

# VOLUME OF SMALL BALLS AND SUB-RIEMANNIAN CURVATURE IN 3D CONTACT MANIFOLDS

DAVIDE BARILARI, IVAN BESCHASTNYI, AND ANTONIO LERARIO

ABSTRACT. We compute the asymptotic expansion of the volume of small sub-Riemannian balls in a contact 3-dimensional manifold, and we express the first meaningful geometric coefficients in terms of geometric invariants of the sub-Riemannian structure.

## 1. INTRODUCTION

Let  $M$  be an  $n$ -dimensional Riemannian manifold and for  $p \in M$  and  $\varepsilon > 0$  let us denote by  $B(p, \varepsilon)$  the Riemannian ball of radius  $\varepsilon$  centered at  $p$ , i.e., the set of points in  $M$  at distance at most  $\varepsilon$  from  $p$ . A classical result allows to write the asymptotic expansion of the Riemannian volume of  $B(p, \varepsilon)$  in terms of the volume  $\beta_n$  of the unit ball in  $\mathbb{R}^n$  and the scalar curvature  $s(p)$  of  $M$  at the point  $p$ :

$$(1) \quad \text{vol}(B(p, \varepsilon)) = \beta_n \varepsilon^n \left( 1 - \frac{s(p)}{6(n+2)} \varepsilon^2 + O(\varepsilon^3) \right).$$

This formula says that to the leading order the volume of  $B(p, \varepsilon)$  coincides with the volume of the  $\varepsilon$ -ball in the model space  $\mathbb{R}^n$  and the first correction term (which is quadratic in  $\varepsilon$ ) depends on the curvature of  $M$  at  $p$ . The purpose of this paper is to derive an analogue formula in the case of a 3-dimensional contact sub-Riemannian manifold.

To be more specific, let  $M$  be a 3-dimensional manifold,  $D \subset TM$  be a contact distribution and  $g$  be a metric on  $D$ . For  $p \in M$  and  $\varepsilon > 0$  let us denote by  $B(p, \varepsilon)$  the *sub-Riemannian*  $\varepsilon$ -ball centered at  $p$ , i.e., the set of points in  $M$  reached by a horizontal curve starting from  $p$  and of length at most  $\varepsilon$ . The model space for  $M$  is the 3-dimensional Heisenberg group  $\mathbb{H}^3$ .

In this framework one can consider a natural volume form on  $M$ , called the *Popp volume*, which is defined as follows. First, we write  $D = \ker(\omega)$  for a one-form  $\omega$  normalized such that  $d\omega|_D$  coincides with the area form of  $g$ . If  $X_0$  denotes the Reeb vector field of  $\omega$  and  $\{X_1, X_2\}$  is an oriented orthonormal basis for  $D$ , the Popp volume form is the unique 3-form  $\mu$  such that  $\mu(X_1, X_2, X_0) = 1$ . We denote by  $\text{vol}(A) = \int_A \mu$ .

In the sub-Riemannian context, we obtain an expansion completely analogous to (1), where now the Riemannian scalar curvature is replaced by a sub-Riemannian curvature term. On a 3-dimensional contact sub-Riemannian manifold one can introduce two curvature functions denoted  $\chi$  and  $\kappa$  (see Section 2). Observe that only the latter appears in the first terms of the asymptotics of the volume of small balls.

**Theorem 1.** *Let  $(M, D, g)$  be a 3-dimensional, contact sub-Riemannian manifold and  $p \in M$ . As  $\varepsilon \rightarrow 0$  the Popp volume of the sub-Riemannian  $\varepsilon$ -ball centered at  $p$  has the following asymptotics:*

$$(2) \quad \text{vol}(B(p, \varepsilon)) = c_0 \varepsilon^4 \left( 1 - c_1 \kappa(p) \varepsilon^2 + O(\varepsilon^3) \right),$$

---

UNIV. PARIS DIDEROT, INSTITUT DE MATHÉMATIQUES DE JUSSIEU-PARIS RIVE GAUCHE, CNRS, SORBONNE UNIVERSITÉ. BÂTIMENT SOPHIE-GERMAIN, CASE 7012, 75205 PARIS CEDEX 13, FRANCE

SISSA, VIA BONOMEA, 265, 34136 TRIESTE TS, ITALIA

SISSA, VIA BONOMEA, 265, 34136 TRIESTE TS, ITALIA

E-mail addresses: [davide.barilari@imj-prg.fr](mailto:davide.barilari@imj-prg.fr), [ibeschas@sissa.it](mailto:ibeschas@sissa.it), [lerario@sissa.it](mailto:lerario@sissa.it).

Date: February 26, 2018.

where  $c_0$  is the volume of the unit ball in the Heisenberg group  $\mathbb{H}^3$ . Explicitly

$$c_0 = \frac{1}{12}(1 + 2\pi\text{Si}(2\pi)) \quad \text{and} \quad c_1 = \frac{1}{c_0 160} \left( 2 + 4\pi\text{Si}(2\pi) - \frac{1}{\pi^2} \right) > 0$$

where  $\text{Si}(x) = \int_0^x \frac{\sin t}{t} dt$  denotes the sine integral function.

The expansion (2) can be used as a definition of scalar curvature for a 3D contact sub-Riemannian structure.

In the general case, it is not straightforward to define such a notion of scalar curvature associated with a given sub-Riemannian structure. A general approach to curvature for general sub-Riemannian structures has been developed in [8, 12, 5, 13]. Here a notion of generalized *sectional* curvature is obtained through horizontal derivatives of the distance functions and a scalar curvature can be built by considering the trace of its horizontal part. In some specific cases, when a canonical connection (in general with non-zero torsion) associated with the metric is available, one can also introduce curvature through this connection, as done for instance in [16, 15]. In the 3D contact case discussed in this paper, all these approaches coincide, and define the same scalar curvature invariant  $\kappa$  (up to a constant).

**1.1. Connection with small time heat kernel asymptotics.** It is interesting to observe that the behavior of the volume of small balls is strictly related to the small time asymptotics of the heat kernel on the manifold. It is well-known, in fact, that for the heat kernel  $e(t, x, y)$  associated with the sub-Riemannian Laplacian, the following estimate holds for  $t > 0$  small enough

$$(3) \quad \frac{C_1}{\text{vol}(B(p, t^{1/2}))} \leq e(t, p, p) \leq \frac{C_2}{\text{vol}(B(p, t^{1/2}))}.$$

This estimate follows from more general off-diagonal Gaussian estimates on the heat kernel, investigated first in the sub-Riemannian setting in [23, 24, 26] and then refined in several subsequent papers in the literature. The estimate (3) roughly says that the main order term of the expansion of  $\text{vol}(B(p, t^{1/2}))^{-1}$  and  $e(t, p, p)$  for small  $t$  is the same.

Actually, in the Riemannian case, one can observe a stronger relationship between the two asymptotics, since the following expansion holds for  $t \rightarrow 0$

$$(4) \quad e(t, p, p) = \frac{1}{(4\pi t)^{n/2}} \left( 1 + \frac{s(p)}{6} t + O(t^2) \right).$$

This shows that even in the first order correction term the expansions (1) and (4) contain the same geometric invariant.

Theorem 1 compared with the results obtained in [10] confirms that the same analogies remain true in the sub-Riemannian setting, at least for 3D contact structures.

Concerning higher-dimensional structures, only partial results are known even the contact case: the small time heat kernel asymptotics for contact structures with symmetries have been obtained in [19, 27] and the first coefficient has been related to the scalar Tanaka-Webster curvature. See also [17, 18, 29] for a recent account on heat kernel asymptotics on higher-dimensional sub-Riemannian model spaces.

**1.2. On the strategy of the proof.** Even if the question of computing the asymptotics of the volume of small sub-Riemannian balls seems very natural, this is the first paper where this question is investigated in the literature. This is related to the fact that the classical ingredients are not available in sub-Riemannian geometry, as we now explain.

A first obstruction is that the sub-Riemannian exponential map, parametrizing arclength geodesics, is defined on a non-compact set (homeomorphic to a cylinder) and is never a local diffeomorphism at zero. As a consequence, balls are not smooth, even for small radii, preventing a uniform description of the injectivity domain for the exponential map in the

cotangent space. In other terms, an information on the cut locus starting from a point (that is always adjacent to the point itself) is necessary to have a correct description of balls through the exponential map.

Another obstacle in the computation of the asymptotic expansion above is that balls are not geodesically homogeneous in sub-Riemannian geometry: if one shrinks a balls to its center along geodesics, one does not obtain a ball of the corresponding radius. More precisely, defining  $\Phi_{p,t}$  the maps that sends  $x \in M$  to the point at time  $t$  along the unique geodesic joining  $p$  with  $x$  in time 1 (this map is well defined for a.e.  $x \in M$  on a contact sub-Riemannian manifold) one has that  $\Phi_{p,t}(B(p, r)) \subsetneq B(p, tr)$ , with strict inclusion. It is possible actually to show that, on every 3-dimensional contact sub-Riemannian manifold, when  $t \rightarrow 0$  the quantity  $\text{vol}(\Phi_{p,t}(B(p, r)))$  goes to zero as  $t^5$ , while  $\text{vol}(B(p, tr))$  tends to zero as  $t^4$ . Hence, even if curvature-like invariants can be extracted by looking at the variation of a smooth volume under the geodesic flow (see [4]), this does not permits to get the volume of small balls.

To overcome these problems we use a perturbative approach: i.e., we describe the original contact structure around a fixed point  $p \in M$  as the perturbation of the Heisenberg sub-Riemannian structure, that is the metric tangent structure at a fixed point. This procedure relies on the so called nilpotent approximation of the sub-Riemannian structure and the use of a version of normal coordinates for the three dimensional exponential map developed in [7, 9, 21]. This permits us to compute the asymptotic expansion of all ingredients that are involved in the computation of the volume of the ball (the exponential map, the cut time for geodesics, the Popp volume) and obtain Theorem 1 without explicit computations of the cut locus.

