SCUOLA INTERNAZIONALE
SUPERIORE di STUDI AVANZATI
International School
for Advanced Studies

Invariants, volumes and heat kernels in sub-Riemannian geometry

Supervisor
Prof. Andrei A. Agrachev

Candidate
Davide Barilari

Thesis submitted for the degree of
"Doctor Philosophiæ"

Academic Year 2010/2011

The essence of mathematics lies in its freedom.
(George Cantor, 1845-1918)

Contents

Introduction

Sub-Riemannian geometry can be seen as a generalization of Riemannian geometry under non-holonomic constraints. From the theoretical point of view, sub-Riemannian geometry is the geometry underlying the theory of hypoelliptic operators (see [32, 57, 70, 92] and references therein) and many problems of geometric measure theory (see for instance [18, 79]). In applications it appears in the study of many mechanical problems (robotics, cars with trailers, etc.) and recently in modern fields of research such as mathematical models of human behaviour, quantum control or motion of self-propulsed micro-organism (see for instance [15, 29, 34])

Very recently, it appeared in the field of cognitive neuroscience to model the functional architecture of the area V1 of the primary visual cortex, as proposed by Petitot in [87, 86], and then by Citti and Sarti in [51]. In this context, the sub-Riemannian heat equation has been used as basis to new applications in image reconstruction (see [35]).

For the richness of its geometrical and analytical properties and the great variety of its applications, during the last decades sub-Riemannian geometry drew an increasing attention, both on mathematicians and engineers.

Formally speaking, a sub-Riemannian manifold $(M, \Delta, \mathbf{g})$ is a smooth differentiable manifold M, endowed with a vector distribution Δ and a Riemannian structure $\mathbf{g}$ on it. From this structure, one derives a distance on M, the so-called *Carnot-Caratheodory metric.* For every pair of points p and q on M one consider the set of all horizontal curves, i.e. curves on the manifold that are tangent to the distribution Δ, that join p to q and define the distance as the infimum of the length of these curves, where the length is computed via the Riemannian structure.

In control language, sub-Riemannian geometry can be thought as a general framework for optimal control systems linear in the control and with quadratic cost. The first question that naturally arises in this context is the problem of *controllability*, i.e. whether it is possible or not to join every pair of points on the manifold by an horizontal curve.

At the end of the 30's, Raschevsky [89] and Chow [50], independently proved that a sufficient condition for controllability is that the Lie algebra generated by the horizontal vector fields generates all the tangent space to the manifold at every point. This condition, usually called *Lie bracket generating* condition, played subsequently a key role in different areas of mathematics.

Forty years later, in the celebrated work of Hörmander [67], the Lie bracket generating condition was proved to be a sufficient condition for the hypoellipticity of the second order differential operator in the form "sum of squares", even if the operator is not elliptic (for this reason it is also common to speak about *Hörmander condition*). Starting from this work, the interplay between sub-Riemannian geometry and the analysis of PDE's become stronger. In fact, many estimates and properties of the fundamental solution of second order degenerate elliptic operators (and on the heat kernel of the relevant heat operator) in terms of the sub-Riemannian distance built starting from the operator have been provided (see for instance [32, 57, 70, 92] and references therein). In some sense these results subserved a deeper investigation on the geometric structure of sub-Riemannian spaces and nowadays sub-Riemannian geometry is a fast-developing field of research on its own.

Sub-Riemannian geometry enjoys major differences from the Riemannian being a generalization of the latter at the same time: not all geodesics are solution of a first order Hamiltonian system, as in the Riemannian case, due to the presence of the so-called abnormal extremals and the cut locus starting from a point is always adjacent to the point itself. Because of these reasons small sub-Riemannian spheres are not smooth (and even the simply connectedness of small balls is still an open problem). The exponential map, which in the sub-Riemannian case is naturally defined on the cotangent space, is never a local diffeomorphism in a neighborhood of the origin. Moreover, the sub-Riemannian distance is even not Lipschitz at a point that is reached by only strictly abnormal minimizer of the distance. There exists a large amount of literature developing sub-Riemannian geometry and some typical general references are [3, 13, 25, 81].

Besides these general results on the geometric structure of sub-Riemannian geometry, the explicit computation of what is called *optimal synthesis*, i.e. the set of all geodesics starting from a given point, together with their optimal time, is in general very difficult to obtain. Usually the steps are the following:

- Apply first order necessary conditions for optimality (which in the case of sub-Riemannian manifolds are given by the Pontryagin Maximum Principle) to reduce the set of candidate optimal trajectories. This first step can be already very difficult since one should find solutions of a Hamiltonian system, which is not integrable in general.

- Use higher order necessary conditions to reduce further the set of optimal trajectories. This step usually leads to the computation of the conjugate locus, i.e. the set of points up to which geodesics are locally optimal.

- Prove that no strict abnormal extremal is optimal (for instance by using conditions such as the so called Goh condition [3, 13]). If one fails to go beyond this step, then one can hardly get an optimal synthesis, since no general technique exists to treat abnormal minimizers.

- Among all solutions of the first order necessary conditions, find the optimal ones. One has to prove that, for each point of a candidate optimal trajectory, there is no other trajectory among the selected ones, reaching that point. The first point after which a first order trajectory loses global optimality is called a *cut point*. The union of all cut points is the *cut locus*.

As a consequence of these difficulties, optimal syntheses in sub-Riemannian geometry have been obtained in few cases.

The most studied cases are those of left invariant sub-Riemannian metrics. The first optimal synthesis was obtained for the Heisenberg group in [63, 64]. Then complete optimal syntheses were obtained for the 3D simple Lie groups $SU(2)$, $SO(3)$, $SL(2)$, with the metric induced by the Killing form in [38, 39]. Recently, Yuri Sachkov also obtained the optimal synthesis for the group of motions of the plane $SE(2)$ (see [80, 95]).

In dimension larger than 3, only nilpotent groups have been attacked. Some results were obtained for the Engel and Cartan groups [93, 94].

When a Lie group structure is not available there are also some results: the optimal synthesis was obtained for a neighborhood of the starting point in the 3D contact case in [8, 10, 53] and in the 4D quasi-contact case in [48]. The optimal synthesis was obtained in the important Martinet nilpotent case, where abnormal minimizers can be optimal (see [5]). They also solved the problem for certain perturbations of this case where strictly abnormal minimizers occur (see [33]).

The heat equation on a sub-Riemannian manifold is a natural model for the description of a non isotropic diffusion process on a manifold. It is defined by the second order PDE

$$\frac{\partial}{\partial t}\psi(t,x) = \mathcal{L}\psi(t,x), \qquad \forall\, t > 0,\, x \in M, \tag{1}$$

where $\mathcal{L}$ is the *sub-Riemannian Laplacian*, also called *sub-Laplacian.* As we said, this is a hypoelliptic, but not elliptic, second order differential operator.

The first tool that one need to relate the geometry of the sub-Riemannian structure with the solution of the heat equation case is a “geometric” definition of the sub-Laplacian, analogously to the Laplace-Beltrami operator defined on a Riemannian manifold. This operator can be intrinsically defined as the divergence of the horizontal gradient. If $f_1, \ldots, f_k$ is a local orthonormal basis for the sub-Riemannian structure, this operator is written in the form “sum of squares” plus a first order part

$$\mathcal{L} = \sum_{i=1}^{k} f_i^2 + a_i f_i, \qquad k = \dim \Delta,$$

where $a_1, \ldots, a_k$ are suitable smooth coefficients which depends on the volume with respect which the divergence is computed.

Hence, the problem of which volume one should use when computing the divergence immediately arise from the very definition of sub-Laplacian. If we want the definition to be intrinsic, in the sense that it does not depend on the coordinate system and the choice of an orthonormal frame, we need a volume which is defined by the geometric structure of the manifold.

Before talking about the sub-Riemannian case, let us briefly discuss the Riemannian one. On a n-dimensional Riemannian manifold there are three common ways of defining an invariant volume. The first is defined through the Riemannian structure and it is the so called Riemannian volume, which in coordinates has the expression $\sqrt{g}\, dx^1 \ldots dx^n$, where g is the determinant of the metric. The second and the third ones are defined via the Riemannian distance and are the n-dimensional Hausdorff measure and the n-dimensional spherical Hausdorff measure. These three volumes are indeed proportional (the constant of proportionality is related to the volume of the euclidean ball, depending on the normalization, see e.g. [49, 55]), hence they are equivalent for the definition of the Laplacian.

In sub-Riemannian geometry, there is an equivalent of the Riemannian volume, the so called Popp's volume $\mathcal{P}$, introduced by Montgomery in his book [81] (see also [6]). The Popp volume is a smooth volume and was used in [6] to define intrinsically the sub-Laplacian on *regular* sub-Riemannian manifolds, i.e. when the dimension of the iterate distributions $\Delta^1 := \Delta, \Delta^{i+1} := \Delta^i + [\Delta^i, \Delta]$ does not depend on the point, for every $i \geq 1$.

Under the regularity assumption, the bracket generating condition guarantees that there exists (a mimimal) $m \in \mathbb{N}$, called *step* of the structure, such that $\Delta_q^m = T_qM$, for all $q \in M$. In [79], Mitchell proved that the Hausdorff dimension of M is given by the formula

$$Q = \sum_{i=1}^{m} ik_i, \qquad k_i := \dim \Delta_q^i - \dim \Delta_q^{i-1}.$$

In particular the Hausdorff dimension is always bigger than the topological dimension of M.

Hence, the Q-dimensional Hausdorff measure (and the spherical one) behave like a volume and are also available to compute the sub-Laplacian. It makes sense to ask if these volumes are equivalent to define the sub-Laplacian, e.g. if these volumes are proportional as in the Riemannian case. This problem was first addressed by Montgomery in his book [81].

For what concerns the relationship between the geometry and the analysis on sub-Riemannian spaces, one of the most challenging problem is to find the relation between the underlying geometric structure of the manifold (topology, curvature, etc.) and the analytical properties of the heat diffusion (e.g. the small time asymptotics of the heat kernel), in the same fashion as in Riemannian geometry. In the Riemannian case there is a well-known relation between the small time asymptotics of the heat kernel and the Riemannian curvature of the manifold (see for instance [91, 28]). Moreover the singularities of the sub-Riemannian distance (in particular the presence of the cut locus) reflects on the kernel of the hypoelliptic heat equation (see [85, 84]).

After Hörmander, many results and estimates on the heat kernel for hypoelliptic heat equations have been proved. Among them, a probabilistic approach to hypoelliptic diffusion can be found in [21, 31, 74], where the existence of a smooth heat kernel for such equations is given.

The existence of an asymptotic expansion for the heat kernel was proved, beside the classical Riemannian case, when the manifold is endowed with a time dependent Riemannian metric in [59], in the sub-Riemannian free case (when $n = k + \frac{k(k-1)}{2}$) in [42]. In [26, 75, 100] the general sub-Riemannian case is considered, using a probabilstic approach, obtaining different expansion depending on the fact that the points that are considered belong to the cut locus or not.

The same method was also applied in [27] to obtain the asymptotic expansion on the diagonal. In particular it was proved, for the sub-Riemannian heat kernel $p(t,x,y)$, that the following expansion holds

$$p(t,x,x) \sim \frac{1}{t^{Q/2}}(a_0 + a_1 t + a_2 t^2 + \ldots + a_j t^j + O(t^{j+1})), \qquad \text{for} \quad t \to 0, \qquad (2)$$

for every $j > 0$, where Q denotes the Hausdorff dimension of M.

Besides these existence results, the geometric meaning of the coefficients in the asymptotic expansion on the diagonal (and out of that) is far from being understood, even in the simplest 3D case, where the heat kernel has been computed explicitly in some cases of left-invariant structures on Lie groups (see [6, 22, 36, 99]). In analogy to the Riemannian case, one would expect that the curvature tensor of the manifold and its derivatives appear in these expansions.

The work presented in this thesis is a first attempt to go in this direction and answer to some of these questions. In particular we considered the problem of classifying of sub-Riemannian structures on three dimensional Lie groups, the Hausdorff volume in sub-Riemannian geometry and its relation with the optimal synthesis in the nilpotent approximation, the geometrically meaningful short-time asymptotic expansion for the heat kernel in the three dimensional contact case. The structure of the thesis is the following:

In Chapter 1 we introduce the basic definitions and some results about sub-Riemannian geometry, with a brief survey on sub-Riemannian geodesic and the nilpotent approximation.

In Chapter 2 we provide a complete classification of left-invariant sub-Riemannian structures on three dimensional Lie groups. Left-invariant structures on Lie groups are the basic models of sub-Riemannian manifolds and the study of such structures is the starting point to understand the general properties of sub-Riemannian geometry. The problem of equivalence for several geometric structures close to left-invariant sub-Riemannian structures on 3D Lie groups were studied in several publications (see [45, 46, 54, 96]).

Here we describe the two functional invariants of a three dimensional contact structure, denoted χ and κ, which plays the analogous role of Gaussian curvature for Riemannian surfaces. Then the classification of left-invariant sub-Riemannian structures on three dimensional Lie groups is provided in terms of these basic invariants (see also Theorem 2.1 and Figure 2.1 for details).

As a corollary of our classification we find a sub-Riemannian isometry between the nonisomorphic Lie groups $SL(2)$ and $A^+(\mathbb{R}) \times S^1$, where $A^+(\mathbb{R})$ denotes the group of orientation preserving affine maps on the real line, which we explicitly compute.

In Chapter 3 we address the problem of the volume in sub-Riemannian geometry described above, and we answer negatively to the Montgomery's open problem. We proved that the Radon-Nikodym derivative of the spherical Hausdorff measure with respect to a smooth volume (e.g. Popp's volume) is proportional to the volume of the unit ball in the nilpotent approximation. It is worth to notice that this result cover also the Riemannian case. Indeed, in that case the nilpotent approximation at different points is always isometric to the standard n-dimensional Euclidean space, hence the volume of the unit ball is constant.

We then prove that the density is always a continuous function and that it is smooth up to dimension 4, as a consequence of the uniqueness of normal forms for the nilpotent approximation for a fixed growth vector (see Theorem 3.12 for details). On the other hand, starting from dimension 5, the nilpotent approximation could depend on the point. We then focused on the corank 1 case, showing that in this case the density is $\mathcal{C}^3$ (and $\mathcal{C}^4$ on every smooth curve) but in general not $\mathcal{C}^5$. In particular the spherical Hausdorff measure and the Popp one are not proportional. These results rely on the explicit computation of the optimal synthesis and the volume of the nilpotent unit ball for these structures.

In Chapter 4 we study nilpotent 2-step, corank 2 sub-Riemannian metrics. We exhibit optimal syntheses for these problems and investigate then its consequences on the regularity of the density of the spherical Haussdorff measure with respect to Popp's one. It turns out that in general, the cut time is not equal to the first conjugate time (that was the case for corank 1 structures) but still has a simple explicit expression. Also we characterize those structures whose cut locus coincide with the first conjugate locus. As a byproduct of this study we get that the spherical Hausdorff measure is $\mathcal{C}^1$ in the case of a generic 6 dimensional, 2-step corank 2 sub-Riemannian metric.

In Chapter 5 we introduce the formal definition of sub-Laplacian, computing its expression in a local orthonormal frame. Then we relate the small time asymptotics for the heat kernel on a sub-Riemannian manifold to its nilpotent approximation, using a perturbative approach. We then explicitly compute, in the case of a 3D contact structure, the first two coefficients of the small time asymptotics expansion of the heat kernel on the diagonal, expressing them in terms of the two basic functional invariants χ and κ defined in Chapter 2.

The research presented in this PhD thesis appears in the following publications:

(B1) A. Agrachev, D. Barilari, *Sub-Riemannian structures on 3D Lie groups.* Journal of Dynamical and Control Systems, vol. 1, 2012.

(B2) A. Agrachev, D. Barilari and U. Boscain, *On the Hausdorff volume in sub-Riemannian geometry.* Calculus of Variations and PDE, 2011.

(B3) D. Barilari, U. Boscain and J. P. Gauthier, *On 2-step, corank 2 sub-Riemannian metrics.* Accepted on SIAM, Journal of Control and Optimization.

(B4) D. Barilari, *Trace heat kernel asymptotics in 3D contact sub-Riemannian geometry.* Accepted on Journal of Mathematical Sciences.

Other material that is related to these topics and that has been part of the research developed during the PhD, but is not presented here, is contained in the following preprints in preparation

(B5) D. Barilari, U. Boscain and R. Neel, *Small time asymptotics at the sub-Riemannian cut locus.* In preparation.

(B6) A. Agrachev and D. Barilari, *Curvature in sub-Riemannian geometry.* In preparation.

In the first paper we investigate the relation between the presence of the cut locus and the behavior of the asymptotics of the sub-Riemannian heat kernel, in the same spirit of [84, 85]. In the second one it is presented a general definition of curvature for sub-Riemannian manifolds, together with some applications.

Finally, an introduction to sub-Riemannian geometry from the Hamiltonian viewpoint is contained in the forthcoming book

(B) A. Agrachev, D. Barilari, and U. Boscain, *Introduction to Riemannian and sub-Riemannian geometry*, Lecture Notes, 179 pp. (2011).
`http://people.sissa.it/agrachev/agrachev_files/notes.html`

Acknowledgments

I owe my deepest gratitude to my supervisor Andrei Agrachev, who kindly supported me during the last five years of studies, always sharing his ideas and willing to be at disposal. The joy and enthusiasm he has for his research was contagious and motivational for me. Jointly with his human qualities, his guide was invaluable.

I am also definitively indebted to Ugo Boscain, not only for our collaborations and the trust he transmitted to me even during tough times in the Ph.D., but also for the friendship we developed. I strongly believe that this thesis would not have been possible without his support.

A very special thank goes to Jean-Paul Gauthier. Throughout my periods spent with him he provided to me an intense collaboration, made of good teaching, good company, and lots of good ideas.

It is a pleasure also to thank Robert Neel for our pleasant and fruitful collaboration during the last months.

I wish to thank also SISSA for providing me a stimulating and fun environment in which to learn and grow. It was a pleasure to share doctoral studies and life with such wonderful people and some of them are very close friends now. Among them, a special mention goes to Antonio Lerario, for the infinitely many mathematical discussions, all my flatmates and all the members of the "Pula team".

I would like to thank all my friends, the ones cited above and the others spread all over the world, for letting me sometimes to forget about mathematics.

Finally, I am really grateful to my parents, for always supporting me in all my decisions, and to Alice, for all her love and encouragement.

CHAPTER 1

Sub-Riemannian geometry

In this chapter we recall some preliminary definitions and results about sub-Riemannian geometry. For a more consistent presentation one can see [3, 13, 81, 25].

1.1 Sub-Riemannian manifolds

We start recalling the definition of sub-Riemannian manifold.

Definition 1.1. A *sub-Riemannian manifold* is a triple $\mathbf{S} = (M, \Delta, \mathbf{g})$, where

(*i*) M is a connected orientable smooth manifold of dimension $n \geq 3$;

(*ii*) Δ is a smooth distribution of constant rank $k < n$ satisfying the *bracket generating condition*, i.e. a smooth map that associates a point $q \in M$ with a k-dimensional subspace Δ_q of T_qM and we have

$$\mathrm{span}\{[X_1, [\dots [X_{j-1}, X_j]]](q) \mid X_i \in \overline{\Delta},\ j \in \mathbb{N}\} = T_qM, \quad \forall\, q \in M, \tag{1.1}$$

where $\overline{\Delta}$ denotes the set of *horizontal smooth vector fields* on M, i.e.

$$\overline{\Delta} = \{X \in \mathrm{Vec}(M) \mid X(q) \in \Delta_q \ \ \forall\ q \in M\}.$$

(*iii*) $\mathbf{g}_q$ is a Riemannian metric on Δ_q which is smooth as function of q. We denote the norm of a vector $v \in \Delta_q$ with $|v|$, i.e. $|v| = \sqrt{\mathbf{g}_q(v, v)}$.

A Lipschitz continuous curve $\gamma : [0, T] \to M$ is said to be *horizontal* (or *admissible*) if

$$\dot\gamma(t) \in \Delta_{\gamma(t)} \qquad \text{for a.e. } t \in [0, T].$$

Given an horizontal curve $\gamma : [0, T] \to M$, the *length of* γ is

$$\ell(\gamma) = \int_0^T |\dot\gamma(t)|\, dt. \tag{1.2}$$

The *distance* induced by the sub-Riemannian structure on M is the function

$$d(q_0, q_1) = \inf\{\ell(\gamma) \mid \gamma(0) = q_0, \gamma(T) = q_1, \gamma \text{ horizontal}\}. \tag{1.3}$$

The hypothesis of connectedness of M and the Hörmander condition guarantees the finiteness and the continuity of $d(\cdot, \cdot)$ with respect to the topology of M (Chow-Rashevsky theorem, see, for instance, [13]). The function $d(\cdot, \cdot)$ is called the *Carnot-Caratheodory distance* and gives to M the structure of metric space (see [25, 65]).

Remark 1.2. It is a standard fact that $\ell(\gamma)$ is invariant under reparameterization of the curve γ. Moreover, if an admissible curve γ minimizes the so-called *action functional*

$$J(\gamma) := \frac{1}{2}\int_0^T |\dot\gamma(t)|^2 dt.$$

with T fixed (and fixed initial and final point), then $|\dot\gamma(t)|$ is constant and γ is also a minimizer of $\ell(\cdot)$. On the other side, a minimizer γ of $\ell(\cdot)$ such that $|\dot\gamma(t)|$ is constant is a minimizer of $J(\cdot)$ with $T = \ell(\gamma)/v$.

Locally, the pair $(\Delta, \mathbf{g})$ can be given by assigning a set of k smooth vector fields spanning Δ and that are orthonormal for $\mathbf{g}$, i.e.

$$\Delta_q = \operatorname{span}\{f_1(q), \dots, f_k(q)\}, \qquad \mathbf{g}_q(f_i(q), f_j(q)) = \delta_{ij}. \tag{1.4}$$

In this case, the set $\{f_1, \dots, f_k\}$ is called a *local orthonormal frame* for the sub-Riemannian structure.

The sub-Riemannian metric can also be expressed locally in "control form" as follows. We consider the control system,

$$\dot q = \sum_{i=1}^m u_i f_i(q), \quad u_i \in \mathbb{R}, \tag{1.5}$$

and the problem of finding the shortest curve minimizing that joins two fixed points $q_0,\ q_1 \in M$ is naturally formulated as the optimal control problem,

$$\int_0^T \sqrt{\sum_{i=1}^m u_i^2(t)}\ dt \to \min, \qquad q(0) = q_0, \quad q(T) = q_1. \tag{1.6}$$

Definition 1.3. Let Δ be a distribution. Its *flag* is the sequence of distributions $\Delta^1 \subset \Delta^2 \subset \dots$ defined through the recursive formula

$$\Delta^1 := \Delta, \qquad \Delta^{i+1} := \Delta^i + [\Delta^i, \Delta].$$

A sub-Riemannian manifold is said to be *regular* if for each $i = 1, 2, \dots$ the dimension of Δ^i_q does not depend on the point $q \in M$.

Remark 1.4. In this paper we always deal with regular sub-Riemannian manifolds. In this case Hörmander condition can be rewritten as follows:

$$\exists \text{ minimal } \ m \in \mathbb{N} \quad \text{such that} \quad \Delta^m_q = T_q M, \quad \forall\, q \in M.$$

The sequence $\mathcal{G}(\mathbf{S}) := (\dim \Delta, \dim \Delta^2, \dots, \dim \Delta^m)$ is called *growth vector*. Under the regularity assumption $\mathcal{G}(\mathbf{S})$ does not depend on the point and m is said the *step* of the structure. The minimal growth is $(k, k+1, k+2, \dots, n)$. When the growth is maximal the sub-Riemannian structure is called *free* (see [81]).

A sub-Riemannian manifold is said to be *corank 1* if its growth vector satisfies $\mathcal{G}(\mathbf{S}) = (n-1, n)$. A sub-Riemannian manifold $\mathbf{S}$ of odd dimension is said to be *contact* if $\Delta = \ker \omega$, where $\omega \in \Lambda^1 M$ and $d\omega|_\Delta$ is non degenerate. A sub-Riemannian manifold M of even dimension is said to be *quasi-contact* if $\Delta = \ker \omega$, where $\omega \in \Lambda^1 M$ and satisfies $\dim \ker d\omega|_\Delta = 1$.

Notice that contact and quasi-contact structures are regular and corank 1.

A sub-Riemannian manifold is said to be *nilpotent* if there exists an orthonormal frame for the structure $\{f_1, \ldots, f_k\}$ and $j \in \mathbb{N}$ such that $[f_{i_1}, [f_{i_2}, \ldots, [f_{i_{j-1}}, f_{i_j}]]] = 0$ for every commutator of length j.

Definition 1.5. A *sub-Riemannian isometry* between two sub-Riemannian manifolds $(M, \Delta, \mathbf{g})$ and $(N, \Delta', \mathbf{g}')$ is a diffeomorphism $\phi : M \to N$ that satisfies

(*i*) $\phi_*(\Delta) = \Delta'$,

(*ii*) $\mathbf{g}(f_1, f_2) = \mathbf{g}'(\phi_* f_1, \phi_* f_2), \quad \forall\ f_1, f_2 \in \overline{\Delta}$.

A *local isometry* between two structures defined by the orthonormal frames $\Delta = \mathrm{span}(f_1, \ldots, f_k)$, $\Delta' = \mathrm{span}(g_1, \ldots, g_k)$ is given by a local diffeomorphism such that

$$\phi : M \to N, \quad \phi_*(f_i) = g_i, \qquad \forall\ i = 1, \ldots, k.$$

Remark 1.6. A sub-Riemannian structure on a Lie group G is said to be *left-invariant* if

$$\Delta_{L_x y} = L_{x*} \Delta_y, \qquad \mathbf{g}_y(v, w) = \mathbf{g}_{L_x y}(L_{x*} v, L_{x*} w), \qquad \forall x, y \in G.$$

where $L_x : y \mapsto xy$ denotes the left multiplication map on the group. In particular, to define a left-invariant structure, it is sufficient to fix a subspace of the Lie algebra $\mathfrak{g}$ of the group and an inner product on it.

We also remark that in this case it is possible to have in (1.4) a global equality, i.e. to select k globally linearly independent orthonormal vector fields.

1.2 Geodesics

In this section we briefly recall some facts about sub-Riemannian geodesics. In particular, we define the sub-Riemannian Hamiltonian.

Definition 1.7. A *geodesic* for a sub-Riemannian manifold $\mathbf{S} = (M, \Delta, \mathbf{g})$ is a curve $\gamma : [0, T] \to M$ such that for every sufficiently small interval $[t_1, t_2] \subset [0, T]$, the restriction $\gamma_{|_{[t_1, t_2]}}$ is a minimizer of $J(\cdot)$. A geodesic for which $\mathbf{g}_{\gamma(t)}(\dot\gamma(t), \dot\gamma(t))$ is (constantly) equal to one is said to be parameterized by arclength.

Let us consider the cotangent bundle T^*M with the canonical projection $\pi : T^*M \to M$, and denote the standard pairing between vectors and covectors with $\langle \cdot, \cdot \rangle$. The Liouville 1-form $s \in \Lambda^1(T^*M)$ is defined as follows: $s_\lambda = \lambda \circ \pi_*$, for every $\lambda \in T^*M$. The canonical symplectic structure on T^*M is defined by the closed 2-form $\sigma = ds$. In canonical coordinates (ξ, x)

$$s = \sum_{i=1}^{n} \xi_i dx_i, \qquad \sigma = \sum_{i=1}^{n} d\xi_i \wedge dx_i.$$

We denote the Hamiltonian vector field associated to a function $h \in C^\infty(T^*M)$ with $\vec{h}$. Namely we have $dh = \sigma(\cdot, \vec{h})$ and in coordinates we have

$$\vec{h} = \sum_i \frac{\partial h}{\partial \xi_i}\frac{\partial}{\partial x_i} - \frac{\partial h}{\partial x_i}\frac{\partial}{\partial \xi_i}$$

The sub-Riemannian structure defines an Euclidean norm $|\cdot|$ on the distribution $\Delta_q \subset T_qM$. As a matter of fact this induces a dual norm

$$\|\lambda\| = \max_{\substack{v\in\Delta_q \\ |v|=1}} \langle \lambda, v\rangle , \qquad \lambda \in T_q^*M,$$

which is well defined on $\Delta_q^* \simeq T_q^*M/\Delta_q^\perp$, where

$$\Delta_q^\perp = \{\lambda \in T_q^*M | \langle \lambda, v\rangle = 0, \forall\, v \in \Delta_q\}$$

is the annichilator of the distribution.

Here $\langle \cdot, \cdot \rangle$ denotes the standard pairing between vectors and covectors.

The *sub-Riemannian Hamiltonian* is the smooth function on T^*M, which is quadratic on fibers, defined by

$$H(\lambda) = \frac{1}{2}\|\lambda\|^2, \qquad \lambda \in T_q^*M.$$

If $\{f_1, \dots, f_k\}$ is a local orthonormal frame for the sub-Riemannian structure it is easy to see that

$$H(\lambda) = \frac{1}{2}\sum_{i=1}^k \langle \lambda, f_i(q)\rangle^2 , \qquad \lambda \in T_q^*M, \quad q = \pi(\lambda).$$

Remark 1.8. The sub-Riemannian Hamiltonian is a smooth function on T^*M which contains all the informations about the sub-Riemannian structure. Indeed it does not depend on the orthonormal frame selected $\{f_1, \dots, f_k\}$, i.e. is invariant for rotations of the frame, and the annichilator of the distribution at a point $\Delta_q^\perp$ can be recovered as the kernel of the restriction of h to the fiber T_q^*M

$$\ker H|_{T_q^*M} = \Delta_q^\perp.$$

It is a standard fact that H is also characterized as follows

$$H(\lambda) = \max_{v\in\Delta_q}\{\langle \lambda, v\rangle - \frac{1}{2}|v|^2\}, \qquad \lambda \in T^*M, \quad q = \pi(\lambda), \tag{1.7}$$

Let $\mathbf{S} = (M, \Delta, \mathbf{g})$ be a sub-Riemannian manifold and fix $q_0 \in M$. We define the *endpoint map* (at time 1) as

$$F : \mathcal{U} \to M, \quad F(\gamma) = \gamma(1),$$

where $\mathcal{U}$ denotes the set of admissible trajectories starting from q_0 and defined in $[0, 1]$. If we fix a point $q_1 \in M$, the problem of finding shortest paths from q_0 to q_1 is equivalent to the following one

$$\min_{F^{-1}(q_1)} J(\gamma), \tag{1.8}$$

where J is the action functional (see Remark 1.2). Then Lagrange multipliers rule implies that any $\gamma \in \mathcal{U}$ solution of (1.8) satisfies one of the following equations

$$\lambda_1 D_\gamma F = d_\gamma J, \tag{1.9}$$
$$\lambda_1 D_\gamma F = 0, \tag{1.10}$$

for some nonzero covector $\lambda_1 \in T^*_{\gamma(1)}M$ associated to γ. The following characterization is a corollary of Pontryagin Maximum Principle (PMP for short, see for instance [13, 37, 71, 88]):

Theorem 1.9. *Let γ be a minimizer. A nonzero covector λ_1 satisfies* (1.9) *or* (1.10) *if and only if there exists a Lipschitz curve $\lambda(t) \in T^*_{\gamma(t)}M$, $t \in [0,1]$, such that $\lambda(1) = \lambda_1$ and*

- *if* (1.9) *holds, then $\lambda(t)$ is a solution of $\dot{\lambda}(t) = \vec{H}(\lambda(t))$ for a.e. $t \in [0,1]$,*
- *if* (1.10) *holds, then $\lambda(t)$ satisfies $\sigma(\dot{\lambda}(t), T_{\lambda(t)}\Delta^\perp) = 0$ for a.e. $t \in [0,1]$.*

The curve $\lambda(t)$ is said to be an extremal associated to $\gamma(t)$. In the first case $\lambda(t)$ is called a normal extremal *while in the second one an* abnormal extremal*.*

Remark 1.10. It is possible to give a unified characterization of normal and abnormal extremals in terms of the symplectic form. Indeed the Hamiltonian H is always constant on extremals, hence $\lambda(t) \subset H^{-1}(c)$ for some $c \geq 0$. Theorem 1.9 can be rephrased as follows: any extremal $\lambda(t)$ such that $H(\lambda(t)) = c$ is a reparametrization of a characteristic curve of the differential form $\sigma|_{H^{-1}(c)}$, where $c = 0$ for abnormal extremals, and $c > 0$ for normal ones.

Also notice that, if $\lambda(t)$ is a normal extremal, then, for every $\alpha > 0$, $\lambda_\alpha(t) := \alpha\, \lambda(\alpha t)$ is also a normal extremal. If the curve is parametrized in such a way that $H(\lambda(t)) = \frac{1}{2}$ then we say that the extremal is arclength parameterized. Trajectories parametrized by arclength corresponds to initial covectors λ_0 belonging to the hypercylinder $\Lambda_{q_0} := T^*_{q_0}M \cap H^{-1}(\frac{1}{2}) \simeq S^{k-1} \times \mathbb{R}^{n-k}$ in $T^*_{q_0}M$.

Remark 1.11. From Theorem 1.9 it follows that $\lambda(t) = e^{t\vec{H}}(\lambda_0)$ is the normal extremal with initial covector $\lambda_0 \in \Lambda_{q_0}$. If $\pi : T^*M \to M$ denotes the canonical projection, then it is well known that $\gamma(t) = \pi(\lambda(t))$ is a geodesic (starting from q_0). On the other hand, in every 2-step sub-Riemannian manifold all geodesics are projection of normal extremals, since there is no strict abnormal minimizer (see Goh conditions, [13]).

The following proposition resumes some basic properties of small sub-Riemannian balls

Proposition 1.12. *Let* **S** *be a sub-Riemannian manifold and $B_{q_0}(\varepsilon)$ the sub-Riemannian ball of radius ε at fixed point $q_0 \in M$. For $\varepsilon > 0$ small enough we have:*

(i) $\forall\, q \in B_{q_0}(\varepsilon)$ *there exists a minimizer that join q and q_0,*

(ii) $\operatorname{diam}(B_{q_0}(\varepsilon)) = 2\varepsilon$.

Claim (i) is a consequence of Filippov theorem (see [13, 40]). To prove (ii) it is sufficient to show that, for ε small enough, there exists two points in $q_1, q_2 \in \partial B_{q_0}(\varepsilon)$ such that $d(q_1, q_2) = 2\varepsilon$.

To this purpose, consider the projection $\gamma(t) = \pi(\lambda(t))$ of a normal extremal starting from $\gamma(0) = q_0$, and defined in a small neighborhood of zero $t \in]-\delta, \delta[$. Using arguments of Chapter 17 of [13] one can prove that $\gamma(t)$ is globally minimizer. Hence if we consider $0 < \varepsilon < \delta$ we have that $q_1 = \gamma(-\varepsilon)$ and $q_2 = \gamma(\varepsilon)$ satisfy the property required, which proves claim (ii).

Definition 1.13. Fix $q_0 \in M$. We define the *exponential map* starting from q_0 as

$$\mathcal{E}_{q_0} : T^*_{q_0}M \to M, \qquad \mathcal{E}_{q_0}(\lambda_0) = \pi(e^{\vec{H}}(\lambda_0)).$$

Using the homogeneity property $H(c\lambda) = c^2 H(\lambda),\ \forall\, c > 0$, we have that

$$e^{\vec{H}}(s\lambda) = e^{s\vec{H}}(\lambda), \qquad \forall\, s > 0.$$

In other words we can recover the geodesic on the manifold with initial covector λ_0 as the image under $\mathcal{E}_{q_0}$ of the ray $\{t\lambda_0, t \in [0,1]\} \subset T^*_{q_0}M$ that join the origin to λ_0.

$$\mathcal{E}_{q_0}(t\lambda_0) = \pi(e^{\vec{H}}(t\lambda_0)) = \pi(e^{t\vec{H}}(\lambda_0)) = \pi(\lambda(t)) = \gamma(t).$$

Next, we recall the definition of cut and conjugate time.

Definition 1.14. Let $\mathbf{S} = (M, \Delta, \mathbf{g})$ be a sub-Riemannian manifold. Let $q_0 \in M$ and $\lambda_0 \in \Lambda_{q_0}$. Assume that the geodesic $\gamma(t) = \mathsf{Exp}_{q_0}(t\lambda_0)$ for $t > 0$, is not abnormal.

(i) The *first conjugate time* is $t_{con}(\lambda_0) = \min\{t > 0,\ t\lambda_0$ is a critical point of $\mathcal{E}_{q_0}\}$.

(ii) The *cut time* is $t_{cut}(\lambda_0) = \min\{t > 0,\ \exists\, \lambda_1 \in \Lambda_{q_0}, \lambda_1 \neq \lambda_0$ s.t. $\mathcal{E}_{q_0}(t_c(\lambda_0)\lambda_0) = \mathcal{E}_{q_0}(t_c(\lambda_0)\lambda_1)\}$.

It is well known that if a geodesic is not abnormal then it loses optimality either at the cut or at the conjugate locus (see for instance [8]).

1.3 The nilpotent approximation

In this section we briefly recall the concept of nilpotent approximation. For details see [11, 25].

Let $\mathbf{S} = (M, \Delta, \mathbf{g})$ be a sub-Riemannian manifold and $(f_1, \ldots, f_k)$ an orthonormal frame. Fix a point $q \in M$ and consider the flag of the distribution $\Delta^1_q \subset \Delta^2_q \subset \ldots \subset \Delta^m_q$. Recall that $k_i = \dim \Delta^i_q - \dim \Delta^{i-1}_q$ for $i = 1, \ldots, m$, and that $k_1 + \ldots + k_m = n$.

Let O_q be an open neighborhood of the point $q \in M$. We say that a system of coordinates $\psi : O_q \to \mathbb{R}^n$ is *linearly adapted* to the flag if, in these coordinates, we have $\psi(q) = 0$ and

$$\psi_*(\Delta^i_q) = \mathbb{R}^{k_1} \oplus \ldots \oplus \mathbb{R}^{k_i}, \qquad \forall\, i = 1, \ldots, m.$$

Consider now the splitting $\mathbb{R}^n = \mathbb{R}^{k_1} \oplus \ldots \oplus \mathbb{R}^{k_m}$ and denote its elements $x = (x_1, \ldots, x_m)$ where $x_i = (x_i^1, \ldots, x_i^{k_i}) \in \mathbb{R}^{k_i}$. The space of all differential operators in $\mathbb{R}^n$ with smooth coefficients forms an associative algebra with composition of operators as multiplication. The differential operators with polynomial coefficients form a subalgebra of this algebra with generators $1, x_i^j, \frac{\partial}{\partial x_i^j}$, where $i = 1, \ldots, m;\ j = 1, \ldots, k_i$. We define weights of generators as

$$\nu(1) = 0, \qquad \nu(x_i^j) = i, \qquad \nu(\frac{\partial}{\partial x_i^j}) = -i,$$

and the weight of monomials

$$\nu(y_1 \cdots y_\alpha \frac{\partial^\beta}{\partial z_1 \cdots \partial z_\beta}) = \sum_{i=1}^{\alpha} \nu(y_i) - \sum_{j=1}^{\beta} \nu(z_j).$$

Notice that a polynomial differential operator homogeneous with respect to ν (i.e. whose monomials are all of same weight) is homogeneous with respect to dilations $\delta_t : \mathbb{R}^n \to \mathbb{R}^n$ defined by

$$\delta_t(x_1, \ldots, x_m) = (tx_1, t^2 x_2, \ldots, t^m x_m), \qquad t > 0. \tag{1.11}$$

In particular for a homogeneous vector field X of weight h it holds $\delta_{t*} X = t^{-h} X$. A smooth vector field $X \in \mathrm{Vec}(\mathbb{R}^n)$, as a first order differential operator, can be written as

$$X = \sum_{i,j} a_i^j(x) \frac{\partial}{\partial x_i^j}$$

and considering its Taylor expansion at the origin we can write the formal expansion

$$X \approx \sum_{h=-m}^{\infty} X^{(h)}$$

where $X^{(h)}$ is the homogeneous part of degree h of X (notice that every monomial of a first order differential operator has weight not smaller than $-m$). Define the filtration of $\mathrm{Vec}(\mathbb{R}^n)$

$$\mathcal{D}^{(h)} = \{X \in \mathrm{Vec}(\mathbb{R}^n) : X^{(i)} = 0, \forall\, i < h\}, \qquad \ell \in \mathbb{Z}.$$

Definition 1.15. Let $\mathbf{S}$ be a sub-Riemannian structure and $f_1, \ldots, f_k$ a local orthonormal frame near the point q. A system of coordinates $\psi : O_q \to \mathbb{R}^n$ defined near q is said *privileged* if these coordinates are linearly adapted to the flag and such that $\psi_* f_i \in \mathcal{D}^{(-1)}$ for every $i = 1, \ldots, k$.

Theorem 1.16. *Privileged coordinates always exists. Moreover there exist $c_1, c_2 > 0$ such that in these coordinates, for all $\varepsilon > 0$ small enough, we have*

$$c_1 \,\mathrm{Box}(\varepsilon) \subset B(q, \varepsilon) \subset c_2 \,\mathrm{Box}(\varepsilon), \tag{1.12}$$

where $\mathrm{Box}(\varepsilon) = \{x \in \mathbb{R}^n, |x_i| \leq \varepsilon^i\}$.

