Growth in the Heisenberg Group: Research Scrapbook

Hang Lu Su (hanglu.su@gmail.com)
under the guidance of Andrew P. Sánchez and Moon Duchin

December 30, 2016

The first section covers the basic vocabulary for this research, while the second is about the concepts needed for my projects. The last sections are descriptions and insights useful to the problems I am working on.

Because of the uncertain and in-development nature of my work, there may be more than a few errors and inconsistencies in this document. I appreciate all feedback, questions, and suggested corrections by email at `hanglu.su@gmail.com`, and I will do my best to answer.

Contents

1 Definitions

1.1 Generating Function

Suppose a sequence satisfies a recursion, for example $a_n = a_{n-1} + a_{n-2}$. Then, its *generating function* is the formal power series

$$A(x) = \sum_{n=0}^{\infty} a_n x^n$$

1.2 Rational Function

If the generating function satisfies a nice polynomial, for example,

$$A(x) - xA(x) - x^2 A(x) = 1$$

giving the nice form

$$A(x) = \frac{1}{1 - x - x^2} \in \mathbb{Q}(x)$$

then we say that the generation function is a *rational function*

1.3 Rational Growth

Sequences whose generating functions are rational have *rational growth*. So our example $a_n = a_{n-1} + a_{n-2}$ has rational growth.

Examples of sequences with rational growth are polynomials $a_n = p(n)$, and exponentials, for example $a_n = 2^n$. Lots of things don't have rational growth though!

1.4 Eventually quasi-polynomial with period N

If there is a threshold $T > 0$ and a list of polynomials $g_1, \ldots, g_N$ such that $g(n) = g_k(n)$ for $k = n \mod N$ and $n > T$. We think of this behavior as cycling through the polynomials $g_1, \ldots, g_N$ with period N.

Note: eventually quasi-polynomials have rational growth: we can sum up with $x^k \cdot A_k(x^N)$ and add a polynomial to adjust the low values of n (need to work this out for myself, but I believe it.)

1.5 Algebraic Growth and Transcendental Growth

Notice that if $A(x) = \frac{p(x)}{q(x)}$ is equivalent to saying that $qA - p = 0$.

If A satisfies any polynomial with polynomial coefficients, then A has *algebraic growth*.

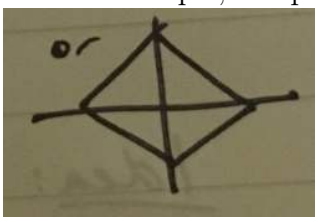
On the other hand, if A satisfies no polynomial with rational growth, then A is *transcendental*.

1.6 Ehrhart Polynomial

Let Ω be a polygon of dimension d, and $n\Omega$ denote its n dilate. Then. Let $\mathbb{Z}^d$ be an integer lattice of dimension d. Then an *Ehrhart polynomial* is a polynomial in n:

$$\begin{aligned} G_\Omega(n) &:= |\mathbb{Z}^d \cap n\Omega| \\ &= a_d n^d + a_{d-1} n^{d-1} + \cdots + a_0 \end{aligned}$$

For example, the polynomial for $\mathbb{Z}^2$ with standard generators, or this shape:



is $\beta(n) = 2n^2 + 2n + 1$.

1.7 Growth Function

Given a group G and a generating set S, the *growth function* β for (G, S) is

$$\beta(n) := |\{\text{words of length } \leq n\}\|.$$

Note that the length of the words is dependent on S.

1.8 Sphere Count

The *sphere count* is $\sigma(n) := \beta(n) - \beta(n-1)$. The intuition for this object is a discrete derivative in n.

1.9 Nilpotent Group

A group is *nilpotent* if for each for every $g \in G$, the sequence of commutators of communutators $\mathrm{ad}_g^1(x) = [g, x], \mathrm{ad}_g^2(x) = [g, [g, x]], \ldots, \mathrm{ad}_g^n(x) = \underbrace{[g, [\ldots [g, x]]]}_{n \text{ times}} \ldots$ terminates after some integer n, i.e. $\mathrm{ad}_g^n(x) = e$ for all $x \in G$.

Note that the idea here is that the group is *almost* commutative because the commutators crash after a finite number of steps. A commutative group is just a group for which $n = 1$.

1.10 Virtually Nilpotent Group

A group is *virtually nilpotent* if there is a subgroup H of finite index such that H is nilpotent.

1.11 Malcev Normal Form

A normal form for certain groups. For example, take the Heisenberg group generated with a, b and $[a, b] = c$. Then, every word is equal to something of the form $a^{n_1}b^{n_2}c^{n_3}$, where $n_i \in \mathbb{Z}$.

Every nilpotent group (recall, that means almost abelian in some sense) admits a Malcev normal form. We are going to write an algorithm to compute this normal form soon!

1.12 Heisenberg Group

Given a ring R, the three-dimensional Heisenberg group is the group of matrices:

$$H(R) := \left\{ \begin{pmatrix} 1 & x & z \\ 0 & 1 & y \\ 0 & 0 & 1 \end{pmatrix} \right\}_{x,y,z \in R}$$

In the case of $H(\mathbb{Z})$, it can be generated by

$$a = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \qquad b = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$$

because

$$a^x = \begin{pmatrix} 1 & x & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \qquad b^y = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & y \\ 0 & 0 & 1 \end{pmatrix}$$

and

$$a^x b^y = \begin{pmatrix} 1 & x & xy \\ 0 & 1 & y \\ 0 & 0 & 1 \end{pmatrix}$$

verify also that if we define $c := [a, b] = aba^{-1}b^{-1}$, then

$$c = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

and

$$c^z = \begin{pmatrix} 1 & 0 & z \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

We can verify by hand that c commutes with a, b. Since c was the commutator, this tells us that the Heisenberg group is nilpotent of step 2 (although we've only implied the result in $H(\mathbb{Z})$, it is true in general). Hence, $H(\mathbb{Z})$ admits a Malcev normal form $a^x b^y c^z$, where we call (x, y, z) the *Malcev coordinates of* $H(\mathbb{Z}) \subset \mathbb{Z}^3$.

If we choose $R = \mathbb{R}$, then $H(\mathbb{R})$ is a Lie group whose behavior resembles that of a "continuous" version of $H(\mathbb{Z})$.

Indeed, $H(\mathbb{R})$ admits a Lie algebra generated by the tangent vectors at the origin

$$A_0 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \qquad B_0 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

If $C_0 = [A_0, B_0]$, where here the bracket is the Lie bracket, then

$$\begin{aligned} C_0 &= A_0 B_0 - B_0 A_0 \\ &= \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} - \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \\ &= \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} - \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \\ &= \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \end{aligned}$$

1.12.1 Exponential Coordinates

In the $H(\mathbb{R})$ in three-dimensions, we can choose $(x, y, z) \in \mathbb{R}^3$ be the *exponential coordinates* representation of $H(\mathbb{R})$:

$$(x, y, z) \leftrightarrow \begin{pmatrix} 1 & x & z + \frac{1}{2}xy \\ 0 & 1 & y \\ 0 & 0 & 1 \end{pmatrix}$$

We call these the exponential coordinates because they come from the map

$$\exp : \text{Lie Algebra} \to \text{Lie Group}$$

This set of coordinates has very nice properties for $H(\mathbb{R})$ that we will explore.

1.13 Carnot Groups

There are *many* equivalent definitions of a Carnot group. Here are a few:

1. If G is a *Carnot group* if it is a simply connected nilpotent group endowed with a one-parameter group of dilations.

2. Alternatively, a *Carnot group* is a simply connected group G such that its Lie Algebra has a Haar measure and a dilation.

For example, $H(\mathbb{R})$ is a Carnot group, as we will see in the next section.

2 Concepts

2.1 Growth in Groups

Relatedly to the word problem, we can ask how many elements in the Cayley graph of a group G are contained within a sphere of radius n? or equivalently, how many words are of length n? Let β_n denote that number, and $B(x) = \sum \beta_n x^n$ be the generating function of that sequence. This function is rational in simple cases, although the exact form of the sequence β_n is dependent on the generators chosen for the Cayley graph. However, whether the growth rate is in the polynomial or exponential range is invariant under change of generators, because changing generators modify the Cayley graph metric in an essentially linear way (I don't really know what this means yet!)

2.2 Gromov's theorem

The groups with growth in the polynomial range are essentially the virtually nilpotent groups.

2.3 Rational Growth, Why is it Important?

If the growth rate of a group is rational, we would have a recursion for the growth values, which would allow us to know how many group elements there are of each length. This way, we could solve the word problem (because we know when we've made progress*), and build the Cayley graph from finitely generated data.

*

However, rationality is pretty delicate, and may depend on the choice of generators.

2.3.1 Cannon, Thurston, Gromov, Benson

If the group is flat (virtually abelian) or hyperbolic, then the growth rate is rational independently of the generators.

2.3.2 Moon, Shapiro

Also true for the Heisenberg group!

2.4 Geodesics and Cayley Graphs

2.4.1 Intuition

Given the Cayley Graph for an infinite group (like the one for the Heisenberg group), we can shrink the edges to make the graph into a smooth manifold, look at the geodesics, and transfer that path back into the discrete graph to find words of shortest length.

