[SQUEAKING]

[RUSTLING]

[CLICKING]

LAWRENCE GUTH: So this is the second day of our little unit on random walks, on groups. And the first day, we introduced it, and the second day, we'll describe some more modern stuff. And the second day will be a little bit more of a survey. I won't prove everything, but I'll try to-- and in the second day, we'll see how it connects to projection theory. OK, so first, let's recall where we're talking about.

So we have a finite group G. And then we have a probability measure on G. And we have an operator T mu. T mu of f is f convolved with mu. And so this does a step of a random walk. So if f is the probability distribution at some moment, then you make a step of the random walk, and then T mu f is the new probability distribution.

So T mu of 1 is 1, the constant function 1. And T mu maps the functions of mean 0. So these are functions who sum to 0 or have mean 0, it maps it to itself. And sigma 1 of T mu is the largest singular value of this.

And we saw that sigma 1 of T mu is related to the mixing behavior of this random walk. So we always have that sigma 1 of T mu is less than or equal to 1. And we don't use a positive number. And sigma 1 of T mu is strictly less than 1. Then that gives some kind of mixing of the random walk. We made that precise last time.

Next, we focused on the group SL2Fp. And one special feature of this group came from representation theory. So we had a proposition that if you have a representation of this group, so that's a unitary representation, and let's say it's non-trivial, then D has to be pretty big.

And we use that to prove a mixing estimate. Mixing estimate says that if mu is a probability measure on this group, then sigma 1 of T mu is bounded by p squared times mu L2 squared. For example, if A is just a subset of SL2Fp, then there's a uniform measure on A, mu A.

And it's straightforward to check that its L2 norm squared, mu A L2 norm squared is 1 over the size of A. So a corollary is that if A is significantly bigger than p squared, then the random walk using A mixes rapidly.

So corollary, if for instance, A is bigger than p to the 2.1, then sigma 1 of TA, so I'll write TA is an abbreviation for T mu A, sigma 1 of TA would be smaller than something like p to the negative 1/10. So it would mix quite rapidly.

So this is a nice proposition. And it's kind of sharp of its type. Because if the set A is a subgroup of SL2Fp, then the random walk will not generate, will not mix at all, and there are subgroups, or easy subgroups that have size around p squared. So we can't do better than this.

So remark is that this corollary is sharp because there exists H, a proper subgroup size of H around p squared. So for instance, H is the upper triangular matrices. Yeah?

AUDIENCE: If you have, for a different group, a similar proposition where you can lower bound the dimension of the nontrivial unitary representation, could you also use that as an upper bound for the sigma 1 for that group?

LAWRENCE GUTH: Yeah, so the question is if we had an analog of this proposition for some other group, say SLDFp, then could we do this whole story for all these other groups? Yeah, you could. That's all that we've used so far. Yeah, OK, cool.

So then also last class, we stated an old theorem, which goes rather-- it's older than this proposition, but it goes rather, it does not obviously follow from it. So theorem, this theorem is essentially due to Selberg, although he didn't say it in quite this way. So here, A, so we pick A is a particular list of generators of SL2Fp.

So this is perhaps the simplest set of generators of SL2Fp. But it is also somewhat special in a way that we'll talk about when we talk about the proof. So then the conclusion is that sigma 1 of TA is less than 1 minus c, where c is positive and c is-- this is true for every p. So c does not depend on p. So these all mix uniformly fast.

So this theorem is not an immediate corollary of this proposition because this is a much smaller set. And so today, a couple of goals in this survey. The first goal is to sketch a proof of this theorem in which this proposition plays a key role. But we'll see some other ideas. And a second goal is to discuss what happens if you replace what was special about this generating set at all, and what happens if we replace it by some other generating set.

OK, so the first proof was due to Selberg, basically. But that proof was quite difficult. It used a bunch of stuff, including the Neyman hypothesis for curves over finite fields. We're going to follow a later proof which is due to Sarnak and Xua, and perhaps simplified or adjusted a bit by various people.

OK, cool. All right, so this set is symmetric, meaning that if a group element is in the set, so is the inverse. And so we could say that mu is symmetric if mu of g is equal to mu of g inverse for every g in G. And T mu, actually, it's a bit nicer if mu is symmetric.

And here's why. So remark, mu is symmetric. That is equivalent to saying that the operator T mu is symmetric as a matrix. And so because-- so if you think of T mu as a matrix, then it has a g1, g2 entry, and the g1, g2 entry is mu of g1, g2 inverse. So if you switch these, you're taking the inverse of that. And if mu is symmetric, that doesn't matter. So it's a symmetric matrix.

So that's nice because a symmetric matrix, you can diagonalize it. And instead of just having singular values, it has eigenvalues. So once T mu is a symmetric matrix, then the first singular value of T mu to the k is the first singular value of T mu to the k. It helps to have a blackboard to say it.

