Notes on SOL and NIL

Rich Schwartz

October 23, 2019

The purpose of these notes is to give some general information about Lie groups and then a bunch of information (not all proved in the notes) about SOL and NIL. These notes have many exercises embedded in them.

1 Generalities

Lie Groups: A *smooth Lie group* is a set G which is simultaneously a smooth manifold and a group, such that the two maps $P : G \times G \to G$ and $I : G \to G$ are smooth. Here $P(a,b) = ab$ and $I(a) = a^{-1}$. For the purposes of these notes, the set G is just $\boldsymbol{R}^3$ and the group law is given by polynomial equations. So, the two examples we discuss qualify as Lie groups in an easy-to-prove way.

Left Multiplication: Given any element $g \in G$ we have the *left multipllication map* $L_g : G \to G$ given by $L_g(h) = gh$. This map is a diffeomorphism from G to G, which is to say that it is a bijection which is smooth and non-singular and whose inverse is smooth and non-singular. (Non-singular means that in local coordinates the matrix of partial derivatives is a linear isomorphism.) In the cases of interest, the map L_g will be an affine transformation – i.e., the composition of an invertible linear transformation and a translation. The invertible linear transformation in question is called the *linear part* of L_g.

Tangent Spaces: Being a manifold, a Lie group has a tangent space at each point. This is a vector space having the same dimension as the manifold. In the cases of interest here, the tangent space to $\boldsymbol{R}^3$ at each point is just a copy of $\boldsymbol{R}^3$. It is convenient to think of the elements of the tangent space at a given point $p \in \boldsymbol{R}^3$ as vectors based at p. Thus, for example, one of the tangent vectors in the space $T_{(1,2,3)}\boldsymbol{R}^3$ is the vector which points from $(1,2,3)$ to $(4,10,10)$. The point $(3,8,7) = (4,10,10) - (1,2,3)$ represents this tangent vector.

In a Lie group G, the left multiplication map $L_g : G \to G$ induces a linear isomorphism dL_g from the tangent space $T_h(G)$ to the tangent space $T_{gh}(G)$. The idea is that the matrix of partial derivatives (which makes sense in local coordinates) acts as linear transformations on the relevant tangent vectors. In the case of SOL and NIL, the map dL_g is just given by the linear part of L_g.

Riemannian Metrics: In general, a *Riemannian metric* on a Lie group is a smoothly varying choice of inner product on each tangent space. In the cases of interest, a Riemannian metric is just a smoothly varying choice of inner product for each tangent space of $\boldsymbol{R}^3$. So, at each point of $\boldsymbol{R}^3$ we specify a (potentially varying) inner product on the relevant copy of $\boldsymbol{R}^3$. A Riemannian metric on a Lie group is *left invariant* if the left translations are isometries with respect to the metric. What this means is that

$$\langle dL_g(V), dL_g(W)\rangle_{gh} = \langle V, W\rangle_h,$$

for any tangent vectors $V, W \in T_h(G)$. Lere $\langle , \rangle_h$ is the Riemannian metric evaluated at h, etc.

Arc Length: Suppose that G is equipped with a Riemannian metric. Given a parameterized curve $\gamma : [0,1] \to G$ is given by

$$\int_0^1 \sqrt{\langle \gamma'(t), \gamma'(t)\rangle_{\gamma(t)}}dt.$$

Here $\gamma'(t)$ is the tangent vector which is derivative of γ at t. In other words, you use the usual arc-length formula but you use the Riemannian metric instead of the dot product to compute speeds.

2 Existence of Geodesics

A *geodesic* in a Riemannian manifold is a curve which locally minimizes the distances between points on it. The general theory of Riemannian manifolds shows that every point of a Riemannian manifold can be connected to any nearby point by a geodesic, and that this geodesic is a smooth path. It is not true in general that any two points in a connected Riemannian manifold can be joined by a geodesic. For instance $\boldsymbol{R}^2 - \{(0,0)\}$ is a connected Riemannian manifold but $(1,1)$ and $(-1,-1)$ cannot be connected by a geodesic.

The metrics on SOL and NIL do not have this pathology. In this section I will sketch a proof of the existence of geodesics in any Riemannian manifold whose underlying space is $\boldsymbol{R}^n$ and whose metric is *proper*. This means that the distance from the origin to any point outside the Euclidean ball of radius R tends to ∞ as R tends to ∞. If you just want to know about SOL and NIL, you can ignore this general existence result and move on to the next section. I thought it would be fun to prove the existence of geodesics without resorting to the usual differential equations proof, which requires a fair amount of background about differential geometry. This is a wierd alternative proof that doesn't use as much background.

