

PROFESSOR: We've been spending quite a few lectures so far discussing Szemerédi's regularity lemma, it's applications in variants of the regularity Lemma. So I want to spend one more lecture, before moving on to a different topic, to tell you about other extensions and other perspectives on Szemerédi's regularity lemma. So hopefully, you all will become experts of the regularity lemma, especially the next homework problem set where there will be plenty of problems for you to practice using the regularity lemma.

So one of the things that I would like to discuss today is a hypergraph extension of the triangle removal lemma. So as we saw, the triangle removal lemma was one of the important application of Szemerédi's graph regularity lemma. OK. So that works inside graphs, but now let's go to hypergraphs. In particular, even in the case of three uniform hypergraphs where-- so let's set some terminology.

So when r uniform hypergraph, or simply abbreviated as r graph, consists of a vertex set and an edge set where the edge set consists of r couples. So the edges are r element subsets of the vertex set. So r equals the 2 corresponds to the graph case. And you can talk about sub-graphs or density, various counts all analogously to how we did it for graphs.

And you can imagine what the hypergraph removal lemma might look like. So let me write down a statement. That for all r graph h and ϵ bigger than zero, there exists a δ such that-- so the last time, there was some complaints about sentences going on too long, so let me try to cut sentences into smaller parts. So if g is an r graph with small number of copies of each, so h is sub-graphs, then g can be made h free.

So far, everything's still the same as in the graph case, but now in the graph case, the number of edges is quadratic, most quadratic. Here, it's a most n to the r , so I want to make this graph, this r graph, h free by removing less than ϵn to the r edges from g . So that's the statement of the hypergraph removal lemma. So it's an extension of the graph removal lemma. Any questions about the statement?

So before discussing the proof, let me show you why you might care about this statement. And we used the triangle removal lemma to deduce Roth's theorem. Remember, there was a graph theoretic set up where start with a 3-AP-free subset. We set up a graph, and then that graph has some nice properties that allows us to use a corollary of the triangle removal

lemma, namely the corollary that says that if you have a graph, where every edge sits on exactly one triangle, then it has a sub quadratic number of edges.

So we can do a similar type of deduction showing that the hypergraph removal lemma implies Szemerédi's theorem. So that's what we'll do. So let's deduce Szemerédi's theorem from the hypergraph removal lemma. So recall Szemerédi's theorem says that for every fixed positive integer k , if A is a subset of 1 through n , that is k -AP free, has no k term arithmetic progressions, then the size of A is sub-linear.

Instead of illustrating how to do this proof for general k , I'm just going to do it for the case of k close to 4. And you can look at the proof and it will be clear how to generalize. So we'll just illustrate the proof for 4-APs. Now, before even showing you what this proof looks like, you might wonder, do we really need the hypergraph removal lemma? Could it be that with the graph removal lemma and a more clever choice of a graph, you could prove Szemerédi's theorem using just that.

So we set up some graph previously where triangles correspond to 3-APs. Maybe you can set up some other graph where some other kinds of structure correspond to 4-APs. And it turns out the answer is emphatically no. So there's a very good reason for this that for-- these are things which we might go into more when we discuss additive combinatorics, but 4-APs is a pattern that's sometimes called complexity two. Where as 3-APs is a pattern which is called complexity one.

I won't go into the precise definitions of what this means, but the message is that you cannot prove the 4-AP theorem with just graph machinery. You really have to use something stronger. That there is a very real sense in which just a graph removal lemma is at least-- or Szemerédi's theorem, Szemerédi's regularity lemma is not enough. And so we really do have to go to hypergraph.

And this extra layer of complexity, like in this word sense of the complexity, introduces also additional difficulties in there. So that there makes significantly harder than graph removal lemma. So I won't even show you anything that's close to a complete proof, but I will illustrate some of the ideas and highlight some of the difficulties today.

But deducing Szemerédi's theorem from the hypergraph removal lemma is actually not so bad. So I will show you how to do that right now. So from the hypergraph removal lemma, even just for three uniform hypergraph, even for tetrahedra-- so into the triangles graph,

tetrahedron, we can have the following corollary analogous to all we have for triangles. So if g is a three graph such that every edge is contained in a unique tetrahedron, then g has sub-cubic number of edges.

So completely analogous to the one we have for triangles and the proof is identical. You read the proof and everything works exactly the same way, once you have the removal lemma. By tetrahedron, I mean the complete graph on four vertices. Complete three graph of four vertices. OK? So you four vertices and you look at all possible triples of vertices. All right.

