YUFEI ZHAO: All right. We are going to continue our discussion of extremal graph theory. Last time, we discussed what happens when we exclude a triangle or more generally a clique. And we wish to find a graph that maximizes the number of edges. And at the end of the lecture, I stated the theorem of Erdos-Stone-Simonovitz, which says, recall, that if you have a fixed H and I wish to understand the maximum number of edges in an n vertex graph that is H-free-- so remember this definition, this is the maximum number of edges in an n vertex H-free graph. So we are going to be looking at this quantity in the next few lectures.

So the Erdos-Stone-Simonovitz theorem tells us that, perhaps, quite surprisingly, this quantity is largely covered by the chromatic number of H, even though H itself might be quite involved. So if you knew the chromatic number, you already know a lot of information about the growth rate of this function. And in particular, as long as it's just not bipartite, so that the chromatic numbers at least 3, we already know the first order asymptotics from the Erdos-Stone-Simonovitz theorem.

However, when H is bipartite so that the chromatic number is exactly 2, then the theorem tells us only that this quantity is little o of n squared. Which is some useful information, but it doesn't tell us the whole story. And in the next several lecturers, I want to explore what more can we say about this quantity here for bipartite graphs H. And it turns out that there is a lot that we do not know, that there are lots of open problems in this area having to do with trying to pin down the growth rate of this function.

And, in particular, for bipartite graphs, there's a bipartite graph that places somewhat special rule, namely the complete bipartite graph. So K st, being the complete bipartite graph, with s vertices on one side and t vertices on the other side-- so this is a very nice bipartite graph. And just to understand, the extremal number for this graph is a famous open problem in this area, and it has the name of Zarankiewicz problem, which is to determine or estimate the extremal numbers for these complete bipartite graphs.

So I'll tell you pretty much all we know about this problem. And there are some interesting things. But we do not know all that much.

Now, every bipartite graph is a subset. It's a subgraph of such a complete bipartite graph. So every bipartite H it is a soft graph of some K st. And we know that if H is a subgraph of this K st, then the extremal number for H if your graph is H-free, then automatically it has to be K s,t-free. So there is this bound between these two extremal numbers.

So in particular, if you have some upper bound on K st, then you have some upper bound on bipartite graphs. Although, for specific by bipartite graphs H, maybe we can do better than using this bound. And we'll see examples of that later in the course as well.

So what can we say about the extremal numbers of these K s,t's? So the most important theorem in this area of this problem is the result due to Kovari-Sos-Turan. So the Kovari-Sos-Turan theorem tells us that for every fixed integers s and t-- s at most t-- there exists some constant C, such that the extremal number is upper-bounded by something which is on the order of n to the 2 minus 1 over s. So it gives you some upper bound, showing you that it's not only subquadratic, but there is a gap from 2 in the exponent.

Now, in combinatorics, as is with many fields of mathematics, it can be somewhat intimidating to newcomers when you see a lot of names. So every theorems seems to be named after a string of mathematicians, some of whom you may have heard of, some you may not. And I agree, it may be difficult to remember who is who and what is what, but I think this theorem has a very nice way to remember, which is that it's a theorem about K st, and this is the K s,t theorem.

[LAUGHTER]

OK. So this is the K s,t theorem. I want to begin by showing you how to prove the K s,t The proof is via a nice and not too difficult double-counting argument. And I show you some applications.

Let's prove this theorem, and it's by double-counting argument. What are we going to count? Well, let's start with the object that we're working with. So there's going to be a graph G that has n vertices, m edges, and K s,t-free. And let's count the number of stars, K s,1 in this graph G. So we're counting configurations like that.

I'll do an upper bound and a lower bound. So let's start with the upper bound. On one hand, every subset of s vertices in the graph G has at most t minus 1 common neighbors. Because if they had t common neighbors, then you get a K st. So that's one down.

On the other hand, let's see what happens to the number of common neighbors if you knew that this graph has a lot of edges. So the number of stars-- well, I can calculate this quantity explicitly by running over all the vertices of G. For each vertex, look at its neighborhood and choose s vertices from its neighborhood. So for each v I need to find a subset of s vertices from its neighborhood.

By convexity, I can lower bound the sum by the average, in some sense, because-- let me first write down the expression. So I can do a convexity argument that gives me this lower bound. Here, I'm abusing notation somewhat and writing binomial coefficients with a real entry on top, where I mean the expression as you would expect, treating this guy as a polynomial in x. And the key fact we're using here is that this function here is convex for x at least and minus 1.

