Random Walks on Hyperbolic Space

Kunal Chawla

November 27, 2020

Contents

1 Introduction

Let's consider the following scenario:

Consider a person, call him Bill, lives at 0 on the number line.

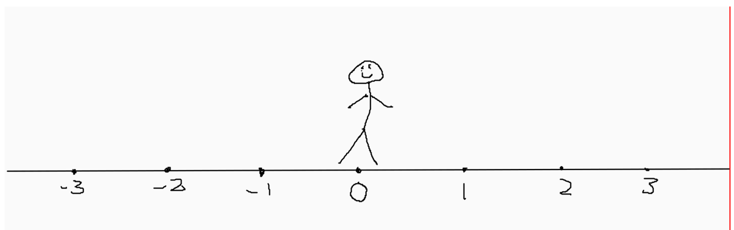


Figure 1: Bill taking a walk

Suppose that Bill walks out his front door, and flips a biased coin with probability $p > 1/2$ of coming up heads. If it comes up heads he walks to the right, and if it comes up tails he walks to the left.

If Bill does this infinitely many times one may wonder: will Bill come home?

There are multiple ways one may answer this question.

Let's see one way. Denote Bill's change in position after the ith time step by X_i. Then this is a random variable such that $\mathbb{P}(X_i = 1) = p$ and $\mathbb{P}(X_i = -1) = 1 - p$. Observe that each step is *i.i.d.*

Computing expectation we have

$$\mathbb{E}X_i = p \cdot 1 + (-1) \cdot (1 - p) = 2p - 1 > 0.$$

Then Bill's position after the nth step is simply the sum $\sum_{i=1}^{n} X_i$.

Therefore by the strong law of large numbers, which tells us that the average

$$\frac{X_1 + X_2 + ...X_n}{n} \to \mathbb{E}X_1 > 0.$$

Therefore Bill eventually leaves every single bounded subset of $\mathbb{Z}$.

Example 1. *The free group F_2*

Suppose Bill moves to Treeland. Since he just moved, he's not sure of his way around, so he's going to try and explore randomly for a while.

In other words, let's consider a random walk on the standard Cayley graph of the free group F_2.

Then observe that at each step of the walk, there are 3 choices to move away from the point and only 1 choice that moves the walker closer to the origin. Let X_i be a random variable taking on the value 1 if the ith step moves further away from the origin and -1 else.

Then Bill's distance from the starting point is equal to $1 + \sum_{i=2}^{n} X_i$ (because at the first step his distance is always increasing). In reality this isn't exactly the case, because if he returns then the distance will always increase at the next step, but either way this gives a lower bound for his distance from the origin.

However, by precisely the same argument we made before, as there is probability $\frac{3}{4} > \frac{1}{2}$ of moving further away, the strong law of large numbers tells us that eventually Bill leaves every bounded set forever.

One may ask - If the walker doesn't return to the origin infinitely often, where do they go?

2 Random Walks on Hyperbolic Spaces

Let us try and translate our two examples to the new setting, the hyperbolic plane $\mathbb{H}^2$.

To do this, we first need to slightly rephrase our previous problems. In the case of $\mathbb{R}$ we considered a sequence of translations, which one can view as simply an element of $Isom(\mathbb{R})$. Likewise in the case F_2 we can think of steps in the random walk as elements of F_2 acting on the Cayley graph.

So one way to translate this over to $\mathbb{H}$ is as follows: consider a group $G \leq Isom(\mathbb{H}^2)$ and a sequence of isometries $g_1, g_2, ...$ in G, then we can consider the walk generated by this sequence as the sequence of points $(w_n x_0)_{n\in\mathbb{N}}$ where $w_n = g_1...g_n$.

To work with this probabilistically we need a probability space to consider. So consider a measure μ on our group G. Then consider the set of sequences $G^{\mathbb{N}}$ equipped with the product measure $\mu^{\mathbb{N}}$. We let this be our probability space.

Observe that in both cases we used some form of the law of large numbers to say something about the asymptotic properties of our walk. To apply these laws we needed to express our change in distance from the basepint x_0 as a function of i.i.d random variables X_i. And the big hypothesis of the strong law is that the expectation $\mathbb{E}|X_i|$ converges absolutely.

