[SQUEAKING]

[RUSTLING]

[CLICKING]

PABLO SHMERKIN: OK, so let's continue discussing the proof of Bourgain's projection theorem and how it's similar and different from the corresponding statement in the finite field setting. So just as a quick review because we are going to keep working with this concept a delta, S, C, d set is a subset of, let's say, the unit ball or maybe the unit cube. It doesn't really matter.

So it has to live in some bounded region of space, such that, if we intersect X with the ball and look at the delta-covering number, this is controlled by, well, first of all, this tolerance constant C, then the radius of the ball to the power S, and then the size of all of X. So it is a spacing condition, and although this is not part of the definition, it is convenient to imagine that the size of X is maybe delta to the minus S, although it doesn't have to be. But it is in the setting of Bourgain's projection theorem, as I stated it.

And speaking of Bourgain's projection theorem as I stated it, let me state the version-- well, not the version, it's really exactly the same theorem I stated last time, but using this language. So Bourgain's projection theorem, again, so it's really exactly the same statement, just using this language and rescaling-- suppose that X is in the unit ball of R2 and is a delta-- oh, sorry, quantifiers. So given t between 0 and 2, S strictly positive, at most, 1, there exist two parameters, epsilon and eta positive, depending on t and S, so that the following happens.

If X is a delta t delta to the minus eta set in R2 with size delta to the minus t, and D the set of directions-- let's say in 0, 1-- is a delta S, delta to the minus eta set, then there exists a direction in our set of directions such that, if you look at the minimum over all subsets of X, which are dense in the sense that the delta-covering number is at least delta to the eta, this tolerance parameter times the delta-covering number of X, and we project the subset, then, no matter how I choose the subset, we get a gain over the, between quotes, "trivial bound," which is less trivial in this setting.

But nevertheless, this is at least delta to the minus t over 2 minus epsilon. So epsilon is how much we gain. And delta 2 to the minus eta is our tolerance.

This is really the same theorem as last time. Last time, X was in the ball of radius R, and it was a unit of unit balls. But we just scaled it down by a scale of 1 over R. And by the equivalences for measuring the delta-covering number that I mentioned last time, this is really the same statement.

So last time, we saw that the following is true, even though we didn't actually prove all of it. Last time, we saw that there exists a polynomial, which is just X1, X2, X3 minus X4, X5, X6. That works. So it exists a polynomial such that, for every S between 0 and 1 on-- OK, I didn't say it this way last time, but maybe let's introduce also the two parameters here. Last time, I stated with just one parameter.

Of course, you can just take epsilon and eta to be the same by taking the minimum. So the whole statement, if you replace epsilon and/or eta by a smaller number, is still true. So we could just take the minimum and have just one parameter. But I sort of want to distinguish them because they play different roles.

So there exists epsilon and eta such that if A is a delta S delta to the minus eta set, it may be contained in 0, 1. And maybe let's assume also that the size of A is delta to the minus S. Then the size grows by applying this polynomial.

And we discussed that this is not enough to iterate, because in order to iterate, we need to know that Q of A set also satisfies these assumptions. And this is not clear at all. And in fact, it's not true. So we need to improve. To Q of A contains A delta S plus epsilon delta to the minus eta set. If we knew this, then we could iterate and keep growing, so keep adding epsilon. And because epsilon is continuous in S, we can get as close to delta to the minus 1 as we want.

So today, we will discuss how to do this. So let me say that it's not going to be the same epsilon and eta. So they're going to-- I mean, they will become much smaller, but same order of quantifiers. So there exists S-- sorry, for every S, there will exist eta and epsilon so that this is true.

So this is what we are going to mostly discuss today and, hopefully, then go briefly over how to finish the proof. But I will say, this is the main difference with respect to the finite field setting, where we don't have to worry about any of these. So before coming back to this, we'll take a detour, and then we'll come back.

So let me state a lemma with some basic properties of delta S, C sets. Suppose that X is a delta S, C set in any dimension. Then we make the following two fairly easy but important observations. So maybe I should have done this last time to get some intuition about what this concept does.

First of all, X is large at all scales. So the rho-covering number of X is at least C to the minus 1 rho to the minus S for every rho between delta and one. I mentioned in passing last time that this is true for rho equals delta. But in fact, it's true for every rho.

And two, this concept is robust under passing to subsets. So if X or maybe Y is a subset of X and the delta-covering number is, let's say, 1 over K times the delta covering number of X, then Y is a delta S C times K set. So the constant C gets worse by the parameter measuring how dense Y is in X, always using delta-covering numbers.