2.4.2 CC Metric and Long Words

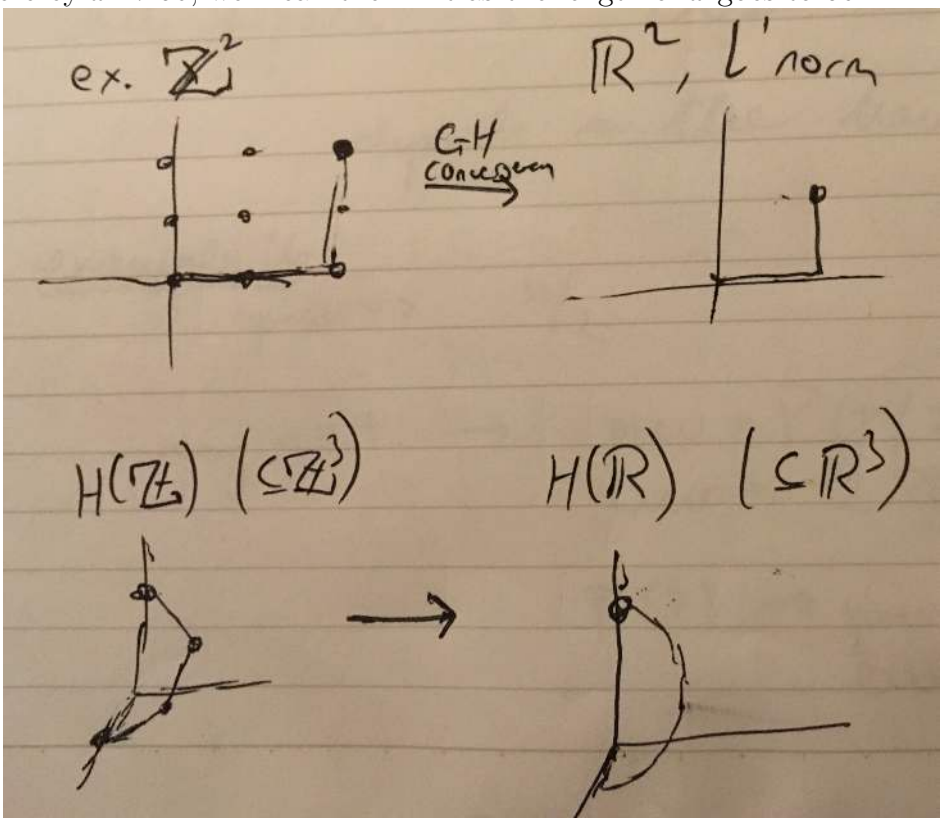
More specifically, given a generating set S and a word x with alphabet in S, we have for some constant k_s

$$d_{CC}(0,x) - k_s \leq |x|_S \leq d_{CC}(0,x) + k_s$$

and

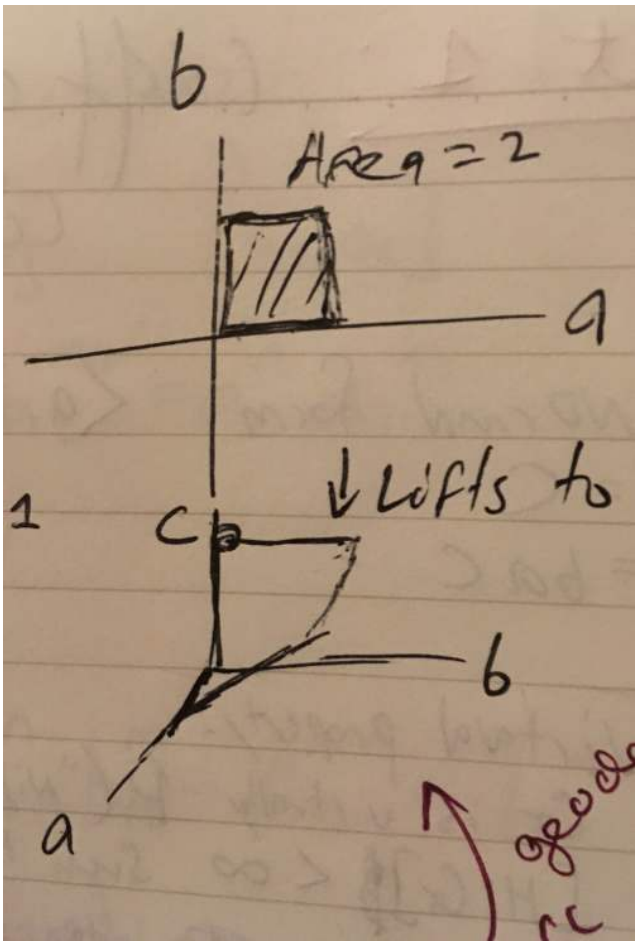
$$\lim_{x\to\infty} \frac{|x|_S}{d_{CC}(0,x)} = 1$$

where by $x \to \infty$, we mean the limit as the length of x goes to ∞.

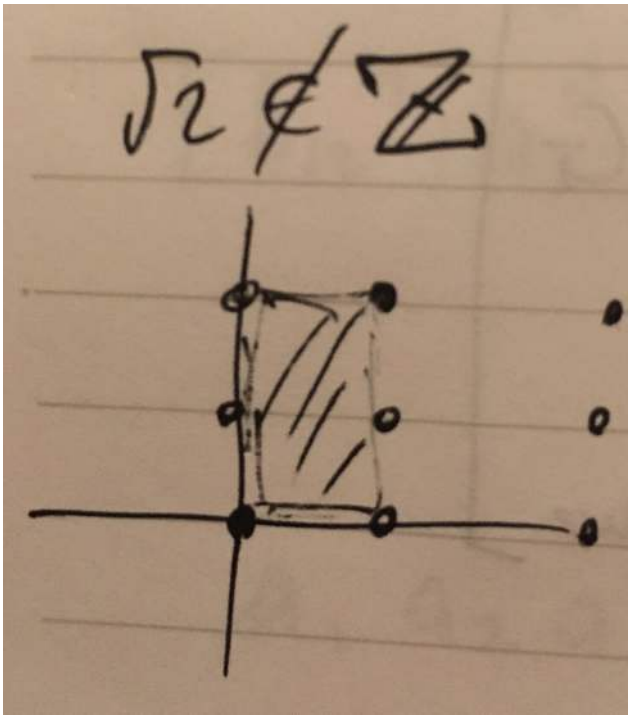


This picture illustrates some examples of limits as you "shrink" the word metric (make the edges smaller and smaller). $\mathbb{Z}^2$ with the word metric converges to $\mathbb{R}^2$ with the L^1-norm. $H(\mathbb{Z}) \subset \mathbb{Z}^3$ with the word metric converges to $H(\mathbb{R}) \subset \mathbb{R}^3$. [I'm not totally clear on what this means yet].

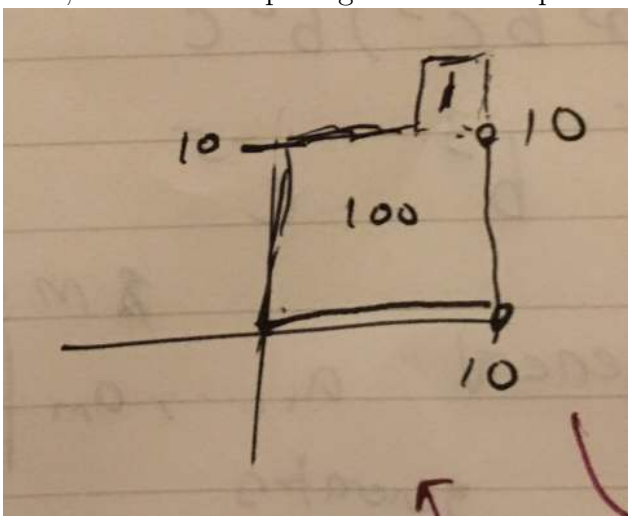
The intuition behind why the metrics are roughly equal for long enough word x is illustrated by the following problem: Given the Heisenberg group generated by a, b and $c = [a, b]$, what is a geodesic to c^2?



Since $c = aba^{-1}b^{-1}, c^{nm} = [a^n, b^m]$. Then, c^2 corresponds to enclosing a rectangle of area 2 in the xy-plane. In the cc-metric space $H(\mathbb{R}) \subset \mathbb{R}^3$, this is simply the problem of finding the rectangle of area 2 with the least perimeter, given by a square of area 2. In other words, the geodesic is given by $a^{\sqrt{2}}b^{\sqrt{2}}a^{-\sqrt{2}}b^{-\sqrt{2}} = c^2$. However, in the Cayley graph, we clearly cannot do this, so we have to settle for $a^2ba^{-2}b = c^2$.



Note, however, that as the power of c grows, for example, for c^101, the more "leeway" we have, because the spacing between the points are smaller relative to the area.



It's then not crazy to think that the word-metric and the cc-metric would be roughly equal for large enough length.