So now let's prove Szemerédi's theorem, at least the 4-AP case. The general case is completely analogous. Of course, you have to go to higher order. Instead of three graphs, you have to look at r graphs. So just as in the proof of Roth's theorem, we're going to set up some particular three graph where let's look at a certain modulus m being $6n$ plus 1. The exact number is not so important here. I really just wanted this number is bigger than $3n$ and it's co-prime to 6, as well as t sum divisibility's by 2 and 3, that will be useful.

So let's build a 4-partite three graph g where the four vertex parts x, y, z, w , they all have m vertices each. And I'll show you what the edges are. So four vertices, little x and big X and so on. So here are the rules for putting the edges. I'll just tell you exactly what they are.

So I put in an edge xyz if and only if the following expression $3x$ plus $2y$ plus z lies in a . I put in the edge xyw if and only if $2x$ plus y minus w lies in a . OK? So xzw if and only if x minus the minus $2w$ lies in a . And finally, yzw if and only if minus y minus $2z$ minus $3w$ lies in a . So these are the rules for putting in edges of this hypergraph.

Of course, you have this hypergraph. And you might be wondering, why did I choose these expressions? So if it's not clear yet, it will soon be very clear why we do this. Just as in the proof of Roth's theorem using triangle removal lemma, let's examine the tetrahedra in this three graph. So what are the tetrahedra? Notice that $xyzw$ is a tetrahedron. So if all three triples are present, if and only if all four of these expressions lie in a .

Well, just like in a proof of Roth's theorem, these four expressions form a 4-AP. And the common difference is minus x minus y minus z minus w . OK? So they were chosen to satisfy this property. But furthermore, notice that I can't just put any expressions in. I put these expressions in with the very nice property that the i -th linear form does not use the i -th variable. So each expression really corresponds to an edge in this three graph. All right.

But we started with a set a that is 4-AP free. It follows that you don't have this kind of configurations unless the common difference is zero. So the only tetrahedra correspond to these trivial 4-APs. And just as in the proof of Roth's theorem that we saw from triangle removal lemma, the conclusion is that every edge lies in exactly one tetrahedron. And therefore, by the corollary, the number of edges is equal to little o of m cubed.

But on the other hand, how many edges are there? So for each of the four conditions here, so for each of these four parts, the number of edges is, well, you get to choose x and y whatever you want. And the last variable has a choices. So this implies that the size of a is little o of m , and m is on the same order as n , and that proves the theorem. OK. Any questions so far? Yeah.

AUDIENCE: Where do we use the m is co-prime to 6?

PROFESSOR: Great. Question's where do we use the condition that m is co-prime to 6? Anyone know the answer?

AUDIENCE: To solve for that last variable, divide by 2 or 3.

PROFESSOR: Great. So to solve for the last variable, we ought to maybe divide by 2 or 3. And to do that, you need to have co-primality with 6. Yeah, so I'm hiding a bit of details here. OK? But it's a great question. So if you work out the details of this statement here, every edge slicing exactly one tetra tetrahedron and also counting the number of edges, but more especially, the first sentence. You need to actually do just a tiny bit of work. Any more questions?

So this deduction is the same deduction as the one that we did four triangles, but you have to come up with a slightly different set of linear forms. And usually, if you're given a specific pattern, you know you can play by hand and try to come up with a set of linear forms. You can think about, also, how to do this generally. And more generally, this hypergraph removal lemma using this type of ideas, it allows you to deduce multi-dimensional Szemerédi theorem.

So if you give me some pattern, then a subset of the integer lattice in the fixed dimension avoiding that pattern must have density going to zero. So we stated this in the very first lecture. And I won't spell all the details, but you can follow this kind of framework and gave you that theorem. And I will post the problem where you asked to do this for a specific pattern, namely that of a geometric square, axis-aligned square. So if you have that pattern in \mathbb{Z}^2 , so it's worth thinking about how would you run this argument for that pattern in \mathbb{Z}^2 . Yes.

AUDIENCE: How close can the small ϵ be?

PROFESSOR: OK. The question's how close can the-- so you're asking what is the rate, this little ϵ gives? Let me address-- OK, so hold onto this question. I will address it once I discuss what is known about hypergraph removal lemma. And that's a great question and there's a lot of mystery surrounding what happens there. OK. Questions? Any others? Great.

