

YUFEI ZHAO: OK. So let's get started. So we spent quite a bit of time with graph theory in the first part of this course, and today I want to move beyond that. So we're going to talk about more central topics and additive combinatorics, starting with the Fourier analytic proof of Roth's theorem.

We discussed Roth's theorem, and we gave a proof, during the course, using similarities, graph regularity lemma, as well as the triangle removal lemma. Today, I want to show you a different approach to proving Roth's theorem that goes through Fourier analysis. So this is a very important proof, and it's one of the main tools in additive combinatorics.

Let me remind you what Roth's theorem says. So Roth proved, in 1953, that if we write $r_3(n)$ to be the maximum size of a 3AP subset of 1 through n , then Roth showed that $r_3(n)$ is little O of n .

So in other words, if you have a positive density subset of the integers, then it must contain a three-term arithmetic progression then. So what I said is equivalent to the statement here. So previously, we gave a proof using regularity. Actually the regularity approach of Szemerédi was only found the '70s so Roth's original proof was through Fourier analysis.

And we'll see that tomorrow. Today, we'll see a toy version of this proof. But it's not really a toy version. It has the same ideas, but in a slightly easier setting that has fewer technicalities. But before showing you that, let me just discuss a bit of history around Roth's theorem. We will show, next time, the next lecture, we'll show the bound.

Also by regularity, we get some bound, which is little O then, but because of the use of regularity lemmas, it's pretty poor dependence. We got something like n over $\log^* n$. Next lecture, and basically Roth's original proof, gives you a bound which is n over $\log\text{-}\log n$. So it's a much more reasonable bound.

The current best upper bound known has the form, essentially, n over $\log n$ raised to 1 plus little 1 , roughly n over $\log n$. We do not know, or even have great guesses on, what the answer should be.

So the best lower bound, and this is a construction that we saw earlier in the course due to Behrend is of the form n over E to the c root $\log n$. It seems it may be very difficult to improve this upper bound without some genuine new ideas.

On the other hand, there is some evidence that the lower bound might be closer to the truth in that there are variants of the Roth problem for which we know that the lower bound is basically the truth. What I want to do today is look at a variant of this problem in what's called a finite field model.

And that basically just means we're going to be looking at Roth's theorem, not in the integers, but in some finite field vector space, specifically F_3 to the n . So we're going to define our sub- 3 of this F_3 to the n to be the maximum size of the $3AP$ subset of this finite field vector space.

So the finite field model is really useful. We're going to see this again later in the course as well. Many of the ideas and techniques that work for the real problem, so to speak, many of those techniques also work in the finite field model, but they are technically simpler to execute.

So this is often-- you view it as a sandbox, a playing ground, for testing out many of the ideas. And once you have those ideas, then you can see if you can bring them to the integer setting. And this is a very successful program, and we'll see one aspect of what happens when we do this.

For this specific problem of Roth's theorem, in F_3 to the n , there are some nice interpretations of what this problem means. So here's a pretty easy fact that for-- so $n \geq 3$ in F_3 to the n for three elements, x, y, z , the following interpretation, so what it means to be a $3AP$, are equivalent. So x, y, z form a $3AP$.

So $3AP$ means the y is x plus D . Z is x plus $2D$. Equivalently, they satisfy this equation, x minus $2y$ plus z equals zero. OK. In F_3 , minus 2 is plus 1 . So it's the same as this even nicer looking equation, x plus y plus z equals 2 .

It also turns out to be the same as saying that x, y, z would lie on a line. But they are aligned.

So the line has three points in F_3 . And if you look at the coordinate for every i , the i -th coordinate of x, y, z are all distinct or all equal. So easy to check all these things are equivalent to each other.

And the last one is a nice interpretation in terms of a game that many of you know, Set. So in the game Set, you have a bunch of cards. There have some number of properties, n properties, like color, the number of symbols, the shape. And you want to form a set being three cards such that every property, they're all same or all different. So that's exactly this

model over here.

So what can we say about this problem? What's the size of the maximum subset of F_3 to the n -th without 3-AP. If you look at the proof that we did earlier in this course, the one using triangle removal lemma, you see the proof works verbatim. Previously, we worked over $\mathbb{Z} \bmod n$. Now, you work over a different group, same proof.