So to have any hope of applying some form of law of large numbers, we need to require that the expected change in distance at each time step is finite. In other words, we require that

$$\int_G d(x_0, gx_0) d\mu(g) < \infty.$$

There is another common term for this, called the *finite first moment* condition.

Now it suffices to express our distance in a nice enough way such that we can apply some law of large numbers.

However, in both previous cases we could easily express the distance at each step as an incredibly simple random variable - which we could do because both the Cayley graphs of $\mathbb{Z}$ and F_2 are trees.

However, for the case of the hyperbolic plane, and some word $w_n = g_1 ... g_n$ that determines the nth step of Bill's walk, the quantity

$$d(x_0, w_n x_0)$$

is not necessarily a linear combination of the steps

$$d(x_0, g_1 x_0), ..., d(x_0, g_n x_0).$$

So how close can we get?

While we cannot necessarily decompose our walk completely as a sum, maybe we can 'partially' decompose it by applying the triangle inequality. Observe that by the triangle inequality, we have

$$d(x_0, w_{m+n} x_0) \leq d(x_0, w_m x_0) + d(w_m x_0, w_{m+n} x_0)).$$

If we write $w_{m+n} = g_1 g_2 ... g_{m+n}$, then as w_m is an isometry the second term in this expression is equal to $d(x_0, g_{m+1} ... g_{m+n} x_0)$.

We can express this more compactly by $(T^m w)_n$ where T is the left shift map.

So this may be helpful. Let's define the function $a(n, w) : \mathbb{N} \times X \to \mathbb{R}$ by

$$a(n, w) = d(x_0, w_n x_0).$$

Then rephrasing the paragraph above, we have shown that

$$a(m + n, w) \leq a(m, w) + a(n, T^m w).$$

Then the finite first moment condition is precisely the statement that $a(1, w)$ is integrable, and by a judicious application of the triangle inequality one can see that each $a(n, w)$ is integrable for any $n \in \mathbb{N}$.

The common term for this is a *subadditive cocycle*. Then we have a nice analogue of the law of large numbers for this function, also known as Kingman's subadditive ergodic theorem.

Theorem 1 (Kingman)**.** *Let $(X, \mathcal{F}, \mu)$ be a probability space and $T : X \to X$ a measure preserving transformation. Further, let $a(n, w)$ be a subadditive cocycle. Then if each $a(n, w)$ is integrable for any fixed n, we have*

$$\lim_{n\to\infty} \frac{a(n, w)}{n} = \hat{a}(w) \geq -\infty$$

for a.e. w, where $\hat{a}(w)$ is T-invarant. Furthermore, if T is ergodic, then $\hat{a}(x)$ is constant a.e..

By the paragraph above, our cocycle satisfies precisely the hypothesis of Kingman's subadditive ergodic theorem. Before we go further, I think it's best to take a quick excursion into ergodicity.

2.1 Ergodicity

Definition 1. *Let $(X, \mathcal{F}, \mu)$ be a probability space and $T : X \to X$ a measure preserving transformation. Then we say that T is ergodic if for any $A \in \mathcal{F}$, $T^{-1}(A) = A$ implies that $\mu(A) = 0$ or 1.*

Intuitively, an ergodic transformation is one that 'mixes' the whole system over time. Here's a good example:

Example 2. *Consider the circle S^1 with the Lebesgue measure, then rotation by an irrational angle α is ergodic.*

The intuition here is that rational rotations, say by $\pi/2$, preserve a proper subset of positive measure by periodicity. Irrational rotations, on the other hand , lack any sort of periodicity. There is a very slick generating-function type proof using this argument. The idea is to take a fourier expansion, which sort of 'counts periodicity' as opposed to standard generating functions which count combinatorial structures.

