

YUFEI ZHAO: All right. Last time we started talking about pseudorandom graphs, and we considered this theorem of Chung, Graham, and Wilson, which, for dense graphs, gave several equivalent notions of quasi-randomness that, at least the phase values, do not appear to be all that equivalent. But they are actually-- you can deduce one from the other. There was one condition at the very end which had to do with eigenvalues. And, basically, it said that if your second largest eigenvalue in absolute value is small, then the graph is pseudorandom. So that's something that I want to explore further today to better understand the relationship between eigenvalues of a graph and the pseudorandomness properties.

For much of-- pretty much all of today, we're going to look at a special class of graphs known as (n, d, λ) graphs. This just means we have n vertices, and we're only going to consider, mostly out of convenience, d regular graphs. So this will make our life somewhat simpler. And the λ stands for that-- if you look at the adjacency matrix, and if you write down the eigenvalues of the adjacency matrix, then, well, what are these eigenvalues?

The top one, because it's d regular, is equal to d . And λ corresponds to the statement that all the other eigenvalues are, at most, λ in absolute value. So the top one is equal to d . All the other ones in absolute value-- so it could be basically the maximum of these two-- is bounded above by λ .

And at the end of last time, we showed this expander mixing lemma, which, in this language, says that if G is (n, d, λ) , then one has the following discrepancy type. So the randomness property, namely that if you look at two vertex sets and look at how many actual edges are between them compared to what you expect if this were a random graph of a similar density, then these two numbers are very similar, and the amount of error is controlled by your λ . In particular, a smaller λ gives you a more pseudorandom graph.

So the second part of today's class, I want to explore the question of how small this λ can be. So what's the optimal amount of pseudorandomness? But, first, I want to show you some examples. So, so far, we've been talking about pseudorandom graphs, and the only example, really, I've talked about is that a random graph is pseudorandom. Which is true. A random graph is pseudorandom with high probability, but some of the spirit of pseudorandomness is to come up with non-random examples, come up with deterministic constructions that give you pseudorandom properties.

So I want to begin today with an example. A lot of examples, especially for pseudorandomness, come from this class of graphs called Cayley graphs, which are built from a group. So we're going to reserve the letter G for graphs, so I'm going to use Γ for a group. And I have a subset S of Γ , and S is symmetric, in that if you invert the elements of S , they remain in S .

Then we define the Cayley graph given by this group and the set S to be the following graph, where V , the set of vertices, is just the set of group elements. And the edges are obtained by taking a group element and multiplying it by S to go to its neighbor. So this is a Cayley graph. And Cayley graphs are-- start with any group, start with any subset of the group, you get a Cayley graph. And this is a very important construction of graphs. They have lots of nice properties

And, in particular, an example of a Cayley graph is a Paley graph. They're not related. So a Paley graph is a special case of a Cayley graph obtained by considering the group, the cyclic group mod p , where p is prime $1 \pmod{4}$. And I'm looking at S being the set of quadratic residues, mod p . It's actually nonzero quadratic residues. So elements of mod p that could be a square.

So we will show in a second that this Paley graph has nice pseudorandom properties by showing that it is an (n, d, λ) graph with λ fairly small compared to the degree. Just a historical note-- so Raymond Paley-- so the Paley graph named after him-- he actually-- he was from the earlier part of 20th century. So from 1907 to 1932. So he died very young at the age of 26, and he actually died in an avalanche when he was skiing by himself in Banff. So Banff is a national park in Alberta in Canada.

And when I was in Banff earlier this year for a math conference-- so there's also a math conference center there-- so I had a chance to go visit the-- Raymond Paley's tomb. So there's a graveyard there where you can find his tomb. And it's very sad that, in his short mathematical timespan, actually he managed to do a lot of amazing mathematical-- find a lot of amazing mathematical discoveries. And there are many important concepts named after him. So things like Paley-Wiener theorem, Paley-Zygmund, Littlewood-Paley, all this important ideas and analysis named after Paley. And Paley graph is also one of his contributions.

So what we'll claim is that this Paley graph has the desired pseudorandom properties, in that if you look at its eigenvalues, then the top eigenvalues, except-- so except for the top

eigenvalue, all the other eigenvalue are quite small. So keep in mind that the size of S is basically half of the group. So $p - 1$ over 2. So for especially larger values of p , p 's eigenvalues are quite small compared to the degree.

