Departments of Mathematics
Montana State University
Prof. Kevin Wildrick

An introduction to non-smooth analysis and geometry

Lecture 12: The Heisenberg group

1. An introduction to the Heisenberg group

The sub-Riemannian Heisenberg group is an excellent example of a non-trivial metric measure space that supports a Poincaré inequality and has an easy to understand Cheeger structure. It is a very interesting space for a variety of reasons, but here is one to get whet your appetite: the sub-Riemannian Heisenber group has a two-dimensional Cheeger structure, is homeomorphic to $\mathbb{R}^3$, and is Ahlfors 4-regular. It shares many properties with Euclidean space, but it does not bi-Lipschitzly embed into Hilbert space.

Let us begin by describing the Heisenberg group itself. The most familiar way to describe the Heisenberg group is as the group of upper triangular 3×3 real matrices with 1's on the diagonal, with respect to matrix multiplication:

$$\begin{pmatrix} 1 & a & c \\ 0 & 1 & b \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & a' & c' \\ 0 & 1 & b' \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & a+a' & c+c'+ba' \\ 0 & 1 & b+b' \\ 0 & 0 & 1 \end{pmatrix}.$$

In otherwords, the group has underlying set $\mathbb{R}^3$ and multiplication

$$(a,b,c)(a',b',c') = (a+a', b+b', c+c+ba').$$

This is clearly a non-commutative group operation, but it lacks a symmetry which will be convenient for us.

To that end, we denote by $\mathbb{H}$ the group $(\mathbb{R}^3, *)$, where

$$(x,y,t) * (x',y',t') = (x+x', y+y', t+t'-2(xy'-yx')).$$

We may also consider the underlying set as $\mathbb{C} \times \mathbb{R}$. Identifying $z = x+iy$, the group operation becomes

$$(z,t) * (z',t') = (z+z', t+t'+2\,\mathrm{im}(z\bar{z})).$$

The factor accounting for the non-commutativity of the group operation is the term $\mathrm{im}(z\bar{z})$. This term can be interpreted as the *signed area* of the parallelogram spanned by z and z', i.e., the determinant of the linear mapping defined by

$$1 \mapsto z \text{ and } i \mapsto z'.$$

Note that the identity element is $(0,0,0)$ and

$$(x,y,t)^{-1} = (-x,-y,-t),$$

just as in the usual group structure on $\mathbb{R}^3$.

The two groups described above are isomorphic via the isomorphism

$$x = a, y = b, t = 4c - 2ab.$$

One of the first things we notice about the group operation in $\mathbb{H}$ is that the formulas describing it are smooth. This fact makes the Heisenberg group an example of a Lie group. A *Lie group* is a smooth manifold M endowed with a group operation $*$ so that the mapping $*\colon M \times M \to M$ defined

$$*(g,h) = g^{-1} * h$$

is smooth. Note that this implies that the mapping $g \mapsto g^{-1}$ is smooth, and for each $g_0 \in M$, the mappings $g \mapsto g_0 * g$ are smooth.

Noticing this, the first thing that comes to mind to try is to differentiate some of these smooth mappings we have hanging around!

Let's pick a point $g_0 = (x_0, y_0, t_0) \in \mathbb{H}$, and consider the mapping $L_{g_0} \colon \mathbb{R}^3 \to \mathbb{R}^3$ defined by

$$L_{g_0}(h) = g_0 * h.$$

Its derivative at a point $p \in \mathbb{H}$ will be a linear mapping from $T_p(\mathbb{R}^3)$ to $T_{g*p}(\mathbb{R}^3)$:

$$D(L_{g_0})_p \colon T_p(\mathbb{R}^3) \to T_{g_0 * p}(\mathbb{R}^3).$$

Let's examine this mapping when $p = \mathbf{0}$. To figure out what it does, we need to evaluate it on a basis of $T_{\mathbf{0}}(\mathbb{R}^3)$. What are elements of $T_{\mathbf{0}}(\mathbb{R}^3)$? They are equivalence classes of smooth curves $\gamma \colon (-\epsilon, \epsilon) \to \mathbb{R}^3$ satisfying $\gamma(0) = \mathbf{0} \in \mathbb{R}^3$, where two such curves γ and β are equivalent if $\gamma'(0) = \beta'(0)$.