Let K be some giant cube in $\boldsymbol{R}^n$. The *Hausdorff distance* between two closed subsets A and B of K is defined to be the smallest $\epsilon > 0$ so that each point of A is within ϵ of a point of B, and *vice versa*. This makes the set of closed subsets of K into a metric space.

Lemma 2.1 *Any sequence $\{A_n\}$ of closed subsets of K converges on a subsequence to another closed subset of K.*

Proof: For any integer M we divide K into a grid of M^n equally spaced cubes. Call this grid Γ_M. There are finitely many possible ways to color some of the cubes of Γ_M black and the some white. Each A_n touches some of our small cubes, and we color these cubes black and the rest white. This gives us a sequence of colorings. Passing to a subsequence, we can assume that each A_n touches precisely the same small cubes. Let Λ_M be the union of these. We start with $m = 1$ and define Λ_1 relative to some subsequence, as just mentioned. We then take a further subsequence and produce Λ_2. Then we produce Λ_4. And so on. By construction, we have $\Lambda_1 \supset \Lambda_2 \supset \Lambda_4 \supset \Lambda_8$... Using the Cantor diagonal trick, we can pass to one final subsequence such that, for sufficiently large n, our set A_n touches a cube in Γ_{2^m} if and only if this cube belongs to Λ_{2^m}. But then the Hausdorff distance from A_n to Λ_{2^m} is at most $2^{-m}D$ where D is the diameter of K. So, the intersection

$$\bigcap_{m=0}^{\infty} \Lambda_{2^m}$$

is the desired limit of A_n with respect to our final subsequence. ♠

Lemma 2.2 *The Hausdorff limit of a sequence $\{A_n\}$ of connected paths is connected.*

Proof: We will argue by contradiction. Suppose not. Let B be the limit. Then there are disjoint open sets U and V such that B is contained in $U \cup V$ and intersects both. But then there is some positive $\eta > 0$ such that each point of $B \cap U$ is at least η from $B - U$. Likewise each point of $B \cap V$ is at least η from $B - V$. (This is a standard result that comes from compactness of closed and bounded subsets of $\boldsymbol{R}^n$.) Then each point of B is at least η from $\boldsymbol{R}^n - U - V$. Since A_n is connected, it must contain a point of $\boldsymbol{R}^n - U - V$. Such a point is at least η from B. Hence the Hausdorff distance from A_n to B is at least η. This a contradiction. ♠

Now, choose two points $p, q \in \boldsymbol{R}^m$. Let $D(p,q)$ denote the infimal length of all curves joining p to q with respect to the given Riemannian metric. Thanks to the properness condition, $D(p,q)$ is positive, and a sequence of curves joining p to q whose length converges to $D(p,q)$ stays within some big cube K. Let A_n be the nth such curve in such a sequence. we can pass to a subsequence and take $B = \lim A_n$ in the Hausdorff metric. We know that B is a connected set containing p and q. We want to see that B is a length minimizing path.

Let $D = D(p,q)$. Consider a point $\beta \in B$. There are points of A_n converging to β. This gives us $D(\beta,p) + D(\beta,q) \leq D$. If it happens that $D(\beta,p) + D(\beta,q) < D$ then we could build a path much shorter than A_n by connecting p to q by concatenating short paths from p to β and from β to q, which nearly realize $D(p,\beta)$ and $D(q.\beta)$. Hence

$$D(\beta,p) + D(\beta,q) = D.$$

This equation, together with the connectivity of B, implies that for each $t \in (0,D)$ there is a unique point $\beta(t) \in B$ such that $d(\beta(t),p) = t$. (If there were two such points, the paths A_n would have to pass one of them first on the way from p to q, and this would be an inefficiency.) Moreover $\beta(t)$ varies continuously with t. So, in fact B is a path from p to q such that $d(\beta(t),p) = t$ for all $t \in D$. By construction, β is a geodesic joining p to q.

This existence proof does not show that geodesics are smooth curves. However, one can see pretty easily that they must be differentiable. if not, then they would make some kind of bend, and on a small scale one could shorten them by cutting across the bend.

3 Numerical Construction of Geodesics

In this section I will explain without giving any proofs a way to numerically construct geodesics in a Lie group with a left invariant metric. Let G be such a Lie group. I learned about the theoretical way of doing this from Matt Grayson's thesis, and then I turned it into an algorithm, presumably in the same way that Grayson did when he made his numerical experiments 40 years ago.

