From Queen Dido to the Heisenberg Group: Control-Theoretic Perspective

Fedor Prikolab

Dec 13 2021

Abstract

This article offers a step-by-step, control-theoretic introduction to the Heisenberg group, emphasizing its role as the simplest setting where non-holonomic constraints naturally appear. We show how the horizontal distribution of the Heisenberg group arises from restricting the motion of a control system, and how this leads to a sub-Riemannian notion of length. Building on these foundations, we derive the geodesics of the Heisenberg group directly from its Hamiltonian formulation, showing how their structure reflects the geometry of the underlying distribution. To clarify the theory further, we revisit the classical proof that abnormal geodesics cannot occur in this setting and highlight the key geometric idea behind the argument. Finally, we connect these concepts to a familiar variational problem—the Queen Dido isoperimetric problem—to illustrate how sub-Riemannian geodesics can be viewed as natural analogues of classical extremals. The goal is to give an intuitive and accessible pathway into the geometry and control theory of the Heisenberg group.

1 Introducing Heisenberg Group

Before introducing the sub-Riemannian geometry of the Heisenberg group, we begin with its realization as a Lie group. We start with a standard definition - the (real) three-dimensional Heisenberg group is defined as the subgroup

$$\mathbb{H} = \left\{ \begin{pmatrix} 1 & a & c \\ 0 & 1 & b \\ 0 & 0 & 1 \end{pmatrix} : a, b, c \in \mathbb{R} \right\},$$

with the group operation given by matrix multiplication. Although this is already a smooth manifold inside $GL_3(\mathbb{R})$, we would like to exhibit explicitly the Lie group structure and the associated Lie algebra.

The Lie Algebra and the Exponential Map

Consider the vector space of strictly upper triangular matrices

$$\mathfrak{h} = \left\{ \begin{pmatrix} 0 & x & z \\ 0 & 0 & y \\ 0 & 0 & 0 \end{pmatrix} : x, y, z \in \mathbb{R} \right\}.$$

This is a three-dimensional nilpotent Lie algebra (in fact, the matrices are nilpotent of order 3). Since any element $A \in \mathfrak{h}$ satisfies $A^3 = 0$, the exponential map truncates after the quadratic term:

$$\exp(A) = I + A + \frac{1}{2}A^2.$$

A direct computation shows that if

$$A = \begin{pmatrix} 0 & x & z \\ 0 & 0 & y \\ 0 & 0 & 0 \end{pmatrix}, \qquad A^2 = \begin{pmatrix} 0 & 0 & xy \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

then

$$\exp(A) = \begin{pmatrix} 1 & x & z + \frac{1}{2}xy \\ 0 & 1 & y \\ 0 & 0 & 1 \end{pmatrix}.$$

Since the exponential map of a nilpotent Lie algebra is a global diffeomorphism, $\exp : \mathfrak{h} \to \mathbb{H}$ is a diffeomorphism, allowing us to identify

$$(x, y, z) \quad \longleftrightarrow \quad \exp \begin{pmatrix} 0 & x & z \\ 0 & 0 & y \\ 0 & 0 & 0 \end{pmatrix}.$$

The Group Law in Exponential Coordinates

Let

$$A = \begin{pmatrix} 0 & x & z \\ 0 & 0 & y \\ 0 & 0 & 0 \end{pmatrix}, \qquad A' = \begin{pmatrix} 0 & x' & z' \\ 0 & 0 & y' \\ 0 & 0 & 0 \end{pmatrix}.$$

Using the Baker–Campbell–Hausdorff formula for step-2 nilpotent Lie algebras,

$$\exp(A)\exp(A') = \exp\bigl(A + A' + \tfrac{1}{2}[A, A']\bigr),$$

and computing

$$[A, A'] = \begin{pmatrix} 0 & 0 & xy' - x'y \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

we obtain the induced group law on $\mathbb{R}^3$:

$$(x, y, z) * (x', y', z') = \bigl(x + x',\ y + y',\ z + z' + \tfrac{1}{2}(xy' - x'y)\bigr).$$

Thus the Heisenberg group is diffeomorphic to $\mathbb{R}^3$ endowed with the above non-commutative multiplication.

Smooth Curves in the Heisenberg Group

Let $\gamma : [0, T] \to \mathbb{H}$ be a smooth curve. In exponential coordinates we may write

$$\gamma(t) = \exp \begin{pmatrix} 0 & x(t) & z(t) \\ 0 & 0 & y(t) \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & x(t) & z(t) + \frac{1}{2}x(t)y(t) \\ 0 & 1 & y(t) \\ 0 & 0 & 1 \end{pmatrix},$$

so that

$$\gamma(t) \sim (x(t), y(t), z(t)) \in \mathbb{R}^3.$$

This representation will be convenient later when we introduce the left-invariant vector fields and describe the associated Lie algebra generators that form the standard basis of the Heisenberg Lie algebra.

A control-theoretic viewpoint on the Heisenberg group

We now explain how the geometry of the Heisenberg group naturally leads to a control system and how the problem of finding geodesics in this group can be interpreted as an optimal control problem.

Let $\gamma : [0, T] \to \mathbb{H}$ be a smooth curve. Writing $\gamma(t)$ as a matrix in $\mathbb{H}$, we have

$$\gamma(t) = \begin{pmatrix} 1 & a(t) & c(t) \\ 0 & 1 & b(t) \\ 0 & 0 & 1 \end{pmatrix},$$

where, in terms of exponential coordinates, we identify

$$a(t) = x(t), \qquad b(t) = y(t), \qquad c(t) = z(t) + \tfrac{1}{2}x(t)y(t).$$

Lemma 1. *For every smooth curve $\gamma : [0,T] \to \mathbb{H}$ there exists a smooth curve $V : [0,T] \to \mathfrak{h}$ such that*

$$\dot\gamma(t) = \gamma(t)\, V(t) \qquad \text{for all } t \in [0,T],$$

where $\mathfrak{h}$ is the Lie algebra of $\mathbb{H}$.