**Acknowledgements.** This research has been supported by the ANR project SRGI ‘‘Sub-Riemannian Geometry and Interactions’’, contract number ANR-15-CE40-0018.

## 2. TECHNICAL INGREDIENTS

A *sub-Riemannian structure* on a manifold  $M$  is a pair  $(D, g)$  where  $D$  is a vector distribution, i.e., a subbundle of the tangent bundle  $TM$ , and  $g$  is a smooth metric defined on  $D$ . It is required that  $D$  is *bracket generating*, i.e., Lie brackets of vector fields tangent to  $D$  span the full tangent space to  $M$  at every point.

Under this assumptions there is a well-defined *sub-Riemannian (or Carnot-Carathéodory) distance*  $d$ , namely  $d(p, q)$  is the infimum of the length of Lipschitz curves joining two points  $p$  and  $q$  and that are tangent to  $D$  (also called horizontal curves). Here the length of the curve is computed with respect to the metric  $g$ . We refer to [3] for a comprehensive presentation.

We will also say that the triplet  $(M, D, g)$  is a sub-Riemannian manifold, when  $(D, g)$  is a sub-Riemannian structure on a smooth manifold  $M$ .

**2.1. Contact sub-Riemannian manifold.** Let  $M$  be a 3-dimensional manifold. A sub-Riemannian structure  $(D, g)$  on  $M$  is said to be *contact* if  $D$  is a contact distribution, i.e.,  $D = \ker \omega$ , where  $\omega \in \Lambda^1(M)$  satisfies  $d\omega \wedge \omega \neq 0$ .

A contact distribution is bracket generating and endows  $M$  with a canonical orientation. It is convenient to recall the following characterization of contact distributions in terms of local frames.

**Lemma 2.** *Let  $M$  be a 3-dimensional smooth manifold,  $\omega \in \Lambda^1(M)$  and set  $D := \ker \omega$ . The following are equivalent:*

- (i)  $D$  is a contact distribution,
- (ii)  $d\omega|_D \neq 0$ ,
- (iii) if  $X_1, X_2$  is a local frame that spans  $D$ , then  $[X_1, X_2]$  is transversal to  $D$ .

Moreover, if  $(D, g)$  is a given sub-Riemannian structure on  $M$ , the contact form  $\omega$  can be selected in such a way that  $d\omega|_D$  coincides with the Euclidean volume defined on  $D$  by  $g$ .

Given a contact sub-Riemannian structure on  $M$ , we always assume that the contact form  $\omega$  is normalized by the metric  $g$  as in Lemma 2.

*Remark 1.* By Lemma 2, it is not restrictive to fix a sub-Riemannian contact structure by a pair of everywhere linearly independent vector fields  $X_1, X_2$  such that  $X_1, X_2, [X_1, X_2]$  is a basis of the tangent space at every point, and declaring  $X_1, X_2$  to be an orthonormal frame for  $g$  on the distribution  $D = \text{span}\{X_1, X_2\}$ .

The *Reeb vector field* associated with the contact structure is the unique vector field  $X_0$  satisfying

$$(5) \quad X_0 \in \ker d\omega, \quad \omega(X_0) = 1.$$

The Reeb vector field depends only on the sub-Riemannian structure, and its orientation.

Given an orthonormal frame  $X_1, X_2$  for the sub-Riemannian structure  $(D, g)$  there exists smooth functions  $c_{ij}^k$  defined on  $M$  such that

$$(6) \quad \begin{aligned} [X_1, X_0] &= c_{01}^1 X_1 + c_{01}^2 X_2, \\ [X_2, X_0] &= c_{02}^1 X_1 + c_{02}^2 X_2, \\ [X_2, X_1] &= c_{12}^1 X_1 + c_{12}^2 X_2 + X_0, \end{aligned}$$

The particular structure of the equations (6) are obtained from the properties (5) of the Reeb vector field by applying Cartan formula.

**Definition 3.** We define the following quantities in terms of the structural equations (6) of the orthonormal frame

- (a) the invariant  $\chi$  defined by

$$\chi = \sqrt{\frac{(c_{01}^2 + c_{02}^2)^2}{4} - c_{01}^1 c_{02}^2},$$

- (b) the invariant  $\kappa$  defined by

$$\kappa = X_2(c_{12}^1) - X_1(c_{12}^2) - (c_{12}^1)^2 - (c_{12}^2)^2 + \frac{c_{01}^2 - c_{02}^1}{2}.$$

Notice that  $\chi$  and  $\kappa$  are smooth functions defined on  $M$ .

*Remark 2.* We list here some properties of the coefficients just introduced. More details are provided in [3] and [8, Section 7.5] (cf. also [7, 1] and references below)

- (i) These coefficients have been first introduced in [6, 7]. A direct calculation shows that  $\chi$  and  $\kappa$  are independent of the orthonormal frame  $X_1, X_2$  on the distribution and are hence local metric invariant. Indeed, Theorem 1 also shows independently that  $\kappa$  is a metric invariant.
- (ii) It is possible to introduce a canonical connection  $\nabla$  on a 3-dimensional contact manifold. The functions  $\chi$  and  $\kappa$  are expressed in terms of  $\nabla$  as follows:

$$\chi = \sqrt{-\det^2 T^\nabla(X_0, \cdot)}, \quad \kappa = R^\nabla(X_1, X_2, X_2, X_1),$$

where  $T^\nabla$  and  $R^\nabla$  represents the torsion and the curvature tensor associated with  $\nabla$ , respectively. More details on the canonical connection are provided in Appendix A.

- (iii) One can show that  $\chi \geq 0$ , and  $\chi$  vanishes everywhere if and only if the flow of the Reeb vector field  $X_0$  is a flow of sub-Riemannian isometries for  $M$ . When  $\chi = 0$  identically, the sub-Riemannian structure can be represented as an isoperimetric problem on a two-dimensional Riemannian manifold  $N$ , and  $\kappa$  represents the Gaussian curvature on  $N$ .
- (iv) The functions  $\chi$  and  $\kappa$  are invariant by local isometries and they are constant functions for left-invariant structures on Lie groups. In particular if  $\chi = \kappa = 0$  the structure is local isometric to the Heisenberg group. See [1, 22] for the classification of left-invariant structures in terms of these invariants.

**2.2. Normal coordinates.** The basic example of contact sub-Riemannian structure in dimension three is the Heisenberg group: this is the sub-Riemannian structure defined by the orthonormal frame in  $\mathbb{R}^3$

$$(7) \quad \widehat{X}_1 = \partial_x - \frac{y}{2}\partial_z, \quad \widehat{X}_2 = \partial_y + \frac{x}{2}\partial_z.$$

Notice that the normalized contact form and the corresponding Reeb vector field for this structure are

$$\widehat{\omega} = -dz - \frac{y}{2}dx + \frac{x}{2}dy, \quad \widehat{X}_0 = -\partial_z.$$

For a general 3-dimensional contact sub-Riemannian structure, 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. In this coordinates a general 3-dimensional contact sub-Riemannian structure is presented as a perturbation of the Heisenberg group.

**Theorem 4** ([9, 21]). *Let  $M$  be a 3-dimensional contact sub-Riemannian manifold and  $X_1, X_2$  a local orthonormal frame. Fix  $p \in M$ . Then there exists a smooth coordinate system  $(x, y, z)$  around  $p$  such that  $p = (0, 0, 0)$  and*

$$(8) \quad X_1 = (\partial_x - \frac{y}{2}\partial_z) + \beta y(y\partial_x - x\partial_y) - \gamma \frac{y}{2}\partial_z,$$

$$(9) \quad X_2 = (\partial_y + \frac{x}{2}\partial_z) - \beta x(y\partial_x - x\partial_y) + \gamma \frac{x}{2}\partial_z,$$

where  $\beta = \beta(x, y, z)$  and  $\gamma = \gamma(x, y, z)$  are smooth functions satisfying the following boundary conditions

$$(10) \quad \beta(0, 0, z) = \gamma(0, 0, z) = \frac{\partial\gamma}{\partial x}(0, 0, z) = \frac{\partial\gamma}{\partial y}(0, 0, z) = 0.$$

Notice that, when  $\beta = \gamma = 0$ , one recovers formulas (7). In the same spirit as Riemannian normal coordinates, the coordinates given by Theorem 4 normalize the zero-order term of the metric and have no first order correction term. For this statement to be formalized, let us introduce the notion of nilpotent approximation.

**2.3. Nilpotent approximation.** In normal coordinates we introduce the family of dilations  $\delta_\varepsilon : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ , for every  $\varepsilon > 0$ , by

$$\delta_\varepsilon(x, y, z) = (\varepsilon x, \varepsilon y, \varepsilon^2 z).$$

For  $i = 1, 2$  we denote by  $X_i^\varepsilon$  the vector fields in  $\mathbb{R}^3$

$$X_i^\varepsilon = \varepsilon(\delta_{\frac{1}{\varepsilon}})_* X_i.$$

For  $\varepsilon > 0$ , we consider the distribution  $D^\varepsilon = \text{span}\{X_1^\varepsilon, X_2^\varepsilon\}$ ; we put a metric  $g^\varepsilon$  on this distribution by declaring  $\{X_1^\varepsilon, X_2^\varepsilon\}$  an orthonormal basis. Observe that  $(\mathbb{R}^3, D^0, g^0)$  with this metric is the Heisenberg group  $\mathbb{H}^3$ . We denote by  $B^\varepsilon(1)$  the unit ball centered at the origin for the sub-Riemannian manifold  $(\mathbb{R}^3, D^\varepsilon, g^\varepsilon)$  and by  $B(\varepsilon) \subset \mathbb{R}^3$  the image of  $B(p, \varepsilon)$  under the normal coordinates map.

