[SQUEAKING]

[RUSTLING]

[CLICKING]

LAWRENCE GUTH: So today is the last day of our additive combinatorics section of the class. And we get to prove one of the core theorems of additive combinatorics, the Balog-Szemeredi-Gowers theorem. And I think it's a wonderful thing and I will try to do it justice. I'm learning myself also. So first let me remind everybody what it says.

All right. So If A and B are in an abelian group, and let's say they have size at most N, and we have a subset of A cross B, so a subset of pairs, little a little b, and it's pretty big. So X is bigger than k inverse n squared. So k is a parameter here. But you should think of k as being significantly smaller than n. That's the regime where there's a meaningful conclusion. And pi 1 of X is at most K So remember that pi 1 of X is the set of a plus b where a comma b is in X.

So this is surprisingly small given how large X is. And it means that in this set there are many pairs, a1, b1, and a2, b2 that they add up to the same number. So there are a lot of additive coincidences happening when we add up the pairs in x. Then the conclusion is that there is a subset A prime in A, a subset B prime in B, which are pretty big. So the way we'll say that they're pretty big, so we could say the following is true that A and B are each at least some power of K times N. So they're each pretty big.

But actually, we'll say something slightly fancier which is what we really used in applications. Namely if X prime is defined to be the part of X in A prime times B prime, then X prime is pretty big.

And also when you take sums of A prime and B prime, not just the ones in X prime, but all of them, those are all pretty small. So A prime plus B prime is less than K to the big O of 1 times N.

So I guess last time we introduced this and talked about examples a little bit. But let me just repeat the example because it helps a lot to digest the statement. So here is one example, A and B are both the numbers up to N with some garbage. And then X could be any subset of N cross N, which is pretty big.

And so now it's not true that A plus B is small, but it is true that a good maybe union a little bit. OK, so pi 1 of X is small because most of X is contained in here, And pi 1 of that is small. So A plus B, it might be large. But pi 1 of X is small. And the only way this can happen is that in this case, a big chunk of X, most of X is contained in a product that actually has a small subset. So A prime equals B prime equals N. Does the statement make sense? All right, cool

I'll tell you a little bit of history of this theorem. So Balog and Szemeredi were the first people to think about this. And they proved a version of this theorem. But in Balog and Szemeredi's theorem, this K to the minus big O of one was much worse. So here there was, in Balog-Szemereri, there was like 1 over some function of K. And over here in Balog-Szemeredi, there was some function of K. And this function of k was really terrible. It was worse than a tower of exponentials.

And it followed from something called Szemeredi's regularity lemma, which is a really important qualitative fact in graph theory and combinatorics, which has many consequences. But it's well known that the bounds in Szemeredi's regularity lemma are really terrible. And that's where these really terrible bounds come from.

Now, in Szemeredi's regularity lemma, those really terrible bounds are the best thing that's true, more or less. But in some of the applications they're not. And Gowers revisited this. And he proved that this thing with the polynomial dependence.

So the proof that I'm going to show you is from the 1990s. So it's really not that old. It's been very influential. And it's not that complicated. It's clever, but it's something that we can do in a class and this is a significant thing. And I think it's something worth seeing and worth knowing about. All right. So any questions or comments before we start the proof?

So X is a subset of A cross B. And we're going to visualize it as a bipartite graph, and visualize X, which is a subset of A cross B. So over here, I have A which I'll visualize as a bunch of dots. Over here I'll have B, which I'll visualize as a bunch of dots. And if A comma B is in X, then I will visualize that by drawing an edge from A to B. So the set X will look something like this. So in this picture, A comma B is in X, if and only if, I have an edge.

And now it turns out to be a nice connection between some graph theory things you can see in this graph and the additive combinatorics of A and B. So let me make a definition that P sub K of a to b is the number of paths that have length K and they go from a to b. And their paths are in this graph, in the graphics.

I had to take P3 of A comma. B is going to be very important in our story. How many ways are there to get from A to B in three steps? So there may not be any in the picture right now. Let me add an edge. So there's one way to get from A to B in three steps. And here's another way to get from A to B in three steps. P3 of AB is counting how many ways there are to do that. All right.

So here's a lemma.