With this perspective, a basis for $T_{\mathbf{0}}(\mathbb{R}^3)$ is given by the equivalence classes of the curves

$$\gamma_x(s) = (s, 0, 0), \;\; \gamma_y(s) = (0, s, 0), \;\; \gamma_t(s) = (0, 0, s).$$

The corresponding vectors in $T_{\mathbf{0}}(\mathbb{R}^3)$ are usually denoted by

$$\begin{aligned} \frac{\partial}{\partial x} &= (1, 0, 0) = [\gamma_x'(0)], \\ \frac{\partial}{\partial y} &= (0, 1, 0) = [\gamma_y'(0)], \\ \frac{\partial}{\partial t} &= (0, 0, 1) = [\gamma_t'(0)]. \end{aligned}$$

To see how $D(L_{g_0})_{\mathbf{0}}$ behaves, we push forward these curves with the mapping L_{g_0}, and then differentiate.

$$\begin{aligned} L_{g_0} \circ \gamma_x(s) &= (x_0, y_0, t_0) * (s, 0, 0) = (x_0 + s, y_0, t_0 + 2sy_0), \\ L_{g_0} \circ \gamma_y(s) &= (x_0, y_0, t_0) * (0, s, 0) = (x_0, y_0 + s, t_0 - 2sx_0), \\ L_{g_0} \circ \gamma_t(s) &= (x_0, y_0, t_0) * (0, 0, s) = (x_0, y_0, t_0 + s). \end{aligned}$$

Hence,

$$\begin{aligned} (L_{g_0} \circ \gamma_x)'(0) &= (1, 0, 2y_0) = \frac{\partial}{\partial x} + 2y_0 \frac{\partial}{\partial t}, \\ (L_{g_0} \circ \gamma_y)'(0) &= (0, 1, -2x_0) = \frac{\partial}{\partial y} - 2x_0 \frac{\partial}{\partial t}, \\ (L_{g_0} \circ \gamma_t)'(0) &= (0, 0, 1) = \frac{\partial}{\partial t}. \end{aligned}$$

Another way to look at this calculation is to consider the full differential of the mapping:

$$L_{g_0}(x, y, t) = (x_0 + x, y_0 + y, t_0 + t - 2(x_0 y - y_0 x)).$$

Differentiating with respect to (x, y, t) and evaluating at $(x, y, t) = \mathbf{0}$ yields the matrix

$$D(L_{g_0})_{\mathbf{0}} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 2y_0 & -2x_0 & 1 \end{pmatrix}.$$

Then

$$D(L_{g_0})_{\mathbf{0}}\begin{pmatrix}1\\0\\0\end{pmatrix} = (1, 0, 2y_0),$$

agreeing with the above computation. Notice that the above matrix would have been the same if we evalutated at *any* value of (x, y, t). Since it is non-singular, this implies that L_{g_0} is a diffeomorphism, and its differential maps bases to bases.

Proposition 1.1. *Let $g = (x, y, t) \in \mathbb{R}^3$. The vector fields*

$$X(g) := \frac{\partial}{\partial x} + 2y\frac{\partial}{\partial t},$$
$$Y(g) := \frac{\partial}{\partial x} - 2x\frac{\partial}{\partial t},$$
$$T(g) := \frac{\partial}{\partial t}$$

on $\mathbb{R}^3$ are invariant *under the action of $\mathbb{H}$, i.e., for any $g_0 \in \mathbb{H}$*

$$D(L_{g_0})_g X(g) = X(L_{g_0}(g)),$$
$$D(L_{g_0})_g Y(g) = Y(L_{g_0}(g)),$$
$$D(L_{g_0})_g T(g) = T(L_{g_0}(g)).$$

Moreover, they form a basis of $T_g(\mathbb{R}^3)$.

Proof. Given a vector $V(\mathbf{0}) \in T_{\mathbf{0}}\mathbb{R}^3$, we may *create* an invariant vector field V by defining

$$V(g) := D(L_g)_{\mathbf{0}}(V(\mathbf{0})).$$

To see this, note that then, for each $g_0 \in \mathbb{H}$, the chain rule yields

$$D(L_{g_0})_g(V(g)) = D(L_{g_0})_g D(L_g)_{\mathbf{0}} V(e) = D(L_{g_0} \circ L_g)_{\mathbf{0}} V(e) = D(L_{g_0 * g})_{\mathbf{0}} V(e) = V(L_{g_0}(g)).$$

This is exactly how we defined X, Y, and T. ☐

The proof shows that a left-invariant vector field is defined by its value at the identity.

Notice that $D(L_{g_0})_g$ in fact has determinant 1. This implies that L_{g_0} is a *volume preserving* transformation; if Ω is an open set, the

$$\mathcal{L}^3(L_{g_0}(\Omega)) = \int_\Omega \det D(L_{g_0})\, d\mathcal{L}^3 = \mathcal{L}^3(\Omega).$$

This quickly implies that for any measurable set $E \subseteq$, the Lebesgue measure of E is the same as the Lebesgue measure of the image of E under any left translation. Thus we have shown

Proposition 1.2. *The Lebesgue measure $\mathcal{L}^3$ is a left-invariant measure on $\mathbb{H}$.*

One can in fact show that $\mathcal{L}^3$ is also right-invariant, and it is, up to scaling, the unique such measure.

In quantum mechanics, *position* and *momentum* correspond non-commutating self-adjoint operators on a complex Hilbert space. The Heisenberg group was first considered in this context - the vector fields X and Y above provide the simplest possible model of this behavior (this can be made precise).

