The regularity problem for sub-Riemannian geodesics

Roberto Monti

Abstract We review some recent results on the regularity problem of sub-Riemannian length minimizing curves. We also discuss a new nontrivial example of singular extremal that is not length minimizing near a point where its derivative is only Hölder continuous. In the final section, we list some open problems.

1 Introduction

One of the main open problems in sub-Riemannian geometry is the regularity of length minimizing curves, see [12, Problem 10.1]. All known examples of length minimizing curves are smooth. On the other hand, there is no regularity theory of a general character for sub-Riemannian geodesics.

It was originally claimed by Strichartz in [15] that length minimizing curves are smooth, all of them being normal extremals. The wrong argument relied upon an incorrect application of Pontryagin Maximum Principle, ingnoring the possibility of *abnormal* (also called *singular*) extremals. In 1994 Montgomery discovered the first example of a singular length minimizing curve [11]. In fact, manifolds with distributions of rank 2 are rich of abnormal geodesics: in [9], Liu and Sussmann introduced a class of abnormal extremals, called *regular abnormal extremals*, that are always locally length minimizing. On the other hand, when the rank is at least 3 the situation is different. In [4], Chitour, Jean, and Trélat showed that for a generic distribution of rank at least 3 every singular curve is of minimal order and of corank 1. As a corollary, they show that a generic distribution of rank at least 3 does not admit (nontrivial) minimizing singular curves.

The question about the regularity of length minimizing curves remains open. The point, of course, is the regularity of abnormal minimizers. Some partial results in

R. Monti (✉)
Dipartimento di Matematica, Università di Padova, via Trieste 63, 35121 Padova
e-mail: monti@math.unipd.it

G. Stefani, U. Boscain, J.-P. Gauthier, A. Sarychev, M. Sigalotti (eds.): *Geometric Control Theory and Sub-Riemannian Geometry*, Springer INdAM Series 5, DOI 10.1007/978-3-319-02132-4_18,

this direction are obtained in [8] and [13]. In this survey, we describe these and other recent results. In Sect. 5.2, we present the classification of abnormal extremals in Carnot groups [6], that was announced at the meeting *Geometric control and sub-Riemannian geometry* held in Cortona in May 2012. The example of nonminimizing singular curve of Sect. 7 is new.

We refer the reader to the monograph [2] for an excellent introduction to Geometric Control Theory, see also the book in preparation [1].

2 Basic facts

Let M be an n-dimensional smooth manifold, $n \geq 3$, let $\mathcal{D}$ be a completely non-integrable (i.e., bracket generating) distribution of r-planes on M, $r \geq 2$, called *horizontal distribution*, and let $g = g_x$ be a smooth quadratic form on $\mathcal{D}(x)$, varying smoothly with $x \in M$. The triple $(M, \mathcal{D}, g)$ is called *sub-Riemannian manifold*.

A Lipschitz curve $\gamma : [0, 1] \to M$ is $\mathcal{D}$-*horizontal*, or simply *horizontal*, if $\dot\gamma(t) \in \mathcal{D}(\gamma(t))$ for a.e. $t \in [0, 1]$. We can then define the length of γ

$$L(\gamma) = \Big(\int_0^1 g_{\gamma(t)}(\dot\gamma(t))\, dt \Big)^{1/2}.$$

For any couple of points $x, y \in M$, we define the function

$$d(x, y) = \inf \Big\{ L(\gamma) : \gamma \text{ is horizontal, } \gamma(0) = x \text{ and } \gamma(1) = y \Big\}. \tag{1}$$

If the above set is nonempty for any $x, y \in M$, then d is a distance on M, usually called Carnot-Carathéodory distance.

By construction, the metric space (M, d) is a *length space*. If this metric space is complete, then closed balls are compact, and by a standard application of Ascoli-Arzelà theorem, the infimum in (1) is attained. Namely, for any given pair of points $x, y \in M$ there exists at least one Lipschitz curve $\gamma : [0, 1] \to M$ joining x to y and such that $L(\gamma) = d(x, y)$. This curve, which in general is not unique, is called a *length minimizing* curve. Its a priori regularity is the Lipschitz regularity. In particular, length minimizing curves are differentiable a.e. on $[0, 1]$.

For our porpouses, we can assume that M is an open subset of $\mathbb{R}^n$ or the whole $\mathbb{R}^n$ itself, and that we have $\mathcal{D}(x) = \operatorname{span}\{X_1(x), \dots, X_r(x)\}$, $x \in \mathbb{R}^n$, where $X_1, \dots, X_r$ are $r \geq 2$ linearly independent smooth vector fields in $\mathbb{R}^n$. With respect to the standard basis of vector fields in $\mathbb{R}^n$, we have, for any $j = 1, \dots, r$,

$$X_j = \sum_{i=1}^n X_{ji} \frac{\partial}{\partial x_i}, \tag{2}$$

where $X_{ji} : \mathbb{R}^n \to \mathbb{R}$ are smooth functions. A Lipschitz curve $\gamma : [0, 1] \to M$ is then horizontal if there exists a vector of functions $h = (h_1, \dots, h_r) \in$

$L^\infty([0,1];\mathbb{R}^r)$, called controls of γ, such that

$$\dot\gamma = \sum_{j=1}^{r} h_j X_j(\gamma), \quad \text{a.e. on } [0,1].$$

We fix on $\mathcal{D}(x)$ the quadratic form g_x that makes $X_1,\dots,X_r$ orthonormal. Any other choice of metric does not change the regularity problem. In this case, the length of γ is

$$L(\gamma) = \Big(\int_0^1 |h(t)|^2 dt\Big)^{1/2}.$$

Let $h = (h_1,\dots,h_r)$ be the controls of a horizontal curve γ. When γ is length minimizing we call the pair (γ, h) an *optimal pair*. Pontryagin Maximum Principle provides necessary conditions for a horizontal curve to be a minimizer.

Theorem 1. *Let (γ, h) be an optimal pair. Then there exist $\xi_0 \in \{0,1\}$ and a Lipschitz curve $\xi : [0,1] \to \mathbb{R}^n$ such that:*

i) $\xi_0 + |\xi| \neq 0$ *on* $[0,1]$;
ii) $\xi_0 h_j + \langle \xi, X_j(\gamma)\rangle = 0$ *on* $[0,1]$ *for all* $j = 1,\dots,r$;
iii) *the coordinates ξ_k, $k = 1,\dots,n$, of the curve ξ solve the system of differential equations*

$$\dot\xi_k = -\sum_{j=1}^{r}\sum_{i=1}^{n} \frac{\partial X_{ji}}{\partial x_k}(\gamma) h_j \xi_i, \quad \text{a.e. on } [0,1]. \tag{3}$$

Above, $\langle \xi, X_j\rangle$ is the standard scalar product of ξ and X_j as vectors of $\mathbb{R}^n$. If we identify the curve ξ with the 1-form in $\mathbb{R}^n$ along γ

$$\xi = \xi_1 dx_1 + \ldots + \xi_n dx_n,$$

then $\langle \xi, X_j\rangle$ is the covector-vector duality.

The proof of Theorem 1 relies upon the open mapping theorem, see [2, Chap. 12]. For any $v \in L^2([0,1];\mathbb{R}^r)$, let γ^v be the solution of the problem

$$\dot\gamma^v = \sum_{j=1}^{r} v_j X_j(\gamma^v), \quad \gamma^v(0) = x_0.$$

The mapping $\mathcal{E} : L^2([0,1];\mathbb{R}^r) \to \mathbb{R}^n$, $\mathcal{E}(v) = \gamma^v(1)$, is called the *end-point mapping* with initial point x_0. The *extended end-point mapping* is the mapping $\mathcal{F} : L^2([0,1];\mathbb{R}^r) \to \mathbb{R}^{n+1}$

$$\mathcal{F}(v) = \Big(\int_0^1 |v|^2 dt, \mathcal{E}(v)\Big).$$

If (γ, h) is an optimal pair with $\gamma(0) = x_0$ then $\mathcal{F}$ is not open at $v = h$ and then its differential is not surjective. It follows that there exists a nonzero vector $(\lambda_0, \lambda) \in \mathbb{R}\times\mathbb{R}^n = \mathbb{R}^{n+1}$ such that for all $v \in L^2([0,1];\mathbb{R}^r)$ there holds $\langle d\mathcal{F}(h)v, (\lambda_0,\lambda)\rangle =$

0. The case $\lambda_0 = 0$ is the case of abnormal extremals, that are precisely the critical points of the end-point mapping $\mathcal{E}$, i. e., points h where the differential $d\mathcal{E}(h)$ is not surjective. In particular, the notion of abnormal extremal is independent of the metric fixed on the horizontal distribution.