AUDIENCE: Can I have a question?

LAWRENCE GUTH: Yeah.

AUDIENCE: Are the [INAUDIBLE] allowed to repeat notes?

LAWRENCE GUTH: They are allowed to repeat notes. Yeah, you're right. So they might have been some I didn't count before. So that is also a path of length 3 from A to B. Yeah. Thank you. Yeah.

So if I have A prime in A and B prime in B, if you look where we want to go, we'd like to eventually prove that for some sets that we haven't found yet that A prime plus B prime is small. So how are we going to prove that?

A prime plus B prime, the size of A prime plus B prime is going to be related to the paths of length 3 in this graph. So if I have this and if for every pair A and B, A and A prime, and B and B prime-- there are lots of paths between them. Then A prime plus B prime isn't very big. It's less than pi 1 of X cubed over p. So let's prove this lemma.

All right. So let's visualize a path of length 3 from A to B. So it looks like this. There's a path of length 3. And what is this telling us? Well, this edge tells me that A comma B1 is in x, right? Hence therefore, A plus B1 is in pi 1 of X. And this edge tells me similarly that A1 plus B1 is in pi 1 of X. And this edge tells me similarly that A1 plus B is in pi 1 of X.

Now I'm interested in A plus B. And A plus B is related to these guys, and so we can write A plus B as A plus B1 minus A1 plus B1 plus A1 plus B. Did I do this right? So A plus B, the A1's cancel and the B1's cancel. And these three characters, they correspond to these three edges. So let's call this Z1, Z2, and Z3,

So we have seen that the number of triples-- Z1, Z2, Z3, and pi 1 of x cubed, so that A plus B is Z1 minus Z2 plus Z3. That number of triples is at least P3 A to B.

So if there are a lot of paths from A to B of length 3, then there are a lot of ways of writing A plus B by combining guys in pi 1 of X. All right. And we're supposing that P3 of a, b is always at least P. So we have this.

And then we can conclude from this that the size of A plus B times P is at most pi 1 of X cubed. Because for every sum in A plus B, I can write this sum as Z1 minus Z2 plus Z3 at least p different ways. And for each Z1, Z2, and Z3, this is only going to give me one number when I do.

AUDIENCE: That's A prime and B prime, right?

LAWRENCE GUTH: Those are A prime and B prime, right.

So I'm digesting along with all of you. And as I've been reading the proof, I've been trying to figure out how did people think to do this and do these different things. So let's try to give some big-picture strategy while we might look at this. So pi 1 of X is small. That tells us that there are a lot of coincidences where you can write a number as A plus B in a lot of different ways. And how might that be helpful?

So the fact that pi 1 of X is small, that tells us that there are going to be a lot of edges. What do I say? So by assumption there are a lot of edges in this graph. That's one of our hypotheses that X is big. So there are a lot of edges in this graph. And therefore there are lots of this paths of length 3 in this graph. We'll figure that out later. And when there are many different paths of length 3 between these two guys, then it means that there are lots of ways of representing A plus B by combining guys in pi 1 of X. On the other hand, if pi 1 of X is small, then that should make this not too large.

OK. I don't think that really explained how somebody might have thought to do this, but it's fairly short, let's take a second and see. I'll let you digest a little bit and see if people have anything to add. Yeah?

AUDIENCE: The idea of the paths of three different things, looks kind of like you're taking A minus A plus A and being similar to the kind of continuous structure idea. Is there a length to those theorems?

LAWRENCE GUTH: Yeah so the question is this felt a little bit like [? plunica ?] and is there a link to the [? plunicane ?] equality. Maybe Yeah?

AUDIENCE: Is there something special about paths of length 3? It seems like this should work for any odd number of paths.

LAWRENCE GUTH: Yeah OK. So the question is, why did we pick paths of length three. There's really nothing special about paths of length 3. It would be some similar estimate for any length. And 3 is the length where the later part of the proof, we managed to get it to work. OK.

So next, let's think a bit about how many edges there are in this graph and how many paths there are of length 2, and how many paths there are of length 3, and stuff. Let's just orient ourselves to this graph. So OK, I'll call this some basic estimates and PK, PK of AB. Actually, let me frame it this way. Let me state a lemma about paths of length 3. And then we'll start to do some things about that.

