UNIVERSITÀ DEGLI STUDI DI PADOVA

Università degli Studi di Padova

DIPARTIMENTO DI MATEMATICA "TULLIO LEVI-CIVITA"

Corso di Laurea Triennale in Matematica

The regularity problem for sub-Riemannian geodesics

Relatori:
Prof. Roberto Monti
Prof. Davide Vittone

Laureando: Giorgio Baglioni
Matricola: 2053469

Anno Accademico 2023/2024

19/04/2024

Contents

Introduction

One of the main problems in sub-Riemannian geometry is the regularity of length-minimizing curves (or geodesics). It has been open since the work of R. S. Strichartz [22] and U. Hamenstädt [8] in the late 1980s. In particular, in [22] it was claimed that every geodesic parametrized by constant speed corresponds to a solution for a suitable Hamiltonian system in the cotangent bundle, hence its expected C^∞ smoothness and overall "good" behavior, in perfect analogy with the Riemannian case. This claim was soon proved wrong, precisely due to the unexpected, and initially only theoretical, possibility of so-called *abnormal* length minimizers, an actual example of which was later displayed by R. Montgomery in [16]. Their existence is the fundamental reason why the problem's complexity far exceeds that of its Riemannian version.

In Sections 2.1 through 2.4 of this thesis, our goal is to introduce the basic concepts of sub-Riemannian geometry and to develop a *first order theory* for *horizontal* curves (i.e., curves whose tangent vectors at any point lie inside the prescribed distribution of tangent planes). This will eventually allow us to distinguish candidate minimizers (*extremals*) between normal and abnormal, a distinction which is closely related to the concept of *dual curve*, and to illustrate the massive difference in their behavior. Notably, it is needed to work around the horizontality constraint, therefore such a straightforward application of variational calculus techniques as in the Riemannian case is impossible. We also discuss higher-order constraints for abnormal extremals, such as the Goh condition.

In 2016, the first general regularity result for sub-Riemannian geodesics was obtained: following on from some partial results in the years before [14] [12], Enrico Le Donne and Eero Hakavuori established that length minimizers in sub-Riemannian manifolds never present corner-type singularities [7]. Chapter 3 of this thesis is mainly devoted to the presentation of their inductive argument, which is carried out by initially performing a reduction to the model case of a Carnot group, particularly favorable because of its additional algebraic structure: nilpotency and stratification are key properties to allow for successfully lifting curves from a quotient Carnot group (Lemma 3.3.3), while the subsequent corrections to the lifted curve are performed by employing conjugation maps (Lemma 3.3.4). In Chapter 1 we establish the necessary Lie theory to thereinafter introduce Carnot groups (Section 2.5) and properly understand them simultaneously as nilpotent Lie groups and as sub-Riemannian manifolds; we also introduce the Baker-Campell-Hausdorff formula, a crucial tool throughout our whole survey as it bridges the gap between group and Lie algebraic operations.

The reasons that make the above mentioned reduction to Carnot groups possible can be traced back to the pivotal role they have in the context of sub-Riemannian geometry: namely, they emerge as “tangent" metric spaces to all sub-Riemannian manifolds that satisfy a technical condition (*equiregularity*), once having adequately generalized the notion of “tangency". Subsection 2.5.1 aims to present this issue in a precise way. However, to effectively transfer the entire problem onto these infinitesimal models of the original spaces, the local approximation (possibly preceded by a “regularization" procedure) needs to preserve length minimizers and other properties of curves. These highly technical issues, addressed in [18] and [10], mostly extend beyond the scope of this work, even though we manage to adequately face some related questions in Subsection 2.5.1 and Section 3.4.

Chapter 1

Basic facts about Lie groups

Refernces for Chapter 1 are [21], [13], [23].

1.1 The structure of Lie groups and the Baker-Campell-Hausdorff formula

Definition 1.1.1. A *Lie Algebra* (over $\mathbb{R}$) is a real vector space $\mathfrak{g}$ endowed with a bilinear map (called *Lie Bracket*)

$$[\cdot,\cdot] : \mathfrak{g} \times \mathfrak{g} \longrightarrow \mathbb{R}$$

such that, for all $x, y, z \in \mathfrak{g}$:

(i) $[x, y] = -[y, x]$;

(ii) $\big[[x, y], z\big] + \big[[y, z], x\big] + \big[[z, x], y\big] = 0$ (*Jacobi's identity*).

A Lie algebra is called *nilpotent* if there exists $k \in \mathbb{N}$ such that the iterated brackets vanish

$$\big[...\big[[x_1, x_2], x_3\big]..., x_k\big] = 0, \tag{1.1}$$

for every $x_1, ..., x_k \in \mathfrak{g}$. The minimal k for which (1.1) holds is called the *step* of the Lie algebra.
A Lie algebra is called *Abelian* if $[x, y] = 0$ for all $x, y \in \mathfrak{g}$, i.e., if it is a nilpotent Lie algebra of step 1.
Observe that the set of vector fields over an open set of $\mathbb{R}^n$ (as well as over any given manifold) is a Lie algebra, with bracket defined as the commutator $[X, Y] = XY - YX$.

Definition 1.1.2. Let Ω be an open subset of $\mathbb{R}^n$, let X be a smooth vector field on Ω. Given $t \in \mathbb{R}$, we define the *exponential* of the field tX as the operator:

$$\exp(tX)f(x) := f(\phi_{X,t}(x)) \quad x \in \Omega,$$

where $\phi_{X,t}$ represents the flow of X at time t. For each $x \in \Omega$, there exists a $\delta(x) > 0$ such that the exponential is well defined for $t \in (-\delta(x), \delta(x))$.

If we take f to be the identity function, then we can also identify the exponential $\exp(tX)$ with an integral curve of X, i.e. $\exp(tX)(x) = \phi_{X,t}(x)$ is the solution to the Cauchy problem $\dot{\gamma}(t) = X(\gamma(t)), \gamma(0) = x$. It is possible to show that the following Taylor expansion holds:

$$\exp(tX)f(x) = \sum_{j=0}^{k} \frac{t^j}{j!} X^j f(x) + O(t^{k+1}). \tag{1.2}$$

Hence, we have

$$\exp(tX) = \sum_{k=0}^{\infty} \frac{t^k}{k!} X^k$$

as formal power series.
One may now observe that the operators $\exp(sX) \circ \exp(tY)$ and $\exp(sX + tY + st[X,Y]/2)$ have the same expansion up to the term of degree 2. This formal procedure can, in fact, be pushed to any degree, constructing a formal power series $S(s,t;X,Y) = sX + tY + \sum_{j,k\geq 1} s^j t^k Z_{j,k}(X,Y)$, where each $Z_{j,k}$ is a fixed combination of iterated commutators of X and Y, containing X j times and Y k times. This series satisfies the formal identity $\exp(sX)\exp(tY) = \exp(S(s,t;X,Y))$, known as the Baker-Campell-Hausdorff formula.

Definition 1.1.3. A *Lie group* is a smooth manifold G, which is also a group $(G, \cdot)$ such that the map from $G \times G$ to G which assigns xy^{-1} to (x,y) is smooth.

For $a \in G$, let L_a be the "left translation" operator, i.e., $L_a f(x) := f(a^{-1}x)$.

Definition 1.1.4. A vector field X on G is called *left-invariant* if

$$L_a(Xf) = X(L_a f)$$

for every $a \in G$.

There is a 1:1 correspondence between tangent vectors at the identity $e \in G$ and left-invariant vector fields on G: for any given vector $v \in T_e G$, which can also be thought of as a real linear functional defined on smooth functions which satisfies Leibnitz's identity, there exists a unique left-invariant vector field X such that $X_e = v$. More specifically, the correspondence is given by $Xf(x) = v(L_{x^{-1}} f)$.