Let $S(G)$ denote the space of unit tangent vectors based at the origin in G. This space is a sphere. A geodesic γ is *normalized* if it is given by a unit speed parametrization $t \to \gamma(t)$ such that $\gamma(0) = 0$. We define

$$\gamma^*(t) = dL_{\gamma(t)^{-1}}(\gamma'(t)). \tag{1}$$

The curve $\gamma^*(t)$ is a curve in $S(G)$. What we are doing is looking at the tangent vectors along $\gamma(t)$ and pulling them back by left multiplication so that they are unit vectors based at the identity element of G.

It turns out that there is a vector field Σ on $S(G)$ so that the curve $\gamma^*(t)$ is always an integral curve to Σ. In other words

$$\frac{d\gamma^*(t)}{dt} = \Sigma(\gamma^*(t)). \tag{2}$$

This fact suggests an algorithm for numerically drawing geodesics. The input is a pair (L, U_0, ϵ) where

- L is the length of the geodesic you want to draw,
- U_0 is the tangent vector to the geodesic at the identity.
- ϵ is the "precision". The smaller ϵ is, the more precise the approximation will be and the longer it will take to draw.

Here is the algorithm.

1. Define $U_{k+1} = \pi(U_k + \epsilon\Sigma(U_k))$. Here π is radial projection to $S(G)$.
2. Define $V_0 = e$, the identity, and $V_{k+1} = L_{\epsilon U_k}(V_k)$.

The vectors $U_0, ..., U_{N-1}$ are successive unit vectors which closely approximate the flow lines to Σ. The points $V_1, ..., V_N$ are closely approximations to evenly spaced points along the geodesic segment you are approximating.

4 SOL

SOL is a Lie group whose underlying space is $\boldsymbol{R}^3$ and whose group law is given by

$$(x,y,z)\cdot(a,b,c) = (e^z a + x, e^{-z} b + y, c + z). \tag{3}$$

Here $e = 2.71828...$ is the base of the natural log.

Exercise 1: Check that this operation really does make SOL into a group. So, you want to check that inverses exist and that the associative law holds.

For this discussion is convenient to think of the *horizontal directions* as those parallel to the XY plane and the *vertical direction* as the one parallel to the Z axis. To get a feel for what this group action is like, consider L_g where $g = (x,y,0)$. This map is just horizontal translation in $\boldsymbol{R}^3$. At the same time, consider L_g when $g = (0,0,1)$. This time,

$$L_g(a,b,c) = (ea, b/e, c+1).$$

This map "translates" $\boldsymbol{R}^3$ vertically by one unit and at the same time applies the diagonal matrix

$$\begin{bmatrix} e & 0 \\ 0 & 1/e \end{bmatrix}$$

to the horizontal directions.

The standard left-invariant metric on SOL is given by

$$\langle (V_1,V_2,V_3),(W_1,W_2,W_3)\rangle_{x,y,z} = e^{2z}V_1W_1 + e^{-2z}V_2W_2 + V_3W_3. \tag{4}$$

In this metric, the horizontal and vertical directions are perpendicular.

Exercise 2: Prove that this metric is indeed left-invariant. Hint: it suffices to consider the two kinds of translations considered above, because SOL is generated as a group by the corresponding elements.

Exercise 3: Prove that the restriction of the metric to any XY plane makes it isometric to the Euclidean plane. Likewise, prove that the restriction of the metric to any XZ plane or any YZ plane makes it isometric to the hyperbolic plane. Hint: for the second part of the problem, you want to find an isometry to the upper half plane model of the hyperbolic plane, and exponentiating one of the coords should do the job.

Exercise 4: Prove that the XZ plane is totally geodesic. This means that the shortest path connecting two points in the ZX plane stays in the ZY plane. The same goes for the YZ planes. Hint: Show that the map $F(X,Y,Z) = (X,Z)$ is distance-non-increasing.

Referring to the material in the previous section, the geodesic field for SOL is given by

$$\Sigma(x,y,Z) = (+xz, -yz, -x^2 + y^2). \tag{5}$$

This vector field vanishes at the 6 points $(0,0,\pm1)$ and $(\pm1,\pm1,0)$. The geodesics in this direction are Euclidean straight lines. When you plug this vector field into the algorithm presented in the previous section, you have an efficient way for numerically constructing the geodesics on SOL. Since $S(SOL)$ is just S^2, we will use this alternate notion. So, Σ is a vector field on S^2.

