YUFEI ZHAO: For the past few lectures, we've been discussing the structure of set addition, and which culminated in the proof of Freiman's theorem. So this was a pretty big and central result in additive combinatorics, which gives you a complete characterization of sets with small doubling. Today, I want to look at a somewhat different issue also related to sets of small doubling, but this time we want to have a somewhat different characterization of what does it mean for a set to have lots of additive structure.

So in today's lecture, we're always going to be working in an Abelian group. Let me define the following quantity. Given sets A and B, we define the additive energy between A and B to be denoted by E of A and B. So A and B are subgroups. They're subsets of this arbitrary Abelian group. So E of A and B is defined to be the number of quadruples, a1, a2, b1, b2, where a1, a2 are elements of A, and b1, b2 are elements of B, such that a1 plus b1 equals to a2 plus b2.

So the additive energy is the number of quadruples of these elements where you have this additive relation. And we would like to understand sets with large additive energy. So, intuitively, if you have lots of solutions to this equation in your sets, then the sets themselves should have lots of internal additive structure. So it's a different way of describing additive structure, and we'd like to understand how does this way of describing additive structure relate to things we've seen before, namely small doubling.

When you have not two sets but just one set-- slightly easier to think about-- we just write E of A. I mean E of A comma A. And these objects are analogous to 4 cycles in graph theory. Because if you about this expression here in a Cayley graph, let's say over F2, then this is the description of a 4 cycle. You go around 4 steps, and you come back to where you started from. So these objects are the analogs of 4 cycles. And we already saw in our discussion of quasi-randomness, and also elsewhere, that 4 cycles play an important role in graph theory. And, likewise, these additive energies are going to play an important role in describing sets with additive structure.

Consider the following quantity. We're going to let r sub A comma B of x to be the number of ways to write x as a plus b. So x equals to a plus b. So r sub A comma B of x is the number of ways I can write x as a plus b, where a comes from big A, little b comes from big B. Then, reinterpreting the formula up there, we see that the additive energy between two sets A and B is simply the sum of the squares of A-- r sub A comma B. As x ranges over all elements of the

group, we only need to take x in the sumset A plus B.

So the basic question, like when we discussed additive combinatorics, in the sense of when we discussed sets of small doubling, there we asked, if you have a set A of a certain size, how big can a plus a be? Here, let's ask the same. If I give you set A of a certain size, how big or how small can the additive energy of the set be? What's the most number of possible number of additive quadruples. What's the least possible number of additive quadruples?

There's some trivial bounds, just like in the case of sumsets. So what are some trivial bounds? On one hand, by taking a1 equal to a2, and b2 equal to b2, we see that the energy is always at least the square of the size of A. On the other hand, if I fix three of the four elements, then the fourth element is determined. So the upper bound is cube of the size of A.

And you convince yourself that, except up to maybe a constant factors, this is the best possible general upper and lower bound. Similar situation with sumsets, where you have lower bound linear, upper bound quadratic. Which is the side with additive structure? So if you have lots of additive structure, you have high energy. So this range is when you have lots of additive structure. And we would like to understand, what can you say about a set with high additive energy?

Well, what are some examples of sets with high additive energy? It turns out that if you have a set that has small doubling, then, automatically, it implies large additive energy. So, in particular, intervals, or GAPs, or a large subset of GAPs, or all these examples that we saw-- in fact, these are all the examples coming from Freiman's theorem. Also, arbitrary groups. You can have subgroups. And so all of these examples have large additive energy. So let me-- I'll you the proof just in a second. It's not hard.

But the real question is, what about the converse? So can you see much in the reverse direction? But, first, let me show you this claim that small doubling implies large additive energy. Well, if you have small doubling, if a plus A is size, at most, k times the size of A, then it turns out the additive energy of A is at least the maximum possible, which is A cubed divided by k. So that's within a constant factor of the maximum. It's pretty large. If you have small doubling, then large additive energy.