So let's discuss this hypergraph removal lemma. And as I had warned you already at the beginning of lecture, this one is very difficult. So I mentioned in the very first lecture that the development of Szemerédi's regularity lemma was a stroke of ingenuity. But this one here-- but we saw that the proof and we did the proof of Szemerédi's graph regularity lemma in one lecture. And once you understand it, it's not so bad. You do the energy increment, conceptually it's not so bad.

But that when there is actually incredibly difficult. It's incredibly difficult both conceptually and technically. But I want to at least illustrate some of the ideas and give you some sense of the difficulty, like why is this difficult. So as we imagine, we have graph regularity. And to prove hypergraph removal, one would develop some kind of a hypergraph regularity method.

And the basic idea in hypergraph regularity or just regularity, in general, is that I give you some arbitrary graph or hypergraph, and you want to find some kind of partitioning, some kind of regularization into some bounded number of pieces, some bound amount of data so that that's a good approximation for the actual graph. Just like in the graph regularity case.

So let's try to do this. So what does this partition even look like? So here's an attempt. And of course, I call it an attempt because eventually, it will not work. But it's a very natural first thing to try. And maybe, I shouldn't call it naive because it's actually not a bad idea to begin with. OK, so let's see. So suppose you were given a three graph g . I'm going to just to help you remember what's the uniformity of each graph. I will denote in parentheses, in the superscript, so just help you remember that this is the three graph.

So in geometry, they also do this with manifolds. So if you put an n on top, it's an n manifold. But this 3 is for three graph. Suppose we partition the vertex set of this three graph similar to proof of Szemerédi's regularity lemma. Think about how the proof of Szemerédi's regularity lemma goes. So you have this partition, but in the proof, there is this iterative refinement. And each step you say, well, I have this notion of regularity. If it doesn't satisfy regularity, I can

keep cutting things up further.

So what's the notion of regularity that might get you some kind of vertex partition? Can anyone think of a notion of a regularity for three uniform hypergraphs? Yep.

AUDIENCE: So the same sort of thing, [INAUDIBLE] variations like [INAUDIBLE].

PROFESSOR: So let me try to rephrase what you're saying. So if we have a notion of regularity, let's say I have three vertex sets, V_1 , V_2 , V_3 . And I want that the density between these three, they do not differ from-- if I restrict these vertex sets to subsets that are not too small. Does this make sense? So here, this d is the fraction of triples that are edges of the hypergraph.

So this is the natural extension, natural generalization of the notion of regularity that we saw earlier for graphs. And indeed, it's a very natural notion. It is a nice notion. And actually, if you use this notion, if you use more or less precisely what I've written, you can run through the entire proof of Szemerédi's graph regularity lemma and produce a regularity theorem that tells you given an arbitrary three uniform hypergraph, you can decompose the vertex set in such a way that most triples of vertex sets have this property.

So the same proof as Szemerédi's regularity lemma implies-- so I won't write down the entire statement, but you get the idea. So for every ϵ , there exists m such that we can partition into the vertex set into at most m parts, even equitable, if you like, so that at most an ϵ fraction of triples or parts are not ϵ regular in the sense that I just said.

So it's literally is the same proof. So you really look at the proof and then you get that. OK? So far everything seems pretty good, pretty easy. OK. So why did I say, initially, that actually, hypergraph regularity is incredibly difficult? So what not good about this one? Also remember, in our application of the regularity method, there were three steps. What are they? Partition, clean, and count.

OK. So partition, OK, you do partition. Clean, well, you do some kind of cleaning. But counting, that's a big thing. And that's something that wasn't so hard. We had to do a little bit of work, but it wasn't so hard. And you can ask, is there a counting lemma associated to this regularity lemma? And the answer is emphatically no. And I want to convince you that for this notion of regularity, there's no counting lemma. OK. Yes.

AUDIENCE: Is this version true, though?

PROFESSOR:

It is true. So you ask, is this version true? So this statement is true with this definition. And you can prove it by literally rerunning the entire proof of Szemerédi's graph regularity lemma. So the regularity statement I've written down is true, but it is not useful. For example, it cannot be used to prove the tetrahedron removal lemma because if you try to run the same regularity proof of the removal lemma, you run to the issue that you do not have a counting lemma.

OK. So why is it that you do not have a counting lemma? So let me show you an example. And keep in mind that the notions of regularity, they are supposed to model the idea of pseudo randomness, which is the topic we'll explore in further length in the next chapter. But the idea of pseudo randomness is that I want some graph which is not random, but in some aspects look random. So this is an important concept in mathematics and computer science, and it's a very important idea.