2.4.3 Rigorous Definition of CC metric

Given a manifold M (say the Lie Group $H(\mathbb{R})$), and its tangent bundle, let $E(M)$ be a sub-bundle of this tangent bundle. Suppose then that we have a norm on $E(M) \subset M$ (and quite possible M), then, let γ be *admissible* if $\gamma'(t) \in E(M)$ for all t. In other word, a curve is admissible if and only if it goes along directions in the sub-bundle. Let Γ be the set of

admissible curves. Then, for $p, q \in M$, the CC metric is given by

$$d_{CC}(p,q) = \inf_{\gamma \in \Gamma} \int_0^1 \|\gamma'(t)\| dt.$$

For example, if we are in $\mathbb{R}^3$, and we select our sub-bundle to be $\mathbb{R}^2$ with the Euclidean norm, then there are no admissible curves from $(0,0,0)$ to $(0,0,1)$ because we are "trapped" in the xy-plane. Thus, we will have to take the infimum of the empty set, which is ∞ in correspondance with our intuition.

Things get more interesting in $H(\mathbb{R})$.

2.5 Heisenberg Geometry

2.5.1 Exponential coordinates in the Heisenberg group

In the $H(\mathbb{R})$ in three-dimensions, let $(x,y,z) \in \mathbb{R}^3$ be the *exponential coordinates* representation of $H(\mathbb{R})$:

$$(x,y,z) \leftrightarrow \begin{pmatrix} 1 & x & z + \frac{1}{2}xy \\ 0 & 1 & y \\ 0 & 0 & 1 \end{pmatrix}$$

Then

$$\begin{aligned}(x_1,y_1,z_1)(x_2,y_2,z_2) &\leftrightarrow \begin{pmatrix} 1 & x_1 & z_1 + \frac{1}{2}x_1y_1 \\ 0 & 1 & y_1 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & x_2 & z_2 + \frac{1}{2}x_2y_2 \\ 0 & 1 & y_2 \\ 0 & 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 1 & x_1+x_2 & z_1+z_2+\frac{1}{2}(x_1y_1+x_2y_2)+x_1y_2 \\ 0 & 1 & y_1+y_2 \\ 0 & 0 & 1 \end{pmatrix} \\ &\leftrightarrow (x_1+x_2, y_1+y_2, z')\end{aligned}$$

now, by definition,

$$\begin{aligned} z' &= \underbrace{z_1 + z_2 + \frac{1}{2}(x_1y_1 + x_2y_2) + x_1y_2}_{z} - \underbrace{\frac{1}{2}(x_1y_1)(x_2y_2)}_{\frac{1}{2}xy} \\ &= z_1 + z_2 + \frac{1}{2}(x_1y_2 - x_2y_1). \end{aligned}$$

here,

$$\boxed{(x_1,y_1,z_1)(x_2,y_2,z_2) = (x_1+x_2, y_1+y_2, z_1+z_2+\frac{1}{2}(x_1y_2 - x_2y_1))}$$

for the exponential coordinate representation of $H(\mathbb{R})$.

It is now easy to that for any n, $(x,y,z)^n = (nx, ny, nz)$.

2.5.2 Lie Algebra and Horizontal Planes

Intuitively, the tangent space of every point in $H(\mathbb{R})$ is a subspace of $\mathbb{R}^n$ in some nice embedding. There are *horizontal planes* at every point in $H(\mathbb{R})$ generated by a push foward of X_0 and Y_0 by left multiplication in $H(\mathbb{R})$ to produce left-invariant vector fields X and Y.

Let $E(M)$ be the tangent bundle of these horizontal planes. Then, since we can only really go in the X and Y directions, given a curve $\gamma(t) = (\gamma_1(t), \gamma_2(t), \gamma_3(t)))$, the curve is admissible if and only if

$$\gamma_3'(t) = \frac{1}{2}(\gamma_1(t)\gamma_2'(t) - \gamma_2(t)\gamma_1'(t))$$

[I NEED TO CHECK THIS. I know it's a remark about the tangent space].

If we use the $\mathbb{R}^3$ coordinate representation for $H(\mathbb{R})$, we may define the projection

$$\pi : H(\mathbb{R}) \to E(M), \qquad (x, y, z) \mapsto (x, y).$$

2.5.3 Theorem: Standard fact from Heisenberg Geometry

For any path $\gamma = (\gamma_1, \gamma_2)$ from $(0, 0)$ to (x, y), there is a unique admissible curve $\bar{\gamma} = (\gamma_1, \gamma_2, \gamma_3)$. The lifted curve connects the origin to the point (x, y, z) where z is equal to the signed Euclidean area of R, the region in the plane enclosed by γ and a straight chord (line) from $(0, 0)$ to (x, y).

Proof. Indeed,

$$\begin{aligned} z &= \int_0^1 \gamma_3'(t)dt \\ &= \int_0^1 \frac{1}{2}(\gamma_1\gamma_2' - \gamma_2'\gamma_1)dt \end{aligned}$$

Let $\lambda = (\lambda_1(t), \lambda_2(t)) = (x(1-t), y(1-t))$ be the straight chord from (x, y) to $(0, 0)$. We observe that

$$\lambda_1'\lambda_2 - \lambda_2'\lambda_1 = -xy(1-t) + x(1-t)y = 0$$

so we may concatenate λ in the integral without changing its value. Let $*$ denote the concatenation of paths. Then, $\gamma * \lambda = \partial R$ and

$$\begin{aligned} z &= \int_\gamma \frac{1}{2}(\gamma_1\gamma_2' - \gamma_2'\gamma_1) \\ &= \int_{\gamma*\lambda} \frac{1}{2}(\gamma_1\gamma_2' - \gamma_2'\gamma_1) \\ &= \int_{\partial R} \frac{1}{2}(\gamma_1\gamma_2' - \gamma_2'\gamma_1) \\ &= \frac{1}{2}\int_R dx \wedge dy. \quad \text{(Stokes' Theorem)} \end{aligned}$$

the last quantity is clearly the signed area of the region. □

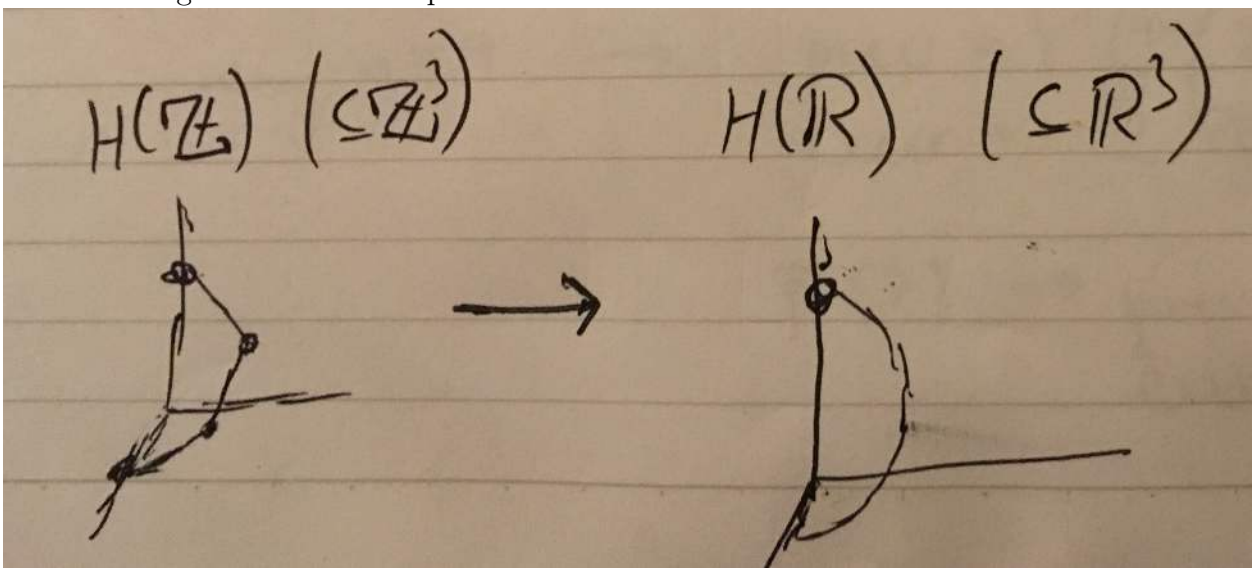
2.5.4 CC metric on $H(\mathbb{R})$

Choose a centrally symmetry convex body $Q \subset E(M)$ with boundary L, and let $\|\cdot\|_L$ denote the norm induced by taking Q as the unit ball, and L as the unit sphere. This induces a CC metric on $H(\mathbb{R})$.

Due to the last theorem, measure the length of an admissible curve in $H(\mathbb{R})$ with respect to the CC metric is equivalent to measuring the length of its projection to M with respect to $E(M)$ with respect to the norm $\|\cdot\|_L$ [ACTUALLY NOT SURE IF I'M BULLSHITTING HERE].

Addressing the last section, one really interesting phenomenon arising from this is when we take the Q to be the unit sphere. This induces a Euclidean metric on $H(\mathbb{R})$, which we will denote $\|\cdot\|$.

This brings us back to this picture:



which draws the analogy between a geodesic with the word metric in $H(\mathbb{Z})$ and a geodesic in $H(\mathbb{R})$ with the CC metric.