A direct computation shows that

$$V(t) = \begin{pmatrix} 0 & \dot a(t) & \dot c(t) - a(t)\dot b(t) \\ 0 & 0 & \dot b(t) \\ 0 & 0 & 0 \end{pmatrix}.$$

Using the relations

$$a = x, \qquad b = y, \qquad c = z + \tfrac{1}{2}xy,$$

we obtain

$$\dot c - a\dot b = \dot z - \tfrac{1}{2}\big(x\dot y - y\dot x\big).$$

Thus, in exponential coordinates (x,y,z), the "velocity" of the curve γ can be written in terms of $(\dot x, \dot y, \dot z)$ and decomposed in the basis of the Lie algebra.

Let us introduce a basis of left-invariant vector fields for the Lie algebra of $\mathbb{H}$. At a point $p = (x,y,z) \in \mathbb{H} \cong \mathbb{R}^3$ we define

$$X_p = \partial_x - \frac{y}{2}\,\partial_z, \qquad Y_p = \partial_y + \frac{x}{2}\,\partial_z, \qquad Z_p = \partial_z.$$

These are the left-invariant vector fields corresponding to the matrices

$$X = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad Y = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}, \quad Z = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

and they satisfy the commutation relations

$$[X,Y] = Z, \qquad [X,Z] = [Y,Z] = 0.$$

Any tangent vector $\dot\gamma(t) \in T_{\gamma(t)}\mathbb{H}$ can be uniquely written in the form

$$\dot\gamma(t) = u_1(t)\, X_{\gamma(t)} + u_2(t)\, Y_{\gamma(t)} + u_3(t)\, Z_{\gamma(t)},$$

for suitable scalar functions $u_1, u_2, u_3 : [0,T] \to \mathbb{R}$. In control theory it is natural to interpret these coefficients as *controls*. Thus we obtain a control system on $\mathbb{H}$:

$$\dot p(t) = u_1(t)\, X_{p(t)} + u_2(t)\, Y_{p(t)} + u_3(t)\, Z_{p(t)}, \qquad p(t) \in \mathbb{H}. \tag{1}$$

In many applications to sub-Riemannian geometry and control, one restricts attention to curves whose velocity lies in the span of X and Y only, i.e. we impose

$$u_3(t) = 0 \quad \text{for all } t.$$

These are called *horizontal curves*. In coordinates $(x(t), y(t), z(t))$, the condition

$$\dot\gamma(t) = u_1(t) X_{\gamma(t)} + u_2(t) Y_{\gamma(t)}$$

is equivalent to the system of differential equations

$$\begin{cases} \dot x(t) = u_1(t), \\ \dot y(t) = u_2(t), \\ \dot z(t) = \dfrac{1}{2}\big(x(t)u_2(t) - y(t)u_1(t)\big). \end{cases} \tag{2}$$

We can regard (x, y, z) as the state of the system and (u_1, u_2) as a pair of controls. Equation (2) is a classical example of a *driftless* control system with nonholonomic constraints, and the Heisenberg group provides one of the simplest nontrivial models of such systems.

Geodesics as optimal controls. Once we endow the distribution spanned by X and Y with a sub-Riemannian metric (see below), we can measure the length of horizontal curves. If we choose the metric so that X and Y form an orthonormal frame, then the length of a horizontal curve γ driven by controls (u_1, u_2) is

$$L(\gamma) = \int_0^T \sqrt{u_1(t)^2 + u_2(t)^2}\, dt.$$

Equivalently, one often considers the energy functional

$$E(\gamma) = \frac{1}{2} \int_0^T \big(u_1(t)^2 + u_2(t)^2\big)\, dt.$$

Fixing the endpoints $p(0) = p$ and $p(T) = q$, the problem of finding shortest horizontal curves between p and q in the Heisenberg group can be formulated as a control problem:

$$\text{Minimize } L(\gamma) \text{ (or } E(\gamma)) \text{ subject to } \quad \dot{p}(t) = u_1 X_{p(t)} + u_2 Y_{p(t)}, \quad p(0) = p,\ p(T) = q.$$

The solutions of this optimal control problem are precisely the *sub-Riemannian geodesics* of the Heisenberg group. In this way, the study of geodesics is directly linked to the theory of optimal control and Pontryagin's Maximum Principle.

Sub-Riemannian structure on the Heisenberg group

We now formalize the sub-Riemannian structure suggested above. Since we identify $\mathbb{H} \cong \mathbb{R}^3$ via exponential coordinates (x, y, z), we define a rank-2 distribution $\mathcal{D} \subset T\mathbb{R}^3$ as the kernel of a 1-form.