But of course, you can generate a pseudo random object by just taking a random object, and it should hopefully satisfy some properties of pseudo randomness. So let's see what this notion of regularity, how it works even for random hypergraph. What's a random hypergraph? So there are different ways to generate a random hypergraph. One way is to have a bunch of triples all appearing uniformly at random, independently at random.

So I have a bunch of possible triples that I can make as edges, each one I flip a coin. But there's a different way, and let me show you a different way to generate a random three graph. Let me give you two parameters, p and q . They are constants between 0 and 1. So let's build first a graph, so a random two graph which I'll call g_2 . So this is just the Erdős-Rényi random graph, g_{np} . The usual one where you flip a coin for each edge so that each edge appears with probability p independently.

And now, I make the actual three graph that want, g_3 , by including every triple, so every triangle of g has an edge. So here, an edge means a triple-- edge is three vertices in the hypergraph-- with probability q . So it's a two-step process. I first generate a random graph, and then I look at the triangles on top of that graph. And each triangle I include as a triple with a probability q . If you like, q be even one. So I do a random graph and my hypergraph is a set of triangles.

And let's compare this construction to the more naive version of a random hypergraph where we look at this hypergraph of graph where I put in each triple appearing independently with probability $p^3 q$. So these are two different constructions of random hypergraph. And you

can check that they have basically the same edge density. So how many edges appear in the first one? While the density of triangles in G_2 is p^3 , and each of those triangles appears an edge further with probability q .

So they have similar edge densities. And furthermore, you can check that this condition here is true for both graphs. So both graphs satisfy this notion of epsilon regularity as justifying with high probability. OK. Great. So if you have the counting lemma, it should give you some prediction as to the number of tetrahedra that come directly from the densities, in particular, they should be the same for these two constructions. But are they the same?

So the density of tetrahedra in the first case, actually, let's do the second case first. In G_2 , so what's the density of tetrahedra? So if I have four vertices, so each of those three edges appear uniformly at random, independently, so the density of tetrahedra is just the edge density raised to the power of 4. What about the first one?

AUDIENCE: 6.

PROFESSOR: So in the first one, to get a tetrahedra, the underlying graph needs to have a K_4 . So p^6 raised to 6, and then on top of that, I want $4q^3$'s. So p^6 raised to 6, q raised to 4. And when p is different, these numbers are different. So this is an example showing why there is no counting lemma because you have two different graphs that have the same type of regularity and densities, but have vastly different densities of tetrahedra. Any questions about this example?

It shows you, at least, why this naive attempt does not work, at least if you follow our regularity recipe. But in any case, it's good for something. So you do not have a counting lemma for tetrahedra, but you can still salvage something. So it turns out there is a counting lemma if your graph H is linear. So linear means every pair of edges intersecting at most one vertex.

So for example, if you look at-- so hypergraph or each line is an edge of triples. So that's a linear hypergraph because each pair intersecting at most one vertex. Tetrahedron is not linear. Two faces of a tetrahedron can intersect in two vertices. OK. So we can try to prove that this is true. And actually, the proof is basically the same as the counting lemma that we saw for graphs. Yes.

AUDIENCE: How many edges can a linear graph have?

PROFESSOR: The question, how many edges can a linear hypergraph have? You mean, given the bounded

number of vertices. OK. I'll leave you to think about it. Any more questions? But for the graph that we really care about, namely tetrahedra which relates to Szemerédi's theorem, this method does not work. So what should we do instead?

Let's come up with a different notion of regularity. And that's somewhat inspired by that example up there where we need to look at not just triple densities between three vertex sets, but also what happens to triples that sit on top of a graph. So we should come up with some notion of an edge density on top of two graphs.

So given a , b , and c being edge sets of a complete graphs, so these are graphs a , b , and c -- you should think of them as graphs-- and a three graph g , we can define this quantity b of abc -- so there's always a hidden g which I'll usually omit-- to be the fraction of triples xyz where xyz are such that they sit on top of abc .

So yz lie in a . xz lie in b . xy lie in c . So diffraction off such triples that are actually triples of g . In the case when abc are the same, it's asking what fraction of triangles are edges of g . But you are allowed to use three different sets. So think of abc as red, green, blue masking for fractions of red, green, blue triangles that are actually triples of the hypergraph. All right.