**Lemma 5.** *For every  $\varepsilon > 0$  small enough we have  $B(\varepsilon) = \delta_\varepsilon(B^\varepsilon(1))$ .*

*Proof.* We prove that  $\gamma : I \rightarrow \mathbb{R}^3$  is a horizontal curve for  $(\mathbb{R}^3, D^1, g^1)$  with length  $\ell(\gamma)$  if and only if  $\gamma_\varepsilon = \delta_{1/\varepsilon} \circ \gamma$  is a horizontal curve for  $(\mathbb{R}^3, D^\varepsilon, g^\varepsilon)$  with length  $\varepsilon^{-1}\ell(\gamma)$ . From this the conclusion follows.

Since  $\gamma : [0, 1] \rightarrow M$  is horizontal, there exists  $u_1, u_2 \in L^\infty(I)$  such that:

$$\dot{\gamma}(t) = u_1(t)X_1(\gamma(t)) + u_2(t)X_2(\gamma(t)) \quad \text{for a.e. } t \in [0, 1].$$

In particular:

$$\begin{aligned} \dot{\gamma}_\varepsilon(t) &= u_1(t)\delta_{\frac{1}{\varepsilon}*}X_1(\gamma_\varepsilon(t)) + u_2(t)\delta_{\frac{1}{\varepsilon}*}X_2(\gamma_\varepsilon(t)) \\ &= \frac{u_1(t)}{\varepsilon}X_1^\varepsilon(\gamma_\varepsilon(t)) + \frac{u_2(t)}{\varepsilon}X_2^\varepsilon(\gamma_\varepsilon(t)) \\ &= u_1^\varepsilon(t)X_1^\varepsilon(\gamma_\varepsilon(t)) + u_2^\varepsilon(t)X_2^\varepsilon(\gamma_\varepsilon(t)), \end{aligned}$$

where we have set  $u_i^\varepsilon = \varepsilon^{-1}u_i$ . In particular  $\gamma_\varepsilon$  is also horizontal and, since by definition  $\{X_1^\varepsilon, X_2^\varepsilon\}$  is an orthonormal frame for  $(\mathbb{R}^3, D^\varepsilon, g^\varepsilon)$ , its length with respect to  $g^\varepsilon$  equals:

$$\ell^\varepsilon(\gamma_\varepsilon) = \int_I \sqrt{u_1^\varepsilon(t)^2 + u_2^\varepsilon(t)^2} dt = \varepsilon^{-1} \int_I \sqrt{u_1(t)^2 + u_2(t)^2} dt = \varepsilon^{-1}\ell(\gamma).$$

□

Next lemma expresses the vector fields  $X_i^\varepsilon$  as perturbations of the vector fields  $\widehat{X}_i$  defining the Heisenberg structure and follows from a direct computation.

Given a smooth function  $F(x, y, z)$  of three variables, we denote by  $F^{[2]}(x, y, z)$  the second order homogeneous part of its Taylor polynomial at zero. Moreover we set  $F^{[2]}(x, y) := F^{[2]}(x, y, 0)$ .

**Lemma 6.** *The following asymptotic expansion holds for  $\varepsilon \rightarrow 0$*

$$\begin{aligned} X_1^\varepsilon &= (\partial_x - \frac{y}{2}\partial_z) - \varepsilon^2 \frac{y}{2} \gamma^{[2]}(x, y) \partial_z + O(\varepsilon^3), \\ X_2^\varepsilon &= (\partial_y + \frac{x}{2}\partial_z) + \varepsilon^2 \frac{x}{2} \gamma^{[2]}(x, y) \partial_z + O(\varepsilon^3), \end{aligned}$$

Moreover, denoting by  $X_0^\varepsilon := \varepsilon^2(\delta_{\frac{1}{\varepsilon}})_*X_0$  and  $(c_{ij}^k)^\varepsilon$  the structure constant satisfying

$$[X_j^\varepsilon, X_i^\varepsilon] = \sum_{k=0}^2 (c_{ij}^k)^\varepsilon X_k^\varepsilon, \quad i, j = 0, 1, 2,$$

we have

$$\begin{aligned} (c_{12}^1)^\varepsilon &= 2\varepsilon^2 \partial_y \gamma + O(\varepsilon^3), & (c_{12}^2)^\varepsilon &= -2\varepsilon^2 \partial_x \gamma + O(\varepsilon^3) \\ (c_{01}^1)^\varepsilon &= -2\varepsilon^2 \partial_{xy}^2 \gamma + O(\varepsilon^3) & (c_{01}^2)^\varepsilon &= 2\varepsilon^2 \partial_x^2 \gamma + O(\varepsilon^3) \\ (c_{02}^1)^\varepsilon &= -2\varepsilon^2 \partial_{xy}^2 \gamma + O(\varepsilon^3) & (c_{02}^2)^\varepsilon &= 2\varepsilon^2 \partial_{xy}^2 \gamma + O(\varepsilon^3). \end{aligned}$$

where the partial derivatives of  $\gamma$  are computed at zero.

*Proof.* The expansion of  $X_i^\varepsilon$  follows directly from the definitions of the vector fields and their explicit form in normal coordinates (8) and (10).

To prove the asymptotics of the structure constants, we note first that the vector fields  $X_1^\varepsilon, X_2^\varepsilon$  for each  $\varepsilon \geq 0$  define a contact structure with  $X_0^\varepsilon$  as the Reeb field. Indeed, it is easy to verify using the definitions that

$$\omega^\varepsilon = \frac{1}{\varepsilon^2} \delta_\varepsilon^* \omega$$

is the one-form defining the distribution. This means that  $X_i^\varepsilon$  satisfy the structure equations (6) with  $(c_{ij}^k)^\varepsilon$  as structure constants, for  $i, j, k = 0, 1, 2$ . Knowing explicitly the Reeb field

$X_0^\varepsilon$ , one could expand  $X_i^\varepsilon$  and  $(c_{ij}^k)^\varepsilon$  into power series of  $\varepsilon$  and solve recursively for coefficients of  $(c_{ij}^k)^\varepsilon$  remembering that for the Heisenberg structure  $c_{12}^0 = -c_{21}^0 = 1$  and  $c_{ij}^k = 0$  otherwise.

To avoid explicit computations, we proceed in a slightly different manner. Instead of considering our original contact structure defined by  $X_i$ , we look at different structure defined by  $\tilde{X}_i$ , such that the asymptotic expansions of vector fields  $X_i^\varepsilon$  and  $\tilde{X}_i^\varepsilon$  agree up to a certain order of  $\varepsilon$ . Then by repeating the previous argument we obtain an asymptotic expansion for  $(\tilde{c}_{ij}^k)^\varepsilon$  that agrees with asymptotic expansion for  $(c_{ij}^k)^\varepsilon$  up to the same order.

To prove our claim we need the asymptotic expansion up to order two. This can be achieved by considering vector fields

$$\begin{aligned}\tilde{X}_1 &= \partial_x - \frac{y(1 + \varepsilon^2 \gamma^{[2]}(x, y))}{2} \partial_z, \\ \tilde{X}_2 &= \partial_y + \frac{x(1 + \varepsilon^2 \gamma^{[2]}(x, y))}{2} \partial_z,\end{aligned}$$

which is just a truncation of the original vector fields. In [10] explicit expressions for the corresponding one-form  $\tilde{\omega}$ , the Reeb vector field  $\tilde{X}_0$  and the structure constants  $(\tilde{c}_{ij}^k)$  we found. In particular

$$\begin{aligned}(\tilde{c}_{12}^1)^\varepsilon &= \frac{2\partial_y \gamma}{1 + 2\gamma}, & (\tilde{c}_{12}^2)^\varepsilon &= -\frac{2\partial_x \gamma}{1 + 2\gamma} \\ (\tilde{c}_{01}^1)^\varepsilon &= -\frac{2((1 + 2\gamma)\partial_{xy}\gamma - 2\partial_y \gamma \partial_x \gamma)}{(1 + 2\gamma)^2}, & (\tilde{c}_{01}^2)^\varepsilon &= \frac{2((1 + 2\gamma)\partial_{xx}\gamma - 2(\partial_y \gamma)^2)}{(1 + 2\gamma)^2} \\ (\tilde{c}_{02}^1)^\varepsilon &= -\frac{2((1 + 2\gamma)\partial_{yy}\gamma + 2(\partial_x \gamma)^2)}{(1 + 2\gamma)^2}, & (\tilde{c}_{02}^2)^\varepsilon &= \frac{2((1 + 2\gamma)\partial_{xy}\gamma + 2\partial_y \gamma \partial_x \gamma)}{(1 + 2\gamma)^2}.\end{aligned}$$

Lemma 6 now is a direct consequence of the following homogeneity property. If  $c_{ij}^k$  are the structure constants associated with the vector fields  $X_i$ , then

$$(c_{ij}^k)^\varepsilon = \varepsilon^{d_i + d_j - d_k} (c_{ij}^k \circ \delta_\varepsilon)$$

where we set  $d_1 = d_2 = 1$  and  $d_0 = 2$ . □

**2.4. Exponential map.** In order to describe a geodesic ball, we need a good description of geodesics. As in Riemannian geometry, one can define an analogue of the exponential map. But unlike the Riemannian case, it is defined as a map from the cotangent bundle to the manifold using Hamiltonian dynamics.