Definition 1. *Define the 1-form*

$$\omega = dz - \frac{1}{2}\big(x\, dy - y\, dx\big).$$

For each point $p = (x, y, z) \in \mathbb{R}^3$, the horizontal distribution *is*

$$\mathcal{D}_p = \ker(\omega_p) \subset T_p\mathbb{R}^3.$$

A direct computation shows that the vector fields X and Y introduced above satisfy

$$\omega(X) = 0, \qquad \omega(Y) = 0,$$

so they span the distribution:

$$\mathcal{D}_p = \operatorname{span}\{X_p, Y_p\} \quad \text{for all } p \in \mathbb{R}^3.$$

Contact structure. We compute

$$d\omega = -\frac{1}{2}(dx \wedge dy + dy \wedge dx) = -dx \wedge dy,$$

and hence

$$\omega \wedge d\omega = \Big(dz - \frac{1}{2}(x\, dy - y\, dx)\Big) \wedge (-dx \wedge dy) = dz \wedge dx \wedge dy \neq 0$$

at every point of $\mathbb{R}^3$. This nonvanishing of $\omega \wedge d\omega$ means that ω is a *contact form*, and $\mathcal{D} = \ker(\omega)$ is a *contact distribution*. The Heisenberg group is the simplest nontrivial example of a contact manifold.

Definition 2. *We define a sub-Riemannian metric $\langle \cdot, \cdot \rangle_{\mathrm{SR}}$ on $\mathcal{D}$ by declaring*

$$\langle X_p, X_p \rangle_{\mathrm{SR}} = \langle Y_p, Y_p \rangle_{\mathrm{SR}} = 1, \qquad \langle X_p, Y_p \rangle_{\mathrm{SR}} = 0,$$

for all $p \in \mathbb{R}^3$, and extending by linearity to all $v, w \in \mathcal{D}_p$. Equivalently, if $\pi : \mathbb{R}^3 \to \mathbb{R}^2$ is the projection $\pi(x, y, z) = (x, y)$, then for $v, w \in \mathcal{D}_p$ we can write

$$\langle v, w \rangle_{\mathrm{SR}} = \langle \pi_* v, \pi_* w \rangle_{\mathbb{R}^2},$$

where $\langle \cdot, \cdot \rangle_{\mathbb{R}^2}$ is the standard Euclidean inner product on $\mathbb{R}^2$.

With this structure, the triple $(\mathbb{H}, \mathcal{D}, \langle\cdot,\cdot\rangle_{\mathrm{SR}})$ is a sub-Riemannian manifold. A piecewise smooth curve $\gamma : [0,T] \to \mathbb{H}$ is called *horizontal* if $\dot\gamma(t) \in \mathcal{D}_{\gamma(t)}$ for almost every t. In our coordinates this is equivalent to $\dot\gamma(t) = u_1(t)X_{\gamma(t)} + u_2(t)Y_{\gamma(t)}$ for some controls u_1, u_2.

Bracket-generating condition and horizontal connectivity. The Lie brackets of the frame $\{X, Y\}$ satisfy

$$[X, Y] = Z = \partial_z.$$

At each point $p \in \mathbb{R}^3$, the vectors

$$\{X_p, Y_p, [X, Y]_p\} = \{X_p, Y_p, Z_p\}$$

span the entire tangent space $T_p\mathbb{R}^3$. This property is known as the *Hörmander* or *bracket-generating* condition.

By the Chow–Rashevskii theorem, any two points $p, q \in \mathbb{H}$ can be joined by a horizontal curve, i.e., there exists a curve $\gamma : [0,T] \to \mathbb{H}$ such that $\gamma(0) = p$, $\gamma(T) = q$, and $\dot\gamma(t) \in \mathcal{D}_{\gamma(t)}$ for almost every t. Such curves are precisely the trajectories of the control system (2).

Sub-Riemannian distance and geodesics. Given the metric $\langle\cdot,\cdot\rangle_{\mathrm{SR}}$, the length of a horizontal curve γ is

$$L(\gamma) = \int_0^T \big\|\dot\gamma(t)\big\|_{\mathrm{SR}}\, dt = \int_0^T \sqrt{u_1(t)^2 + u_2(t)^2}\, dt,$$

where $\dot\gamma(t) = u_1(t)X_{\gamma(t)} + u_2(t)Y_{\gamma(t)}$. For $p, q \in \mathbb{H}$, the *sub-Riemannian distance* between p and q is defined by

$$d_{\mathrm{SR}}(p, q) = \inf\big\{L(\gamma) : \gamma \text{ horizontal},\ \gamma(0) = p,\ \gamma(T) = q\big\}.$$

Curves that realize this infimum (or are locally minimizing) are the sub-Riemannian geodesics. As explained above, they can be characterized as solutions of an optimal control problem for the Heisenberg control system. This makes the Heisenberg group a fundamental example in both sub-Riemannian geometry and geometric control theory.

2 Finding geodesics in the Heisenberg group

In this section we set up the Hamiltonian formalism for the subRiemannian geodesics on the Heisenberg group $\mathbb{H}$, and explain why, in this case, all length-minimizing geodesics are *normal*. We end with a short argument (following Corollary 4.39 in [1]) that in a 3-dimensional rank–2 subRiemannian manifold all abnormal geodesics are contained in the Martinet set.

2.1 The cometric and Hamiltonian

The subriemannian structure is uniquely determined by its Hamiltonian. For introducing the Hamiltonian for Heisenberg group, we first define a *cometric*

$$\beta_q : T_q^*\mathbb{H} \longrightarrow T_q\mathbb{H}, \qquad q \in \mathbb{H},$$

defined by the following properties (cf. [1]):

1. $\mathrm{im}(\beta_q) = \mathcal{D}_q$ for every $q \in \mathbb{H}$;
2. for every $p \in T_q^*\mathbb{H}$ and every $v \in \mathcal{D}_q$ we have

$$p(v) \;=\; \langle \beta_q(p), v\rangle_{SR},$$

   where $\langle\cdot,\cdot\rangle_{SR}$ is the subRiemannian inner product on $\mathcal{D}_q$ induced by g_{SR}.

