CHAPTER 1

Geometric Mechanics on the Heisenberg Group

The Heisenberg group and its sub-Laplacian are at the cross-roads of many domains of analysis and geometry: nilpotent Lie groups theory, hypoelliptic second order partial differential equations, strongly pseudoconvex domains in complex analysis, probability theory of degenerate diffusion process, subRiemannian geometry, control theory and semiclassical analysis of quantum mechanics, see *e.g.,* [**13**], [**16**], [**30**], [**31**], and [**44**].

Here we give a detailed discussion of the behavior of the subRiemannian geodesics on the Heisenberg group, which is the paradigm of the theory. Research on the subRiemannian geometry induced by the sub-Laplacian and its analytic consequences has been studied extensively in the past ten years (see *e.g.,* [**3**], [**13**], [**4**], [**19**], [**18**] and [**26**]). To the best of our knowledge, the first time these geodesics have been computed by Koranyi, see [**39**], [**38**], and [**40**]. In this chapter we shall provide a systematic method to handle problems along this direction.

The Heisenberg group is the prototype for the subRiemannian manifolds of step 2, which makes this case very special in the class of subRiemannian manifolds.

There are a few ways to introduce the Heisenberg group. In the first section we shall show that all of these are equivalent.

1.1. Definitions for the Heisenberg group

Let $\mathbf{G}$ be a noncommutative group. If $h, k \in \mathbf{G}$ are two elements, define the commutator of h and k by $[h,k] = hkh^{-1}k^{-1} = hk(kh)^{-1}$. If $[h,k] = e$, we say that h and k commute (e denotes the unit element of $\mathbf{G}$). The set of elements which commute with all other elements is called the center of the group $Z(\mathbf{G}) = \{g \in \mathbf{G}; [g,k] = e, \forall k \in G\}$. If $\mathbf{K} \lhd \mathbf{G}$ is a subgroup of $\mathbf{G}$ then let $[\mathbf{K}, \mathbf{G}]$ be the group generated by all commutators $[k,g]$ with $k \in \mathbf{K}$ and $g \in \mathbf{G}$. When $\mathbf{K} = \mathbf{G}$, then $[\mathbf{G}, \mathbf{G}]$ is called the commutator subgroup of $\mathbf{G}$.

Definition 1.1. Let $\mathbf{G}$ be a group. Define the sequence of groups $\big(\Gamma_n(\mathbf{G})\big)_{n\geq 1}$ by $\Gamma_0(\mathbf{G}) = \mathbf{G}$, $\Gamma_{n+1}(\mathbf{G}) = [\Gamma_n(\mathbf{G}), \mathbf{G}]$. $\mathbf{G}$ is called nilpotent if there is an $n \in \mathbb{N}$ such that $\Gamma_n(\mathbf{G}) = e$. The smallest integer n with the above property is called the class of nilpotence of $\mathbf{G}$.

The subset of $\mathcal{M}_3(\mathbb{R})$ given by

$$\mathbf{G} = \left\{ \begin{pmatrix} 1 & a & c \\ 0 & 1 & b \\ 0 & 0 & 1 \end{pmatrix} ; a, b, c \in \mathbb{R} \right\} \tag{1.1}$$

defines a noncommutative group with the usual matrix multiplication. Consider the matrices

$$A = \begin{pmatrix} 1 & a_1 & a_3 \\ 0 & 1 & a_2 \\ 0 & 0 & 1 \end{pmatrix}, \qquad B = \begin{pmatrix} 1 & b_1 & b_3 \\ 0 & 1 & b_2 \\ 0 & 0 & 1 \end{pmatrix}. \tag{1.2}$$

Then

$$AB = \begin{pmatrix} 1 & a_1 + b_1 & a_3 + b_3 + a_1 b_2 \\ 0 & 1 & a_2 + b_2 \\ 0 & 0 & 1 \end{pmatrix}, \tag{1.3}$$

$$A^{-1} = \begin{pmatrix} 1 & -a_1 & a_1 a_2 - a_3 \\ 0 & 1 & -a_2 \\ 0 & 0 & 1 \end{pmatrix}, \qquad B^{-1} = \begin{pmatrix} 1 & -b_1 & b_1 b_2 - b_3 \\ 0 & 1 & -b_2 \\ 0 & 0 & 1 \end{pmatrix} \tag{1.4}$$

The commutator

$$[A, B] = ABA^{-1}B^{-1} = \begin{pmatrix} 1 & 0 & a_1 b_2 - b_1 a_2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

and hence the commutator subgroup is

$$\Gamma_1(\mathbf{G}) = [\mathbf{G}, \mathbf{G}] = \langle [A, B]; A, B \in \mathbf{G} \rangle = \left\{ \begin{pmatrix} 1 & 0 & k \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}; k \in \mathbb{R} \right\}.$$

Let

$$C = \begin{pmatrix} 1 & 0 & k \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Then

$$AC = \begin{pmatrix} 1 & a & c + k \\ 0 & 1 & b \\ 0 & 0 & 1 \end{pmatrix} = CA,$$

and therefore $[A, C] = AC(AC)^{-1} = I_3$. Hence $\Gamma_2(\mathbf{G}) = [\Gamma_1(\mathbf{G}), \mathbf{G}] = I_2 = e$, and the group $\mathbf{G}$ is nilpotent of class 2. $\mathbf{G}$ is called the Heisenberg group with 3 parameters.

The nilpotence class measures the noncommutativity of the group. In the following we shall associate with this group a noncommutative geometry of step 2. This geometry will have the Heisenberg principle built in.

The bijection $\phi : \mathbb{R}^3 \to \mathcal{M}_3(\mathbb{R})$,

$$\phi(x_1, x_2, t) = \begin{pmatrix} 1 & x_1 & t \\ 0 & 1 & x_2 \\ 0 & 0 & 1 \end{pmatrix}$$

induces a noncommutative group law structure on $\mathbb{R}^3$

$$(x_1, x_2, t) \circ (x_1', x_2', t') = (x_1 + x_1', x_2 + x_2', t + t' + x_1 x_2'). \tag{1.5}$$

The zero element is $e = (0, 0, 0)$ and the inverse of (x_1, x_2, t) is $(-x_1, -x_2, x_1 x_2 - t)$.

$\mathbb{R}^3$ together with the group law (1.5) will be called the **nonsymmetric 3-dimensional Heisenberg group**. This group can be regarded also as a Lie group. The left translation $L_a : \mathbf{G} \to \mathbf{G}$, $L_a g = ag$, $\forall g \in \mathbf{G}$ is an analytic diffeomorphism with inverse $L_a^{-1} = L_{a^{-1}}$. A vector field X on $\mathbf{G}$ is called left invariant if

$$(L_a)_*(X_g) = X_{ag}, \qquad \forall a, g \in \mathbf{G}.$$

The set of all left invariant vector fields form the Lie algebra of $\mathbf{G}$, denoted by $L(\mathbf{G})$. The Lie algebra of $\mathbf{G}$ has the same dimension as $\mathbf{G}$ and it is isomorphic to the tangent space $T_e\mathbf{G}$. We shall use this result in the following proposition in order to compute a basis for the Lie algebra of the Heisenberg group.

Proposition 1.2. The vector fields

$$X = \partial_{x_1}, \quad Y = \partial_{x_2} + x_1\partial_t, \quad T = \partial_t \tag{1.6}$$

are left invariant with respect to the Lie group law (1.5) on $\mathbb{R}^3$.

Proof. Consider the notation $x_3 = t$. In this case the left translation is

$$L_{(a_1,a_2,a_3)}(x_1, x_2, x_3) = (a_1 + x_1, a_2 + x_2, a_3 + x_3 + a_1x_2).$$

Let X be a left invariant vector field. Then $\forall a = (a_1, a_2, a_3) \in \mathbf{G}$, $X_a = (L_a)_*X_e$. In local coordinates $X_a = \sum_i X_a^i \partial_{x_i}$. The components are given by

$$X_a^i = X_a(x_i) = (L_a)_*X_e(x_i) = X_e(x_i \circ L_a), \tag{1.7}$$

where x_i is the i-th coordinate and X_e is the value of the vector field X at the origin. Let $b = (b_1, b_2, b_3) \in \mathbf{G}$. Then

$$\begin{aligned}
(x_1 \circ L_a)(b) &= x_1(L_ab) = x_1(ab) = a_1 + b_1 = x_1(a) + x_1(b),\\
(x_2 \circ L_a)(b) &= x_2(L_ab) = x_2(ab) = a_2 + b_2 = x_2(a) + x_2(b),\\
(x_3 \circ L_a)(b) &= x_3(L_ab) = x_3(ab) = a_3 + b_3 + a_1b_2 = x_3(a) + x_3(b) + x_1(a)x_2(b).
\end{aligned}$$

Dropping b, yields

$$\begin{aligned}
x_1 \circ L_a &= x_1(a) + x_1,\\
x_2 \circ L_a &= x_2(a) + x_2,\\
x_3 \circ L_a &= x_3(a) + x_3 + x_1(a)x_2.
\end{aligned}$$

Substituting in equation (1.7) and using $X_e = \xi^1\partial_{x_1} + \xi^2\partial_{x_2} + \xi^3\partial_{x_3}$, yields

$$\begin{aligned}
X_a^1 &= X_e\big(x_1(a) + x_1\big) = \xi^1,\\
X_a^2 &= X_e\big(x_2(a) + x_2\big) = \xi^2,\\
X_a^3 &= X_e\big(x_3(a) + x_3 + x_1(a)x_2\big) = \xi^3 + x_1(a)\xi^2.
\end{aligned}$$

Hence, the left invariant vector field X depends on the parameters ξ^i

$$\begin{aligned}
X &= \xi^1\partial_{x_1} + \xi^2\partial_{x_2} + (\xi^3 + x_1\xi^2)\partial_{x_3}\\
&= \xi^1\partial_{x_1} + \xi^2(\partial_{x_2} + x_1\partial_{x_3}) + \xi^3\partial_{x_3}\\
&= \xi^1X + \xi^2Y + \xi^3T,
\end{aligned}$$

and the Lie algebra is generated by the linear independent vector fields X, Y and T. □

In the following we shall show that the nonsymmetric model can always be reduced to a symmetric model, using a coordinate transformation.

PROPOSITION 1.3. Under the change of coordinates

$$y_1 = x_1, \quad y_2 = x_2, \quad \tau = 4t - 2x_1x_2,$$

the vector fields

$$X = \partial_{x_1}, \quad Y = \partial_{x_2} + x_1\partial_t, \quad T = \partial_t,$$

are transformed into

$$X = \partial_{y_1} - 2y_2\partial_\tau, \quad Y = \partial_{y_2} + 2y_1\partial_\tau, \quad T = 4\partial_\tau. \tag{1.8}$$

PROOF. The proof follows from the following relationships

$$\begin{aligned}\partial_t &= 4\partial_\tau,\\ \partial_{x_2} &= \partial_{y_2} - 2y_1\partial_\tau,\\ \partial_{x_1} &= \partial_{y_1} - 2y_2\partial_\tau.\end{aligned}$$

□

Making $y_1 = x_1$, $y_2 = x_2$, $\tau = -t$ in (1.8) we consider the vector fields $X_1 = \partial_{x_1} + 2x_2\partial_t$, $X_2 = \partial_{x_2} - 2x_1\partial_t$, $X_3 = \partial_t$ on $\mathbb{R}^3 = \mathbb{R}^2_x \times \mathbb{R}_t$. We are interested in a Lie group law on $\mathbb{R}^3$ such that X_1, X_2 and X_3 are left invariant. This shall be done using the Campbell-Hausdorff formula. The constants of structure are denoted by c^k_{ij} and are defined by

$$[X_i, X_j] = \sum_{k=1}^{3} c^k_{ij} X_k.$$

From $[X_1, X_2] = -4\partial_t$ and $[X_1, \partial_t] = [X_2, \partial_t] = 0$, the constants of structure are

$$c^1_{12} = c^2_{12} = 0, \qquad c^3_{12} = -4,$$

$$c^j_{13} = c^j_{23} = 0, \qquad j = 1, 2, 3.$$

If $x = (x_1, x_2, x_3)$ and $y = (y_1, y_2, y_3)$, a locally Lie group structure is given by the Campbell-Hausdorff formula:

$$(x \circ y)_i = x_i + y_i + \frac{1}{2}\sum_{j,k} c^i_{jk}\, x_j\, y_k + \frac{1}{12}\sum_{k,s,p,j} x_p\, y_j\, x_s\, c^k_{pj}\, c^i_{ks} + \dots \tag{1.9}$$

In our case, we obtain a globally defined group structure

$$\begin{aligned}(x \circ y)_1 &= x_1 + y_1,\\ (x \circ y)_2 &= x_2 + y_2,\\ (x \circ y)_3 &= x_3 + y_3 - \frac{1}{2} \cdot 4(x_1y_2 - x_2y_1),\end{aligned}$$

because the term $x_p\, y_j\, x_s\, c^k_{pj}\, c^3_{ks} = 0$. We arrived at the following result.

PROPOSITION 1.4. The vector fields $X_1 = \partial_{x_1} + 2x_2\partial_t$, $X_2 = \partial_{x_2} - 2x_1\partial_t$ are left invariant with respect to the following Lie group law on $\mathbb{R}^3$

$$(x_1, x_2, t) \circ (x_1', x_2', t') = \big(x_1 + x_1', x_2 + x_2', t + t' - 2(x_1x_2' - x_2x_1')\big). \tag{1.10}$$

The Lie group $\mathbf{H}_1 = (\mathbb{R}^3, \circ)$ is called the symmetric three dimensional **Heisenberg group**. The unit element is $e = (0,0,0)$ and the inverse is $(x_1, x_2, t)^{-1} = (-x_1, -x_2, -t)$.

We shall study the subRiemannian geometry associated with this model. The geometry has the Heisenberg uncertainty principle

$$[X_1, X_2] = -4\partial_t \tag{1.11}$$

built in. This brings the hope, that Heisenberg manifolds (step 2 subRiemannian manifolds) will play a role for Quantum Mechanics in the future, similar to the role played by the Riemannian manifolds for Classical Mechanics. (see *e.g.* [**2**], [**1**], [**14**]). In Quantum Mechanics, the states of a quantum particle (position, momentum) are described by differential operators. It is known that two states, which cannot be measured simultaneously correspond to operators, which do not commute (see [**45**]).

For instance, if $\mathbf{x}$ and $\mathbf{p}$ are the position and the momentum of a particle, then one cannot measure them simultaneously and hence, we write $[\mathbf{x}, \mathbf{p}] \neq 0$. The state of the particle is measured using a radiation beam sent towards the particle. The radiation is reflected partially back. Using the variation of frequency between the sent and reflected beams, the Doppler-Fizeau formula will provide the speed of the particle. This method will provide accurate results if the radiation will not significantly change the speed of the particle *i.e.,* its kinetic energy $K = mv^2/2$. This means the radiation has a low energy and hence, a low frequency, because $E_{radiation} = h\nu$. Therefore the wave length of the radiation $\lambda = 1/\nu$ will be large. In this case the position of the particle cannot be measured accurately (see [**32**]).

In order to measure the position accurately, the radiation wave length has to be as small as possible. In this case the frequency ν is large as will be the energy $E_{radiation}$. This will change the kinetic energy of the particle and hence its velocity. Hence, one cannot measure accurately both the position and the speed of the particle.

The Heisenberg uncertainty principle, fundamental in the study of quantum particles, can be found also in other examples at the large scale structure. Let's assume that you are watching high-street traffic from an airplane. You will notice the position of the cars but you cannot say too much about their speed. They look like they are not moving at all. A policeman on the road will see the picture completely differently. For him, the speed of the cars will make more sense than their position. The latter is changing too fast to be noticed accurately.