So the proof of this is fairly simple. So let's do it. We have to prove something. Well, we need to show that the rho-covering number is something. So the definition of the rho-covering number is the smallest number of rho balls needed to cover the set. So let's cover the set by rho balls.

And let's see how many balls we need. So let's say we cover the set by. M balls of radius rho. Well, then the delta-covering number of A-- so the delta-covering number is subadditive, something which will come up repeatedly today. And this is just because if we have a delta cover-- sorry, a rho cover-- sorry, a delta cover. We are covering by [INAUDIBLE] delta now.

If you have a delta cover of each of these, we can put them together, and we get a delta cover of this. So for this reason, the delta-covering number is subadditive. So the delta-covering number of A is at most the sum of the delta covering numbers of A intersected with these balls.

But by definition of delta-covering set, each of these is at most C rho to the S times the delta-covering number of A. So we have a uniform upper bound that only depends on rho. So the center doesn't matter. So this is just the definition of A being a delta S, C set.

So this is at most M C rho to the S delta-covering number of A. So we see that we can cancel out the delta-covering number of A, and we get the claim one. So we get that M is at least what is on the right-hand side there. But the smallest possible M is, by definition, the rho-covering number. OK?

And II is even easier. Well, we have to show that Y is a delta D, C, of course, would be S. So we have to Y satisfies the definition of delta-covering number, so of delta, S set. So we intersect with the ball of radius rho.

Well, Y is contained in X. So this is at most the same thing with X instead of rho. But X is a delta S, C set. So this is at most C times rho to the S times the delta-covering number of X.

But the delta-covering number of X is at most K times by the delta covering number of Y, by assumption. So this is at most CK rho to the S delta-covering number of Y. So we have checked that Y is a delta, S, CK set.

So what information have we gained regarding our task? So we need to show that Q of A contains a delta S plus epsilon delta to the minus eta set. Well if it contains such a set, then Q of A must satisfy the conclusion of the first part of the lemma. So the rho-covering number of Q of A has to be large.

We have proved that the delta-covering number is large. It is not sufficient, but it is necessary to show that Q of A grows at every scale. We want to show that it contains a delta t over 2 plus epsilon something set. So it needs to be large at all scales.

So we see that we need Q of A at scale rho to be larger than rho to the minus t over 2 minus epsilon for every rho between delta and 1 And a priori, we only know it for delta. So what we discussed last time is that this is true for delta.

And again, this is not sufficient. Yes?

AUDIENCE: Does that say, rho to the minus S minus epsilon.

PABLO SHMERKIN: Rho to the-- so now-- OK, so this was a general statement about delta S sets. And now I'm going back to what we are trying to achieve. Oh, sorry, S-- yeah, so there is-- sorry. Yeah. So this is still not Bourgain's projection theorem. We are working with just one setting, 0, 1. Thank you.

We need this to be true. And OK, so one we could think that we have it for free because why can't we just apply what we know at scale rho? So this is a question for you. So why-- so again, this is necessary. It's not sufficient. But OK, let's try to focus on getting this first.

So why don't we get it for free? So what is the difference between rho and delta? So that's the assumption that we have over there. Yes?

AUDIENCE: Because we don't know that A is rho, S [INAUDIBLE]

PABLO SHMERKIN: Well, OK. So the assumption is that A is a delta S something set. And you're right. We don't know that A is a rho S something set. There is another thing that could be a bit worrying-- I would say it's less worrying-- which is that delta to the minus epsilon is not rho to the minus epsilon. We would need rho to the minus epsilon. Delta to the minus epsilon is much bigger.

This last thing is not so serious because we can always-- the last thing is not so serious. Neither of them is so serious in some sense. But the second one, maybe we'll come back to discussing later. But it is less serious. The main issue is that-- so we don't get this for free, because A needs not be a delta rho delta to the minus epsilon set.

And we also would need rho here, but let's ignore this for a moment. And my parameters are all wrong. Rho S delta to the minus epsilon set. And also, even if we did, we still have the issue that this is necessary but not sufficient. So this condition is necessary, but not sufficient for Q of A to contain the delta S plus epsilon delta to the minus eta set.

So we are going to overcome these two issues by passing to what is known as uniform sets. So now we make a new detour, and we are going to study uniform sets. And so uniform sets and uniformization come up all the time in studying these type of problems, not only in this proof. So maybe we will come up again later in the course. But it is what we are going to use to overcome these two issues.

So let's start by defining what a uniform set is. So definition-- so maybe we've heard definition. Let's do some notation so we can write the definition. So this is from last time. But recall that I was denoting the delta to be the delta mesh cubes. So we split our D into cubes of side length delta, and one of them has a vertex at the origin.

