ESAIM: COCV 21 (2015) 625–634
DOI: 10.1051/cocv/2014041

ESAIM: Control, Optimisation and Calculus of Variations
www.esaim-cocv.org

CORNERS IN NON-EQUIREGULAR SUB-RIEMANNIAN MANIFOLDS

Enrico Le Donne[1], Gian Paolo Leonardi[2], Roberto Monti[3] and Davide Vittone[3]

Abstract. We prove that in a class of non-equiregular sub-Riemannian manifolds corners are not length minimizing. This extends the results of [G.P. Leonardi and R. Monti, *Geom. Funct. Anal.* **18** (2008) 552–582]. As an application of our main result we complete and simplify the analysis in [R. Monti, *Ann. Mat. Pura Appl.* (2013)], showing that in a 4-dimensional sub-Riemannian structure suggested by Agrachev and Gauthier all length-minimizing curves are smooth.

Mathematics Subject Classification. 53C17, 49K21, 49J15.

Received March 10, 2014.
Published online May 1, 2015.

1. Introduction

One of the major open problems in sub-Riemannian geometry is the regularity of length-minimizing curves. Indeed, no example of a non-smooth minimizer is known, and even the possibility of minimizers with singularities of corner-type has not yet been excluded in full generality (see Problem II in [1] and the discussion in Sect. 4 of [7]).

In [4], the second and third-named authors introduced a shortening technique specifically designed for showing the non-minimality of curves with corner-type singularities (see also the developments in [8]). This technique works for a class of equiregular sub-Riemannian manifolds satisfying the technical condition (1.2) below.

In this paper, we prove the non-minimality of corners in a class of sub-Riemannian manifolds of non-equiregular type. Namely, we show that if the horizontal distribution satisfies condition (1.3) below at the corner point, then the curve is not length minimizing. In this case, the construction of a competitor shorter than the corner is simpler than the one in [4] and relies on the Nagel−Stein−Wainger estimates [9] for the Carnot−Carathèodory distance.

Keywords and phrases. Sub-Riemannian geometry, regularity of geodesics, corners.

[1] University of Jyväskylä, Department of Mathematics and Statistics, P.O. Box 35, 40014 Jyväskylä, Finland. ledonne@msri.org

[2] Università di Modena e Reggio Emilia, Dipartimento di Scienze Fisiche, Informatiche e Matematiche, via Campi 213/b, 41100 Modena, Italy. gianpaolo.leonardi@unimore.it

[3] Università di Padova, Dipartimento di Matematica, via Trieste 63, 35121 Padova, Italy. monti@math.unipd.it; vittone@math.unipd.it

Article published by EDP Sciences

Let M be an n-dimensional differentiable manifold and let $\mathcal{D} \subset TM$ be a smooth subbundle of rank m, for some $2 \leq m \leq n$. Then, $\mathcal{D}(x) \subset T_xM$ is an m-dimensional subspace of the tangent space T_xM, for all $x \in M$. Let $X_1, \ldots, X_m$ be a frame of smooth vector fields that form a basis for $\mathcal{D}(x)$, that is $\mathcal{D}(x) = \operatorname{span}\{X_1(x), \ldots, X_m(x)\}$ for each $x \in M$. This frame always exists locally. The assumption that the fiber $\mathcal{D}(x)$ has a constant dimension on M plays no role in our argument and can be dropped. Here and in the following, the word "smooth" refers to C^∞ regularity.

We denote by $\mathcal{I} = \bigcup_{i \geq 1} \{1, \ldots, m\}^i$ the set of admissible multi-indices. For any $\beta = (\beta_1, \ldots, \beta_i) \in \mathcal{I}$, for some $i \geq 1$, let us define the iterated commutator

$$X_\beta = [X_{\beta_i}, [X_{\beta_{i-1}}, \ldots [X_{\beta_2}, X_{\beta_1}] \ldots]]. \tag{1.1}$$

We say that $\operatorname{length}(\beta) = i$ is the length of the multi-index β. Analogously, we say that $\operatorname{length}(X_\beta) = i$ is the length of the iterated commutator X_β. For any point $x \in M$ and $i \geq 1$, let

$$\mathcal{D}_i(x) = \operatorname{span}\Big\{ X_\beta(x) \in T_xM \,:\, \operatorname{length}(X_\beta) = i \Big\}.$$

Finally, we let $\mathcal{L}_i(x) = \mathcal{D}_1(x) + \ldots + \mathcal{D}_i(x)$ for $i \geq 1$, and we also agree that $\mathcal{L}_0(x) = \{0\}$. We assume that $\mathcal{D}$ is bracket generating, *i.e.*, for any $x \in M$ there exists an index $i \in \mathbb{N}$ such that $\mathcal{L}_i(x) = T_xM$.

An absolutely continuous curve $\gamma : [0,1] \to M$ is said to be *horizontal* with respect to the distribution $\mathcal{D}$ (or simply $\mathcal{D}$-horizontal) if there exist bounded measurable functions $h_1, \ldots, h_m : [0,1] \to \mathbb{R}$ such that

$$\dot\gamma(t) = \sum_{j=1}^m h_j(t)\, X_j(\gamma(t)), \quad \text{for almost every } t \in [0,1].$$

Let $g(x;\cdot)$ be a positive quadratic form (metric) on $\mathcal{D}(x)$, $x \in M$. The length of γ in the sub-Riemannian manifold $(M, \mathcal{D}, g)$ is defined as

$$L(\gamma) = \int_0^1 \sqrt{g(\gamma(t); \dot\gamma(t))}\, \mathrm{d}t,$$

and the *sub-Riemannian distance* between two points $x, y \in M$ is defined as

$$d(x,y) := \inf \Big\{ L(\gamma) :\ \gamma \in AC([0,1]; M) \text{ horizontal},\ \gamma(0) = x,\ \gamma(1) = y \Big\}.$$

When M is connected, the above set is always nonempty because the distribution $\mathcal{D}$ is bracket-generating, and d is a distance on M. Finally, we say that a horizontal curve γ joining x to y minimizes the sub-Riemannian length (*i.e.*, it is a length minimizer) if $L(\gamma) = d(x,y)$.