The space of left-invariant vector fields over G is a Lie algebra $Lie(G)$ (if equipped with the commutator operation as bracket), called the Lie algebra of G.
In the nilpotent case, an *a posteriori* procedure to construct a Lie group with a prescribed (nilpotent) Lie algebra is also possible: given such a Lie algebra $\mathfrak{g}$, it is possible to show that if we equip $\mathfrak{g}$ with the composition law

$$x \cdot y = S(1,1;x,y) = x + y + \frac{1}{2}[x,y] + \frac{1}{12}[[x,y],x] - \frac{1}{12}[[x,y],y] + ...$$

then we obtain a Lie group structure with $\mathfrak{g}$ as its underlying manifold. Moreover, if we denote this group by G, then $Lie(G) = \mathfrak{g}$, and G is the unique connected and simply connected Lie group, up to isomorphism, with $\mathfrak{g}$ as its Lie algebra (*Ado's theorem*), see

[23], p.31. More specifically, if the Lie algebra is real and finite-dimensional, then there is a bijective correspondence between connected Lie subgroups of a Lie group and subalgebras of its Lie algebra.
Any nilpotent Lie algebra is a solvable group when equipped with the group operation above: indeed, a group is solvable if and only if every normal subgroup is strictly contained in its normalizer. Hence the associated Lie group is also solvable; this fact will be useful later.

Definition 1.1.5. Let G and H be Lie groups. A *Lie group homomorphism* is a smooth map $F : G \longrightarrow H$ which is also a group homomorphism. If it is also a diffeomorphism, then it is called a *Lie group isomorphism.*

Fact 1.1.6. Any Lie group homomorphism is smooth.

Given a group homomorphism $\phi : G \longrightarrow H$, thanks to the correspondence between Lie algebras and tangent spaces at the identity, the tangent map at the identity $d\phi_e : T_eG \longrightarrow T_eH$ induces a (linear) map $\phi_* : \mathfrak{g} \longrightarrow \mathfrak{h}$ such that the following diagram commutes:

$$\begin{tikzcd} G \arrow[r, "\varphi"] \arrow[d, "\exp"'] & H \arrow[d, "\exp"] \\ \mathfrak{g} \arrow[r, "\varphi_*"] & \mathfrak{h} \end{tikzcd}$$

Definition 1.1.7. Let G be a Lie group, and let H be a normal subgroup of G. Then H is also a Lie group and so is the quotient G/H, as its group operation is again continuous (G/H has the quotient topology). We denote the canonical projection homomorphism by $\pi : G \longrightarrow G/H$, and the induced map between the Lie algebras by π_* .

1.2 Stratified groups

In the following G will be a Lie group, and $\mathfrak{g}$ will be the Lie algebra $Lie(G)$ of G.

Definition 1.2.1. A *stratified group* G is a nilpotent, connected and simply connected Lie group which admits a *stratification*, i.e., a decomposition

$$\mathfrak{g} = V_1 \oplus V_2 \oplus ... \oplus V_s \quad , \quad V_s \neq 0, \tag{1.3}$$

with the property that, for $i = 1, ..., s-1$, $V_i = \left[V_1, V_{i-1}\right]$, and $\left[V_1, V_s\right] = \{0\}$.

The integer $r = dimV_1$ is called the *rank* of the Lie algebra/Lie group.

The definition of rank we just gave is equivalent to the definition of rank found in [9]. The observations below are straightforward.

- $\mathfrak{g}$ is nilpotent of step s;
- $\left[V_i, V_j\right] \subset V_{i+j}$, whenever $i + j \leq s$;

- $\left[V_i, V_j\right] = \{0\}$, whenever $i + j > s$.

Let us fix an adapted basis of $\mathfrak{g}$, i.e. a basis $X_1, ..., X_n$, $n = dim\mathfrak{g}$, whose order is coherent with the stratification:

$$\underbrace{X_1, .., X_r}_{V_1}, \underbrace{X_{r+1}, ..., X_{r_2}}_{V_2}, \underbrace{X_{r_2+1} ...,}_{V_3} ...X_n.$$

In the sequel, we will employ many times the so-called *coordinates of the second kind.* That is, we locally identify G with $\mathbb{R}^n$ by means of

$$\mathbb{R}^n \ni (x_1, ..., x_n) \leftrightarrows \exp(x_n X_n) \circ ... \circ \exp(x_1 X_1)(0) =: \exp(x_1 X_1) \cdot ... \cdot \exp(x_n X_n) \in G.$$

In the nilpotent case, this actually serves as a global coordinate system.

The stratification allows for the existence of a family of intrinsic dilations on G: first, for any $i = 1, ..., n$, we define the *degree* of i, $d \in \{1, ..., s\}$, as $d(i) = j$ if and only if $X_i \in V_j$.
Because of the stratification assumption, for each $r > 0$ there is a (unique) linear automorphism δ_r of the Lie algebra $\mathfrak{g}$ such that

$$\delta_r(X_i) = r^{d_i} X_i.$$

These induce dilations on G (which are also group automorphisms): in exponential coordinates of the second kind we have

$$\delta_r(x_1, ..., x_n) := (rx_1, ..., r^{d(i)} x_i, ..., r^s x_n).$$

Chapter 2

Introduction to Sub-Riemannian Manifolds and First order theory

2.1 Sub-Riemannian manifolds

A sub-Riemannian manifold is a smooth, connected $n-$dimensional manifold M, endowed with a smooth sub-bundle $\Delta \subset TM$ of constant rank r and which is also *bracket generating* (more on this later), called the horizontal sub-bundle, and with a smooth metric g on Δ. Without loss of generality, we can localize the analysis by setting $M = \mathbb{R}^n$, and assuming that Δ is generated by smooth orthonormal (with respect to g) vector fields $X_1, ..., X_r$.

In this chapter, we denote the scalar product between $x, y \in \mathbb{R}^d$ as $x \cdot y$.

A Lipschitz continous curve $\gamma : [0,1] \longrightarrow M = \mathbb{R}^n$ is said to be horizontal if $\dot{\gamma}(t) \in \Delta_{\gamma(t)}$ for a.e. $t \in [0,1]$. This happens iff there exist uniquely defined functions $h = (h_1, ..., h_r) \in L^\infty([0,1], \mathbb{R}^r)$, called the *controls* associated to γ, such that:

$$\dot{\gamma}(t) = \sum_{j=1}^{r} h_j(t) X_j(t) = h \cdot X(t). \tag{2.1}$$

We define the *length* of γ as:

$$L(\gamma) := \int_0^1 |h(t)| dt. \tag{2.2}$$

Definition 2.1.1 (Carnot-Carathéodory distance)**.** The Carnot-Carathéodory (CC) distance between $x, y \in \mathbb{R}^n$ is defined as

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

If d is an actual distance (i.e., if we are taking the infimum over a non-empty set), the resulting metric space (M, d) is called a *Carnot-Carathéodory space*. The structure induced by such a distance is then called *sub-Riemannian*, as intuitively the “allowed" directions in x are a subspace of T_xM.

Additionally, upon defining $L_2(\gamma) := \left(\int_0^1 |h(t)|^2\right)^{\frac{1}{2}}$ and

$$d_2(x, y) = \inf\{L_2(\gamma) \mid \gamma \text{ is horizontal}, \gamma(0) = x, \gamma(1) = y\},$$

we have that, since $L(\gamma)$ is independent of the parametrization of γ, then $d(x, y) = d_2(x, y)$ by applying the Cauchy-Schwarz inequality, for which curves parametrized at constant speed are equality cases (i.e. $L(\gamma) = L_2(\gamma)$). The quantity $L_2(\gamma) = (2E(\gamma))^{\frac{1}{2}}$ is closely related to the *energy* $E(\gamma)$.

2.2 The Hörmander condition

Consider a point $p \in \mathbb{R}^n$ and a vector field X on $\mathbb{R}^n$. By the Baker-Campell-Hausdorff formula, $\exp(sX)\exp(tY)$ and $\exp(sX + tY + st[X, Y]/2)$ have the same expansion up to the term of degree 2. As a consequence, for small $t \in \mathbb{R}$ the following holds:

$$\exp(-tY)\exp(-tX)\exp(tY)\exp(tX)(p) = \exp(t^2[X, Y])(p) + o(t^2) \tag{2.3}$$

Roughly speaking, if X, Y are in the set of admissible directions along which we can move, then so is the commutator $[X, Y]$. Applying this to iterated commutators should suggest the validity of the following:

Theorem 2.2.1 (Chow-Rashevski Theorem)**.** If the bracket generating condition

$$rank\mathscr{L}(X_1, ..., X_r)(p) = n \tag{2.4}$$

holds at any $p \in \mathbb{R}^n$, then for any $x, y \in \mathbb{R}^n$ there exists an horizontal curve joining x and y (i.e., the CC distance is an actual distance).