So let's define D delta a set X to be the cubes. This is in any dimension, so the cubes that are hit by X. So these are the Q cubes of side length delta which hit X.

And let's define a box-counting version of a delta-covering number by looking at this. So let's define x delta star to distinguish it from the delta-covering number to be the cardinality of this set. How many delta cubes the set hits-- and we discussed last time that up to a multiplicative constant, that could depend on the ambient dimension, but we don't care. This is really the same as the delta-covering number.

So in some sense, they are the same. But in order to work with uniform sets, this version is better, as we will see. So we are counting how many of these are hit by the set X. And we can compare cubes with balls. And this is why these two things are comparable, as we discussed last time.

Now we can give a definition of uniform sets. So a set X-- and now our reference-bounded part of space is going to be the unit ball-- not the unit ball, the unit cube. Not the unit ball, the unit cube. So in 0, 1 to the d is delta m uniform. What is delta is a parameter in 0, 1 So it's a number in 0, 1. And what is M? A natural number.

And you should think-- so I will maybe come back to this. But you should think that delta is fixed, but M is not. So we fixed delta and let M go to infinity, something like this. So we're going to fix delta, and then M will grow. But this is just a definition.

If the following happens, so for every J between 0 and M minus 1, and for every cube of size delta to the J that intersects X, we look at how many cubes of the next generation hit X. And the next generation is delta to the J plus 1 cubes, so for every scale delta to the J.

So we're going to look at scales delta to the J. For every cube of one of these scales that intersects the set, we look at how many-- sorry, star-- how many cubes of the next-generation scale hit X. And this is some number that depends on the scale J, but it doesn't depend on the particular cube Q. Let's draw a picture. So this picture's to understand what's going on.

And I guess we actually want delta to be not just a number in 0, 1, but actually, a dyadic number. This is a definition. But in practice, we want to get a dyadic number so things are nested.

So let's draw a picture to understand what's going on. Suppose that delta is 1/2, and we are in dimension 2. So we look at-- OK, we have a set X. Let's say that X hits this 3 dyadic squares of first level, but it doesn't hit this one. Then R0 will be 3.

There is only one cube of first level, which is the unit cube, by the assumption that Q lives in the unit cube. So the unit cube has size delta to the 0. And X hits just one of those because it's contained in the unit cube. Now, we split each of the surviving squares, again, dyadically.

And what uniformity means is that the number in each of these three cubes that are hit by X is the same for all three cubes. But the actual cubes could be different. So for example, X could hit this one, this one, this one, this one, this one, and this one. So this is a possibility. And in this case, R1 will be 2.

So every of these three cubes that survived in the first generation hit two cubes of the second generation. And what is important is that it is 2, 2, and 2, the same for all the cubes. This is what makes it uniform.

Now we could keep going. So maybe in the next generation, every one of these will hit just one. It could be anywhere. Then R2 would be 1. Is it clear what a uniform set is?

So we can see the set X as a tree where the root is the unit cube the vertices of first level. So the children of the unit cube are the squares of size 1/2 hitting the set, the cubes of size 1/2 squared. So delta squared will be the children of the vertices of the first level, and so on. Yes?

AUDIENCE: How does this work when delta is the reciprocal of an integer, and you have this kind of [INAUDIBLE]?

PABLO SHMERKIN: Otherwise, things don't line up exactly. It's not the end of the world. But in practice, we will see that we are free to pick delta satisfying certain-- well, at this stage, delta could be anything. But yeah, we are going to pick delta dyadic, not just reciprocal of an integer, but reciprocal of a power of 2, just in case. Yeah, otherwise, things are not exactly nested, and this picture is not as neat.

So these numbers are J. We call them-- so we can view as a tree, X as a tree. And for this reason, we call them branching numbers. OK, I'm having a very bad blackboard management. So let's show this again for a second.

So in general, if we have a rho S delta-- sorry, a delta S, C set, it does not need to be a rho S, C set. But if the set is uniform, it is. So this is one big advantage of working with uniform sets. So let's say-- so this will overcome this issue and also, actually, this issue.

If the set is uniform, then this condition is not only necessary, but sufficient for the set to be-- yeah, so if we knew that Q of A is uniform-- which we don't, but OK-- we have to work a lot to actually make this work for us. But if we have a uniform set, then both of these issues are overcome. So if the set is uniform, and it is a delta S, C set, it is a rho, S, slightly worse C set. And if the set satisfies that it's large at every scale, then it is a delta S, C set. So both of these issues are overcome by working with uniform sets.

OK, let me state how it works. So let lemma-- so let X be delta M uniform. And let delta be big delta to the M. So again, at some point, we are going to fix big delta. And then this is not a restriction. This is not a restriction because delta will be a constant.