Here, the x-axis represents the power of a, the generator, the y-axis the power of b, and the z-axis the power of $c = aba^{-1}b^{-1}$. Since the CC geodesic represents in this case the "smooth" version of the word geodesic, we can intuit that if x is a word in $H(\mathbb{Z})$ with standard generators $S = \{a, b\}$ and $|\cdot|$ is the word metric, then

$$d_{CC}(0, x) \leq |x|_S$$

so we have a way of easily upper-bounding the CC metric.

Now, how does the CC metric and the Euclidean one (we'll denote it d) compare? Suppose we want to go from $(0, 0, 0)$ to $(0, 0, 1)$. Then since $c = aba^{-1}b^{-1}$, we have that

$$1 < d_{CC}((0, 0, 0), (0, 0, 1)) \leq 4, \qquad d((0, 0, 0), (0, 0, 1)) = 1$$

so we might ask, what's the use of travelling in the CC metric? Here's the punchline. Suppose we want to go from $(0, 0, 0)$ to $(0, 0, 10000)$. Then, $d((0, 0, 0), (0, 0, 10000)) = 10000$ of course. However, because of the formula $c^{-mn} = a^n b^m a^{-n} b^{-m}$, we have that $c^{10000} = a^{-100} b^{-100} a^{100} b^{100}$ and

$$d_{CC}((0, 0, 0), (0, 0, 10000)) \leq 400.$$

For large N,

$$d_{CC}((0,0,0),(0,0,N)) \leq 4\sqrt{N}.$$

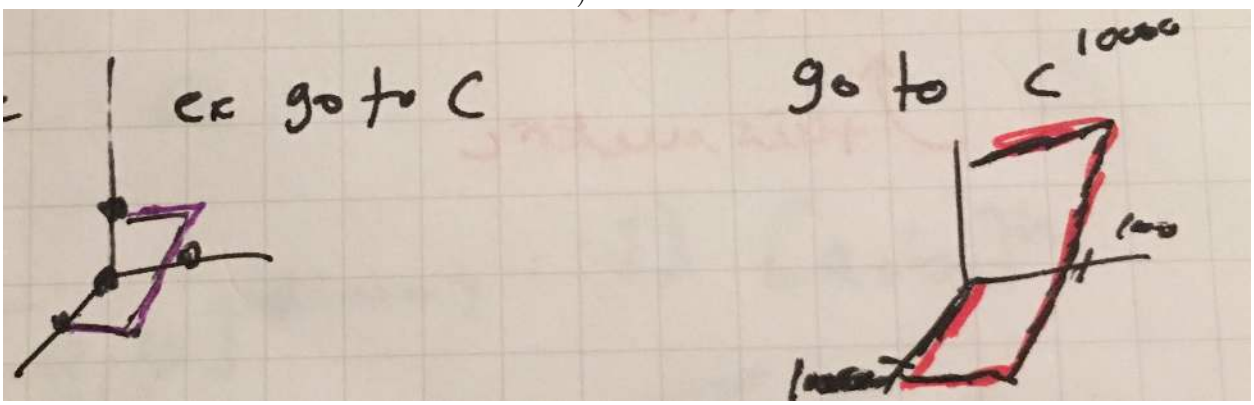
Crazy!

2.6 Gromov-Hausdorff Limit

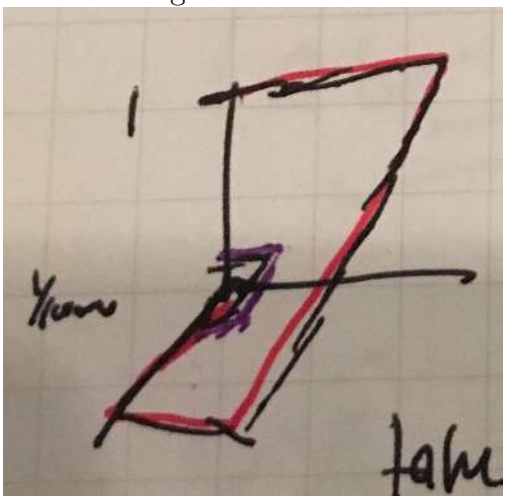
2.6.1 Intuition

The way to, intuitively, take the Gromov-Hausdorff limit (so for example, have $H(\mathbb{Z}) \to H(\mathbb{R})$) is to look at large-scale behaviors of $H(\mathbb{Z})$. Because $H(\mathbb{Z})$ is self-similar in the sense that it admits a dilation, we can do the following:

1. For every path in the $H(\mathbb{Z})$, take a large dilation $\delta_s, s \gg 1$ of the path (say go from $c = aba^{-1}b^{-1} \mapsto c^{10000} = a^{100}b^{100}a^{-100}b^{-100}$).



2. Rescale using the same kind of dilation map $\delta_{s'}$, this time with $0 < s' \ll 1$.



3. Keep doing this until all the rational points are filled.

4. To get the whole $H(\mathbb{R})$, take the limit of the rational points.

5. Every word geodesic becomes a d_{CC} geodesic.

This means, in particular, that $(H(\mathbb{Z}), |\cdot|_S)$ (where $|\cdot|_S$ is the word metric with generator set $S = \{a, b, c\}$) converges to $(H(\mathbb{R}), d_{CC}(\cdot))_{L^1}$, where $d_{CC}(\cdot))_{L^1}$ is the CC metric with L^1-norm.

Important point: $H(\mathbb{Z})$ does *not* converge to $H(\mathbb{R})$ with the Riemannian (unit sphere) norm in the CC metric as was misleading pictured in the previous section on the CC metric. [Ugh!]

$(H(\mathbb{R}), \|\cdot\|)$ and $(H(\mathbb{R}), \|\cdot\|_1)$ come from totally different regimes!

$(H(\mathbb{R}), \|\cdot\|)$ comes from an analysis / differential geometry background, whereas $(H(\mathbb{R}), \|\cdot\|_1)$ comes from the geometric group theory background.

So remember the setup we had where

$$\lim_{|x|\to\infty} \frac{|x|_S}{d_{CC}(x)} = 1$$

in the sense of the structure theorem of Pansu?

That d_{CC} is induced by the L^1 norm, *not* the Riemannian one. I repeat, SO NOT the Riemannian (unit sphere) one.

2.7 Dilations

A *dilation* is a metric space (X, d) with metric d is a function $f : X \to X$ such that

$$d(f(x), f(y)) = rd(x, y)$$

with $r > 0$. A dilation is also a *similarity* of space because it is a bijective map respecting the property

$$d(f(x), f(y)) = rd(x, y), \qquad r > 0$$

as stated above (the two definitions are extremely similar).

In $\mathbb{R}^n$, the dilation map

$$\delta_s(x_1, x_2, \ldots, x_n) = (sx_1, \ldots, sx_n)$$

scales volumes of subsets $E \subset \mathbb{R}^n$ in the following way:

$$\lambda(\delta_s(E)) = s^n \lambda(E),$$

where λ is the Lebesgue measure.

In $H(\mathbb{R})$, the dilation map with respect to the exponential coordinates is:

$$\delta_s(x, y, z) = (sx, sy, s^2 z)$$

is a map with the property that for $E \subset H(\mathbb{R})$,

$$\lambda(\delta_s(E)) = s^4 \lambda(E).$$

Indeed, this is because we've proved in the last section that in the CC metric, the length of a segment in z is the area formed by xy. Therefore, let E be a cube with lengths x, y, z. Then the volume of $\delta_s(E)$ is

$$\begin{aligned}
\lambda(\delta_s(E)) &= d_{CC}((0,0,0),(sx,0,0)) \cdot d_{CC}((0,0,0),(0,sy,0)) \cdot d_{CC}((0,0,0),(0,0,s^2 z)) \\
&= sx \cdot sy \cdot sxsy \\
&= s^4(x \cdot y \cdot xy) \\
&= s^4 \cdot \lambda(E)).
\end{aligned}$$

2.7.1 Exponential Coordinates

In the exponential coordinates, δ_s is a group homomorphism (over the Heisenberg group) under those coordinates as well as a similarity of space when viewing $H(\mathbb{R})$ as a subset of $\mathbb{R}^3$. Indeed, $\delta_s((x_1, y_1, z_1)(x_2, y_2, z_2)) = \delta_s(x_1, y_1, z_1)\delta_s(x_2, y_2, z_2)$ as we check:

$$\begin{aligned}
&\delta_s((x_1, y_1, z_1)(x_2, y_2, z_2)) \\
=&\delta_s(x_1 + x_2, y_1 + y_2, z_1 + z_2 + \frac{1}{2}(x_1y_2 - x_2y_1)) \\
=&(s(x_1 + x_2), s(y_1 + y_2), s^2(z_1 + z_2 + \frac{1}{2}(x_1y_2 - x_2y_1)))
\end{aligned}$$

$$\begin{aligned}
&\delta_s(x_1, y_1, z_1)\delta_s(x_2, y_2, z_2) \\
=&(sx_1, sy_1, s^2z_1)(sx_2, sy_2, s^2z_2) \\
=&(sx_1 + sx_2, sy_1 + sy_2, s^2z_1 + s^2z_2 + \frac{1}{2}(sx_1sy_2 - sx_2sy_1)) \\
=&(s(x_1 + x_2), s(y_1 + y_2), s^2(z_1 + z_2 + \frac{1}{2}(x_1y_2 - x_2y_1))).
\end{aligned}$$

2.7.2 Hausdorff Dimension

[This section needs more work]
In the previous example, n and 4 were the Hausdorff dimension of $\mathbb{R}^n$ and $H(\mathbb{R})$ respectively.