Here $\mathscr{L}(X_1, ..., X_r)(p)$ denotes the Lie algebra generated by $X_1, ..., X_r$ together with the bracket operation, evaluated at the point p. This theorem was proved independently by W.L.Chow [6] and P.K.Rashevski [20].

In the theory of Partial differential equations, requirement (2.4) is also known as the Hörmander condition. From now on we assume it is always verified.

2.3 Length minimizers and First-order necessary conditions

Definition 2.3.1. An horizontal curve $\gamma : [0, 1] \longrightarrow \mathbb{R}^n$ is a *length minimizer* if $L(\gamma) = d(\gamma(0), \gamma(1))$.

As a consequence of the Ball-box theorem (see [19]), the topology induced by the CC distance is precisely the Euclidean topology of $\mathbb{R}^n$. Thanks to this fact, it is possible to prove the local (in the sense of the Euclidean topology) existence of length minimizers:

Theorem 2.3.2. For every $x \in \mathbb{R}^n$ there exist $\rho > 0$ such that if $d(x,y) < \rho$, then there exists a length minimizer connecting x and y.

However, contrary to the case of Riemannian geometry, length minimizers are in general not unique, not even locally.
To study length minimizers, we would like to derive necessary conditions for a horizontal curve to be length-minimizing. To do this, we begin with some definitions.
Given a system of vector fields $X = (X_1, ..., X_r)$, fix a length minimizer $\gamma : [0,1] \longrightarrow \mathbb{R}^n$ (defined on the whole interval [0,1]), with associated controls h. Without loss of generality assume $\gamma(0) = 0$ and $|h| = c$ (the curve is parametrized by constant speed). For any $x \in \mathbb{R}^n$, we define $\gamma_x : [0,1] \longrightarrow \mathbb{R}^n$ as the solution of

$$\begin{cases} \dot{\gamma}_x = h \cdot X(\gamma_x) \\ \gamma_x(0) = x \end{cases}.$$

It is then natural to define, for $t \in [0,1]$, the diffeomorphism (also known as *optimal flow*) $F : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ as $F_t(x) = \gamma_x(t)$.
On the other hand, for controls $k \in L^\infty([0,1], \mathbb{R}^r)$ denote by q_k the horizontal curve solving

$$\begin{cases} \dot{q}_k = k \cdot X(q_k) \\ q_k(0) = 0. \end{cases}$$

Finally, we define the *endpoint map* $End : L^\infty([0,1], \mathbb{R}^r) \longrightarrow \mathbb{R}^n$

$$End(k) := q_k(1),$$

and, for a given $v \in L^\infty([0,1], \mathbb{R}^r)$, the *variation map* $\phi_v : \mathbb{R} \longrightarrow \mathbb{R}^{n+1} = \mathbb{R}^n \times \mathbb{R}$ as

$$\phi_v(s) := \left(F_1^{-1}(q_{h+sv}(1)), \int_0^1 (h+sv)^2 \right) = (F_1^{-1} \circ End(h+sv), L_2(h+sv)^2).$$

In other words, the first component of $\phi_v(s)$ is the starting point of the curve with controls k, whose endpoint is the same as the curve with controls $h + sv$ starting at the origin. Denote $\frac{\partial \phi_v(s)}{\partial s}$ by $\phi'_v(s)$. We now prove the following lemma:

Lemma 2.3.3. If γ is a length minimizer parametrized by constant speed, then there exists a $\overline{\xi} \in \mathbb{R}^n \setminus \{0\}$ such that:

$$\langle \overline{\xi}, \phi'_v(0) \rangle = 0 \quad \forall v \in L^\infty([0,1], \mathbb{R}^r). \tag{2.5}$$

Proof. Suppose not, i.e., there exist $v_1, ..., v_{n+1}$ such that $\phi'_{v_1}(0), ..., \phi'_{v_{n+1}}(0)$ are linearly independent. Then the map

$$\Phi : \mathbb{R}^{n+1} \longrightarrow \mathbb{R}^{n+1}$$

$$\Phi(s_1, ..., s_{n+1}) := \left(F_1^{-1}(q_{h+s\cdot v}(1)), L_2(q_{h+s\cdot v})^2\right)$$

has inverible Jacobian $\nabla\Phi(0)$ because $\frac{\partial \Phi}{\partial s_i}(0) = \phi_{v_i}(0)$. By the inverse function theorem, Φ is an open map in a neighborhood of 0. Hence, there exist a vector $\overline{s} = (\overline{s_1}, ..., \overline{s_n})$ such that, by defining $\overline{h} = h + \overline{s} \cdot v$:

$$\begin{cases} F_1^{-1}(q_{\overline{h}}(1)) = F_1^{-1}(q_h(1)) \\ L_2(q_{\overline{h}})^2 < L_2(q_h)^2. \end{cases}$$

Since F_1 is a diffeomorphism, if $\overline{s}$ is sufficiently close to 0, then by local injectivity of F_1 $q_h(1) = q_{\overline{h}}(1)$, contradicting the minimality of $q_h = \gamma$. □

Lemma 2.3.4. If $v \in L^\infty([0,1], \mathbb{R}^r)$ and ϕ_v is the variation map associated with v, then

$$\phi_v'(0) = \left(\int_0^1 JF_t(0)^{-1}(v \cdot X(\gamma(t)))dt, 2\int_0^1 \langle h(t), v(t) \rangle \ dt \right) \in \mathbb{R}^n \times \mathbb{R} \tag{2.6}$$

Proof. Let $s \in \mathbb{R}$ be fixed. For any $t \in [0,1]$ we define $x_{h+sv}(t) = F_t^{-1}(q_{h+sv}(t))$. In particular, the first n components in $\phi_v(s)$ form the vector $x_{h+sv}(1)$. We may differentiate the following identity:

$$q_{h+sv}(t) = F_t(x_{h+sv}(t))$$

to obtain:

$$\begin{aligned} (h + sv) \cdot X(q_{h+sv}) &= h \cdot X(q_{h+sv}) + JF_t(x_{h+sv})\dot{x}_{h+sv} \\ \dot{x}_{h+sv} &= sJF_t(x_{h+sv})^{-1}[v \cdot X(q_{h+sv})], \end{aligned}$$

from which it follows that

$$x_{h+sv}(t) = s\int_0^t JF_\tau(x_{h+sv}(\tau))^{-1}[v \cdot X(F_\tau(x_{h+sv}(\tau)))]d\tau,$$

i.e.,

$$\begin{aligned} \phi_v'(0) &= \left(\left. \frac{\partial x_{h+sv}(1)}{\partial s} \right|_{s=0}, 2\int_0^1 \langle h(\tau), v(\tau) \rangle \, d\tau \right) \\ &= \left(\int_0^1 JF_\tau(x_h(\tau))^{-1}[v \cdot X(F_\tau(x_h(\tau)))]d\tau, 2\int_0^1 \langle h(\tau), v(\tau) \rangle \, d\tau \right). \end{aligned}$$

The desired identity (2.6) is easily achieved after observing that $x_h(\tau) = F_\tau^{-1}(q_h(\tau)) = F_\tau^{-1}(\gamma(\tau)) = 0$. □

We now state the main result of this section:

Theorem 2.3.5. (First-order necessary conditions) Let $\gamma : [0,1] \in \mathbb{R}^n$, $\gamma(0) = 0$ be a length minimizer, and let h be the associated controls. Assume γ is parametrized by constant speed, i.e. $|\gamma'| = |h| \equiv c$. Then there exists $\xi_0 \in \{0,1\}$, $\xi \in Lip([0,1], \mathbb{R}^n)$ such that

(i) $(\xi(t), \xi_0) \neq 0$ for any $t \in [0,1]$;

(ii) for any $j = 1, ..., r$, $\xi_0 h_j + \xi \cdot X_j(\gamma)$=0 holds a.e. on $[0,1]$;

(iii) $\dot{\xi} = -\left(\sum_{j=1}^{r} h_j JXj(\gamma)\right)^T \xi$ a.e. on $[0,1]$.

