MODIFIED ACTION AND DIFFERENTIAL OPERATORS ON THE 3-D SUB-RIEMANNIAN SPHERE

DER-CHEN CHANG, IRINA MARKINA, AND ALEXANDER VASIL'EV

Abstract. Our main aim is to present a geometrically meaningful formula for the fundamental solution to a second order sub-elliptic differential operator and the heat kernel associated with this operator in the sub-Riemannian geometry on the unit sphere $\mathbb{S}^3$. Our method is based on the Hamiltonian approach, where the corresponding Hamitonian system is solved with mixed boundary conditions. A closed form of the modified action which play the role of distance on $\mathbb{S}^3$, and which is a sub-Riemannian invariant, is presented.

1. Introduction

The unit 3-sphere centered on the origin is the set of $\mathbb{R}^4$ defined as

$$\mathbb{S}^3 = \{(x_1, x_2, x_3, x_4) \in \mathbb{R}^4 : \ x_1^2 + x_2^2 + x_3^2 + x_4^2 = 1)\}.$$

It is often convenient to regard $\mathbb{R}^4$ as the space of quaternions $\mathbb{H}$. The unit 3-sphere is then given by

$$\mathbb{S}^3 = \{q \in \mathbb{H} : \ |q|^2 = 1)\}.$$

This description represents the sphere $\mathbb{S}^3$ as a set of unit quaternions and it can be considered as the spin group $Sp(1)$, where the group operation is just a multiplication of quaternions. Let us identify $\mathbb{R}^3$ with pure imaginary quaternions. The conjugation $qh\bar{q}$ of a pure imaginary quaternion h by a unit quaternion q defines rotation in $\mathbb{R}^3$, and since $|qh\bar{q}| = |h|$, the map $h \mapsto qh\bar{q}$ defines a two-to-one homomorphism $Sp(1) \to SO(3)$. The Hopf map $\pi : \mathbb{S}^3 \to \mathbb{S}^2$ can be defined by

$$\mathbb{S}^3 \ni q \mapsto qi\bar{q} = \pi(q) \in \mathbb{S}^2.$$

The Hopf map defines a principle circle bundle also known as the Hopf bundle.

The sub-Riemannian structure on S^3 comes naturally from the non-commutative group structure of the sphere in the sense that two vector fields span the smoothly varying distribution of the tangent bundle and their commutator generates the missing direction. The missing direction is also can be obtained as an integral line of the Hopf vector field related to the Hopf fibration. The sub-Riemannian geometry on S^3 was studied in [10, 12, 17]. Explicit formulas for al geodesics were given in [12]. Let us mention

2000 *Mathematics Subject Classification.* Primary: 53C17; Secondary: 70H05 .

Key words and phrases. Sub-Riemannian geometry, action, sub-Laplacian, heat kernel, geodesic, Hamiltonian system, optimal control.

The first author is partially supported by a research grant from the United State Army Research Office and a Hong Kong RGC competitive earmarked research grant #600607. The second and the third authors have been supported by the grant of the Norwegian Research Council #177355/V30, and by the European Science Foundation Research Networking Programme HCAA.

that the word 'geodesic' in our terminology stands for solutions to a Hamiltonian system, that is a good generalization of the notion of geodesic to the sub-Riemannian manifolds, see for instance [24, 28]. The Lagrangian approach was applied in [10] and [17] in order to characterize and find the shortest geodesics.

Our main aim in this paper is to deduce a geometrically meaningful formula for the Green function for a second order sub-elliptic differential operator and the heat kernel associated with this operator in the sub-Riemannian geometry on the unit sphere $\mathbb{S}^3$. There is a vast amount of literature studying sub-elliptic operators. Here we give only a very few of possible references [14, 16, 18, 19, 27, 29]. Our method is based on the Hamiltonian approach. Analogously to Hadamard's method for strictly hyperbolic operators our method essentially uses two important ingredients:

- Solution of the Hamiltonian system with mixed boundary conditions and construction of modified action on solutions to this systems;
- Solution of the corresponding transport equations and deduction of the volume elements.
- Integration of the function of modified action over the characteristic variety with respect the measure defined by the volume element.

This method was realized for the two step nilpotent groups, for instance, in series of papers [1, 2, 3] where the geometrical meaning of the fundamental solutions was revealed. In other geometries see, for example [5, 4, 11]. The case of 3-sphere reveals new features and being non-nilpotent group it is not a direct analog of previous considerations.

The structure of the paper is as follows. The classical setup for the heat kernels in the Riemannian case is presented in Section 2. In Section 3 we define the horizontal distribution and sub-Riemannian metric. The Hamiltonian system is derived in the fourth section. In the next section we treat the problem of finding geodesics as an optimal control problem. Symmetries of the Hamiltonain system are discussed. In Section 6 we solve the Hamiltonian system to find geodesics and to solve the boundary value problem. The number of geodesics connecting two fixed points on $\mathbb{S}^3$ is found. Both cartesian and hyperspherical coordinates were used. At the end of this section we define the modified action and investigate its properties. Special directions in the cotangent bundle given by the Hamiltonian system are revealed clearly in the hyperspherical coordinates contrasting with the cartesian ones. We use these directions and construct the modified action making use of the Hamiltonian system with non-standard mixed boundary values. The modified action satisfies a generalized Hamilton-Jacobi equation (Section 7). It is a sub-Riemannian invariant of $\mathbb{S}^3$ and it is used for construction of a distance function (Section 8). The distance function is involved into the fundamental solutions to the sub-Laplacian equation and the heat equation associated to the sub-Laplacian. The concluding Section 9 is concerned with the finding of the volume element. The sub-Laplacian and the heat operator associated with this sub-Laplacian are not elliptic, they degenerate along a singular manifold of dimension one in the cotangent space. The fundamental solutions to this equations can be obtained by integrating the distance function over this one-dimensional singular manifold which is the characteristic manifold of the corresponding Hamiltonian with respect of a special measure with the density called the *volume element*. Unlike the case of the nilpotent groups the volume element depends on the phase variables which does not permit to find its explicit form. Instead we present differential equations, called the *transport equations* which solutions give the necessary volume elements.

The paper was initiated when the authors visited the National Center for Theoretical Sciences and National Tsing Hua University during May 2008. They would like to express their profound gratitude to Professor Jing Yu for the invitation and for the warm hospitality of the staff extended to them during their stay in Taiwan.

2. Heat kernel in $\mathbb{R}^n$

Let us present some simple calculations in R^n for the heat operator motivating further generalizations for the case of sub-Riemannian geometry of $\mathbb{S}^3$. Let $\Delta = \frac{1}{2}\sum_{j=1}^{n}\frac{\partial}{\partial x_j}$ be the Laplace operator. Then the kernel $P_u(x,x_0)$ for the operator $\Delta - \frac{\partial}{\partial u}$ is given by

$$P_u(x,x_0) = \frac{1}{(2\pi u)^{\frac{n}{2}}}e^{-\frac{|x-x_0|^2}{2u}}.$$

If we write $f = \frac{1}{2}|x-x_0|^2$, then it is easy to see that the function $\frac{f}{u}$ satisfies the Hamilton-Jacobi equation

$$\frac{\partial}{\partial u}\Big(\frac{f}{u}\Big) + \frac{1}{2}\sum_{j=1}^{n}\Big(\frac{\partial}{\partial x_j}\Big(\frac{f}{u}\Big)\Big)^2 = 0, \qquad \frac{1}{2}\sum_{j=1}^{n}\Big(\frac{\partial}{\partial x_j}\Big(\frac{f}{u}\Big)\Big)^2 = H\Big(\nabla\Big(\frac{f}{u}\Big)\Big),$$

where H is a Hamiltonian function associated with the Laplace operator Δ. In the standard theory, the function $S = \frac{f}{u}$ is the classical action related to the Hamiltonian H.

In the case of a general second order elliptic operator defined by smooth linearly independent vector fields X_j, $j = 1,\dots,n$ in $\mathbb{R}^n$, the heat kernel $P_u(x,x_0)$ for the operator

$$\Delta_X - \frac{\partial}{\partial u}, \quad \text{where} \quad \Delta_X = \frac{1}{2}\sum_{j=1}^{n}X_j$$

admits the form

$$P_u(x,x_0) = \frac{1}{(2\pi u)^{\frac{n}{2}}}e^{-\frac{|x-x_0|^2}{2u}}(v_0 + v_1 u + v_2 u^2 + \dots),$$

where the function $\frac{|x-x_0|^2}{2u}$ still satisfies the Hamilton equation with respect to the vector fields X_j. Associated Hamiltonian is degenerating only at one point of $\mathbb{R}^n \times \mathbb{R}^n$ and the constants v_l is chosen so that the delta function supported at x_0 is seen.

Let us turn to the vector fields $X_1,\dots,X_k$ satisfying the Chow (or bracket generating) condition on n-dimensional manifold M, $k < n$. In this case the operator $\Delta_X = \frac{1}{2}\sum_{j=1}^{k}X_j$ is sub-elliptic and degenerates over a set of positive measure. Previous studies (see, e.g., [1, 2, 3, 4, 5, 11]) show that it is reasonable to expect the heat kernel $P_u(x,x_0)$ for the operator associated with the sub-Laplacian Δ_X in the form

$$P_u(x,x_0) = \frac{C}{u^q}\int\limits_{chv(H)_{x_0}(\tau)} e^{-\frac{f(x,x_0,\tau)}{u}}v(x,u,\tau)\,d\tau.$$

Here $chv(H)_{x_0}$ is the characteristic variety of the Hamilton function at x_0 associated with the sub-Laplacian Δ_X defined by

$$chv(H) = \big\{(x,\xi) \in T^*M : \ H(x,\xi) = 0\big\}.$$

The characteristic variety represents the singular set of the sub-elliptic operator. The function $f(x, x_0, \tau)$ plays the role of the square of the distance between the points x_0 and x on the manifold M and satisfies the generalized Hamilton-Jacobi equation

$$\tau \nabla_\tau f + H(x, \nabla_x f) = f.$$

The function f is a modified action associated with the degenerating Hamiltonian. The term $v(x,\tau)$ is a suitable measure on the characteristic variety $chv(H)_{x_0}$ at x_0 making the integral convergent. It is called volume element and it can be found from a differential equation known as the transport equation.

3. Horizontal distribution on $\mathbb{S}^3$

Let us turn to sub-Riemannian geometry of $\mathbb{S}^3$. In order to calculate left-invariant vector fields we use the definition of $\mathbb{S}^3$ as a set of unit quaternions equipped with the following non-commutative multiplication '$\circ$': if $x = (x_1, x_2, x_3, x_4)$ and $y = (y_1, y_2, y_3, y_4)$, then

$$\begin{aligned} x \circ y = (x_1, x_2, x_3, x_4) \circ (y_1, y_2, y_3, y_4) \ = \ & \Big((x_1y_1 - x_2y_2 - x_3y_3 - x_4y_4), \\ & (x_2y_1 + x_1y_2 - x_4y_3 + x_3y_4), \\ & (x_3y_1 + x_4y_2 + x_1y_3 - x_2y_4), \\ & (x_4y_1 - x_3y_2 + x_2y_3 + x_1y_4)\Big). \end{aligned} \tag{3.1}$$

The rule (3.1) gives us the left translation $L_x(y)$ of an element $y = (y_1, y_2, y_3, y_4)$ by the element $x = (x_1, x_2, x_3, x_4)$. The left-invariant basis vector fields are defined as $X(x) = (L_x(y))_* X(0)$, where $X(0)$ are the basis vectors at the unity of the group. Calculating the action of $(L_x(y))_*$ in the basis of unit vectors of $\mathbb{R}^4$ we get four left-invariant vector fields

$$\begin{aligned} X_1(x) &= x_1\partial_{x_1} + x_2\partial_{x_2} + x_3\partial_{x_3} + x_4\partial_{x_4}, \\ X_2(x) &= -x_2\partial_{x_1} + x_1\partial_{x_2} + x_4\partial_{x_3} - x_3\partial_{x_4}, \\ X_3(x) &= -x_3\partial_{x_1} - x_4\partial_{x_2} + x_1\partial_{x_3} + x_2\partial_{x_4}, \\ X_4(x) &= -x_4\partial_{x_1} + x_3\partial_{x_2} - x_2\partial_{x_3} + x_1\partial_{x_4}. \end{aligned} \tag{3.2}$$

It is easy to see that the vector $X_1(x)$ is the unit normal to $\mathbb{S}^3$ at x with respect to the usual inner product $\langle\cdot,\cdot\rangle$ in $\mathbb{R}^4$, hence, we denote $X_1(x)$ by N. Moreover, the vector fields $X_2(x)$, $X_3(x)$, $X_4(x)$ form an orthonormal basis of the tangent space $T_x\mathbb{S}^3$ with respect to $\langle\cdot,\cdot\rangle$ at any point $x \in \mathbb{S}^3$. Let us denote these vector fields by

$$X_3 = X, \quad X_4 = Y, \quad X_2 = Z.$$

The vector fields possess the following commutation relations

$$[X, Y] = XY - YX = 2Z, \quad [Z, X] = 2Y, \quad [Y, Z] = 2X.$$

Let $\mathcal{D} = \text{span}\{X, Y\}$ be the distribution generated by the vector fields X and Y. Since $[X, Y] = 2Z \notin \mathcal{D}$, it follows that $\mathcal{D}$ is not involutive. The distribution $\mathcal{D}$ will be called *horizontal*. Any curve on the sphere with the velocity vector contained in the distribution $\mathcal{D}$ will be called a *horizontal curve*. Since $T_x\mathbb{S}^3 = \text{span}\{X, Y, Z = 1/2[X, Y]\}$, the distribution is bracket generating at each point $x \in S^3$. We define the metric on the distribution $\mathcal{D}$ as the restriction of the metric $\langle\cdot,\cdot\rangle$ to $\mathcal{D}$, and the same notation $\langle\cdot,\cdot\rangle$ will be used. The manifold $(\mathbb{S}^3, \mathcal{D}, \langle\cdot,\cdot\rangle)$ is a step two sub-Riemannian manifold.