And so that suggests a-- this is helpful. This isn't true for non-symmetric matrices. But it's true for symmetric matrices. And it suggests a strategy of how to get a handle on sigma 1 of T mu. So we can say the following thing.

So we have sigma 1. Yeah, let's also think about T mu to the k. So T mu to the k of f is f convolved with mu and then convolved with mu and so on k times, which is f convolved with mu convolved with itself k times.

So that's an operator of our type. But with a new measure, the measure is mu convolved with itself k times. So therefore, sigma 1 of T mu to the k, sigma 1 T mu to the k, which is sigma 1 of T mu to the k.

So if you try to explain this without a blackboard, you have to have really good ability to evoke things with your voice. All right, so mu might be the uniform measure on this set of four elements A. And if we apply this proposition directly to that set of four elements, it gives us nothing.

But we can use this. And then we can try to apply the proposition to mu to the k. So if we put together what we have, we can make the following corollary, sigma of T mu to the k-- so OK, so this is if mu is a probability measure on SL2 of Fp, then sigma 1 of T mu to the k is bounded by, well, this thing. And for this thing, we can apply our proposition here.

So we get p squared times mu tensor k L2 squared. So for the original, if you take the Selberg measure A then for the original mu A, this thing would be much too big to be interesting. But perhaps we could find a high power k where this L2 norm is actually smaller than 1 over p squared. And that would give us an interesting bound.

Cool, so all of the stuff that we'll talk about today is based on this corollary. So you can think about it like this, you start off with a probability measure mu, and we want to know, does it get really evenly mixed if you convolve it with itself a lot of times? And what we'll see here is that maybe if you convolve it with itself a bunch of times, it will get fairly well mixed, not spectacularly well mixed, but just fairly well mixed to guarantee that this L2 norm is smaller than 1 over p squared.

So if it manages to be kind of roughly evenly distributed around p to the 2.1 elements in this group, that would be good enough that this would be less than 1. Once that happens, we can use this corollary and we can bound sigma 1. And once we bound sigma 1, then we'll see that that, in turn, tells us what happens as we keep mixing. Then we'll be able to say it gets really very well mixed at some later time.

OK, but how can we-- so how can we bound this? How can we show that mu tensor k is decently well mixed, fairly well mixed? So there are two different approaches that we'll talk about today. So one approach, which is special for this theorem, involves either-- let me say it involves lifting to SL2Z and/or it involves hyperbolic geometry. I'll describe it in a couple of slightly different ways.

So first let's talk about lifting to SL2Z. So if you reduce mod p, it gives a group homomorphism from SL2Z to SL2Fp. So let's suppose that M is a probability measure on SL2Z. And mu is the pushforward pi p of M. So that's a probability measure here.

And this setup with group homomorphisms commutes, plays nicely with convolution. So if I take the pushforward of M convolved with itself k times, I will get mu convolved with itself k times. So one approach to try to understand mu convolved with itself a bunch of times is to first try to understand M convolved with itself a bunch of times and then try to understand what happens when we perform this push forward.

So here's our plan, study M convolved with itself k times to study the push forward operation. So this is a reasonable strategy because of the fact, or to the extent that we feel that SL2Z is easier to get our hands on then SL2Fp.

And there are a couple of reasons that SL2Z is easier to get our hands on then SL2Fp. Let me call this good features of SL2Z. We won't necessarily use all of these, but let me mention them.

So the first feature is that SL2Z is virtually free. So that virtually means is that there is a subgroup H in SL2Z which is a finite index free subgroup. So in general, if you say that a group is virtually blah, it means that there's a finite index subgroup, which is blah. So it's pretty close to being free.

That's not true at all for SL2Fp. You take SL2Fp, take the simplest generators you can think of, try to write down all the relations, it's not super easy. I don't personally know how to do it, although I think people know how to do it.

OK, number two, SL2Z is sitting inside of SL2R. It's kind of a discrete approximation of SL2R, so it's pretty closely related to a Lie group. This is what we'll actually use. And number three is that SL2Z acts nicely-- maybe I should call it 2B, it's quite closely related to two.

It acts nicely on the hyperbolic plane. So recall that the isometries of the hyperbolic plane is PSL2R. So SL2R, basically SL2R, not quite. So SL2R acts by isometries on the hyperbolic plane. And SL2Z is sitting inside there, so it acts by isometries on the hyperbolic plane. And that action is captures a lot of SL2Z. You can also use this to help understand SL2Z. And so all three of these things are much more complicated for SL2Fp.

All right, so I don't-- so one of the things that I won't completely prove is some thing about doing convolutions on SL2Z. But I want to give some intuition and rough statement about it. So intuition about convolution on SL2Z.