So if delta, small delta, is not big delta to the M, we sandwich it between big delta to the M And big delta to the M plus 1 and work with those powers. So this is not a big restriction. Then if the size of X is large, let's say, 1 over C rho to the minus S, for every rho, but in fact, we don't even need every rho. Let's say, for every rho in 1, delta, delta to the M. So we only need this to be true for powers of delta.

But the point is, if we know that the size of X is large at every scale, then X is a delta S and not quite C set, but-- well, up to some constant that depends on delta. But because delta will be fixed at some point, this is harmless. So for uniform sets, it is sufficient to be large at all scales for the set to be a delta S, C set.

And what is the other issue that we wanted to overcome with sets? Oh, OK, and this one, as well. And if X-- so X is a uniform set. If x is a delta S, C set, then X is a rho, S, and again, not quite C, but we lose something that depends on delta, so we are happy set for every rho between delta and one.

Maybe an exercise for you is to think about why both of these claims are false if we don't assume that the set is uniform. It does have to be uniform. It's not true in general. Any questions? Yes?

AUDIENCE: I was a little bit confused. If X in a ball that was covered, like you said, like [INAUDIBLE] X, it's not a rho S something set. If X was like rho S set, why does that imply Q X as the [INAUDIBLE]?

PABLO SHMERKIN: OK, so what we showed last time is that, if A is a delta S, C set, forget about what is the C, then Q of A at scale delta grows. Well, that is true for every delta. So if the original A was also a rho, S, C set, then we could apply what we did last time with Rho instead of delta and get this conclusion.

AUDIENCE: It looks like different rho with different polynomial.

PABLO SHMERKIN: The polynomial is fixed. So in the order of quantifiers-- so maybe this is still on some board. Yeah. You see, the polynomial comes first. The polynomial is fixed.

So it's actually A, A, A minus A, A, A.

AUDIENCE: Oh.

PABLO SHMERKIN: The polynomial doesn't depend on S, doesn't depend on rho, doesn't depend on C. It doesn't depend on anything. It's a completely universal polynomial. So it's really A, A, A minus A, A, A. Or there are other options.

AUDIENCE: Thank you.

PABLO SHMERKIN: Yeah. OK. So now we are seeing both the lemma, and you sort of see why the lemma will be helpful. Although, one has to be careful. And we will be careful because in order to overcome this, we need A to be a rho, S, something set. So if A was uniform, we would be happy. But for the second one, we need Q of A to be uniform-- not A, but Q of A.

So to overcome one issue, we need A to be uniform. To overcome the other issue, we need Q of A to be uniform. So we still will have work to do. But OK, let's go step by step and let's prove this lemma. And the proof of the lemma will also help us understand and work with uniform sets.

So the idea is to do everything just at the scales powers of delta. And any other scale we can sandwich between powers of delta because delta is a constant. The sandwiching will only lose a constant depending on delta. So for the first-- OK, maybe let's give it the letters A and B.

So for A, well, let's consider rho, some power of delta. Then what is-- so we have to look at the intersection of A with a ball of radius rho. But instead of a ball of radius rho, let's use a cube of size rho. And again, we can always pass between cubes and balls using only constants. And this makes sense because we want to use the assumption that the set is uniform.

So let's take a Q in delta D delta to the J of-- I guess I called it X. We fix one of the scales delta to the J. We fix a cube that intersects the set. And what is this? Delta is delta to the M.

So here, we are jumping from scale delta to the J to the delta to the M. When we jump from scale delta to the J to delta to the J plus 1, we have R J plus 1 children. Each of those children has R to the J plus 2 children and so on, up to R M minus 1, with a star. So you see what I have. I wanted to have this star. Without the star, this would be true up to constants.

So every time we go down one scale, every cube splits into R, depending on the scale number of offspring. And the point is, it doesn't matter where the set is with the number of offspring is fixed, but the assumption that the set is uniform. So let's write this as R1, R M minus 1, divided by R1, RJ.

What is the numerator? Sorry 0. We start at 0. What is the numerator? Yes?

AUDIENCE: Is it just a delta-covering number of X?

PABLO SHMERKIN: The delta are not covering number of X, because we start at the top and we go all the way down to the bottom. So the numerator is the delta covering number of A. And what is the denominator? Sorry, actually, we start at J because of the way we define things. So we stop at J minus 1. So J, J plus 1, I think, now is correct. Yes?

AUDIENCE: Is the denominator just a capital delta to the J covering of A?

PABLO SHMERKIN: Exactly. This is the delta to the J covering number, I guess with star-- because we split into R0 children, then R1 and so on. And if we stop here, we stopped at the scale delta to the J, big delta to the J.