Existence of privileged coordinates is proved in [11, 14, 25, 30]. In the regular case the construction of privileged coordinates was also done in the context of hypoelliptic operators (see [92]). The second statement is known as *Ball-Box theorem* and a proof can be found in [25]. Notice however that privileged coordinates are not unique.

Definition 1.17. Let $\mathbf{S} = (M, \Delta, \mathbf{g})$ be a regular sub-Riemannian manifold and $(f_1, \ldots, f_k)$ a local orthonormal frame near a point q. Fixed a system of privileged coordinates, we define the *nilpotent approximation of* $\mathbf{S}$ *near* q, denoted by $\widehat{\mathbf{S}}_q$, the sub-Riemannian structure on $\mathbb{R}^n$ having $(\widehat{f}_1, \ldots, \widehat{f}_k)$ as an orthonormal frame, where $\widehat{f}_i := (\psi_* f_i)^{(-1)}$.

Remark 1.18. It is well known that under the regularity hypothesis, $\widehat{\mathbf{S}}_q$ is naturally endowed with a Lie group structure whose Lie algebra is generated by left-invariant vector fields $\widehat{f}_1, \ldots, \widehat{f}_k$. Moreover the sub-Riemannian distance $\widehat{d}$ in $\widehat{\mathbf{S}}_q$ is homogeneous with respect to dilations δ_t, i.e. $\widehat{d}(\delta_t(x), \delta_t(y)) = t\,\widehat{d}(x, y)$. In particular, if $\widehat{B}_q(r)$ denotes the ball of radius r in $\widehat{\mathbf{S}}_q$, this implies $\delta_t(\widehat{B}_q(1)) = \widehat{B}_q(t)$.

Theorem 1.19. *The nilpotent approximation* $\widehat{\mathbf{S}}_q$ *of a sub-Riemannian structure* $\mathbf{S}$ *near a point* q *is the metric tangent space to* M *at point* q *in the sense of Gromov, that means*

$$\delta_{1/\varepsilon} B(q, \varepsilon) \longrightarrow \widehat{B}_q, \tag{1.13}$$

where $\widehat{B}_q$ *denotes the sub-Riemannian unit ball of the nilpotent approximation* $\widehat{\mathbf{S}}_q$.

Remark 1.20. Convergence of sets in (1.13) is intended in the Gromov-Hausdorff topology [25, 66]. In the regular case this theorem was proved by Mitchell in [78]. A proof in the general case can be found in [25].

CHAPTER 2

Classification of sub-Riemannian structures on 3D Lie groups

In this chapter we give a complete classification of left-invariant sub-Riemannian structures on three dimensional Lie groups in terms of the basic functional invariants of a 3D contact structure. As a corollary we explicitly find a sub-Riemannian isometry between the nonisomorphic Lie groups $SL(2)$ and $A^+(\mathbb{R})\times S^1$, where $A^+(\mathbb{R})$ denotes the group of orientation preserving affine maps on the real line.

2.1 Introduction

A sub-Riemannian structure on a n-dimensional smooth manifold is said to be *contact* if its distribution is defined as the kernel of a contact differential one form ω, i.e. $n = 2l + 1$ and $\left(\bigwedge^l d\omega\right) \wedge \omega$ is a nonvanishing n-form on M.

In this chapter we focus on the three dimensional case. Three dimensional contact sub-Riemannian structures have been deeply studied in the last years (for example see [4, 8, 10]) and they possess two basic differential invariants χ and κ (see Section 2.2 for the precise definition and [3, 8] for their role in the asymptotic expansion of the sub-Riemannian exponential map).

The invariants χ and κ are smooth real functions on M. It is easy to understand, at least heuristically, why it is natural to expect exactly *two* functional invariants. Indeed, in local coordinates the sub-Riemannian structure is defined by its orthonormal frame, i.e. by a couple of smooth vector fields on $\mathbb{R}^3$ or, in other words, by 6 scalar functions on $\mathbb{R}^3$. One function can be normalized by the rotation of the frame within its linear hull and three more functions by smooth change of variables. What remains are two scalar functions.

We exploit these local invariants to provide a complete classification of left-invariant structures on 3D Lie groups. Recall that a sub-Riemannian structure on a Lie group is said to be *left-invariant* if its distribution and the inner product are preserved by left translations on the group and a left-invariant distribution is uniquely determined by a two dimensional subspace of the Lie algebra of the group. The distribution is bracket generating (and contact) if and only if the subspace is not a Lie subalgebra.

Left-invariant structures on Lie groups are the basic models of sub-Riemannian manifolds and the study of such structures is the starting point to understand the general properties of sub-Riemannian geometry. In particular, thanks to the group structure, in some of these cases it is also possible to compute explicitly the sub-

Riemannian distance and geodesics (see in particular [64] for the Heisenberg group, [38] for semisimple Lie groups with Killing form and [80, 95] for a detailed study of the sub-Riemannian structure on the group of motions of a plane).

The problem of equivalence for several geometric structures close to left-invariant sub-Riemannian structures on 3D Lie groups were studied in several publications (see [45, 46, 54, 96]). In particular in [96] the author provide a first classification of symmetric sub-Riemannian structures of dimension 3, while in [54] is presented a complete classification of sub-Riemannian homogeneous spaces (i.e. sub-Riemannian structures which admits a transitive Lie group of isometries acting smoothly on the manifold) by means of an adapted connection. The principal invariants used there, denoted by τ_0 and K, coincide up to a normalization factor with our differential invariants χ and κ.

A standard result on the classification of 3D Lie algebras (see, for instance, [68]) reduce the analysis on the Lie algebras of the following Lie groups:

H_3, the Heisenberg group,

$A^+(\mathbb{R})\oplus\mathbb{R}$, where $A^+(\mathbb{R})$ is the group of orientation preserving affine maps on $\mathbb{R}$,

$SOLV^+, SOLV^-$ are Lie groups whose Lie algebra is solvable and has 2-dim square,

$SE(2)$ and $SH(2)$ are the groups of orientation preserving motions of Euclidean and Hyperbolic plane respectively,

$SL(2)$ and $SU(2)$ are the three dimensional simple Lie groups.

Moreover it is easy to show that in each of these cases but one all left-invariant bracket generating distributions are equivalent by automorphisms of the Lie algebra. The only case where there exists two non-equivalent distributions is the Lie algebra $\mathfrak{sl}(2)$. More precisely a 2-dimensional subspace of $\mathfrak{sl}(2)$ is called *elliptic* (*hyperbolic*) if the restriction of the Killing form on this subspace is sign-definite (sign-indefinite). Accordingly, we use notation $SL_e(2)$ and $SL_h(2)$ to specify on which subspace the sub-Riemannian structure on $SL(2)$ is defined.

For a left-invariant structure on a Lie group the invariants χ and κ are constant functions and allow us to distinguish non isometric structures. To complete the classification we can restrict ourselves to *normalized* sub-Riemannian structures, i.e. structures that satisfy

$$\chi = \kappa = 0, \qquad \text{or} \qquad \chi^2 + \kappa^2 = 1. \tag{2.1}$$

Indeed χ and κ are homogeneous with respect to dilations of the orthonormal frame, that means rescaling of distances on the manifold. Thus we can always rescale our structure in such a way that (2.1) is satisfied.

To find missing discrete invariants, i.e. to distinguish between normalized structures with same χ and κ, we then show that it is always possible to select a canonical orthonormal frame for the sub-Riemannian structure such that all structure constants of the Lie algebra of this frame are invariant with respect to local isometries.

Then the commutator relations of the Lie algebra generated by the canonical frame determine in a unique way the sub-Riemannian structure.

Collecting together these results we prove the following

Theorem 2.1. *All left-invariant sub-Riemannian structures on 3D Lie groups are classified up to local isometries and dilations as in Figure 2.1, where a structure is identified by the point (κ, χ) and two distinct points represent non locally isometric structures.*

Moreover

(i) If $\chi = \kappa = 0$ then the structure is locally isometric to the Heisenberg group,

(ii) If $\chi^2 + \kappa^2 = 1$ then there exist no more than three non isometric normalized sub-Riemannian structures with these invariants; in particular there exists a unique normalized structure on a unimodular Lie group (for every choice of χ, κ).

(iii) If $\chi \neq 0$ or $\chi = 0, \kappa \geq 0$, then two structures are locally isometric if and only if their Lie algebras are isomorphic.

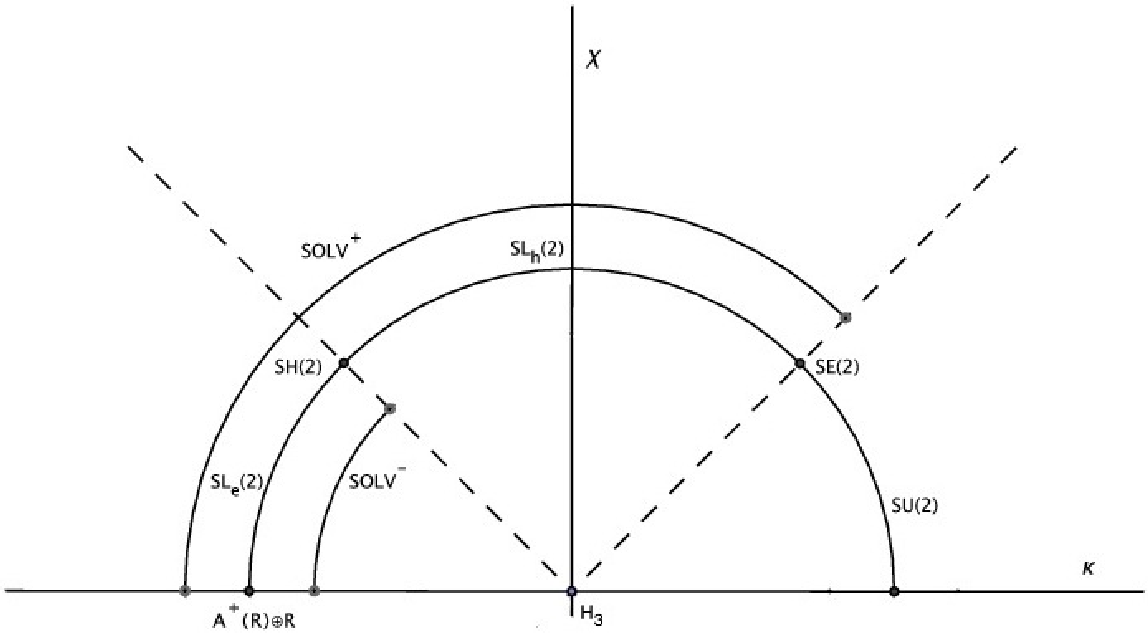

Figure 2.1: Classification

In other words every left-invariant sub-Riemannian structure is locally isometric to a normalized one that appear in Figure 2.1, where we draw points on different circles since we consider equivalence classes of structures up to dilations. In this way it is easier to understand how many normalized structures there exist for some fixed value of the local invariants. Notice that unimodular Lie groups are those that appear in the middle circle (except for $A^+(\mathbb{R}) \oplus \mathbb{R}$).

From the proof of Theorem 2.1 we get also a uniformization-like theorem for "constant curvature" manifolds in the sub-Riemannian setting:

Corollary 2.2. *Let M be a complete simply connected 3D contact sub-Riemannian manifold. Assume that $\chi = 0$ and κ is costant on M. Then M is isometric to a left-invariant sub-Riemannian structure. More precisely:*

(i) if $\kappa = 0$ it is isometric to the Heisenberg group H_3,

(ii) if $\kappa = 1$ it is isometric to the group $SU(2)$ with Killing metric,

(iii) if $\kappa = -1$ it is isometric to the group $\widetilde{SL}(2)$ with elliptic type Killing metric,

where $\widetilde{SL}(2)$ is the universal covering of $SL(2)$.

Another byproduct of the classification is the fact that there exist non isomorphic Lie groups with locally isometric sub-Riemannian structures. Indeed, as a consequence of Theorem 2.1, we get that there exists a unique normalized left-invariant structure defined on $A^+(\mathbb{R}) \oplus \mathbb{R}$ having $\chi = 0, \kappa = -1$. Thus $A^+(\mathbb{R}) \oplus \mathbb{R}$ is locally isometric to the group $SL(2)$ with elliptic type Killing metric by Corollary 2.2.

This fact was already noted in [54] as a consequence of the classification. In this paper we explicitly compute the global sub-Riemannian isometry between $A^+(\mathbb{R})\oplus\mathbb{R}$ and the universal covering of $SL(2)$ by means of Nagano principle. We then show that this map is well defined on the quotient, giving a global isometry between the group $A^+(\mathbb{R})\times S^1$ and the group $SL(2)$, endowed with the sub-Riemannian structure defined by the restriction of the Killing form on the elliptic distribution.

The group $A^+(\mathbb{R})\oplus\mathbb{R}$ can be interpreted as the subgroup of the affine maps on the plane that acts as an orientation preserving affinity on one axis and as translations on the other one[1]

$$A^+(\mathbb{R}) \oplus \mathbb{R} := \left\{ \begin{pmatrix} a & 0 & b \\ 0 & 1 & c \\ 0 & 0 & 1 \end{pmatrix},\ a > 0, b, c \in \mathbb{R} \right\}.$$

The standard left-invariant sub-Riemannian structure on $A^+(\mathbb{R}) \oplus \mathbb{R}$ is defined by the orthonormal frame $\Delta = \operatorname{span}\{e_2, e_1 + e_3\}$, where

$$e_1 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad e_2 = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad e_3 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix},$$

is a basis of the Lie algebra of the group, satisfying $[e_1, e_2] = e_1$.

The subgroup $A^+(\mathbb{R})$ is topologically homeomorphic to the half-plane $\{(a, b) \in \mathbb{R}^2, a > 0\}$ which can be descirbed in standard polar coordinates as $\{(\rho, \theta) | \rho > 0, -\pi/2 < \theta < \pi/2\}$.

[1]We can recover the action as an affine map identifying $(x, y) \in \mathbb{R}^2$ with $(x, y, 1)^T$ and

$$\begin{pmatrix} a & 0 & b \\ 0 & 1 & c \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = \begin{pmatrix} ax + b \\ y + c \\ 1 \end{pmatrix}.$$

Theorem 2.3. *The diffeomorphism* $\Psi : A^+(\mathbb{R}) \times S^1 \longrightarrow SL(2)$ *defined by*

$$\Psi(\rho, \theta, \varphi) = \frac{1}{\sqrt{\rho \cos \theta}} \begin{pmatrix} \cos \varphi & \sin \varphi \\ \rho \sin(\theta - \varphi) & \rho \cos(\theta - \varphi) \end{pmatrix}, \tag{2.2}$$

where $(\rho, \theta) \in A^+(\mathbb{R})$ *and* $\varphi \in S^1$, *is a global sub-Riemannian isometry.*

Using this global sub-Riemannian isometry as a change of coordinates one can recover the geometry of the sub-Riemannian structure on the group $A^+(\mathbb{R}) \times S^1$, starting from the analogous properties of $SL(2)$ (e.g. explicit expression of the sub-Riemannian distance, the cut locus). In particular we notice that, since $A^+(\mathbb{R}) \times S^1$ is not unimodular, the canonical sub-Laplacian on this group is not expressed as a sum of squares. Indeed if X_1, X_2 denotes the left-invariant vector fields associated to the orthonormal frame, the sub-Laplacian is expressed as follows

$$\mathcal{L} = X_1^2 + X_2^2 + X_1.$$

Moreover in the non-unimodular case the generalized Fourier transform method, used in [6], cannot apply . Hence the heat kernel of the corresponding heat equation cannot be computed directly. On the other hand one can use the map (3.1) to express the solution in terms of the heat kernel on $SL(2)$.

2.2 Sub-Riemannian invariants

In this section we study a contact sub-Riemannian structure on a 3D manifold and we give a brief description of its two invariants (see also [8]). We start with the following characterization of contact distributions.

Lemma 2.4. *Let* M *be a 3D manifold,* $\omega \in \Lambda^1 M$ *and* $\Delta = \ker \omega$. *The following are equivalent:*

(*i*) Δ *is a contact distribution,*

(*ii*) $d\omega\big|_\Delta \neq 0$,

(*iii*) $\forall f_1, f_2 \in \overline{\Delta}$ *linearly independent, then* $[f_1, f_2] \notin \overline{\Delta}$.

Moreover, in this case, the contact form can be selected in such a way that $d\omega\big|_\Delta$ *coincide with the Euclidean volume form on* Δ.

By Lemma 2.4 it is not restrictive to assume that the sub-Riemannian structure satisfies:

$$\begin{gathered} (M, \omega) \text{ is a 3D contact structure,} \\ \Delta = \operatorname{span}\{f_1, f_2\} = \ker \omega, \\ \mathbf{g}(f_i, f_j) = \delta_{ij}, \quad d\omega(f_1, f_2) = 1. \end{gathered} \tag{2.3}$$

We stress that in (2.3) the orthonormal frame f_1, f_2 is not unique. Indeed every rotated frame (where the angle of rotation depends smoothly on the point) defines the same structure.

The sub-Riemannian Hamiltonian (1.7) is written

$$h = \frac{1}{2}(h_1^2 + h_2^2).$$

Definition 2.5. In the setting (2.3) we define the *Reeb vector field* associated to the contact structure as the unique vector field f_0 such that

$$\begin{aligned} \omega(f_0) &= 1, \\ d\omega(f_0, \cdot) &= 0. \end{aligned} \tag{2.4}$$

From the definition it is clear that f_0 depends only on the sub-Riemannian structure (and its orientation) and not on the frame selected.

Condition (2.4) is equivalent to

$$\begin{aligned} &[f_1, f_0], [f_2, f_0] \in \overline{\Delta}, \\ &[f_2, f_1] = f_0 \pmod{\overline{\Delta}}. \end{aligned}$$

and we deduce the following expression for the Lie algebra of vector fields generated by f_0, f_1, f_2

$$\begin{aligned} [f_1, f_0] &= c_{01}^1 f_1 + c_{01}^2 f_2, \\ [f_2, f_0] &= c_{02}^1 f_1 + c_{02}^2 f_2, \\ [f_2, f_1] &= c_{12}^1 f_1 + c_{12}^2 f_2 + f_0, \end{aligned} \tag{2.5}$$

where c_{ij}^k are functions on the manifold, called structure constants of the Lie algebra.

If we denote with (ν_0, ν_1, ν_2) the basis of 1-form dual to (f_0, f_1, f_2), we can rewrite (2.5) as:

$$\begin{aligned} d\nu_0 &= \qquad \nu_1 \wedge \nu_2, \\ d\nu_1 &= c_{01}^1 \nu_0 \wedge \nu_1 + c_{02}^1 \nu_0 \wedge \nu_2 + c_{12}^1 \nu_1 \wedge \nu_2, \\ d\nu_2 &= c_{01}^2 \nu_0 \wedge \nu_1 + c_{02}^2 \nu_0 \wedge \nu_2 + c_{12}^2 \nu_1 \wedge \nu_2, \end{aligned} \tag{2.6}$$

Let $h_0(\lambda) = \langle \lambda, f_0(q) \rangle$ denote the Hamiltonian linear on fibers associated with the Reeb field f_0. We now compute the Poisson bracket $\{h, h_0\}$, denoting with $\{h, h_0\}_q$ its restriction to the fiber $T_q^* M$.

Proposition 2.6. *The Poisson bracket $\{h, h_0\}_q$ is a quadratic form. Moreover we have*

$$\{h, h_0\} = c_{01}^1 h_1^2 + (c_{01}^2 + c_{02}^1) h_1 h_2 + c_{02}^2 h_2^2, \tag{2.7}$$

$$c_{01}^1 + c_{02}^2 = 0. \tag{2.8}$$

In particular, $\Delta_q^\perp \subset \ker \{h, h_0\}_q$ and $\{h, h_0\}_q$ is actually a quadratic form on $T_q^* M / \Delta_q^\perp = \Delta_q^*$.

Proof. Using the equality $\{h_i, h_j\}(\lambda) = \langle \lambda, [f_i, f_j](q) \rangle$ we get

$$\begin{aligned}\{h, h_0\} &= \frac{1}{2}\{h_1^2 + h_2^2, h_0\} = h_1\{h_1, h_0\} + h_2\{h_2, h_0\} \\ &= h_1(c_{01}^1 h_1 + c_{01}^2 h_2) + h_2(c_{02}^1 h_1 + c_{02}^2 h_2) \\ &= c_{01}^1 h_1^2 + (c_{01}^2 + c_{02}^1) h_1 h_2 + c_{02}^2 h_2^2.\end{aligned}$$

Differentiating the first equation in (2.6) we find:

$$\begin{aligned}0 = d^2 \nu_0 &= d\nu_1 \wedge \nu_2 - \nu_1 \wedge d\nu_2 \\ &= (c_{01}^1 + c_{02}^2)\nu_0 \wedge \nu_1 \wedge \nu_2.\end{aligned}$$

which proves (2.8). □

Being $\{h, h_0\}_q$ a quadratic form on the Euclidean plane Δ_q (using the canonical identification of the vector space Δ_q with its dual Δ_q^* given by the scalar product), it is a standard fact that it can be interpreted as a symmetric operator on the plane itself. In particular its determinant and its trace are well defined. From (2.8) we get

$$\operatorname{trace} \{h, h_0\}_q = 0.$$

It is natural then to define our *first invariant* as the positive eigenvalue of this operator, namely:

$$\chi(q) = \sqrt{-\det\{h, h_0\}_q}. \tag{2.9}$$

which can also be written in terms of structure constant of the Lie algebra as follows

$$\chi(q) = \sqrt{-\det C}, \qquad C = \begin{pmatrix} c_{01}^1 & (c_{01}^2 + c_{02}^1)/2 \\ (c_{01}^2 + c_{02}^1)/2 & c_{02}^2 \end{pmatrix}, \tag{2.10}$$

Remark 2.7. Notice that, by definition $\chi \geq 0$, and it vanishes everywhere if and only if the flow of the Reeb vector field f_0 is a flow of sub-Riemannian isometries for M.

The *second invariant*, which was found in [8] as a term of the asymptotic expansion of conjugate locus, is defined in the following way

$$\kappa(q) = f_2(c_{12}^1) - f_1(c_{12}^2) - (c_{12}^1)^2 - (c_{12}^2)^2 + \frac{c_{01}^2 - c_{02}^1}{2}. \tag{2.11}$$

where we refer to notation (2.5). A direct calculation shows that κ is preserved by rotations of the frame f_1, f_2 of the distribution, hence it depends only on the sub-Riemannian structure.

χ and κ are functions defined on the manifold; they reflect intrinsic geometric properties of the sub-Riemannian structure and are preserved by the sub-Riemannian isometries. In particular, χ and κ are constant functions for left-invariant structures on Lie groups (since left translations are isometries).

2.3 Canonical Frames

In this section we want to show that it is always possible to select a canonical orthonormal frame for the sub-Riemannian structure. In this way we are able to find missing discrete invariants and to classify sub-Riemannian structures simply knowing structure constants c_{ij}^k for the canonical frame. We study separately the two cases $\chi \neq 0$ and $\chi = 0$.

We start by rewriting and improving Proposition 2.6 when $\chi \neq 0$.

Proposition 2.8. *Let M be a 3D contact sub-Riemannian manifold and $q \in M$. If $\chi(q) \neq 0$, then there exists a local frame such that*

$$\{h, h_0\} = 2\chi h_1 h_2. \tag{2.12}$$

In particular, in the Lie group case with left-invariant stucture, there exists a unique (up to a sign) canonical frame (f_0, f_1, f_2) such that

$$\begin{aligned} [f_1, f_0] &= c_{01}^2 f_2, \\ [f_2, f_0] &= c_{02}^1 f_1, \\ [f_2, f_1] &= c_{12}^1 f_1 + c_{12}^2 f_2 + f_0. \end{aligned} \tag{2.13}$$

Moreover we have

$$\chi = \frac{c_{01}^2 + c_{02}^1}{2}, \qquad \kappa = -(c_{12}^1)^2 - (c_{12}^2)^2 + \frac{c_{01}^2 - c_{02}^1}{2}. \tag{2.14}$$

Proof. From Proposition 2.6 we know that the Poisson bracket $\{h, h_0\}_q$ is a non degenerate symmetric operator with zero trace. Hence we have a well defined, up to a sign, orthonormal frame by setting f_1, f_2 as the orthonormal isotropic vectors of this operator (remember that f_0 depends only on the structure and not on the orthonormal frame on the distribution). It is easily seen that in both of these cases we obtain the expression (2.12). □

Remark 2.9. Notice that, if we change sign to f_1 or f_2, then c_{12}^2 or c_{12}^1, respectively, change sign in (2.13), while c_{02}^1 and c_{01}^2 are unaffected. Hence equalities (2.14) do not depend on the orientation of the sub-Riemannian structure.

If $\chi = 0$ the above procedure cannot apply. Indeed both trace and determinant of the operator vanish, hence we have $\{h, h_0\}_q = 0$. From (2.7) we get the identities

$$c_{01}^1 = c_{02}^2 = 0, \quad c_{01}^2 + c_{02}^1 = 0. \tag{2.15}$$

so that commutators (2.5) simplify in (where $c = c_{01}^2$)

$$\begin{aligned} [f_1, f_0] &= c f_2, \\ [f_2, f_0] &= -c f_1, \\ [f_2, f_1] &= c_{12}^1 f_1 + c_{12}^2 f_2 + f_0. \end{aligned} \tag{2.16}$$

We want to show, with an explicit construction, that also in this case there always exists a rotation of our frame, by an angle that smoothly depends on the point, such that in the new frame κ is the only structure constant which appear in (2.16).

Lemma 2.10. *Let f_1, f_2 be an orthonormal frame on M. If we denote with $\widehat{f}_1, \widehat{f}_2$ the frame obtained from the previous one with a rotation by an angle $\theta(q)$ and with $\widehat{c}_{ij}^k$ structure constants of rotated frame, we have:*

$$\begin{aligned}\widehat{c}_{12}^1 &= \cos\theta(c_{12}^1 - f_1(\theta)) - \sin\theta(c_{12}^2 - f_2(\theta)),\\ \widehat{c}_{12}^2 &= \sin\theta(c_{12}^1 - f_1(\theta)) + \cos\theta(c_{12}^2 - f_2(\theta)).\end{aligned}$$

Now we can prove the main result of this section.

Proposition 2.11. *Let M be a 3D simply connected contact sub-Riemannian manifold such that $\chi = 0$. Then there exists a rotation of the original frame $\widehat{f}_1, \widehat{f}_2$ such that:*

$$\begin{aligned}[][\widehat{f}_1, f_0] &= \kappa \widehat{f}_2,\\ [\widehat{f}_2, f_0] &= -\kappa \widehat{f}_1,\\ [\widehat{f}_2, \widehat{f}_1] &= f_0.\end{aligned} \tag{2.17}$$

Proof. Using Lemma 2.10 we can rewrite the statement in the following way: there exists a function $\theta : M \to \mathbb{R}$ such that

$$f_1(\theta) = c_{12}^1, \quad f_2(\theta) = c_{12}^2. \tag{2.18}$$

Indeed, this would imply $\widehat{c}_{12}^1 = \widehat{c}_{12}^2 = 0$ and $\kappa = c$.

Let us introduce simplified notations $c_{12}^1 = \alpha_1,\ c_{12}^2 = \alpha_2$. Then

$$\kappa = f_2(\alpha_1) - f_1(\alpha_2) - (\alpha_1)^2 - (\alpha_2)^2 + c. \tag{2.19}$$

If (ν_0, ν_1, ν_2) denotes the dual basis to (f_0, f_1, f_2) we have

$$d\theta = f_0(\theta)\nu_0 + f_1(\theta)\nu_1 + f_2(\theta)\nu_2.$$

and from (2.16) we get:

$$\begin{aligned}f_0(\theta) &= ([f_2, f_1] - \alpha_1 f_1 - \alpha_2 f_2)(\theta)\\ &= f_2(\alpha_1) - f_1(\alpha_2) - \alpha_1^2 - \alpha_2^2\\ &= \kappa - c.\end{aligned}$$

Suppose now that (2.18) are satisfied, we get

$$d\theta = (\kappa - c)\nu_0 + \alpha_1\nu_1 + \alpha_2\nu_2 =: \eta. \tag{2.20}$$

with the r.h.s. independent from θ.

To prove the theorem we have to show that η is an exact 1-form. Since the manifold is simply connected, it is sufficient to prove that η is closed. If we denote $\nu_{ij} := \nu_i \wedge \nu_j$ dual equations of (2.16) are:

$$\begin{aligned}d\nu_0 &= \nu_{12},\\ d\nu_1 &= -c\nu_{02} + \alpha_1\nu_{12},\\ d\nu_2 &= c\nu_{01} - \alpha_2\nu_{12}.\end{aligned}$$

and differentiating we get two nontrivial relations:

$$f_1(c) + c\alpha_2 + f_0(\alpha_1) = 0, \tag{2.21}$$
$$f_2(c) - c\alpha_1 + f_0(\alpha_2) = 0. \tag{2.22}$$

Recollecting all these computations we prove the closure of η

$$\begin{aligned}
d\eta &= d(\kappa - c) \wedge \nu_0 + (\kappa - c)d\nu_0 + d\alpha_1 \wedge \nu_1 + \alpha_1 d\nu_1 + d\alpha_2 \wedge \nu_2 + \alpha_2 d\nu_2 \\
&= -dc \wedge \nu_0 + (\kappa - c)\nu_{12} + \\
&\qquad + f_0(\alpha_1)\nu_{01} - f_2(\alpha_1)\nu_{12} + \alpha_1(\alpha_1\nu_{12} - c\nu_{02}) \\
&\qquad\quad + f_0(\alpha_2)\nu_{02} + f_1(\alpha_2)\nu_{12} + \alpha_2(c\nu_{01} - \alpha_2\nu_{12}) \\
&= (f_0(\alpha_1) + \alpha_2 c + f_1(c))\nu_{01} \\
&\qquad + (f_0(\alpha_2) - \alpha_1 c + f_2(c))\nu_{02} \\
&\qquad\quad + (\kappa - c - f_2(\alpha_1) + f_1(\alpha_2) + \alpha_1^2 + \alpha_2^2)\nu_{12} \\
&= 0.
\end{aligned}$$

where in the last equality we use (3.13) and (2.21)-(2.22). □

2.4 Classification

Now we use the results of the previous sections to prove Theorem 2.1.

In this section G denotes a 3D Lie group, with Lie algebra $\mathfrak{g}$, endowed with a left-invariant sub-Riemannian structure defined by the orthonormal frame f_1, f_2, i.e.

$$\Delta = \operatorname{span}\{f_1, f_2\} \subset \mathfrak{g}, \qquad \operatorname{span}\{f_1, f_2, [f_1, f_2]\} = \mathfrak{g}.$$

Recall that for a 3D left-invariant structure to be bracket generating is equivalent to be contact, moreover the Reeb field f_0 is also a left-invariant vector field by construction.

From the fact that, for left-invariant structures, local invariants are constant functions (see Remark 4.17) we obtain a necessary condition for two structures to be locally isometric.

Proposition 2.12. *Let G, H be 3D Lie groups with locally isometric sub-Riemannian structures. Then $\chi_G = \chi_H$ and $\kappa_G = \kappa_H$.*

Notice that this condition is not sufficient. It turns out that there can be up to three mutually non locally isometric normalized structures with the same invariants χ, κ.

Remark 2.13. It is easy to see that χ and κ are homogeneous of degree 2 with respect to dilations of the frame. Indeed assume that the sub-Riemannian structure $(M, \Delta, \mathbf{g})$ is locally defined by the orthonormal frame f_1, f_2, i.e.

$$\Delta = \operatorname{span}\{f_1, f_2\}, \qquad \mathbf{g}(f_i, f_j) = \delta_{ij}.$$

Consider now the dilated structure $(M, \Delta, \widetilde{\mathbf{g}})$ defined by the orthonormal frame $\lambda f_1, \lambda f_2$

$$\Delta = \operatorname{span}\{f_1, f_2\}, \qquad \widetilde{\mathbf{g}}(f_i, f_j) = \frac{1}{\lambda^2}\delta_{ij}, \qquad \lambda > 0.$$

If χ, κ and $\widetilde{\chi}, \widetilde{\kappa}$ denote the invariants of the two structures respectively, we find

$$\widetilde{\chi} = \lambda^2 \chi, \qquad \widetilde{\kappa} = \lambda^2 \kappa, \qquad \lambda > 0.$$

A dilation of the orthonormal frame corresponds to a multiplication by a factor $\lambda > 0$ of all distances in our manifold. Since we are interested in a classification by local isometries, we can always suppose (for a suitable dilation of the orthonormal frame) that the local invariants of our structure satisfy

$$\chi = \kappa = 0, \qquad \text{or} \qquad \chi^2 + \kappa^2 = 1,$$

and we study equivalence classes with respect to local isometries.

Since χ is non negative by definition (see Remark 2.7), we study separately the two cases $\chi > 0$ and $\chi = 0$.

2.4.1 Case $\chi > 0$

Let G be a 3D Lie group with a left-invariant sub-Riemannian structure such that $\chi \neq 0$. From Proposition 2.8 we can assume that $\Delta = \text{span}\{f_1, f_2\}$ where f_1, f_2 is the canonical frame of the structure. From (2.13) we obtain the dual equations

$$\begin{aligned} d\nu_0 &= \nu_1 \wedge \nu_2, \\ d\nu_1 &= c_{02}^1 \nu_0 \wedge \nu_2 + c_{12}^1 \nu_1 \wedge \nu_2, \\ d\nu_2 &= c_{01}^2 \nu_0 \wedge \nu_1 + c_{12}^1 \nu_1 \wedge \nu_2. \end{aligned} \tag{2.23}$$

Using $d^2 = 0$ we obtain structure equations

$$\begin{cases} c_{02}^1 c_{12}^2 = 0, \\ c_{01}^2 c_{12}^1 = 0. \end{cases} \tag{2.24}$$

We know that the structure constants of the canonical frame are invariant by local isometries (up to change signs of c_{12}^1, c_{12}^2, see Remark 2.9). Hence, every different choice of coefficients in (2.13) which satisfy also (2.24) will belong to a different class of non-isometric structures.

Taking into account that $\chi > 0$ implies that c_{01}^2 and c_{02}^1 cannot be both non positive (see (2.14)), we have the following cases:

(*i*) $c_{12}^1 = 0$ and $c_{12}^2 = 0$. In this first case we get

$$\begin{aligned} [f_1, f_0] &= c_{01}^2 f_2, \\ [f_2, f_0] &= c_{02}^1 f_1, \\ [f_2, f_1] &= f_0, \end{aligned}$$

and formulas (2.14) imply

$$\chi = \frac{c_{01}^2 + c_{02}^1}{2} > 0, \qquad \kappa = \frac{c_{01}^2 - c_{02}^1}{2}.$$

In addition, we find the relations between the invariants

$$\chi + \kappa = c_{01}^2, \qquad \chi - \kappa = c_{02}^1.$$

We have the following subcases:

(a) If $c_{02}^1 = 0$ we get the Lie algebra $\mathfrak{se}(2)$ of the group $SE(2)$ of the Euclidean isometries of $\mathbb{R}^2$, and it holds $\chi = \kappa$.

(b) If $c_{01}^2 = 0$ we get the Lie algebra $\mathfrak{sh}(2)$ of the group $SH(2)$ of the Hyperbolic isometries of $\mathbb{R}^2$, and it holds $\chi = -\kappa$.

(c) If $c_{01}^2 > 0$ and $c_{02}^1 < 0$ we get the Lie algebra $\mathfrak{su}(2)$ and $\chi - \kappa < 0$.

(d) If $c_{01}^2 < 0$ and $c_{02}^1 > 0$ we get the Lie algebra $\mathfrak{sl}(2)$ with $\chi + \kappa < 0$.

(e) If $c_{01}^2 > 0$ and $c_{02}^1 > 0$ we get the Lie algebra $\mathfrak{sl}(2)$ with $\chi+\kappa > 0, \chi-\kappa > 0$.

(ii) $c_{02}^1 = 0$ and $c_{12}^1 = 0$. In this case we have

$$\begin{aligned} [f_1, f_0] &= c_{01}^2 f_2, \\ [f_2, f_0] &= 0, \\ [f_2, f_1] &= c_{12}^2 f_2 + f_0, \end{aligned} \tag{2.25}$$

and necessarily $c_{01}^2 \neq 0$. Moreover we get

$$\chi = \frac{c_{01}^2}{2} > 0, \qquad \kappa = -(c_{12}^2)^2 + \frac{c_{01}^2}{2},$$

from which it follows

$$\chi - \kappa \geq 0.$$

The Lie algebra $\mathfrak{g} = \text{span}\{f_1, f_2, f_3\}$ defined by (2.25) satisfies $\dim [\mathfrak{g}, \mathfrak{g}] = 2$, hence it can be interpreted as the operator $A = \text{ad } f_1$ which acts on the subspace $\text{span}\{f_0, f_2\}$. Moreover, it can be easily computed that

$$\text{trace } A = -c_{12}^2, \qquad \det A = c_{01}^2 > 0,$$

and we can find the useful relation

$$2\frac{\text{trace}^2 A}{\det A} = 1 - \frac{\kappa}{\chi}. \tag{2.26}$$

(iii) $c_{01}^2 = 0$ and $c_{12}^2 = 0$. In this last case we get

$$\begin{aligned} [f_1, f_0] &= 0, \\ [f_2, f_0] &= c_{02}^1 f_1, \\ [f_2, f_1] &= c_{12}^1 f_1 + f_0, \end{aligned} \tag{2.27}$$

and $c_{02}^1 \neq 0$. Moreover we get

$$\chi = \frac{c_{02}^1}{2} > 0, \qquad \kappa = -(c_{12}^1)^2 - \frac{c_{02}^1}{2},$$

from which it follows

$$\chi + \kappa \leq 0.$$

As before, the Lie algebra $\mathfrak{g} = \operatorname{span}\{f_1, f_2, f_3\}$ defined by (2.27) has two-dimensional square and it can be interpreted as the operator $A = \operatorname{ad} f_2$ which acts on the plane $\operatorname{span}\{f_0, f_1\}$. It can be easily seen that it holds

$$\operatorname{trace} A = c_{12}^1, \qquad \det A = -c_{02}^1 < 0,$$

and we have an analogous relation

$$2\frac{\operatorname{trace}^2 A}{\det A} = 1 + \frac{\kappa}{\chi}. \tag{2.28}$$

Remark 2.14. Lie algebras of cases (ii) and (iii) are *solvable* algebras and we will denote respectively $\mathfrak{solv}^+$ and $\mathfrak{solv}^-$, where the sign depends on the determinant of the operator it represents. In particular, formulas (2.26) and (2.28) permits to recover the ratio between invariants (hence to determine a unique normalized structure) only from intrinsic properties of the operator. Notice that if $c_{12}^2 = 0$ we recover the normalized structure (i)-(a) while if $c_{12}^1 = 0$ we get the case (i)-(b).

Remark 2.15. The algebra $\mathfrak{sl}(2)$ is the only case where we can define two nonequivalent distributions which corresponds to the case that Killing form restricted on the distribution is positive definite (case (d)) or indefinite (case (e)). We will refer to the first one as the *elliptic* structure on $\mathfrak{sl}(2)$, denoted $\mathfrak{sl}_e(2)$, and with *hyperbolic* structure in the other case, denoting $\mathfrak{sl}_h(2)$.

2.4.2 Case $\chi = 0$

A direct consequence of Proposition 2.11 for left-invariant structures is the following

Corollary 2.16. *Let G, H be Lie groups with left-invariant sub-Riemannian structures and assume $\chi_G = \chi_H = 0$. Then G and H are locally isometric if and only if $\kappa_G = \kappa_H$.*

Thanks to this result it is very easy to complete our classification. Indeed it is sufficient to find all left-invariant structures such that $\chi = 0$ and to compare their second invariant κ.

A straightforward calculation leads to the following list of the left-invariant structures on simply connected three dimensional Lie groups with $\chi = 0$:

- H_3 is the Heisenberg nilpotent group; then $\kappa = 0$.
- $SU(2)$ with the Killing inner product; then $\kappa > 0$.
- $\widetilde{SL}(2)$ with the elliptic distribution and Killing inner product; then $\kappa < 0$.
- $A^+(\mathbb{R}) \oplus \mathbb{R}$; then $\kappa < 0$.

Remark 2.17. In particular, we have the following:

(i) All left-invariant sub-Riemannian structures on H_3 are locally isometric,

(ii) There exists on $A^+(\mathbb{R}) \oplus \mathbb{R}$ a unique (modulo dilations) left-invariant sub-Riemannian structure, which is locally isometric to $SL_e(2)$ with the Killing metric.

Proof of Theorem 2.1 is now completed and we can recollect our result as in Figure 2.1, where we associate to every normalized structure a point in the (κ,χ) plane: either $\chi=\kappa=0$, or (κ,χ) belong to the semicircle

$$\{(\kappa,\chi)\in\mathbb{R}^2, \chi^2+\kappa^2=1, \chi>0\}.$$

Notice that different points means that sub-Riemannian structures are not locally isometric.

2.5 Sub-Riemannian isometry

In this section we want to write explicitly the sub-Riemannian isometry between $SL(2)$ and $A^+(\mathbb{R})\times S^1$.