Proof. Let A be a set of positive measure that is invariant under this rotation, and consider the indicator 1_A. Then consider the fourier expansion

$$1_A(x) = \sum_{n \in \mathbb{Z}} a_n e^{nx}$$

Rotation by α is like multiplying this indicator by $e^{2\pi i \alpha}$, so as 1_A is invariant this tells us that

$$\sum_{n \in \mathbb{Z}} a_n e^{nx} = \sum_{n \in \mathbb{Z}} a_n e^{n(x+\alpha)}.$$

However as α is irrational this implies that every coefficient other than a_0 is equal to 0. Therefore the indicator 1_A is constant a.e.. As it is 1 on a set of positive measure then 1_A is equal to 1 almost everywhere. Therefore $\mu(A) = 1$. □

Another very pertinent example is the shift transformation T, defined above. This transformation is ergodic.

Before proving this, we use the following useful characterization of ergodicity:

Lemma 1. *Let $(X, \mathcal{F}, \mu)$ be a probability space and $T : X \to X$ a measure preserving transformation. Then T is ergodic if and only if for all measurable $A, B \in \mathcal{F}$ we have*

$$\lim_{n \to \infty} \frac{1}{n} \sum_{i=0}^{n-1} \mu(T^{-i} A \cap B) = \mu(A)\mu(B).$$

To see why this implies ergodicity, observe that if A is a measurable set with $A = T^{-1}(A)$ then each term in the left hand side becomes $\mu(A)$. Therefore the statement becomes

$$\mu(A)\mu(A) = \lim_{n \to \infty} \frac{1}{n} \sum_{i=0}^{n-1} \mu(T^{-i} A \cap A) = \lim_{n \to \infty} \frac{1}{n} \sum_{i=0}^{n-1} \mu(A) = \mu(A).$$

Therefore $\mu(A)^2 = \mu(A)$ so that A has measure 0 or 1.

Intuitively, the statement is about independence. If A is an event that occurs with positive probability, then $T^{-i}A$ sort of gets 'mixed around' as i goes to infinity. Then averaging out the probability that both B and $T^{-i}A$ over all i, one should expect this to be $\mu(A)\mu(B)$.

Now we can show that the shift map is ergodic.

Proof. It suffices to show that this is true if A and B are cylinder sets, because the σ-algebra is generated by cylinders.

Then in this case there exist some k such that for all $k' > k$ that $T^{-k'}(A)$ is all of G only in coordinates where B is not, and vice versa. Because we are taking the product measure on our space, then for such a k' we have $\mu(T^{-k'}A \cap B) = \mu(A)\mu(B)$. Therefore

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} \mu(T^{-i}A \cap B) = \mu(A)\mu(B),$$

so by the characterization of ergodicity we are done. $\square$

.

Now that we've covered some of the basic notions of ergodicity, we can return to the question of random walks

2.2 Applying the subadditive ergodic theorem

As the shift map T is ergodic, we know by the subadditive ergodic theorem that the distance

$$\lim_{n\to\infty} \frac{d(x_0, w_n x_0)}{n}$$

converges almost everywhere to some constant $\hat{a} \geq -\infty$.

As our cocycle is defined in terms of distances, it is in particular nonnegative. So either $\hat{a} = 0$ or $\hat{a} > 0$.

Consider the case where $\hat{a} = 0$. Here we cannot necessarily say much about the random walk.

However, we can give some examples of groups G and measures μ in which $\hat{a} = 0$.

First, consider the case where our measure μ is simply the dirac mass δ_g where g is some finite-order elliptic. Then any random walk must go through at most finitely many points, so we clearly have

$$\lim_{n\to\infty} \frac{d(x_0, w_n x_0)}{n} = 0.$$

Second, consider $\mathbb{H}$ as the upper half plane,where our measure μ is the uniform measure on $\{g, g^{-1}\}$ and g is the transformation $z \to z + 1$.

Then this is simply a random walk on $\mathbb{Z}$ simply embedded into $\mathbb{H}^2$. In this case, our expected distance from x_0 is 0, so we have $\hat{a} = 0$.

Observe that in both of these examples, the group generated by the support of μ is an elementary subgroup of $Isom(\mathbb{H}^2)$.

Now consider the second case, where $\hat{a} > 0$. Then knowing that

$$\lim_{n\to\infty} \frac{d(x_0, w_n x_0)}{n} > 0$$

for almost every sample path tells us that eventually Bill leaves every bounded set.