And now we use the assumption. In part A, the assumption is that-- sorry, I keep switching around A and X. The assumption is that this is large. So the fraction is small.

So the assumption is that, so this is smaller up to a constant because now I'm going to go back to standard delta-covering numbers. This is C, rho to the S delta-covering number of A-- X, sorry. So we are done. So we took a cube of size delta to the J. Sorry, delta to the J. So this is rho S.

So we took a cube of size delta to the J. So delta to the J is our Rho. And we show that the delta-covering number of that is a constant, rho to the S, and the delta-covering number of X. So this shows X is a delta, S, C set at scales 1, delta, delta to the M.

So it satisfies the condition of being a delta S, maybe 0, 1-- sorry, 0 of C, 0 of 1 times C at these scales. But then we sandwich a scale rho in delta 1 between two of these. And we are going to lose some power of delta by the sandwiching. We are going to lose a power of delta in the constant. So this is why the constant in the claim depends on delta.

An exercise is to fill in these details. But hopefully, this convinces you that the claim is true. Any questions?

And the second part of the lemma says that, if the original set was a delta S, C set, then it is also a rho, S, slightly worse C set. It is a very similar argument. So it's really the same argument. So we do exactly the same thing again, except that now, instead of delta, we have to write rho. OK, let's do it again, but it's basically the same argument.

b, so assume that rho is delta to the J. Otherwise, we sandwich as before. Maybe delta to the J 0. And then we have to look at X intersection-- so now we take a J which is at most J 0 because we have to see what happens when we intersect X with cubes or balls of size larger than rho now. And let's pick a cube, well, of size larger than rho.

Then we intersect X with Q. And now we look at the rho-covering number. But it is exactly the same story as before. So what is this? This is R J0, blah, blah, blah, R J minus 1.

AUDIENCE: Does it start at RJ and goes to R J0 minus 1?

PABLO SHMERKIN: Yes. OK. And this is R1 R J0 minus 1-- sorry, R 0, R 0, R 0 R J minus 1. And this is the covering number at scale delta to the J0 divided by the delta-covering number at scale delta to the J.

And because the set is a delta S, C set, this is at most-- I'll do a constant because I'm going to now delete the star. This is at most C-- let's write it like rho to the-- sorry. Let's try to do this correctly.

So what is this. So here, we start at the top and go up to scale delta to the J0, which is rho. So this is actually A the original set at scale rho. We start at the top. We go down to scale delta to the J0

AUDIENCE: So is this A or X?

PABLO SHMERKIN: Sorry. [LAUGHS] Sorry about that, X. I should just call it A and then-- but if I call it in the statement, then I will write X. So it's not going to work. It's X. There is just one set which is X.

And the denominator is a delta to the J covering number of X. The numerator I just write the rho-covering number of X without the star. And for the denominator, we use the previous lemma. So we previously see I think it's not in any board anymore. Or maybe it is. It is there.

We saw that, if we have a delta S, C set, then the rho-covering number is at least rho to the minus S divided by C. So we use that at this scale. So this is our R.

So this has size R equals delta to the J. So we have shown what we wanted. So now, if we replace cubes by balls, and we interpolate scales, so we use just powers of delta, we get that X intersection a ball of radius R rho-covering number is at most a constant, at most a constant that depends on delta that comes from the sandwiching. The original constant that is here.

Then this is actually rho to the minus S in the denominator, R to the minus S in the denominator. So we get R to the S and then the rho-covering number of X. And that's the definition of being a rho, S, slightly worse C set. Questions?

So we have seen that uniform sets are much better behaved. We can first of all, if it is a good set of scale delta, it is a good set at coarser scales as well. And also, it's enough to know that the set is large at all scales to know that it has this spacing condition. So how do we find uniform sets? So this is the content of the next lemma, which is simple but extremely important and has lots and lots of applications all over the place-- but in particular, for what we are trying to do.

Let. little delta be big delta to the M. And let X be a subset of 0, 1 to the d. So we assume nothing about X. And what the lemma says is that x contains a "dense" which encodes a uniform subset.

So what does "dense" mean? Well, it means different things in different settings, so I'm going to prove a slightly more general version. So let mu be a subadditive set function finitely-- set function, like a measure, but the only axiom I assume is that mu of a plus B is less than or equal than mu of A plus mu of B, nothing else.

And a very good example to keep in mind, which is one of the examples where this is useful, but not the only one-- this is why I state the more general version-- is that mu of B can be the delta-covering number of B. So this is not a measure, because it is not additive. It is extremely far from additive, but it is subadditive.