So to finish the proof of Balog-Szemeredi-Gowers, there's a key lemma. And it's just a lemma in graph theory. So all of the additive structure went into writing this and noticing that things cancel. And now we have our graph theory lemma about paths of length 3 in a graph. So the main lemma says that if X is in. A cross B that X is bigger than K inverse times the size of A times the size of B.

Then there are big sets with lots of paths between them. So then there exists A prime in A and B prime in B so that if we define X prime is the part of X in A prime cross B prime, then X prime is pretty big. And for every A and A prime and B and B prime the number of length 3 paths from A to B is at least K to the minus O of 1 A B.

So if you combine this lemma and this key lemma, and just put them together, you get the Balog-Szemeredi-Gowers theorem. And the reason that 3 appears here is it's the smallest number where this key lemma is true.

So I'll let you in your mind just fill in this is our P, and fill it in for P over here, and you do a little algebra, and that's the Balog-Szemeredi-Gowers theorem.

So if it's OK with everybody, I'm going to erase this and the rest of the class, we just focus on proving this lemma from graph theory, with the rest of the proof of BSG.

So let's orient ourselves a little bit about how many paths there are in this graph between different points, and prove some simple things first before we prove this lemma. So I'll call this a simple bounds about PK a, b.

All right. So first let's think about how many edges there are in this graph. So the number of edges is the size of X, which is at least K inverse A B. And then let's think about how many paths there are of different lengths. So if I write PK without an A or B, this is the number of paths of length K, and they could start or end anywhere. But I will just make the restriction that they start in A. It turns out it's in a bipartite graph It's simpler to count things like this.

So P1 is just the number of edges in the graph, paths of length 1 are just individual edges. So this is the size of X, which is at least K inverse A B. Now a helpful thing to talk about is the neighborhood of any vertex. So if I have a vertex A over here. This is A and this is B. I have vertex A over here. There are some edges coming out of it. And this set here is called the neighborhood of A. And you can do it in the other directions also. It's the neighborhood of A. Yeah?

AUDIENCE: There's two different K's up here, right?

LAWRENCE GUTH: There are two different K's. Yeah there are two different K's. I apologize. Yeah. Yeah. Let's do this. So P2, so that's the size of P1. And then if I were to take the average A and A of P1 of A to anywhere, that would be P1 divided by the size of A. And so that would be at least K inverse times the size of P. So this is the average degree of one of these vertices on the left. And it's at least a decent fraction of all the vertices on the right.

All right. Now let's think about paths of length 2. So P2, a path of length 2 looks like this. So there's some guy B that it goes through here, and then there has to be somebody who goes to B and somebody who comes from B. So this is the sum over b of the size of neighborhood of b squared.

Well, if I just were to take the sum over b of the size of neighborhood of b, that would count the edges. So whenever I'm interested in the sum of somebody squared and I have a nice formula for the sum of somebody, then it makes sense to apply Cauchy-Schwarz. So if you apply Cauchy-Schwarz here, you get the sum of N of b quantity squared over the size of b. So that's at least K inverse size of A size of B bar all squared over the size of B. And that works out to k to the minus 2 A squared B.

Now that kind of makes sense. So for reference, if we had the complete graph with all the edges, the number of paths of length 2 would be A squared B. You pick somebody in A, you pick somebody in B, you pick somebody in A. And so this says that the fraction of them is at least K to the minus 2. And that makes sense because the fraction of edges is K to the minus 1. And we want two different edges to both be in our path. So that makes K to the minus 2 a good reference point. And it's not necessarily equal to that. It's greater than or equal to that. And the reason is our edges may not be distributed evenly.

So if some vertices have way more edges than others, that will cause the number of paths of length 2 to go up. And again, we can take the average. So if I were to take the average over A1 and A2 of the number of paths from a1 to a2, that would be at least I would divide by A squared. So K to the minus 2 times B.

As an exercise you can keep going to longer lengths. So exercise is that P3 is at least K to the minus 3 A squared B squared, A squared B squared would be the number of paths of length 3 in the complete graph, and at least K to the minus 3, a fraction of them are paths in our smaller graph.