Let $\gamma : [0,1] \to M$ be a $\mathcal{D}$-horizontal curve. When they exist, we denote by $\dot\gamma_L(t)$ and $\dot\gamma_R(t)$ the left and right derivative of γ at the point $t \in (0,1)$. We say that γ has a corner at the point $x = \gamma(t) \in M$, if the left and right derivatives at t, do exist and are linearly independent. In [4], it is shown that if the distribution $\mathcal{D}$ is equiregular (*i.e.*, for every $i \geq 1$ the dimension of $\mathcal{D}_i(x)$ is constant on M) and satisfies the condition

$$[\mathcal{D}_i, \mathcal{D}_j] \subset \mathcal{L}_{i+j-1}, \quad \text{for } i,j \geq 2 \text{ such that } i + j > 4, \tag{1.2}$$

then corners in $(M, \mathcal{D}, g)$ are not length minimizing. In this paper, we prove that if the distribution $\mathcal{D}$ satisfies at some point $x \in M$ the condition

$$\mathcal{L}_i(x) \neq \mathcal{L}_{i-1}(x) \quad \Rightarrow \quad \mathcal{L}_{i+1}(x) = \mathcal{L}_i(x), \qquad \text{for all } i \geq 2, \tag{1.3}$$

then corners at x are not length minimizing.

Theorem 1.1. *Let $\gamma : [0,1] \to M$ be a horizontal curve with a corner at the point $x = \gamma(t) \in M$, for $t \in (0,1)$. If the distribution $\mathcal{D}$ satisfies (1.3) at x, then γ is not length minimizing in $(M, \mathcal{D}, g)$.*

The proof of Theorem 1.1, the main result of this paper, is presented in Section 2. After a blow-up argument, we can assume that $M = \mathbb{R}^n$, that $\mathcal{D}$ is a 2-dimensional distribution of planes in $\mathbb{R}^n$, and that $\gamma : [-1,1] \to \mathbb{R}^n$ is a corner at the point $0 \in \mathbb{R}^n$ of the type

$$\gamma(t) = \begin{cases} -tx & \text{if } t \in [-1,0] \\ ty & \text{if } t \in (0,1], \end{cases}$$

where $x, y \in \mathbb{R}^n$ are linearly independent. We prove the non-minimality of γ by an inductive argument on the dimension $n \geq 2$. In the inductive step, we use assumption (1.3) and known estimates on the sub-Riemannian distance to find a competitor shorter than the corner.

We found the basic idea of the Proof of Theorem 1.1 starting from a question raised by Agrachev and Gauthier during the meeting *Geometric control and sub-Riemannian geometry* held in Cortona in May 2012. They suggested the following situation, in order to find a *nonsmooth* length-minimizing curve. On the manifold $M = \mathbb{R}^4$, let Δ be the distribution of 2-planes spanned point-wise by the vector fields

$$X_1 = \frac{\partial}{\partial x_1} + 2x_2 \frac{\partial}{\partial x_3} + x_3^2 \frac{\partial}{\partial x_4}, \qquad X_2 = \frac{\partial}{\partial x_2} - 2x_1 \frac{\partial}{\partial x_3}. \tag{1.4}$$

The distribution Δ satisfies (1.3). We fix on Δ the quadratic form g making X_1 and X_2 orthonormal.

Let $\alpha > 0$ be a parameter and consider the initial and final points $y = (-1, \alpha, 0, 0) \in \mathbb{R}^4$ and $x = (1, \alpha, 0, 0) \in \mathbb{R}^4$, respectively. Agrachev and Gauthier asked whether the corner $\gamma : [-1,1] \to \mathbb{R}^4$ joining y to x

$$\gamma_1(t) = t, \quad \gamma_2(t) = \alpha|t|, \quad \gamma_3(t) = 0, \quad \gamma_4(t) = 0, \quad t \in [-1,1] \tag{1.5}$$

is, for small $\alpha > 0$, a length minimizer in $(\mathbb{R}^4, \Delta, g)$. The presence of the variable x_3 in the coefficients of the vector field X_1 in (1.4) is the technical obstruction for the application of the results of [4].

In [6], the curve γ in (1.5) was shown not to be length minimizing for $\alpha \neq 1$, by the explicit construction of a shorter competitor. This answered the above question in the negative. The case $\alpha = 1$, however, was left open.

In Section 3, as an application of Theorem 1.1, we prove the following result.

Theorem 1.2. *Let g be any smooth metric on Δ. In the sub-Riemannian manifold $(\mathbb{R}^4, \Delta, g)$ all length minimizing curves are smooth and, in particular, no corner is length minimizing.*

The Proof of Theorem 1.2 relies on Theorem 1.1. The inductive base is provided by the regularity of geodesics in the first Heisenberg group. This proof covers in particular the case $\alpha = 1$ in (1.5) and is simpler than the one in [6].