There is one important fact about Σ not mentioned in Grayson's thesis. Define $F(x,y,z) = xy$. The restriction of F to S^2 gives a function on the sphere. The *symplectic gradient* X_F is defined by taking the gradient of this function (on the sphere) and rotating it 90 degrees counterclockwise. Up to sign $X_F = \Sigma$. By construction, the flow lines of Σ lie in the level sets of F.

Exercise 5: Prove that the normalized geodesic $\gamma(t) = (t,t,0)/\sqrt{2}$ is not a distance minimizer when t is sufficiently large, by explicitly constructing a shorter path joining $(0,0,0)$ to $(N,N,0)$ once N is large enough.

Exercise 6: This exercise is kind of a doozy. We know the geodesics in the upper half plane model of $\boldsymbol{H}^2$, and from Exercise 3 we know what the geodesics are in the XZ plane though the origin. Using the explicit parameterization of these geodesics, show that the vector field Σ is correct, at least when restricted to the great circle you get by intersecting S^2 with the XZ plane.

Define a *slab* to be a region of SOL bounded by two horizontal planes. Say that a geodesic in SOL is *ordinary* if it does not lie in an XZ or YZ plane. The next result is the key to understanding the large scale geometry of geodesics in SOL.

Lemma 4.1 *Every ordinary geodesic lies in a slab.*

Proof: (Sketch) Let $\Phi(X, Y, Z) = Z$ be the projection onto the vertical axis. Any left translation map L_g has the property that

$$\Phi \circ L_g = T_g \circ \Phi,$$

where T_g is just an isometric translation of $\boldsymbol{R}$. In other words, if we just think about what left translation does to the vertical lines, it acts as a translation. Any two such vertical translations commute (even though SOL is itself non-abelian.)

The flow lines of Σ corresponding to ordinary geodesics are all periodic. Call a segment σ of a normalized geodesic γ *perfect* if the corresponding segment $\sigma^*(t)$ winds exactly once around a flow line. The flow lines are symmetric with respect to the XY plane. For every point of $\sigma^*(t)$ below the XY plane, there is a corresponding point of $\sigma^*(t)$ above it at the same distance from the XY plane. Call these points *partner points*.

Since the left translations acting on the vertical direction all commute, we can compute the change in vertical height of σ by composing these left translations in any order. (Really we are composing an infinite number of infinitesimal left translations, which we can think of some kind of integral, but we can think of this as being approximated by a finite sequence of compositions of small left translations, as in our numerical scheme.) Since we can compute the vertical changes in any order, the vertical displacements corresponding to left translations at partner points exactly cancel out. Hence, the endpoints of σ lie in the same horizontal plane.

The whole geodesic γ is composed of a bi-infinite sequence of perfect segments, all isometric to σ. Each one starts and ends in the same horizontal plane, and a suitable horizontal translation maps any one of them to σ. So, all of these perfect segments lie in the same slab. Therefore, the whole geodesic lies in a slab. ♠

According to Grayson, all the ordinary geodesics in fact lie in cylinders whose axes are parallel to the XY planes. Each such cylinder is the product of a convex oval and $\boldsymbol{R}$. When the axis of such a cylinder is contained in the plane $Z = 0$, the axis is parallel to either the line $\{(t, t, 0)|\ t \in \boldsymbol{R}\}$ or the line $\{(t, -t, 0)|\ t \in \boldsymbol{R}\}$. Translation along the axis of a Grayson cylinder is both a Euclidean and a SOL isometry.

In general, once have one geodesic on the cylinder, we have an entire decomposition of the cylinder into parallel geodesics that are translates of each other both in the Euclidean sense and in the SOL sense. You just take one of these geodesics and add a suitable vector of the form $(a, b, 0)$ to get the rest.

Exercise 7: Prove that an ordinary geodesic in SOL cannot remain a distance minimizer forever. This is a generalization Exercise 5. Of course, you want to use the property that such a geodesic is confined to a slab. You don't need the cylinder result for this.

Exercise 7 says something very optimistic about large spheres in SOL. Fix a radius R and let S_R denote the set of all points of SOL which are a distance exactly R from the origin. Which tangent vectors in S^2 are directions of geodesics which actually make it to S_R in a distance minimizing way? Let Ω denote the union of the two great circles obtained by intersecting S^2 with the XZ and YZ planes. If we choose a vector in S^2 that is far from Ω, then the corresponding geodesic lies in a fairly thin slab. But then such a geodesic will fail to be minimizing long before reaching S_R. So, there is some $\epsilon(R)$ such that a unit vector in S^2 actually corresponds to a point on S_R if this vector is within $\epsilon(R)$ of Ω. Moreover $\lim_{R\to\infty} \epsilon(R) = 0$. So, the large spheres in SOL are really created by drawing the geodesics emanating from a tiny portion of S^2! Most of the geodesics in SOL crash into each other in a very complicated way, but perhaps the ones which actually are involved in the construction of S_R are fairly well behaved.