So the main way to show that Cayley graphs like that have small eigenvalues is to just compute what the eigenvalues are. And this is actually not so hard to do for Cayley graphs, so let me do this explicitly. So I will tell you very explicitly a set of eigenvectors. And they are-- the first eigenvector is just the all 1's vector. The second eigenvector is the vector coming from 1, ω , ω squared, so ω to the $p - 1$, where ω is a parameter of p -th root of unity.

The next one is 1, ω square, ω fourth, all the way to ω p -- ω to the 2 times $p - 1$. And so on. So I want to have-- yes, so OK. So I make this list, and I have p of them. So these are my eigenvectors. And let me check that they are actually eigenvectors. And then we can also compute their eigenvalues.

So the top eigenvector corresponds to d . So the all 1's in a d regular graph is always an eigenvector with eigenvalue d . And the other ones, we'll just do this computation. So instead of getting confused with indices, let me just compute, as an example, the j -th coordinate of the adjacency matrix times V_2 . So the j -th coordinate, so what it comes to, is the following sum. If I run over S , then ω raised to j plus s . So S is symmetric, so I don't have to worry so much about plus or minus. So I say j plus s .

So if you think about what this Cayley graph, how it is defined, if you hit this vector with that matrix, the j -th coordinate is that sum there. But I can rewrite the sum by taking out this common factor ω to j . And you see that this is the j -th coordinate of V_2 . And this is true for all j . So this number here is λ_2 .

And, more generally, λ_k is the following sum, for k from 0-- so from k being 1 through p . So when you plug in k equals to 1, you just get d . And the others are sums of these exponential sums. Now, this is a pretty straightforward computation. And, in fact, we're not using anything about quadratic residues. This is a generic fact about Cayley graphs of $z \pmod p$. So this is true for all Cayley graphs S , not necessarily for quadratic residues. And the basic reason is that, here, you have this set of eigenvectors, and they do not depend on S . So you might know this concept from other places, such as circular matrices and whatnot, but this is true in this simple computation.

So now we have the values of λ_k explicitly. I can now compute their sizes. I want to know how big this λ_k is. Well, the first one, when k equals to 1, it's exactly d , the degree, which is $p - 1$ over 2. But what about the other ones? So, for the other ones, we can do a computation as follows. So note that I can rewrite λ_k by noting that if I take twice it and plus 1, then I obtain the following sum. Because here I am using the S as a set of quadratic residues. So if I consider this sum here, every quadratic residue gets counted twice, except for 0, which gets counted once.

And now I would like to evaluate the size of this sum, this exponential sum. And this is something that's known as a Gauss sum. So, basically, a Gauss sum is what happens when you have something that's like a quadratic, an exponential sum with a quadratic dependence in the exponent. And the trick here is to consider the square of the sum. So the magnitude squared.

Now if I expand the square-- so squaring is a common feature of many of the things we do in this course. It really simplifies your life. You do the square, you expand the sum. You can re-parameterize one of the summands like that. So do two steps at once. I'm re-parameterizing and I'm expanding. But now you see, if I expand the exponent, we find-- so that's just algebra.

And now you notice that this sum here, the sum over a is equal to-- when b is nonzero, I claim that this sum is 0. And when b is nonzero, then I'm summing over some permutations of the roots of unity. So here I'm assuming that k is bigger than-- let's say here k is not 0. So I'm re-parameterizing k a little bit. So k is not 0. Then when b is not 0, the sum over a is 0. And otherwise it equals to p . So the sum over here equals to p .

And, therefore, λ_k , λ_{k+1} -- how about if I-- so what should I change that to? So if I-- k is 0, then I want this to be λ_{k+1} . Then λ_{k+1} is equal to $p - 1$ over 2, for all λ_k not equal to 0.

So, really, except for the top eigenvalue, which is just the degree, all the other ones are one of these two values, and they're all quite small. So this is an explicit computation showing you that this Paley graph is indeed a pseudorandom graph. It's an example of a quasi-random graph. Yes.

AUDIENCE: Do we know what the sign is?

YUFEI ZHAO: The question is, do we know what the sign is? So we actually-- so here I am not telling you

what the sign is, but you can look up. Actually, people have computed exactly what the sign should be. And this is something that you can find in a number theory textbook, like Aaron and Rosen. Any more questions?

There is a concept here I just want to bring out, that you might recognize sums like this. So this kind of sum. That's a Fourier coefficient. So if you have some Fourier transform, I mean, this is exactly what Fourier transforms look like. And it is indeed the case that, in general, if you have an Abelian group, then the eigenvalues and the spectral information of the corresponding Cayley graph corresponds to Fourier coefficients.