In the orthonormal frame $\{X, Y\}$, the cometric is simply

$$\beta_q(p) \;=\; p(X)\,X + p(Y)\,Y, \qquad p \in T_q^*\mathbb{H}.$$

Let us introduce coordinates on $T^*\mathbb{H}$. Write a covector p at (x, y, z) as

$$p = p_x\,dx + p_y\,dy + p_z\,dz.$$

Then

$$P_X := p(X) = p_x - \frac{y}{2}\,p_z, \qquad P_Y := p(Y) = p_y + \frac{x}{2}\,p_z.$$

The subRiemannian Hamiltonian is defined as

$$H(q,p) = \frac{1}{2}\,\|\beta_q(p)\|^2_{SR} = \frac{1}{2}\big(P_X^2 + P_Y^2\big).$$

2.2 Hamiltonian system and normal geodesics

A curve $\zeta(t) = (\gamma(t), p(t)) \in T^*\mathbb{H}$ is an integral curve of the Hamiltonian vector field X_H if it satisfies Hamilton's equations

$$\dot{x}_i = \frac{\partial H}{\partial p_i}, \qquad \dot{p}_i = -\frac{\partial H}{\partial x_i},$$

where $(x_1, x_2, x_3) = (x, y, z)$ and $(p_1, p_2, p_3) = (p_x, p_y, p_z)$. The projection $\gamma(t) = \pi(\zeta(t))$ of such a curve to $\mathbb{H}$ is called a *normal subRiemannian geodesic.*

Writing H in coordinates,

$$H(x, y, z, p_x, p_y, p_z) = \frac{1}{2}\Big((p_x - \tfrac{y}{2}p_z)^2 + (p_y + \tfrac{x}{2}p_z)^2\Big) = \frac{1}{2}\big(P_X^2 + P_Y^2\big),$$

we obtain the Hamiltonian system

$$\begin{aligned}
\dot{x} &= \frac{\partial H}{\partial p_x} = P_X, & \dot{y} &= \frac{\partial H}{\partial p_y} = P_Y,\\
\dot{z} &= \frac{\partial H}{\partial p_z} = -\frac{y}{2}P_X + \frac{x}{2}P_Y, & \dot{p}_x &= -\frac{\partial H}{\partial x} = -\frac{1}{2}p_z\,P_Y,\\
\dot{p}_y &= -\frac{\partial H}{\partial y} = \frac{1}{2}p_z\,P_X, & \dot{p}_z &= -\frac{\partial H}{\partial z} = 0.
\end{aligned}$$

Any solution to this system with $H \equiv \frac{1}{2}$ (corresponding to arc-length parametrization) projects to a horizontal curve $\gamma(t)$ which is a normal geodesic for the Heisenberg subRiemannian metric.

More generally, on an arbitrary subRiemannian manifold $(M, \mathcal{D}, g)$, every length-minimizing geodesic is either:

- *normal*, i.e. it arises as the projection of a solution of the Hamiltonian system above; or
- *abnormal*, i.e. it is minimizing but does not solve the Hamiltonian system associated with the subRiemannian Hamiltonian.

Abnormal geodesics are characterized by the Pontryagin Maximum Principle using covectors that annihilate the distribution $\mathcal{D}$.

In the Heisenberg case, we will see below that the geometry is everywhere *contact*, and hence abnormal geodesics cannot exist.

2.3 Abnormal geodesics and the Martinet set

Let $(M, \mathcal{D}, g)$ be a rank–2 subRiemannian manifold of dimension 3. Assume $\mathcal{D} = \ker\omega$ for some 1-form ω on M. The *Martinet set* is defined as

$$\mathbb{M} = \big\{q \in M \;\big|\; (\omega \wedge d\omega)_q = 0\big\}.$$

At points $q \notin \mathbb{M}$ we have $(\omega \wedge d\omega)_q \neq 0$, and $\mathcal{D}$ is a *contact distribution* near q. At points in $\mathbb{M}$ the growth of the distribution fails to be maximal in a precise sense, and these are exactly the points where abnormal geodesics may occur.

We now give a short argument (cf. Corollary 4.39 in [1]) that all abnormal geodesics are contained in $\mathbb{M}$.

Theorem 1. *Let $(M, \mathcal{D}, g)$ be a rank–2 subRiemannian manifold of dimension 3, with $\mathcal{D} = \ker\omega$ for some smooth 1-form ω. Then any abnormal geodesic of $(M, \mathcal{D}, g)$ is contained in the Martinet set*

$$\mathbb{M} = \{q \in M \mid (\omega \wedge d\omega)_q = 0\}.$$

Equivalently, no abnormal geodesic can pass through a point where $\omega \wedge d\omega \neq 0$ (i.e. through a contact point).

Proof. Let $\gamma : [0,T] \to M$ be an abnormal geodesic. By the Pontryagin Maximum Principle (PMP), there exists a non-zero covector curve $\lambda(t) \in T^*_{\gamma(t)}M$ such that:

(i) $\lambda(t)$ annihilates the distribution:

$$\lambda(t)|_{\mathcal{D}_{\gamma(t)}} = 0 \qquad \forall\, t.$$

(ii) For every horizontal vector field $V \in \Gamma(\mathcal{D})$, one has

$$\lambda(t)\big([u(t), V]\big) = 0,$$

where $u(t) = \dot\gamma(t) \in \mathcal{D}_{\gamma(t)}$ is the horizontal velocity of γ.