2. Proof of Theorem 1.1

The first step of the proof is a blow-up argument that closely follows [4].

Let $\gamma : [-1,1] \to M$ be a horizontal curve with a corner at the point $x = \gamma(0) \in M$. We can choose smooth and linearly independent vector fields $X_1, X_2 \in \mathcal{D}$ such that $X_1(x) = \dot{\gamma}_R(0)$ and $X_2(x) = -\dot{\gamma}_L(0)$ and we complete X_1, X_2 to a (local) frame $X_1, \dots, X_m$ for $\mathcal{D}$. Then we complete $X_1, \dots, X_m$ to a frame $X_1, \dots, X_n$ for TM in the following way. We choose iterated commutators $X_{m+1}, \dots, X_n \in \{X_\beta : \beta \in \mathcal{I}, \text{length}(\beta) \geq 2\}$ such that $X_1(x), \dots, X_n(x)$ are linearly independent. This choice is possible because $\mathcal{D}$ is bracket generating at x. We can also assume that $j \leq k$ implies $\text{length}(X_j) \leq \text{length}(X_k)$.

In a neighbourhood of $x \in M$, we fix exponential coordinates of the first type induced by the frame $X_1, \dots, X_n$ starting from x. Then we can identify M with $\mathbb{R}^n$, $X_1, \dots, X_n$ with vector fields on $\mathbb{R}^n$, and x with $0 \in \mathbb{R}^n$. The fact that we have exponential coordinates of the first type means that for $x = (x_1, \dots, x_n) \in \mathbb{R}^n$ (in fact, for x belonging to a neighbourhood of $0 \in \mathbb{R}^n$) we have

$$x = \exp\left(\sum_{i=1}^{n} x_i X_i\right)(0). \tag{2.1}$$

Here, the exponential mapping is defined by $\exp(X)(0) = \gamma(1)$ where γ is the solution of $\dot{\gamma} = X(\gamma)$ and $\gamma(0) = 0$.

We assign to the coordinate x_i the weight $w_i = \text{length}(X_i)$, $i = 1, \dots, n$. Then we have $w_1 = \ldots = w_m = 1$. The natural dilations on $\mathbb{R}^n$ adapted to the frame $X_1, \dots, X_n$ are

$$\delta_\lambda(x) = (\lambda^{w_1} x_1, \lambda^{w_2} x_2, \dots, \lambda^{w_n} x_n), \quad x \in \mathbb{R}^n, \ \lambda > 0. \tag{2.2}$$

Let $X = X_\beta$ be any iterated commutator of the vector fields $X_1, \dots, X_m$. Then we have

$$X = \sum_{i=1}^{n} a_i(x) \frac{\partial}{\partial x_i}, \tag{2.3}$$

where $a_i \in C^\infty(\mathbb{R}^n)$, $i = 1, \dots, n$, are smooth functions that have the structure described in the following proposition.

Proposition 2.1. *There exist polynomials $p_i : \mathbb{R}^n \to \mathbb{R}$ and functions $r_i : \mathbb{R}^n \to \mathbb{R}$, $i = 1, \dots, n$, such that:*

(i) $a_i(x) = p_i(x) + r_i(x)$, $x \in \mathbb{R}^n$;
(ii) $p_i(\delta_\lambda(x)) = \lambda^{w_i - \text{length}(X)} p_i(x)$;
(iii) $\lim_{\lambda \to \infty} \lambda^{w_i - \text{length}(X)} r_i(\delta_{1/\lambda}(x)) = 0$, $x \in \mathbb{R}^n$.

Proposition 2.1 can be proved as in ([5], p. 306). We omit the details, here. For $\lambda > 0$, we let

$$X^\lambda(x) = \sum_{i=1}^{n} \lambda^{w_i - \text{length}(X)} a_i(\delta_{1/\lambda}(x)) \frac{\partial}{\partial x_i}, \quad x \in \mathbb{R}^n. \tag{2.4}$$

The mapping $X \mapsto X^\lambda$ is bracket-preserving. Namely, for any multi-index $\beta \in \mathcal{I}$ and for $i = 1, \dots, m$ we have

$$[X_i, X_\beta]^\lambda = [X_i^\lambda, X_\beta^\lambda], \quad \lambda > 0. \tag{2.5}$$

We let $\mathcal{D}^\lambda = \text{span}\{X_1^\lambda, \dots, X_m^\lambda\}$, $\mathcal{D}_i^\lambda = \text{span}\{X_\beta^\lambda : \text{length}(\beta) = i\}$, and $\mathcal{L}_i^\lambda = \mathcal{D}_1^\lambda + \ldots + \mathcal{D}_i^\lambda$. By (2.5), from (1.3) we deduce that at the point $x = 0$ we have

$$\mathcal{L}_i^\lambda \neq \mathcal{L}_{i-1}^\lambda \quad \Rightarrow \quad \mathcal{L}_{i+1}^\lambda = \mathcal{L}_i^\lambda, \quad \text{for } i \geq 2. \tag{2.6}$$

By Proposition 2.1, for any iterated commutator $X = X_\beta$ as in (2.3), we can define the vector field X^∞ in $\mathbb{R}^n$

$$X^\infty(x) = \lim_{\lambda \to \infty} X^\lambda(x) = \sum_{i=1}^{n} p_i(x) \frac{\partial}{\partial x_i}, \quad x \in \mathbb{R}^n,$$

where p_i, $i = 1, \dots, n$, are polynomials such that $p_i \circ \delta_\lambda = \lambda^{w_i - \mathrm{length}(X)} p_i$. In particular, if $w_i < \mathrm{length}(X)$ then $p_i = 0$. Passing to the limit as $\lambda \to \infty$ in (2.5), we see that also the mapping $X \mapsto X^\infty$ is bracket-preserving, *i.e.*, $[X_i, X_\beta]^\infty = [X_i^\infty, X_\beta^\infty]$. Then at the point $x = 0$, condition (2.6) holds also for $\lambda = \infty$.

