Università degli Studi di Padova

Dipartimento di Matematica "Tullio Levi-Civita"
Corso di Laurea Magistrale in Matematica

Height estimate and Lipschitz-graph approximation of length minimizers

Tesi di Laurea Magistrale

Relatore:
Prof. Roberto Monti

Candidato:
Michele Zaccaron
Matricola 1132126

Sessione di Laurea del 20 aprile 2018
Anno Accademico 2017/2018

Contents

1 Introduction

The problem of the regularity of minimizers is a natural question that arises in the general field of calculus of variations. For many classical differential and variational problems it is not always wise to attack the issue with classical methods, but it is more useful to set the question in a weaker ambient, and to find a solution in this settings, where for example it can be shown relatively easily that a minimizer exists. Of course, once this minimizer is obtained, the successive step is to show that, due to its property of minimizing, it has more regularity than the ambient space in which it was found, trying to show that it is in fact regular enough to be a classical solution of the problem in the initial formulation.

The content of this work is focused in particular on the length-minimizers, or geodesics, in sub-Riemannian manifolds, particularly in Carnot groups, Lie groups with a stratified nilpotent Lie algebra. In a sub-Riemannian setting, some first results on such curves were shown by Strichartz in [15], who showed that under an assumption called the *strong bracket generating hypothesis*, every geodesic is locally a length minimizer. In the same paper Strichartz erroneously claimed also that every length minimizer was also a geodesic, or a regular minimizer, and concluded that, as in the case of Riemannian geometry, they were smooth. But this claim derived from a wrong use of the Pontryagin Maximum Theorem, specifically an omission of the special case of the theorem. Nonetheless, in [16] he corrected himself and the result holds if one assumes the strong bracket generating hypothesis. But the question whether there can exist "abnormal", non smooth, sub-Riemannian length minimizer in general assumptions remained open.

The answer to it arrived in 1991 when Montgomery in [10] showed an example of such an abnormal (although still smooth) length minimizer. In the following years numerous other examples emerged and it was finally understood that such non-regular minimizers are in fact not only existing but also not pathological, in the sense that are very common. This whole aspect and the events that lead to it is greatly discussed in [17], where they also show the ubiquity of non-smooth minimizers.

The question whether length minimzers in sub-Riemannian manifolds, that a priori are only Lipschitz regular, are in fact $\mathcal{C}^\infty$ smooth is still unanswered, but there have been various improvements in the understanding of what regularity they possess.

In the special case of Carnot groups, a particular type of sub-Riemannian manifolds, it is known that if the distribution is at most of step two, then the (constant-speed) minimizers are indeed smooth. Moreover, in [7] the problem in Carnot groups of step three is also faced through the use of certain polynomials, drawing the same conclusion.

In [18] Sussmann shows that in the setting of real-analytic manifolds, and consequently also Carnot groups, length minimizing arcs parameterized by arc-length are real-analytic on a open dense subset of times in their domains of definition.

Lately Le Donne and Hakavuori showed a general result in the most general setting of sub-Riemannian manifolds stating that length minimizers do not have *corner-type singularities* (See [5]). The paper relies on previous works made by Monti and Leonardi in [8]. Further results are [13], [4] and [2]. Other useful material that discuss the general problem of regularity of length minimizers in sub-Riemannian manifolds or in the special case of Carnot groups is found in [12], [19],

This paper inserts itself in this context of work, trying to give some results in the difficult problem of the characterization of length minimizers' regularity. It is partly inspired by the methods used in [9] through pages 290-319. It comes within the general framework discussed in [13], where Monti, Pigati and Vittone show that in Carnot-Carathéodory spaces, without any further assumpion on the space or on the length-minimizing curves, minimizers possess at any point at least one tangent curve.

This is why, in this work, we will explore the problem when there is a fixed direction (that will be denoted by $X_1 = \partial/\partial x_1$), and we will eventually arrive to the conclusion that, along this fixed direction, the length minimizer "looks like" the graph of a Lipschitz function.

The thesis will initially present the Carnot-Carathéodory metric, the natural metric of sub-Riemannian manifolds, and we will prove some of its properties that will be useful later in the dissertation. We shall see what admissible curves are, and these curves will be the main object of the results presented. Thereafter we will give the notion of Carnot group, and we will see how to endow $\mathbb{R}^n$ with such a structure. We will visit some useful properties and in particular the Hall method construction of a basis for this type of groups.

The first result is the so-called Height estimate, that gives an upper bound for a quantity of the curve in terms of its excess. In $\mathbb{R}^n$, we set the m vector fields

$$
\begin{aligned}
X_1(x) &= \frac{\partial}{\partial x_1}\\
X_2(x) &= \frac{\partial}{\partial x_2} + \sum_{j>2} p_{2,j}(x)\frac{\partial}{\partial x_j}\\
&\vdots\\
X_m(x) &= \frac{\partial}{\partial x_m} + \sum_{j>m} p_{m,j}(x)\frac{\partial}{\partial x_j}.
\end{aligned}
$$

where $p_{i,j}$ are monomials precisely described in Section 3. These m vector fields form a system of generators that endows $\mathbb{R}^n$ with a structure through which it is isomorphic to the nilpotent Lie algebra of step s on m generators. This structure also gives a "weight" $1 \leq w_i \leq s$ to every coordinate x_i, for all $i = 1, \dots, n$.
Call $X = (X_1, \dots, X_m)$; we say that a Lipschitz continuous curve $\gamma : [0,1] \to \mathbb{R}^n$ is X-admissible if $\dot{\gamma} = \sum_{j=1}^m h_j X_j(\gamma)$ a.e., where $h_j \in \mathrm{L}^\infty(0,1)$ for $j = 1, \dots, m$. We introduce the notion of excess for an admissible curve. Let $\langle \cdot, \cdot \rangle$ be the scalar product on the span of $X_1, \dots, X_m$ that makes the m vector fields orthonormal, and let $\gamma : [0,1] \to \mathbb{R}^n$ be a Lipschitz curve.

Definition. The *parametric excess* of γ at a point $\eta \in [0,1]$, at a scale $r > 0$ such that $\eta + r \leq 1$, in direction X_1 is

$$E(\gamma;\eta;r;X_1) := \frac{1}{r}\int_\eta^{\eta+r} \langle \dot{\gamma}(\theta) - X_1, \dot{\gamma}(\theta) - X_1 \rangle d\theta.$$

Theorem 1.1 (Height estimate)**.** *Let $X_1, \dots, X_m$ be defined as above. Let $\gamma : [0,1] \to \mathbb{R}^n$ be an X-admissible curve parameterized by arc-length, with $\gamma(0) = 0$. Let $r \leq 1$. Then for all $i = 2, ..., n$ there exist α_i, β_i positive integers such that:*

1) $\alpha_i + \beta_i + 1 = w_i$;

2) $\left(\dfrac{|\gamma_i(t)|}{|t|^{\alpha_i}}\right)^{\frac{1}{\beta_i+1}} \leq t\sqrt{E(\gamma;0;r;X_1)}$ *for all* $0 < t \leq r$.

An X-admissible curve γ is called a *length-minimizer* if for any other admissible curve ζ with same domain and same extremes, the length or total variation of γ is always equal or inferior to the one of ζ. We will reformulate this theorem through the definition of the (geometric) excess, or simply *excess*, of a curve, in the case that the curve is in fact a length-minimizer. Call Γ the *support* of the curve and τ_Γ its *unit tangent vector.*

Definition. The *excess* of Γ at the point $x \in \Gamma$, at a scale $r > 0$, in direction X_1 is

$$E(\Gamma; x; r; X_1) = \fint_{\Gamma \cap B_r(x)} \langle \tau_\Gamma - X_1, \tau_\Gamma - X_1 \rangle d\mathscr{H}^1.$$

where $\mathscr{H}^1$ is the 1-dimensional Hausdorff measure built from the Carnot-Carathéodory metric arising in $\mathbb{R}^n$ from the m vector fields $X_1, \ldots, X_m$.

Using this estimate we then prove the next result: the Lipschitz approximation of length minimizing curves. We set the projection map $\pi : \mathbb{R}^n \to \mathbb{R}$, $\pi(x_1, \ldots, x_n) = x_1$.

Theorem 1.2 (Lipschitz approximation). *Let $\gamma : [-1, 1] \to \mathbb{R}^n$ be a length minimizer parameterized by arch-length, with $\gamma(0) = 0$ and support Γ. For any $\varepsilon > 0$ there exist a closed set $I \subset \pi(\Gamma \cap B_{1/4})$ and a curve $\bar{\gamma} : I \to \mathbb{R}^n$ with support $\bar{\Gamma}$ such that:*

i) $\bar{\Gamma} \subset \Gamma$;

ii) $\bar{\gamma}_1(t) = t$ *for* $t \in I$, *i.e.* $\bar{\gamma}$ *is a graph along* X_1;

iii) $\left| \left(\bar{\gamma}(s)^{-1} \cdot \bar{\gamma}(t) \right)_i \right|^{1/w_i} \leq \varepsilon |t - s|$ *for* $s, t \in I$, $i = 2, ..., n$;

iv) $\mathscr{H}^1 \left(B_{1/4} \cap \bar{\Gamma} \smallsetminus \Gamma \right) \leq C(\varepsilon, \alpha_i, \beta_i) E(\Gamma; 0; 1; X_1)$;

v) $\mathscr{L}^1 \left(\pi \left(\Gamma \cap B_{1/4} \right) \smallsetminus I \right) \leq C(\varepsilon, \alpha_i, \beta_i) E(\Gamma; 0; 1; X_1)$.