The existence of a dilation makes $H(\mathbb{R})$ a Carnot group. It's worth clarifying that because we've shown there is a dilation, the dilation works for *all* CC metrics on $H(\mathbb{R})$. (So all choice of subbundle followed by a choice of norm on the subbundle). In particular, this implies that the Hausdorff dimension of $H(\mathbb{R})$ does not depend on the metric, which makes sense given that it's also defined by its lower central series for $H(\mathbb{R})$, as we will see shortly in the section about lower central series.

2.8 Lower Central Series and Homogenous Dimension

The *lower central series* of a group G is the descending series of subgroups

$$G = G_1 \trianglerighteq G_2 \trianglerighteq \ldots \trianglerighteq G_n \trianglerighteq \ldots$$

where $G_2 = [G, G]$ the subgroup generated by all the commutators of G, $G_3 = [[G, G], G]$ and so on ($G_{n+1} = [G_n, G]$).

Nilpotent groups are precisely the groups for which this lower central series terminates.

For example, for $H(\mathbb{Z})$, the lower central series would be:

$$H(\mathbb{Z}) \trianglerighteq \langle c \rangle \cong \mathbb{Z} \trianglerighteq 1,$$

and for the nilpotent group of step 2 with 3 generators, $N_{2,3}$, we have:

$$N_{2,3} = \langle a_1, a_2, a_3 | [[a_i, a_j], a_k] = 1 \rangle$$

where (i, j, k) is any permutation of $(1, 2, 3)$. Then, since we have 3 distinct combinations (as $[a_i, a_j] = [a_j, a_i]^{-1}$), $G_2 \cong \mathbb{Z}^3$, yielding the following lower central series:

$$N_{2,3} \trianglerighteq \mathbb{Z}^3 \trianglerighteq 1.$$

2.8.1 Homogenous dimension

For a lower central series of depth n, define the following quantity called the *homogenous dimension*:

$$\sum_{i=1}^{n-1} i \cdot \text{rank of } i\text{-th quotient.}$$

For example, let's calculate the homogenous dimension of $H(\mathbb{Z})$:

$$\underbrace{H(\mathbb{Z})}_{\text{depth}=1} \trianglerighteq \underbrace{\mathbb{Z}}_{\text{depth}=2} \trianglerighteq \underbrace{1}_{\text{depth}=3}$$

The first quotient is: $H(\mathbb{Z})/[a, b]$, which is the abelianization of $H(\mathbb{Z})$, which has rank 2. The second $\mathbb{Z}/1$ has rank 1.

Putting this together, the homogenous dimension of $H(\mathbb{Z})$ is:

$$1 \cdot 2 + 2 \cdot 1 = 4$$

which is also the Hausdorff dimension of $H(\mathbb{R})$.

This is because $H(\mathbb{R})$ is a Carnot group and the lower central series of $H(\mathbb{R})$ is the same as that of $H(\mathbb{Z})$.

2.8.2 Well-known fact about Carnot groups

If G is a Carnot group, then its homogenous dimension is equal to its Hausdorff dimension.

This in fact intuitive, because the exponents of the dilation coefficients are additive when we multiply the coordinates together to get the volume. Let's be more precise:

for $H(\mathbb{R})$, the dilation is given by

$$\partial_s(x, y, z) = (sx, sy, s^2 z)$$

Since x, y are abelian, and the plane in which the curves are admissible, the $1 \cdot 2$ term corresponds to a scaling of s^1 for two generators:

$$d_{CC}((0,0,0),(sx,0,0)) \cdot d_{CC}((0,0,0),(0,sy,0)) = s^1 x \cdot s^1 y = s^2 xy,$$

whereas the $2 \cdot 1$ factor corresponds to a scaling of s^2 for 1 generator:

$$d_{CC}((0,0,0),(0,0,s^2 z)) = s^2 xy.$$

Which adds up in the exponent to

$$\underbrace{s^1 x \cdot s^1 y}_{1\cdot 2} \cdot \underbrace{s^2 xy}_{2\cdot 1} = \underbrace{s^4}_{=4}(x^2 y^2).$$

[Might give a more formal argument at some point.]

2.9 Some Lie Groups concepts

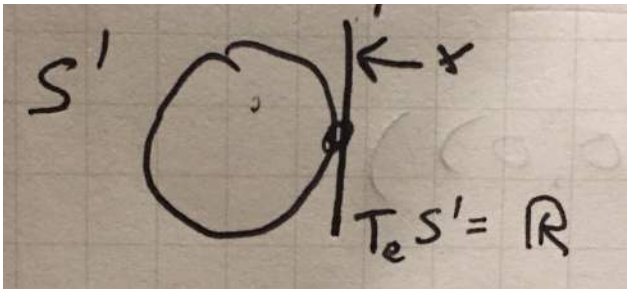
2.9.1 Exponential Map

Given a Lie group G, and its tangent space at the identity T_eG an exponential map is a map:

$$\exp : \text{Lie Algebra} \to \text{Lie Group}$$

$$\mathrm{T}_eG \mapsto G$$

Suppose we have the Lie group of rotationals aroudn the circle, S^1. Then, $T_eS^1 = \mathbb{R}$, as illustrated:



if x is a point in $\mathbb{R} = T_eG$, then we have the exponential map:

$$\exp : \mathbb{R} \to S^1$$

$$x \mapsto e^{ix}$$

Intuitively, we "wrap around" $\mathbb{R}$ to map x back on S^1. However, x was taken from the tangent space at *one* point! This can only works if S^1 is self-similar. But that's precisely the definition of a Lie group: a space that is self-similar under its own action.

3 Project 1: Malcev Normal Form

3.1 Statement of the Problem

It is a fact that every nilpotent group (every group whose commutators of commutators of ... become the identity element) admits a Malcev normal form. We would like to make an algorithm which, given a presentation (for example $N_{2,2}$, the nilpotent group with two generators whose commutator of commutator is the identity, $\langle a, b | [a, b] = c \text{ central } \rangle$) and a word in $\{a, b, c\}^*$, returns its Malcev normal form $a^x b^y c^z$.

3.2 Strategy

3.2.1 $N_{2,2}$ case

Suppose we are in $N_{2,2}$. Note that, $aba^{-1}b^{-1} = c \iff ba = abc^{-1}$ and that c is central. Then,

$$b^2a = (ba)bc^{-1} = abc^{-1}bc^{-1} = ab^2c^{-2}$$
$$ba^2 = abc^{-1}a = a(ba)c^{-1} = a^2bc^{-2}$$

we can check that by induction

$$b^m a^n = a^n b^m c^{-(mn)}$$

Suppose we want to reduce the word aba^2b^2a, then, reading from right to left, we are going to replace every instance of $b^m a^n$ with $a^n b^m c^{-mn}$, move c at the end, then freely reduce and repeat. For example:

$$\begin{aligned}
&= aba^2(b^2a) \\
&= aba^2(ab^2 \underbrace{c^-2}) \\
&= aba^2ab^2 \underbrace{c^-2}) \\
&= ab(a^2a)b^2c^-2 \\
&= ab(a^3)b^2c^-2 \\
&= aba^3b^2c^-2 \\
&= a(ba^3)b^2c^{-2} \\
&= a(a^3b \underbrace{c^{-3}})b^2c^{-2} \\
&= aa^3bb^2c^{-2} \underbrace{c^{-3}} \\
&= (aa^3)(bb^2)c^{-2}c^{-3} \\
&= a^4b^3c^{-5}
\end{aligned}$$

3.3 $N_{2,3}$ (2-generated, 3 steps) case

Suppose we are given generators $\langle a_1, a_2 | b_{12} = [a_1, a_2] | c_{121} = [b_{12}, a_1], c_{122} = [b_{12}, a_2] \rangle$, then similarly as before

$$[a_2^m, a_1^n] = b_{12}^{-mn}$$
$$\Rightarrow a_1^m a_2^n b_{12}^{-mn} = a_2^n a_1^m$$

and

$$[b_{12}^m, a_1^n] = c_{121} = a_1^n b_{12}^m c_{121}^{-mn} = b_{12}^m a_1^n$$
$$[b_{12}^m, a_2^n] = c_{122} = a_2^n b_{12}^m c_{122}^{-mn} = b_{12}^m a_2^n$$

Keeping in mind only the c's are central.

3.4 $N_{3,2}$ (3-generated, 2 step case)

The presentation is the following:

$$\langle a, b, c, [a, b] = d, [a, c] = e, [b, c] = f | d, e, f \text{ central} \rangle$$

Then adapting directly from the $N_{2,2}$ case, we have the following:

$$b^m a^n = a^n b^m d^{-mn}$$
$$c^m a^n = a^n c^m e^{-mn}$$
$$c^m b^n = b^n c^m f^{-mn}$$

The algorithm is as follows: for every instance of $b^m a^n$, replace it with $a^n b^m d^{-mn}$. If there are no e, f's at the end, move d there. If there is, move d before e and f. Freely reduce. Repeat process for $c^m a^n$ and $c^m b^n$, moving e after d but before f, and move f at the very end.