Proof. Let $\overline{\xi} \in \mathbb{R}^{n+1}\{0\}$ be as in Lemma 2.3.3, and express it as $\overline{\xi} := (\xi(0), \xi_0/2)$. Using Lemma 2.3.4, we deduce from (2.5) the following necessary condition:

$$\begin{aligned} 0 &= \int_0^1 \sum_{j=1}^{r} v_j(t)\{\xi(0), JF_t(0)^{-1}X_j(\gamma(t))) + \xi_0 h_j dt \\ &= \int_0^1 \sum_{j=1}^{r} v_j(t)\{\langle [JF_t(0)^{-1}]^T\xi(0), X_j(\gamma(t)))\rangle + \xi_0 h_j\}dt, \end{aligned} \tag{2.7}$$

For all $v \in L^\infty([0,1],\mathbb{R}^r$. By setting $\xi(t) = [JF_t(0)^{-1}]^T\xi(0)$ we automatically get (ii) thanks to the Fundamental lemma of the Calculus of Variations. (i) is trivial if $\xi_0 \neq 0$ (in which case we can normalize $(\xi(t), \xi_0)$ dividing by ξ_0, and (ii) still holds). Conversely, when $\xi_0 = 0$ we must have, by definition of $\overline{\xi} \neq 0$, that $\xi(0) \neq 0$, hence for every $t \in [0,1]$, $\xi(t) \neq 0$ as $JF_t(0)^{-1}$ is invertible. We are left to prove (iii). Differentiating $\xi(t) = [JF_t(0)^{-1}]^T\xi(0)$ with respect to t, we get:

$$0 = \left(\frac{d}{dt}JF_t(0)^T\right)\xi(t) + JF_t(0)^T\dot{\xi}(t), \quad \text{a.e. on } [0,1], \tag{2.8}$$

which leads to:

$$\dot{\xi}(t) = -(JF_t(0)^{-1})^T\left(\frac{d}{dt}JF_t(0)^T\right)\xi(t) \tag{2.9}$$

therefore we just need to compute:

$$\begin{aligned} \frac{d}{dt}JF_t(0) = J\frac{d}{dt}F_t(x)\Big|_{x=0} &= J\left(\sum_{j=1}^{r} h_j(t)X_j(F_t(x))\right)\Big|_{x=0} = \sum_{j=1}^{r} h_j(t)JX_j(F_t(0))JF_t(0) \\ &= \left(\sum_{j=1}^{r} h_j(t)JX_j(\gamma(t))\right)JF_t(0) \text{ a.e. on } [0,1]. \end{aligned} \tag{2.10}$$

Hence (iii) holds:

$$\dot{\xi}(t) = -\left(\sum_{j=1}^{r} h_j(t)JX_j(\gamma(t))\right)^T\xi(t). \tag{2.11}$$

□

Definition 2.3.6. A horizontal curve $\gamma : [0,1] \longrightarrow \mathbb{R}^n$ with $\gamma(0) = 0$ and associated controls h is an *extremal* if there exist $\xi_0 \in \{0,1\}$ and $\xi \in Lip([0,1],\mathbb{R}^n)$ such that the conditions (i), (ii), (iii) hold.
If $\xi_0 = 1$, we say that γ is a *normal* exremal.
If $\xi_0 = 0$, we say that γ is an *abnormal* extremal.

The curve $\xi \in Lip([0,1], \mathbb{R}^n)$ given by Theorem 2.3.5 is called a *dual curve* (or *dual variable*) of γ.
In general, the dual curve is not unique, and in particular an extremal might be both normal and abnormal. We will therefore call *strictly normal* (resp. *strictly abnormal*) the extremals which are normal but not abnormal (resp. abnormal but not normal).

Theorem 2.3.5 also possesses an Hamiltonian formulation: let us define the Hamiltonian function $H : \mathbb{R}^n \times \mathbb{R}^n \longrightarrow \mathbb{R}$

$$H(x, \xi) := \sum_{j=1}^{r} \langle X_j(x), \xi \rangle^2, \quad x, \xi \in \mathbb{R}^n. \tag{2.12}$$

Theorem 2.3.7. If γ is a normal extremal, with dual curve ξ, then (γ, ξ) is C^∞ smooth and solves the system of Hamiltonian equations:

$$\begin{cases} \dot{\gamma} = -\frac{1}{2}\frac{\partial H}{\partial \xi}(\gamma, \xi) \\ \dot{\xi} = \frac{1}{2}\frac{\partial H}{\partial x}(\gamma, \xi). \end{cases} \tag{2.13}$$

If γ is an abnormal extremal, with dual curve ξ, then $H(\gamma, \xi) = 0$.

This further emphasizes the duality between γ and ξ.

Proof. If γ is normal, then the following chains of equalities hold thanks to conditions (ii) and (iii) in Theorem 2.3.5:

$$\frac{\partial H}{\partial \xi}(\gamma, \xi) = 2\sum_{j=1}^{r} \langle X_j(\gamma), \xi \rangle X_j(\gamma) = 2\sum_{j=1}^{r}(-\xi_0 h_j) X_j(\gamma) = 2\sum_{j=1}^{r} -h_j X_j(\gamma) = -2\dot{\gamma}$$

$$\frac{\partial H}{\partial x}(\gamma, \xi) = 2\sum_{j=1}^{r} \langle X_j(\gamma), \xi \rangle \xi_j J X_j(\gamma) = 2\left(-\sum_{j=1}^{r} h_j J X_j(\gamma)\right)^T \xi,$$

which together prove the desired result.
If γ is abnormal, then by (ii) in Theorem 2.3.5:

$$H(\gamma, \xi) = \sum_{j=1}^{r} \langle X_j(\gamma), \xi \rangle^2 = \sum_{j=1}^{r}(-\xi_0 h_j)^2 = 0.$$

$\square$

With this in mind, it is possible to prove that normal extremals must be parametrized at constant speed, while in the case of an abnormal extremal γ, with dual curve ξ, quick calculations show that for every increasing Lipschitz continous homeomorphism $f : [0,1] \longrightarrow [0,1]$, a new parametrization of the same path $\tilde{\gamma} = \gamma \circ f$ is also an abnormal extremal, as $\tilde{\xi} = \xi \circ f$ is an associated dual curve. For these reasons, we do not request extremals to be parametrized at constant speed.

2.4 Properties of normal and abnormal extremals

2.4.1 Normal extremals

We begin with two fundamental results for normal extremals:

Proposition 2.4.1. Normal extremals are C^∞ smooth.

Proof. Suppose $\gamma : [0,1] \longrightarrow \mathbb{R}^n$ is a normal extremal with dual curve ξ and controls h. By (ii) in Theorem 2.3.5 we know that if ξ, γ are Lipschitz continuous, then so are the controls h. On the other hand, by (2.1) and (iii) in Theorem 2.3.5, we know that if h is Lipschitz continuous then so are $\dot{\gamma}, \dot{\xi}$. Iterating the argument, we have the following two implications:

$$\gamma, \xi \in C^i \implies h \in C^i([0,1], \mathbb{R}^r), \tag{2.14}$$

$$h \in C^i([0,1], \mathbb{R}^r) \implies \gamma, \xi \in C^{i+1}, \tag{2.15}$$

from which the claim follows by induction. □

Theorem 2.4.2. Any normal extremal $\gamma : [0,1] \longrightarrow \mathbb{R}^n$ is locally length minimizing, and in particular there exists $\delta > 0$ such that whenever $s, s' \in [0,1]$, $0 < s' - s < \delta$, we have that $\gamma|_{[s,s']}$ is a length minimizer.

The second part follows from the first by an easy compactness argument. This theorem, which is proved in [23], pp.27-29, is a special case of more general results in the context of Optimal Control theory.

Observe that in the Riemannian case $r = n$, we have that every extremal is strictly normal because of (i) and (ii) in Theorem 2.3.5. Moreover, (iii) is equivalent to the ODE of Riemannian geodesics, hence every extremal is a geodesic.