Consider the Lie algebra $\mathfrak{sl}(2)=\{A\in M_2(\mathbb{R}),\ \mathrm{trace}(A)=0\}=\mathrm{span}\{g_1,g_2,g_3\}$, where

$$g_1=\frac{1}{2}\begin{pmatrix}1&0\\0&-1\end{pmatrix},\quad g_2=\frac{1}{2}\begin{pmatrix}0&1\\1&0\end{pmatrix},\quad g_3=\frac{1}{2}\begin{pmatrix}0&1\\-1&0\end{pmatrix}.$$

The sub-Riemannian structure on $SL(2)$ defined by the Killing form on the elliptic distribution is given by the orthonormal frame

$$\Delta_{\mathfrak{sl}}=\mathrm{span}\{g_1,g_2\},\qquad \text{and}\qquad g_0:=-g_3, \tag{2.29}$$

is the Reeb vector field. Notice that this frame is already canonical since equations (2.17) are satisfied. Indeed

$$[g_1,g_0]=-g_2=\kappa g_2.$$

Recall that the universal covering of $SL(2)$, which we denote $\widetilde{SL}(2)$, is a simply connected Lie group with Lie algebra $\mathfrak{sl}(2)$. Hence (2.29) define a left-invariant structure also on the universal covering.

On the other hand we consider the following coordinates on the Lie group $A^+(\mathbb{R})\oplus\mathbb{R}$, that are well-adapted for our further calculations

$$A^+(\mathbb{R})\oplus\mathbb{R}:=\left\{\begin{pmatrix}-y&0&x\\0&1&z\\0&0&1\end{pmatrix},\quad y<0,x,z\in\mathbb{R}\right\}. \tag{2.30}$$

It is easy to see that, in these coordinates, the group law reads

$$(x,y,z)(x',y',z')=(x-yx',-yy',z+z'),$$

and its Lie algebra $\mathfrak{a}(\mathbb{R})\oplus\mathbb{R}$ is generated by the vector fields

$$e_1=-y\partial_x,\quad e_2=-y\partial_y,\quad e_3=\partial_z,$$

with the only nontrivial commutator relation $[e_1,e_2]=e_1$.

The left-invariant structure on $A^+(\mathbb{R})\oplus\mathbb{R}$ is defined by the orthonormal frame

$$\begin{aligned}\Delta_{\mathfrak{a}}&=\mathrm{span}\{f_1,f_2\},\\ f_1&:=e_2=-y\partial_y,\\ f_2&:=e_1+e_3=-y\partial_x+\partial_z.\end{aligned} \tag{2.31}$$

With straightforward calculations we compute the Reeb vector field $f_0 = -e_3 = -\partial_z$.

This frame is not canonical since it does not satisfy equations (2.17). Hence we can apply Proposition 2.11 to find the canonical frame, that will be no more left-invariant.

Following the notation of Proposition 2.11 we have

Lemma 2.18. *The canonical orthonormal frame on $A^+(\mathbb{R}) \oplus \mathbb{R}$ has the form:*

$$\begin{aligned} \widehat{f}_1 &= \quad y \sin z \, \partial_x - y \cos z \, \partial_y - \sin z \, \partial_z, \\ \widehat{f}_2 &= -y \cos z \, \partial_x - y \sin z \, \partial_y + \cos z \, \partial_z. \end{aligned} \tag{2.32}$$

Proof. It is equivalent to show that the rotation defined in the proof of Proposition 2.11 is $\theta(x, y, z) = z$. The dual basis to our frame $\{f_1, f_2, f_0\}$ is given by

$$\nu_1 = -\frac{1}{y} dy, \qquad \nu_2 = -\frac{1}{y} dx, \qquad \nu_0 = -\frac{1}{y} dx - dz.$$

Moreover we have $[f_1, f_0] = [f_2, f_0] = 0$ and $[f_2, f_1] = f_2 + f_0$ so that, in equation (2.20) we get $c = 0, \alpha_1 = 0, \alpha_2 = 1$. Hence

$$d\theta = -\nu_0 + \nu_2 = dz.$$

□

Now we have two canonical frames $\{\widehat{f}_1, \widehat{f}_2, f_0\}$ and $\{g_1, g_2, g_0\}$, whose Lie algebras satisfy the same commutator relations:

$$\begin{aligned} &[\widehat{f}_1, f_0] = -\widehat{f}_2, && [g_1, g_0] = -g_2, \\ &[\widehat{f}_2, f_0] = \widehat{f}_1, && [g_2, g_0] = g_1, \\ &[\widehat{f}_2, \widehat{f}_1] = f_0, && [g_2, g_1] = 0. \end{aligned} \tag{2.33}$$

Let us consider the two control systems

$$\begin{aligned} \dot{q} &= u_1 \widehat{f}_1(q) + u_2 \widehat{f}_2(q) + u_0 f_0(q), \quad q \in A^+(\mathbb{R}) \oplus \mathbb{R}, \\ \dot{x} &= u_1 g_1(x) + u_2 g_2(x) + u_0 g_0(x), \quad x \in \widetilde{SL}(2). \end{aligned}$$

and denote with $x_u(t), q_u(t),\ t \in [0, T]$ the solutions of the equations relative to the same control $u = (u_1, u_2, u_0)$. Nagano Principle (see [13] and also [83, 97, 98]) ensure that the map

$$\widetilde{\Psi} : A^+(\mathbb{R}) \oplus \mathbb{R} \to \widetilde{SL}(2), \qquad q_u(T) \mapsto x_u(T). \tag{2.34}$$

that sends the final point of the first system to the final point of the second one, is well-defined and does not depend on the control u.

Thus we can find the endpoint map of both systems relative to constant controls, i.e. considering maps

$$\widetilde{F} : \mathbb{R}^3 \to A^+(\mathbb{R}) \oplus \mathbb{R}, \qquad (t_1, t_2, t_0) \mapsto e^{t_0 f_0} \circ e^{t_2 \widehat{f}_2} \circ e^{t_1 \widehat{f}_1}(1_A), \tag{2.35}$$

$$\widetilde{G} : \mathbb{R}^3 \to SL(2), \qquad (t_1, t_2, t_0) \mapsto e^{t_0 g_0} \circ e^{t_2 g_2} \circ e^{t_1 g_1}(1_{SL}). \tag{2.36}$$

where we denote with 1_A and 1_{SL} identity element of $A^+(\mathbb{R}) \oplus \mathbb{R}$ and $\widetilde{SL}(2)$, respectively.

The composition of these two maps makes the following diagram commutative

$$\begin{array}{ccc} A^+(\mathbb{R}) \oplus \mathbb{R} & \xrightarrow{\widetilde{\Psi}} & \widetilde{SL}(2) \\ \Big\downarrow{\scriptstyle \widetilde{F}^{-1}} & \searrow^{\Psi} & \Big\downarrow{\scriptstyle \pi} \\ \mathbb{R}^3 & \xrightarrow{\widetilde{G}} & SL(2) \end{array} \tag{2.37}$$

where $\pi : \widetilde{SL}(2) \to SL(2)$ is the canonical projection and we set $\Psi := \pi \circ \widetilde{\Psi}$.

To simplify computation we introduce the rescaled maps

$$F(t) := \widetilde{F}(2t), \qquad G(t) := \widetilde{G}(2t), \qquad t = (t_1, t_2, t_0),$$

and solving differential equations we get from (2.35) the following expressions

$$F(t_1, t_2, t_0) = \left(2e^{-2t_1} \frac{\tanh t_2}{1 + \tanh^2 t_2},\ -e^{-2t_1} \frac{1 - \tanh^2 t_2}{1 + \tanh^2 t_2},\ 2(\arctan(\tanh t_2) - t_0)\right). \tag{2.38}$$

The function F is globally invertible on its image and its inverse

$$F^{-1}(x, y, z) = \left(-\frac{1}{2} \log \sqrt{x^2 + y^2},\ \operatorname{arctanh}\left(\frac{y + \sqrt{x^2 + y^2}}{x}\right),\ \arctan\left(\frac{y + \sqrt{x^2 + y^2}}{x}\right) - \frac{z}{2}\right).$$

is defined for every $y < 0$ and for every x (it is extended by continuity at $x = 0$).

On the other hand, the map (2.36) can be expressed by the product of exponential matrices as follows[2]

$$G(t_1, t_2, t_0) = \begin{pmatrix} e^{t_1} & 0 \\ 0 & e^{-t_2} \end{pmatrix} \begin{pmatrix} \cosh t_2 & \sinh t_2 \\ \sinh t_2 & \cosh t_2 \end{pmatrix} \begin{pmatrix} \cos t_0 & -\sin t_0 \\ \sin t_0 & \cos t_0 \end{pmatrix}. \tag{2.39}$$

To simplify the computations, we consider standard polar coordinates (ρ, θ) on the half-plane $\{(x, y), y < 0\}$, where $-\pi/2 < \theta < \pi/2$ is the angle that the point (x, y) defines with y-axis. In particular, it is easy to see that the expression that appear in F^{-1} is naturally related to these coordinates:

$$\xi = \xi(\theta) := \tan \frac{\theta}{2} = \begin{cases} \dfrac{y + \sqrt{x^2 + y^2}}{x}, & \text{if} \quad x \neq 0, \\ 0, & \text{if} \quad x = 0. \end{cases}$$

Hence we can rewrite

$$F^{-1}(\rho, \theta, z) = \left(-\frac{1}{2} \log \rho,\ \operatorname{arctanh} \xi,\ \arctan \xi - \frac{z}{2}\right).$$

[2]since we consider left-invariant system, we must multiply matrices on the right.

and compute the composition $\Psi = G \circ F^{-1} : A^+(\mathbb{R}) \oplus \mathbb{R} \longrightarrow SL(2)$. Once we substitute these expressions in (2.39), the third factor is a rotation matrix by an angle $\arctan\xi - z/2$. Splitting this matrix in two consecutive rotations and using standard trigonometric identities $\cos(\arctan\xi) = \frac{1}{\sqrt{1+\xi^2}}$, $\sin(\arctan\xi) = \frac{\xi}{\sqrt{1+\xi^2}}$, $\cosh(\operatorname{arctanh}\xi) = \frac{1}{\sqrt{1-\xi^2}}$, $\sinh(\operatorname{arctanh}\xi) = \frac{\xi}{\sqrt{1-\xi^2}}$, for $\xi \in (-1,1)$, we obtain:

$$\Psi(\rho,\theta,z) =$$

$$= \begin{pmatrix} \rho^{-1/2} & 0 \\ 0 & \rho^{1/2} \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{1-\xi^2}} & \frac{\xi}{\sqrt{1-\xi^2}} \\ \frac{\xi}{\sqrt{1-\xi^2}} & \frac{1}{\sqrt{1-\xi^2}} \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{1+\xi^2}} & -\frac{\xi}{\sqrt{1+\xi^2}} \\ \frac{\xi}{\sqrt{1+\xi^2}} & \frac{1}{\sqrt{1+\xi^2}} \end{pmatrix} \begin{pmatrix} \cos\frac{z}{2} & \sin\frac{z}{2} \\ -\sin\frac{z}{2} & \cos\frac{z}{2} \end{pmatrix}.$$

Then using identities: $\cos\theta = \dfrac{1-\xi^2}{1+\xi^2}$, $\sin\theta = \dfrac{2\xi}{1+\xi^2}$, we get

$$\Psi(\rho,\theta,z) = \begin{pmatrix} \rho^{-1/2} & 0 \\ 0 & \rho^{1/2} \end{pmatrix} \begin{pmatrix} \frac{1+\xi^2}{\sqrt{1-\xi^4}} & 0 \\ \frac{2\xi}{\sqrt{1-\xi^4}} & \frac{1-\xi^2}{\sqrt{1-\xi^4}} \end{pmatrix} \begin{pmatrix} \cos\frac{z}{2} & \sin\frac{z}{2} \\ -\sin\frac{z}{2} & \cos\frac{z}{2} \end{pmatrix}$$

$$= \sqrt{\frac{1+\xi^2}{1-\xi^2}} \begin{pmatrix} \rho^{-1/2} & 0 \\ 0 & \rho^{1/2} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ \frac{2\xi}{1+\xi^2} & \frac{1-\xi^2}{1+\xi^2} \end{pmatrix} \begin{pmatrix} \cos\frac{z}{2} & \sin\frac{z}{2} \\ -\sin\frac{z}{2} & \cos\frac{z}{2} \end{pmatrix}$$

$$= \frac{1}{\sqrt{\rho\cos\theta}} \begin{pmatrix} 1 & 0 \\ 0 & \rho \end{pmatrix} \begin{pmatrix} 1 & 0 \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} \cos\frac{z}{2} & \sin\frac{z}{2} \\ -\sin\frac{z}{2} & \cos\frac{z}{2} \end{pmatrix}$$

$$= \frac{1}{\sqrt{\rho\cos\theta}} \begin{pmatrix} \cos\frac{z}{2} & \sin\frac{z}{2} \\ \rho\sin(\theta-\frac{z}{2}) & \rho\cos(\theta-\frac{z}{2}) \end{pmatrix}.$$

Lemma 2.19. *The set $\Psi^{-1}(I)$ is a normal subgroup of $A^+(\mathbb{R}) \oplus \mathbb{R}$.*

Proof. It is easy to show that $\Psi^{-1}(I) = \{F(0,0,2k\pi), k \in \mathbb{Z}\}$. From (2.38) we see that $F(0,0,2k\pi) = (0,-1,-4k\pi)$ and (2.30) implies that this is a normal subgroup. Indeed it is enoough to prove that $\Psi^{-1}(I)$ is a subgroup of the centre, that follows from the identity

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 4k\pi \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} -y & 0 & x \\ 0 & 1 & z \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} -y & 0 & x \\ 0 & 1 & z+4k\pi \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} -y & 0 & x \\ 0 & 1 & z \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 4k\pi \\ 0 & 0 & 1 \end{pmatrix}.$$

□

Remark 2.20. With a standard topological argument it is possible to prove that actually $\Psi^{-1}(A)$ is a discrete countable set for every $A \in SL(2)$, and Ψ is a representation of $A^+(\mathbb{R}) \oplus \mathbb{R}$ as universal covering of $SL(2)$.

By Lemma 2.19 the map Ψ is well defined isomorphism between the quotient

$$\frac{A^+(\mathbb{R}) \oplus \mathbb{R}}{\Psi^{-1}(I)} \simeq A^+(\mathbb{R}) \times S^1,$$

and the group $SL(2)$, defined by restriction of Ψ on $z \in [-2\pi, 2\pi]$.

If we consider the new variable $\varphi = z/2$, defined on $[-\pi, \pi]$, we can finally write the global isometry as

$$\Psi(\rho, \theta, \varphi) = \frac{1}{\sqrt{\rho \cos\theta}} \begin{pmatrix} \cos\varphi & \sin\varphi \\ \rho\sin(\theta - \varphi) & \rho\cos(\theta - \varphi) \end{pmatrix}, \tag{2.40}$$

where $(\rho, \theta) \in A^+(\mathbb{R})$ and $\varphi \in S^1$.

Remark 2.21. In the coordinate set defined above we have that $1_A = (1, 0, 0)$ and

$$\Psi(1_A) = \Psi(1, 0, 0) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = 1_{SL}.$$

On the other hand Ψ is not a homomorphism since in $A^+(\mathbb{R}) \oplus \mathbb{R}$ it holds

$$\Big(\frac{\sqrt{2}}{2}, \frac{\pi}{4}, \pi\Big)\Big(\frac{\sqrt{2}}{2}, -\frac{\pi}{4}, -\pi\Big) = 1_A,$$

while it can be easily checked from (2.40) that

$$\Psi\Big(\frac{\sqrt{2}}{2}, \frac{\pi}{4}, \pi\Big)\Psi\Big(\frac{\sqrt{2}}{2}, -\frac{\pi}{4}, -\pi\Big) = \begin{pmatrix} 2 & 0 \\ 1/2 & 1/2 \end{pmatrix} \neq 1_{SL}.$$

CHAPTER 3

The Hausdorff volume in sub-Riemannian geometry

In this chapter we address the problem of the volume in sub-Riemannian geometry, that naturally arise in the definition of the invariant sub-Riemannian Laplacian (see also the Introduction). For a regular sub-Riemannian manifold, we study the Radon-Nikodym derivative of the spherical Hausdorff measure with respect to a smooth volume (e.g. Popps volume) is proportional to the volume of the unit ball in the nilpotent approximation. We then prove that up to dimension 4 it is smooth, while starting from dimension 5, in corank 1 case, it is $\mathcal{C}^3$ (and $\mathcal{C}^4$ on every smooth curve) but in general not $\mathcal{C}^5$. These results answer to a question addressed by Montgomery about the relation between two intrinsic volumes that can be defined in a sub-Riemannian manifold, namely the Popp and the Hausdorff volume. If the nilpotent approximation depends on the point (that may happen starting from dimension 5), then they are not proportional, in general.

3.1 Introduction

A sub-Riemannian manifold $\mathbf{S} = (M, \Delta, \mathbf{g})$ has a natural structure of metric space, with the Carnot-Caratheodory distance d. Hence, for every $\alpha > 0$ one can define the α-dimensional Hausdorff measure on M, and compute the Hausdorff dimension of M.

Under the hypothesis that the sub-Riemannian manifold is regular, i.e. if the dimension of Δ_q^i, $i = 1, \ldots, m$ do not depend on the point, the Hörmander condition guarantees that there exists (a mimimal) $m \in \mathbb{N}$, called *step* of the structure, such that $\Delta_q^m = T_qM$, for all $q \in M$. The sequence

$$\mathcal{G}(\mathbf{S}) := (\underset{\underset{k}{\shortparallel}}{\dim \Delta}, \dim \Delta^2, \ldots, \underset{\underset{n}{\shortparallel}}{\dim \Delta^m})$$

is called *growth vector* of the structure.

In this case, the graded vector space associated to the filtration $\Delta_q \subset \Delta_q^2 \subset \ldots \subset \Delta_q^m = T_qM$,

$$\mathrm{gr}_q(\Delta) = \bigoplus_{i=1}^m \Delta_q^i/\Delta_q^{i-1}, \qquad \text{where} \quad \Delta_q^0 = 0.$$

is well defined. Moreover, the Hausdorff dimension of M is given by the formula (see [79])

$$Q = \sum_{i=1}^{m} i k_i, \qquad k_i := \dim \Delta_q^i / \Delta_q^{i-1}.$$

In particular the Hausdorff dimension is always bigger than the topological dimension of M.

Moreover, the Q-dimensional Hausdorff measure (denoted by $\mathcal{H}^Q$ in the following) behaves like a volume. More precisely, in [79] Mitchell proved that if μ is a smooth volume[1] on M , then $d\mu = f_{\mu\mathcal{H}} d\mathcal{H}^Q$, where $f_{\mu\mathcal{H}}$ is a positive measurable function that is locally bounded and locally bounded away from zero, that is the Radon-Nikodym derivative of μ with respect to $\mathcal{H}^Q$. According to Mitchell terminology, this means that the two measures are *commensurable* one with respect to the other.

Hausdorff measure on sub-Riemannian manifolds has been intensively studied, see for instance [65, 79]. A deep study of the Hausdorff measure for hypersurfaces in sub-Riemannian geometry, in particular in the context of Carnot groups, can be found in [16, 19, 20, 32, 44, 58, 77, 82] and references therein. Hausdorff measures for curves in sub-Riemannian manifolds were also studied in the problem of motion planning and complexity, see [60, 61, 62, 69].

Let us recall that there are two common non-equivalent definitions of Hausdorff measure. The standard Hausdorff measure, where arbitrary coverings can be used, and the spherical Hausdorff measure, where only ball-coverings appear (see Definition 3.8).

However it is well known that, if $\mathcal{S}^Q$ denotes the Q-dimensional spherical Hausdorff measure, then $\mathcal{H}^Q$ is commensurable with $\mathcal{S}^Q$.[2] As a consequence, $\mathcal{S}^Q$ is commensurable with μ, i.e.

$$d\mu = f_{\mu\mathcal{S}} d\mathcal{S}^Q,$$

for a positive measurable function $f_{\mu\mathcal{S}}$ that is locally bounded and locally bounded away from zero. In this paper, we are interested to the properties of the function $f_{\mu\mathcal{S}}$. In particular, we would like to get informations about its regularity.

The reason why we study the spherical Hausdorff measure and not the standard Hausdorff measure is that the first one appears to be more natural. Indeed, as explained later, $f_{\mu\mathcal{S}}$ is determined by the volume of the unit sub-Riemannian ball of the nilpotent approximation of the sub-Riemannian manifold, that can be explicitly described in a certain number of cases (see Theorem 3.1 below). On the other hand nothing is known on how to compute $f_{\mu\mathcal{H}}$. We conjecture that $f_{\mu\mathcal{H}}$ is given by the μ-volume of certain isodiametric sets, i.e. the maximum of the μ-volume among all sets of diameter 1 in the nilpotent approximation (see [76, 90] and reference therein for a discussion on isodiametric sets). This quantity is not very natural in sub-Riemannian geometry and is extremely difficult to compute.

Our interests in studying $f_{\mu\mathcal{S}}$ comes from the following question:

Q1 How can we define an intrinsic volume in a sub-Riemannian manifold?

[1] In the following by a smooth volume on M we mean a measure μ associated to a smooth non-vanishing n-form $\omega_\mu \in \Lambda^n M$, i.e. for every measurable subset $A \subset M$ we set $\mu(A) = \int_A \omega_\mu$.

[2] Indeed they are absolutely continuous one with respect to the other. In particular, for every $\alpha > 0$, we have $2^{-\alpha}\mathcal{S}^\alpha \leq \mathcal{H}^\alpha \leq \mathcal{S}^\alpha$ (see for instance [55]).

Here by intrinsic we mean a volume which depends neither on the choice of the coordinate system, nor on the choice of the orthonormal frame, but only on the sub-Riemannian structure.

This question was first pointed out by Brockett, see [41], and by Montgomery in his book [81]. Having a volume that depends only on the geometric structure is interesting by itself, however, it is also necessary to define intrinsically a Laplacian in a sub-Riemannian manifold. We recall that the Laplacian is defined as the divergence of the gradient and the definition of the divergence needs a volume since it measures how much the flow of a vector field increases or decreases the volume.

Before talking about the question **Q1** in sub-Riemannian geometry, let us briefly discuss it in the Riemannian case. In a n-dimensional Riemannian manifold there are three common ways of defining an invariant volume. The first is defined through the Riemannian structure and it is the so called Riemannian volume, which in coordinates has the expression $\sqrt{g}\, dx^1 \dots dx^n$, where g is the determinant of the metric. The second and the third ones are defined via the Riemannian distance and are the n-dimensional Hausdorff measure and the n-dimensional spherical Hausdorff measure. These three volumes are indeed proportional (the constant of proportionality depending on the normalization, see Remark 3.16 and e.g. [49, 55]).

For what concern sub-Riemannian geometry, a regular sub-Riemannian manifold is a metric space, hence it is possible to define the Hausdorff volume $\mathcal{H}^Q$ and the spherical Hausdorff volume $\mathcal{S}^Q$. Also, there is an equivalent of the Riemannian volume, the so called Popp's volume $\mathcal{P}$, introduced by Montgomery in his book [81] (see also [6]). The Popp volume is a smooth volume and was used in [6] to define intrinsically the Laplacian (indeed a sub-Laplacian) in sub-Riemannian geometry.

In his book, Montgomery proposed to study whether these invariant volumes are proportional as it occours in Riemannian geometry. More precisely, he addressed the following question:

Q2 Is Popp's measure equal to a constant multiple (perhaps depending on the growth vector) of the Hausdorff measure?

Mongomery noted that the answer to this question is positive for left-invariant sub-Riemannian structures on Lie groups, since the Hausdorff (both the standard and spherical one) and the Popp volumes are left-invariant and hence proportional to the left Haar measure. But this question is nontrivial when there is no group structure.

One of the main purpose of our analysis is to answer to question **Q2** for the spherical Hausdorff measure, i.e. to the question if the function $f_{\mathcal{PS}}$ (defined by $d\mathcal{P} = f_{\mathcal{PS}} d\mathcal{S}^Q$) is constant or not. More precisely, we get a positive answer for regular sub-Riemannian manifolds of dimension 3 and 4, while a negative answer starting from dimension 5, in general.

Once a negative answer to **Q2** is given, it is natural to ask

Q3 What is the regularity of $f_{\mathcal{PS}}$?

This question is important since the definition of an intrinsic Laplacian via $\mathcal{S}^Q$ require $f_{\mathcal{PS}}$ to be at least $\mathcal{C}^1$.

Notice that since the Popp measure is a smooth volume, then $f_{\mu\mathcal{S}}$ is $\mathcal{C}^k, k = 0, 1, \dots, \infty$ if and only if $f_{\mathcal{PS}}$ is as well.

We prove that $f_{\mu\mathcal{S}}$ is a continuous function and that for $n \leq 4$ it is smooth. In dimension 5 it is $\mathcal{C}^3$ but not smooth, in general. Moreover, we prove that the same result holds in all corank 1 cases.

Our main tool is the nilpotent approximation (sometimes also called the symbol) of the sub-Riemannian structure. Recall that, under the regularity hypothesis, the sub-Riemannian structure $\mathbf{S} = (M, \Delta, \mathbf{g})$ induces a structure of nilpotent Lie algebra on $\mathrm{gr}_q(\Delta)$. The nilpotent approximation at q is the nilpotent simply connected Lie group $\mathrm{Gr}_q(\Delta)$ generated by this Lie algebra, endowed with a suitable left-invariant sub-Riemannian structure $\widehat{\mathbf{S}}_q$ induced by $\mathbf{S}$, as explained in Section 1.3.

Recall that there exists a canonical isomorphism of 1-dimensional vector spaces (see [6] for details)

$$\bigwedge^n (T_q^* M) \simeq \bigwedge^n (\mathrm{gr}_q(\Delta)^*). \tag{3.1}$$

Given a smooth volume μ on M, we define the induced volume $\widehat{\mu}_q$ on the nilpotent approximation at point q as the left-invariant volume on $\mathrm{Gr}_q(\Delta)$ canonically associated with $\omega_\mu(q) \in \wedge^n(T_q^* M)$ by the above isomorphism.

The first result concerns an explicit formula for $f_{\mu\mathcal{S}}$.

Theorem 3.1. *Let $\mathbf{S} = (M, \Delta, \mathbf{g})$ be a regular sub-Riemannian manifold. Let μ be a volume on M and $\widehat{\mu}_q$ the induced volume on the nilpotent approximation at point $q \in M$. If A is an open subset of M, then*

$$\mu(A) = \frac{1}{2^Q} \int_A \widehat{\mu}_q(\widehat{B}_q)\, d\mathcal{S}^Q, \tag{3.2}$$

where $\widehat{B}_q$ is the unit ball in the nilpotent approximation at point q, i.e.

$$f_{\mu\mathcal{S}}(q) = \frac{1}{2^Q} \widehat{\mu}_q(\widehat{B}_q).$$

Starting from this formula we prove our first result about regularity of the density:

Corollary 3.2. *Let $\mathbf{S} = (M, \Delta, \mathbf{g})$ be a regular sub-Riemannian manifold and let μ be a smooth volume on M. Then the density $f_{\mu\mathcal{S}}$ is a continuous function.*

Theorem 3.1, specified for the Popp measure $\mathcal{P}$, permits to answer the Montgomery's question. Indeed, the measure $\widehat{\mathcal{P}}_q$ induced by $\mathcal{P}$ on the nilpotent approximation at point q coincides with the Popp measure built on $\widehat{\mathbf{S}}_q$, as a sub-Riemannian structure. In other words, if we denote $\mathcal{P}_{\widehat{q}}$ the Popp measure on $\widehat{\mathbf{S}}_q$, we get

$$\widehat{\mathcal{P}}_q = \mathcal{P}_{\widehat{q}}. \tag{3.3}$$

Hence, if the nilpotent approximation does not depend on the point, then $f_{\mathcal{P}\mathcal{S}}$ is constant. In other words we have the following corollary.

Corollary 3.3. *Let $\mathbf{S} = (M, \Delta, \mathbf{g})$ be a regular sub-Riemannian manifold and $\widehat{\mathbf{S}}_q$ its nilpotent approximation at point $q \in M$. If $\widehat{\mathbf{S}}_{q_1}$ is isometric to $\widehat{\mathbf{S}}_{q_2}$ for any $q_1, q_2 \in M$, then $f_{\mathcal{P}\mathcal{S}}$ is constant. In particular this happens if the sub-Riemannian structure is free.*

For the definition of free structure see [81].

Notice that, in the Riemannian case, nilpotent approximations at different points are isometric, hence the Hausdorff measure is proportional to the Riemannian volume (see [49, 55]).

When the nilpotent approximation contains parameters that are function of the point, then, in general, $f_{\mathcal{PS}}$ is not constant. We have analyzed in details all growth vectors in dimension less or equal than 5:

- dimension 3: (2,3),
- dimension 4: (2,3,4), (3,4),
- dimension 5: (2,3,5), (3,5), (4,5) and the non generic cases (2,3,4,5), (3,4,5).

In all cases the nilpotent approximation is unique, except for the (4,5) case. As a consequence, we get:

Theorem 3.4. *Let $\mathbf{S} = (M, \Delta, \mathbf{g})$ be a regular sub-Riemannian manifold of dimension $n \leq 5$. Let μ be a smooth volume on M and $\mathcal{P}$ be the Popp measure. Then*

(i) if $\mathcal{G}(\mathbf{S}) \neq (4,5)$, then $f_{\mathcal{PS}}$ is constant. As a consequence $f_{\mu\mathcal{S}}$ is smooth.

(ii) if $\mathcal{G}(\mathbf{S}) = (4,5)$, then $f_{\mu\mathcal{S}}$ is $\mathcal{C}^3$ (and $\mathcal{C}^4$ on smooth curves) but not $\mathcal{C}^5$, in general.

Actually the regularity result obtained in the (4,5) case holds for all corank 1 structures, as specified by the following theorem.

Theorem 3.5. *Let $\mathbf{S} = (M, \Delta, \mathbf{g})$ be a regular corank 1 sub-Riemannian manifold of dimension $n \geq 5$. Let μ be a smooth volume on M. Then $f_{\mu\mathcal{S}}$ is $\mathcal{C}^3$ (and $\mathcal{C}^4$ on smooth curves) but not $\mathcal{C}^5$, in general.*

Recall that for a corank 1 structure one has $\mathcal{G}(\mathbf{S}) = (n-1, n)$.

Notice that Theorem 3.5 apply in particular for the Popp measure. The loss of regularity of $f_{\mu\mathcal{S}}$ is due to the presence of what are called *resonance points*. More precisely, the parameters appearing in the nilpotent approximation are the eigenvalues of a certain skew-symmetric matrix which depends on the point. Resonances are the points in which these eigenvalues are multiple.

To prove Theorem 3.5, we have computed explicitly the optimal synthesis (i.e. all curves that minimize distance starting from one point) of the nilpotent approximation and, as a consequence, the volume of nilpotent balls $\widehat{B}_q$.

Another byproduct of our analysis is

Proposition 3.6. *Under the hypothesis of Theorem 3.5, if there are no resonance points then $f_{\mu\mathcal{S}}$ is smooth.*

The structure of the Chapter is the following. In Section 3.2 we provide normal forms for nilpotent structures in dimension less or equal than 5. In Section 3.3 we prove Theorem 3.1 and its corollaries, while in Section 3.4 we study the differentiability of the density for the corank 1 case. In the last Section we prove Theorem 3.4.

3.1.1 Hausdorff measures

In this section we recall definitions of Hausdorff measure and spherical Hausdorff measure. We start with the definition of smooth volume.

Definition 3.7. Let M be a n-dimensional smooth manifold, which is connected and orientable. By a *smooth volume* on M we mean a measure μ on M associated to a smooth non-vanishing n-form $\omega_\mu \in \Lambda^n M$, i.e. for every subset $A \subset M$ we set

$$\mu(A) = \int_A \omega_\mu.$$

The Popp volume $\mathcal{P}$, which is a smooth volume in the sense of Definition 3.7, is the volume associated to a n-form $\omega_{\mathcal{P}}$ that can be intrinsically defined via the sub-Riemannian structure (see [6, 81]).

Let (M, d) be a metric space and denote with $\mathcal{B}$ the set of balls in M.

Definition 3.8. Let A be a subset of M and $\alpha > 0$.

The *α-dimensional Hausdorff measure* of A is

$$\mathcal{H}^\alpha(A) := \lim_{\delta \to 0} \mathcal{H}^\alpha_\delta(A),$$

where

$$\mathcal{H}^\alpha_\delta(A) := \inf\{\sum_{i=1}^{\infty} \operatorname{diam}(A_i)^\alpha, A \subset \bigcup_{i=1}^{\infty} A_i, \operatorname{diam}(A_i) < \delta\}.$$

The *α-dimensional spherical Hausdorff measure* of A is

$$\mathcal{S}^\alpha(A) := \lim_{\delta \to 0} \mathcal{S}^\alpha_\delta(A),$$

where

$$\mathcal{S}^\alpha_\delta(A) := \inf\{\sum_{i=1}^{\infty} \operatorname{diam}(B_i)^\alpha, A \subset \bigcup_{i=1}^{\infty} B_i, B_i \in \mathcal{B}, \operatorname{diam}(B_i) < \delta\}.$$

These two measures are commensurable since it holds (see [55])

$$2^{-\alpha}\mathcal{S}^\alpha(A) \leq \mathcal{H}^\alpha(A) \leq \mathcal{S}^\alpha(A), \qquad \forall\, A \subset M. \tag{3.4}$$

The *Hausdorff dimension* of A is defined as

$$\inf\{\alpha > 0, \mathcal{H}^\alpha(A) = 0\} = \sup\{\alpha > 0, \mathcal{H}^\alpha(A) = +\infty\}. \tag{3.5}$$

Formula (3.4) guarantees that Hausdorff dimension of A does not change if we replace $\mathcal{H}^\alpha$ with $\mathcal{S}^\alpha$ in formula (3.5).

It is a standard fact that the Hausdorff dimension of a Riemannian manifold, considered as a metric space, coincides with its topological dimension. On the other side, we have the following

Theorem 3.9. *Let $(M, \Delta, \mathbf{g})$ be a regular sub-Riemannian manifold. Its Hausdorff dimension as a metric space is*

$$Q = \sum_{i=1}^{m} i k_i, \qquad k_i := \dim \Delta^i - \dim \Delta^{i-1}.$$

Moreover $\mathcal{S}^Q$ is commensurable to a smooth volume μ on M, i.e. for every compact $K \subset M$ there exists $\alpha_1, \alpha_2 > 0$ such that

$$\alpha_1 \mathcal{S}^Q \leq \mu \leq \alpha_2 \mathcal{S}^Q. \tag{3.6}$$

This theorem was proved by Mitchell in [79]. In its original version it was stated for the Lebesgue measure and the standard Hausdorff measure.

Definition 3.10. Let $\mathbf{S} = (M, \Delta, \mathbf{g})$ be a regular sub-Riemannian manifold and $\widehat{\mathbf{S}}_q$ its nilpotent approximation near q. If μ a smooth volume on M, associated to the smooth non-vanishing n-form ω_μ, we define the *induced volume* $\widehat{\mu}_q$ at the point q as the left-invariant volume on $\widehat{\mathbf{S}}_q$ canonically associated with $\omega_\mu(q) \in \wedge^n(T_q^*M)$ (cf. isomorphism (3.1)).

From Theorem 1.19 and the relation[3] $\mu(\delta_\varepsilon A) = \varepsilon^Q \widehat{\mu}_q(A) + o(\varepsilon^Q)$ when $\varepsilon \to 0$, one gets

Corollary 3.11. *Let μ be a smooth volume on M and $\widehat{\mu}_q$ the induced volume on the nilpotent approximation at point q. Then, for $\varepsilon \to 0$, we have*

$$\mu(B(q, \varepsilon)) = \varepsilon^Q \widehat{\mu}_q(\widehat{B}_q) + o(\varepsilon^Q).$$

3.2 Normal forms for nilpotent approximation in dimension ≤ 5

In this section we provide normal forms for the nilpotent approximation of regular sub-Riemannian structures in dimension less or equal than 5. One can easily shows that in this case the only possibilities for growth vectors are:

- $\dim(M) = 3$: $\quad \mathcal{G}(\mathbf{S}) = (2, 3)$,
- $\dim(M) = 4$: $\quad \mathcal{G}(\mathbf{S}) = (2, 3, 4)$ or $\mathcal{G}(\mathbf{S}) = (3, 4)$,
- $\dim(M) = 5$: $\quad \mathcal{G}(\mathbf{S}) \in \{(2, 3, 4, 5), (2, 3, 5), (3, 5), (3, 4, 5), (4, 5)\}$.

We have the following.

Theorem 3.12. *Let $\mathbf{S} = (M, \Delta, \mathbf{g})$ be a regular sub-Riemannian manifold and $\widehat{\mathbf{S}}_q$ its nilpotent approximation near q. Up to a change of coordinates and a rotation of the orthonormal frame we have the following expression for the orthonormal frame of $\widehat{\mathbf{S}}_q$:*

[3] Notice that this formula is meaningful in privileged coordinates near q.

Case $n = 3$ • $\mathcal{G}(\mathbf{S}) = (2, 3)$. *(Heisenberg)*

$$\begin{aligned} \widehat{X}_1 &= \partial_1, \\ \widehat{X}_2 &= \partial_2 + x_1 \partial_3. \end{aligned}$$

Case $n = 4$ • $\mathcal{G}(\mathbf{S}) = (2, 3, 4)$. *(Engel)*

$$\begin{aligned} \widehat{X}_1 &= \partial_1, \\ \widehat{X}_2 &= \partial_2 + x_1 \partial_3 + x_1 x_2 \partial_4. \end{aligned}$$

• $\mathcal{G}(\mathbf{S}) = (3, 4)$. *(Quasi-Heisenberg)*

$$\begin{aligned} \widehat{X}_1 &= \partial_1, \\ \widehat{X}_2 &= \partial_2 + x_1 \partial_4, \\ \widehat{X}_3 &= \partial_3. \end{aligned}$$

Case $n = 5$ • $\mathcal{G}(\mathbf{S}) = (2, 3, 5)$. *(Cartan)*

$$\begin{aligned} \widehat{X}_1 &= \partial_1, \\ \widehat{X}_2 &= \partial_2 + x_1 \partial_3 + \frac{1}{2} x_1^2 \partial_4 + x_1 x_2 \partial_5. \end{aligned}$$

• $\mathcal{G}(\mathbf{S}) = (2, 3, 4, 5)$. *(Goursat rank 2)*

$$\begin{aligned} \widehat{X}_1 &= \partial_1, \\ \widehat{X}_2 &= \partial_2 + x_1 \partial_3 + \frac{1}{2} x_1^2 \partial_4 + \frac{1}{6} x_1^3 \partial_5. \end{aligned}$$

• $\mathcal{G}(\mathbf{S}) = (3, 5)$. *(Corank 2)*

$$\begin{aligned} \widehat{X}_1 &= \partial_1 - \frac{1}{2} x_2 \partial_4, \\ \widehat{X}_2 &= \partial_2 + \frac{1}{2} x_1 \partial_4 - \frac{1}{2} x_3 \partial_5, \\ \widehat{X}_3 &= \partial_3 + \frac{1}{2} x_2 \partial_5. \end{aligned}$$

• $\mathcal{G}(\mathbf{S}) = (3, 4, 5)$. *(Goursat rank 3)*

$$\begin{aligned} \widehat{X}_1 &= \partial_1 - \frac{1}{2} x_2 \partial_4 - \frac{1}{3} x_1 x_2 \partial_5, \\ \widehat{X}_2 &= \partial_2 + \frac{1}{2} x_1 \partial_4 + \frac{1}{3} x_1^2 \partial_5, \\ \widehat{X}_3 &= \partial_3. \end{aligned}$$

- $\mathcal{G}(\mathbf{S}) = (4,5)$. *(Bi-Heisenberg)*

$$\begin{aligned}\widehat{X}_1 &= \partial_1 - \frac{1}{2}x_2\partial_5,\\ \widehat{X}_2 &= \partial_2 + \frac{1}{2}x_1\partial_5,\\ \widehat{X}_3 &= \partial_3 - \frac{\alpha}{2}x_4\partial_5, \qquad \alpha \in \mathbb{R},\\ \widehat{X}_4 &= \partial_4 + \frac{\alpha}{2}x_3\partial_5.\end{aligned} \tag{3.7}$$

Proof. It is sufficient to find, for every such a structure, a basis of the Lie algebra such that the structural constants[4] are uniquely determined by the sub-Riemannian structure. We give a sketch of the proof for the $(2,3,4,5)$ and $(3,4,5)$ and $(4,5)$ cases. The other cases can be treated in a similar way.

(i). Let $\widehat{\mathbf{S}} = (G, \Delta, \mathbf{g})$ be a nilpotent $(3,4,5)$ sub-Riemannian structure. Since we deal with a left-invariant sub-Riemannian structure, we can identify the distribution Δ with its value at the identity of the group Δ_{id}. Let $\{e_1, e_2, e_3\}$ be a basis for Δ_{id}, as a vector subspace of the Lie algebra. By our assumption on the growth vector we know that

$$\dim \operatorname{span}\{[e_1,e_2],[e_1,e_3],[e_2,e_3]\}/\Delta_{id} = 1. \tag{3.8}$$

In other words, we can consider the skew-simmetric mapping

$$\Phi(\cdot,\cdot) := [\cdot,\cdot]/\Delta_{id} : \Delta_{id} \times \Delta_{id} \to T_{id}G/\Delta_{id}, \tag{3.9}$$

and condition (3.8) implies that there exists a one dimensional subspace in the kernel of this map. Let $\widehat{X}_3$ be a normalized vector in the kernel and consider its orthogonal subspace $D \subset \Delta_{id}$ with respect to the Euclidean product on Δ_{id}. Fix an arbitrary orthonormal basis $\{X_1, X_2\}$ of D and set $\widehat{X}_4 := [X_1, X_2]$. It is easy to see that $\widehat{X}_4$ does not change if we rotate the base $\{X_1, X_2\}$ and there exists a choice of this frame, denoted $\{\widehat{X}_1, \widehat{X}_2\}$, such that $[\widehat{X}_2, \widehat{X}_4] = 0$. Then set $\widehat{X}_5 := [\widehat{X}_1, \widehat{X}_4]$. Therefore we found a canonical basis for the Lie algebra that satisfies the following commutator relations:

$$[\widehat{X}_1, \widehat{X}_2] = \widehat{X}_4, \qquad [\widehat{X}_1, \widehat{X}_4] = \widehat{X}_5,$$

and all other commutators vanish. A standard application of the Campbell-Hausdorff formula gives the coordinate expression above.