Definition 1.3. Given vector fields V and W on $\mathbb{R}^3$, their *Lie bracket* or *commutator* or *Lie derivative* is defined by

$$[V, W] := VW - WV.$$

The above definition requires some discussion about *what exactly vector fields are.* If one takes the view that vector fields are functions $\mathbb{R}^3 \to \mathbb{R}^3$; the above definition *cannot be properly understood.* By VW, we *do not mean composition of functions.* We can better understand this by identifying vector fields with *derivations* or as *derivatives of curves*, or better yet as *differential equations.*

Given a vector field $V\colon \mathbb{R}^3 \to \mathbb{R}^3$, we can define an ordinary differential equation as follows. A solution to the differential equation is a path $\gamma\colon (-\epsilon, \epsilon) \to \mathbb{R}^3$ that satisfies

$$\gamma'(s) = V(\gamma(s)).$$

If V is smooth enough (and ours always will be), then solutions to this differential equation exist and are unique after specification of $\gamma(0)$.

Let us try understand $[V, W]$ in this way. Let $g \in \mathbb{R}^3$ and let $s_0 > 0$. Let γ_1 solve

$$\gamma_1'(s) = V(\gamma_1(s)),\ \gamma_1(0) = g.$$

Let γ_2 solve

$$\gamma_2'(s) = W(\gamma_2(s)),\ \gamma_2(0) = \gamma_1(s_0).$$

Let γ_3 solve

$$\gamma_3'(s) = -V(\gamma_3(s)),\ \gamma_3(0) = \gamma_2(s_0).$$

Let γ_4 solve

$$\gamma_4'(s) = -W(\gamma_4(s)),\ \gamma_4(0) = \gamma_3(s_0).$$

Now define $g(s_0)$ to be the end point of the concatenation of $\gamma_1, \gamma_2, \gamma_3$ and γ_4.

We define

$$[V, W](g) = \frac{1}{2}\frac{d^2}{dt^2} g(s)\bigg|_{s=0}.$$

The idea is the following: we start at g, flow with V for time s, flow with W for time s, flow backwards in time with V for time s, and then flow backwards in time with W for time s. Now, we look at how much we now differ from g and take the second derivative.

Let's compute $[X, Y](e)$.

First, we flow with X from e for time s_0. The first coordinate of e will grow with speed 1, the second won't change, and the third will grow with speed 2 times the second coordinate. But, the second coordinate is stuck at 0, so

$$\gamma_1(s) = (s, 0, 0),\ \text{and } \gamma_1(s_0) = (s_0, 0, 0)$$

Now we flow with Y from $(s, 0, 0)$ for time s. The first coordinate doesn't change, the second grows with speed 1, and the third will grow with speed -2 times the first coordinate (which is constant):

$$\gamma_2(s) = (s_0, s, -2s_0 s).\ \text{and } \gamma_2(s_0) = (s_0, s_0, -2s_0^2),$$

Now reverse

$$\gamma_3(s) = (s_0 - s, s_0, -2s_0^2 - 2s_0 s) \text{ and } \gamma_3(s_0) = (0, s_0, -4s_0^2).$$

Finally

$$\gamma_4(s) = (0, s_0 - s, -6s_0^2),\ \text{and } \gamma_4(s_0) = (0, 0, -4s_0^2).$$

Thus

$$[X, Y](e) = -4T.$$

It's a good exercise to show that if V and W are left-invariant vector fields, then so is their commutator $[V, W]$. It's also fun to check that if we treat

$$\frac{\partial}{\partial x} = (1, 0, 0), \frac{\partial}{\partial y} = (0, 1, 0), \frac{\partial}{\partial t} = (0, 0, 1)$$

as differential operators on the space of smooth functions on $\mathbb{R}^3$, e.g.

$$X(x_0, y_0, z_0)f = \frac{\partial f}{\partial x}(x_0, y_0, z_0) + 2y_0 \frac{\partial f}{\partial y}(x_0, y_0, z_0),$$

then the commutator can be viewed as iterated differentiation (this is because the tangent bundle has an alternate description as the space of derivations).

It's not hard to see that $[X, T] = 0 = [Y, T]$ using a similar method of computation. Thus, we have seen that X and Y do not commute, but instead their commutator is a vector field that commutes with all others! This is the original motivation for considering the Heisenberg group coming from quantum mechanics and is at the heart of the Heisenberg uncertainty principle.

Definition 1.4. The *Lie algebra* of the Heisenberg group is the three-dimensional real vector space generated by X, Y, and T, and equipped with the operation $[\cdot, \cdot]$.

Representations of the Heisenberg algebra in the Hilbert space of "states" play an important role in quantum mechanics.

2. The Carnot-Carathéodory metric on the Heisenberg group

The fact that the vector fields X and Y and their commutator $[X, Y] = -4T$ generate the entire Heisenberg algebra makes the situation very different from the standard orthonormal frame for the tangent bundle: the vector fields $\frac{\partial}{\partial x}$ and $\frac{\partial}{\partial y}$ commute, and so they do not generate (in the sense of Lie algebras) the entire tangent bundle.