So triangle removal lemma, it tells you that this r_3 is always little o of the size of the space. But we would like to do better. So this gives you something like $\log^* n$. It's not very good dependence. So we would like to do better.

So what we will show today, so this theorem is attributed to Meshulam. So in this case, the order of history is somewhat reversed. So we'll see the finite field toy model, but it historically actually came afterwards.

But you'll see that the Fourier analytic proof that we'll see today, it's basically the same proof in the two settings. So r_3 is F_3 , we will prove a bound, well, just like that. So much better than what you get from the regularity method.

In terms of-- OK, so let me tell you a bit more about the history of this problem in terms of what we know in terms of upper bounds and lower bounds. So let me say more about F_3 . So what's the best that you might hope for? So the best lower bound is due to Edel. And it's some construction, some very specific construction, which gives you a bound that's something like 2.21^n . And the upper bound is 3^n on the-- 3^n minus little 1 to the n -th.

So for a long time, it was open whether the answer should be basically roughly like 3^n to the n -th or some constant less than 3 to the n -th. And improvements on the upper bound were very slow, and/or some very difficult works that nudge that down below just a little bit.

And then a couple years ago, a few years ago, there was this incredible breakthrough, where in this paper that was just a couple pages long, they managed to significantly improve the upper bound to basically 2.76^n to the n -th. So this was an incredible breakthrough that happened just a few years ago. And we'll talk about this proof in a couple of lectures.

It turns out this proof, which uses what's now called the polynomial method-- so not Fourier analytic, but a different method-- unfortunately does not seem to generalize to the original Roth's theorem. In fact, you shouldn't expect it to generalize in a straightforward way, because up there you we know that you do not have a power saving, whereas here you have a power

saving. So the exponent goes down. OK, so this is roughly the history of this theorem. Any questions?

AUDIENCE: Do we have to have [INAUDIBLE]?

YUFEI ZHAO: Yeah. So I'll-- So I can tell you this is known as a Croot-Lev-Pach. So I'll say more about it in a couple lectures. But this for F_3 is due to Ellenberg and Gijswijt. So I'll tell you more about it in a couple lectures.

What I want to focus on today is the Fourier analytic nature of the proof that gives you this bound up there, 3 to the n over n . And it may seem like a completely different topic compared to what we've been doing so far in the course, which is more about graph theory. But I want you to think about what are the relationships between what we'll see today and what we've seen so far.

And there are lots of connections. So even though the proof may superficially look quite different, many of these ideas about quasirandomness versus structure will come up. And I want to present the proof in a way that highlights the similarities between what we did previously and this Fourier analytic proof.

So let's talk about the strategy. In the proof of the Szemerédi graph regularity lemma, we had the strategy that we called the energy increment strategy. So you start-- you want to find a good partition. You start doing partitioning. And you keep track of this thing called the energy-- must go up at every step, cannot go up forever, so has a bounded number of steps.

This strategy for Roth's theorem is also an important strategy. It's a variant of energy increment, but now density increment. So we start with a set-- A subset of F_3 to the n -th, and we would like to understand something about its structure versus pseudorandomness in a way that is similar to when we discussed the similar issue for graphs.

In particular, there will be this dichotomy that if A is in some sense pseudorandom-- so earlier, we saw what it means for a graph to be pseudorandom. So now, what does it mean for a subset of F_3 to the n -th to be pseudorandom. So we'll address that today.

If A pseudorandom-- OK, so the short answer is that it is Fourier uniform-- in other words, all Fourier coefficients small. That's what pseudorandom will refer to. So then there is a counting lemma. And the counting lemma will in particular imply that A has lots of 3 -APs. So then you

find your 3-AP.

If this is not the case, then-- so what's the opposite of Fourier uniform-- is that A now has some large Fourier coefficient. And what we'll do is to use this Fourier coefficient to extract some codimension 1 affine subspace-- it's also called a hyperplane-- where the density of A goes up significantly, if you restrict to that sub hyperplane.