And this is the connection that we'll see also later on in the course when we consider additive combinatorics and giving a Fourier analytic proof of Roth's theorem. And there Fourier analysis will play a central role. But this is actually-- this analogy, as I've written it, is only for Abelian groups. If you try to do the same for non-Abelian groups, you will get something somewhat different.

So for non-Abelian groups, you do not have this nice notion of Fourier analysis, at least in the versions that generalizes what's above in a straightforward way. But, instead, you have something else, which many of you have seen before but under a different name. And that's representation theory, which, in some sense, is Fourier analysis, except, instead of one-dimensional objects and complex numbers, we're looking at higher-dimensional representations. So I just want to point out this connection, and we'll see more of it later on. Any questions?

So let's talk more about Cayley graphs. So, last time, we mentioned these notions of quasi-randomness. And I said at the end of the class that many of these equivalences between quasi-random graphs, they fail for sparse graphs. If your density, if your x density is a constant, then the equivalences no longer hold.

But what about for Cayley graphs? And, in particular, I would like to consider two specific notions that we discussed last time and try to understand how they relate to each other for Cayley graphs. So for dense Cayley graphs, it's a special case of what we did yesterday. So I'm really interested in sparser Cayley graphs, even down the degree. So even down the degree. So that's much sparser than the regime we were looking at last time.

And the main result I want to tell you is that the DISC condition is, in a very strong sense,

actually equivalent to the eigenvalue condition for all Cayley graphs, including non-Abelian Cayley graphs. So before telling you what the statement is, I first want to give an example showing you that this equivalence is definitely not true if you remove the assumption of Cayley graphs.

For example, if you-- so example that this is false for non-Cayley. Because if you take, let's say, a large-- so let's say d regular graph. So let's say a large random d regular graph. d here can be a constant or growing with n , but this is a pretty robust example. And then I add to it an extra destroying copy of $k \leq d + 1$ that's much smaller in terms of number of vertices.

The big, large random graph, well, by virtue of being a random graph, has the discrepancy property. And because we're only adding in a very small number of vertices, it does not destroy the discrepancy property. The discrepancy property, if you're just adding a small number of vertices, it doesn't change much. So this whole thing has discrepancy.

However, what about the eigenvalues? Claim that the top two eigenvalues are in fact both equal to d . And that's because you have two eigenvectors, one which is the all 1's vector on this graph, another which is the all 1's vector on that graph. These two DISC components each give you a top eigenvector of d , so you get d twice. And, in particular, the second eigenvalue is not small. So the implication from DISC to eigenvalue really fails for non-Cayley graph for general graphs.

The implication, the other direction is actually OK. In fact, the eigenvalue implies DISC is actually the content of the expander mixing lemma. So this follows by expander mixing lemma. And that's because, if you look at the expander mixing lemma for a Cayley graph-- or for a-- not for Cayley graph-- for-- if you have the eigenvalue condition, then, automatically, you would find that these two guys here are at most n . So if λ is quite small compared to the degree, then you still have the desired type of quasi-randomness. So I'll make the statements more precise in a second.

So the question is, how can we certify, how can we show that, in fact, DISC, which is seemingly weaker property, implies a stronger property of eigenvalue for Cayley graphs. And what is special about Cayley graphs that would allow to do this, that the statement is generally false for non-Cayley graphs? So let me define-- so let me first tell you the result. So this is the result due to David Conlon and myself two years ago.

So many of you may not have been to many seminar talks, where there's this convention in

mathematics talks where you don't write out your full name, only by the initial. Although some kind of false modesty. But, of course, we all love talking about our own results, but somehow we don't like to write our own name for some reason.

So here's the theorem. So I start with a finite group, Γ . And let me consider a subset S of Γ that is symmetric. And consider G the Cayley graph. Let me write n as the number of vertices, and d the size of S . So this is a d regular graph. Let me define the following properties.

The first property, I'll call DISC with ϵ . So I give you an explicit parameter. The number of edges between x and y differs from the number of edges that you would expect. So as in the expander mixing lemma. So the DISC property is that this quantity is small relative to the total number of edges.

The second property, which we'll call the eigenvalue property, EIG, is that G is an n, d, λ graph, with λ , at most, ϵd . So λ is quite small as a function of d . The conclusion of the theorem is that, up to a small change of parameters, these two properties are equivalent. In particular, eigenvalue implies ϵ implies DISC of ϵ . And DISC of ϵ implies EIG of 8ϵ and this is the more interesting direction-- it implies EIG. Well, you lose a little bit, but, at most, a constant factor. EIG of 8ϵ . Any questions about the statement so far?

