau_1).$$

This contradicts the positive invariance of $C_{\vartheta_0}(v_+(t_0))$ for the equation $\dot x = A\frac{x}{|x|}$ on (τ_0, τ_1). □

In the case $\det A(t_0) = 0$ the proof follows the steps of the construction of Section 6.3. In particular, let $P \in \mathrm{GL}(2,\mathbb{R})$ be such that

$$P^{-1}A(t)P = \begin{pmatrix} -a(t) & 1-c(t) \\ c(t) & a(t) \end{pmatrix},$$

where a, b, c are affine combinations of h_{2112}, h_{2212}, and h_{1112} with $a \to 0$, and $c \to 0$ as $t \to t_0$. Let

$$P^{-1}h(t) = r(t)e^{i\omega(t)}, \tag{24}$$

with $\omega(t)$ uniquely defined modulus 2π only when $h(t) \neq 0$.

The crucial point is the following counterpart to Lemma 16, whose proof can be obtained using exactly the same arguments.

Lemma 20. *We have the following dichotomy:*

(i) *for any $0 < \varepsilon < \pi/2$, for η small enough, $|\sin(\omega(t))| < \varepsilon$ for all $t \in (t_0 - \eta, t_0)$ such that $h(t) \neq 0$;*

(ii) *for any $0 < \varepsilon < \pi/2$ there exists an increasing sequence $(t_n)_{n\in\mathbb{N}}$ in $(0, t_0)$ tending to t_0 and such that*

$$\omega(t_{2n}) = \pi - \varepsilon \mod 2\pi, \qquad \omega(t_{2n+1}) = \varepsilon \mod 2\pi,$$
$$h(t) \neq 0,\ \sin(\omega(t)) > 0,\ \dot\omega(t) < 0 \quad \forall t \in [t_{2n}, t_{2n+1}]$$

or

$$\omega(t_{2n}) = -\varepsilon \mod 2\pi n, \qquad \omega(t_{2n+1}) = \varepsilon - \pi \mod 2\pi n,$$
$$h(t) \neq 0,\ \sin(\omega(t)) < 0,\ \dot\omega(t) < 0 \quad \forall t \in [t_{2n}, t_{2n+1}]$$

holds true.

Case (ii) can be excluded by similar computations as in Section 6.3, since it contradicts the optimality of γ.

Consider now case (i). Let $v_\star, w_\star \in \mathbb{R}^2$ be two orthogonal unit vectors such that $v_\star$ is parallel to $P(1,0)$. According to (24), if $\sin(\omega(t)) = 0$ and $r(t) \neq 0$, then $u(t) = h(t)/|h(t)|$ is equal to $v_\star$ or $-v_\star$. For every $\eta \in (0, t_0)$ we set

$$I_\eta^+ = \{t \in (t_0 - \eta, t_0) \mid v_\star^T u(t) > 0\}, \qquad I_\eta^- = \{t \in (t_0 - \eta, t_0) \mid v_\star^T u(t) < 0\}.$$

Property (i) implies that, for η small, $I_\eta^+ \cup I_\eta^-$ contains $\{t \in (t_0 - \eta, t_0) \mid h(t) \neq 0\}$. Moreover, if t_0 is a density point for $I = I_\eta^+ \cap \{t \in (t_0 - \eta, t_0) \mid h(t) \neq 0\}$ (respectively, $I = I_\eta^- \cap \{t \in (t_0 - \eta, t_0) \mid h(t) \neq 0\}$), then,

$$\lim_{t\in I, t\to t_0} u(t) = v_\star \qquad \text{(respectively, } \lim_{t\in I, t\to t_0} u(t) = -v_\star).$$

Let $\Phi_\eta = \{t \in (t_0 - \eta, t_0) \mid h(t) = 0\}$. For almost every $t \in \Phi_\eta$, $u(t)$ is in the kernel of $A(t)$ and $|u(t)| = 1$. Notice that, if t_0 is a density point for $J = \{t \in (0, t_0) \mid \ker A(t) \neq (0)\}$, then the kernel of $A(t)$ converges to the kernel of $A(t_0)$ as $t \in J,\ t \to t_0$. By construction of P, moreover, $\ker(A(t_0)) = \mathrm{span}(P(1,0)) = \mathrm{span}(v_\star)$. Hence, for η small enough, almost every $t \in \Phi_\eta$ is in $I_\eta^+ \cup I_\eta^-$.