So, In particular, if you think about this function here, you have a lot of 0's, and then it becomes convex afterwards. So you can even think of extending this function as 0 to the left of n minus 1. So this is a convex function, and you can apply convexity to deduce this inequality.

But you know the sum of the degrees. That's just essentially twice-- I mean, that is the twice the number of edges. So we have this expression right here. So you have an upper bound on the number of s stars coming from K s,t-freeness and lower bound on the number of s stars coming from just having lots of edges and applying convexity.

Putting these two things together, we find that there's this inequality here. Here, we are thinking of s and t as fixed, and we are trying to understand how m and n depend on each other as they get large. So it will be helpful to use the asymptotics that n choose s grows like n to the s divided by s factorial for a fixed s. So looking at that expression and applying this asymptotics to both sides, we have this inequality here, So I've eliminated s factorial from both sides.

So now rearrange, clean things up. We find the following upper bound-- m the number of edges in G. And that's the expression. So for fixed s and t it grows like n to the 2 minus 1 over s. Any questions? Yeah.

AUDIENCE: Where does the right side of the inequality come from? Are you counting that from the different cycles [INAUDIBLE]?

YUFEI ZHAO: So the question is where the right side of inequality-- which inequality? This one here?

AUDIENCE: Yeah.

YUFEI ZHAO: Right. So here we are counting the number of K s1's.

AUDIENCE: Oh, OK.

YUFEI ZHAO: So here I am upper and lower-bounding the number of K s1's using what we derived earlier. Yes.

AUDIENCE: Do you actually care that s is less than or equal t in the argument?

YUFEI ZHAO: The question is, do we care that s is less than or equal to t in the argument? The argument doesn't care. But, of course, if you want the better asymptotics for fixed s and t as n gets large, you should take s to be less than or equal to t. Question?

AUDIENCE: What happens when t equals 1?

YUFEI ZHAO: What happens when t equals to 1. Well, that's a great question. I'll leave you to think about it. If you know that your graph has maximum degree n most t, what can you tell me about the number of edges in the graph?

[LAUGHTER]

OK. Any more questions? We'll come back to this theorem. In fact, this will occupy us for at least a couple of lectures. Basically, is this theorem tight? And it is conjectured to be, although that is a major open problem extremal graph theory. We only know a small number of values.

Well, for most values of s and t, in particular when s and t are both equal to 4, we do not know if this bound is tight. But there are some values of s and t for which we do know that it is tight. For example, 2 and 2, 3 and 3, 4 and 7, s and if t is really, really large. And I will show you some constructions later on, creating graphs G that are K s,t-free for those parameters that matches this K st bound up to a constant factor. Yes?

AUDIENCE: For a fixed s, what's the bound on t-- if there a bound on t for t with equality cases?

YUFEI ZHAO: Right. So the question is, for a fixed value of s, is there some bound on t for which we get equality cases? There is a conjecture that this bound is sharp up to a constant factor for every s and t. But we only know how to prove that conjecture-- and I will tell you much more about it later on-- when t is much larger compared to s. So there is a lot of unexplored territory on this problem. Any more questions?

Before diving into the Kovari-Sos-Turan further, I just want to show you some neat applications. And I want to begin with a geometric application. There is a classic problem asked by Erdos, back in the '40s, called the unit distance problem. Which asks, what is the maximum number of unit distances formed by n points in the plane.

Let me give you some examples. If you have three points, you put them in equilateral triangle, and all three distances are unit. Great. If you have four points, you cannot place them so that all six distances are units if you're staying in the plane. So the best thing you can do is something like this-- so I'm drawing all the edges that are unit distances.

If you have one more point, it turns out that's the best thing you can do with five points. With six points there are some more possibilities. Let me draw for you some possibilities. You can extend the previous configuration by adding more triangles. And there are many different ways that you can attach an extra triangle.

There's actually one more way to do it with six points. Namely, I can put them like a projection of a prism. So all of these configurations have the maximum number of unit distances obtained if you draw six points.

If you draw seven points, it turns out this is the best way to do it. So you can go on for a while. And people have. So you can try to tabulate for every n what's the maximum number of unit distances you can generate having n points in the plane.

And the question is, what is the answer for n points? And it turns out, for this problem, and for many problems like it in combinatorics, you can have a lot of fun with playing with small examples, but they are often misleading. It doesn't really tell you what the overall structure should be like. And it turns out this is the major opium problem for which we do not understand what the structure is like for large values of n.