And that means that if I take the average on A and B of the number of paths of length 3 from a to b, that's at least K to the minus 3 times A times B. So if I pick two random vertices and I count the number of paths of length 3, on average, it's at least this big.

Now let's compare that to the key lemma. The key Lemma says that I can find some substantial subsets, so that within these substantial subsets, the number of paths of length 3 is always K to the minus 100 size of A size of B. So this number here is a small fraction of the average. I'm looking for pairs where the number of paths of length 3 is at least a small fraction of the average.

But the thing that makes this tricky is I don't want just like some A's and B's or most of the A's and B's to be at least a small fraction of the average. I want actually, for every A and B in our subset, I want them to be at least some fraction of the average.

So now we're eventually going to go up to paths of length 3. Yeah?

AUDIENCE: For this one, we don't need them to be all of them? If 80% of A cross B had a lot of paths between them, then could we still not have this [INAUDIBLE] hold for just like a worse bound?

LAWRENCE GUTH: Great, OK. So the question is in this Lemma as written, we have assumed that for every A and B in A prime and B prime, we have a lot of paths, which then matches the key lemma. But the question is, did we really need that in this lemma? Or could we have relaxed a little bit over here?

So suppose instead of this lemma that we only had this for most A's and B's. So let's do a little experiment. So in blue, I'll put an experiment. And the experiment is that instead of this, I'm going to say for 90% of pairs, a prime, b prime and A prime cross B prime. OK. So now let's see what would happen.

If I take an element of this sum before I knew that-- so let's say a prime plus b prime is in this sum. And then I can say that then I'm going to use this thing.

Now if I knew that 99% of the guys in this sum could be represented by a good pair, a prime, b prime, I would be happy. But that's not what I have written in my experiment. So my anxiety is 90% of the a primes and b primes are good. They have lots of paths. 10% of them are bad. But then I take the sum set, and the 90% produce a few sums, but the 10% produce a ton of sums. I'll make a picture.

So here is A prime cross B prime. This part of it is good. Good means that. And here's the 10% That's bad. Then I take the sum operation. But now life changes. So pi 1 of good is tiny. Pi one of bad is huge. All right.

And now I go down here. And for each of these sums, if the sums are good, they get counted many times. If sums are bad, they don't. But this isn't going to work if pi 1 is bad.

AUDIENCE: At this point, if instead of assuming for 90% if we assume that the average is bigger than p, then at this step, if we added together every p in AB, then would we get the same? Will the sum of the three AB over each AB would be less than the guys in pi 1 of X cubed, and then the sum of them will be the number of times the average?

LAWRENCE GUTH: Yeah so there's another suggestion, but I think it's similar. The other suggestion was instead of a 90% we could take the average over a prime b prime. Yeah. But the issue is that when we look sometimes at A prime cross B prime and sometimes we look at a prime plus b prime and their notions of average may be very different from each other, because there may be a big chunk of A prime cross B prime, which is compressed a lot when you add, and another chunk which is not.

So when you look at a prime plus b prime, we don't need every sum to have a lot of representatives. We just need most of them. But it's hard to keep track of the way A prime cross B prime is related to a prime plus b prime, and it seems to be the case that the easiest way to know that most of the sums are good is to actually know that all of A prime cross B prime is good. Yeah. OK. Great questions. Yeah?

AUDIENCE: Could you get around this issue by requiring that the bad subset is asymptotically small, that the size is 1 over the square root of the size of A? Because that way when you project down, you'll square it, but it will still be relatively small.

LAWRENCE GUTH: Right. OK, so the comment was that if we were to replace 90% by 1 minus epsilon, and if we tuned epsilon to be really, really small, eventually this would work. And that sounds right to me. We'll keep our eye out for the rest of the proof. Perhaps, that will allow us to simplify things. But it might not. Yeah. OK, great.

So let's remember where we are. We're trying to prove the key lemma that I can choose big subsets, so that for every pair a and b, there are lots of length 3 paths between them. And for every, maybe we could relax to almost every, but anyway, it's not that hard. It's actually not that hard to prove these theorems.