(ii). Let us assume now that $\widehat{\mathbf{S}}$ is a nilpotent $(2,3,4,5)$ sub-Riemannian structure. As before we identify the distribution Δ with its value at the identity and consider any orthonormal basis $\{e_1, e_2\}$ for the 2-dimensional subspace Δ_{id}. By our assumption on $\mathcal{G}(\mathbf{S})$

$$\begin{gathered}\dim \operatorname{span}\{e_1, e_2, [e_1,e_2]\} = 3\\ \dim \operatorname{span}\{e_1, e_2, [e_1,e_2], [e_1,[e_1,e_2]], [e_2,[e_1,e_2]]\} = 4.\end{gathered} \tag{3.10}$$

[4]Let $X_1, \ldots, X_k$ be a basis of a Lie algebra $\mathfrak{g}$. The coefficients c_{ij}^{ℓ} that satisfy $[X_i, X_j] = \sum_{\ell} c_{ij}^{\ell} X_{\ell}$ are called structural constant of $\mathfrak{g}$.

As in (i), it is easy to see that there exists a choice of the orthonormal basis on Δ_{id}, which we denote $\{\widehat{X}_1, \widehat{X}_2\}$, such that $[\widehat{X}_2, [\widehat{X}_1, \widehat{X}_2]] = 0$. From this property and the Jacobi identity it follows $[\widehat{X}_2, [\widehat{X}_1, [\widehat{X}_1, \widehat{X}_2]]] = 0$. Then we set $\widehat{X}_3 := [\widehat{X}_1, \widehat{X}_2]$, $\widehat{X}_4 = [\widehat{X}_1, [\widehat{X}_1, \widehat{X}_2]]$ and $\widehat{X}_5 := [\widehat{X}_1, [\widehat{X}_1, [\widehat{X}_1, \widehat{X}_2]]]$. It is easily seen that (3.10) implies that these vectors are linearly independent and give a canonical basis for the Lie algebra, with the only nontrivial commutator relations:

$$[\widehat{X}_1, \widehat{X}_2] = \widehat{X}_3, \qquad [\widehat{X}_1, \widehat{X}_3] = \widehat{X}_4, \qquad [\widehat{X}_1, \widehat{X}_4] = \widehat{X}_5.$$

(iii). In the case $(4,5)$ since $\dim T_{id}G/\Delta_{id} = 1$, the map (3.9) is represented by a single 4×4 skew-simmetric matrix L. By skew-symmetricity its eigenvalues are purely imaginary $\pm ib_1, \pm ib_2$, one of which is different from zero. Assuming $b_1 \neq 0$ we have that $\alpha = b_2/b_1$. Notice that the structure is contact if and only if $\alpha \neq 0$ (see also Section 3.4.1 for more details on the normal form). □

Remark 3.13. Notice that, in the statement of Theorem 3.12, in all other cases the nilpotent approximation does not depend on any parameter, except for the $(4,5)$ case. As a consequence, up to dimension 5, the sub-Riemannian structure induced on the tangent space, and hence the Popp measure $\mathcal{P}$, does not depend on the point, except for the $(4,5)$ case.

In the $(4,5)$ case we have the following expression for the Popp's measure

$$\mathcal{P} = \frac{1}{\sqrt{b_1^2 + b_2^2}}\, dx_1 \wedge \ldots \wedge dx_5,$$

where b_1, b_2 are the eigenvalues of the skew-simmetric matrix that represent the Lie bracket map.

Since the normal forms in Theorem 3.12 do not depend on the point, except when $\mathcal{G}(\mathbf{S}) \neq (4,5)$, we have the following corollary

Corollary 3.14. *Let* $\mathbf{S} = (M, \Delta, \mathbf{g})$ *be a regular sub-Riemannian manifold such that* $\dim(M) \leq 5$ *and* $\mathcal{G}(\mathbf{S}) \neq (4,5)$. *Then if* $q_1, q_2 \in M$ *we have that* $\widehat{\mathbf{S}}_{q_1}$ *is isometric to* $\widehat{\mathbf{S}}_{q_2}$ *as sub-Riemannian manifolds.*

3.3 The density is the volume of nilpotent balls

In this section we prove Theorem 3.1, i.e.

$$f_{\mu\mathcal{S}}(q) = \frac{1}{2^Q}\widehat{\mu}_q(\widehat{B}_q). \tag{3.11}$$

It is well known that, being μ absolutely continuous with respect to $\mathcal{S}^Q$ (see Theorem 3.9), the Radon-Nikodym derivative of μ with respect to $\mathcal{S}^Q$, namely $f_{\mu\mathcal{S}}$, can be computed almost everywhere as

$$\lim_{r\to 0} \frac{\mu(B(q,r))}{\mathcal{S}^Q(B(q,r))}.$$

By Corollary 3.11 we get

$$\frac{\mu(B(q,r))}{\mathcal{S}^Q(B(q,r))} = \frac{r^Q\widehat{\mu}_q(\widehat{B}_q)+o(r^Q)}{\mathcal{S}^Q(B(q,r))} = \frac{\widehat{\mu}_q(\widehat{B}_q)}{2^Q}\frac{2^Qr^Q}{\mathcal{S}^Q(B(q,r))} + \frac{o(r^Q)}{\mathcal{S}^Q(B(q,r))}.$$

Then we are left to prove the following

Lemma 3.15. *Let A be an open subset of M. For $\mathcal{S}^Q$-a.e. $q\in A$ we have*

$$\lim_{r\to 0}\frac{\mathcal{S}^Q(A\cap B(q,r))}{(2r)^Q} = 1. \tag{3.12}$$

Proof. In the following proof we make use of Vitali covering lemma [5] and we always assume that balls of our covering are small enough to satisfy property (ii) of Proposition 1.12.

We prove that the set where (3.12) exists and is different from 1 has $\mathcal{S}^Q$-null measure.
(i). First we show

$$\mathcal{S}^Q(E_\delta)=0, \qquad \forall\, 0<\delta\leq 1,$$

where

$$E_\delta := \{q\in A : \mathcal{S}^Q(A\cap B(q,r)) < (1-\delta)(2r)^Q, \forall\, 0<r<\delta\}$$

Let $\{B_i\}$ a ball covering of E_δ with $\mathrm{diam}(B_i)<\delta$ and such that

$$\sum_i \mathrm{diam}(B_i)^Q \leq \mathcal{S}^Q_\delta(E_\delta)+\varepsilon \leq \mathcal{S}^Q(E_\delta)+\varepsilon.$$

Then we have

$$\begin{aligned}\mathcal{S}^Q(E_\delta) &\leq \mathcal{S}^Q(A\cap \bigcup B_i)\\ &\leq \sum \mathcal{S}^Q(A\cap B_i)\\ &\leq (1-\delta)\sum \mathrm{diam}(B_i)^Q\\ &\leq (1-\delta)(\mathcal{S}^Q(E_\delta)+\varepsilon).\end{aligned}$$

then $\varepsilon\to 0$ and $1-\delta<1$ implies $\mathcal{S}^Q(E_\delta)=0$.
(ii). Next we prove that

$$\mathcal{S}^Q(E_t)=0, \quad \forall\, t>1,$$

where

$$E_t := \{q\in A : \mathcal{S}^Q(A\cap B(q,r)) > t(2r)^Q, \forall\, r \,\text{small enough}\}.$$

[5] **Theorem.**(Vitali covering lemma, [55, 17]) Let E be a metric space, $B\subset E$ and $\alpha>0$ such that $\mathcal{H}^\alpha(B)<\infty$, and let $\mathcal{F}$ be a fine covering of B. Then there exist a countable disjoint subfamily $\{V_i\}\subset\mathcal{F}$ such that

$$\mathcal{H}^\alpha(B\setminus \bigcup V_i)=0.$$

We recall that $\mathcal{F}$ is a *fine covering* of B if for every $x\in B$ and $\varepsilon>0$ there exists $V\in\mathcal{F}$ such that $x\in V$ and $\mathrm{diam}(V)<\varepsilon$.

Now let U be an open set such that $E_t \subset U$ and $\mathcal{S}^Q(A \cap U) < \mathcal{S}^Q(E_t) + \varepsilon$. We define

$$\mathcal{F} := \{B(q,r) : q \in E_t, B(q,r) \subset U, \operatorname{diam} B(q,r) \leq \delta\}.$$

Now we can apply Vitali covering lemma to $\mathcal{F}$ and get a family $\{B_i\}$ of disjoint balls such that $\mathcal{S}^Q(E_t \setminus \bigcup_i B_i) = 0$. Then we get

$$\begin{aligned}
\mathcal{S}^Q(E_t) + \varepsilon &> \mathcal{S}^Q(A \cap U) \\
&\geq \mathcal{S}^Q(A \cap \bigcup B_i) \\
&\geq t \sum \operatorname{diam}(B_i)^Q \\
&\geq t\, \mathcal{S}^Q_\delta(E_t \cap \bigcup B_i) \\
&\geq t\, \mathcal{S}^Q_\delta(E_t).
\end{aligned}$$

Letting $\varepsilon, \delta \to 0$ we have an absurd because $t > 1$. □

Since A is open, from this lemma follows formula (3.11).

Remark 3.16. Notice that, for a n-dimensional Riemannian manifold, the tangent spaces at different points are isometric. As a consequence the Riemannian volume of the unit ball in the tangent space is constant and one can show that it is $C_n = \pi^{\frac{n}{2}}/\Gamma(\frac{n}{2}+1)$. Formula (3.2), where $\mu = \text{Vol}$ is the Riemannian volume, implies the well-known relation between Vol and the (spherical) Hausdorff measure

$$d\text{Vol} = \frac{C_n}{2^n}\, d\mathcal{S}^n = \frac{C_n}{2^n}\, d\mathcal{H}^n.$$

3.3.1 Continuity of the density

In this section we prove Corollary 3.2. More precisely we study the continuity of the map

$$f_{\mu\mathcal{S}} : q \mapsto \widehat{\mu}_q(\widehat{B}_q). \tag{3.13}$$

To this purpose, it is sufficient to study the regularity under the hypothesis that $\widehat{\mu}_q$ does not depend on the point. Indeed it is easily seen that the smooth measure μ, which is defined on the manifold, induces on the nilpotent approximations a smooth family of measures $\{\widehat{\mu}_q\}_{q \in M}$. In the case $\mu = \mathcal{P}$, this is a consequence of equality (3.3). In other words we can identify all tangent spaces in coordinates with $\mathbb{R}^n$ and fix a measure $\widehat{\mu}$ on it.

We are then reduced to study the regularity of the volume of the unit ball of a smooth family of nilpotent structures in $\mathbb{R}^n$, with respect to a fixed smooth measure. Notice that this family depends on an n-dimensional parameter.

To sum up, we have left to study the regularity of the map

$$q \mapsto \mathcal{L}(\widehat{B}_q), \qquad q \in M, \tag{3.14}$$

where $\widehat{B}_q$ is the unit ball of a family of nilpotent structures $\widehat{\mathbf{S}}_q$ in $\mathbb{R}^n$ and $\mathcal{L}$ is the standard Lebesgue measure.

Let us denote $\widehat{d}_q$ the sub-Riemannian distance in $\widehat{\mathbf{S}}_q$ and $\rho_q := \widehat{d}_q(0,\cdot)$. Following this notation $\widehat{B}_q = \{x \in \mathbb{R}^n |\, \rho_q(x) \leq 1\}$ and the coordinate expression (1.11) implies that

$$\mathcal{L}(\delta_\alpha(\widehat{B}_q)) = \alpha^Q \mathcal{L}(\widehat{B}_q), \qquad \forall\, \alpha > 0. \tag{3.15}$$

Notice that, since our sub-Riemannian structure is regular, we can choose privileged coordinates $\psi_q : O_q \to \mathbb{R}^n$ smoothly with respect to q. Let now $q' \neq q$, there exists $\alpha = \alpha(q, q')$ such that (see Remark 1.18)

$$\delta_{\frac{1}{\alpha}} \widehat{B}_{q'} \subset \widehat{B}_q \subset \delta_\alpha \widehat{B}_{q'}. \tag{3.16}$$

Using (3.15), (3.16) and monotonicity of the volume we get

$$\left(\frac{1}{\alpha^Q} - 1\right) \mathcal{L}(\widehat{B}_{q'}) \leq \mathcal{L}(\widehat{B}_q) - \mathcal{L}(\widehat{B}_{q'}) \leq (\alpha^Q - 1)\mathcal{L}(\widehat{B}_{q'}).$$

Then it is sufficient to show that $\alpha(q, q') \to 1$ when $q' \to q$. This property follows from the next

Lemma 3.17. *The family of functions $\rho_q|_K$ is equicontinuous for every compact $K \subset \mathbb{R}^n$. Moreover $\rho_{q'} \to \rho_q$ uniformly on compacts in $\mathbb{R}^n$, as $q' \to q$.*

In the case in which $\{\rho_t\}_{t>0}$ is the approximating family of the nilpotent distance $\widehat{\rho}$, this result is proved in [7]. See also [3] for a more detailed proof, using cronological calculus. With the same arguments one can extend this result to any smooth family of regular sub-Riemannian structures. The key point is that we can construct a basis for the tangent space to the structure with bracket polynomials of the orthonormal frame where the structure of the brackets does not depend on the parameter.

3.4 Differentiability of the density in the corank 1 case

In this section we prove Theorem 3.5, We start by studying the contact case. Then we complete our analysis by reducing the quasi-contact case and the general case to the contact one.

3.4.1 Normal form of the nilpotent contact case

Consider a 2-step nilpotent sub-Riemannian manifold in $\mathbb{R}^n$ of rank k.

Select a basis $\{X_1, \ldots, X_k, Z_1, \ldots, Z_{n-k}\}$ such that

$$\begin{cases} \Delta = \operatorname{span}\{X_1, \ldots, X_k\}, \\ [X_i, X_j] = \displaystyle\sum_{h=1}^{n-k} b_{ij}^h Z_h, \qquad i,j = 1, \ldots, k, \quad \text{where} \quad b_{ij}^h = -b_{ji}^h, \\ [X_i, Z_j] = [Z_j, Z_h] = 0, \qquad i = 1, \ldots, k, \qquad j,h = 1, \ldots, n-k. \end{cases} \tag{3.17}$$

Hence the Lie bracket can be considered as a map

$$[\cdot,\cdot] : \Delta \times \Delta \longrightarrow TM/\Delta, \tag{3.18}$$

and is represented by the $n-k$ skew-simmetric matrices $L^h = (b^h_{ij}), h = 1, \dots, n-k$.

In the contact case we have $(k, n) = (2\ell, 2\ell+1)$ and our structure is represented by one non degenerate skew-symmetric matrix L. Take coordinates in such a way that L is normalized in the following block-diagonal form

$$L = \begin{pmatrix} B_1 & & \\ & \ddots & \\ & & B_\ell \end{pmatrix}, \qquad \text{where} \qquad B_i := \begin{pmatrix} 0 & -b_i \\ b_i & 0 \end{pmatrix}, \qquad b_i > 0.$$

with eigenvalues $\pm ib_1, \dots, \pm ib_\ell$. Hence we can find a basis of vector fields $\{X_1, \dots, X_\ell, Y_1, \dots, Y_\ell, Z\}$ such that relations (3.17) reads

$$\begin{cases} \Delta = \text{span}\{X_1, \dots, X_\ell, Y_1, \dots, Y_\ell\}, & \\ [X_i, Y_i] = -b_i Z, & i = 1, \dots, \ell, \\ [X_i, Y_j] = 0, & i \neq j, \\ [X_i, Z] = [Y_i, Z] = 0, & i = 1, \dots, \ell. \end{cases} \tag{3.19}$$

In the following we call $b_1, \dots, b_\ell$ *frequences* of the contact structure.

We can recover the product on the group by the Campbell-Hausdorff formula. If we denote points $q = (x, y, z)$, where

$$x = (x_1, \dots, x_\ell) \in \mathbb{R}^\ell, \qquad y = (y_1, \dots, y_\ell) \in \mathbb{R}^\ell, \qquad z \in \mathbb{R},$$

we can write the group law in coordinates

$$q \cdot q' = \left(x + x', y + y', z + z' - \frac{1}{2} \sum_{i=1}^{\ell} b_i (x_i x_i' - y_i y_i') \right). \tag{3.20}$$

Finally, from (3.20), we get the coordinate expression of the left-invariant vector fields of the Lie algebra, namely

$$\begin{aligned} X_i &= \partial_{x_i} + \frac{1}{2} b_i y_i \partial_z, \qquad i = 1, \dots, \ell, \\ Y_i &= \partial_{y_i} - \frac{1}{2} b_i x_i \partial_z, \qquad i = 1, \dots, \ell, \\ Z &= \partial_z. \end{aligned} \tag{3.21}$$

In this expression one of frequences b_i can be normalized to 1.

3.4.2 Exponential map in the nilpotent contact case

Now we apply the PMP to find the exponential map $\mathcal{E}_{q_0}$ where q_0 is the origin. Define the hamiltonians (linear on fibers)

$$h_{u_i}(\lambda) = \langle \lambda, X_i(q) \rangle, \qquad h_{v_i}(\lambda) = \langle \lambda, Y_i(q) \rangle, \qquad h_w(\lambda) = \langle \lambda, Z(q) \rangle.$$

Recall from Section 1.2 that $q(t)$ is a normal extremal if and only if there exists $\lambda(t)$ such that

$$\begin{cases} \dot{u}_i = -b_i w v_i \\ \dot{v}_i = b_i w u_i \\ \dot{w} = 0 \end{cases} \qquad \begin{cases} \dot{x}_i = u_i \\ \dot{y}_i = v_i \\ \dot{z} = \frac{1}{2} \sum_i b_i (u_i y_i - v_i x_i) \end{cases} \tag{3.22}$$

where

$$u_i(t) := h_{u_i}(\lambda(t)), \qquad v_i(t) := h_{v_i}(\lambda(t)), \qquad w(t) := h_w(\lambda(t)).$$

Remark 3.18. Notice that from (3.22) it follows that the sub-Riemannian length of a geodesic coincide with the Euclidean length of its projection on the horizontal subspace $(x_1, \ldots, x_n, y_1, \ldots, y_n)$.

$$l(\gamma) = \int_0^T \left(\sum_i (u_i^2(t) + v_i^2(t)) \right)^{\frac{1}{2}} dt.$$

Now we solve (3.22) with initial conditions (see also Remark 1.10)

$$(x^0, y^0, z^0) = (0, 0, 0), \tag{3.23}$$

$$(u^0, v^0, w^0) = (u_1^0, \ldots, u_\ell^0, v_1^0, \ldots, v_\ell^0, w^0) \in S^{2\ell-1} \times \mathbb{R}. \tag{3.24}$$

Notice that $w \equiv w^0$ is constant on geodesics. We consider separately the two cases:

(*i*) If $w \neq 0$, we have (denoting $a_i := b_i w$)

$$\begin{aligned} u_i(t) &= u_i^0 \cos a_i t - v_i^0 \sin a_i t, \\ v_i(t) &= u_i^0 \sin a_i t + v_i^0 \cos a_i t, \\ w(t) &= w. \end{aligned} \tag{3.25}$$

From (3.22) one easily get

$$\begin{aligned} x_i(t) &= \frac{1}{a_i}(u_i^0 \sin a_i t + v_i^0 \cos a_i t - v_i^0), \\ y_i(t) &= \frac{1}{a_i}(-u_i^0 \cos a_i t + v_i^0 \sin a_i t + u_i^0), \\ z(t) &= \frac{1}{2w^2}(wt - \sum_i \frac{1}{b_i}((u_i^0)^2 + (v_i^0)^2)) \sin a_i t). \end{aligned} \tag{3.26}$$

(*ii*) If $w = 0$, we find equations of straight lines on the horizontal plane in direction of the vector (u^0, v^0):

$$x_i(t) = u_i^0 t \qquad y_i(t) = v_i^0 t \qquad z(t) = 0.$$

Remark 3.19. To recover symmetry properties of the exponential map it is useful to rewrite (3.26) in polar coordinates, using the following change of variables

$$u_i = r_i \cos \theta_i, \qquad v_i = r_i \sin \theta_i, \qquad i = 1, \ldots, \ell. \tag{3.27}$$

In these new coordinates (3.26) becomes

$$\begin{aligned} x_i(t) &= \frac{r_i}{a_i}(\cos(a_i t + \theta_i) - \cos \theta_i), \\ y_i(t) &= \frac{r_i}{a_i}(\sin(a_i t + \theta_i) - \sin \theta_i), \\ z(t) &= \frac{1}{2w^2}(wt - \sum_i \frac{r_i^2}{b_i} \sin a_i t). \end{aligned} \tag{3.28}$$

and the condition $(u^0, v^0) \in S^{2\ell-1}$ implies that $r = (r_1, \ldots, r_\ell) \in S^\ell$.

From equations (3.28) we easily see that the projection of a geodesic on every 2-plane (x_i, y_i) is a circle, with period T_i, radius ρ_i and center C_i where

$$T_i = \frac{2\pi}{b_i w}, \qquad \rho_i = \frac{r_i}{b_i w} \qquad C_i = -\frac{r_i}{b_i w}(\cos\theta_i, \sin\theta_i), \qquad \forall\, i = 1, \ldots, \ell \tag{3.29}$$

Moreover (3.22) shows that the z component of the geodesic at time t is the weighted sum (with coefficients b_i) of the areas spanned by the vectors $(x_i(t), y_i(t))$ in $\mathbb{R}^2$.

Lemma 3.20. *Let $\gamma(t)$ be a geodesic starting from the origin and corresponding to the parameters (r_i, θ_i, w). The cut time t_{cut} along γ is equal to first conjugate time and satisfies*

$$t_{cut} = \frac{2\pi}{w \max_i b_i}, \tag{3.30}$$

with the understanding that $t_{cut} = +\infty$, if $w = 0$.

Proof. We divide the proof into two steps. Recall that a geodesic lose optimality either at a cut time or at a conjugate time, see Definition 1.14. First we prove that (4.35) is a conjugate time and then that for every $t < t_{cut}$ our geodesic is optimal.

The case $w = 0$ is trivial. Indeed the geodesic is a straight line and, by Remark 3.18, we have neither cut nor conjugate points. Then it is not restrictive to assume $w \neq 0$. Moreover, up to relabeling indices, we can also assume that $b_1 \geq b_2 \geq \ldots \geq b_\ell \geq 0$.

(i) Since, by assumption, $b_1 = \max_i b_i$, from (3.28) it is easily seen that projection on the (x_1, y_1)-plane satisfies

$$x_1(t_{cut}) = y_1(t_{cut}) = 0.$$

Next consider the one parametric family of geodesic with initial condition

$$(r_1, r_2, \ldots, r_\ell, \theta_1 + \phi, \theta_2, \ldots, \theta_\ell, w), \qquad \phi \in [0, 2\pi].$$

It is easily seen from equation (3.28) that all these curves have the same endpoints. Indeed neither (x_i, y_i), for $i > 1$, nor z depends on this variable. Then it follows that t_c is a critical time for exponential map, hence a conjugate time.

(ii) Since $w \neq 0$, our geodesic is non horizontal (i.e. $z(t) \not\equiv 0$). By symmetry, we can focus on the case $w > 0$. We know that, for every i, the projection of every non horizontal geodesic on on the plane (x_i, y_i) is a circle. Moreover for the i-th projected curve, the distance from the origin is easily computed

$$\eta_i(t) = \sqrt{x_i(t)^2 + y_i(t)^2} = r_i\, t\, \mathrm{sin}_c\Big(\frac{b_i w t}{2}\Big), \qquad \text{where} \qquad \mathrm{sin}_c(x) = \frac{\sin x}{x}.$$

Let now $\bar{t} < t_c$, we want to show that there are no others geodesics $\widetilde{\gamma}(t)$, starting from the origin, that reach optimally the point $\gamma(\bar{t})$ at the same time $\bar{t}$. Assume that $\widetilde{\gamma}(t)$ is associated to the parameters $(\widetilde{r}_i, \widetilde{\theta}_i, \widetilde{w})$, where $(\widetilde{r}_1, \ldots, \widetilde{r}_\ell) \in S^\ell$, and let us argue by contradiction. If $\gamma(\bar{t}) = \widetilde{\gamma}(\bar{t})$ it follows that $\eta_i(\bar{t}) = \widetilde{\eta}_i(\bar{t})$ for every i, that means

$$r_i\, \bar{t}\, \mathrm{sin}_c\Big(\frac{b_i w \bar{t}}{2}\Big) = \widetilde{r}_i\, \bar{t}\, \mathrm{sin}_c\Big(\frac{b_i \widetilde{w} \bar{t}}{2}\Big), \qquad i = 1, \ldots, \ell. \tag{3.31}$$

Notice that, once $\widetilde{w}$ is fixed, $\widetilde{r}_i$ are uniquely determined by (3.31) (recall that $\bar{t}$ is fixed). Moreover, $\widetilde{\theta}_i$ also are uniquely determined by relations (3.29). Finally, from the assumption that $\widetilde{\gamma}$ also reach optimally the point $\widetilde{\gamma}(\bar{t})$, it follows that

$$\bar{t} < t_c = \frac{2\pi}{b_1\widetilde{w}} \qquad \Longrightarrow \qquad \frac{b_i\widetilde{w}\bar{t}}{2} < \pi \quad \forall\, i = 1, \dots, \ell. \tag{3.32}$$

Since $\sin_c(x)$ is a strictly decreasing function on $[0, \pi]$ it follows from (3.31) that, if $\widetilde{w} \neq w$ we have

$$\begin{aligned} \widetilde{w} > w \quad &\Rightarrow \quad \widetilde{r}_i > r_i \ \ \forall\, i = 1, \dots, \ell \quad \Rightarrow \quad \sum_i \widetilde{r}_i^2 > \sum_i r_i^2 = 1, \\ \widetilde{w} < w \quad &\Rightarrow \quad \widetilde{r}_i < r_i \ \ \forall\, i = 1, \dots, \ell \quad \Rightarrow \quad \sum_i \widetilde{r}_i^2 < \sum_i r_i^2 = 1, \end{aligned}$$

which contradicts the fact that $(\widetilde{r}_1, \dots, \widetilde{r}_\ell) \in S^\ell$.

□

Consider now the exponential map from the origin

$$(r, \theta, w) \mapsto \mathsf{Exp}(r, \theta, w) = \begin{cases} x_i = \dfrac{r_i}{b_i w}(\cos(b_i w + \theta_i) - \cos\theta_i) \\ y_i = \dfrac{r_i}{b_i w}(\sin(b_i w + \theta_i) - \sin\theta_i) \\ z = \dfrac{1}{2w^2}(w|r|^2 - \sum_i \dfrac{r_i^2}{b_i}\sin b_i w) = \sum_i \dfrac{r_i^2}{2b_i w^2}(b_i w - \sin b_i w). \end{cases} \tag{3.33}$$

By Lemma 3.20, the set D where geodesics are optimal and their length is less or equal to 1, is characterized as follows

$$D := \{(r, \theta, w), |r| \leq 1, |w| \leq 2\pi/\max b_i\}.$$

Thus the restriction of the exponential map to the interior of D gives a regular parametrization of the nilpotent unit ball $\widehat{B}$, and we can compute its volume with the change of variables formula

$$\mathcal{L}(\widehat{B}) = \int_{\widehat{B}} d\mathcal{L} = \int_D |\det J|\, R\, dr d\theta dw, \tag{3.34}$$

where J denotes the Jacobian matrix of the exponential map (3.33). Notice that we have to integrate with respect to the measure $dudvdw = R\, dr d\theta dw$, where $R = \prod_i r_i$ because of the change of variables (3.27).

Lemma 3.21. *The Jacobian of Exponential map* (3.33) *is given by the formula*

$$\det J(r, \theta, w) = \frac{4^\ell R}{B^2 w^{2\ell+2}} \sum_{i=1}^{\ell} \left(\prod_{j \neq i} \sin^2(\frac{b_j w}{2}) \right) \sin(\frac{b_i w}{2}) \left(\frac{b_i w}{2}\cos(\frac{b_i w}{2}) - \sin(\frac{b_i w}{2}) \right) r_i^2, \tag{3.35}$$

where we denote with $B = \prod_i b_i$.

Proof. We reorder variables in the following way

$$(r_1, \theta_1, \dots, r_\ell, \theta_\ell, w), \qquad (x_1, y_1, \dots, x_\ell, y_\ell, z),$$

in such a way that the Jacobian matrix J of the exponential map (3.33) is

$$J = \begin{pmatrix} Q_1 & & & & W_1 \\ & Q_2 & & & W_2 \\ & & \ddots & & \vdots \\ & & & Q_\ell & W_\ell \\ Z_1 & Z_2 & \dots & Z_\ell & \partial_w z \end{pmatrix} \tag{3.36}$$

where we denote

$$Q_i = \begin{pmatrix} Q_i^r & Q_i^\theta \end{pmatrix} := \begin{pmatrix} \partial_{x_i} r_i & \partial_{x_i} \theta_i \\ \partial_{y_i} r_i & \partial_{y_i} \theta_i \end{pmatrix}, \qquad i = 1, \dots, \ell, \tag{3.37}$$

$$W_i := \begin{pmatrix} \partial_{x_i} w \\ \partial_{y_i} w \end{pmatrix}, \qquad Z_i := \begin{pmatrix} \partial_z r_i & \partial_z \theta_i \end{pmatrix}.$$

Notice that x_i, y_i depend only on r_i, θ_i.

To compute the determinant of J, we write

$$z = z(r, w) = \sum_i z_i(r_i, w), \qquad z_i := \frac{r_i^2}{2b_i w^2}(b_i w - \sin b_i w),$$

and we split the last column of J as a sum

$$\begin{pmatrix} W_1 \\ W_2 \\ \vdots \\ W_\ell \\ \partial_w z \end{pmatrix} = \begin{pmatrix} W_1 \\ 0 \\ \vdots \\ 0 \\ \partial_w z_1 \end{pmatrix} + \begin{pmatrix} 0 \\ W_2 \\ \vdots \\ 0 \\ \partial_w z_2 \end{pmatrix} + \dots + \begin{pmatrix} 0 \\ 0 \\ \vdots \\ W_\ell \\ \partial_w z_\ell \end{pmatrix}. \tag{3.38}$$

Notice that in the i-th column only the i-th variables appear.

By multilinearity of determinant, $\det J$ is the sum of the determinants of ℓ matrices, obtained by replacing each time last column with one of vectors appearing in the sum (3.38). If we replace it, for instance, with the first term, we get

$$J_1 = \begin{pmatrix} Q_1 & & & & W_1 \\ & Q_2 & & & 0 \\ & & \ddots & & \vdots \\ & & & Q_\ell & 0 \\ Z_1 & Z_2 & \dots & Z_\ell & \partial_w z_1 \end{pmatrix}$$

Now with straightforward computations (notice that $\partial_\theta z_1 = 0$ in Z_1), we get

$$\begin{aligned} \det J_1 &= \det Q_2 \cdots \det Q_\ell \cdot \det \begin{pmatrix} Q_1 & W_1 \\ Z_1 & \partial_w z_1 \end{pmatrix} \\ &= \det Q_2 \cdots \det Q_\ell \cdot \det \begin{pmatrix} Q_1^r & Q_1^\theta & W_1 \\ \partial_{r_1} z_1 & 0 & \partial_w z_1 \end{pmatrix}. \end{aligned}$$

Setting

$$A_i = \begin{pmatrix} Q_i^\theta & W_i \end{pmatrix} \qquad i = 1, \dots, \ell,$$

we find

$$\det J_1 = \det Q_2 \cdots \det Q_\ell \cdot (\partial_w z_1 \det Q_1 + \partial_{r_1} z_1 \det A_1).$$

Similarly, we find analogous expressions for $J_2, \dots, J_\ell$. Then

$$\det J = \sum_{i=1}^{\ell} \det J_i = \sum_{i=1}^{\ell} (\prod_{j\neq i} \det Q_j)(\partial_w z_i \det Q_i + \partial_{r_i} z_i \det A_i).$$

From (3.33), by direct computations, it follows

$$\det Q_j = \frac{4r_j}{b_j^2 w^2} \sin^2(\frac{b_j w}{2}),$$

Moreover, with some computations, one can get

$$\begin{aligned}
\partial_w z_i &= \frac{r_i^2}{w^2} \sin^2(\frac{b_i w}{2}) - \frac{r_i^2}{b_i w^3}(b_i w - \sin b_i w) \\
\partial_{r_i} z_i &= \frac{r_i}{b_i w^2}(b_i w - \sin b_i w) \\
\det A_i &= -\frac{4r_i^2}{b_i^2 w^3} \sin(\frac{b_i w}{2}) \left(\frac{b_i w}{2} \cos(\frac{b_i w}{2}) - \sin(\frac{b_i w}{2}) \right)
\end{aligned}$$

and finally, after some simplifications

$$\partial_w z_i \det Q_i + \partial_{r_i} z_i \det A_i = -\frac{4r_i^3}{b_i^2 w^4} \sin(\frac{b_i w}{2}) \left(\frac{b_i w}{2} \cos(\frac{b_i w}{2}) - \sin(\frac{b_i w}{2}) \right).$$

from which we get (3.35). □

From the explicit expression of the Jacobian (3.35) we see that integration with respect to horizontal variables (r_i, θ_i) does not involve frequences, providing a costant C_ℓ. Hence, the computation of the volume reduces to a one dimensional integral in the vertical variable w:

$$V = \int_{-\frac{2\pi}{\max b_i}}^{\frac{2\pi}{\max b_i}} \frac{C_\ell}{B^2 w^{2\ell+2}} \sum_{i=1}^{\ell} \left(\prod_{j\neq i} \sin^2(\frac{b_j w}{2}) \right) \sin(\frac{b_i w}{2}) \left(\frac{b_i w}{2} \cos(\frac{b_i w}{2}) - \sin(\frac{b_i w}{2}) \right) dw.$$

Using simmetry property of the jacobian with respect to w and making the change of variable $2s = w$, the volume become (reabsorbing all costants in C_ℓ)

$$V = \int_0^{\frac{\pi}{\max b_i}} \frac{C_\ell}{B^2 s^{2\ell+2}} \sum_{i=1}^{\ell} \left(\prod_{j\neq i} \sin^2(b_j s) \right) \sin(b_i s)(b_i s \cos(b_i s) - \sin(b_i s)) ds. \quad (3.39)$$

Remark 3.22. In the case $\ell = 1$ (Heisenberg group) we get

$$V = \frac{1}{12}(1 + 2\pi \operatorname{Si}(2\pi)) \simeq 0.8258, \qquad \operatorname{Si}(x) := \int_0^x \frac{\sin t}{t} dt.$$

Notice that this is precisely the value of the constant $f_{\mathcal{PS}}$, since in this case Popp's measure coincide with the Lebesgue measure in our coordinates.

3.4.3 Differentiability properties: contact case

Let us come back to the differentiability of the map

$$q \mapsto \mathcal{L}(\widehat{B}_q). \tag{3.40}$$

A smooth family of sub-Riemannian structures is represented by a smooth familly of skew symmetric matrices $L(q)$ (see Section 3.4.1). Recall that, for a smooth family of skew-symmetric matrices that depend on a n-dimensional parameter, the eigenvalue functions $q \mapsto b_i(q)$ exists and are Lipschitz continuous with respect to q (see [72]).

Thus, if we denote with $V(q)$ the volume of the nilpotent unit ball corresponding to frequences $b_1(q), \ldots, b_\ell(q)$, formula (3.39) can be rewritten as

$$V(q) = \int_0^{a(q)} G(q,s)ds, \tag{3.41}$$

where

$$a(q) := \frac{\pi}{\max b_i(q)},$$

$$G(q,s) := \frac{1}{s^{2\ell+2}B^2} \sum_{i=1}^{\ell} \left(\prod_{j\neq i} \sin^2(b_j(q)s) \right) \sin(b_i(q)s)(b_i(q)s\cos(b_i(q)s) - \sin(b_i(q)s)). \tag{3.42}$$

Notice that we have dropped the constant C_ℓ that appear in (3.39) since it does not affect differentiability of the volume.

Remark 3.23. Since the family of sub-Riemannian structures $q \mapsto L(q)$ is smooth, the exponential map smoothly depends on the point q. As a consequence the integrand $G(q,s)$, being the Jacobian of the exponential map, is a smooth function of its variables.

In addition, altough $q \to b_i(q)$ is only Lipschitz, it is easy to see that the function $a(q) = \max b_i(q)$ is semiconvex with respect to q (see also [43]). In particular $a(q)$ admits second derivative almost everywhere.

If for all q all $b_i(q)$ are different (i.e. there are no resonance points), then the eigenvalue functions can be chosen in a smooth way. As a consequence the volume $V(q)$ is smooth, since all functions that appear in (3.41) are smooth. This argument provides a proof of Proposition 3.6 in the case of contact structures.

On the other hand, we prove that, along a curve where $a(q)$ is not smooth, i.e. when the two bigger frequences cross, $V(q)$ is no longer smooth at that point.

Proof of Theorem 3.5. In this proof by a resonance point we mean a point q_0 where the two (or more) biggest frequences of $L(q_0)$ coincide. More precisely, if we order the frequences as

$$b_1(q) \geq b_2(q) \geq \ldots \geq b_\ell(q),$$

a resonance point is defined by the relation $b_1(q_0) = b_2(q_0)$. As we noticed, V is smooth at non-resonance points.

We divide the proof into two steps: first we prove that $V \in \mathcal{C}^3$ and then we show that, in general, it is not smooth (but $\mathcal{C}^4$ on smooth curves).

(i). We have to show that V is $\mathcal{C}^3$ in a neighborhood of every resonance point q_0. First split the volume as follows

$$V(q) = \int_0^{a(q_0)} G(q,s)ds + \int_{a(q_0)}^{a(q)} G(q,s)ds \tag{3.43}$$

The first term in the sum is smooth with respect to q, since it is the integral of a smooth function on a domain of integration that does not depend on q. We are then reduced to the regularity of the function

$$W(q) := \int_{a(q_0)}^{a(q)} G(q,s)ds \tag{3.44}$$

We have the following key estimate

Lemma 3.24. *Let $q_0 \in M$ be a resonance point. Then, for any neighborhood of q_0, there exists $C > 0$ such that*

$$\left| \int_{a(q_0)}^{a(q)} G(q,s)ds \right| \leq C|q - q_0|^4 \tag{3.45}$$

Proof. It is sufficient to prove that every derivative up to second order of G vanish at $(q_0, a(q_0))$. Indeed, being G smooth, computing its Taylor polynomial at $(q_0, a(q_0))$ only terms with order greater or equal than three appear (both in $q - q_0$ and $s - a(q_0)$). Thus, integrating with respect to s and using that $|a(q) - a(q_0)| = O(|q - q_0|)$, we have the desired result.

From the explicit formula (3.42) it is easy to see that $G(q, a(q)) \equiv 0$ for every $q \in M$. In particular $G(q_0, a(q_0)) = 0$. Moreover, since at a resonance point q_0 at least the two bigger eigenvalues coincide, say $b_1(q_0) = b_2(q_0) = \beta$, we have $a(q_0) = \pi/\beta$ and in a neighborhood of $(q, s) = (q_0, a(q_0))$

$$\sin^2(b_1(q)\frac{\pi}{a(q)}) \sin(b_2(q)\frac{\pi}{a(q)}) = O(|b_1(q) - b_2(q)|^3) = O(|q - q_0|^3). \tag{3.46}$$

due to the Lipschitz property of $b_j(q)$, for $j = 1, 2$.

From (3.42) and (3.46) one can easily get that every derivative of G up to second order (in both variables q and s) vanish at $(q_0, a(q_0))$.

□

To show that $V \in \mathcal{C}^3$ we compute the first three derivatives of W at a non resonant points q, and we show that, when q tend to a resonance point q_0, they tends to zero. We then conclude the continuity of the first three derivatives by Lemma 3.24.

In the following, for simplicity of the notation, we will denote by $\frac{\partial}{\partial q}$ the partial derivative with respect to some coordinate function on M. For instance $\frac{\partial^2 W}{\partial q^2}$ denote some second order derivative $\frac{\partial^2 W}{\partial x_i \partial x_j}$.