The curve ξ, sometimes called *dual curve* of γ, is obtained in the following way. Let h be the controls of an optimal trajectory γ starting from x_0. For $x \in \mathbb{R}^n$, let γ_x be the solution to the problem

$$\dot{\gamma}_x = \sum_{j=1}^{r} h_j X_j(\gamma_x) \quad \text{and} \quad \gamma_x(0) = x.$$

The *optimal flow* is the family of mappings $P_t : \mathbb{R}^n \to \mathbb{R}^n$, $P_t(x) = \gamma_x(t)$ with $t \in \mathbb{R}$. We are assuming that the flow is defined for any $t \in \mathbb{R}$. Let $(\lambda_0, \lambda) \in \mathbb{R} \times \mathbb{R}^n$ be a vector orthogonal to the image of $d\mathcal{F}(h)$. At the point x_0 we have the 1-form $\xi(0) = \lambda_1 dx_1 + \ldots + \lambda_n dx_n$, where $(\lambda_1, \ldots, \lambda_n)$ are the coordinates of λ. Then the curve $t \mapsto \xi(t)$ given by the pull-back of $\xi(0)$ along the optimal flow at time t, namely the curve

$$\xi(t) = P_{-t}^*(x_0)\xi(0), \tag{4}$$

satisfies the *adjoint Eq.* (3).

We can use i)–iii) in Theorem 1 to define the notion of *extremal*. We say that a horizontal curve $\gamma : [0, 1] \to \mathbb{R}^n$ is an *extremal* if there exist $\xi_0 \in \{0, 1\}$ and $\xi \in \mathrm{Lip}([0, 1]; \mathbb{R}^n)$ such that i), ii), and iii) in Theorem 1 hold. We say that γ is a *normal extremal* if there exists such a pair (ξ_0, ξ) with $\xi_0 \neq 0$. We say that γ is an *abnormal extremal* if there exists such a pair with $\xi_0 = 0$. We say that γ is a *strictly abnormal extremal* if γ is an abnormal extremal but not a normal one.

If γ is an abnormal extremal with dual curve ξ, then by ii) we have, for any $j = 1, \ldots, r$,

$$\langle \xi, X_j(\gamma) \rangle = 0 \quad \text{on } [0, 1]. \tag{5}$$

Further necessary conditions on abnormal extremals can be obtained differentiating identity (5). In fact, one gets for any $j = 1, \ldots, r$,

$$\sum_{i=1}^{r} h_i \langle \xi, [X_i, X_j](\gamma) \rangle = 0 \quad \text{a.e. on } [0, 1]. \tag{6}$$

When the rank is $r = 2$, from (6) along with the free assumption $|h| \neq 0$ a.e. on $[0, 1]$ we deduce that

$$\langle \xi, [X_1, X_2](\gamma) \rangle = 0 \quad \text{on } [0, 1]. \tag{7}$$

In the case of strictly abnormal minimizers, necessary conditions analogous to (7) can be obtained also for $r \geq 3$.

Theorem 2. *Let $\gamma : [0, 1] \to \mathbb{R}^n$ be a strictly abnormal length minimizer. Then any dual curve $\xi \in \mathrm{Lip}([0, 1], \mathbb{R}^n)$ of γ atisfies*

$$\langle \xi, [X_i, X_j](\gamma) \rangle = 0 \quad \text{on } [0, 1] \tag{8}$$

for any $i, j = 1, \ldots, r$.

Condition (8) is known as *Goh condition*. Theorem 2 can be deduced from second order open mapping theorems. We refer to [2, Chap. 20] for a systematic treatment of the subject. See also the work [3].

The Goh condition naturally leads to the notion of Goh extremal. A horizontal curve $\gamma : [0, 1] \to \mathbb{R}^n$ is a *Goh extremal* if there exists a Lipschitz curve $\xi : [0, 1] \to \mathbb{R}^n$ such that $\xi \not\equiv 0$, ξ solves the adjoint Eq. (3) and $\langle \xi, X_i(\gamma) \rangle = \langle \xi, [X_i, X_j](\gamma) \rangle = 0$ on $[0, 1]$ for all $i, j = 1, \ldots, r$.

3 Known regularity results

In this section, we collect some regularity results for extremal and length minimzing curves. Other results are discussed in Sect. 4. The case of normal extremal is clear and classical.

Theorem 3. *Let $(M, \mathcal{D}, g)$ be any sub-Riemannian manifold. Normal extremals are C^∞ curves that are locally length minimzing.*

In fact, with the notation of Sect. 2, if γ is a normal extremal with controls h and dual curve ξ, by condition ii) in Theorem 1 we have, for any $j = 1, \ldots, r$,

$$h_j = -\langle \xi, X_j(\gamma) \rangle \quad \text{a.e. on } [0, 1]. \tag{9}$$

This along with the adjoint Eq. (3) implies that the pair (γ, ξ) solves a.e. the system of Hamilton's equations

$$\dot\gamma = \frac{\partial H}{\partial \xi}(\gamma, \xi), \quad \dot\xi = -\frac{\partial H}{\partial x}(\gamma, \xi), \tag{10}$$

where H is the Hamiltonian function

$$H(x, \xi) = -\frac{1}{2} \sum_{j=1}^{r} \langle \xi, X_j(x) \rangle^2.$$

This implies that $\dot\gamma$ and $\dot\xi$ are Lipschitz continuous and thus $\gamma, \xi \in C^{1,1}$. By iteration, one deduces that $\gamma, \xi \in C^\infty$.

The fact that normal extremals are locally length minimizing follows by a calibration argument, see [9, Appendix C]. Indeed, using the Hamilton's Eq. (10), the

1-form ξ along γ can be locally extended to an *exact* 1-form ξ satisfying

$$\sum_{j=1}^{r} \langle \xi, X_j(x) \rangle^2 = 1.$$

This 1-form provides the calibration.

The distribution $\mathcal{D} = \operatorname{span}\{X_1, \dots, X_r\}$ on M is said to be *bracket-generating of step* 2 if for any $x \in M$ we have

$$\dim\big(\operatorname{span}\{X_j(x), [X_i, X_j](x) : i, j = 1, \dots, r\}\big) = n, \tag{11}$$

where $n = \dim(M)$. For distributions of step 2, Goh condition (8) implies the smoothness of any minimizer.

Theorem 4. *Let $(M, \mathcal{D}, g)$ be a sub-Riemannian manifold where $\mathcal{D}$ is a distribution that is bracket generating of step* 2. *Then any length minimizing curve in $(M, \mathcal{D}, g)$ is of class C^∞.*

In fact, if γ is a strictly abnormal length minimizing curve with dual curve ξ then by (5), (8), and (11) it follows that $\xi = 0$ and this is not possible. In other words, there are no strictly abnormal minimizers and this implies the claim made in Theorem 4.

When the step of the distribution is at least 3, then there can exist strictly abnormal extremals. When the step is precisely 3, the regularity question is clear within the setting of Carnot groups. Let $\mathfrak{g}$ be a stratified nilpotent n-dimensional real Lie algebra with

$$\mathfrak{g} = \mathfrak{g}_1 \oplus \dots \oplus \mathfrak{g}_s, \quad s \geq 2,$$

where $\mathfrak{g}_{i+1} = [\mathfrak{g}_1, \mathfrak{g}_i]$ for $i \leq s-1$ and $\mathfrak{g}_i = \{0\}$ for $i > s$.