As in the spirit of [9], where the results were concerning minimal surfaces, we hope that theorems 1.1 and 1.2 will be the starting point for further results and regularity theorems for length minimizing curves.

2 Brief introduction to Carnot-Carathéodory spaces

2.1 Carnot-Carathéodory metrics

Consider m vector fields $X_1, ..., X_m$ in $\mathbb{R}^n$, and assume they are pointwise linearly independent. Each vector field can be identified as an n-tuple, since we have the basis of the tangent bundle $\partial_1 = \partial/\partial x_1, \ldots, \partial_n = \partial/\partial x_n$, thus we can write $X_j(x) = \sum_{i=1}^n a_{ij}(x)\partial_i = (a_{1j}(x), ..., a_{nj}(x))$; we also assume that $a_{ij} \in C^\infty$ for $j = 1, .., m$ and $i = 1, ..., n$. We shall write the coefficients a_{ij} in the $n \times m$ matrix

$$\mathcal{A}(x) = \begin{pmatrix} a_{11}(x) & \ldots & a_{1m}(x) \\ \vdots & \ddots & \vdots \\ a_{n1}(x) & \ldots & a_{nm}(x) \end{pmatrix}.$$

We define its norm as

$$\|\mathcal{A}\| := \sup_{v \in \mathbb{R}^m, |v| \leq 1} |\mathcal{A}v|.$$

Definition 2.1. A Lipschitz continuous curve $\gamma : [0, T] \to \mathbb{R}^n$, $T \geq 0$, is said to be *X-admissible* if there exists a vector of measurable funtions $h = (h_1, ..., h_m) : [0, T] \to \mathbb{R}^m$ such that:

(i) $h_j \in \mathrm{L}^\infty(0, T)$ for all $j = 1, ..., m$;

(ii) $\dot{\gamma}(t) = \mathcal{A}(\gamma(t))h(t) = \sum_{j=1}^m h_j(t)X_j(\gamma(t))$ for a.e. $t \in [0, T]$.

In particular, the curve γ is said to be *X-subunit* if it is X-admissible and $\|h\|_\infty = \left\| \sqrt{\sum_{j=1}^m h_j^2} \right\|_\infty \leq 1$.

Notice that, since the vector fields are linearly independent in every point, the vector h is unique.

We now define the metric on Ω.

Definition 2.2. The Carnot-Carathéodory metric $d : \mathbb{R}^n \times \mathbb{R}^n \to [0, +\infty]$ on $\mathbb{R}^n$ is defined as follows:

$$\begin{aligned} d(x, y) = \inf \{ T \geq 0 : & \text{ there exists a } X\text{-subunit path } \gamma : [0, T] \to \Omega \\ & \text{such that } \gamma(0) = x \text{ and } \gamma(T) = y \} \end{aligned}$$

and if the above set is empty, then we set $d(x, y) = +\infty$.

It can be shown that if $d(x, y) < +\infty$ for all $x, y \in \mathbb{R}^n$, then d is indeed a metric on $\mathbb{R}^n$.

2.2 Properties of the metric d

Lemma 2.3. *Let $x_0 \in \mathbb{R}^n$ and $r > 0$. Define $B = B(x_0, r) = \{x \in \mathbb{R}^n : |x - x_0| < r\}$. Let $M = \sup_{x \in B} \|\mathcal{A}(x)\|$ and $\gamma : [0, T] \to \mathbb{R}^n$ be a X-subunit curve such that $\gamma(0) = x_0$. If $MT < r$ then $\gamma(t) \in B$ for all $t \in [0, T]$.*

Proof. Assume by contradiction that

$$\bar{t} := \inf\{t \in [0, T] : \gamma(t) \notin B\} \le T.$$

Then

$$\begin{aligned}
|\gamma(\bar{t}) - x_0| &= \left| \int_0^{\bar{t}} \dot{\gamma}(\theta) d\theta \right| = \left| \int_0^{\bar{t}} \mathcal{A}(\gamma(\theta)) h(\theta) d\theta \right| \\
&\le \int_0^{\bar{t}} |\mathcal{A}(\gamma(\theta)) h(\theta)| \, d\theta \le \int_0^{\bar{t}} \|\mathcal{A}(\gamma(\theta))\| |h(\theta)| \, d\theta \\
&\le \bar{t} M \le TM < r,
\end{aligned}$$

and therefore $\gamma(\bar{t}) \in B$, which is open. This is in contradiction with the definition of $\bar{t}$. □

Proposition 2.4. *Let $K \subset \mathbb{R}^n$ be a compact set and $X = (X_1, ..., X_m)$ the vector fields giving rise to the metric d. Then there exists a constant $\beta > 0$, depending on K and X, such that*

$$d(x, y) \ge \beta |x - y| \tag{2.1}$$

for all $x, y \in K$.

Proof. Let $\varepsilon > 0$, $K_\varepsilon = \{x \in \mathbb{R}^n : \min_{y \in K} |x - y| \le \varepsilon\}$, and $M = \sup_{x \in K_\varepsilon} \|\mathcal{A}(x)\|$. Take $x, y \in K$ and set $r = \min\{\varepsilon, |x - y|\}$. Let $\gamma : [0, T] \to \mathbb{R}^n$ be a X-subunit curve such that $\gamma(0) = x$ and $\gamma(T) = y$. Since $|\gamma(T) - \gamma(0)| = |x - y| \ge r$, by Lemma 2.3 we have $TM \ge r$. If $r = \varepsilon$ then

$$T \ge \frac{\varepsilon}{M} \ge \frac{\varepsilon}{MD} |x - y|,$$

where $D := \sup_{x, y \in K} |x - y|$. If $r = |x - y|$ then $T \ge |x - y| / M$. Since the subunit curve γ is arbitrary, by the definition of d we get

$$d(x, y) \ge \min\left\{\frac{1}{M}, \frac{\varepsilon}{MD}\right\} |x - y|.$$

□

With the above result we can show that d is indeed a metric.

Proposition 2.5. *If $d(x,y) < +\infty$ for all $x, y \in \mathbb{R}^n$, then $(\mathbb{R}^n, d)$ is a metric space.*

Proof. The symmetry property follows from the fact that if $\gamma : [0,T] \to \mathbb{R}^n$ is X-subunit then $\bar{\gamma}(t) = \gamma(T-t)$ is X-subunit too.

Moreover, if $\gamma_1 : [0,T_1] \to \mathbb{R}^n$ and $\gamma_2 : [0,T_2] \to \mathbb{R}^n$ are X-subunit curves such that $\gamma_1(0) = x, \gamma_1(T_1) = z, \gamma_2(0) = z$ and $\gamma_2(T_2) = y$ then

$$\gamma(t) = \begin{cases} \gamma_1(t), & \text{if } t \in [0,T_1] \\ \gamma_2(t-T_1), & \text{if } t \in [T_1, T_1+T_2], \end{cases}$$

is a X-subunit curve such that $\gamma(0) = x$ and $\gamma(T_1, T_1+T_2) = y$. Taking the infimum one finds the triangle inequality $d(x,y) \le d(x,z) + d(z,y)$.

Finally, $d(x,x) = 0$ and if $x \neq y$, from Proposition 2.4 it follows that $d(x,y) > 0$. □

We now turn to a different definition of d. Let $\gamma : [0,1] \to \mathbb{R}^n$ be an X-admissible curve with canonical vector of coordinates $h \in \mathrm{L}^\infty(0,1)^m$. Define

$$\mathrm{length}_1(\gamma) = \|h\|_1 = \int_0^1 \left|h(t)\right| dt,$$

and

$$d_1(x,y) = \inf\{\mathrm{length}_1(\gamma) : \gamma : [0,1] \to \mathbb{R}^n \text{ is an } X\text{-admissible curve such that } \gamma(0) = x \text{ and } \gamma(1) = y\}.$$

If the above set is empty put $d_1(x,y) = +\infty$.

Theorem 2.6. *For all $x, y \in \mathbb{R}^n$ the equality $d(x,y) = d_1(x,y)$ holds.*

Proof. See [12, page 20]. □

Now let $\gamma : [0,T] \to (\mathbb{R}^n, d)$ be an X-admissible curve with $\dot{\gamma}(t) = \mathcal{A}(\gamma(t))h(t)$, $h \in \mathrm{L}^\infty(0,T)^m$. The *total variation* of γ is

$$\mathrm{Var}(\gamma) = \sup_{0 \le t_1 < \ldots < t_k \le T} \sum_{i=1}^{k-1} d(\gamma(t_{i+1}), \gamma(t_i))$$

with the supremum taken over all finite partitions of [0,T].