Let $g(x; \cdot)$ be a metric on $\mathcal{D}(x)$. On the distribution $\mathcal{D}^\lambda$, $\lambda > 0$, we introduce the metric $g^\lambda(x; \cdot)$ defined by

$$g^\lambda(x; X^\lambda) = g(\delta_{1/\lambda}(x); X), \quad x \in \mathbb{R}^n,$$

and on $\mathcal{D}^\infty = \mathrm{span}\{X_1^\infty, \dots, X_m^\infty\}$ we introduce the metric $g^\infty(x; \cdot)$ defined by

$$g^\infty(x; X^\infty) = \lim_{\lambda \to \infty} g^\lambda(x; X^\lambda) = g(0; X), \quad x \in \mathbb{R}^n.$$

We blow up the curve γ at the corner point $0 \in \mathbb{R}^n$. For $\lambda > 0$ and $t \in [-\lambda, \lambda]$, let $\gamma^\lambda(t) = \delta_\lambda \gamma(t/\lambda)$. Because $\dot\gamma_R(0) = X_1(0)$, $\dot\gamma_L(0) = -X_2(0)$, we obtain the limit curve $\gamma^\infty = \lim_{n \to \infty} \gamma^\lambda$,

$$\gamma^\infty(t) = \begin{cases} \mathrm{e}_1 t, \ t \in [0, 1] \\ -\mathrm{e}_2 t, \ t \in [-1, 0), \end{cases} \tag{2.7}$$

where $\mathrm{e}_1 = (1, 0, \dots, 0)$ and $\mathrm{e}_2 = (0, 1, 0, \dots, 0)$.

Proposition 2.2. *If the curve γ is length minimizing in $(M, \mathcal{D}, g)$ then the curve γ^∞ is length minimizing in $(\mathbb{R}^n, \mathcal{D}^\infty, g^\infty)$.*

Proposition 2.2 is proved in [4], Proposition 2.1. Our goal is to prove that the corner γ^∞ is not length minimizing in $(\mathbb{R}^n, \mathcal{D}^\infty, g^\infty)$. Thus, we can without loss of generality assume that $M = \mathbb{R}^n$, $\mathcal{D} = \mathcal{D}^\infty$, and $\gamma = \gamma^\infty$. Since γ is contained in the orbit of the distribution $\mathrm{span}\{X_1, X_2\}$, we can also assume that $m = 2$. Finally, we can pass to exponential coordinates of the second type associated with $X_1, \dots, X_n$. Namely, we can assume that for all $x = (x_1, \dots, x_n) \in \mathbb{R}^n$ belonging to a neighbourhood of $0 \in \mathbb{R}^n$ we have

$$x = \exp(x_1 X_1) \circ \dots \circ \exp(x_n X_n)(0).$$

Then we can also assume that X_1 and X_2 are vector fields in $\mathbb{R}^n$ of the form

$$X_1 = \frac{\partial}{\partial x_1} \quad \text{and} \quad X_2 = \frac{\partial}{\partial x_2} + \sum_{i=3}^n p_i(x) \frac{\partial}{\partial x_i}, \tag{2.8}$$

where $p_i : \mathbb{R}^n \to \mathbb{R}$, $i = 3, \dots, n$, are polynomials of the variable $x = (x_1, \dots, x_n) \in \mathbb{R}^n$ such that $p_i(\delta_\lambda(x)) = \lambda^{w_i - 1} p_i(x)$.

Condition (2.6) passes to the limit as $\lambda \to \infty$. Then, assumption (1.3) at the point $x = 0$ reads

$$\mathcal{L}_i(0) \neq \mathcal{L}_{i-1}(0) \quad \Rightarrow \quad \mathcal{L}_{i+1}(0) = \mathcal{L}_i(0), \quad \text{for all } i \geq 2. \tag{2.9}$$

Condition $\mathcal{L}_{i+1}(0) = \mathcal{L}_i(0)$ is equivalent to $\mathcal{D}_{i+1}(0) \subset \mathcal{L}_i(0)$. As $\mathcal{D} = \mathcal{D}^\infty$, (2.9) is equivalent to

$$\mathcal{L}_i(0) \neq \mathcal{L}_{i-1}(0) \quad \Rightarrow \quad \mathcal{D}_{i+1}(0) = \{0\}, \quad \text{for all } i \geq 2. \tag{2.10}$$

We prove this claim. Let X_β be an iterated commutator such that $\mathrm{length}(\beta) = i + 1$, *i.e.*, $X_\beta \in \mathcal{D}_{i+1}$. According to Proposition 2.1, we have

$$X_\beta = \sum_{j=1}^n p_j(x) \frac{\partial}{\partial x_j},$$

where p_j are polynomials satisfying $p_j(\delta_\lambda(x)) = \lambda^{w_j-i-1}p_j(x)$. Then the sum above ranges over indices j such that $w_j \geq i+1$. On the other hand, $\mathcal{L}_i(0) = \operatorname{span}\left\{\frac{\partial}{\partial x_j} : w_j \leq i\right\}$ and thus if $X_\beta(0) \in \mathcal{L}_i(0)$ we conclude that $X_\beta(0) = 0$. This proves (2.10).