At non resonance points q we have

$$\begin{aligned}\frac{\partial W}{\partial q}(q) &= \underbrace{G(q,a(q))}_{=0}\frac{\partial a}{\partial q}(q) + \int_{a(q_0)}^{a(q)} \frac{\partial G}{\partial q}(q,s)ds \\ &= \int_{a(q_0)}^{a(q)} \frac{\partial G}{\partial q}(q,s)ds\end{aligned}$$

which tends to zero for $q \to q_0$. Using Lemma 3.24 we conclude $V \in \mathcal{C}^1$. Next let us compute

$$\frac{\partial^2 W}{\partial q^2}(q) = \frac{\partial G}{\partial q}(q,a(q))\frac{\partial a}{\partial q}(q) + \int_{a(q_0)}^{a(q)} \frac{\partial^2 G}{\partial q^2}(q,s)ds \tag{3.47}$$

Since $\frac{\partial G}{\partial q}(q_0,a(q_0)) = 0$ (see proof of Lemma 3.24) and $\frac{\partial a}{\partial q}$ is bounded by Lipschitz continuity of the maximum eigenvalue, it follows that $\frac{\partial^2 W}{\partial q^2}$ tends to zero as $q \to q_0$, and using again Lemma 3.24, we have that $V \in \mathcal{C}^2$. In analogous way one can compute the third derivative

$$\begin{aligned}\frac{\partial^3 W}{\partial q^3}(q) = \frac{\partial^2 G}{\partial q \partial s}(q,a(q))\left(\frac{\partial a}{\partial q}(q)\right)^2 + 2\frac{\partial^2 G}{\partial q^2}(q,a(q))\frac{\partial a}{\partial q}(q) + \\ + \frac{\partial G}{\partial q}(q,a(q))\frac{\partial^2 a}{\partial q^2}(q) + \int_{a(q_0)}^{a(q)} \frac{\partial^3 G}{\partial q^3}(q,s)ds\end{aligned} \tag{3.48}$$

Using again that every second derivative of G vanish at $(q_0,a(q_0))$ and that $\frac{\partial a}{\partial q}$ is bounded, it remains to check that $\frac{\partial G}{\partial q}(q,a(q))\frac{\partial^2 a}{\partial q^2}(q)$ tends to zero as $q \to q_0$. From (3.46) one can see that $\frac{\partial G}{\partial q} = O(|b_1(q) - b_2(q)|^2)$. Hence it is sufficient to prove that $\frac{\partial^2 a}{\partial q^2} = O(1/|b_1(q) - b_2(q)|)$, which is a consequence of the following lemma.

Notice that, for every skew-symmetric matrix A, with eigenvalues $\pm i\lambda_j$, with $j = 1,\dots,n$, the matrix iA is an Hermitian matrix which has eigenvalues $\pm\lambda_j$, $j = 1,\dots,n$.

Lemma 3.25. *Let A, B be two $n \times n$ Hermitian matrices and assume that for every t the matrix $A + tB$ has a simple eigenvalue $\lambda_j(t)$. Then the the following equation is satisfied*

$$\ddot{\lambda}_j = 2\sum_{k\neq j} \frac{|\langle Bx_j, x_k\rangle|^2}{\lambda_j - \lambda_k} \tag{3.49}$$

where $\{x_k(t)\}_{k=1,\dots,n}$ is an orthonormal basis of eigenvectors and $x_j(t)$ is the eigenvector associated to $\lambda_j(t)$.

Proof. In this proof we endow $\mathbb{C}^n$ with the standard scalar product $\langle z,w\rangle = \sum_{k=1}^n z_k\overline{w_k}$. Since $\lambda_j(t)$ is simple for every t, both $\lambda_j(t)$ and the associated eigenvector $x_j(t)$ can be choosen smoothly with respect to t. By definition

$$(A + tB)x_j(t) = \lambda_j(t)x_j(t), \qquad |x_j(t)| = 1.$$

Then we compute the derivative with respect to t of both sides

$$(A+tB)\dot{x}_j(t) + Bx_j(t) = \dot{\lambda}_j(t)x_j(t) + \lambda_j(t)\dot{x}_j(t), \tag{3.50}$$

and computing the scalar product with $x_j(t)$ we get

$$\dot{\lambda}_j(t) = \langle Bx_j(t), x_j(t)\rangle, \qquad \text{hence} \qquad \ddot{\lambda}_j(t) = 2\,\mathrm{Re}\,\langle Bx_j(t), \dot{x}_j(t)\rangle. \tag{3.51}$$

using that $A+tB$ is Hermitian and $\langle \dot{x}_k(t), x_k(t)\rangle = 0$. On the other hand, the scalar product of (3.50) with x_k, with $k \neq j$ gives

$$\langle \dot{x}_j, x_k\rangle = \frac{\langle Bx_j, x_k\rangle}{\lambda_j - \lambda_k}.$$

Substituting $\dot{x}_j = \sum_{k=1}^n \langle \dot{x}_j, x_k\rangle\, x_k$ in (3.51) we have (3.49). □

Then the third derivative goes to zero for $q \to q_0$, and using again Lemma 3.24 we conclude that $V \in \mathcal{C}^3$.

(ii). Now we study the restriction of the map (3.40) along any smooth curve on the manifold, and we see that, due to a simmetry property, V is $\mathcal{C}^4$ on every curve but in general is not $\mathcal{C}^5$. Notice that (3.45) gives no information on the fourth derivative but the fact that it is bounded. Indeed it happens that it is continuous on every smooth curve but its value depend on the curve we choose.

From now on, we are left to consider a smooth *one-parametric* family of sub-Riemannian structure, i.e. of skew symmetric matrices.

Remark 3.26. An analytic family of skew-simmetric matrices $t \mapsto L(t)$ depending on one parameter, can be simultaneously diagonalized (see again [72]), in the sense that there exists an analytic (with respect to t) family of orthogonal changes of coordinates and analytic functions $b_i(t) > 0$ such that

$$L = \begin{pmatrix} B_1(t) & & \\ & \ddots & \\ & & B_\ell(t) \end{pmatrix}, \qquad \text{where} \qquad B_i(t) := \begin{pmatrix} 0 & -b_i(t) \\ b_i(t) & 0 \end{pmatrix}. \tag{3.52}$$

In the case of a $\mathcal{C}^\infty$ family $t \mapsto L(t)$, we can apply the previuos result to the Taylor polynomial of this family. As a consequence we get an approximate diagonalization for $L(t)$, i.e. for every $N > 0$ there exists a smooth family of orthogonal changes of coordinates and smooth functions $b_i(t) > 0$ such that every entry out of the diagonal in $L(t)$ is $o(t^N)$. Namely

$$L(t) = \begin{pmatrix} B_1(t) & & o(t^N) \\ & \ddots & \\ o(t^N) & & B_\ell(t) \end{pmatrix}, \qquad \text{where} \qquad B_i(t) := \begin{pmatrix} o(t^N) & -b_i(t) \\ b_i(t) & o(t^N) \end{pmatrix}. \tag{3.53}$$

Since we are interested, in the study the $\mathcal{C}^k$ regularity of (3.40), for k finite, in what follows we can ignore higher order terms and assume that $L(t)$ can be diagonalized as in the analytic case (3.52).

From the general analysis we know that V is $\mathcal{C}^3$. To prove that $t \mapsto V(t)$ is actually $\mathcal{C}^4$ we discuss first the easiest case $\ell = 2$ (i.e. the contact (4,5) case) and then generalize to any ℓ.

(i) Case $\ell = 2$. To start, assume that $b_1(t)$, $b_2(t)$ cross transversally at $t = 0$. This means that for the volume $V(t)$ we have the expression

$$V(t) = \begin{cases} \displaystyle\int_0^{\frac{\pi}{b_1(t)}} G(t,s)\,ds, & \text{if } t > 0, \\ \displaystyle\int_0^{\frac{\pi}{b_2(t)}} G(t,s)\,ds, & \text{if } t < 0, \end{cases} \qquad \text{and} \qquad \begin{array}{c} b_1(0) = b_2(0) \\ b_1'(0) \neq b_2'(0) \end{array} \tag{3.54}$$

Since the regularity of the volume does not depend on the value $b_1(0) = b_2(0)$, we can make the additional assumption

$$b_i(t) = 1 + t\,c_i(t), \qquad i = 1,2$$

for some suitable functions $c_1(t), c_2(t)$. Notice that $a'(t)$ is discontinuous at $t = 0$ and the left and right limits are

$$a'_+ := \lim_{t\to 0+} a'(t) = -\pi c_1(0), \qquad a'_- := \lim_{t\to 0-} a'(t) = -\pi c_2(0),$$

From the explicit expression of G it is easy to compute that

$$\frac{\partial^3 G}{\partial t^3}(0,a(0)) = \frac{6}{\pi^2} c_1 c_2 (c_1 + c_2),$$

$$\frac{\partial^3 G}{\partial t^2 \partial s}(0,a(0)) = \frac{2}{\pi^3}(c_1^2 + 4c_1c_2 + c_2^2), \qquad \frac{\partial^3 G}{\partial t \partial s^2}(0,a(0)) = \frac{6}{\pi^4}(c_1 + c_2),$$

where we denote for simplicity $c_i := c_i(0)$.

Let us compute $\frac{\partial^4 W}{\partial t^4}$. To this purpose let us differentiate with respect to t formula (3.48) (where q is replaced by t). Using the fact that all second derivatives of G vanish at $(t,s) = (0,a(0))$ (see the proof of Lemma 3.24) we have that the 4-th derivative of W at $t = 0$ is computed as follows

$$\lim_{t\to 0+} W^{(4)}(t) = 3\frac{\partial^3 G}{\partial t^3} a'_+ + 3\frac{\partial^3 G}{\partial t^2 \partial s}(a'_+)^2 + \frac{\partial^3 G}{\partial t \partial s^2}(a'_+)^3,$$
$$\lim_{t\to 0-} W^{(4)}(t) = 3\frac{\partial^3 G}{\partial t^3} a'_- + 3\frac{\partial^3 G}{\partial t^2 \partial s}(a'_-)^2 + \frac{\partial^3 G}{\partial t \partial s^2}(a'_-)^3,$$

where W is defined in (4.46). It is easily checked that $W^{(4)}$ is continuous (but does not vanish!). Indeed we have

$$\lim_{t\to 0+} W^{(4)}(t) = \lim_{t\to 0-} W^{(4)}(t) = -\frac{12}{\pi} c_1^2 c_2^2.$$

The same argument produce an example that, in general, $V(t)$ is not $\mathcal{C}^5$. Assuming

$$b_i(t) = 1 + t\,c_i, \qquad c_1 \neq c_2 \quad \text{constant}, \qquad i = 1,2,$$

a longer computation, but similar to the one above, shows that

$$\lim_{t\to 0+} W^{(5)}(t) = -\frac{2}{\pi}\, c_1^3(13c_1^2 - 29c_1c_2 + 22c_2^2)$$
$$\lim_{t\to 0-} W^{(5)}(t) = -\frac{2}{\pi}\, c_2^3(13c_2^2 - 29c_1c_2 + 22c_1^2)$$

and the 5-th derivatives do not coincide.

Remark 3.27. Notice that the assumption of transversality on $b_1(t)$ and $b_2(t)$ at $t = 0$ is not restrictive. Indeed if $b_1(t) - b_2(t) = O(t^k)$ for some $k > 1$, then from the proof of Lemma 3.24 (see in particular (3.46)) it follows that at least $2k+1$ derivatives of G vanish at $(t, s) = (0, a(0))$, increasing the regularity of W.

(ii) General case. We reduce to case (i).

We can write $G(t, s) = \sum_{i=1}^{\ell} G_i(t, s)$ and $V(t) = \sum_{i=1}^{\ell} V_i(t)$ where we set

$$G_i(t,s) := \frac{1}{s^{2\ell+2}} \left(\prod_{j\neq i} \sin^2(b_j(t)s)\right) \sin(b_i(t)s)(b_i(t)s\cos(b_i(t)s) - \sin(b_i(t)s)), \tag{3.55}$$

$$V_i(t) := \int_0^{a(t)} G_i(t,s)ds, \qquad i = 1, \dots, \ell.$$

Assume that b_1, b_2 are the bigger frequences and that they cross at $t = 0$, i.e.

$$b_i(t) < b_2(t) < b_1(t), \qquad \forall\, t < 0, \quad \forall\, i = 3, \dots, n.$$
$$b_i(t) < b_1(t) < b_2(t), \qquad \forall\, t > 0, \quad \forall\, i = 3, \dots, n.$$

From the explicit expression above it is easy to recognise that for G_1 and G_2 we can repeat the same argument used in (i). Indeed if we denote with $\widetilde{G}(t, s)$ the integrand of the $(4, 5)$ case we can write G_1+G_2 as the product of a smooth function and $\widetilde{G}$

$$G_1(t,s) + G_2(t,s) = \left(\frac{1}{s^{2\ell-4}} \prod_{j=3}^{\ell} \sin^2(b_j(t)s)\right) \widetilde{G}(t,s),$$

which implies that that $V_1 + V_2$ is a $\mathcal{C}^4$ function.

Moreover it is also easy to see that $V_3, \dots, V_n$ are $\mathcal{C}^4$. Indeed from the fact that $b_1(t)$ and $b_2(t)$ both appear in $\sin^2$ terms in $G_i(t, s)$ for $i > 3$, it follows that in this case

$$G_i(t,a(t)) \equiv \frac{\partial G_i}{\partial t}(t,a(t)) \equiv 0, \quad \frac{\partial^2 G_i}{\partial t^2}(0,a(0)) = 0, \quad \frac{\partial^3 G_i}{\partial t^3}(0,a(0)) = 0, \qquad i = 3, \dots, n,$$

and we can apply the same argument used in (i) to the function $V_i'(t) = \int_0^{a(t)} \frac{\partial G_i}{\partial t}(t, s)ds$, for $i = 3, \dots, n$, showing that it is $\mathcal{C}^3$, that means $V_i \in \mathcal{C}^4$. □

Remark 3.28. As we said the value of the 4-th derivative depend on the curve we chose, hence we cannot conclude that $V \in \mathcal{C}^4$ in general. Moreover, we explicitly proved that $V \notin \mathcal{C}^\infty$ since in general is not $\mathcal{C}^5$, even when restricted on smooth curves.

Moreover from the proof it also follows that, if more than two frequences coincide at some point (for instance if we get a triple eigenvalue), we have a higher order regularity for every V_i, and the regularity of V increases.

3.4.4 Extension to the quasi-contact case

Recall that in the quasi contact case the dimension of the distribution is odd and the kernel of the contact form is one dimensional. Hence, applying the same argument used in Section 3.4.1, we can always normalize the matrix L in the following form:

$$L = \begin{pmatrix} B_1 & & & \\ & \ddots & & \\ & & B_\ell & \\ & & & 0 \end{pmatrix}, \qquad B_i := \begin{pmatrix} 0 & -b_i \\ b_i & 0 \end{pmatrix}, \quad b_i > 0.$$

In other words we can select a basis $\{X_1, \dots, X_\ell, Y_1, \dots, Y_\ell, K, Z\}$ such that

$$\begin{cases} \Delta = \operatorname{span}\{X_1, \dots, X_\ell, Y_1, \dots, Y_\ell, K\}, & \\ [X_i, Y_i] = -b_i Z, & i = 1, \dots, \ell \\ [X_i, Y_j] = 0, & i \neq j \\ [X_i, K] = [Y_i, K] = 0, & i = 1, \dots, \ell \\ [X_i, Z] = [Y_i, Z] = 0, & i = 1, \dots, \ell \end{cases} \tag{3.56}$$

where the new vector field K is in the kernel of the bracket mapping, i.e. it commutes with all others elements. Since abnormal extremals are never optimal in quasi contact case (see Remark 1.11), we are reduced to compute the exponential map to find geodesics. With analogous computations of contact case we get the following expression for the exponential map from the origin

$$\begin{aligned} x_i(t) &= \frac{r_i}{b_i w}(\cos(b_i w t + \theta_i) - \cos\theta_i), \\ y_i(t) &= \frac{r_i}{b_i w}(\sin(b_i w t + \theta_i) - \sin\theta_i), \\ x_{2\ell+1}(t) &= u_{2\ell+1} t, \\ z(t) &= \frac{1}{2w^2}(|r|^2 w t - \sum_i \frac{r_i^2}{b_i} \sin b_i w t). \end{aligned} \tag{3.57}$$

From (3.57) it is easily seen that the jacobian of exponential map has exactly the same expression as in contact case (3.33). Since zero is always an eigenvalue of L, but is never the maximum one, we can proceed as in the contact case and all the regularity results extend to this case.

3.5 Extension to general corank 1 case

We start this section with the proof of Theorem 3.4, after that we extend the result to the general corank 1 case.

Proof of Theorem 3.4. Let $\mathbf{S}$ be a sub-Riemannian structure such that $\dim M \leq 5$.

(i). If $\mathcal{G}(\mathbf{S}) \neq (4,5)$. From Theorem 3.12 we know that at every point $q \in M$, the nilpotent approximation $\widehat{\mathbf{S}}_q$ has a unique normal form, hence by Corollary 3.14 all nilpotent approximations are isometric. From this property it easily follows that $f_{\mathcal{P}\mathcal{S}}$, the Popp volume of the unit ball, is constant (recall that Popp measure is intrinsic for the sub-Riemannian structure). This also implies that for a smooth volume μ the density $f_{\mu\mathcal{S}}$ is smooth.

(ii). If $\mathcal{G}(\mathbf{S}) = (4,5)$ by Theorem 3.12 it is sufficient to consider the case when the family of nilpotent structure has the normal form (3.7), where $\alpha = \alpha(q)$ depends on the point. Notice that the formula (3.39) for the volume of the unit ball is still valid, where now $b_1 = 1$ and $b_2 = |\alpha|$.

Theorem 3.5 proves that the density is $\mathcal{C}^3$ at points where $|\alpha| > 0$, i.e. in the contact case. We are then reduced to the study of the volume near a point where the eigenvalue α crosses zero. In particular we show that the volume is smooth at these points. Since the eigenvalue α is approaching zero, it is not restrictive to assume $|\alpha| < 1$. Let us consider then the function defined on the interval $(-1,1)$

$$W(\alpha) = \int_0^{\pi} \frac{1}{\alpha^2 s^6} \left(\sin^2(\alpha s)\sin s(s\cos s - \sin s) + \sin^2 s \sin(\alpha s)(\alpha s\cos(\alpha s) - \sin(\alpha s))\right) ds. \tag{3.58}$$

Note that $V(\alpha) = W(|\alpha|)$, where $V(\alpha)$ denotes the volume of the nilpotent ball relative to frequences 1 and $\alpha < 1$. It is easy to see that both

$$\frac{\sin^2 \alpha s}{\alpha^2} \qquad \text{and} \qquad \frac{1}{\alpha^2}\sin(\alpha s)(\alpha s\cos(\alpha s) - \sin(\alpha s)) \tag{3.59}$$

are smooth as functions of α (also at $\alpha = 0$). Hence W is a smooth function (for $\alpha \in (-1,1)$). Moreover it is easy to see that W is an even smooth function of α. Thus W it is smooth also as a function of $|\alpha|$, which completes the proof.

The same argument applies to prove that the $\mathcal{C}^3$ regularity holds in the general corank 1 case. Indeed, from (3.42) and the fact that (5.26) are smooth functions at $\alpha = 0$, it follows that the integrand $G(q,s)$ is smooth as soon as one of the eigenvalue $b_i(q)$ is different from zero (recall that $b_i \geq 0$ by definition). Since the structure is regular (i.e. the dimension of the flag do not depend on the point) and bracket generating, we have that $\max_i b_i(q) > 0$ for every q, hence the conclusion. □

CHAPTER 4

Nilpotent corank 2 sub-Riemannian metrics

In this chapter we study nilpotent 2-step, corank 2 sub-Riemannian metrics that are nilpotent approximations of general sub-Riemannian metrics. We exhibit optimal syntheses for these problems. It turns out that in general the cut time is not equal to the first conjugate time but has a simple explicit expression. As a byproduct of this study we get some smoothness properties of the spherical Hausdorff measure in the case of a generic 6 dimensional, 2-step corank 2 sub-Riemannian metric.

4.1 Introduction

In this chapter, we assume that the structure is 2-step bracket generating, i.e.

$$T_qM = \Delta_q + [\Delta, \Delta]_q, \quad \text{for every } q \in M.$$

and we quote a 2-step sub-Riemannian metric by its rank and its dimension, i.e. with the pair (k, n). The quantity $m = n - k$ is called the *corank* of the structure.

It follows from Theorem 3.9 that the Hausdorff dimension of M, as a metric space, is $Q = k + 2m > n$. In this paper we focus on the case $(k, k+2)$.

Nilpotent approximation

For our convenience, we give here another construction of the nilpotent approximation, that makes sense for any sub-Riemannian metric, but it coincides with the standard one (see Chapter 1) in the 2-step bracket generating case only.

The tensor bilinear mapping

$$[\cdot, \cdot] : \Delta_q \times \Delta_q \to T_qM/\Delta_q, \tag{4.1}$$

is skew symmetric. Then for every $Z^* \in (T_qM/\Delta_q)^*$, we have

$$Z^*([X, Y] + \Delta_q) = \langle A_{Z^*}(X), Y\rangle_{\mathbf{g}},$$

for some $\mathbf{g}$-skew symmetric endomorphism A_{Z^*} of Δ_q.

Remark 4.1 (Notation). We denote by $\mathcal{L}_q$ the m-dimensional space of skew symmetric endomorphisms of Δ_q obtained by taking the union of all the A_{Z^*} at q. This notation is used in the Section 4.4.

The space $L_q = \Delta_q \oplus T_qM/\Delta_q$ is endowed with the structure of a 2-step nilpotent Lie-algebra by setting

$$[(V_1, W_1), (V_2, W_2)] = (0, [V_1, V_2] + \Delta_q).$$

The associated simply connected nilpotent Lie group is denoted by G_q and the exponential mapping Exp: $L_q \to G_q$ is one to one and onto. By translation, the metric $\mathbf{g}_q$ over Δ_q allows to define a left-invariant sub-Riemannian metric over G_q, called the nilpotent approximation of $(M, \Delta, \mathbf{g})$ at q.

Any k dimensional vector sub-space $\mathcal{V}_q$ of T_qM, transversal to Δ_q allows to identify L_q and G_q to $T_qM \simeq \Delta_q \oplus T_qM/\Delta_q$.

Fix $q_0 \in M$. We can chose coordinates x in Δ_{q_0} such that the metric $\mathbf{g}_{q_0}$ is the standard Euclidean metric, and for any linear coordinate system y in $\mathcal{V}_{q_0}$, there are skew symmetric matrices $L_1, \dots, L_m \in so(k)$ such that the mapping (4.1) writes

$$[X, Y] + \Delta_{q_0} = \begin{pmatrix} X'L_1Y \\ \vdots \\ X'L_mY \end{pmatrix}.$$

where X' denotes the transpose of the vector X. Then the nilpotent approximation written in control form is

$$\begin{cases} \dot{x}_i = u_i, \qquad i = 1, \dots, k, \\ \dot{y}_1 = \frac{1}{2}x'L_1u, \\ \quad\vdots \\ \dot{y}_m = \frac{1}{2}x'L_mu. \end{cases} \tag{4.2}$$

Proposition 4.2. *The distribution is 2-step bracket generating if and only if the endomorphisms of Δ_q, L_i, $i = 1, \dots, m$ (respectively the matrices L_i when coordinates y in $\mathcal{V}_q$ are chosen) are independent.*

In the 2-step bracket generating case these linear coordinates y in T_qM/Δ_q may be chosen in such a way that the endomorphisms L_i, $i = 1, \dots, m$ are orthonormal with respect to the Hilbert-Schmidt norm $\langle L_i, L_j\rangle = \frac{1}{k}\mathrm{trace}(L_i'L_j)$. This choice defines a canonical Euclidean structure in T_qM/Δ_q and a corresponding volume in T_qM/Δ_q. Then using the Euclidean structure over Δ_q we get a canonical Euclidean structure over $\Delta_q \oplus T_qM/\Delta_q$. The choice of the vector subspace $\mathcal{V}_q$ induces an Euclidean structure on T_qM which depends on the choice of $\mathcal{V}_q$, but the associated volume on T_qM is independent on this choice.

This volume form on M coincide with the Popp measure, that is a smooth volume form.

Statement of the results

The main purpose of this chapter is to build the optimal synthesis for $(k, k+2)$ nilpotent sub-Riemannian metrics, i.e. the set of all trajectories starting from the identity of the group and realizing the minimum of the distance, with a precise description of their cut time.

Remark 4.3 (Notation). In the case of our nilpotent approximations, covectors in T_q^*M can be identified with vectors in T_qM via the Euclidean structure of T_qM given by the choice of $\mathcal{V}_q$. In our coordinates (x, y), these covectors/vectors are typically denoted by (u_0, r).

For nilpotent $(k, k+1)$ sub-Riemannian metrics that are nilpotent approximations of general sub-Riemannian metrics, the control systems can be written as follows (see Section 3.4)

$$\begin{cases} \dot{x}_i = u_i, & i = 1, \dots, k, \\ \dot{y} = \frac{1}{2} x' L u, & L \text{ skew symmetric}, \end{cases} \tag{C1}$$

Denote by $\sigma(L)$ the set of all moduli of eigenvalues of the matrix L. In Chapter 3 it is proved the following result:

Theorem 4.4. *Arclength geodesics of system* (C1)*, starting from the origin, are parametrized by an initial covector* $\lambda_0 = (u_0, r) \in S^{k-1} \times \mathbb{R}$*, and they are optimal until time*

$$t_{cut}(\lambda_0) = \frac{2\pi}{|r| \max \sigma(L)},$$

with the understanding $t_{cut} = +\infty$ *if* $r = 0$*. Moreover* $t_{cut}(\lambda_0) = t_{con}(\lambda_0)$.

The proof of this result is based on the fact that geodesics can be expressed in terms of usual trigonometric functions and, thanks to a certain monotonicity property, the cut locus can be explicitly computed and is exactly equal to the conjugate locus.

Optimal synthesis for the nilpotent $(k, k+2)$ case

The main result of this paper is the optimal synthesis in the case of a nilpotent approximation in the $(k, k+2)$ case. In this case the control system can be written in coordinates $q = (x_1, \dots, x_k, y_1, y_2)$ as

$$\begin{cases} \dot{x}_i = u_i, & i = 1, \dots, k, \\ \dot{y}_1 = \frac{1}{2} x' L_1 u, \\ \dot{y}_2 = \frac{1}{2} x' L_2 u, & L_1, L_2 \text{ skew symmetric}. \end{cases} \tag{C2}$$

Set $r_1 = |r| \cos\theta$, $r_2 = |r| \sin\theta$, and $L_\theta = \cos(\theta) L_1 + \sin(\theta) L_2$.

Theorem 4.5. *Arclength geodesics of system* (C2)*, starting from the origin, are parametrized by an initial covector* $\lambda_0 = (u_0, r) \in S^{m-1} \times \mathbb{R}^2$*, and they are optimal until time*

$$t_{cut}(\lambda_0) = \frac{2\pi}{\max \sigma(r_1 L_1 + r_2 L_2)} = \frac{2\pi}{|r| \max \sigma(L_\theta)},$$

with the understanding $t_{cut} = +\infty$ *if* $r = 0$*. Moreover, in general,* $t_{cut}(\lambda_0) \neq t_{con}(\lambda_0)$.

The reason why the corank 2 case is more difficult than the corank 1 case is precisely the fact that the cut locus is not equal to the conjugate locus. (The latter we are not able to compute explicitly.)

Explicit expression of geodesics for this optimal synthesis are given in Section 4.2.

The nilpotent $(4,6)$ case

In the nilpotent $(4,6)$ case our first result is the following:

Theorem 4.6. *The following properties are equivalent:*

(P1) The first conjugate locus is equal to the cut locus.

(P2) The linear coordinates y in T_qM/Δ_q can be chosen in such a way that the pair (L_1, L_2) of 4×4 skew symmetric matrices belongs to the set $(\mathcal{Q} \cup \widehat{\mathcal{Q}})^2$.

Here $\mathcal{Q}$ (resp. $\widehat{\mathcal{Q}}$) denotes the set of pure quaternions (resp. pure skew quaternions), see Appendix 4.4.1.

Our second result is a continuation of the results contained in Chapter 3 for corank 1 structures (see Theorem 3.5). Here we show the following result

Theorem 4.7. *For a generic $(4,6)$ sub-Riemannian metric*[1]*, the Radon-Nykodym derivative of the spherical Hausdorff measure with respect to the Popp measure is $\mathcal{C}^1$.*

In the previous chapter it is shown that the Radon-Nikodym derivative of the spherical Hausdorff measure with respect to the Popp measure is inversely proportional (as a function of q) to the volume of the unit sub-Riemannian ball of the nilpotent approximation at q. Then Theorem 4.7 is a byproduct of the optimal synthesis given here.

Note that in the corank 1 case, the higher differentiability of the Radon-Nikodym derivative is due to the fact that the conjugate locus is equal to the cut locus, which is not the case here.

Due to the complexity of the computations even in this low dimensional case, it is not easy to determine the real degree of differentiability of Hausdorff measure. This is still an interesting open question.

4.1.1 Organization of the chapter

Section 4.2 is devoted to the construction of the optimal synthesis for $(m, m+2)$ nilpotent sub-Riemannian metrics and, as a consequence, to the proof of Theorem 4.5. In Sections 4.2.1 and 4.2.2 we compute the exponential map. In Section 4.2.3 we prove that geodesics are optimal up to t_{cut}. Finally in Section 4.2.4 we show that the cut time does not coincide, in general, with the first conjugate time. In Section 4.3 we give the proofs of Theorems 4.6 and 4.7.

In Section 4.4 we recall basic facts about quaternions, we prove a technical Lemma, and applying an Abraham's transversality theorem, we prove that, generically, for the $(4,6)$ case, a certain "bad set" is made of isolated points, which permits to conclude about the differentiability of the Radon-Nikodym derivative (Theorem 4.7).

[1] which means for an open and dense subset of all $(4,6)$ sub-Riemannian metrics, endowed with the Whitney topology.

4.2 Exponential map and synthesis

4.2.1 Hamiltonian equations in the (k, n) case

The purpose of this section is to compute the exponential map, i.e. the set of all geodesics, parametrized by length, starting form the origin of the control system (4.2), i.e. the system

$$\begin{cases} \dot{x}_i = u_i, & i = 1, \dots, k, \\ \dot{y}_h = \frac{1}{2} x' L_h u, & h = 1, \dots, m. \end{cases} \tag{4.3}$$

Let $L_h = (b_{ij}^h)$, for $h = 1, \dots, m$. Then the control system can be written in the form $\dot{q} = \sum_{i=1}^k u_i X_i(q)$ where $q = (x, y)$ and

$$X_i = \partial_{x_i} + \frac{1}{2} \sum_{j,h} b_{ij}^h x_j \partial_{y_h}, \qquad i = 1, \dots, k.$$

Setting $Y_h = \partial_{y_h}$, for $h = 1, \dots, m$, the commutation relations are

$$[X_i, X_j] = \sum_{h=1}^m b_{ij}^h Y_h, \qquad i, j = 1, \dots, k, \tag{4.4}$$

$$[X_i, Y_j] = [Y_j, Y_h] = 0, \qquad i = 1, \dots, k, \qquad j, h = 1, \dots, m. \tag{4.5}$$

Define the functions on T^*M, that are linear on fibers,

$$u_i(\lambda, q) = \langle \lambda, X_i(q) \rangle, \qquad i = 1, \dots, k, \tag{4.6}$$

$$r_h(\lambda, q) = \langle \lambda, Y_h(q) \rangle, \qquad h = 1, \dots, m. \tag{4.7}$$

These functions can be treated as coordinates on the fiber of T^*M to solve the Hamiltonian system given by the Pontryagin Maximum Principle, see Section 1.2. This Hamiltonian system is associated with the Hamiltonian

$$H(\lambda, q) = \frac{1}{2} \sum_{i=1}^k \langle \lambda, X_i(q) \rangle^2 = \frac{1}{2} \sum_{i=1}^k u_i^2(\lambda, q), \qquad \lambda \in T_q^* M. \tag{4.8}$$

Remark 4.8. The geodesics parametrized by length correspond to the level set $\{H = 1/2\}$. Notice that, for systems of type $\dot{q} = \sum_{i=1}^k u_i X_i(q)$, with fixed initial and final points, the problem of finding length-parametrized curves minimizing the length, is equivalent to the problem of minimizing time with the constraint $\{\|u\| \leq 1\}$.

For a function $a \in C^\infty(T^*M)$ we have that, along the sub-Riemannian flow

$$\dot{a} = \{a, H\} = \sum_{i=1}^k \{a, u_i\} u_i, \tag{4.9}$$

where $\{a, b\}$ denotes the Poisson bracket of two functions in T^*M. The following Lemma gives a way of computing the covector $\lambda(t)$, solution of the Hamiltonian system associated with (4.8) in the coordinates (u, r).

Lemma 4.9. *If $u(t)$ and $r(t)$ are solution of (4.9) corresponding to level set $\{H = 1/2\}$, then they satisfy*

$$\begin{cases} \dot{u}(t) = (r_1 L_1 + \ldots + r_m L_m)u(t), \qquad u(0) = u_0, \qquad \|u_0\| = 1, \\ \dot{r}(t) = 0. \end{cases}$$

Proof. Remind that, if $a_i(\lambda, q) = \langle \lambda, Z_i(q) \rangle$, for some vector fields Z_i, $i = 1, 2$, then

$$\{a_1, a_2\} = \langle \lambda, [Z_1, Z_2] \rangle .$$

Applying (4.9) for $a = r_h$ and using (4.5) we get

$$\dot{r}_h = \sum_{i=1}^{k} \{r_h, u_i\} u_i = 0 \quad \Rightarrow \quad r_h = \text{const}.$$

Similarly, using (4.4), one find

$$\dot{u}_i = \sum_{i=1}^{k} \{u_i, u_j\} u_j = \sum_{i=1}^{k} b_{ij}^k r_k u_j.$$

□

Remark 4.10. In the following geodesics are parametrized by the initial covector $\lambda(0) = (p(0), r(0)) = (u_0, r)$, since $r = \textit{const}$ and $u_i = \langle \lambda, X_i \rangle$ and at the starting point we have $X_i(0) = \partial_{x_i}$.

4.2.2 Exponential map in the corank 2 case

From now on we focus on the case $(m, m+2)$, i.e. when the corank k is equal to 2. We can write the equation of geodesics starting from the origin as follows

$$\begin{cases} x(t) = \displaystyle\int_0^t e^{s(r_1 L_1 + r_2 L_2)} u_0 ds, \quad x(0) = 0, \\ y_1(t) = \dfrac{1}{2} \displaystyle\int_0^t x(s)' L_1\, u(s) ds, \quad y_1(0) = 0, \\ y_2(t) = \dfrac{1}{2} \displaystyle\int_0^t x(s)' L_2\, u(s) ds, \quad y_2(0) = 0. \end{cases} \tag{4.10}$$

Remark 4.11 (Notation). In the following we denote by $E_{L_1,L_2}^{u_0,r_1,r_2}(t)$ the geodesic, parametrized by the length, and starting from the origin, defined by equations (4.10), associated with L_1, L_2.

Definition 4.12. The matrices L_1, L_2 being fixed, the *exponential map* is the map $\mathcal{E} : \mathbb{R}^+ \times \Lambda \to \mathbb{R}^n$ defined by

$$\mathcal{E}(t, u_0, r_1, r_2) = E_{L_1,L_2}^{u_0,r_1,r_2}(t), \qquad \Lambda = \{(u_0, r_1, r_2), u_0 \in S^{k-1}, r_i \in \mathbb{R}\}$$

Remark 4.13. The optimal control problem

$$\begin{cases} \dot{x}_i = u_i, & i = 1, \dots, k, \\ \dot{y}_1 = \frac{1}{2} x' L_1 u, \\ \dot{y}_2 = \frac{1}{2} x' L_2 u, \end{cases} \tag{4.11}$$

is invariant with respect to the following change of coordinates

(a) orthogonal changes of coordinates in the x space,

(b) linear changes of coordinates in the y space.

Indeed, let M be a nonsingular orthogonal matrix ($M^{-1} = M'$) and define the new coordinates $\widetilde{x} = Mx$. Then

$$\dot{\widetilde{x}} = M\dot{x} = Mu =: \widetilde{u},$$

and

$$\dot{y}_i = x' L_i u = (Mx)' M L_i M' (Mu) = \widetilde{x}' M L_i M' \widetilde{u}.$$

Hence, in the new coordinates, L_i is changed for $\widetilde{L}_i := M L_i M'$.

Also, it is easy to see that the change of coordinates

$$\widetilde{y}_1 = \alpha_1 y_1 + \alpha_2 y_2, \qquad \widetilde{y}_2 = \beta_1 y_1 + \beta_2 y_2, \tag{4.12}$$

corresponds to the change

$$\widetilde{L}_1 = \alpha_1 L_1 + \alpha_2 L_2, \qquad \widetilde{L}_2 = \beta_1 L_1 + \beta_2 L_2.$$

In other words we can change L_1 and L_2 up to congruence and linear combinations.

Using these arguments one immediately gets

Lemma 4.14. *Let $(r_1, r_2) =: (r\cos\theta, r\sin\theta)$ and $L_\theta := \cos\theta L_1 + \sin\theta L_2$, $\widetilde{L}_\theta := -\sin\theta L_1 + \cos\theta L_2$. Consider the rotation matrix $R_\theta = \left(\begin{smallmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{smallmatrix}\right)$ and the orthogonal matrix M such that $M L_\theta M'$ is block diagonal.*

Denote $\Omega = \left(\begin{smallmatrix} M & 0 \\ 0 & R_\theta \end{smallmatrix}\right)$ and $\widetilde{u}_0 = M u_0$. We have the equality

$$\Omega\, E^{u_0, r_1, r_2}_{L_1, L_2}(t) = E^{\widetilde{u}_0, r, 0}_{L_\theta, \widetilde{L}_\theta}(t). \tag{4.13}$$

Thanks to Lemma 4.14, one can always restrict to geodesics of the type

$$\begin{cases} x(t) = \displaystyle\int_0^t e^{srL_\theta} u_0 ds, \\ y_1(t) = \displaystyle\frac{1}{2}\int_0^t x(s)' L_\theta\, u(s) ds, \\ y_2(t) = \displaystyle\frac{1}{2}\int_0^t x(s)' \widetilde{L}_\theta\, u(s) ds, \end{cases} \tag{4.14}$$

where L_θ is in the block-diagonal form

$$L_\theta = \begin{pmatrix} 0 & a_1 & & & \\ -a_1 & 0 & & & \\ & & \ddots & & \\ & & & 0 & a_\ell \\ & & & -a_\ell & 0 \end{pmatrix}, \qquad \text{or} \qquad \begin{pmatrix} 0 & a_1 & & & & \\ -a_1 & 0 & & & & \\ & & \ddots & & & \\ & & & 0 & a_\ell & \\ & & & -a_\ell & 0 & \\ & & & & & 0 \end{pmatrix},$$

depending on the fact that m is even ($m = 2\ell$) or odd ($m = 2\ell + 1$), and where the geodesic is associated with the covector $(r_1, r_2) = (r, 0)$.

Remark 4.15. When we deal with a fixed sub-Riemannian metric we can assume also that the coordinates in the x space are chosen in such a way that $a_1 \geq a_i$, for every i. In this case

$$\frac{2\pi}{r \max(\sigma(L_\theta))} = \frac{2\pi}{a_1 r}.$$

4.2.3 Computation of the cut time

In this section we prove Theorem 4.5, i.e. we compute the last time at which a geodesic parametrized by length is optimal.

We first consider the case $r = 0$. In this case equations (4.10) can be easily integrated and gives the straight lines

$$\begin{cases} x(t) = u_0 t, \\ y_i(t) = 0. \end{cases}$$

This trajectory is optimal for any time (i.e. $t_{cut} = +\infty$) since the sub-Riemannian length of a geodesic coincides with the Euclidean length of its projection on the horizontal subspace $(x_1, \ldots, x_m)$, as follows from formula (1.6).

In what follows we use the notation $A = rL_\theta$, $\widetilde{A} = r\widetilde{L}_\theta$ and we focus on the case when A is even dimensional (i.e. $m = 2\ell$) and invertible. The case A non invertible (in particular A odd dimensional) needs an obvious modification of the proof.

With this notation the system (4.14) is rewritten as

$$\begin{cases} x(t) = A^{-1}(e^{tA} - I)u_0, \\ y_1(t) = \dfrac{1}{2}\displaystyle\int_0^t x(s)' A\, u(s) ds, \\ y_2(t) = \dfrac{1}{2}\displaystyle\int_0^t x(s)' \widetilde{A}\, u(s) ds, \end{cases} \tag{4.15}$$

Maxwell points

Consider $\gamma(t) = E^{u_0, r, 0}(t)$, the geodesic associated with the problem (4.15) and with initial covector $(u_0, r, 0), r > 0$. Let us first show that there exists another geodesic reaching the point $\gamma(T^*)$ in time $T^* = 2\pi/(a_1 r)$. Using Arnol'd's terminology, points reached in the same time by more than one geodesic are called Maxwell points. At the end of this section we prove that γ cannot be optimal after T^*.

Set $u_0 = (u_1, u_2, u_3, \dots, u_m)$ and consider the following variation of the horizontal covector

$$u_0^\omega = (\cos\omega\, u_1 + \sin\omega\, u_2, -\sin\omega\, u_1 + \cos\omega\, u_2, u_3, \dots, u_m), \qquad \omega \in [0, 2\pi].$$

Denote $\gamma^\omega(t) = (x^\omega, y_1^\omega, y_2^\omega) := E^{u_0^\omega, r, 0}(t)$ the geodesic associated with this variation.

Claim: There exists $\omega \neq 0$ such that $\gamma(T^*) = \gamma^\omega(T^*)$.