Theorem 2.7. *Let $\gamma : [0,1] \to (\mathbb{R}^n, d)$ be a Lipschitz curve with canonical coordinates $h \in L^\infty(0,1)^m$. Then*

$$\lim_{\delta\to 0} \frac{d\big(\gamma(t+\delta),\gamma(t)\big)}{|\delta|} = \big|h(t)\big| \tag{2.2}$$

for a.e. $t \in [0,1]$. Therefore

$$\mathrm{Var}(\gamma) = \int_0^1 \big|h(t)\big|\,dt. \tag{2.3}$$

Proof. See [12, page 26]. □

Thanks to this result it can be seen that a curve is parameterized by arch-length in the metric space $(\mathbb{R}^n, d)$ if and only if $\big|h(t)\big| = 1$ for a.e. $t \in [0,T]$.

2.3 Lie groups, Lie algebras and Carnot groups

Definition 2.8 (Lie algebra)**.** A *Lie algebra* $\mathfrak{g}$ over $\mathbb{R}$ is a real vector field toghether with a $\mathbb{R}$-bilinear mapping $[\cdot,\cdot] : \mathfrak{g} \to \mathfrak{g} \times \mathfrak{g}$, called the Lie bracket, that satifies:

- $[X,X] = 0$ for all $X \in \mathfrak{g}$;
- $[X,[Y,Z]] + [Z,[X,Y]] + [Y,[Z,X]] = 0$ for all $X,Y,Z \in \mathfrak{g}$ (the Jacobi identity).

A consequence of these properties is that the Lie bracket is anticommutative, i.e., $[X,Y] = -[Y,X]$ for all elements $X,Y \in \mathfrak{g}$.

A linear subspace $\mathfrak{a}$ of $\mathfrak{g}$ is said to be a *Lie subalgebra* if it is closed under the Lie bracket, that is $[\mathfrak{a},\mathfrak{a}] \subseteq \mathfrak{a}$. If $\mathfrak{a}$ satisfies the stronger condition that $[\mathfrak{g},\mathfrak{a}] \subseteq \mathfrak{a}$ then it is called an *ideal* of $\mathfrak{g}$. Given a subset S of $\mathfrak{g}$, the Lie subalgebra generated by S is the smallest subalgebra that contains S. If a Lie algebra $\mathfrak{g}$ can be generated by m of its elements $E_1,\dots,E_m$, and if any other Lie algebra generated by m other elements $F_1,\dots,F_m$ is a homomorphic image of $\mathfrak{g}$ under the map $E_i \mapsto F_i$, we say that it is the *free* Lie algebra on m generators. The *lower central series* of a Lie algebra $\mathfrak{g}$ is the sequence of subalgebras recursively defined as follows:

$$\mathfrak{g}_i = [\mathfrak{g},\mathfrak{g}_{i-1}]$$

for $i > 1$, and $\mathfrak{g}_1 = \mathfrak{g}$. We have that $\mathfrak{g}_i \subseteq \mathfrak{g}_{i-1}$ for all $i > 1$, so this sequence is decreasing, and, since $[\mathfrak{g},\mathfrak{g}_i] \subseteq \mathfrak{g}_i$, it is a sequence of ideals. We say that the Lie algebra is *nilpotent of step* $s \in \mathbb{N}$ if $\mathfrak{g}_s \neq \{0\}$ and $\mathfrak{g}_{s+1} = \{0\}$. The free nilpotent Lie algebra $\mathfrak{g}_{m,s}$ on m generators of step s is the quotient of the free Lie algebra by the ideal $\mathfrak{g}_{s+1}$.

Definition 2.9 (Lie group)**.** A *Lie group* is a smooth differentiable manifold G that is also a group, and such that the group operations of product $\cdot : G \times G \to G, (x,y) \mapsto x \cdot y$ and inversion ${}^{-1} : G \to G, x \mapsto x^{-1}$ are smooth maps.

In the sequel we will always assume G to be connected and simply connected. If $g \in G$, let $\tau_g : G \to G$ be the left translation $x \mapsto g \cdot x$. We can associate to any Lie group a canonical Lie algebra $\mathfrak{g}$, which is the set of left invariant vector fields $X \in \Gamma(TG)$, i.e. X is a section of the tangent bundle of G such that

$$(Xf)(\tau_g(x)) = X(f \circ \tau_g)(x)$$

for all $x, g \in G$ and for all $f \in C^\infty(G)$. This set is a vector space and endowed with the usual commutator as Lie bracket, it becomes a Lie algebra, since commutator of left invariant vector fields is still a left invariant vector field.

Let $X \in \mathfrak{g}$ and consider the one-parameter subgroup $\gamma_X : \mathbb{R} \to G$ which is solution to the Cauchy problem

$$(\mathrm{CP}) \begin{cases} \dot{\gamma_X}(t) = X(\gamma_x(t)) \\ \gamma_X(0) = 0 \in G. \end{cases}$$

The integral curve γ_X is defined for all $t \in \mathbb{R}$ since left invariant vector fields are complete. The *exponential map* $\exp : \mathfrak{g} \to G$ is defined by $\exp(X) = \gamma_X(1)$. Define analogously $\exp(X)(g)$ taking $g \in G$ as the initial datum instead of the origin. Since X is left invariant, it can be seen that the integral curve of X with initial datum $g \in G$ is exactly the image of the curve γ solution of (CP) via the left translation map τ_g. In particular $\exp(X)(g) = \tau_g\big(\exp(X)\big) = g \cdot \exp(X)$, thus we obtain

$$\exp(Y) \cdot \exp(X) = \exp(X)(\exp(Y))$$

for all $X, Y \in \mathfrak{g}$.

Let $X \in \mathfrak{g}$ and define the map $\mathrm{ad}X : \mathfrak{g} \to \mathfrak{g}$ in the following way: $\mathrm{ad}X(Y) = [X,Y]$ for all $Y \in \mathfrak{g}$. If $\alpha = (\alpha_1, \dots, \alpha_k)$ is a multi-index of non negative integers, define $|\alpha| = \alpha_1 + \ldots + \alpha_k$ and $\alpha! = \alpha_1! \cdots \alpha_k!$. If α and β are multi-indices set

$$D_{\alpha\beta}(X,Y) = \begin{cases} (\mathrm{ad}X)^{\alpha_1}(\mathrm{ad}Y)^{\beta_1} \cdots (\mathrm{ad}X)^{\alpha_k}(\mathrm{ad}Y)^{\beta_k - 1} Y & \text{if } \beta_k \neq 0 \\ (\mathrm{ad}X)^{\alpha_1}(\mathrm{ad}Y)^{\beta_1} \cdots (\mathrm{ad}X)^{\alpha_k - 1} X & \text{if } \beta_k = 0, \end{cases}$$

and

$$c_{\alpha\beta} = \frac{1}{|\alpha + \beta| \alpha! \beta!}.$$

The Campbell-Hausdorff formula states that

$$\exp(X) \cdot \exp(Y) = \exp(P(X,Y)),$$

where

$$P(X,Y) = \sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k} \sum_{\alpha_j+\beta_j \geq 1} c_{\alpha\beta} D_{\alpha\beta}(X,Y).$$

The inner sum ranges over all $\alpha = (\alpha_1, \ldots, \alpha_k)$ and $\beta = (\beta_1, \ldots, \beta_k)$ such that $\alpha_i + \beta_i \geq 1$. It can be checked by direct computation that

$$P(X,Y) = X + Y + \frac{1}{2}[X,Y] + \frac{1}{12}[X,[X,Y]] + \frac{1}{12}[Y,[Y,X]] + R(X,Y) \quad (2.4)$$

where $R(X,Y)$ is a formal series of commutators with at least 4 times iterated brackets.

We say that the Lie group G is *nilpotent of step* $s \in \mathbb{N}$ if its Lie algebra $\mathfrak{g}$ it is. A nilpotent Lie group G is *stratified* if its Lie algebra $\mathfrak{g}$ can be written in the following way:

$$\mathfrak{g} = V_1 \oplus ... \oplus V_s$$

with $V_1, ..., V_s$ linear subspaces of $\mathfrak{g}$ such that $V_i = [V_1, V_{i-1}]$ for $i = 2,,, s$ and $V_{s+1} = \{0\}$. V_1 generates the whole algebra $\mathfrak{g}$ by iterated brackets. A *Carnot group* is simply a stratified Lie group.

Now suppose we have a Carnot group G with Lie algebra $\mathfrak{g}$. We can transport the Carnot group structure onto $\mathbb{R}^n$, where n is the dimension of $\mathfrak{g}$ as a vector space over the field $\mathbb{R}$.

