[SQUEAKING] [RUSTLING] [CLICKING]

**PABLO SHMERKIN:** OK. So hi, everyone. I'm Pablo, I'm teaching Larry's lecturer this week. I hope we are all here to learn about prediction theory. The plan for this week is to talk about Bourgain's prediction theorem, which is an extremely important theorem with applications in many areas of mathematics, including some pretty striking recent ones. And I think Larry has told you at least a bit about it. Well, we won't be able to do a full proof because it would take way too long.

But what I'm going to do is explain-- OK, so you've seen the analog of this theorem in the finite field setting. I'm going to state it again to so as our starting point. And then the plan is to go through the steps in the strategy in the finite field setting and see what goes wrong and what are the ideas that one uses to overcome what goes wrong.

OK, so recall how this theorem looks like in the finite field setting. And by finite field, I mean prime finite field, prime order, because otherwise, it's not true. The statement was, hopefully, something like this, given t between 0 and 2 and s between 0 and 1. t, I want it to be neither 0 nor 2 because there wouldn't be anything to say. s has to be positive. Could be 1. There exists some epsilon that depends on these two parameters and is positive with the following properties.

If X is in Fp squared and has size p to the s. So s dimensional set in some sense. And D is a set of directions, let's say a subset of p. And it has size p to the t, so dimension t. Then so I use the notation pi theta of X to be the projection. And we can write the projection in this way, I guess, just x1 plus theta x2. So from Fp squared to Fp.

Then we can find a-- sorry, I messed up the parameters already. So the size of X is between p to the 0 and p squared, and t is between 0 and 2. So this should be t. This should be s. OK. Then there is a "large," between quotes, projection. Then the maximum over theta in D of the size of the projection of X is at least p to the t over 22 plus epsilon.

So we're getting an epsilon over t over 2. And p to the t over 2 is the obvious lower bound. And I'm going to come back to this obvious lower bound in a minute. In fact, there is a stronger version. I think you proved this in detail, and you discussed the stronger version. So let me state the stronger version because, well, first of all, it's important in itself for applications of the finite field result. But I would say it's even more important in the Euclidean case.

Not only we can find-- so this says we can find the direction so that the projection is large, the projection of the whole set. But moreover, we can find the direction, so that if we project any dense subset of X, the projection is still large. So if we take the maximum over now the minimum-- so we look at the worst possible subset which is dense. So we take the minimum over subsets of X which are dense in the sense that they have relative density p to the minus epsilon.

And then we look at the projection of Y. So we try to make the projection of a dense subset of X as small as possible. But we can't even if we have this additional robustness of looking at all subsets of Y, we still have a gain over t over 2. So this is Bourgain's projection theorem in the finite field setting.

And you have proved at least this version, and I think discussed this version. If anything I say is not correct, or if I say something that you are not familiar with or you don't understand anything, please just ask. So now we want to-- OK, so before moving to the Euclidean case, let me make this remark on the trivial bound. So the trivial bound is p to the t over 2 just without the epsilon.

So this theorem is not true if we replace p by p squared or anything which is not a prime. Because if you look at Fp squared, for example, then Fp is sitting there. And if we take f to be Xp and D to be Fp sitting in Fp squared, we don't escape Fp because Fp is a f So this is not true.

But without the epsilon, it's true in any field, and the argument is trivial. OK, so one should never use the word "trivial," but I will use it in this case. So everything is trivial once you've understood it. And nothing is trivial before you've understood it. But I will still use the word "trivial" here.

So if D has at least two elements-- so same setting as above. So if D has more than one element, so let's allow p to the s where s is positive, just two elements, then the maximum over theta in D of the projection of X is at least the size of X to the 1/2. So if the size of X is p to the t would be at least p to the t over 2. And again, we will see that this doesn't use the factor p is prime. So it's much more general.

This is trivial because just take two elements. OK. Again, trivial. I will explain it. And then once you've understood it, you will agree that it's trivial OK, so know that. So just take two elements. So D has at least two elements by-- so let's just take two elements, phi times theta prime in D. So the observation that makes this trivial is that if you map X to the projection in direction theta comma the projection in direction theta prime-- so this is small X, an element of Xp squared. This is injective. So we can recover a vector from two projections by linear algebra.

So this means that the image of X under this map has size at least the domain. Because you have injective map, the image has the size of the domain. OK, so this implies that the size of X is at least the size of the projection in direction theta of X times the size of the projection in direction theta prime. So at least one of these two has to be at least square root of the size of X.

So this is the proof of the trivial bound. And I hope you will agree that it's very easy. And it uses literally nothing. Well, it uses that this map is injective, but this is a very general fact. In particular, this is true on any finite field. It doesn't have to be prime. But for this, we need it to be prime.

So now we want to have a Euclidean version of this statement. And we will see that-- OK, so maybe let's discuss the Euclidean setting. So I will start with the setting, which is the way in which Euclidean setting has appeared in this course before. But then I will rescale things to a setting that I like better. But let's start with the following. So we have a parameter R, which is the scale, or maybe 1 over R is the scale.

And then we have our set X is contained in the ball of radius R. Let's say centered at the origin or really any ball of radius R. It doesn't make any difference. And let's suppose that X is a union of unit balls. And D, let's say in 0, 1. Maybe 1, 2 would be slightly more natural, but it doesn't really matter much, is a 1 over R separated set.

OK, so D is a finite set of directions. They are separated by the scale 1 over R. And X is a unit of unit balls in BR. So we measure X by Lebesgue measure. We measure D by cardinality and use bars in both cases. I think this is the convention you'll be using in this course. And well, we would like the following.