Remark 1. Notice that the choice of the horizontal distribution is not unique. The relations $[Z, X] = 2Y$ and $[Y, Z] = 2X$ imply possible choices $\mathcal{D} = \text{span}\{X, Z\}$ or $\mathcal{D} = \text{span}\{Y, Z\}$. The geometries defined by different horizontal distributions are cyclically symmetric, so we restrict our attention to the $\mathcal{D} = \text{span}\{X, Y\}$.

We also can define the distribution as a kernel of the following one form

$$\omega = -x_2dx_1 + x_1dx_2 + x_4dx_3 - x_3dx_4$$

on $\mathbb{R}^4$. One can easily check that

$$\omega(X) = 0, \quad \omega(Y) = 0, \quad \omega(Z) = 1 \neq 0, \quad \omega(N) = 0.$$

Hence, the horizontal distribution $\mathcal{D}_x$ at $x \in S^3$ can be written as $\ker \omega_x \cap T_x\mathbb{S}^3$. The one form ω has the following geometric meaning. It is the difference of two independent area forms $\alpha = -x_2dx_1 + x_1dx_2$ in (x_1, x_2)-plane and $\beta = -x_4dx_3 + x_3dx_4$ in (x_3, x_4)-plane.

Let $\gamma(s) = (x_1(s), x_2(s), x_3(s), x_4(s))$ be a curve on $\mathcal{S}^3$. Then the velocity vector, written in the left-invariant basis, is

$$\dot\gamma(s) = a(s)X(\gamma(s)) + b(s)Y(\gamma(s)) + c(s)Z(\gamma(s)),$$

where

$$\begin{aligned} a &= \langle\dot\gamma, X\rangle = -x_3\dot x_1 - x_4\dot x_2 + x_1\dot x_3 + x_2\dot x_4, \\ b &= \langle\dot\gamma, Y\rangle = -x_4\dot x_1 + x_3\dot x_2 - x_2\dot x_3 + x_1\dot x_4, \\ c &= \langle\dot\gamma, Y\rangle = -x_2\dot x_1 + x_1\dot x_2 + x_4\dot x_3 - x_3\dot x_4. \end{aligned} \tag{3.3}$$

The following proposition holds.

Proposition 1. *Let $\gamma(s) = (x_1(s), x_2(s), x_3(s), x_4(s))$ be a curve on $\mathbb{S}^3$. The curve γ is horizontal, if and only if,*

$$c = \langle\dot\gamma, Z\rangle = \langle\dot\gamma, X\rangle = -x_2\dot x_1 + x_1\dot x_2 + x_4\dot x_3 - x_3\dot x_4 = 0. \tag{3.4}$$

If we take into account the geometric meaning of the one form ω, then we can reformulate Proposition 1 in the following way. Let us denote by A the area swept by the projection of the horizontal curve γ onto the (x_1, x_2)-plane and the straight line connecting its ends, and by B we denote the area swept by the projection of the horizontal curve onto the (x_3, x_4)-plane and the straight line connecting the ends of horizontal curve.

Proposition 2. *Let $\gamma(s) = (x_1(s), x_2(s), x_3(s), x_4(s))$ be a curve on $\mathbb{S}^3$ and A, B be as introduced above. Then the curve γ is horizontal, if and only if, $A = B$.*

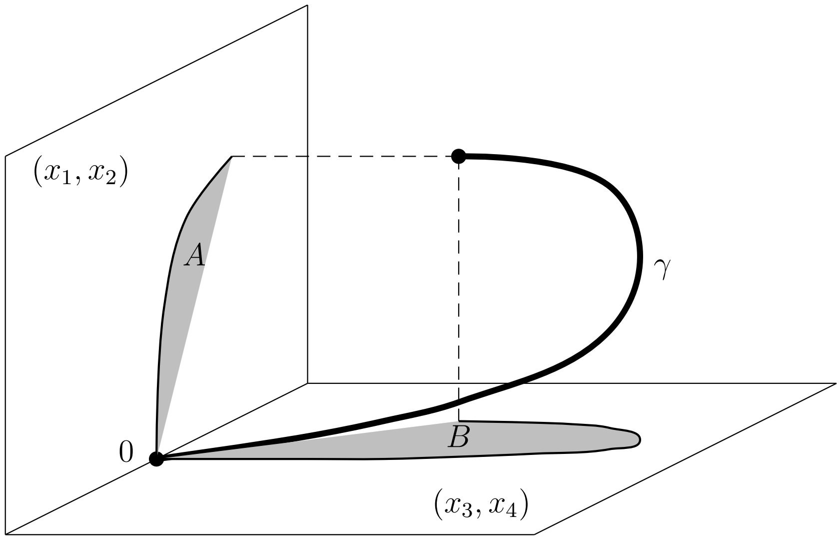

FIGURE 1. Projections of γ to the planes (x_1, x_2) and (x_3, x_4) in Proposition 2

The manifold $\mathbb{S}^3$ is connected and it satisfies the bracket generating condition. By the Chow theorem [13], there exist piecewise C^1 horizontal curves connecting two arbitrary points on $\mathbb{S}^3$. In fact, smooth horizontal curves connecting two arbitrary points on $\mathbb{S}^3$ were constructed in [10, 12].

Proposition 3. *The horizontality property is invariant under the left translation.*

Proof. It can be shown that (3.3) does not change under the left translation. This implies the conclusion of the proposition. ☐

4. Hamiltonian system

Ones we have a system of curves, in our case the system of horizontal curves, we can define their length as in the Riemannian geometry. Let $\gamma : [0, 1] \to \mathbb{S}^3$ be a horizontal curve such that $\gamma(0) = x$, $\gamma(1) = y$, then the length $l(\gamma)$ of γ is defined as follows

$$l(\gamma) = \int_0^1 \langle \dot{\gamma}, \dot{\gamma} \rangle^{1/2}\, dt = \int_0^1 \big(a^2(t) + b^2(t)\big)^{1/2}\, dt. \tag{4.1}$$

Now we are able to define the distance between two points x and y by minimizing integral (4.1) or the corresponding energy integral $\int_0^1 \big(a^2(t) + b^2(t)\big)\, dt$ under the non-holonomic constraint (3.4). This is the Lagrangian approach. The Lagrangian formalism was applied to study sub-Riemannian geometry of $\mathbb{S}^3$ in [10, 17]. In the Riemannian geometry the minimizing curve locally coinsides with the geodesic, but it is not the case for sub-Riemannian manifolds. Interesting examples and discussions can be found, for instance in [. Given the sub-Riemannian metric we can form the Hamiltonian function defined on the cotangent bundle of $\mathbb{S}^3$. A geodesic on a sub-Riemannian manifold is defined as the projection of the solution to the corresponding Hamiltonian system onto the manifold. It is a good generalization of the Riemannian case in the following sense. The Riemannian geodesics (which are defined as curves with vanishing acceleration) can be lifted to the solutions of the Hamilton system on the cotangent bundle.

In the present paper we are interested in the construction of sub-Riemannian geodesics on $(\mathbb{S}^3, \mathcal{D}, \langle\cdot,\cdot\rangle)$. Let us write the left-invariant vector fields X, Y, Z, using the matrices

$$I_1 = \begin{bmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}, \quad I_2 = \begin{bmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}, \quad I_3 = \begin{bmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{bmatrix}.$$

Then

$$X = \langle I_1 x, \nabla x\rangle, \quad Y = \langle I_2 x, \nabla x\rangle, \quad Z = \langle I_3 x, \nabla x\rangle.$$

The Hamiltonian is defined as

$$H = \frac{1}{2}(X^2 + Y^2) = \frac{1}{2}\Big(\langle I_1 x, \xi\rangle^2 + \langle I_2 x, \xi\rangle^2\Big),$$

or

$$H = \frac{1}{2}(-x_3\xi_1 - x_4\xi_2 + x_1\xi_3 + x_2\xi_4)^2 + \frac{1}{2}(-x_4\xi_1 + x_3\xi_2 - x_2\xi_3 + x_1\xi_4)^2, \tag{4.2}$$

where $\xi = \nabla x$. Then the Hamiltonian system follows as

$$\begin{aligned} \dot{x} &= \frac{\partial H}{\partial \xi} \quad \Rightarrow \quad \dot{x} = \langle I_1 x, \xi\rangle \cdot (I_1 x) + \langle I_2 x, \xi\rangle \cdot (I_2 x) \\ \dot{\xi} &= -\frac{\partial H}{\partial x} \quad \Rightarrow \quad \dot{\xi} = \langle I_1 x, \xi\rangle \cdot (I_1 \xi) + \langle I_2 x, \xi\rangle \cdot (I_2 \xi). \end{aligned} \tag{4.3}$$

As it was mentioned, a geodesic is the projection of a solution to the Hamiltonian system onto the x-space. We obtain the following properties.

1. Since $\langle I_1 x, x\rangle = \langle I_2 x, x\rangle = \langle I_3 x, x\rangle = 0$, multiplying the first equation of (4.3) by x we get
$$\langle \dot{x}, x\rangle = 0 \quad \Rightarrow \quad |x|^2 = const.$$
We conclude that *any solution to the Hamiltonian system belongs to the sphere.* Taking the constant equal to 1 we get geodesics on $\mathbb{S}^3$.
2. Multiplying the first equation of (4.3) by $I_3 x$, we get
$$\langle \dot{x}, I_3 x\rangle = 0, \tag{4.4}$$
by the rule of multiplication for I_1, I_2, and I_3. The reader easily recognizes the horizontality condition $\langle \dot{x}, Z\rangle = 0$ in (4.4). It means that *any solution to the Hamiltonian system is a horizontal curve.*
3. Multiplying the first equation of (4.3) by $I_1 x$, and then by $I_2 x$, we get
$$\langle \xi, I_1 x\rangle = \langle \dot{x}, I_1 x\rangle, \qquad \langle \xi, x I_2\rangle = \langle \dot{x}, x I_2\rangle.$$
From the other side, we know that $\langle \dot{x}, I_1 x\rangle = a$ and $\langle \dot{x}, x I_2\rangle = b$. The Hamiltonian function can be written in the form
$$H = \frac{1}{2}\Big(\langle I_1 x, \xi\rangle^2 + \langle I_2 x, \xi\rangle^2\Big) = \frac{1}{2}\Big(\langle I_1 x, \dot{x}\rangle^2 + \langle I_2 x, \dot{x}\rangle^2\Big) = \frac{1}{2}\Big(a^2 + b^2\Big).$$

Thus, *the Hamiltonian function gives the kinetic energy* $H = \frac{|\dot q|^2}{2}$ *and it is a constant along the geodesics.*

4. If we multiply the first equation of (4.3) by $\dot x$, then we get

$$|\dot x|^2 = \langle I_1 x, \xi\rangle^2 + \langle I_2 x, \xi\rangle^2 = \langle I_1 x, \dot x\rangle^2 + \langle I_2 x, \dot x\rangle^2 = a^2 + b^2 = 2H.$$

Therefore

$$|\dot x|^2 = a^2 + b^2. \tag{4.5}$$

The following theorem was proved in [12] and [17].