For given indices $j, k = 1, \dots, n$, we say that $j \prec k$ if there exists $i \geq 2$ such that $X_j(0) \in \mathcal{D}_i(0)$ but $X_k(0) \notin \mathcal{L}_i(0)$. From (2.10), we deduce that the weights $w_1, \dots, w_n$ satisfy the following condition:

$$j \prec k \quad \Rightarrow \quad w_j + 2 \leq w_k. \tag{2.11}$$

We are ready to prove that the corner $\gamma = \gamma^\infty$ in (2.7) is not a length minimizer in $(\mathbb{R}^n, \mathcal{D}, g)$; this will give a contradiction and prove Theorem 1.1.

We can without loss of generality assume that X_1 and X_2 in (2.8) are orthonormal with respect to the metric g. This is because two different metrics are locally equivalent and the equivalence constants do not affect our estimates below. Then the length of a $\mathcal{D}$-horizontal curve $\gamma : [0,1] \to \mathbb{R}^n$ is

$$L(\gamma) = \int_0^1 \sqrt{\dot\gamma_1(t)^2 + \dot\gamma_2(t)^2}\mathrm{d}t. \tag{2.12}$$

The proof is by induction on the dimension n of $\mathbb{R}^n$. In order to fix the base of induction we distinguish two cases:

(1) We have $\mathcal{L}_2(0) = \mathcal{L}_1(0)$. In this case, the base of induction is $n = 2$. On $\mathbb{R}^2$ we have the standard Euclidean metric and corners are not length minimizing.
(2) We have $\mathcal{L}_2(0) \neq \mathcal{L}_1(0)$. In this case, the base of induction is $n = 3$. On $\mathbb{R}^3$ we have the Heisenberg group structure. We know that corners are not length minimizing for any sub-Riemannian metric in the Heisenberg group.

We assume that the claim holds for $n-1$ with $n \geq 3, 4$ in the two cases, and we prove it for n.

Let $\pi : \mathbb{R}^n \to \mathbb{R}^{n-1}$ be the projection $\pi(x_1, \dots, x_n) = (x_1, \dots, x_{n-1})$, and define the vector fields $\widehat{X}_1 = \pi_* X_1$ and $\widehat{X}_2 = \pi_* X_2$. Recall that for each $j = 3, \dots, n$ the polynomial p_j appearing in X_2 in (2.8) satisfies $p_j \circ \delta_\lambda = \lambda^{w_j-1}p_j$ and thus it depends only on the variables $x_1, \dots, x_{j-1}$. In particular, for each $j = 3, \dots, n$ the polynomial p_j does not depend on x_n. It follows that

$$\widehat{X}_1(x) = \frac{\partial}{\partial x_1}, \quad \widehat{X}_2(x) = \frac{\partial}{\partial x_2} + \sum_{j=3}^{n-1} p_j(x)\frac{\partial}{\partial x_j}, \quad \text{where } x = (x_1, \dots, x_{n-1}) \in \mathbb{R}^{n-1}.$$

We let $\widehat{\mathcal{D}} = \operatorname{span}\{\widehat{X}_1, \widehat{X}_2\}$ and we denote by $\widehat{g}$ the metric on $\widehat{\mathcal{D}}$ that makes $\widehat{X}_1$ and $\widehat{X}_2$ orthonormal. The distribution $\widehat{\mathcal{D}}$ satisfies (1.3).

The projection of the curve γ in (2.7) to $\mathbb{R}^{n-1}$, the curve $\widehat{\gamma} = \pi(\gamma) = (\gamma_1, \dots, \gamma_{n-1})$, is a corner at $0 \in \mathbb{R}^{n-1}$. By the inductive assumption, this curve is not length minimizing in $(\mathbb{R}^{n-1}, \widehat{\mathcal{D}}, \widehat{g})$. Then there exists a $\widehat{\mathcal{D}}$-horizontal curve $\widehat{\sigma} = (\widehat{\sigma}_1, \dots, \widehat{\sigma}_{n-1})$ in $\mathbb{R}^{n-1}$ joining the point $\mathrm{e}_2 \in \mathbb{R}^{n-1}$ to the point $\mathrm{e}_1 \in \mathbb{R}^{n-1}$ and satisfying

$$k := L(\widehat{\gamma}) - L(\widehat{\sigma}) > 0.$$

Let $\sigma := (\sigma_1, \dots, \sigma_n)$ be the $\mathcal{D}$-horizontal lift to $\mathbb{R}^n$ of the plane curve $(\widehat{\sigma}_1, \widehat{\sigma}_2)$ starting from the initial point e_2. Clearly, we have $\sigma_i = \widehat{\sigma}_i$ for $i = 1, \dots, n-1$ and

$$L(\sigma) = \int_0^1 \sqrt{\dot\sigma_1(t)^2 + \dot\sigma_2(t)^2}\mathrm{d}t = L(\widehat{\sigma}).$$

Finally, the end-point of σ is of the form $\mathrm{e}_1 + h\mathrm{e}_n \in \mathbb{R}^n$, for some $h \in \mathbb{R}$.