So let's see the proof. So you can often tell how hard a proof is by how simple the statement is, although that's not always the case, as we've seen with some of our theorems, like Plunnecke's inequality. But in this case, it turns out to be fairly simple. So we see that r sub A

comma A is supported on A plus A. So we use Cauchy-Schwarz to write-- so, first, we write additive energy in terms of the sum of the squares of these r's. And now, by Cauchy-Schwarz, we find that you can replace the sum of the squared r's by the sum of the r's. But now the key point here is that we take out this factor coming from Cauchy-Schwarz, which is only A plus A. So if the support size is small, we gain in this step.

But what is the sum of r's? I mean, r of x is just number of ways to write x as little a1 plus little ab-- little a2. So if I sum over all x, I'm just looking at different two ways-- we're just looking at ways of picking an ordered pair from A. So this last expression is equal to the size of A to power 4 divided by A plus A. And now we use that A has small doubling to conclude that the final quantity is at least A cubed divided by k. So we see small doubling implies large additive energy. And this kind of makes sense. If your set doesn't expand, then there are many collisions of sums. And so you must have lots of solutions to that equation up there.

But what about the converse? If I give you a set with large additive energy, must it necessarily have small doubling? Oh. Let me show you an example. So, well-- so a large additive energy, does it imply small doubling? So consider the following example, where you take a set A which is a combination, is a union of a set with small doubling plus a bunch of elements without additive structure.

So I take a set with small doubling plus a bunch of elements without additive structure. Then it has large additive energy, just coming from this interval itself. So the energy of A is order N cubed. N is the number of elements. What about A plus A? Well, for A plus A, this part doesn't-- that's the part that contributes, or the part of this A without additive structure. And we see that the size of A plus A is quadratic in the size of A.

So, unfortunately, the converse fails. So you can have sets that have large additive energy and also large doubling. But, you see, the reason why this has large additive energy is because there is a very highly structured additively structured piece of it. And, somehow, we want to forget about this extra garbage. And that's part of the reason why the converse is not true.

So we would like a statement that says that if you have large additive energy, then it must come from some highly structured piece that has small doubling. And that is true, and that's the content of the Balog-Szemeredi-Gowers theorem, which is the main topic today. So the Balog-Szemeredi-Gowers theorem says that if you have a set-- so we're working always in

some arbitrary Abelian group. If you have a set with large energy, then there exists some subset A prime of A such that A prime is a fairly large proportion of A. And here, by large I mean up to polynomial changes in the error parameters.

So this A prime is such that A prime has small doubling. If you have large additive energy, then I can pick out a large piece with small doubling constant, and I only loose a polynomial in the error factors. So that's the Balog-Szemeredi-Gowers theorem, and it describes this example up here. Any questions about the statement?

So what I will actually show you is a slight variant, actually a more general statement, where, instead of having one set, we're going to have two sets. So here's Balog-Szemeredi-Gowers theorem version 2, where now we have two sets. Again, A and B are-- I'm not going to write any-- I'm not going to write it in this lecture, but A and B are always subsets of some arbitrary Abelian group. So A and B both have size of, at most, n, and the energy between A and B is large. Then there exists a subset A prime of A, B prime of B such that both A prime and B prime are large fractions of their parent set, and such that A prime plus B prime is not too much bigger than n.

It's not so obvious why the second version implies the first version. So you can say, well, take A and B to be the same. But then the conclusion gives you possibly two different subsets, A prime and B prime. But the first version, I only want one subset that has small doubling. So, fortunately, the second version does imply the first version. So let's see why.

The second version implies the first version because, if we-- so there's a tool that we introduced early on when we discussed Freiman's theorem, and this is the Ruzsa triangle inequality. So the spirit of Ruzsa triangle inequality is it allows you to relate, to sort of go back and forth between different sumsets in different sets. So by Ruzsa triangle inequality, if we apply the second version with A equals to B, then-- and we pick out this A prime and B prime, then we see that A prime plus A prime is, at most, A prime plus B prime squared over B prime.