And so, as I mentioned, this is completely false if you consider non-Cayley graphs. And we also, using expander mixing lemma, using that implication up there, this direction follows. One of the main reasons I want to show you a proof of this theorem is that it uses this tool which I think is worth knowing. And this is an important inequality known as Grothendieck's inequality.

So many of you probably know Grothendieck as this famous French mathematician who reinvented modern algebraic geometry and spent the rest of his life writing tomes and tomes of text that have yet to be translated to English. But he also did some important foundational work in functional analysis before he became an algebraic geometry nerd. And this is one of the important results in that area that he-- so Grothendieck's inequality tells us that there exists some absolute constant k such that for every matrix A -- so a real-valued matrix-- we have that the-- so we have that, if you-- so here's the idea.

Let's consider the supremum-- so let's consider the following quantity. This is a bilinear form. So this is a bilinear form. This is basically a-- so bilinear form, if you hit it by a vector x and y

from the two sides. And I'm interested in what is the maximum value of this bilinear form if you are allowed to take x and y to be plus/minus 1-valued real numbers?

So this is an important quantity, and it gives you a matrix. And it's basically asking you, you get a sign of plus or minus to each row and column, and I want to maximize this number here. This is an important quantity that we'll see actually much more in the next chapter on graph limits. But, for now, just take my word. This is a very important quantity.

And this is actually a quantity that is very difficult to evaluate. If I give you a very large matrix and ask you to compute this number here, there is no good algorithm for it. And it's believed that there is no good algorithm for it. On the other hand, there is a relaxation of this problem, which is the following. It's still a sum, but now, instead of considering the bilinear form there, let's consider the x_i 's and y_i 's. Not-- take them not from real numbers, but take vectors.

So let's consider the sum where I'm taking a similar-looking sum, except that x_i 's and y_i 's come from a unit ball in some vector space with an inner product, where B is the unit ball in some \mathbb{R}^m , where here the dimension is actually not so relevant. The dimension is arbitrary. If you like, you can make m or $2n$ because you only have that many vectors.

So this quantity here, just by very definition, is a relaxation of the right-hand of this quantity here. So it's at least this large. So, in particular, if you have whatever plus/minus, you can always look at the same quantity with m equal to 1, and you obtain this quantity here. But this quantity may be substantially larger. So the x and y 's have more room to put themselves in to maximize the sum.

And Grothendieck's inequality tells us that the left-hand side actually cannot be too much larger than the right-hand side. It exceeds it by, at most, a constant factor. So, in other words, the left-hand side, which is known as a semi-definite relaxation, you are not losing by more than a constant factor compared to the original problem. And this is important in computer science because the left-hand side turns out to be a Semidefinite Program, an SDP, which does have efficient algorithms to compute. So you can give a constant factor approximation to this difficult compute but important quantity by using semidefinite relaxation. And Grothendieck's inequality promises us that it is a good relaxation.

You might ask, what is the value of k ? So I said there exists some constant k . So this is actually a mystery. So the current proofs have been improved over time. And Grothendieck

himself proved this theorem, but it constantly has been improved over time. And, currently, the best-known result is something along the lines of k roughly 1.78 works. But the optimal value, which is known as Grothendieck's constant, is unknown.

So this is Grothendieck's constant. Actually, this, what I've written down is what's called the real Grothendieck's constant. Because you can also write a version for complex numbers and complex vectors, and that's the complex Grothendieck's constant. Yes.

AUDIENCE: Is there a lower bound that's known [INAUDIBLE] greater than 1?

YUFEI ZHAO: Is there a lower bound that is known? Yes. It's known that it's strictly bigger than 1.

AUDIENCE: Do we know [INAUDIBLE]?

YUFEI ZHAO: So there are some specific numbers, but I forget what they are. You can look it up. Any more questions? So we'll leave Grothendieck's inequality. We'll use it as a black box. So if you wish to learn the proof, I encourage you to do so. There are some quite nice proofs out there. And we'll use it to prove this theorem here about quasi-random Cayley graphs.

So let's suppose DISC holds. So what would we like to-- what do we like to show? We want to show that this eigenvalue condition holds. And we'll use the-- some min-max characterization of eigenvalues. But, first, some preliminaries.