Recall the stratification $\mathfrak{g} = V_1 \oplus ... \oplus V_s$, and set $m_i = \dim_{\mathbb{R}} V_i$. Fix a vector basis $X_1, \ldots, X_n$ of $\mathfrak{g}$ adapted to the stratification , i.e., if $M_i = m_1 + \ldots + m_i$ for all $i = 1, \ldots, s$ and $M_0 = 0$, then

$$X_{M_{j-1}+1}, \ldots, X_{M_j} \quad \text{is a base of } V_j \text{ for every } j = 1, \ldots, s.$$

Every basis element X_i will have its own weight $w_i \in \{1, \ldots, s\}$ defined as follows: if $X_i \in V_j$ for some $1 \leq j \leq s$, then $w_i = j$.

If $X, Y \in \mathfrak{g}$, then $X = \sum_{i=1}^n x_i X_i$ and $Y = \sum_{i=1}^n y_i Y_i$ for some $x = (x_1, \ldots, x_n), y = (y_1, \ldots, Y_n) \in \mathbb{R}^n$. We introduce a group law in $\mathbb{R}^n$, denoted by $\cdot$ (and not to be confused with the usual scalar product), in the following way:

$$x \cdot y = z$$

where $z = (z_1, \ldots, z_n)$ is the only n-tuple such that $P(X,Y) = \sum_{i=1}^n z_i X_i$. Equipped with this product, $(\mathbb{R}^n, \cdot)$ becomes a Lie group isomorphic to the Carnot group G, and whose Lie algebra is isomorphic to $\mathfrak{g}$. The identity is 0, and we have that $0 \cdot x = x \cdot 0 = x$ for all $0 \in \mathbb{R}^n$.

Finally, we now show some results that will be later used in Section 5. Define for all $\lambda > 0$ the group dilations $\delta_\lambda : \mathbb{R}^n \to \mathbb{R}^n$, $\delta_\lambda(x) = (\lambda^{w_1}x_1, \dots, \lambda^{w_n}x_n)$. Let now $X_1, \dots, X_n$ be a basis for $\mathfrak{g}$, write

$$X_j(x) = \sum_{i=1}^{n} a_{ij}(x)\partial_i, \quad j = 1, \dots, n,$$

and assume $X_j(0) = \partial_j$. The coefficients $a_{ij} \in \mathcal{C}^\infty(\mathbb{R}^n)$ and the product $x \cdot y = P(x,y)$ are linked in the following way. Let $\gamma : (-\eta, \eta) \to \mathbb{R}^n$ be a $\mathcal{C}^1$ curve such that $\gamma(0) = 0$ and $\dot\gamma(0) = \partial_j$. Since X_j is left invariant, if $f \in \mathcal{C}^1(\mathbb{R}^n)$ then

$$\begin{aligned} X_j f(x) = X_j(f \circ \tau_x)(0) &= \lim_{t \to 0} \frac{f(P(x,\gamma(t))) - f(P(x,0))}{t} \\ &= \frac{\partial f}{\partial x}(x)\frac{\partial P}{\partial y}(x,0)\dot\gamma(0) = \frac{\partial f}{\partial x}(x)\frac{\partial P}{\partial y_j}(x,0). \end{aligned}$$

The vector fields have polynomial coefficients $a_{ij}(x)$, and precisely

$$X_j(x) = \sum_{i=1}^{n} \frac{\partial P_i}{\partial y_j}(x,0)\partial_i.$$

As a consequence the following homogeneity property holds

$$a_{ij}(\delta_\lambda(x)) = \lambda^{w_i - w_j} a_{ij}(x), \tag{2.5}$$

where w_i and w_j are the weights of x_i and x_j respectively.

Proposition 2.10. *For all $x, y, z \in \mathbb{R}^n$ and $\lambda > 0$*

(i) $d(z \cdot x, z \cdot y) = d(x,y)$;

(ii) $d(\delta_\lambda(x), \delta_\lambda(y)) = \lambda(x,y)$.

Proof. Statement (i) follows from the fact that $\gamma : [0,T] \to \mathbb{R}^n$ is a subunit curve if and only if $z\cdot\gamma$ is a subunit curve joining $z\cdot x$ to $z\cdot y$, where $(z\cdot\gamma)(t) := z \cdot (\gamma(t))$.

Let now $\gamma : [0,T] \to \mathbb{R}^n$ be a subunit curve joining x to y.

$$\dot\gamma(t) = \sum_{j=1}^{m} h_j(t) X_j(\gamma(t)) = \sum_{i=1}^{n} \left(\sum_{j=1}^{m} h_j(t) a_{ij}(\gamma(t)) \right) \partial_i.$$

Define $\gamma_\lambda : [0, \lambda T] \to \mathbb{R}^n$ by $\gamma_\lambda(t) = \delta_\lambda(\gamma(t/\lambda))$. Then, by (2.5) with $w_j = 1$ if $j = 1, \dots, m$

$$\begin{aligned}\dot{\gamma}_\lambda(t) &= \sum_{i=1}^{n} \lambda^{w_i - 1} \left(\sum_{j=1}^{m} h_j(t/\lambda) a_{ij}(\gamma(t/\lambda)) \right) \partial_i \\ &= \sum_{i=1}^{n} \left(\sum_{j=1}^{m} h_j(t/\lambda) a_{ij}(\gamma_\lambda(t)) \right) \partial_i = \sum_{j=1}^{m} h_j(t/\lambda) X_j(\gamma_\lambda(t)).\end{aligned}$$

As $\gamma_\lambda(0) = \delta_\lambda(x)$, $\gamma_\lambda(\lambda T) = \delta_\lambda(y)$ and γ_λ is subunit, it follows that $d(\delta_\lambda(x), \delta_\lambda(y)) \le \lambda T$. Since γ was arbitrary, $d(\delta_\lambda(x), \delta_\lambda(y)) \le \lambda d(x, y)$; the converse inequality can be obtained in the same way. □

If $x \in \mathbb{R}^n$, introduce the homogeneous norm

$$\|x\| = \max \left\{ |x_i|^{1/w_i} : i = 1, ..., n \right\}.$$

Proposition 2.11. *There exist a constant $C_1 > 0$ such that for all $x, y \in \mathbb{R}^n$ we have*

$$\frac{1}{C_1} \left\| y^{-1} \cdot x \right\| \le d(x, y) \le C_1 \left\| y^{-1} \cdot x \right\|. \tag{2.6}$$

Proof. Consider $K = \left\{ x \in \mathbb{R}^n : d(x, 0) = 1 \right\}$. Since in $\mathbb{R}^n$ all norms are equivalent, K is closed and bounded, and then by Harel-Borel theorem, compact. Therefore the continuous function $\| \cdot \| : K \to [0, +\infty)$ admits minimum and maximum: for all $x \in K$, $q_1 \le \|x\| \le q_2$, with $0 < q_1 \le q_2$. If $0 \ne x \in \mathbb{R}^n$, $\left\| \delta_\lambda(x) \right\| = \lambda \|x\|$ for all $\lambda > 0$. Consequently, setting $\lambda = d(x, 0)$ and $\hat{x} = \delta_{\lambda^{-1}}(x)$, we have that $\hat{x} \in K$ thanks to property (ii) of (2.10) and therefore $q_1 d(x, 0) \le \|x\| \le q_2 d(x, 0)$. Using now the property (i) of the same proposition, we get

$$q_1 d(x, y) \le \left\| y^{-1} \cdot x \right\| \le q_2 d(x, y) \quad \text{for all } x, y \in \mathbb{R}^n,$$

from which the thesis is obtained by setting $C_1 = \max\{q_2, 1/q_1\}$. □

2.4 Chow-Hörmander condition

Given a real smooth manifold n-dimensional M and two smooth vector fields $X, Y \in \Gamma(TM)$, thought as derivations on the set $C^\infty(M)$, their Lie bracket, or Lie product, is defined as

$$[X, Y](f) = X(Y(f)) - Y(X(f)) \quad \text{for all } f \in C^\infty(M).$$

In coordinates, if $X = \sum_{i=1}^{n} a_i(x)\partial_i$ and $Y = \sum_{j=1}^{n} b_j(x)\partial_j$, then

$$[X,Y] = \sum_{i=1}^{n}\sum_{j=1}^{n}\left(a_j(x)\partial_j b_i(x) - b_j(x)\partial_j a_i(x)\right)\partial_i.$$

Endorsed with the Lie brackets, the real vector space $\Gamma(TM)$ of all vector fields on M has a structure of Lie algebra.

Now, starting from smooth vector fields $X_1, ..., X_m \in C^\infty(\mathbb{R}^n; \mathbb{R}^n)$, we can proceed to find new vector fields by iterating the Lie brackets. The Lie algebra generated with this procedure shall be denoted by $\mathcal{L}(X_1, ...X_m)$; for each $x \in \mathbb{R}^n$ this Lie algebra is a vector space $\mathcal{L}(X_1, ...X_m)(x)$. If

$$\operatorname{rank} \mathcal{L}(X_1, ...X_m) = n \qquad \text{for all } x \in \mathbb{R}^n \tag{2.7}$$

the vector fields $X_1, ..., X_m$ are said to satisfy the Chow-Hörmander condition. This means that at every point, the vector fields $X_1, ...X_m$ and their iterated Lie brackets generate the whole tangent space.