Well, actually, this uses the-- vice versa it uses a slightly stronger version that we had to use Plunnecke-Ruzsa key lemma to prove. But you can come up-- I mean, if you don't care about the precise loss in the polynomial factors, you can also use the basic Ruzsa triangle inequality to deduce a similar statement. This is easier to deduce. So you have that.

And now, the second version tells you that the numerator is, at most, poly kn, and the denominator is, at most-- at least, n divided by poly k. Remember, over here, to get this

hypothesis, we automatically have that the size of A and B are not two much smaller than n. Or else this cannot be true. So putting all these estimates together, we get that. So these two versions, they are equivalent to each other. Second version implies the first. The second one is stronger. The first one is slightly more useful. They're not necessarily equivalent, but the second one is stronger. Any questions? All right.

So this is a Balog-Szemeredi-Gowers theorem. So the content of today's lecture is to show you how to prove this theorem. A remark about the naming of this theorem. So you might notice that these three letters do not coming in alphabetical order. And the reason is that this theorem was initially approved by Balog and Szemeredi, but using a more involved method that didn't give polynomial high bounds. And Gowers, in his proof of Szemeredi's theorem, his new proof of Szemeredi's theorem with good bounds, he required-- well, he looked into this theorem and gave a new proof that resulted in this polynomial type bounds. And it is that idea that we're going to see today.

So this course is called graph theory and additive combinatorics. And the last two topics of this course-- today being Balog-Szemeredi-Gowers, and tomorrow we're going to see sum-product problem-- are both great examples of problems in additive combinatorics where tools from graph theory play an important role in their solutions. So it's a nice combination of the subject where we see both topics at the same time.

So I want to show you the proof of Balog-Szemeredi-Gowers, and the proof goes via a graph analog. So I'm going to state for you a graphical version of the Balog-Szemeredi-Gowers theorem. And it goes like this. If G is a bipartite graph between vertex sets A and B-- and here A and B are still subsets of the Abelian group-- we define this restricted sumset, A plus sub G of B, to be the set of sums where I'm only taking sums across edges in g.

So, in particular, if G is the complete bipartite graph, then this is the usual sumset. But now I may allow G to be a subset of the complete bipartite graph. So only taking some but not all of the-- only taking-- yes, some of this sums but not all of them.

The graphical version of Balog-Szemeredi-Gowers says that if you have A and B be subsets of an Abelian group, both having size, at most, n, and G is a bipartite graph between A and B, such that G has lots of edges, has at least n squared over k edges. If the restricted sumset between A and B is small-- So here we're not looking at all the sums but a large fraction of the possible pairwise sums. If that sumset has small size, this is kind of like a restricted doubling

constant. Then there exists A prime, subset of A, B prime, subset of B, with A prime and B prime both fairly large fractions of their parent set, and such that the unrestricted sumset between A prime and B prime is not too large.

So let me say it again. So we have a fairly dense-- so a constant fraction edge density, a fairly dense bipartite graph between A and B. A and B are subsets of the Abelian group. Then-- and such that the restricted sumset is small. Then I can restrict A and B to subsets, fairly large subsets, so that the complete sumset between the subsets A prime and B prime is small. Let me show you why the graphical version of BSG implies the version of BSG I stated up there.

But, so why do we care about this graphical version? Well, suppose we-- so we have all of these hypotheses. Let's write-- so we have all of those hypotheses up there. So let's write r to be r sub A comma B, so I don't have to carry the subscripts all around. What do you think-- so I start with A and B up there, and I need to construct that graph G. So what should we choose as our graph?

Let's consider the popular sums. So the popular sums are going to be elements in the complete sumset such that it is represented as a sum in many different ways. And we're going to take edges that correspond to these popular sums. So let's consider bipartite graph G such that A comma B is an edge if and only A plus B is a popular sum.

So let's verify some of the hypotheses. So we're going to assume graph BSG, and let's verify the hypothesis in graph BSG. On one hand, because each element of S is a popular sum, if we consider its multiplicity, we find that the size of S multiplied by n over 2k, lower bound be size of A times the size of B. So if you think about all the different pairs in A and B, each sum here, each popular sum, contributes this many times to this A cross B.