We start by defining the basis Hamiltonian  $h_i : T^*M \rightarrow M$  as linear on fibers functions

$$h_i(\lambda) = \langle \lambda, X_i \rangle, \quad \lambda \in T^*M, \quad i = 0, 1, 2.$$

We are going to use  $h_i(\lambda)$  as coordinate functions on fibers of  $T^*M$  and therefore from now on we do not indicate explicitly the dependence on  $\lambda$ .

It is well known that geodesics on a rank 2 sub-Riemannian manifold are projections of solutions of a Hamiltonian system with a quadratic Hamiltonian [3]

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

So we can define the exponential map  $\exp : T^*M \rightarrow M$  associated to the Hamiltonian  $H$  as:

$$\exp(\lambda) = \pi(e^{\vec{H}}(\lambda)),$$

where  $\vec{H}$  is the corresponding Hamiltonian vector field. So if we wish to restrict only to geodesics that go out from a point  $p \in M$ , we have to consider  $\lambda \in T_p^*M$ , and we define the map  $\exp_p : T_p^*M \rightarrow M$  as a restriction

$$\exp_p = \exp|_{T_p^*M}.$$

Since we are interested only in the behaviour of small balls around  $p$ , one can use the usual Darboux coordinates on the cotangent bundle and the standard Poisson bracket to write down explicitly the Hamiltonian system. But it is better to take a slightly more invariant approach and consider the Lie-Poisson bracket on  $T^*M$ . The Lie-Poisson bracket of two basis Hamiltonians  $h_i, h_j$  is defined as

$$\{h_j, h_i\}(\lambda) = \langle \lambda, [X_j, X_i] \rangle = \sum_{k=0}^2 c_{ij}^k h_k(\lambda).$$

A bracket of any two smooth functions on the fibers of  $T^*M$  can be defined via linearity and Leibnitz rule. Then our Hamiltonian system can be written as

$$(11) \quad \begin{cases} \dot{p} = h_1 X_1(p) + h_2 X_2(p) \\ \dot{h}_i = \{H, h_i\} \end{cases}$$

*Remark 3.* It is interesting that the invariant  $\chi$  can be obtained directly from the Hamiltonian system. If we denote by

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

the corresponding quadratic form in  $h_1, h_2$ , then  $\text{trace}\{H, h_0\} = 0$  since  $e_*^{tX_0} D = D$ . The other invariant

$$\sqrt{-\det\{H, h_0\}} \geq 0$$

is exactly  $\chi$  and it is zero when  $(e_*^{tX_0})^* g = g$ .

It is well known that solutions of a Hamiltonian system lie on a level set of the corresponding Hamiltonian. In our case the projections of the level sets to fibers of  $T^*M$  are cylinders. So we can introduce *cylindrical coordinates*  $(\rho, \theta, w)$  on  $T_p^*M$  as

$$\begin{aligned} h_1 &= \rho \cos \theta \\ h_2 &= \rho \sin \theta \\ h_0 &= -w \end{aligned}$$

It follows immediately that  $\rho$  is constant along solutions of (11) and it is equal to the speed of the corresponding geodesics.

We are interested in the study of small balls  $B(\varepsilon)$ . Lemma 5 gives an explicit relation between  $B(\varepsilon)$  and the unit ball  $B^\varepsilon(1)$  of the dilated system. We can describe the ball  $B^\varepsilon(1)$  and its volume using the exponential map of the dilated system, that we denote by  $\exp_p^\varepsilon$ . We also have a different set of cylindrical coordinates, but they are related as can be seen from the following lemma.

**Lemma 7.** *For every  $\varepsilon > 0$  let  $\tau_\varepsilon : T_p^*M \rightarrow T_p^*M$  be the map defined in cylindrical coordinates by  $\tau_\varepsilon(\rho, \theta, w) = (\rho\varepsilon, \theta, w)$ . Then for every  $\varepsilon > 0$  the following diagram is commutative:*

$$\begin{array}{ccc} T_p^*M & \xrightarrow{\tau_\varepsilon} & T_p^*M \\ \exp_p \downarrow & & \downarrow \exp_p^\varepsilon \\ \mathbb{R}^3 & \xrightarrow{\delta_\varepsilon} & \mathbb{R}^3 \end{array}$$

Here we identify  $\mathbb{R}^3$  with an open neighborhood of  $p$  on which normal coordinates are defined.

*Proof.* Since both  $\delta_\varepsilon$  and  $\tau_\varepsilon$  are diffeomorphisms, we prove the equivalent statement:

$$\delta_\varepsilon \circ \exp_p^\varepsilon = \exp_p \circ \tau_\varepsilon.$$

We start by recalling the definition of  $\exp_p^\varepsilon = \pi \circ e^{\overrightarrow{H^\varepsilon}}$ , where  $H^\varepsilon : T^*M \rightarrow \mathbb{R}$  is the Hamiltonian:

$$(12) \quad H^\varepsilon = \frac{1}{2} \left( (h_1^\varepsilon)^2 + (h_2^\varepsilon)^2 \right).$$

By definition the hamiltonians  $h_i^\varepsilon$  are given by:

$$h_i^\varepsilon(\lambda) = \left\langle \lambda, \varepsilon(\delta_{1/\varepsilon})_* X_i \right\rangle = \varepsilon \left\langle (\delta_{1/\varepsilon})^* \lambda, X_i \right\rangle = \varepsilon h_i(\alpha_\varepsilon(\lambda)),$$

where we have defined the diffeomorphisms  $\alpha_\varepsilon \doteq (\delta_{1/\varepsilon})^* : T^*M \rightarrow T^*M$ . Notice that  $\alpha_\varepsilon$  lifts  $\delta_\varepsilon : M \rightarrow M$ , hence it is a symplectomorphisms. As a consequence  $H^\varepsilon = \varepsilon^2 H \circ \alpha_\varepsilon$ . In particular we can write (we use simple identities that can be easily verified by the reader, referring for example to [3] for a detailed proof):

$$\begin{aligned} \delta_\varepsilon \circ \exp^\varepsilon &= \delta_\varepsilon \circ \pi \circ e^{\varepsilon^2 \overrightarrow{H \circ \alpha_\varepsilon}} \\ &= \delta_\varepsilon \circ \pi \circ e^{\varepsilon^2 (\alpha_\varepsilon^{-1})_* \overrightarrow{H}} \quad (\text{by [3, Proposition 4.52]}) \\ &= \delta_\varepsilon \circ \pi \circ \alpha_\varepsilon^{-1} \circ e^{\varepsilon^2 \overrightarrow{H}} \circ \alpha_\varepsilon \quad (\text{by [3, Lemma 2.20]}) \\ &= \pi \circ e^{\varepsilon^2 \overrightarrow{H}} \circ \alpha_\varepsilon \quad (\text{because } \alpha_\varepsilon^{-1} \text{ lifts } \delta_{1/\varepsilon}) \\ &= \pi \circ \varepsilon^{-2} \circ e^{\overrightarrow{H}} \circ \varepsilon^2 \circ \alpha_\varepsilon \quad (\text{by [3, Lemma 8.33]}) \\ &= \pi \circ e^{\overrightarrow{H}} \circ \varepsilon^2 \circ \alpha_\varepsilon \quad (\text{because } \varepsilon^{-2} \text{ preserves the fibers of } \pi). \end{aligned}$$

It remains to verify that  $\varepsilon^2 \circ \alpha_\varepsilon|_{T^*_p M} = \tau_\varepsilon$ . Recalling the definition of  $\alpha_\varepsilon = (\delta_{1/\varepsilon})^*$  we see that  $\alpha_\varepsilon|_{T^*_q M}$  is given in cylindrical coordinates by:

$$\alpha_\varepsilon(\rho, \theta, w) = \left( \frac{\rho}{\varepsilon}, \theta, \frac{w}{\varepsilon^2} \right),$$

and consequently  $\varepsilon^2 \circ \alpha_\varepsilon(\rho, \theta, w) = (\varepsilon\rho, \theta, w) = \tau_\varepsilon(\rho, \theta, w)$ . This concludes the proof.  $\square$

The next proposition gives the necessary asymptotics of the Jacobian of the exponential maps  $\exp_p^\varepsilon$ .

**Proposition 8.** *The Jacobians  $\det(J \exp^\varepsilon)$  of the family of exponential maps  $\exp_p^\varepsilon$  have the following expansion in cylindrical coordinates  $(\rho, \theta, w)$  as  $\varepsilon \rightarrow 0$ :*

$$(13) \quad \det(J \exp_p^\varepsilon) = \det(J \exp_p^0) + v_2(\rho, \theta, w)\varepsilon^2 + O(\varepsilon^3),$$

where  $\exp_p^0 : \mathbb{R}^3 \rightarrow \mathbb{R}^3$  is the exponential map for the Heisenberg group and

$$v_2(\rho, \theta, w) = \rho^5 \left( \frac{\kappa(p)}{2} g_0(w) + g_c(w) \cos 2\theta + g_s(w) \sin 2\theta \right)$$

with  $g_0(w), g_c(w), g_s(w)$  smooth functions of  $w$  variable only. Moreover the functions  $\det(J \exp_p^0)$  and  $g_0$  have the following expression:

$$(14) \quad \det(J \exp_p^0)(\rho, \theta, w) = \rho^3 \frac{(2 - 2 \cos w - w \sin w)}{w^4}$$

$$g_0(w) = \frac{(16 - 3w^2) \cos w + 2 \cos 2w + 13w \sin w + w \sin 2w - 18}{w^6}.$$

*Remark 4.* We observe that the crucial information contained in Proposition 8 that we use later is

$$v_2(\rho, \theta, w) = \kappa(p) f_1(\rho, w) + f_2(\rho, \theta, w) \quad \text{with} \quad \int_0^{2\pi} f_2(\rho, \theta, w) d\theta = 0.$$