This condition is necessary for an important result concerning the relation between Carnot-Carathéodory metric and the usual Euclidean one.

Theorem 2.12. *Suppose $X_1, ...X_m \in C^\infty(\mathbb{R}^n; \mathbb{R}^n)$ satisfy the Chow-Hörmander condition. Let $K \subset \mathbb{R}^n$ be a compact set and assume that for all $x \in K$ condition* (2.7) *is guaranteed by iterated commutators of length less than or equal to s. Then there exists a constant $C > 0$, depending on the compact set K, such that*

$$d(x,y) \le C\,|x-y|^{1/s} \tag{2.8}$$

for all $x, y \in K$.

Proof. See [12, page 33]. □

This theorem, also known as the Chow–Rashevskii theorem, toghether with proposition (2.4), tells us that if $K \subset \mathbb{R}^n$ is a compact set and the family of vector fields $X = (X_1, ..., X_m)$ characterizing the metric d satisfies the above requirements, then for all $x, y \in K$

$$\frac{1}{C}\,|x-y| \le d(x,y) \le C\,|x-y|^{1/s}, \tag{2.9}$$

there exists a constant $C > 0$ depending on K and X. Therefore the euclidean topology of $\mathbb{R}^n$ and the one induced by the Carnot-Carathéodory metric are equivalent. Another important fact descending from this theorem is the connectivity of such a Carnot-Carathéodory space, i.e. for any two points x, y there exist an X-admissible curve connecting them.

3 Hall Basis

In this section we will show a procedure to construct a basis of the Lie algebra $\mathfrak{g}_{m,s}$, due to Hall.

Let $E_1, ..., E_m$ be the m generators of are elements of $\mathfrak{g}_{m,s}$, and let them be basis elements of weight 1. The rest of the basis is defined recursively: if we have defined basis elements of weights $1, ..., r-1$, they are simply ordered so that $E < F$ if weight(E)<weight(F). Also, if weight$(E) = q$ and weight$(F) = t$ and $r = q + t$, then $[E, F]$ is a basis element of weight r if:

1. E and F are basis elements and $E > F$;

2. if $E = [G, H]$, then $F \geq H$.

Fix now the number of generators m and the step $s \geq 1$ of the nilpotent Lie algebra $\mathfrak{g}_{m,s}$, and let n denote its dimension. Let $E_1, ..., E_m, ..., E_n$ be the Hall basis of $\mathfrak{g}_{m,s}$, and consider the linear subspaces V_i, for $i = 1, \ldots, s$, each one the span of the basis elements of weight i respectively. We now have a grading

$$\mathfrak{g}_{m,s} = V_1 \oplus V_2 \oplus .... \oplus V_s.$$

This induces a grading on the isomorphic vector space $\mathbb{R}^n$, by sending each $E_i \mapsto e_i$, where $\{e_1, \ldots, e_n\}$ is the canonical basis of $\mathbb{R}^n$, and consequently equipping it with a "graded coordinates system". Indeed for any element $x \in \mathbb{R}^n$, $x = \sum_{i=1}^m x_i e_i$ uniquely, and therefore, identifying any x element of $\mathbb{R}^n$ with a n-tuple $(x_1, ..., x_n)$, and remembering that every e_i has an associated weight $w_i = \text{weight}(E_i)$, we have

$$x = (\underbrace{x_1, ..., x_m}_{w=1}, \underbrace{x_{m+1}, ..x_t}_{w=2}, ..., \underbrace{x_k, ..., x_n}_{w=s}).$$

Number the basis elements for the Lie algebra by ordering them as explained above, i.e., $E_{m+1} = [E_2, E_1], E_{m+2} = [E_3, E_1], E_{m+3} = [E_3, E_2], E_{m+4} = [E_4, E_1]$, etc. Consider now a basis element E_i and write it as a bracket of lower order basis elements, $E_i = [E_{j_1}, E_{k_1}]$, where $j_1 > k_1$. Repeat this process of writing the left-most element as a bracket of always further lower basis elements, until we obtain

$$E_i = [[\cdots[[E_{j_p}, E_{k_p}]E_{k_{p-1}}], \cdots, E_{k_2}], E_{k_1}], \tag{3.1}$$

where $k_p < j_p \leq m$, and $k_{l+1} \leq k_l$ for $1 < l < p-1$. This expansion involves p brackets, and we write $\ell(i) = p$ and define $\ell(1) = ... = \ell(m) = 0$. We also associate to this expansion a multi-index $I(i) = (a_1, ..., a_n)$, with a_q defined by $a_q = \#\{t : k_t = q\}$. For the first m basis elements, their associate multi-index is $(0, ..., 0)$. We say that E_i is a *direct descendant* of each E_{j_t}, and we indicate this by writing $j_t \prec i$. Note that $\prec$ is a partial ordering. Moreover, to any index i we can associate another index $\Lambda_i \in \{1, ..., m\}$, being the index of the (unique) generator that has i as a direct descendant, that is $\Lambda_i \prec i$; if $i \in \{1, ..., m\}$ already, then set $\Lambda_i = i$. Notice that if $\Lambda_i = 1$ if and only if $i = 1$. If $E_i = [E_j, E_k]$, then $\Lambda_i = \Lambda_j$, $\ell(i) = \ell(j) + 1$ and each entry in $I(i)$ is at least as large as the corresponding entry in $I(j)$.

For every pair i and j with $j \prec i$, we define the monomial $p_{i,j}$ by

$$p_{i,j}(x) = \frac{(-1)^{\ell(i)-\ell(j)}}{(I(i) - I(j))!} x^{I(i)-I(j)} \tag{3.2}$$

The next theorem gives the connection between the abstract Lie algebra $\mathfrak{g}_{m,s}$ and the vector space $\mathbb{R}^n$

Theorem 3.1. *Fix $s \geq 1$ and $m \geq 2$ and let n denote the dimension of the free, nilpotent Lie algebra on m generators of step s. Then the derivations*

$$\begin{aligned}
E_1 &= \frac{\partial}{\partial x_1} \\
E_2 &= \frac{\partial}{\partial x_2} + \sum_{j \succ 2} p_{j,2} \frac{\partial}{\partial x_j} \\
&\vdots \\
E_m &= \frac{\partial}{\partial x_m} + \sum_{j \succ m} p_{j,m} \frac{\partial}{\partial x_j}
\end{aligned}$$

have the following properties:

1. *They are homogeneous of weight one with respect to the grading*

$$\mathbb{R}^n = V_1 \oplus \ldots \oplus V_s;$$

2. *the Hall basis elements E_i they generate satisfy $E_i(0) = \frac{\partial}{\partial x_i}$; in other words, E_1 through E_m are free to step s at 0;*

3. *the graded Lie algebra they generate is isomorphic to $\mathfrak{g}_{m,s}$.*

Proof. See [3]. □

Now we show a lemma that will be used in the next section.

Lemma 3.2. *Consider the whole Hall basis $E_1, \dots, E_m, \dots, E_n$ and suppose that for $i \in \{m+1, .., n\}$, the corresponding basis element E_i is of the form $E_i = [E_j, E_q]$ for some $1 \le q < j < i$. Then*

$$p_{i,\Lambda_i}(x) = -\frac{p_{j,\Lambda_i}(x)x_q}{\big(I(i)\big)_q}. \tag{3.3}$$

In particular $\left|p_{i,\Lambda_i}(x)\right| \le \left|p_{j,\Lambda_i}(x)x_q\right|$.

Proof. Indeed if we consider $E_i = [E_j, E_q]$ and we remember its decomposition as in (3.1), we have that

$$E_i = [[\cdots[[E_{j_p}, E_{k_p}]E_{k_{p-1}}], \cdots, E_{k_2}], E_{k_1}],$$

therefore $E_q = E_{k_1}$ and $E_j = [[\cdots[[E_{j_p}, E_{k_p}]E_{k_{p-1}}], \cdots, E_{k_2}]$. Moreover $\Lambda_i = \Lambda_j$ and $\ell(i) = \ell(j) + 1$ and all entries of $I(i)$ are equal to those of $I(j)$ except for the q-th one, $\big(I(i)\big)_q = \big(I(j)\big)_q + 1$. The thesis now follows immediately just by looking at (3.2). □

4 Height estimate

Consider $\mathbb{R}^n$ endowed with the structure of graded Lie algebra as shown in Section 3, so that we have $0 \leq m \leq n$ independent vector fields in $\mathbb{R}^n$ $X_1, ..., X_m$ defined as follows

$$\begin{aligned}
X_1 &= \frac{\partial}{\partial x_1} \\
X_2 &= \frac{\partial}{\partial x_2} + \sum_{j>2} p_{j,2}(x) \frac{\partial}{\partial x_j} \\
&\vdots \\
X_m &= \frac{\partial}{\partial x_m} + \sum_{j>m} p_{j,m}(x) \frac{\partial}{\partial x_j}.
\end{aligned}$$

Call $X = (X_1, ..., X_m)$ and let $\gamma : [0,1] \to \mathbb{R}^n$ be an X-admissible curve parameterized by arc-length in the metric space $(\mathbb{R}^n, d)$, such that for a.e. $\theta \in [0,1]$

$$\dot{\gamma}(\theta) = \sum_{i=1}^{m} h_i(\theta) X_i(\gamma(\theta)),$$

where $h_1, ..., h_m \in L^\infty\big([0,1]\big)$. Since γ is arc-length parameterized, we have that

$$h_1^2(\theta) + ... + h_m^2(\theta) = 1 \quad \text{for a.e } \theta \in [0,1]. \tag{4.1}$$

Let $\langle \cdot, \cdot \rangle$ be the scalar product on the span of $X_1, ..., X_m$ making them orthonormal. From now on we will use the norm induced by this scalar product, when dealing with vectors of the tangent bundle.