So, as a result, because size of A and size of B are both, at most, n, we find that the size of S is, at most, 2kn. And if you think about what G is, then this implies also that the restricted sumset of A and B across this graph G-- which only requires the popular sums. So the restricted sumset is precisely the popular sums. So restricted sumset is not too large. OK, good. So we got one of the conditions, that the restricted sumset is not too large.

And now we want to show that this graph has lots of edges. It has lots of edges. And here's where we would need to use the hypothesis that, between A and B, originally there is large additive energy. And the point here is that these unpopular sums cannot contribute very much to the additive energy in total, because each one of them is unpopular. So the dominant

contributions to the additive energy are going to come from the popular sums, and we're going to use that to show that G has lots of edges.

So let's lower bound the number of edges of G by first showing that-- so we'll show that the unpopular sums contribute very little to the additive energy between A and B. Indeed, the sums of the squares of the r's, if for x not in popular sums, is upper bounded by-- well, claim that it is upper bounded by the following quantity, that n over 2k times n squared. Because I can take out one factor r, upper bound by this number, just by definition, and the sums of the r's is n squared.

So you have this additive energy between A and B. I know that it is large by hypothesis. Whereas, I also know that I can write it as a sum of the squares of the r's, which I can break into the popular contributions and the unpopular contributions. And, hopefully, this should all be somewhat reminiscent of basically all these proofs that we did so far in this course, where we separate a sum into the dominant terms and the minor terms. This came up in Fourier analysis in particular. So we do this splitting, and we upper bound the unpopular contributions by the estimate from just now.

So, as a result, bringing this small error term, it doesn't cancel much of the energy. So we still have a lower bound on the sum of the squares of the r's in the popular sums. But I can also give a fairly trivial upper bound to a single r, namely it cannot be bigger than n. And so the number of edges of G-- so what's the number of edges of G? Look at that. Each x here contributes rx many edges. So the number of edges of G is simply the sums of these rx's. Which is quite large.

So the hypothesis of graph BSG are satisfied. And so we can use the conclusion of graph BSG, which is the conclusion that we're looking for in BSG. Any questions? Good. So the remaining task is to prove the graphical version of BSG. So let's take a quick break, and when we come back we'll focus on this theorem, and it has some nice graph theoretic arguments.

OK, let's continue. We've reduced the proof of the Balog-Szemeredi-Gowers theorem to the following graphical result. Well, it's not just graphical, right? Still-- we're still inside some an Abelian group, still looking at some set in some Abelian group, but, certainly, now it has a graph attached to it. Let me show this theorem through several steps. First, something called a path of length 2 lemma.

So the path of length 2 lemma, the statement is that you start with a graph G which is a

bipartite graph between vertex sets A and B. And now A and B no longer need-- they're just sets. They're just vertex sets. We're not going to have sums. And the number of edges is at least a constant fraction of the maximum possible. Then the conclusion is that there exists some U, a subset of A, such that U is fairly large. And between most pairs of elements of U-- so between 1 minus epsilon fraction of pairs of U-- there are lots of common neighbors. So at least epsilon delta squared B over 2 common neighbors.

So you start with this bipartite graph A and B. Lots of edges. And we would like to show that there exists a pretty large subset U such that between most pairs-- all but an epsilon fraction-- of ordered pairs-- they could be the same, but it doesn't really matter-- the number of paths of length 2 between these two vertices is quite large. So they have lots of common neighbors. Where have we seen something like this before? There's a question?

AUDIENCE: Is there a [INAUDIBLE] epsilon?

YUFEI ZHAO: Ah, yes. So for every epsilon and every delta. So let epsilon, delta be parameters. Where have we seen something like this before? So in a bipartite graph with lots of edges, I want to find a large subset of one of the parts so that every pair of elements, or almost every pair of elements, have lots of common neighbors. Yes.

AUDIENCE: [INAUDIBLE].