So think n large. What are some possible ways to generate many unit distances? Yep.

AUDIENCE: Drawing triangles?

YUFEI ZHAO: Great. So one way is to draw lots of triangles. So extend this figure forward. So let's do that. Well, actually, let me give you something even simpler to begin with. I can just put the n points on a line, equal spaced, OK, I get n minus 1 distances-- n minus 1 unit distances.

If I put them like that, how many unit distances do I get?

AUDIENCE: 2 times n minus 1.

YUFEI ZHAO: Good. 2 times n minus 1. So each new point gives you two extra new distances. And you can try to think about how to do better. Well, if you follow that sequence of examples, maybe you're limited to such ideas. Yes?

AUDIENCE: Are we allowed to put points in the same place?

YUFEI ZHAO: You are not allowed to put points in the same place. That's a great question. But you're not allowed to put points on the same place. Yes?

AUDIENCE: Is for something to keep doubling the number of points, the degrees get big?

YUFEI ZHAO: So I didn't quite understand-- so you say if you put--

AUDIENCE: [INAUDIBLE] something.

YUFEI ZHAO: Yeah.

AUDIENCE: Like keep translating in different directions.

YUFEI ZHAO: So you want to take this example and keep translating in some direction I think you'll run out of room pretty quickly.

AUDIENCE: No, I mean just copy it and put it over [INAUDIBLE].

YUFEI ZHAO: I see. You want to take this configuration and translate it in some unit direction. Well, then you double the number of points, but you don't actually increase the number of units distances all that much.

AUDIENCE: Don't we get each point having one extra unit distance added?

YUFEI ZHAO: So the suggestion is we take some configuration, let's say this graph G, and I form two copies of G by translating G in some unit directions-- some generic unit direction. OK, great. So what happens to the number of vertices? So n goes to 2m and the number of edges goes from m to 2m plus n. Is that right?

OK, great. So if you do that, what do you get?

AUDIENCE: Log n.

YUFEI ZHAO: n log. OK. That's much better than before. OK, good. Good. Very nice. Yes?

AUDIENCE: That picture [INAUDIBLE]. Let's start with the construction for n over 3 and then replace each point with a triangles-- equilateral triangle.

YUFEI ZHAO: Right. So the suggestion is to start with some graph G and then replacing-- so let's, as an example, look at that one-- each vertex here by an equilateral triangle. But I want to maintain the same unit distance. So the-- uh-huh.

AUDIENCE: So you just choose [INAUDIBLE]. Or you choose [INAUDIBLE]

YUFEI ZHAO: Is this similar to taking this graph G and then translating it in two different directions that form an equilateral triangle? I think it's-- yeah. So it's a very similar idea. And maybe you do a little bit better in terms of constant. All great suggestions. Actually, this is really nice.

AUDIENCE: [INAUDIBLE] two times [INAUDIBLE].

YUFEI ZHAO: Two times-- you want to make a constant correction? OK, let me just do that.

[LAUGHTER]

Any more suggestions? OK. Yeah, all of this is really nice. So let me tell you what is the best construction that people are aware of, and this is A construction due to Erdos. And the idea is to think big, not build from small examples.

So what Erdos did is to consider a square grid of root n by root n. When I see something that's root n that's non-integer, I just think round down to the nearest integer. I have a square grid. Now, if you take these distances as unit distances, well, you get something which is linear in m. So you don't gain that much.

But you can take any specific distance as your unit distance. So what we can do is take a distance that is represented many times in this grid. So let's take the unit distance to be a distance root r where r is some integer that can be represented as a sum of two squares in many different ways.

So, for example, if we-- so we can take some r so that it has many appearances. And if you know some elementary number theory, then you might know that the best way to do this is to take r to be a product of primes that are 1 mod 4. In any case, you can do this and you can use some analytic number theory to calculate if you choose the best possible value of r, how many distances do you get. So that calculation was done, and it turns out you get n raised to the power of 1 plus some constant C over log log n unit distances. So this is better than the constructions we've seen before.

So what can we say about this problem? So this is a construction. Well, then you want to understand some upper bounds. What can we say about the upper bounds to this problems?

And let me show you a fairly easy upper bound, which can be deduced very quickly from the Kovari-Sos-Turan theorem, that every set of n points in the plane has at most on the order off n to the 3/2 unit distances. So not quite this bound, but it's some bound. So the trivial upper bound is n choose 2. So it's much better than the trivial upper bound.