OK. So the reason we use 3 here is that, well, I think that this is false with 2, if we replaced it with 2. I don't quite have a counterexample. Buy still we're going to warm up by talking about paths of length 2 because they're simpler than paths of length 3. And we'll prove some weaker version for paths of length 2 that we'll use as a stepping stone.

So paths of length 2, so lemma, If X is a subset of A cross B and it's pretty big, and we have a parameter epsilon bigger than 0, Then I can choose subsets. It's actually just one subset, just a subset A prime. So that A prime is pretty big and P2 of a1, a2, is at least epsilon K to the minus 2 times b.

Now, not for every pair a1, a2. We're going to weaken that a little bit, but with an exceptional set of only about an epsilon fraction. So actually the people who set this up had the thought that was similar to yours, which is that let's allow an epsilon fraction of bad pairs, and we'll make that a parameter, and then we'll see what we can do.

And actually we put an epsilon in two different places. And so, all right, so let's process this first. Remember that the average number of paths from a1 to a2. If we average over all the choices a1 and a2 is at least this big. So we're going to be at least an epsilon fraction of the average. And that's going to be true not all the time but except for about an epsilon fraction of pairs. And A prime is a decent size it.

All right. So before we try to prove this, let me ask a question to people. So question, Can we just take A prime to be all of A? So then we would just keep this part. Is it plausible that the number of paths is at least an epsilon fraction of the average most of the time? It doesn't sound so crazy. Is it possible that that's always true?

Yeah, so the question is, what's the maximum that this could be? The maximum it could be is B because there are paths that start at a1, they'll go across to somebody in b. They'll come back to a2. So they're at most b. So there's a decent amount of range there. And that is important.

Oh, the answer to this question is no. And here's an example. Yeah, was there a comment?

AUDIENCE: And then could you maybe have two points in A connect them by a every point in B, and then every other point in a is not very connected?

LAWRENCE GUTH: Yeah, OK. So the suggestion was I take two points in A and I connect them to all the points in B. And then the rest of them are not very connected. Now two of them might not be enough, but it will take a few of them. Here is B. Here this A. This is a small piece. I guess this could be a small part of A. And then over here, there are no edges going out of there at all, right. So how big would this have to be? Well, we need a fraction 1 over k of all the edges in our graph. So this small part of A would have size 1 over k times all of A. And now I have a 1 over k fraction of all the edges.

So if I pick a1 and a2 in here, I have a ton of paths between them. But if either one of them is out here, there are no paths between them. And this set is actually much bigger than that set. So for almost all the pairs in A, there are no paths between them. So this is a counterexample. Good.

And there's a second counterexample. So in this counterexample, the suspicious feature is that these vertices have way fewer edges than those vertices. Could I also make a counter example where every vertex on both sides has the same degree? Yes, you can. And it works by dividing into clumps. So I have a first clump. I have all the edges in here. And then I have a second clump. And I have all the edges here. And so on.

So if you do this and you have k clumps of the same size, then you'll see you have about 1 over k of all the edges. And then if you pick a1 and a2, if you pick them in different clumps, then there are no paths between them. If you pick them in the same clump, there are lots of paths. So because of these two examples, we cannot just take A prime to be A. All right.

So we have to understand, we need to pick A prime. And now let's think about how are we going to pick A prime. And in both these examples, the way to find a plausible A prime is to pick some b and then look at its neighborhood. So here I pick some b and I look at its neighborhood. That's the set. So this should be A prime right? And here I pick some b anywhere, and I look at its neighborhood. And this is a reasonable choice for A prime.

So we will take A prime to be the neighborhood of b for some b.

So now, so let me make a definition. I'll say a pair a1, a2 is called epsilon bad if the number of paths from a1 to A2 of length 2 is smaller than our threshold there. All right. So the main thing we have to think about is I'm going to pick different b's and look at their neighborhoods. And for each neighborhood, I want to count how many bad pairs there are. So BP epsilon of B is the number of pairs a1, a2 that are both in the neighborhood of b, so that a1, a2 is epsilon bad.

All right. So we'd like to know about how big that is. And the easiest thing to compute is its expectation. So lemma, the expectation over B of the size of this thing is not too crazy. So let's see what that means.