To put this in context, a related important result from differential geometry is Frobenius' theorem. Let M be a smooth manifold, and let $\mathcal{B} \subseteq TM$ be a subbundle of the tangent bundle. Then $\mathcal{B}$ arises from a foliation of M if and only $\mathcal{B}$ is closed under taking commutators.

We can see this in action here: the subbundle generated by $\frac{\partial}{\partial x}$ and $\frac{\partial}{\partial y}$ arises from the foliation of $\mathbb{R}^3$ by planes parallel to the $x - y$ plane. However, try as you might, you will never find a foliation of $\mathbb{R}^3$ by smooth surfaces whose tangent planes lie in the subbundle generated by X and Y.

Definition 2.1. The *horizontal tangent bundle* $HT\mathbb{H}$ of $\mathbb{H}$ is the subbundle of $\mathbb{R}^3$ generated by the vector fields X and Y.

Of course, since the vector fields X and Y are left-invariant, the horizontal tangent bundle is also left-invariant:

$$(DL_g)_p HT_p\mathbb{H} = HT_{g*p}\mathbb{H}.$$

There is a general result, called the Chow-Rashevskii theorem, that is a sort of converse to Frobenius' theorem: if a subbundle generates (by taking commutators) the entire tangent bundle, then any two points of the manifold can be connected by paths whose tangents lie in the subbundle.

Definition 2.2. Let $\gamma\colon [0, 1] \to \mathbb{R}^3$ be an absolutely continuous path. We say that γ is *(Heisenberg) horizontal* if for almost every $s \in [0, 1]$, the tangent $\gamma'(s)$ lies in the horizontal tangent space $HT\mathbb{H}_{\gamma(s)}$.

Notice that the notion of a horizontal path is left-invariant under the action of the Heisenberg group: if $g \in \mathbb{H}$ and $\gamma\colon [0, 1] \to \mathbb{H}$ horizontal, then for almost every $s \in [0, 1]$,

$$(L_g \circ \gamma)'(s) = (DL_g)_{\gamma(s)}\gamma'(s) \in (DL_g)_{\gamma(s)} HT_{\gamma(s)}\mathbb{H} = HT_{L_g \circ \gamma(s)}\mathbb{H}.$$

When we calculated that $[X, Y] = -4T$, we did so by creating a path that connected the origin to a point on the T axis only by flowing with multiples of X and Y. We can use a similar construction to show that

Theorem 2.3. *Any two points of* $\mathbb{R}^3$ *can be connected by a horizontal path.*

Proof. It suffices to connect any point to the origin. First, note that for any point $(x_0, y_0, 0) \in \mathbb{H}$, the path

$$\gamma(s) = s(x_0, y_0, 0)$$

is a horizontal path connecting it to the origin. Combining this path with a flow along X, Y, $-X$, and $-Y$ (or its negation) as before gives the result. $\square$

For such paths, there is a preferred notion of length that is adapted to this structure. Namely, let $\gamma\colon [0,1] \to \mathbb{R}^3$ be a horizontal path. Then, for almost every $s \in [0,1]$, we have the vector equation

$$\gamma'(s) = (\gamma_1'(s), \gamma_2'(s), \gamma_3'(s)) = a_\gamma(s)X(\gamma(s)) + b_\gamma(s)Y(\gamma(s)).$$

This implies that

$$\begin{cases} a(s) = \gamma_1'(s), \\ b(s) = \gamma_2'(s). \end{cases} \tag{2.1}$$

Moreover,

$$\gamma_3'(s) = 2(\gamma_1'(s)\gamma_2(s) - \gamma_2'(s)\gamma_1(s)). \tag{2.2}$$

The equality (2.2) is an instance of a *contact equation.* For an absolutely continuous curve $\gamma\colon [0,1] \to \mathbb{R}^3$, it is *equivalent* to the horizontality of the curve. It implies that the third coordinate of γ is *determined* by the first two! More concretely, if $(\gamma_1, \gamma_2)\colon [0,1] \to \mathbb{R}^2$ is an absolutely continuous planar path, there is a *unique* way to lift it to a horizontal path $\gamma\colon [0,1] \to \mathbb{R}^3$, once $\gamma(0)$ is specified:

$$\gamma_3(s_0) = \gamma_3(0) + 2\int_0^{s_0} \gamma_1'(s)\gamma_2(s) - \gamma_2'(s)\gamma_1(s)\; ds. \tag{2.3}$$

Now, notice that for $g = (x_0, y_0, t_0)$,

$$\begin{aligned}(L_g \circ \gamma)'(s) &= \frac{d}{ds}(x_0 + \gamma_1(s), y_0 + \gamma_2(s), t_0 + \gamma_3(s) + 2(x_0\gamma_2(s) - y_0\gamma_1(s)) \\ &= (\gamma_1'(s), \gamma_2'(s), 2(x_0\gamma_2'(s) - y_0\gamma_1'(s)).\end{aligned}$$

This shows that the standard Euclidean coordinates of γ' are *not* (in general) invariant under the action of $\mathbb{H}$, but the X,Y coordinates are invariant - exactly because the $\frac{\partial}{\partial x}, \frac{\partial}{\partial y}$ coordinates are invariant.