Theorem 1. *The set of geodesics with constant velocity coordinates form the unit sphere* $\mathbb{S}^2$ *in* $\mathbb{R}^3$

5. Optimal control viewpoint

The above Hamiltonian system and calculation of geodesics admits the optimal control interpretation. The interplay of Control theory and sub-Riemannian geometry has been well known since early 80s. One of the pioneering contributions was made by Brockett [8]. He considered a time optimal control problem leading to the sub-Riemannian geometry in $\mathbb{R}^3$, or to the Heisenberg group. His results then were generalized in several ways, see e.g., [21]. Several results, already known by this time due to the fundamental Gaveau's work [15], were rediscovered and the problem of finding normal and abnormal geodesics was formulated in terms of the optimal control. Pontryagin's maximum principle provides such optimal controls. Interesting features of such Hamiltonian systems are symmetries given by the first integrals although such systems generally are not (Frobenius) integrable because of singular geometric background, i.e., constraints on the velocities can not be re-written in terms of the configuration coordinates. A good reference to the control theory viewpoint is [6].

Let us consider the following time optimal control problem given by the system

$$\begin{array}{rcl} \dot x_1 &=& -ux_3 - vx_4, \\ \dot x_2 &=& -ux_4 + vx_3, \\ \dot x_3 &=& ux_1 - vx_2, \\ \dot x_4 &=& ux_2 + vx_1, \end{array}$$

with the cost functional

$$E = \frac{1}{2}\int_0^t \langle \mathbf{u}, \mathbf{u}\rangle ds,$$

where $\mathbf{u} = (u, v)$. The functional E represents the total kinetic energy. The system is encoded in the kernel of the contact 1-form

$$\omega = -x_2 dx_1 + x_1 dx_2 + x_4 dx_3 - x_3 dx_4.$$

The Hamiltonian for this system admits the form

$$H = -\frac{1}{2}(u^2 + v^2) + u(-x_3\xi_1 - x_4\xi_2 + x_1\xi_3 + x_2\xi_4) + v(-x_4\xi_1 + x_3\xi_2 - x_2\xi_3 + x_1\xi_4),$$

and the system for covectors becomes

$$\begin{aligned}
\dot{\xi}_1 &= -u\xi_3 - v\xi_4, \\
\dot{\xi}_2 &= -u\xi_4 + v\xi_3, \\
\dot{\xi}_3 &= u\xi_1 - v\xi_2, \\
\dot{\xi}_4 &= u\xi_2 + v\xi_1.
\end{aligned}$$

This system for the phase coordinates may be rewritten in the following form

$$\begin{aligned}
u &= -x_3\dot{x}_1 - x_4\dot{x}_2 + x_1\dot{x}_3 + x_2\dot{x}_4, \\
v &= -x_4\dot{x}_1 + x_3\dot{x}_2 - x_2\dot{x}_3 + x_1\dot{x}_4, \\
0 &= -x_2\dot{x}_1 + x_1\dot{x}_2 + x_4\dot{x}_3 - x_3\dot{x}_4, \\
0 &= x_1\dot{x}_1 + x_2\dot{x}_2 + x_3\dot{x}_3 + x_4\dot{x}_4,
\end{aligned}$$

which has a clear geometric meaning. Indeed, u and v are the coefficients of the velocity vector $uX + vY$, the third equation is just the horizontality condition and the fourth means that the trajectory belongs to a sphere.

From the Hamiltonian system one derives four first integrals

$$\begin{aligned}
J_1 &= x_1\xi_1 + x_2\xi_2 + x_3\xi_3 + x_4\xi_4, \\
J_2 &= -x_2\xi_1 + x_1\xi_2 + x_4\xi_3 - x_3\xi_4, \\
J_3 &= -x_3\xi_1 + x_4\xi_2 + x_1\xi_3 - x_2\xi_4, \\
J_4 &= -x_4\xi_1 - x_3\xi_2 + x_2\xi_3 + x_1\xi_4.
\end{aligned}$$

The Poisson structure is given by the Poisson brackets

$$[F, G] = \sum_{k=1}^{4} \frac{\partial F}{\partial x_k}\frac{\partial G}{\partial \xi_k} - \frac{\partial G}{\partial x_k}\frac{\partial F}{\partial \xi_k}.$$

The integrals J_1 and J_2 represent natural symmetries (following two natural geometric conditions: J_1 is the normal covector and J_2 gives the horizontality condition) and J_3, J_4 give hidden symmetries. All first integrals are involutive in pairs $[J_k, J_m] = 0$, $k, m = 1, \ldots, 4$, which implies Liouville integrabilty of the above Hamiltonain system. Observe, that the Hamiltonian system for the Heisenberg group is not Liouville integrable as well as the Hamiltonian system corresponding to sub-Riemannian geometry on $SO(n)$ for $n \geq 4$, see [7, 26]. Let us remark that the optimal control problem in sub-Riemannian geometry on $SO(n)$ can be viewed as the problem of optimal laser-induced population transfer in n-level quantum systems, see [7].

In order to find geodesics we can use the Pontryagin Maximum Principle [25] which states that any normal geodesic is a solution to the above Hamiltonian system on the cotangent bundle with the control $\mathbf{u}^*$ which maximizes the Hamiltonian H, i.e., satisfies the equation

$$\frac{\partial H}{\partial u} = \frac{\partial H}{\partial v} = 0.$$

This problem is equivalent to the geometric problem of minimizing the Carnot-Carathéodory distance (or, equivalently, sub-Riemannian energy) in the optimal control problem for our control-linear system. The optimal control admits the form

$$u^* = -x_3\xi_1 - x_4\xi_2 + x_1\xi_3 + x_2\xi_4, \quad v^* = -x_4\xi_1 + x_3\xi_2 - x_2\xi_3 + x_1\xi_4.$$

Substituting $\mathbf{u}^*$ in the Hamiltonian system we obtain the geodesic equation. The importance of integrability of the sub-Riemannian geodesic equation was argued by Brockett and Dai [9], who showed the explicit integrability in some special cases in terms of elliptic functions and discussed applications to controllability problems. But the question of integrabilty of Hamiltonain systems associated with nonholonomic distributions has a long history, see the survey [30] for the historical account.

As it was shown in [22], abnormal geodesics are not geometrically relevant for step 2 groups. Nevertheless, we give here independent treatment of abnormal geodesics from the Pontryagin Maximum Principle viewpoint. The Hamiltonian in this case becomes

$$H_0 = u(-x_3\xi_1 - x_4\xi_2 + x_1\xi_3 + x_2\xi_4) + v(-x_4\xi_1 + x_3\xi_2 - x_2\xi_3 + x_1\xi_4) = uJ_3 + vJ_4.$$

The Pontryagin Maximum Principle implies that H_0 vanishes along the extremal. We can assume that the velocity coordinates u and v do not vanish simultaneously. After differentiating J_3 and J_4 along the extremal we obtain

$$0 = \dot{J}_3 = [J_3, H_0] = -2vJ_2,$$

$$0 = \dot{J}_4 = [J_4, H_0] = 2uJ_2.$$

Let us suppose that u does not vanish on some time interval $s \in U$. Then, $J_2 = 0$ on this interval, and being the first integral, is vanishing everywhere. Then we obtain

$$0 = \dot{J}_2 = [J_2, H_0] = -2uJ_4,$$

and J_4 is identically 0 by the same reason. Therefore, $J_3 \equiv 0$. Solving the system $J_k = 0$, $k = 1, \dots, 4$, with respect to x_k we see that the discriminant of this system is 1. Fixing initial conditions for the Hamiltonian system we deduce that $\xi_k = 0$, $k = 1, \dots, 4$, and only stationary solution is valid.

For normal geodesics we have that along the extremal

$$H = \frac{1}{2}(-x_3\xi_1 - x_4\xi_2 + x_1\xi_3 + x_2\xi_4)^2 + \frac{1}{2}(-x_4\xi_1 + x_3\xi_2 - x_2\xi_3 + x_1\xi_4)^2.$$

6. Geodesics and modified action

6.1. **Cartesian coordinates.** Fix the initial point $x^{(0)} = (1, 0, 0, 0)$. It is convenient to introduce complex coordinates $z = x_1 + ix_2$, $w = x_3 + ix_4$, $\varphi = \xi_1 + i\xi_2$, and $\psi = \xi_3 + i\xi_4$. Hence, the Hamiltonian H from the above section admits the form $H = \frac{1}{2}|\bar{w}\varphi - z\bar{\psi}|^2$ (compare with (4.2)). The corresponding Hamiltonian system becomes

$$\begin{aligned}
\dot{z} &= w(\bar{w}\varphi - z\bar{\psi}), && z(0) = 1, \\
\dot{w} &= -z(w\bar{\varphi} - \bar{z}\psi), && w(0) = 0, \\
\dot{\bar{\varphi}} &= \bar{\psi}(w\bar{\varphi} - \bar{z}\psi), && \bar{\varphi}(0) = A - iB, \\
\dot{\bar{\psi}} &= -\bar{\varphi}(\bar{w}\varphi - z\bar{\psi}), && \bar{\psi}(0) = C - iD.
\end{aligned}$$

Here the constants B, C, and D have the following dynamical meaning: $\dot{w}(0) = C + iD$, and $B = -i\ddot{w}(0)/2\dot{w}(0)$ or if we write in real variables, $C = \dot{x}_3(0)$, $D = \dot{x}_4(0)$, $B = \frac{1}{2}(\dot{x}_3(0)\ddot{x}_4(0) - \dot{x}_4(0)\ddot{x}_3(0))/(\dot{x}_3^2(0) + \dot{x}_4{}_4^2(0))$. This complex Hamiltonian system has the first integrals

$$\begin{aligned}
z\psi - w\varphi &= C + iD, \\
z\bar{\varphi} + w\bar{\psi} &= A - iB,
\end{aligned}$$

and we have $|z|^2+|w|^2=1$ as a normalization. Therefore,

$$\begin{aligned}\varphi &= z(A+iB)-\bar{w}(C+iD),\\ \psi &= \bar{z}(C+iD)+w(A+iB).\end{aligned}$$

Let us introduce an auxiliary function $p=\bar{w}/z$. Then substituting φ and ψ in the Hamiltonian system we get the equation for p as

$$\dot{p}=(C+iD)p^2-2iBp+(C-iD),\quad p(0)=0.$$

The solution is

$$p(s)=\frac{(C-iD)\sin(s\sqrt{B^2+C^2+D^2})}{\sqrt{B^2+C^2+D^2}\cos(s\sqrt{B^2+C^2+D^2})+iB\sin(s\sqrt{B^2+C^2+D^2})}.$$

Taking into account that $\dot{z}\bar{z}=-w\dot{\bar{w}}$, we get the solution

(6.1)
$$z(s)=\left(\cos(s\sqrt{B^2+C^2+D^2})+i\frac{B}{\sqrt{B^2+C^2+D^2}}\sin(s\sqrt{B^2+C^2+D^2})\right)e^{-iBs},$$

and

$$w(s)=\frac{C+iD}{\sqrt{B^2+C^2+D^2}}\sin(s\sqrt{B^2+C^2+D^2})e^{iBs}. \tag{6.2}$$

Remark 2. Let us consider three limiting cases. If $B=0$, then we get the solutions with constant horizontal velocity coordinates

$$z(s)=\cos s,\qquad w(s)=(\dot{x}_3(0)+i\dot{x}_4(0))\sin s$$

which lies on the horizontal 2-sphere and a geodesic joining two given points on it is unique. If $C^2+D^2=0$, then the only solution $w(s)$ to the Hamiltonian system is $w(s)\equiv 0$. The horizontality condition in this case is read as $x_2\dot{x}_1=x_1\dot{x}_2$ and the solution is a straight line which contradicts the condition $|z|^2=1$. So $H=\frac{1}{2}(C^2+D^2)>0$.

Now we want to define geodesics joining two given points.

Theorem 2. *Let A be a point of the vertical line, i. e. $Q=(\cos\omega,\sin\omega,0,0)$, $\omega\in(-\pi,0)\cup(0,\pi)$, then there are countably many geometrically different geodesics γ_n connecting $P=(1,0,0,0)$ with Q. They have the following parametric equations*

$$\begin{aligned} z_n(s) &= \Big(\cos(s\frac{\pi n}{t})-i\frac{\omega}{\pi n}\sin(s\frac{\pi n}{t})\Big)e^{-\frac{is\omega}{t}},\\ w_n(s) &= (C+iD)\frac{t}{\pi n}\sin(s\frac{\pi n}{t})e^{\frac{is\omega}{t}}, \end{aligned} \tag{6.3}$$

$n\in\mathbb{Z}\setminus\{0\}$, $s\in[0,t]$, and the length of geodesics γ_n is given as $l_n=\frac{1}{\sqrt{2}}\sqrt{(\pi n)^2-\omega^2}$.