So here's the proof. So let's consider the unit distance graph, which is basically the graphs I've been drawing, where you have the vertices, the points, and I join an edge between two vertices if and only if their distance is exactly 1. I claim this graph is K 2,3 free

Why is that? Because if you have two points, what are their common neighbors? Their common chambers must all be at distance 1 from each of the two points. So their common neighbors must land in the intersections of the unit circles centered at these two points, and they meet at most two points. So it's K 2,3 free.

Therefore, the Kovari-Sos-Turan theorem tells us that the number of edges in this G is upper bounded by n to the 3/2. And that's it. That gives you some upper bound. Any questions?

So it's an application of Kovari-Sos-Turan, where we design the graph that has to be K 2,3, free and the extremal number tells us some information about this graph. What is the best bound that we know? So it turns out we do know how to do a little bit better, but not anywhere close to what we believe to be the truth.

So the current best bound is on the order of n to the 4/3. And we'll actually see a proof of this later in the course, once we've developed something called the crossing numbers inequality. It's also a very short, very neat proof.

I want to tell you about another problem which looks superficially similar and also a really interesting geometric problem. And this is the Erdos distinct distances problem, which asks for the maximum number of distinct distances among n points in the plane. The minimum-- thank you. The minimum number of distinct distances. So maximum would be very easy.

Again, you have lots of examples. In fact, you can look at basically all the examples that we've given earlier I mean, these two problems are very much related to each other. If you want lots of repeat distances or if you want very few distinct distances, whether they seem like they are very much related. And, of course, you can also write down the inequality that relates the two of them, because the number of distinct distances is at least n choose 2 divided by the maximum number of unit distances-- in other words, the number of distances that can be repeated in any single configuration.

So each of the examples that I just erased-- so, for example, if you put points on the line, you have n minus 1 distinct distances. And we already saw the square grid, so that might be a good candidate to look at as well. And there it takes some more number theory to figure out how many distinct distances there are, and this roughly corresponds to the number of integers that can be written as the sum of two squares.

So this has been calculated, and the result is n divided by square root of log n. So that's the order. And Erdos conjectured that this example should be more or less the best you can do.

So these problems look very similar. But, actually, this problem here was solved in a spectacular fashion about a decade ago by an important paper of Guth and Katz. So Larry Guth is in our department. And they showed that every set of n points in the plane generates at least on the order of n over log n distinct distances.

So not quite what Erdos conjectured, but nearly there. And, in fact, all the previous results were much worse. So they were off in the exponent. But this one more or less got to the truth.

So this is a very sophisticated result that used lots of amazing techniques, ranging from the polynomial method in combinatorics to some algebraic geometry. And that's-- I just wanted to mention it for cultural value. OK. Any questions? Yes.

AUDIENCE: For the unit distance problem, what is believed to be the bound?

YUFEI ZHAO: Yeah. So for the unit distance problem, and actually for the distinct distances problem, the question is, what was is believed to be the bound? And it's the square grid. So maybe you can do slightly better, but it is conjecture that you cannot do much better, let's say, by more than a constant factor compared to the square grid. Yes.

AUDIENCE: What is Katz's first name?

YUFEI ZHAO: Sorry?

AUDIENCE: What is Katz's first name?

YUFEI ZHAO: So this is Nets Katz. He's at CalTech. I want to take a short break now. And when we come back, I want to show you some ways to generate lower bounds to the extremal number.

In the Kovari-Sos-Turon theorem, we understood some upper bounds for the extremal number. So I want to turn our attention now to lower bounds. Which means constructing, in some sense, graphs that have lots of edges, at the same time being K s,t-free. There are many classes of construction, so there are several techniques for doing this. And I want to introduce you to some of these techniques. And the first one comes from the probabilistic method, where we use a randomized construction by picking a graph at random, modifying it a little bit in a way that's to our desire.

This method is very powerful. It is very general. It is applicable in a lot of situations. But, unfortunately, for the problem of getting tight bound, it really allows you to get tight bounds. So it is very robust, but somehow it's often not sharp. And another class of constructions that are algebraic-- all the sharp examples come from algebraic constructions. And there, we're able to use some nice ideas from algebra or from algebraic geometry to get tight constructions.

And you can wonder, is there some way to combine the best of both worlds? And it turns out-- this was an important reason development that was found just several years ago-- where one can combine ideas from the two and obtain a randomized algebraic construction. And that leads to a new source of constructions for large H-free graphs with many edges.