So as a warm up, I want to think about convolution on Z. So let's say nu is a probability measure on Z. And I take nu convolved with itself k times. So nu is some fixed simple probability measure. Maybe let's take nu of x is 1/2 if x is plus minus 1 and 0 otherwise.

So what happens when we convolve it with itself k times? Well, the central limit theorem tells us that this thing is almost a Gaussian. Gaussian, also closely related to the heat kernel on R.

So in analogy with this, what might we hope to happen when we take a measure on SL2Z and we convolve it with itself a lot? So you might hope-- so we have M, measure, probability measure on SL2Z. So you might hope to have some kind of central limit theorem for matrices. And it would say that M convolved with itself a lot is sort of like the heat kernel, which needs to be defined, on SL2R.

OK, so this seems like a natural thing in the intersection of probability theory and matrix theory and has been studied, although I'm not an expert on it, but cf work of Furstenberg and. So I believe that I know roughly how it works out, although I don't have a precise statement. But let me tell you roughly how I think it works out. And then we'll see how this would be helpful to understand SL2Fp.

So first of all, we'll have some kind of balls in SL2R. So BT is the set of matrices A, B, C, D in SL2R with the property that a squared plus b squared plus c squared plus d squared is at most R squared-- is at most T squared.

So this is some of rough-- some sort of notion of a ball in the group SL2R. You could imagine a fancier notion where you start with the Lie algebra, you put a ball in the Lie algebra, and you take the exponential map. And that's probably a better thing to do than this. But this is a very low brow definition. And it is approximately describes balls.

And then if we're interested in SL2Z, we would take the integer points in this. BT of Z, that's defined to be BT intersected with SL2Z. So these are integers. So for reference, what is the cardinality of BT of Z?

So little lemma, cardinality of BT of Z is roughly T squared. Sketch, really choose a and d minus T, T. So that's T squared choices. And then we have to solve b times c is a times d minus 1. I think I did that right. So here, your SL2R, the determinant is 1, so ad minus bc is 1.

So we have this integer of size around T squared. And we want to know how many ways are there to factor it into two integers of size around T. OK, so not every integer will factor like this. It could be prime. But mostly, they will factor like that. And how many ways will they factor? Well, not too many ways, at most, T to the epsilon ways.

So number of choices for b and c is on the order of 1. So the level of our back of the envelope computation, there should be about T squared group elements in there. All right, so there is a vague statement that if you fix M, then M star to the k is roughly equidistributed on the ball of radius T for T around exponential of some constant depending on M times k. So it's not stated precisely, but this is roughly what I think the central limit theorem should say in SL2Z.

In a little bit we'll do the hyperbolic geometry point of view, where I'll be able to state things a little bit better if we want. OK, so that's step one of our plan, study what the convolution looks like in upstairs in SL2Z. Next, we need to think about what happens when we project from SL2Z down to SL2Fp.

Why don't we pause there, actually? I mean, I did some things not completely precisely, but let me see if people have any questions, or comments, or are bothered by things. Yeah?

AUDIENCE: Question about [INAUDIBLE] to the existence of the heat kernel that-- the heat kernel on R, and, in general, kind of Euclidean looking spaces is kind of closely related to the Fourier transform. But in a non-abelian group, is there anything analogous to that that makes this some sort of heat kernel-ish thing work?

LAWRENCE GUTH: Yeah so the question is, how does the heat kernel work in a non-abelian group? Let me punt on that until we get to the hyperbolic space version. And we'll talk about the heat kernel there, which is a little bit easier, at least for me, to make it precise. So let's do the second step, where we pass from SL2F-- from SL2Z to SL2Fp.

So remember our setup that M is a probability measure on SL2Z And then mu is the pushforward of M, which is probability measure on SL2Fp. Also, let's mention that gamma p could be the kernel of pi p. So that's a normal subgroup of SL2Z. And it's gamma p is the set of a, b, c, d in SL2Z, so that a, d, b, c is congruent to the identity mod p.

OK, great, so now our goal is to understand the L2 norm of some convolution of this. And here's a cute trick, little lemma. So if mu is symmetric, so for instance, if M is symmetric, then mu tensor k L2 squared is mu tensor 2k of the identity.

So here's the proof. Let's start on the other side. Mu tensor 2k of the identity is, well, mu tensor k convolved with-- so I was saying the word "tensor." That was not the right word. I mean convolved. Mu convolved k convolved with mu convolved k, the identity. So that's the sum over g in our group, mu convolved k of g mu convolved k of g inverse. That's how you convolve two functions.

But what's nice is that mu is symmetric is that these are the same as each other. So we just get the sum on g mu convolved k of g squared. So that's equal to that. OK, cool, so that's convenient, because now we just have to estimate this one thing. And mu convolved 2k of the identity is-- that's something that lives here, but we can relate it to something that lives here. It's the sum g in gamma p of M convolved 2k of g.