Since $\mathcal{D} = \ker\omega$ and $\dim\mathcal{D} = 2$, the annihilator $\mathcal{D}^\perp$ is one-dimensional and generated by ω. Hence for some scalar function $a(t)$, not identically zero,

$$\lambda(t) = a(t)\,\omega_{\gamma(t)}.$$

Now assume for contradiction that there exists t_0 with $q = \gamma(t_0)$ such that $(\omega \wedge d\omega)_q \neq 0$. Thus $\mathcal{D}$ is a contact distribution at q, and therefore $d\omega|_{\mathcal{D}_q}$ is a non-degenerate 2-form. Since $\dim\mathcal{D}_q = 2$, non-degeneracy means:

$$\text{for any } u \in \mathcal{D}_q \setminus \{0\},\ \exists\, V \in \mathcal{D}_q \text{ such that } d\omega_q(u,V) \neq 0.$$

Because γ is not a constant curve, we have $u(t_0) = \dot\gamma(t_0) \neq 0$, and we may choose a horizontal vector field V such that $d\omega_q(u(t_0), V(q)) \neq 0$.

To use the PMP condition, recall Cartan's identity:

$$d\omega(U,V) = U(\omega(V)) - V(\omega(U)) - \omega([U,V]).$$

If $U, V \in \Gamma(\mathcal{D})$, then $\omega(U) = \omega(V) = 0$, and so

$$d\omega(U,V) = -\omega([U,V]).$$

Applying this to $U = u(t)$ and our chosen horizontal V, at t_0 we obtain

$$\omega([u(t_0), V]) = -\,d\omega(u(t_0), V).$$

Using $\lambda = a\,\omega$, condition (ii) of the PMP gives:

$$0 = \lambda(t_0)([u(t_0), V]) = a(t_0)\,\omega([u(t_0), V]) = -\,a(t_0)\,d\omega(u(t_0), V).$$

By construction, $d\omega(u(t_0), V) \neq 0$, hence

$$a(t_0) = 0.$$

But this is impossible: the PMP forbids the abnormal multiplier $\lambda(t)$ from vanishing at any time, since vanishing at a single time would imply $\lambda \equiv 0$, contradicting the definition of an abnormal extremal. Thus our assumption that $\gamma(t_0)$ lies at a contact point must be false.

Consequently, an abnormal geodesic cannot pass through any point where $\omega \wedge d\omega \neq 0$. Hence every abnormal geodesic is contained entirely in the Martinet set

$$\mathbb{M} = \{\, q \in M : (\omega \wedge d\omega)_q = 0 \,\}.$$

□

2.4 Absence of abnormal geodesics in the Heisenberg group

For the Heisenberg group $\mathbb{H}$ with $\omega = dz + \frac{1}{2}(x\,dy - y\,dx)$, one checks directly that

$$(\omega \wedge d\omega)_q \neq 0 \quad \text{for all } q \in \mathbb{H}.$$

In particular, the Martinet set is empty:

$$\mathbb{M} = \varnothing.$$

By the theorem above, this implies that $\mathbb{H}$ does not admit abnormal geodesics. Hence every length-minimizing geodesic in the Heisenberg group is normal, and is obtained as the projection of a solution of the Hamiltonian system associated with

$$H = \frac{1}{2}\big(P_X^2 + P_Y^2\big) = \frac{1}{2}\Big((p_x - \tfrac{y}{2}p_z)^2 + (p_y + \tfrac{x}{2}p_z)^2\Big).$$

2.5 Explicit integration of the Hamiltonian system

Since the Heisenberg group is a contact sub-Riemannian manifold, all geodesics are normal and therefore arise as projections of integral curves of the Hamiltonian system associated with

$$H(q,p) = \tfrac{1}{2}\big(P_X^2 + P_Y^2\big),$$

where

$$P_X = p(X) = p_x - \tfrac{y}{2}p_z, \qquad P_Y = p(Y) = p_y + \tfrac{x}{2}p_z, \qquad P_Z = p(Z) = p_z.$$

Note that P_Z does not contribute to the cometric but enters the Hamiltonian dynamics through the Lie bracket $[X,Y] = Z$.

We adopt the Poisson–Lie convention

$$\{P_X, P_Y\} = -P_{[X,Y]} = -P_Z, \qquad \{P_X, P_Z\} = \{P_Y, P_Z\} = 0.$$

2.5.1 Hamiltonian equations

Hamilton's equations $\dot{f} = \{f, H\}$ together with the above Poisson relations give:

1. $\dot{x} = \dfrac{\partial H}{\partial p_x} = P_X,$
2. $\dot{y} = \dfrac{\partial H}{\partial p_y} = P_Y,$
3. $\dot{z} = \dfrac{\partial H}{\partial p_z} = -\dfrac{1}{2}yP_X + \dfrac{1}{2}xP_Y,$
4. $\dot{P}_X = \{P_X, H\} = \tfrac{1}{2}\{P_X, P_Y^2\} = \{P_X, P_Y\}P_Y = -P_Z P_Y,$
5. $\dot{P}_Y = \{P_Y, H\} = -\tfrac{1}{2}\{P_X^2, P_Y\} = -\{P_X, P_Y\}P_X = P_Z P_X,$
6. $\dot{P}_Z = \{P_Z, H\} = 0.$

Thus $P_Z = p_z$ is constant along the geodesic.