The Heisenberg group is also closely linked with theory of functions in several complex variables. Let us start with the unit disc $\{w \in \mathbb{C} : |w| \leq 1\}$. Consider the Cayley transform:

$$w = \frac{i - z}{i + z}. \tag{1.12}$$

It is known that the transformation (1.12) maps the real line $\mathrm{Im}(z) = 0$ together with the point "∞" into the unit circle $|w| = 1$, and the upper half plane $\{z = x + iy \in \mathbb{C} : \mathrm{Im}(z) > 0\}$ into the disc $|w| < 1$. From the point of view of Fourier analysis, there are two groups of holomorphic self-mappings on it which are of importance to us. The first one is the translation group, $z = x + iy \rightarrow (x + h) + iy$,

$h \in \mathbb{R}$, which is isomorphic to $\mathbb{R}$. This group acts transitively on each level set $\{z = x + iy \in \mathbb{R}^2_+ : y = c, c \geq 0\}$, which is responsible for the fact that many important operators, *e.g.,* Hilbert transform and Cauchy operator, are convolution operators. The second one is the dilation group, $z \to \delta z$, $\delta > 0$, which determines the homogeneity of the kernels of these operators. The analogues of these groups are equally significant.

A complex analogue of the upper-half plane $\mathbb{R}^2_+$ is the Siegel upper-half space

$$\mathcal{U} = \Big\{(z_1, z_2) \in \mathbb{C}^2 : \operatorname{Im}(z_2) > |z_1|^2\Big\}.$$

Its boundary is the "paraboloid"

$$\partial\mathcal{U} = \big\{(z_1, z_2) \in \mathbb{C}^2 : \operatorname{Im}(z_2) = |z_1|^2\big\}.$$

The domain $\mathcal{U}$ is biholomorphically equivalent with the unit ball

$$B_2 = \big\{(w_1, w_2) \in \mathbb{C}^2 : |w_1|^2 + |w_2|^2 < 1\big\}$$

via the generalized Cayley transform:

$$w_1 = \frac{2iz_1}{i + z_2}, \quad w_2 = \frac{i - z_2}{i + z_2}.$$

In this correspondence the boundary $\partial\mathcal{U}$, together with the "point at ∞" maps onto the unit sphere in $\mathbb{C}^2$.

It is known that the Heisenberg group $\mathbf{H}_1$ gives the translations of the domain $\mathcal{U}$. The action of an element $(z, t) \in \mathbf{H}_1$ on the domain $\mathcal{U}$ is given by

$$(\xi_1, \xi_2) \to \big(\xi_1 + z, \xi_2 + t + i(2\xi_1 \cdot \bar{z}_1 + |z_1|^2)\big). \tag{1.13}$$

Since $|\xi_1 + z_1|^2 - |z_1|^2 = \operatorname{Im}\{i(2\xi_1 \cdot \bar{z}_1 + |z_1|^2)\}$, this mapping preserves each "level surface"

$$\operatorname{Im}(\xi_2) - |\xi_1|^2 = \text{constant}.$$

It follows that the mapping (1.13) is simply transitive on the boundary $\partial\mathcal{U}$: the boundary $\partial\mathcal{U}$ therefore may be identified as the orbit of the "origin" $(0, 0)$ under the action of the group $\mathbf{H}_1$:

$$\mathbf{H}_1 \in (z, t) \mapsto (z, t + i|z|^2) \in \partial\mathcal{U}.$$

The second group consists of dilations. The action on the domain $\mathcal{U}$ is given by

$$(\xi_1, \xi_2) \to (\delta\xi_1, \delta^2\xi_2), \qquad \delta > 0.$$

It is easy to see that these are automorphisms of the group structure of $\mathbf{H}_1$. Summarizing our discussion, the group $\mathbf{H}_1$ is isomorphic to the boundary $\partial\mathcal{U}$. Similar to analysis in $\mathbb{R}^2$, we may use the Heisenberg group and the Siegel upper-half space to "approximate" a bounded strongly pseudoconvex domain in $\mathbb{C}^{n+1}$. This phenomena indeed happened in many situations, see Folland-Stein [**24**] , Phong-Stein [**49**], [**48**], and Greiner-Stein [**28**].

1.2. The horizontal distribution

Unlike on Riemannian manifolds, where one may measure the velocity and distances in all directions, on Heisenberg manifolds there are directions where we cannot say anything using direct methods.

On the Heisenberg group, an important role is played by the distribution $\mathcal{H}$ generated by the linearly independent vector fields X_1 and X_2:

$$x \to \mathcal{H}_x = \text{span}_x\{X_1, X_2\},$$

called the *horizontal distribution.* As $[X_1, X_2] = -4\partial_t \notin \mathcal{H}$, the horizontal distribution $\mathcal{H}$ is not involutive, and hence, by Frobenius theorem, it is not integrable, *i.e.,* there is no surface locally tangent to $\mathcal{H}$.

A vector field V on $\mathbb{R}^3$ is called *horizontal* if and only if $V_x \in \mathcal{H}_x$, $\forall x$. A curve $c : [0,1] \to \mathbb{R}^3$ is called horizontal if the velocity vector $\dot{c}(s)$ is a horizontal vector field along $c(s)$. Horizontality is a constraint on the velocities and it is called a *non-holonomic constraint.*

In this chapter we shall construct many horizontal objects, *i.e.,* geometrical objects which can be constructed directly from the horizontal distribution and the subRiemannian metric defined on it. The main goal is to recover the external structure of the space, such as the missing direction ∂_t by means of horizontal objects.

PROPOSITION 1.5. A curve $c = (x_1, x_2, t)$ is horizontal if and only if

$$\dot{t} = 2(\dot{x}_1 x_2 - x_1 \dot{x}_2). \tag{1.14}$$

PROOF. The velocity vector can be written as

$$\begin{aligned} \dot{c} &= \dot{x}_1 \partial_{x_1} + \dot{x}_2 \partial_{x_2} + \dot{t}\partial_t \\ &= \dot{x}_1\big(\partial_{x_1} + 2x_2\partial_t\big) - 2\dot{x}_1 x_2 \partial_t \\ &\quad + \dot{x}_2\big(\partial_{x_2} - 2x_1\partial_t\big) + 2x_1\dot{x}_2\partial_t + \dot{t}\partial_t \\ &= \dot{x}_1 X_1 + \dot{x}_2 X_2 + \big(\dot{t} + 2x_1\dot{x}_2 - 2x_2\dot{x}_1\big)\partial_t. \end{aligned}$$

Hence, $\dot{c} \in \mathcal{H}$ iff the coefficient of ∂_t vanishes *i.e.*, (1.14) holds. □

COROLLARY 1.6. A curve $c = (x_1, x_2, t)$ is horizontal if and only if

$$\dot{c} = \dot{x}_1 X_1 + \dot{x}_2 X_2. \tag{1.15}$$

In the following we shall give a geometrical interpretation for the t component of a horizontal curve. This will be used in the proof of the connectivity theorem later. Using the polar coordinates $x_1 = r\cos\phi$, $x_2 = r\sin\phi$, equation (1.14) becomes

$$\begin{aligned} \dot{t} &= 2(\dot{x}_1 x_2 - x_1 \dot{x}_2) \\ &= -2r^2\dot{\phi}(\sin^2\phi + \cos^2\phi) \\ &= -2r^2\dot{\phi}. \end{aligned}$$

In differential notation

$$dt = -2r^2\, d\phi. \tag{1.16}$$

Let $r = r(\phi)$ be the equation in polar coordinates of the projection of the horizontal curve on the x-plane. The area of an infinitesimal triangle with vertices at the origin,

$(r(\phi), \phi)$ and $(r(\phi + d\phi), \phi + d\phi)$ is $\frac{1}{2}r(\phi)r(\phi + d\phi)d\phi \approx \frac{1}{2}r^2(\phi)d\phi$. Integrating, we obtain the area swept by the vectorial radius between the initial angle ϕ_0 and ϕ, see Figure 1.1.

$$\mathcal{A} = \frac{1}{2}\int_{\phi_0}^{\phi} r^2(\phi)\, d\phi.$$

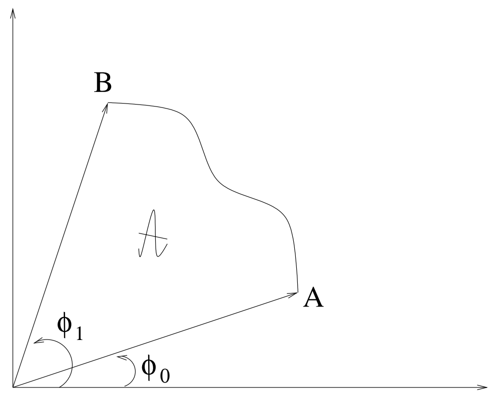

Figure 1.1 : The area swept by the vectorial radius between two points in the plane.

Taking the derivative, we obtain

$$\frac{d\mathcal{A}}{d\phi} = \frac{1}{2}r^2(\phi),$$

or

$$d\mathcal{A} = \frac{1}{2}r^2\, d\phi. \tag{1.17}$$

Dividing the equations (1.16) and (1.17) yields

$$\frac{dt}{d\mathcal{A}} = -4, \tag{1.18}$$

which says that the t-component is roughly the area swept by the vectorial radius on the x-plane, up to a multiplication factor. The negative sign in (1.18) shows that when t is increasing, the rotation in the x-plane is done clock-wise. The equation (1.18) is valid only if t is not constant.

The *areal velocity* is defined as

$$\alpha = \frac{d\mathcal{A}}{ds}.$$

Areal velocity also appears in Kepler's laws: all of the planets have plane trajectories, which are ellipses with the sun in one of the focuses. The areal velocity α is constant along the motion.

THEOREM 1.7. *A curve c in $\mathbb{R}^3$ is horizontal if and only if the rate of change of the t-component is equal to 4α, i.e., $\dot{t} = 4\alpha$.*

PROOF. $c = (x_1, x_2, t)$ is a horizontal curve iff

$$\dot{t} = 2(\dot{x}_1 x_2 - x_1 \dot{x}_2),$$

which in polar coordinates is

$$\dot{x}_1 x_2 - x_1 \dot{x}_2 = r^2 \dot{\phi},$$

and hence

$$\alpha = \frac{1}{2}(\dot{x}_1 x_2 - x_1 \dot{x}_2) = \frac{1}{4}\dot{t}.$$

□

The characterization of horizontal curves with t constant is given in the following result.

PROPOSITION 1.8. A smooth curve $c(s)$ is horizontal with $t(s) = t$ constant if and only if $c(s) = (as, bs, t)$, with $a, b \in \mathbb{R}$, $a^2 + b^2 \neq 0$.

PROOF. If $c(s)$ is horizontal with t constant, the equation $\dot{t} = -2r^2\dot{\phi}$ yields ϕ =constant. Hence the projection on the x-space is a line which passes through the origin.
If $c(s) = (as, bs, t)$, with t constant, then

$$2(\dot{x}_1 x_2 - x_1 \dot{x}_2) = 2(abs - abs) = 0 = \dot{t}$$

and hence the horizontality condition (1.14) holds. □

The following proposition shows that the left translation of a horizontal curve is a horizontal curve.

PROPOSITION 1.9. If $c(s)$ is a horizontal curve, then $\overline{c}(s) = L_a c(s)$ is a horizontal curve, for any $a \in \mathbf{H_1}$.

PROOF. If $c = (c_1, c_2, c_3)$ and $\overline{c} = (\overline{c}_1, \overline{c}_2, \overline{c}_3)$, then

$$\begin{aligned} \overline{c}_1 &= a_1 + c_1 \Longrightarrow \dot{\overline{c}}_1 = \dot{c}_1 \\ \overline{c}_2 &= a_2 + c_2 \Longrightarrow \dot{\overline{c}}_2 = \dot{c}_2 \\ \overline{c}_3 &= a_3 + c_3 - 2(a_1 c_2 - a_2 c_1) \Longrightarrow \dot{\overline{c}}_3 = \dot{c}_3 - 2(a_1 \dot{c}_2 - a_2 \dot{c}_1). \end{aligned} \tag{1.19}$$

Using that c is horizontal, equation (1.14) yields

$$\begin{aligned} \dot{\overline{c}}_3 &= \dot{c}_3 - 2(a_1 \dot{c}_2 - a_2 \dot{c}_1) \\ &= 2(\dot{c}_1 c_2 - c_1 \dot{c}_2) - 2(a_1 \dot{c}_2 - a_2 \dot{c}_1) \\ &= 2\big(\dot{c}_1(a_2 + c_2) - \dot{c}_2(a_1 + c_1)\big) \\ &= 2\big(\dot{\overline{c}}_1 \overline{c}_2 - \dot{\overline{c}}_2 \overline{c}_1\big). \end{aligned}$$

From equation (1.14) it follows that $\overline{c}(s)$ is horizontal. □

COROLLARY 1.10. The velocity of the horizontal curve $\overline{c}(s) = L_a c(s)$ is

$$\dot{\overline{c}}(s) = \big(L_a\big)_* \dot{c}(s) = \dot{c}_1(s) X_{1|\overline{c}(s)} + \dot{c}_2(s) X_{2|\overline{c}(s)}. \tag{1.20}$$

PROOF. As $\overline{c}(s)$ is horizontal, from equations (1.15) and (1.19) we have

$$\begin{aligned} \dot{\overline{c}}(s) &= \dot{\overline{c}}_1(s) X_{1|\overline{c}(s)} + \dot{\overline{c}}_2(s) X_{2|\overline{c}(s)} \\ &= \dot{c}_1(s) X_{1|\overline{c}(s)} + \dot{c}_2(s) X_{2|\overline{c}(s)} \\ &= \dot{c}_1(s) \big(L_a\big)_* X_{1|c(s)} + \dot{c}_2(s) \big(L_a\big)_* X_{2|c(s)} \\ &= \big(L_a\big)_* \big(\dot{c}_1(s) X_{1|c(s)} + \dot{c}_2(s) X_{2|c(s)}\big) \\ &= \big(L_a\big)_* \dot{c}(s). \end{aligned}$$

□

1.3. Horizontal connectivity theorem

On the Heisenberg group $\mathbf{H_1}$, the vector fields X_1, X_2 and $[X_1, X_2]$ generate the tangent space of $\mathbb{R}^3$ at every point. Such a subRiemannian manifold is called step 2. In the case of a general subRiemannian manifold, the number of brackets needed to generate all directions $+1$ is called the step of the manifold. The higher the step, the more noncommutative and the harder it will be to study the geometry. The step 1 corresponds to Riemannain geometry, which is the commutative case. The step condition was used independently by Chow [**22**] and Hörmander [**33**] to study connectivity and hypoellipticity, respectively. Using Hörmander's theorem, the Heisenberg Laplacian $\Delta_H = \frac{1}{2}(X_1^2 + X_2^2)$ is hypoelliptic, *i.e.,* $\Delta_H u = f \in C^\infty \Longrightarrow u \in C^\infty$.

In the following we shall prove Chow's connectivity theorem (see [**22**]) in the particular case of the Heisenberg group $\mathbf{H_1}$.

Proposition 1.11. Any two points in $\mathbf{H_1}$ can be joined by a piece-wise horizontal curve, *i.e.,* a curve tangent to the horizontal distribution.

Proof. Let P and Q be two points in $\mathbb{R}^3$. Let t_P and t_Q be the t-coordinates of P and Q. We distinguish between the following two cases:
case (i): $t_P \neq t_Q$
Consider the number $\alpha = t_P - t_Q \neq 0$. Let P_1 and Q_1 be the projections on the x-plane of the points P and Q. Consider in the x-plane a curve $\bar{\phi} : [0, 1] \to \mathbb{R}^2$ which joins P_1 and Q_1, such that the area situated between the graph of $\bar{\phi}$ and the segments OP_1 and OQ_1 is equal to $\alpha/4$. The area will be considered positive in the case of a counter clock-wise rotation of the curve $\bar{\phi}$ between P_1 and Q_1, see Figure 1.2. If $\bar{\phi}(s) = (x_1(s), x_2(s))$, then consider the function

$$t(s) = t_P + 2\int_0^s \Big(x_2(u)\dot{x}_1(u) - x_1(u)\dot{x}_2(u)\Big)\,du. \tag{1.21}$$

We claim that $\phi : [0, 1] \to \mathbb{R}^3$ defined as $\phi(s) = \big(\bar{\phi}(s), t(s)\big)$ is a horizontal curve joining P and Q. Differentiating in (1.21) we obtain the horizontality condition $\dot{t}(s) = 2\big(x_2(s)\dot{x}_1(s) - x_1(s)\dot{x}_2(s)\big)$ and hence, $\phi(s)$ is a horizontal curve. We shall check that ϕ joins the points P and Q.

$$\phi(0) = (\bar{\phi}(0), t(0)) = (x(P_1), t_P) = P.$$

Using $t(0) = t_P$, integrating between 0 and 1 in equation (1.18) yields

$$\begin{aligned} t(1) &= t(0) - 4\big(\mathcal{A}(1) - \mathcal{A}(0)\big) \\ &= t(0) - \alpha = t_P - (t_P - t_Q) \\ &= t_Q, \end{aligned}$$

and hence, $\phi(1) = (\bar{\phi}(1), t(1)) = (x(Q_1), t_Q) = Q$.