Well, we would like that, right? So except that now X is not in Fp squared, but in BR, and D is no longer in Fp but in 0, 1. So pi theta is still the same. So pi theta of x1, x2 is x1 plus theta x2. We could look at orthogonal projections but up to a reparametrization is actually easier to work with. It's not important.

OK, so we can ask, Does the same happen? So just all the bars mean different things now. So when we project X, we are looking at the measure, the big measure one-dimensional Lebesgue measure of the image, and so on. So do we have any hope of proving the same theorem in this generality? So does anyone want to guess? Could this theorem be true in this setting?

**AUDIENCE:** No, it's not true because the balls are closely packed into kind of a ball of radius blue R.

**PABLO SHMERKIN:** OK. Great. That's a great example. But that great example, then I will come back to this, still satisfies the trivial bound. So can the trivial bound happen in this setting as written? Because, yeah, if everything is packed into a smaller ball, if you project, you still have an interval of size the square root. So you're right, that fails. I'm going to come back to your example. That fails, but this still holds. Can the trivial bound be true in this setting without assuming anything else?

**AUDIENCE:** Do you see how only two interactions are--

**PABLO SHMERKIN:** OK. Yeah, two directions maybe is asking too much. Let's say there are-- OK. What is p to the t, and what is p to the s? It would be rho to the t, and rho to the s. Not rho, R-- R to the t, R to the s. Let's say there are R to the s directions. Let's make it a little bit more. Yeah, with two directions maybe-- OK. Let's say there are R to the s directions, which R1 over R2 separated. Can the trivial bound hold? OK, what is an example of a set that has one very small projection? So what is a set in the plane that has one very small projection?

**AUDIENCE:** Called the Gaussian line?

**PABLO SHMERKIN:** A line. OK, you cannot literally be a line because it's a union of balls. But we put the balls in a long segment. So an example that should make us kind of scared is a 1 times R rectangle. So we pack with balls in this way as someone suggested. Then if we project literally down, we get something of size 1, which is much, much, much less than square root of R.

But if we perturb the vertical projection by R to the s, the projection will have size R to the s. And if s is less than 1/2, it will be less than square root of R. So instead of projecting in this direction, if we project in a nearby direction, this is still bad. So if this is X and-- OK. So I think actually, I chose a bad-- so this projection cannot be written in this way. Oh yes, it can be written in this way, just theta is 0. It can be very well written in this way.

If theta is, let's say, minus R to the s, R to the s, then the projection of X will have size of the order of R to the s. Well, s could be positive but less than 1/2. So this is very bad. So we don't want-- this is not theta. This is D. So this is very bad. So we don't want D to look like an interval or to be very concentrated in an interval or something like this. So we want to avoid such a situation. So this is bad. It makes even the trivial estimate hopeless.

So maybe we'll come back to this if there is time, maybe not. But maybe as an exercise for you, not exercise for marks or anything, but you can think of it. Not now. Now, pay attention. But you can think later. What goes wrong with this argument? Because in R2, it's still true that two projections, even if the angles are very close, determine the original vector. So how is this compatible? So you can think about this. And then you will understand many of the-- well, not many, one of one of the important reasons why things are much more complicated in this Euclidean world.

OK. So now suppose that we avoid this in some way. Can we have a hope of having the theorem? And the answer, as you said, is still no. So this is one example we should be worried about. Here is another example we should be worried about, where X is, essentially, let's say a ball of radius square root of R. It can be really any radius between 0 and 1, but just to-- so X should be a union of disjoint unit balls. So we just pack, in a very dense way, unit balls in a ball of radius R to the 1/2.

So the volume of X the area of X is roughly R. But if we project it in any direction-- so we have to get rid of bad sets like this. But let's get rid of bad sets like this in the strongest possible way. Let's suppose that D is everything. So here, we are focusing on what can go wrong with X. So let D be absolutely everything. Well, here, every projection is going to be a segment of length roughly R to the 1/2.

So if you look at the projection of X, this has size roughly square root of the size of X. And this is for every theta. So in fact, if you avoid this in a suitable way, we recover the trivial bound, but it is no longer trivial. Maybe we'll come back to this on Thursday depending on how we are doing with time. So the trivial bound is no longer trivial even in the best case.

But if we want to go beyond the trivial bound, we also need to assume something about X. So we need to avoid this, and we need to avoid this. And you see that in both cases, what's going on-- so the geometric picture is different. But in both cases, D is an interval. X is a ball. So we want to avoid D looking like an interval or X looking like a box. If we do that in a suitable way, then we can indeed get a similar statement. So let me actually state it.

OK, let me say that Bourgain didn't state it in this way, but-- it follows from what he proved. Suppose that-- same setting as before. So we are still assuming this. Would have been faster to just write everything down again. OK, but we need to add some assumptions to avoid these bad examples.

OK, so first of all, we need some quantifiers. So given 0, t. So t is strictly between 0 and 2 and s larger than 0. There exist now two parameters. So I could do with one parameter. So you see that even here, I'm using epsilon for two different things-- for the growth and for the robustness in passing two subsets. Here, I'm going to use two different letters because they denote different things. So there exists epsilon and eta, both, depending on s and t, and both positive such that the following holds.