To summarize, for η small enough, $I_\eta^+ \cup I_\eta^-$ has full measure in $(t_0 - \eta, t_0)$. Moreover,

$$\lim_{t\in I_\eta^+,\, t\to t_0} u(t) = v_\star, \qquad \lim_{t\in I_\eta^-,\, t\to t_0} u(t) = -v_\star. \tag{25}$$

We next prove that u converges either to $v_\star$ or to $-v_\star$ as $t \to t_0$ by showing that, for η small enough, either I_η^+ or I_η^- has measure zero.

Suppose by contradiction that there exists a sequence of intervals (τ_0^n, τ_1^n) in $(0, t_0)$ such that $\tau_0^n, \tau_1^n \to t_0$ as $n \to \infty$ and both $|(\tau_0^n, \tau_1^n) \cap I^+|$ and $|(\tau_0^n, \tau_1^n) \cap I^-|$ are positive, where

$|\cdot|$ denotes the Lebesgue measure and $I^{\pm} = \{t \in (0, t_0) \mid \pm v_\star^T u(t) > 0\}$. Moreover, up to restricting (τ_0^n, τ_1^n), we can assume that

$$|(\tau_0^n, \tau_1^n) \cap I^+| = |(\tau_0^n, \tau_1^n) \cap I^-| > 0. \tag{26}$$

This can be seen, for instance, by considering a continuous deformation of an interval around a Lebesgue point of $(\tau_0^n, \tau_1^n) \cap I^+$ towards an interval around a Lebesgue point of $(\tau_0^n, \tau_1^n) \cap I^-$.

For every $n \in \mathbb{N}$, let $u_n \in L^\infty([0,1], \mathbb{R}^2)$ be defined by $u_n(\tau) = u(\tau_0^n + \tau(\tau_1^n - \tau_0^n))$. Up to extracting a subsequence, u_n weakly-$\star$ converges to some $u_\star$. Condition (26) and the limits in (25) imply that

$$\int_0^1 u_\star(t)dt = 0. \tag{27}$$

Thanks to (25) we also have that $w_\star^T u_n$ L^∞-converges to zero as $n \to \infty$. In particular, $w_\star^T u_\star \equiv 0$. By Lemma 8, $u_\star$ is optimal and $v_\star^T u_\star$ has values in $\{-1, 1\}$. Hence, by Lemma 21 in the appendix, $v_\star^T u_\star$ is constantly equal to $+1$ or -1. This contradicts (27) and the proof is concluded. □

Appendix A. An elementary lemma

Lemma 21. *Let (M, D, g) be a sub-Riemannian manifold. Let V be a Lipschitz continuous vector field on T^*M such that $\pi_* V(\lambda) \in D_{\pi(\lambda)} \setminus \{0\}$ for every $\lambda \in T^*M$. Let $\lambda : [0,T] \to T^*M$ satisfy $\dot\lambda(t) = \sigma(t)V(\lambda(t))$ with $\sigma \in L^\infty([0,T], [-1,1])$. Assume that $\gamma = \pi \circ \lambda : [0,T] \to M$ is a length-minimizer. Then σ has constant sign, i.e., either $\sigma \geq 0$ a.e. on $[0,T]$ or $\sigma \leq 0$ a.e. on $[0,T]$.*

Proof. Set $\kappa = \int_0^T \sigma(t)dt$ and notice that $\lambda(T) = e^{\kappa V}(\lambda(0))$. If σ does not have constant sign, then $[0,1] \ni t \mapsto \pi \circ e^{t\kappa V}(\lambda(0))$ is a curve connecting $\gamma(0)$ to $\gamma(T)$ and having length smaller than γ. □

A particular case of the lemma occurs when $V = \vec{H}$ is the Hamiltonian vector field on T^*M associated with the Hamiltonian $\lambda \mapsto \langle \lambda, X(\pi(\lambda))\rangle$, where X is a smooth horizontal never-vanishing vector field on M. This means that if a solution of $\dot\gamma(t) = \sigma(t)X(\gamma(t))$ is a length-minimizer then σ has constant sign.

References

[ABB17] A. Agrachev, D. Barilari, and U. Boscain. Introduction to Riemannian and sub-Riemannian geometry. Lecture notes available at http://webusers.imj-prg.fr/∼davide.barilari/Notes.php, preprint SISSA 09/2012/M. Version Nov 17, 2017.

[AGM15] L. Ambrosio, R. Ghezzi, and V. Magnani. BV functions and sets of finite perimeter in sub-Riemannian manifolds. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, 32(3):489–517, 2015.