Then there exists Y, which is contained in X, which is-- well, two things happen. One is that Y is delta M uniform. And the second thing is that Y is large measured by mu. Mu of Y is large compared to what? Compared to mu of X.

So mu of X could be anything. We don't know what it is. Mu is something very general. And we do lose something. We lose something like 2 delta log 1 over big delta 2d. Sorry, 2d is the ambient dimension to the n, to the minus n.

So here, the point is that the scale is delta to the M. And here, we have log of 1 over delta to the M. So we lose something exponential. But if delta is small, it's a very small exponential.

Maybe let's write it in this equivalent way. This is delta to the minus sigma M mu of X, where sigma is logarithm of 2d logarithm of 1 over delta-- so it's a double log-- divided by log of 1 over delta. So if we make delta small enough, we can make sigma, arbitrarily, close to 0.

So in practice, at some point, we are going to choose this delta so that this is small enough for our purposes. And from that point on, delta will be fixed, but M will not. So M will go to infinity.

Proof of the lemma--

AUDIENCE: Where it says delta to the minus sigma M, is it just delta to the minus sigma?

PABLO SHMERKIN: Plus sigma M, sorry. Yeah. Yeah, so the measure-- it doesn't have to be a measure, but the measure of Y is not much smaller than the measure of X. But not much smaller depends on your point of view. So we do lose some exponential thing. But the sigma is as small as you want by choosing delta small enough. Yes?

AUDIENCE: So you said, we fix sigma, that is in delta, and then we let M go to infinity. So does that blow up?

PABLO SHMERKIN: Yeah, it blows up. So it's not a-- yeah, it blows up. But it blows up like a power. But the power is arbitrarily small. Sorry. Ah. Sorry, bad day with typos. Delta to the sigma. So small delta is big delta to the M.

So small delta is big delta to the M. So this-- OK, sorry about this. So this is delta to the sigma by definition of sigma. And actually, it is delta to the minus sigma. [LAUGHS] Sorry. This is delta to the minus sigma. By definition, this is the definition of sigma.

So this is big delta to the sigma M. Big delta to the sigma M is small delta to the sigma. So yes, as M goes to infinity, delta goes to 0. This goes to 0, but it goes to 0 as a power, which is arbitrarily small. So this is the way we should think about this. Sorry. Is it correct now? OK. So the proof is some clever, clever pigeonholing, starting from the bottom of the tree. So the proof is trim the tree starting from the bottom.

So we have this at X, and let's call it XM. And we are going to construct-- by pigeonholing from the bottom, we are going to construct a nested sequence of sets all the way down to X0. And this is going to be our Y that we are looking for. OK, so let's explain. I'm going to explain how to construct XM minus 1, and then we just continue the same way by induction. So we have XM, which is X, really.

So we have all the cubes of size delta to the M minus 1 that intersect the set. But the branching numbers could be different. So this could be 4, this could be 1, this could be 2, 2, and so on. And we want we want them all to be the same. So we are just going to pigeonhole the most popular value. Because there are too many values, we are going to do it dyadically. So that's basically the proof.

So let's write it down. So let XM minus 1 l be the union of all the queues which are squares of level M minus 1 that intersects the original set such that the branching numbers, or the number of descendants of generation M, so Q intersection X delta to the M star covering number is roughly 2 to the l.

And what are the possible values of l? Well, it could be 0 because there is at least one-- because the cube intersects X, there is at least one descendant. And the largest it could possibly be-- so the largest 2 to the l can be is delta to the minus D. So each time we go down one generation, we subdivide into relative scale 1 over delta, and we have D dimensions.

So l ranges from 0 to the logarithm of-- well, D logarithm of 1 over delta. So we pigeonhole. Mu is subadditive. So you see that this is what we use. So mu of X is this. These sets cover X.

So mu of X is at most the sum over l of mu XM. So it is X, but let's write it as XM so it's clear how to continue inductively, X minus 1 l, and l takes D log 1 over delta possible values. So there exists some l such that mu of x minus 1 l has measured at least-- maybe let's put a tilde just in case-- 1 over the log of 1 over delta times the measure of X.

And so what the situation now? Now we know that the number of offspring is constant up to a factor of 2. So here there could be 7. Here there could be 8. Here there could be five. Well, in that case, we are going to keep the 4 with the largest mass, which, again, we can do by subadditivity of mu.

So inside each Q which is a subset of these, so the ones that survived, keep 2 to the l cubes-- exactly 2 to the l. So there are between 2 to the l and 2 to the l plus 1. We give exactly 2 to the l cubes of size delta to the M.

By doing this, we lose another factor of 2. So this is why we have 2d log of 1 over delta. Maybe it should be 4 or something like this. But of course, it doesn't matter. Let X M minus 1 be the union of all of those.