So suppose that we have X union of unit balls in the R, 1 over R separated set of directions. And then we need to assume something because otherwise it's not true. So we need to assume that X and D are not very concentrated in small balls or small intervals. So this is what we are going to assume. If we intersect X with any ball of radius r, we see only a small piece of X And the smaller they are, the smaller the piece. The range of this little r is between 1 and big R, because X is a union of unit balls, and it's a subset of the R ball.

So this should be less than or equal than what? Well, on the one hand, we want to compare this to the natural scale, which is R. So this tells us how much smaller the reference scale is compared to the global scale. And let's put the t here, which is motivated by the fact that-- OK. So X has area R to the t And D has cardinality R to the s. This t matches this t. I'm going to have the size of X.

So this is saying if R is much smaller than big R-- so if R is big R, we have all of X. So we cannot get any gain because X is contained in the ball of radius R. But if we are intersecting X with a smaller ball, we have a big decay, sort of power decay with respect to-- so only a very small part of X can be in a small ball compared to the whole of X. It turns out that it's very important for applications to have some leeway here. And this leeway is given by this parameter eta.

So this leeway says that actually, we don't need to assume anything if little r is very close to big R. Because if r is very close to big R, this will still be bigger than 1, and then we have a trivial inequality. So this is saying the global scales don't matter. OK. A very similar story with D. If we intersect D-- OK, both here are intervals, but I'm going to-- steal the notation for balls. This is at most. Again, we have our leeway R to the epsilon.

And then rho is between 0 and 1. So here, the right thing to write is rho to the s times-- OK, so the size of X is actually R to the t. But still, I want to write it in this way so it's clear that when we intersect with the ball, we are seeing only a small piece of X. And the same thing here. OK, so these are the conditions that we impose to avoid these bad counterexamples.

You could say that they are overkill. And we are avoiding not only this but-- OK. Put it in a different way, there is a whole universe between this and this. And, yeah, this is true. In fact, these conditions can be weakened. Maybe I will make some remark about this. Anyway, so these are the assumptions. Then there exists one direction. And once there is one direction, there are many directions, such that we have the same conclusion.

So if we look at the infimum over all dense subsets-- and "dense" means with respect to this parameter eta. So eta is like a tolerance parameter. It gives us some leeway at different places to play with. So Y is a dense subset of X with this leeway R to the minus epsilon. Then we project the subset, and we get the gain. So the "trivial" now is no longer trivial but sort of natural. Lower bound is t over 2. And we beat it by epsilon. Yes.

**AUDIENCE:** Does Y also have to be a union of unit balls?

**PABLO SHMERKIN:** That's a very good question. I don't think so. But, yeah, just in case, yes. Just in case I'm too tired, yes. Yeah. Let's say yes. Let's say yes. So we're going to restate this in a different language soon, where we will get rid of balls. But, yeah, let's say Y is also union of unit balls So you see that there are lots of parameters. It's kind of a technical statement. The assumption looks a bit strange if you have never seen this before, but nevertheless is, again, an extremely important theorem with an unbelievable range of applications. We keep discovering applications.

Something which is also important to remark is that all the applications really use the statement as a black box. Sometimes people also get inspired by the proof. But it's not a situation where the proof gets used, really the statement gets used. So it's a very, very powerful statement. And I guess in the rest of the course with Larry, you will see some of the ways in which it is very powerful.

How does one prove this? By the way, I'm going to, as I say, sketch the proof, skipping pretty much all details. But the proof I'm going to sketch is not Bourgain's proof. It's maybe a new proof that I'm writing together with Hong Wang that takes ideas from different places. So it's not Bourgain's original proof.

The general scheme will follow the steps that you've seen in the finite field setting, but there are very significant challenges. So what I'm going to do is go over each of the steps and explain what the challenges are and what can one do to fix them. And then you will have to trust me that, indeed, they can be fixed. So I won't go into the details.

Before doing so, you saw that if you want to go from finite field to Euclidean settings, sometimes you have to add more assumptions, otherwise things fail because of some easy examples. But this is not always the case. So there are some statements in the finite setting that extend a pretty straightforward way to this Euclidean setting. So in order to do this, I'm going to change the language a little bit and maybe rescale this. So instead of looking at unions of unit balls in the R ball, I'm going to look at-- so basically, I need to rescale the set down by 1 over R.

So X is going to live in the unit ball. And then it will be a union of 1 over R balls. And of course, if you rescale you scale, you have to scale everywhere else in the statement. But instead of looking at the volume of the set X, I'm going to use the delta-covering number. For the directions, instead of assuming that they are disjoint, I'm also going to use delta-covering number. So I bring both-- so basically X and D, I want to use the same language and same scale to make things more homogeneous.

So let me spend a couple of minutes talking about delta-covering number. Suppose that X is a subset of really any metric space, but let's say Rd, then-- OK, so this is, I'm guessing, definition or notation X bar delta. So bar with the subscript delta is the smallest. number of delta balls needed to cover X. Could be infinite, but in all this business, we are only looking at bounded sets, in which case this is finite. So everything we are going to see is going to be a finite number.

Basically, our measuring resolution is delta. We cannot distinguish anything that happens under scale delta. So this is a natural way of measuring the size of the set. Just a definition, but some observations-- well, the first one I think I should never use the word "trivial," but I hope I'm justified in saying that this is trivial if X is, to make it really trivial, 2 delta separated, then the delta-covering number is just the cardinality, because for each point in the set, we need a different ball.