[Agr14] A. A. Agrachev. Some open problems. In *Geometric control theory and sub-Riemannian geometry*, volume 5 of *Springer INdAM Ser.*, pages 1–13. Springer, Cham, 2014.

[AS95] A. A. Agrachëv and A. V. Sarychev. Strong minimality of abnormal geodesics for 2-distributions. *J. Dynam. Control Systems*, 1(2):139–176, 1995.

[AS99] A. A. Agrachev and A. V. Sarychev. Sub-Riemannian metrics: Minimality of abnormal geodesics versus subanalyticity. *ESAIM Control Optim. Calc. Var.*, 4, 1999.

[BdSRar] A. Belotto da Silva and L. Rifford. The Sard conjecture on Martinet surfaces. *Duke Math. J.*, to appear.

[Bel96] A. Bellaïche. The tangent space in sub-Riemannian geometry. In *Sub-Riemannian geometry*, volume 144 of *Progr. Math.*, pages 1–78. Birkhäuser, Basel, 1996.

[CJT06] Y. Chitour, F. Jean, and E. Trélat. Genericity results for singular curves. *J. Differential Geom.*, 73(1):45–73, 2006.

[HL16] E. Hakavuori and E. Le Donne. Non-minimality of corners in subriemannian geometry. *Invent. Math.*, pages 1–12, 2016.

[Jea14] F. Jean. *Control of Nonholonomic Systems: from Sub-Riemannian Geometry to Motion Planning*. Springer International Publishing, SpringerBriefs in Mathematics, 2014.

[LDLMV13] E. Le Donne, G. P. Leonardi, R. Monti, and D. Vittone. Extremal curves in nilpotent Lie groups. *Geom. Funct. Anal.*, 23(4):1371–1401, 2013.

[LDMO+16] E. Le Donne, R. Montgomery, A. Ottazzi, P. Pansu, and D. Vittone. Sard property for the end-point map on some Carnot groups. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, 33(6):1639–1666, 2016.

[LM08] G. P. Leonardi and R. Monti. End-point equations and regularity of sub-Riemannian geodesics. *Geom. Funct. Anal.*, 18(2):552–582, 2008.

[LS95] W. S. Liu and H. J. Sussmann. Shortest paths for sub-Riemannian metrics of rank two distributions. *Memoirs AMS*, 118(564), 1995.

[Mon02] R. Montgomery. *A tour of subriemannian geometries, their geodesics and applications*, volume 91 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 2002.

[Mon14] R. Monti. The regularity problem for sub-Riemannian geodesics. In *Geometric control theory and sub-Riemannian geometry*, volume 5 of *Springer INdAM Ser.*, pages 313–332. Springer, Cham, 2014.

[MPV17] R. Monti, A. Pigati, and D. Vittone. On tangent cones to length minimizers in Carnot–Carathéodory spaces. *Preprint CVGMT*, 2017.

[Rif17] L. Rifford. Singulières minimisantes en géométrie sous-Riemannienne. *Astérisque*, (390):Exp. No. 1113, 277–301, 2017. Séminaire Bourbaki. Vol. 2015/2016. Exposés 1104–1119.

[TY13] K. Tan and X. Yang. Subriemannian geodesics of Carnot groups of step 3. *ESAIM Control Optim. Calc. Var.*, 19(1):274–287, 2013.

[ZZ95] I. Zelenko and M. Zhitomirskii. Rigid paths of generic 2 -distributions on 3 -manifolds. *Duke Math. J.*, 79(2):281–307, 08 1995.

Univ. Paris Diderot, Institut de Mathématiques de Jussieu-Paris Rive Gauche, CNRS, Sorbonne Université. Bâtiment Sophie-Germain, case 7012, 75205 Paris cedex 13, France
E-mail address: davide.barilari@imj-prg.fr

L2S, CentraleSupélec, Gif-sur-Yvette, France
E-mail address: yacine.chitour@l2s.centralesupelec.fr

Unité de Mathématiques Appliquées, ENSTA ParisTech, Université Paris-Saclay, 91120 Palaiseau, France
E-mail address: frederic.jean@ensta-paristech.fr

CRNS, L2S, CentraleSupélec, Gif-sur-Yvette, France
E-mail address: dario.prandi@l2s.centralesupelec.fr

Inria & Laboratoire Jacques-Louis Lions, Sorbonne Université, 75005 Paris, France
E-mail address: mario.sigalotti@inria.fr