Thus, we may define a left-invariant notion of length for a horizontal path.

Definition 2.4. Let $\gamma\colon [a,b] \to \mathbb{R}^3$ be a horizontal path. The *Carnot-Carathéodory* length of γ is defined to be

$$\operatorname*{length}_{cc}(\gamma) = \int_a^b \left(\gamma_1'(s)^2 + \gamma_2'(s)^2\right)^{1/2}\; ds.$$

In otherwords, we have defined $\{X, Y\}$ to be an *orthonormal frame* of the horizontal tangent bundle $HT\mathbb{H}$. It's interesting to note the following fact: The *cc*-length of a horizontal curve is equal to the length of its projection to $\mathbb{R}^2 \times \{0\}$. Moreover, the change in the height of a horizontal curve is given by (2.3):

$$\gamma_3(s_0) - \gamma_3(0) = 2\int_0^{s_0} \gamma_1'(s)\gamma_2(s) - \gamma_2'(s)\gamma_1(s)\; ds.$$

Some of you calc-3 teachers might recall the formula for the signed area enclosed by a closed planar curve $(\gamma_1, \gamma_2) : [0, s_0] \to \mathbb{R}^2$. To see this, recall that the signed area of an open set $U \subseteq \mathbb{R}^2$ is given by the integral of the volume form:

$$\int_U dx \wedge dy.$$

By Stoke's theorem, this is equal to the integral over the boundary of U of a differential one form whose exterior derivative is $dx \wedge dy$. This one-form is

$$\alpha = \frac{1}{2}(xdy - ydx),$$

since

$$d\alpha = \frac{1}{2}(dx \wedge dy - dy \wedge dx) = dx \wedge dy.$$

Thus

$$2\int_U dx \wedge dy = \int_{\partial U} xdy - ydx.$$

If the boundary ∂U is parameterized by γ, then the change of variables formula gives

$$\int_{\partial U} xdy - ydx = \int_{[0,s_0]} \gamma_1(s)\gamma_2'(s) - \gamma_2(s)\gamma_1'(s)\ ds.$$

In particular,

$$-2\int_U dx \wedge dy = \int_\gamma ydx - xdy = \int_0^{s_0} \gamma_1'(s)\gamma_2(s) - \gamma_2'(s)\gamma_1(s)\ ds.$$

Since γ is a closed curve, the horizontality condition (2.3) now shows

$$0 = \gamma_3(s_0) - \gamma_3(0) = 2\int_0^{s_0} \gamma_1'(s)\gamma_2(s) - \gamma_2'(s)\gamma_1(s)\ ds = -4\int_U dx \wedge dy.$$

Thus, we have shown:

Theorem 2.5. *Let γ be a closed horizontal curve in $\mathbb{H}$. Then the signed area of the projection of γ to $\mathbb{R}^2 \times \{0\}$ is 0.*

We can also give a second proof that any two points can be connected by a horizontal curve: Let $p = (x_0, y_0, t_0)$, and let $\alpha(s) = s(x_0, y_0)$ be the planar segment connecting the origin to the projection of p. If β is any absolutely continuous planar path connecting (x_0, y_0) to 0, then the concatenation is a closed path and by our argument above, the signed area enclosed is

$$\int_U dx \wedge dy = \frac{1}{2}\int_{\beta\circ\alpha} xdy - ydx.$$

However, α contributes noting to this integral, since it is horizontal:

$$\alpha_1(s)\alpha_2'(s) - \alpha_2(s)\alpha_1'(s) = 0.$$

Thus

$$-2\int_U dx \wedge dy = \int_\beta ydx - xdy.$$

Choose β so that the signed area enclosed by $\beta \circ \alpha$ is $t_0/4$, and lift β to a horizontal curve γ starting at $p = (x_0, y_0, t_0)$. We claim that γ ends at the origin. To see this, note that by definition

$$\gamma_3(1) - \gamma_3(0) = 2\int_\beta ydx - xdy = -4\int_U dx \wedge dy = -t_0.$$

Thus $\gamma_3(1) = 0$, as desired.

We are now ready to define a left-invariant metric on $\mathbb{H}$.