So this will be the case for D. D is delta separated, but delta separated, 2 delta separated. Doesn't make a big difference. So if the set is delta separated, it's morally just cardinality. If X is a union of delta balls, then the delta-covering number is-- well, you use the balls to cover the set. You need to normalize to pass between measure and counting how many balls there are. So I think it should be delta to the minus D. So D is the ambient dimension where X leaves times the Lebesgue measure of X. Did I get this right? It's not actually equal. It's comparable. Yes.

**AUDIENCE:** Afterward, does X have to be a disjoint union of delta balls?

**PABLO SHMERKIN:** That's a very good question. No. It's easier to believe if it is a disjoint union of delta balls. And it is also true if it's a non-disjoint union of delta balls. The reason is that if you have a non-disjoint union of delta balls, you can cover it by finitely overlapping. OK, if you have a disjoint union of delta balls-- I hope this is clear, because the best way to cover it by using those balls.

If you have a finite overlapping collection of delta balls, then you just cover by defining the overlapping family and you lose. Any union of balls, you can cover by finitely overlapping. So I guess you need some covering theorem to make it precise. Anyway, it's true, and we are not going to use it. But you can convince yourself that it's true. Again, easier to believe if the balls are disjoint, but also true if the balls are disjoint. It's equal if the balls are not disjoint. There is a wiggle. In fact, the wiggle, I guess, maybe depends on the ambient dimension D.

I'm just making these remarks to explain how this relates to our previous setting. Because now, from now on, we're going to measure everything by the delta-covering number. The point I want to make is that this recovers the volume and the cardinality when the sets are delta-separated or union of delta balls.

Another way of thinking about delta-covering-number is as a box-counting number. So let D delta be the covering of D by delta cubes. So this is delta k plus 0 delta. Maybe let's make it half open so that this is a partition to the d, where k is in Z to the d. So I'm just doing a dyadic covering. Delta doesn't have to be dyadic, but I'm-- so this is-- how do you say? The tiling of Rd by cubes of size delta. I make them half open so they are disjoint, but it's not very important

OK, so the delta-covering number is up to a multiplicative constant, the number of cubes in the delta that are hit by the set. So the cardinality of the cubes in D delta which intersect the set. So this is why this is called box counting number as well. Why is this true is because if you cover by cubes in an efficient way-- so if you know how many cubes you are hitting, well, you can cover each cube by a finite number of balls of radius delta.

Again, this wiggle may depend on d. So d is a fixed parameter. It will be 1 or 2 for protection theorem. So I consider it as a constant. So you can cover each of these cubes by a finite number of delta balls and vice versa. If you have a delta ball, you can cover it by a finite number of cubes. So that's the proof, proof by picture.

One nice consequence of this is that this allows us to snap any set to a grid. And then this allows us to apply Euclidean-- not Euclidean, integer results. Discrete results. So let's define-- so what is X superscript bracket delta? It's the set that is obtained by snapping points to the grid. So if X intersects this cube, then this point will be in X delta. And this way, I snap it to the lattice.

OK. Formally, this is all the case such that the cube delta k plus 0 delta to the d intersects X. So anytime that we see something here, we look at this point. And then because of this, the delta-covering number is roughly up to a multiplicative constant, the same as the cardinality. This is now a finite set. It's a finite set in a lattice. So it's not just the lattice. It's just the integer lattice scaled down by delta, but it's just the integer lattice.

By using this, one can extend many results that are true EZD to delta-covering numbers. So using this, in particular, the following three things, which are three very important tools in the proof of the finite field case of recurrence projection theorem and also in the Euclidean case, the Ruzsa triangle inequality. The Plunnecke-Ruzsa inequalities. And the Balog-Szemerédi-Gowers theorem.

They all hold for delta-covering numbers by putting bar there. So when you see bar, you put delta. You use multiplicative constants because of this. So where you have equality, you no longer have equality. But this is very harmless. For example, I'm going to state that Ruzsa triangle inequality as a model. And then you can imagine how the other two behave. So we have three sets in the same Rd, let's say, bounded to make sure that the delta-covering number is finite.

Then let's see if I get this right. If not, you just let me know. OK. In the original, so in the discrete Ruzsa triangle inequality, here we have inequality. Here, we have inequality up to a constant. So this is the only thing we have to change. OK. Did I get this right? Is this the correct statement? OK, so this is an example. Just put delta, delta, delta, delta, and change equality by equality up to a constant. And how do you prove this? Well, you snap it to the grid.

You have to be a little bit careful because you can snap A to the grid. You can snap B to the grid. And then you can take A minus B for the snapped versions. That's not going to be exactly the same as snapping a minus b to the grid. So there is something that doesn't exactly commute, but it almost commutes. So almost commutes. So I'm not going to-- another exercise is for you to convince yourself that this is true.

So you see that for the projection theorem, we really need to impose assumptions if we want it to hold in the Euclidean setting. But this is not always true. Here, we have three very important tools. They basically work in the same way. So now we can start going over the steps in the proof of the theorem, the finite field setting and seeing what the obstacles are and what one can do to overcome the obstacles if one wants to follow the same scheme of proof in the Euclidean setting. Any questions so far?

In the finite field setting, what was the first step in the proof? It was to prove expansion using a polynomial for just one set. So let's recall what this was. So if we have A in Fp-- OK, I guess let's say the quantifiers first. So given S in 0, 1, there exists an epsilon such that the following holds. If we have a subset of our prime field-- so here, it has to be a prime field. These are ready fails if it is not a prime field.