So now, we can try to come up with some notion of regularity. And as you might expect, at this point, it's not sufficient to partition the vertex set. Instead, we'll go further. We'll partition the edge set of the complete graph. So we'll partition the set of pairs of vertices. Let's partition each set of the complete graph as a union of graphs such that we would like a similar type of regularity condition but for those types of densities.

Such that for most ijk such that we have lots of triangles on top of these three graphs in the partition. So for most ijk , such that this is the case, this partition, so this triple is regular in the sense that for all subgroups with not too few copies, so not too few triangles, on top of these a 's. One has that the density, the triple density, among the g 's is similar to the triple density on top of the a 's.

So I'm doing some kind of partition where g_1 is like that. g_2 like that, and g_3 that. And what I'm saying is that if you take subset of g_1 , g_2 , g_3 so that there are still lots of triangles, and that's analogous to this condition of a i 's now being too small, then counting the number of triples to the fraction of those triangles that are edges of g . That fraction is roughly the same when you pass down to sub-graphs.

Don't worry about the specific details, and I'm not going to try to give you the specific details. But think about the analogy to instead of partitioning the vertex set, we are partitioning the edge set of a complete graph. But actually, hypergraph regularity involves one more step, namely that we need to further regularize these g 's via partitioning the vertex set. Similar to what happens in Szemerédi's graph regularity lemma, but actually more similar to the strong regularity lemma that we discussed last time.

So the data of hypergraph regularity is not simply a partition of the vertex set, but it's twofold. One is a partition of the edge set of the complete graph, so partition of the vertex pairs, into pseudo random graphs, so in into graphs, so that the hypergraph g sits pseudo randomly on top. And furthermore, there's also a partition of the vertex set of g so that the graphs in part one are extremely pseudo random with respect to this partition.

And this idea of extremely random we saw in the last lecture, you have some the sequence of epsilons that depend on how many parts you have in the first step of the regularity. OK. Any questions? Yes.

AUDIENCE: What happens with triples g 's that don't have a lot of triangles?

PROFESSOR: The question is, what happens to triples of g 's that do not have lots of triangles? So they are similar to in graphs, you have these small sets of vertices. So you have to deal with them somehow, but I'm, again, leaving out all these technical details. And in fact, I am writing down a very sketchy version of hypergraph regularity. You could write down a more precise version. You can find it in literature. In fact, you can find more than one version of the statement of hypergraph regularity in literature.

And they're not all obviously equivalent. It actually takes a lot of work even to show that different versions of the statement are equivalent to each other. And it's still somewhat mysterious as to what is the right, the most natural formulation of hypergraph regularity. That's something that I think we still do not yet have a satisfactory answer.

There was a question earlier about bounds. OK. So what kind of bounds do you get for hypergraph regularity? So let me address that issue now. So what kind of bounds do you get? Well for Szemerédi's graph regularity lemma, the bound is a power function because we have to iterate the exponential which comes out of the partitioning.

And in hypergraph regularity, because of this extremely pseudo random-- so you are doing

some kind of partitioning the first stage, and then you are iterating that on top for the second stage. Similar to how we did strong regularity in the last lecture. So the bounds for hypergraph regularity is also iterated power which we saw last time, and this is known as a Wowzer. So it's even worse than graph regularity.

And just like in the case of graph regularity, this was Wowzer type bound is necessary, at least for most statements, any of these useful statements of hypergraph regularity. What about the applications? So applications to multi-dimensional Szemerédi's theorem-- OK, so first of all to Szemerédi theorem, well, you can prove Szemerédi theorem this way and you would get inverse Wowzer type bounds, which is it's not so great. But there are better proofs. So there are more efficient proofs quantitatively.

So for Szemerédi's theorem, the best result for general k is due to Gowers' which tells you that a must be, at most, n over something that's $\log \log n$ raised to power some constant c depending on k . That's for k equals 3 and 4, you can do somewhat better before general k . This is the best bounds so far.

But for multi-dimensions, for multi-dimensional patterns, it turns out that-- well, historically, the first proof of the multi-dimensional Szemerédi's theorem was done using Ergodic theory which has even worse bounds compared to this approach in that the Ergodic theoretic proof gives no bounds because it has to use compactness arguments, so they actually give no quantitative bounds. And one of the motivations for this hypergraph regularity method, the removal lemma, is to produce quantitative proof of multi-dimensional Szemerédi's theorem.