So first, we do dyadic pigeonholing because in dyadic pigeonholing, things are constant up to a factor of 2. Then I trim again to make it exactly constant. Then I only lose a factor of 2. So by doing this, now we have our set X M minus 1.

And the size-- sorry, the size in terms of the measure. So the measure could be the delta-covering number, but it could be other things. So the measure of X M minus 1 is bigger than 1 over to the log 1 over delta times the measure of X and D set is now uniform at the last scale. So every Q that intersects this new set has some fixed number-- let's call it R minus 1-- descendants.

So this will be 2 to the l for the l that we pigeonholed. And now we take X M minus 1, and we keep going up until all the way up to the top, until we get X0. And then we have to convince ourselves that X 0 works.

OK, so one thing is clear. Every time we iterate, we lose this factor. So we lose this factor M times. And that was the second part of the claim.

And the second claim is that the set Y is uniform. Well, every time we do this, it becomes uniform from one scale to the next. And the only thing that needs to be observed is that this is preserved. So when we go up from scale M minus 1 to scale M minus 2, we don't mess up the uniformity that we already had. And let me do a proof by picture.

Let's say that, in the first stage, we chose these three cubes. And each of them has the same number of descendants. In the next stage, we might remove some of these. But it doesn't matter. Let's say that we remove this one in the next stage. So the ones we keep still have the uniformity. So we don't destroy the uniformity that we already obtained at finer scale scales when we go to coarser scales. So it is uniform.

So this is the proof that, given any set, we can find a dense in this set, dense up to delta to the sigma such that it is also uniform. And well, by taking mu to be the delta-covering number, it is literally dense in the standard way. But sometimes it's convenient, as we will see or maybe not, depending on time. It is convenient to take other mu's. This is why I stated it in this way.

Any questions? Again, all of the tools that we have been discussing today are useful not just in this proof, but all over the place in projection theory. Yes?

AUDIENCE: I'm a bit confused by this construction. Yeah, I [INAUDIBLE] on level 1, everything's-- the degree is the same. But if you do induction for each subtree, you can only guarantee that the degree inside each subtree are the same. You cannot guarantee the degrees between different subtrees are equal, right?

PABLO SHMERKIN: What do you mean, different subtrees?

AUDIENCE: The first cube you divide this [INAUDIBLE] cube into four cubes and [INAUDIBLE] three yellow ones. And these three yellow ones have the same number of smaller boxes. But then-- yeah, as my understanding, you kind of do an inductive argument for each of those three yellow boxes.

PABLO SHMERKIN: No, we start from the bottom. So--

AUDIENCE: [INAUDIBLE]

PABLO SHMERKIN: Yeah, what is confusing you, you're thinking about starting from the top.

AUDIENCE: Yeah, yeah, right.

PABLO SHMERKIN: You cannot start from the top, precisely because of what is confusing you. It wouldn't-- I mean, you cannot make it work. But we start from the bottom. And this is exactly why we start from the bottom.

AUDIENCE: OK

PABLO SHMERKIN: When We go from the bottom to the top, everything, every uniformity that we gain is preserved when we keep iterating, moving in the up direction.

AUDIENCE: OK, yeah, that makes sense. Thanks.

PABLO SHMERKIN: Mm-hmm. So in the remaining 14 minutes, let's try to go back and see how all of this helps us deal with the issue that we had. Maybe let's keep the lemma

So let's go back. So now we go back to our goal. So A is a delta S. So this C is delta to the minus eta. But that's not where the issue lies. On the other hand, if you want to write the things correctly, you have to keep track of many parameters and how they change. And I'm not going to be able to do that today, so let's ignore this. So we know that Q of A has large size at scale delta.

So the first thing we do is we replace-- so Q of A is A, A, A minus A, A, A. So by using similar techniques or basically the same techniques that you are already used in the finite field setting, we can get the following conclusion. There exists one A in A, such that A plus little A, A grows.

It's not the same epsilon. So every time I write epsilon, it's a different epsilon. Is that OK, or do you want me to write the epsilon 1? We lose a lot. So this epsilon, we lose a lot.

So this is similar to the finite field setting. This uses a Ruzsa's triangle inequality. I guess the main thing it uses is Plünnecke-Ruzsa. And as I explained last time, all of these things are still true for delta-covering numbers. So uses Plünnecke-Ruzsa, Ruzsa's triangle inequality, and I guess, maybe some double counting. So I'm not saying this is trivial, but I hope you did something similar in the finite field setting. OK, great.