And you can repeat this process. Now, restrict to this hyperplane and ask yourself the same question. Is A , when restricted to this hyperplane, pseudorandom? In which case, we find APs. Or is A restricted to this hyperplane, does it have a large Fourier coefficient? In which case, we restrict further.

And each time you iterate, you obtain a density increment. And the density increment cannot go on forever, because your total density is at most 1. So the number of steps must be bounded. So that's the strategy.

So this should remind you somewhat of the energy increment strategy from Szemerédi's regularity lemma, although there are some fundamental differences. We're not doing partitionings. Any questions about this strategy? OK.

I want to tell you about Fourier analysis. So probably, all of you have seen some version of Fourier analysis, maybe in your calculus class with Fourier series and whatnot. So you play with formulas, and solve some differential equations. So I want to give you more than just a bunch of ways about handling Fourier coefficients, a way to think about Fourier analysis. So think of this as a crash course about Fourier analysis from the perspective of combinatorics.

And Fourier analysis, I think, it's much easier if you work in a finite group, in a finite abelian group, which is what we're doing here. Many of the technicalities go away. So we'll be looking specifically at Fourier analysis in F_3 to the n -th, although the 3 can be any prime. So it's really the same.

So the main actors in Fourier analysis are the Fourier characters. The Fourier characters are denoted γ_r . And they're characters on the group, meaning that they are maps which-- so they turn out, happen to be homomorphisms for the multiplicative group under-- so \mathbb{C} under multiplication. And they're indexed by r , which also elements of F_3 to the n -th.

So I'm going to be fairly concrete here. There are ways to do this more abstractly. But I'll be fairly concrete. So it's defined by γ_r evaluated on x equals to ω raised to r dot

product x , where here ω is a third root of unity and the dot is a dot product. So--

So that's the definition of the Fourier transform-- sorry, that's the definition of the Fourier characters. And once you have the Fourier characters, you can have this Fourier transform, just defined as follows. If you start with a function-- let's say, a complex-valued function on your space-- then I define the Fourier transform to be another function, like that, defined by the following formula.

So that's the formula for the Fourier transform. It is basically the inner product between F and the Fourier character. So let me make the comment here, I think this is actually a pretty important comment, about the normalization.

Now, when you first learn Fourier transforms, usually in the reals, there are all these questions about what number to put in the exponent. Is it 2π ? Is it $\sqrt{2\pi}$? Is it some other thing? And somehow, one answer is better than the others. And the same thing is true here in groups.

So we'll stick with the following convention-- and I want all of you to stick with this convention, otherwise, we'll confuse ourselves to no end-- is that for a finite group-- actually, let me, start of a board. So the convention is that, in a finite group, the Fourier transform is defined-- and more generally, anything you do in the physical space, we always use the averaging measure. Don't sum, always average in the physical space. And in the frequency space, always use sums, use the counting measure.

Keep this in mind. Any of these questions about normalization, if you stick with this convention, things will become much easier. So there won't be any of these questions about when you take the inverse Fourier transform, do I put an extra factor in front or not, if you stick with this correct convention. So with that convention in mind, what the Fourier transform really is is inner product between F and a Fourier character.

There are some important properties of the Fourier transform. So let me go through a few of the key properties that we'll need. The first one is pretty easy. What is the meaning of the 0-th Fourier coefficient? You plug it in, and you see that it is just the average of F . So 0-th coefficient is the average of F .

The second fact goes under one of two names, and they're often used interchangeably-- Plancherel or Parseval. And it says that if you look at the inner product in the physical space, then this product is preserved, if you take the Fourier transform. But now, of course, you're in

the frequency space, so you should sum instead of doing the inner product.

So this identity can be proved in a fairly straightforward way by plugging in what the definition is for the Fourier transform. This is a straightforward computation I'm not going to do on the board, but I highly encourage you to actually do at home, just to do it once to make sure you understand how it goes.

But there is also a more conceptual way to understand this identity. And that's because-- now, this is also important to understand what the Fourier transform is. It's not just some magical formula somebody wrote down, like this is a very natural operation. It's because the characters, the set of characters, is an orthonormal basis. So the Fourier characters form an orthonormal basis.