Suppose you have vectors x and y which have plus/minus 1 coordinate values. Then, by letting-- so let's consider the following vectors, where I split up x and y according to where they're positive and where they're negative. So, here, these are such that x plus is equal to-- so if I evaluate it on a coordinate g , then it's 1. So if x sub g is plus 1, and 0 otherwise. x sub minus is 1 if x sub g is minus 1. 0 otherwise. So x splits into x plus minus x minus, and y splits into y plus minus y minus.

Let's consider a matrix A where the g comma h entry of A is the following quantity. I have the set S , and I look at whether g inverse h lies in S . And I can consider an indicator of that. So it's 1 or 0. And then subtract d over n so that this value has mean 0. So this is a matrix.

And now if I consider the bilinear form, hit A from left and right with x and y , then the bilinear form splits according to the plus and minuses of the x 's. And I claim that each one of these terms is controlled because of DISC. So, for example, the first term is, if you expand out what this guy is-- so here's an indicator vector. That's an indicator vector.

And if you look at the definition, then this is precisely the number of edges between x plus and y plus minus d over n times the size of x plus times the size of y plus, where x plus is the set of group elements such that $x \text{ sub } g$ is 1, and so on. All right. So the punchline up there is that this quantity-- so this quantity is, at most, by discrepancy, ϵn . So this sum here, by triangle inequality, is, at most, $4 \epsilon n$.

All right. So, so far, we've reinterpreted the discrepancy property. And what we really want to show is that this graph satisfies eigenvalue condition. So what does that actually mean to satisfy the eigenvalue condition? So by the min-max characterization of eigenvalues, it follows that the maximum of these two eigenvalues, which is the quantity that we would like to control, is equal to the following. It is equal to the supremum of this bilinear form when x and y are unit-length vectors.

And this is simply because A is the matrix-- it's not the adjacency matrix. A is not the adjacency matrix. A is the matrix obtained by essentially taking the adjacency matrix and subtracting that constant there. And subtracting that constant gets rid of the top eigenvalue. And what you remained is whatever that's left. And you want to show that whatever you remained has small spectral radius. So we would like to show that this quantity here is quite small.

Well, let's do it. So give me a pair of vectors, x and y . And let's set the following quantities, where I take a twist on this x vector by rotating the coordinates, setting $x \text{ super } s \text{ sub } g$, the coordinate g , to be $x \text{ sub } sg$. So x is a vector indexed by the group elements, and then rotating this indexing of the group elements by s . So that's what I mean by superscript s . And, likewise, $y \text{ superscript } s$ is defined similarly.

So I claim that these twists, these rotations, do not change the norm of these vectors. And that should be pretty clear, because I'm simply relabeling the coordinates in a uniform way. And, likewise, same for y .

So I would like to show this quantity up here is small. So let's consider two unit vectors. And consider this bilinear form. If I expand out this bilinear form, it looks like that. I'm just writing it out. But now let me just throw in an extra variable of summation. What we'll do is essentially look at the same sum, but now I add in an extra s , and put this s over here. So convince yourself that this is the same sum. So it's simply re-parameterizing the sum. So this is the same sum.

But now, if you look at the definition of A , there's this cancellation. So the two s 's cancel out. So let's rewrite the sum. $\frac{1}{n}$, then g, h, s , all group elements. Then-- now, if I bring this summation of s , now I bring it inside, and then you see that what's inside is simply the inner product between the two vectors, x sub g -- between the two vectors. So this is-- so what's inside is simply the product, inner product, between these two.

So I may need to redefine. Yes. So when you're looking at-- when you're talking about non-Abelian groups, it's always a question of which side should you multiply things by. And you guys are OK? Or I need to change this s to over here. But anyway, it should work. Yes, question.

AUDIENCE: yh [INAUDIBLE].

YUFEI ZHAO: yh. Thank you. Yes, I think-- OK. Question.

AUDIENCE: [INAUDIBLE]

YUFEI ZHAO: Great. So maybe I need to switch the definition here, but, in any case, some version of this should be OK. Yes. So figure it out later in the notes. But now-- OK. So you have this-- we have this here. And if you look at this quantity here, it is the kind of quantity that comes up in Grothendieck's inequality. So this is basically the left-hand side of Grothendieck's inequality.

What about the right-hand side of Grothendieck's inequality? Well, we already controlled that. We already controlled that because we said, whenever you have up there little x and little y -- so the conclusion of this board was that-- let me erase over here. So the conclusion of this board was that this bilinear form is bounded by, at most, $4\epsilon^d$, for all x and y being plus/minus 1 coordinate valued.