*Proof.* We start by writing the Hamiltonian system for the dilated structure. The Hamiltonian  $H^\varepsilon$  is given by (12) and we can write the Hamiltonian system (11) explicitly using the Lie-Poisson bracket. We get

$$(15) \quad \begin{cases} \dot{p} = h_1^\varepsilon X_1(p) + h_2^\varepsilon X_2(p) \\ \dot{h}_1^\varepsilon = \{H^\varepsilon, h_1^\varepsilon\} = \{h_2^\varepsilon, h_1^\varepsilon\} h_2^\varepsilon \\ \dot{h}_2^\varepsilon = \{H^\varepsilon, h_2^\varepsilon\} = \{h_1^\varepsilon, h_2^\varepsilon\} h_1^\varepsilon \\ \dot{h}_0^\varepsilon = \{H^\varepsilon, h_0^\varepsilon\} = \{h_1^\varepsilon, h_0^\varepsilon\} h_1^\varepsilon + \{h_2^\varepsilon, h_0^\varepsilon\} h_2^\varepsilon \end{cases}$$

To rewrite our system in cylindrical coordinates, we make a change of variables on the fibers of  $T^*M$

$$\begin{aligned} h_1^\varepsilon &= \rho \cos \theta \\ h_2^\varepsilon &= \rho \sin \theta \\ h_0^\varepsilon &= -w \end{aligned}$$

We also introduce the following functions

$$\begin{aligned} a^\varepsilon(\theta) &= \frac{1}{\rho^2} \{H^\varepsilon, h_0^\varepsilon\} = (c_{01}^1)^\varepsilon \cos(\theta)^2 + ((c_{01}^2)^\varepsilon + (c_{02}^1)^\varepsilon) \cos(\theta) \sin(\theta) + (c_{02}^2)^\varepsilon \sin(\theta)^2, \\ b^\varepsilon(\theta) &= -\frac{1}{\rho} (\{h_1^\varepsilon, h_2^\varepsilon\} + h_0^\varepsilon) = (c_{12}^1)^\varepsilon \cos \theta + (c_{12}^2)^\varepsilon \sin \theta. \end{aligned}$$

Then, after various simplifications, we obtain the Hamiltonian system

$$\begin{cases} \dot{p} = \rho \cos(\theta) X_1^\varepsilon(p) + \rho \sin(\theta) X_2^\varepsilon(p) \\ \dot{\rho} = 0 \\ \dot{\theta} = w - \rho b^\varepsilon(\theta) \\ \dot{w} = -\rho^2 a^\varepsilon(\theta). \end{cases}$$

Now we expand the right-hand side and the phase variables in series of powers of  $\varepsilon$ . This will give us a number of ordinary differential equations on the coefficients, that we are going to solve. Since the Hamiltonian system (15) is smooth, depends smoothly on  $\varepsilon$ , and we are interested only in the behaviour for small  $\varepsilon$ , the resulting asymptotics is going to be uniform.

We fix normal coordinates  $(x, y, z)$  around  $p \in M$ . In this coordinates  $p = (0, 0, 0)$ . Then we fix an initial covector  $(\bar{\rho}, \bar{\theta}, \bar{w})$  and look at how the corresponding geodesic changes as  $\varepsilon$  goes to zero. Thus our asymptotic expansions are

$$\begin{cases} x(t) = x_0(t) + x_1(t)\varepsilon + x_2(t)\frac{\varepsilon^2}{2} + O(\varepsilon^3) \\ y(t) = y_0(t) + y_1(t)\varepsilon + y_2(t)\frac{\varepsilon^2}{2} + O(\varepsilon^3) \\ z(t) = z_0(t) + z_1(t)\varepsilon + z_2(t)\frac{\varepsilon^2}{2} + O(\varepsilon^3) \end{cases} \quad \begin{cases} w(t) = w_0(t) + w_1(t)\varepsilon + w_2(t)\frac{\varepsilon^2}{2} + O(\varepsilon^3) \\ \theta(t) = \theta_0(t) + \theta_1(t)\varepsilon + \theta_2(t)\frac{\varepsilon^2}{2} + O(\varepsilon^3) \end{cases}$$

Since the initial covector  $(\bar{\rho}, \bar{\theta}, \bar{w})$  is independent of  $\varepsilon$ , we have the following boundary conditions

$$\begin{aligned} x_i(0) &= y_i(0) = z_i(0) = 0, & \forall i \in \mathbb{N}_0 \\ \theta_i(0) &= w_i(0) = 0, & \forall i \in \mathbb{N} \\ \theta_0(0) &= \bar{\theta}, \quad w_0(0) = \bar{w}. \end{aligned}$$

Let us look at the principal and first order terms of the asymptotics. First of all we note that from Lemma 6 it follows that all the structure constants are  $O(\varepsilon^2)$ . Thus functions  $a^\varepsilon$

and  $b^\varepsilon$  are  $O(\varepsilon^2)$  as well. Using the asymptotics of  $X_i^\varepsilon$  from the same lemma, we then obtain a system for the zero-order term

$$\begin{cases} \dot{x}_0 = \bar{\rho} \cos \theta_0 \\ \dot{y}_0 = \bar{\rho} \sin \theta_0 \\ \dot{z}_0 = \frac{\bar{\rho}}{2}(x_0 \sin \theta_0 - y_0 \cos \theta_0) \end{cases} \quad \begin{cases} \dot{\theta}_0 = w_0 \\ \dot{w}_0 = 0 \end{cases}$$

But this is nothing but the geodesic equations on the Heisenberg group whose solutions are explicit

$$(16) \quad \begin{cases} x_0(t) = \frac{\bar{\rho}(\sin(\bar{w}t + \bar{\theta}) - \sin \bar{\theta})}{\bar{w}} \\ y_0(t) = -\frac{\bar{\rho}(\cos(\bar{w}t + \bar{\theta}) - \cos \bar{\theta})}{\bar{w}} \\ z_0(t) = \frac{\bar{\rho}^2(\bar{w}t - \sin t\bar{w})}{2\bar{w}^2} \end{cases} \quad \begin{cases} \theta_0(t) = \bar{w}t + \bar{\theta} \\ w_0(t) = \bar{w} \end{cases}$$

Thus we see that as  $\varepsilon \rightarrow 0$  geodesics of the dilated system converge to the geodesics of the Heisenberg group as expected. Moreover, setting  $t = 1$  in (16) and differentiating with respect to  $(\bar{\rho}, \bar{\theta}, \bar{w})$  we immediately obtain (14).

Next we write the system of order one. We obtain

$$\begin{cases} \dot{x}_1 = -\bar{\rho}\theta_1 \sin \theta_0 \\ \dot{y}_1 = \bar{\rho}\theta_1 \cos \theta_0 \\ \dot{z}_1 = \frac{\bar{\rho}}{2}(-y_1 \cos \theta_0 + x_1 \sin \theta_0 + x_0\theta_1 \cos \theta_0 + y_0\theta_1 \sin \theta_0) \end{cases} \quad \begin{cases} \dot{\theta}_1 = w_1 \\ \dot{w}_1 = 0 \end{cases}$$

Using the zero boundary conditions we get

$$w_1(t) = \theta_1(t) = x_1(t) = y_1(t) = z_1(t) = 0, \quad \forall t.$$

From here it immediately follows that the zero-order term in the expression is the Heisenberg term and the first order term is identically zero.

We continue this procedure. At each next step we integrate expression involving only terms from the previous steps. Then we can plug all asymptotic expansions into the Jacobian and after various simplifications, we obtain the result. The second order term in the asymptotics of the exponential map is a result of similar but rather long computations. The simplification of the expression for the Jacobian becomes a tedious exercise after applying various trigonometric identities.  $\square$

**2.5. The Popp volume and curvature invariants in normal coordinates.** On a contact sub-Riemannian manifold it is possible to define a canonical volume that depends only on the sub-Riemannian structure, called Popp volume. Here we recall its construction only in the 3-dimensional case, the interested reader is referred to [25] and [11] for the general construction and its explicit expression in terms of an adapted frame.

Given an orthonormal frame  $X_1, X_2$  for the sub-Riemannian structure and the corresponding Reeb vector field  $X_0$ , let us denote by  $\nu_1, \nu_2, \nu_0$  the dual basis of 1-forms. The Popp volume  $\mu$  is defined as the three-form  $\mu = \pm \nu_1 \wedge \nu_2 \wedge \nu_0$ . The sign is chosen in such a way that the volume is positive.

Recall that we denote by  $F^{[2]}(x, y, z)$  the second order homogeneous part of a smooth function  $F(x, y, z)$  of three variables and  $F^{[2]}(x, y) := F^{[2]}(x, y, 0)$ .

**Lemma 9.** *Using normal coordinates (introduced in Section 2.2) the Popp volume form can be written as  $\mu = \psi dx \wedge dy \wedge dz$ , where  $\psi : \mathbb{R}^3 \rightarrow \mathbb{R}$  is a smooth function such that:*

$$(17) \quad \psi(x, y, z) = 1 - 2\gamma^{[2]}(x, y) + O(\|(x, y, z)\|^3).$$

where

$$(18) \quad \gamma^{[2]}(x, y) = x^2 \partial_x^2 \gamma + 2xy \partial_{xy}^2 \gamma + y^2 \partial_y^2 \gamma$$

where the partial derivatives of  $\gamma$  in (18) are computed at zero.