YUFEI ZHAO: Dependent random choice. So in the very first chapter of this course, when we did extremal graph theory forbidding bipartite subgraphs, there was a technique for proving the extremal number, upper bounds, for bipartite graphs of bounded degree. And there we used something called dependent random choice that had a conclusion that was very similar flavor. So there, we had every pair-- so a fairly large, but not as large as this-- a fairly large subset where every pair of elements had lots of common neighbors. For every couple, every k couple of vertices, have lots of common neighbors. So it's very similar. In fact, it's the same type of technique that we'll use to prove this lemma over here. So who remembers how dependent random choice goes?

So the idea is that we are going to choose U not uniformly at random. So that's not going to work. Going to choose it in a dependent random way. So I want elements of U to have lots of common neighbors, typically. So one way to guarantee this is to choose U to be a neighborhood from the right. So pick a random element in B and choose U to be its

neighborhood. So let's do that. So we're going to use dependent random choice. See, everything in the course comes together.

So let's pick v an element of B uniformly at random. And let U be the neighborhood v. So, first of all, by linearity of expectations, the size of U is at least delta of A. So because the average degree from the right from B is at least delta A just based on the number of edges. If you have two vertices a and a prime in A with a small number of common neighbors, then the size of-- so sorry. Let me-- I skipped ahead a bit.

So if a and a prime have a small number of common neighbors, then the probability that a and a prime both lie in U should be quite small. Because if they both had-- if a and a prime have a small number of common neighbors, in order for a and a prime to be included in this U, you must have chosen-- so suppose this were their common neighbor. Then in order that a and a prime be contained in U, it must have chosen this v to be inside the common neighborhood of a and a prime. Which is unlikely if a and a prime had a small number of common neighbors. So this probability is, at most, epsilon delta squared over 2. Just think about how U is constructed.

So if we let x be the number of a and a primes in U cross U with, at most, epsilon delta squared over 2 times B common neighbors, then, by linearity of expectations, the expectation of x is-- well, by summing up all of these probabilities of a and a prime, both being in U-- so this is, at most, epsilon delta squared over 2 times size of A squared. So, typically, at least in expectation, you do not expect very many pairs of elements in U with few common neighbors.

But we can also turn such an estimate into a specific instance. And the way to do this is to consider the quantity size of U squared minus x over epsilon. Well, first of all, we can lower bound this quantity, because the size of second moment of U is at least the first moment of U squared. And we also know that the size of x in expectation is not very large. So the whole expression can be lower bounded by delta squared over 2 times the size of A squared. So this is epsilon, sorry.

Therefore, there is some concrete instance of this randomness resulting in some specific U such that this inequality holds. So there exists some U such that this inequality holds. And, in particular, we find that the size of U is at least-- just forget about this minus term-- is at least that right-hand side, square root. So, in particular, the size of U is at least delta over 2 times the size of A. And, just looking at the left-hand side, which must be a non-negative quantity

because the right-hand side is non-negative, we find that x is, at most, an epsilon fraction of U squared.

So putting these together, we arrive at the path of length 2 lemma. So let me go through it again. So this is the dependent random choice method, where we're going to-- we want to find this U, where most pairs of vertices in U have lots of common neighbors. So we start from the right side. We start from B, pick a uniform random vertex, which you call v, and let U be the neighborhood of v. And I claim that this U, typically, should have the desired property.

And the reason is that, if you have a pair of vertices on the left that do not have many common neighbors, then I claim it is highly unlikely that these two vertices both appear in U. Because for them to both appear in U, your v have been selected inside the common neighborhood of a and a prime, which is unlikely if a and a prime have few common neighbors. So, as a result, the expected number of pairs in U with small number of common neighbors is small.

And, already, that's a very good indication that we're on the right track. And, to finish things off, we look at this expression, which we can lower bound by convexity. And we know the size of U in expectation is large. And, also, the size of x, that we just saw, is small in expectation. So you have this inequality over here.

And because there's an expectation, it implies that there's some specific instance such that, without the expectation, the inequality holds. So take that specific instance. We obtain some U such that this inequality is true, which simultaneously implies that U is large and x, the number of bad pairs, is small. So that was dependent random choice. Any questions? All right.