Let's do an example:

$$\begin{aligned}
& ba^2cabc \\
=&(ba^2)cabc \\
=&(a^2b\underbrace{d^{-2}})cabc \\
=&a^2b(ca)bc\underbrace{d^{-2}} \\
=&a^2b(ac\underbrace{e^{-1}})bcd^{-2} \\
=&a^2bacbcd^{-2}\underbrace{e^{-1}} \\
=&a^2(ba)cbcd^{-2}e^{-1} \\
=&a^2(ab\underbrace{d^{-1}})cbcd^{-2}e^{-1} \\
=&a^2abcbcd^{-2}\underbrace{d^{-1}}e^{-1} \\
=&(a^2a)bcbc(d^{-2}d^{-1})e^{-1} \\
=&a^3bcbcd^{-3}e^{-1} \\
=&a^3b(cb)cd^{-3}e^{-1} \\
=&a^3b(bc\underbrace{f^{-1}})cd^{-3}e^{-1} \\
=&a^3(bb)(cc)d^{-3}e^{-1}f^{-1} \\
=&a^3b^2c^2d^{-3}e^{-1}f^{-1}
\end{aligned}$$

3.5 General case $N_{k,s}$

We would like to, given a word in $\langle a_1, \ldots, a_k | b_{12}, \ldots, b_{\binom{k}{2}} | c_{121} \underbrace{\ldots}_{s\text{-steps}} \rangle$ return the Malcev normal form.

3.6 Applications

Other than *it would be great to have a Malcev normal form calculator*, this allows to determine how many words there are of a certain length. Take $H(\mathbb{Z})$ again, with generators a, b and $c = [a, b]$. Then, if we want to calculate how many words there are of length n, we need to be able to determine if two words are equal in order to not overcount. (This is basically solving the word problem).

In the case of nilpotent groups, we can simply take the Malcev normal form of both words. As far as the length in concerned, we want to take the geodesic length. For example $ba = abc^{-1}$ is of length 2. This brings up an important point: *the Malcev normal form is SO NOT a geodesic*! [Thanks David Yang for pointing it out.]

4 Project 2: Detecting Polynomials

4.1 Statement of the Problem

We would like to write an algorithm to generate the growth functions and sphere counts for the Heisenberg group $H(\mathbb{Z})$ with generating sets $S = \{a, b\}$, $S = \{a, b, c = [a, b]\}$, and $S_{\text{hex}} = \{a, b, ab\}$.

4.2 Recurrence and Sphere Count

Notice that for the sphere count $\sigma(n) := \beta(n) - \beta(n-1)$, if

$$\beta(n) = a_d n^d + a_{d-1} n^{d-1} + a_{d-2} n^{d-2} + \cdots + a_0,$$

then, recalling the identity $(n-1)^p = \sum_{k=0}^{p} \binom{p}{k} n^{p-k} (-1)^k$,

$$\beta(n-1) = \underbrace{a_d (n-1)^d}_{(1)} + \underbrace{a_{d-1}(n-1)^{d-1}}_{(2)} + d - 2(n-1)^{d-2} + \cdots + a_0$$

One way to determine $\sigma(n)$ is then to group the terms of same power together. Let $\sigma(n) = \sigma_d n^d + \cdots + \sigma_0$. Then,

$$\sigma_d = a_d - \underbrace{a_d}_{(1)} = 0$$

$$\sigma_{d-1} = a_{d-1} - \left[\underbrace{-d a_d}_{(1)} + \underbrace{a_{d-1}}_{(2)} \right] = d a_d$$

So $\sigma(n) = a_d d n^{d-1} + O(n^{d-2})$.

Intuitively, this makes sense because $\sigma(n)$ works like a derivative on the ball $\beta(n)$ of radius n.

Indeed, we define a derivation D on functions of $n \in \mathbb{N}$, by

$$\begin{aligned}
&D^0[f(n)] := f(n) \\
&D^1[f(n)] := f(n) - f(n-1) \\
&D^2[f(n)] = D[f(n) - f(n-1)] = f(n) - f(n-1) - (f(n-1) - f(n-2))
\end{aligned}$$

This is indeed a derivation because it is linear:

$$\begin{aligned}
D[af(n) + bg(n)] &= af(n) + bg(n) - (af(n-1) + bg(n-1)) \\
&= a(f(n) - f(n-1)) + b(g(n) - g(n-1)) \\
&= aD[f(n)] + bD[g(n)]
\end{aligned}$$

and satisfies the Leibnitz rule $D[f(n)g(n)] = D[f(n)]g(n-1) + f(n)D[g(n)]$, which we check:

$$\begin{aligned}
D[f(n)]g(n-1) + f(n)g(n) &= [f(n) - f(n-1)]g(n-1) + f(n)[g(n) - g(n-1)] \\
&= f(n)g(n-1) - f(n-1)g(n-1) + f(n)g(n) - f(n)g(n-1) \\
&= f(n)g(n) - f(n-1)g(n-1) \\
&= D[f(n)g(n)]
\end{aligned}$$

So we can equivalently define the sphere count as $\sigma(n) = D[\beta(n)]$, and now we are free to think of the sphere count as a derivative of $\beta(n)$.

4.2.1 Proposition

For $m \leq d$, $D^m[n^d] = \frac{d!}{(d-m)!}n^{d-m} + O(n^{d-m-1})$.

Proof. Base case: $D^0[n^d] = \frac{d!}{d!}n^d + O(n^d - 1) = n^d$. ✓.
Induction:

$$
\begin{aligned}
D[D^{m-1}n^d] &= D\left[\frac{d!}{(d-m+1)!}n^{d-m+1} + O(n^{d-m})\right] \\
&= \frac{d!}{(d-m+1)!}D[n^{d-m+1}] + O(D[n^{d-m}]) \\
&= \frac{d!}{(d-m+1)!}\left(n^{d-m+1} - (n-1)^{d-m+1}\right) + O(n^{d-m} - n^{d-m} + O(n^{d-m-1})) \\
&= \frac{d!}{(d-m+1)!}\left[n^{d-m+1} - \sum_{k=0}^{d-m+1}\binom{d-m+1}{k}n^k(-1)^{d-m+1-k}\right] + O(n^{d-m-1}) \\
&= \frac{d!}{(d-m+1)!}\left[\underbrace{n^{d-m+1} - n^{d-m+1}}_{=0} - \sum_{k=0}^{d-m}\binom{d-m+1}{k}n^k(-1)^{d-m+1-k}\right] + O(n^{d-m-1}) \\
&= \frac{d!}{(d-m+1)!}\left[\sum_{k=0}^{d-m}\binom{d-m+1}{k}n^k(-1)^{d-m+2-k}\right] + O(n^{d-m-1}) \\
&= \frac{d!}{(d-m+1)!}\left[(d-m+1)n^{d-m} + \underbrace{\sum_{k=0}^{d-m-1}\binom{d-m+1}{k}n^k(-1)^{d-m+2-k}}_{O(n^{d-m-1})}\right] + O(n^{d-m-1}) \\
&= \frac{d!}{(d-m)!}n^{d-m} + O(n^{d-m-1})
\end{aligned}
$$

□

4.2.2 Corollary

If $\beta(n) = a_d n^d + \cdots + a_0$, then

$$D^d[\beta(n)] = D^d[a_d n^d + \cdots + a_0] = d!a_d$$

4.2.3 Theorem 1

If $D^d[\beta(n)] = c$ for some constant c (constant in n), then $\beta(n)$ satisfies a recurrence relation.

Proof. We will prove something stronger. We'll prove that

$$D^d[\beta(n)] = \sum_{k=0}^{d} (-1)^k \binom{d}{k} \beta(n-k)$$

by induction on d. The base case is trivial:

$$D^0[\beta(n)] = (-1)^0 \binom{d}{0} \beta(n-0) = \beta(n).$$

Now using the hypothesis,

$$\begin{aligned}
D[D^{d-1}\beta(n)] &= D\left[\sum_{k=0}^{d-1} (-1)^k \binom{d-1}{k} \beta(n-k)\right] \\
&= \sum_{k=0}^{d-1} (-1)^k \binom{d-1}{k} D[\beta(n-k)] \\
&= \sum_{k=0}^{d-1} (-1)^k \binom{d-1}{k} [\beta(n-k) - \beta(n-k-1)] \\
&= \sum_{k=0}^{d-1} (-1)^k \binom{d-1}{k} \beta(n-k) - \underbrace{\sum_{k=0}^{d-1} (-1)^k \binom{d-1}{k} \beta(n-k-1)}_{(*)}.
\end{aligned}$$

Let's focus on this $(*)$-term, letting $k' = k+1$

$$\begin{aligned}
(*) &= -\sum_{k=0}^{d-1} (-1)^k \binom{d-1}{k} \beta(n-k-1) \\
&= -\sum_{k=0}^{d-1} (-1)^k \binom{d-1}{k} \beta(n-(k+1)) \\
&= -\sum_{k'=1}^{d} (-1)^{k'-1} \binom{d-1}{k'-1} \beta(n-k') \\
&= +\sum_{k=1}^{d} (-1)^k \binom{d-1}{k-1} \beta(n-k)
\end{aligned}$$