The Lie algebra $\mathfrak{g}$ is the Lie algebra of a connected and simply-connected Lie group G that is diffeomorphic to $\mathbb{R}^n$. Such a Lie group is called *Carnot group*. The horizontal distribution $\mathcal{D}$ on G is induced by the first layer $\mathfrak{g}_1$ of the Lie algebra. In fact, $\mathcal{D}$ is spanned by a system of r linearly independent left-invariant vector fields. By nilpotency, the distribution is bracket-generating. So any quadratic form on $\mathfrak{g}_1$ induces a left-invariant sub-Riemannian metric on G. The number $r = \dim(\mathfrak{g}_1)$ is the *rank* of the group. The number $s \geq 2$ is the *step* of the group.

Theorem 5. *Let G be a Carnot group of step $s = 3$ with a smooth left-invariant quadratic form g on the horizontal distribution $\mathcal{D}$. Any length minimizing curve in $(G, \mathcal{D}, g)$ is of class C^∞.*

This theorem is proved in [16]. A short and alternative proof, given in [6, Theorem 6.1], relies upon the fact that a strictly abnormal length minimizing curve must be contained in (the lateral of) a proper Carnot subgroup. Then a reduction argument on the rank of the group reduces the analysis to the case $r = 2$, where abnormal extremals are easily shown to be integral curves of some horizontal left-invariant vector field.

When the step is $s = 4$, there is a regularity result only for Carnot groups of rank $r = 2$, see [8, Example 4.6].

Theorem 6. *Let G be a Carnot group of step $s = 4$ and rank $r = 2$ with a smooth left invariant quadratic form g on the horizontal distribution $\mathcal{D}$. Then any length minimizing curve in $(G, \mathcal{D}, g)$ is of class C^∞.*

The proof of this result relies upon two facts. First, one proves that the horizontal coordinates of any abnormal extremal are contained in the zero set of a quadratic polynomial in two variables. This shows that the only singularity that abnormal extremals can have is of corner type. Then using a general theorem proved in [8] (see Sect. 4) one concludes that extremal curves with corners are not length minimizing.

When the rank is $r = 2$ and the step s is larger than 4, the best regularity known for minimizers is the $C^{1,\delta}$ regularity.

Theorem 7. *Let G be a Carnot group of rank $r = 2$, step $s > 4$ and with Lie algebra $\mathfrak{g} = \mathfrak{g}_1 \oplus \ldots \oplus \mathfrak{g}_s$ satisfying*

$$[\mathfrak{g}_i, \mathfrak{g}_j] = 0 \quad \text{for all } i, j \geq 2 \text{ such that } i + j > 4. \tag{12}$$

Then any length minimizing curve in $(G, \mathcal{D}, g)$, where g is a smooth left-invariant metric on the horizontal distribution $\mathcal{D}$, is of class $C^{1,\delta}$ for any

$$0 \leq \delta < \min\left\{\frac{2}{s-4}, \frac{1}{4}\right\}. \tag{13}$$

This theorem is proved in [37, Theorem 10.1]. It is a byproduct of a technique that is used to analyse the length minimality properties of extremals of class C^1 whose derivative is only δ-Hölder continuos for some $0 < \delta < 1$. We give an example of such techniques in Sect. 7. The restriction $\delta < 2/(s-4)$ is a technical one. The estimates developed in [13], however, show that the restriction $\delta < 1/4$ is deeper. We shall discuss (12) it in the next section.

4 Analysis of corner type singularities

Let M be a smooth manifold with dimension $n \geq 3$, and let $\mathcal{D}$ be a completely non-integrable distribution on M. Let $\mathcal{D}_1 = \mathcal{D}$ and $\mathcal{D}_i = [\mathcal{D}_1, \mathcal{D}_{i-1}]$ for $i \geq 2$, i. e., $\mathcal{D}_i$ is the linear span of all commutators $[X, Y]$ with $X \in \mathcal{D}_1$ and $Y \in \mathcal{D}_{i-1}$. We also let $\mathcal{L}_0 = \{0\}$ and $\mathcal{L}_i = \mathcal{D}_1 + \ldots + \mathcal{D}_i$, $i \geq 1$. By the nonintegrability condition, for any $x \in M$ there exists $s \in \mathbb{N}$ such that $\mathcal{L}_s(x) = T_x M$, the tangent space of M at x. Assume that $\mathcal{D}$ is equiregular, i. e., assume that for each $i = 1, \ldots, s$

$$\dim\big(\mathcal{L}_i(x)/\mathcal{L}_{i-1}(x)\big) \text{ is constant for } x \in M. \tag{14}$$

In [8], Leonardi and the author proved the following theorem.

Theorem 8. *Let* $(M, \mathcal{D}, g)$ *be a sub-Riemannian manifold, where* g *is a metric on the horizontal distribution* $\mathcal{D}$. *Assume that* $\mathcal{D}$ *satisfies* (14) *and*

$$[\mathcal{L}_i, \mathcal{L}_j] \subset \mathcal{L}_{i+j-1}, \quad i, j \geq 2,\ i + j > 4. \tag{15}$$

Then any curve in M *with a corner is not length minimizing in* $(M, \mathcal{D}, g)$.

A "curve with a corner" is a $\mathcal{D}$-horizontal curve $\gamma : [0, 1] \to M$ such that at some point $t \in (0, 1)$ the left and right derivatives $\dot{\gamma}_L(t) \neq \dot{\gamma}_R(t)$ exist and are different. The proof of Theorem 8 is divided into several steps.

1) First one blows up the manifold M, the distribution $\mathcal{D}$, the metric g, and the curve γ at the corner point $x = \gamma(t)$. The blow-up is in the sense of the nilpotent approximation of Mitchell, Margulis and Mostow (see e. g. [10]). The limit structure is a Carnot group and the limit curve is the union of two half-lines forming a corner.

2) The limit curve is actually contained in a subgroup of rank 2, and after a suitable choice of coordinates one can assume that the manifold is $M = \mathbb{R}^n$ with a 2-dimensional distribution $\mathcal{D} = \text{span}\{X_1, X_2\}$ spanned by the vector fields in $\mathbb{R}^n$

$$X_1 = \frac{\partial}{\partial x_1} \quad \text{and} \quad X_2 = \frac{\partial}{\partial x_2} + \sum_{j=3}^{n} f_j(x) \frac{\partial}{\partial x_j}, \tag{16}$$

where $f_j : \mathbb{R}^n \to \mathbb{R}$, $j = 3, \ldots, n$, are polynomials with certain properties. The curve obtained after the blow-up is $\gamma : [-1, 1] \to \mathbb{R}^n$

$$\gamma(t) = \begin{cases} -t\mathrm{e}_2, \ t \in [-1, 0], \\ t\mathrm{e}_1, \ \ t \in [0, 1], \end{cases} \tag{17}$$

where $\mathrm{e}_1, \ldots, \mathrm{e}_n$ is the standard basis of $\mathbb{R}^n$. If the limit curve is not length minimizing in the limit structure, then the original curve is not length minimizing in the original structure.

3) At this stage, one uses (15). If the original distribution satisfies (15), then the limit Lie algebra satisfies (12) and the polynomials f_j only depends on the variables x_1 and x_2. This makes possible an effective and computable way to prove that the curve γ in (17) is not length minimizing. One cuts the corner of γ in the x_1x_2 plane gaining some length. The new planar curve must be lifted to get a horizontal curve, changing in this way the end-point. One can use several different devices to bring the end-point back to its original position. To do this, we can use a total amount of length that is less than the length gained by the cut. This adjustment is in fact possible, and the entire construction is the main achievement of [8].

The restriction (15) has a technical character. The problem of dropping this restriction is adressed in [14] (see also Sect. 6.2). The cut-and-adjust technique introduced in [8] is extended in [13] to the analysis of curves having singularities of higher order. In Sect. 7, we study a nontrivial example of such a situation.

5 Classification of abnormal extremals

The notion of abnormal extremal is rather indirect or implicit. There is a differential equation, the differential Eq. (3), involving the dual curve and the controls of the extremal. Even though this equation can be translated into some better form (see Theorem 2.6 in [6]), nevertheless the carried information is not transparent. In this section, we present some attempts to describe abnormal extremals is a more geometric or algebraic way.