So we have a success! Using an analogue of the law of large numbers, we have shown the same positive drift phenomenon that we have for our other examples.

However, we can say more. Observe that $\mathbb{H}^2 \cup \partial\mathbb{H}^2$ is topologically a disk, and therefore compact. Then our random walk, which is in particular an infinite sequence, must have a convergent subsequence somewhere in the disk.

However, as Bill eventually leaves every bounded set, we know that such a limit point cannot lie in $\mathbb{H}^2$, so we can deduce that if $\hat{a} > 0$ then almost every sample path has a limit point in the Gromov boundary!

To make this argument, we only used the fact that $\mathbb{H}^2 \cup \partial\mathbb{H}^2$ is compact, which is a result that holds for arbitrary locally compact δ-hyperbolic space. For the sake of simplicity, however, we continue to make our argument in $\mathbb{H}^2$.

So we now have a limit point in the boundary $y \in \partial\mathbb{H}^2$. We sketch a proof that such a point y is almost always unique.

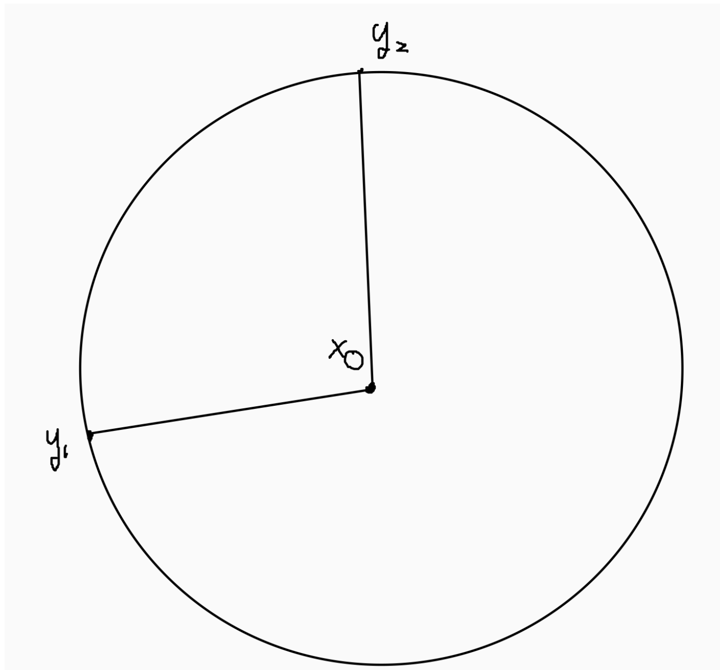
Indeed, suppose that we have two limit points in the boundary:

Then consider the geodesics from x_0 to each limit point y_1 and y_2. Then we have two convergent subsequences x_k and x'_k to each limit point, which must eventually be within bounded distance to each of our geodesics.

Recall that as $\hat{a} > 0$ and Bill eventually leaves every bounded set in $\mathbb{H}^2$. Also recall one important property of δ-hyperbolic spaces, that geodesics which diverge must diverge exponentially fast.

Therefore if Bill walks from one geodesic to another outside of every ball around x_0, then he must make exponentially longer and longer treks back and forth. However by the finite first moment condition, this implies that he must make an infinite sequence of treks such that the probability of each treks decays exponentially.

Therefore by the Borel-Cantelli lemma, the probability that he makes infinitely many treks is 0.



To recap, we have shown that using some tools from dynamics we can get interesting geometric results about random walks on hyperbolic spaces. Here we have a correspondence between geometry and random walks, and also a correspondence between dynamical systems and probability. This, in my opinion, is a really cool thing. Just to give a hint of what else one can show, we state a theorem of Maher and Tiozzo:

Theorem 2. *Let G be a countable group of isometries of a separable Gromov hyperbolic space X, and let μ be a non-elemetary probability distribution on G. Then for any basepoint $x_0 \in X$, almost every sample path $(w_n x_0)_{n\in\mathbb{N}}$ converges to a point $\omega_+ \in \partial X$. The resulting hitting measure is non-atomic, and is the unique μ-stationary measure on ∂X.*

3 Sublinear tracking

One can go even further than showing convergence to the boundary. Observe that in the case of a biased random walk on $\mathbb{Z}$, one could rephrase the strong law as saying that

$$\frac{|X_1 + ... + X_n - \gamma(\lambda n)|}{n} \to 0,$$

where $\lambda = \mathbb{E}X_1$ is a constant and $\gamma(t) = t$ is a unit-speed geodesic.