Proof of the Claim. Denote by $M_A(t) := A^{-1}(e^{tA} - I)$ and notice that

$$M_A(t) = A^{-1}(e^{tA} - I) = \begin{pmatrix} \frac{\sin a_1 rt}{a_1 r} & \frac{1-\cos a_1 rt}{a_1 r} & & & \\ \frac{-1+\cos a_1 rt}{a_1 r} & \frac{\sin a_1 rt}{a_1 r} & & & \\ & & \ddots & & \\ & & & \frac{\sin a_\ell rt}{a_\ell r} & \frac{1-\cos a_\ell rt}{a_\ell r} \\ & & & \frac{-1+\cos a_\ell rt}{a_\ell r} & \frac{\sin a_\ell rt}{a_\ell r} \end{pmatrix}.$$

In other words we can write

$$M_A(t) = \begin{pmatrix} D_1(t) & & \\ & \ddots & \\ & & D_\ell(t) \end{pmatrix},$$

where

$$D_i(t) = \begin{pmatrix} \frac{\sin a_i rt}{a_i r} & \frac{1-\cos a_i rt}{a_i r} \\ \frac{-1+\cos a_i rt}{a_i r} & \frac{\sin a_i rt}{a_i r} \end{pmatrix} = 2\frac{\sin(a_i rt/2)}{a_i r} \begin{pmatrix} \cos(a_i rt/2) & \sin(a_i rt/2) \\ -\sin(a_i rt/2) & \cos(a_i rt/2) \end{pmatrix}. \tag{4.16}$$

We prove our claim by steps.
(i). From (4.16) it is easy to see that

$$x^\omega(T^*) - x(T^*) = A^{-1}(e^{T^*A} - I)(u_0^\omega - u_0) = M_A(T^*)(u_0^\omega - u_0) = 0, \qquad \forall\, \omega \in [0, 2\pi],$$

since $e^{T^*A} - I$ (and so $M_A(T^*)$) has its first 2×2 block equal to zero.

(ii). Now we show that $y_1^\omega(T^*) = y_1(T^*)$ for all $\omega \in [0, 2\pi]$. Indeed from (4.15), we get

$$\begin{aligned} y_1(t) &= -\frac{1}{2} u_0' \int_0^t (e^{-sA} - I) e^{sA} ds\, u_0 \\ &= \frac{1}{2} u_0' \int_0^t (e^{sA} - I) ds\, u_0 \\ &= \frac{1}{2} \langle M_A(t) u_0, u_0 \rangle + \frac{1}{2} t \|u_0\|^2, \end{aligned}$$

and

$$y_1^\omega(T^*) - y_1(T^*) = \frac{1}{2}\left(\langle M_A(T^*)u_0^\omega, u_0^\omega\rangle - \langle M_A(T^*)u_0, u_0\rangle\right) + \frac{T^*}{2}(\|u_0^\omega\|^2 - \|u_0\|^2).$$

First notice that

$$\|u_0^\omega\|^2 = \|u_0\|^2, \qquad \forall\, \omega \in [0, 2\pi].$$

Moreover, setting $u_0^\omega = u_0 + v^\omega$ we get (we omit T^* in the argument of M_A)

$$\langle M_A u_0^\omega, u_0^\omega\rangle - \langle M_A u_0, u_0\rangle = \langle M_A v^\omega, v^\omega\rangle + \big\langle (M_A + M_A')v^\omega, u_0\big\rangle = 0$$

since the first 2×2 block of M_A that is zero at T^* and v^ω has nonzero component only in the first two entries.

Remark 4.16. Note that (i) and (ii) are just the manifestation of the fact that, forgetting about the second vertical component y_2, we are facing the corank 1 case, for which T^* is a cut time and there is a rotational symmetry that implies that it is also a conjugate time.

(iii). Now one can proceed in a similar way and compute

$$\begin{aligned} y_2(t) &= \frac{1}{2}\int_0^t x(s)'\widetilde{A}\,u(s)ds \\ &= \frac{1}{2}u_0'\int_0^t (A^{-1}(e^{sA} - I))'\widetilde{A}\,e^{sA}ds\,u_0 \\ &= -\frac{1}{2}u_0'\int_0^t (e^{-sA} - I)A^{-1}\widetilde{A}\,e^{sA}ds\,u_0 \\ &= \langle C(t)u_0, u_0\rangle, \end{aligned}$$

where we set

$$C(t) = \frac{1}{2}\int_0^t (e^{-sA} - I)A^{-1}\widetilde{A}\,e^{sA}ds. \tag{4.17}$$

Since all matrices appearing in (4.17) but $\widetilde{A}$ are 2×2 block diagonal, the first 2×2 diagonal block of

$$K(s) := (e^{-sA} - I)A^{-1}\widetilde{A}\,e^{sA},$$

is the product of the respective blocks. A direct computation shows that it is

$$\frac{\alpha_0}{a_1}\begin{pmatrix} 1-\cos(a_1 rs) & -\sin(a_1 rs) \\ \sin(a_1 rs) & 1-\cos(a_1 rs)\end{pmatrix},$$

where $\begin{pmatrix} 0 & \alpha_0 \\ -\alpha_0 & 0\end{pmatrix}$ denotes the first 2×2 block of $\widetilde{A}$. Integrating from 0 to T^* one obtains for the first block of $C(T^*)$

$$\frac{\pi\alpha_0}{a^2r^2}\begin{pmatrix} 1 & 0 \\ 0 & 1\end{pmatrix}. \tag{4.18}$$

As before, we set $u_0^\omega = u_0 + v^\omega$ and we get (omiting T^* in the argument of C)

$$y_2^\omega(T^*) - y_2(T^*) = \langle Cv^\omega, v^\omega\rangle + \big\langle (C + C')v^\omega, u_0\big\rangle. \tag{4.19}$$

Using (4.18) and

$$\begin{aligned}\|v^\omega\|^2 &= ((\cos\omega - 1)u_1 + \sin\omega\, u_2)^2 + (-\sin\omega\, u_1 + (\cos\omega - 1)u_2)^2 \\ &= 4(u_1^2 + u_2^2)\sin^2(\omega/2),\end{aligned}$$

one gets that (4.19) is linear with respect to the variables

$$\cos\omega - 1 = 2\sin^2(\omega/2), \qquad \sin\omega = 2\cos(\omega/2)\sin(\omega/2).$$

In other words, if we prescribe the expression (4.19) to be zero, we get

$$C_0 \sin(\omega/2)(C_1\cos(\omega/2) + C_2\sin(\omega/2)) = 0, \tag{4.20}$$

for some suitable constants C_0, C_1, C_2 that do not depend on ω. The Claim is proved since equation (4.20) has always a nontrivial solution $\widetilde{\omega} \in [0, 2\pi]$. □

Let us now show that $\gamma(t)$ cannot be optimal after T^*. From the previous computation we have $\dot\gamma(T^*) \neq \dot\gamma^{\widetilde{\omega}}(T^*)$. By contradiction if γ is optimal after time T^* then the concatenation of $\gamma^{\widetilde{\omega}}|_{[0,T^*]}$ and $\gamma|_{[T^*,T^*+\varepsilon]}$ (for some $\varepsilon > 0$) is optimal as well, which is impossible since all optimal trajectories are projections of the Hamiltonian system associated with (4.8) and they are smooth.

Optimality of geodesics

In this section we prove that $\gamma(t) = E^{u_0,r,0}(t)$, $r > 0$, is optimal up to its first Maxwell time $T^* = 2\pi/(a_1 r)$.
To this extent, consider the following auxiliary optimal control problem:

P. Let $T < T^*$ and set $(\bar x, \bar y_1, \bar y_2) = \gamma(T)$. Find a length-parametrized trajectory of the system

$$\begin{cases} \dot x_i = u_i, & i = 1, \dots, k, \\ \dot y_1 = \frac{1}{2} x' L_1 u, \\ \dot y_2 = \frac{1}{2} x' L_2 u, \end{cases} \tag{4.21}$$

starting from the origin, and reaching the hyperplane $\{x = \bar x\}$ in time T, maximizing the y_1 coordinate.

Remark 4.17. Notice that $\bar y_1 > 0$ since $r > 0$ implies that the trajectory is not contained in the hyperplane $\{y_1 = 0\}$.

Lemma 4.18. *The following assertions hold: (i) There exists a solution γ^* of the problem* **P**. *(ii) γ^* is a length minimizer. (iii) $\gamma^*(t) = E^{\widetilde{u}_0,\widetilde{r},0}(t)$ for some $(\widetilde{u}_0, \widetilde{r})$.*

Proof. Let us prove (i). In problem **P**, since we deal with length-parametrized trajectories, we can assume that the set of controls in (4.21) is $U = \{\|u\| \le 1\}$. The existence of a solution of **P** can be obtained with standard arguments using the compactness and convexity of the set of admissible velocities (see [13, 47]).

To prove (ii) assume by contradiction, that there exists a trajectory of (4.21) reaching the point $(\bar x, y_1^*, y_2^*) = \gamma^*(T)$ in time $T_0 < T$. By small time controllability there exists a trajectory of system (4.21) reaching in time T the point $(\bar x, \widehat{y}_1, \widehat{y}_2)$, with $\widehat{y}_1 > y_1^*$ contradicting the fact that γ^* maximize the y_1 coordinate. The fact that γ^* is also a length minimizer follows from Remark 4.8.

To prove (iii) observe that γ^* satisfies the Pontryagin Maximum Principle (see again [13]) for the problem of minimizing $-y_1 = -\int_0^T \dot{y}_1 dt = -\int_0^T x' L_1 u \, dt$, i.e. with the Hamiltonian

$$
\begin{aligned}
\mathcal{H}_u &= \sum_{i=1}^{k} \langle \lambda, u_i X_i \rangle + \nu x' L_1 u \qquad (4.22)\\
&= pu + r_1 x' L_1 u + r_2 x' L_2 u + \nu x' L_1 u.
\end{aligned}
$$

where $\lambda = (p, r_1, r_2)$ are the dual variables to (x, y_1, y_2) in T^*M. In formula (4.22) ν is a nonnegative constant. The Hamiltonian equations give

$$
\begin{cases}
\dot{r}_1 = -\frac{\partial \mathcal{H}_u}{\partial y_1} = 0,\\
\dot{r}_2 = -\frac{\partial \mathcal{H}_u}{\partial y_2} = 0,\\
\dot{p}' = -\frac{\partial \mathcal{H}_u}{\partial x} = -(r_1 L_1 + r_2 L_2 + \nu L_1) u.
\end{cases}
$$

Since the final point is constrained on the set $\{x = \bar{x}\}$, the transversality conditions give $r_1 = 0, r_2 = 0$. Hence we have

$$
\begin{aligned}
\mathcal{H}_u &= (p + \nu x' L_1) u,\\
\dot{p}' &= -\nu L_1 u. \qquad (4.23)
\end{aligned}
$$

Notice that actually $\nu > 0$, otherwise the trajectory is a straight line contained in the plane $\{y_1 = y_2 = 0\}$, see Remark 4.17. The maximality condition and the condition that the final time is fixed in such a way that trajectory are parametrized by length give

$$
\mathcal{H}_{u(t)}(x(t), y_1(t), y_2(t), p(t), r_1, r_2) = \max_v \mathcal{H}_v(x(t), y_1(t), y_2(t), p(t), r_1, r_2) = 1,
$$

$$
u(t) = \frac{p'(t) - \nu L_1 x(t)}{\|p'(t) - \nu L_1 x(t)\|} = p'(t) - \nu L_1 x(t). \qquad (4.24)
$$

Notice that a geodesic for the problem (4.21) associated with the covector (u_0, r_1, r_2) corresponds to a control

$$
u(t) = p' - r_1 L_1 x - r_2 L_2 x, \qquad (4.25)
$$

where

$$
\dot{p}' = (-r_1 L_1 - r_2 L_2)(p' - r_1 L_1 x - r_2 L_2 x). \qquad (4.26)
$$

Comparing equations (4.23) - (4.24) with (4.25) - (4.26) it follows that γ^* is a geodesic for the problem (4.21) corresponding to an initial covector $(\widetilde{u}_0, \nu, 0)$, for some $\widetilde{u}_0$, with $\|\widetilde{u}_0\| = 1$. Then (iii) is proved for $\widetilde{r} = \nu$. □

We have the following

Claim. $\gamma^* = \gamma$, i.e. $\widetilde{u}_0 = u_0$ and $\widetilde{r} = r$.

Proof of the Claim. It is enough to prove that the parameters u_0, r such that a geodesic $\gamma(t) = (x(t), y_1(t), y_2(t)) = E^{u_0,r,0}$ satisfies $x(T) = \bar{x}$ with $T < T^*$ are unique.

From the computations in Sections 4.2.3 we know that

$$x(t) = M_A(t)u_0, \qquad \text{where} \quad A = rL_1, \quad \text{and} \quad M_A(t) = A^{-1}(e^{tA} - I).$$

In particular, using the non singularity of A, the equality at $t = T$ gives

$$u_0 = M_A^{-1}(T)\bar{x}. \tag{4.27}$$

Computing the norm of vectors in equality (4.27), it follows

$$1 = \|u_0\|^2 = \sum_{i=1}^{\ell} \frac{\rho_i(\bar{x})^2}{T^2} \frac{a_i rT/2}{\sin(a_i rT/2)}, \qquad \text{where} \quad \rho_i(x) = (x_{2i-1}^2 + x_{2i}^2)^{1/2}. \tag{4.28}$$

Notice that the right hand side of (4.28) is the sum of monotonic functions with respect to the variable rT, on the segment $[0, 2\pi/a_1]$ ($T < T^*$ implies $rT \leq 2\pi/a_1$).

Moreover since the curve is length-parametrized we have $\|x(T)\| \leq T$. As a consequence there exists a unique solution rT of equation (4.28) in the segment $[0, 2\pi/a_1]$. In particular r is uniquely determined and u_0 is uniquely recovered from equation (4.27). □

Since $\gamma = \gamma^*$ and γ^* is length-minimizer for every $T < T^*$, it follows that $t_{cut} = T^*$.

4.2.4 First conjugate time

In this section we prove that in the corank 2 case, the cut time is not equal to the first conjugate time, in general. This is deeply different from the corank 1 case, where the cut locus always coincides with the first conjugate locus.

It is enough to show that the cut time is not conjugate in the $(4, 6)$ case. Define the Jacobian of the exponential map

$$J_{\mathcal{E}}(t, u_0, r_1, r_2) := \det\left(\frac{\partial \mathcal{E}}{\partial t}, \frac{\partial \mathcal{E}}{\partial u_0}, \frac{\partial \mathcal{E}}{\partial r_1}, \frac{\partial \mathcal{E}}{\partial r_2}\right). \tag{4.29}$$

Remark 4.19. Recall that the first conjugate time t_{con} for the geodesic corresponding to the covector (u_0, r_1, r_2) is the first time $t > 0$ for which we have

$$J_{\mathcal{E}}(t, u_0, r_1, r_2) = 0. \tag{4.30}$$

We have to prove that equation (4.30) is not satisfied when $t = t_{cut}$.

To compute $J_{\mathcal{E}}$ we use the following trick. Let $\mathfrak{L} = pdx + r_1 dy_1 + r_2 dy_2$ be the Liouville form.

Lemma 4.20. *We have* $\mathfrak{L}\left(\frac{\partial \mathcal{E}}{\partial t}\right) = 1$, $\mathfrak{L}\left(\frac{\partial \mathcal{E}}{\partial u_0}\right) = \mathfrak{L}\left(\frac{\partial \mathcal{E}}{\partial r_i}\right) = 0.$

Proof. The first equality follows from the fact that the Hamiltonian is homogeneous of degree 2. Indeed set $\lambda = (p, r_1, r_2)$ and $q = (x, y_1, y_2)$, we have $\mathfrak{L} = \lambda dq$ and

$$\mathfrak{L}\left(\frac{\partial \mathcal{E}}{\partial t}\right) = \left\langle \lambda, \frac{\partial q}{\partial t}\right\rangle = \left\langle \lambda, \frac{\partial H}{\partial \lambda}\right\rangle = 2H = 1,$$

since length-parametrized trajectory belong to the set $\{H = 1/2\}$. The second and the third identities follow from the fact that the Liouville form is preserved by the Hamiltonian flow, hence the values of $\frac{\partial \mathcal{E}}{\partial u_0}$ and $\frac{\partial \mathcal{E}}{\partial r_i}$ are constant with respect to t. In particular at $t = 0$ they are annichilated by the Liouville form. □

If we compute the exponential map in a neighborhood of a geodesic with $(r_1, r_2) = (r, 0)$, with $r \neq 0$, using the identity $r_1 dy_1 = \mathfrak{L} - p dx - r_2 dx_2$ and Lemma 4.20 we get

$$\begin{aligned} J_{\mathcal{E}} &= dx \wedge dy_1 \wedge dy_2 \left(\frac{\partial \mathcal{E}}{\partial t}, \frac{\partial \mathcal{E}}{\partial u_0}, \frac{\partial \mathcal{E}}{\partial r_1}, \frac{\partial \mathcal{E}}{\partial r_2}\right) \\ &= \frac{1}{r_1} dx \wedge dy_2 \left(\frac{\partial \widetilde{\mathcal{E}}}{\partial u_0}, \frac{\partial \widetilde{\mathcal{E}}}{\partial r_1}, \frac{\partial \widetilde{\mathcal{E}}}{\partial r_2}\right), \end{aligned} \tag{4.31}$$

where $\widetilde{\mathcal{E}}(t, u_0, r_1, r_2) = (x(t, u_0, r_1, r_2), y_2(t, u_0, r_1, r_2))$ denote the exponential map where y_1 is removed. More precisely, (4.31) is the function of (t, u_0, r_1, r_2) given by

$$J_{\mathcal{E}} = \frac{1}{r_1} \det \begin{pmatrix} \frac{\partial x}{\partial u_0} v_1 & \frac{\partial x}{\partial u_0} v_2 & \frac{\partial x}{\partial u_0} v_3 & \frac{\partial x}{\partial r_1} & \frac{\partial x}{\partial r_2} \\ \frac{\partial y_2}{\partial u_0} v_1 & \frac{\partial y_2}{\partial u_0} v_2 & \frac{\partial y_2}{\partial u_0} v_3 & \frac{\partial y_2}{\partial r_1} & \frac{\partial y_2}{\partial r_2} \end{pmatrix},$$

where v_1, v_2, v_3 are 3 independent tangent vectors to the 3-sphere $\{u_0 \in \mathbb{R}^4, \|u_0\| = 1\}$. We select

$$v_1 = \begin{pmatrix} -u_2 \\ u_1 \\ 0 \\ 0 \end{pmatrix}, \quad v_2 = \begin{pmatrix} 0 \\ 0 \\ -u_4 \\ u_3 \end{pmatrix}, \quad v_3 = \begin{pmatrix} -u_4 \\ 0 \\ 0 \\ u_1 \end{pmatrix}. \tag{4.32}$$

From the computation of Section 4.2.3 one easily gets

$$\frac{\partial x}{\partial u_0} = M_A(t) = A^{-1}(e^{tA} - I),$$

$$\begin{aligned} \frac{\partial x}{\partial r_1} &= -A^{-1}A(A^{-1}(e^{tA} - I) + t\, e^{tA})u_0 \\ &= -(M_A(t) + t\, e^{tA})u_0, \end{aligned}$$

$$\begin{aligned} \frac{\partial x}{\partial r_2} &= -A^{-1}\widetilde{A}(A^{-1}(e^{tA} - I) + t\, e^{tA})u_0 \\ &= -A^{-1}\widetilde{A}(M_A(t) + t\, e^{tA})u_0. \end{aligned}$$

Moreover (see again Section 4.2.3)

$$y_2(t) = \langle C(t)u_0, u_0 \rangle \,, \tag{4.33}$$

where

$$C(t) = -\frac{1}{2}\int_0^t (e^{-sA} - I)A^{-1}\widetilde{A}\, e^{sA} ds. \tag{4.34}$$

The function $y_2(t)$ from (4.33), being a quadratic form with respect to u_0, gives

$$\frac{\partial y_2}{\partial u_0} v_i = \big\langle (C(t) + C'(t))u_0, v_i \big\rangle \,.$$

Now we compute these derivatives at $t = t_{cut} = \frac{2\pi}{ar}$ where $a > b$ are the moduli of the eigenvalues of A.

It is easily seen that

$$B := M_A(t_{cut}) = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{\sin(2\pi b/a)}{br} & \frac{2\sin^2(\pi b/a)}{br} \\ 0 & 0 & -\frac{2\sin^2(\pi b/a)}{br} & \frac{\sin(2\pi b/a)}{br} \end{pmatrix},$$

from which it follows that

$$\begin{pmatrix} Bv_1 & Bv_2 & Bv_3 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & u_3 \frac{2\sin^2(\pi b/a)}{br} - u4 \frac{\sin(2\pi b/a)}{br} & u_1 \frac{2\sin^2(\pi b/a)}{br} \\ 0 & u_4 \frac{2\sin^2(\pi b/a)}{br} + u_3 \frac{\sin(2\pi b/a)}{br} & u_1 \frac{\sin(2\pi b/a)}{br} \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & M \end{pmatrix},$$

where the last identity defines the matrix M.

The determinant of the exponential map computed at $t = t_{cut}$ is then expressed as follows

$$
\begin{aligned}
J_{\mathcal{E}}(t_{cut}) &= \det \begin{pmatrix} \frac{\partial x}{\partial u_0} v_1 & \frac{\partial x}{\partial u_0} v_2 & \frac{\partial x}{\partial u_0} v_3 & \frac{\partial x}{\partial r_1} & \frac{\partial x}{\partial r_2} \\ \frac{\partial y_2}{\partial u_0} v_1 & \frac{\partial y_2}{\partial u_0} v_2 & \frac{\partial y_2}{\partial u_0} v_3 & \frac{\partial y_2}{\partial r_1} & \frac{\partial y_2}{\partial r_2} \end{pmatrix} \\
&= \det \begin{pmatrix} \vdots & \vdots & \vdots & \vdots & \vdots \\ Bv_1 & Bv_2 & Bv_3 & \frac{\partial x}{\partial r_1} & \frac{\partial x}{\partial r_2} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \frac{\partial y_2}{\partial u_0} v_1 & \frac{\partial y_2}{\partial u_0} v_2 & \frac{\partial y_2}{\partial u_0} v_3 & \frac{\partial y_2}{\partial r_1} & \frac{\partial y_2}{\partial r_2} \end{pmatrix} \\
&= \det \begin{pmatrix} 0 & 0 & 0 & \frac{\partial x_1}{\partial r_1} & \frac{\partial x_1}{\partial r_2} \\ 0 & 0 & 0 & \frac{\partial x_2}{\partial r_1} & \frac{\partial x_2}{\partial r_2} \\ 0 & M_{11} & M_{12} & \frac{\partial x_3}{\partial r_1} & \frac{\partial x_3}{\partial r_2} \\ 0 & M_{21} & M_{22} & \frac{\partial x_4}{\partial r_1} & \frac{\partial x_4}{\partial r_2} \\ \langle Cu_0, v_1 \rangle & \langle Cu_0, v_2 \rangle & \langle Cu_0, v_3 \rangle & \frac{\partial y_2}{\partial r_1} & \frac{\partial y_2}{\partial r_2} \end{pmatrix}, \qquad (4.35)
\end{aligned}
$$

where we use the notation

$$
M = \begin{pmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{pmatrix}.
$$

From (4.35) it follows

$$
J_{\mathcal{E}}(t_{cut}) = \langle Cu_0, v_1 \rangle \cdot \det M \cdot \det N, \qquad (4.36)
$$

where N is the matrix

$$
N = \begin{pmatrix} \frac{\partial x_1}{\partial r_1} & \frac{\partial x_1}{\partial r_2} \\ \frac{\partial x_2}{\partial r_1} & \frac{\partial x_2}{\partial r_2} \end{pmatrix}.
$$

It is easy to see from the explicit expression of the geodesics that, in the general case when $a \neq b$, the three factors in (4.36) do not vanish identically in u_0, since the matrix $\widetilde{A}$ is arbitrary. The proof of Theorem 4.5 is then completed.

Remark 4.21. Notice that M is the zero matrix when $a = b$. Hence, in the $(4,6)$ case, $t_{cut} = t_{con}$ for those θ such that L_θ has double eigenvalue. Moreover in this case the rank of the Jacobian matrix drops by 2, since the first three columns are proportional.

4.3 The nilpotent $(4,6)$ case

In this section we restrict to the (4,6) case. By the previous discussion the geodesics of the sub-Riemannian metric can be written as follows

$$\begin{cases} x(t) = A^{-1}(e^{tA} - I)u_0 \\ y_1(t) = \dfrac{1}{2}\displaystyle\int_0^t x(s)' A\, u(s) ds, \\ y_2(t) = \dfrac{1}{2}\displaystyle\int_0^t x(s)' \widetilde{A}\, u(s) ds \end{cases} \qquad A = rL_\theta,\ \widetilde{A} = r\widetilde{L}_\theta, \tag{4.37}$$

and we can assume the matrix L_θ to be diagonal

$$L_\theta = \begin{pmatrix} 0 & a & 0 & 0 \\ -a & 0 & 0 & 0 \\ 0 & 0 & 0 & b \\ 0 & 0 & -b & 0 \end{pmatrix}, \qquad a \geq b, \tag{4.38}$$

while $\widetilde{L}_\theta$ is an arbitrary skew-symmetric matrix

$$\widetilde{L}_\theta = \begin{pmatrix} 0 & \alpha_0 & \alpha_1 & \alpha_2 \\ -\alpha_0 & 0 & \alpha_3 & \alpha_4 \\ -\alpha_1 & -\alpha_3 & 0 & \alpha_5 \\ -\alpha_2 & \alpha_4 & -\alpha_5 & 0 \end{pmatrix}.$$

4.3.1 Proof of Theorem 4.6

Recall that the cut time t_{cut} coincide with t_{con} if and only if t_{cut} is a time that satisfies the equation

$$J_{\mathcal{E}}(t, u, \theta)\big|_{t=t_{cut}} = 0. \tag{4.39}$$

$(P2) \Rightarrow (P1)$. We consider separately the two cases:

(a) L_1, L_2 both belong to the same subspace, either $\mathcal{Q}$ or $\widehat{\mathcal{Q}}$. Then it is not restrictive to assume that $L_1, L_2 \in \mathcal{Q}$. In this case all linear combination of L_1, L_2 belong to $\mathcal{Q}$, i.e. $L_\theta = \cos(\theta)L_1 + \sin(\theta)L_2 \in \mathcal{Q}$ for every θ. In particular L_θ has a double eigenvalue for every $\theta \in [0, 2\pi]$. From the computation of Section 4.2 it is easily seen that $a = b$ implies $M = 0$, hence from (4.36) it follows that $t_{cut} = t_{con}$.

(b) $L_1 \in \mathcal{Q}$ and $L_2 \in \widehat{\mathcal{Q}}$ (L_1 and L_2 plays the same role). By (4.51) we have $[L_1, L_2] = 0$. Let us prove then that this property implies $(P1)$.

Indeed every two commuting skew-symmetric matrices can be block diagonalized simultaneously in the same basis. Hence we can assume that, choosing an appropriate coordinate system y_1, y_2, that both L_1, L_2 are diagonal. As a consequence L_θ and $\widetilde{L}_\theta$ are also diagonal. Moreover from (4.34) it is easily seen that, if both L_θ and $\widetilde{L}_\theta$ are diagonal, C is 2×2 block diagonal, with the first block equal to cI, for some constant c (see also (4.18)).

In particular it follows that $\langle Cu_0, v_1\rangle = 0$ and again (4.36) implies $t_{cut} = t_{con}$.

$(P1) \Rightarrow (P2)$. By assumption the identity

$$J_{\mathcal{E}}(t,u_0,\theta)\big|_{t=t_{cut}(\theta)} = \langle C(\theta)u_0, v_1\rangle \cdot \det M(u_0,\theta) \cdot \det N(u_0,\theta) = 0, \qquad (4.40)$$

holds for every u_0 (the horizontal part of the initial covector) and every θ. Since the exponential map is linear with respect to u_0 in the x-variable, and quadratic with respect to u_0 in the y_i-variables, it follows that (4.40) is an analytic expression of (u_0,θ) (it is polynomial with respect to u_0 and trigonometric in θ). In particular one of the three factors in (4.40) must vanish identically.

Assume that $\det M(u_0,\theta) \equiv 0$. Then from the explicit expression it is computed that

$$\det M(u_0,\theta) = \frac{4u_1u_3}{b(\theta)^2 r^2} \sin^2\left(\pi \frac{b(\theta)}{a(\theta)}\right),$$

and since $a(\theta) \geq b(\theta)$ by assumption, $\det M \equiv 0$ implies $a(\theta) = b(\theta)$, for all θ.

From this it easily follows that L_θ has double eigenvalue for all θ, i.e. if we write

$$L_\theta = q(\theta) + \widehat{q}(\theta),$$

it follows that one of $\|q(\theta)\|$ and $\|\widehat{q}(\theta)\|$ is identically zero (it is not a restriction to assume $\|\widehat{q}(\theta)\| \equiv 0$). Hence $L_\theta \in \mathcal{Q}$ for all θ, that implies in particular that $L_1, L_2 \in \mathcal{Q}$.

It is not restrictive now to assume that $a(\bar\theta) \neq b(\bar\theta)$ for some $\bar\theta \in [0,2\pi]$. We show that the identities

(a) $\det N(u_0,\bar\theta) = 0$,

(b) $\langle C(\bar\theta)u_0, v_1\rangle = 0$,

both imply that there exists a choice of the coordinates such that $L_1 \in \mathcal{Q}$, $L_2 \in \widehat{\mathcal{Q}}$. We give details only for case (b), the other one is similar. Considering (b) as an equation in the variables $\alpha_1,\alpha_2,\alpha_3,\alpha_4$ (the non diagonal entries of the matrix $\widetilde{L}_\theta$) it is easy to see that the identity (b) can be written as an equation

$$F(\alpha_i, u_i) = 0,$$

where F is a quadratic form in the u_i whose coefficients depend linearly on α_i. Since these equation should be satisfied for all $u_0 = (u_1,\ldots,u_4)$, choosing values

$$u_0 \in \{(1,0,1,0),(1,0,0,1),(0,1,1,0),(0,1,0,1)\},$$

one gets the set of 4 linear equations:

$$\begin{cases} (a\alpha_4 + b\alpha_1)\cos\eta + (a\alpha_3 - b\alpha_2)\sin\eta = 0, \\ (a\alpha_3 - b\alpha_2)\cos\eta - (a\alpha_4 + b\alpha_1)\sin\eta = 0, \\ (a\alpha_2 - b\alpha_3)\cos\eta + (a\alpha_1 + b\alpha_4)\sin\eta = 0, \\ (a\alpha_1 + b\alpha_4)\cos\eta - (a\alpha_2 - b\alpha_3)\sin\eta = 0, \end{cases} \qquad (4.41)$$

where we set $\eta = \pi b/a$, and for simplicity of the notation we denote $a = a(\theta_0), b = b(\theta_0)$.

It is easy to show, using the fact that $a \neq b$, that this system has the unique solution

$$\alpha_1 = \ldots = \alpha_4 = 0,$$

which means that L_1 and L_2 are both diagonal. Due to this fact they can be written, as pure quaternions (see Appendix 4.4.1), as a linear combination of $i, \widehat{i}$

$$L_1 = \alpha i + \widehat{\alpha}\,\widehat{i}, \qquad L_2 = \beta i + \widehat{\beta}\,\widehat{i}. \tag{4.42}$$

Performing the change of variables

$$\begin{pmatrix} \widetilde{y}_1 \\ \widetilde{y}_2 \end{pmatrix} = \begin{pmatrix} \alpha & \widehat{\alpha} \\ \beta & \widehat{\beta} \end{pmatrix}^{-1} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix},$$

we find a system of coordinates such that $L_1 = i$, and $L_2 = \widehat{i}$, i.e. that satisfies (P2).

4.3.2 Proof of Theorem 4.7

In Chapter 3 it is proved that, on a regular sub-Riemannian manifold M with Hausdorff dimension Q, the Radon-Nikodym derivative of the spherical Hausdorff measure $\mathcal{S}^Q$ with respect to the Popp's measure μ, denoted $f_{\mathcal{S}\mu}$, is given by the volume of the unit ball in the nilpotent approximation, namely

$$f_{\mathcal{S}\mu}(q) = \frac{2^Q}{\widehat{\mu}_q(\widehat{B}_q)}, \qquad q \in M, \tag{4.43}$$

where $\widehat{\mu}_q$ is the Popp's measure defined on the nilpotent approximation G_q of the structure at the point q. Note that $\widehat{\mu}_q$ is the left-invariant measure on G_q that coincide with the Popp's measure of the original sub-Riemannian metric at the point q.

Remark 4.22. Notice that in our $(k, k+2)$ case the structure is automatically regular since, by assumption, the distribution has constant rank and with one bracket we get all the tangent space.

Remark 4.23. Recall that $f_{\mathcal{S}\mu}$ is a continuous function, which is bounded and bounded away from zero, in restriction to compact sets.

Remark 4.24. For the analysis of the regularity of (4.43) is convenient to parametrize the nilpotent unit ball via the exponential map, as a function defined on the whole fiber in the cotangent space. In other words we do not restrict to the set $\{\|u_0\| = 1\}$ and define for every $\lambda_0 = (u_0, r) \in \mathbb{R}^6$

$$\mathcal{E}(\lambda_0) = \pi(e^{\vec{H}}(\lambda_0)),$$

where H is the Hamiltonian defined in (4.8) and $e^{t\vec{H}}$ denotes the flow in T^*M of the Hamiltonian vector field associated with H. Using the homogeneity property $H(c\lambda) = c^2 H(\lambda)$, $\forall\, c > 0$, we have that

$$e^{\vec{H}}(s\lambda) = e^{s\vec{H}}(\lambda), \qquad \forall\, s > 0.$$

In other words we can recover the geodesic on the manifold with initial covector λ_0 as the image of the ray $\{t\lambda_0, t \in [0,1]\} \subset T^*_{q_0}M$ that join the origin to λ_0.

$$\mathcal{E}(t\lambda_0) = \pi(e^{\vec{H}}(t\lambda_0)) = \pi(e^{t\vec{H}}(\lambda_0)) = \gamma(t).$$

Due to the previous analysis and thanks to Remark 4.24, we can express the volume of the unit ball of the nilpotent approximation as follows

$$V = \int_0^{2\pi} \int_0^{A(\theta)} \int_B J_{\mathcal{E}}(u,\theta,r)dudrd\theta, \tag{4.44}$$

where $J_{\mathcal{E}}$ is the Jacobian of the exponential map, expressed in the new variables, $B = \{u_0 = (u_1, \dots, u_4), \|u_0\| \leq 1\}$ is the 4-dimensional unit ball and

$$A(\theta) = \frac{2\pi}{\max \sigma(L_\theta)}.$$

The problem of the regularity of the function (4.43) is then reduced to the regularity of the function

$$p \mapsto V(p),$$

where p is a (6-dimensional) parameter. Since the family of sub-Riemannian metrics is smooth with respect to p, the exponential map smoothly depends on the parameter p. As a consequence the integrand in (4.44), being the Jacobian of the exponential map, is a smooth function of its variables.

In addition, the function $p \mapsto A(\theta,p)$ is Lipschitz, being the inverse of the "maximum moduli of eigenvalues" function, which is Lipschitz (see [73]). In particular $A(\theta,p)$ admits bounded first derivative almost everywhere with respect to (θ,p).

Definition 4.25. Define the following sets

- Σ is the set of p such that $\exists\,\theta$ for which $L_\theta(p)$ has a double eigenvalue,
- Σ_0 is the set of p such that $\exists$ *a finite number* of θ for which $L_\theta(p)$ has a double eigenvalue,
- Σ_∞ is the set of p such that $\forall\,\theta$, $L_\theta(p)$ has a double eigenvalue.

Thanks to Lemma 4.30, for a generic sub-Riemannian metric, the set of points $p \in \Sigma_\infty$ is a union of isolated points. Moreover, due to the expression (4.50) of the eigenvalues in terms of the quaternions given in the Appendix 4.4.1, the fact that L_θ has a double eigenvalue for all θ is written as $\|q(\theta)\| = 0$ for all θ (or the same for $\widehat{q}$). This condition is equivalent to the equation $\|q(\theta)\|^2 = 0$, that is analytic in θ. In particular this equation, if it is not identically satisfied, has a finite number of solution in $[0, 2\pi]$.

Remark 4.26. Notice that the expression (4.50) for the eigenvalues, provides a crucial obstruction for the generalization of the result to $m > 4$.

From this it follows that, for a generic sub-Riemannian metric, the set of critical points Σ is the disjoint union $\Sigma = \Sigma_0 \cup \Sigma_\infty$. Moreover the set of points where $p \mapsto A(\cdot, p)$ is not smooth is contained in Σ.

Let us write the volume function, depending on the parameter p, as follows

$$V(p) = \int_{\theta=0}^{2\pi} \int_{r=0}^{A(\theta,p)} f(\theta, r, p) dr d\theta, \tag{4.45}$$

where we denote by

$$f(\theta, r, p) = \int_B J_{\mathcal{E}}(u, \theta, r, p) du.$$

Recall that f is smooth as a function of all its variables, while $A(\theta, p)$ is Lipschitz with respect to the parameters (θ, p). In particular it has bounded derivatives.

We want to prove that V is C^1 at any point p_0. To this extent, let us write

$$V(p) = \int_0^{2\pi} \int_0^{A(\theta,p_0)} f(\theta, r, p) dr d\theta + \int_0^{2\pi} \int_{A(\theta,p_0)}^{A(\theta,p)} f(\theta, r, p) dr d\theta.$$

The function

$$p \mapsto \int_0^{2\pi} \int_0^{A(\theta,p_0)} f(\theta, r, p) dr d\theta,$$

is always smooth since it is the integral of a smooth function (with respect to p) on a fixed domain. Denote now

$$W(p) := \int_{\theta=0}^{2\pi} \int_{A(\theta,p_0)}^{A(\theta,p)} f(\theta, r, p) dr d\theta. \tag{4.46}$$

We are left to prove that W is C^1 around p_0. Notice that, by definition, $W(p_0) = 0$.

Assume that $p_0 \notin \Sigma$. Then, since both functions A and f in (4.46) are smooth, W is C^1 at p_0 and the derivative at a point p (in a neighborhood of p_0), is computed as follows

$$\frac{\partial W}{\partial p_i}(p) = \int_0^{2\pi} \int_{A(\theta,p_0)}^{A(\theta,p)} \frac{\partial f}{\partial p_i}(\theta, r, p) dr d\theta + \int_0^{2\pi} \frac{\partial A}{\partial p_i}(\theta, p) f(\theta, A(\theta, p), p) dr d\theta. \tag{4.47}$$

Assume now that $p_0 \in \Sigma_0$. The first term in (4.47) is continuous. Moreover, since at p_0 there are only a finite number of θ such that L_θ has double eigenvalue we have

$$\frac{\partial A}{\partial p_i}(\theta, p) \underset{p \to p_0}{\longrightarrow} \frac{\partial A}{\partial p_i}(\theta, p_0), \qquad \text{a.e. } \theta \in [0, 2\pi]. \tag{4.48}$$

Since $\frac{\partial A}{\partial p_i}$ is bounded and f is smooth, by Lebesgue's dominated convergence we have that the second term is also continuous.

Finally, consider the case when $p_0 \in \Sigma_\infty$. Since p_0 is an isolated point, the partial derivatives are defined and continuous in $N_{p_0} \setminus \{p_0\}$, where N_{p_0} is a neighborhood of p_0. We claim that

$$\frac{\partial W}{\partial p_i}(p) \longrightarrow 0, \qquad \text{when} \quad p \to p_0 \in \Sigma_\infty. \tag{4.49}$$

Indeed, by definition of Σ_∞, the cut time $A(\theta, p_0)$ coincides with the conjugate time, i.e. it satisfies the identity

$$J_{\mathcal{E}}(u, \theta, A(\theta, p_0), p_0) = 0, \qquad \forall\, \theta \in [0, 2\pi].$$

From this it follows that $f(\theta, A(\theta, p), p) \to 0$ for all θ, that easily implies (4.49). Applying Lemma 4.27 to W, the theorem is proved.

4.4 Proof of some technical lemmas

In this section we collect some technical results that have been used in the previous sections.

4.4.1 Quaternions

The Lie algebra $so(4)$ of 4×4 skew-symmetric matrices is the direct sum

$$so(4) = \mathcal{Q} \oplus \widehat{\mathcal{Q}},$$

where $\mathcal{Q}$ is the space of pure quaternions and $\widehat{\mathcal{Q}}$ is the set of pure skew quaternions.

The space $\mathcal{Q}$ (resp. $\widehat{\mathcal{Q}}$) is generated by the three matrices i, j, k (respectively $\widehat{i}, \widehat{j}, \widehat{k}$)

$$i = \begin{pmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{pmatrix}, \quad j = \begin{pmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix}, \quad k = \begin{pmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix},$$

and

$$\widehat{i} = \begin{pmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{pmatrix}, \quad \widehat{j} = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix}, \quad \widehat{k} = \begin{pmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}.$$

If we endow $so(4)$ with the Hilbert-Schmitd scalar product

$$\langle L_1, L_2 \rangle = \frac{1}{4}\mathrm{trace}(L_1' L_2),$$

then $i, j, k, \widehat{i}, \widehat{j}, \widehat{k}$ is an orthonormal basis.

The eigenvalues ω_1, ω_2 of $A = q + \widehat{q}$ satisfy:

$$-(\omega_{1,2})^2 = (\|q\| \pm \|\widehat{q}\|)^2. \tag{4.50}$$

As a consequence an element $A \in so(4)$ has a double eigenvalue if and only if $A \in \mathcal{Q} \cup \widehat{\mathcal{Q}}$.