As a result, what the Fourier transform is is a unitary change of basis. You can check. It's very straightforward to check that the Fourier characters, indeed, form an orthonormal basis, because-- well, you can evaluate the inner product between two Fourier characters. So remember, in the physical space, we're always doing averaging.

And so now, I'll just write down first what I mean by the inner product. So that's the inner product. And by the definition of the Fourier character, you have that. So think about what this expectation is-- unless r equals to s , in which case, this expectation is 1. Unless that is the case, you always have some coordinate of x in the exponent. So as you average over all possibilities, they average out to 0.

So this calculation shows you that the Fourier characters form an orthonormal basis. And a basic fact you know from linear algebra is that if you do a change of basis, if you do a unitary change of basis, then inner product is preserved. It's like a rotation. It's the same-- so you're not changing the inner product. So the inner product is preserved under this change of basis. And that's why Plancherel is true.

Another important thing is what's known as the Fourier inversion formula. The Fourier transform tells you how to go from a function to the Fourier transform. Well, now, if you are given the Fourier transform, how do you go back?

There's a formula which tells you that you can go back by the following formula there. So that's the Fourier inversion formula. It allows you to do this inversion.

And again, it's one of these formulas where I encourage you to try it out yourself by plugging in the formula and expanding. And it's pretty easy to check. It's much easier in the finite field setting, by the way.

So if you use the usual Fourier transform on the real line, there are some technicalities even to prove the Fourier inversion. But in finite groups, it's almost trivial. You expand, and then you'll see. So it's very easy to prove.

But you can also see this Fourier inversion formula more conceptually, because you're in a unitary change of basis. So to go back, well, think about what it means in linear algebra to revert a unitary transformation. You simply multiply the coefficients with the coordinates. Orthogonal, orthonormal change of basis.

Finally, Fourier transform behaves well under convolution. So by convolution, we define the convolution of two functions, f and g , using the following formula. And so then the claim is that the Fourier transform behaves very well under convolution. It's basically multiplicative under convolution.

So what this means is, if I put in-- so it's pointwise true everywhere. Again, very easy proof, because I just evaluate the left hand side, see what-- plug in the formula for the Fourier transform. And I find it's that.

And now, I plug in the formula for convolution. So now, you can do a change of variables. And then you-- it's not hard to see. You eventually end up at the right hand side.

So these are some of the properties. So there are important properties of the Fourier transform. So this is something that, whenever you learn about Fourier transform, you always see these few properties. And so we'll use them. But we'll also need another property that is specific to the analysis of 3-term arithmetic progressions.

So what does Fourier transform have to do with 3-APs? So we want to use it to prove Roth's theorem. So we better have some tool that allows us to analyze the number 3-APs. And here is a key identity relating Fourier with 3-APs. And it's that, if you have three functions, then the following quantity, which relates the number of 3-APs--

So this function basically counts the number 3-APs, if your f , g , and h are indicator functions of a set. I want to express this formula in terms of the Fourier transforms of these functions. The formula turns out to be fairly simple, that it is simply that. So it's a single sum over the r 's of f

hat of r , g hat of ω , and h hat of r .

You might wonder why I put a ω here, because ω is r , and it looks certainly much nicer with just r in there. And that is true. This formula as written is true for over any group. And our proof will show it. So it's not really about F_3 at all, but any group.

So let me prove this for you in a couple of different ways. So the first proof is basically a straightforward no thinking involved proof, as in we apply these formula for using either Fourier inversion or the inverse Fourier transform, and plug it in, and expand, and check. So it's worth doing at least this once. So let's do this together at least once. But this something that is a fairly straightforward computation.

The left hand side can be expanded using Fourier inversion. so $r_1 f$ hat $r_1 \omega$ to the minus $r_1 \cdot x$, and sum over $r_2 g$ hat $r_2 \omega$ to the minus $r_2 \cdot x$ plus y , and then finally sum over $r_3 h$ hat of $r_3 \omega$ to the minus $r_3 \cdot x$ plus $2y$. So I'm using Fourier inversion, replace f , g , and h by their Fourier transforms. Oh, sorry, there should be-- yeah, so no minus.