So in other words, we want to know how much mass-- what is the measure, using this measure M convolved 2k, what is the measure of the subgroup gamma p. All right, now M convolved 2k is approximately evenly distributed on some ball. So by the vague statement, this is equivalent to estimating what fraction of the ball is in gamma p.

So we have some big ball in our Lie group. What fraction of it is in the subgroup gamma p? So remark, the index of gamma p is the cardinality of SL2Fp, which is around p cubed. So if this were evenly distributed, then you would expect this quotient to be about 1 over p cubed.

And that's basically true. So here's another lemma. BT intersected with gamma p is bounded by T squared over p cubed, which is also like the size of BT of Z over p cubed. So it is pretty evenly distributed.

All right, so here's the proof. So we're looking for a, b, c, d, where each entry is size smaller than T. And they're integers. And we want this to be congruent to 1, 0, 0, 1, mod p.

So first of all, let's try to do a back of the envelope calculation. So if I erase this congruence condition, then I'm just asking how many integer points are there in BT. And we said before, that's about T squared.

Now, I'd like a to be congruent to 1 mod p. So maybe I should divide by p for that. And I'd like b to be congruent to 0, so maybe I should divide by p again. And I have four conditions. So maybe I should divide by p four times, T squared divided by P to the fourth. Is anybody suspicious?

OK, so that's the wrong answer. It should be closer to T squared divided by p cubed. And the reason that was not really OK to do is that those four conditions are not independent of each other. We already know that ad minus bc is 1 because we're in-- I didn't write this properly. This is supposed to be in SL2Z, which means that ad minus bc equals 1.

And I think that probably implies that once I have three of these congruence conditions, then I would get the fourth. But it also brings up the point that I shouldn't be too fast and loose. That it's not super obvious that all of these things are independent of each other.

So there is actually a kind of algebra trick to make this work that allows us to keep things more independent of each other. And here is the algebra trick. OK, so b is divisible by p. And c is divisible by p. So p divides b, and p divides C. Therefore, p divides bc. And bc is ad minus 1-- sorry, p squared divides bc. And so p squared divides ad minus 1. These are equal to each other.

All right, similarly, p plus 1 divides a-- sorry, p divides a minus 1. So a is congruent to 1 mod p. And p divides d minus 1. So p squared divides a minus 1 times d minus 1. All right, the leading term here is also a times d. So we can play these off against each other. And we conclude that p squared divides a plus d minus 2.

All right, now let's start choosing things. So we're going to choose a. And a has an integer of size at most T. And it's congruent to 1 mod p. So I have less than T over p choices. Next, I'm going to choose d.

I've already chosen a. So I now know the value of d modulo p squared because of this sneaky algebra trick. So I get, at most, T over p squared choices. This was only really OK if T is at least p squared. So let me say T is at least p squared.

And now, once I've chosen a and d, now I have to solve for b and c. And I forget about all the congruences, just like above, they're at most around 1 choices. And so now, they are less than 1 choice, choose b and c. That's because bc is ad minus 1 as integers. Yeah?

AUDIENCE: When you're saying there's kind of at most one factorization of ad minus 1, would that actually be a log of, to be precise, the number of factors?

LAWRENCE GUTH: Yeah, so the question is, how many factors are there if a number of some size? And it is a little super constant. So if this number has size n, It's actually more than log n. It could potentially could be worse than that. But it's less than n to the epsilon for any epsilon. And that's what this symbol was supposed to be.

AUDIENCE: A whole lot of symbols being slightly different versions of [INAUDIBLE].

LAWRENCE GUTH: Yeah, I know. So any other questions or comments about the lemma? So those are the ingredients of the Sarnak-Xua proof of theorem one. All the ingredients are rigorous and precise, except for the vague statement about how random walks behave on SL2Z. And if we're willing to believe that, then we can put them together and prove theorem one.

OK, so proof of theorem one. So M will be MA for this set A SL2Z. Then by the vague statement, M convolves k times is roughly equidistributed on BT, where T is around x of some constant times k. So now, we will choose k. So we'll choose T to be a little more than p squared. So that we can use the lemma at the end, which means that k is around log p.

So take around round log p steps, and after that, M convolves k is roughly evenly distributed on a ball of size T p squared. And I guess here we can put 2k, I guess. So then M star 2k of gamma p will be around BT intersect gamma p over BT, which is smaller than p to the minus 3 by our lemma. And here, we're losing a little power. So if I wanted to be more clear about what I was saying, that would be something like this.

Therefore, mu tensor k L2 squared, that's the same thing as m tensor 2k of gamma p. And so that would be around p to the minus 3 plus epsilon. In particular, significantly smaller than p to the minus 2. And then we use the corollary. That gives us a bound for this. And we get that sigma 1 of t mu is less than 1 minus c, where the c is independent of p.