5.1 Rank 2 distributions

We consider first the case when $M = \mathbb{R}^n$ and $\mathcal{D}$ is a rank 2 distribution in M spanned by vector fields X_1 and X_2 as in (16), where $f_3, \ldots, f_n \in C^\infty(\mathbb{R}^2)$ are functions depending on the variables x_1, x_2. We fix on $\mathcal{D}$ the quadratic form g making X_1 and X_2 orthonormal. Let $K : \mathbb{R}^{n-2} \times \mathbb{R}^2 \to \mathbb{R}$ be the function

$$K(\lambda, x) = \sum_{i=1}^{n-2} \lambda_i \frac{\partial f_{i+2}}{\partial x_1}(x), \tag{18}$$

where $\lambda = (\lambda_1, \ldots, \lambda_{n-2}) \in \mathbb{R}^{n-2}$ and $x \in \mathbb{R}^2$.

In this special situation, Pontryagin Maximum Principle can be rephrased in the following way (see Propositions 4.2 and 4.3 in [8]).

Theorem 9. *Let $\gamma : [0, 1] \to M$ be a $\mathcal{D}$-horizontal curve that is length minimizing in $(M, \mathcal{D}, g)$. Let $\kappa = (\gamma_1, \gamma_2)$ and assume that $|\dot\kappa| = 1$ almost everywhere. Then one (or both) of the following two statements holds:*

1) *there exists $\lambda \in \mathbb{R}^{n-2}$, $\lambda \neq 0$, such that*

$$K(\lambda, \kappa(t)) = 0, \quad \text{for all } t \in [0, 1]; \tag{19}$$

2) *the curve γ is smooth and there exists $\lambda \in \mathbb{R}^{n-2}$ such that κ solves the system of differential equations*

$$\ddot\kappa = K(\lambda, \kappa)\dot\kappa^\perp, \tag{20}$$

where $\kappa^\perp = (-\kappa_2, \kappa_1)$.

The geometric meaning of the curvature Eq. (20) was already noticed by Montgomery in [11].

The interesting case in Theorem 9 is the case 1): the curve κ, i. e., the horizontal coordinates of γ, is in the zero set of a nontrivial explicit function.

5.2 Stratified nilpotent Lie groups

In free stratified nilpotent Lie groups (free Carnot groups) there is an algebraic characterization of extremal curves in terms of an algebraic condition analogous to (19).

Let G be a *free nilpotent* Lie group with Lie algebra $\mathfrak{g}$. Fix a Hall basis $X_1, \ldots, X_n$ of $\mathfrak{g}$ and assume that the Lie algebra is generated by the first r elements $X_1, \ldots, X_r$.

We refer to [5] for a precise definition of the Hall basis. The basis determines a collection of *generalized structure constants* $c_{i\alpha}^k \in \mathbb{R}$, where $\alpha = (\alpha_1, \ldots, \alpha_n) \in \mathbb{N}^n$ is a multi-index and $i, k \in \{1, \ldots, n\}$. These constants are defined via the identity

$$[X_i, X_\alpha] = \sum_{k=1}^{n} c_{i\alpha}^k X_k, \tag{21}$$

where the iterated commutator X_α is defined via the relation

$$[X_i, X_\alpha] = [X_i, [\underbrace{X_1, \ldots, [X_1}_{\alpha_1 \text{ times}}, \ldots, \underbrace{[X_n, \ldots, X_n]}_{\alpha_n \text{ times}} \ldots] \ldots]]. \tag{22}$$

Using the constants $c_{i\alpha}^k$, for any $i = 1, \ldots, n$ and for any multi-index $\alpha \in \mathbb{N}^n$, we define the linear mappings $\phi_{i\alpha} : \mathbb{R}^n \to \mathbb{R}$

$$\phi_{i\alpha}(v) = \frac{(-1)^{|\alpha|}}{\alpha!} \sum_{k=1}^{n} c_{i\alpha}^k v_k, \quad v = (v_1, \ldots, v_n) \in \mathbb{R}^n. \tag{23}$$

Finally, for each $i = 1, \ldots, n$ and $v \in \mathbb{R}^n$, we introduce the polynomials $P_i^v : \mathbb{R}^n \to \mathbb{R}$

$$P_i^v(x) = \sum_{\alpha \in \mathbb{N}^n} \phi_{i\alpha}(v) x^\alpha, \quad x \in \mathbb{R}^n, \tag{24}$$

where we let $x^\alpha = x_1^{\alpha_1} \cdots x_n^{\alpha_n}$.

The group G can be identified with $\mathbb{R}^n$ via exponential coordinates of the second type induced by the basis $X_1, \ldots, X_n$. For any $v \in \mathbb{R}^n$, $v \neq 0$, we call the set

$$Z_v = \left\{ x \in \mathbb{R}^n : P_1^v(x) = \ldots = P_r^v(x) = 0 \right\}$$

an *abnormal variety* of G *of corank* 1. For linearly independent vectors $v_1, \ldots, v_m \in \mathbb{R}^n$, $m \geq 2$, we call the set $Z_{v_1} \cap \ldots \cap Z_{v_m}$ an *abnormal variety* of G *of corank* m. Recall that the property of having corank m for an abnormal extremal γ means that the range of the differential of the end-point map at the extremal curve is $n - m$ dimensional.

The main result of [6] is the following theorem.

Theorem 10. *Let $G = \mathbb{R}^n$ be a free nilpotent Lie group and let $\gamma : [0, 1] \to G$ be a horizontal curve with $\gamma(0) = 0$. The following statements are equivalent:*

A) *the curve γ is an abnormal extremal of corank $m \geq 1$;*
B) *there exist m linearly independent vectors $v_1, \ldots, v_m \in \mathbb{R}^n$ such that $\gamma(t) \in Z_{v_1} \cap \ldots \cap Z_{v_m}$ for all $t \in [0, 1]$.*

A stronger version of Theorem 10 holds for Goh extremals. If $\mathfrak{g} = \mathfrak{g}_1 \oplus \mathfrak{g}_2 \oplus \cdots \oplus \mathfrak{g}_s$, we let $r_1 = \dim(\mathfrak{g}_1)$ and $r_2 = \dim(\mathfrak{g}_2)$. Then, for $v \in \mathbb{R}^n$ with $v \neq 0$ we define the zero set

$$\Gamma_v = \left\{ x \in \mathbb{R}^n : P_i^v(x) = 0 \text{ for all } i = r_1 + 1, \ldots, r_1 + r_2 \right\}.$$

Theorem 11. *Let $G = \mathbb{R}^n$ be a free nilpotent Lie group and let $\gamma : [0, 1] \to G$ be a horizontal curve such that $\gamma(0) = 0$. The following statements are equivalent:*

A) *the curve γ is a Goh extremal;*
B) *there exists $v \in \mathbb{R}^n$, $v \neq 0$, such that $\gamma(t) \in \Gamma_v$ for all $t \in [0, 1]$.*

The zero set Γ_v is always nontrivial for $v \neq 0$ and, moreover, there holds $v_i = 0$ for all $i = 0, \ldots, r_1 + r_2$. See Remark 4.12 in [6].

These results are obtained via an explicit integration of the adjoint Eq. (3). Some work in progress [7] shows that Theorems 10 and 11 also hold in *nonfree* stratified nilpotent Lie groups.

6 Some examples

In this section, we present two examples. In the first one, we exhibit a Goh extremal having no regularity beyond the Lipschitz regularity. In the second example, there are extremals with corner in a sub-Riemannian manifold violating (15).