So now, we exchange sums and expectations, do a switch in the order of summation, so r_1 , r_2 , r_3 and f hat of $r_1 g$ hat of $r_2 h$ hat of r_3 . And you have this expectation over x and y and ω of $x \cdot r_1$ plus r_2 plus r_3 . In fact, I can even write the x and y separately, $y \omega$ to the $y \cdot r_2$ plus $2r_3$. So just rearranging.

And now, you see that as you take expectation over x , this expectation is equal to 1, if r_1 plus r_2 plus r_3 is equal to 0, and 0 otherwise. And likewise, the third expectation is either 1 or 0, depending on the sums of r_2 and r_3 . So the only terms that remain, after you take out these 0's, are cases where both these two equations are satisfied.

And then you see that the only remaining terms are basically the ones given in the sum on the right hand side. OK? So that's the proof. Pretty straightforward, you plug in Fourier inversion.

I want to show you a different proof that hopefully will be more familiar and more conceptual. Now, it doesn't involve carrying through this calculation, even though this is not at all hard calculation. But first, let me rewrite the formula up there.

So in F_3 , it will be convenient, and so the formula is actually slightly easier to interpreting F_3 , in F_3 , the identity says that, if you look at the quantity-- I need. So let me give you a second proof that works just in F_3 , but you can modify it to work in other groups. But in F_3 , it's

particularly nice.

The left hand side, you see, the left hand side, I can rewrite it as the following form, where I sum over-- well, I take expectation over all triples x, y, z that sum to 0. Because a 3-AP is the same as three elements, three points in the vector space summing to 0. But now, you see that this quantity is the same as the convolution evaluated at 0-- so if you extend the definition of convolution to more than one func-- more than two functions.

But now, we apply Fourier inversion. And we find that-- OK. So by Fourier inversion, you have that. But now, by the identity that relates the Fourier transform and inversion, you have that. And that's the proof, because minus $2r$ is the same as r . So it's shorter, because we're using some properties here about convolution and-- yeah, so about the convolution.

That formula up there is, of course, related to counting 3-APs. Because if $f, g,$ and h are all indicators of some set, then the left hand side is the same as basically the number of triples of elements in A whose sum is equal to 0. And the right hand side is the sum of the third power of the Fourier coefficients.

And this formula should look somewhat familiar, because we also used this kind of formula back when we discussed spectral graph theory. And remember, the third moment of the eigenvalues is the trace of the third power, which counts closed walks in Cayley graph. So this is actually the same formula. So in the case if A is symmetric, let's say, then this is the same as the formula that counts closed walks of length three in the Cayley graph.

The point of this comment is just to tell you that Fourier transform is somehow it's not this brand new concept that we've never seen before. It is intimately tied to many of the things that we have seen earlier in this course but in disguise. So it is related to the spectral graph theory that we discussed at length earlier in this course.

Now that we have the Fourier transform, I want to develop some machinery to prove Roth's theorem following the strategy up there. So let's take a quick break. And then when we come back, we'll prove Roth's theorem.

Any questions so far?

AUDIENCE:

So for this one, you said-- is it like we only used the fact that it's F_3 to the n -th at the end of the proof, like in that last step?

YUFEI ZHAO: OK. So question is, where do we use that in F_3 to the n -th? This formula here holds in every finite abelian group, if you use the correct definition of Fourier transform with the averaging normalization. So in the other formula where you replace minus 2 by 1, that requires F_3 .

But you can-- I mean, you can follow the proof and come up with a similar formula for every equation. So there's a general principle here, which I'll discuss more at length in a bit, that for patterns that are governed by a single equation-- in this case, 3-APs, x minus $2y$ plus z equal to 0-- patterns that can be governed by a single equation can be controlled by a Fourier transform.

So let's begin our proof, the Fourier analytic proof of Roth's theorem in F_3 to the n -th.

AUDIENCE: So at the end, you said it was going to be connected with counting the [INAUDIBLE] graph. Does this mean that the Fourier transform of the indicator of A , those are exactly the eigenvalues of the Cayley graph? Or is it like [INAUDIBLE]?

YUFEI ZHAO: OK, so you're asking about the final step, where we're talking about-- so I mentioned that there was this connection between counting walks in graphs and spectral graph theory. So you can check that, if you have a subset A of an abelian group, then the Fourier transforms of A are exactly the eigenvalues of the Cayley graph.