Proof. The geodesics are parametrized by the time interval $s\in[0,t]$. If the point $A=(z(t),w(t))=(z,w)$ belongs to the vertical line starting at $P=(1,0,0,0)$, then $|z|=1$ and $|w|=0$ provided that $-Bt=\omega$. It implies

$$\cos^2(t\sqrt{B^2+C^2+D^2})+\frac{B^2}{B^2+C^2+D^2}\sin^2(t\sqrt{B^2+C^2+D^2})=1,$$
$$\sin(t\sqrt{B^2+C^2+D^2})=0,\qquad -Bt=\omega.$$

These equations imply

$$t = \frac{\pi n}{\sqrt{B^2 + C^2 + D^2}} > 0, \quad Bt = \omega. \tag{6.4}$$

The latter relations give

$$B = B_n \equiv \frac{\omega \sqrt{C^2 + D^2}}{\sqrt{(\pi n)^2 - \omega^2}}.$$

Substituting (6.4) in the solutions to the hamiltonian system we come to the parametric representation given in the formulation of the theorem. The first relation of (6.4) yields

$$\sqrt{C^2 + D^2} = \frac{1}{t}\sqrt{(\pi n)^2 - \omega^2}.$$

The length of each geodesic is given as

$$l_n = t\sqrt{H} = \frac{t}{\sqrt{2}}\sqrt{C^2 + D^2} = \frac{1}{\sqrt{2}}\sqrt{(\pi n)^2 - \omega^2}.$$

This finishes the proof. □

Remark 3. In the formulation of the theorem the words 'geometrically different' mean that due to the rotation of the argument of $C + iD$ in $w(s)$, there exist uncountably many geodesics.

So far we have had a clear picture of trivial geodesics whose velocity has constant coordinates. They are essentially unique (up to periodicity). The situation with geodesics joining the point $(1, 0, 0, 0)$ with the points of the vertical line A has been described in the preceding theorem. Let us consider the general position of the right endpoint (z, w), $z = re^{i\xi_1}$, $w = \rho e^{i\xi_2}$ on $\mathbb{S}^3$.

Remark 4. First we consider three limiting cases. If $\rho = 1$, then $r = 0$, which implies $B = 0$, and the point lies on the horizontal 2-sphere. If $\rho = 0$, then $\sin(t\sqrt{B^2 + C^2 + D^2}) = 0$ and the point $z = \cos(Bt) - i\sin(Bt)$, $w = 0$ belongs to the vertical line. If $\arg z = 0$, then $z = \pm r$, $w = \pm\sqrt{1 - r^2}(\cos\xi_2 + i\sin\xi_2)$ is a point is on the horizontal 2-sphere.

In other situations we have the following theorem.

Theorem 3. *Given an arbitrary point $(z, w) \in \mathbb{S}^3$ which neither belongs to the vertical line A nor to the horizontal sphere $\mathbb{S}^2$, there is a finite number of geometrically different geodesics joining the initial point (north pole) $P = (1, 0, 0, 0) \in \mathbb{S}^3$ with $Q = (z, w)$.*

Proof. Let us denote

$$z = re^{i\xi_1}, \quad w = \rho e^{i\xi_2}, \quad C + iD = \sqrt{C^2 + D^2}e^{i\theta}.$$

Then from (6.1) and (6.2) we have that

$$\rho^2 = \frac{C^2 + D^2}{B^2 + C^2 + D^2}\sin^2(t\sqrt{B^2 + C^2 + D^2}), \quad \text{and} \quad \xi_2 = Bt + \theta, \tag{6.5}$$

where t is the right end of the time interval $s \in [0, t]$ at which the endpoint Q is reached. We suppose for the moment that the angles $s\sqrt{B^2 + C^2 + D^2}$ and tB are from the first quadrant. Other cases are treated similarly. Then we have

$$z = \Big(\sqrt{1 - \frac{B^2 + C^2 + D^2}{C^2 + D^2}\rho^2} + i\frac{B\rho}{\sqrt{C^2 + D^2}}\Big)e^{i(\theta - \xi_2)},$$

and

$$\xi_1 = \theta - \xi_2 + \arctan\frac{B\rho}{\sqrt{C^2 + D^2 - (B^2 + C^2 + D^2)\rho^2}}.$$

The first expression in (6.5) leads to the value of the length parameter t as

$$t = \frac{1}{\sqrt{B^2 + C^2 + D^2}}\arcsin\left(\rho\sqrt{1 + \frac{B^2}{C^2 + D^2}}\right),$$

and the second to

$$\xi_2 = \theta + \frac{B}{\sqrt{B^2 + C^2 + D^2}}\arcsin\left(\rho\sqrt{1 + \frac{B^2}{C^2 + D^2}}\right).$$

Substituting θ in the latter equation we come to an equation which depends only on

$$k = \frac{B}{\sqrt{C^2 + D^2}},$$

which we rewrite as

$$\sin\left(\sqrt{1 + \frac{1}{k^2}}\left[\arctan\Big(\frac{k\rho}{1 - (1 + k^2)\rho^2}\Big) - \xi_1\right]\right) = \rho\sqrt{1 + k^2}, \tag{6.6}$$

as an equation for the parameter k, which is the curvature of the geodesic at the initial moment. Observe that $\theta - \xi_2 = \alpha - \arctan\left(\frac{k\rho}{1-(1+k^2)\rho^2}\right)$ is a bounded function and $\lim_{B\to 0}\theta(B) \neq 0$. Indeed, if the latter limit were vanishing, then the value of given ξ_2 would be zero and the solution of the problem would be only $k = B = 0$ which is the trivial case excluded from the theorem. So the left-hand side of equation (6.6) is a function of k which is bounded by 1 in absolute value and fast oscillating about the point $k = 0$. Observe, that $\xi_1 = 0$ corresponds to the horizontal sphere which also was excluded from the theorem. The right-hand side of (6.6) is an even function increasing for $k > 0$, see Figure (2). Therefore, there exists a countable number of non-vanishing different solutions $\{k_n\}$ of the equation (6.6) within the interval $|k| \leq \sqrt{\frac{1}{\rho^2} - 1} = \frac{|z|}{|w|}$ with a limit point at the origin.

However, in order do define all parameters B, C, and D we need to solve the equations (6.5), (6.6), and not all k_n satisfy all three equations. Let us consider positive k_n. We calculate the argument of z as

$$\xi_1 = -Bt + \arctan\left[\frac{B}{\sqrt{B^2 + C^2 + D^2}}\tan\Big(t\sqrt{B^2 + C^2 + D^2}\Big)\right]$$

$$= -Bt + \arctan\left[\frac{k_n\rho}{\sqrt{1 - (1 + k_n^2)\rho^2}}\right]$$

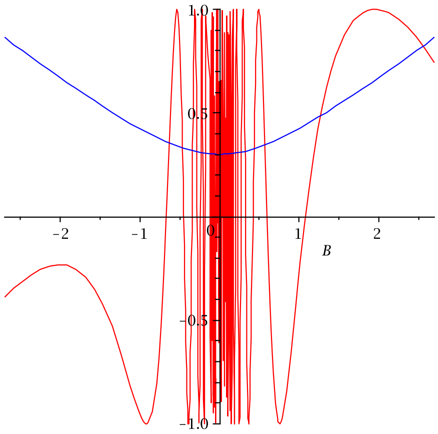

FIGURE 2. Solutions to the equation (6.6)

$$< -k_n t\sqrt{C^2+D^2} + \frac{k_n\rho}{\sqrt{1-(1+k_n^2)\rho^2}}.$$

On the other hand, we have

$$\sqrt{C^2+D^2} = \frac{\arcsin(\rho\sqrt{1+k_n^2})}{t\sqrt{1+k_n^2}} > \frac{\rho}{t}.$$

Observe that due the remark before this theorem, $\xi_1 > 0$ and $0<\rho<1$. Therefore, we deduce the inequality

$$\xi_1 < k_n\rho\frac{1-\sqrt{1-\rho^2(1+k_n^2)}}{\sqrt{1-\rho^2(1+k_n^2)}},$$

or

$$k_n\rho > \xi_1\frac{\sqrt{1-\rho^2(1+k_n^2)}}{1-\sqrt{1-\rho^2(1+k_n^2)}}. \tag{6.7}$$

The right-hand side of the inequality (6.7) decreases with respect to $k_n>0$.

Set $\varepsilon = \frac{1+\rho^2}{2}$. If $\varepsilon < \rho^2(1+k_n^2) < 1$, then immediately we have the inequality $k_n^2 > \frac{1}{2}(\frac{1}{\rho^2}-1) > 0$. If $0<\rho^2(1+k_n^2)\leq\varepsilon$, then the inequality (6.7) implies that

$$k_n > \xi_1\frac{\sqrt{1-\varepsilon}}{\rho(1-\sqrt{1-\varepsilon})} = \xi_1\frac{\sqrt{1-\rho^2}}{\rho(\sqrt{2}-\sqrt{1-\rho^2})} > 0.$$

Finally, we obtain

$$k_n > \min\left\{\xi_1\frac{\sqrt{1-\rho^2}}{\rho(\sqrt{2}-\sqrt{1-\rho^2})}, \sqrt{\frac{1}{2}(\frac{1}{\rho^2}-1)}\right\} \equiv b(\xi_1,\rho) > 0.$$

This proves that all positive solutions to the equation (6.6) must belong to the interval $(b(\xi_1,\rho), \sqrt{\frac{1}{\rho^2}-1})$, hence there are only finite number of such k_n. The same arguments are applied for negative values of k_n. □

Remark 5. If ρ is approaching 0, the point Q is approaching the vertical line and the value of k_n becomes

$$k_n = \frac{\pm\xi_1}{\sqrt{(\pi n)^2 - \xi_1^2}},$$

and the solution is reduced to the case considered in Theorem 2 with $\omega = \xi_1$, i.e., the number of geodesics is increasing infintely.

6.2. **Hyperspherical coordinates.** Let us use now the hyperspherical coordinates

$$\begin{aligned} x_1 + ix_2 &= e^{i\xi_1}\cos\eta, \\ x_3 + ix_4 &= e^{i\xi_2}\sin\eta, \qquad \eta \in (0, \pi/2), \quad \xi_1, \xi_2 \in [0, 2\pi), \end{aligned} \tag{6.8}$$

to write the Hamiltonian system.

The horizontal coordinates are written as

$$\begin{aligned} a &= \dot\eta\cos(\xi_1 - \xi_2) + (\dot\xi_1 + \dot\xi_2)\sin(\xi_1 - \xi_2)\frac{\sin 2\eta}{2}, \\ b &= -\dot\eta\sin(\xi_1 - \xi_2) + (\dot\xi_1 + \dot\xi_2)\cos(\xi_1 - \xi_2)\frac{\sin 2\eta}{2}, \\ c &= \dot\xi_1\cos^2\eta - \dot\xi_2\sin^2\eta. \end{aligned}$$

The horizontality condition in hyperspherical coordinates becomes

$$\dot\xi_1\cos^2\eta - \dot\xi_2\sin^2\eta = 0.$$

The horizontal 2-sphere is obtained from the parametrization (6.8), if we set $\xi_1 = 0$, $\xi_2 = \psi$, $\eta = s$. We get

$$a^2 + b^2 = 1 = \dot\eta^2 \quad \Longrightarrow \quad a = \cos\psi, \quad b = \sin\psi.$$

The vertical line is obtained from the parametrization (6.8) setting $\eta = 0$, $\xi_1 = s$.