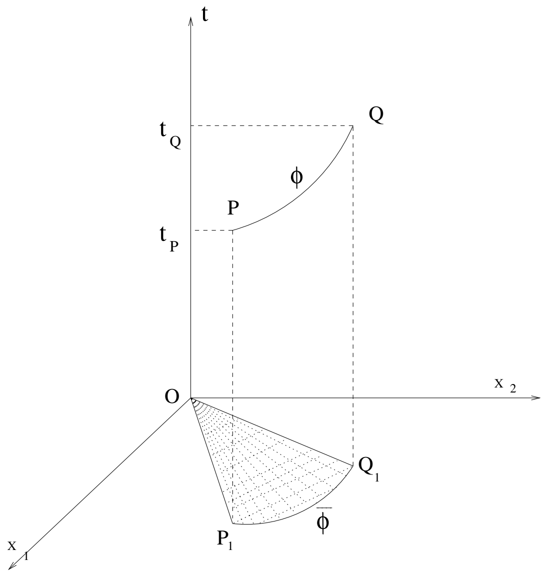

Figure 1.2: The projection of a horizontal curve.

case (ii): $t_P = t_Q = t$
Let $R = (0, 0, t)$. From Proposition 1.8 the curves $c_i : [0, 1] \to \mathbb{R}^3$

$$c_1(s) = \big(sx_1(P), sx_2(P), t\big),$$

$$c_2(s) = \big(sx_1(Q), sx_2(Q), t\big)$$

are horizontal and join the points R with P and R with Q, respectively. Then the piece-wise defined curve

$$\phi(s) = \begin{cases} c_1(1-2s), & 0 \le s \le \frac{1}{2} \\ c_2(2s-1), & \frac{1}{2} \le s \le 1 \end{cases}$$

is horizontal with $\phi(0) = c_1(1) = P$ and $\phi(1) = c_2(1) = Q$. □

The piece-wise condition in the above proposition can be dropped. The proof can be modified such that any two given points can be connected by a horizontal smooth curve. In order to do that, we should take advantage of the group law.

Given the points $P(x_1, y_1, t_1)$ and $Q(x_2, y_2, t_2)$, a translation to the left by $(-x_1, -y_1, -t_1)$ will transform them into $O(0, 0, 0)$ and $S(x', y', t')$, with

$$x' = x_2 - x_1, \quad y' = y_2 - y_1, \quad t' = t_2 - t_1 - 2(y_1x_2 - x_1y_2).$$

If $t' \neq 0$, applying case (i) of the previous proposition, we get a smooth horizontal curve $c(s)$ joining O and S.
If $t' = 0$, then $c(s) = (sx', sy', 0)$ is horizontal and joins O and S, see Proposition 1.8. Translating to the left by (x_1, y_1, t_1), the points O and S are sent into P and Q, respectively. Applying Proposition 1.9, the curve $c(s) = (c_1(s), c_2(s), c_3(s))$ is sent into the horizontal smooth curve between P and Q

$$(1.22)\quad (x_1, y_1, t_1) \circ c(s) = \Big(x_1 + c_1(s), y_1 + c_2(s), t_1 + c_3(s) - 2(x_1c_2(s) - y_1c_1(s)\Big).$$

COROLLARY 1.12. By a left translation, the t-axis $c(s) = (0, 0, s)$ is transformed into

$$L_{(x_1,y_1,t_1)}c(s) = (x_1, y_1, t_1 + s).$$

PROOF. It is an obvious consequence of equation (1.22). □

In the following we shall find an explicit smooth horizontal curve between the origin and a given point $P(\mathbf{x_1}, \mathbf{x_2}, \mathbf{t})$.

LEMMA 1.13. *For any $a, b, \alpha, \beta \in \mathbb{R}$, the curve $c(s) = \Big(x_1(s), x_2(s), t(s)\Big)$,*

$$\begin{aligned} x_1(s) &= as + bs^2 \\ x_2(s) &= \alpha s^2 + \beta s^3 \\ t(s) &= -\frac{2}{3}\alpha a s^3 - a\beta s^4 - \frac{2}{5}b\beta s^5 \end{aligned}$$

is horizontal and passes through the origin.

PROOF. A computation shows that

$$\begin{aligned} \dot{x}_1 x_2 - x_1 \dot{x}_2 &= (a + 2bs)(\alpha s^2 + \beta s^3) - (as + bs^2)(2\alpha a + 3\beta s^2) \\ &= -\alpha a s^2 - 2a\beta s^3 - b\beta s^4 = \frac{1}{2}\dot{t} \end{aligned}$$

and hence the curve is horizontal. For $s = 0$ the curve passes through the origin. □

We shall show that given $\mathbf{x_1}, \mathbf{x_2}, \mathbf{t}$, we can chose the constants a, b, α, β such that the curve given by Lemma 1.13 passes through the point $P(\mathbf{x_1}, \mathbf{x_2}, \mathbf{t})$ for $s = 1$. This means that a, b, α and β are solutions of the following system

$$\begin{aligned} \mathbf{x_1} &= a + b \\ \mathbf{x_2} &= \alpha + \beta \\ \mathbf{t} &= -\frac{2}{3}\alpha a - a\beta - \frac{2}{5}b\beta. \end{aligned}$$

By computation

$$\begin{aligned} 15\mathbf{t} &= -10\alpha a - 15a\beta - 6b\beta = -10a(\alpha + \beta) - 5a\beta - 6b\beta \\ &= -10a\mathbf{x_2} - (5a + 6b)\beta = -10a\mathbf{x_2} - (5\mathbf{x_1} + b)\beta \\ &= -10(\mathbf{x_1} - b)\mathbf{x_2} - 5\mathbf{x_1}\beta - b\beta = -10\mathbf{x_1}\mathbf{x_2} + 10b\mathbf{x_2} - (5\mathbf{x_1} + b)\beta. \end{aligned}$$

Hence

$$15\mathbf{t} + 10\mathbf{x_1}\mathbf{x_2} - 10b\mathbf{x_2} + (5\mathbf{x_1} + b)\beta = 0.$$

Choose b such that $5\mathbf{x_1} + b \neq 0$. Then

$$\beta = \frac{-15\mathbf{t} - 10\mathbf{x_1}\mathbf{x_2} + 10b\mathbf{x_2}}{5\mathbf{x_1} + b},$$

and

$$\alpha = \mathbf{x_2} - \beta, \qquad a = \mathbf{x_1} - b.$$

Hence we have constructed a family of horizontal curves between the origin and P, which depends on the parameter b.

1.4. Hamiltonian formalism on the Heisenberg group

The Heisenberg group is a good environment to apply the Hamiltonian formalism. Consider the Hamiltonian $H : T^*\mathbb{R}^3_{(x,t)} \to \mathbb{R}$ given by

$$H(\xi, \theta, x, t) = \frac{1}{2}(\xi_1 + 2x_2\theta)^2 + \frac{1}{2}(\xi_2 - 2x_1\theta)^2, \tag{1.23}$$

which is the principal symbol of the Heisenberg Laplacian

$$\Delta_H = \frac{1}{2}(X_1^2 + X_2^2), \tag{1.24}$$

where $X_1 = \partial_{x_1} + 2x_2\partial_t$, $X_2 = \partial_{x_2} - 2x_1\partial_t$. In Quantum Mechanics, the procedure of obtaining the operator (1.24) from the Hamiltonian (1.23) is called *quantization.*

It is natural to consider the Hamiltonian system

$$\begin{aligned} \dot{x} &= \partial H/\partial \xi \\ \dot{t} &= \partial H/\partial \theta \\ \dot{\xi} &= -\partial H/\partial x \\ \dot{\theta} &= -\partial H/\partial t. \end{aligned} \tag{1.25}$$

The solutions $c(s) = \big(x(s), t(s), \xi(s), \theta(s)\big)$ of the system (1.25) are called *bicharacteristics.*

DEFINITION 1.14. Given two points $P(x_0, t_0)$, $Q(x_1, t_1) \in \mathbb{R}^3$, a geodesic between P and Q is the projection on the (x, t)-space of a bicharacteristic $c : [0, \tau] \to \mathbb{R}^3$, which satisfies the boundary conditions:

$$(x(0), t(0)) = (x_0, t_0), \qquad (x(\tau), t(\tau)) = (x_1, t_1). \tag{1.26}$$

The most basic questions we shall answer in this chapter are:

- Given any two points, can we join them by a geodesic?
- How many geodesics are between any two given points?

We shall start by proving the following result.

PROPOSITION 1.15. Any geodesic is a horizontal curve.

PROOF. Let $c(s) = \big(x_1(s), x_2(s), t(s)\big)$ be a geodesic. From the Hamiltonian system (1.25)

$$\dot{x}_1 = \xi_1 + 2x_2\theta, \qquad \dot{x}_2 = \xi_2 - 2x_1\theta,$$

and then

$$\begin{aligned} \dot{t} &= \frac{\partial H}{\partial \theta} \\ &= 2x_2(\xi_1 + 2x_2\theta) - 2x_1(\xi_2 - 2x_1\theta) \\ &= 2x_2\dot{x}_1 - 2x_1\dot{x}_2, \end{aligned}$$

which is the horizontality condition (1.14). Hence, any geodesic is a horizontal curve. $\square$

Solving the Hamiltonian system. We shall solve the Hamiltonian system explicitly. We start with the observation that H does not depend on t. Then

$$\dot{\theta} = -\frac{\partial H}{\partial t} = 0$$

and hence, $\theta =$ constant. The equations

$$\dot{x}_1 = \frac{\partial H}{\partial \xi_1}, \qquad \dot{x}_2 = \frac{\partial H}{\partial \xi_2}$$

become

$$\dot{x}_1 = \xi_1 + 2x_2\theta, \qquad \dot{x}_2 = \xi_2 - 2x_1\theta. \tag{1.27}$$

Differentiating, yields

$$\ddot{x}_1 = \dot{\xi}_1 + 2\dot{x}_2\theta, \qquad \ddot{x}_2 = \dot{\xi}_2 - 2\dot{x}_1\theta. \tag{1.28}$$

Using $\dot{\xi} = -\partial H/\partial x$ and the system (1.27) we obtain

$$\dot{\xi}_1 \;=\; 2\theta(\xi_2 - 2x_1\theta) = 2\theta\dot{x}_2 \tag{1.29}$$

$$\dot{\xi}_2 \;=\; -2\theta(\xi_1 + 2x_2\theta) = -2\theta\dot{x}_1. \tag{1.30}$$

From systems (1.28) and (1.29)

$$\ddot{x}_1 = 4\theta\dot{x}_2, \qquad \ddot{x}_2 = -4\theta\dot{x}_1 \tag{1.31}$$

with constant θ. The system (1.31) can be written as

$$\ddot{x}(s) = 4\theta J\dot{x}(s), \tag{1.32}$$

where

$$J = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

and $x = \begin{pmatrix} x_1 & x_1 \end{pmatrix}$. The equation (1.32) describes the projection of the geodesic on the x-space. We shall show that this is a circle.

With the substitution $y(s) = \dot{x}(s)$ equation (1.32) becomes

$$\dot{y}(s) = 4\theta Jy(s), \tag{1.33}$$

with the solution

$$y(s) = e^{4\theta Js}y(0).$$

Therefore $\dot{x}(s) = e^{4\theta Js}y(0)$. Integrating and using that J and $e^{4\theta Js}$ commute, yields

$$\begin{aligned} x(s) &= x(0) + \int_0^s e^{4\theta Ju}y(0)\,du \\ &= x(0) + \frac{1}{4\theta}J^{-1}\, e^{4\theta Js}y(0)\Big|_{u=0}^{u=s} \\ &= x(0) - \frac{1}{4\theta}Je^{4\theta Js}y(0) + \frac{1}{4\theta}\mathcal{J}^{-1}y(0) \\ &= e^{4\theta Js}K + C, \end{aligned} \tag{1.34}$$

where $K = -Jy(0)/(4\theta)$ and $C = x(0) + K$. We need the following result.

LEMMA 1.16. *If* $J = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ *then* $e^{4\theta Js} = R_{4\theta s}$, *where* R_α *denotes the rotation by angle* α *in the* x*-plane.*

PROOF. A computation yields

$$\begin{aligned} e^{4\theta Js} &= \sum_{n=0}^{\infty} \frac{(4\theta s)^n J^n}{n!} = I \sum_{k=0}^{\infty} \frac{(4\theta s)^{4k}}{(4k)!} + J \sum_{k=0}^{\infty} \frac{(4\theta s)^{4k+1}}{(4k+1)!} \\ &\quad - I \sum_{k=0}^{\infty} \frac{(4\theta s)^{4k+2}}{(4k+2)!} - J \sum_{k=0}^{\infty} \frac{(4\theta s)^{4k+3}}{(4k+3)!} \\ &= \sum_{k=0}^{\infty} \begin{pmatrix} \frac{(4\theta s)^{4k}}{(4k)!} - \frac{(4\theta s)^{4k+2}}{(4k+2)!} & \frac{(4\theta s)^{4k+1}}{(4k+1)!} - \frac{(4\theta s)^{4k+3}}{(4k+3)!} \\ -\frac{(4\theta s)^{4k+1}}{(4k+1)!} + \frac{(4\theta s)^{4k+3}}{(4k+3)!} & \frac{(4\theta s)^{4k}}{(4k)!} - \frac{(4\theta s)^{4k+2}}{(4k+2)!} \end{pmatrix} \\ &= \begin{pmatrix} \cos(4\theta s) & \sin(4\theta s) \\ -\sin(4\theta s) & \cos(4\theta s) \end{pmatrix} \\ &= R_{4\theta s}. \end{aligned}$$

□

Using Lemma 1.4, the equation (1.34) becomes

$$x(s) = R_{4\theta s} K + C. \tag{1.35}$$

As $|x(s) - C| = |R_{4\theta s} K| = |K| = \text{constant}$, $x(s)$ will describe a circle centered at C with the radius

$$|K| = \left| \frac{-Jy(0)}{4\theta} \right| = \frac{|y(0)|}{4|\theta|} = \frac{|\dot{x}(0))|}{4|\theta|}. \tag{1.36}$$

PROPOSITION 1.17. Consider a geodesic which joins the points $P(x_0, t_0)$ and $Q(x_1, t_1)$, with $t_0 \neq t_1$.
(i) The projection of the geodesic on the x-space is a circle or a piece of a circle with end points x_0 and x_1.
(ii) If the projection is one complete circle, with $x_0 = x_1$, let σ be its area. Then

$$\sigma = \frac{|t_1 - t_0|}{4}.$$

PROOF. (i) comes from the solution of the Hamiltonian system discussed above. (ii) By Proposition 1.15, any geodesic is horizontal. From equation (1.18), the area of the projection on the x-plane and the t-component of a horizontal curve are related by

$$4\, d\mathcal{A} = -dt.$$

Integrating, yields

$$\begin{aligned} 4\sigma &= \int_0^\tau 4\, d\mathcal{A} = -\int_0^\tau dt \\ &= t_0 - t_1, \end{aligned}$$

where $x(\tau) = x_1$, $t(\tau) = t_1$. □

COROLLARY 1.18. The radius of the projection circle is $R = \sqrt{\frac{|t_1 - t_0|}{4\pi}}$.

If we let $\theta \to 0$ in (1.36), the radius of the circle $|K| \to \infty$. This corresponds to a circle with infinite radius, which is a line. This case is covered by the next result.

PROPOSITION 1.19. If $\theta = 0$, the projection of the geodesic onto the x-space is a line.