Definition 2.6. Let p and q be points of $\mathbb{H}$. Then

$$d_{cc}(p,q) = \inf\{\underset{cc}{\operatorname{length}}(\gamma) : \gamma(0) = p, \gamma(1) = q, \gamma \text{ horizontal}\}.$$

A horizontal curve γ is called a *geodesic* if

$$d_{cc}(\gamma(0), \gamma(1)) = \underset{cc}{\operatorname{length}}\, \gamma.$$

Exericse 2.7. *Let $p = (x_0, y_0, t_0)$ be a point of $\mathbb{H}$. We will try to calculate $d_cc(0,p)$ and find the horizontal curve that realizes it.*

- *Let γ be any horizontal curve connecting p to $0 \in \mathbb{H}$, and let $\beta = \pi_{\mathbb{R}^2}\gamma$. Let α be the line segment connecting 0 to p in $\mathbb{R}^2$. Denote by U the interior of $\beta \circ \alpha$. Show that*
$$t_0 = 4 \int_U dx \wedge dy.$$
- *Show (note?) that*
$$\underset{cc}{\operatorname{length}}(\gamma) = \underset{\mathbb{R}^2}{\operatorname{length}}\, \beta.$$
- *Find the shortest curve β in $\mathbb{R}^2$ connecting 0 to (x_0, y_0) so that the concatentation with α has signed area $t_0/4$. (hint: it is an arc of a circle).*
- *Show that for this choice of β, it's lift γ starting at p is a Heisenberg geodesic connecting p to 0.*
- *When is the geodesic connecting p to 0 unique?*

Exericse 2.8. *The mapping $d_{cc} \colon \mathbb{R}^3 \times \mathbb{R}^3 \to [0, \infty)$ is continuous.*

3. Scaling in the Heisenberg group

In $\mathbb{R}^3$, we have an action of $[0, \infty)$ by scaling that preserves the space and interacts well with both metric and measure. In the Heisenberg group, this exists but is a bit more delicate.

What we would like is a one-parameter family of mappings $\{\delta_\tau \colon \mathbb{H} \to \mathbb{H}\}_{\tau \geq 0}$ so that

$$d_{cc}(\delta_\tau(p), \delta_\tau(q)) = \tau d_{cc}(p,q).$$

In order to prove such an equality, we would ideally want δ_τ to preserve the family of horizontal paths (since they define the metric), and to transform the length of a horizontal path by a factor of τ. Since the length of a horizontal path is simply the length of the projection to the plane, it follows that δ_τ must be the standard scaling by τ on the first two coordinates:

$$\delta_\tau(x,y,t) = (\tau x, \tau y, ???).$$

However, we can deduce how the dilation should behave on the third coordinate if it is to preserve the class of horizontal curves. Suppose that γ and $\delta_\tau \circ \gamma$ are both horizontal curves. Then

$$\begin{aligned}(\delta_\tau \circ \gamma_3)'(s) &= (\delta_\tau \circ \gamma_1'(s))(\delta_\tau \circ \gamma_2(s)) - (\delta_\tau \circ \gamma_2'(s))(\delta_\tau \circ \gamma_1(s)) \\ &= \tau^2(\gamma_1'(s)\gamma_2(s) - \gamma_2'(s)\gamma_1(s)) \\ &= \tau^2 \gamma_3'(s).\end{aligned}$$

Hence, we see that the correct definition should be

$$\delta_\tau(x,y,t) = (\tau x, \tau y, \tau^2 t).$$

It is now easy to check the following statement:

Proposition 3.1. *For each $\tau > 0$, the mapping $\delta_\tau \colon \mathbb{H} \to \mathbb{H}$ preserves the class of horizontal paths, and satisfies*
$$d_{cc}(\delta_\tau(p), \delta_\tau(q)) = \tau d_{cc}(p, q)$$
for each $p, q \in \mathbb{H}$.

Now we see the strange nature of Heisenberg geometry - the scaling of the metric behaves differently in different directions!

For example, let $\gamma(s) = (s, 0, 0)$. Then γ is a horizontal curve, and
$$\gamma'(s) = 1X(\gamma(s)) + 0Y(\gamma(s)).$$
Thus
$$\underset{cc}{\text{length}}(\gamma|_{[0,s_0]} = s_0.$$
In otherwords, $\gamma|_{[0,s_0]}$ is isometric to the line segment $[0, s_0]$ in $\mathbb{R}$. Accordingly,
$$d_{cc}(\gamma(s), 0) = d_{cc}((s, 0, 0), (0, 0, 0) = s d_{cc}((1, 0, 0), (0, 0, 0).$$
However, if $\alpha(s) = (0, 0, s)$, then
$$d_{cc}(\alpha(s), 0) = d_{cc}(\delta_{\sqrt{s}}(0, 0, 1), (0, 0, 0)) = \sqrt{s} d_{cc}((0, 0, 1), (0, 0, 0)).$$
This means that $\alpha|_{[0,s_0]}$ is a scaled copy of $\mathbb{R}$ equipped with the square root metric!

4. The Koranyi norm

Now that the we know how Heisenberg metric scales, we can easily construct another left-invariant metric that scales the same way, but is much easier to compute.