*Proof.* For notational convenience, let us introduce  $X_3 := [X_2, X_1]$  and denote by  $(x_1, x_2, x_3)$  the coordinates  $(x, y, z)$ . Let  $\nu_1, \nu_2, \nu_3$  be the dual basis of 1-forms to  $X_1, X_2, X_3$ . Notice that the Popp volume  $\mu$  is written as  $\mu = \nu_1 \wedge \nu_2 \wedge \nu_3$  (up to the choice of positive sign), as a consequence of the relation  $[X_2, X_1] = X_3 \pmod{D}$  (cf. (6)). Considering the coordinate expression of the vector fields and the basis of 1-forms

$$X_i = \sum_{j=1}^3 a_{ij} \partial_j, \quad i = 1, 2, 3, \quad \text{and} \quad \nu_k = \sum_{l=1}^3 b_{kl} dx_l, \quad k = 1, 2, 3,$$

for some smooth functions  $a_{ij}, b_{kl}$ . Then the matrices  $A = (a_{ij})$  and  $B = (b_{kl})$  satisfy the relation  $B = (A^T)^{-1}$ . In particular

$$(19) \quad \mu = |\det(B)| dx \wedge dy \wedge dz = |\det(A)|^{-1} dx \wedge dy \wedge dz.$$

From the explicit expression of the vector fields (8)-(9) and boundary conditions (10), it is easy to check that every coefficient of the vector fields containing  $\beta$  gives a contribution of order at least three in the expansion of the determinant. Hence, to compute the expansion of (19) up to second order, it is not restrictive to assume that  $\beta = 0$ . Under this assumption, one computes

$$[X_1, X_2] = \left( 1 + \gamma + \frac{1}{2}(x\partial_x\gamma + y\partial_y\gamma) \right) \partial_z,$$

which implies

$$\begin{aligned} |\det(A)| &= 1 + \gamma(x, y, z) + \frac{1}{2}(x\partial_x\gamma(x, y, z) + y\partial_y\gamma(x, y, z)) + O\left(\|(x, y, z)\|^3\right) \\ &= 1 + 2\gamma^{[2]}(x, y) + O\left(\|(x, y, z)\|^3\right), \end{aligned}$$

where in the last equality we used the boundary conditions (10). Taking the inverse and combining with (19), the proof is completed.  $\square$

Along the same lines of the proof of Lemma 9 one obtains the following result. A proof is contained in [10, Lemma 4].

**Lemma 10.** *In normal coordinates (introduced in Section 2.2) writing*

$$\gamma^{[2]}(x, y) = ax^2 + 2bxy + cy^2,$$

*we have the following expression for the curvature-like invariants at the origin*

$$\kappa(p) = 2(a + c), \quad \chi(p) = 2\sqrt{b^2 + (c - a)^2}.$$

**2.6. Cut-time asymptotic.** Let  $\gamma : [0, T] \rightarrow M$  be a horizontal curve. We say that  $\gamma$  is a *length-minimizer* if  $d(\gamma(0), \gamma(T)) = \ell(\gamma)$ . Notice that this notion is independent on the parametrization of the curve.

Fix now a horizontal curve  $\gamma : [0, T] \rightarrow M$  parametrized by arc-length. Then we define

$$t_{cut}(\gamma) = \sup\{t > 0 : \gamma|_{[0,t]} \text{ is a length-minimizer}\}.$$

If  $\gamma$  is a geodesic parametrized by arclength on a 3-dimensional contact manifold, then it is well-known that  $t_{cut}(\gamma) > 0$ . This is related with the fact that there are no abnormal minimizers, see for instance [3, Chapter 8] and [14, Appendix].

Parametrizing geodesics as in Section 2.4 with covectors in cylindrical coordinates  $(\rho, \theta, w)$  we have a well-defined cut time associated with every initial covector with  $\rho = 1$ .

We give here the asymptotic expansion for the cut time of geodesics. This result is obtained combining [7, Theorem 4.2] and [7, Theorem 5.2], covering the case  $\chi(p) \neq 0$  and  $\chi(p) = 0$ , respectively.<sup>1</sup>

<sup>1</sup>A note on the reference: to recover the cut time, denoted  $\ell_*(\theta; \nu)$  in [7, Theorem 5.2], one needs the formula of  $\ell_1(\theta; \nu)$  of [7, Theorem 5.1], whose expression contains a typo. Indeed the second summand of its expression is  $-\pi\kappa(q_0)\nu^{-3}$  (and not  $-\pi\kappa(q_0)\nu^{-2}$ ) as it can be directly checked from the proof.

**Proposition 11.** *We have the following asymptotic expansion for the cut time from  $p \in M$ .*

$$(20) \quad t_{cut}(1, \theta, w) = \frac{2\pi}{|w|} - \frac{\pi(\kappa(p) + 2\chi(p) \sin^2 \theta)}{|w|^3} + O\left(\frac{1}{|w|^4}\right), \quad w \rightarrow \pm\infty.$$

Thanks to this result we get an asymptotic description of the set of initial parameters mapped on the ball of radius  $\varepsilon$  through the exponential map.

**Corollary 12.** *For  $\varepsilon > 0$  small enough there exists a set  $\Omega(\varepsilon) \subset T_p^*M$ , whose description in cylindrical coordinates is given by*

$$(21) \quad \Omega(\varepsilon) = \left\{ |\rho| \leq \varepsilon, \theta \in [0, 2\pi], w \in [-2\pi + \rho^2 f(\theta) + O(\rho^3), 2\pi - \rho^2 f(\theta) + O(\rho^3)] \right\},$$

where  $f(\theta)$  is a smooth function of  $\theta$ . Moreover the following properties holds:

- (i)  $\overline{B}(p, \varepsilon) = \exp(\Omega(\varepsilon))$ ;
- (ii)  $\exp|_{\text{int}(\Omega(\varepsilon) \setminus \{\rho=0\})}$  is injective with injective differential;
- (iii)  $\exp(\text{int}(\Omega(\varepsilon) \setminus \{\rho=0\}))$  has full measure in  $B(p, \varepsilon)$ .

We observe that the existence of a set  $\Omega(\varepsilon)$  satisfying conditions (i)-(iii) above is true as soon as the sub-Riemannian structure does not contain non-trivial abnormal minimizers. This condition is, in particular, satisfied for contact sub-Riemannian manifolds. The crucial fact in Corollary 12 is the asymptotic description given in (21).

*Proof.* First we discuss the existence of the set  $\Omega(\varepsilon)$  satisfying properties (i)-(iii). For  $\varepsilon > 0$  small enough, the closure of the ball  $B(p, \varepsilon)$  is compact, hence for every point  $x \in B(p, \varepsilon)$  there exists a length-minimizer  $\gamma_\lambda : [0, 1] \rightarrow M$ , associated with initial covector  $\lambda = (\rho, \theta, w) \in T_p^*M$  joining  $p$  with  $x$ . Recall that contact sub-Riemannian structures have no non-trivial abnormal minimizers, thus  $\gamma_\lambda$  does not contain any abnormal segment and cannot be minimizing after its first conjugate time. Moreover, under the assumption that there are no non-trivial abnormal minimizers, a cut time is either the first conjugate time or a point where two optimal geodesics intersect. For a proof of these statements one can see [3, Chapter 3, Chapter 8] or [14, Appendix A].

For every unit initial covector  $\lambda = (1, \theta, w)$  we have  $\gamma_\lambda(t) = \exp_p(t(1, \theta, w)) = \exp_p(t, \theta, tw)$  and this trajectory is by definition a length-minimizer up to the corresponding cut time  $t_{cut}(1, \theta, w)$ . Notice that  $\exp_p(0, \theta, w) = p$  for every  $\theta, w$ . We stress that for every  $s \in ]0, t_{cut}(\lambda)[$ , the restriction  $\gamma_\lambda|_{[0, s]}$  is a length-minimizer hence does not contain neither cut points (by construction), nor conjugate points (by length-minimality).

Let us introduce the star-shaped set in  $T_p^*M$

$$A = \{(s, \theta, sw) \in T_p^*M \mid \theta \in [0, 2\pi], w \in \mathbb{R}, 0 \leq s \leq t_{cut}(1, \theta, w)\},$$

and set  $\Omega(\varepsilon) := A \cap \{0 \leq \rho \leq \varepsilon\}$ . It follows by construction that  $\exp_p : \Omega(\varepsilon) \rightarrow \overline{B}(p, \varepsilon)$  is onto. Moreover

$$\exp_p(\text{int}(\Omega(\varepsilon) \setminus \{\rho=0\})) = B(p, \varepsilon) \setminus \text{Cut}(p),$$

which has full measure in  $B(p, \varepsilon)$ . The fact that  $\exp_p$  is injective with injective differential on the open set  $\text{int}(\Omega(\varepsilon) \setminus \{\rho=0\})$ , is a consequence of the fact that length minimizers do not contain neither cut nor conjugate points.

To complete the proof of the statement, we compute the asymptotic description of the set  $\Omega(\varepsilon)$  in the cotangent space  $T_p^*M$ . Let us rewrite the set  $A$  as follows, in cylindrical coordinates<sup>2</sup>

$$(22) \quad A = \{(\rho, \theta, w) \in T_p^*M \mid \theta \in [0, 2\pi], \rho \geq 0, |w| \leq \rho t_{cut}^{-1}(1, \theta, \rho)\},$$

<sup>2</sup>One can use the following identity, for  $\varphi : [0, +\infty) \rightarrow [0, +\infty)$  an invertible function

$$\{(s, sw) \in \mathbb{R}^2 \mid w \in \mathbb{R}, 0 \leq s \leq \varphi(|w|)\} = \{(x, y) \in \mathbb{R}^2 \mid x \geq 0, |y| \leq x\varphi^{-1}(x)\}.$$

where  $t_{cut}^{-1}(1, \theta, \cdot)$  means the inverse function of the map  $w \mapsto t_{cut}(1, \theta, w)$ , for a fixed  $\theta \in [0, 2\pi]$ . Notice that the function  $t_{cut}(1, \bar{\theta}, \cdot)$  is smooth at infinity, for fixed  $\bar{\theta} \in [0, 2\pi]$ , with derivative approaching a positive constant, and therefore it is invertible close to infinity.