So in general, still the best bounds come from this removal lemma, so hypergraph removal lemma. Although in special cases, and really, not that many special cases, but really just the case of a corner as we saw earlier you have somewhat better bounds. So for corner, you have bounds, which have density like polylog log . But even for a geometric square, we do not know any Fourier analytic methods, we do not know other methods. And this is basically the best bound coming out of hypergraph regularity.

And there are serious obstructions for trying to use Fourier methods to do other patterns such as geometric square. OK? Any questions? Yes.

AUDIENCE: So what do the bounds look like for higher degree uniformity. Are they still just Wowzers?

PROFESSOR: OK. So what are the bounds like for higher degree uniformity? So this is Wowzers for free

uniform, and for full uniform, you iterate Wowzer. So you go up in Ackermann hierarchy. You iterate Wowzer, you get a four uniform hypergraph regularity lemma as so. Let's take a short break.

So the second topic I want to discuss today is a different approach to proving Szemerédi's graph regularity lemma. And this is a good segue into our next topic, the next lecture, which is about pseudo random graphs, in particular, the idea of the spectrum eigenvalues, in particular, play a central role. So I want to consider a spectral approach giving an alternative way to prove the Szemerédi regularity lemma.

And if you're already sick of the regularity lemma at this point, this will be the last topic on regularity lemma for now, although it will come up again later in this course when we discuss graph limits. But for now, this is the last thing I want to say. And just like the discussion about hypergraph regularity, it will be somewhat sketchy. So this idea has appeared in literature in the past, but it was popularized by many good things in life by Terry Tao's blog. So it's a good place to look up a discussion of what I'm about to say.

OK. So we saw the proof of regularity lemma via this iterated partitioning and keeping track of our progress through the use of an energy. But here's a different perspective, namely if we start with a graph G , I can look at the adjacency matrix A of G . So this is the n by n matrix where n is a number of vertices whose (i, j) -th entry is zero if i is not adjacent to j , and 1 if i is adjacent to j . So this is a pretty standard thing to look at to associate a graph to this matrix.

So this graph here would be like that and so on. It's a real symmetric matrix and that's always pretty nice. Symmetric matrices have lots of great properties that will be convenient to use. In fact, if you're like myself, if you're too used to working with symmetric matrices, you forget that some of these properties actually do not apply in general to non-symmetric matrices. But it is symmetric, so we're happy. So for symmetric matrices, we have a set of real eigenvalues.

We have real eigenvalues and eigenvectors. And for now, let me enumerate the eigenvalues by λ_1 through λ_n , so multiplicity included, and I sort them according to the size of their absolute value. So the spectral theorem tells us a decomposition.

So here again, we're using the A as real symmetric. So it tells us that this matrix A can be written as the sum coming from the eigenvalues and eigenvectors where the u_i 's are the eigenvectors, but I can choose them so that they form an orthogonal basis orthonormal basis, so they're all unit vectors. So when I say the spectrum, I mean this data also, specifically, this

set of eigenvalues.

All right. So let's go through some basic properties of the spectrum. So first, how big can the lambdas be? So I claim that-- so first of all, the sum of the squares of these lambdas is-- let me not even call this a lemma, so it's just an observation. So the sum of the squares is this one. So this is the trace of a squared, which is also the sum of the squares of the entries. So here, I'm always using that A is real symmetric-- sum of squares of entries of A .

And the case when you have A being an adjacency matrix, this is simply twice the number of edges, which is at most n^2 . So that's always a good thing to remember. OK? So as a result, the i -th eigenvalue cannot be bigger than what? So you have i eigenvalues. So they're sorted in decreasing order. So the i -th eigenvalue cannot be too large, particular it cannot be larger than n/\sqrt{i} . Because otherwise, the sum of the first i eigenvalue squared would exceed n^2 . So these things, they do decay.

So second observation is that if you have some ϵ and an arbitrary function, so this is known as a growth function. That's just a name, don't worry about it. So which we'll call f . So its function from the positive integers, the positive integers, and for convenience, I'm going to assume that $f(j)$ is always at least j . For every j there exists some c which depends only on your ϵ and this growth function. So this growth function plays the same role as the sequence of decaying ϵ 's in these strong regularity lemma.

So there exists some constant bound such that for every graph G and A_G as above, so associated with the lambdas and u 's, there exists a j less than c such that if I sum up the eigenvalues squared for eigenvalues i index between j and c of j , the sum is fairly small. It's at most ϵn^2 .