I'm going to come back to this point several times. Suppose that it has size P to the s. Then there exists a polynomial-- actually, there exists a polynomial, which is fixed for all s, for all epsilon, and for all a. It can be x1, x2, x3 minus x4, x5, x6. So that polynomial works, but it doesn't really matter which polynomial it is. So let's call it Q. If we apply Q to A, we get expansion.

OK, so I think what you did with Larry gave the following polynomial. But I think it was clear from what you did with Larry that this really doesn't matter which polynomial it is, as long as you have some polynomial. Because at a later stage, one is going to apply Plunnecke-Ruzsa to bring down the polynomial the simplest possible polynomial. So that was a later step. At this step, any polynomial works. So this one works, but it's not important, which one it is.

How did you prove this? So we can see to what extent things go wrong in the Euclidean setting, and to what extent we can recycle some of the ideas. So idea of proof. If I claim that you saw something with Larry and you didn't, please let me know. He told me and showed me notes, but I could still be wrong about something. What was the idea of proof behind these? So one looks at, maybe let's call it X. A minus A over A minus A.

Here, we skip 0. We don't want to divide by 0. And then there are two possibilities. Either X is everything, and then one can conclude in some way-- I know this is not-- I'm not going to recall how to do it. The other possibility is that X is not everything. If it's not everything, there exists some little X in X such that X plus 1 is not in X. Because it's not everything. And then one can also conclude.

So in both cases, one has to work. I'm not saying it's trivial, but this was the structure of the proof. These extremely innocent step is what distinguishes Fp from Fp squared. If you were working in Fp squared, we would not be able to do this because Fp squared has a subgroup which is Fp. So you cannot escape Fp by adding 1. So this extremely innocent adding plus 1, distinguishes Fp and Fp squared, let's say.

We want to do something similar. So let's say we have an A. Well, first of all, it's very likely that we will need to assume something about A. In fact, we need to assume something about in the Euclidean setting, because if A is a segment, like any polynomial check, it doesn't grow. So if X is 1, 1 plus R, 1, 1 plus delta to the 1 minus t, something like this, just a small segment, if you add it to itself multiplied by itself, you get slightly longer segments. There is no real growth.

It's exactly the same issue as we had when we were projecting an X which was to concentrate in a large ball. But now we know how to get rid of that. We need to assume that A is not too concentrated inside any segment. But even assuming that we do that, we see that we have lots of problems if we try to do something like this. What are the problems? So can you tell me some problems? So some challenges in trying to do something like this. If now A is a subset of let's say 0, 1 or 1, 2-- let's say 0, 1. It doesn't matter. So what things you should be worried about? Yes.

**AUDIENCE:** X should be unbounded.

**PABLO SHMERKIN:** X could be unbounded. Yes. So in the finite field, we need to avoid 0 in the denominator. But in the Euclidean case, we don't only need to assume 0. We should be very worried about small denominators, not 0 but small denominators, because this is going to blow up. Yeah, this is certainly an issue. Are there any other issues one should worry about?

**AUDIENCE:** What does it mean to equal Fp in this case?

**PABLO SHMERKIN:** That's right. Well, I mean, what does it mean to recall Fp? It's not clear. I would say what is even much less clear is, What does it mean to add 1? Because I don't know, R is certainly not-- I mean, R is invariant than adding 1. I mean, this doesn't live in 0, 1 anymore. If we force it to live in 0, 1 by intersecting with 0, 1, there could still be something. I mean, if we add 1, we could still land here. So adding 1 doesn't seem like a good idea.

From a conceptual point of view, there is something that is even more worrying, which is here, it is crucial that Fp doesn't have any additive subgroup. So it's a simple group. It doesn't have any non-trivial subgroups. Because if it had a subgroup, this wouldn't be true. Now does R have any subfields or subgroups or subrings? Well, first of all, it does have subgroups, additive subgroups. It does of any dimension. So, of course, it does. All of our rational numbers, number fields.

But number fields are countable. So we are interested in larger things. But it does have some groups of intermediate size. It does not have subrings of intermediate size. But this is a difficult theorem. In fact, it's a consequence of projection theorem. So we cannot really use it. So here, we are escaping because X is not a subring. But it's not so clear if there is a subring type structure here in R.

All of these things should make it very scary, and a normal person would give up at this stage. But there is a paper by Larry Guth and Josh Zahl, my colleague at UBC, and Etzkowitz, where they managed to extend this idea to the Euclidean setting, not to prove Bourgain's projection theorem-- to prove some different related theorem. So let me tell you what they did to overcome these obstacles. But first of all, I should define something that I should have defined before.

So we have these non-concentration conditions. So we call these non-concentration conditions because it tells us that X and D are not concentrated in both, non-concentration conditions. So looking at this, we are going to write a definition, which is inspired by this so that we don't have to repeat these sorts of conditions every time.

By the way, even in the finite field setting, what Bourgain, Katz, and Tao did was also not this. So this idea came later. It originated from Gaurav, which is after Bourgain, Katz, and Tao proved something like this. So Bourgain this is something much, much more complicated. OK. Definition. A subset of the unit ball in Rd-- we just want to fix some region of space, so why not fix the unit ball? Let delta be between 0 and 1, not 0. And maybe also not 1.

So we say that this is a delta. And then we have a parameter t, which is between 0 and d. And then we have a constant, which is positive, let's say at least 1. We say that X is a delta S, C, d. So this d tells us the ambient dimension, and I will probably forget to write it down, or it will be clear from context. Set if the following holds. If we intersect X-- so I'm not assuming X is a union of vaults. I'm not assuming X is delta-separated. I'm not assuming anything, and I'm not assuming anything because I'm using to use the delta-covering number to measure things.