So the plan is to show you how these constructions work. First, let's begin with randomized constructions. As I mentioned, this construction is very general, it's very robust, and we can use it to obtain H-free graphs for every graph H. Of course, when H is not bipartite, then we saw from the Turon graph that that's pretty much the right thing to do.

So you should think of bipartite graphs H. So fix a graph H with at least two edges. Otherwise, the problem is trivial. The claim is that there is some constant C such that for every n-- you should think of and this being large-- there exists an H-free graph on n vertices and having lots of edges.

And we will show that you can obtain the following number of edges. So 2 minus number of vertices of H minus 2 divided by the number of edges of H minus 1. Don't worry about this expression in the exponent, it will come out of the proof. In other words, the extremal number is at least this quantity here.

How good is this now? So it's some scary looking expression in the exponent. So let me just give you some special cases for comparison to the Kovari-Sos-Turon theorem. For K s,t's, the bound, this construction gives you a lower bound which is of the form m to the 2 minus s plus t minus 2 divided by s,t minus 1. So I used this symbol, this bigger tilde to denote dropping constant factors.

In particular, setting s and t to be the same, we find that the K s,s extremal number is lower-bounded by this quantity here. So how does that compare to Kovari-Sos-Turon? In Kovari-Sos-Turon we saw that K s,s is at most 2 to the n to the 2 minus 1 over s. So there's a bit of a gap, and in particular even for 2 and 2.

These results tell you a lower bound on the order of n to the 4/3 and the upper bound on the order of 3/2. So I'm just doing this explicitly to show you that even in sometimes the simplest case there is a gap between these two bounds. Later on, we'll see a construction that shows that the right-hand side is tight. So this is the truth. Randomized construction gives you something, but it doesn't give you the truth.

On the other hand, I want you to notice that if t gets very large as a function of s, this exponent here approaches 1 over s as t goes to infinity. So for very large values of t, at least in the exponent, it's not that far off. Although, you never get the right exponent for any specific essential.

So this is some limitation of the randomized method, but it's variable robust. And, in fact, you can use this result to bootstrap it to a slightly better one for some graphs H. So, for example, one might do better by replacing this H by a subgraph. Because if you have a subgraph, H prime, and you construct your G to be H prime-free, then it is automatically H-free. But maybe this theorem actually gives you a better construction when restricted to H prime.

And what can you do better? Well, you can do better if h prime is such that whatever the quantity that comes up in the exponent-- so let me write it like this-- is superior to the exponent you would obtain by just looking at H.

So let me give the name to this notion here. So let me call it the 2-density of the graph to be whatever quantity here that would be the maximum if you allow to pass down to subgraphs. So the 2-density of a graph H-- so denoted in literature often as m sub 2-- to be the maximum where you are allowed to look at subgraphs, let's say, on at least three vertices to avoid degeneracies of this ratio that comes up in the expressions. So then the theorem-- that construction there-- implies that the extremal number is at least not just the expression I've written, but maybe sometimes you can do slightly better by passing to a subgraph.

So let me give you a concrete example. Suppose my H is this graph here. So you can run the calculation and find what these numbers are. So the number of vertices, edges, and this ratio here to be-- so 5 vertices, 8 edges, and 7/3. And the idea here is that it's more helpful. So I want to create something which is H-free, but dense things are easier to avoid.

So if I have something which has a fairly dense core, that's easier to avoid. So maybe it's better to, instead of looking at the whole H, look at this K 4. So if you can avoid K 4, of course, you avoid H, and maybe it's easier to just avoid K 4 and not worry about some of the extra decorations.

So if you look at this H prime and go through the parameters, you find that it is like that. And it is somewhat denser in this sense, compared to looking at the whole graph H. So you can improve on this theorem for some graphs H where you can pass to a denser core. Any questions so far? Yes.

AUDIENCE: Why is this method called the 2-density?

YUFEI ZHAO: The questions is, why is this measure called a 2-density? So that's a name given in literature. It partly has to do with these extra ratios. So there are other notions of densities that actually we'll see later on. So this term we'll only see today. So it's more of an ad hoc term for the purpose of this course. Later on, we'll see notions of density that are more relevant for our discussions. Any more questions?

So let me show you how to prove this theorem. The proof is very intuitive. The idea is you take something at random and then you fix it. That's it.

Let's consider a random graph. The Erdos-Rényi random graph-- so whenever I say random graph, I almost always refer to this one here. And the Erdos-Rényi random graph is obtained by considering n vertices and each possible edge appearing independently and uniformly with probability p. And we're going to decide this p later on. So let me not tell you what p is for now.