Second, we can assume that a is uniform. So first, we assume that A is uniform. And second, we do this. Maybe this is better. So A is uniform. So we have to choose, at this stage, delta small enough. So here, the mu is just the delta-covering number. So we apply the uniformization lemma.

So we start with the delta S, C set. Now we pass to a uniform subset. The uniform subset is dense up to a factor delta to the epsilon-- not epsilon, delta to the sigma, where sigma is as small as I want. Then subsets of delta, S sets are still delta, S sets by the lemma that we saw before. This is why we can assume this.

Now, this constant is going to become worse by a factor of delta to the minus sigma because that's the density. So C gets worse by delta to the minus sigma. But sigma is arbitrarily small. So I'm going to ignore it. So you have to keep track of all of these parameters. But they are not the real issue. You have to be careful about that. They are not the real issue.

Then, this A still satisfies the assumptions of what we did last time. So this is really the first step. And so we know that a for the uniform one A, A, A minus A, A, A grows. So already, this grows. But A is uniform. So in fact, because A is uniform, by the lemma that we used about uniformity, this is true at every scale. So in fact, we have this at every scale.

The third thing that is going to be important is that, a priori, we only have one A in A. But in fact, once we have it for 1A in we have it for almost all in a. And maybe you have seen this sort of argument in the finite field setting. but this is true for almost all a and A because we define A bad to be the points in A such that the conclusion we want-- so the second step is not true. It doesn't hold. And if the set is large, we can apply the same argument to this set and get a contradiction.

So a dense subset of a delta S, C set is a delta S, slightly worse C set. How much worse a constant is depends on how dense A bad is. So if A bad is too dense, we get a contradiction. So it cannot be too dense. And this means that very few A's are bad. So almost all A are not bad. Well, the step just 2 to A bad, and we get a contradiction.

The first step is maybe the key step. So we are going to upgrade the second step to not just saying that A plus little A, A is large, but-- so what is A plus little A, A? It's pi A of A times A. So we are going to upgrade this to say that this is still true if we pass to a dense subset of A times A.

So first step, if X is contained in A plus A, and it is dense, in the sense that the delta covering number doesn't-- so at some point, there should be a parameter eta that I am completely ignoring. Maybe at this step, we want something like eta square dense. Then if we project X, we still have growth. And in fact, at all scales, or almost all scales. Maybe instead of epsilon, the growth will be epsilon over 2. But there will still be growth.

Let me try to explain. The first step in the last five minutes. So this is really, as we will see, the absolutely key step. We need to make this estimate in the second step robust underpass into dense subsets. And we will see, maybe next time, why this is what we really need to do in order to conclude what we want. So suppose that one is in A, and this is just for notational simplicity.

So there are two possibilities. One-- OK, so we want to show that there exists an A with this property, one a in A with this property. I guess step 3 should maybe be a later step. Or maybe we want to apply it at many steps. So many steps, we prove that something exists. But once something exists, actually, almost all elements in A have this property.

OK, let's say we want to prove that-- So at this stage we want to get one in little A in A so that this is true. So maybe we are lucky. And the first thing we choose-- and let's call it 1-- works. Well, if it works, it works. So first case-- A equals 1 works. So it has this property. Well, if it works, it works. We are done. So let's assume that 1 doesn't work.

What does it mean that A equals 1 doesn't work? It means that there exists some X in A times A which has size bigger than delta to the eta squared or something like this times the size of eta times A. But the projection is small.

We have to do this for every rho. But let's do it one rho at a time, and next time, I will-- let's do it for delta. Next time, I will explain why it's true for every delta, in some sense. OK, let's assume that, for the particular scale delta, there is no growth.

Well, this reminds you of what? So there is a theorem that you've seen before that should remind you of. So we are projecting-- sorry, an A is 1. Actually, A is 1. So this is just the sum.

So we are-- so if 1-- we assume one is in A. Either it is good-- if it is good, we have proved what we want because something is good, or it's not good. If it's not good, this is what it means, by definition of not being good.

B is G. So this is exactly the assumption to apply Balog-Szemeredi-Gowers. So A has size delta to the minus S. Here we have a dense subset of the product set which has small partial sum set. The pi 1 image is small. It doesn't grow. So these are exactly the assumptions that we need to apply Balog-Szemeredi-Gowers.

I've run out of time. So we are in the middle of the most important argument. So maybe on Tuesday, sorry, you will see me again. I will start with step 4 from the top. But maybe as an exercise, you have five days. You can try to imagine how this will continue.

So at least try to well, just write down what is the outcome of Balog-Szemeredi-Gowers, and try to think how we can use that to eventually get what we want, at least this claim. Then I will explain how this claim, together with uniformization lemma, allows us to conclude what we want.