So this is at most epsilon fraction K to the minus 2 A squared. All right. So let's digest this. So we'll prove, of course. But let's digest the right-hand side. So these bad pairs, they're sitting in just the set of all pairs. So for reference, we should think about how big is the set of all pairs. So let's take the expected value over b of just the size of its neighborhood squared. That's at least K to the minus 2 A squared. This has an easy proof that we'll do in a second.

All right. So if you compare these, you see that the fraction of epsilon bad pairs is at most a small fraction of the total pairs. Fraction of epsilon bad pairs is at most epsilon, on average. And that's good. That's the thing we want to prove. We want to prove there's only about an epsilon fraction of bad pairs.

So let's call this lemma P1 and lemma P2. This one is real easy. So I can keep it on this board. And then we'll prove lemma P1 which is the most important thing. So Lemma P2 proof, if I take the sum on b of N of b squared.

When you see this, what does it make you want to do? Actually, we already did it. So this is the sum on b N of b quantity squared over the size of B. That's at least C inverse A B all squared over the size of B. That's at least k to the minus 2 a squared B. And to turn the sum into an average, we divide by B, by the size of B [INAUDIBLE].

Actually we could say an even simpler thing. So what's the average size of just N of B? The average size of N of B is exactly K inverse times the size of A. And so by Cauchy-Schwarz, then if you square it, it's at least

So the heart of the matter for this lemma is that the expected number of bad pairs is a lot less than the expected number of total pairs. So let's prove this lemma.

All right. So proof of Lemma P1, so again, let's think about the sum on B of the size of BP epsilon of b. All right. So this is a set of pairs a1, a2, so this is a number of things. This is a double counting argument. It's a number of a1, a2, and b, so that a few things hold. So first of all, we have this connection pattern. And also a1 and a2 are bad, which means the number of paths from a1 to a2 is at most epsilon k to the minus 2 B.

All right. So how should we count this? Initially, we put the B on the outside. But to count this quantity, it's a lot easier to pick a1 and a2 first, because then after we fix a1 and a2, we're counting how many b's we can put there. And this hypothesis is just a bound on the number of b's that go there. So this is less than A squared. Pick any a1 and a2. And then if a1 and a2 obey this condition, there are at most this many choices for B. And that's it. Now I should convert this sum to an average by dividing by B. Then what's left is epsilon K to the minus 2 A squared.

So we are now almost done proving this lemma about paths of length 2. And I can fit the rest of it in the little space under the statement of the lemma.

So on average, N of b is pretty big my lemma 2. On average, N of b is bigger than this. So proof, A prime is going to be N of n for some b. And on average, this true because of lemma P1. So we just need to be a little careful about how exactly to combine these two things. And the way to do it is to look at the average value of N of b squared minus 1 over 2 epsilon times the number of bad pairs.

So why do we combine things by taking a linear combination? We do that because that's nice. Because of linearity of expectation, I can compute this directly from the two lemmas computing these things individually. So that's why we'd want to add them. And what do we get? Well, just plugging in the bounds from lemma P1 and lemma P2, we get that this is at least 1/2 half K to the minus 2 A squared.

So now I choose b so that N of b squared minus 1 over 2 epsilon BP epsilon of b is greater than or equal to its mean. And I set A prime to be this N of b. All right. And then I claim we can read off the two things we care about. The first thing we care about is that N of b is pretty big. So we can ignore this negative thing. N of b squared is at least this big. And so N of b is at least around k inverse A. That is bullet point one.

And bullet point two is that there are only a small fraction of bad pairs a1, a2. Well, for this thing is positive, so 1 over this one I'll write down. So we get the number of bad pairs is at most-- bring this over here and multiply, 2 epsilon times the number of total pairs. Only a 2 epsilon fraction of the total pairs are bad. That's this statement.

Any questions or comments about this lemma?

AUDIENCE: So at the end, all the A's we choose are the neighbors of one b?

LAWRENCE GUTH: Yeah, that's right. So our set A prime is just the set of neighbors of one b. That's right.