2.4.2 Abnormal extremals

The results in Subsection 2.4.1 underline many common features between normal extremals and Riemannian geodesics. Abnormal extremals, on the other hand, satisfy weaker conditions, that in general only provide Lipschitz regularity.

By (ii) in Theorem 2.3.5, for abnormal extremals we have

$$\langle \xi, X_j(\gamma) \rangle = 0 \text{ on } [0,1], \text{ for all } j = 1, ..., r. \tag{2.16}$$

If we request the extremal to be length minimizing, a stronger condition holds:

Theorem 2.4.3. (Goh condition) If γ is a *strictly* abnormal length minimizer, then any associated dual curve ξ satisfies:

$$\langle \xi, [X_i, X_j](\gamma) \rangle = 0 \text{ on } [0,1], \text{ for all } i, j = 1, ..., r. \tag{2.17}$$

The proof is found in [2].
Abnormal extremals are often introduced in the literature as the (curves whose controls are) singular points of the endpoint map *End*: more information about this characterization is in [24], pp.12-13.

Recently, a generalization of the Goh condition to longer iterated commutators was found in [5].

Theorem 2.4.3 admits the following consequence:

Corollary 2.4.4. If the horizontal distribution $X_1, ..., X_r$ has step 2, i.e., if

$$\dim span\{X_i, [X_j, X_k] : i, j, k \in \{1, ..., r\}\}(x) = n \quad \forall x \in \mathbb{R}^n,$$

then any length minimizer is C^∞ smooth.

Proof. Assume by contradiction that there is a length minimizer γ which is not C^∞ smooth; then, by Proposition 2.4.1, γ must be strictly abnormal. Because of Theorem 2.4.3 and (2.16), any dual variable ξ is orthogonal (along γ) to X_i, $[X_i, X_j]$ for any $i, j \in 1, ..., r$. Our assumption is that the former vector fields generate the whole tangent space at any point: this means that ξ is identically zero, which is absurd. □

The following is another remarkable fact about abnormal extremals, which is proved in [1]:

Theorem 2.4.5. Suppose the horizontal vectors $X_1, ..., X_r$ are analytic. Then the set of points Σ in $\mathbb{R}^n$ which can be connected to the origin by an abnormal length minimizer is a closed set with empty interior.

It is still unknown whether Σ always has measure zero (Morse-Sard problem).

2.5 Carnot groups

Consider a stratified group G, together with its Lie algebra $\mathfrak{g}$. The stratification of $\mathfrak{g} = V_1 \oplus ... \oplus V_s$, implies that the first stratum V_1 is generating for the bracket operation. This means that the subspace of T_eG corresponding to V_1 generates T_eG through the operation on the tangent vectors induced by the bracket. By left-invariancy, the same happens in T_pG, for every $p \in G$. This is equivalent to saying that $V_1 = \Delta$ is a generating sub-bundle of the tangent bundle of the manifold G (which we identify with $\mathbb{R}^n$ thanks to exponential coordinates of the second kind). In other words, the Hörmander condition is satisfied and (G, Δ) is a Carnot-Carathéodory structure with well-defined CC distance d.

For any $p, x, y \in G$ and any $r > 0$, there holds

$$d(p \cdot x, p \cdot y) = d(x, y) \quad \text{and} \quad d(\delta_r(x), \delta_r(y)) = rd(x, y).$$

By left-invariancy and the fact that the horizontal sub-bundle coincides with the first stratum, hence if γ is a horizontal curve s.t. $\gamma(0) = x, \gamma(1) = y$, then $\overline{\gamma}(t) = \delta_r(\gamma(t))$ is also a horizontal curve, and $L(\overline{\gamma}) = r \cdot L(\gamma)$.

As sub-Riemannian manifolds, Carnot groups also enjoy a purely metric characterization [11]:

Theorem 2.5.1. Carnot groups are the only metric spaces X which are:

(i) Locally compact;

(ii) Geodesic ($\forall x, y \in X$ there is a geodesic joining x and y);

(iii) Isometrically homogeneous ($\forall x, y \in X$ there is an isometry sending x to y);

(iv) Self-similar (admitting a dilatation).

Of course, $\mathbb{R}^n$ together with the vector fields $X_i := \partial_{x_i}$, $i = 1, ..., n$ is an Abelian Carnot group. However, the most popular example of non-Abelian Carnot group is the Heisenberg group: the Carnot group (therefore the only simply connected Lie group) associated with the Lie algebra of step 2 and rank 2 $\mathfrak{g} = V_1 \oplus V_2$ where $V_1 = \text{span}\{X_1, X_2\}$, $V_2 = \text{span}\{X_3\}$, and with

$$[X_2, X_1] = X_3 \quad [X_1, X_3] = 0 \quad [X_2, X_3] = 0$$

One of its representations is as $\mathbb{R}^3$ and $X_1 = \partial_{x_1}$, $X_2 = \partial_{x_2} - x_1\partial_{x_3}$, $X_3 = \partial_{x_3}$.

Clearly, the Heisenberg group is (up to isomorphism) the only stratified group of rank 2 and step 2.
We report part of a characterization found independently by Dimixier and Saito in 1957, which will be useful later:

Theorem 2.5.2. Let G be a real finite-dimensional Lie group, and let $\mathfrak{g}$ be its Lie algebra. The following are equivalent:

(i) exp is injective;

(ii) exp is a real analytic diffeomorphism;

(iii) G is solvable, simply connected and any quotient of $\mathfrak{g}$ does not admit a subalgebra isomorphic to $\mathfrak{e}$.

Here $\mathfrak{e}$ is the Lie algebra of the isometries of three-dimensional space $\mathbb{R}^3$.

Clearly, a Carnot group is solvable (it is nilpotent) and simply connected, and one can prove that $\mathfrak{e}$ cannot appear as a subalgebra of some quotient. Hence in a Carnot group the exponential map is injective.

2.5.1 Carnot groups as tangent spaces for sub-Riemannian manifolds

Our interest in Carnot groups is motivated by the role that they play in the context of sub-Riemannian geometry, similar to the one of Euclidean space in Riemannian geometry: they appear as "tangent spaces" (in a more general sense which we will briefly illustrate; for reference, see [23], pp.39-40) for a much larger class of metric spaces than just Riemannian manifolds.

Definition 2.5.3. The *Hausdorff distance* between two non-empty subsets of a metric space (X, d) is

$$d_H(A, B) := \max\left\{\sup_{x\in X}\inf_{y\in Y} d(x, y), \sup_{y\in Y}\inf_{x\in X} d(x, y)\right\}$$

Actually, d_H is not a distance ($d_H(A, \overline{A})$=0, moreover d_H can take the value $+\infty$), but it is when we restrict ourselves to compact subsets of X.

Definition 2.5.4. The *Gromov-Hausdorff distance* between two (non-empty) metric spaces is defined as the infimum:

$$d_{GH}(X, Y) := \inf d_H(i(X), j(Y))$$

taken among all possible isometric embeddings $i : X \longrightarrow Z$, $j : Y \longrightarrow Z$ in some metric space Z.

This quantity is well-defined (take $Z = i(X) \sqcup j(Y)$) and, again, it is not really a distance, but it becomes one if we reduce to the class of all non-empty compact metric spaces (up to isometry) [3].

Definition 2.5.5. A family $(X_\lambda, x_\lambda)_{\lambda>0}$ of pointed metric spaces is said to *converge* at (X, x) if, for every $R > 0$: $d_{GH}(B_R^{X_\lambda}(x_\lambda), B_R^X(x)) \longrightarrow 0$ as $\lambda \longrightarrow +\infty$.

Defining the *dilated metric space* $(\lambda X, d_{\lambda X})$ as $\lambda X := X$ and $d_{\lambda X} := \lambda d_X$ we are able to extend the notion of tangent space:

Definition 2.5.6. A pointed metric space (Y, y) is a *tangent space* for X at $x \in X$ if the family $(\lambda X, x)_{\lambda>0}$ converges to (Y, y).