For $p = (z, t) \in \mathbb{H}$, define the *Koranyi norm* of p by
$$||p|| = (|z|^4 + |t|^2)^{1/4}.$$
It is a computation to show that
$$d_K(p, q) = ||p^{-1} * q||$$
is a metric on $\mathbb{H}$; the triangle inequality for this metric is exactly why we needed the factor two in the definition of the Heisenberg group law (otherwise we would have some inconvenient constants involved in the norm).

It's trivial to check that d_K is left-invariant:
$$d_K(L_g(p), L_g(q)) = ||(g * p)^{-1} * (g * q)|| = ||p^{-1} * g^{-1} * g * q|| = ||p^{-1} * q|| = d_K(p, q).$$
Moreover,
$$d_K(\delta_\tau(p), \delta(\tau)(q)) = \tau d_K(p, q).$$
We can immediately determine the topology of this metric from its explicit form.

Exericse 4.1. *The metric space $(\mathbb{H}, d_K)$ is homeomorphic as a topological space to $\mathbb{R}^3$.*

The Koranyi norm and distance are *much* easier to compute than the cc-distance. Moreover, we see directly that the t-axis is a snowflake of $\mathbb{R}$. However, the Koranyi distance does not reflect the sub-Riemannian structure of the Heisenberg group. Despite this, the two metrics are closely related.

Theorem 4.2. *The identity mapping $(\mathbb{H}, d_{cc}) \to (\mathbb{H}, d_K)$ is bi-Lipschitz.*

We can use the above result to show the following:

Theorem 4.3. *Let $E \subseteq \mathbb{R}^3$ be a compact set. Then there is a constant $C_E \geq 1$ depending on E such that for any pair of points $p, q \in E$,*
$$\frac{||p - q||_{\mathbb{R}^3}}{C_E} \leq d_{cc}(p, q) \leq C_E ||p - q||_{\mathbb{R}^3}^{1/2}.$$

And moreover,

Theorem 4.4. *There is a constant $C \geq 1$ such that for any ball B in $(\mathbb{H}, d_{cc})$ of radius r,*

$$\frac{r^4}{C} \leq \mathcal{L}^3(B) \leq Cr^4.$$

Proof of Theorem 4.2. The set $S = \{p \in \mathbb{H} : ||p||_K = 1\}$ is compact, and so the continuous function

$$p \mapsto d_{cc}(p, 0)$$

has a minimum m and a maximum M on S. Since $0 \notin S$, the minimum m is positive. In otherwords, for any $p \in S$,

$$0 < m \leq d_{cc}(p, 0) \leq M < \infty.$$

Now, if $p \in \mathbb{H}$ is arbitrary, then $\delta_{||p||_K^{-1}}(p) \in S$. Hence

$$m \leq ||p||_K^{-1} d_{cc}(p, 0) \leq M,$$

as desired. The statement now follows by left-invariance of the metrics. □

Proof of Theorem 4.3. By our construction of Heisenberg geodesics, we know that there is a compact set $E' \supseteq E$ such that if $p, q \in E$, then the Heisenberg geodesic γ_{pq} connecting p to q is in E'. Now, note that

$$d_{cc}(p, q) = \int \left(|(\gamma'_{pq})_1|^2 + |(\gamma'_{pq})_2|^2\right)^{1/2}$$

while

$$||p - q||_{\mathbb{R}^3} \leq \int \left(|(\gamma'_{pq})_1|^2 + |(\gamma'_{pq})_2|^2 + |(\gamma'_{pq})_3|^2\right)^{1/2}.$$

Since γ_{pq} is horizontal we may write

$$(\gamma'_{pq})_3 = (\gamma'_{pq})_1(\gamma_{pq})_2 - (\gamma'_{pq})_2(\gamma'_{pq})_1.$$

Since γ_{pq} is contained in E', there is a number $M \geq 1$ such that

$$|(\gamma'_{pq})_3| \leq M(|(\gamma'_{pq})_1| + |(\gamma'_{pq})_2)|.$$

This implies that there is some M' satisfying

$$\left(|(\gamma'_{pq})_1|^2 + |(\gamma'_{pq})_2|^2 + |(\gamma'_{pq})_3|^2\right)^{1/2} \leq M' \left(|(\gamma'_{pq})_1|^2 + |(\gamma'_{pq})_2|^2\right)^{1/2}.$$

Integrating now show that

$$||p - q||_{\mathbb{R}^3} \leq M' d_{cc}(p, q).$$

For the upper bound on $d_{cc}(p, q)$, by Theorem 4.2 it suffices to show that

$$||p^{-1} * q||_K \leq C_E ||p - q||_{\mathbb{R}^3}^{1/2}.$$

Let $p = (z_p, t_p)$ and $q = (z_q, t_q)$. Then

$$||p^{-1} * q||_K \leq C(|z_p - z_q| + (|t_p - t_q| + 2|\operatorname{im}(z_p\overline{z_q})|)^{1/2}).$$