I'm interested in avoiding H. So this random graph may have some copies of H. Let me count the number of copies of H.

We can compute the expected number of copies of H by linearity of expectations. For every possible placement, look at what's the probability that that placement generates an H. Namely, look at all different possibilities for choosing the possible vertices of H. I need to divide it by a factor that accounts for the number of automorphisms of H. But just a constant factor-- don't worry about it. And for each of these possible placements, H appears with probability exactly p to the number of edges of H, just by linearity of expectations. And I can upper-bound this quantity very crudely by e to the number of edges of H times n, which is the number of vertices of G raised to the number of vertices of H.

On the other hand, I also want a graph G that has lots of edges, because that's what we're trying to do. We're trying to generate a graph with lots of edges that's H-free. So the number of edges of G, that's also easy to compute. It's a binomial distribution, and it has expectation p times n choose 2.

And, basically, I want this quantity to be much larger than that quantity. So I choose an appropriate p. Namely, by comparing these two quantities, we can choose p to be, let's say 1/2-- the 1/2 is not important-- times n to 2 the v of H minus 2 divided by e of H minus 1. So the exponent comes out of comparing these two expressions. So you see a 1 and a 2.

Once you have this p, see that the number which is the difference-- so take the number of edges of G minus the number of copies of each. I know the expectation of both. I can look at their difference with this value of p. We find that it is at least 1/2 of the number of edges. So on expectation, you don't lose too much. So p is chosen so that this inequality is true.

So you set what p is. We find that this quantity here is at least some constant times n to the 2 minus v of H minus 2 divided by e the H minus 1. We're still working with a random graph, and we know that this quantity here is on expectation at least that number-- so not too small. Therefore, there exists some instance in this randomness, some G such that the quantity above for the specific random instance is at least its expectation.

So this gives us a graph G which on one hand has lots of edges, but also has very few copies of H relative to the number of edges. So we can now get rid of all the copies of H by removing one edge from each copy of H in G to remove all copies of H. And now, we obtain an H-free graph.

How many edges are there in this graph? Well, we removed at most one edge for each copy of H. So the number of edges is at least this quantity here, which is what we wrote just now. And that's it. So now we've obtained our graph on n vertices with lots of edges with the claimed bound that's H-free.

So this is the probabilistic method. You start with something random. You try to fix it. And this method sometimes has the name of the alteration method. And this is a very important idea and one of the key ideas in the probabilistic method, which I encourage you to go and learn more about. We'll also see this method later on when we discuss the randomized algebraic construction.

AUDIENCE: Just to clarify, none of the copies of H [INAUDIBLE].

YUFEI ZHAO: So the questions is-- yes-- what do I mean by the number of copies of H? I mean every instance of H you see. So there could be intersectings. I'm not asking for destroying copies.

AUDIENCE: Sure.

YUFEI ZHAO: So a complete graph on n vertices has n choose 3 triangles. Any more questions? OK.

So now we've seen that the probabilistic method gives you know some bound. And it's not too hard to apply, but it doesn't give you the right bound. It doesn't give you the truth. So now, I want to show you a different type of constructions, namely algebraic constructions that do allow you to get the truth, but they work in only a small number of cases. And so it's more magical, but they work better when the magic happens.

So let's discuss algebraic constructions. In particular, I want to show you how to obtain the type bound on the extremal number for K 2,2, namely a fourth cycle. So this is a result due to Erdos-Rényi-Sos, and it tells us that the extremal number for K 2,2 is at least 1/2 basically up to asymptotics times n to the 3/2.

Actually, if you look at the constant that came out of the proof of the Kovari-Sos-Turon theorem, It is also 1/2. So as a corollary, we see that the extremal number is like that. So this is one of extremely few cases where we know the extremal number so well. So if you go back to the proof of Kovari-Sos-Turon, you see that the constant actually there is 1/2.

So I want to construct for you a graph that has no fourth cycles and has lots of edges. So that's the name of the game. And I'll describe this graph for you explicitly. So this graph has a name. It's called a polarity graph.

Let's suppose that n is a number such that 1 bigger than this number is a square of a prime. So our construction will use some finite fields. I'll explain a bit. If n is not of this form, then you can change n to a number of very close to of this form, and everything will be OK.