I'll let you ponder that for a second. So choose your favorite growth function. It can be as quickly growing as you can. It can be exponential, power, or whatever. And it's saying that I can look up to a bounded point so that this stretch of spectrum squared is, at most, ϵn^2 . Question.

AUDIENCE: What is c of j ?

PROFESSOR: f of j . Thank you. Well, the statement hopefully will become clearer once I show you the proof. OK. So here's how you would prove it. So you first let j_1 equal to 1, and I obtained the subsequent j 's by applying f to j . So I claim that one cannot have this inequality violated for too

many of these j 's. So one cannot have the sum going between j_k and j_{k+1} for all k from 1 to $1/\epsilon$. Let's change this to zero.

But you cannot have this because if you had this then you sum up all of these inequalities you would get that the total sum would exceed ϵ^2 , which would violate the inequality about sum of the squares of the spectrum. And so therefore, so thus to the claimed inequality, so this is true star holds for $\sum_{j=1}^k j_i$ so j_k where k is less than $1/\epsilon$. And this j , in particular, is less than-- well, whatever it is, it's bounded. So it's less than f applied to itself at most $1/\epsilon$ times--

OK. So this should look somewhat familiar. And I'll ask you to think about, later on, how this proof of spectral proof of Szemerédi's graph regularity lemma compare to the proof that we saw earlier. And you should see where the analogous step is here. This is that density. This is the energy increment step. All right. OK.

So what's the regularity decomposition? So I give you this graph. I give you this adjacency matrix. And I want be able to find a partition, but there's a different way to view a partition. So this is, I think, a important idea which, again, is popularized by Terry Tao, that instead of looking at things as a regularity partition, we can view these ideas as regularity decompositions.

Namely, pick j as in the lemma and I now write my adjacency matrix a sub g as a sum of three matrices, which we'll call a structured plus a small plus a pseudo random. Where a structured equals to the sum for basically that sum, this sum here, so this spectral de-composition but only for the first $j-1$ eigenvalues.

So those of you coming from or who have taken classes in something like Statistics might recognize this as a principal component analysis. So this has many names. It's a very powerful idea. You look at the top spectral data, and that should describe you most of the information that you care about about a graph or a matrix, in general. The small piece is the sum but only for i between j and f of j . And the pseudo random piece is for i at least f of j . OK.

So we decompose this adjacency matrix into these three pieces. And the question now is, what does this have to do with Szemerédi's graph regularity lemma? So what do the individual components correspond to in the version of the regularity lemma that you've seen and are now familiar with? So here is what's going on. So I want to show you that this structured piece roughly corresponds to the partition. So this is the bounded partition.

And the small piece roughly corresponds to the small fraction of irregular pairs. And the pseudo random piece roughly corresponds to the idea of pseudo randomness between pairs. First, to understand what the spectral data have anything to do with partitions, let me remind you a basic fact about how the spectrum, how the eigenvalues of a real symmetric matrix relate to other properties of this matrix. And namely, this notion of a spectral radius or sometimes called spectral norm.

So far I'm only going to discuss real symmetric matrices. So many of the things I will say are not true for if you're not in a real symmetric case. So the spectral radius spectral norm of A is the largest eigenvalue of A in absolute value. And this quantity turns out to be equal to the operator norm which is the norm of this A as a linear operator, namely it is the max or super, it turns out to be a max, of $\|Av\|$ over-- length of Av divided by length of v . So if you hit it with a unit vector, how far can you go?

So it's also equal to this bi-linear form. If you hit it from left and right by unit vectors, how big can you get? So for real symmetric matrices, these quantities are equal to each other. And that will be an essential fact for relating the spectral data with combinatorial quantities. All right. So if you give me this de-composition, how can I produce for you a partition?

Basically, you can look at a structure which has its state in its data a bounded number of eigenvectors. And by rounding, we can basically round these guys so when you round the individual values by rounding the coordinate values-- So let's pretend that they take only a bounded, let's see, a small number of values.

So just to simplify things in your mind, pretend for a second-- well, of course, this is far from the truth-- pretend for a second that these guys are 0 comma 1 valued. Of course, that's not going to be the case but 0 comma 1 or plus/minus 1, if you like. OK. So this is definitely not true. But for the purpose of exposition, let's pretend this is the case. And you can more or less achieve it by rounding the individual values to their nearby closest multiple of something.