Definition 4.1. The *parametric excess* of γ at a point $\eta \in [0,1]$, at a scale $r > 0$ such that $\eta + r \leq 1$, in direction X_1 is

$$E(\gamma; \eta; r; X_1) := \frac{1}{r} \int_{\eta}^{\eta+r} \langle \dot{\gamma}(\theta) - X_1, \dot{\gamma}(\theta) - X_1 \rangle d\theta.$$

Notice that

$$\begin{aligned}
\langle \dot{\gamma}(\theta) - X_1, \dot{\gamma}(\theta) - X_1 \rangle &= \langle \dot{\gamma}(\theta), \dot{\gamma}(\theta) \rangle - 2\langle \dot{\gamma}(\theta), X_1 \rangle + 1 \\
&= 1 + 1 - 2\, h_1(\theta) = 2(1 - h_1(\theta)).
\end{aligned}$$

From (4.1) we deduce that

- $|h_i| \leq 1$ for all $i = 1, ..., m$;

– for all $i \neq 1$, $h_i^2 \leq 1 - h_1^2 = (1-h_1)(1+h_1) \leq 2(1-h_1)$;

– for $t \in [0,1]$ and for all $i \neq 1$, $\int_0^t |h_i(\theta)| d\theta \leq t\sqrt{\frac{1}{t}\int_0^t h_i(\theta)^2 d\theta} \leq$
$\leq t\sqrt{\frac{1}{t}\int_0^t 2(1-h_1(\theta))d\theta} = t\sqrt{E(\gamma;0;t;X_1)}$.

Theorem 4.2 (Height estimate). *Let $X_1,\dots,X_m$ vector fields on $\mathbb{R}^n$ defined as above. Let $\gamma : [0,1] \to \mathbb{R}^n$ be an X-admissible curve parameterized by arc-length, with $\gamma(0) = 0$. Let $r \leq 1$. Then for all $i = 2,...,n$ there exist α_i, β_i positive integers such that:*

1) $\alpha_i + \beta_i + 1 = w_i$;

2) $$\left(\frac{|\gamma_i(t)|}{|t|^{\alpha_i}}\right)^{\frac{1}{\beta_i+1}} \leq t\sqrt{E(\gamma;0;r;X_1)} \quad \textit{for all } 0 < t \leq r. \tag{4.2}$$

Proof. For semplicity, we shall use the notation $E(t) = E(\gamma;0;t;X_1)$. We will prove this result by induction on the weights of indices, and we will show that 1), 2) and

$$|p_{i,\Lambda_i}(\gamma(\theta))| \leq t^{\alpha_i}\left(t\sqrt{E(t)}\right)^{\beta_i} \qquad \text{for all } 0 \leq \theta \leq t \tag{4.3}$$

hold for every $i \in \{2,\dots,n\}$.

Initial step: For all $i = 2,..,m$, i.e. indices of weight $w_i = 1$,

$$|\gamma_i(t)| \leq \int_0^t |h_i(\theta)|\, d\theta \leq t\sqrt{E(t)} \iff \left(\frac{|\gamma_i(t)|}{|t|^0}\right)^{\frac{1}{1+0}} \leq t\sqrt{E(t)},$$

so 2) holds with $\alpha_i = 0$ and $\beta_i = 0$, and indeed we have that $\alpha_i + \beta_i + 1 = 0 + 0 + 1 = 1 = w_i$. Moreover if $i \in \{1,\dots,m\}$ then $p_{i,\Lambda_i} = p_{i,i} \equiv 1$, thus also (4.3) holds. Thus the initial step is proved.

Inductive step: Let i be of weight $w_i \geq 2$. Following Hall basis construction, i will be of the form $i = [j,q]$ for some $1 \leq q < j < i$ with weights w_j and w_q such that $w_j + w_q = w_i$. Λ_i is the (only) index in $\{2,...,m\}$ that has i as a direct descendant. Notice that Λ_i can't be 1 since $i \neq 1$, and that $\Lambda_i = \Lambda_j$.

By Lemma 3.2 we have that $|p_{i,\Lambda_i}(x)| \leq |p_{j,\Lambda_i}(x)x_q|$. Therefore, using (4.2) on γ_q and (4.3) on p_{j,Λ_i} inductively, we have that there exist $\alpha_q, \beta_q, \alpha_j, \beta_j$

positive integers, with $\alpha_j + \beta_j + 1 = w_j$ and $\alpha_q + \beta_q + 1 = w_q$, such that

$$\begin{aligned}
\left|p_{i,\Lambda_i}(\gamma(\theta))\right| &\le \left|p_{j,\Lambda_i}(\gamma(\theta))\right|\left|\gamma_q(\theta)\right| \\
&\le \begin{cases} t^{\alpha_j}\left(t\sqrt{E(t)}\right)^{\beta_j} t^{\alpha_q}\left(t\sqrt{E(t)}\right)^{\beta_q+1} & \text{if } q > 1 \\ t^{\alpha_j}\left(t\sqrt{E(t)}\right)^{\beta_j} t & \text{if } q = 1 \end{cases} \\
&\le t^{\alpha_i}\left(t\sqrt{E(t)}\right)^{\beta_i},
\end{aligned}$$

where in the first case we have set $\alpha_i := \alpha_j + \alpha_q$ and $\beta_i := \beta_j + \beta_q + 1$, and in the second one $\alpha_i := \alpha_j + 1$ and $\beta_i := \beta_j$. Notice that in both cases $\alpha_i + \beta_i + 1 = w_j + w_q = w_i$, as we wanted. And we therefore proved the induction step for (4.3).

At this point we have that $\dot{\gamma}_i = h_{\Lambda_i} p_{i,\Lambda_i}$ and so

$$\gamma_i(t) = \int_0^t h_{\Lambda_i}(\theta) p_{i,\Lambda_i}(\gamma(\theta)) d\theta.$$

Hence using estimate (4.3) (that we already proved to be true at this step) we obtain

$$\begin{aligned}
\left|\gamma_i(t)\right| &\le t^{\alpha_i}\left(t\sqrt{E(t)}\right)^{\beta_i} \int_0^t \left|h_{\Lambda_i}(\theta)\right| d\theta \\
&\le t^{\alpha_i}\left(t\sqrt{E(t)}\right)^{\beta_i} t\sqrt{E(t)} \\
&= t^{\alpha_i}\left(t\sqrt{E(t)}\right)^{\beta_i+1}
\end{aligned} \tag{4.4}$$

that becomes

$$\left(\frac{\left|\gamma_i(t)\right|}{t^{\alpha_i}}\right)^{\frac{1}{\beta_i+1}} \le t\sqrt{E(t)}.$$

□

Now denote the support of γ by $\Gamma = \gamma([0,1]) \subset \mathbb{R}^n$, and its unit tangent vector by $\tau_\Gamma = \dot{\gamma} \in \operatorname{span}\{X_1, \dots, X_m\}$.

For any subset $U \subseteq \mathbb{R}^n$ we call

$$\operatorname{diam}(U) = \sup\{d(x,y) : x, y \in U\}$$

the *diameter* of U, where d is the Carnot-Carathéodory metric introduced in Section 2; by definition we set $\operatorname{diam}(\varnothing) = 0$. Let S be a subset of $\mathbb{R}^n$ and $\delta > 0$ a real number, and define

$$\mathscr{H}^1_\delta(S) = \inf\left\{\sum_{i=1}^{\infty} \operatorname{diam}(U_i) : S \subseteq \bigcup_{i=1}^{\infty} U_i,\ \operatorname{diam}(U_i) < \delta\right\}.$$

It can be proved that each $\mathscr{H}^1_\delta$ is an outer measure. Since the map $\delta \mapsto \mathscr{H}^1_\delta(S)$ is increasing, the limit

$$\mathscr{H}^1(S) := \lim_{\delta \downarrow 0} \mathscr{H}^1_\delta(S) = \sup_{\delta > 0} \mathscr{H}^1_\delta(S)$$

exists (although it may be infinite). $\mathscr{H}^1$ is a measure, upon restriction onto the Carathéodory-measurable sets, and it is called the *Hausdorff 1-dimensional measure* of $(\mathbb{R}^n, d)$.