Before stating our fundamental result, we need to establish one more concept. Given a sub-Riemannian manifold M, with generating sub-bundle $\Delta = \Delta^0 = \{X_1, ..., X_r\}$ for each $s \in \mathbb{N}$ we define the set of vector fields $\Delta^{s+1} = \{[X_i, X_j] : X_i \in \Delta^s, X_j \in \Delta\}$. Since the Hörmander condition holds, at each point p of M we have a flag of subspaces:

$$\Delta^0(p) \subset \Delta^1(p) \subset ... \subset \Delta^{g-1}(p) \subsetneq \Delta^g(p) = T_pM.$$

Set $d(i) = \dim \Delta^i(p)$. Then $(d(0), d(1), ..., d(g))$ is called the *growth vector* at p.

Definition 2.5.7. A point $p \in M$ is *equiregular* (w.r.t. Δ) if the growth vector is constant in a neighborhood of p.
A sub-Riemannian manifold is *equiregular* if each of its points is equiregular.

Theorem 2.5.8. (Mitchell [15], 1985) For each equiregular point $p \in M$, the unique tangent space to (M, d) at p is a Carnot group, where d is the Carnot-Carathéodory distance induced by Δ.

The situation for non-equiregular (also known as singular) points with respect to Δ is more complicated: a similar theorem which generalizes [15] was later proved by Bellaïche in 1994 [4].

Chapter 3

Non minimality of corners in sub-Riemannian geometry

3.1 Outline

In 2016, Enrico Le Donne and Eero Hakavuori provided a short proof [7] for the following statement:

Theorem 3.1.1. Length-minimizing curves in sub-Riemannian manifolds do not have corner-type singularities.

Previously, some partial results were known: in [14] (2008), the statement was proved for a certain class of equiregular manifolds, which satisfy an additional technical condition. These results were then extended in [12] (2015). More recently, a stronger result which implies Theorem 3.1.1 (namely, that any point of any length minimizer possesses at least one straight line as a tangent curve) was proven in [17] (2018).

As we will later explain, up to a desingularization, blow up and reduction argument, to prove Theorem 3.1.1 it suffices to verify its validity for a Carnot group. Moreover, as we will see, one may assume that the Carnot group has rank 2.

The proof is by induction on the step s of the group.
We will first show that the statement holds for $s = 2$. For $s \geq 3$ we are able to project the corner into a group of step $s - 1$ (Lemma 3.3.3); by inductive hypothesis, we can find a shorter curve in the group of step $s - 1$, which we can lift back (Lemma 3.3.4) to the original group modulo an error in the endpoint of degree s.
We will then correct the error by a system of curves placed along the corner, with endpoints in the subspace of degree $s - 1$ in Lemma 3.3.4, as a crucial consequence of the fact that the space is a nilpotent stratified group.
The conclusion will follow from the fact that the order s with which the error scales is strictly larger than the order $s-1$ with which the correction scales. Hence, after considering the situation at smaller scales, we will eventually show that an opportunely corrected version of the lifted curve has a length that differs from the length of the corner by a

quantity of the form:

$$-a\epsilon + b\epsilon^{s/(s-1)}$$

for some positive constants a, b. For a sufficiently small ϵ the new curve is therefore strictly shorter than the angle. The correction is achieved by modifying the initial corner with an ϵ-dilation of the lifted curve and suitable dilations of the correcting curves.

3.2 Preliminary definitions

Let $\gamma : [-1,1] \longrightarrow M$ be an absolutely continuous curve on a manifold M. We say that γ has a corner-type singularity at time 0 if the left and right derivatives at 0 exist and are linearly independent.
Let G be a Lie group. We say that a curve $\gamma : [1,1] \longrightarrow G$ is a *corner* if there exist linearly independent vectors X_1, X_2 in the Lie algebra of G such that

$$\gamma(t) = \begin{cases} \exp(-tX_1) & \text{if} \quad t \in [-1,0] \\ \exp(tX_2) & \text{if} \quad t \in (0,1] \end{cases}$$

In such a case, we will say that γ is the corner from $\exp(X_1)$ to $\exp(X_2)$. The length of such a corner is clearly $|X_1| + |X_2|$. Notice that at 0 the left derivative of γ is X_1, while the right derivative is X_2. Hence, a corner has a corner-type singularity at 0.

We call a norm $|\cdot|$ *strictly convex* if its unit ball contains no non-trivial segments. In other words, if $|x| = |y| = 1$ and $|x+y| = 2$, then $x = y$.

3.3 Preliminary lemmas

In the following, we are going to prove a slightly stronger statement, namely:

Theorem 3.3.1. Length-minimizing curves in every Carnot group equipped with a CC distance which comes from a strictly convex norm (on V_1) do not have corner-type singularities

Indeed, in the sub-Riemannian case the CC distance comes from the inner product of $\mathbb{R}^n$, which evidently induces a strictly convex norm. This will imply that length-minimizers do not have corner-type singularities in any CC space.

Lemma 3.3.2. *Let G be a step-2 Carnot group with a distance d associated to a strictly convex norm. Then in (G,d) no corner is length minimizing.*

Proof. Let X_1, X_2 be linearly independent vectors of the first layer V_1. Fix $\epsilon > 0$, and consider the points:

$$g_1 = \exp((\epsilon - 1)X_1) \qquad g_2 = \exp(\epsilon(X_2 - X_1)) \qquad g_3 = \exp((\tfrac{1}{2} - \epsilon)X_2)$$

$$g_4 = \exp(-\epsilon^2 X_1) \qquad g_5 = \exp(\tfrac{1}{2}X_2) \qquad g_6 = \exp(\epsilon^2 X_1)$$

Since G has step 2, the Baker-Campell-Hausdorff formula reduces to $\exp(X)\exp(Y) = \exp(X + Y + \frac{1}{2}[X, Y])$. Therefore it is easy to verify that $\exp(X_2) = \exp(X_1)g_1...g_6$. We may assume $|X_1| = |X_2| = 1$. Let $D = d(e, \exp(X_2 - X_1)) = |X_2 - X_1|$. By left invariancy and triangular inequality, we obtain $D = d(\exp(X_1), \exp(X_2)) = d(e, g_1...g_6) \leq \sum_{j=1}^{6} d(e, g_j)$, which we explictly compute as:

$$\sum_{j=1}^{6} d(e, g_j) = (1 - \epsilon) + \epsilon D + (\frac{1}{2} - \epsilon) + \epsilon^2 + \frac{1}{2} + \epsilon = 2 - (2 - D)\epsilon + 2\epsilon^2.$$

Crucially, by strict convexity of $|\cdot|$, we have that $-(2 - D) < 0$, hence taking ϵ small enough, we get that the corner is not length minimizing, as its length is 2. □

Remark that if $|\cdot|$ was the norm coming from the inner product in $\mathbb{R}^n$, then Corollary 2.4.4 would have already been a proof for our base case.

The next lemma encapsulates the argument's main idea. Geometrically, we can interpret its statement as the fact that curves in a quotient group can be isometrically lifted to the original group. This is the foundation of our inductive step, since we are able to lift a geodesic in the quotient and obtain a curve which only has error in the last layer.