PROOF. If $\theta = 0$ the equation (1.32) yields $\ddot{x} = 0$, which corresponds to a line. □

Proposition 1.15 claims that geodesics are horizontal curves. We shall show that the converse is false.

PROPOSITION 1.20. There are horizontal curves which are not geodesics.

PROOF. Consider $c(s) = \big(s^2/2, s, s^3/3\big)$. The curve is horizontal, because the horizontality condition holds

$$\dot{t}(s) = s^2 = 2\big(s^2 - s^2/2\big) = 2\big(\dot{x}_1 x_2 - x_1 \dot{x}_2\big).$$

On the other hand, the system (1.31) becomes $4\theta = 1$ and $0 = -4\theta s$, which leads to a contradiction. □

The t-component

Using the Hamiltonian equation $\dot{t} = \partial H/\partial\theta$, and (1.36) yields

$$\begin{aligned}
\dot{t}(s) &= 2\big(x_2(s)\dot{x}_1(s) - x_1(s)\dot{x}_2(s)\big) \\
&= 2\langle x(s), J(\dot{x}(s))\rangle \\
&= 2\langle e^{4\theta J s}K + C,\ J(4\theta J e^{4\theta J s}K)\rangle \\
&= 2\langle e^{4\theta J s}K,\ -4\theta e^{4\theta J s}K\rangle + 2\langle C,\ -4\theta e^{4\theta J s}K\rangle \\
&= -8\theta|K|^2 - 8\theta\langle C, e^{4\theta J s}K\rangle.
\end{aligned}$$

Integrating we obtain

$$t(s) = \int \Big(-8\theta|K|^2 - 8\theta\langle C, e^{4\theta J s}K\rangle\Big)\, ds.$$

As

$$\begin{aligned}
\frac{d}{ds}\langle JC, e^{4\theta J s}K\rangle &= \langle JC, \frac{d}{ds}e^{4\theta J s}K\rangle = \langle JC, 4\theta J e^{4\theta J s}K\rangle \\
&= 4\theta\langle J^T JC, e^{4\theta J s}K\rangle = 4\theta\langle C, e^{4\theta J s}K\rangle
\end{aligned}$$

then

$$\int \langle C, e^{4\theta J s}K\rangle\, ds = \frac{1}{4\theta}\langle JC, e^{4\theta J s}K\rangle + const$$

Hence

$$t(s) = -8\theta|K|^2 s - 2\langle JC, e^{4\theta J s}K\rangle + C_1$$

where $C_1 = t(0) + 2\langle JC, K\rangle$.

The conservation of energy. Let $E = \frac{1}{2}\big(\dot{x}_1^2 + \dot{x}_2^2\big)$ be the kinetic energy. One may show that the Hamiltonian is equal to E along the geodesics, and hence E is a first integral for the Hamiltonian system. In the following proposition we shall give a direct proof.

PROPOSITION 1.21. The kinetic energy is preserved along the geodesics.

PROOF. Using equation (1.32)

$$\begin{aligned} \frac{dE}{ds} &= \frac{d}{ds}\frac{\dot{x}_1^2 + \dot{x}_2^2}{2} = \dot{x}_1\ddot{x}_1 + \dot{x}_2\ddot{x}_2 \\ &= \langle \dot{x}, \ddot{x}\rangle = 4\theta\, \langle \dot{x}, \mathcal{J}\dot{x}\rangle = 0. \end{aligned}$$

□

In the following we shall show that the projection of the solution is a circle, using the conservation of energy. Let $2E = R^2$ with $R > 0$ constant. For conservative systems with $\dot{x}_1(s)^2 + \dot{x}_2(s)^2 = R^2$, there is a function $\alpha = \alpha(s)$ such that

$$\dot{x}_1(s) = R\cos\alpha(s), \qquad \dot{x}_2(s) = R\sin\alpha(s). \tag{1.37}$$

Then

$$\ddot{x}_1(s) = -R\sin\alpha(s)\,\dot{\alpha}(s), \qquad \ddot{x}_2(s) = R\cos\alpha(s)\,\dot{\alpha}(s),$$

and the system (1.31)

$$\begin{aligned} \ddot{x}_1 &= 4\theta\dot{x}_2 \\ \ddot{x}_2 &= -4\theta\dot{x}_1 \end{aligned}$$

becomes

$$\begin{aligned} \sin\alpha(s)\big(4\theta + \dot{\alpha}\big) &= 0 \\ \cos\alpha(s)\big(4\theta + \dot{\alpha}\big) &= 0. \end{aligned}$$

Adding the squares, yields $\dot{\alpha}(s) = 4\theta$. Integrating, we obtain

$$\alpha(s) = -4\theta s - \alpha_0.$$

Hence equations (1.37) can be integrated

$$\begin{aligned} x_1(s) &= R\int_0^s \cos(-4\theta s - \alpha_0)\, ds \\ &= \frac{R}{4\theta}\sin(4\theta s + \alpha_0) + x_1(0). \\ x_2(s) &= R\int_0^s \sin(-4\theta s - \alpha_0)\, ds \\ &= \frac{R}{4\theta}\cos(4\theta s + \alpha_0) + x_2(0), \end{aligned}$$

where

$$x_1(0) = -\frac{R\sin\alpha_0}{4\theta}, \qquad x_2(0) = -\frac{R\cos\alpha_0}{4\theta}.$$

We obtain the solution

$$x(s) = \frac{R}{4\theta}\Big(\sin(4\theta s + \alpha_0), \cos(4\theta s + \alpha_0)\Big) + x(0),$$

which is the parametric equation of a circle of radius $\frac{R}{4\theta}$ centered at $x(0)$.

1.5. The connection form

Let $x \to \mathcal{H}_x = \text{span}_x\{X_1, X_2\}$ be the horizontal distribution on $\mathbb{R}^3$. A connection 1-form is a non-vanishing form $\omega \in T^*\mathbb{R}^3$ such that $\ker_x \omega = \mathcal{H}_x$. The form ω is unique up to a multiplicative factor. In this chapter we shall choose the *standard* 1-form with the property $\omega(\partial_t) = 1$, which is

$$\omega = dt - 2(x_2 dx_1 - x_1 dx_2). \tag{1.38}$$

DEFINITION 1.22. The curvature 2-form of the distribution $\mathcal{H}$ is defined as $\Omega : \mathcal{H} \times \mathcal{H} \to \mathcal{F}(\mathcal{R}^3)$

$$\Omega(U, V) = d\omega(U, V). \tag{1.39}$$

In our case $\Omega = 4dx_1 \wedge dx_2$. If the horizontal distribution $\mathcal{H}$ belongs to the intrinsic subRiemannian geometry, the form Ω describes the extrinsic geometry of the Heisenberg group. In general, the 2-form Ω describes the non-integrability of the horizontal distribution.

DEFINITION 1.23. The pair $(\mathbb{R}^3, \omega)$ is called a contact manifold if $\omega \wedge \Omega$ never vanishes.

In our case $\omega \wedge \Omega = 4dt \wedge dx_1 \wedge dx_2$ and hence, the Heisenberg group becomes a contact manifold. The following theorem shows that, locally, all contact manifolds are the same as the Heisenberg group, see Cartan [**21**].

THEOREM 1.24. *(Darboux) Each point p of a contact manifold admits a local coordinate system t, x_1, x_2 in a neighborhood U such that $\omega = dt - 2(x_2 dx_1 - x_1 dx_2)$.*

1.5.0.1. *The osculator plane.* Let $c(s) = \big(x_1(s), x_2(s), t(s)\big)$ be a curve. The *osculator* plane at $c(s)$ is defined by $\text{span}\{\dot{c}(s), \ddot{c}(s)\}$.

PROPOSITION 1.25. Let $c(s)$ be a curve. Then the curve $c(s)$ is horizontal if and only if the osculator plane at $c(s)$ is the horizontal plane $\mathcal{H}_{c(s)}$, for any s.

PROOF. If $\text{span}\{\dot{c}(s), \ddot{c}(s)\} = \mathcal{H}_{c(s)}$, then $\dot{c}(s) \in \mathcal{H}_{c(s)}$. Hence, the curve is horizontal.
If the curve $c(s)$ is horizontal, $\dot{c}(s) \in \mathcal{H}_{c(s)}$. It suffices to show $\ddot{c}(s) \in \mathcal{H}_{c(s)}$. From the horizontality condition (1.14),

$$\dot{t} = 2x_2\dot{x}_1 - 2x_1\dot{x}_2. \tag{1.40}$$

Differentiating in equation (1.40) yields

$$\begin{aligned} \ddot{t} &= 2\dot{x}_2\dot{x}_1 + 2x_2\ddot{x}_1 - 2\dot{x}_1\dot{x}_2 - 2x_1\ddot{x}_2 \\ &= 2x_2\ddot{x}_1 - 2x_1\ddot{x}_2. \end{aligned} \tag{1.41}$$

The acceleration vector along $c(s)$ is

$$\begin{aligned} \ddot{c} &= \ddot{x}_1\partial_{x_1} + \ddot{x}_2\partial_{x_2} + \ddot{t}\partial_t \\ &= \ddot{x}_1\big(\partial_{x_1} + 2x_2\partial_t\big) - 2x_2\ddot{x}_1\partial_t \\ &\quad + \ddot{x}_2\big(\partial_{x_2} - 2x_1\partial_t\big) + 2x_1\ddot{x}_2\partial_t + \ddot{t}\partial_t \\ &= \ddot{x}_1 X_1 + \ddot{x}_2 X_2 + \big(\ddot{t} - 2x_2\ddot{x}_1 + 2\ddot{x}_2 x_1\big)\partial_t \\ &= \ddot{x}_1 X_1 + \ddot{x}_2 X_2, \end{aligned}$$

since we used equation (1.41). Hence, $\ddot{c} \in \mathcal{H}_c$ and the osculator plane is horizontal. □

COROLLARY 1.26. For a horizontal curve c

$$\ddot{c} = \ddot{x}_1 X_1 + \ddot{x}_2 X_2, \tag{1.42}$$

$$\ddot{t} = 2x_2\ddot{x}_1 - 2\ddot{x}_2 x_1. \tag{1.43}$$

DEFINITION 1.27. Let $J : \mathcal{H} \to \mathcal{H}$ be defined by $J(X_1) = -X_2$, $J(X_2) = X_1$. J is called the complex structure of the horizontal plane.

We shall use J in order to write the equations for the geodesics on the Heisenberg group. The following result shows that the geodesics satisfy a Newton type equation. The left side is the acceleration, while the right side is the force, which keeps the distribution bent. As before, θ is a constant.

PROPOSITION 1.28. A curve c is a geodesic on the Heisenberg group if and only if
(i) c is a horizontal curve and
(ii) c satisfies

$$\ddot{c} = 4\theta J\dot{c}. \tag{1.44}$$

PROOF. If $c(s)$ is a geodesic, by Proposition 1.15, $c(s)$ is horizontal. Using Corollary 1.26 and the system (1.31),

$$\begin{aligned} \ddot{c} &= \ddot{x}_1 X_1 + \ddot{x}_2 X_2 \\ &= 4\theta\dot{x}_2 X_1 - 4\theta\dot{x}_1 X_2 \\ &= 4\theta\dot{x}_2 J(X_2) + 4\theta\dot{x}_1 J(X_1) \\ &= 4\theta J\big(\dot{x}_1 X_1 + \dot{x}_2 X_2\big) \\ &= 4\theta J(\dot{c}). \end{aligned}$$

We shall now prove the converse: if (i) and (ii) hold, then c is a geodesic. We shall use the definition 1.14. The horizontality condition (i) can be written as $\dot{t} = \partial H/\partial\theta$, which is the Hamiltonian equation for t. Using a similar computation as in the first part, equation (1.44) written in components becomes the system (1.31). Let $x_1(s)$ and $x_2(s)$ be solutions of this system. Define the following curve in the cotangent space

$$\gamma(s) = \big(x_1(s), x_2(s), t(s), \xi_1(s), \xi_2(s), \theta\big),$$

where

$$\xi_1 = \dot{x}_1 - 2x_2(s)\theta, \quad \xi_2 = \dot{x}_2 + 2x_1\theta,$$

with θ constant. Then $\gamma(s)$ satisfies the bicharacteristics system (1.25) for the Hamiltonian (1.23). Then the projection on the (x,t)-space is a geodesic and hence $c(s)$ is a geodesic. □

The subRiemannian metric.

DEFINITION 1.29. A non-degenerate, positive definite bilinear form $g_x : \mathcal{H}_x \times \mathcal{H}_x \to \mathcal{F}(\mathbb{R})$ at any point $x \in \mathbb{R}^3$, is called a subRiemannian metric.

We shall consider the subRiemannian metric in which the vector fields X_1, X_2 are orthonormal. The subRiemannian metric will become a Kaehler metric on the horizontal distribution, as we shall explain later. The following definitions can be found, for instance, in Kobayashy and Nomizu, see [**37**].

DEFINITION 1.30. A Hermitian metric on a real vector space V with a complex structure J is a non-degenerate, positive definite inner product h such that

$$h(JX, JY) = h(X, Y), \qquad \text{for } X, Y \in V.$$

We associate with each Hermitian metric of a vector space V a skew-symmetric bilinear form on V.

DEFINITION 1.31. The fundamental 2-form Φ is defined by

$$\Phi(X, Y) = h(X, JY), \quad \text{for all vector fields } X \text{ and } Y.$$

A Hermitian metric on a vector space V with a complex structure J is called a Kaehler metric if its fundamental 2-form is closed.

The relationship with the subRiemannian metric is given in the following proposition.

PROPOSITION 1.32. The subRiemannian metric g in which $\{X_1, X_2\}$ are orthonormal is a Kaehler metric on $\mathcal{H}_x$, for any $x \in \mathbb{R}^3$. The fundamental 2-form satisfies $4\Phi = \Omega$. Hence

$$\Omega(U, V) = 4g(U, JV), \quad \text{for all horizontal vectors } U \text{ and } V.$$

PROOF. Consider $V = \mathcal{H}_x$. We shall show first that g is a Hermitian metric. Let $U = U^1 X_1 + U^2 X_2$ and $V = V^1 X_1 + V^2 X_2$ be two horizontal vector fields. Using $JX_1 = -X_2$ and $JX_2 = X_1$, yields

$$JU = -U^1 X_2 + U^2 X_1,$$

$$JV = -V^1 X_2 + V^2 X_1.$$

Using the orthonormality of X_1 and X_2 we obtain

$$\begin{aligned} g(JU, JV) &= g(U^2 X_1 - U^1 X_2, V^2 X_1 - V^1 X_2) \\ &= U^1 V^1 + U^2 V^2 \\ &= g(U^1 X_1 + U^2 X_2, V^1 X_1 + V^2 X_2) \\ &= g(U, V), \end{aligned}$$

and hence g is invariant by J. The 2-form Ω is closed because it is exact $\Omega = d\omega$.

$$\begin{aligned}
\Omega(U,V) &= \Omega(U^1X_1 + U^2X_2, V^1X_1 + V^2X_2) \\
&= (U^1V^2 - U^2V^1)\Omega(X_1, X_2) \\
&= (U^1V^2 - U^2V^1)\big(X_1(\omega(X_2)) - X_2(\omega(X_1)) - \omega([X_1, X_2])\big) \\
&= 4(U^1V^2 - U^2V^1)\omega(\partial_t) \\
&= 4(U^1V^2 - U^2V^1) \\
&= 4g(U^1X_1 + U^2X_2, V^2X_1 - V^1X_2) \\
&= 4g(U, JV).
\end{aligned}$$

□

Using the skew-symmetry of Ω we obtain:

COROLLARY 1.33.

$$g(U, JU) = 0, \qquad \text{for any horizontal vector } U.$$

The geometrical interpretation of Ω is given below.

PROPOSITION 1.34. Let $\pi : \mathbb{R}^3_{(x,t)} \to \mathbb{R}^2_x$ be the projection which sends the horizontal plane onto the x-plane

$$\pi_*(X_1) = \partial_{x_1}, \qquad \pi_*(X_2) = \partial_{x_2}.$$

Then, for any horizontal vectors U and V, the area of the parallelogram generated by $\pi_*(U)$ and $\pi_*(V)$ is equal to $|\Omega(U,V)|/4$.