The graph will be constructed as follows-- the vertex set will be the plane over Fp. Let's remove the origin. So it has n points exactly. And the edges are such that I put an edge between x,y and a,b, if and only if the equation ax plus by equals to 1 holds. And this equation is meant to be read in Fp. So that's the graph. That's an explicit description of this graph.

So I need to show you two things-- one, that has lots of edges, the claimed number of edges. And two, it has no fourth cycles. So let's start with not having fourth cycles. So y is a K 2,2-free

So what would the K 2,2, be? So let's consider two points. And I want to understand the number of common neighbors of these two points.

Well, look at the description for the edges. What are the neighbors? The neighbors correspond to solutions to this system of equations. And the basic claim is that there is at most one solution-- x,y. So it's basic fact linear algebra. You have to be slowly careful in case a,b is a multiple of a prime b prime. But, actually, in that case, you have no solutions anyway.

The second claim then is that this graph has lots of edges. Well, actually, that's not too hard to show either. So I claim that every vertex has degree-- so how many edges come out of every vertex? I give you a common b, so how many x comma y satisfy that equation up there? Basically, for-- so 1 of x and a and b is non-zero.

So let's say a is non-zero. Therefore, whatever value of y you set, I can find a unique x that solves the equation. I have to be slightly careful, because I don't allow loops in my graph. So I might lose one edge because of that. In any case, every vertex has to agree exactly P or P minus 1. Just solve that equation in x and y.

So the P minus 1 comes from no loops. Therefore, the number of edges is equal to the claimed bound. So this finishes the proof in case when n has that special number theoretic form. But we can extend to all values of n like this-- if n doesn't have that form, then I can take a prime P. It may not necessarily be exactly satisfying that inequality, but I can always take a prime pretty close to it. So I can always take a prime which is up to a negligible multiplicative error what I want.

And then we use-- and such that P squared minus 1 is at most n. And use the above construction and add isolated vertices to finish the job to get exactly n vertices. And the reason that I can always take a prime very close to it is because there's a theorem in number theory that tells us that for n large enough, I can always find the prime which is slightly less than n but no more than a negligible multiplicative factor of n. The best result of this form is-- I'm just telling you something in number theory for cultural reasons-- due to Baker-Harman-Pintz. And so, this is the question regarding how large can gaps between primes be.

So you might know the [Bertrand's postulate]] theorem that tells you there is always a prime between n and 2n. So what about between n and n plus root n. Actually, we don't know that. So the best result of the form is that there is always a prime-- so for n sufficiently large there exists a prime between n minus n to the exponent 0.525 and n.

In any case, this number here, whatever it is is little n, and that's enough for our purpose. So it suffices to look at n of a special number theoretic form where you're allowed to use primes. So that's the construction there. Let me show you a interpretation of that construction which I think is may be helpful to think about, and that's that you can view it as the incidence graph between points and lines in projective space-- in projective plane.

So I start with a projective plane. So I can view a bipartite version of that construction. It can be viewed as the point-line incidence graph of a projective plane over a finite field. And by this, I mean put as one vertex set the points of the projective plane and on the other side the lines. And I put in an edge between a point and a line if and only if the point lies on the line.

So you can do this more explicitly in coordinates if you view points and lines as coordinates. And so the equation for getting a point to be on the line is like that. So now why is there no fourth cycle? A fourth cycle would correspond to two points lying on two different lines, which is not possible in this geometry. So that's the reason for that construction up there.

So no two points in two lines. Any questions about this polarity constructions?

AUDIENCE: Why is it called a polarity construction.

YUFEI ZHAO: The question is, why is it called a polarity construction? So it relates points and their polars, which are lines. Yeah.

AUDIENCE: Why does this not have-- like, on your two P squared vertices, looking at one vertex for every [INAUDIBLE], one vertex for [INAUDIBLE]?

YUFEI ZHAO: OK. So the question has to do with the number of vertices here. It's true. Here, I double the number of vertices, and so I don't get that constant there. But what that graph up there-- that's not a bipartite graph. It is identifying the points and the lines and overlaying the two parts into one. But if you don't care about the constants, this graph here may be conceptually easier to think about. Yes.

AUDIENCE: [INAUDIBLE] to generalize the polarities bound to K [INAUDIBLE].

YUFEI ZHAO: Great. The question is, can you generalize this polarity graph to K 3,3 and higher? So that's what we're about to do next. So for K 3,3, what can you do?