So that was the path of length 2 lemma. So it tells us I can take a large set with lots of paths of length 2 between most pairs of vertices. Let's upgrade this lemma to a path of length 3 lemma. So, in the path of length 3 lemma, we start with a bipartite graph, as before, between A and B. So G is a bipartite between A and B. And, as before, we have a lot of edges between A and B. It's the delta fraction of all possible edges. Then the conclusion is that there exists A prime in A and B prime subset of B such that A prime and B prime are both large fractions of their parent set.

And now, the-- and, furthermore, every pair between A prime and B prime is joined by many paths of length 3. So a path of length 3 means there's 3 edges. And, here, this eta is basically the original error term up to a polynomial change. So starting with this bipartite graph that's fairly dense, the lemma tells us that we can find some large A prime and large B prime so that

between every vertex in A prime and every vertex in B prime, there are lots of paths of length 3 between them. Every time.

So we should think about all of these constants as-- plus you only make polynomial changes in the constants, we're happy. Here, eta is a polynomial change in the delta. There's a convention which I like which is not universal, but it's often solved, unlike this convention. It's the difference between the little c and the big C is that a little c is better if you make it smaller, and a big C is better-- I mean, it's better in the sense that if this is true for little c and big C, and you make little c smaller and big C bigger, then it is still true. So big C is a sufficiently large constant, and little c is a sufficiently small constant. Just a--

So let's see the path of length 3 lemma, see it's proof. We're going to use the path of length 2 lemma, but we need a bit of preparation first. So the proof has some nice ideas, but it's also-- some parts of it are slightly tedious, so bear with me. So we're going to construct a chain of subsets A-- inside A. So A1, A2, A3. And this is just because there's a few cleaning up steps that need to be done.

Let's call two vertices in A friendly if they have lots of common neighbors. And, precisely, we're going to say they're friendly if they have more than delta squared over 80 times the size of B common neighbors. Let me construct this sequence of subsets as follows. First, let A1 be all the vertices in A with degree not too small. So this is in preparation. So it will make our life quite a bit easier later on. Let's just trim all the really small degree vertices so that we don't have to think about them.

So you trim all the small degree vertices. And think about how many edges you trim. You cannot trim so many edges, because each time you trim such a vertex, you only get rid of a small number of edges. So, in the end, at least half of the original set of edges must remain. And, as a result, the size of A1 is at least a delta over 2 fraction of the original vertex set. Otherwise, you could not have contained half of the original set of edges. So this is the first trimming step. So we got rid of some edges, but we got rid of fewer than half of the original edges. And because now you have a minimum degree on A1, the number of edges between A1 and B is quite large, still quite large. So think about passing down to A1 now.

In the second step, we are going to apply the path of length 2 lemma to this A1. So A2 is going to be constructed from-- so using the path of length 2 lemma, specifically with parameter epsilon being delta over 10. Although, remember, now the density of the graph went from

delta to delta over 2. Again, if you don't care about the specific numbers, they're all polynomials in delta. So don't worry about them. Everything's poly delta.

So we're going to apply the path of length 2 lemma to find this subset A2. And it has the property that A2 is quite large, and all but a small fraction of pairs in A2 are friendly. So we passed down to, first, trimming small degree vertices, and then passed down further to A2, where all but a small fraction of elements in A2, or all but a small fraction of the pairs are friendly to each other, meaning they have lots of common neighbors.

And now let's look at the other side. Let's look at B. So we're in this situation now where you have-- so we're now in a situation where you've passed down to A2 and in B, where, because of what we did initially, every vertex in here have large degree. So there's this minimum degree condition from every vertex on the left. So the average degree is still very high. As a result, the average degree from B is going to be quite high. So let's focus on the B side and pick out vertices in B that have high degree.

So let's B1 denote vertices in B such that the degree from B to A2 is at least half of what you expect based on average degree. And, as before, the same logic as the A1 step. We see that B1 has large size, is a large fraction of B. And now we pass down to this B1 set.