To be more precise, there is a function $\zeta : S^2 \to (0, \infty)$ which I will call the *utility function*. The value $\zeta(U)$ is the largest value for which the normalized geodesic tangent to U is a global distance minimizer. (For the expert: one can probably compute $\zeta(U)$ numerically by approximating the associated Jacobi fields and looking numerically for conjugate points.) The previous discussion says that the geodesics which contribute to S_R are precisely those contained in the *superlevel set* $\zeta^{-1}[R, \infty)$.

So, drawing the level sets of ζ seems the key to understanding the geometry of large spheres in SOL. Just for the hell of it, I conjecture that each level set of ζ (if nonempty) is a union of 4 simple closed loops. (There must be a multiple of 4 by symmetry.) If this is the case, then the spheres in SOL are topological spheres, obtained simply by cutting out the disks bounded by these level sets and sewing together their boundaries in a two-to-one fashion.

5 The Isoperimetric Inequality

The isoperimetric inequality is crucial for the understanding of NIL. Let's just prove the isometric inequality for piecewise differentiable simple closed curves. Such a curve is made from finitely many differentiable arcs, joined vertex to vertex in a cyclic fashion. The isoperimetric inequality says that a piecewise differentiable loop bounds a region of maximal area precisely when it is a circle.

Let us first prove that if there is a piecewise differentiable loop which bounds a region of maximal area then this loop must be a circle. Call such a loop a *winner*.

Exercise 8: Given the hypothesis above, prove that there is also a winner which is symmetric with respect to reflection through the origin. (Hint: First show that a winner is convex and then show that a line which bisects the region bounded by the winner must also bisect the arc length of the winner.)

Exercise 9: Choose 3 points on the winner A, B, C such that A and B are diametrically opposed. That is, $A = -B$. Let $D = -C$. The points A, B, C, D are vertices of a parallelogram. Prove that $ABCD$ must be a rectangle. (Hint: if not, suitable flex the picture so as to keep the arc length the same but increase the area.)

Exercise 10: Exercise 9 proves the isoperimetric inequality under the assumption that there is a winner – i.e. that the maximum is actually attained. Use the same kinds of arguments to show that a near-winner must be nearly a circle – i.e., contained in a thin neighborhood of a circle. (Hint: basically just put the word *near* in front of all the arguments you used for Exercises 8 and 9.) Then conclude that the isoperimetric inequality is always true.

Here is one consequence of the isoperimetric inequality. Suppose we have an arc γ joining two points $p, q \in \boldsymbol{R}^2$ and we consider the loop $\gamma \cup \overline{pq}$, where $\overline{pq}$ is the segment joining p and q. If we fix the area of this loop and try to minimize the length of γ, then γ is an arc of a circle. If not, then we could replace a suitable arc of a circle by a better curve and violate the isoperimetric inequality.

6 NIL (or maybe HEIS)

The space underlying NIL is again $\boldsymbol{R}^3$ but it is better to work with $\boldsymbol{C} \times \boldsymbol{R}$. The most symmetric version of the group law for NIL is given by

$$.(z_1, t_1) \cdot (z_2, t_2) = (z_1 + z_2, t_1 + t_2 - \mathrm{Im}(z_1 \overline{z}_2)). \tag{6}$$

Remark: I want to address a question that came up in class. It is worth pointing out that there are other ways to present a group that is isomorphic to NIL which lead to a mildly different geometry. An alternate definition is

$$(z_1, t_1) \cdot (z_2, t_2) = (z_1 + z_2, t_1 + t_2 + x_1 y_2).$$

Here $z_i = z_i + iy_i$. This is the definition used in my computer code, and it leads to balls which are not rotationally symmetric. I think that perhaps some people would call the rotationally symmetric version of NIL coming from Equation 6 by the name *the Heisenberg group* and abbreviate it as HEIS. So, maybe I am really talking about HEIS and not NIL. But I'm just going to stick with calling the rotationally symmetric version NIL, as I did in class.

Let's work with the definition in Equation 6. A *left invariant plane field* is a smooth choice of plane at each point of NIL which is invariant under left multiplication.