So this is nothing about the delta-covering number. You don't need to assume anything as long as you remember to write the bad delta. So if I intersect X with the ball and measure it by the delta-covering number, this is at most C. So C is the tolerance, the leeway that I have, times R to the s. So s measures the degree of concentration of non-concentration, the degree of non-concentration. So the larger the s-- so R is going to be between 0 and 1, or between delta and one. So the larger s, the smaller this is, and the stronger this condition becomes times the size of X. So again, this says that a ball cannot contain too much of X. And this is for every center and for every radius between delta and 1. Yes.

**AUDIENCE:** The s is supposed to be t unless you use both s and t.

**PABLO SHMERKIN:** Sorry, yes. But it could be s or t, but one has to fix one. Yeah, s. Thank you. OK, so I hope you agree that in the statement of Bourgain's projection theorem, this is the condition we have on x and on d. On X, we have a t, and on d, we have an s. But this is the condition that we had. And so we can bring them both under the same umbrella using the delta-covering number.

OK. This is saying X is not too concentrated inside small balls. And the strength of concentration is given by s, and there is a tolerance given by C. C could depend on delta. So C is anything. In particular, it could depend on delta. And in the statement of Bourgain projection theorem-- so in Bourgain projection theorem, this, in fact, is going to be delta to the minus eta. So we have a leeway that increases as a power with delta, is very small power but nevertheless a power.

So now we can state something. I don't know. Lemma. So it is the analog of the first step in the proof. So it's the analog of this. Given an s between 0 and 1, there exists an epsilon, and all these epsilons can be taken to depend-- they depend continuously on s. So this is important. OK. Well, I started saying the sentence I'm going to finish. So in the finite field setting, one does this. And then what is the next step? One iterates this.

And then one keeps growing. In order to know that you keep growing, you need to know that this epsilon remains bounded away from 0 as long as the exponent remains bounded away from 1. OK, so the same is going to be true here, although as we will see next time-- I think next time is much, much more complicated to iterate in this case. But let's focus on this for today.

Given s, there is an epsilon such that the following holds. Suppose A is in 0, 1. The delta-covering number is delta to the minus s, and A is a delta s is delta to the epsilon minus epsilon set. So it's very non-concentrated because s matches the size. But on the other hand, there is quite a bit of leeway in this delta to the minus epsilon. OK, so there exists a polynomial Q. It could be the same Q, so that Q if you want.

Any polynomial. So really, it doesn't matter because you can always apply Plunnecke-Ruzsa. But let's say that polynomial. But any polynomial will do. Then we have growth. I'm just putting delta-covering number everywhere. And we get growth. So we get delta minus s, which was the size of A. But we get a delta to the minus epsilon grow over delta.

It looks very, very similar to that. To what? To this. Except that we need to add this non-concentration assumption. We certainly need to add something because otherwise, if A is just an interval it's not true, it's not clear between interval and this what happens. But OK, with this assumption, it is true. OK, so how can one prove that? So we want to use this strategy, but we saw that there are many problems.

So I'm not going to do the whole proof. But I'm going to tell you how to solve two of the problems. To some degree, I will tell you how to solve two of the problems. One is really cheating, but OK. So I guess what I'm going to tell you is that cheating doesn't always work, but it does work in this case, in this particular case.

I should finish at 2:25. Is that right? OK, plenty of time. The idea of proof. I'm not going to prove it, but-- in the finite field setting, this works. Here, we have the problem that the denominator can be very small, and then this could be huge. So well, we are going to do two things at twice to try to at least reduce the issue. So we're going to define B in the following way. It is essentially A minus A over A minus A.

So we pick some gamma. And the gamma in the proof has to be picked carefully so that things work out. I'm going to go into the proof, but you would have to pretend that there is a gamma so that will work. OK, so first of all, I restrict the denominators to being larger than delta to the gamma. In this way, well, I got rid of the worst possible case where the expansion is too much. But even then, this could escape the interval 0, 1. I don't want to escape the interval 0, 1. So I intersect it with the interval 0, 1.

So this B is going to play the role of X. By doing this, again, I'm avoiding two related but not exactly the same issues. One is that this could escape our reference interval. And the other is that the denominators could be too small. They can still be quite small because-- so gamma is not zero, so this can still be small. So there is a trade-off. If we allow two small denominators denominator, things blow up too much. But we still have to allow small denominators. Otherwise, we are missing too much of A minus A over A minus A. So this gamma gives us the trade off.

So that's one thing we do. But the most important thing to understand is, What do we do instead of this? So really, adding 1 doesn't really make sense in this setting. So what do we do? Here comes a very easy but absolutely critical lemma. So in the same way that the method in Fp doesn't work in Fp squared, what works in R-- this is a remark I should have made earlier, but I don't think we have the statement anymore.

We don't have the statement of Bourgain's projection theorem anymore. But Bourgain projection theorem is not true Over C, because over C, you could take X to be essentially R squared, leaving in C squared. You could take D to be a piece of R, leaving in C. Because R is a subring of C, you don't escape. So there is no growth. So it's exactly the same reason why it doesn't work in Fp squared. Because Fp squared has an intermediate algebraic structure. C has an intermediate algebraic structure which is R. And because of this, it doesn't work.