Now, finally, let's consider A3 to be vertices in A2 where a is friendly. So vertices a in A2 such that a is friendly to at least 1 over delta over-- so 1 minus delta over 5 fraction of A2. So we saw that, in A2, most pairs of vertices are friendly. So most, meaning all but a delta over 10 fraction.

So if we consider vertices which are unfriendly to many other vertices in A2, there aren't so many of them. If there were many of them, you couldn't have had that. So that's why I constructed this set A3 consisting of elements in A2 that are friendly to many elements. And the size of A3 is at least half of that of A2. So we have this A3 inside. All right.

And now I claim that we can take A3 and B as our final sets, and that between every vertex in A3 and every vertex in B1, I claim there must be lots of paths of length 3. But, first, let's check their sizes. I mean, the sizes all should be OK, because we never lost too much at each step. If you only care about polynomial factors, well, you already see that we never lost anything more than a polynomial factor. But just to be precise, the size of A3 is at least-- so if you count up the factor lost at each step, so it's 1/2 delta over 4 delta over 2. So it's at least delta squared over 16 fraction of the original set A.

And now, if we consider a comma b to be an arbitrary pair in A3 cross B1, I claim that there must be many paths. Because by using-- so what properties do we know? We know that b is adjacent to a large fraction. So here large means at least delta over 4-- so bounded below-- a large fraction of A2. Yes. So I apologize. When I say the word large, depending on context it can mean bigger than delta, or it could mean at least 1 minus delta. So you look at what I write down. So b is adjacent to at least delta over 4 fraction of A2.

At the same time, we know that a is friendly to at least 1 minus delta over 5 fraction of A2. So these two sets, they must overlap by at least a delta over 20 fraction. So let's take a vertex b. So you-- so it's adjacent to many vertices here. And if you look at a vertex in A, it's friendly to a large fraction. So, in particular, it's friendly to all these elements over here.

So, to finish off, what does it mean for a-- this is-- this vertex is a. This vertex is b. What does it mean for a to be friendly to all of these shaded elements? It means that there are lots of paths from a to each of these elements. And then you can finish off the paths going back to b. Yes.

AUDIENCE: The shaded stuff is allowed to be outside of A3?

YUFEI ZHAO: No. the shaded-- the question is, is the shaded stuff allowed to be outside of A3? No. The shaded things are inside A3. So we're looking at intersections within A3. No, sorry. Actually, no, you're right. So the shaded things can be outside A3. So shaded things can be outside A3. I apologize. So everything now is in A2.

So b is adjacent to a large fraction of A2. And a here is friendly to some part of the neighbors of b. So you can complete paths like that. Yes. So only the starting and ending points have to be in A prime and B prime. Everything else, they can go outside of the A prime and B prime. Yes, thank you.

So the number of paths from a to B to A2 back to b is-- let's see if I can stay within B1-- so is at least-- yes. So it's-- sorry. This is B. So it's at least delta over 20 times A2 times delta over delta squared over 80 times B. So if you don't care about polynomial factors in delta, then you see that-- the point is there's a large fraction of-- there are a lot of paths. So there are a lot of paths between each little a and each little b by the construction we've done.

So let me just do a recap. So there were quite a few details in this proof, and some of them have to do with cleaning up. Because it's not so nice to work with graphs that just have large

average degree. It's much nicer to work with graphs with large minimum degree. So there are a couple of steps here to take care of vertices with small degrees. So we started with, between A and B, lots of edges. And we trim vertices from A with small degree. So we get A1.

And then we apply the path of length 2 lemma to get A2. So inside A2, most pairs of vertices have lots of common neighbors, but not all. We then go back to B to get B1, which has large minimum degree to A2. And then A3 looks at vertices in A with many friendly companions in A2. And A3 is large, and I claim that between every vertex in A3 and every vertex in B, you have many paths of length 3. Because if you start with a vertex in A3, it has many friendly companions.