OK, let me tell you just briefly about the hyperbolic geometry point of view. And then we'll switch to other choices of generators. OK, hyperbolic geometry, all right, so we mentioned that SL2Z or SL2R acts on the hyperbolic plane H2.

Let's say that X of p is H2 modulo gamma p. So X of 1 is the hyperbolic plane modulo SL2Z. And X of p is a cover of it. X of p is a cover of X of 1. And the group of deck transformations is SL2Fp.

OK, so I would like to try to make a picture of this. So X of 1 is not quite a manifold, but it's not super important for this except for being a technical irritation. So I'm not going to try to draw that. So it is almost a non-compact manifold of genus 1. It looks sort of like this. The non-compactness is annoying, but it's maybe important.

So then X of p is a cover of this. So it's going to look sort of like so. And I guess it has a cusp for each. It also has some cusps, not sure exactly how many. This is X of p. And this is a cover. So how to make that look nice?

Well, you could kind of chop this into fundamental domains. And then each fundamental domain is a bijection onto X of 1. And if you do this nicely, so then the different fundamental domains, they correspond to the group of deck transformations, which is SL2Fp.

And if you do this nicely, you could put a dot in each fundamental domain. And you could put an edge if the fundamental domains touch each other. And the graph that I just drew would be a Cayley graph of SL2Fp. So the geometry of these two-dimensional surfaces is closely related to the geometry of SL2Fp.

So then, we talk about the spectrum of the Laplacian On xp, our subject of study. And 0 is in the spectrum, so the Laplacian of a constant function is 0. But after 0, there's a gap.

So if this is the real line, and here's 0, that's part of the spectrum. Then there's a gap. And this is called lambda 1 of X of p. And the rest of the spectrum lives here, is contained in.

So that's what lambda 1 is, the smallest non-zero part of the spectrum. All right, so the theorem that Selberg actually proved in the 1950s is that lambda 1 of X of p is at least 3/16 for all p. And the conjecture is that lambda 1 of X of p is at least a quarter, maybe minus little o of 1, I'm not sure if you need it, but certainly if somebody proved that, that would be a big deal, for all p.

This may not look immediately that it matters very much. A quarter is a significant number because it's lambda 1 of hyperbolic plane. So I would say, it matches that. Anyway, so this is the theorem that Selberg actually proved. With modern techniques, it's not that difficult to get from this theorem to that theorem. It's much more difficult to prove either one of them.

OK instead of the random walk, we could use here, the heat flow on X of p. And there was a question earlier, what is the heat flow? So the question earlier was what is the heat flow on SL2R? But the analysis question here is, what is the heat flow on this hyperbolic manifold? Well, you can think of it as solving the heat equation.

So the heat equation would be del t of u equals Laplacian of u. And the Laplacian of k would take a while to write down, but it's a second order differential operator on your hyperbolic manifold. All right, OK, and the solution has a formula using a heat flow operator.

So let's say u of dot comma t would be Ht of u. So this is now an operator Ht, which plays the role of our operator T mu or T mu to the k. So everything we talked about with a little bit more analysis in PDE to set things up, everything we talked about applies here.

So for instance, SL2 of Fp acts isometrically on X of p. And therefore, it acts on the eigenspaces of Ht. They have multiplicity. And also, you can do lifting. So Ht tilde is the heat operator on the universal cover, the hyperbolic plane. And then Ht of x1, x2 is the sum over g in gamma p of Ht tilde x1 tilde gx2 tilde.

So, in other words, the new heat operator is-- take the heat operator on the hyperbolic plane and push it forward. And that follows basically because this is a local equation. So solution to the heat equation, you pull back, you get a solution to the heat equation.

And Ht tilde, the heat kernel on the hyperbolic plane, it has an explicit formula. And that's the equivalent that plays the role of the vague statement. So this hopefully is a dictionary of most of the elements of our proof. They all have natural analogs for this hyperbolic surface.

And the one that was the most troubling has a less troubling analog, or the one that was not stated precisely would have a precisely stated analog. So without doing all the details, if you then translate the proof that we did into the setting of hyperbolic surfaces, then that would give a proof of a theorem like this one with a constant that's not quite as good. But usually, it doesn't matter. Yeah?

AUDIENCE: I have a question about the heat flow. Since the spaces X of p are not compact, does there have to be some of decay condition to make the heat flow well-defined in these spaces?

LAWRENCE GUTH: Yeah. So the question is, since X of p is not compact, should we have some kind of decay condition to make the heat flow work? Yeah, I think the answer is yes. So these spaces have finite area. So it's not as bad as it would be in some other settings. But I think we should say something about the solution not going crazy near infinity.