Writing the vector fields N, Z, X, Y in the hyperspherical coordinates we get

$$N = -2\,\mathrm{cotan}\,2\eta\partial_\eta, \quad Z = \partial_{\xi_1} - \partial_{\xi_2},$$

$$X = \sin(\xi_1 - \xi_2)\tan\eta\partial_{\xi_1} + \sin(\xi_1 - \xi_2)\,\mathrm{cotan}\,\eta\partial_{\xi_2} + 2\cos(\xi_1 - \xi_2)\partial_\eta,$$

$$Y = \cos(\xi_1 - \xi_2)\tan\eta\partial_{\xi_1} + \cos(\xi_1 - \xi_2)\,\mathrm{cotan}\,\eta\partial_{\xi_2} - 2\sin(\xi_1 - \xi_2)\partial_\eta.$$

In this parametrization the similarity with the Heisenberg group can be shown. The commutator of two horizontal vector fields X, Y gives the constant vector field Z which is orthogonal to the horizontal vector fields at each point of the manifold. In hyperspherical coordinates it is easy to see that the form $\omega = \cos^2\eta d\xi_1 - \sin^2\eta d\xi_2$, that defines the horizontal distribution is contact because

$$\omega \wedge d\omega = \sin(2\eta)\,d\eta \wedge d\xi_1 \wedge d\xi_2 = 2dV,$$

where dV is the volume form. The sub-Laplacian is defined as

$$\frac{1}{2}(X^2 + Y^2) = \frac{1}{2}(\tan^2\eta\partial_{\xi_1}^2 + \mathrm{cotan}^2\,\eta\partial_{\xi_2}^2 + 4\partial_\eta^2 + 2\partial_{\xi_1}\partial_{\xi_2}).$$

The Hamiltonian becomes

$$H(\xi_1, \xi_2, \eta, \psi_1, \psi_2, \theta) = \frac{1}{2}(\tan^2\eta\psi_1^2 + \mathrm{cotan}^2\,\eta\psi_2^2 + 4\theta^2 + 2\psi_1\psi_2).$$

It gives the Hamiltonian system

$$
\begin{array}{rcl}
\dot{\xi}_1 & = & \frac{\partial H}{\partial \psi_1} = \psi_1 \tan^2 \eta + \psi_2 \\
\dot{\xi}_2 & = & \frac{\partial H}{\partial \psi_2} = \psi_2 \operatorname{cotan}^2 \eta + \psi_1 \\
\dot{\eta} & = & \frac{\partial H}{\partial \theta} = 4\theta \\
\dot{\psi}_1 & = & -\frac{\partial H}{\partial \xi_1} = 0 \\
\dot{\psi}_2 & = & -\frac{\partial H}{\partial \xi_2} = 0 \\
\dot{\theta} & = & -\frac{\partial H}{\partial \eta} = -\psi_1^2 \frac{\tan \eta}{\cos^2 \eta} + \psi_2^2 \frac{\operatorname{cotan} \eta}{\sin^2 \eta}
\end{array}
\tag{6.9}
$$

6.3. **Geodesics in hyperspherical coordinates.** Let us find geodesics

$$\gamma(s) = (\xi_1(s), \xi_2(s), \eta(s))$$

(solutions to system (6.9)) joining the points $P = \gamma(0) = (\xi_1^0, \xi_2^0, \eta^0)$ and $Q = \gamma(t) = (\xi_1, \xi_2, \eta)$.

Observe that the system (6.9) is coupled and the system

$$
\begin{array}{rcl}
\dot{\eta} & = & \frac{\partial H}{\partial \theta} = 4\theta \\
\dot{\theta} & = & -\frac{\partial H}{\partial \eta} = -\psi_1^2 \frac{\tan \eta}{\cos^2 \eta} + \psi_2^2 \frac{\operatorname{cotan} \eta}{\sin^2 \eta}
\end{array}
\tag{6.10}
$$

with the boundary conditions $\eta(0) = \eta_0$, $\eta(t) = \eta$ is independent.

Multiplying the third and the last equations in the Hamiltonain system crosswise we obtain

$$4\theta\dot{\theta} = \left(-\psi_1^2 \frac{\tan \eta}{\cos^2 \eta} + \psi_2^2 \frac{\operatorname{cotan} \eta}{\sin^2 \eta}\right) \dot{\eta},$$

or

$$\frac{d}{dt}(2\theta^2) = \frac{d}{dt}(-\frac{1}{2}\psi_1^2 \tan^2 \eta - \frac{1}{2}\psi_2^2 \cot^2 \eta).$$

Therefore,

$$4\theta^2(s) = 4\theta_0^2 - \psi_1^2 \tan^2 \eta(s) - \psi_2^2 \cot^2 \eta(s) + \psi_1^2 \tan^2 \eta_0 + \psi_2^2 \cot^2 \eta_0. \tag{6.11}$$

Let us substitute the expression for $\theta(s)$ in the third equation of the Hamiltonian system. We obtain

$$\dot{\eta} = 4\theta = \pm 2\sqrt{4\theta_0^2 + \psi_1^2 \tan^2 \eta_0 + \psi_2^2 \cot^2 \eta_0 - \psi_1^2 \tan^2 \eta(s) - \psi_2^2 \cot^2 \eta(s)}.$$

Observe that the expression under the square root is non-negative for all $s \in [0, t]$. Let us consider the case of increasing η and $(+)$ in front of the square root. Negative case will be treated later. Change variables $u = \sin^2 \eta \in [0, 1]$. Then we arrive at

$$\frac{1}{4}\dot{u} = \sqrt{u(1-u)(4\theta_0^2 + (\psi_1 \tan \eta_0 + \psi_2 \cot \eta_0)^2) - (\psi_1 u + \psi_2(1-u))^2}.$$

The square polynomial under the root is reduced to

$$\left(\frac{\psi_1^2}{\cos^2 \eta_0} + \frac{\psi_2^2}{\sin^2 \eta_0} + 4\theta_0^2\right) \left\{ \frac{\left(\frac{\psi_1^2}{\cos^2 \eta_0} + \frac{\psi_2^2}{\sin^2 \eta_0} + 4\theta_0^2 + \psi_2^2 - \psi_1^2\right)^2}{4\left(\frac{\psi_1^2}{\cos^2 \eta_0} + \frac{\psi_2^2}{\sin^2 \eta_0} + 4\theta_0^2\right)^2} - \frac{\psi_2^2}{\frac{\psi_1^2}{\cos^2 \eta_0} + \frac{\psi_2^2}{\sin^2 \eta_0} + 4\theta_0^2} - \right.$$

$$-\left[u-\frac{\frac{\psi_1^2}{\cos^2\eta_0}+\frac{\psi_2^2}{\sin^2\eta_0}+4\theta_0^2+\psi_2^2-\psi_1^2}{2(\frac{\psi_1^2}{\cos^2\eta_0}+\frac{\psi_2^2}{\sin^2\eta_0}+4\theta_0^2)}\right]^2\Bigg\}.$$

The polynomial is non-negative for all $u\in[0,1]$, as it was mentioned before. Therefore,

$$\frac{\left(\frac{\psi_1^2}{\cos^2\eta_0}+\frac{\psi_2^2}{\sin^2\eta_0}+4\theta_0^2+\psi_2^2-\psi_1^2\right)^2}{4\left(\frac{\psi_1^2}{\cos^2\eta_0}+\frac{\psi_2^2}{\sin^2\eta_0}+4\theta_0^2\right)^2}-\frac{\psi_2^2}{\frac{\psi_1^2}{\cos^2\eta_0}+\frac{\psi_2^2}{\sin^2\eta_0}+4\theta_0^2}$$

is non-negative too. Integrating gives us

$$\arcsin\frac{\sin^2\eta(s)-\frac{1}{2}\left(1+\frac{\psi_2^2-\psi_1^2}{A}\right)}{\sqrt{\frac{1}{4}\left(1+\frac{\psi_2^2-\psi_1^2}{A}\right)^2-\frac{\psi_2^2}{A}}}=4s\sqrt{A}+const, \tag{6.12}$$

where

$$A=\frac{\psi_1^2}{\cos^2\eta_0}+\frac{\psi_2^2}{\sin^2\eta_0}+4\theta_0^2>0,$$

Now we consider the sign (-) in front of the square root. Finally, our solution is written as

$$\frac{\sin^2\eta(s)-\frac{1}{2}\left(1+\frac{\psi_2^2-\psi_1^2}{A}\right)}{\sqrt{\frac{1}{4}\left(1+\frac{\psi_2^2-\psi_1^2}{A}\right)^2-\frac{\psi_2^2}{A}}}=\sin(\pm 4s\sqrt{A}+const), \tag{6.13}$$

where

$$const=\arcsin\frac{\sin^2\eta_0-\frac{1}{2}\left(1+\frac{\psi_2^2-\psi_1^2}{A}\right)}{\sqrt{\frac{1}{4}\left(1+\frac{\psi_2^2-\psi_1^2}{A}\right)^2-\frac{\psi_2^2}{A}}}+2\pi n.$$

Substituting $s=t$ gives us implicit functions $\theta_0^{(n)}=\theta_0^{(n)}(\eta_0,\eta,\psi_1,\psi_2,t)$. Observe that $\theta_0^{(n)}$ satisfies the stretching condition $\theta_0^{(n)}(\eta_0,\eta,\psi_1,\psi_2,t)=\lambda\theta_0^{(n)}(\eta_0,\eta,\lambda\psi_1,\lambda\psi_2,\frac{t}{\lambda})$. For us it is more convenient to keep A instead of θ_0, so we define $A=A^{(n)}(\psi_1\psi_2,\eta_0,\eta,t)$ is a solution to the equation

$$\frac{\sin^2\eta-\frac{1}{2}\left(1+\frac{\psi_2^2-\psi_1^2}{A}\right)}{\sqrt{\frac{1}{4}\left(1+\frac{\psi_2^2-\psi_1^2}{A}\right)^2-\frac{\psi_2^2}{A}}}=\sin\left(\pm 4t\sqrt{A}+\arcsin\frac{\sin^2\eta_0-\frac{1}{2}\left(1+\frac{\psi_2^2-\psi_1^2}{A}\right)}{\sqrt{\frac{1}{4}\left(1+\frac{\psi_2^2-\psi_1^2}{A}\right)^2-\frac{\psi_2^2}{A}}}\right). \tag{6.14}$$

Then,

$$\tan^2\eta(s)=-1+\frac{1}{1-\frac{1}{2}\left(1+\frac{\psi_2^2-\psi_1^2}{A}\right)\mp D_0\sin(4s\sqrt{A}+D_1)},$$

where D_0 and D_1 are the constants

$$D_0=\sqrt{\frac{1}{4}\left(1+\frac{\psi_2^2-\psi_1^2}{A}\right)^2-\frac{\psi_2^2}{A}},$$

$$D_1 = \arcsin \frac{\sin^2 \eta_0 - \frac{1}{2}\left(1 + \frac{\psi_2^2 - \psi_1^2}{A}\right)}{\sqrt{\frac{1}{4}\left(1 + \frac{\psi_2^2 - \psi_1^2}{A}\right)^2 - \frac{\psi_2^2}{A}}},$$

and

$$\cot^2 \eta(s) = -1 + \frac{1}{\frac{1}{2}\left(1 + \frac{\psi_2^2 - \psi_1^2}{A}\right) \pm D_0 \sin(4s\sqrt{A} + D_1)}.$$

Integrating two first equations of the Hamiltonian system (6.9) gives

$$\xi_1(s) - \xi_1^0 = s\psi_2 + \psi_1 \int_0^s \tan^2 \eta(s) ds$$

$$= s(\psi_2 - \psi_1) + \int_0^s \frac{\psi_1 ds}{1 - \frac{1}{2}\left(1 + \frac{\psi_2^2 - \psi_1^2}{A}\right) \mp D_0 \sin(4s\sqrt{A} + D_1)}$$

$$= s(\psi_2 - \psi_1) + \frac{1}{2} \arctan \frac{\sqrt{A}}{\psi_1} \left\{ \left[1 - \frac{1}{2}\left(1 + \frac{\psi_2^2 - \psi_1^2}{A}\right)\right] \tan\left(2s\sqrt{A} + \frac{D_1}{2}\right) \mp D_0 \right\}$$

$$-\frac{1}{2} \arctan \frac{\sqrt{A}}{\psi_1} \left\{ \left[1 - \frac{1}{2}\left(1 + \frac{\psi_2^2 - \psi_1^2}{A}\right)\right] \tan \frac{D_1}{2} \mp D_0 \right\}.$$

Analogously,

$$\xi_2(s) - \xi_2^0 = s\psi_1 + \psi_2 \int_0^s \cot^2 \eta(s) ds$$

$$= s(\psi_1 - \psi_2) + \frac{1}{2} \arctan \frac{\sqrt{A}}{\psi_2} \left\{ \frac{1}{2}\left(1 + \frac{\psi_2^2 - \psi_1^2}{A}\right) \tan\left(2s\sqrt{A} + \frac{D_1}{2}\right) \pm D_0 \right\}$$

$$-\frac{1}{2} \arctan \frac{\sqrt{A}}{\psi_2} \left\{ \frac{1}{2}\left(1 + \frac{\psi_2^2 - \psi_1^2}{A}\right) \tan \frac{D_1}{2} \pm D_0 \right\}.$$

In order to solve the boundary value problem we express implicitly ψ_1 and ψ_2 as functions $\psi_1 = \psi_1^{(n)}(\xi_1^0, \xi_2^0, \eta_0, \xi_1, \xi_2, \eta)$ and $\psi_2 = \psi_2^{(n)}(\xi_1^0, \xi_2^0, \eta_0, \xi_1, \xi_2, \eta)$ by the equations

$$\xi_1 - \xi_1^0 = t(\psi_2 - \psi_1) + \frac{1}{2} \arctan \frac{\sqrt{A}}{\psi_1} \left\{ \left[1 - \frac{1}{2}\left(1 + \frac{\psi_2^2 - \psi_1^2}{A}\right)\right] \tan\left(2t\sqrt{A} + \frac{D_1}{2}\right) \mp D_0 \right\}$$

$$-\frac{1}{2} \arctan \frac{\sqrt{A}}{\psi_1} \left\{ \left[1 - \frac{1}{2}\left(1 + \frac{\psi_2^2 - \psi_1^2}{A}\right)\right] \tan \frac{D_1}{2} \mp D_0 \right\},$$

and

$$\xi_2 - \xi_2^0 = t(\psi_1 - \psi_2) + \frac{1}{2} \arctan \frac{\sqrt{A}}{\psi_2} \left\{ \frac{1}{2}\left(1 + \frac{\psi_2^2 - \psi_1^2}{A}\right) \tan\left(2t\sqrt{A} + \frac{D_1}{2}\right) \pm D_0 \right\}$$

$$-\frac{1}{2} \arctan \frac{\sqrt{A}}{\psi_2} \left\{ \frac{1}{2}\left(1 + \frac{\psi_2^2 - \psi_1^2}{A}\right) \tan \frac{D_1}{2} \pm D_0 \right\},$$

where A is a solution $A = A^{(n)}(\psi_1\psi_2, \eta_0, \eta, t)$ to the equation (6.14).