There is a connection between the total variation of the curve γ and the 1-dimensional Hausdorff measure of its support Γ.

Lemma 4.3. *Let (M, ν) be a metric space and let $\mathcal{H}^1$ be the 1-dimensional Hausdorff measure in the metric ν. If $\gamma : [a, b] \to M$ is continuous, then*

$$\mathcal{H}^1\left(\gamma([a, b])\right) > \nu\left(\gamma(a), \gamma(b)\right).$$

Proof. Define the auxiliary function $\varphi(x) := \nu\left(x, \gamma(a)\right)$. Then $\varphi : M \to \mathbb{R}$ is 1-Lipschitz, hence

$$\mathcal{H}^1\Big(\varphi\left(\gamma([a, b])\right)\Big) \leq \mathcal{H}^1\left(\gamma([a, b])\right).$$

On the other hand, since $\mathcal{H}^1$ coincides on $\mathbb{R}$ with outer Lebesgue measure, and $\varphi\left(\gamma([a, b])\right)$ is an interval of the kind $[0, \sigma]$, we obtain

$$\mathcal{H}^1\Big(\varphi\left(\gamma([a, b])\right)\Big) = \sup_{t \in [a,b]} \varphi\left(\gamma(t)\right) = \sup_{t \in [a,b]} \nu\left(\gamma(t), \gamma(a)\right) \geq \nu\left(\gamma(a), \gamma(b)\right)$$

and the proof is completed. □

Theorem 4.4. *Let (M, ν) be a metric space and suppose that $\gamma : [a, b] \to M$ is a Lipschitz curve with support Γ. Then*

$$\mathscr{H}^1(\Gamma) \leq \mathrm{Var}(\gamma), \tag{4.5}$$

and equation holds if γ is injective.

Proof. Due to the Reparametrisation Theorem ([1, page 63]), we can assume that $|\dot{\gamma}| = 1$ a.e. and $a = 0$, $b = \mathrm{Var}(\gamma)$. Let $\delta > 0$, choose $k \in \mathbb{N}$ such that $\mathrm{Var}(\gamma)/k < \delta$, and set $\rho := \mathrm{Var}(\gamma)/k$, $J_i := [i\rho, (i+1)\rho], i = 0, \ldots, k-1$.

Since γ is 1-Lipschitz, thus $\mathrm{diam}(\gamma(J_i)) \leq \mathrm{diam}(J_i) < \delta$, therefore

$$\mathscr{H}^1_\delta(\Gamma) \leq \sum_{i=0}^{k-1} \mathrm{diam}(J_i) = \mathrm{Var}(\gamma),$$

and (4.5) follows since $\delta > 0$ was arbitrary.

Now suppose that γ is injective, and choose $a \le t_0 < \cdots < t_k \le b$. We have from Lemma 4.3

$$\sum_{i=0}^{k-1} \nu\left(\gamma(t_{i+1}, \gamma(t_i)\right) \le \sum_{i=0}^{k-1} \mathcal{H}^1\left(\gamma([t_i, t_{i+1}])\right) \le \mathcal{H}^1\left(\gamma([a,b])\right),$$

where the last inequality rely on the injectivity of γ and the additivity of the Hausdorff measure. Since the partition $\{t_i\}$ was arbitrary, we deduce that

$$\mathrm{Var}(\gamma) \le \mathcal{H}^1\left(\gamma([a,b])\right),$$

hence the equality in (4.5) holds. □

This, combined with Theorem 2.7 allows us to see that if γ is an X-admissible curve with support Γ, for any $K \subset \Gamma$ compact,

$$\mathscr{H}^1(K) = \int_{\gamma^{-1}(K)} \left|h(\theta)\right| d\theta.$$

Definition 4.5. The *excess* of Γ at the point $x \in \Gamma$, at a scale $r > 0$, in direction X_1 is

$$E(\Gamma; x; r; X_1) = \fint_{\Gamma \cap B_r(x)} \langle \tau_\Gamma - X_1, \tau_\Gamma - X_1 \rangle d\mathscr{H}^1.$$

An X-admissible curve $\gamma : [a,b] \to (\mathbb{R}^n, d)$ is called a *length-minimizer* if for any other X-admissible curve ζ such that $\zeta(a) = \gamma(a)$ and $\zeta(b) = \gamma(b)$, then $\mathrm{Var}(\gamma) \le \mathrm{Var}(\zeta)$, i.e. the total variation of the curve γ coincides with the Carnot-Carathéodory distance $d(\gamma(a), \gamma(b))$ between its extremes. It is easy to see that this implies that if γ is a length-minimizer and we consider a point $x = \gamma(t) \in \Gamma$, if there exist times $t_1 < t < t_2$ such that $\gamma(t_1), \gamma(t_2) \in \partial B_r(x)$, where $B_r(x)$ is the ball centered in x of radius r in the metric d, then $\mathcal{H}^1(\Gamma \cap B_r(x)) = 2r$.

Corollary 4.6 (to Height estimate). *Let $\gamma : [0,1] \to \mathbb{R}^n$ be an X-admissible length minimizer parameterized by arc-length, with $\gamma(0) = 0$ and support Γ. Let $r \le 1$. Then for all $i = 2, \ldots, n$ there exist α_i, β_i positive integers such that:*

1) $\alpha_i + \beta_i + 1 = w_i$;

2) $$\left(\frac{\left|\gamma_i(t)\right|}{|t|^{\alpha_i}}\right)^{\frac{1}{\beta_i+1}} \le t\sqrt{E(\Gamma; 0; r; X_1)} \quad \text{for all } 0 < t \le r. \tag{4.6}$$

Proof. Observe that if Γ is a length minimizer and $d(0,\gamma(1)) > r$, then $\mathscr{H}^1(\Gamma \cap B_r(x)) = r$. Using this observation one can see that the two definitions of excess coincide, and then conclude. $\square$

Remark. The proof of the last corollary shows that in fact in order to obtain such a result we only need Γ to satisfies certain density estimates, without necessarily being a length minimizer; if there exist two constants $0 < c_1 \leq c_2$ such that

$$c_1 r \leq \mathscr{H}^1(\Gamma \cap B_r(x)) \leq c_2 r,$$

then (4.6) holds provided we insert an adequate constant in the right hand-side of the inequality.

5 Lipschitz approximation

Consider the group $(\mathbb{R}^n, \cdot)$ with the Carnot-Carathéodory metric d already discussed. Recall the homogeneous norm in $\mathbb{R}^n$

$$\|x\| = \max\left\{ |x_i|^{1/w_i} : i = 1, ..., n \right\}$$

and the group dilations

$$\delta_\lambda(x) = (\lambda^{w_1} x_1, \dots, \lambda^{w_n} x_n).$$

Define the projection $\pi : \mathbb{R}^n \to \mathbb{R}$ letting $\pi(x) = x_1$. Then it can be easily deducted from (2.4) that $\pi(x \cdot y) = (x \cdot y)_1 = x_1 + y_1 = \pi(x) + \pi(y)$. This means that $\pi : (\mathbb{R}^n, \cdot) \to (\mathbb{R}, +)$ is a group homomorphism. Moreover, we have

$$\left|\pi(x) - \pi(y)\right| = |x_1 - y_1| \le d(x, y).$$

Thus π is 1-Lipschitz from $(\mathbb{R}^n, d)$ to $\mathbb{R}$.

Theorem 5.1 (Lipschitz approximation)**.** *Let $\gamma : [-1, 1] \to \mathbb{R}^n$ be a length minimizer parameterized by arch-length, with $\gamma(0) = 0$ and support Γ. For any $\varepsilon > 0$ there exist a closed set $I \subset \pi(\Gamma \cap B_{1/4})$ and a curve $\bar\gamma : I \to \mathbb{R}^n$ with support $\bar\Gamma$ such that:*

i) $\bar\Gamma \subset \Gamma$*;*

ii) $\bar\gamma_1(t) = t$ *for* $t \in I$*, i.e.* $\bar\gamma$ *is a graph along* X_1*;*

iii) $\left|\left(\bar\gamma(s)^{-1} \cdot \bar\gamma(t)\right)_i\right|^{1/w_i} \le \varepsilon|t - s|$ *for* $s, t \in I$*,* $i = 2, ..., n$*;*

iv) $\mathscr{H}^1\left(B_{1/4} \cap \Gamma \setminus \bar\Gamma\right) \le C(\varepsilon, \alpha_i, \beta_i) E(\Gamma; 0; 1; X_1)$*;*

v) $\mathscr{L}^1\left(\pi\left(\Gamma \cap B_{1/4}\right) \setminus I\right) \le C(\varepsilon, \alpha_i, \beta_i) E(\Gamma; 0; 1; X_1)$.