Putting the $(*)$-term back into the $D^d[\beta(n)]$ sum, and noting the binomial identity $\binom{d}{k} =$

$\binom{d}{k-1} + \binom{d-1}{k-1}$,

$$\begin{aligned}
D^d[\beta(n)] &= \sum_{k=0}^{d-1}(-1)^k\binom{d-1}{k}\beta(n-k) + \sum_{k=1}^{d}(-1)^k\binom{d-1}{k-1}\beta(n-k) \\
&= \beta(n) + \sum_{k=1}^{d-1}(-1)^k\left[\binom{d-1}{k} + \binom{d-1}{k-1}\right]\beta(n-k) + (-1)^d\beta(n-d) \\
&= \beta(n) + \sum_{k=1}^{d-1}(-1)^k\binom{d}{k}\beta(n-k) + (-1)^d\beta(n-d) \\
&= \sum_{k=0}^{d}(-1)^k\binom{d}{k}\beta(n-k)
\end{aligned}$$

which finishes our subproof. □

The overall proof follows immediately because if $D^d[\beta(n)] = c$, then

$$\begin{aligned}
&\sum_{k=0}^{d}(-1)^k\binom{d}{k}\beta(n-k) = c \\
\iff &\beta(n) = c - \sum_{k=1}^{d}(-1)^k\binom{d}{k}\beta(n-k)
\end{aligned}$$

is a recurrence relation for $\beta(n)$. □

4.2.4 Theorem 2

If $\beta(n)$ is a polynomial, $\beta(n) = \sum_{i=0}^{d} a_i n^i$, then $D^d[\beta(n)] = c$ for some $d \in \mathbb{Z}$.

Proof. This is simply because we have proven previously that $D^d[\beta(n)] = d!a_d$ □

4.2.5 Summary

Recalling that $\beta(n)$ satisfying a recurrence relation means that G (for which $\beta(n)$ is the sphere count) has rational growth (from notices), this gives us the following summary:

1. polynomial $\subset$ recurrence $\iff$ rational growth.
2. polynomial $\Rightarrow D^d(\beta(n)) = c$ for some d
3. rational $\iff$ recurrence $\Leftarrow D^d[\beta(n)] = c$ for some d

4.3 Detecting Polynomials

Let $\beta(n)$ be eventually quasi-polynomial with degree d. Recall that this means that there is a $T > 0$ and a list of polynomials $\beta_1, \ldots, \beta_N$ such that $\beta(n) = \beta_k(n)$ for $k = n \mod N$, with $n > T$. Then, for $n > T$,

$$D^d[\beta(n)] = \beta_i(0)$$

or equivalently, $D^d[\beta(n)]$ is eventually periodic.

It is known that this is the behavior of $H(\mathbb{Z})$ with generating set $S = \{a, b\}$. We want to know more about the growth behavior of $H(\mathbb{Z})$ with $S = \{a, b, c\}$, and $S = \{a, b, ab\}$. In particular, we want to be able to generate these numbers $\beta(n)$, and take their derivation to get the $\beta_i(0)$'s, and look for a periodic behavior up to some finite n.

Note that this is a well-defined question because polynomial growth is a well-defined property of groups, and therefore knowing the behavior of $H(\mathbb{Z})$ with generating set $S = \{a, b\}$ gives us the coarse behavior of $H(\mathbb{Z})$ with any other generating set.

4.3.1 Philosophical Question

How many coefficients $\beta_i(0)$ does it take for us to be convinced that we have found the period of $\beta(n)$?

Current bets given the guess that the period is 60:
Moon: 200
Andrew: 1000
Hang Lu: 20 000

Note that the number seems to be growing exponentially with relative novelty to the problem.

4.4 A Polynomial Time Algorithm For Sphere Count

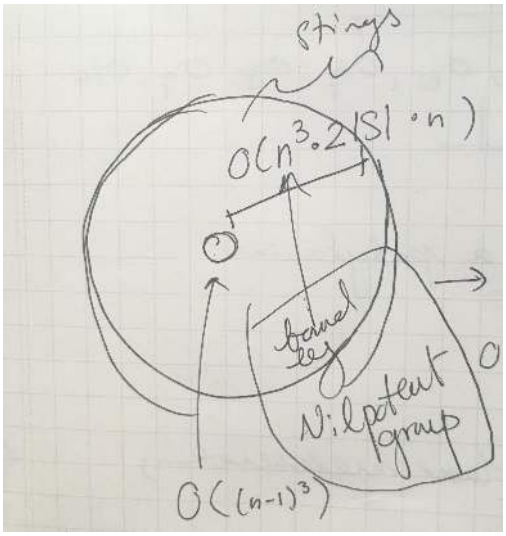
Naively, given a generating set S, and a ball of radius n, we expect the ball algorithm to take at least $O((2|S|)^n n)$ in time-complexity, and $O((2|S|)^n)$ in space-complexity.

Indeed, for each string of length n, we can take for each letter a generator $g \in S$ or its inverse, g^{-1}. Since there are n letters to pick, this gives $(2|S|)^n$ possible strings. Now we need to put each string in its word equivalence class by Malcifying it.

Optimally, Malcifying is $O(n)$ because we need to at least read the string to process it. Since we *have such an optimal algorithm(!!!)* (see Fast Malcifier document), the total run time is at least $O((2|S|)^n n)$, assuming insertion of each word in the right data-structure is $O(1)$.

However, because we are dealing with a *nilpotent group* the work of Bass and Guivarc'h [1] have shown that the growth is in the polynomial range. We can exploit this characteristic by Malcifying every string in every shell. Suppose we have done this for the shell of radius r. Then, there should be polynomially many words found in that shell of radius r, which we can multiply by $2|S|$ letters to get to the relevant strings of length at most $r + 1$ (I say at most because if we have something like $a \cdot a^{-}1$ in the shell of radius 2, this is clearly a string of length 0, not 2).

So if we Malcify the words, we know there will be *polynomially many of them* in each shell! Since we know the Heisenberg group has $O(n^3)$ for $n >> 1$ of them in each shell, which we Malcify optimally in time $O(n)$, this gives us $O(n^4)$ in time complexity and $O(n^3)$ in time complexity for each shell. Since there are n shells in a ball of radius n, we get $O(n^5)$ in time-complexity and $O(n^4)$ in space-complexity. Not bad!



4.5 Polynomial Interpolation

Suppose we have a sphere count of period N, and want to find the quasipolynomial involved. Since we know that each polynomial is of degree 3 [1], all we have to do is solve the following systems of equations, given by

$$p_i(n) = a_3^{(i)} n^3 + a_2^{(i)} n^2 + a_1^{(i)} n + c_i = \sigma_n$$

where i is taken $\mod N$. Or to be more concrete:

$$\begin{aligned}
p_i(n) &= a_3^{(i)} n^3 + a_2^{(i)} n^2 + a_1^{(i)} n + c_i = \sigma_n \\
p_i(n+N) &= a_3^{(i)} (n+N)^3 + a_2^{(i)} (n+N)^2 + a_1^{(i)} (n+N) + c_i = \sigma_{n+N} \\
p_i(n+2N) &= a_3^{(i)} (n+2N)^3 + a_2^{(i)} (n+2N)^2 + a_1^{(i)} (n+2N) + c_i = \sigma_{n+2N} \\
p_i(n+3N) &= a_3^{(i)} (n+3N)^3 + a_2^{(i)} (n+3N)^2 + a_1^{(i)} (n+3N) + c_i = \sigma_{n+3N}
\end{aligned}$$

gives us 4 equations to solve for the 4 unknowns $a_3^{(i)}, a_2^{(i)}, a_1^{(i)}, c_i$. We need N such systems of equations to find all the p_i's, $i \in [1, N]$.

Thus, for an N periodic quasipolynomial, we need the set

$$\{\sigma_1, \ldots, \sigma_{4N}\}$$

to interpolate the quasipolynomial.

A similar approach for the sphere count of the Heisenberg group with standard generators (a, b) yield's Shapiro's polynomial:

$$\sigma_n = \frac{1}{18}\left(31n^3 - 57n^2 + 105n + c_n\right)$$

where $c_n = -7, -14, 9, -16, -23, 18, -7, -32, 9, 2, -23, 0$ repeating with period 12.

The technique has been modified from above by knowing that only the constant term oscillates. Then, we have instead only one system of equation, with 3 unknowns being the a's, and 12 being the c_n's. Thus, 15 data points were sufficient to recover this polynomial.

4.5.1 Wait a Minute!

To find the period of the hex generator, we can't really interpolate the polynomial, since *we don't know the period to start with*!

What follows is the heuristics we are going to use to come up with a conjecture for what this polynomial is.

4.6 Guessing the Period

Here is an algorithm to guess the period:

Let N be the period. State an upper bound K for the period. Start with $N = 1$. Interpolate with the technique above. Then do $N = 2$...until $N = K$

Then look at the unused sample points to see how well each quasipolynomial interpolates with these.

Note: the larger K is, the more sample points we need. Also note that it may be that the actual period is larger than our K, which we might never notice if we fix a finite amount of unused sample points.