2.5.2 Solution for $P_X(t)$ and $P_Y(t)$

For $p_z \neq 0$, equations (4) and (5) can be written in matrix form:

$$\begin{pmatrix} \dot{P}_X \\ \dot{P}_Y \end{pmatrix} = \begin{pmatrix} 0 & -p_z \\ p_z & 0 \end{pmatrix} \begin{pmatrix} P_X \\ P_Y \end{pmatrix}.$$

Thus

$$\begin{pmatrix} P_X(t) \\ P_Y(t) \end{pmatrix} = \begin{pmatrix} \cos(p_z t) & -\sin(p_z t) \\ \sin(p_z t) & \cos(p_z t) \end{pmatrix} \begin{pmatrix} P_X(0) \\ P_Y(0) \end{pmatrix}.$$

To obtain explicit expressions, let us choose the initial horizontal momentum in polar form:

$$\begin{pmatrix} P_X(0) \\ P_Y(0) \end{pmatrix} = a \begin{pmatrix} -\sin\theta \\ \cos\theta \end{pmatrix}, \qquad a \geq 0.$$

Then

$$\begin{pmatrix} P_X(t) \\ P_Y(t) \end{pmatrix} = \begin{pmatrix} -a\sin(p_z t + \theta) \\ a\cos(p_z t + \theta) \end{pmatrix}.$$

2.5.3 Horizontal components $x(t)$ and $y(t)$

Using $\dot{x} = P_X$ and $\dot{y} = P_Y$, and assuming the geodesic starts at the origin, we obtain:

$$\begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = \frac{a}{p_z} \begin{pmatrix} \cos(p_z t + \theta) - \cos\theta \\ \sin(p_z t + \theta) - \sin\theta \end{pmatrix}.$$

Thus the projection of the geodesic onto the (x, y)–plane is a circle of radius $\frac{a}{p_z}$ centered at $\left(\frac{a}{p_z}\cos\theta, \frac{a}{p_z}\sin\theta\right)$.

2.5.4 Vertical component $z(t)$

The third Hamilton equation gives

$$\dot{z}(t) = -\tfrac{1}{2} y(t) P_X(t) + \tfrac{1}{2} x(t) P_Y(t) = \tfrac{1}{2}\{x\dot{y} - y\dot{x}\},$$

which is the signed area swept by the horizontal curve. Substituting $x(t), y(t), P_X(t), P_Y(t)$ and integrating yields:

$$z(t) = \frac{a^2}{2p_z^2}\big(p_z t - \sin(p_z t)\big).$$

2.5.5 The case $p_z = 0$

If $p_z = 0$, then P_X and P_Y are constant, and the system reduces to

$$\dot{x} = P_X(0), \qquad \dot{y} = P_Y(0), \qquad \dot{z} = 0.$$

Hence the geodesics are horizontal straight lines:

$$(x(t), y(t), z(t)) = \big(tP_X(0),\, tP_Y(0),\, 0\big).$$

2.5.6 Summary

Normal geodesics in the Heisenberg group consist of:

- horizontal straight lines when $p_z = 0$;
- lifts of horizontal circles when $p_z \neq 0$, with height determined by

$$z(t) = \frac{a^2}{2p_z^2}\big(p_z t - \sin(p_z t)\big).$$

This gives a complete classification of all sub-Riemannian geodesics in $\mathbb{H}$.

3 Some Features of Geodesics in $\mathbb{H}^3$

Let $\gamma(t)$ be a sub-Riemannian geodesic in the Heisenberg group $\mathbb{H} = \mathbb{R}^3$ (equipped with its standard horizontal distribution) starting at the origin $\mathcal{O} = (0, 0, 0)$. If the initial covector is

$$P_0 = \big(P_X(0), P_Y(0), p_z\big)^T, \qquad p_z \neq 0,$$

then the corresponding geodesic has the explicit form

$$\gamma(t) = \left(\frac{a}{p_z}\big(\cos(p_z t + \theta) - \cos\theta\big),\ \frac{a}{p_z}\big(\sin(p_z t + \theta) - \sin\theta\big),\ \frac{a}{2p_z}\big(p_z t - \sin(p_z t)\big)\right),$$

where $a > 0$ is the horizontal speed and $\theta \in [0, 2\pi)$ encodes the direction of the initial horizontal momentum. The projection of γ to the (x, y)-plane is a circle of radius

$$R = \frac{a}{p_z},$$

centered at

$$\vec{R} = \frac{a}{p_z}(\cos\theta, \sin\theta).$$

Thus, for $p_z \neq 0$, the geodesic is a helix around this circle.

If $p_z = 0$, then the canonical equations show that γ is a straight horizontal line,

$$\gamma(t) = (-at\sin\theta,\ at\cos\theta,\ 0).$$

These geodesics are globally minimizing.

Metric Balls in $\mathbb{H}$

Let $\vec{Q} \in \partial\mathcal{B}_R = \{q \in \mathbb{H} : d_{SR}(0,q) = R\}$. If $\gamma : [0,T] \to \mathbb{H}$ is a minimizing geodesic from 0 to $\vec{Q}$, then its horizontal speed is constant:

$$\|\dot{\gamma}\|_{SR} = a > 0, \qquad R = aT.$$

Thus $T = R/a$, and substituting in the geodesic formula gives

$$\vec{Q} = \gamma\left(\frac{R}{a}\right) = \frac{a}{p_z}\Big(\cos(\lambda+\theta) - \cos\theta,\ \sin(\lambda+\theta) - \sin\theta,\ \frac{1}{2}\big(\lambda - \sin\lambda\big)\Big),$$

where

$$\lambda = \frac{p_z R}{a}.$$

For fixed $R > 0$ and varying $\theta \in [0, 2\pi]$, these points describe the entire sphere $\partial\mathcal{B}_R$, showing that the Carnot–Carathéodory distance is cylindrically symmetric.