By our choice of the basis $X_1, \dots, X_n$, there exists a multi-index $\beta = (\beta_1, \dots, \beta_i) \in \mathcal{I}$, $i \geq 3$, such that $X_n = X_\beta$. Since we are in exponential coordinates and also using Proposition 2.1, we deduce that $X_n = \partial/\partial x_n$. Thus, we have

$$\frac{\partial}{\partial x_n} = X_\beta = [X_{\beta_i}, [X_{\beta_{i-1}}, \dots [X_{\beta_2}, X_{\beta_1}] \dots]]. \tag{2.13}$$

The integer $w_n = i$ is the length of the multi-index. We define the multi-index $\widehat{\beta} = (\beta_1, \dots, \beta_{i-1})$, that has length $i - 1 = w_n - 1$, and we define the corresponding iterated commutator

$$Z = X_{\widehat{\beta}} = [X_{\beta_{i-1}}, \dots [X_{\beta_2}, X_{\beta_1}] \dots] = \sum_{j=1}^{n} b_j(x) \frac{\partial}{\partial x_j},$$

where $b_j \in C^\infty(\mathbb{R}^n)$ are suitable functions, and, in fact, polynomials. By Proposition 2.1, these polynomials are homogeneous:

$$b_j(\delta_\lambda(x)) = \lambda^{w_j - w_n + 1} b_j(x), \quad x \in \mathbb{R}^n.$$

Thus, when $w_j - w_n + 1 < 0$ the polynomial b_j vanishes identically, $b_j = 0$, and the vector field Z has the form

$$Z = \sum_{w_j \geq w_n - 1} b_j(x) \frac{\partial}{\partial x_j}.$$

If $w_j = w_n - 1$ then $b_j(x)$ has homogeneous degree 0 and thus it is constant. On the other hand, we have $\partial/\partial x_n \in \mathcal{D}_i(0)$ and thus $\mathcal{D}_i(0) \neq \{0\}$. From (2.10) it follows that $\mathcal{L}_{i-1}(0) = \mathcal{L}_{i-2}(0)$, that is $\mathcal{D}_{i-1}(0) = \{0\}$. Because we have $Z \in \mathcal{D}_{i-1}$, then $Z(0) = 0$ and we conclude that $b_j = 0$ when $w_j = w_n - 1$ and Z is, in fact, of the form

$$Z = \sum_{w_j = w_n} b_j(x) \frac{\partial}{\partial x_j},$$

with $b_j(\delta_\lambda(x)) = \lambda b_j(x)$. Therefore we have

$$b_j(x) = c_{j1}\, x_1 + c_{j2}\, x_2 \tag{2.14}$$

for all j such that $w_j = w_n$, and for suitable constants c_{j1}, c_{j2}. Since the coefficients of X_2 (and X_1) in (2.8) do not contain the variables x_j such that $w_j = w_n$, we infer that

$$\frac{\partial}{\partial x_n} = [X_{\beta_i}, Z] = \sum_{w_j = w_n} \partial_{\beta_i} b_j(x) \frac{\partial}{\partial x_j},$$

and this implies that

$$\frac{\partial}{\partial x_{\beta_i}} b_n(x) = 1 \qquad \text{and} \qquad \frac{\partial}{\partial x_{\beta_i}} b_j(x) = 0, \quad j \neq n, \tag{2.15}$$

where either $\beta_i = 1$ or $\beta_i = 2$. We conclude that either $c_{n1} = 1$ or $c_{n2} = 1$ (or both).

Assume that $c_{n1} = 1$. The proof in the case $c_{n2} = 1$ is analogous. By our choice of the basis $X_1, \dots, X_n$, for any $j = 3, \dots, n-1$ there exists a multi-index $\beta^j \in \mathcal{I}$ such that

$$X_j = X_{\beta^j} = \frac{\partial}{\partial x_j} + \sum_{k=j+1}^{n} p_{jk}(x) \frac{\partial}{\partial x_k},$$

for suitable polynomials p_{jk}. Thus, at the point $x = \mathrm{e}_1 \in \mathbb{R}^n$, the vectors

$$X_1(x), X_2(x), \dots, X_{n-1}(x), Z(x)$$

are linearly independent, *i.e.*, they form a basis of $T_x\mathbb{R}^n$. In particular, the vector field Z is an iterated commutator of X_1 and X_2 with length $w_n - 1$. By the Nagel–Stein–Wainger estimate for the Carnot–Carathèodory distance (see [9] and, in particular, Thm. 4), there exist a neighbourhood U of $x = \mathrm{e}_1$ and a constant $C > 0$ such that

$$d(x, \exp(tZ)(x)) \leq Ct^{\frac{1}{w_n - 1}} \quad \text{for all } \exp(tZ)(x) \in U. \tag{2.16}$$