The expansion of  $t_{cut}^{-1}(1, \theta, \cdot)$  at zero is then obtained from the one of  $t_{cut}(1, \theta, \cdot)$  at infinity (20) as follows

$$(23) \quad t_{cut}^{-1}(1, \theta, \rho) = \frac{2\pi}{\rho} - \frac{(\kappa + 2\chi \sin^2 \theta)}{4\pi} \rho + O(\rho^2), \quad \rho \rightarrow 0.$$

Multiplying (23) by  $\rho$  and combining with (22), one gets the statement by setting

$$f(\theta) = \frac{\kappa + 2\chi \sin^2 \theta}{4\pi} \quad \square$$

### 3. PROOF OF MAIN THEOREM

We compute the volume of the ball  $B(p, \varepsilon)$  in normal coordinates. Recall that in this coordinate chart the ball is denoted simply  $B(\varepsilon)$  and the Popp volume writes  $\mu = \psi dx dy dz$ . We have

$$\begin{aligned} \text{vol}(B(p, \varepsilon)) &= \int_{B(\varepsilon)} \psi(x, y, z) dx dy dz \\ &= \varepsilon^4 \int_{B^\varepsilon(1)} \psi(\varepsilon x, \varepsilon y, \varepsilon^2 z) dx dy dz \quad (\text{using Lemma 5}) \\ &= \varepsilon^4 \left( \int_{B^\varepsilon(1)} (1 - 2\varepsilon^2 \gamma^{[2]}(x, y, z)) dx dy dz + O(\varepsilon^3) \right) \quad (\text{using Lemma 9}) \end{aligned}$$

Using again Lemma 5, we can write

$$\begin{aligned} B^\varepsilon(1) &= \delta_{\frac{1}{\varepsilon}}(B(\varepsilon)) \\ &= \delta_{\frac{1}{\varepsilon}}(\exp(\Omega(\varepsilon))) \quad (\text{by property (i) from Corollary 12}) \\ &= \exp^\varepsilon(\tau_{\frac{1}{\varepsilon}}(\Omega(\varepsilon))) \quad (\text{by Lemma 7}) \end{aligned}$$

Observe also that  $\Omega^\varepsilon(1) \doteq \tau_{\frac{1}{\varepsilon}}(\Omega(\varepsilon))$  has the following description in cylindrical coordinates:

$$\tau_{\frac{1}{\varepsilon}}(\Omega(\varepsilon)) = \{|\rho| < 1, \theta \in [0, 2\pi], w \in [-2\pi + \varepsilon^2 \rho^2 f(\theta) + O(\varepsilon^3), 2\pi - \varepsilon^2 \rho^2 f(\theta) + O(\varepsilon^3)]\}.$$

In particular we can write:

$$(24) \quad \text{vol}(B(p, \varepsilon)) = \varepsilon^4 \left( \int_{\exp^\varepsilon(\Omega^\varepsilon(1))} (1 - 2\varepsilon^2 \gamma^{[2]}(x, y, z)) dx dy dz + O(\varepsilon^3) \right).$$

Observe now that properties (ii) and (iii) from Corollary 12 remains true if we compose the various maps with a diffeomorphism, after considering the images of the corresponding sets under the diffeomorphism itself. In particular, since both  $\delta_{\frac{1}{\varepsilon}}$  and  $\tau_{\frac{1}{\varepsilon}}$  are diffeomorphisms, we can apply the change of variable formula and compute the integral in (24) as:

$$(25) \quad \text{vol}(B(p, \varepsilon)) = \varepsilon^4 \left( \int_0^1 \int_0^{2\pi} \int_{-2\pi + \varepsilon^2 \rho^2 f(\theta) + O(\varepsilon^3)}^{2\pi - \varepsilon^2 \rho^2 f(\theta) + O(\varepsilon^3)} u(\rho, \theta, w) dw d\theta d\rho + O(\varepsilon^3) \right),$$

where

$$u(\rho, \theta, w) = (1 - 2\varepsilon^2 \gamma^{[2]}(\exp^\varepsilon(\rho, \theta, w)) | \det(J \exp^\varepsilon)(\rho, \theta, w) |).$$

We compute now the expansion in  $\varepsilon$  of the various terms involved. Let us start with  $u$ , which using the expansion  $\exp_p^\varepsilon = \exp_p^0 + O(\varepsilon)$  and (13) we can write as

$$u(\rho, \theta, w) = \det(J \exp_p^0)(\rho, \theta, w) +$$

$$\begin{aligned}
 & + \varepsilon^2 \left( -2\gamma^{[2]}(\exp_p^0(\rho, \theta, w)) \det(J \exp_p^0)(\rho, \theta, w) + v_2(\rho, \theta, w) \right) + O(\varepsilon^3) \\
 & = u_0(\rho, \theta, w) + \varepsilon^2 u_2(\rho, \theta, w) + O(\varepsilon^3)
 \end{aligned}$$

Observe now also that:

$$\begin{aligned}
 \int_{-2\pi+\varepsilon^2\rho^2f(\theta)+O(\varepsilon^3)}^{2\pi-\varepsilon^2\rho^2f(\theta)+O(\varepsilon^3)} u(\rho, \theta, w) dw & = \int_{-2\pi+\varepsilon^2\rho^2f(\theta)+O(\varepsilon^3)}^{2\pi-\varepsilon^2\rho^2f(\theta)+O(\varepsilon^3)} u_0(\rho, \theta, w) + \varepsilon^2 u_2(\rho, \theta, w) dw + O(\varepsilon^3) \\
 & = \int_{-2\pi}^{2\pi} u_0(\rho, \theta, w) + \varepsilon^2 u_2(\rho, \theta, w) dw + \\
 & \quad - 2\varepsilon^2 \rho^2 f(\theta) (u_0(\rho, \theta, -2\pi) + u_0(\rho, \theta, 2\pi)) + O(\varepsilon^3) \\
 (26) \qquad \qquad \qquad & = \int_{-2\pi}^{2\pi} u_0(\rho, \theta, w) dw + \varepsilon^2 \int_{-2\pi}^{2\pi} u_2(\rho, \theta, w) dw + O(\varepsilon^3),
 \end{aligned}$$

where in the last line we have used the crucial fact that  $u_0(\rho, \theta, 2\pi) = u_0(\rho, \theta, -2\pi) = 0$ , as it can be immediately verified from (14). Recall that  $u_0 = \det(J \exp_p^0)$  is the Jacobian determinant in the Heisenberg group. In more geometric terms, the last equality is saying that the cut time coincides also with the first conjugate time in the Heisenberg group.

Consider now the fixed domain  $\Omega = \{|\rho| \leq 1, \theta \in [0, 2\pi], w \in [-2\pi, 2\pi]\}$ . Plugging (26) into (25), we obtain:

$$(27) \qquad \qquad \qquad \text{vol}(B(p, \varepsilon)) = \varepsilon^4 \left( \int_{\Omega} u_0 + \varepsilon^2 \int_{\Omega} u_2 + O(\varepsilon^3) \right).$$

From the definition of  $u_0 = \det(J \exp_p^0)$ , we immediately recognize

$$c_0 := \int_{\Omega} u_0 = \text{volume of the unit ball in the Heisenberg group.}$$

For the integral of  $u_2$ , we proceed analyzing the various functions appearing in its definition

$$u_2 = -2\gamma^{[2]}(\exp_p^0(\rho, \theta, w)) \det(J \exp_p^0)(\rho, \theta, w) + v_2(\rho, \theta, w).$$

Writing  $\gamma^{[2]}(x, y, z) = ax^2 + 2bxy + cy^2$  as in (17) and using (16), we have:

$$\begin{aligned}
 -2\gamma^{[2]}(\exp_p^0) \det(J \exp_p^0) & = 4(a+c) \frac{\rho^5 \sin(\frac{w}{2})^2 (2 \cos w + w \sin w - 2)}{w^6} + \\
 & \quad - \frac{4\rho^2 \sin(\frac{w}{2})^2 ((a-c) \cos(2\theta+w) + b \sin(2\theta+w))}{w^2} \det(J \exp_p^0) \\
 (28) \qquad \qquad \qquad & = 2\kappa(p) \frac{\rho^5 \sin(\frac{w}{2})^2 (2 \cos w + w \sin w - 2)}{w^6} + \\
 & \quad + \cos(2\theta+w) g_1(\rho, w) + \sin(2\theta+w) g_2(\rho, w),
 \end{aligned}$$

where in the last line we have used the fact that  $2(a+c) = \kappa(p)$  and that  $\det(J \exp_p^0)$  only depends on  $(\rho, w)$  (see the explicit expression (14)). Note in particular that, exchanging the order of integration and using the fact that for every fixed  $w \in [0, 2\pi]$  the integrals  $\int_0^{2\pi} \cos(2\theta+w) d\theta$  and  $\int_0^{2\pi} \sin(2\theta+w) d\theta$  vanish, (28) implies:

$$\begin{aligned}
 \int_{\Omega} -2\gamma^{[2]}(\exp_p^0) \det(J \exp_p^0) & = 2\kappa(p) \int_0^1 \int_{-2\pi}^{2\pi} 2\pi \frac{\rho^5 \sin(\frac{w}{2})^2 (2 \cos w + w \sin w - 2)}{w^6} dw d\rho \\
 (29) \qquad \qquad \qquad & = \kappa(p) \int_{-2\pi}^{2\pi} \frac{2\pi \sin(\frac{w}{2})^2 (2 \cos w + w \sin w - 2)}{3 w^6} dw.
 \end{aligned}$$