PROOF. We have

$$\begin{aligned}
\Omega(U,V) &= 4\big(dx_1 \wedge dx_2\big)(U,V) = 4\begin{vmatrix} dx_1(U) & dx_1(V) \\ dx_2(U) & dx_2(V) \end{vmatrix} \\
&= 4\begin{vmatrix} U(x_1) & V(x_1) \\ U(x_2) & V(x_2) \end{vmatrix} = 4\begin{vmatrix} U^1 & V^1 \\ U^2 & V^2 \end{vmatrix}.
\end{aligned}$$

Using the interpretation of the determinant as an area, the proof is complete. □

PROPOSITION 1.35. Let $c(s)$ be a geodesic curve. Then

$$\theta\Omega(U, \dot c) = g(U, \ddot c) \qquad \text{for any horizontal vector } U.$$

PROOF. Using the Kaehler property of the metric g and the geodesics equation (1.44)

$$\theta\Omega(U, \dot c) = 4\theta g(U, J\dot c) = g(U, \ddot c).$$

□

If $\theta = 0$, then $g(U, \ddot c) = 0$ for any horizontal vector U and then $\ddot c = 0$. Hence $\ddot c_1(s) = \ddot c_2(s) = 0$. We obtain the following result:

COROLLARY 1.36. Let $c(s)$ be a geodesic for which the momentum θ vanishes. Then

$$c(s) = (as + a_0, bs + b_0, t_0),$$

with a, a_0, b, b_0 and t_0 constants.

PROOF. The fact that $c_1(s) = as + a_0$ and $c_2(s) = bs + b_0$ is obvious. We shall show that $t(s)$ is constant. We have

$$\begin{aligned} \dot{t}(s) &= 2(x_2\dot{x}_1 - x_1\dot{x}_2) \\ &= 2(c_2\dot{c}_1 - c_1\dot{c}_2) \\ &= 2\Big((bs+b_0)a - (as+a_0)b\Big) \\ &= 0, \end{aligned}$$

i.e., t is constant. □

The following proposition deals with the metric properties of the velocity and acceleration.

PROPOSITION 1.37. Let $c(s)$ be a geodesic. Then
(i) The velocity $\dot{c}$ and the acceleration $\ddot{c}$ vector fields are perpendicular in the subRiemannian metric.
(ii) The magnitude of $\dot{c}$ in the subRiemannian metric is constant along the geodesic.
(iii) The magnitude of $\ddot{c}$ in the subRiemannian metric is constant along the geodesic.

PROOF. (i) For $U = \dot{c}$, Proposition 1.35 yields

$$g(\dot{c}, \ddot{c}) = \theta\Omega(\dot{c}, \dot{c}) = 0.$$

(ii) Differentiating and using (i) yields

$$\frac{d}{ds} g(\dot{c}, \dot{c}) = g(\ddot{c}, \dot{c}) + g(\dot{c}, \ddot{c}) = 0,$$

i.e. $g(\dot{c}(s), \dot{c}(s))$ is constant.
(iii) The vector field $\ddot{c}$ is horizontal, and using equation (1.44) and (ii) we have

$$g(\ddot{c}, \ddot{c}) = g(4\theta J\dot{c}, 4\theta J\dot{c}) = 16\theta^2 g(J\dot{c}, J\dot{c}) = 16\theta^2 g(\dot{c}, \dot{c}) = \text{ constant}.$$

□

In the classical theory of three dimensional curves, the curvature becomes along a curve $c(s)$ is a function defined by $k(s) = |\dot{T}(s)|$, where $T(s) = \dot{c}(s)/|\dot{c}(s)|$ is the unit tangent vector. If s is the arc length parameter, $|\dot{c}(s)| = 1$ and the curvature becomes $k(s) = |\ddot{c}(s)|$.

PROPOSITION 1.38. The curvature of a geodesic curve is constant, $k(s) = 4|\theta|$.

PROOF. Consider the geodesic $c(s)$ parametrized by the arc length. Then

$$k(s)^2 = |\ddot{c}|^2 = g(\ddot{c}, \ddot{c}) = 16\theta^2 g(\dot{c}, \dot{c}) = 16\theta^2.$$

□

The curvature of a geodesic depends on the momentum θ. As we shall show later, θ is a Lagrange multiplier which describes the number of rotations of the geodesics around the t-axis. Hence, we shall be able to prove a Gauss-Bonett type theorem for the geodesic curves.

The following proposition provides the complex structure of the horizontal distribution in function of Ω and the basic horizontal vector fields.

PROPOSITION 1.39. For any horizontal vector field U, we have

$$J(U) = \frac{1}{4}\Big(\Omega(X_1, U)X_1 + \Omega(X_2, U)X_2\Big).$$

PROOF. As both sides are linear, it suffices to check the relation only for the basic vector fields X_1 and X_2. Using $\Omega(X_1, X_2) = -\Omega(X_2, X_1) = 4$, and $\Omega(X_i, X_i) = 0$, yields

$$JX_1 = -X_2 = \frac{1}{4}\Omega(X_2, X_1)X_2 = \frac{1}{4}\Big(\Omega(X_1, X_1)X_1 + \Omega(X_2, X_1)X_2\Big)$$
$$JX_2 = X_1 = \frac{1}{4}\Omega(X_1, X_2)X_1 = \frac{1}{4}\Big(\Omega(X_1, X_2)X_1 + \Omega(X_2, X_2)X_2\Big).$$

□

1.6. Lagrangian formalism on the Heisenberg group

In this section we shall find the subRiemannian geodesics and characterize their lengths. The horizontality condition is a constraint on velocities, *i.e.*, they are *non-holonomic*. The Lagrangian which describes the geodesics has a non-holonomic constraint. This constraint can be expressed using the 1-form ω.

PROPOSITION 1.40. If $\phi(s)$ is a horizontal curve, then

$$\int_\phi \omega = 0. \tag{1.45}$$

PROOF. As $\phi^*\omega$ is a 1-form on $\mathbb{R}$, then $\phi^*\omega$ and ds are proportional

$$\phi^*\omega(s) = h(s)\, ds,$$

where the proportionality function $h(s)$ is

$$h(s) = \phi^*_{(s)}\omega\Big(\frac{d}{ds}\Big).$$

Let $\phi(s)$ be defined on $[0, 1]$. Then

$$\begin{aligned}\int_\phi \omega &= \int_0^1 \phi^*\omega = \int_0^1 h(s)\, ds = \int_0^1 \phi^*_{(s)}(\omega)\Big(\frac{d}{ds}\Big)\, ds \\ &= \int_0^1 \omega\Big(\phi_*\Big(\frac{d}{ds}\Big)\Big)\, ds = \int_0^1 \omega\big(\dot\phi(s)\big) = 0,\end{aligned}$$

because $\dot\phi \in \mathcal{H}$. □

We shall associate a Lagrangian $L : T\mathbb{R}^3 \to \mathbb{R}$ with the Hamiltonian (1.23). This will be done using a formal Legendre transform in $(\dot x, \dot t)$, see [**14**]. In the case of a quadratic Hamiltonian, the Lagrangian is given by the maximal distance between the hyperplane $< \xi, \dot x > + \theta \dot t$ and the convex surface given by the Hamiltonian in $\mathbb{R}^6$:

$$\begin{aligned}L(x, t, \dot x, \dot t) &= \max_{\xi,\theta}\Big(\xi_1\dot x_1 + \xi_2\dot x_2 + \theta\dot t - H(\xi, \theta, x, t)\Big) \\ &= \max_{\xi,\theta} F(x, t, \dot x, \dot t, \xi, \theta).\end{aligned}$$

In our case the Hamiltonian is degenerate, so that the equivalence between the Lagrangian and the Hamiltonian formalism of Classical Mechanics might not hold. We shall still proceed formally, and set

$$\frac{\partial F}{\partial \xi} = 0, \quad \frac{\partial F}{\partial \theta} = 0 \tag{1.46}$$

as in Classical Mechanics, where the partial derivatives vanish when the maximum is reached.

The equations (1.46) can be written as

$$\dot{x}_i = \frac{\partial H}{\partial \xi_i}, \quad \dot{t} = \frac{\partial H}{\partial \theta} = 0, \tag{1.47}$$

$$\dot{x}_1 = \xi_1 + 2x_2\theta, \quad \dot{x}_2 = \xi_2 - 2x_1\theta, \quad \dot{t} = 2x_2\dot{x}_1 - 2x_1\dot{x}_2. \tag{1.48}$$

Using relations (1.48), we formally define the Lagrangian by

$$\begin{aligned} L(x,t,\dot{x},\dot{t}) &= \xi_1\dot{x}_1 + \xi_2\dot{x}_2 + \theta\dot{t} - \frac{1}{2}\big(\xi_1 + 2x_2\theta\big)^2 - \frac{1}{2}\big(\xi_2 - 2x_1\theta\big)^2 \\ &= (\dot{x}_1 - 2x_2\theta)\dot{x}_1 + (\dot{x}_2 + 2x_1\theta)\dot{x}_2 + \theta\dot{t} - \frac{1}{2}(\dot{x}_1^2 + \dot{x}_2^2) \\ &= (\dot{x}_1^2 + \dot{x}_2^2) + \theta(\dot{t} - 2x_2\dot{x}_1 + 2x_1\dot{x}_2) - \frac{1}{2}(\dot{x}_1^2 + \dot{x}_2^2) \\ &= \frac{1}{2}(\dot{x}_1^2 + \dot{x}_2^2) + \theta(\dot{t} - 2x_2\dot{x}_1 + 2x_1\dot{x}_2). \end{aligned}$$

Using the 1-connection form ω, we write the Lagrangian as

$$L(c,\dot{c}) = \frac{1}{2}(\dot{x}_1^2 + \dot{x}_2^2) + \theta\omega(\dot{c}), \tag{1.49}$$

where $c = (x_1, x_2, t)$. Let the action integral be

$$S(c,\tau) = \int_0^\tau L(c,\dot{c})\,ds = \int_0^\tau \frac{1}{2}(\dot{x}_1^2 + \dot{x}_2^2)\,ds + \theta\int_c \omega. \tag{1.50}$$

The action has two parts: the kinetic energy and the non-holonomic constraint. In this case the constant θ is a Lagrange multiplier. The curves c, which are critical points for the action $S(c,\tau)$, satisfy the Euler-Lagrange equation

$$\frac{d}{ds}\Big(\frac{\partial L}{\partial \dot{c}}\Big) = \frac{\partial L}{\partial c}. \tag{1.51}$$

The partial equivalence between the Lagrangian and the Hamiltonian formalism is given in the following result.

Proposition 1.41. A solution of the Euler-Lagrange equation (1.51) is a geodesic if and only if it is a horizontal curve.

Proof. As any geodesic is a horizontal curve, it suffices to show that a horizontal solution of (1.51) is a geodesic. If $c = (x_1, x_2, t)$, the equation (1.51) becomes on components

$$\ddot{x}_1 = 4\theta\dot{x}_2, \quad \ddot{x}_2 = -4\theta\dot{x}_1, \quad \dot{\theta} = 0 \Longrightarrow \theta = \text{constant}.$$

Proposition 1.28 completes the proof. □

REMARK 1.42. The Euler-Lagrange equation (1.51) does not imply horizontality of solutions, which corresponds to the Hamiltonian equation $\dot{t} = \partial H/\partial\theta$. Hence the Hamiltonian and the Euler-Lagrange systems are not equivalent. This behavior is specific to subRiemannian geometry and has no analog in the Riemannian case.

Lagrangian symmetries. The symmetry of the Lagrangian influences the symmetry of the solutions of the Euler-Lagrange system and hence the geodesics. The study of the Lagrangian symmetries makes easier the understanding of geodesic behavior.

PROPOSITION 1.43. The Lagrangian (1.49) is left invariant with respect to the Heisenberg Lie group structure, *i.e.,* $L(\overline{c}, \dot{\overline{c}}) = L(c, \dot{c})$, where $\overline{c} = L_a(c)$, for any $a \in \mathbf{H}_1$.

PROOF. If $c = (c_1, c_2, c_3)$, $\overline{c} = (\overline{c}_1, \overline{c}_2, \overline{c}_3)$ and $a = (a_1, a_2, a_3)$, then

$$\begin{aligned}
\overline{c}_1 &= a_1 + c_1 \Longrightarrow \dot{\overline{c}}_1 = \dot{c}_1 \\
\overline{c}_2 &= a_2 + c_2 \Longrightarrow \dot{\overline{c}}_2 = \dot{c}_2 \\
\overline{c}_3 &= a_3 + c_3 - 2(a_1 c_2 - a_2 c_1) \Longrightarrow \dot{\overline{c}}_3 = \dot{c}_3 - 2(a_1 \dot{c}_2 - a_2 \dot{c}_1)
\end{aligned}$$

Then the kinetic energy is left invariant

$$\frac{1}{2}(\dot{c}_1^2 + \dot{c}_2^2) = \frac{1}{2}(\dot{\overline{c}}_1^{\,2} + \dot{\overline{c}}_2^{\,2}).$$

The horizontal constraint is also invariant

$$\begin{aligned}
\dot{\overline{c}}_3 - 2\overline{c}_2\dot{\overline{c}}_1 + 2\overline{c}_1\dot{\overline{c}}_2 &= \dot{c}_3 - 2(a_1\dot{c}_2 - a_2\dot{c}_1) - 2(a_2 + c_2)\dot{c}_1 + 2(a_1 + c_1)\dot{c}_2 \\
&= \dot{c}_3 - 2c_2\dot{c}_1 + 2c_1\dot{c}_2.
\end{aligned}$$

Hence, the Lagrangian is preserved by left translations. □

COROLLARY 1.44. The solutions of the Euler-Lagrange equation (1.51) are left invariant by the translations on $\mathbf{H}_1$.

In the following we shall use Noether's theorem approach to find first integrals of motion for the Lagrangian (1.49)

$$L(x, t, \dot{x}, \dot{t}) = \frac{1}{2}(\dot{x}_1^2 + \dot{x}_2^2) + \theta(\dot{t} - 2x_2\dot{x}_1 + 2x_1\dot{x}_2).$$

The following theorem can be found in Arnold [**2**]. For its generalizations on manifolds one may consult [**14**].

THEOREM 1.45. *(Noether) To every one-parameter group of diffeomorphisms $(h_s)_s$ of the coordinate space M of a Lagrangian system which preserves the Lagrangian, corresponds a first integral of the Euler-Lagrange equations of motion $I : TM \to \mathbb{R}$ given by*

$$I(q, \dot{q}) = \Big\langle \frac{\partial L}{\partial \dot{q}}, \frac{dh_s(q)}{ds}\Big|_{s=0} \Big\rangle, \tag{1.52}$$

where $\langle , \rangle$ denotes the standard inner product on $\mathbb{R}^3$.

In the case of the Heisenberg group, $M = \mathbb{R}^3$, $q = (x,t)$ and

$$\frac{\partial L}{\partial \dot q} = \Big(\frac{\partial L}{\partial \dot x_1}, \frac{\partial L}{\partial \dot x_2}, \frac{\partial L}{\partial \dot t}\Big) = \Big(\dot x_1 - 2\theta x_2, \dot x_2 + 2\theta x_1, \theta\Big).$$

Consider the following three independent one-parameter groups of diffeomorphisms

$$h_s(x,t) = L_{a(s)}(x,t) = (a_1(s) + x_1, a_2(s) + x_2, a_3(s) + t - 2(a_1(s)x_2 - a_2(s)x_1)),$$

with $a(s) \in \{(0,0,s),(0,s,0),(s,0,0)\}$. The associated vector field is

$$\frac{dh_s(q)}{ds}\Big|_{s=0} = \begin{cases} (0,0,1), & \text{for } a(s) = (0,0,s) \\ (0,1,2x_1), & \text{for } a(s) = (0,s,0) \\ (1,0,-2x_2), & \text{for } a(s) = (s,0,0). \end{cases}$$

Hence, formula (1.52) provides the following three functional independent first integrals

$$\begin{aligned} I_1 &= \theta = \text{constant}, \\ I_2 &= \dot x_2 + 4\theta x_1 = \text{constant}, \\ I_3 &= \dot x_1 - 4\theta x_2 = \text{constant}. \end{aligned}$$

Differentiating, we obtain the Euler-Lagrange system

$$\begin{cases} \ddot x_1 = 4\theta \dot x_2 \\ \ddot x_2 = -4\theta \dot x_1 \\ \theta = \text{constant}, \end{cases}$$

which is equivalent to (1.51).