4.7 Back-of-the Envelope Calculations and Bad News

We stated above that we had a polynomial algorithm for the sphere count. After running said algorithm, we realized that *polynomial does not mean good!* Indeed, the program would keep crashing around $n \in [20, 30]$ on my Macbook from the lack of memory. Let's see why:

The words (there were $O(n^4)$ of them) were stored in their Malcev coordinates, so as `(int,int,int)` which takes 80 bytes. For each word, we would generate new strings from it by right-multiplying by a generator in the generating set S. Thus,

$$\texttt{Memory} > 2|S| \times 80 \times n^4 \ \texttt{bytes}$$

making the approximation $2|S| \approx 10$ and $80 \approx 100$, we get

$$\texttt{Memory} \geq 1000 \times n^4 \ \texttt{bytes}.$$

Since my Mac runs on 4 GB RAM, $\texttt{Memory} = 4 \times 10^9 \ \texttt{bytes}$ in this case. Thus

$$\begin{aligned} & 4 \times 10^9 \ \texttt{bytes} \geq 1000 \times n^4 \ \texttt{bytes} \\ \iff & n \leq 45 \end{aligned}$$

Assuming my RAM did not dedicate all of its memory to this process, and that I did not account for all of the memory use, this is pretty accurate.

Now, suppose we wanted to interpolate the hex-quasipolynomial, presumable of period 60. Then, we would need $n = 4 \times 60 = 240$ to at least interpolate such a quasipolynomial. Now,

$$1000 \times (240)^4 \texttt{bytes} = 3.3 \texttt{ TB}$$

Yup. That's right. Thats 3300 `GB`. Of RAM. That's a lot.

4.7.1 Key Point From This

So remember when we said that polynomial-time was good? Well, there are different kinds of polynomials. For n large, $O(n)$ is very far from $O(n^2)$, let alone $O(n^4)$. In industry, higher than $O(n^2)$ is actually considered not-efficient, and we're starting to see why.

Another point to note is $O(n^4)$ *memory space is so bad*. Indeed, it doesn't just take longer; the program literally stops running after it runs out of memory. In other words given the choice between suboptimal time and suboptimal memory, remember this: *We can't physically run out of time, but we can run out of memory. And when we do, the program crashes.*

4.8 Silver-Lining: an $O(n^3)$ memory-space algorithm

Depending on the set of generators, we can keep only certain layers of the sphere to avoid double-counting. For hex and standard, the data suggests that for S_{n+1}, S_n, S_{n-1} may be enough.

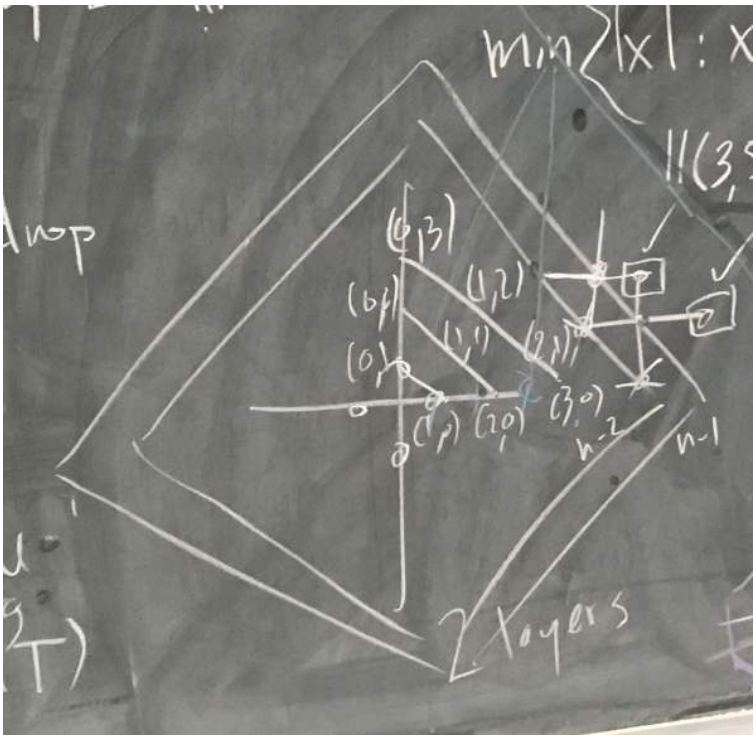
$$\begin{aligned} \texttt{Memory} &= 2|S| \times 3n^3 \times 80 \texttt{ bytes} \\ &\approx 3000n^3 \texttt{ bytes} \end{aligned}$$

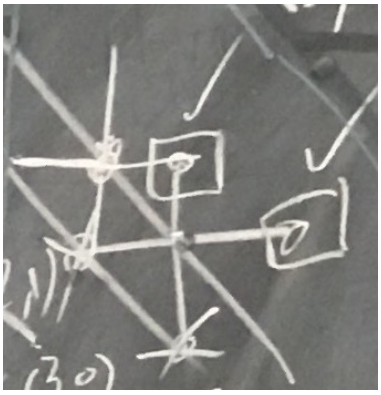
Plugging in $n = 240$, this gives us we need > 41.2 `GB` of RAM. Still a lot, but not nearly as much.

But how does the number of layers we can keep depend on the generating set? This is the next question we want to investigate.

4.9 Geometric Intuition

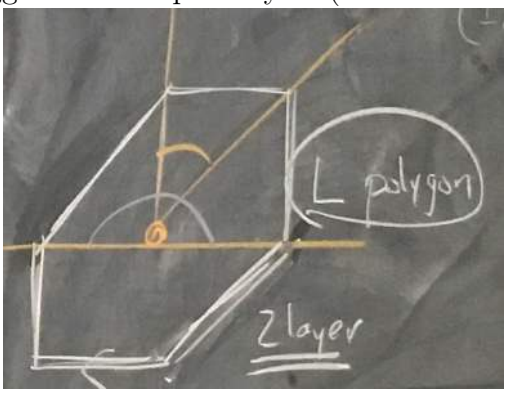
We can imbed the convex hull of the generators in $\mathbb{Z}^2$ (suppose we're not dealing with c's at the moment). Then, for the standard generating set, it looks something like this:





The zoomed in part shows how the generators pushes the words of the S_n sphere. This corroborates with the data.

Another example is the convex hull generated by the hex generators. The data also suggests that 2 past layers (hence 3 in total are sufficients.

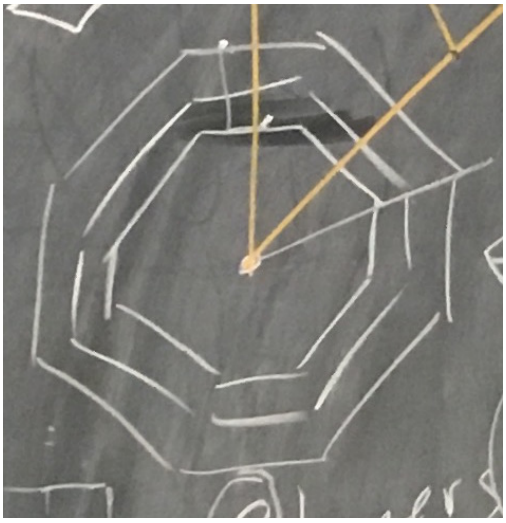


An interesting feature of both of these is that $|x|_S = \|x\|_L$, i.e. the word metric is equal to the norm induced by taking the convex-hull as the unit sphere. More precisely,

$$\|x\|_L = \lambda : x \in \lambda L$$

Note that this is unique because L has the thickness of *line* and expands in all directions :P.

An example of when $|x|_S \neq \|x\|_L$ is for the chess generators:



Indeed, the $\|(1,0)\|_S = 3$ (it takes 3 moves to get to $(1,0)$), but $\|(1,0)\|_L = \frac{1}{2}$

It also appears that we need 3 past layers (so 4 in total) to keep track of chess. Interesting!

Now, we would like to be able to quantify this for all generators. I mean literally, it could be just the max of the coordinates it takes the neutral element to. So like, for the chess it would be $2+1$ because chess is generated by $(\pm 1, \pm 2), (\pm 2, \pm 1)$, whereas for hex it's 1 because $a = (1,0), b = (1,0), ab = (1,1)$. Similarly for standard. We will see.

Another useful thing to note is that these spheres (convex hulls) admit symmetries. For certain, we only need to record one word, and multiply by two for the inverse. But further than that, these convex hulls admit many more symmetries that would be useful to exploit.

4.10 Guessing the Period: A Heuristic

If σ_n, σ_{n-1} are part of a constance sequence, then $\sigma_n - \sigma_{n-1} = 0$. Similarly, the σ_n, σ_{n-1} are part of some linear polynomial $p(n) = a_1 n - a_0$, then $\sigma_n - \sigma_{n-1} = a_1$ and so on.

If we have a quasi-polynomial, we want to keep cancelling things together and see where we get 0. How? Up to me to figure out this week!

References

[1] Moon Duchin. Counting in groups: Fine asymptotic geometry. *Notices of the AMS , Volume 63 Issue 8*, September 2016.

[2] Moon Duchin and Christopher Mooney. Fine asymptotic geometry in the heisenberg group, 2011.