Let us fix a positive parameter $\epsilon > 0$ and let $(\gamma_1^\epsilon, \gamma_2^\epsilon)$ be the planar curve obtained by the concatenation of the following three curves: the line segment from $(0,1)$ to $(0,\epsilon)$, the curve $(\epsilon\sigma_1, \epsilon\sigma_2)$, and the line segment from $(\epsilon, 0)$ to $(1,0)$. When $a = 0$ and $b = 1$ we consider the same curve but starting from $(1,0)$. Let $\gamma^\epsilon = (\gamma_1^\epsilon, \dots, \gamma_n^\epsilon)$ be the $\mathcal{D}$-horizontal lift of this curve to $\mathbb{R}^n$, starting from the point e_2 (starting from e_1, when $a = 0$ and $b = 1$). Notice that the $\mathcal{D}$-horizontal lift of $(\epsilon\sigma_1, \epsilon\sigma_2)$ is the curve $\delta_\epsilon \circ \sigma = (\epsilon^{w_1}\sigma_1, \dots, \epsilon^{w_n}\sigma_n)$, by (2.2), and hence the end-point of $\delta_\epsilon \circ \sigma$ is the point $\epsilon \mathrm{e}_1 + \epsilon^{w_n} h \mathrm{e}_n$. Moving along the vector field X_1 does not change the nth coordinate x_n, hence we conclude that the final point of γ^ϵ is $x^\epsilon = \mathrm{e}_1 + \epsilon^{w_n} h \mathrm{e}_n$. By (2.14) and (2.15) we have $\mathrm{e}_1 + \epsilon^{w_n} h \mathrm{e}_n = \exp(\epsilon^{w_n} hZ)(\mathrm{e}_1)$. Since $x^\epsilon \to x = \mathrm{e}_1$ as $\epsilon \to 0$, by (2.16) and for ϵ small enough we have

$$d(x, x^\epsilon) \leq Ch^{\frac{1}{w_n - 1}} \epsilon^{\frac{w_n}{w_n - 1}}. \tag{2.17}$$

The sub-Riemannian length in $(\mathbb{R}^n, \mathcal{D}, g)$ of γ^ϵ is

$$\begin{aligned} L(\gamma^\epsilon) &= (1 - \epsilon) L(\gamma) + L(\delta_\epsilon \circ \sigma) \\ &= (1 - \epsilon) L(\gamma) + \epsilon L(\sigma) \\ &= L(\gamma) - \epsilon(L(\gamma) - L(\sigma)) \\ &= L(\gamma) - \epsilon(L(\widehat{\gamma}) - L(\widehat{\sigma})) \\ &= L(\gamma) - \epsilon k. \end{aligned} \tag{2.18}$$

Thus, from (2.17) and (2.18) we obtain (below we let $y = \mathrm{e}_2$)

$$\begin{aligned} d(y, x) &\leq d(y, x^\epsilon) + d(x^\epsilon, x) \\ &\leq L(\gamma^\epsilon) + C\epsilon^{\frac{w_n}{w_n - 1}} h^{\frac{1}{w_n - 1}} \\ &= L(\gamma) - \epsilon k + C\epsilon^{\frac{w_n}{w_n - 1}} h^{\frac{1}{w_n - 1}}. \end{aligned}$$

Since $k > 0$, there exists an $\epsilon > 0$ such that $Ch^{\frac{1}{w_n - 1}} \epsilon^{\frac{1}{w_n - 1}} < k/2$ and hence

$$d(x, y) < L(\gamma) - \epsilon k/2 < L(\gamma).$$

This proves that γ is not length minimizing in $(\mathbb{R}^n, \mathcal{D}, g)$. This also concludes the proof by induction of Theorem 1.1.

3. Proof of Theorem 1.2

Let $\Delta = \mathrm{span}\{X_1, X_2\}$ be the distribution of planes in $\mathbb{R}^4$ spanned by the vector fields X_1 and X_2 in (1.4). We fix the metric g on Δ making X_1, X_2 an orthonormal frame for Δ. Length minimizers for the sub-Riemannian distance are *extremals* in the sense of Geometric Control Theory, *i.e.*, they satisfy certain necessary conditions given by Pontryagin Maximum Principle. Extremals may be either normal or abnormal. Normal extremals are always smooth. The following proposition classifies abnormal nonsmooth extremals.

Proposition 3.1. *In the structure $(\mathbb{R}^4, \Delta)$, the only nonsmooth abnormal extremals are the curves*

$$\gamma(t) = \begin{cases} (-tx_1, -tx_2, 0, a) & \text{if } t \in [-1, 0] \\ (ty_1, ty_2, 0, a) & \text{if } t \in (0, 1], \end{cases} \tag{3.1}$$

where $a \in \mathbb{R}$ and $(x_1, x_2), (y_1, y_2) \in \mathbb{R}^2$ are linearly independent.

Proof. Let $\gamma : [0,1] \to \mathbb{R}^4$ be an abnormal extremal of the distribution Δ. By Pontryagin Maximum Principle, there exists an absolutely continuous curve $\xi : [0,1] \to \mathbb{R}^4$ solving almost everywhere the system of differential equations

$$\dot{\xi} = \Big(2\dot{\gamma}_2\xi_3,\ -2\dot{\gamma}_1\xi_3,\ -2\gamma_3\dot{\gamma}_1\xi_4,\ 0\Big). \tag{3.2}$$

See, *e.g.* ([2], Thm. 2.1) for a formulation of Pontryagin Maximum Principle. Moreover, we have $\langle X_1(\gamma), \xi\rangle = \langle X_2(\gamma), \xi\rangle = \langle [X_1, X_2], \xi\rangle = 0$, where $\langle\cdot,\cdot\rangle$ is the standard scalar product of $\mathbb{R}^4$. The last equation $\langle [X_1, X_2], \xi\rangle = 0$ is Goh condition, that holds automatically true in the rank 2 case (see, *e.g.*, [10]). Namely, the curve ξ also solves the system of equations

$$\begin{aligned} &\xi_1 + 2\gamma_2\xi_3 + \gamma_3^2\xi_4 = 0\\ &\xi_2 - 2\gamma_1\xi_3 = 0\\ &\xi_3 - \gamma_1\gamma_3\xi_4 = 0. \end{aligned} \tag{3.3}$$