6.1 Purely Lipschitz Goh extremals

Let G be the free nilpotent Lie group of rank $r = 3$ and step $s = 4$. This group is diffeomorphic to $\mathbb{R}^{32}$. By Theorem 11, Goh extremals of G starting from 0 are precisely the horizontal curves γ in G contained in the algebraic set

$$\Gamma_v = \left\{x \in \mathbb{R}^{32} : P_4^v(x) = P_5^v(x) = P_6^v(x) = 0\right\},$$

for some $v \in \mathbb{R}^{32}$ such that $v \neq 0$ and $v_1 = \ldots = v_6 = 0$. The structure constants $c_{i\alpha}^k$ are determined by the relations of the Lie algebra of G. Using (24), we can then compute the polynomials defining Γ_v (for details, see [6]). These are

$$\begin{aligned}
P_4^v(x) &= -x_1 v_7 - x_2 v_8 - x_3 v_9 + x_5 v_{30} + x_6 v_{31} \\
&\quad + \frac{x_1^2}{2} v_{15} + x_1 x_2 v_{16} + x_1 x_3 v_{17} + \frac{x_2^2}{2} v_{18} + x_2 x_3 v_{19} + \frac{x_3^2}{2} v_{20} \\
P_5^v(x) &= -x_1 v_{10} - x_2 v_{11} - x_3 v_{12} - x_4 v_{30} + x_6 v_{32} \\
&\quad + \frac{x_1^2}{2} v_{21} + x_1 x_2 v_{22} + x_1 x_3 v_{23} + \frac{x_2^2}{2} v_{24} + x_2 x_3 v_{25} + \frac{x_3^2}{2} v_{26} \\
P_6^v(x) &= x_1(v_9 - v_{11}) - x_2 v_{13} - x_3 v_{14} - x_4 v_{31} - x_5 v_{32} \\
&\quad + x_1^2\left(-\frac{1}{2} v_{17} + \frac{1}{2} v_{22} + v_{30}\right) \\
&\quad + x_1 x_2(-v_{19} + v_{24} + v_{31}) + x_1 x_3(-v_{20} + v_{25}) + \frac{x_3^2}{2} v_{29}.
\end{aligned}$$

Theorem 12. *For any Lipschitz function $\phi : [0, 1] \to \mathbb{R}$ with $\phi(0) = 0$, the horizontal curve $\gamma : [0, 1] \to G = \mathbb{R}^{32}$ such that $\gamma(0) = 0$, $\gamma_1(t) = t^2$, $\gamma_2(t) = t$, and $\gamma_3(t) = \phi(t)$ is a Goh extremal.*

With the choice $v_7 = 1$, $v_{18} = 2$, and $v_j = 0$ otherwise, the relevant polynomials are $P_4^v(x) = x_2^2 - x_1$, $P_5^v(x) = P_6^v(x) = 0$. Then, the curve γ is contained in the zero set Γ_v and, by Theorem 11, it is a Goh extremal. The Lipschitz function ϕ is arbitrary. It would be interesting to understand the length minimality properties of γ depending on the regularity of ϕ.

6.2 A family of abnormal curves

During the meeting *Geometric control and sub-Riemannian geometry*, A. Agrachev and J. P. Gauthier suggested the following situation, in order to find a *nonsmooth* length-minimizing curve.

In $M = \mathbb{R}^4$, consider the vector fields

$$X_1 = \frac{\partial}{\partial x_1} + 2x_2\frac{\partial}{\partial x_3} + x_3^2\frac{\partial}{\partial x_4}, \quad X_2 = \frac{\partial}{\partial x_2} - 2x_1\frac{\partial}{\partial x_3}, \tag{25}$$

and denote by $\mathcal{D}$ the distribution of 2-planes in $\mathbb{R}^4$ spanned pointwise by X_1 and X_2. Fix a parameter $\alpha > 0$ and consider the initial and final points $L = (-1, \alpha, 0, 0) \in \mathbb{R}^4$ and $R = (1, \alpha, 0, 0) \in \mathbb{R}^4$. Let $\gamma : [-1, 1] \to \mathbb{R}^4$ be the curve

$$\gamma_1(t) = t, \quad \gamma_2(t) = \alpha|t|, \quad \gamma_3(t) = 0, \quad \gamma_4(t) = 0, \quad t \in [-1, 1]. \tag{26}$$

The curve γ is horizontal and joins L to R. Moreover, it can be easily checked that γ is an abonormal extremal.

This situation is interesting because the distribution $\mathcal{D}$ violates condition (15) with $i = 2$ and $j = 3$. In fact, we have

$$[[X_2, X_1], [[X_2, X_1], X_1]] = 48\frac{\partial}{\partial x_4}.$$

That condition (15) is violated is also apparent from the fact that the nonhorizontal variable x_3 do appear in the coefficients of the vector field X_1 in (25). The fact that the distribution $\mathcal{D}$ is not equiregular, is not relevant.

Agrachev and Gauthier asked whether the curve γ is length minimizing or not, especially for small $\alpha > 0$. The results of [8] cannot be used, because of the failure of (15). In [14], we answered in the negative to the question, at least when $\alpha \neq 1$.

Theorem 13. *For any $\alpha > 0$ with $\alpha \neq 1$, the curve γ in* (26) *is not length minimizing in* $(\mathbb{R}^4, \mathcal{D}, g)$*, for any choice of metric g on $\mathcal{D}$.*

The proof is a lengthy adaptation of the cut-and-adjust technique of [8]. When $\alpha = 1$ the construction of [13] does not work and, in this case, the length minimality property of γ remains open.

7 An extremal curve with Hölder continuous first derivative

On the manifold $M = \mathbb{R}^5$, let $\mathcal{D}$ be the distribution spanned by the vector fields

$$X_1 = \frac{\partial}{\partial x_1}, \qquad X_2 = \frac{\partial}{\partial x_2} + x_1 \frac{\partial}{\partial x_3} + x_1^5 \frac{\partial}{\partial x_4} + x_1 x_2^3 \frac{\partial}{\partial x_5}. \tag{27}$$

We look for abnormal curves passing through $0 \in \mathbb{R}^5$. In view of Theorem 9, case 1), we consider the function $K : \mathbb{R}^3 \times \mathbb{R}^2 \to \mathbb{R}$, defined as in (18),

$$K(\lambda, x) = \lambda_1 + 5\lambda_2 x_1^4 + \lambda_3 x_2^3.$$

With the choice $\lambda_1 = 0$, $\lambda_2 = 1/5$, and $\lambda_3 = -1$, the equation $K(\lambda, x) = 0$ reads $x_1^4 - x_2^3 = 0$. Thus the curve $\kappa : [0, 1] \to \mathbb{R}^2$, $\kappa(t) = (t, t^{4/3})$, is in the zero set of K. It can be checked that the horizontal curve $\gamma : [0, 1] \to M$ such that $(\gamma_1, \gamma_2) = \kappa$ is an abnormal extremal with dual curve $\xi : [0, 1] \to \mathbb{R}^5$,

$$\xi(t) = \big(0, \frac{4}{5}t^5, 0, \frac{1}{5}, -1\big).$$

Notice that we have, for any $t \in [0, 1]$,

$$\langle \xi, [X_1, [X_1, X_2]](\gamma) \rangle = 4t^3,$$
$$\langle \xi, [X_2, [X_1, X_2]](\gamma) \rangle = -3t^{8/3}.$$

Then, when $t > 0$ the curve γ is a *regular* abnormal extremal, in the sense of Definition 14 on page 36 of [9]. By Theorem 5 on page 59 of [9], the curve γ is therefore locally (uniquely) length minimizing on the set where $t > 0$.

The curve γ fails to be *regular* abnormal at $t = 0$. Moreover, there holds $\gamma \in C^{1,1/3}([0, 1]; \mathbb{R}^5)$ with no further regularity at $t = 0$. In this section, we show that γ is not length minimizing.

Theorem 14. *Let g be any metric on the distribution $\mathcal{D}$. The horizontal curve $\gamma : [0, 1] \to M$ defined above is not length minimizing in $(M, \mathcal{D}, g)$ at $t = 0$.*

Proof. For any $0 < \eta < 1$, let $T_\eta \subset \mathbb{R}^2$ be the set

$$T_\eta = \Big\{(x_1, x_2) \in \mathbb{R}^2 : x_1^{4/3} < x_2 < \eta^{1/3} x_1,\ 0 < x_1 < \eta\Big\}.$$

The boundary ∂T_η is oriented counterclockwise. Let $\kappa^\eta : [0, 1] \to \mathbb{R}^2$ be the curve $\kappa^\eta(t) = (t, \eta^{1/3} t)$ for $0 \le t \le \eta$ and $\kappa^\eta(t) = (t, t^{4/3})$ for $\eta \le t \le 1$, and let $\gamma^\eta : [0, 1] \to \mathbb{R}^5$ be the horizontal curve such that $(\gamma_1^\eta, \gamma_2^\eta) = \kappa^\eta$.