Yeah, maybe at the end, we'll try to look back over the Balog-Szemeredi-Gowers theorem and the Balog-Szemeredi-Gowers theorem, the conclusion is there exists A prime and B prime, so that A prime plus B prime is small. And it's a natural question, like who are A prime plus B prime. How do we find them? This one I actually probably should have name something different. This is not the final A prime in the end of Balog-Szemeredi-Gowers. It's a stepping stone. But we can keep track of it along the way and see, how do we find A prime and B prime.

So here, most of the pairs are good. And we're going to upgrade that in a trivial way. If you fix a1, we could ask whether most of the a2 are good. And that's not necessarily going to work. There could be a few bad a1's where many a2's are bad, but we just cross them out and make A prime a bit smaller, and then we could say for every a1, most of the a2's are good. So let me write that down.

So I think this is the best thing that I know how to say about paths of length 2.

OK, so let me call this lemma 2 for paths of length 2. So if x is in A cross B, the size of X is at least K inverse A B and epsilon is bigger than 0. Then there exists A2 in A so that it's pretty big. Not quite as big as before, but I don't know, 1/4 K inverse A. And for every a and a2, a fraction of 1 minus-- it won't be 1 minus 2 epsilon, but maybe 1 minus 10 epsilon times A2 Of a2.

Let me say it this way. At most, 10 epsilon times the size of A2 little a2 and A2, there are at most there, so that the pair little a1, a2, is epsilon bad. OK, so proof sketch is just use the previous lemma. And A2 is a prime. Take away the A's that have too many bad neighbors.

Yeah, actually, let me write this in a better way. And then I can draw the picture of it. So for every in A our set, we can break up the set as a bad part with a good part. And what's the feature of each of these, the bad part is small. And what's good about the good part?

The good part, there are many paths from a to a2 in the good part. Yeah?

AUDIENCE: a comma a2 is epsilon bad? or it's a1?

LAWRENCE GUTH: This a comma a2, correct.

So this is the final word in today's story about what we can say about paths of length 2. And now we're going to bootstrap that and say something even better about paths of length 3, which is the main one.

Now we can prove the key lemma. OK there's a little bit at the beginning. There is one example that was mentioned earlier that there might be a few elements in A that are in no edges or in very few edges, and it's helpful to just eliminate them from the beginning. It's easy to reduce to the case that doesn't happen. So step one, so A1 is defined to be the set of and a and A so that the number of edges from A to B, let me see how I wrote that.

Oh Yeah. That's just the size of the neighborhood of A. The size of the neighborhood of A is at least a small fraction of the typical value. This is the average value of the size of the neighborhood. We get rid of the guys who have way below average number of neighbors. And then it's straightforward to check that. So if I write X of A prime, B prime that's defined to be the set of pairs a, b and A prime cross B prime and in X.

So if I take the edges X from A1 to B, that is still most of the edges. And also since I removed-- A1 comma B-- since I removed low density edges, now the typical density has gone up. So this is still true. So basically, without loss of generality, I can assume that all the edges have a decent size.

All right. So next I use Lemma 2. And it tells me that I have this set A2 in A that has this good property, and A prime is just going to be this A2. All right. So here's my A2. This is A2. And maybe there's some more of A up here. And if I pick a particular a in here, then I can break this up into a small bad part and a big good part. The bad part is small. And for the good part, there are lots of edges to the good part, which I'll illustrate like this.

So now I would like to choose a B. I'd like to choose B prime in B. I'm make some more vertices here. So over here is B. Who should I include in B prime? And I want to know that if I start in any A, I have many paths of length 3 to everybody in B prime. Now, one thing that's a bit tricky is that as I vary A, G of A and B of A aren't going to change, and I don't know much about them. But it seems like a good idea to take vertices B that have a lot of edges to A2.

And if vertex B has no edges to A2, this is not going to be helpful. So that's how we're going to define B prime. B prime is the set of b in B with the property that the number of edges from B to A prime is at least 20 epsilon times the size of A prime. All right.

Now, let's compare that with what we know about here. We know that the bad set of a is at most 10 epsilon times the size of A prime. So this was selected to be bigger than that. So now I can say every A and A prime and B and B prime, the number of paths of length 3 from A to B is at least-- OK, I start in A. Then for every good a, I can have a lot of paths to that good a. So it's at least that was epsilon K to the minus 2 B. That's how many paths I have to a good a times the size of N of b intersected with the grid a's, the guys that were good for a.