In order to simplify the study of geodesics joining P and Q we consider a special case of P when $\xi_1^0 = \xi_2^0$ and $\eta_0 = \pi/4$. Observe that the chosen parametrization does not give us a chart about the north pole $(1,0,0,0)$. So we can shift the considerations by a left-invariant group action to any point, e.g., $(1/\sqrt{2},0,1/\sqrt{2},0)$.

6.4. **Modified action.** We see that ψ_1 and ψ_2 are the first integrals of the system. We solve this Hamiltonian system for the following mixed boundary conditions:

$$\begin{aligned}&\eta(0)=\eta_0,\ \eta(t)=\eta,\ \xi_1(t)=\xi_1,\ \xi_2(t)=\xi_2,\\&\psi_1(0)=\psi_1,\ \psi_2(0)=\psi_2,\ \theta(0)=\tfrac{\dot\eta(0)}{4}=\theta_0.\end{aligned}\tag{6.15}$$

In the classical case the action is defined on arbitrary smooth curves which join two given points. In our case the non-classical (or modified) action is defined on solutions to the above Hamiltonian system where only the coordinate η is given in two ending points of the interval $[0,t]$. The coordinates ξ_1 and ξ_2 are given only at the right-hand endpoint t. We are looking for the modified action in the form

$$\begin{aligned}g(\xi_1,\xi_2,\eta_0,\eta,\psi_1,\psi_2)&=\psi_1(0)\xi_1(0)+\psi_2(0)\xi_2(0)+\int_0^t(\psi_1\dot\xi_1(s)+\psi_2\dot\xi_2(s)+\theta\dot\eta(s)-H)ds\\&=\psi_1\xi_1+\psi_2\xi_2+\int_0^t(\theta\dot\eta(s)-H)ds.\end{aligned}$$

The function H does not depend on s on the solutions to the Hamiltonian system (independently on boundary conditions), there fore we have

$$g=\psi_1\xi_1+\psi_2\xi_2-\frac{t}{2}((\tan\eta_0\psi_1+\operatorname{cotan}\eta_0\psi_2)^2+4\theta_0^2)+\int_0^t 4\theta^2(s)ds.$$

In order to calculate our modified action we have to

- calculate the integral;
- represent θ_0 in terms of η.

Now let us calculate the integral in the modified action. The action g admits the form

$$\begin{aligned}g=\psi_1\xi_1+\psi_2\xi_2&+\frac{t}{2}((\psi_1\tan\eta_0+\psi_2\cot\eta_0)^2+4\theta_0^2)-2t\psi_1\psi_2-\int_0^t(\psi_1^2\tan^2\eta(s)+\psi_2^2\cot^2\eta(s))ds\\&=\psi_1\xi_1+\psi_2\xi_2+\frac{t}{2}((\psi_1\tan\eta_0+\psi_2\cot\eta_0)^2+4\theta_0^2)+t(\psi_1-\psi_2)^2\\&-\int_0^t\left(\frac{\psi_1^2}{1-\frac12\left(1+\frac{\psi_2^2-\psi_1^2}{A}\right)\mp D_0\sin(4s\sqrt{A}+D_1)}+\frac{\psi_2^2}{\frac12\left(1+\frac{\psi_2^2-\psi_1^2}{A}\right)\pm D_0\sin(4s\sqrt{A}+D_1)}\right)ds.\end{aligned}$$

We observe that

$$\frac14\left(1+\frac{\psi_2^2-\psi_1^2}{A}\right)^2-D_0^2=\frac{\psi_2^2}{A}>0,$$

and

$$\left(1-\frac12\left(1+\frac{\psi_2^2-\psi_1^2}{A}\right)\right)^2-D_0^2=\frac{\psi_1^2}{A}>0.$$

Thus, integrating gives us

$$g=\psi_1\xi_1+\psi_2\xi_2+\frac{t}{2}A+\frac{t}{2}(\psi_1-\psi_2)^2$$

$$-\frac{1}{2}\psi_1 \arctan\frac{\sqrt{A}}{\psi_1}\left\{\left[1-\frac{1}{2}\left(1+\frac{\psi_2^2-\psi_1^2}{A}\right)\right]\tan\left(2t\sqrt{A}+\frac{D_1}{2}\right)\mp D_0\right\}$$
$$+\frac{1}{2}\psi_1 \arctan\frac{\sqrt{A}}{\psi_1}\left\{\left[1-\frac{1}{2}\left(1+\frac{\psi_2^2-\psi_1^2}{A}\right)\right]\tan\frac{D_1}{2}\mp D_0\right\}$$
$$-\frac{1}{2}\psi_2 \arctan\frac{\sqrt{A}}{\psi_2}\left\{\frac{1}{2}\left(1+\frac{\psi_2^2-\psi_1^2}{A}\right)\tan\left(2t\sqrt{A}+\frac{D_1}{2}\right)\pm D_0\right\}$$
$$+\frac{1}{2}\psi_2 \arctan\frac{\sqrt{A}}{\psi_2}\left\{\frac{1}{2}\left(1+\frac{\psi_2^2-\psi_1^2}{A}\right)\tan\frac{D_1}{2}\pm D_0\right\}.$$

7. Generalized Hamilton-Jacobi equation

Theorem 4. *Let ξ_1, ξ_2, and $\eta=\eta$ be fixed, and let $\xi_1(s)=\xi_1(s;\psi_1,\psi_2,\xi_1,\xi_2,\eta,\eta_0,t)$ $\xi_2(s)=\xi_2(s;\psi_1,\psi_2,\xi_1,\xi_2,\eta,\eta_0,t)$, and $\eta(s)=\eta(s;\psi_1,\psi_2,\eta,\eta_0,t)$ be solutions to the Hamiltonian system* (6.9) *with the mixed boundary conditions* (6.15) *and with the Hamiltonian $H(\xi_1(s),\xi_2(s),\eta(s),\psi_1,\psi_2,\theta(s))$. The modified action*

$$g=\psi_1\xi_1+\psi_2\xi_2+\int_0^t(\theta\dot\eta(s)-H)ds$$

satisfies the Hamilton-Jacobi equation

$$\frac{\partial g}{\partial t}+H(\xi_1,\xi_2,\eta,\nabla g)=0,$$

where the gradient ∇ is taken with respect to the coordinates of the endpoint (ξ_1,ξ_2,η).

Proof. Let us calculate the derivatives of g with respect to ξ_1, ξ_2, and η explicitly.

$$\frac{\partial g}{\partial\xi_1}=\psi_1+\int_0^t\left(\frac{\partial\theta(s)}{\partial\xi_1}\dot\eta+\theta(s)\frac{d}{ds}\frac{\partial\eta(s)}{\partial\xi_1}-\frac{\partial H}{\partial\eta(s)}\frac{\partial\eta(s)}{\partial\xi_1}-\frac{\partial H}{\partial\theta(s)}\frac{\partial\theta(s)}{\partial\xi_1}\right)ds$$
$$=\psi_1+\int_0^t\left(\frac{d}{ds}\Big[\theta(s)\frac{\partial\eta(s)}{\partial\xi_1}\Big]\right)ds=\psi_1.$$

We used that $\frac{\partial H}{\partial\theta(s)}=\dot\eta$, and $\eta(s)$ does not depend on ξ_1. Here $\dot\eta$ means the derivative with respect to s. Analogously,

$$\frac{\partial g}{\partial\xi_2}=\psi_2.$$

We continue with

$$\frac{\partial g}{\partial\eta}=\int_0^t\left(\frac{\partial\theta(s)}{\partial\eta}\dot\eta+\theta(s)\frac{d}{ds}\frac{\partial\eta(s)}{\partial\eta}-\frac{\partial H}{\partial\eta(s)}\frac{\partial\eta(s)}{\partial\eta}-\frac{\partial H}{\partial\theta(s)}\frac{\partial\theta(s)}{\partial\eta}\right)ds$$
$$=\int_0^t\left(\frac{d}{ds}\Big[\theta(s)\frac{\partial\eta(s)}{\partial\eta}\Big]\right)ds=\theta(s)\frac{\partial\eta(s)}{\partial\eta}\Bigg|_{s=0}^{s=t}.$$

The last derivative is taken with respect to t as

$$\frac{\partial g}{\partial t}=(\theta(s)\dot\eta-H)_{s=t}+\int_0^t\left(\frac{\partial\theta(s)}{\partial t}\dot\eta+\theta(s)\frac{\partial}{\partial t}\dot\eta(s)-\frac{\partial H}{\partial\eta(s)}\frac{\partial\eta(s)}{\partial t}-\frac{\partial H}{\partial\theta(s)}\frac{\partial\theta(s)}{\partial t}\right)ds$$

$$= (\theta(s)\dot{\eta} - H)_{s=t} + \theta(s)\frac{\partial \eta(s)}{\partial t}\Big|_{s=0}^{s=t}.$$

We need to calculate

$$\frac{\partial \eta(s)}{\partial \eta}\Big|_{s=0}^{s=t} \quad \text{and} \quad \frac{\partial \eta(s)}{\partial t}\Big|_{s=0}^{s=t}.$$

Let us rewrite the equation (6.12) in the form

$$P(\eta(s), A) = 4s\sqrt{A},$$

where $A = A(\psi_1, \psi_2, \eta_0, \eta, t)$ satisfies the equation

$$P(\eta(t), A) = 4t\sqrt{A}.$$

Then

$$\frac{\partial P(\eta(s), A)}{\partial \eta} = \frac{\partial P(\eta(s), A)}{\partial \eta(s)}\frac{\partial \eta(s)}{\partial \eta} + \frac{\partial P(\eta(s), A)}{\partial A}\frac{\partial A}{\partial \eta} = \frac{2s}{\sqrt{A}}\frac{\partial A}{\partial \eta},$$

$$\frac{\partial P(\eta, A)}{\partial \eta} = \frac{\partial P(\eta, A)}{\partial \eta} + \frac{\partial P(\eta, A)}{\partial A}\frac{\partial A}{\partial \eta} = \frac{2t}{\sqrt{A}}\frac{\partial A}{\partial \eta}.$$

Substituting $s = t$ in the first equation and subtracting the second one we obtain

$$\frac{\partial \eta(s)}{\partial \eta}\Big|_{s=t} = 1.$$

Substituting $s = 0$ in the first equation we get

$$\frac{\partial \eta(s)}{\partial \eta}\Big|_{s=0} = -\frac{\partial P(\eta(s), A)}{\partial A}\frac{\partial A}{\partial \eta}\left(\frac{\partial P(\eta(s), A)}{\partial \eta(s)}\right)^{-1}\Big|_{s=0} = 0.$$