The rotational symmetry of the Lagrangian will provide a non obvious first integral. The Lagrangian is invariant by the one-parameter group of rotations $h_s(x,t) = (R_s x, t)$. Using $R_s(x) = e^{Js}x$, we have $\frac{dR_s}{ds} = Je^{Js}$ and hence, the vector field generated by the rotation is

$$\frac{dh_s(q)}{ds}\Big|_{s=0} = \big(Jx, 0\big) = (x_2, -x_1, 0).$$

The first integral associated to the rotation vector field is the kinetic momentum with respect to the t-axis

$$\begin{aligned} I &= (\dot x_1 - 2\theta x_2)x_2 + (\dot x_2 + 2\theta x_1)(-x_1) \\ &= (\dot x_1 x_2 - \dot x_2 x_1) - 2\theta |x|^2. \end{aligned}$$

Using the horizontality condition $\dot t = 2\dot x_1 x_2 - 2\dot x_2 x_1$, we get the first integral

$$2I = \dot t - 4\theta|x|^2 = \text{ constant}. \tag{1.53}$$

Proposition 1.46. If c is a geodesic, for any $a \in \mathbf{H}_1$ we have
(i) the left translation $\overline{c} = L_a c$ is also a geodesic,
(ii) the geodesics c and $\overline{c}$ have the same length.

Proof. (i) Let c be a geodesic and let $\overline{c} = L_a c$. From Proposition 1.41 the curve c is horizontal and solves the equation (1.51). From Corollary 1.44 and Proposition 1.9, the curve $\overline{c}$ is a solution of the Euler-Lagrange equation (1.51) and it is horizontal. Using Proposition 1.41, we find that $\overline{c}$ is a geodesic.
(ii) From Corollary 1.10

$$\dot{\overline{c}} = \dot c_1 X_{1|\overline{c}} + \dot c_2 X_{2|\overline{c}}$$

and hence $|\dot{c}|^2 = |\dot{\overline{c}}|^2$. Proposition 1.37, (*ii*) yields

$$\ell(c) = \int_0^1 |\dot{c}|\, ds = |\dot{c}| = |\dot{\overline{c}}| = \int_0^1 |\dot{\overline{c}}|\, ds = \ell(\overline{c}).$$

□

Connectivity by geodesics. Consider a geodesic joining the points P and Q. By a left translation, the point P can be transformed into $(0,0,0)$. By Proposition 1.46, the geodesic is transformed into another geodesic of the same length, which starts at the origin. Hence, it makes sense to study the metric properties and the connectivity only for geodesics starting from the origin.

Because of the Lagrangian rotational symmetry, it is useful to use polar coordinates $x_1 = r\cos\phi$, $x_2 = r\sin\phi$. The Lagrangian becomes

$$L = \frac{1}{2}(\dot{r}^2 + r^2\dot{\phi}^2) + \theta(\dot{t} + 2r^2\dot{\phi}). \tag{1.54}$$

The symmetries for the Lagrangian will provide symmetries for the geodesics. The Lagrangian (1.54) is invariant under the symmetry

$$S : (r, \phi, t; \theta) \to (r, -\phi, -t; -\theta). \tag{1.55}$$

If $(r(s), \phi(s), t(s))$ is a geodesic corresponding to θ, then $(r(s), -\phi(s), -t(s))$ is a geodesic corresponding to $-\theta$. Consequently, whatever statement is made for the geodesics joining the origin with the point $(r(s), \phi(s), t(s))$, when $\theta > 0$, it can also be made for the geodesics between the origin and $(r(s), -\phi(s), -t(s))$, when $\theta < 0$. This allows us to do the analysis only for the case $\theta > 0$ and $\theta = 0$.

1.6.0.2. *Euler-Lagrange equations.* A computation shows

$$\frac{d}{ds}\frac{\partial L}{\partial \dot{r}} = \ddot{r} \qquad \frac{\partial L}{\partial r} = r\dot{\phi}(\dot{\phi} + 4\theta), \tag{1.56}$$

$$\frac{\partial L}{\partial \dot{\phi}} = r^2\dot{\phi} + 2\theta r^2, \qquad \frac{\partial L}{\partial \phi} = 0, \tag{1.57}$$

$$\frac{\partial L}{\partial \dot{t}} = \theta, \qquad \frac{\partial L}{\partial t} = 0, \tag{1.58}$$

and hence $r(s)$, $\phi(s)$ and θ satisfy the Euler-Lagrange system

$$\begin{cases} \ddot{r} = r\dot{\phi}(\dot{\phi} + 4\theta) \\ r^2(\dot{\phi} + 2\theta) = C(\text{constant}) \\ \theta = \theta_0(\text{constant}). \end{cases} \tag{1.59}$$

If a geodesic starts at the origin, then $r(0) = 0$ and hence $C = 0$. The second equation of (1.59) yields $\dot{\phi} = -2\theta$. With the assumption $\theta > 0$, the argument angle ϕ will rotate clock-wise. The Euler-Lagrange system becomes

$$\begin{cases} \ddot{r} = -4\theta^2 r \\ \dot{\phi} = -2\theta \\ \theta = \theta_0(\text{constant}). \end{cases} \tag{1.60}$$

When $\theta = 0$, the system (1.60) becomes

$$\begin{cases} \ddot{r} = 0 \\ \dot{\phi} = 0 \\ \theta = 0. \end{cases} \tag{1.61}$$

PROPOSITION 1.47. Given a point $P(x, 0)$, there is a unique geodesic between the origin and P. It is a straight line in the plane $\{t = 0\}$ of length $|x|$, and it is obtained for $\theta = 0$.

PROOF. In the case $\theta = 0$, the Hamiltonian is $H(\xi) = \frac{1}{2}\xi_1^2 + \frac{1}{2}\xi_2^2$. From the Hamiltonian equation $\dot{t} = \partial H/\partial\theta = 0$, and hence $t(s)$ is constant. Since $t(0) = 0$, it follows that $t(s) = 0$ and the solution belongs to the x-plane. Using the equation $\ddot{r} = 0$, it follows the solution is a straight line. □

From now on, unless otherwise stated, we shall assume $\theta > 0$. From the system (1.60) we can arrive at a first integral of energy.

PROPOSITION 1.48. (i) The function $I(r, \dot{t}, \theta) = \dot{r}^2 + 4\theta^2 r^2$ is a first integral for the system (1.60).
(ii) $\frac{1}{2}I(r, \dot{t}, \theta)$ is equal to the energy of the system, *i.e.*, $I = \dot{x}_1^2 + \dot{x}_2^2$.

PROOF. (*i*) Differentiating

$$\frac{d}{ds} I(r, \dot{t}, \theta) = 2\dot{r}(\ddot{r} + 4\theta^2 r) = 0.$$

(*ii*) In polar coordinates, $\dot{x}_1^2 + \dot{x}_2^2 = \dot{r}^2 + r^2\dot{\phi}^2$. Using $\dot{\phi} = -2\theta$ completes the proof. □

Let $c(s) = \big(x_1(s), x_2(s), t(s)\big)$ be a geodesic. Since $c(s)$ is a horizontal curve,

$$\dot{c} = \dot{x}_1 X_1 + \dot{x}_2 X_2$$

Let g denote the subRiemannian metric in which X_1, X_2 are orthonormal. Then

$$|\dot{c}(s)|_g^2 = g(\dot{c}(s), \dot{c}(s)) = \dot{x}_1^2(s) + \dot{x}_2^2(s).$$

If s is the arc length along the geodesic, the curve $c(s)$ becomes unit speed

$$\dot{x}_1^2(s) + \dot{x}_2^2(s) = 1.$$

Proposition 1.48 yields

$$\dot{r}^2(s) + 4\theta^2 r^2(s) = 1. \tag{1.62}$$

Separating the variables, we have

$$\frac{dr}{\sqrt{1 - 4\theta^2 r^2}} = \pm ds,$$

and integrating, we obtain

$$\frac{1}{2\theta} \arcsin(2\theta r(s)) = \pm s,$$

which yields

$$r(s) = \pm\frac{1}{2\theta}\sin(2\theta s), \tag{1.63}$$

where we consider

$$\begin{array}{ll} + & \text{sign for } 2n\pi \le 2\theta s \le (2n+1)\pi, \text{ and} \\ - & \text{sign for } (2n+1)\pi \le 2\theta s \le (2n+2)\pi. \end{array}$$

This yields a circle of diameter $1/(2\theta)$ which passes through the origin.

LEMMA 1.49. *If* $\phi(0) = \phi_0$, *then*

$$r(\phi)^2 = \Big(\frac{1}{2\theta}\Big)^2 \sin^2(\phi - \phi_0), \tag{1.64}$$

$$t(\phi) - t(\phi_0) = \frac{1}{4\theta^2} \frac{\sin 2(\phi - \phi_0) - 2(\phi - \phi_0)}{2}. \tag{1.65}$$

PROOF. From the second equation of the system (1.60) we obtain $\phi(s) - \phi(0) = -2\theta s$. Substituting in (1.63) yields (1.64). One of the Hamiltonian equations yields

$$\dot t = \frac{\partial H}{\partial \theta} = 2(\dot x_1 x_2 - x_1 \dot x_2) = -2r^2 \dot\phi.$$

Integrating between ϕ_0 and ϕ and using equation (1.64) we have

$$\begin{aligned} t(\phi) - t(\phi_0) &= -2\int_{\phi_0}^{\phi} r^2(\phi)\, d\phi = -2\Big(\frac{1}{2\theta}\Big)^2 \int_{\phi_0}^{\phi} \sin^2(\phi - \phi_0)\, d\phi \\ &= -\frac{2}{4\theta^2}\int_0^{\phi-\phi_0} \sin^2 u\, du = \frac{1}{4\theta^2}\frac{\sin 2(\phi - \phi_0) - 2(\phi - \phi_0)}{2}. \end{aligned}$$

□

1.6.0.3. *Boundary conditions.* We are interested in connecting the origin with a point $P(x, t)$ by a geodesic $c : [0, s_f] \to \mathbb{R}^3$. Consider the following boundary conditions:

$$x(0) = 0, \quad t(0) = 0, \quad \phi(0) = \phi_0, \tag{1.66}$$

$$\|x(s_f)\| = R, \quad t(s_f) = t, \quad \phi(s_f) = \phi_f. \tag{1.67}$$

Because of the rotational invariance around the t-axis, we may choose $\phi_0 = 0$. From $\dot t = -2r^2\dot\phi$ and $\dot\phi = -2\theta$, we get $\dot t = 4\theta r^2 > 0$. Hence $t(s)$ is increasing and if $t(0) = 0$, then $t(s_f) > 0$.

LEMMA 1.50. *The following relations among the boundary conditions take place:*

$$\phi_f = -2\theta s_f, \tag{1.68}$$

$$(\sin\phi_f)^2 = 4\theta^2 R^2, \tag{1.69}$$

$$t = \frac{1}{4\theta^2}\frac{\sin(2\phi_f) - 2\phi_f}{2}, \tag{1.70}$$

$$\frac{t}{R^2} = -\mu(\phi_f) = \mu(2\theta s_f), \tag{1.71}$$

where

$$\mu(x) = \frac{x}{\sin^2 x} - \cot x. \tag{1.72}$$

PROOF. Integrating in the second equation of the system (1.60) and using the boundary conditions yields (1.68). Equations (1.64) and (1.65) together with the boundary conditions (1.66)-(1.67) yield equations (1.69) and (1.70). Eliminating $4\theta^2$ from equations (1.69) and (1.70) yields (1.71). □

The behavior of the function μ given by (1.72) is very important in understanding the subRiemannian geometry of the Heisenberg group. The function μ was first used by Beals, Gaveau and Greiner in [**3**]. The graph of μ for $x \geq 0$ is given in Figure 1.2.

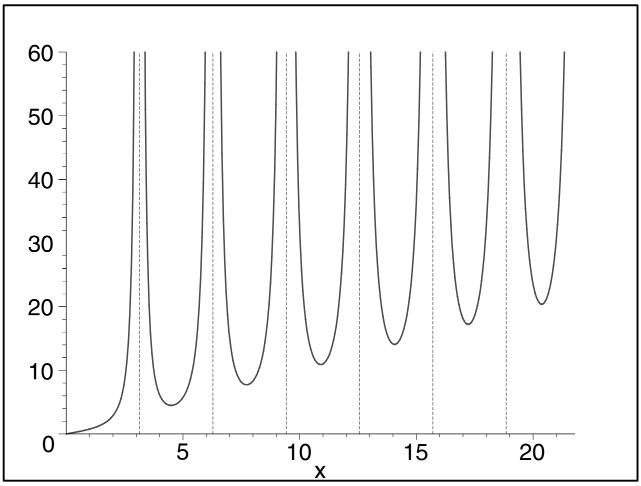

Figure 1.2: The graph of $\mu(x)$ for $x \geq 0$.

LEMMA 1.51. *μ is a monotone increasing diffeomorphism of the interval $(-\pi, \pi)$ onto $\mathbb{R}$. On each interval $(m\pi, (m+1)\pi)$, $m = 1, 2, \ldots,$ μ has a unique critical point x_m. On this interval μ decreases strictly from $+\infty$ to $\mu(x_m)$ and then increases strictly from $\mu(x_m)$ to $+\infty$. Moreover*

$$\mu(x_m) + \pi < \mu(x_{m+1}), \quad m = 1, 2, \ldots \tag{1.73}$$

$$0 < \Big(m + \frac{1}{2}\Big)\pi - x_m < \frac{1}{m\pi}. \tag{1.74}$$

PROOF. As μ is an odd function, it suffices to show that it is a monotone increasing diffeomorphism of the interval $(0, \pi)$ onto $(0, +\infty)$. We note that $\sin x - x\cos x$ vanishes at $x = 0$ and it is increasing on $(0, \pi)$. Then

$$\frac{1}{2}\mu'(x) = \frac{\sin x - x\cos x}{\sin^3 x} = \begin{cases} = 1/3, & x = 0, \\ > 1/3, & x \in (0, \pi). \end{cases}$$

The first identity holds as an application of the l'Hospital rule:

$$\begin{aligned} \lim_{x \to 0} \frac{\sin x - x\cos x}{\sin^3 x} &= \lim_{x \to 0} \frac{\cos x - \cos x + x\sin x}{3\sin^2 x} \\ &= \frac{1}{3} \lim_{x \to 0} \frac{x}{\sin x} = \frac{1}{3}. \end{aligned}$$

The second inequality holds because

$$\frac{1}{2}\mu''(x) = \frac{x + 2x\cos^2 x - 3\cos x \sin x}{\sin^4 x} > 0.$$

The numerator vanishes at $x = 0$, and its derivative is

$$4\sin x(\sin x - x\cos x) > 0, \quad x \in (0, \pi).$$

Therefore μ is a diffeomorphism of the interval $(0, \pi)$ onto $(0, \infty)$. In the interval $(m\pi, (m+1)\pi)$ the function μ approaches $+\infty$ at the endpoints. In order to find the critical points, we set

$$\frac{1}{2}\mu'(x) = \frac{\sin x - x\cos x}{\sin^3 x} = \frac{1 - x\cot x}{\sin^4 x} = 0.$$