Yeah, so this is the end of our discussion of the Selberg theorem for this particular set of generators. So take a moment to see if you have questions or comments about it. And then we'll explore what would happen if we didn't have that particular set of generators. Yeah, anything on people's minds? Yeah?

AUDIENCE: Can you give some intuition why the support grows exponentially?

LAWRENCE GUTH: Yeah, so the question was, if we go back to-- we were thinking about, we had a probability measure on SL2Z, we convolve it with itself a lot of times, and then we discovered that the support is roughly evenly distributed on a ball of size T, But T grows exponentially with k.

OK, yeah, so this is a kind of different behavior from what happens on the real line. And let's think about what's going on. So M, let's say, is just evenly distributed between a few matrices, maybe the matrices in the Selberg thing.

So now, I'm going to take k of those matrices, and I'm going to multiply them together. Now, if I multiply k matrices together, the biggest that the entries could be at the end is exponential in k. On the other hand, you might wonder, is it really that big? Or is there some cancellation?

Suppose I had 1 by 1 matrices, and I had a measure, at 50% I was 1/2, and 50%, it was 2. Take k of those, and I multiply them together. What would happen? Well, there would a lot of cancellation. There are a lot of cancellations between the 1/2s and the 2s

It would get clearer to us if we took the logarithm of everything. And the size of that would actually end up being exponential of the square root of k. So the 2 by 2 case is different from that. The actual size of this is typically exponential in k. There's not as much cancellation of that kind. And let's see, how could we get some intuition about that?

AUDIENCE: Is it in some sense because of the [INAUDIBLE]

LAWRENCE GUTH: Yeah, that's a good comment. So the comment was SL2Z is virtually free. So it's not very likely to go out in SL2Z and then come back. And that's a related fact. Yeah, so it's also related to hyperbolic geometry.

So you do a random walk in the hyperbolic plane, if you take s steps, you'll typically be around a distance s from the origin. And that's because the hyperbolic plane is branching out. And so at every moment, you're not equally likely to be going back towards the origin or away, you're kind of 3/4 of the directions are away. And that's what life is like inside of the matrices. Yeah, great question. Anything else on people's minds?

OK, all right, so let's look back at this Selberg theorem, now that we understand it better. And we had this particular A listed here. What did we really use about A? What was special about A? Yeah?

AUDIENCE: We used the fact that it was a generator of SL2Fp.

LAWRENCE GUTH: Yeah, so we used that as a generator of SL2 of Fp. And what I'd like us all to think about is if we had any set of generators of SL2Fp, do we think that this would still be true? And do we think that this proof would still work? I'll write that down while you think. it.

So question, can we replace A by any set of generators of SL2Fp? Yeah?

AUDIENCE: So at minimum, they need to be symmetric, right?

LAWRENCE GUTH: Yeah, let's say they're symmetric. Let's focus on the symmetric case. That makes things technically easier. And I would be happy with any symmetric set of generators. Yeah?

AUDIENCE: Well, does the set of generators still need to be kind of at least close to virtually free?

LAWRENCE GUTH: Does it still need to be close to virtually free? Yeah, that's a good question. So although I mentioned being virtually free, we didn't actually use virtually free in the proof. That was for intuition. But we used something in the proof. Yeah, well, it's--

AUDIENCE: They're conjugate to their own powers or something like that, right?

LAWRENCE GUTH: They're conjugate. So a nifty feature of these guys, so the comment was these guys are conjugate to their own powers. Yeah, so that's important. But we didn't use that either. So when we proved that every representation of SL2Fp is big, we considered these elements. And we used the fact that they're conjugate to their own powers.

But once we knew abstractly that every representation of SL2Fp is big, we just used it as a black box. So it's not important that these elements in the statement of Selberg's theorem are the same elements we used in that proof. Yeah, good. But well, so one part of the answer is that we used that A is a set of generators of SL2Z.

So in the proof where we lifted things to SL2Z, we took this set A, we lifted it to SL2Z, and then we did a random walk on SL2Z. And we needed to know that thing was kind of evenly distributed. And the only way we could hope for that to be true is if this was a set of generators of SL2Z. Yeah?

AUDIENCE: Are there any even straightforward sets of generators of SL2Fp that would not be generators of SL2Z?

LAWRENCE GUTH: OK, so the question is, could it happen that we have generators of SL2Fp, but they don't generate SL2Z? And the answer is yes. That happens fairly commonly. And let me show you an example.

So here's an example. Maybe we won't prove everything, but let's say that Ak-- no, I should use a different letter. Am is the set of 1 plus or minus m, 1, 0, and 1 plus or minus m, 1, 0. So this is still a set of four elements. Size of A equals 4.

And if m is 1 or 2, it generates SL2Z. But if m is at least 3, it does not generate-- does not generate SL2Z. On the other hand, for every m, for all almost every p, so for all but finitely many p, this Am generates SL2Fp.