So if I look at the guys that are good for a, they neighbor b, so here's b. This is good for a, and they neighbor b. Actually, let me do this. So this is b. These guys here are good for a and they neighbor b. And there are pretty many paths to each guy that's good. And then from here, I can get to b.

And then this is bigger than epsilon K to the minus 2 B times-- well, there are at least 20 epsilon A prime guys and N of b intersected A. Let's do this. N of b intersected A prime minus the number of bad guys. And then this whole thing is at least 10 epsilon times the size of A.

So the key point is that I'm looking at b's that have a lot of edges to A prime, many more than the size of the bad set. So there's still lots of edges that go to the good set. So I put this together. This is like epsilon squared K to the minus to d A. So A prime is also about the size of A. A prime is greater than k inverse A. We have something like this.

We're not done yet because we haven't proven that B prime, we haven't proven that it's non-empty. We need to prove that it's pretty big. But for each a in A prime and b and B prime we do have a lot of edges. We do have the right number of edges. We also have to choose epsilon. But as long as I choose epsilon to be 1 over a power of k, that will fit the bill of what we're trying to prove.

So the last thing is to check that B prime is pretty big, or a little more carefully to check that there's still a lot of edges between and A prime B prime.

OK. So last step, we have to check that the number of edges from A prime to B prime is at least K to the minus O of 1 times the original number of edges. Or you could put times A B, but those are the same. All right.

So remember we proved earlier that A prime is at least around 1 over K times A. And we also know, so why did we do this pruning at the beginning? A prime is contained in A1. We did this initial pruning. So Na is greater than 1/10 M k inverse B for every a in A prime, which is in A. So they're all pretty big. So there are lots of edges that leave A prime.

So the number of edges between A prime and all of B, that's at least K to the minus 2 times AB. Now, to make B prime, B prime are the guys that have pretty many edges to A prime. So we're going to cross out the guys that have very few edges to A prime. So if I count edges from A prime to B, take away B prime, so each one of these has only 20 epsilon edges to A prime. So that's at most sorry 20 epsilon A prime edges times B.

So that's less than 20 epsilon A B. OK, now we have paused about when to choose epsilon. We need it to be at least some small power, some K to the minus O of 1. But we can make it small enough that this is way bigger than this. So we now choose epsilon to be, say, 1 over a million times K to the minus 2. And what that tells us is that the number of edges from A prime to the complement of B prime is way smaller than the number of edges from A prime to B.

So the conclusion is that the number of edges from A prime to B prime is about the same as the number of edges from A prime to B, which is at least K to the minus O of 1. So we have checked the last thing. We still have a lot of edges.

OK. Do people have any last questions or comments about this proof? OK. I have some last picture questions or comments about this proof. So this proof is not that hard and not that long, but it's clever. And I think it came as a surprise to people. And it certainly is surprising to me. So first of all, we have this problem about additive combinatorics. And there are different tools like Fourier analysis is very well suited to additive combinatorics. So you can think of it as something about convolutions and Fourier analysis doesn't seem to help much to prove this thing.

And then we have this graph. This graph, you think of it as an adjacency matrix. And if you want to count paths from a1 to a2 you've taken, so the adjacency matrix is A. We took A, A star. If you want to take paths of length 3, It's A, A star, A. So this is a theorem about powers of a matrix. So you might think that linear algebra would be helpful to prove this theorem. I thought about that. I could totally be missing something, but as far as I can see, linear algebra is not helpful to prove this theorem about the powers of a matrix.

And on the other hand, this elementary double counting, clever, but elementary double counting and graph theory really makes it work. So that's something that I find rather surprising.

I mentioned last time that this theorem is really used a lot. It's worth knowing about. It's important. And so hopefully, you've seen a little bit of that. But as the class goes on, actually, we might have some more time to comment on how this Balog-Szemeredi-Gowers theorem. When I first saw it, it seemed a little bit technical. There's two at first, rather similar sounding ways a set of having additive structure. And they're related to each other. But this is an extremely useful and important thing. And we talked about it, but we'll keep talking about it.