Since p and q are in the compact set E, we may find C_E so that

$$|z_p - z_q| \leq C_E |z_p - z_q|^{1/2}$$

and

$$\operatorname{im}(z_p\overline{z_q})^{1/2} \leq C_E |z_p - z_q|^{1/2}.$$

This completes the proof. □

Proof of Theorem 4.4. By translation invariance of the Lebesgue measure, it suffices to consider balls centered at the origin. Theorem 4.2 implies that it suffices to consider balls in the Koranyi metric. However, there is a universal constant $c \geq 1$ such that

$$\{(x, y, t) : |x| < c^{-1}r, |y| < c^{-1}r, |t| < (c^{-1}r)^2\} \subseteq B_K(0, r)$$

and

$$B_K(0, r) \subseteq \{(x, y, t) : |x| < cr, |y| < cr, |t| < (cr)^2\}.$$

This implies the result. □

5. The Poincaré inequality in the Heisenberg group

Theorem 5.1. *The Heisenberg group $(\mathbb{H}, d_{cc}, \mathcal{L}^3)$ supports a 1-Poincaré inequality.*

Let $u\colon \mathbb{H} \to \mathbb{R}$ be a Lipschitz function and let g be an upper gradient of u. Fix a ball $B = B(p_0, r) \subseteq \mathbb{H}$. Since left-translation is an isometry, we may assume without loss of generality that p_0 is the origin.

For $q \in \mathbb{H}$, set $d_q = d_{cc}(0, q)$, and let $\gamma_q\colon [0, d_q] \to \mathbb{H}$ be the arclength parameterized Heisenberg geodesic connecting 0 to q. Then for any $p \in \mathbb{H}$, $p * \gamma_q$ is the arclength parameterized Heisenberg geodesic connecting p to $p * q$. Thus the upper gradient inequality implies that

$$|u(p) - u(p * q)| \leq \int_0^{d_q} g(p * \gamma_q(s))\, ds.$$

Now, the left-invariance of $\mathcal{L}^3$ allows us to change variables as in the Euclidean case, and apply the above inequality:

$$\begin{aligned}
\int_B |u(p) - u_B|\, d\mathcal{L}^3(p) &\leq \frac{1}{\mathcal{L}^3(B)} \int_B \int_B |u(p) - u(q)|\, d\mathcal{L}^3(q) d\mathcal{L}^3(q) \\
&\leq \frac{1}{\mathcal{L}^3(B)} \int_{\mathbb{H}} \int_{\mathbb{H}} \chi_B(p)\chi_B(p * q)|u(p) - u(p * q)|\, d\mathcal{L}^3(p) d\mathcal{L}^3(q) \\
&\leq \frac{1}{\mathcal{L}^3(B)} \int_{\mathbb{H}} \int_{\mathbb{H}} \int_0^{d_q} \chi_B(p)\chi_B(p * q) g(p * \gamma_q(s))\, ds\, d\mathcal{L}^3(p)\, d\mathcal{L}^3(q) \\
&\leq \frac{1}{\mathcal{L}^3(B)} \int_{\mathbb{H}} \int_0^{d_q} \int_{\mathbb{H}} \chi_B(p)\chi_B(p * q) g(p * \gamma_q(s))\, d\mathcal{L}^3(p)\, ds\, d\mathcal{L}^3(q)
\end{aligned}$$

Now, we change variables again: $w = p * \gamma_q(s)$. Using the right-invariance of the measure:

$$\int_{\mathbb{H}} \chi_B(p)\chi_B(p * q) g(p * \gamma_q(s))\, d\mathcal{L}^3(p) = \int_{\mathbb{H}} \chi_{B*\gamma_q(s)}(w)\chi_{B*q^{-1}*\gamma_q(s)}(w) g(w)\, d\mathcal{L}^3(w).$$

Now, if the integrand above is non-zero, then

$$w = a * \gamma_q(s) = b * q^{-1} * \gamma_q(s)$$

for some $a, b \in B$. Thus

$$q = b * a^{-1} \in 2B.$$

As a result,

$$\int_{\mathbb{H}} \chi_{B*\gamma_q(s)}(w)\chi_{B*q^{-1}*\gamma_q(s)}(w) g(w)\, d\mathcal{L}^3(w) \leq \chi_{2B}(q) \int_{2B} g(w)\, d\mathcal{L}^3(w).$$

Plugging this in, we get

$$\begin{aligned}\int_B |u(p) - u_B| \, d\mathcal{L}^3(p) &\leq \frac{1}{\mathcal{L}^3(B)} \int_{\mathbb{H}} \int_0^{d_q} \chi_{2B}(q) \int_{2B} g(w) \, d\mathcal{L}^3(w) ds d\mathcal{L}^3(q) \\ &\leq \frac{1}{\mathcal{L}^3(B)} \int_{2B} d_q \int_{2B} g(w) d\mathcal{L}^3(w) d\mathcal{L}^3(q) \\ &\leq Cr \int_{2B} g(w) d\mathcal{L}^3(w).\end{aligned}$$

This implies the result.