Lemma 3.3.3. *Let G be a Carnot group of step s. Assume there are no minimizing corners in any Carnot group of step $s - 1$ with first layer isometric to the first layer of G. For all linearly independent $X_1, X_2 \in V_1$ there exists an $h \in \exp(V_s)$ such that:*

$$d(h \cdot \exp(X_1), \exp(X_2)) < |X_1| + |X_2|. \tag{3.1}$$

Proof. $H = \exp(V_s)$ is a central subgroup of G, as $[V_s, \mathfrak{g}] = \{0\}$ and so for $x \in H, y \in G$, we have $\exp(x)\exp(y) = \exp(x + y) = \exp(y + x) = \exp(x)\exp(y)$ by the Baker-Campell-Hausdorff formula. Therefore we can take the quotient group G/H, which is also a stratified group, but of step $s - 1$: the canonical projection π_* between the Lie algebras sends a generic element $x_1X_1 + ... + x_{r_{s-1}}X_{r_{s-1}} + x_{r_{s-1}+1}X_{r_{s-1}+1} + ... + x_nX_n$ of $\mathfrak{g}$ in $x_1X_1 + ... + x_{r_{s-1}}X_{r_{s-1}}$. We endow G/H with the unique norm that makes the canonical projection an isometry, thus making G/H a Carnot group. Its first layer is $\pi_*(V_1)$, which is isometric to V_1. Hence, by the inductive hypothesis, in G/H corners are never length-minimizers. For the same reason, if X_1, X_2 are independent, then so are $\pi_*(X_1)$ and $\pi_*(X_2)$, and by hypothesis the corner from $\exp(\pi_*(X_1))$ to $\exp(\pi_*(X_2))$, which has length $|X_1| + |X_2|$, is not length-minimizing. We have that $\exp(\pi_*(X)) = \pi(\exp(X))$ as π is a Lie group homomorphism. Hence

$$d(\pi(\exp(X_1)), \pi(\exp(X_2))) < |X_1| + |X_2|$$

But at the same time, by left-invariance of d:

$$d(\pi(\exp(X_1), \pi(\exp(X_2))) = d(H\exp(X_1), H\exp(X_2)) = \inf_{h \in H} d(h \cdot \exp(X_1), \exp(X_2)).$$

such an $h \in \exp(V_s)$ is an error of degree s. □

The next lemma is instead the technical core of the argument. We take care of the error coming from the previous lemma, by using vectors in the layer $s-1$. We also quantify how corrections vary by scaling the error.
In the following, we consider the conjugation map $C_p(q) = pqp^{-1}$

Lemma 3.3.4. Let G be a Carnot group of step $s \geq 3$ and let X_1, X_2 be vectors spanning V_1. Then for any $h \in \exp(V_s)$, there exist $Y_1, Y_2, Y_3 \in V_{s-1}$ such that

$$C_{\exp(X_1)}(\epsilon^s Y_1) \cdot C_{\exp(\frac{1}{2}X_2)}(\exp(\epsilon^s Y_2)) \cdot C_{\exp(X_2)}(\exp(\epsilon^s Y_3)) = \delta_\epsilon(h)$$

holds for all $\epsilon > 0$

Proof. For a fixed $Z \in V_s$, first look at the equation:

$$C_{\exp(X_1)}(Y_1) \cdot C_{\exp(\frac{1}{2}X_2)}(\exp(Y_2)) \cdot C_{\exp(X_2)}(\exp(Y_3)) = \exp(Z). \tag{3.2}$$

In the variables $Y_1, Y_2, Y_3 \in V_{s-1}$. Since G has step s, each of the conjugations can be rewritten using the Baker-Campell-Hausdorff formula:

$$C_{\exp(X)}(\exp(Y)) = \exp(X)\exp(Y)\exp(-X) = \exp(Y + [X,Y]).$$

From $s \geq 3$, we get $2(s-1) > s$ and so $\exp(V_{s-1}) \oplus \exp(V_s)$ is abelian, thanks to the Baker-Campell-Hausdorff formula: $\exp(X)\exp(Y) = \exp(X+Y)$. This means that exp is a group homomorphism from $\mathfrak{g}$ to G. At the same time, since G is a Carnot group, by Theorem 2.5.2 exp is injective, so from (3.2) we obtain a linear equation:

$$Y_1 + Y_2 + Y_3 + [X_1, Y_1] + [X_2, \frac{1}{2}Y_2 + Y_3] = Z. \tag{3.3}$$

Now we exploit the fact that V_1 is spanned by X_1 and X_2: the point is that, since $V_s = [V_1, V_{s-1}]$ any $Z \in V_s$ can be expressed as:

$$Z = [X_1, W_1] + [X_2, W_2].$$

With this in mind, our strategy to find solutions for (3.3) is to solve the system:

$$\begin{cases} Y_1 + Y_2 + Y_3 = 0 \\ Y_1 = W_1 \\ \frac{1}{2}Y_2 + W_2 = W_3, \end{cases}$$

whose (only) solution is $Y_1 = W_1$, $Y_2 = -2W_1 - 2W_2$, $Y_3 = W_1 + 2W_2$.
Now that we ensured the existence of a solution (Y_1, Y_2, Y_3) for any given $Z \in V_s$ to (3.2), suppose we are given some $h \in \exp(V_s)$. If $Z \in V_s$ is such that $h = \exp(Z)$, then $\forall \epsilon > 0$ we have $\delta_\epsilon(h) = \exp(\delta_\epsilon(Z)) = \exp(\epsilon^s Z)$. As (3.3) is linear, $(\epsilon^s Y_1, \epsilon^s Y_2, \epsilon^s Y_3)$ is a solution for (3.3) when we replace Z with $\epsilon^s Z$, resulting in the statement of the lemma. $\square$

3.4 Reduction to Carnot groups

We are now ready to prove Theorem 3.1.1.

Recall that the tangent space to a sub-Riemannian manifold (M, Δ) at one of its equiregular points is a Carnot group by Theorem 2.5.8. In this case, we are able to perform a "blow-up" procedure (as described in [18]) and reduce ourselves to considering the same problem in the tangent space (which is a Carnot group), because "blow-ups" of length minimizing curves with a corner-type singularity also are length minimizing and with a corner-type singularity, see [10], p.39. However, if some points of M are singular, then we first need a desingularization procedure.
Let γ be a curve in M, and suppose that Δ is generated by the orthonormal vector fields $X_1, ..., X_r$ near $\gamma(0)$. As shown in [10], p.49, there exist an equiregular sub-Riemannian manifold N, with an orthonormal frame $\xi_1, ..., \xi_r$ and a map $\pi : N \longrightarrow M$ onto a neighborood of $\gamma(0)$ such that $\pi_* \xi_i = X_i$ for $i = 1, ..., r$.

Assume that γ is length minimizing, it has a corner-type singularity at 0 and it is contained in $\pi(N)$. Write $\dot\gamma = \sum_{j=1}^r u_j X_j$ a.e. for suitable functions u_j, and let be σ a curve in N such that $\dot\sigma = \sum_{j=1}^r u_j \xi_j$: hence $\gamma = \pi \circ \sigma$ and the curves σ and γ have the same length (see [10],p.39-40, [18]). We are now done as:

- Since π does not alter distances, we deduce that σ also has to be a length minimizer.
- Considering that the projection map π is C^∞ smooth, and $\gamma = \pi \circ \sigma$, since by assumption γ has a corner-type singularity, then so has σ, which however has values in an equiregular manifold N.

For this reasons we can reduce to the case of a Carnot group.

3.5 The inductive non-minimality argument

As we saw it is sufficient to consider the case of a Carnot group G. We can also assume that G has rank 2: indeed, any corner is contained in a Carnot subgroup of rank 2, and if a corner is length minimizing among all the horizontal curves with image in G (with the same endpoints), then it is among the curves with image in any subgroup of G.

As we anticipated, we prove the result by induction on the step s of G. Lemma 3.3.2 is the base of the induction.

Let G a Carnot group be a rank-2 Carnot group of step s with a Carnot–Carathéodory distance coming from a strictly convex norm. Consider the corner from $\exp(X_1)$ to $\exp(X_2)$, for some linearly independent $X_1, X_2 \in V_1$, $|X_1| = |X_2| = 1$.
As in Lemma 3.3.3, we take the quotient with respect to the central subgroup $\exp(V_s)$, and we get a new Carnot group of step $s - 1$, whose first layer is isometric to the first

layer of G. Since the projection of our corner is still a corner, by Lemma 3.3.3 there is an $h \in V_s$ such that:

$$d(h \cdot \exp(X_1), \exp(X_2)) < |X_1| + |X_2| = 2.$$

By Lemma 3.3.4, there exist $Y_1, Y_2, Y_3 \in V_{s-1}$ satisfying:

$$(\delta_\epsilon(h))^{-1} C_{\exp(X_1)}(\epsilon^s Y_1) \cdot C_{\exp(\frac{1}{2}X_2}(\exp(\epsilon^s Y_2)) \cdot C_{\exp(X_2)}(\exp(\epsilon^s Y_3)) = e. \tag{3.4}$$