Let us look now at the integral of the function  $v_2$ . Using its explicit expression and integrating the  $\theta$ -variable first, we obtain:

$$\begin{aligned}
\int_{\Omega} v_2 &= \int_{\Omega} \rho^5 \left( \kappa(p) \frac{g_0(w)}{2} + g_c(w) \cos 2\theta + g_s(w) \sin 2\theta \right) dw d\theta d\rho \\
(30) \quad &= \kappa(p) \int_0^1 \int_{-2\pi}^{2\pi} \rho^5 2\pi \frac{g_0(w)}{2} dw d\rho \\
&= \kappa(p) \int_{-2\pi}^{2\pi} \frac{\pi}{6} g_0(w) dw.
\end{aligned}$$

Combining (29) and (30) we obtain:

$$\begin{aligned}
\int_{\Omega} u_2 &= \kappa(p) \int_{-2\pi}^{2\pi} \left( \frac{2\pi \sin(\frac{w}{2})^2 (2 \cos w + w \sin w - 2)}{3 w^6} + \frac{\pi}{6} g_0(w) \right) dw \\
&= \kappa(p) \int_{-2\pi}^{2\pi} \frac{\pi (5w \sin w - (w^2 - 8) \cos w - 8)}{2 w^6} dw \\
&= \kappa(p) \frac{1}{160} \left( \frac{1}{\pi^2} - 2 - 4\pi \text{Si}(2\pi) \right).
\end{aligned}$$

Together with (27) this finally gives:

$$\text{vol}(B(p, \varepsilon)) = \varepsilon^4 c_0 \left( 1 - \kappa(p) c_1 \varepsilon^2 + O(\varepsilon^3) \right),$$

where:

$$c_1 = \frac{1}{c_0 160} \left( 2 + 4\pi \text{Si}(2\pi) - \frac{1}{\pi^2} \right) > 0 \quad \text{and} \quad c_0 = \frac{1}{12} (1 + 2\pi \text{Si}(2\pi)).$$

The explicit formula of  $c_0$ , which is the volume of the unit ball in the Heisenberg group, coincides with the one obtained in [2, Remark 39].

#### APPENDIX A. REMARKS ON CURVATURE COEFFICIENTS

In this appendix we relate the geometric invariants  $\chi$  and  $\kappa$  defined in Section 2.1 with the canonical connection that is available on a 3-dimensional contact manifold.

We extend the sub-Riemannian metric  $g$  on  $D$  to a global Riemannian structure (that we denote with the same symbol  $g$ ) by promoting  $X_0$  to a unit vector orthogonal to  $D$ .

We define the *contact endomorphism*  $J : TM \rightarrow TM$  by:

$$g(X, JY) = d\omega(X, Y), \quad \forall X, Y \in \Gamma(TM).$$

Clearly  $J$  is skew-symmetric w.r.t. to  $g$ . In the 3-dimensional case, the previous condition forces  $J^2 = -\mathbb{I}$  on  $D$  and  $J(X_0) = 0$ .

**Theorem 13** (canonical connection, [20, 28, 22]). *There exists a unique linear connection  $\nabla$  on  $(M, \omega, g, J)$  such that*

- (i)  $\nabla \omega = 0$ ,
- (ii)  $\nabla X_0 = 0$ ,
- (iii)  $\nabla g = 0$ ,
- (iv)  $T(X, Y) = d\omega(X, Y)X_0$  for any  $X, Y \in \Gamma(D)$ ,
- (v)  $T(X_0, JX) = -JT(X_0, X)$  for any vector field  $X \in \Gamma(TM)$ ,

where  $T$  is the torsion tensor of  $\nabla$ .

In terms of this connection we have  $\kappa = R^\nabla(X_1, X_2, X_2, X_1)$ , where  $R^\nabla$  is the curvature associated with  $\nabla$ .

If  $X$  is a horizontal vector field, so is  $T(X_0, X)$ . As a consequence, if we define  $\tau(X) = T(X_0, X)$ ,  $\tau$  is a symmetric horizontal endomorphism which satisfies  $\tau \circ J + J \circ \tau = 0$ , by property (v).

Notice that  $\text{trace}(\tau) = 0$ . It holds  $\chi = \sqrt{-\det(\tau)^2}$ . A contact structure is  $K$ -type iff  $X_0$  is a Killing vector field or, equivalently, if  $\tau = 0$ .

## REFERENCES

- [1] A. Agrachev and D. Barilari. Sub-Riemannian structures on 3D Lie groups. *J. Dyn. Control Syst.*, 18(1):21–44, 2012.
- [2] A. Agrachev, D. Barilari, and U. Boscain. On the Hausdorff volume in sub-Riemannian geometry. *Calc. Var. Partial Differential Equations*, 43(3-4):355–388, 2012.
- [3] A. Agrachev, D. Barilari, and U. Boscain. Introduction to Riemannian and sub-Riemannian geometry (Lecture Notes). <http://webusers.imj-prg.fr/~davide.barilari/notes.php>. 2016. v20/11/16.
- [4] A. Agrachev, D. Barilari, and E. Paoli. Volume geodesic distortion and Ricci curvature for Hamiltonian dynamics. *ArXiv e-prints*, Feb. 2016.
- [5] A. Agrachev, D. Barilari, and L. Rizzi. Sub-Riemannian curvature in contact geometry. *J. Geom. Anal.*, 27(1):366–408, 2017.
- [6] A. A. Agrachev. Methods of control theory in nonholonomic geometry. In *Proceedings of the International Congress of Mathematicians, Vol. 1, 2 (Zürich, 1994)*, pages 1473–1483. Birkhäuser, Basel, 1995.
- [7] A. A. Agrachev. Exponential mappings for contact sub-Riemannian structures. *J. Dynam. Control Systems*, 2(3):321–358, 1996.
- [8] A. A. Agrachev, D. Barilari, and L. Rizzi. Curvature: a variational approach. *Memoirs of the AMS, in press*, 2013. <https://arxiv.org/pdf/1306.5318.pdf>.
- [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)*, volume 25 of *CMS Conf. Proc.*, pages 29–78. Amer. Math. Soc., Providence, RI, 1998.
- [10] D. Barilari. Trace heat kernel asymptotics in 3D contact sub-Riemannian geometry. *J. Math. Sci. (N.Y.)*, 195(3):391–411, 2013. Translation of *Sovrem. Mat. Prilozh.* No. 82 (2012).
- [11] D. Barilari and L. Rizzi. A formula for Popp’s volume in sub-Riemannian geometry. *Anal. Geom. Metr. Spaces*, 1:42–57, 2013.
- [12] D. Barilari and L. Rizzi. Comparison theorems for conjugate points in sub-Riemannian geometry. *ESAIM Control Optim. Calc. Var.*, 22(2):439–472, 2016.
- [13] D. Barilari and L. Rizzi. On Jacobi fields and a canonical connection in sub-Riemannian geometry. *Archivum Mathematicum*, 53(2):77–92, 2017.
- [14] D. Barilari and L. Rizzi. Sub-Riemannian interpolation inequalities: ideal structures. *ArXiv e-prints*, May 2017.
- [15] F. Baudoin. Sub-Laplacians and hypoelliptic operators on totally geodesic Riemannian foliations. In *Geometry, analysis and dynamics on sub-Riemannian manifolds. Vol. 1*, EMS Ser. Lect. Math., pages 259–321. Eur. Math. Soc., Zürich, 2016.
- [16] F. Baudoin and N. Garofalo. Curvature-dimension inequalities and Ricci lower bounds for sub-Riemannian manifolds with transverse symmetries. *J. Eur. Math. Soc. (JEMS)*, 19(1):151–219, 2017.
- [17] F. Baudoin and J. Wang. The subelliptic heat kernel on the CR sphere. *Math. Z.*, 275(1-2):135–150, 2013.
- [18] F. Baudoin and J. Wang. The subelliptic heat kernels of the quaternionic Hopf fibration. *Potential Anal.*, 41(3):959–982, 2014.
- [19] R. Beals, P. C. Greiner, and N. K. Stanton. The heat equation on a CR manifold. *J. Differential Geom.*, 20(2):343–387, 1984.
- [20] D. E. Blair. *Riemannian geometry of contact and symplectic manifolds*, volume 203 of *Progress in Mathematics*. Birkhäuser Boston, Inc., Boston, MA, second edition, 2010.
- [21] E.-H. C. El-Alaoui, J.-P. Gauthier, and I. Kupka. Small sub-Riemannian balls on  $\mathbf{R}^3$ . *J. Dynam. Control Systems*, 2(3):359–421, 1996.
- [22] E. Falbel and C. Gorodski. Sub-Riemannian homogeneous spaces in dimensions 3 and 4. *Geom. Dedicata*, 62(3):227–252, 1996.
- [23] D. Jerison and A. Sánchez-Calle. Subelliptic, second order differential operators. In *Complex analysis, III (College Park, Md., 1985–86)*, volume 1277 of *Lecture Notes in Math.*, pages 46–77. Springer, Berlin, 1987.
- [24] 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(4):835–854, 1986.
- [25] R. Montgomery. *A tour of subriemannian geometries, their geodesics and applications*, volume 91 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 2002.

- [26] A. Sánchez-Calle. Fundamental solutions and geometry of the sum of squares of vector fields. *Invent. Math.*, 78(1):143–160, 1984.
- [27] N. K. Stanton and D. S. Tartakoff. The heat equation for the  $\bar{\partial}_b$ -Laplacian. *Comm. Partial Differential Equations*, 9(7):597–686, 1984.
- [28] S. Tanno. Variational problems on contact Riemannian manifolds. *Trans. Amer. Math. Soc.*, 314(1):349–379, 1989.
- [29] J. Wang. The subelliptic heat kernel on the anti-de Sitter space. *Potential Anal.*, 45(4):635–653, 2016.