From (3.2), we see that that ξ_4 is constant. This constant is nonzero, otherwise (3.3) would imply $\xi = 0$, and this is not possible for abnormal extremals. By linearity we can assume that $\xi_4 = 1$, and thus (3.3) trasforms into the system

$$\xi = \Big(-2\gamma_1\gamma_2\gamma_3 - \gamma_3^2, 2\gamma_1^2\gamma_3, \gamma_1\gamma_3, 1\Big), \tag{3.4}$$

and the system (3.2) becomes

$$\dot{\xi} = \Big(2\gamma_1\gamma_3\dot{\gamma}_2, -2\gamma_1\gamma_3\dot{\gamma}_1, -2\gamma_3\dot{\gamma}_1, 0\Big). \tag{3.5}$$

Differentiating the second equation in (3.4), we find $\dot{\xi}_2 = 4\gamma_1\gamma_3\dot{\gamma}_1 + 2\gamma_1^2\dot{\gamma}_3$, and comparing with the second equation in (3.5), we deduce that $\gamma_1^2(3\dot{\gamma}_1\gamma_3 + \gamma_1\dot{\gamma}_3) = 0$. This in turn implies that the function $\phi(t) = \gamma_1(t)^3\gamma_1(t)$ is a constant $c \in \mathbb{R}$.

Now there are two cases.

First case. $c = 0$. In this case, the equation $\gamma_1^3\gamma_3 = 0$ implies that γ is either a line or a corner of the form (3.1).

Second case. $c \neq 0$. In this case, by differentiating the identity $\gamma_3 = c/\gamma_1^3$ and using the horizontality condition $\dot{\gamma}_3 = 2\dot{\gamma}_1\gamma_2 - 2\dot{\gamma}_2\gamma_1$, we deduce that

$$\Big\langle (3c + 2\gamma_1^4\gamma_2, -2\dot{\gamma}_1^5), (\dot{\gamma}_1, \dot{\gamma}_2)\Big\rangle = 0,$$

where $\langle\cdot,\cdot\rangle$ is the standard scalar product of $\mathbb{R}^2$. In other words, the planar curve (γ_1, γ_2) is, up to reparameterization, an integral curve of the vector field in the plane

$$\frac{\partial}{\partial x_1} + \frac{3c + 2x_1^4x_2}{2x_1^5}\frac{\partial}{\partial x_2}, \quad x_1 \neq 0.$$

Thus the curve γ is

$$\gamma(t) = \Big(t, bt - \frac{3}{10}ct^{-4}, ct^{-3}, -\frac{1}{5}c^2t^{-5} + d\Big), \qquad \text{with } t \neq 0, \tag{3.6}$$

for some $b, c, d \in \mathbb{R}$. All such curves are C^∞.

We conclude that the nonsmooth abnormal extremals in $(\mathbb{R}^4, \Delta)$ are precisely the corners (3.1). □

Remark 3.2. In the proof of Proposition 3.1, we have the formula (3.4) for the dual curve ξ of an abnormal extremal γ. The coordinates of ξ are polynomial functions of the coordinates of γ. This is analogous to the results obtained in [2,3] for stratified nilpotent groups. In such groups, dual curves can be reconstructed using a special family of polynomials, called *extremal polynomials*, and abnormal extremals are always contained in the level sets of extremal polynomials.

We conclude with the Proof of Theorem 1.2.

Proof of Theorem 1.2. Thanks to Proposition 3.1, it is enough to prove the non-minimality of corners in $(\mathbb{R}^4, \Delta)$ at $x = 0$. Since the distribution Δ satisfies the assumption (1.3) at $x = 0$, we can use Theorem 1.1 and obtain the desired conclusion. □

References

[1] A. Agrachev, Some open problems, Geometric Control Theory and sub-Riemannian Geometry. Edited by G. Stefani, U. Boscain, J.-P. Gauthier, A. Sarychev, M. Sigalotti. Vol. 5 of *Springer INdAM Series* (2014) 1–14.

[2] E. Le Donne, G.P. Leonardi, R. Monti and D. Vittone, Extremal Curves in Nilpotent Lie Groups. *Geom. Funct. Anal.* **23** (2013) 1371–1401.

[3] E. Le Donne, G.P. Leonardi, R. Monti and D. Vittone, Extremal polynomials in stratified groups. Preprint `ArXiv:1307.5235` (2013).

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

[5] G.A. Margulis and G.D. Mostow, Some remarks on the definition of tangent cones in a Carnot-Carathéodory space. *J. Anal. Math.* **80** (2000) 299–317.

[6] R. Monti, A family of nonminimizing abnormal curves. *Ann. Mat. Pura Appl.* (2013).

[7] R. Monti, The regularity problem for sub-Riemannian geodesics, Geometric Control Theory and sub-Riemannian Geometry. Edited by G. Stefani, U. Boscain, J.-P. Gauthier, A. Sarychev, M. Sigalotti. Vol. 5 of *Springer INdAM Series* (2014) 313–332.

[8] R. Monti, Regularity results for sub-Riemannian geodesics. *Calc. Var. Partial Differ. Eqs.* **49** (2014) 549–582.

[9] A. Nagel, E.M. Stein and S. Wainger, Balls and metrics defined by vector fields. I. Basic properties. *Acta Math.* **155** (1985) 103–147.

[10] D. Vittone, The regularity problem for sub-Riemannian geodesics. Preprint (2013). Available at `http://cvgmt.sns.it/`.