We assume without loss of generality that g is the quadratic form on $\mathcal{D}$ that makes X_1 and X_2 orthonormal. The *gain of length* in passing from γ to γ^η is

$$\Delta L(\eta) = \int_0^1 |\dot\kappa|\, dt - \int_0^1 |\dot\kappa^\eta|\, dt = \frac{1}{30}\eta^{5/3} + o(\eta^{5/3}). \tag{28}$$

On the generic monomial $x_1^{i+1}x_2^j$ with $i, j \in \mathbb{N}$, the cut T_η produces the error T_η^{ij} given by the formula

$$\begin{aligned} T_\eta^{ij} &= \int_{\partial T_\eta} x_1^{i+1}x_2^j dx_2 = (i+1)\int_{T_\eta} x_1^i x_2^j dx_1 dx_2 \\ &= \frac{i+1}{j+1}\Big[\frac{1}{i+j+2} - \frac{1}{i+\frac{4}{3}(j+1)+1}\Big]\eta^{i+\frac{4}{3}(j+1)+1}. \end{aligned} \tag{29}$$

We are interested in this formula when $i = j = 0$, when $i = 4$ and $j = 0$, when $i = 0$ and $j = 3$. The initial error produced by the cut T_η is the vector of $\mathbb{R}^3$

$$\begin{aligned} \mathcal{E}_0(\eta) &= \big(T_\eta^{0,0}, T_\eta^{4,0}, T_\eta^{0,4}\big) \\ &= \Big(\frac{1}{14}\eta^{7/3}, \frac{5}{114}\eta^{19/3}, \frac{1}{95}\eta^{19/3}\Big). \end{aligned} \tag{30}$$

Only the exponents $7/3$ and $19/3$ of η are relevant, not the coefficients.

Our first step is to correct the error of order $\eta^{7/3}$ on the third coordinate. For fixed parameters $b > 0$, $\lambda > 0$, and $\varepsilon > 0$, let us define the curvilinear rectangle

$$R_{b,\lambda}(\varepsilon) = \big\{(x_1, x_2) \in \mathbb{R}^2 : b < x_1 < b + |\varepsilon|^\lambda,\ x_1^{4/3} < x_2 < x_1^{4/3} + \varepsilon\big\}. \tag{31}$$

When $\varepsilon < 0$, we let

$$R_{b,\lambda}(\varepsilon) = \big\{(x_1, x_2) \in \mathbb{R}^2 : b < x_1 < b + |\varepsilon|^\lambda,\ x_1^{4/3} + \varepsilon < x_2 < x_1^{4/3}\big\}. \tag{32}$$

The boundary $\partial R_{b,\lambda}(\varepsilon)$ is oriented counterclockwise if $\varepsilon > 0$, while it is oriented clockwise when $\varepsilon < 0$. The curve κ^η is deviated along the boundary of this rectangle and then it is lifted to a horizontal curve. The effect of $R_{b,\lambda}(\varepsilon)$ on the generic monomial $x_1^{i+1}x_2^j$ is

$$\begin{aligned} R_{b,\lambda}^{ij}(\varepsilon) &= \int_{\partial R_{b,\lambda}^{ij}(\varepsilon)} x_1^{i+1}x_2^j dx_2 \\ &= \frac{i+1}{j+1}\sum_{k=0}^{j}\binom{j+1}{k}\frac{\varepsilon^{j+1-k}}{i+\frac{4}{3}k+1}\Big[(b+|\varepsilon|^\lambda)^{i+\frac{4}{3}k+1} - b^{i+\frac{4}{3}k+1}\Big]. \end{aligned} \tag{33}$$

The *cost of length* of $R_{b,\lambda}(\varepsilon)$ is

$$\Lambda\big(R_{b,\lambda}(\varepsilon)\big) = 2|\varepsilon|. \tag{34}$$

When $i = j = 0$, formula (33) reads $R_{b,\lambda}^{0,0}(\varepsilon) = \varepsilon|\varepsilon|^\lambda$, whereas

$$\begin{aligned} R_{b,\lambda}^{4,0}(\varepsilon) &= 5\varepsilon\big[(b+|\varepsilon|^\lambda)^5 - b^5\big] \\ &= 5\varepsilon\sum_{k=0}^{4}\binom{5}{k}|\varepsilon|^{\lambda(5-k)}b^\kappa, \end{aligned} \tag{35}$$

and

$$\begin{aligned} R^{0,3}_{b,\lambda}(\varepsilon) &= \frac{1}{4}\sum_{k=0}^{3}\binom{4}{k}\frac{\varepsilon^{4-k}}{\frac{4}{3}k+1}\Big[(b+|\varepsilon|^{\lambda})^{\frac{4}{3}k+1}-b^{\frac{4}{3}k+1}\Big] \\ &= \frac{\varepsilon}{5}\big[(b+|\varepsilon|^{\lambda})^{5}-b^{5}\big]+\frac{1}{4}\sum_{k=0}^{2}\binom{4}{k}\frac{\varepsilon^{4-k}}{\frac{4}{3}k+1}\Big[(b+|\varepsilon|^{\lambda})^{\frac{4}{3}k+1}-b^{\frac{4}{3}k+1}\Big] \\ &= \frac{1}{25}R^{4,0}_{b,\lambda}(\varepsilon)+\widehat{R}^{0,3}_{b,\lambda}(\varepsilon), \end{aligned} \tag{36}$$

where $\widehat{R}^{0,3}_{b,\lambda}(\varepsilon)$ is defined via the last identity.

We choose $b=\eta$. The parameter $0<\lambda<1$ will be fixed at the end of the argument. To correct the error on the third coordinate, we solve the equation

$$R^{0,0}_{\eta,\lambda}(\varepsilon)+T^{0,0}_{\eta}=0,$$

in the unknown ε. In fact, this equation is $\varepsilon|\varepsilon|^{\lambda}+\frac{1}{14}\eta^{7/3}=0$ and the solution is

$$\varepsilon=-c_0\eta^{\beta},\quad \text{where } c_0=\frac{1}{14^{1/(1+\lambda)}} \text{ and } \beta=\frac{7}{3(1+\lambda)}.$$

The choice $b=\eta$ is not relevant, here. By (28) and (34), the cost of length is admissible if $\eta>0$ is small enough and we have

$$\frac{7}{3(1+\lambda)}>\frac{5}{3}\quad\Leftrightarrow\quad\lambda<\frac{2}{5}. \tag{37}$$

This is our first restriction on λ.

The rectangle $R_{\eta,\lambda}(\varepsilon)$ produces new errors on the fourth and fifth coordinates. Namely, by (35) we have

$$R^{4,0}_{\eta,\lambda}\big(-c_0\eta^{\beta}\big)=-5c_0\eta^{\beta}\sum_{k=0}^{4}\binom{5}{k}\big(c_0\eta^{\beta}\big)^{\lambda(5-k)}\eta^{k}. \tag{38}$$

When $\lambda<3/4$, condition implied by (37), the leading term in η in the sum above is obtained for $k=0$.

By (36), the error produced on the last coordinate is

$$R^{0,3}_{\eta,\lambda}\big(-c_0\eta^{\beta}\big)=\frac{1}{4}\sum_{k=0}^{3}\binom{4}{k}\big(-c_0\eta^{\beta}\big)^{4-k}\frac{3}{4k+3}\Big[\Big(\eta+\big(c_0\eta^{\beta}\big)^{\lambda}\Big)^{\frac{4}{3}k+1}-\eta^{\frac{4}{3}k+1}\Big]. \tag{39}$$

When $\lambda<3/4$, the bracket $[\ldots]$ in the sum over k above is

$$\big(c_0\eta^{\beta}\big)^{\lambda(\frac{4}{3}k+1)}\Big[1+\Big(\frac{4}{3}k+1\Big)c_0^{-\lambda}\eta^{1-\frac{7\lambda}{3(1+\lambda)}}+\frac{2k}{3}\Big(\frac{4}{3}k+1\Big)c_0^{-2\lambda}\eta^{2-\frac{14\lambda}{3(1+\lambda)}}+\ldots\Big].$$

The leading term in the sum in (39) is obtained for $k=3$, and the second leading term is obtained for $k=2$.