Then the level sets of these top eigenvectors, they partition the vertex set into a bounded number of parts. So if you, for example, in the simplified version where you're only have plus minus 1 values for this eigenvectors, then you have, at most, 2^j parts. But you may get some more because some epsilons, but for the purpose of illustration, let's not worry about that.

And this is basically the regularity partition. I want to show that this set here has very nice properties that they basically behave like the regularity partition we've gotten previously. So what I would like to show is that the other two parts, they do not contribute very much in the sense of our regularity partition. So for example, if you look at the pseudo random piece, if I hit it left and right with indicator vectors of vertex sets, how big can this number get?

So this number here is, at most, while the norm of u times the norm of this w which is just-- so let me write down-- so the norm of indicator of u , norm indicator of w multiplied by the operator norm of the pseudo random part of a . But these two guys here, so they're, at most, root n each. So this number here is, at most, n but we know from our hypothesis on the pseudo random part of a that the spectral norm is no more than this quantity over here.

And by choosing f appropriately large, I can make sure that this number is extremely small. f to be large compared to the number of parts in b partition so this quantity is small. And this is basically the notion of epsilon regularity that you saw in the usual version or version of Szemerédi's regularity lemma I presented in the first version, in the very first lecture that we discussed regularity.

This quantity here is something which measures the difference between-- so for now, if for a second ignore the middle piece. If you ignore the small piece, then this is precisely the difference between the actual densities between u and w and the predicted density between u and w .

AUDIENCE: Why is there a square root here?

PROFESSOR: The question is, why is there a square root here? There should not be a square root here. Good. It then becomes a square root of the-- yeah, so there is no square root, but the length of this vector is the square root of the size of u which is, at most, n . Yeah.

AUDIENCE: Did you say to be small in general or just small like you compare it to f squared? Because I guess, isn't it like going to be this constant function that you choose before [INAUDIBLE]?

PROFESSOR: OK. Question is do, how small do we want this f of j to be? So I want this quantity to be quite a bit smaller than, let's say-- so basically, I want it to be less than f of n squared, but f of n over the number of parts square. Because this quantity is, let's say, the sizes of each part.

So let me just be not precise and say much less than. So this quantity here is the size of each part. And I want to think about the case when u and w they lie inside each part. In which case,

I want the difference to be much less than epsilon times the size of the part squared. Yeah.

AUDIENCE: [INAUDIBLE] j is different based on the graph?

PROFESSOR: Question is, is the j different based on the graph? Yes. And that's also the case for Szemerédi's regularity lemma. In Szemerédi's regularity lemma, you don't know when you stop. But you know that you stop before a certain point. OK.

And finally, what's happening with a small part of a . So in a small, the sum of the squares of entries-- so this also has a convenient name. It's called the Hilbert-Schmidt norm. So the sum of the squares of the entries. We basically saw this calculation earlier, it's the sum of the squares of the eigenvalues in which case we've truncated all the other eigenvalues. So the only eigenvalues left are between j and index between j and f of j . And we chose j so that this number is small.

So a small a as in a bunch of noise, but no adversarial noise, if you will, into your graph, but only a very small amount, at most, epsilon amount. So it might destroy the epsilon regularity for, let's say, around epsilon fraction of pairs. But that's all it could do. So all but epsilon fraction of your pairs will still be epsilon regular. And that is the consequence of Szemerédi's graph regularity lemma that we saw earlier. Yeah.

AUDIENCE: Doesn't large F have to be special?

PROFESSOR: OK. So question, does a large F have to be special? The F should be chosen-- if you want to achieve Szemerédi's graph regularity lemma, you should find this F so that basically this inequality is true. So f should be quite a bit larger than the number of parts. But if you choose even bigger values of f , you can achieve more regularity. And this is akin to what happened with strong regularity. So there's this idea if you iterate one version regularity, you can get a strong version of regularity. And there's some iteration happening over there.

So if you choose your f to be a much bigger function of j , you can achieve a much stronger notion irregularity which is similar and, perhaps, even equivalent to strong regularity that we discussed last time. So you get to choose what f you want to put in here. Yeah.

AUDIENCE: How do you make it equitable?

PROFESSOR: The question is, how do you make it equitable? OK. So let me now discuss that. So in this case, you can also do very similar things to what we've done before, but to massage the

partitions. It's not entirely clear from this formulation. But the message here is that there's this equivalence between operator norm on one hand and combinatorial discrepancy on the other hand. And we'll explore this notion further in the next several lectures.