So many here means at least 1 minus delta over 5 fraction. Whereas every vertex in B1 has lots of neighbors in A2, where lots means at least delta over 4. So there's necessarily an overlap of at least delta over 20. And for that overlap, we can create lots of paths going through this overlap from A to B. Any questions? OK, great.

So let's put everything together to prove the graphical version of Balog-Szemeredi-Gowers. So we'll prove the graphical version of Balog-Szemeredi-Gowers. So by-- so, first, note that the hypothesis of Balog-Szemeredi-Gowers already implies that the size of A and the size of B are not too small. Because, otherwise, you couldn't have had n squared over k edges to begin with.

So by the path of length 3 lemma, there exists A prime in A and B prime in B with the following properties. That A prime has a large fraction of-- so A prime and B prime are both large in size. And for all vertices a in A prime and vertices b in B prime, there are lots of paths of length 3 between these vertices. So there are at least k to the minus little o1-- to the minus big O1 times n squared pairs of intermediate vertices a1, b1 in A cross B, such that a b1 a1 b is a path in G.

So let me draw the situation for you. So we have A and B. And so inside A and B, we have this fairly large A prime and B prime, such that for every little a and little b, there are many paths like that going to b1 and a2. Let me set-- so let me set x to be a plus b1, that sum, y to be a1 plus b1, and z to be a1 plus b.

So now notice that we can write this a plus b in at least k to the minus big O1 times n squared ways as x minus y plus z by following this path, where x, y, and z all lie in the restricted sumset, because that's how the restricted sumset is defined. So if you have an edge, then the

sum of the elements across on the two ends, by definition, lies in the restricted sumset.

So the path of length 3 lemma tells us that every pair a and b, their sum can be written in many different ways as this combination. As a result, we see that A prime plus B prime-- so this sum, if we consider sum along with its multiplicity-- so now we're really looking at all the different sums as well as ways of writing the sum as this combination-- we see that it is bounded above by the restricted sumset raised to the third power. Because each of these choices, x, y, and z, they come from the restricted sumset.

But the hypothesis of Balog-Szemeredi-Gowers, the graphical version, is that the restricted sumset is small in size. So we can now upper bound the restricted sumset by, basically, the-- within a constant, within a factor of the maximum possible. And now we are done, because we have deduced that the complete sumset between A prime and B prime is, at most, a constant factor with change in constant by a polynomial. So a constant factor more than the maximum possible. So it's, at mostly, k to the big O1 poly k times n.

So that proves the graphical version of Balog-Szemeredi-Gowers. And because we showed earlier that the graphical version of Balog-Szemeredi-Gowers implies Balog-Szemeredi-Gowers, this shows the Balog-Szemeredi-Gowers theorem. So let me recap some of the ideas we saw today. And so the whole point of Balog-Szemeredi-Gowers and all of these related lemmas and theorems and variations is that you start with something that has a lot of additive structure. Well, after we passed down to graphs just a lot of edges.

So you start with a situation where you have kind of 1% goodness. And you want to show that you can restrict to fairly large subsets, so that you have perfection. So you have complete goodness between these two sets. And this is what's going on in both the graphical version and the additive version. So back to the graph path of length 3 lemma. So we were able to boost the path of length 2 lemma, which tells us something about 99% of the pairs having lots of common neighbors, to 100% of the pairs having lots of path of length 3.

And in the additive setting, we saw that by starting with a situation where the hypothesis is somewhat patchy, so like a 1% type hypothesis, we can pass down to fairly large sets, where the complete sumset, starting with just the restricted sumset being small, can pass down to large sets where the complete sumset is small. And this is an important principle, that, often, when we have some typicality by an appropriate argument-- and, here, it's not at all a trivial argument. So there's some cleverness involved, that by doing some kind of argument, we may

be able to pass down to some fairly large set where it's not typically good, but everything's perfectly good. That's the spirit here of the Balog-Szemeredi-Gowers theorem. So, next time, for the last lecture of this course, I will tell you about the sum-product problem, where the-- there are also some graph-- very nice graph theoretic inputs.