Also pure quaternions and pure skew quaternions commute:

$$[q, \widehat{q}] = 0, \quad q \in \mathcal{Q}, \quad \widehat{q} \in \widehat{\mathcal{Q}}. \tag{4.51}$$

4.4.2 A technical Lemma

Lemma 4.27. *Let f be a function germ at $(\mathbb{R}^n, 0)$ satisfying the following conditions:* **i)** *f is continuous,* **ii)** *f is $\mathcal{C}^1$ out of zero,* **iii)** *the limits* $\lim_{x\to 0,\ x\neq 0} \frac{\partial f}{\partial x_i}(x) = 0$, *$i = 1, \dots, n$. Then f is a $\mathcal{C}^1$ germ.*

Proof. We have to show that for every $\varepsilon > 0$, there exists $\delta > 0$ such that

$$\frac{|f(x) - f(0)|}{\|x\|} \leq \varepsilon, \qquad \forall\, x \neq 0,\ \|x\| \leq \delta.$$

Define the quantity $A(x) := \frac{f(x)-f(0)}{\|x\|}$ and write

$$|A(x)| \leq \frac{|f(x) - f(\theta(x)x)|}{\|x\|} + \frac{|f(\theta(x)x) - f(0)|}{\|x\|},$$

where $0 < \theta(x) \leq 1$ is small enough to satisfy $|f(\theta(x)x) - f(0)| \leq \frac{\varepsilon}{2}\|x\|$ (such a function exists by continuity of f). By assumption (ii), for every x, the function $g_x(t) := f(tx)$, $\theta(x) \leq t \leq 1$ is C^1 in its domain, and its derivative is computed

$$g_x'(t) = \sum_{i=1}^{n} x_i \frac{\partial f}{\partial x_i}(tx).$$

In particular

$$|f(x) - f(\theta(x)x)| = |g_x(1) - g_x(\theta(x))| \leq |g_x'(c_x)|(1 - \theta(x)), \quad \text{for some } c_x \in [\theta(x), 1]. \tag{4.52}$$

Now, by assumption (iii), choose δ small enough such that the uniform estimate holds

$$\left| \frac{\partial f}{\partial x_i}(cx) \right| \leq \frac{\varepsilon}{2\sqrt{n}}, \quad \forall\, c < 1, \quad \forall\, x,\ \|x\| \leq \delta. \tag{4.53}$$

and from (4.52) and (4.53) it follows

$$\begin{aligned} |A(x)| &\leq \frac{1 - \theta(x)}{\|x\|} |g_x'(c_x)| + \frac{\varepsilon}{2} \\ &\leq \frac{1}{\|x\|} |\left\langle \nabla f(c_x x), x \right\rangle| + \frac{\varepsilon}{2} \\ &\leq \|\nabla f(c_x x)\| + \frac{\varepsilon}{2} < \varepsilon. \end{aligned}$$

□

4.4.3 Transversality Lemmas

Let S be the set of (k, n) smooth sub-Riemannian metrics over M, equipped with the Whitney topology. Due to the $\mathcal{C}^\infty$ structure, we have the existence of smooth bump functions and the results in this section are essentially local. Then we can assume that S is the set of m-tuples $F = (f_1, \dots f_k)$ of smooth independent vector fields on some open subset M of $\mathbb{R}^n$, satisfying

$$T_q M = \Delta_q + [\Delta, \Delta]_q, \quad \text{for every } q \in M.$$

The vector fields $f_1, \dots, f_k$ form an orthonormal basis for the sub-Riemannian metric $\mathbf{g}$ they specify.

Let B be the bundle over M whose fiber at $q \in M$ is the variety of m-dimensional vector space of $\mathbf{g}$-skew symmetric endomorphisms of Δ_q.

Let us consider the mapping

$$\begin{array}{rcl} \rho : S \times M & \to & B \\ (F, q) & \to & \mathcal{L}_q^F \end{array}$$

where $\mathcal{L}_q^F$ has been defined in Remark 4.1. It is clear that ρ is $\mathcal{C}^\infty$.

Let us fix a point $F_0 \in S$, a point $q_0 \in M$ and coordinates (x, y) in M such that $q_0 = (0, 0)$ and the nilpotent approximation of F_0 reads in control form

$$\begin{cases} \dot{x}_i = u_i, \qquad i = 1, \dots, k, \\ \dot{y}_1 = \frac{1}{2} x' L_1 u, \\ \quad \vdots \\ \dot{y}_m = \frac{1}{2} x' L_m u, \end{cases} \tag{4.54}$$

In the coordinates y, the space $\mathcal{L}_q^{F_0}$ is the vector subspace spanned by the matrices $L_1, \dots L_k$. We have a natural gradation in formal power series of (x, y) induced by setting that the x_i have weight 1 and the y_i have weight 2. This induces a formal gradation on formal vector fields on M_{q_0} in which $\frac{\partial}{\partial x_i}$ have weight -1 and $\frac{\partial}{\partial y_i}$ have weight -2. The vector fields of the nilpotent approximation (4.54) have weight -1.

In control form the sub-Riemannian metric F_0 itself reads

$$\begin{pmatrix} \dot{x} \\ \dot{y}_1 \\ \vdots \\ \dot{y}_m \end{pmatrix} = \begin{pmatrix} u \\ \frac{1}{2} x' L_1 u \\ \vdots \\ \frac{1}{2} x' L_m u \end{pmatrix} + H \tag{4.55}$$

where H is a term of order > -1 as a formal u-dependent vector field. Then we take a smooth bump function $b(x, y)$ which is compactly supported in M and which is 1 in a neighborhood of $q_0 = (0, 0)$. We consider the affine space $\mathcal{A}$ of variations of F_0 of the form

$$\begin{pmatrix} \dot{x} \\ \dot{y}_1 \\ \vdots \\ \dot{y}_m \end{pmatrix} = \begin{pmatrix} u \\ \frac{1}{2} x' L_1 u \\ \vdots \\ \frac{1}{2} x' L_m u \end{pmatrix} + \begin{pmatrix} 0 \\ \frac{1}{2} x' \delta L_1 u \\ \vdots \\ \frac{1}{2} x' \delta L_m u \end{pmatrix} b(x, y) + H. \tag{4.56}$$

This defines new sub-Riemannian metrics $F_0 + \delta F$. Since S is open in the set of all rank k smooth sub-Riemannian metrics over M, then, for a small perturbation δF, we have $F_0 + \delta F \in S$.

To show that ρ_{q_0}, defined by $\rho_{q_0}(F) := \rho(q_0, F)$, is a submersion at q_0 on the fiber B_{q_0}, it is enough to observe that

$$\begin{array}{rcl} \widehat{\rho}_{q_0} : \mathcal{A} & \to & so(4)^m \\ \delta F & \mapsto & (L_1 + \delta L_1, \dots, L_k + \delta L_k) \end{array}$$

is an affine submersion. Then we have proven the following Lemma.

Lemma 4.28. *The map ρ is transversal to points*

Now let us restrict to the (4,6) case.

Definition 4.29. We say that a point $q_0 \in M$ is critical for a sub-Riemannian metric F if all elements of the subspace $\mathcal{L}_{q_0}^F$ (from Remark 4.1) have a double eigenvalue.

By formula 4.50, this means that, whatever the coordinates y, the matrices $L_1(q_0), L_2(q_0)$ both belongs either to $\mathcal{Q}$ or to $\widehat{\mathcal{Q}}$.

The dimension d_1 of the fiber of the bundle B is the dimension of the Grassmannian $G(2,6)$ of 2-subspaces of $\mathbb{R}^6$, i.e. $d_1 = 8$.

The dimension of the set of pairs L_1, L_2 that both belong to $\mathcal{Q}$ (respectively $\widehat{\mathcal{Q}}$), is the dimension d_2 of the Grassmannian $G(2,3)$, i.e $d_2 = 2$.

Let us define now the partially algebraic "wrong set" $W \subset B$ as follows: the fiber W_{q_0} is the set of 2-subspaces of the $\mathbf{g}$-skew symmetric endomorphisms of Δ_{q_0}, whose elements have a double eigenvalues. The codimension of W in B is $d_1 - d_2 = 6$.

The next Lemma follows from Lemma 4.28 and a non-compact version of Abraham's parametric transversality Theorems ([1]).

Lemma 4.30 ($(4,6)$ case)**.** *The set of sub-Riemannian metrics that have only isolated critical points is open and dense in S.*

CHAPTER 5

The sub-Laplacian and the heat equation

In this chapter we study the small time asymptotics for the heat kernel on a sub-Riemannian manifold, using a perturbative approach. We explicitly compute, in the case of a 3D contact structure, the first two coefficients of the small time asymptotics expansion of the heat kernel on the diagonal, expressing them in terms of the two basic functional invariants χ and κ defined in Chapter 2 for a 3D contact structure.

5.1 Introduction

The heat equation on a sub-Riemannian manifold is a natural model that describes a non isotropic diffusion process on a manifold. It is defined by the second order PDE

$$\frac{\partial}{\partial t}\psi(t,x) = \mathcal{L}_f\psi(t,x), \qquad \forall\, t>0,\, x\in M, \tag{5.1}$$

where $\mathcal{L}_f$ is the *sub-Riemannian Laplacian*, also called *sub-Laplacian*, which is a hypoelliptic, but not elliptic, second order differential operator. Locally this operator can be written in the form "sum of squares" plus a first order part

$$\mathcal{L}_f = \sum_{i=1}^{k} f_i^2 + a_i f_i, \qquad k<n.$$

where $f_1,\ldots,f_k$ is an orthonormal basis for the sub-Riemannian structure and $a_1,\ldots,a_k$ are suitable smooth coefficients (see Section 5.2 for a precise definition).

From the analytical point of view, these operators, and their parabolic counterpart $\partial_t - \mathcal{L}_f$, have been widely studied, starting from the well known work of Hörmander [67]. A probabilistic approach to hypoelliptic diffusion equation can be found in [21, 31, 74], where the existence of a smooth heat kernel for such equations is given.

On the other hand, a "geometric" definition of the Laplacian is needed if one want to find some relations between the analytical properties of the heat kernel (e.g. the small time asymptotics) and the geometric properties of the manifold, like in the Riemannian case (see [91, 28] for the relation between the heat kernel and the Riemannian curvature of the manifold, and [85] for a characterization of the cut locus via the heat kernel).

As it was pointed out in [6, 81], to have an intrinsic definition of the sub-Laplacian $\mathcal{L}_f$ (i.e. that depends only on the geometric structure) it is necessary to build an instrinsic volume for the structure.

In the sub-Riemannian case there are two intrinsic volumes that are defined, namely the Popp's volume (that is the analogue of the Riemannian volume form in Riemannian geometry) and the Hausdorff volume. In [4] it is proved that, starting from dimension 5, in general they are not proportional. On the other hand in the 3D contact case they coincide.

The existence of an asymptotic expansion for the heat kernel was proved, beside the classical Riemannian case, when the manifold is endowed with a time dependent Riemannian metric in [59], in the sub-Riemannian free case (when $n = k + \frac{k(k-1)}{2}$) in [42]. In [75, 26, 100] the general sub-Riemannian case is considered, using a probabilstic approach, obtaining different expansion depending on the fact that the points that are considered belong to the cut locus or not.

The same method was also applied in [27] to obtain the asymptotic expansion on the diagonal respectively. In particular it was proved that, for the sub-Riemannian heat kernel $p(t, x, y)$ the following expansion holds

$$p(t, x, x) \sim \frac{1}{t^{Q/2}}(a_0 + a_1 t + a_2 t^2 + \ldots + a_k t^k + O(t^{k+1})), \qquad \text{for} \quad t \to 0, \tag{5.2}$$

where Q denotes the Hausdorff dimension of M.

Besides these existence results, the geometric meaning of the coefficients of the expansion on the diagonal (and out of that) is far from being understood, even in the simplest 3D case, where the heat kernel has been computed explicitly in some cases of left-invariant structures on Lie groups (see [6, 22, 99]). In analogy to the Riemannian case, one would expect that the curvature tensor of the manifold and its derivatives appear in these expansions.

In this Chapter we compute the first terms (precisely a_0 and a_1, referring to (5.2)), for every 3D contact structure. Our main tool is the nilpotent approximation (or the symbol) of the sub-Riemannian structure. Under the regularity hypothesis, the metric tangent cone (in the Gromov sense) of sub-Riemannian structure $\mathbf{S} = (M, \Delta, \mathbf{g})$ at a point $q \in M$, is endowed itself with a left-invariant sub-Riemannian structure $\widehat{\mathbf{S}}_q$ on a so-called Carnot group (i.e. homogeneous nilpotent and simply connected Lie group) whose Lie algebra is generated by the nilpotent approximation of a basis of the Lie algebra of $\mathbf{S}$, as explained in Section 1.3.

Recall that, if $\mathbf{S}$ is a 3D contact sub-Riemannian manifold, then the nilpotent approximation $\widehat{\mathbf{S}}_q$ of $\mathbf{S}$ at every point $q \in M$ is isometric to the Heisenberg group, i.e. the sub-Riemannian structure on $\mathbb{R}^3$ (where coordinates are denoted by $q = (x, y, w)$) defined by the global orthonormal frame

$$\widehat{f}_1 = \partial_x + \frac{y}{2}\partial_w, \qquad \widehat{f}_2 = \partial_y - \frac{x}{2}\partial_w. \tag{5.3}$$

Using this approach, we reduce the problem of computing the small time asymptotics of the heat kernel of the original sub-Riemannian structure to the problem of studying the heat kernel of a family of approximating structures at the fixed time $t = 1$. In such a way we present the original structure (and as a consequence its heat kernel) as a perturbation of the nilpotent one. With the perturbative method then we compute the term a_0, which reflects the properties Heisenberg group case (which is the nilpotent approximation of every 3D contact structure), and the term a_1, where we find the local invariant κ.

The main result is stated as follows:

Theorem 5.1. *Let M be a 3D contact sub-Riemannian structure, with local invariants χ and κ. Let $p(t,x,y)$ denotes the heat kernel of the sub-Riemannian heat equation. Then the following small time asymptotic expansion hold*

$$p(t,x,x) \sim \frac{1}{16t^2}(1+\kappa(x)t+O(t^2)), \qquad \text{for} \quad t \to 0.$$

Notice that the Hausdorff dimension of a 3D contact structure is $Q = 4$.

5.2 The sub-Laplacian in a sub-Riemannian manifold

In this section we compute the intrinsic hypoelliptic Laplacian on a regular sub-Riemannian manifold $(M, \Delta, \mathbf{g})$, also called *sub-Laplacian*, writing its expression in a local orthonormal frame. In particular we find its explicit expression in the 3D contact sub-Riemannian case in terms of the structure constant appearing in (2.5).

The sub-Laplacian is the natural generalization of the Laplace-Beltrami operator $\mathcal{L}$ defined on a Riemannian manifold, that is $\mathcal{L}\phi = \operatorname{div}(\nabla\phi)$, where ∇ is the unique operator from $C^\infty(M)$ to $\operatorname{Vec}(M)$ satisfying

$$\mathbf{g}(\nabla\phi, X) = d\phi(X), \qquad \forall\, X \in \operatorname{Vec}(M).$$

Here $\mathbf{g}$ denotes the Riemannian metric, and the divergence of a vector field X is the unique function $\operatorname{div} X$ satisfying

$$L_X \mu = (\operatorname{div} X)\mu, \tag{5.4}$$

where μ is the Riemannian volume form and L_X denotes the Lie derivative.

In the sub-Riemannian case these definitions are replaced by the notions of horizontal gradient and of divergence with respect to the Popp measure, which is well defined in the regular case (see [6]).

Definition 5.2. Let M be a sub-Riemannian manifold and $\phi \in C^\infty(M)$. The *horizontal gradient* of ϕ is the unique horizontal vector field $\nabla\phi \in \overline{\Delta}$ that satisfies

$$\mathbf{g}(\nabla\phi, X) = d\phi(X), \qquad \forall\, X \in \overline{\Delta}, \tag{5.5}$$

Given a local orthonormal frame $\{f_1, \dots, f_k\}$ for the sub-Riemannian structure, it is easy to see that the horizontal gradient $\nabla\phi \in \overline{\Delta}$ of a function is computed as folows

$$\nabla\phi = \sum_{i=1}^{k} f_i(\phi) f_i, \qquad \phi \in C^\infty(M), \tag{5.6}$$

where the vector field acts on functions as a differential operator.

Notation. In what follows we will denote by $\mathcal{L}_f$ the sub-Laplacian associated to the sub-Riemannian structure defined by the local orthonormal frame $f_1, \dots, f_k$. Actually this definition does not depend on the choice of the orthonormal frame (see also Proposition 5.6).

In the sub-Riemannian regular case, even if there is no scalar product defined in T_qM, we can still define an intrinsic volume, called Popp volume, by means of the Lie bracket of the horizontal vector fields (see [6]). Here we recall a convenient definition of Popp measure only for the 3D contact case.

Definition 5.3. Let M be an orientable 3D contact sub-Riemannian structure and $\{f_1, f_2\}$ a local orthonormal frame. Let f_0 be the Reeb vector field and ν_0, ν_1, ν_2 the dual basis of 1-form, i.e. $\langle \nu_i, f_j \rangle = \delta_{ij}$. The *Popp volume* on M is the volume associated with the 3-form[1] $\mu := \nu_0 \wedge \nu_1 \wedge \nu_2$.

Using the formula

$$\operatorname{div}(aX) = Xa + a \operatorname{div} X, \qquad \forall\, a \in C^\infty(M),\ X \in \operatorname{Vec}(M), \tag{5.7}$$

it is easy to find the expression of the sub-Laplacian with respect to any volume

$$\begin{aligned} \operatorname{div}(\nabla\phi) &= \sum_{i=1}^k \operatorname{div}(f_i(\phi) f_i) \\ &= \sum_{i=1}^k f_i(f_i(\phi)) + (\operatorname{div} f_i) f_i(\phi). \end{aligned}$$

Thus

$$\mathcal{L}_f = \sum_{i=1}^k f_i^2 + (\operatorname{div} f_i) f_i. \tag{5.8}$$

Remark 5.4. Here we collect few properties of the sub-Laplacian that immediately follows from the definition:

(i) The sub-Laplacian is always presented as sum of squares of the horizontal vector fields plus a first order horizontal part (see (5.8)), whose coefficients heavily depends on the choice of the volume. Moreover $\mathcal{L}_f$ is the sum of squares if and only if all the vector fields of the orthonormal frame are divergence free.

(ii) From (5.8) and (5.7) it easily follows that the sub-Laplacian is a homogeneous differential operator of degree two with respect of dilations of the metric structure. More precisely, if we consider the dilated structure (denoted by λf) where all vector fields of the orthonormal frame are multiplied by a positive constant $\lambda > 0$, we have

$$\mathcal{L}_{\lambda f} = \lambda^2 \mathcal{L}_f. \tag{5.9}$$

(iii) Define the bilinear form in $C_0^\infty(M)$

$$(\phi, \varphi)_2 = \int_M \phi\varphi \, d\mu, \qquad \forall\, \phi, \varphi \in C_0^\infty(M)$$

Given a vector field $X \in \operatorname{Vec}(M)$, its formal adjoint X^* is the differential operator that satisfies the identity

$$(X\phi, \varphi)_2 = (\phi, X^*\varphi)_2, \qquad \forall\, \phi, \varphi \in C_0^\infty(M).$$

[1] In [6] the dual basis of the frame $\{f_1, f_2, [f_1, f_2]\}$ was considered to build the Popp volume. From (2.5) it follows that these two constructions agree each other.

It is easily computed that $X^* = -X - \operatorname{div} X$. In particular it follows that the sub-Laplacian $\mathcal{L}_f$ is rewritten as

$$\mathcal{L}_f = -\sum_{i=1}^{k} f_i^* f_i,$$

and that satisfies the identities

$$(\mathcal{L}_f \phi, \varphi)_2 = (\phi, \mathcal{L}_f \varphi)_2, \qquad (\phi, \mathcal{L}_f \phi)_2 \leq 0, \qquad \forall\, \phi, \varphi \in C_0^\infty(M). \tag{5.10}$$

From this we can explicitly compute the sub-Laplacian in the 3D contact case (see also [6], formula (8)).

Proposition 5.5. *Let M be a 3D sub-Riemannian manifold and f_1, f_2 be an orthonormal frame and f_0 be the Reeb vector field. Then the sub-Laplacian is expressed as follows*

$$\mathcal{L}_f = f_1^2 + f_2^2 + c_{12}^2 f_1 - c_{12}^1 f_2, \tag{5.11}$$

where c_{12}^1, c_{12}^2 are the structure constant appearing in (2.5).

Proof. Using (5.8), it is enough to compute the functions $a_i := \operatorname{div}_\mu f_i$, where $\mu = \nu_1 \wedge \nu_2 \wedge \nu_0$ is the Popp measure.

To this purpose let us compute the quantity $L_X \mu$, for every vector field X. Recall that the action of the Lie derivative on a differential 1-form is defined as

$$L_X \nu = \left.\frac{d}{dt}\right|_{t=0} e^{tX*}\nu, \qquad \nu \in \Lambda^1(M).$$

where e^{tX} denotes the flow on M generated by the vector field X. Using the fact that L_X is a derivation we get

$$L_X(\nu_0 \wedge \nu_1 \wedge \nu_2) = L_X\nu_0 \wedge \nu_1 \wedge \nu_2 + \nu_0 \wedge L_X\nu_1 \wedge \nu_2 + \nu_0 \wedge \nu_1 \wedge L_X\nu_2. \tag{5.12}$$

Moreover, for every $i = 0, 1, 2$, we can write

$$L_X \nu_i = \sum_{j=0}^{2} a_{ij} \nu_j,$$

The coefficients a_{ij} can be computed evaluating $L_X \nu_i$ on the dual basis

$$\begin{aligned} a_{ij} &= \langle L_X \nu_i, f_j \rangle \\ &= \left\langle \left.\frac{d}{dt}\right|_{t=0} e^{tX*}\nu_i, f_j \right\rangle \\ &= \left\langle \nu_i, \left.\frac{d}{dt}\right|_{t=0} e_*^{tX} f_j \right\rangle \\ &= \langle \nu_i, [f_j, X] \rangle . \end{aligned}$$

Plugging these coefficients into (5.12) we get

$$L_X \mu = (\langle \nu_1, [f_1, X] \rangle + \langle \nu_2, [f_2, X] \rangle + \langle \nu_0, [f_0, X] \rangle)\mu,$$

from which it follows

$$\operatorname{div} X = \langle \nu_1, [f_1, X] \rangle + \langle \nu_2, [f_2, X] \rangle + \langle \nu_0, [f_0, X] \rangle . \tag{5.13}$$

Using (2.5) it is easy to see that $\langle \nu_0, [f_0, X] \rangle = 0$ for every horizontal vector field. Thus, applying (5.13) with $X = f_i$, with $i = 1, 2$ one gets

$$\operatorname{div} f_1 = \langle \nu_2, [f_2, f_1] \rangle = c_{12}^2, \qquad \operatorname{div} f_2 = \langle \nu_1, [f_1, f_2] \rangle = -c_{12}^1.$$

Then (5.11) easily follows from (5.8). □

From the above construction it is clear that the sub-Laplacian depends only on the sub-Riemannian structure and not on the frame selected, i.e. it is invariant for rotations of the orthonormal frame. Here we give also a direct proof of this fact.

Proposition 5.6. *The sub-Laplacian is invariant with respect to rotation of the orthonormal frame.*

Proof. Let us consider an orthonormal frame f_1, f_2 and the rotated one

$$\widetilde{f}_1 = \cos\theta f_1 + \sin\theta f_2, \tag{5.14}$$

$$\widetilde{f}_2 = -\sin\theta f_1 + \cos\theta f_2, \tag{5.15}$$

where $\theta = \theta(q)$ is a smooth function on M. If we denote by $\widetilde{c}_{ij}^k$ the structure constants computed in the rotated frame, from the formula

$$[\widetilde{f}_1, \widetilde{f}_2] = [f_1, f_2] - f_1(\theta) f_1 - f_2(\theta) f_2, \tag{5.16}$$

it is easy to prove that the new structure constant are computed according to the formulas

$$\widetilde{c}_{12}^1 = \cos\theta(c_{12}^1 - f_1(\theta)) - \sin\theta(c_{12}^2 - f_2(\theta)),$$
$$\widetilde{c}_{12}^2 = \sin\theta(c_{12}^1 - f_1(\theta)) + \cos\theta(c_{12}^2 - f_2(\theta)).$$

From these relations one can easily compute

$$\widetilde{f}_1^2 + \widetilde{f}_2^2 = f_1^2 + f_2^2 + f_1(\theta) f_2 - f_2(\theta) f_1, \tag{5.17}$$

and

$$-\widetilde{c}_{12}^1 \widetilde{f}_2 + \widetilde{c}_{12}^2 \widetilde{f}_1 = -(c_{12}^1 - f_1(\theta)) f_2 + (c_{12}^2 - f_2(\theta)) f_1. \tag{5.18}$$

Combining (5.17) and (5.18) one gets, denoting $\mathcal{L}_{\widetilde{f}}$ the laplacian defined by the rotated frame,

$$\begin{aligned}
\mathcal{L}_{\widetilde{f}} &= \widetilde{f}_1^2 + \widetilde{f}_2^2 - \widetilde{c}_{12}^1 \widetilde{f}_2 + \widetilde{c}_{12}^2 \widetilde{f}_1 \\
&= f_1^2 + f_2^2 + f_1(\theta) f_2 - f_2(\theta) f_1 - \widetilde{c}_{12}^1 \widetilde{f}_2 + \widetilde{c}_{12}^2 \widetilde{f}_1 \\
&= f_1^2 + f_2^2 + f_1(\theta) f_2 - f_2(\theta) f_1 - (c_{12}^1 - f_1(\theta)) f_2 + (c_{12}^2 - f_2(\theta)) f_1 \\
&= f_1^2 + f_2^2 - c_{12}^1 f_2 + c_{12}^2 f_1 = \mathcal{L}_f.
\end{aligned}$$

□

Notice from (5.17) that the sum of squares is not an intrinsic operator of the sub-Riemannian structure.

Remark 5.7. The same argument provides a proof of the fact that, on a 2-dimensional Riemannian manifold M with local orthonormal frame f_1, f_2 that satisfies

$$[f_1, f_2] = a_1 f_1 + a_2 f_2,$$

the Laplace-Beltrami operator is locally expressed as

$$\mathcal{L} = f_1^2 + f_2^2 + a_1 f_2 - a_2 f_1.$$

5.3 Nilpotent approximation and normal coordinates

Given a regular sub-Riemannian manifold $\mathbf{S} = (M, \Delta, \mathbf{g})$ and a local orthonormal frame $\{f_1, \ldots, f_k\}$ near a point q we can consider a system of privileged coordinates near q. As usual we denote by $\widehat{\mathbf{S}}_q$ the *nilpotent approximation of* $\mathbf{S}$ *near* q, i.e. the sub-Riemannian structure on $\mathbb{R}^n$ having $\{\widehat{f}_1, \ldots, \widehat{f}_k\}$ as an orthonormal frame, where $\widehat{f}_i := (\psi_* f_i)^{(-1)}$ (see Section 1.3) .

Under the regularity assumption, $\widehat{\mathbf{S}}_q$ is naturally endowed with a Lie group structure whose Lie algebra is generated by left-invariant vector fields $\widehat{f}_1, \ldots, \widehat{f}_k$. Moreover the sub-Riemannian distance $\widehat{d}$ in $\widehat{\mathbf{S}}_q$ is homogeneous with respect to dilations δ_t, i.e. $\widehat{d}(\delta_t(x), \delta_t(y)) = t\,\widehat{d}(x, y)$. In particular, if $\widehat{B}_q(r)$ denotes the ball of radius r in $\widehat{\mathbf{S}}_q$, this implies $\delta_t(\widehat{B}_q(1)) = \widehat{B}_q(t)$.

The following Lemma shows in which sense the nilpotent approximation is the first order approximation of the sub-Riemannian structure.

Lemma 5.8. *Let M be a sub-Riemannian manifold and $X \in \mathrm{Vec}(M)$. Fixed a system of privileged coordinates, we define $X^\varepsilon := \varepsilon \delta_{\frac{1}{\varepsilon}*} X$. Then*

$$X^\varepsilon = \widehat{X} + \varepsilon Y^\varepsilon, \qquad \textit{where } Y^\varepsilon \textit{ is smooth w.r.t. } \varepsilon.$$

Proof. Since we work in a system of priviliged coordinates, in the homogeneous expansion of X only terms of order ≥ 1 appear. Hence we can write

$$X \simeq X^{(-1)} + X^{(0)} + X^{(1)} + \ldots$$

Applying the dilation and using property (**??**) we get

$$\delta_{\frac{1}{\varepsilon}*} X \simeq \frac{1}{\varepsilon} X^{(-1)} + X^{(0)} + \varepsilon X^{(1)} + \ldots$$

Multiplying by ε and using that, by definition, $\widehat{X} = X^{(-1)}$, we have

$$X^\varepsilon \simeq \widehat{X} + \varepsilon X^{(0)} + \varepsilon^2 X^{(1)} + \ldots \tag{5.19}$$

□

In other words the nilpotent approximation of a vector field at a point q is the first meaningful term that appears in the expansion when one consider the blow up coordinates near the point q, with rescaled distances.

Remark 5.9. If $\mathbf{S} = (M, \Delta, \mathbf{g})$ is a 3D contact sub-Riemannian manifold, then $\dim \Delta_q = 2$ and $\dim \Delta_q^2 = 3$ for all $q \in M$. Under this assumption the nilpotent approximation $\widehat{\mathbf{S}}_q$ of $\mathbf{S}$ at every point $q \in M$ is isometric to the Heisenberg group, since this is the only nilpotent left-invariant structure with $\mathcal{G}(\mathbf{S}) = (2,3)$ (see e.g. [2] for a classification of left-invariant structures on 3D Lie groups).

The sub-Riemannian structure on the Heisenberg group is defined by the global orthonormal frame on $\mathbb{R}^3$ (where coordinates are denoted by $q = (x, y, w)$)

$$\widehat{f}_1 = \partial_x + \frac{y}{2}\partial_w, \qquad \widehat{f}_2 = \partial_y - \frac{x}{2}\partial_w. \tag{5.20}$$

Notice that the Lie algebra $\mathrm{Lie}\{\widehat{f}_1, \widehat{f}_2\}$ is nilpotent since

$$[\widehat{f}_1, [\widehat{f}_1, \widehat{f}_2]] = [\widehat{f}_2, [\widehat{f}_1, \widehat{f}_2]] = 0.$$

Moreover the Reeb vector field is $\widehat{f}_0 = \partial_w$ and the local invariants of the structure are identically zero $\chi = \kappa = 0$.

5.3.1 Normal coordinates

In the 3D contact case there exists a smooth normal form of the sub-Riemannian structure (i.e. of its orthonormal frame) which is the analogue of normal coordinates in Riemannian geometry. This normal form is crucial for the study of the heat kernel of the sub-Laplacian with a perturbative approach, since it presents the sub-Riemannian structure of a general 3D contact case as a perturbation of the Heisenberg (nilpotent) case.

Theorem 5.10 ([9, 12])**.** *Let M be a 3D contact sub-Riemannian manifold and f_1, f_2 a local orthonormal frame. There exists a smooth coordinate system (x, y, w) such that*

$$f_1 = (\partial_x + \frac{y}{2}\partial_w) + \beta y(y\partial_x - x\partial_y) + \gamma y \partial_w,$$
$$f_2 = (\partial_y - \frac{x}{2}\partial_w) - \beta x(y\partial_x - x\partial_y) + \gamma x \partial_w,$$

where $\beta = \beta(x, y, w)$ and $\gamma = \gamma(x, y, w)$ are smooth functions that satisfy the following boundary conditions

$$\beta(0,0,w) = \gamma(0,0,w) = \frac{\partial \gamma}{\partial x}(0,0,w) = \frac{\partial \gamma}{\partial y}(0,0,w) = 0.$$

Remark 5.11. Notice that the normal coordinate system is privileged at 0. Indeed from the explicit expression of the frame it immediately follows that these coordinates are linearly adapted at 0 since

$$\Delta_0 = \mathrm{span}\{f_1(0), f_2(0)\} = \mathrm{span}\{\partial_x, \partial_y\} = \mathbb{R}^2.$$

Moreover the weights of the coordinates (x, y, w) at the origin are

$$\nu(x) = \nu(y) = 1, \qquad \nu(w) = 2,$$

and every homogeneous term of the vector fields f_1, f_2 has degree ≥ -1.

Finally, notice also that, when $\beta = \gamma = 0$, we recover the Heisenberg group structure (5.20).

5.4 Perturbative method

In this section we consider the sub-Riemannian heat equation associated to a sub-Riemannian structure f on a complete sub-Riemannian manifold M, i.e. the initial value problem

$$\begin{cases} \dfrac{\partial \psi}{\partial t}(t,x) = \mathcal{L}_f \psi(t,x), & \text{in } (0,\infty) \times M, \\ \psi(0,x) = \varphi(x), & x \in M, \quad \varphi \in C_0^\infty(M). \end{cases} \tag{5.21}$$

where $\psi(0,x) = \lim_{t \to 0} \psi(t,x)$ and the limit is meant in the distributional sense.

Recall that a differential operator $\mathcal{L}$ is said to be hypoelliptic on a subset $U \subset M$ if every distributional solution to $\mathcal{L}u = \phi$ is $C^\infty(U)$, whenever $\phi \in C^\infty(U)$. The following well-known Hörmander Theorem gives a sufficient condition for the hypoellipticity of a second order differential operator.

Theorem 5.12 (Hörmander,[67])**.** *Let $\mathcal{L}$ be a differential operator on a manifold M, that locally in a neighborhood U is written as*

$$\mathcal{L} = \sum_{i=1}^{k} X_i^2 + X_0,$$

where $X_0, X_1, \dots, X_k \in \mathrm{Vec}(M)$. If $\mathrm{Lie}_q\{X_0, X_1, \dots, X_k\} = T_q M$ for all $q \in U$, then $\mathcal{L}$ is hypoelliptic.

From this Theorem and the bracket generating condition it follows that the sub-laplacian $\mathcal{L}_f$ is hypoelliptic. Moreover, since M is complete and the sub-Laplacian is symmetric and negative with respect to the Popp's measure (see (5.10)), it follows that $\mathcal{L}_f$ is essentially self-adjoint on $C_0^\infty(M)$ (see also (iii) in Remark 5.4).

As a consequence the operator $\mathcal{L}_f$ admits a unique self-adjoint extension on $L^2(M)$ and the heat semigroup $\{e^{t\mathcal{L}_f}\}_{t \geq 0}$ is a well-defined one parametric family of bounded operators on $L^2(M)$. Moreover the heat semigroup is contractive on $L^2(M)$ (see [96]).

The problem (5.21) has a unique solution, for every initial datum $\varphi \in L^2(M)$, namely $\psi(t,x) := e^{t\mathcal{L}_f}\varphi$. Moreover, due to the hypoellipticity of $\mathcal{L}_f$, the function $(t,x) \mapsto e^{t\mathcal{L}_f}\varphi(x)$ is smooth on $(0,\infty) \times M$ and

$$e^{t\mathcal{L}_f}\varphi(x) = \int_M p(t,x,y)\varphi(y)dy, \qquad \varphi \in C_0^\infty(M),$$

where $p(t,x,y)$ is the so-called *heat kernel* associated to $e^{t\mathcal{L}_f}$, that satisfies the following properties

(i) $p(t,x,y) \in C^\infty(\mathbb{R}^+ \times M \times M)$,

(ii) $p(t,x,y) = p(t,y,x), \quad \forall\, t > 0, \forall\, x,y \in M$,

(iii) $p(t,x,y) > 0, \quad \forall\, t > 0, \forall\, x,y \in M$.

A probabilistic approach to hypoelliptic diffusion equation can be found in [21, 31, 74], where the existence of a smooth heat kernel for such equations is given.

In particular, since $e^{t\mathcal{L}_f}\varphi$ satisfies the initial condition, it holds

$$\lim_{t\to 0}\int_M p(t,x,y)\varphi(y)dy = \varphi(x).$$

and the heat kernel $p(t,x,y)$ is a solution of the problem (5.21) with the initial condition $\psi(0,x)=\delta_y(x)$, where δ_y denotes the Dirac delta function.

For a more detailed discussion on the analytical properties of the sub-Riemannian heat equation and its heat kernel one can see [96, 23].

To study the asymptotics of the heat kernel associated to the sub-Riemannian structure defined by f near a point $q \in M$, we consider the approximation of the sub-Riemannian structure (cfr. also Lemma 5.8).

Definition 5.13. Let $f_1,\dots,f_k$ be an orthonormal frame for a sub-Riemannian structure on M and fix a system of privileged coordinates around the point $q \in M$. The ε-approximated system at q is the sub-Riemannian structure induced by the orthonormal frame $f_1^\varepsilon,\dots,f_k^\varepsilon$ defined by

$$f_i^\varepsilon := \varepsilon\delta_{1/\varepsilon *}f_i, \qquad i=1,\dots,k.$$

Remark 5.14. Notice that in the definition of approximated structure we have to perturbate the basis of the distribution only. Their Lie brackets are changed accordingly to the formula

$$[f_i^\varepsilon, f_j^\varepsilon] = \varepsilon^2\delta_{1/\varepsilon *}[f_i,f_j].$$

In particular, in the 3D contact case, the Reeb vector field f_0^ε of the ε-approximated structure is related to the unperturbed one by $f_0^\varepsilon = \varepsilon^2\delta_{1/\varepsilon *}f_0$.

The following Lemma shows the relation between the heat kernel associated with f and the one defined by f^ε, whose existence is guaranteed by the above results.

Lemma 5.15. *Let M be a sub-Riemannian manifold and fix a set of privileged coordinates in a neighborhood N of q. Denote by f the sub-Riemannian structure and by f^ε its ε-approximation at q. If we denote by $p(t,x,y)$ and $p^\varepsilon(t,x,y)$ the heat kernels respectively of the equations*

$$\frac{\partial\psi}{\partial t}(t,x) = \mathcal{L}_f\psi(t,x), \qquad \frac{\partial\psi}{\partial t}(t,x) = \mathcal{L}_{f^\varepsilon}\psi(t,x),$$

we have that

$$p^\varepsilon(t,x,y) = \varepsilon^Q p(\varepsilon^2 t, \delta_\varepsilon x, \delta_\varepsilon y), \qquad \forall\, x,y \in N,$$

where Q denotes the Hausdorff dimension of M.

Proof. Recall that, if $f_1,\dots,f_k$ is an orthonormal frame for the sub-Riemannian structure, the approximated system is defined by vector fields

$$f_i^\varepsilon := \varepsilon\delta_{1/\varepsilon *}f_i, \qquad i=1,\dots,k.$$

Then the proof follows from the following facts

(i). If we perform the change of coordinates $x' = \delta_{1/\varepsilon}x$ and we denote by $q^\varepsilon(t,x,y)$ the heat kernel written in the new coordinate system (which depends on ε), we have the equality

$$q^\varepsilon(t,x,y) = \varepsilon^Q p(t,\delta_\varepsilon x,\delta_\varepsilon y)$$

Indeed it is easy to see that $q^\varepsilon(t,x,y)$ is a solution of the equation written in the new variables $x' = \delta_{1/\varepsilon}x$ and the factor ε^Q comes from the fact that $|\det\delta_\varepsilon| = \varepsilon^Q$. More precisely if we set

$$\psi^\varepsilon(t,x) = \int_M q^\varepsilon(t,x,y)\varphi^\varepsilon(y)dy,$$

where $\varphi^\varepsilon(x) = \varphi(\delta_\varepsilon x)$, one can see that the initial condition is satisfied:

$$\begin{aligned}\lim_{t\to 0}\int_M q^\varepsilon(t,x,y)\varphi^\varepsilon(y)dy &= \lim_{t\to 0}\int_M \varepsilon^Q p(t,\delta_\varepsilon x,\delta_\varepsilon y)\varphi(\delta_\varepsilon y)dy\\ &= \lim_{t\to 0}\int_M p(t,\delta_\varepsilon x,z)\varphi(z)dz \qquad (z=\delta_\varepsilon y)\\ &= \varphi(\delta_\varepsilon x) = \varphi^\varepsilon(x).\end{aligned}$$

(ii). Since $\delta_{\frac{1}{\varepsilon}*}f_i = \frac{1}{\varepsilon}f_i^\varepsilon$, for every $i=1,\dots,k$, using (5.9) we get

$$\frac{\partial\psi}{\partial t} = \mathcal{L}_{\frac{1}{\varepsilon}f^\varepsilon}\psi = \frac{1}{\varepsilon^2}\mathcal{L}_{f^\varepsilon}\psi.$$

This equality can be rewritten as

$$\varepsilon^2\frac{\partial\psi}{\partial t} = \mathcal{L}_{f^\varepsilon}\psi,$$

and performing the change of variable $t=\varepsilon^2\tau$ we get

$$\frac{\partial\psi}{\partial\tau} = \mathcal{L}_{f^\varepsilon}\psi,$$

from which it follows

$$p^\varepsilon(t,x,y) = q^\varepsilon(\varepsilon^2 t,x,y) = \varepsilon^Q p(\varepsilon^2 t,\delta_\varepsilon x,\delta_\varepsilon y).$$

□

This result is useful for the study of the asymptotics on the diagonal since, by construction, the intial point is fixed for the dilation δ_ε. In particular we can recover the small time behaviour of the original heat kernel from the approximated one

Corollary 5.16. *Following the notations introduced above*

$$p^\varepsilon(1,0,0) = \varepsilon^Q p(\varepsilon^2,0,0).$$

As a corollary we recover an homogeneity property of the heat kernel for a nilpotent structure.