Proof. Let $B_{1/4} = \{x \in \mathbb{R}^n : d(x, 0) < 1/4\}$. For $\eta > 0$ consider the set

$$\bar\Gamma = \left\{x \in \Gamma \cap B_{1/4} : E(\Gamma; x; r; X_1) \le \eta \text{ for all } 0 \le r \le 1/2\right\} \subset \Gamma.$$

Take points $x \in \Gamma \cap B_{1/4}$ and $y \in \bar\Gamma$, with $x \ne y$, and define $\lambda = d(x, y) > 0$. By the triangle inequality we have $\lambda \le 1/2$. The set

$$\Gamma_\lambda = \delta_{\frac{1}{\lambda}}(y^{-1} \cdot \Gamma)$$

is the support of a length-minimizing curve and $0 \in \Gamma_\lambda$. The point $z = \delta_{1/\lambda}(y^{-1}\cdot x)$ is in Γ_λ and by the Proposition 2.10, $d(z,0) = \frac{1}{\lambda}d(x,y) = 1$. By the height estimate (4.6) we have that for any $i \geq 2$

$$\left(\frac{|z_i|}{d(z,0)^{\alpha_i}}\right)^{\frac{1}{\beta_i+1}} \leq \sqrt{E(\Gamma_\lambda;0;1;X_1)} = \sqrt{E(\Gamma;y;\lambda;X_1)} \leq \sqrt{\eta}.$$

We also used the invariance properties of the excess. By (2.6), this in turn gives

$$\left|(y^{-1}\cdot x)_i\right| \leq \eta^{\beta_i/2+1/2} d(x,y)^{w_i} \leq C_1^{w_i}\eta^{\beta_i/2+1/2}\left\|y^{-1}\cdot x\right\|^{w_i}. \tag{5.1}$$

Depending on $\varepsilon > 0$, we choose $\eta > 0$ so small that for all $i = 2,...,n$ we have

$$C_1^{w_i}\eta^{\beta_i/2+1/2} \leq \min\left\{\varepsilon^{w_i}, \frac{1}{2}\right\} = \varepsilon^{w_i}. \qquad \text{(assume this on } \varepsilon) \tag{5.2}$$

In this way, the maximum norm is given by

$$\left\|y^{-1}\cdot x\right\| = \max_{j=1,\dots,n}\left|(y^{-1}\cdot x)_j\right|^{1/w_j} = \left|(y^{-1}\cdot x)_1\right|^{1/w_1} = |x_1 - y_1|,$$

and (5.1) becomes

$$\left|(y^{-1}\cdot x)_i\right|^{1/w_i} \leq \varepsilon|x_1 - y_1|, \quad i = 2,...,n. \tag{5.3}$$

The projection $\pi : \bar{\Gamma} \to \mathbb{R}$ is injective because $\pi(x) = \pi(y)$ means $x_1 = y_1$ and thus, by (5.3), we have $\left|(y^{-1}\cdot x)_i\right| = 0$ for all $i \geq 2$. This implies $y^{-1}\cdot x = 0$ and so $x = y$. Let $I = \pi\left(\bar{\Gamma}\right)$ and denote by $\pi^{-1} : I \to \bar{\Gamma}$ the inverse of the projection. We define the curve $\bar{\gamma} : I \to \mathbb{R}^n$ letting

$$\bar{\gamma}(t) = \pi^{-1}(t), \quad t \in I.$$

The support of $\bar{\gamma}$ is $\bar{\Gamma} \subset \Gamma$. This is i). Then we have $\bar{\gamma}_1(t) = \pi\left(\pi^{-1}(t)\right) = t$ for all $t \in I$. This is ii). Claim iii) follows from (5.3).

Next, we prove claim iv). For any point $x \in B_{1/4} \cap \Gamma \setminus \bar{\Gamma}$ there exist a radius $0 < r_x \leq 1/2$ such that

$$\frac{1}{2r_x}\int_{\Gamma\cap B_{r_x}(x)} |\tau_\Gamma - X_1|^2 \, d\mathscr{H}^1 = E(\Gamma;x;r_x;X_1) > \delta.$$

Then we have

$$B_{1/4}\cap\Gamma\setminus\bar{\Gamma} \subset \bigcup_{x\in B_{1/4}\cap\Gamma\setminus\bar{\Gamma}} B_{r_x/5}(x)\cap\Gamma.$$

By the 5-covering lemma there exists a sequence of points $x_k \in B_{1/4} \cap \Gamma \setminus \bar{\Gamma}$ such that, letting $r_k = r_{x_k}$, we have

$$B_{1/4} \cap \Gamma \setminus \bar{\Gamma} \subset \bigcup_{k \in \mathbb{N}} B_{r_k}(x_k) \cap \Gamma,$$

and the balls $B_{r_k/5}(x_k)$ are pair-wise disjoint. Thus obtain

$$\begin{aligned}
\mathscr{H}^1 \left(B_{1/4} \cap \Gamma \setminus \bar{\Gamma} \right) &\le \sum_{k \in \mathbb{N}} \mathcal{H}^1 \left(B_{r_k}(x_k) \cap \Gamma \right) = \sum_{k \in \mathbb{N}} 2 r_k \\
&\le \sum_{k \in \mathbb{N}} \frac{1}{\delta} \int_{\Gamma \cap B_{r_k}(x_k)} \left| \tau_\Gamma - X_1 \right|^2 d\mathcal{H}^1 \\
&\le \frac{1}{\delta} \int_{\Gamma \cap B_1} \left| \tau_\Gamma - X_1 \right|^2 d\mathscr{H}^1 = \frac{2}{\delta} E(\Gamma; 0; 1; X_1).
\end{aligned}$$

Finally, claim v) follows from iv) and the fact that the projection π is 1-Lipschitz. The set I may assumed to be closed, because all the claims are stable passing to the closure.

$\square$

References

[1] Luigi Ambrosio, Paolo Tilli, *Selected topics on "analysis in metric spaces"*, Scuola Normale Superiore, 2000.

[2] D. Barilari, Y. Chitour, F. Jean, D. Prandi, M. Sigalotti, *On the regularity of abnormal minimizers for rank 2 sub-riemannian structures*, preprint.

[3] Matthew Grayson, Robert Grossmann, *Models for free nilpotent Lie algebras*, Journal Algebra, Volume 35, 1990, pp. 177-191.

[4] Eero Hakavuori, Enrico Le Donne, *Blowups and Blowdowns of geodesics in Carnot groups*, preprint.

[5] Eero Hakavuori, Enrico Le Donne, *Non-minimality of corners in subriemannian geometry*, Inventiones mathematicae (2016), pp. 1-12.

[6] Enrico Le Donne, Gian Paolo Leonardi, Roberto Monti, Davide Vittone, *Corners in non-equiregularsub-Riemannian manifolds*, ESAIM Control Optim. Calc. Var. 21 (2015), pp.625-634

[7] E. Le Donne, G. P. Leonardi, R. Monti, D. Vittone, *Extremal polynomials in stratified groups*, Communications in Analysis and Geometry, accepted 2016.

[8] Gian Paolo Leonardi, Roberto Monti, *End-Point Equations and Regularity of Sub-Riemannian Geodesics*, Geometric and Functional Analysis 18.2 (2008), pp. 552-582.

[9] Francesco Maggi, *Sets of finite perimeter and geometric variational problems. An introduction to geometric measure theory.* Cambridge Studies in Advanced Mathematics, 135, Cambridge University Press, Cambridge, 2012.

[10] R. Montgomery, *Abnormal Minimizers*, SIAM Journal on Control and Optimization 32.6 (1994), pp. 1605-1620.

[11] Roberto Monti, *Distances, boundaries and surface measures in Carnot-Carathéodory spaces.* UTM PhDTS - Novembre 2001.

[12] Roberto Monti, *The regularity problem for sub-Riemannian geodesics*, Geometric Control Theory and sub-Riemannian Geometry, Ed. by Gianna Stefani et al. Cham: Springer International Publishing, 2014, pp. 313-332.

[13] R.Monti, A.Pigati, D.Vittone, *Existence of tangent lines to Carnot-Carathéodory geodesics*, preprint.

[14] R.Monti, A.Pigati, D.Vittone, *On tangent cones to length minimizers in Carnot-Carathéodory spaces*, preprint.

[15] Robert S. Strichartz, *Sub-Riemannian geometry*, J. Differential Geom. 24.2 (1986), pp. 221-263.

[16] Robert S. Strichartz, *Corrections to: "Sub-Riemannian geometry"*, J. Differential Geom. 30.2 (1989), pp. 595-596 .

[17] H. J. Sussmann, Whensheng Liu, *Shortest paths for sub-Riemannian metrics on rank 2 distributions*, Mem. Amer. Math. Soc. 118(564), 1995.

[18] H. J. Sussmann, *A regularity theorem for minimizers of real-analytic subriemannian metrics*, 53rd IEEE Conference on Decision and Control, Dec. 2014, pp. 4801-4806.

[19] Davide Vittone, *The regularity problem for sub-Riemannian geodesics*, Geometric Measure Theory and Real Analysis, Ed. by Luigi Ambrosio, Pisa: Scuola Normale Superiore, 2014, pp. 193-226.