So combining them by Grothendieck, we have an upper bound, which is the Grothendieck constant times $4\epsilon^d$ -- so $4\epsilon^d n$. There's a-- sorry. There's an n missing here. And, therefore, because the Grothendieck constant is less than 2, we have a bound of $8\epsilon^d$. And this shows that this variational problem, which characterizes the largest eigenvalue in absolute value, is, at most, $8\epsilon^d$, thereby implying the eigenvalue property.

So the main takeaway from this proof, two things. One is Grothendieck's inequality is a nice thing to know. So it's a semidefinite relaxation that changes the problem, which is initially somewhat intractable, to a semidefinite problem which is both, from a computer science point of view, algorithmically tractable, but also has nice mathematical properties. And for this

application here, there's this nice trick in this proof where I'm symmetrizing the coordinates using the group symmetries. And that allows me to obtain this characterization showing that eigenvalue condition and this discrepancy condition are equivalent for Cayley graphs. Let's take a quick break.

Any questions so far? So we've been talking about n , d , λ graphs. So d regular graphs. And the next question I would like to address is, In an n , d , λ graph, how small can λ be? So smaller λ corresponds to a more pseudorandom graph. So how small can this be?

And the right kind of setting that I want you to think about is think of d as a constant. So think of d as a constant, and n getting large. So how small can λ be. And it turns out there is a limit to how small it can be. And it is known as the Alon-Boppana bound, which tells you that if you have a fixed d -- and so G is an n -vertex graph with adjacency matrix eigenvalues λ_1 through λ_n , sorted in non-increasing order.

Then the second largest eigenvalue has to be at least, basically, $2\sqrt{d} - 1$ minus a small error term, little $o(1)$, where the little $o(1)$ goes to 0 as n goes to infinity. So the Alon-Boppana bound tells you that the λ_2 cannot be below this quantity here. And I want to explain what is the significance of this quantity, and you will see it in the proof. And this quantity is the best possible. And it also says what do we know about the existence of graphs which have λ_2 close to this number. So this is the optimal number you can put here. Question.

AUDIENCE: Does it say anything about how negative λ_n can be?

YUFEI ZHAO: Question-- does it say how negative λ_n can be? So I'll address that in a second, but, essentially, if you have a bipartite graph and λ_n equals to minus λ_1 .

AUDIENCE: [INAUDIBLE]

YUFEI ZHAO: More questions? So I want to show you a proof and, time permitting, a couple of proofs of Alon-Boppana bound. And they're all quite simple to execute, but the-- I think it's a good way to understand how these special techniques work. So, first, as with all of the proofs that we did concerning-- or most of them-- concerning eigenvalues, we're looking at the Courant-Fischer characterization of eigenvalues. It suffices to show, to exhibit some vector z -- so a nonzero vector-- such that z is orthogonal to the all 1's vector and this quotient is at least the claimed

bound.

So by the Courant-Fischer characterization of the second eigenvalue, if you vary over all such d that are orthogonal to the unit vector, then the maximum value this quantity attains is equal to λ_2 . So to show the λ_2 is large, it suffices to exhibit such a z . So let me construct such a z for you.

So let r be a positive integer. And let's pick an arbitrary vertex v . So v is a vertex in the graph. And let V_i denote vertices at distance exactly i from v . From-- yes, from v . So, in particular, V_0 is equal to $\{v\}$ and I can just draw you a picture. So you have V_0 , and then the neighbors of V_0 , and each of them have more neighbors. Like that. So I'm calling V_0 this stuff, V_1 . And then V_2 , and so on.

So I'm going to define a vector, which I'll eventually make into z , by telling you what is the value of this vector on each of these vertices. I will do this by setting very explicitly-- so set x to be a vector with value x_u to be w_i , where w_i is $d^{-1/2}$ raised to power $d - i$ over 2 whenever u lies in set V_i . So u is distance exactly i from v . I set it to this number. So notice that they decrease as you get further away from v .

And I do this for all distances less than r . So this is my x vector. And I set all the other vertices to be 0 if the distance between u and v is at least r . So that gives you this vector. And I would like to compute that quotient over there for this vector. And I claim that this quotient here is at least the following quantity.

But this is a computation, so let's just do it. So why is this true? Well, if you compute the norm of x -- so I'm just taking the sum of the squares of these coordinates. Well, that comes from adding up these values. So for each element in the i -th neighborhood. So I have w_i^2 .

And if I look at that quantity up there, so what is this? A is the adjacency matrix. So over here, A is the adjacency matrix. So this quantity, I can write it as a sum over all vertices u . And I look at x_u , and now I sum again over all neighbors of u , and consider x_u' . It's that sum there.