We have the new vector of errors

$$\mathcal{E}_1(\eta) = \Big(0, R^{4,0}_{\eta,\lambda}\big(-c_0\eta^\beta\big) + T^{4,0}_\eta, R^{0,3}_{\eta,\lambda}\big(-c_0\eta^\beta\big) + T^{0,3}_\eta\Big). \tag{40}$$

When $\lambda < 3/4$, the errors on the fourth and fifth coordinates produced by the rectangle $R_{\eta,\lambda}$ dominate the errors produced by the cut, see (30). In fact, we have

$$\beta(1+5\lambda) < \frac{19}{3} \quad \Leftrightarrow \quad \lambda < \frac{3}{4}.$$

Also the second leading term in $R^{0,3}_{\eta,\lambda}\big(-c_0\eta^\beta\big)$ dominates $T^{0,3}_\eta$. In fact, we have

$$\beta\Big(2+\frac{11}{3}\lambda\Big) < \frac{19}{3} \quad \Leftrightarrow \quad \lambda < \frac{3}{4}.$$

Now we use a rectangle $R_{b,\mu}(\varepsilon)$ to correct the error on the fourth coordinate. Here, $\frac{1}{2} < b < 3/4$ is position parameter and $\mu > 0$ is small enough. Conceptually, we could take $\mu = 0$. The parameter $\mu > 0$ is only needed to confine the construction in a bounded region. We solve the equation

$$R^{4,0}_{b,\mu}(\varepsilon) + R^{4,0}_{\eta,\lambda}\big(-c_0\eta^\beta\big) = 0 \tag{41}$$

in the unknown ε. By the formulas computed above, we deduce that the solution $\varepsilon = \bar\varepsilon$ is

$$\bar\varepsilon = c_1\eta^{\beta\frac{1+5\lambda}{1+5\mu}} + \dots,$$

where $c_1 > 0$ is an explicit constant and the dots stand for lower order terms in η. The cost of length of the rectangle $R_{b,\mu}(\bar\varepsilon)$ is admissible for any $\mu > 0$ close to 0, because $\beta(1+5\lambda) > 5/3$.

By (36) and (41), we have the identity

$$R^{0,3}_{b,\mu}(\bar\varepsilon) + R^{0,3}_{\eta,\lambda}\big(-c_0\eta^\beta\big) = \widehat R^{0,3}_{b,\mu}(\bar\varepsilon) + \widehat R^{0,3}_{\eta,\lambda}\big(-c_0\eta^\beta\big),$$

and, therefore, the new vector of errors is

$$\mathcal{E}_2(\eta) = \Big(R^{0,0}_{b,\mu}(\bar\varepsilon), T^{4,0}_\eta, \widehat R^{0,3}_{b,\mu}(\bar\varepsilon) + \widehat R^{0,3}_{\eta,\lambda}\big(-c_0\eta^\beta\big) + T^{0,3}_\eta\Big), \tag{42}$$

where we have

$$R^{0,0}_{b,\mu}(\bar\varepsilon) = \bar\varepsilon^{1+\mu} = c_2\eta^{\beta\frac{(1+5\lambda)(1+\mu)}{1+5\mu}} + \dots, \tag{43}$$

with the coefficient $c_2 = c_1^{1+\mu}$.

In the next step, we correct simultaneously the errors on the fourth and fifth coordinates. We need curvilinear squares. Let $0 < b < 1$ be a position parameter. For any $\varepsilon \in (-1,1)$, we let

$$Q_b(\varepsilon) = R_{b,1}(|\varepsilon|). \tag{44}$$

The parameter λ of the rectangle is set to $\lambda = 1$. Set-theoretically, the definition is the same for positive and negative ε. However, when $\varepsilon > 0$ the boundary $\partial Q_b(\varepsilon)$ of the square is oriented clockwise; when $\varepsilon < 0$ the boundary is oriented counterclockwise. The *cost of length* $\Lambda(Q_b(\varepsilon))$ of the square is the sum of the length of the four sides. For some constant $C > 0$ independent of b and ε we have

$$\Lambda(Q_b(\varepsilon)) \leq C|\varepsilon|. \tag{45}$$

By (33), when $\varepsilon > 0$ the effect $Q_b^{ij}(\varepsilon)$ of the square on the monomial $x_1^{i+1}x_2^j$ is

$$Q_b^{ij}(\varepsilon) = \frac{i+1}{j+1}\sum_{k=0}^{j}\binom{j+1}{k}\varepsilon^{j+1-k}\frac{1}{i+\frac{4}{3}k+1}\Big[(b+\varepsilon)^{i+\frac{4}{3}k+1} - b^{i+\frac{4}{3}k+1}\Big].$$

When $\varepsilon < 0$, we have $Q_b^{ij}(\varepsilon) = -Q_b^{ij}(|\varepsilon|)$.

Let $3/4 < b_1 < b_2 < 1$ be position parameters and let $\mu > 0$ be close to 0. We solve the system of equations

$$\begin{cases} Q_{b_1}^{4,0}(\varepsilon_1) + R_{b_2,\mu}^{4,0}(\varepsilon_2) + T_\eta^{4,0} = 0 \\ Q_{b_1}^{0,3}(\varepsilon_1) + R_{b_2,\mu}^{0,3}(\varepsilon_2) + \widehat{R}_{b,\mu}^{0,3}(\bar{\varepsilon}) + \widehat{R}_{\eta,\lambda}^{0,3}\big(-c_0\eta^\beta\big) + T_\eta^{0,3} = 0 \end{cases}$$

in the unknowns $\varepsilon_1, \varepsilon_2$. Subtracting the first equation from the second one and using (36), we get the equivalent system

$$\begin{cases} Q_{b_1}^{4,0}(\varepsilon_1) + R_{b_2,\mu}^{4,0}(\varepsilon_2) + T_\eta^{4,0} = 0 \\ \widehat{Q}_{b_1}^{0,3}(\varepsilon_1) + \widehat{R}_{b_2,\mu}^{0,3}(\varepsilon_2) + \mathcal{E}(\eta) = 0 \end{cases} \tag{46}$$

where

$$\begin{aligned} \mathcal{E}(\eta) &= \widehat{R}_{b,\mu}^{0,3}(\bar{\varepsilon}) + \widehat{R}_{\eta,\lambda}^{0,3}\big(-c_0\eta^\beta\big) + T_\eta^{0,3} - T_\eta^{4,0} \\ &= c_3\eta^{\beta(2+\frac{11}{3}\lambda)} + \dots \end{aligned}$$

for some $c_3 > 0$. The dots stand for lower order terms in η. In fact, the leading term in $R_{\eta,\lambda}^{0,3}\big(-c_0\eta^\beta\big)$ dominates the remaining terms. Using a notation consistent with (35), we also let

$$\begin{aligned} \widehat{Q}_{b_1}^{0,3}(\varepsilon_1) &= \operatorname{sgn}(\varepsilon_1)\frac{1}{4}\sum_{k=0}^{2}\binom{4}{k}\frac{|\varepsilon_1|^{4-k}}{\frac{4}{3}k+1}\Big[(b_1+|\varepsilon_1|)^{\frac{4}{3}k+1} - b_1^{\frac{4}{3}k+1}\Big] \\ &= c_4\varepsilon_1^3 + ... \end{aligned}$$

Above, $c_4 > 0$ is a constant and the dots stand for negligible terms. Notice that we have control on the sign of the leading term.