Analogously, calculating the derivative

$$\frac{\partial P(\eta(s), A)}{\partial t},$$

and substituting $s = 0$ we get $\frac{\partial \eta(s)}{\partial t}\Big|_{s=0} = 0$. Finally, we have

$$0 = \frac{\eta(t; \psi_1, \psi_2, \eta_0, \eta, t)}{\partial t} = \left(\dot{\eta} + \frac{\eta(s; \psi_1, \psi_2, \eta_0, \eta, t)}{\partial t}\right)_{s=t},$$

which implies that $\frac{\partial g}{\partial t} = -H$. Now we are able to calculate $H(\xi_1, \xi_2, \eta, \nabla g)$ as

$$H(\xi_1, \xi_2, \eta, \nabla g) = \frac{1}{2}\Big[\left(\frac{\partial g}{\partial \xi_1}\right)^2 \tan^2\eta + \left(\frac{\partial g}{\partial \xi_2}\right)^2 \cot^2\eta + 2\frac{\partial g}{\partial \xi_1}\frac{\partial g}{\partial \xi_2} + 4\left(\frac{\partial g}{\partial \eta}\right)^2\Big]$$

$$= \frac{1}{2}(\psi_1^2 \tan^2\eta + \psi_2^2 \cot^2\eta + 2\psi_1\psi_2 + 4\theta^2)_{s=t} = H = -\frac{\partial g}{\partial t}.$$

This finishes the proof. □

We observe that the modified action satisfies the stretching property

$$g(\xi_1,\xi_2,\eta_0,\eta,\psi_1,\psi_2,t)=\lambda g(\xi_1,\xi_2,\eta_0,\eta,\frac{\psi_1}{\lambda},\frac{\psi_2}{\lambda},\lambda t).$$

Let us construct the function

$$f(\xi_1,\xi_2,\eta_0,\eta,\psi_1,\psi_2)=g(\xi_1,\xi_2,\eta_0,\eta,\psi_1,\psi_2,1).$$

Theorem 5. *f is a solution to the generalized Hamilton-Jacobi equation*

$$\psi_1\frac{\partial f}{\partial\psi_1}+\psi_2\frac{\partial f}{\partial\psi_2}+H(\xi_1,\xi_2,\eta,\nabla f)=f.$$

Proof. By the above stretching property of the function g we have

$$g(\xi_1,\xi_2,\eta_0,\eta,\psi_1,\psi_2,t)=\frac{1}{t}g(\xi_1,\xi_2,\eta_0,\eta,t\psi_1,t\psi_2,1)=\frac{1}{t}f(\xi_1,\xi_2,\eta_0,\eta,\psi_1,\psi_2).$$

Therefore,

$$\frac{\partial g}{\partial t}=-\frac{1}{t^2}f+\frac{1}{t}\left(\psi_1\frac{\partial f}{\partial\psi_1}+\psi_2\frac{\partial f}{\partial\psi_2}\right),$$

for any t, in particular for $t=1$. On the other hand g satisfies the Hamilton-Jacobi equation. This finishes the proof. □

8. Modified action as a distance function

Theorem 5 implies that if ψ_1 and ψ_2 are critical points for the modified action f, then f is equal to the Hamiltonian $H(\xi_1,\xi_2,\eta,\nabla f)$. This allows us to interpret f as a distance function.

Theorem 6. *Suppose that the right endpoint ξ_1, ξ_2 and η does not belong to the vertical line passing through the initial point $P=(0,0,\eta_0)$. Among all critical points $\psi_1=\psi_1^{(n)}(\xi_1,\xi_2,\eta_0,\eta)$ and $\psi_2=\psi_2^{(n)}(\xi_1,\xi_2,\eta_0,\eta)$ of the modified action $f(\xi_1,\xi_2,\eta_0,\eta,\psi_1,\psi_2)$, i.e., those points satisfying the equations $\partial f/\partial\psi_1=0$ and $\partial f/\partial\psi_2=0$, there exist exactly one such that the action f evaluated at this critical point is a distance function from the point $P=(0,0,\eta_0)$ to $Q=(\xi_1,\xi_2,\eta)$. Every such critical value $\psi_1^{(n)}$ and $\psi_2^{(n)}$ defines a unique geodesic joining P and Q.*

Proof. Let us calculate the derivative

$$\frac{\partial f}{\partial\psi_1}=\xi_1+\int_0^1\left(\frac{\partial\theta(s)}{\partial\psi_1}\dot\eta+\theta(s)\frac{d}{ds}\frac{\partial\eta(s)}{\partial\psi_1}-\frac{\partial H}{\psi_1}-\frac{\partial H}{\partial\eta(s)}\frac{\partial\eta(s)}{\partial\psi_1}-\frac{\partial H}{\partial\theta(s)}\frac{\partial\theta(s)}{\partial\psi_1}\right)ds$$

$$=\xi_1+\theta(s)\frac{\partial\eta(s)}{\partial\psi_1}\Big|_{s=0}^{s=t}-\int_0^t\frac{\partial H}{\partial\psi_1}ds.$$

We need to calculate

$$\frac{\partial\eta(s)}{\partial\psi_1}\Big|_{s=0}\quad\text{and}\quad\frac{\partial\eta(s)}{\partial\psi_1}\Big|_{s=t}.$$

Similarly to the previous section we use the derivatives of the function $P(\eta(s),A)$ as

$$\frac{\partial P(\eta(s),A)}{\partial\psi_1}=\frac{\partial P(\eta(s),A)}{\partial\psi_1}+\frac{\partial P(\eta(s),A)}{\partial\eta(s)}\frac{\partial\eta(s)}{\partial\psi_1}+\frac{\partial P(\eta(s),A)}{\partial A}\frac{\partial A}{\partial\psi_1}=\frac{2s}{\sqrt{A}}\frac{\partial A}{\partial\psi_1}.$$

Substituting $s = 0$ in the latter equation we get

$$\frac{\partial \eta(s)}{\partial \psi_1}\bigg|_{s=0} = 0.$$

Consider now the derivative

$$\frac{\partial P(\eta, A)}{\partial \psi_1} = \frac{\partial P(\eta, A)}{\partial \psi_1} + \frac{\partial P(\eta, A)}{\partial A}\frac{\partial A}{\partial \psi_1} = \frac{2t}{\sqrt{A}}\frac{\partial A}{\partial \psi_1}.$$

Then

$$\frac{\partial P(\eta(s), A)}{\partial \psi_1}\bigg|_{s=t} - \frac{\partial P(\eta, A)}{\partial \psi_1} = \frac{\partial P(\eta(s), A)}{\partial \eta(s)}\frac{\partial \eta(s)}{\partial \psi_1}\bigg|_{s=t} = 0.$$

Hense,

$$\frac{\partial \eta(s)}{\partial \psi_1}\bigg|_{s=t} = 0.$$

Therefore,

$$\frac{\partial f}{\partial \psi_1} = \xi_1 - \int_0^1 \dot{\xi}_1 ds = \xi_1^0 = 0. \tag{8.1}$$

Analogously,

$$\frac{\partial f}{\partial \psi_2} = \xi_2 - \int_0^1 \dot{\xi}_2 ds = \xi_2^0 = 0. \tag{8.2}$$

On the other hand, the equations (8.1) and (8.2) can be rewritten in terms of the function $\eta(s)$ as

$$\frac{\partial f}{\partial \psi_1} = \xi_1 - \int_0^1 (\psi_1 \tan^2 \eta(s) + \psi_2) ds = 0, \quad \frac{\partial f}{\partial \psi_2} = \xi_2 - \int_0^1 (\psi_2 \cot^2 \eta(s) + \psi_1) ds = 0.$$

The critical points $\psi_1 = \psi_1^{(n)}(\xi_1, \xi_2, \eta_0, \eta)$ and $\psi_2 = \psi_2^{(n)}(\xi_1, \xi_2, \eta_0, \eta)$ are defined as solutions to the above system of two equations, which is the same as the system which defines reparametrization of geodesics for $\xi_1^0 = \xi_2^0 = 0$. Thus, every critical point $\psi_1 = \psi_1^{(n)}(\xi_1, \xi_2, \eta_0, \eta)$ and $\psi_2 = \psi_2^{(n)}(\xi_1, \xi_2, \eta_0, \eta)$ defines a unique geodesic starting at the point P and ending at the point Q as stated in the theorem. Since the point Q does not belong to the vertical line, there is a finite number of geodesics joining P and Q, see Theorem 3. We choose the geodesic which minimizes the Hamiltonian ($C^2 + D^2$ in terms of Theorem 3). Thus, the function f evaluated in its critical points corresponding to this geodesic satisfies all the properties of a distance. □

We consider now the modified action f restricted to the characteristic manifold in the cotangent bundle defined by $\psi_1 = \tau$, $\psi_2 = -\tau \tan^2 \eta_0$, $\theta = 0$. This manifold is singular because the Hamiltonian H vanishes there. The modified action does not depend on the variable θ of the tangent bundle. Choosing the initial point $(0, 0, \pi/4)$ in the phase manifold with coordinates (ξ_1, ξ_2, η), we set $\psi_1 = \tau$, $\psi_2 = -\tau$. Let us suppose the sign (+) in all formulas for the modified action. This simplify significantly the analytic expression for f which becomes of the form

$$f(\psi_1,\psi_2,\xi_1,\xi_2,\eta,\eta_0) = f(\tau,-\tau,\xi_1,\xi_2,\eta,\pi/4)$$
$$= \tau(\xi_1-\xi_2) + \frac{1}{2}A + 2\tau^2 - \frac{1}{2}\tau\arctan\left(4\frac{\tau}{\sqrt{A}}\frac{\tan 2\sqrt{A}}{1-\tan^2 2\sqrt{A}}\right)$$
$$= \tau(\xi_1-\xi_2) + \frac{1}{2}A + 2\tau^2 - \frac{1}{2}\tau\arctan\left(\frac{2\tau}{\sqrt{A}}\tan 4\sqrt{A}\right),$$

where $A > 4\tau^2$, and satisfies the equation

$$\sin^2\eta = \frac{1}{2} + (\sin 4\sqrt{A})\sqrt{\frac{1}{4}-\frac{\tau^2}{A}}.$$

We see that the action $f(\tau,-\tau,\xi_1,\xi_2,\eta,\pi/4)$ has singularities at all points

$$\tau = \pm\frac{1}{16}(\pi+2\pi n)\sin 2\eta.$$

However, the term τ^2 is dominating and the function $\exp(-f)$ exponentially decays as $\tau \to \pm\infty$ and integrable. This contrasts the case of non-compact sub-Riemannian manifolds considered earlier in, e.g., [1, 2, 3, 11], where the authors had to avoid non-integrable singularities introducing complex modified action.

9. Volume element

The aim of this section is to deduce the transport equation for the volume element v in the case of 3-D sphere. We shall use the notation of Sections 4 and 6. We choose the point $(0,0,\frac{\pi}{4})$ as a center for the heat kernel just to have simpler formulas. The characteristic variety is a straight line given by

$$\begin{cases} \psi_1 &= \tau\,\mathrm{cotan}\,\eta \\ \psi_2 &= -\tau\tan\eta \\ \theta &= 0 \end{cases}$$

Let f be the modified action studied at Section 6. We remind that it can be considered as a square of the distance from the point $(0,0,\eta_0)$, and particularly for $\eta_0=\frac{\pi}{4}$, to some point (ξ_1,ξ_2,η). We want to integrate the function $e^{-\frac{f}{u}}$ over the characteristic variety at $(0,0,\frac{\pi}{4})$ with respect to a measure $v(\xi_1,\xi_2,\eta,\tau)$. The characteristic variety at $(0,0,\frac{\pi}{4})$ gives us the relation between ψ_1 and ψ_2, namely, $\psi_1=-\psi_2=\tau$. Since f satisfies the generalized Hamilton-Jacobi equation for any ψ_1,ψ_2, we have the equation

$$2\tau\frac{\partial f}{\partial\tau} + H(\xi_1,\xi_2,\eta,\tau,\nabla_X f) = f(\xi_1,\xi_2,\eta,\tau).$$

We look for the heat kernel in the form

$$P_u(\xi_1,\xi_2,\eta) = \frac{C}{u^q}\int_{-\infty}^{+\infty} e^{-\frac{f(\xi_1,\xi_2,\eta,\tau)}{u}} v(\xi_1,\xi_2,\eta,\tau)\,d\tau. \tag{9.1}$$

Let us deduce the transport equation for $v(\xi_1, \xi_2, \eta, \tau)$. We should have

$$(\Delta_X - \frac{\partial}{\partial u})P_u(\xi_1, \xi_2, \eta) = C\int_{-\infty}^{+\infty} (\Delta_X - \frac{\partial}{\partial u})\Big(u^{-q}e^{-\frac{f(\xi_1,\xi_2,\eta,\tau)}{u}}v(\xi_1, \xi_2, \eta, \tau)\Big)\, d\tau.$$

Calculate

$$\frac{\partial}{\partial u}\Big(u^{-q}e^{-\frac{f(\xi_1,\xi_2,\eta,\tau)}{u}}v(\xi_1, \xi_2, \eta, \tau)\Big) = \Big(\frac{e^{-\frac{f}{u}}v}{u^{q+1}}\Big)\Big(-q + \frac{f}{u}\Big). \tag{9.2}$$