Define

$$r(R,\lambda) = \sqrt{x^2+y^2} = \frac{2R}{\lambda}\sin\left(\frac{\lambda}{2}\right), \qquad z(R,\lambda) = \frac{1}{2}\left(\frac{R}{\lambda}\right)^2\big(\lambda - \sin\lambda\big).$$

Rotating the curve $(r(R,\lambda),\ z(R,\lambda))$ around the z-axis yields the surface of the sub-Riemannian ball of radius R. Its well-known shape resembles that of an *apple*.

Geodesics and the Queen Dido Problem

To better understand how geodesics behave in the Heisenberg group, it is helpful to begin with a classical variational problem dating back to antiquity: the *Queen Dido problem*. This problem provides a clean two-dimensional picture that turns out to mirror the three-dimensional geometry of horizontal curves in $\mathbb{H}^1$.

The classical Queen Dido problem

According to legend, Queen Dido was granted as much land as she could enclose with a fixed length of rope. This leads to the classical question:

> *Among all planar curves of fixed length joining two points, which curve encloses the greatest possible area?*

The famous (and surprisingly rigid) answer is:

> *The maximizing curve is an arc of a circle.*

Together with the straight segment connecting the endpoints, the circular arc encloses the maximal possible area. This is the geometric origin of the circle in many classical isoperimetric results.

Why a minimization version also appears

In modern variational calculus and in control theory, it is often more convenient to place the constraint on the *area* instead of the length. This simply reverses the optimization:

Fix the enclosed area and minimize the length.

The two formulations are equivalent: fixing the length and maximizing the area leads to the same extremals as fixing the area and minimizing the length. The latter matches the structure of sub-Riemannian problems, in which one typically minimizes length under geometric constraints.

This motivates the following formulation:

Problem 1 (Queen Dido, variational form). Fix $A \in \mathbb{R}$ and $q \in \mathbb{R}^2$. Among all smooth curves $\xi : [0, T] \to \mathbb{R}^2$ satisfying

$$\xi(0) = 0, \qquad \xi(T) = q, \qquad \int_\xi \tfrac{1}{2}(x\,dy - y\,dx) = A,$$

find one that minimizes the length $l(\xi)$.

As in the classical version, the minimizing curve is an arc of a circle (orientation being the only freedom).

Lifting the solution into the Heisenberg group

Now we describe why this two-dimensional problem encodes the geometry of Heisenberg geodesics.

Given a minimizing planar curve $\xi(t) = (x(t), y(t))$, define its *lift* to the Heisenberg group by

$$\gamma(t) = \big(x(t), y(t), z(t)\big), \qquad z(t) = \int_0^t \tfrac{1}{2}\big(x(s)\dot{y}(s) - y(s)\dot{x}(s)\big)\,ds.$$

Recall that the horizontal distribution of $\mathbb{H}^1$ is

$$\ker\omega, \qquad \omega = dz - \tfrac{1}{2}(x\,dy - y\,dx).$$

A direct computation shows that this lifted curve satisfies

$$\dot{\gamma}(t) \in \ker\omega \qquad \forall\, t,$$

so γ is automatically horizontal.

Moreover, the sub-Riemannian length of γ depends only on its (x, y) projection. Thus minimizing the length of ξ in the plane is equivalent to minimizing the length of γ in the Heisenberg group.

This gives rise to:

Problem 2 (Sub-Riemannian Lifting). Find $\gamma : [0, T] \to \mathbb{H}^1$ with

$$\gamma(0) = 0, \qquad \gamma(T) = (q, A), \qquad \dot{\gamma} \in \ker\omega,$$

that minimizes the sub-Riemannian length $l(\gamma)$.

Thus:

A Heisenberg geodesic is length-minimizing if and only if its projection to the plane solves the Queen Dido problem.

Consequences for Heisenberg geodesics

This correspondence provides an intuitive picture of how Heisenberg geodesics behave.

A geodesic in $\mathbb{H}^1$ with $p_t \neq 0$ projects to an arc of a circle. Since the Dido problem completely characterizes which circular arcs are minimizing, we deduce:

> *A helical geodesic in the Heisenberg group is length-minimizing exactly until the moment when its planar projection finishes one full turn of its circle.*

After that time, the circular projection fails to be the minimizing arc, and the geodesic stops being optimal. Geometrically, this loss of minimality occurs when the projected circular arc returns to intersect its center (the z-axis) for the second time.

In contrast:

> *Straight-line geodesics (those with $p_t = 0$) remain minimizing for all time.*

This gives a simple geometric rule: helical Heisenberg geodesics minimize only for a finite interval, while horizontal straight lines minimize globally.

4 Geodesics in $\mathbb{H}^n$

The Heisenberg group generalizes naturally to higher dimensions. For $n \geq 1$ we identify

$$\mathbb{H}^n \cong \mathbb{C}^n \times \mathbb{R} \cong \mathbb{R}^{2n+1}, \qquad (z,t) = (x_1, \dots, x_n, y_1, \dots, y_n, t),$$

with group operation

$$(z,t) * (z',t') = (z + z',\, t + t' + \tfrac{1}{2}\operatorname{Im} \textstyle\sum_{j=1}^n z_j \overline{z'_j}),$$

or, in real coordinates,

$$(x,y,t) * (x',y',t') = \big(x + x',\, y + y',\, t + t' + \tfrac{1}{2}\sum_{j=1}^n (x'_j y_j - x_j y'_j)\big).$$

Horizontal distribution and contact structure

The standard horizontal distribution is

$$\mathcal{D} = \ker(\Omega), \qquad \Omega = dt - \tfrac{1}{2}\sum_{j=1}^n (x_j\, dy_j - y_j\, dx_j).$$

A direct computation yields

$$\Omega \wedge d\Omega = \sum_{j=1}^n dt \wedge dy_j \wedge dx_j + \Theta,$$

where Θ is a nonzero $(2n+1)$–form. Thus

$$(\Omega \wedge d\Omega)_p \neq 0 \qquad \forall\, p \in \mathbb{H}^n,$$

so $\mathcal{D}$ is a contact distribution everywhere. Consequently, $\mathbb{H}^n$ admits only *normal* geodesics (no abnormal extremals).