All right, so why does this happen? All right, so notice that this guy is 1, 0, 1, 1 to the plus or minus m right. So we assume that these guys generate SL2Z. It's not that hard to show.

Take a word in these guys, it's only included in the group generated by these guys if it has a rather special structure that each letter is raised to the plus or minus n-th power, so maybe m is 10.

So now it's clear that most words do not look like that. Now, it's not immediately clear that we don't have the whole group, because an element in the group might be represented by many words, but we mentioned before that there has a finite index subgroup which is free. So there are not that many ways to represent most words.

So this tiny subset of the words, we don't expect it to cover very much of the group. And life is different in SL2Fp. And a little later, we'll talk about who are the subgroups of SL2Fp. But there aren't a whole lot of them. And once you have enough elements that they're not trapped in a subgroup, then they must generate the whole group.

OK, all right, so there are other choices of generators for SL2Z. Any choice of generators would have worked fine in our argument, but it's actually quite common that you have some generators for SL2Fp, and they don't generate SL2Z.

In this case, our argument doesn't work very well. Let's say that Gm is the group generated by Am. This group is a lot harder to understand than SL2Z. So for instance, if you were to write down a matrix and then you'd like to know is it in Gm, that's not super easy to figure out.

SL2Z are all the integer matrices where the entries satisfy the equation ad minus bc equals 1. Whether or not a matrix is in this group does not have a simple criterion like that. As far as I know, the best thing-- well, I don't know anything faster to do than to take the generators and multiply them together in different ways and see if you get this matrix. Yeah?

AUDIENCE: How close is Gm to gamma m? Do they have anything in common? Because at first, it sort of look like things congruent to the identity mod m.

LAWRENCE GUTH: Right, so Gm is a subgroup. So OK, so the question was how close is Gm to gamma m? Remember that gamma m was the subgroup of matrices in SL2Z that are congruent to the identity mod m. So these guys are congruent to the identity mod m. So therefore, Gm is a subgroup of gamma m.

However, this group has infinite index In SL2Z for m at least 3. Whereas, gamma m had finite index because it was the kernel of a map to a finite group. So gamma m, which is already not a super easy thing to understand if m is large, this is well-- objectively, this is a smaller group than gamma m, and subjectively, I think it is a harder group to understand than gamma m.

So this whole method has not really worked. And in fact, it goes in the opposite direction. So namely, for generators of SL2Z, we first understand SL2Z really well, and then we can use that to understand SL2Fp with those generators.

And with a situation like this, there's some progress. And it goes, first, you try to understand SL2Fp with those generators, and you use that to try to understand G of m. There's actually a big literature now about these groups G of m. And first, you try to figure out things about SL2Fp with those generators.

So that's one part of the answer to this question. And what we said so far is that our proof doesn't usually work for other sets of generators, but that's not an answer to the question whether the theorem is true. And as far as we know, the theorem may well be true. So there's a conjecture you could take any set of generators of SL2Fp, you just always have a spectral gamma. That's not proven yet, but there's some significant progress. And I will tell you something about it. Yeah?

AUDIENCE: What's the equivalent spectral statement?

LAWRENCE GUTH: What is the equivalent spectral statement? So if you take-- so I'm not sure if this is your question, but let me say something, and then-- so if we take another set of generators here, we have a different operator, and we can ask what is sigma 1 and so on.

But so I think what you were wondering is if we took another set of generators there, would that be related to the spectrum of the Laplacian of something? Is that your question? OK. OK, so not as far as I know.

And so you might say this theorem is great because it connects to this hyperbolic geometry, and it tells us really interesting estimate for the spectrum of some surface. With other generators, maybe it doesn't exactly do that. But on the other hand, it says some interesting stuff about this group, which is interesting in some other problems.

Yeah, let's leave that. All right, so suppose we're just given any set A inside of SL2Fp, and we'd like to figure out if there's a spectral gap. One thing that we know could go wrong is that A could be contained in a subgroup.

So if A is contained in H, which is a proper subgroup, then sigma 1 of TA is 1. All right, so that's elementary to see. But in order to really use this, we would actually-- it helps to know what are all the proper subgroups. So let's talk a little bit about the classification of subgroups of SL2Fp.

So the classification was done by someone named Dickson in 1901. And it's not that hard, but it's not that easy either. And I was a little surprised to learn that if you'd like to classify the subgroups of SLDFp, that was only done-- first of all, it's only done approximately. And it was only done around the year 2000. So some of these problems actually are really pretty hard. So it's done by Dixon in 1901.

OK, well, let's think of some subgroups. We could have the diagonal subgroup, which is typically called T for torus. There's this guy. We have the upper triangular, unipotent subgroup, like that. Let me put a little 0 here. And we have the more general upper triangular guys, which have that.