And here, we use that Fp doesn't have an intermediate algebraic structure. So actually, this lemma is very crucial in the sense that it is what distinguishes R from C or for many other things, but in particular, from C. So there are many ways in which R is different from C. We will see that it is the order structure, what distinguishes R from C in the lemma, in the proof of the lemma. Actually, the lemma can be proved in many ways, and some of them do not use the order structure. But in my proof, I will use the order structure.

OK, so here is the lemma. Suppose that B. Eventually, we are going to apply to this B. That's why I call it B. But a priori is any B, is contained in 0, 1. And suppose that 0 is in B. 0 is in this B because a1 could be equal to a2. And then let rho be the largest gap in B. What is the largest gap? So a gap is a connected component of the complement. So we are in R. So connected components are segments.

So the complement is a union of segments. Well, it could be points. I mean, it could be dense. But let's imagine B is closed. I think it doesn't have to be closed. Maybe let's say closed. It's not really important. If it's closed and the complement is a finite or countable union of open intervals. Each of them has a length, and we take the supremum. Everything is going to be finitary in our application. But let's say the largest in the sense of the supremum of the length of the gaps in B.

So is it clear what this is? We just look at the complementary intervals, and we look at the largest one. Then there exists B in B such that either the distance from B over 2 to B is at least rho over 4 or the distance from B plus 1 over 2 to B is at least rho over 4. So proof. Let B prime be B over 2 union B plus 1 over 2. This means what you expect it to mean. So you take little b and B divide it by 2. And you do that for every little b and B. And then you take little b plus 1 over 2 for every little b and B. So I hope it's clear what this is.

This is containing 0, 1 as well, because B is containing 0, 1. So this is contained in 0, 1/2. This is containing 1/2, 1. So we have 0. We have 1. We have 1/2. And 1/2 is in B prime because 0 is in B. You will see that this is important. So 0 is in B. So 1/2 is in B prime. Now what is the largest gap in B prime? So by definition, the largest gap in B is rho. So that's the definition of rho. So what is the largest gap in B prime?

**AUDIENCE:** Rho over 2.

**PABLO SHMERKIN:** Rho over 2. Here, the largest gap is rho over 2 because I'm scaling down by a factor of 2. And here, the largest gap is rho over 2 because I'm translating a scaling down by-- yes.

**AUDIENCE:** So we know actually there'll be a gap of size rho if B has a big gap at the ends of the intervals?

**PABLO SHMERKIN:** So I guess maybe let's assume that 1 is also in B or something like this. So I mean, there could be an infinite gap. Is that what you're worried about?

**AUDIENCE:** Yeah, there's a gap from 1 minus rho to 1 or something, but if 1 is in B then.

**PABLO SHMERKIN:** Well, let's say 0 and 1 are in B. So 1 is also here because you could take a3 and a4 to be a1 and a2. So 1 is in B. If you define the gap in a way that avoids this issue, you don't need 1 to be in B, but it's not important. Why is it important that 1/2 is in B? Because here, the largest gap has size rho over 2. Here, the largest gap has size rho over 2. But if 1/2 wasn't here, they could be combined to form a larger gap going across. But it cannot go across because 1/2 is there.

Now let's draw the largest gap or the gap that approximates the largest gap. So this one has size, let's say, rho minus epsilon. OK. So this is the center of the gap. So this is the gap in B, not in B prime. In B. Now let's draw an interval in the middle of length rho over 2. Maybe a closed interval. This closed interval contains an element of B prime, because the largest gap of B prime is rho over 2. That means that any interval of length rho over 2 contains a point in B prime.

So maybe let's write this down. So largest gap-- so this interval contains a point in B prime. But this point is at distance at least rho over 4 from B because this gap is double the size. So even if the point was at one of the endpoints-- so the point could be here. But even if it was one of the endpoints, this is rho over 4. This is rho over 4.

So maybe let's do a larger picture. So this is rho minus epsilon gap in B. This is a rho interval. So it contains some point in B prime. Because the largest gap in B prime is rho. And this interval has length rho. So it cannot be completely empty. Otherwise, it would be a larger gap. Whichever point is in this interval is at large distance from B because there is nothing in B here. So that's the proof. So you do the calculation is at least rho over 4. Maybe rho over 5, because you need to take into account this epsilon. Maybe rho over 5, just in case. It's not important.

So you see that this is a simple geometric argument, but it's very, very strongly uses the order structure of R. So this argument doesn't work in C because there is no order. In fact, it cannot work because what we are going to prove using this is not true over C. Any questions about this? You can prove this lemma in other ways. So both Katz and Zahl in their paper, I think they don't do this proof, but I think this proof is easier and shows how the order structure is fundamental.

Hopefully, I have proved the lemma. Of course, a point in B prime has either this form or this form because that's the definition of B prime. So here, we have a B prime in B prime, which is far from B But B prime is either this or this because that's the definition of B prime. So that's how one concludes the proof. Any questions again? OK, so now what? The idea is to use B over 2 or B plus 1 over 2 instead of B plus 1.

So why does B plus 1 work? So what is important about B plus 1 or X plus 1? So what is important is that, again, it's a polynomial. In this case, it's a linear polynomial that takes something in X and lands it not in X. Well, the same is true for B over 2 and B plus 1 over 2. You take something in B, and it lands far from B. It's a first step of proof, so I'm not going to prove. Well, we have to pick a rho in 0, 1 carefully.

In fact, it's going to be a power of delta as usual, some power. Don't remember which power. One has to choose this power carefully. And then there are two cases that morally correspond to those two cases. B now is that B. That B up there, not an arbitrary. That B, is rho dense. Rho dense means that the largest gap is at most rho. Or if you take the rho neighborhood, you cover everything. Everything is 0, 1. So any questions about what rho dense means?