So the main observation here is that two lines intersecting at most one point. But there are other geometric facts of that form. So we're going to use one of them to get K 3,3-free graph. And this construction is due to Brown, that the extremal number for K 3,3 is also at least a factor 1/2 minus total 1 times-- so now, what's the exponent? What is predicted by Kovari-Sos-Turon is 2 minus 1/3, and Brown obtains the correct exponent.

It turns out this is also the right constant, this 1/2, although it doesn't follow from the Kovari-Sos-Turon theorem I stated. One needs to do a little bit extra work. But it turns out it is true that this is also the correct constant. And that's actually pretty much all the cases where we know the correct constant. And there are other cases where we know the correct exponent, but these things tend to be hard to come by.

So let me show you how to construct this graph. So it's based on a similar idea as the polarity graph. It has some more technicalities. So I'm not going to do the full proof and just give you the sketch.

As earlier, I'm using the same trick. We can assume that n has a special form. Here, let me assume the n is a cube of a prime.

I'm going to put s edges. So first, the vertices of my graph are going to be points in the affine plane over Fp. Previously, the edges had to do with lines. And now, let's use spheres.

So the edges of the form where I join two vertices like this if and only if they're-- well, it's not really a distance, but it's something that looks like the equation of a sphere, where u is some fixed non-zero element of Fp. You may have to be somewhat careful in choosing this u, but let me not worry too much about it. So you fixed some so-called distance, even though it's not a distance, and I join the vertices whenever they satisfy that equation having that not distance.

What's the intuition here? The intuition is that I want to avoid-- so how do I know that this graph has no K 3,3? Well, first, let's think about what happens in real space. So intuition in the real space-- well, here, I have this graph that, let's say, the unit distance graph in r3. So the neighborhood of each point is a unit sphere. And what I want to know is that if you have three spheres, three unit spheres, how many common intersection points can they have?

Two spheres intersecting a circle. And that circle cannot lie on the third circle. That you should think about. So that circle intersects the third circle in at most two points. So three unit spheres have at most two common points. And so the unit distance graph in r3 is K 3,3-free.

That entire argument, even though I expressed it geometrically, it's an algebraic argument. You can write down equations on intersections between two spheres. It's the intersection of the sphere with its coaxial plane.

I have a couple of these colossal planes. They get me a coaxial line. That line has to intersect the sphere in almost two points.

You should actually do this algebra if you want to do the proof, because there are funny things that can happen you find fields. For example, maybe the sphere contains a line. But you choose your parameters correctly, and these things don't happen.

And that's the intuition. And if you actually work this out, you'll find that this graph here is indeed K 3,3-free. So I'm skipping the details. But you should do the algebra if you want to have a proof.

On the other hand, it also has lots of edges. And that's basically the same reason as before. But I can count the number of edges by fixing some x, y, and z, and look at how many abc's satisfy that equation. And that's, again, something that needs to be checked, but the point is that this graph-- so every vertex has either a P square-- so it has close to P squared degree.

So lots of edges, and combining with basically the same idea as before, you get the construction. Any questions?

So where can we go from here? To construct the K 2,2-- by the way, if you construct K 2,2-

free, the same construction works for K 2,3-free. If it's K 2,2-free, then, it's also K 2,3-free or K 2,4-free. Here, likewise, this is also K 3,4-free.

So now, what about higher K s,t's? And you might think, well, let's take these geometric objects and try to extend them further. But that actually seems kind of difficult. We do not really know how to do it. We do not know how to obtain a construction which is of this form that works for K 4,4. In fact, there are even some evidence that that might be even impossible.

As I mentioned, K 4,4 is a major open problem. It is an open problem to determine the order of the extremal number of K 4,4. But in any case, this construction, this idea of using algebraic constructions, is very enlightening, that we should look at ways to get large K s,t-free graphs by coming up with clever algebraic constructions.

And next, time I will show you a couple of very nice ideas where you can get-- you can come up with a different kind of algebraic construction which has some superficial similarities to what we've seen today but that's really of a different nature. So, next time, we will see the following theorem, which is obtained in a sequence of two papers, union of authors, Alon, Rényi and Sazbó, that shows that if t is much larger than s-- so minus 1 factorial plus 1-- then the extremal number for K s,t is on the same order as the upper bound determined in Kovari-Sos-Turon theorem.

Just to be more explicit about what these s and t are, if you plug in what's the smallest t that this theorem gives for various values of s, find 2,2, and 3,3. And the next one is 4, 7. And then, it gets worse from there. So these constructions are based on-- I mean, they are algebraic constructions. So we'll see next time what happens.