AUDIENCE: So then I guess, have we done anything so far that could have been done in a spectral way yet? Well, I guess, where is the Fourier analysis better than the spectral [INAUDIBLE]?

YUFEI ZHAO: OK. Question, where is the Fourier analysis better than the spectral posts? Well, let's see the proof first. And then you'll see, yeah. So there's no graphs anymore. So we're going to work inside F_3 to the n -th.

But just like the proof of regularity in counting, we're going to have a counting lemma. So all of these are analytic. And at this point, they should be very familiar to you. They may come in a different form. They may be dressed in different clothing. But it's still a counting lemma. So let's see.

The counting lemma, in this case, says that if you are in the setting of A in F_3 to the n -th-- and I'm going to throughout write the density of A as α , then let me write, let me define this λ_3 of A to be the function which basically counts 3-APs in A but with the averaging normalization. So this is-- we saw this earlier.

So the counting lemma says that this normalized number of 3-APs in A -- so including trivial 3-APs; that's why this is a nice analytic expression-- differs from what you might guess based on density alone. This difference should be small if all the nonzero Fourier coefficients of A are small.

So in this strategy, I said that if-- so the counting lemma tells you, if A is Fourier uniform, then it is pseudorandom. And this is where it comes in. If A has all small Fourier coefficients, then you have a counting lemma, which tells you that the counts of A should not be so different from the guess based on density.

So the proof is very short. It's based on the identity that we saw earlier. The 3-AP count of A by the identity earlier is simply the third power of the Fourier transforms. And all of these calculations should be reminiscent, because we've done these kind of calculations in some form or another earlier in this course. So we're going to separate out the main term and subsequent terms.

So the main term is the one corresponding to r equals to 0. So that's the density. And all the other terms I'm going to lump together into this sum. So we now know that the difference we're trying to bound is upper bounded by the third moment of the absolute values of the Fourier transform. And I want to upper bound this quantity here, assuming that all of these Fourier coefficients are small.

We've also done this kind of calculations before. So where have we seen this before? We saw this calculation earlier in the class with the 3 replaced by a 4. So in counting four cycles in a proof equivalent of quasirandomness, we said that, if all the eigenvalues other than the top one are small, then you can count four cycles. It's the same proof.

And remember, in that proof, there was a important trick, where you do not uniformly bound each term by the max, because then you lose. You lose by an extra factor of n that you don't want. So you only take out one factor.

So you take out one factor. And you keep the rest in there. In fact, I can be more generous and even throw the r equal to 0 term back in.

And now, by Plancherel-- so by Plancherel/Parseval, this here is equal to the expectation of the indicator function of A squared. So you take this-- you go back to physical space, and that's simply the density. So then that proves the theorem.

So the moral of the counting lemma is the same as the one that we've seen before when we discussed graphs. If you're in pseudorandom, then you have good counting. And here, pseudorandom means having small Fourier coefficients, uniformly small Fourier coefficients.

So now, let's begin the proof of Roth's theorem. The Roth's theorem proof will have three steps. The first step, we will observe that if you're in the 3-AP-free set, then there exists a large Fourier coefficient.

Throughout, I'm going to use uppercase N to denote the size of the ambient group. And specifically, we will prove the following. And also throughout, A is a subset of F_3 to the n -th and with density α . So I'll keep this convention throughout this proof.

We will show that if A is 3-AP-free and N is at least 2 to the α to the minus 2 -- so N is at least somewhat large-- then there exists a nonzero r such that the r -th Fourier coefficient is at least α squared over 2 . If you're 3-AP-free, free then provided that you're working in a large enough ambient space, you always have some large Fourier coefficient.

So the proof is essentially-- well, this claim is essentially a corollary of counting lemma, by the counting lemma. And using the fact that in a 3-AP-free set, what is the quantity λ_3 of A ? Up there, you only have the trivial 3-APs present. So this quantity λ_3 must then be the size of A divided by N squared or α over N , which are precisely counting the trivial 3-APs.