Hence the critical point x_m is the solution of the equation $x = \tan x$ on the interval $(m\pi, (m+1)\pi)$. Note that

$$\begin{aligned}\mu(x+\pi) &= \frac{x+\pi}{\sin^2(x+\pi)} - \cot(x+\pi) \\ &= \frac{x}{\sin^2(x+\pi)} - \cot(x+\pi) + \frac{\pi}{\sin^2 x} \\ &= \mu(x) + \frac{\pi}{\sin^2 x},\end{aligned}$$

so the successive minimum values increase by more than π. From Figure 1.3 we have

$$m\pi < x_m < m\pi + \frac{\pi}{2} = (m + \frac{1}{2})\pi. \tag{1.75}$$

Using $x_m = \tan x_m$, yields

$$\cot x_m = \frac{1}{x_m} < \frac{1}{m\pi}. \tag{1.76}$$

Let $f(x) = \cot x$. As $f'(x) = -\frac{1}{\sin^2 x} < -1$, there is a ξ between x and y such that

$$f(x) - f(y) = f'(\xi)(x - y) < -(x - y).$$

Hence $x - y < f(y) - f(x)$. Choosing $x = m\pi + \frac{\pi}{2}$, $y = x_m$ and using

$$f(m\pi + \frac{\pi}{2}) = \frac{\cos(m\pi + \frac{\pi}{2})}{\sin(m\pi + \frac{\pi}{2})} = 0,$$

and (1.76) yields

$$0 < (m + \frac{1}{2})\pi - x_m < \cot x_m < \frac{1}{m\pi}.$$

□

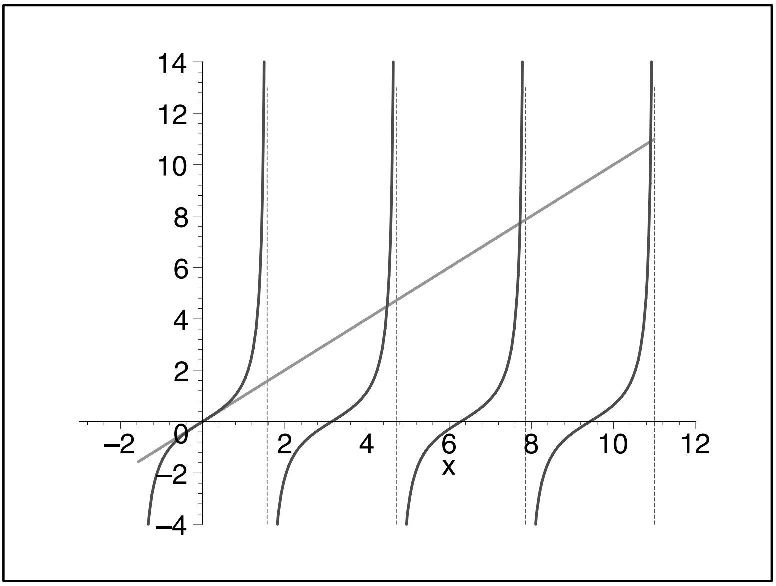

Figure 1.3: *Critical points of* μ *are solutions of* $\tan x = x$.

The number of geodesics which join the origin with an arbitrary given point is given in the following theorems, which can be found in [**3**].

THEOREM 1.52. *There are finitely many geodesics that join the origin to* (x, t) *if and only if* $x \neq 0$. *These geodesics are parametrized by the solutions* ζ *of*

$$\frac{|t|}{\|x\|^2} = \mu(\zeta). \tag{1.77}$$

There is exactly one such geodesic if and only if

$$|t| < \mu(x_1)\|x\|^2, \tag{1.78}$$

where x_1 *is the first critical point of* μ. *The number of geodesics increase without bound as* $|t|/\|x\|^2 \to \infty$. *Let* $\zeta_1 < \zeta_2 < \cdots < \zeta_N$ *be the solutions of (1.77). The square of the length of the geodesic associated with the solution* ζ_m *is*

$$s_m^2 = \Big(\frac{\zeta_m}{\sin \zeta_m}\Big)^2 \|x\|^2. \tag{1.79}$$

PROOF. The enumeration of geodesics follows from Lemma 1.51. The line $y = |t|/\|x\|^2$ intersects the graph of μ finitely many times. There is only one intersection if and only if inequality (1.78) holds. See the graph of μ in Figure 1.2. As the geodesic was parametrized by arc length, its length is given by the value of the parameter s_f. Dividing the square of the equation (1.68) by equation (1.69) yields the square of the length

$$s_f^2 = \Big(\frac{\phi_f}{\sin \phi_f}\Big)^2 R^2.$$

ϕ_f satisfies equation (1.71). For each $\phi_f = -\zeta_m$ we obtain the length described by formula (1.79). $\square$

The natural dilations of the Heisenberg Laplacian operator

$$\frac{1}{2}\Big(\partial_{x_1} + 2x_2\partial_t\Big)^2 + \frac{1}{2}\Big(\partial_{x_2} - 2x_1\partial_t\Big)^2$$

are

$$(x_1, x_2, t) \to (\lambda x_1, \lambda x_2, \lambda^2 t), \quad \lambda > 0.$$

We are looking for a formula for the length of geodesics, different than (1.79), such that s_f is homogeneous of degree 1 with respect to the above dilations. This is done in the next result.

THEOREM 1.53. *Consider the geodesics joining the origin with the point (x, t), $x \neq 0$. Let $\zeta_1, \dots, \zeta_N$ be the solutions of equation (1.77). Then the square of the lengths is given by*

$$s_m^2 = \nu(\zeta_m)\Big(|t| + \|x\|^2\Big), \tag{1.80}$$

where

$$\nu(x) = \frac{x^2}{x + \sin^2 x - \sin x \cos x}.$$

PROOF. Equation (1.77) yields

$$|t| + \|x\|^2 = \mu(\zeta_m)\|x\|^2 + \|x\|^2 = (1 + \mu(\zeta_m))\|x\|^2,$$

and hence

$$\|x\|^2 = \frac{1}{1 + \mu(\zeta_m)}\Big(|t| + \|x\|^2\Big).$$

Using equation (1.79) and the definition of μ given in (1.72)

$$\begin{aligned} s_m^2 &= \Big(\frac{\zeta_m}{\sin \zeta_m}\Big)^2 \|x\|^2 = \Big(\frac{\zeta_m^2}{\sin^2 \zeta_m} \cdot \frac{1}{1 + \mu(\zeta_m)}\Big)\Big(|t| + \|x\|^2\Big) \\ &= \frac{\zeta_m^2}{\sin^2 \zeta_m + \zeta_m - \sin^2 \zeta_m \cot \zeta_m}\Big(|t| + \|x\|^2\Big) = \nu(\zeta_m)\Big(|t| + \|x\|^2\Big). \end{aligned}$$

For the graph of ν see Figure 1.4. □

PROPOSITION 1.54. The projection of the geodesics joining the origin and (x, t), $x \neq 0$ are circles or arcs of circle with diameters of at least $\|x\|$

$$\frac{1}{2\theta_m} = \frac{\|x\|}{|\sin \zeta_m|} \geq \|x\|. \tag{1.81}$$

PROOF. From the first equation of Lemma 1.49, the diameter of the circle is $1/2\theta$. Using relation (1.69) yields

$$\frac{1}{2\theta} = \frac{R}{|\sin \phi_f|}.$$

Replacing R by $\|x\|$ and ϕ_f by ζ_m, we arrive at relation (1.81). □

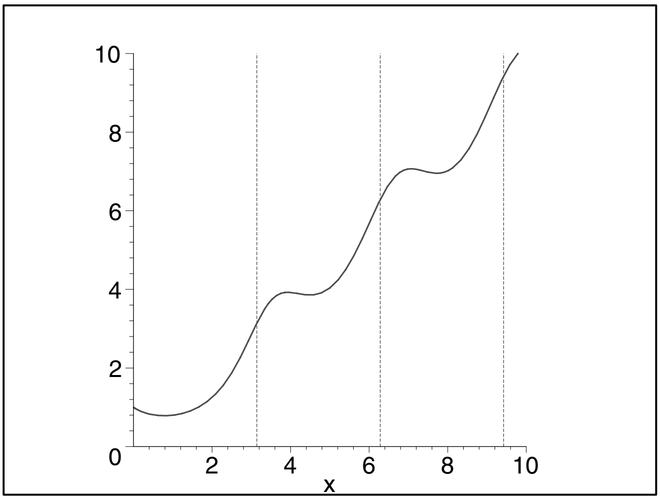

Figure 1.4*: The graph of* $\nu(x)$.

1.6.0.4. *Geodesics between the origin and points on the* t*-axis.* We have seen that the number of geodesics increases without bound as $|t|/\|x\|^2 \to \infty$. The following theorem deals with the limit case $\|x\| = 0$. This result was found by Gaveau (see [**26**] and [**3**]).

THEOREM 1.55. *The geodesics that join the origin to a point* $(0, 0, t)$ *have lengths* $s_1, s_2 \ldots,$ *where*

$$s_m^2 = m\pi|t|. \tag{1.82}$$

For each length s_m*, the geodesics of that length are parametrized by the circle* $\mathbb{S}^1$.

PROOF. We shall treat the problem as the limit case $\|x\| \to 0$ of Theorem 1.53. Then the solutions $\zeta_m \to m\pi$. Relation (1.80) becomes

$$s_m^2 = \nu(m\pi)|t|,$$

with the coefficient given by

$$\nu(m\pi) = \frac{m^2\pi^2}{m\pi + \sin^2 m\pi - \sin m\pi\, \cos m\pi} = m\pi.$$

□

The above proof is using Theorem 1.53. In the following we shall provide a direct proof. The approach will use polar coordinates and the fact that the solution is given by

$$r(\phi) = r_{max}|\sin(\phi - \phi_0)|.$$

Consider the solution γ parameterized by $[0, 1]$, with the end points $\gamma(0) = (0, 0, 0)$ and $\gamma(1) = (0, 0, t)$. Using that $|\dot\gamma|$ is constant along the solution, we have the identity in Cauchy's inequality

$$\ell(\gamma) = \int_0^1 |\dot\gamma(s)|\, ds = \Big(\int_0^1 ds\Big)^{1/2}\Big(\int_0^1 |\dot\gamma(s)|^2\Big)^{1/2} = \sqrt{2E},$$

where E is the constant value of the energy along the solution

$$E = \frac{1}{2}|\dot\gamma|^2 = \frac{1}{2}(\dot x_1^2 + \dot x_2^2).$$

1.6.0.5. *Quantization of* θ. θ is constant along the solution and it is an eigenvalue for the following boundary value problem

$$\begin{aligned} \ddot r &= -4\theta^2 r \\ r(0) &= 0,\ r(1) = 0. \end{aligned}$$

The solution is $r(s) = A\sin(2\theta s)$. The boundary conditions yield

$$2\theta = m\pi, \tag{1.83}$$

with m integer. Each value of θ will determine a certain length.

1.6.0.6. *The energy and* t. Consider the boundary condition $t(1) = t$. One has $\dot\phi = -2\theta$ and then $\phi(s) = \phi_0 - 2\theta s$, where $\phi_0 = \phi(0)$ and $\phi_1 = \phi(1) = \phi_0 - 2\theta = \phi_0 - m\pi$. One may integrate between ϕ_0 and $\phi_1 = \phi_0 - m\pi$ in the equation $\dot t = -2r^2\dot\phi$.

$$\begin{aligned} t &= -2\int_{\phi_0}^{\phi_0 - m\pi} r^2(\phi)\,d\phi = -2\,r_{max}^2\int_{\phi_0}^{\phi_0-m\pi}\sin^2(\phi-\phi_0)\,d\phi \\ &= -2\,r_{max}^2\int_0^{-m\pi}\sin^2 u\,du = 2\,r_{max}^2\int_0^{m\pi}\sin^2 u\,du \\ &= 2\,r_{max}^2(\frac{1}{2}u - \frac{1}{4}\sin 2u)\Big|_0^{m\pi} = r_{max}^2 m\pi = 2\theta r_{max}^2. \end{aligned}$$

From the conservation of energy

$$\frac{1}{2}(\dot r^2 + r^2\dot\phi^2) = \frac{1}{2}(\dot r^2 + 4r^2\theta^2) = E.$$

When $r = r_{max}^2$, then $\dot r = 0$. Hence

$$r_{max}^2 = \frac{E}{2\theta^2},$$

and using $t = 2\theta r_{max}^2$, one obtains

$$t = \frac{E}{\theta}.$$

The lengths are

$$(\ell_m)^2 = 2E = 2\theta t = m\pi t\ ,\quad m = 1, 2, 3\ldots$$

PROPOSITION 1.56. The equations of the geodesics starting at the origin are

$$\begin{aligned} r_m(\phi) &= \sqrt{\frac{|t|}{m\pi}}\,\sin(\phi - \phi_0) \\ t_m(\phi) &= \frac{|t|}{2m\pi}\Big(\sin(2(\phi-\phi_0)) - 2(\phi - \phi_0)\Big),\ \ m \ge 1. \end{aligned}$$

PROOF. Using

$$r_{max} = \sqrt{\frac{E}{2\theta^2}} = \sqrt{\frac{|t|}{2\theta}} = \sqrt{\frac{|t|}{m\pi}}$$

then

$$r_m(\phi) = \sqrt{\frac{|t|}{m\pi}} \sin(\phi - \phi_0).$$

For computing $t_m(\phi)$ one integrates $dt = -2r^2\, d\phi$ between ϕ_0 and ϕ

$$t(\phi) = -2r_{max}^2 \int_{\phi_0}^{\phi} \sin^2(\phi - \phi_0)\, d\phi = -2\frac{|t|}{m\pi} \int_0^{\phi-\phi_0} \sin^2 u\, du.$$

Hence

$$t_m(\phi) = \frac{|t|}{2m\pi}\Big(\sin(2(\phi - \phi_0)) - 2(\phi - \phi_0)\Big).$$

□

COROLLARY 1.57. The projection of the m-th geodesic on the x-plane is a circle with the radius

$$R_m = \frac{1}{2}\sqrt{\frac{|t|}{m\pi}} \tag{1.84}$$

and area

$$\sigma_m = \frac{|t|}{4m}.$$

We note that r_m and t_m depend on ϕ_0 but the corresponding lengths $\ell_m = \sqrt{m\pi|t|}$ do not. This is because of the rotational symmetry around the t-axis.

1.6.0.7. *A Milnor-type property.* In the classical theory of three dimensional curves one defines the total curvature of a curve $c : [0, \tau] \to \mathbb{R}^3$ as $\int_0^\tau k(u)\, du$, where $k(s)$ is the curvature along the curve c. Milnor's theorem states

THEOREM 1.58. *The total curvature of a closed curve in $\mathbb{R}^3$ is $2\pi\chi$, where χ is an integer.*

It is known that χ represents the number of knots of the curve c. In the case of a plane curve, $\chi = 1$. This result was proved independently by Fechnel. For these proofs the reader may consult Millman and Parker [**46**].

We shall show a similar property for the geodesics between the origin and points on the t-axis. We shall prove the following:

PROPOSITION 1.59. The total curvature of a geodesic between the origin and the point $(0, 0, t)$ is $2\pi m$, where $m = 1, 2, \dots$ is an integer which gives the number of rotations of the geodesic.