Corollary 5.17. *Assume that the sub-Riemannian structure is regular and nilpotent. Then the heat kernel $p(t,x,y)$ of the heat equation satisfies the homogeneity property*

$$p(t,x,y) = \lambda^Q p(\lambda^2 t,\delta_\lambda x,\delta_\lambda y), \qquad \forall\,\lambda>0.$$

This result was already well-known (see e.g. [56]). In our notations, it is a direct consequence of the fact that, if $f=\widehat{f}$, then $f = f^\varepsilon = \widehat{f}$ for every $\varepsilon>0$.

5.4.1 General method

In this section we briefly recall the perturbative method for the heat equation presented in Chapter 3 of [91]. For further discussions one can see also [28, 52] and [101].

Let X, Y be operators on a Hilbert space of functions. Moreover assume that X and $X+Y$ (where Y is treated as a perturbation of X) have well defined heat operators e^{tX}, $e^{t(X+Y)}$, i.e. a semigroup of one parameter family of bounded self-adjoint operators satisfying

$$(\partial_t - X)e^{tX}\varphi = 0, \qquad \lim_{t\to 0} e^{tX}\varphi = \varphi,$$

and similarly for $X+Y$. Given $A(t), B(t)$ two operators on the Hilbert space, if we denote their convolution as

$$(A * B)(t) = \int_0^t A(t-s)B(s)ds,$$

then the classical Duhamel formula

$$e^{t(X+Y)} = e^{tX} + \int_0^t e^{(t-s)(X+Y)}Ye^{sX}ds,$$

can be rewritten as follows

$$e^{t(X+Y)} = e^{tX} + e^{t(X+Y)} * Ye^{tX}.$$

Iterating this construction one gets the expansion

$$e^{t(X+Y)} = e^{tX} + e^{tX} * Ye^{tX} + e^{t(X+Y)} * (Ye^{tX})^{*2}. \tag{5.22}$$

where $A^{*2} = A * A$ denotes the iterated convolution product.

If $A(t)$ and $B(t)$ have heat kernels $a(t,x,y)$ and $b(t,x,y)$ respectively, then $(A * B)(t)$ has kernel (see again [91, 101])

$$(a * b)(t,x,y) := \int_0^t \int_M a(s,x,z)b(t-s,z,y)dzds,$$

Interpreting (5.22) at the level of kernels, denoting by $p(t,x,y)$ the heat kernel for the operator X and by $p^Y(t,x,y)$ the kernel of the perturbed operator $X+Y$ we can write the expansion

$$p^Y(t,x,y) = p(t,x,y) + (p * Yp)(t,x,y) + (p^Y * Yp * Yp)(t,x,y) \tag{5.23}$$

5.5 Proof of Theorem 5.1

In this section we compute the first terms of the small time asymptotics of the heat kernel. To this extent we compute the sub-Laplacian associated to the approximated sub-Riemannian structure and we use the pertubative method of Section 5.4.1 to compute this terms using the explicit expression of the heat kernel in the Heisenberg group.

Remark 5.18. The sub-Laplacian on the Heisenberg group H^3 is written as the sum of squares (cfr. also Remarks 5.4 and 5.9)

$$\mathcal{L}_{\widehat{f}} = \widehat{f}_1^2 + \widehat{f}_2^2 = (\partial_x - \frac{y}{2}\partial_w)^2 + (\partial_y + \frac{x}{2}\partial_w)^2.$$

The heat kernel for $\Delta_{\widehat{f}}$ has been computed explicitly for the first time in [63]. Here we use the expression given in [6] in the same coordinate set. Denote by $q = (x, y, w) \in \mathbb{R}^3$ a point in the Heisenberg group. The heat kernel $H(t, q, q')$, is presented as

$$H(t, q, q') = h_t(q' \circ q^{-1}), \tag{5.24}$$

where

$$h_t(x, y, w) = \frac{1}{2(2\pi t)^2} \int_{\mathbb{R}} \frac{s}{\sinh s} \exp\left(-\frac{s(x^2 + y^2)}{4t \tanh s}\right) \cos(\frac{ws}{t}) ds, \tag{5.25}$$

and $\circ$ denotes the group law in H_3

$$(x, y, w) \circ (x', y', w') = (x + x', y + y', w + w' + \frac{1}{2}(x'y - xy')).$$

Notice that the inverse of an element with respect to $\circ$ is

$$(x, y, w)^{-1} = (-x, -y, -w).$$

For a discussion on the convergence of the integral (5.25) one can see [24].

5.5.1 Local invariants

In this section we compute the invariants χ and κ at the origin of the sub-Riemannian manifold. By Theorem 5.10 we can assume that the orthonormal frame has the form

$$\begin{aligned} f_1 &= (\partial_x + \frac{y}{2}\partial_w) + \beta y(y\partial_x - x\partial_y) + \gamma y \partial_w, \\ f_2 &= (\partial_y - \frac{x}{2}\partial_w) - \beta x(y\partial_x - x\partial_y) + \gamma x \partial_w, \end{aligned} \tag{5.26}$$

where β and γ are smooth functions near $(0, 0, 0)$ that satisfy

$$\beta(0, 0, w) = \gamma(0, 0, w) = \frac{\partial \gamma}{\partial x}(0, 0, w) = \frac{\partial \gamma}{\partial y}(0, 0, w) = 0. \tag{5.27}$$

Moreover, since we are interested up to second order terms in the expansion (5.29), we can assume the following

Lemma 5.19. *We can assume that the orthonormal frame has the form*

$$\begin{aligned} f_1 &= \partial_x - \frac{y}{2}(1 + \gamma)\partial_w, \\ f_2 &= \partial_y + \frac{x}{2}(1 + \gamma)\partial_w, \end{aligned} \tag{5.28}$$

where γ is a quadratic polynomial of the form $\gamma(x, y) = ax^2 + bxy + cy^2$, for some $a, b, c \in \mathbb{R}$.

Proof. Recall that if we expand a vector field X in homogeneous components (when written in a privileged coordinate system)

$$X \simeq X^{(-1)} + X^{(0)} + X^{(1)} + X^{(2)} + \dots$$

its ε-approximation X^ε has the following expansion

$$X^\varepsilon \simeq \widehat{X} + \varepsilon X^{(0)} + \varepsilon^2 X^{(1)} + \varepsilon^3 X^{(2)} + \dots \tag{5.29}$$

It is then sufficient to consider, in the Taylor expansion of the orthonormal frame near the origin, only the homogeneous term up to weight one, since every other term with weight ≥ 2 gives a contribution $o(\varepsilon^2)$ when one compute the expansion of $\mathcal{L}_{f^\varepsilon}$. Hence a contribution $o(\varepsilon^2)$ in the heat kernel due to (5.23).

Moreover the boundary condition (5.27) implies that the following derivatives of the coefficients of (5.26) vanish at the origin

$$\frac{\partial \beta}{\partial w}(0,0,0) = \frac{\partial \gamma}{\partial w}(0,0,0) = \frac{\partial^2 \gamma}{\partial w \partial x}(0,0,0) = \frac{\partial^2 \gamma}{\partial w \partial y}(0,0,0) = 0, \tag{5.30}$$

together with all higher order derivatives with respect to w.

Since $\nu(\partial_x) = \nu(\partial_y) = 1$ and $\beta(0,0,0) = 0$, the terms $\beta y(y\partial_x - x\partial_y)$ and $\beta x(y\partial_x - x\partial_y)$ have weight ≥ 2. Moreover $\nu(\partial_w) = -2$ implies that $x\partial_w$ and $y\partial_w$ have weight -1. The only terms that we need in the expansion of γ are those of weight less or equal than one. Since the terms of order zero vanish by (5.27) and (5.30), the only meaningful term in the expansion of γ is

$$\gamma(x,y,w) \sim \frac{\partial^2 \gamma}{\partial x^2} x^2 + \frac{\partial^2 \gamma}{\partial xy} xy + \frac{\partial^2 \gamma}{\partial y^2} y^2$$

where derivatives are computed at the origin (0,0,0). □

Now we express the invariants χ and κ in terms of the perturbation (5.28).

Lemma 5.20. *Assume that the orthonormal frame of the sub-Riemannian structure has the form* (5.28). *Then value of the invariants at the origin are*

$$\chi = 2\sqrt{b^2 + (c-a)^2}, \qquad \kappa = 2(a+c). \tag{5.31}$$

Proof. To compute the invariants we need to compute the Reeb vector field f_0 and the structure constant of the Lie algebra $\mathrm{Lie}\{f_0, f_1, f_2\}$. Every contact form for the structure is a multiple of

$$\widetilde{\omega} = dz - \frac{x}{2}(1+\gamma)dy + \frac{y}{2}(1+\gamma)dx,$$

whose differential is computed as follows

$$d\widetilde{\omega} = -(1+2\gamma)dxdy.$$

Since $d\widetilde{\omega}(f_1, f_2) = -(1+2\gamma)$, the normalized contact form ω that satisfies $d\omega(f_1, f_2) = 1$ (see Remark **??** and (2.3)) is

$$\omega := -\frac{1}{1+2\gamma}\widetilde{\omega} = -\frac{1}{1+2\gamma}(dz - \frac{x}{2}(1+\gamma)dy + \frac{y}{2}(1+\gamma)dx).$$

Notice that every contact form vanishes on the distribution. Thus

$$d(\phi\omega)(f_1, f_2) = \phi\, d\omega(f_1, f_2), \qquad \forall\, \phi \in C^\infty(M).$$

Next we compute the differential of the normalized contact form

$$\begin{aligned} d\omega &= \frac{d(1+2\gamma)\wedge\widetilde{\omega}}{(1+2\gamma)^2} + dxdy \\ &= \frac{2\partial_x\gamma}{(1+2\gamma)^2}dxdw + \frac{2\partial_y\gamma}{(1+2\gamma)^2}dydw + \left(1 - \frac{2\gamma(1+\gamma)}{(1+2\gamma)^2}\right)dxdy, \end{aligned}$$

The Reeb vector field is, by definition, the kernel of $d\omega$ (normailzed in such a way that $\omega(f_0) = 1$). From this one gets

$$f_0 = \frac{2\partial_x\gamma}{1+2\gamma}\partial_y - \frac{2\partial_y\gamma}{1+2\gamma}\partial_x + \left(\frac{2\gamma(1+\gamma)}{1+2\gamma} - (1+2\gamma)\right)\partial_w.$$

The commutator between horizontal vector fields is computed as follows

$$\begin{aligned} [f_2, f_1] &= \left[\partial_y + \frac{x}{2}(1+\gamma)\partial_w, \partial_x - \frac{y}{2}(1+\gamma)\partial_w\right] \\ &= -(1+2\gamma)\partial_w, \end{aligned}$$

and writing $[f_2, f_1] = f_0 + c_{12}^1 f_1 + c_{12}^2 f_2$ we find the structure constants

$$c_{12}^1 = \frac{2\partial_y\gamma}{1+2\gamma}, \qquad c_{12}^2 = -\frac{2\partial_x\gamma}{1+2\gamma}.$$

Moreover, a longer computation for $[f_1, f_0]$ and $[f_2, f_0]$ shows that

$$c_{0i}^j = -\frac{2}{(1+2\gamma)^2}\widetilde{c}_{0i}^j,$$

where we set

$$\begin{aligned} \widetilde{c}_{01}^1 &= (1+2\gamma)\partial_{xy}\gamma - 2\partial_y\gamma\partial_x\gamma, \\ \widetilde{c}_{01}^2 &= -(1+2\gamma)\partial_{xx}\gamma + 2(\partial_x\gamma)^2, \\ \widetilde{c}_{02}^1 &= (1+2\gamma)\partial_{yy}\gamma - 2(\partial_y\gamma)^2, \\ \widetilde{c}_{02}^2 &= -(1+2\gamma)\partial_{xy}\gamma + 2\partial_y\gamma\partial_x\gamma. \end{aligned}$$

Recalling that at the origin $\partial_x\gamma = \partial_y\gamma = 0$, while $\partial_{xx}\gamma = 2a, \partial_{xy}\gamma = b, \partial_{yy}\gamma = 2c$ it follows from (2.10) that

$$\chi = 2\sqrt{-\det\begin{pmatrix} b & c-a \\ c-a & -b \end{pmatrix}} = 2\sqrt{b^2 + (c-a)^2}.$$

and

$$\kappa = f_2(c_{12}^1) - f_1(c_{12}^2) - (c_{12}^1)^2 - (c_{12}^2)^2 + \frac{c_{01}^2 - c_{02}^1}{2} = 2(a+c).$$

□

5.5.2 Asymptotics

In this section we compute the Laplacian $\mathcal{L}_{f^\varepsilon}$ up to second order in ε. First notice that

$$f_1^\varepsilon = (\partial_x - \frac{y}{2}\partial_w) - \varepsilon^2(\frac{y}{2}\gamma\partial_w) + o(\varepsilon^2),$$
$$f_2^\varepsilon = (\partial_y + \frac{x}{2}\partial_w) + \varepsilon^2(\frac{x}{2}\gamma\partial_w) + o(\varepsilon^2).$$

Moreover, defining $f_0^\varepsilon := \varepsilon^2\delta_{\frac{1}{\varepsilon}*}f_0$ (see Remark 5.14), from the formula

$$[f_2^\varepsilon, f_1^\varepsilon] = f_0^\varepsilon + (c_{12}^1)^\varepsilon f_1^\varepsilon + (c_{12}^2)^\varepsilon f_2^\varepsilon,$$

we get the following expansion

$$(c_{12}^1)^\varepsilon = \varepsilon\frac{2\varepsilon(\partial_y\gamma)}{1+2\varepsilon^2\gamma} = 2\varepsilon^2\partial_y\gamma + o(\varepsilon^2), \tag{5.32}$$
$$(c_{12}^2)^\varepsilon = -\varepsilon\frac{2\varepsilon(\partial_x\gamma)}{1+2\varepsilon^2\gamma} = -2\varepsilon^2\partial_x\gamma + o(\varepsilon^2). \tag{5.33}$$

Thus we can compute every term defining the sub-Laplacian

$$\begin{aligned}(f_1^\varepsilon)^2 &= (\partial_x - \frac{y}{2}\partial_w)^2 - \varepsilon^2\Big((\partial_x - \frac{y}{2}\partial_w)(\frac{y}{2}\gamma\partial_w) + (\frac{y}{2}\gamma\partial_w)(\partial_x - \frac{y}{2}\partial_w)\Big) + o(\varepsilon^4)\\ &= (\widehat{f}_1)^2 - \varepsilon^2(y\gamma\partial_{wx} - \frac{y^2}{2}\gamma\partial_w^2 + \frac{y}{2}\partial_x\gamma\partial_w) + o(\varepsilon^4),\end{aligned}$$

$$\begin{aligned}(f_2^\varepsilon)^2 &= (\partial_y + \frac{x}{2}\partial_w)^2 + \varepsilon^2\Big((\partial_y + \frac{x}{2}\partial_w)(\frac{x}{2}\gamma\partial_w) + (\frac{x}{2}\gamma\partial_w)(\partial_y + \frac{x}{2}\partial_w)\Big) + o(\varepsilon^4)\\ &= (\widehat{f}_2)^2 + \varepsilon^2(x\gamma\partial_{wx} + \frac{x^2}{2}\gamma\partial_w^2 + \frac{x}{2}\partial_y\gamma\partial_w) + o(\varepsilon^4),\end{aligned}$$

where $\widehat{f}_1, \widehat{f}_2$ denote the orthonormal frame of the Heisenberg group (see again Remark 5.9). Moreover, from (5.32) and (5.33) one easily gets

$$-(c_{12}^1)^\varepsilon f_2^\varepsilon = -2\varepsilon^2\partial_y\gamma(\partial_y + \frac{x}{2}\partial_w) + o(\varepsilon^4),$$

$$(c_{12}^2)^\varepsilon f_1^\varepsilon = -2\varepsilon^2\partial_x\gamma(\partial_x - \frac{y}{2}\partial_w) + o(\varepsilon^4).$$

Recollecting all the terms we find that

$$\mathcal{L}_{f^\varepsilon} = \mathcal{L}_{\widehat{f}} + \varepsilon^2\mathcal{Y} + o(\varepsilon^2),$$

where

$$\mathcal{L}_{\widehat{f}} = \widehat{f}_1^2 + \widehat{f}_2^2 = (\partial_x - \frac{y}{2}\partial_w)^2 + (\partial_y + \frac{x}{2}\partial_w)^2,$$

is the sub-Laplacian on the Heisenberg group, and $\mathcal{Y}$ denotes the second order differential operator

$$\mathcal{Y} = \frac{\gamma}{2}(x^2+y^2)\partial_w^2 + \gamma(x\partial_{wy} - y\partial_{wx}) - \frac{1}{2}(x\partial_y\gamma - y\partial_x\gamma)\partial_w - 2(\partial_x\gamma\partial_x + \partial_y\gamma\partial_y).$$

Specifying expansion (5.23) to our case, where $Y = Y(\varepsilon)$ is a smooth perturbation which expands with respect to ε as follows

$$Y(\varepsilon) = \varepsilon^2 \mathcal{Y} + o(\varepsilon^2)$$

we find that

$$p^\varepsilon = H + \varepsilon^2 (H * \mathcal{Y}H) + \varepsilon^4 (p^\varepsilon * \mathcal{Y}H * \mathcal{Y}H). \tag{5.34}$$

Since p^ε is the heat kernel of a contraction semigroup (for every $\varepsilon > 0$), one can see that the last term in (5.34) is bounded and

$$p^\varepsilon(1,0,0) = H(1,0,0) + \varepsilon^2 H * \mathcal{Y}H(1,0,0) + O(\varepsilon^4),$$

where

$$\mathcal{Y}H(t,q,q') = \mathcal{Y}_q H(t,q,q'),$$

means that $\mathcal{Y}$ acts as a differential operator on the first spatial variable.

From the explicit expression (5.25) it immediately follows that

$$H(1,0,0) = \frac{1}{16t^2}.$$

Thus, denoting by $K_1 := H * \mathcal{Y}H(1,0,0)$ from Corollary 5.16 we have the expansion of the original heat kernel

$$p(t,x,x) \sim \frac{1}{16t^2}(1 + K_1 t + O(t^2)).$$

We are left to computation of the convolution between H and YH, namely

$$\begin{aligned} H * \mathcal{Y}H(1,0,0) &= \int_0^t \int_{\mathbb{R}^3} H(s,0,q)\mathcal{Y}H(1-s,q,0)\, dq ds \\ &= \int_0^t \int_{\mathbb{R}^3} h_s(q)\mathcal{Y}h_{1-s}(q)\, dq ds, \qquad q = (x,y,w). \end{aligned} \tag{5.35}$$

Computing derivatives under the integral sign one gets

$$\begin{aligned} \mathcal{Y}h_t(x,y,w) = &-\frac{1}{(4\pi t)^2} \int_{\mathbb{R}} \frac{r}{\sinh r} \exp\left(-\frac{r(x^2+y^2)}{2t \tanh r}\right) \frac{r}{t^2} \\ &\times \left[\gamma(x,y) \cos(\frac{rw}{t}) \left(r(x^2+y^2) - \frac{4t}{\tanh r}\right) + t\,\gamma'(x,y) \sin(\frac{rw}{t})\right] dr, \end{aligned}$$

where

$$\gamma(x,y) = ax^2 + bxy + cy^2, \qquad \gamma'(x,y) = 2(a-c)xy + b(y^2 - x^2).$$

Notice that we are interested only in computing how the integral (5.35) depends on the constants a, b, c, and the perturbation of the metric is on the variables x, y only. Hence, using that the integrand has exponential decay with respect to x, y, we can exchange the order of integration in (5.35) and integrate first with respect to these variables.

Using that, for $\alpha > 0$

$$\iint_{\mathbb{R}^2} x^2 e^{-\alpha(x^2+y^2)} dxdy = \frac{c_1\pi}{\alpha}, \qquad \iint_{\mathbb{R}^2} x^4 e^{-\alpha(x^2+y^2)} = \frac{c_2\pi}{\alpha^3},$$
$$\iint_{\mathbb{R}^2} xy e^{-\alpha(x^2+y^2)} dxdy = 0,$$

for some constants $c_1, c_2 > 0$, it is easily seen that integrating (5.35) we get an expression of the kind

$$H * \mathcal{Y}H(1,0,0) = C_0(a+c) = C_0\kappa, \tag{5.36}$$

where C_0 is a universal constant that does not depend on the sub-Riemannian structure. Hence the value of C_0 can be computed from some explicit formula of the heat kernel in the non nilpotent case. Using the expression given in [22] for the heat kernel on $SU(2)$, where the value of local invariants are constant $\chi = 0, \kappa = 1$ (see also [2]) we get that

$$p_{SU(2)}(t,0,0) = \frac{e^t}{16t^2} \sim \frac{1}{16t^2}(1 + t + O(t^2)),$$

where we renormalized the constants in order to fit into our setting. Hence $C_0 = 1$ in (5.36), and the Theorem is proved.

Remark 5.21. The same method applies to get a quick proof of the following-well known result (see e.g. [28, 91]): on a 2-dimensional Riemannian manifold M, the heat kernel $p(t,x,y)$ satisfies an asymptotic expansion on the diagonal

$$p(t,x,x) \sim \frac{1}{4\pi t}(1 + \frac{K(x)}{6}t + O(t^2)), \qquad \text{for } t \to 0,$$

where $K(x)$ denotes the gaussian curvature at the point $x \in M$. Indeed one can use the normal coordinates on M to write the orthonormal frame in the following way

$$f_1 = \partial_x + \beta y(y\partial_x - x\partial_y),$$
$$f_2 = \partial_y - \beta x(y\partial_x - x\partial_y),$$

where β is, a priori, a smooth function $\beta = \beta(x,y)$. Reasoning as in Lemma 5.19, β can be chosen as a constant since we are interested only in first order term. In this case it is also easily seen that the Gaussian curvature at the origin is computed via the parameter β as $K = 6\beta$.

Bibliography

[1] R. Abraham and J. Robbin, *Transversal mappings and flows*, An appendix by Al Kelley, W. A. Benjamin, Inc., New York-Amsterdam, 1967.

[2] A. Agrachev and D. Barilari, *Sub-Riemannian structures on 3D Lie groups*, to appear on J. Dyn. and Contr. Syst., (2011).

[3] A. Agrachev, D. Barilari, and U. Boscain, *Introduction to Riemannian and sub-Riemannian geometry*, Lecture Notes, (2011). http://people.sissa.it/agrachev/agrachev_files/notes.html.

[4] ——, *On the Hausdorff volume in sub-Riemannian geometry*, Calculus of Variations and Partial Differential Equations, (2011), pp. 1–34. 10.1007/s00526-011-0414-y.

[5] A. Agrachev, B. Bonnard, M. Chyba, and I. Kupka, *Sub-Riemannian sphere in Martinet flat case*, ESAIM Control Optim. Calc. Var., 2 (1997), pp. 377–448 (electronic).

[6] A. Agrachev, U. Boscain, J.-P. Gauthier, and F. Rossi, *The intrinsic hypoelliptic Laplacian and its heat kernel on unimodular Lie groups*, J. Funct. Anal., 256 (2009), pp. 2621–2655.

[7] A. Agrachev and J.-P. Gauthier, *On the subanalyticity of Carnot-Caratheodory distances*, Ann. Inst. H. Poincaré Anal. Non Linéaire, 18 (2001), pp. 359–382.

[8] A. A. Agrachev, *Exponential mappings for contact sub-Riemannian structures*, J. Dynam. Control Systems, 2 (1996), pp. 321–358.

[9] A. A. Agrachev, E.-H. Chakir El-A., and J. P. Gauthier, *Sub-Riemannian metrics on* $\mathbf{R}^3$, in Geometric control and non-holonomic mechanics (Mexico City, 1996), vol. 25 of CMS Conf. Proc., Amer. Math. Soc., Providence, RI, 1998, pp. 29–78.

[10] A. A. Agrachev, G. Charlot, J. P. A. Gauthier, and V. M. Zakalyukin, *On sub-Riemannian caustics and wave fronts for contact distributions in the three-space*, J. Dynam. Control Systems, 6 (2000), pp. 365–395.

[11] A. A. Agrachev, R. V. Gamkrelidze, and A. V. Sarychev, *Local invariants of smooth control systems*, Acta Appl. Math., 14 (1989), pp. 191–237.

[12] A. A. Agrachev and J.-P. A. Gauthier, *On the Dido problem and plane isoperimetric problems*, Acta Appl. Math., 57 (1999), pp. 287–338.

[13] A. A. Agrachev and Y. L. Sachkov, *Control theory from the geometric viewpoint*, vol. 87 of Encyclopaedia of Mathematical Sciences, Springer-Verlag, Berlin, 2004. Control Theory and Optimization, II.

[14] A. A. Agrachev and A. V. Sarychev, *Filtrations of a Lie algebra of vector fields and the nilpotent approximation of controllable systems*, Dokl. Akad. Nauk SSSR, 295 (1987), pp. 777–781.

[15] F. Alouges, A. DeSimone, and A. Lefebvre, *Optimal strokes for low Reynolds number swimmers: an example*, J. Nonlinear Sci., 18 (2008), pp. 277–302.

[16] L. Ambrosio, F. Serra Cassano, and D. Vittone, *Intrinsic regular hypersurfaces in Heisenberg groups*, J. Geom. Anal., 16 (2006), pp. 187–232.

[17] L. Ambrosio and P. Tilli, *Topics on analysis in metric spaces*, vol. 25 of Oxford Lecture Series in Mathematics and its Applications, Oxford University Press, Oxford, 2004.

[18] Z. M. Balogh, M. Rickly, and F. Serra Cassano, *Comparison of Hausdorff measures with respect to the Euclidean and the Heisenberg metric*, Publ. Mat., 47 (2003), pp. 237–259.

[19] ——, *Comparison of Hausdorff measures with respect to the Euclidean and the Heisenberg metric*, Publ. Mat., 47 (2003), pp. 237–259.

[20] Z. M. Balogh, J. T. Tyson, and B. Warhurst, *Sub-Riemannian vs. Euclidean dimension comparison and fractal geometry on Carnot groups*, Adv. Math., 220 (2009), pp. 560–619.

[21] F. Baudoin, *An introduction to the geometry of stochastic flows*, Imperial College Press, London, 2004.

[22] F. Baudoin and M. Bonnefont, *The subelliptic heat kernel on* SU(2)*: representations, asymptotics and gradient bounds*, Math. Z., 263 (2009), pp. 647–672.

[23] F. Baudoin and N. Garofalo, *Curvature-dimension inequalities and Ricci lower bounds for sub-Riemannian manifolds with transverse symmetries*, arXiv:1101.3590v1, (2011).

[24] R. Beals, B. Gaveau, and P. Greiner, *The Green function of model step two hypoelliptic operators and the analysis of certain tangential Cauchy Riemann complexes*, Adv. Math., 121 (1996), pp. 288–345.

[25] A. Bellaïche, *The tangent space in sub-Riemannian geometry*, in Sub-Riemannian geometry, vol. 144 of Progr. Math., Birkhäuser, Basel, 1996, pp. 1–78.

[26] G. Ben Arous, *Développement asymptotique du noyau de la chaleur hypoelliptique hors du cut-locus*, Ann. Sci. École Norm. Sup. (4), 21 (1988), pp. 307–331.

[27] G. Ben Arous, *Développement asymptotique du noyau de la chaleur hypoelliptique sur la diagonale*, Ann. Inst. Fourier (Grenoble), 39 (1989), pp. 73–99.

[28] M. Berger, P. Gauduchon, and E. Mazet, *Le spectre d'une variété riemannienne*, Lecture Notes in Mathematics, Vol. 194, Springer-Verlag, Berlin, 1971.

[29] B. Berret, C. Darlot, F. Jean, T. Pozzo, C. Papaxanthis, and J. P. Gauthier, *The inactivation principle: mathematical solutions minimizing the absolute work and biological implications for the planning of arm movements*, PLoS Comput. Biol., 4 (2008), pp. e1000194, 25.

[30] R. M. Bianchini and G. Stefani, *Graded approximations and controllability along a trajectory*, SIAM J. Control Optim., 28 (1990), pp. 903–924.

[31] J.-M. Bismut, *Large deviations and the Malliavin calculus*, vol. 45 of Progress in Mathematics, Birkhäuser Boston Inc., Boston, MA, 1984.

[32] A. Bonfiglioli, E. Lanconelli, and F. Uguzzoni, *Stratified Lie groups and potential theory for their sub-Laplacians*, Springer Monographs in Mathematics, Springer, Berlin, 2007.

[33] B. Bonnard, M. Chyba, and E. Trelat, *Sub-Riemannian geometry, one-parameter deformation of the Martinet flat case*, J. Dynam. Control Systems, 4 (1998), pp. 59–76.

[34] U. Boscain, T. Chambrion, and J.-P. Gauthier, *On the $K+P$ problem for a three-level quantum system: optimality implies resonance*, J. Dynam. Control Systems, 8 (2002), pp. 547–572.

[35] U. Boscain, J. Duplaix, J.-P. Gauthier, and F. Rossi, *Anthropomorphic image reconstruction via hypoelliptic diffusion*, arXiv:1006.3735v3 [math.OC], (2011).

[36] U. Boscain, J.-P. Gauthier, and F. Rossi, *Hypoelliptic heat kernel over 3-step nilpotent lie groups*, to appear on Proceedings of the Steklov Mathematical Institute, (2010).

[37] U. Boscain and B. Piccoli, *Optimal syntheses for control systems on 2-D manifolds*, vol. 43 of Mathématiques & Applications (Berlin) [Mathematics & Applications], Springer-Verlag, Berlin, 2004.

[38] U. Boscain and F. Rossi, *Invariant Carnot-Caratheodory metrics on S^3, SO(3), SL(2), and lens spaces*, SIAM J. Control Optim., 47 (2008), pp. 1851–1878.

[39] ——, *Projective Reeds-Shepp car on S^2 with quadratic cost*, ESAIM Control Optim. Calc. Var., 16 (2010), pp. 275–297.

[40] A. Bressan and B. Piccoli, *Introduction to the mathematical theory of control*, vol. 2 of AIMS Series on Applied Mathematics, American Institute of Mathematical Sciences (AIMS), Springfield, MO, 2007.

[41] R. W. Brockett, *Nonlinear control theory and differential geometry*, in Proceedings of the International Congress of Mathematicians, Vol. 1, 2 (Warsaw, 1983), Warsaw, 1984, PWN, pp. 1357–1368.

[42] R. W. Brockett and A. Mansouri, *Short-time asymptotics of heat kernels for a class of hypoelliptic operators*, Amer. J. Math., 131 (2009), pp. 1795–1814.

[43] P. Cannarsa and C. Sinestrari, *Semiconcave functions, Hamilton-Jacobi equations, and optimal control*, Progress in Nonlinear Differential Equations and their Applications, 58, Birkhäuser Boston Inc., Boston, MA, 2004.

[44] L. Capogna, D. Danielli, S. D. Pauls, and J. T. Tyson, *An introduction to the Heisenberg group and the sub-Riemannian isoperimetric problem*, vol. 259 of Progress in Mathematics, Birkhäuser Verlag, Basel, 2007.

[45] É. Cartan, *Sur la géométrie pseudo-conforme des hypersurfaces de l'espace de deux variables complexes II*, Ann. Scuola Norm. Sup. Pisa Cl. Sci. (2), 1 (1932), pp. 333–354.

[46] ——, *Sur la géométrie pseudo-conforme des hypersurfaces de l'espace de deux variables complexes*, Ann. Mat. Pura Appl., 11 (1933), pp. 17–90.

[47] L. Cesari, *Optimization—theory and applications*, vol. 17 of Applications of Mathematics (New York), Springer-Verlag, New York, 1983. Problems with ordinary differential equations.

[48] G. Charlot, *Quasi-contact S-R metrics: normal form in* $\mathbf{R}^{2n}$*, wave front and caustic in* $\mathbf{R}^4$, Acta Appl. Math., 74 (2002), pp. 217–263.

[49] I. Chavel, *Isoperimetric inequalities*, vol. 145 of Cambridge Tracts in Mathematics, Cambridge University Press, Cambridge, 2001. Differential geometric and analytic perspectives.

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

[51] G. Citti and A. Sarti, *A cortical based model of perceptual completion in the roto-translation space*, J. Math. Imaging Vision, 24 (2006), pp. 307–326.

[52] H. L. Cycon, R. G. Froese, W. Kirsch, and B. Simon, *Schrödinger operators with application to quantum mechanics and global geometry*, Texts and Monographs in Physics, Springer-Verlag, Berlin, study ed., 1987.

[53] E.-H. C. El-Alaoui, J.-P. Gauthier, and I. Kupka, *Small sub-Riemannian balls on* $\mathbf{R}^3$, J. Dynam. Control Systems, 2 (1996), pp. 359–421.

[54] E. Falbel and C. Gorodski, *Sub-Riemannian homogeneous spaces in dimensions* 3 *and* 4, Geom. Dedicata, 62 (1996), pp. 227–252.

[55] H. Federer, *Geometric measure theory*, Die Grundlehren der mathematischen Wissenschaften, Band 153, Springer-Verlag New York Inc., New York, 1969.

[56] G. B. Folland, *Subelliptic estimates and function spaces on nilpotent Lie groups*, Ark. Mat., 13 (1975), pp. 161–207.

[57] G. B. Folland and E. M. Stein, *Estimates for the $\bar{\partial}_b$ complex and analysis on the Heisenberg group*, Comm. Pure Appl. Math., 27 (1974), pp. 429–522.

[58] B. Franchi, R. Serapioni, and F. Serra Cassano, *Regular hypersurfaces, intrinsic perimeter and implicit function theorem in Carnot groups*, Comm. Anal. Geom., 11 (2003), pp. 909–944.

[59] N. Garofalo and E. Lanconelli, *Asymptotic behavior of fundamental solutions and potential theory of parabolic operators with variable coefficients*, Math. Ann., 283 (1989), pp. 211–239.

[60] J.-P. Gauthier, B. Jakubczyk, and V. Zakalyukin, *Motion planning and fastly oscillating controls*, SIAM J. Control Optim., 48 (2010), pp. 3433–3448.

[61] J.-P. Gauthier and V. Zakalyukin, *On the one-step-bracket-generating motion planning problem*, J. Dyn. Control Syst., 11 (2005), pp. 215–235.

[62] J.-P. Gauthier and V. Zakalyukin, *On the motion planning problem, complexity, entropy, and nonholonomic interpolation*, J. Dyn. Control Syst., 12 (2006), pp. 371–404.

[63] B. Gaveau, *Principe de moindre action, propagation de la chaleur et estimées sous elliptiques sur certains groupes nilpotents*, Acta Math., 139 (1977), pp. 95–153.

[64] V. Gershkovich and A. Vershik, *Nonholonomic manifolds and nilpotent analysis*, J. Geom. Phys., 5 (1988), pp. 407–452.

[65] M. Gromov, *Carnot-Carathéodory spaces seen from within*, in Sub-Riemannian geometry, vol. 144 of Progr. Math., Birkhäuser, Basel, 1996, pp. 79–323.

[66] M. Gromov, *Metric structures for Riemannian and non-Riemannian spaces*, Modern Birkhäuser Classics, Birkhäuser Boston Inc., Boston, MA, english ed., 2007. Based on the 1981 French original, With appendices by M. Katz, P. Pansu and S. Semmes, Translated from the French by Sean Michael Bates.

[67] L. Hörmander, *Hypoelliptic second order differential equations*, Acta Math., 119 (1967), pp. 147–171.

[68] N. Jacobson, *Lie algebras*, Interscience Tracts in Pure and Applied Mathematics, No. 10, Interscience Publishers (a division of John Wiley & Sons), New York-London, 1962.

[69] F. Jean, *Entropy and complexity of a path in sub-Riemannian geometry*, ESAIM Control Optim. Calc. Var., 9 (2003), pp. 485–508 (electronic).

[70] D. S. Jerison and A. Sánchez-Calle, *Estimates for the heat kernel for a sum of squares of vector fields*, Indiana Univ. Math. J., 35 (1986), pp. 835–854.

[71] V. Jurdjevic, *Geometric control theory*, vol. 52 of Cambridge Studies in Advanced Mathematics, Cambridge University Press, Cambridge, 1997.

[72] T. Kato, *Perturbation theory for linear operators*, Classics in Mathematics, Springer-Verlag, Berlin, 1995. Reprint of the 1980 edition.

[73] K. Kurdyka and L. Paunescu, *Hyperbolic polynomials and multiparameter real-analytic perturbation theory*, Duke Math. J., 141 (2008), pp. 123–149.

[74] S. Kusuoka and D. Stroock, *Applications of the Malliavin calculus. II*, J. Fac. Sci. Univ. Tokyo Sect. IA Math., 32 (1985), pp. 1–76.

[75] R. Léandre, *Développement asymptotique de la densité d'une diffusion dégénérée*, Forum Math., 4 (1992), pp. 45–75.

[76] G. P. Leonardi, S. Rigot, and D. Vittone, *Isodiametric sets in Heisenberg groups*, arXiv:1010.1133v1 [math.MG], (2011).

[77] V. Magnani, *Spherical Hausdorff measure of submanifolds in Heisenberg groups*, Ricerche Mat., 54 (2005), pp. 607–613 (2006).

[78] J. Mitchell, *A local study of Carnot-Carathéodory metrics*, PhD Thesis, (1982).

[79] ——, *On Carnot-Carathéodory metrics*, J. Differential Geom., 21 (1985), pp. 35–45.

[80] I. Moiseev and Y. L. Sachkov, *Maxwell strata in sub-Riemannian problem on the group of motions of a plane*, ESAIM Control Optim. Calc. Var., 16 (2010), pp. 380–399.

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

[82] R. Monti and F. Serra Cassano, *Surface measures in Carnot-Carathéodory spaces*, Calc. Var. Partial Differential Equations, 13 (2001), pp. 339–376.

[83] T. Nagano, *Linear differential systems with singularities and an application to transitive Lie algebras*, J. Math. Soc. Japan, 18 (1966), pp. 398–404.

[84] R. Neel, *The small-time asymptotics of the heat kernel at the cut locus*, Comm. Anal. Geom., 15 (2007), pp. 845–890.

[85] R. Neel and D. Stroock, *Analysis of the cut locus via the heat kernel*, in Surveys in differential geometry. Vol. IX, Surv. Differ. Geom., IX, Int. Press, Somerville, MA, 2004, pp. 337–349.

[86] J. Petitot, *Neurog?eom'etrie de la vision - mod'eles math?ematiques et physiques des architectures fonctionnelles*, Les E?ditions de lE?cole Polytechnique, (2008).

[87] J. Petitot and Y. Tondut, *Vers une neurogéométrie. Fibrations corticales, structures de contact et contours subjectifs modaux*, Math. Inform. Sci. Humaines, (1999), pp. 5–101.

[88] L. S. Pontryagin, V. G. Boltyanskii, R. V. Gamkrelidze, and E. F. Mishchenko, *The mathematical theory of optimal processes*, Translated from the Russian by K. N. Trirogoff; edited by L. W. Neustadt, Interscience Publishers John Wiley & Sons, Inc. New York-London, 1962.

[89] P. Rashevsky, *Any two points of a totally nonholonomic space may be connected by an admissible line*, Uch. Zap. Ped Inst. im. Liebknechta, 2 (1938), pp. 83–84.

[90] S. Rigot, *Isodiametric inequality in carnot groups*, arXiv:1004.1369v1 [math.MG], (2010).

[91] S. Rosenberg, *The Laplacian on a Riemannian manifold*, vol. 31 of London Mathematical Society Student Texts, Cambridge University Press, Cambridge, 1997. An introduction to analysis on manifolds.

[92] L. P. Rothschild and E. M. Stein, *Hypoelliptic differential operators and nilpotent groups*, Acta Math., 137 (1976), pp. 247–320.

[93] Y. L. Sachkov, *Symmetries of flat rank two distributions and sub-Riemannian structures*, Trans. Amer. Math. Soc., 356 (2004), pp. 457–494 (electronic).

[94] Y. L. Sachkov, *Complete description of Maxwell strata in the generalized Dido problem*, Mat. Sb., 197 (2006), pp. 111–160.

[95] Y. L. Sachkov, *Conjugate and cut time in the sub-Riemannian problem on the group of motions of a plane*, ESAIM Control Optim. Calc. Var., 16 (2010), pp. 1018–1039.

[96] R. S. Strichartz, *Sub-Riemannian geometry*, J. Differential Geom., 24 (1986), pp. 221–263.

[97] H. J. Sussmann, *An extension of a theorem of Nagano on transitive Lie algebras*, Proc. Amer. Math. Soc., 45 (1974), pp. 349–356.

[98] H. J. Sussmann, *Lie brackets, real analyticity and geometric control*, in Differential geometric control theory (Houghton, Mich., 1982), vol. 27 of Progr. Math., Birkhäuser Boston, Boston, MA, 1983, pp. 1–116.

[99] T. Taylor, *A parametrix for step-two hypoelliptic diffusion equations*, Trans. Amer. Math. Soc., 296 (1986), pp. 191–215.

[100] T. J. S. Taylor, *Off diagonal asymptotics of hypoelliptic diffusion equations and singular Riemannian geometry*, Pacific J. Math., 136 (1989), pp. 379–399.

[101] Y. Yu, *The index theorem and the heat equation method*, vol. 2 of Nankai Tracts in Mathematics, World Scientific Publishing Co. Inc., River Edge, NJ, 2001.