Sub-Riemannian metric

For $v, w \in T_q\mathbb{H}^n$, we define the sub-Riemannian metric by

$$\langle v, w \rangle_{\mathrm{SR}} = \langle \pi(v), \pi(w) \rangle = \sum_{j=1}^n (v_{x_j} w_{x_j} + v_{y_j} w_{y_j}),$$

where the projection

$$\pi : \mathbb{R}^{2n+1} \to \mathbb{R}^{2n}$$

forgets the t–component.

The associated left-invariant frame at $p = (x, y, t)$ is

$$X_j = \partial_{x_j} - \tfrac{1}{2} y_j \partial_t, \qquad Y_j = \partial_{y_j} + \tfrac{1}{2} x_j \partial_t, \qquad Z = \partial_t,$$

with Lie brackets

$$[X_j, Y_j] = Z, \qquad [X_j, X_k] = [Y_j, Y_k] = [X_j, Y_k] = 0 \ (j \neq k).$$

Hamiltonian system

Let $p = (p_{x_1}, \ldots, p_{y_n}, p_t)$ be a covector. Define the momenta

$$P_{X_j} = p_{x_j} - \tfrac{1}{2} y_j p_t, \qquad P_{Y_j} = p_{y_j} + \tfrac{1}{2} x_j p_t.$$

The sub-Riemannian Hamiltonian becomes

$$H = \tfrac{1}{2} \sum_{j=1}^{n} \big(P_{X_j}^2 + P_{Y_j}^2 \big).$$

From Hamilton's equations we obtain:

$$\dot{x}_j = \frac{\partial H}{\partial p_{x_j}} = P_{X_j}, \qquad \dot{y}_j = \frac{\partial H}{\partial p_{y_j}} = P_{Y_j},$$

$$\dot{t} = \frac{\partial H}{\partial p_t} = \sum_{j=1}^{n} \left(-\tfrac{1}{2} y_j P_{X_j} + \tfrac{1}{2} x_j P_{Y_j} \right),$$

and using Poisson brackets equivalent to the Lie algebra relations,

$$\dot{P}_{X_j} = -p_t \, P_{Y_j}, \qquad \dot{P}_{Y_j} = p_t \, P_{X_j}, \qquad \dot{p}_t = 0.$$

Hence the vertical momentum p_t is a constant, denoted λ.

Geodesics

For each j, the pair (P_{X_j}, P_{Y_j}) satisfies the harmonic system

$$\dot{P}_{X_j} = -\lambda P_{Y_j}, \qquad \dot{P}_{Y_j} = \lambda P_{X_j},$$

whose solution is a rotation in the (P_{X_j}, P_{Y_j})–plane:

$$\begin{pmatrix} P_{X_j}(t) \\ P_{Y_j}(t) \end{pmatrix} = \begin{pmatrix} \cos(\lambda t) & -\sin(\lambda t) \\ \sin(\lambda t) & \cos(\lambda t) \end{pmatrix} \begin{pmatrix} P_{X_j}(0) \\ P_{Y_j}(0) \end{pmatrix}.$$

Integrating $\dot{x}_j = P_{X_j}$ and $\dot{y}_j = P_{Y_j}$ gives

$$\begin{pmatrix} x_j(t) \\ y_j(t) \end{pmatrix} = \int_0^t \begin{pmatrix} \cos(\lambda s) & -\sin(\lambda s) \\ \sin(\lambda s) & \cos(\lambda s) \end{pmatrix} \begin{pmatrix} P_{X_j}(0) \\ P_{Y_j}(0) \end{pmatrix} ds.$$

Finally,

$$t(t) = \int_0^t \sum_{j=1}^{n} \big(-\tfrac{1}{2} y_j(s) P_{X_j}(s) + \tfrac{1}{2} x_j(s) P_{Y_j}(s) \big) \, ds.$$

Thus a geodesic starting at the origin decomposes as the product

$$\gamma(t) = \big(Q_{x_1 y_1}(t), \ldots, Q_{x_n y_n}(t), t(t) \big),$$

where each $Q_{x_j y_j}(t)$ is the planar curve associated to the jth pair (X_j, Y_j). Depending on the initial momenta $P_{X_j}(0), P_{Y_j}(0)$, each component is either: - a straight line (when $\lambda = 0$), or - a helix-like projection generated by uniform rotation (when $\lambda \neq 0$).

Hence geodesics in $\mathbb{H}^n$ are obtained as products of the classical $\mathbb{H}^1$ geodesics in each (x_j, y_j)–plane, together with the vertical component determined by the Hamiltonian flow.

References

[1] A. Agrachev, D. Barilari, U. Boscain, *A Comprehensive Introduction to Sub-Riemannian Geometry*, Cambridge Studies in Advanced Mathematics, Cambridge Univ. Press.

[2] R. Montgomery, *A Tour of Subriemannian Geometries, Their Geodesics and Applications*, Mathematical Surveys and Monographs, vol. 91, AMS.