The system (46) can thus be approximated in the following way

$$\begin{cases} \operatorname{sgn}(\varepsilon_1)\varepsilon_1^2 + c_5\varepsilon_2|\varepsilon_2|^{5\mu} + c_6\eta^{\frac{19}{3}} + \ldots = 0 \\ \varepsilon_1^3 + c_7|\varepsilon_2|^{2+\frac{11}{3}\mu} + c_8\eta^{\beta(2+\frac{11}{3}\lambda)} + \ldots = 0, \end{cases} \tag{47}$$

where $c_5, \ldots, c_8 > 0$ are constants and the dots stand for negligible terms. We can compute ε_2 as a function of ε_1 from the first equation and replace this value into the second equation. This operation produces lower order terms. Thus the second equation reads

$$\varepsilon_1^3 + c_8\eta^{\beta(2+\frac{11}{3}\lambda)} + \ldots = 0,$$

and there is a solution $\varepsilon_1 < 0$ satisfying

$$\varepsilon_1 = -c_9\eta^{\beta(2+\frac{11}{3}\lambda)/3} + \ldots = -c_9\eta^{\frac{7(2+\frac{11}{3}\lambda)}{9(1+\lambda)}} + \ldots,$$

where $c_9 > 0$ and the dots stand for lower order terms in η. As a consequence, from the first equation in (47) we deduce that

$$\varepsilon_2 = -c_{10}\eta^{\frac{19}{3(1+5\mu)}} + \ldots.$$

The cost of length of the rectangle $R_{b_2,\mu}(\varepsilon_2)$ is $2|\varepsilon_2|$, and it is admissible because for $\mu > 0$ close to 0 we have

$$\frac{19}{3(1+5\mu)} > \frac{5}{3}.$$

By (45), the cost of lenght of the square $Q_{b_1}(\varepsilon_1)$ is at most $C|\varepsilon_1|$, and, for small η, it is admissible if and only if

$$\frac{7(2+\frac{11}{3}\lambda)}{9(1+\lambda)} > \frac{5}{3} \quad \Leftrightarrow \quad \lambda > \frac{3}{32}. \tag{48}$$

Here, we have a nontrivial restriction for λ. This restriction is compatible with (37). Now the parameter λ is fixed once for all in such a way that

$$\frac{3}{32} < \lambda < \frac{2}{5}. \tag{49}$$

The device $Q_{b_1}(\varepsilon_1)$ produces an error on the third coordinate of the order $|\varepsilon_1|^2$, that is of the order $\eta^{14(2+\frac{1}{3}\lambda)/9}$. The device $R_{b_2,\mu}(\varepsilon_2)$ produces an error on the third coordinate of the order $|\varepsilon_2|^{1+\mu}$, that is of the order $\eta^{19(1+\mu)/3}$. These errors are negligible with respect to the error $R^{0,0}_{b,\mu}(\bar{\varepsilon})$ appearing in (42)–(43). Eventually, after our last correction we have the vector of errors

$$\mathcal{E}_3(\eta) = \big(c_2\eta^{\frac{7}{3}\varrho} + \ldots, 0, 0\big), \quad \text{where } \varrho = \frac{(1+5\lambda)(1+\mu)}{(1+\lambda)(1+5\mu)}, \tag{50}$$

the dots stand for lower order terms and the number ϱ satisfies the key condition $\varrho > 1$, provided that $0 < \mu < \lambda$. Now also μ is fixed.

Comparing the initial error $\mathcal{E}_0(\eta)$ in (30) and the error $\mathcal{E}_3(\eta)$ in (50), we realize that the initial error $\eta^{7/3}$ on the third coordinate decreased by a geometric factor $\varrho > 1$. Now we can iterate the entire construction to set to zero all the three components of the error. Here, we omit the details of this standard part of the argument. This finishes the proof.

Remark 1. The curve γ studied in Theorem 14 is of class $C^{1,1/3}$. The curves considered in Theorem 7 are at most $C^{1,1/4}$. There is a gap between the two cases. In the proof of Theorem 14, the key step is the choice of λ made in (49). In particular, there is a very delicate bound from below for λ. In the proof of Theorem 7, there is no such a bound from below.

8 Final comments

Concerning the question about the regularity of length minimizing curves in sub-Riemannian manifolds, there are two possibilities. Either, in any sub-Riemannian manifold every length minimizing curve is C^∞ smooth (answer in the positive); or, there is some sub-Riemannian manifold with nonsmooth (non C^1, non C^2, etc.) length minimizing curves (answer in the negative). The author has no clear feeling on which of the two answers to bet.

Theorem 5 on step 3 Carnot groups suggests that, in sub-Riemannian manifolds of step 3, any length minimizing curve is C^∞ smooth. This seems to be the first question to investigate in view of an answer in the positive. In the same spirit, Theorem 6 suggests that in sub-Riemannian manifolds of rank 2 and step 4 any length minimizing curve is C^∞ smooth.

On the other hand, the first example to investigate in order to find a length minimizer with a corner type singularity is the one of Sect. 6.2 with the choice $\alpha = 1$. Moreover, Theorem 7 and the computations made in Sect. 7 suggest to look for nonsmooth length minimizing curves in the class of $C^{1,\delta}$ abnormal extremals with $0 < \delta < 1$ sufficiently close to 1. One interesting example could be the manifold $M = \mathbb{R}^5$ with the distribution spanned by the vector fields

$$X_1 = \frac{\partial}{\partial x_1}, \quad X_2 = \frac{\partial}{\partial x_2} + x_1 \frac{\partial}{\partial x_3} + x_1^{2m} \frac{\partial}{\partial x_4} + x_1 x_2^m \frac{\partial}{\partial x_5}, \tag{51}$$

for $m \in \mathbb{N}$ large.

Finally, the example of a purely Lipschitz Goh extremal of Sect. 6.1 proves that the first and second order necessary conditions for strictly abnormal extremals do not imply, in general, any further regularity beyond the given Lipschitz regularity. New and deeper techniques are needed in order to develop the regularity theory.

References

1. Agrachev, A., Barilari, D., Boscain, U.: Introduction to Riemannian and Sub-Riemannian geometry, http://people.sissa.it/ agrachev/agrachev_files/notes.html
2. Agrachev, A., Sachkov, Y.L.: Control Theory from the Geometric Viewpoint. Encyclopaedia of Mathematical Sciences **87**. Control Theory and Optimization, II. Springer-Verlag, Berlin Heidelberg New York (2004)
3. Agrachev, A., Sarychev, A.: Abnormal sub-Riemannian geodesics: Morse index and rigidity. Ann. Inst. Henri Poincaré, **13**(16), 635–690 (1996)
4. Chitour, Y., Jean, F., Trélat, E.: Genericity results for singular curves. J. Differential Geom. **73**(1), 45–73 (2006)
5. Grayson, M., Grossman, R.: Models for free nilpotent Lie algebras. J. Algebra **135**(1), 177–191 (1990)
6. Le Donne, E., Leonardi, G.P., Monti, R., Vittone, D.: Extremal curves in nilpotent Lie groups. Geom. Funct. Anal. **23**(4), 1371–1401 (2013)
7. Le Donne, E., Leonardi, G.P., Monti, R., Vittone, D.: Extremal polynomials in stratified groups (forthcoming 2013)
8. Leonardi, G.P., Monti, R.: End-point equations and regularity of sub-Riemannian geodesics. Geom. Funct. Anal. **18**(2), 552–582 (2008)
9. Liu, W., Sussmann, H.: Shortest paths for sub-Riemannian metrics on rank-two distributions, Mem. Amer. Math. Soc. **118**, x+104 (1995)
10. Margulis, G.A., Mostow, G.D.: Some remarks on the definition of tangent cones in a Carnot-Carathéodory space, J. Anal. Math. **80**, 299–317 (2000)
11. Montgomery, R.: Abnormal minimizers , SIAM J. Control Optim., **32**, 1605–1620 (1994)
12. Montgomery, R.: A Tour of Sub-Riemannian Geometries, Their Geodesics and Applications, AMS (2002)
13. Monti, R.: Regularity results for sub-Riemannian geodesics, Calc. Var. 2013 (to appear). doi: 10.1007/s00526-012-0592-2s
14. Monti, R.: A family of nonminimizing abnormal curves, Ann. Mat. Pura Appl. 2013 (to appear). doi: 10.1007/s10231-013-0344-8
15. Strichartz, R.S.: Sub-Riemannian geometry, J. Differential Geom., **24**, 221–263 (1986) [Corrections to "Sub-Riemannian geometry", J. Differential Geom., **30**, 595–596 (1989)]
16. Tan, K., Yang, X.: Subriemannian geodesics of Carnot groups of step 3. http://arxiv.org/pdf/1105.0844v1.pdf