So by the counting lemma, then we have that the upper bound on the right hand side, which we now write $\alpha \max$ over nonzero r 's, is at least α cubed minus the λ_3 term, which should be α over N . So provided that N is large enough-- big N is large enough-- then the trivial 3-APs should not contribute very much. So I can lower bound the right hand side by, let's say, α^3 -- α cubed over 2 . So then you deduce the conclusion.

I want you to think about how this proof is related to Szemerédi's graph regularity lemma. The analogy will break down at some point. But we've seen this step before as well. From lack of 3-APs, you extract some useful information and from this will extract some structure.

And the structure here-- and this is where the proof now diverges from that of regularity-- having a large Fourier coefficient will now imply a density increment on a hyperplane. Specifically, if you have-- so keeping the same convention as before, if the Fourier coefficient of A at r is at least δ for some nonzero r , then A has density at least $\alpha + \delta$ over 2 when restricted to a hyperplane.

So if you have a large Fourier coefficient, then I can pass down to a smaller part of the space where the density of A goes up significantly. To see why this is true, let's go back to the definition of the Fourier coefficient, the Fourier transform. So recall that Fourier transform is given by the following formula, where I'm looking at this expectation over points in F_3 to the n -th together with indicator of A multiplied by this Fourier character.

And you see that this function here, it is constant on cosets of the hyperplane defined by the orthogonal complement of r . So the value of this dot product is this constant on the three hyperplanes. So I can rewrite this expectation simply as $1/3$ of α_0 plus $\alpha_1 \omega$ plus $\alpha_2 \omega^2$, where α_0 , α_1 , α_2 are the densities of A on the three cosets of r perp. So group this expectation into these three hyperplanes.

So now, you see that if this guy is large, then you should expect that α_0 , α_1 , and α_2 are not all too close to each other. So if they were all equal to each other, you should get 0. But you should not expect them to be too close to each other.

In particular, we would want to say that one of them goes up-- is much bigger than α . So one of these, they must be much bigger than α . That's an elementary inequality.

This is something that I'm sure I give you five minutes you can figure out. But let me show you a small trick to show this. And the reason for this trick is because in next lecture, when we look at Roth's theorem over the integers, we'll need this extra trick.

And the trick here is this. We now know that because of the hypothesis, 3δ is lower bound to the absolute value of α_0 plus $\alpha_1 \omega$ plus $\alpha_2 \omega^2$. OK, so note here that the average of the three alphas is equal to the original α by definition of density.

So this inner sum I can read write like that. So the sum of the three groups of unity add up to 0. And now, I apply triangle inequality to extract the terms.

So now, you should already deduce that one of the alphas i 's has to be significantly larger than α . And has to be significantly different, but there are only three times. One of them has to be significantly larger.

But let me do this one extra trick, which we'll need next time, which is that let me add an extra term like that, which sums out to 0. But now, you see that each summand is always nonnegative. So one of the-- so there exists some j such that δ lower bounds the j -th

summand.

And if you look at what that means, then α lower bounds the j -th summand. So in particular, this sum here should be nonnegative. So you have that. Good. So we obtained a density increment of this hyperplane.

And finally, I want to iterate this density increment. So I want to iterate this density increment-- so summary so far is that we have-- so if A is 3-AP-free with density α and N at least $2/\alpha^2$, then A has density at least $\alpha + \alpha^2/4$ on some hyperplane. So you combine step one and step two, we obtain this conclusion.

Well, I can now repeat this operation. So I can repeat by restricting A to this hyperplane. If A is originally 3-AP-free, I restrict it to a hyperplane, it's still 3-AP-free. So I can keep going.

I can keep going provided that my space is still large enough, because I still need this lower bound on F . So don't forget this one here. So I can keep going as long as N is-- so I'm using N_j to denote what happens after the j -th step. I can keep going as long as this is still satisfied.

But of course, you cannot keep on going forever, because the density is bounded. So density cannot exceed 1. So these two will give you a bound on the total dimension. So let's work this out.

So let α_i denote the density after step i in this iteration. And we see from over here that you start with density α , and each step you go up by an increment, which is basically what-- so you go up by some increment. And you want to know, if you start with α , how many steps at most can you take before you blow out 1.