This is essentially saying that we can track the random walk by a geodesic with sublinear error term. This is called *sublinear tracking.*

Observe that we can extend this to more hyperbolic spaces. Let's consider the same random walks on $\mathbb{H}^2$. Unlike before, where we consider infinite sequences $G^{\mathbb{N}}$ we consider bi-infinite sequences in $G^{\mathbb{Z}}$.

Similar to our previous construction we consider a forward random walk and a backwards random walk. Our forward walk is given by

$$w_n = g_1 ... g_n$$

while our backwards random walk is given by $w_{-n} = g_{-1}...g_{-n}$.

Let consider a function $\varphi(w)$ of our sample path, as given by the distance between our basepoint x_0 and the bi-infinite geodesic between the limit points ω_+ and ω_-.

Then suppose that we have convergence to the boundary just like before, for both directions.

Then our picture is as in figure 2 on page 12

We want to say that we can track this sequence by a bi-infinite geodesic with sublinear error term. Then the only possible candidate must be the unique geodesic between the two endpoints ω_+ and ω_-.

One very nice proof of this result is in Giulio Tiozzo's paper 'Sublinear deviation between geodesics and sample paths'. We include the proof here with added commentary, partly because it aids my understanding and partly because I think it's just a really cool argument.

We can prove this using some ergodic theory. Before summarizing the argument, we recall the Birkhoff ergodic theorem:

Theorem 3. *If $(X, \mathcal{F}, \mu)$ is a probability space and We recall the Birkhoff ergodic theorem, that if $T : X \to X$ is an ergodic transformation on a probability space X. Then if we have any integrable function f, the time average*

$$\frac{1}{n}\sum_{i=0}^{n-1} f(T^i(x))$$

converges almost everywhere to the space average $\int_X f d\mu$.

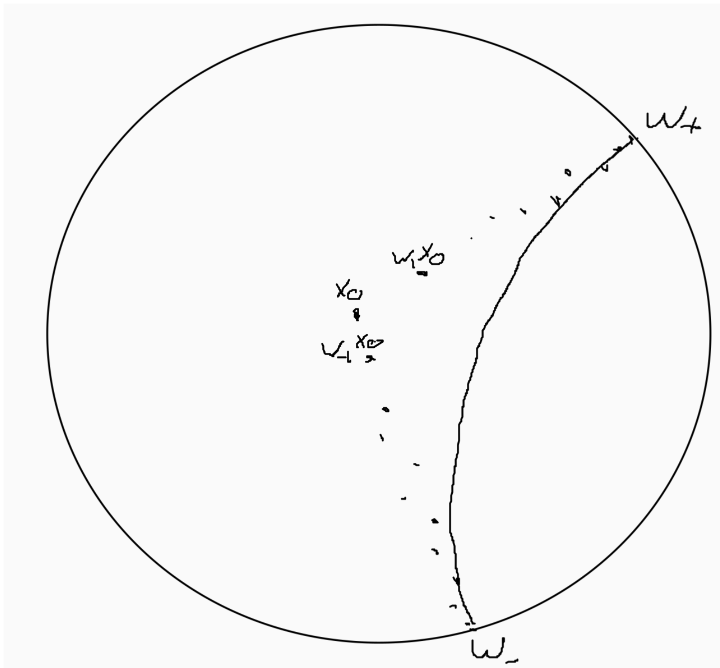


Figure 2: A bi-infinite sample path with two limit points ω_- and ω_+, as well as the geodesic between these two points as infinity.

This can be seen as a vast generalization of the strong law of large numbers. It is an incredibly strong result. One reason it is so strong is that it applies to *any* integrable function, even the most pathological. Therefore many applications it suffices to choose the right function, and then apply the Birkhoff ergodic theorem.