But this sum, I have some control over, because it is-- so what's happening here? I claim it has at least the following quantity. Consider where u is. So u could be-- I mean, it's only nonzero if u lies in the $r - 1$ th neighborhood. So in that neighborhood, I have V_i possible choices for the vertex u . For that choice, this x_u is w_i .

But what about all its neighbors? So it could have neighbors, well, in the same set going left. But there's-- so there's one neighbor going left, and all the other neighbors are-- maybe it's in the same set, maybe it's in the next set. But, in any case, I have the following inequality. There's one neighbor in the same-- in the left, if you look at that picture just now. And then all the remaining neighbors have $x_{u'} \geq w_i + 1$, because these weights are decreasing. So I can-- the worst case, so to speak, is if you-- all the neighbors point to the next set. So I had that inequality there.

There's an issue. Because if you go to the very last set, if you go to the very last set and think about what happens, when i is in that very last set, I'm overcounting neighbors that no longer has weights. So I need to take them out. So I should subtract $d - 1$ times-- and so this is the maximum possible weight sum I could have-- maximum possible overcount. So each product here has $t - 1$ neighbors at most. All right.

So this is-- should be pretty straightforward if you do the counting correctly. But now let's plug in what these weights are. And you'll find that this sum here, this quantity, is equal to-- so the key point here is that this thing simplifies very nicely if you consider what this is. So what ends up happening is that you get this extra factor of $2\sqrt{d-1}$. And then the sum minus $1/2$ of V sub [INAUDIBLE]. It's pretty straightforward computation using the specific weights that we have.

And one more thing is that notice that this-- so notice that the sizes of each neighborhood cannot expand by more than a factor of $d - 1$, because, well, you only have $d - 1$ outward edges going forward at each step. And, as a result, I can bound this guy. And so what you find is that this whole thing here is that least $2\sqrt{d-1}$. The main term is the sum. And this here is less than each individual summand. So I can do $1 - \frac{1}{2r}$.

Putting these two together, you find the claim. All right. So I've exhibited this vector x , which has that quotient property. But that's not quite enough, because we need a vector-- so it's called z up here-- that is orthogonal to the all 1's vector. And that you can do, because if the number of vertices is quite a bit larger than-- compared to the degree, then I claim that there exists u and v vectors-- vertices that are at distance at least $2r$.

So if I let-- this is the size of this tree. So if you have-- everything is within distance r -- distance $2r$ from a vertex, then they all lie on this tree edge. If you count the number of vertices in that tree, it's what I have-- the sum I've written here. So if I consider these two vectors-- so be-- so

x be the vector obtained above, which is, in some sense-- and I'm being somewhat informal here-- centered at v . And if I let y be the vector but I center it now at the vector-- at u , then I claim that, essentially, x and y are supported on disjoint vertex sets that have no edges even between them.

So, in particular, this inner product-- this bilinear form-- not inner product but this bilinear form-- is equal to 0, since no edge between the supports of x and y . So now I have two vectors that do not interact, but both have this nice property above. And now I can take a linear combination.

Let me choose a constant c -- so it's a real constant-- such that this z equal to x minus cy has-- and I can choose this constant. So x and y are both non-negative entries. They're both nonzero, and I can choose this constant c so that it is-- this z is orthogonal to the all 1's vector. And I now I have this extra property I want.

But what about the inner products? Well, these two vectors, x and y , they do not interact at all. So their inner products split just fine, and the bilinear form splits just fine. So you have this inequality here, as desired. And r , notice that I can take r going to infinity as n going to infinity, because d is fixed. So if n goes to infinity, then r can go to infinity, roughly a logarithmic n .

And that proves the Alon-Boppana bound. And just to recap, to prove this bound, we needed to exhibit by the Courant-Fischer some vector with a nice-- this quotient such that this quotient is large. And we exhibit this quotient by constructing the vector explicitly around the vertex and finding two such vertices that are far away from constructing these two vectors, taking the appropriate linear combination so that the final vector is orthogonal to the unit vector, to the all 1's vector, and then showing that the corresponding bilinear form has-- is large enough. Any questions?

I want to show you a different proof which gives you a slightly worse result, but the proof is conceptually nice. So let me give you a second proof which is slightly weakening. And just that we'll show-- so we'll show that-- so the earlier proof showed that λ_2 is quite large. But, next, we'll show that the max of λ_2 and the λ_n is large. So not that the second largest eigenvalue is large, but the second largest eigenvalue in absolute value is large. So it's slightly weaker, but, for all intents and purposes, it's the same spirit. So I'll show this one here.