Similarly as to what we did proving Lemma 3.3.2, we claim that a set of points $g_1, ..., g_7$ is such that both

$$\exp(X_2) = \exp(X_1) g_1 ... g_7 \tag{3.5}$$

and

$$\sum_{j=1}^{7} d(e, g_j) < 2. \tag{3.6}$$

If (3.5) and (3.6) hold, the thesis would follow by left invariancy and triangular inequality as for proving 3.3.2. We fix $\epsilon > 0$, and we choose the following points:

$$\begin{aligned} g_1 &= \exp(\epsilon^s Y_1) = \delta_{\epsilon^{s/(s-1)}}(\exp(Y_1)), & g_2 &= \exp((-1-\epsilon)X_1) = \delta_{1-\epsilon}(\exp(-X_1)), \\ g_3 &= \exp(-\epsilon X_1)\delta_\epsilon(h)^{-1}\exp(X_2), & g_4 &= \exp((\frac{1}{2}-\epsilon)X_2) = \delta_{\frac{1}{2}-\epsilon}(X_2), \\ g_5 &= \exp(\epsilon^s Y_2) = \delta_{\epsilon^{s/(s-1)}}(\exp(Y_2)), & & \\ g_7 &= \exp(\epsilon^s Y_3) = \delta_{\epsilon^{s/(s-1)}}(\exp(Y_3)), & g_6 &= \exp(\frac{1}{2}X_2) = \delta_{\frac{1}{2}}(\exp(X_2)). \end{aligned}$$

The elements indicized by odd numbers correspond to movements in the directions of the Y_i, $i = 1, 2, 3$, while the other ones are scalings of X_1 and X_2. We first show (3.5). An explicit calculation yields:

$$\begin{gathered} \exp(X_1) g_1 ... g_7 = \exp(X_1)\exp(\epsilon^s Y_1)\exp(-(1-\epsilon)X_1)\exp(-\epsilon X_1)\delta_\epsilon(h)^{-1} \cdot \\ \exp(\epsilon X_2)\exp((\tfrac{1}{2}-\epsilon)X_2)\exp(\epsilon^s Y_2)\exp(\tfrac{1}{2}X_2)\exp(\epsilon^s Y_3) \cdot [\exp(-X_2)\exp(X_2)]. \end{gathered}$$

(3.7)

Recalling that $h \in Z(G)$ (and so is $\delta_\epsilon(h)$), thanks to the Baker-Campell-Hausdorff formula and the fact that $\forall X \in \mathfrak{g}$ we have $[X, X] = 0$, we are able to rewrite (3.7) in terms of conjugations as:

$$\delta_\epsilon(h)^{-1} C_{\exp(X_1)}(\exp(\epsilon^s Y_1)) C_{\exp(\frac{1}{2}X_2)}(\exp(\epsilon^s Y_2)) C_{\exp(X_2)}(\epsilon^s Y_3)\exp(X_2). \tag{3.8}$$

Now we exploit how we defined the Y_i, $i = 1, 2, 3$, in the beginning: (3.8) simplifies to $\exp(X_2)$ and (3.5) is proved. To prove (3.6), we use the fact that the points g_j, $j = 1, ..., 7$, are all dilations of some fixed points: we have

$$
\begin{aligned}
d(e, g_1) &= \epsilon^{s/(s-1)} d(e, \exp(Y_1)), \\
d(e, g_2) &= 1 - \epsilon, \\
d(e, g_3) &= \epsilon d(e, \exp(-X_1) h^{-1} \exp(X_2)) = \epsilon d(h \exp(X_1), \exp(X_2)), \\
d(e, g_4) &= \frac{1}{2} - \epsilon, \\
d(e, g_5) &= \epsilon^{s/(s-1)} d(e, \exp(Y_2)), \\
d(e, g_6) &= \frac{1}{2}, \\
d(e, g_7) &= \epsilon^{s/(s-1)} d(e, \exp(Y_3)).
\end{aligned}
$$

Summing all the above distances we obtain

$$
\sum_{j=1}^{7} d(e, g_j) = 2 - (2 - D)\epsilon + o(\epsilon), \tag{3.9}
$$

where $D = d(h \cdot \exp(X_1), \exp(X_2))$.

By our choice of h, we have that $2 - D > 0$. Hence letting $\epsilon \to 0^+$ we can make the quantity in (3.9) strictly less than 2. Formula (3.6) is now proved, and so is Theorem 3.1.1. $\square$

Bibliography

[1] A. Agrachev. "Any sub-Riemannian Metric has Points of Smoothness". In: (2008), pp.4–5. arXiv: `0808.4059 [math.DG]`.

[2] A. A. Agrachev and Y. L. Sachkov. "Control theory from the geometric viewpoint". In: *Encyclopaedia of Mathematical Sciences* (2004), Chapter 20.

[3] L. Ambrosio and P. Tilli. *Topics on analysis in metric spaces*. Oxford University Press, 2004.

[4] André Bellaïche. "The tangent space in sub-riemannian geometry". In: *Journal of mathematical sciences 83.4, pp. 461-476* (1994).

[5] F. Boarotto, R. Monti, and A. Socionovo. "Higher Order Goh Conditions for Singular Extremals of Corank 1". In: *Archive for Rational Mechanics and Analysis* (2024).

[6] W. L. Chow. "Über Systeme von linearen partiellen Differentialgleichungen erster Ordnung". In: *Math. Ann. 117* (1939), pp. 98–105.

[7] Eero Hakavuori and Enrico Le Donne. "Non-minimality of corners in subriemannian geometry". In: *Inventiones mathematicae* 206.3 (Apr. 2016), pp. 693–704.

[8] U. Hamenstädt. "Some regularity theorems for Carnot-Carathéodory metrics". In: *J.Differ.Geom 32* (1990), pp. 819–850.

[9] James. E. Humphreys. "Introduction to Lie Algebras and Representation Theory". In: Graduate Texts in Mathematics (1972).

[10] Frédéric Jean. *Control of Nonholonomic Systems: from Sub-Riemannian Geometry to Motion Planning*. 2014.

[11] Enrico Le Donne. *A metric characterization of Carnot groups*. 2014. arXiv: `1304.7493 [math.MG]`.

[12] Enrico Le Donne et al. "Corners in non-equiregular sub-Riemannian manifolds". In: *ESAIM Control Optim. Calc. Var.* 21.3 (2015), pp.625–634.

[13] John M. Lee. "Introduction to Smooth Manifolds". In: (2003), chapter 7.

[14] G. P. Leonardi and R. Monti. "End-point equations and regularity of sub-Riemannian geodesics". In: *Geom. Funct. Anal. 18, no.2, 552-582* (2008).

[15] John Mitchell. "On Carnot-Carathéodory metrics". In: *Journal of differential geometry 21.1* (1985).

[16] Richard Montgomery. "Abnormal Minimizers". In: *SIAM Journal on Control and Optimization* 32.6 (1994), pp. 1605–1620.

[17] Roberto Monti, Alessandro Pigati, and Davide Vittone. “Existence of tangent lines to Carnot-Carathéodory geodesics”. In: *Calc. Var. Partial Differential Equations 57, no.3, Art.75* (2018), p.18.

[18] Roberto Monti, Alessandro Pigati, and Davide Vittone. “On tangent cones to length minimizers in Carnot-Carathéodory spaces”. In: (Society for Industrial and Applied Mathematics, pp. 3351-3368. 2018).

[19] A. Nagel, E. M. Stein, and S. Wainger. “Balls and metrics defined by vector fields I: basic properties”. In: (1985), pp. 103–147.

[20] P. K. Rashevsky. “Any two points of a totally nonholonomic space may be connected by an admissible line”. In: *Uch. Zap. Ped. Inst. im. Liebknechta, Ser. Phys.Math., 2* (1938), pp. 83–94.

[21] Fulvio Ricci. “Harmonic Analysis on Nilpotent groups”. In: (2017/2018).

[22] R. S. Strichartz. “Sub-Riemannian geometry”. In: *J.Differ.Geom. 24(2)* (1986), pp.221–263.

[23] Alessandro Pigati (Master’s thesis). “New regularity results for sub-Riemannian geodesics”. In: (2016).

[24] Davide Vittone. “The regularity problem for sub-Riemannian geodesics”. In: *Part of the book series: Publications of Scuola Normale Superiore (CRMSNS, volume 17)* (2014).