Exercise 11: Let $\Pi_{z,t}$ denote the plane that is spanned by the vectors $(z, 0)$ and $(iz, \|z\|^2)$. Prove that Π is a left-invariant plane field on NIL. (Hint: Using the cylindrical symmetry of the group law, it suffices to check this for L_g where $g = (t, 0)$ and $t \in (0, \infty)$. This is just a direct calculation.) Note also that possibly I have the sign wrong, so maybe you will get something close to the answer. In this case, just adjust the group law to make it work exactly. I mentioned in class that this might happen.

We define the left invariant metric on NIL so that it coincides with the dot product at the origin. At the origin, the vertical line is perpendicular to $\Pi_{0,0}$, which the XY-plane. By invariance, this holds at all points of NIL. Also at the origin the projection from $\Pi_{0,0}$ to $\boldsymbol{C}$ is an isometry from the metric restricted to this plane. By invariance, this is true at all points of NIL. Thus, the left invariant metric on NIL is such that the verticals are always perpendicular to the invariant plane field and the inner product of two vectors contained in the invariant plane is obtained by projecting these vectors to $\boldsymbol{C}$ and taking their dot product.

Exercise 12: Work out a formula for the left invariant Riemannian metric based on the information given above.

There is a different left invariant metric on NIL, called the *Carnot-Caratheorody metric*, which has more symmetry. We'll call this the CC metric, for short. Define a path in NIL to be *integral* if it is always tangent to the invariant plane field. The length of such a path, with respect to the left invariant Riemannian metric, is just the length of the projection of the path to $\boldsymbol{C}$. The CC distance between any two points in NIL is defined to be the infimal length of integral paths joining the points.

Exercise 13: Prove that the CC distance is well defined, which means that any two points of NIL are joined by an integral path.

The CC metric on NIL is different from the metric on NIL coming from the left invariant Riemannian metric. The difference is that, in computing the latter, you are allowed to infimize lengths over *all* curves joining two points, not just the integral ones. So, the CC distance is potentially larger than the distance derived from the Riemannian metric.

Exercise 14: Show that the two metrics are not the same. (Hint: Try connecting $(0,0)$ to $(0,\epsilon^2)$ for small ϵ and observe that the shortest integral path has length about ϵ whereas the shortest path has length about ϵ^2.)

A *lift* of a path $\gamma \subset \boldsymbol{C}$ is an integral path $\widetilde{\gamma}$ in NIL such that $\pi(\widetilde{\gamma}) = \gamma$. Here $\pi(z,t) = z$. Note that if $\widetilde{\gamma}$ is a lift, then so is $\widetilde{\gamma} + (0,t)$ for any $t \in \boldsymbol{R}$. Once we pin down the endpoint, the lift is unique. We define the *gain* of γ to be the difference in the heights of the endpoints of $\widetilde{\gamma}$. This is independent of the choice of lift.

Exercise 15: Suppose that γ joins points $p,q \in \boldsymbol{C}$. Prove that the gain of γ is twice the signed area of the loop $\gamma \cup \overline{pq}$. (Sketch: The tangent vectors to $\widetilde{\gamma}$ are all annihilated by the 1-form $xdy - ydx - dz$. This means that the integram of $xdy - ydx$ along γ computes the change in height. Now apply

Green's theorem. In this sketch, it is possible that I've made some sign errors...)

Given Exercise 15, we see that if we fix the gain and try to minimize the length of $\widetilde{\gamma}$, we are precisely in the situation discussed at the end of the previous section. Hence, the integral curves of minimal length project either to straight line segments or to arcs of circles.

Exercise 16: Show that the lift of an arc of a circle is contained in a helix. Hence the geodesic in NIL in the CC metric are helices and horizontal straight lines.

There is much more to say about NIL, but I'm just going to say a little more. The very small balls look like pancakes: little $\epsilon \times \epsilon \times \epsilon^2$ ellipsoids, though I do not think that they are literally ellipsoids. They are rotationally symmetric but may have a funny kind of profile.

Exercise 17: Use the pancake property to show that the open subsets of NIL have Hausdorff dimension (or packing dimension, or covering dimension, or box dimension...) equal to 4. If you like, just prove this for your favorite open subset.

The very large balls look like $N \times N \times N^2$ ellipsoids. The map

$$(z, t) \to (\lambda z, |\lambda|^2 t)$$

is a similarity of the CC metric, expanding all CC distances by $|\lambda|$. So, the balls are permuted by this map. This is the easiest way to understand the shapes of the balls.