Here, we use an argument, which is similar to the case X minus X over X. Sorry, it's not X. It's A minus A over A minus A equals Fp to get growth, which I'm not going to show. This is the easier case. I guess here, what's also the easier case. The other possibility is that there exists a B in B such that either B over 2 is far from B or B plus 1 over 2 is far from B. In the sense-- so there are two cases by the lemma.

And then we use over 2 or B plus 1 over 2 in a similar way to the way that X plus 1 was used in the finite field setting. To get expansion. It works. It's much more complicated. So the actual details are much more complicated than in the finite field setting. But it works. I would say the overall idea is very similar, but the details are much more complicated.

OK. So far, so good. Modulo, lots of details. But this was the first step in the finite field setting, and we can replicate it. So it seems that things are going well. But now we have a much bigger issue. Maybe I'm going to explain the much bigger issue and talk about how to overcome it next time, because this is maybe the biggest issue, the biggest difference. You can see that this was already a big issue and a big difference in the proof. But now comes, I would say, a bigger issue.

This was the first step in the finite field setting. What is the second step in the finite field setting? Iterating the first step. And you just iterate. Nobody is preventing you from iterating. So you get that Q of A growth. So you can do Q of Q of A. So here, maybe let's call it. The second step. You look at Q of Q of Q of Q of A. And this is very large. So you can make it as close to p as you want in terms of powers. I don't know. I don't know, eta.

So for any eta, if we iterate enough times, we keep growing, and we can reach any power we want. In fact, eventually, in hindsight, we get all of the finite field after iterating finitely many times. I don't know if you saw this, but anyway. So this is clear because we just keep iterating. So now let's go back to-- sorry, did I ever state-- oh, here it is.

Here is the lemma that I told you we can prove using this idea. So this is the lemma. So we have a delta s delta to the minus epsilon set of size delta to the minus s. Then we apply Q, and the size grows. So what's the problem? We would like to iterate, but we cannot. So why can we not iterate? Well, a priori we cannot. I'm not saying-- in the end, we will iterate. But why is it completely unclear whether we can iterate?

**AUDIENCE:** There's no guarantee that Q of A will satisfy the spacing condition?

**PABLO SHMERKIN:** Exactly. That's the issue. And it is a big issue. So we know that the size of Q of A grows. But here, we have this additional assumption and we need it, or we need some assumption. We need the assumption that A is a delta s delta to the minus epsilon set. So in order to iterate, we would need to know that not only Q of A has grown in size, but we would need to know that Q of is A delta, s plus epsilon delta to the minus epsilon set. And we don't know that a priori.

If we go over the proof, the proof doesn't show that. So when we-- so I haven't given the proof of the lemma, only gave a very rough sketch. But you have to trust me that if you go through the proof, we do not get any non-concentration for Q of A. We literally only get growth at scale delta. So let me write this somewhere.

Big issue. We know Q of A grows in delta-covering number, but the proof does not show that Q of A is delta s plus epsilon could be a different epsilon. That's what we did. The same epsilon, but we need some epsilon if we want to iterate here.

So we are going to talk about some of the things we do to

overcome this in the next lecture. But let me say that I think, in fact, it is not true that Q of A has to be a delta s plus epsilon delta to the minus epsilon set. It is not even true. But what is true is that Q of A contains delta s plus epsilon delta to the minus epsilon set.

This is good enough because we just take the subset and use that to iterate. And then we can iterate. And the second step will work in the same way, but this is a big issue. Somehow, we have to go through the whole steps of the proof in the finite field setting just to get this. And then we have to do the full-- so one has to cycle through all the steps twice in this proof.

The first cycle is to solve this issue. So we have to use one of so many hours, for example, to solve this issue. That comes much later in the proof in the finite field setting. And then we have to use it again once we have expanded a lot by iterating. So I'll explain the ideas behind this on Thursday. Yes.

**AUDIENCE:** After showing that Q of A contains a delta X plus epsilon delta minus epsilon set, does that set have to be kind of very close to the size of A, and unlike like an s plus epsilon or delta spacing level?

**PABLO SHMERKIN:** Well, it needs to be much-- something maybe I should have said, but I didn't say is that a delta TC-- ignore this for a moment. So delta TC set has size at least delta to the minus t. Because you apply the definition of delta TC set at scale delta. This is something I should have mentioned. I will try to remember to mention it next time. So this set that is contained actually has already size bigger than the size of A. But in addition to having larger size, it has good spacing.

But one question is that I'm going to answer next time is, How can we be sure that something contains? So this notion of delta s is set is bad in the sense that it is not monotone. So if you take a set which has this spacing condition or non-concentration condition and make it larger, you may lose the non-concentration condition. And that's kind of bad. So we'll see how to get around this next time. Any other questions? Yes.

**AUDIENCE:** I wonder in the two cases you wrote, so in the first case, we said that it's dense, so somehow it is large. But in the second case, what's the intuition behind, OK, I have one element that is far away from the set, and then we can get the answer?

**PABLO SHMERKIN:** It's really very similar to what you did in the finite field setting in the case when X is in A minus A over A minus A, and X plus 1 is not. You use that to show that some projection of the original set is large.

**AUDIENCE:** OK.

**PABLO SHMERKIN:** This is a very similar idea, except that you have to fight with small denominators. But morally, it's a very similar idea.

**AUDIENCE:** OK