PROOF. Consider the curve $c : [0, s_f] \to \mathbb{R}^3$ joining $(0, 0, 0)$ and $(0, 0, t)$, parametrized by the arc length. Then $|\dot{c}(u)| = 1$, and by Proposition 1.38, the curvature is $k(u) = |\ddot{c}(u)| = 4|\theta|$. Hence, the total curvature is

$$\int_0^{s_f} k(u)\, du = 4|\theta| s_f.$$

From the boundary value problem

$$\begin{aligned} \ddot{r} &= -4\theta^2 r \\ r(0) &= 0,\ r(s_f) = 0 \end{aligned}$$

it follows that $2\theta s_f = m\pi$. Therefore, the total curvature is equal to $2\pi m$, $m = 1, 2, \dots$ □

Geodesics between any two arbitrary points. Given two points $P_1(x_1, y_1, t_1)$ and $P_2(x_2, y_2, t_2)$, we shall study the connectivity by geodesics in the cases $(x_1, y_1) = (x_2, y_2)$ and $(x_1, y_1) \neq (x_2, y_2)$. Making a left translation by $(-x_1, -y_1, -t_1)$, the point P_1 will become the origin and the point P_2 will have the coordinates

$$\big(x_2 - x_1, y_2 - y_1, t_2 - t_1 - 2(y_1 x_2 - x_1 y_2)\big).$$

By Proposition 1.46 the left translation of a geodesic is a geodesic of the same length. Then theorems 1.52 and 1.53 become:

THEOREM 1.60. *Given the points $P_1(x_1, y_1, t_1)$ and $P_2(x_2, y_2, t_2)$, there are finitely many geodesics between P_1 and P_2 if and only if $(x_1, y_1) \neq (x_2, y_2)$. These geodesics are parametrized by the solutions ζ of*

$$\frac{|t_2 - t_1 - 2(y_1 x_2 - x_1 y_2)|}{(x_2 - x_1)^2 + (y_2 - y_1)^2} = \mu(\zeta). \tag{1.85}$$

There is exactly one such geodesic if and only if

$$|t_2 - t_1 - 2(y_1 x_2 - x_1 y_2)| < \mu(c_1)[(x_2 - x_1)^2 + (y_2 - y_1)^2], \tag{1.86}$$

where c_1 is the first critical point of μ. The number of geodesics increase without bound as

$$\frac{|t_2 - t_1 - 2(y_1 x_2 - x_1 y_2)|}{[(x_2 - x_1)^2 + (y_2 - y_1)^2]} \to \infty.$$

Let $\zeta_1 < \zeta_2 < \cdots < \zeta_N$ be the solutions of (1.85). *The square of the length associated with the solution ζ_m is*

$$\begin{aligned} s_m^2 &= \Big(\frac{\zeta_m}{\sin \zeta_m}\Big)^2 [(x_2 - x_1)^2 + (y_2 - y_1)^2] \\ &= \nu(\zeta_m)\Big(|t_2 - t_1 - 2(y_1 x_2 - x_1 y_2)| + [(x_2 - x_1)^2 + (y_2 - y_1)^2]\Big), \end{aligned}$$

where

$$\nu(x) = \frac{x^2}{x + \sin^2 x - \sin x \cos x}.$$

Theorem 1.55 becomes:

THEOREM 1.61. *Given the points $P_1(x_1, y_1, t_1)$ and $P_2(x_2, y_2, t_2)$, with $x_1 = x_2$ and $y_1 = y_2$, the geodesics that join P_1 and P_2 have lengths*

$$s_m^2 = m\pi |t_2 - t_1|.$$

1.7. Carnot-Carathéodory distance

Recall that a curve $c(s) = \big(x_1(s), x_2(s), t(s)\big)$ is horizontal if $\dot{c}(s) \in \mathcal{H}_{c(s)}$, *i.e.,*

$$\dot{t} - 2x_2\dot{x}_1 + 2x_1\dot{x}_2 = 0.$$

By Proposition 1.11, any two points P and Q can be joined by a horizontal curve. We have shown that the curve can be considered smooth. Hence the set

$$\{c;\ c(0) = P, c(1) = Q, c \text{ horizontal curve}\} \neq \emptyset.$$

The length of a horizontal curve c is

$$\ell(c) = \int_0^1 \sqrt{g(\dot{c}(s), \dot{c}(s))}\, ds,$$

where $g : \mathcal{H} \times \mathcal{H} \to \mathcal{F}$ is the subRiemannian metric. The Carnot-Carathéodory distance is defined as $d_C : \mathbb{R}^3 \times \mathbb{R}^3 \to [0, \infty)$,

$$d_C(P, Q) = \inf\{\ell(c); c \in S\}. \tag{1.87}$$

One may verify that d_C verifies the distance axioms:
(i) $d_C(P, P) = 0$,
(ii) $d_C(P, Q) = d(Q, P)$,
(iii) $d_C(P, R) \leq d_C(P, Q) + d(Q, R)$.

The main result of this section is to show that the Carnot-Carathéodory distance d_C is complete, *i.e.,* any Cauchy sequence with respect to d_C is convergent. We shall prove this using the equivalence between the geodesic completeness and the completeness as a metric space.

Definition 1.62. If for any point P, any geodesic $c(t)$ emanating from P is defined for all $t \in \mathbb{R}$, the geometry is called geodesically complete.

The following theorem can be found in Strichartz [**50**], [**51**]:

Theorem 1.63. *Let M be a connected step 2 subRiemannian manifold.*
(a) If M is complete, then any two points can be joined by a geodesic.
(b) If there exists a point P such that every geodesic from P can be indefinitely extended, then M is complete.
(c) Every nonconstant geodesic is locally a unique length minimizing curve.
(d) Every length minimizing curve is a geodesic.

It is known that geodesics starting from the origin on the Heisenberg group are infinitely extendable. Using (b) of Theorem 1.63 we obtain:

Theorem 1.64. *The Carnot-Carathéodory metric d_C is complete.*

Computing the Carnot-Carathéodory distance. In the following we shall describe two ways to obtain the Carnot-Carathéodory distance.

First method:
From (c) and (d) of Theorem 1.63 we obtain a way to compute the Carnot-Carathéodory distance as

$$d_C(P, Q) = \{\ell(c),\ \text{where } c \text{ is the shortest geodesic joinning P and Q}\}.$$

Using Theorem 1.53 with $m = 1$, the Carnot-Carathéodory distance from the origin to (x, t) is

$$d_C(0, (x, t)) = \nu(\zeta_1)\big(|t| + \|x\|^2\big).$$

In particular, if the points are on the same vertical line, the Carnot-Carathéodory distance squared is proportional to the Euclidian distance

$$d_C\big((x, t), (x, t')\big)^2 = \pi|t' - t| = \pi d_E\big((x, t), (x, t')\big).$$

Second method:
Another way to obtain the Carnot-Carathéodory distance is using the complex action. We shall discuss this issue in more detail later in the Complex Hamiltonian Mechanics chapter, (ch 5).
Consider the modified complex action function

$$f(x, t, \tau) = \tau g(x, t, \tau) = -i\tau t + \tau \coth(2\tau)\|x\|^2,$$

where

$$\begin{aligned} g(x, t, \tau) &= -it + \int_0^\tau \{\langle \dot{x}, \xi \rangle - H\}\, ds \\ &= -it + \coth(2\tau)\|x\|^2 \end{aligned}$$

is the complex action. Like the classical action, the complex action g satisfies the Hamilton-Jacobi equation, see [**3**]

$$\frac{\partial g}{\partial \tau} + H\Big(x, \frac{\partial g}{\partial x}\Big) = 0.$$

Using Theorem 1.66 of [**3**], there is a unique critical point with respect to τ of the modified complex action function f in the strip $\{|\mathrm{Im}(\tau)| < \pi/2\}$ given by $\tau_c(x, t) = i\theta_c(x, t)$, where θ_c is the solution of

$$t = \mu(2\tau\theta)\|x\|^2$$

in this interval. At the critical point

$$f(x, t, \tau_c) = \frac{1}{2} d_C(0, (x, t)).$$

This works only in the case $x \neq 0$.

Application to the fundamental solution. We plan to investigate whether we can use the subRiemannian geometry to construct the fundamental solution of the operator

$$\begin{aligned} \Delta_X &= X_1^2 + X_2^2 \\ &= \Big(\partial_{x_1} + 2x_2\partial_t\Big)^2 + \Big(\partial_{x_2} - 2x_1\partial_t\Big)^2 \\ &= \partial_{x_1}^2 + \partial_{x_2}^2 + 4\partial_t\Big(x_2\partial_{x_1} - x_1\partial_{x_2}\Big) + 4(x_1^2 + x_2^2)\partial_t^2. \end{aligned}$$

It is expected that the subRiemannian distance (Carnot-Carathéodory distance) will play a role similar to the one the Euclidian distance plays for the Laplacian, *i.e.,* the fundamental solution is the inverse of the Carnot-Carathéodory distance

$$K(x, t; 0, 0) = \frac{C}{d_C(x, t)} = \frac{C}{(|x|^4 + t^2)^{1/2}}.$$

We need the following lemma.

LEMMA 1.65. *For any functions $f, h \in C^2(\mathbb{R}^3)$, we have*

$$\Delta_X(fh) = f\Delta_X h + h\Delta f + 2\sum_{i=1}^{2} X_i(f)X_i(h) \tag{1.88}$$

$$\Delta_X(f^2) = 2f\Delta_X f + 2\sum_{i=1}^{2} X_i(f)^2 \tag{1.89}$$

$$\Delta_X\Big(\frac{1}{f}\Big) = \frac{2}{f^3}\sum_{i=1}^{2}(X_i f)^2 - \frac{1}{f^2}\Delta_X f. \tag{1.90}$$

PROOF. For $i = 1, 2$ we have

$$\begin{aligned} X_i^2(fh) &= X_i\Big(X_i(fh)\Big) = X_i\Big(f\ X_i f + h\ X_i f\Big) \\ &= X_i\Big(f\ X_i h\Big) + X_i\Big(h\ X_i f\Big) \\ &= f\ X_i^2 h + h\ X_i^2 f + 2X_i(f)X_i(h). \end{aligned}$$

Summing in the above relation with $i = 1, 2$ we obtain relation (1.88). Making $f = h$ we obtain the second relation. By Lemma 1.65, we have

$$0 = \Delta_X\Big(f\frac{1}{f}\Big) = f\,\Delta_X\Big(\frac{1}{f}\Big) + \frac{1}{f}\Delta_X f + 2\sum_{i=1}^{2} X_i(f)X_i\Big(\frac{1}{f}\Big). \tag{1.91}$$

Using

$$X_i\Big(f\frac{1}{f}\Big) = 0$$

we get

$$X_i\Big(\frac{1}{f}\Big) = -\frac{1}{f^2}X_i(f)$$

and substituting in (1.91) yields

$$\Delta_X\Big(\frac{1}{f}\Big) = \frac{2}{f^3}\sum X_i(f)^2 - \frac{1}{f^2}\Delta_X f. \tag{1.92}$$

□

LEMMA 1.66. *If d is the Carnot-Carathéodory distance from the origin, $d_C(x,t) = \sqrt{|x|^4 + t^2}$, then*

$$\sum X_i(d)^2 = 4|x|^2.$$

PROOF. From $X_1(d^2) = 2dX_1(d)$ we get

$$X_1(d) = \frac{1}{2d}X_1(d^2). \tag{1.93}$$

The right side of (1.93) can be computed explicitly

$$\begin{aligned} X_1(d^2) &= (\partial_{x_2} + 2x_1\partial_t)\big((x_1^2 + x_2^2)^2 + t^2\big) \\ &= 4|x|^2 x_2 + 4x_1 t. \end{aligned}$$

Hence

$$X_1(d) = \frac{2}{d}\Big(|x|^2 x_2 + x_1 t\Big). \tag{1.94}$$

Similarly,

$$X_2(d^2) = 4\Big(|x|^2 x_1 - x_2 t\Big),$$

$$X_2(d) = \frac{2}{d}\Big(|x|^2 x_1 - x_2 t\Big). \tag{1.95}$$

Summing the squares of (1.93) and (1.95) yields

$$\begin{aligned} X_1(d)^2 + X_2(d)^2 &= \frac{4}{d^2}\Big((|x|^2 x_2 + x_1 t)^2 + (|x|^2 x_1 - x_2 t)^2\Big) \\ &= \frac{4}{d^2}\Big(|x|^4(x_1^2 + x_2^2) + t^2(x_1^2 + x_2^2)\Big) \\ &= 4|x|^2. \end{aligned}$$

□

Lemma 1.67. *We have*

$$\Delta_X(d^2) = 24|x|^2.$$

Proof. Using the formula for $X_1(d^2)$ from the previous proof

$$\begin{aligned} X_1^2(d^2) &= (\partial_{x_2} + 2x_1\partial_t)(4(x_1^2 + x_2^2)x_2 + 4x_1 t) \\ &= 4\big((x_1^2 + x_2^2) + 2x_2^2 + 2x_1^2\big) \\ &= 12|x|^2. \end{aligned}$$

Similarly,

$$X_2^2(d^2) = 12|x|^2.$$

Summing the squares of the last two relations yields the result. □

Lemma 1.68. *We have*

$$d\Delta_X d = 8|x|^2.$$

Proof. Substitute the formulas of Lemma 1.66 and Lemma 1.67 into the relation

$$\Delta_X d^2 = 2d\Delta_X d + 2\sum (X_i d)^2.$$

□

Proposition 1.69. For $(x, t) \neq (0, 0)$, we have

$$\Delta_X\Big(\frac{1}{d}\Big) = 0.$$

PROOF. Choosing $f = d$ in the third formula of Lemma 1.65 and using the relations given by Lemma 1.67 and Lemma 1.68 yields

$$\begin{aligned}\Delta_X\Big(\frac{1}{d}\Big) &= \frac{2}{d^3}\sum X_i(d)^2 - \frac{1}{d^2}\Delta_X d\\ &= \frac{2}{d^3}4|x|^2 - \frac{1}{d^2}\frac{8|x|^2}{d}\\ &= 0.\end{aligned}$$

□

One may show that in the sense of distributions

$$\Delta_X\Big(\frac{1}{d}\Big) = \delta_{(0,0)}.$$

For details see Folland [**25**].

1.8. Exercises

EXERCISE 1.1. Let $\delta_\lambda : \mathbb{R}^3 \to \mathbb{R}^3$ given by $\delta_\lambda(x,t) = (\lambda x, \lambda^2 t)$.

i) Show that $(\delta_\lambda)_{\lambda\in\mathbb{R}}$ form a group, called the one-parameter group of dilations.

ii) Show that $(\delta_\lambda)^* X_1 = \lambda^{-1}X_1$, $(\delta_\lambda)^* X_2 = \lambda^{-1}X_2$, where X_1, X_2 are the Heisenberg vector fields.

iii) The length of a horizontal curve is multiplied by $|\lambda|$ under the action of δ_λ.

iv) Show that $d_C(\delta_\lambda p, \delta_\lambda q) = |\lambda| d_C(p,q)$, for any two points $p, q \in \mathbb{R}^3$.

v) Prove the following estimation

$$\frac{1}{3}\Big(|x| + |t|^{1/2}\Big) \le d_C(0,(x,t)) \le 4\Big(|x| + |t|^{1/2}\Big).$$

EXERCISE 1.2. Define the unit ball centered at p as

$$\mathbf{B}(p,1) = \{(x,t)\in\mathbb{R}^3; d_C(x,t) \le 1\},$$

where $d_c(x,t) = (|x|^4 + t^2)^{1/2}$.

i) Draw the unit ball centered at the origin.

ii) Find the volume of the unit ball.

iii) How is the volume of the unit ball changing under left translations?

iv) How is the volume of the unit ball changing under dilations?

v) Draw the unit ball centered at a point (x_0, t_0) in the cases

$$a)\quad x_0 = 0, \qquad\qquad b)\quad x_0 \neq 0.$$

vi) Find two constants $C_1, C_2 > 0$ such that

$$C_1[-\epsilon,\epsilon]\times[-\epsilon,\epsilon]\times[-\epsilon^2,\epsilon^2] \subset \mathbf{B}(0,\epsilon) \subset C_2[-\epsilon,\epsilon]\times[-\epsilon,\epsilon]\times[-\epsilon^2,\epsilon^2].$$

vii) Show that

$$Vol(B(p,\epsilon)) = \epsilon^4\, Vol(B(0,1)), \qquad \forall p\in\mathbb{R}^3,$$

where Vol denotes the usual volume defined by the Lebesgue measure.

EXERCISE 1.3. (see [**20**]) Let $c = (c_1, c_2, c_3) : I \to \mathbf{H}_1$ be a C^1-curve, such that 0 is in the interior of I and $c(0) = g$. Show that if $c_3'(0) \neq 0$, then there exists $\epsilon > 0$ such that

$$\mathbf{B}(g_0, \epsilon) \subset \bigcup_{s \in I} H_{c(s)},$$

where $\mathbf{B}(g_0, \epsilon) = \{g \in \mathbf{H} : d_C(g_0^{-1} g) \leq \epsilon\}$, and H_p denotes the horizontal plane through the point p.

EXERCISE 1.4. Find the expression for the exponential map of the Lie group $\mathbf{H}_1$ as subgroup of $GL(3, \mathbb{R})$.

EXERCISE 1.5. Show that on every vertical line in $\mathbb{R}^3$ the Carnot-Carathéodory metric is locally equivalent to $\sqrt{\text{Euclidean metric}}$.