Then

$$\Delta_X\Big(u^{-q}e^{-\frac{f}{u}}v\Big) = \frac{e^{-\frac{f}{u}}}{u^{q+1}}\Big(v\Delta_X f + \frac{v}{u}(f - 2\tau\frac{\partial f}{\partial \tau}) - \nabla_X f\cdot\nabla_X v + u\Delta_X v\Big), \tag{9.3}$$

where we denote by $\nabla_X f = (X_1 f, \dots, X_k f)$, and use the Hamilton-Jacobi equation

$$\frac{1}{2}|\nabla_X f|^2 = H = f - 2\tau\frac{\partial f}{\partial \tau}.$$

Summing (9.2) and (9.3) we get

$$(\Delta_X - \frac{\partial}{\partial u})\Big(\frac{e^{-\frac{f}{u}}v}{u^q}\Big) = \Big(\frac{e^{-\frac{f}{u}}}{u^{q+1}}\Big)\Big((q - \Delta_X f)v - \frac{1}{u}2\tau v\frac{\partial f}{\partial \tau} - \nabla_X f\cdot\nabla_X v + u\Delta_X\Big).$$

Substituting the term $-\frac{1}{u}2e^{-\frac{f}{u}}\tau v\frac{\partial f}{\partial \tau}$ from the equality

$$-\frac{1}{u}2e^{-\frac{f}{u}}\tau v\frac{\partial f}{\partial \tau} = \frac{\partial}{\partial \tau}\Big(2e^{-\frac{f}{u}}\tau v\Big) - 2ve^{-\frac{f}{u}} - 2\tau e^{-\frac{f}{u}}\frac{\partial v}{\partial \tau},$$

we get

$$\begin{aligned}(\Delta_X - \frac{\partial}{\partial u})\Big(\frac{e^{-\frac{f}{u}}v}{u^q}\Big) &= \Big(\frac{e^{-\frac{f}{u}}}{u^{q+1}}\Big)\Big((q - 2 - \Delta_X f)v - 2\tau\frac{\partial v}{\partial \tau} - \nabla_X f\cdot\nabla_X v + u\Delta_X v\Big)\\ &+ \frac{1}{u^{q+1}}\frac{\partial}{\partial \tau}\Big(2e^{-\frac{f}{u}}\tau v\Big).\end{aligned}$$

If we were suppose that $2e^{-\frac{f}{u}}\tau v$ tends to 0 as $\tau \to \pm\infty$, and the function v were satisfy the transport equation

$$\big(q - 2 - \Delta_X f\big)v - 2\tau\frac{\partial v}{\partial \tau} - \nabla_X f\cdot\nabla_X v + u\Delta_X v = 0, \tag{9.4}$$

then the heat kernel $P_u(\xi_1, \xi_2, \eta)$ would be found as in (9.1).

Let us present the coefficient $\Delta_X f$. Since $\frac{\partial f}{\partial \xi_1} = \psi_1$, $\frac{\partial f}{\partial \xi_2} = \psi_2$, $\frac{\partial f}{\partial \eta} = \theta$, we have $\Delta_X f = 2\frac{\partial^2 f}{\partial \eta^2} = 2\frac{\partial \theta}{\partial \eta}$. For the function θ we have the expression

$$\theta^2(\eta) = \theta_0^2 - \tau^2\operatorname{cotan}^2 2\eta, \tag{9.5}$$

where θ_0 can be found from

$$\cos 2\eta = \sqrt{\frac{\theta_0^2}{\tau^2 + \theta_0^2}}\sin 8\sqrt{\tau^2 + \theta_0^2}, \tag{9.6}$$

where we took into consideration the sigh (+) in the solution (6.13). Differentiating (9.5) we find

$$\frac{\partial\theta(\eta)}{\partial\eta}=\frac{1}{2\theta(\eta)}\Big(\frac{\partial\theta_0^2}{\partial\eta}+4\tau^2\frac{\cos 2\eta}{\sin^3 2\eta}\Big).$$

Expression (9.6) gives

$$\frac{\partial\theta_0^2}{\partial\eta}=\frac{4(\tau^2+\theta_0^2)\sin 2\eta}{\frac{\tau^2}{|\theta_0|\sqrt{\tau^2+\theta_0^2}}\sin 8\sqrt{\tau^2+\theta_0^2}+8|\theta_0|\cos 8\sqrt{\tau^2+\theta_0^2}}.$$

Observe that

$$\frac{\sin 8\sqrt{\tau^2+\theta_0^2}}{\sqrt{\tau^2+\theta_0^2}}=-\frac{\cos 2\eta}{|\theta_0|},$$

and

$$|\theta_0|\cos 8\sqrt{\tau^2+\theta_0^2}=\pm\cos 2\eta\sqrt{\theta_0^2\tan^2 2\eta-\tau^2}.$$

This implies

$$\frac{\partial\theta_0^2}{\partial\eta}=\frac{4(\tau^2+\theta_0^2)\tan 2\eta}{\pm 8\sqrt{\theta_0^2\tan^2 2\eta-\tau^2}-\frac{\tau^2}{\theta_0^2}},$$

which leads to the sub-Laplacian

$$\Delta_X f=\frac{1}{\theta(\eta)}\Big(4\tau^2\frac{\cos 2\eta}{\sin^3 2\eta}+\frac{4(\tau^2+\theta_0^2)\tan 2\eta}{\pm 8\sqrt{\theta_0^2\tan^2 2\eta-\tau^2}-\frac{\tau^2}{\theta_0^2}}\Big),$$

where $\theta(\eta)$ is given by (9.5) and θ_0 can be found from (9.6). Since the coefficients of transport equation given by $\Delta_X f$ depends on η, we, unlikely the case of nilpotent groups, can not assume that the volume element v depends only on τ. In fact, it depends also on η, and we have to take into consideration the term $\nabla_X f\cdot\nabla_X v$ in the transport equation. Moreover, since

$$\begin{aligned} Xf &= -4\tau\sin(\xi_1-\xi_2)\frac{\cos 2\eta}{\sin^2(2\eta)}+2\cos(\xi_1-\xi_2)\theta(\eta),\\ Yf &= -4\tau\cos(\xi_1-\xi_2)\frac{\cos 2\eta}{\sin^2(2\eta)}-2\sin(\xi_1-\xi_2)\theta(\eta),\end{aligned}$$

we need to assume that the volume element v depends on all variables ξ_1,ξ_2,η. This assumption makes the problem of finding volume element on S^3 much more difficult, than in the case of the nilpotent groups.

References

[1] R. Beals, B. Gaveau, P. Greiner, *The Green function of model step two hypoelliptic operators and the analysis of certain tangential Cauchy Riemann complexes*, Adv. Math. **121** (1996), 288-345.

[2] R. Beals, B. Gaveau, P. Greiner, *Complex Hamiltonian mechanics and parametrices for subelliptic Laplacians I,II,II.* Bull. Sci. Math. **121** (1997), no. 1,2,3, 1–36, 97–149, 195–259.

[3] R. Beals, B. Gaveau, P. Greiner, *Hamilton-Jacobi theory and the heat kernel on the Heisenberg group.* J. Math. Pur. Appl. **79** (2000), no. 7, 633–689.

[4] R. Beals, B. Gaveau, P. Greiner, *Green's functions for some highly degenerate elliptic operators.* J. Funct. Anal. **165** (1999), no. 2, 407–429.

[5] R. Beals, B. Gaveau, P. Greiner, Y. Kannai, *Exact fundamental solutions for a class of degenerate elliptic operators.* Comm. Partial Differential Equations **24** (1999), no. 3-4, 719–742.
[6] A. Bloch, *Nonholonomic mechanics and control*, Springer-Verlag, New York, 2003.
[7] U. Boscain, G. Charlot, *Resonance of minimizers for n-level quantum systems with an arbitrary cost*, ESAIM Control Opim. Calc. Var. **10** (2004), no. 4, 593–614.
[8] R. W. Brockett, *Control theory and singular Riemannian geometry*, New Directions in Applied Math. (Cleveland, Ohio, 1980), Springer, New York–Berlin, 1982, 11–27.
[9] R. W. Brocket, L. Dai, *Nonholonomic kinematics and the role of elliptic functions in constructive controllability*, Nonholonomic Motion Planning, ed. Z. Li and J. F. Canny, vol. 122, Springer, 1993, 1–22.
[10] O. Calin, D.-Ch. Chang, I. Markina, *Sub-Riemannian geometry on the sphere S^3*, to appear in Canadian J. Math.
[11] O. Calin, D.-Ch. Chang, I. Markina, *Generalized Hamilton-Jacoby equation and heat kernel on step two nilpotent Lie groups*, to appear in Analysis and Math. Phys. Proceedings of the conference 'New trends in harmonic and complex analysis', Voss, 2007, Birchauser-Verlag, 2009.
[12] D.-Ch. Chang, I. Markina, A. Vasil'ev, *Sub-Riemannian geodesics on the 3-D sphere*, Compl. Anal. Oper. Theory, (to appear).
[13] W. L. Chow. *Uber Systeme von linearen partiellen Differentialgleichungen erster Ordnung*, Math. Ann., **117** (1939), 98-105.
[14] G. B. Folland, E. M. Stein, *Estimates for the $\bar{\partial}_b$ complex and analysis on the Heisenberg group.* Comm. Pure Appl. Math. **27** (1974), 429–522.
[15] B. Gaveau, *Principe de moindre action, propagation de la chaleur et estimées sous elliptiques sur certains groupes nilpotents*, Acta Math. **139** (1977), no. 1–2, 95–153.
[16] D. Geller, *Analytic pseudodifferential operators for the Heisenberg group and local solvability.* Mathematical Notes, **37**. Princeton University Press, Princeton, NJ, 1990, 495 pp.
[17] A. Hurtado, C. Rosales, *Area-stationary surfaces inside the sub-Riemannian tree-sphere*, Math. Ann. **340** (2008), 675–708.
[18] L. Hörmander, *Hypoelliptic second order differential equations.* Acta Math. **119** (1967), 147–171.
[19] J. J. Kohn, H. Rossi, *On the extension of holomorphic functions from the boundary of a complex manifold.* Ann. of Math. **81** (19650, no. 2, 451–472.
[20] W. Liu, H. J. Sussman, *Shortest paths for sub-Riemannian metrics on rank-two distributions.* Mem. Amer. Math. Soc. /bf 118 (1995), no. 564, 104 pp.
[21] F. Monroy-Pérez, A. Anzaldo-Meneses, *Optimal control on the Heisenberg group*, J. Dynam. Contr. Sys. **5** (1999), no. 4, 473–499.
[22] R. Montgomery, Abnormal minimizers. SIAM J. Control Optim. **32** (1994), no. 6, 1605–1620.
[23] R. Montgomery, *Survey of singular geodesics.* Sub-Riemannian geometry, 325–339, Progr. Math., 144, Birkhäuser, Basel, 1996.
[24] R. Montgomery, *A tour of subriemannian geometries, their geodesics and applications.* Mathematical Surveys and Monographs, **91**. American Mathematical Society, Providence, RI, 2002. 259 pp.
[25] L. S. Pontryagin, V. Boltyanski, R. Gamkrelidze, E. Mischenko, *The mathematical theory of optimal processes*, John Willey and Sons Inc., New York-London, 1962.
[26] W. Respondek, *Integrability and non-integrability of sub-Riemannian problems*, Workshop on Nonholonomic Dynamics and Integrability, Jan. 28–Feb. 2, 2007, abstract of the talk.
[27] L. P. Rothschild, E. M. Stein, *Hypoelliptic differential operators and nilpotent groups.* Acta Math. **137** (1976), no. 3-4, 247–320.
[28] R. S. Strichartz, *Sub-Riemannian geometry*, J. Differential Geom. **24** (1986) 221–263; Correction, ibid. **30** (1989) 595-596.
[29] M. E. Taylor, *Noncommutative microlocal analysis. I.* Mem. Amer. Math. Soc. **52** (1984), no. 313, 182 pp.

[30] A. M. Vershik, V. Ya. Gershkovich, *Nonholonomic dynamical systems. Geometry of distributions and variational problems.* Encyclopaedia Math. Sci. **16**, Dynamical Systems VII, Springer-Verlag, 1994, 4–79.

Department of Mathematics, Georgetown University, Washington D.C. 20057, USA
E-mail address: chang@georgetown.edu

Department of Mathematics, University of Bergen, Johannes Brunsgate 12, Bergen 5008, Norway
E-mail address: irina.markina@uib.no
E-mail address: alexander.vasiliev@uib.no