Now we consider the function $\varphi(w)$ that is given by the distance between x_0 and thelimit points of the random walk given by w. We want to try and apply the ergodic theorem to show that $\varphi(T^n w)$ grows sublinearly.

So what is the value of $\varphi(T^n w)$? Observe that ω_+ is equal to the limit of the sequence $g_1 g_2 ... g_n x_0$, so the new endpoint is given by $g_i g_{i+1} ... g_{i+n} x_0$, which is precisely the same as $w_i^{-1}\omega_+$.

Therefore we have

$$\varphi(T^i(w)) = d(x_0, [w_i^{-1}\omega_-, w_i^{-1}\omega_+]).$$

However as w_i is an isometry this is equal to

$$d(w_i x_0, [\omega_-, \omega_+]).$$

Therefore geometrically we know that $\varphi(T^i w)$ is precisely the distance between the ith point in our random walk and our candidate geodesic.

So to prove the thing we want it suffices to show that

$$\frac{\varphi(T^n w)}{n}$$

goes to 0 as $n \to \infty$.

We can use the following lemma:

Lemma 2. *Let $f : X \to \mathbb{R}$ be a non-negative, measurable from the probability space X, and let $T : X \to X$ be measure preserving and ergodic. Now suppose that the difference $g(w) = f(T(w)) - f(w)$ is integrable. Then the limit*

$$\frac{f(T^n(w)}{n}$$

tends to 0 almost surely as $n \to \infty$.

Proof. We consider the Birkhoff sum

$$S_n g := g(w) + g(Tw) + ... + g(T^{n-1}w).$$

Then as g is integrable, the Birkhoff ergodic theorem tells us that

$$\lim_{n\to\infty} \frac{S_n g}{n} = \int_X g d\mu.$$

However, observe that the sum $S_n g$ telescopes to give

$$S_n g = f(T^n w) - f(w)$$

by the definition of g.

Then we can rephrase the previous limit as

$$\lim_{n\to\infty} \frac{f(T^n w)}{n} - \frac{f(w)}{n} = \int_X g d\mu.$$

However as the second term on the LHS dies out we know that

$$\lim_{n\to\infty} \frac{f(T^n w)}{n} \to \int_X g d\mu$$

.

In particular, the limit $\lim_{n\to\infty} f(T^n w)/n$ exists.

On the other hand, there exists some constant C such that the set $O_C = \{w | f(w) \leq C\}$ has positive measure. Then as T is measure preserving and X has finite total measure, the Poincare Recurrence theorem tells us that there are infinitely many n such that $f(T^n w) \leq C$.

Therefore

$$\lim_{n\to\infty} \frac{f(T^n(w))}{n} = \liminf_{n\to\infty} \frac{f(T^n(w))}{n} \leq \lim_{n\to\infty} \frac{C}{n} = 0.$$

□

The power behind the argument is that ergodic theorems tell you that limits exist. After obtaining this fact, one can throw their toolbox at the limit prove whatever estimates are required, by considering lim inf's and lim sup's.

We can prove the claim using lemma. To do so, we must show that the function $f(Tw) - f(w)$ is integrable.

This quantity is precisely

$$d(w_1 x_0, [\omega_-, \omega_+]) - d(x_0, [\omega_-, \omega_+]).$$

Therefore by the reverse triangle inequality we can bound this by $d(x_0, w_1 x_0)$

However this function is integrable, because we already assumed that the function had finite first moment. Therefore by the lemma we know that the error term is sublinear, which is precisely what we wanted to show!

With some further arguments one can actually say more about the 'tracking', getting something similar to that expression $\gamma(\lambda n)$ at the beginning of the section, but here we just show that the distance from the geodesic is sublinear.

The lesson to take home here is that ergodic theory gives powerful tools through which we can interesting geometric statements. By viewing some quantity we are interested in as some function $f(T^i x)$ where T is ergodic, we can prove very strong statements about its asymptotic behaviour. Essentially, ergodic theorems and laws of large numbers tell us that nice enough random processes act the way one should expect them to, and give nice machinery for proving these statements.