So can you give me some bound? So what's the maximum-- at most, how many steps? So we know that the density cannot exceed 1.

AUDIENCE: $4/\alpha^2$.

YUFEI ZHAO: So you see that you have at most $4/\alpha^2$ steps, because density is at most 1. And if you plug this in, you get something which is not quite what I stated. So it turns out that if you plug this in, you find that α is-- you find that the size of A is at most $3/\alpha^2$.

So let me do a little bit better, so then simply seeing that this term here is at least α^2 over 4. And the point is that when you increment, you increment faster and faster. So I can use that to give a better bound on the number of steps.

And here's the way to see it. So let me-- we can do better. So starting at α , I then now ask, how many steps do you need to take before it doubles?

AUDIENCE: [INAUDIBLE]

YUFEI ZHAO: It goes up by α^2 over 4. So it doubles after at most $4/\alpha$ steps. And at which point this α , new α , becomes at least the original-- twice the original α ? But now, you keep going. How many times does it take to double again? $2/\alpha$, because the α became twice as much.

So it doubles again after at most $2/\alpha$ steps. And then you keep going. The next iteration is $1/\alpha$. So in total-- so you see that we must stop after at most $8/\alpha$ steps. So the number of times it doubles, actually it decreases by at least half each time.

So now, we know that-- we see that the-- so you keep on going. So you must stop after at most $8/\alpha$ steps. What is the final density when you have to stop, because when are you forced to stop? You are forced to stop if you run out of space. So you're forced to stop when you run out of space.

So if the process terminates after m steps-- so we're at density α^m , so then the final subspace has size less than $2^m \alpha^m$, which is-- So now, I use this bound α . So the initial N is upper bounded by what? So how many steps did you take? You took at most $8/\alpha$ steps.

Each of those steps, you pass down to codimension what? You lose a dimension for each step. And the final subspace has at least this-- has at most that much space. So the final dimension is this, basically $\log 1/\alpha$. So put them together, we see that the size of the space originally is at most $1/\alpha$. Yeah.

AUDIENCE: Should this be a lower case n ?

YUFEI ZHAO: Thank you. Yeah. This should be a lower case n , so the dimension. Good. And OK, so then that's the conclusion, that the density α is big O of $1/n$.

That proves the main theorem for today, so Roth's theorem over F_3 to the n -th. So we went through this Fourier analytic proof. Next lecture, we will see the same proof again but done in the integers for interval.

And there, there are some difficulties that we don't see over here. Because in the finite field space, in the finite field model, there's this very nice idea of looking at subspaces, so looking at hyperplanes. Each Fourier coefficient gets you down to one dimension less. But when you're working in the integers, there are no subspaces you can use. So we'll be looking at ways to get around the lack of subspaces.

And this is why I said in the beginning that the finite field model is often a very good playground for additive combinatorics type techniques, especially Fourier analytic techniques. Because in the additive-- in all of these techniques, they just come out to be much cleaner. If you're working in a finite field setting, you have nice subspaces, you have Fourier transform in a very clean way.

The Fourier transform always takes, in this case, one of three values. Everything's very clean. Everything's very simple. And you get to see the idea here. You get to see the sense of the increment argument.

But once you understand those ideas and you're willing to do more work, then oftentimes, you can bring those ideas to other settings, to other abelian groups, to the integers, for instance, but with more work in the-- there are some extra ingredients that you need to use.

I mentioned that there was a bound-- OK, so initially-- So next time, we'll see that. Next time, we'll see what happens over the integers. Any questions? Yes.

AUDIENCE: [INAUDIBLE]

YUFEI ZHAO: OK, great. So question is why the process must stop after at most $8/\alpha$ steps? So you know that the density doubles after this many steps, doubles again after that many steps. So eventually, if it keeps on doubling, it cannot keep on doubling forever.

So this process cannot keep on doubling forever. So it must stop-- so cannot double more than $\log_2(1/\alpha)$ times. And that point, you have to stop. So how many steps have you taken? Well, you sum this geometric series.

So this-- and the next thing is that you sum this geometric series. And that geometric series

sums to 8 over α . Great. So let's finish here. So next time, we'll see Roth's proof of Roth's theorem.