And this is a nice illustration of what's called a trace method, sometimes also a moment method. Here's the idea. As we saw in the proof relating the quasi-randomness of C_4 and

eigenvalues, well, C_4 's are-- eigenvalues are related to counting closed walks in a graph. And so we'll use that counting closed walks in a graph. And, specifically, the $2k$ -th moment of the spectrum is equal to the trace of the $2k$ -th power, which counts the number of closed walks of length exactly $2k$.

So to lower bound the left-hand side, we want to lower-bound the right-hand side. So let's consider closed walks starting at a fixed vertex. So the number of closed walks of length exactly $2k$ starting at a fixed vertex v . Here we're in a d regular graph. So here we are in a d regular graph. I claim, whatever this number is-- it maybe different for each d -- it is at least the same quantity if I do this walk in an infinite d regular tree. So infinite d regular tree is what? This is an infinite d regular tree. We just start with the vertex, and go out d regular.

So why is this true? So think about how you walk. So let me just explain. This is, I think, pretty easy once you see things the right way. So start with a vertex v . Think about how you walk. And whatever way you can walk, well, you can walk the same way on the infinite d regular tree. Well, I mean, sorry. Whatever walk you can do on an infinite d regular tree, if you label the first vertex, the first edge, second edge, if you do a corresponding labeling on your original graph, you can do that walk on your original graph. Although the original graph may have some additional walks, namely things that involve cycles, that are not available on your tree. But, certainly, every walk, you can do. Every closed walk you can do on a tree, you can do the same walk on your graph.

So you can make this more formal. So you can write down a bijection or injection to make this more formal, but it should be fairly convincing that this inequality is true. But this is just a number. So this is a number of $2k$ walks in a d regular tree starting on the vertex. And this number has been well studied, and we don't need to know the precise number. We just need to know some good lower bound. And here is one lower bound, which is that there's at least a Catalan number, the k -th Catalan number, times $d - 1$ to the k , where C_k is the k -th Catalan number, which is equal to-- so k choose k divided by $k + 1$.

So let me remind you what this is. A Catalan number has many combinatorial interpretations, and it's a fun exercise to do bijections between them. But, in particular, C_3 is equal to 5, which counts the number of up and down walks of length 6 that never dip below the horizontal line where you start. So, then, this corresponds to going away from the root versus coming back to the root. So, then, you have at least that many ways. And when you are moving away from the root, you have $d - 1$ choices on which branch to go to. OK, good.

Given that, the right-hand side is at least, then, n , the number of vertices, times the quantity above related to Catalan numbers. On the other hand, the left-hand side is at most-- here we're using that $2k$ is an even number-- is at most d to the $2k$ plus all the other eigenvalues that are most λ in absolute value. So let me call this quantity λ .

Rearranging this inequality, we find that λ to the $2k$ is at least this number here. Just here, I'm changing $n - 1$ to n . So we have that. And now what can we do? We let n go to infinity and k go to infinity slowly enough. So if k goes to infinity and n goes to-- so k goes to infinity with n , but not too quickly. But k is little $o(\log n)$. And we find that this quantity here is essentially 2 to the k -- 2 to the $2k$. And this guy here is little $o(1)$. So λ is at least $2^{\sqrt{k}}$ minus 1 minus little $o(1)$.

That proves, essentially, the Alon-Boppana bound, although a small weakening because we are-- this big eigenvalue, you might find might actually be very negative instead of very positive. But that's OK. For applications, this is not such a big deal. These are two different proofs. And now, we think about, are they really the same proof? Are they different proofs? Are they related to each other? So it's worth thinking about. They look very different, but how are they related to each other?

And one final remark. You already saw two different proofs as to-- I mean, that shows you this number, and you see where this number comes from. And let me just offer one final remark on where that number really comes from. And it really comes from this infinite d regular tree. So it turns out that $2^{\sqrt{k}}$ minus 1 exactly is the spectral radius of the infinite d regular tree. And that is the reason, in some sense, that this is the correct number occurring Alon-Boppana bound.

This is-- if you've seen things like algebraic topology or topology, this is a universal cover for d regular graphs. So I won't talk more about it, but just some general remarks, and you already saw two different proofs. So beginning of next time, I want to wrap this up and to show you-- to explain some-- what we know about are there graphs for which this bound is tight? And the answer is yes, and there are lots of major open problems as well related to what happens there. And then, after that, I would like to start talking about graph limits. So that's the next chapter of this course. OK, good.