Now, if you take a subgroup, you can conjugate it, and you can get another subgroup. So you can conjugate all of these. And so let me write B, that means a conjugate of B0. So those are a bunch of subgroups. Those are the only ones that I was able to think of without effort.

There are a few other small subgroups, like for instance, apparently the icosahedral group appears in SL2Fp for some p's. And the statement of the classification is a lot easier if we don't say exactly what groups appear, but just kind of roughly and ignore small groups. So there's a theorem about this.

The rough theorem would say if H in SL2Fp is a proper subgroup, then there exists a Borel subgroup B, one of these guys, and B intersect H is a large fraction of H. So this c bigger than 0 is a universal constant.

So more or less, H is not quite in B, but almost all. H has a very large part that's in B. And B is a conjugate of B0. And the fine print is conjugate of B0 conjugated by not necessarily someone in SL2Fp, but maybe the matrix we conjugate by might involve the algebraic closure. Anyway, it's fine.

Now, let's make a remark. The analog of this is not true. For SL2Fq, q is a prime power because SL2Fp is a subgroup of SL2Fq. but SL2Fp is not contained in any upper triangular or something.

So there's an analogous theorem for SL2Fq, but it needs to be stated a little bit different way to take account of the subfields. So this may already start to remind us of something that we've seen, that projection theory is different over Fp and Fq.

OK, so now, I can state the second main theorem that I wanted to say about random walks on groups, which is a generalization of Selberg's theorem. And it allows to take any set of generators where it's not trapped in a Borel subgroup, but it requires a little quantification, which is perhaps not quite as strong as one might fantasize about.

All right, so for every epsilon bigger than 0, there are a couple of positive constants every p. So if I have a subset of SL2Fp, and the subset is not contained in a Borel subgroup, and also it doesn't concentrate too much in a Borel subgroup, so for every B Borel, if I take A intersected with B, this is, at most, p to the minus epsilon size of B and with this small constant.

So a smallish-- sorry, size of A. So only a small fraction of A is in this Borel subgroup. And it's a p to the epsilon kind of fraction where you-- that's a pretty small fraction. You might hope to do a little better, but anyway, that's what it says.

And here, I wasn't positive-- so when I proved it, I needed to say coset, but I'm not sure if it's coset or subgroup, anyway. So if it doesn't concentrate in something closely related to a subgroup, then it expands. So then it has a mixed signal 1 of TA is less than 1 minus c of epsilon.

So up to a little bit of quantification, if A is contained in a Borel subgroup, definitely we won't have a spectral gap. And if A doesn't intersect a Borel subgroup too much, then we will have a spectral gap. It's almost if and only if, although you might hope to improve this a little bit.

So this is kind of its own story to talk about. So since we're almost done, let me just make a couple of high level comments about it. Yeah?

AUDIENCE: What's a Borel subgroup?

LAWRENCE GUTH: A Borel-- so I don't know why-- so Borel is involved in this. A Borel subgroup is these guys B. Take an upper triangular matrix, that's a Borel subgroup, or conjugate that. So one comment about this theorem is that we had better know this theorem. You better know something like this.

And so the first part of this theorem is to prove that if you don't have a spectral gap, then there must be a subset that's kind of approximately a subgroup that intersects A a lot. If you don't have a spectral gap, there must be a subset that's kind of approximately a subgroup where maybe A convolved with itself many times is heavily concentrated in this approximate subgroup.

And that's a very nice argument. And it turns out to use the tools from additive combinatorics. It uses like Balog-Szemeredi-Gowers and things like that in a non-commutative version. And then the second part of the proof is to say, is kind of a robust version of this theorem. This is a classification of actual subgroups. We want to make it a little bit stronger and have a classification of approximate subgroups.

So not only any actual proper subgroup would have to intersect a Borel subgroup a lot, but actually, any kind of approximate subgroup, so any subset where when you multiply it by itself two or three times, it doesn't get very much bigger, the only way that can happen is if your subset heavily intersects a Borel subgroup.

Now, that second part, you'll notice, is not true over Fp-- sorry, it's not true over Fq. And so the second part will draw on some product theory and our theorems that we worked very hard on that give a projection estimate over Fp that's not true over Fq. We will need that.

And so I think this is a nice example of how those theorems are only an epsilon improvement on a trivial estimate where epsilon is very small. But it still is meaningful. There's an important difference between epsilon being zero or not zero. And they're used to prove kind of this c of epsilon is pretty small.

But there's just a really important difference between whether you had zero here, which is like a completely vacuous statement, and whether you had 1 minus epsilon, which even if epsilon is a small constant, is a qualitatively sharp statement about how these random walk mixes. OK, cool. So we'll talk more about those things on Thursday.