

**YUFEI ZHAO:** So today we are going to start a new chapter on graph limits. So graph limits is a relatively new subject in graph theory. So as the name suggests, we're looking at some kind of an analytic limit of graphs, which sounds kind of like a strange idea because you think of graphs as fundamentally discrete objects.

But let me begin with an example to motivate, at least pure mathematical motivation for graph limits. There are several other ways you can motivate graphs limits, especially coming from more applied perspectives. But let me stick with the following story.

So suppose you lived in ancient Greece and you only knew rational numbers. You didn't know about real numbers. But you understand perfectly rational numbers. And we wish to maximize. So we wish to then minimize the following polynomial,  $x^3 - x$ , let's say for  $x$  between 0 and 1.

So you can do this. And suppose also the Greeks knew calculus and take the derivative and all of that. So you find that.

You have a problem because we know-- so given our advanced state of mathematics, we know that the maximum-- so the minimizer is at  $x$  equals to  $1/\sqrt{3}$ . But that number doesn't exist in the real numbers. So how might a civilization that only knew rational numbers express this answer? They could say, the minimum occurs not in  $\mathbb{Q}$ . So there's not minimized in  $\mathbb{Q}$ -- but not minimized by a single number, but by a sequence.

And this is a sequence that a more advanced civilization would know, a sequence that converges to  $1/\sqrt{3}$ . But I can give you this sequence through some other means. And this is one of the ways of defining the complete set of real numbers, for instance. But you can define explicitly a sequence of real numbers that converges.

But of course, this is all quite cumbersome if you have to actually write down this sequence of real numbers to express this answer. It will be much better if we knew the real numbers. And we do. And the real numbers, in some sense, in the very rigorous sense, is a completion of the rational numbers.

That's the story that we're all familiar with. But now let's think about graphs which are some kind of a discrete set of objects, akin to the rational numbers. And the story now is, among

graphs, suppose I have a fixed  $p$  between 0 and 1. And the problem now is to minimize the 4-cycle density among graphs with density, with edge density  $p$ .

So this is some kind of optimization problem. So I don't restrict the number of vertices. You can use as many vertices as you like. And I would like to minimize the 4-cycle density.

Now, we saw a few lectures ago this inequality that tells us that-- so we saw a few lectures ago that this density is always at least  $p$  to the fourth. So in the lecture on quasirandomness, so we saw this inequality. And we also saw that this minimum is approached by a sequence of quasirandom graphs.

And in some sense, that-- so the answer is  $p$  to the fourth. And there's not a specific graph. There's no one graph that minimizes. This 4-cycle density is minimized by a sequence.

And just like in the story with the rational numbers and the real numbers, it would be nice if we didn't have to write out the answer in this cumbersome, sequential way, but just have a single graphical-like object that depicts what the minimizer should be. And graph limits provides a language for us to do this. So one of the goals of the graph limits-- this gives us a single object for this minimizer instead of taking a sequence.

So roughly that is the idea that you have a sequence of graphs. And I would like some analytic object to capture the behavior of the sequence in the limit. And these graph limits can be written actually in a fairly concrete form.

And so now let me begin with some definitions. The main object that we'll look at is something called a graphon. So it merges the two words graph, function. A graphon is by definition a symmetric, measurable function, often denoted by the letter  $W$  from the unit square to the 0, 1 interval. And here being symmetric means that if you exchange the two argument variables, this function remains the same.

So that's it. So that's the definition of a graphon. And these are the objects that will play the role of limits for sequences of graphs. And I will give you lots of examples in a second. So that's the definition.

This is the form of the graphons that we'll be looking at mostly. But just to mention a few remarks, that the domain can be instead any product of any square of a probability measure space-- so instead of taking the 0, 1 interval, I could also use any probability measure space. So it's only slightly more general. So there are some general theorems in measure theory that

tells us that most probability measure spaces, if they're nice enough, they are in some sense equivalent or can be captured by this interval.

So I don't want you to worry too much about all the measure where there are technicalities. I think they are not so important for the discussion of graph limits. But there are some subtle issues like that just lurking behind.

But I just don't want to really talk about them. So for the most part, we'll be looking at graphons of this one. And also the-- so instead of the domain, so the values-- so instead of 0, 1 interval, you could also take a more general space, for example, the real numbers or even the complex numbers.

I'm going to use the word graphon to reserve this word for when the values are between 0 and 1. And if it's in  $\mathbb{R}$ , let me call this just a kernel, although that will not come up so much. So when I say graphon, I just mean the values between 0 and 1. Although if you do look up papers in the literature, sometimes they don't use these words so consistently. So be careful what they mean by a graphon.

So that's the definition. But now let me give you some examples on how do we think of graphons and what do they have to do with graphs. So if we start with a graph, I want to show you how to turn it into a graphon.

So let's start with this graph, which you've seen before. This is the half graph. So from this graph, I can label the vertices and form an adjacency matrix of this graph, where I label the rows and columns by the vertices and put in zeros and ones according to whether the edges are adjacent. So that's the adjacency matrix.

And now I want you to view this matrix as a black and white picture. So think one of these pixelated images, where I turn the ones into black boxes. Of course, on the blackboard, black is white and white is black. So I turn the ones into black boxes. And I leave the zeros as empty white space.

So I get this image. And I think of this image as a function. And this is the function going from 0, 1, squared to 0, 1, interval, taking only 0 and 1 values. So that's a function on the square.

But now, so this is a single graph. So for any specific graph, I can turn it into a graphon like this. But now imagine you have a sequence of graphs. And in particular, consider a sequence

of half graphs.

So here is  $H_3$ . So  $H_n$  is the general half graph. And you can imagine that, as  $n$  gets large, this picture looks like-- instead of the staircase you just have a straight line connecting the two ends. And indeed, this function here, this graphon, is the limit of the sequence of half graphs as  $n$  goes to infinity.

So one way you can think about graphons is you have a sequence of graphs. You look at their adjacency matrix. You view it as a picture, a pixelated image, black and white according to the zeros and ones in its adjacency matrix.

And as you take a sequence, you make your eyes a little bit blurry. And then you think about what the sequence of images converges to. So the resulting limit is the limit of this sequence of graphs. So that's an informal explanation. So I haven't done anything precisely.

And in fact, one needs to be somewhat careful with this depiction because let me give you another example. Suppose I have a sequence of random or quasirandom graphs with edge density  $1/2$ . So what does this look like?

And I have this picture here. And I have a lot of-- so I have a lot of-- one-half of the pixels are black. And the other half pixels are white. And you can think, from far away, I cannot distinguish necessarily which ones are black and which ones are white.

And in the limit, it looks like a grayscale image, with a grayscale being one-half density. And indeed, it converges to the constant function,  $1/2$ . So the limit represented by this problem up here is the constant graphon with the constant value  $p$ .

But now let me give you a different example. Consider a checkerboard. So here is a checkerboard, where I color the squares according to, in this alternating black and white manner, according to a usual checkerboard.

And as the number of squares goes to infinity, what should this converge to? By the story I just told you, you might think that if you zoom out, everything looks density  $1/2$ . And so you might guess that the image, the limit, is the  $1/2$  constant.

But what is this graph? It's a complete bipartite graph. It is a complete bipartite graph between all the even rows. And there's a different way to draw the complete bipartite graph-- namely, that picture, just by permuting the rows and columns. And it's much more reasonable that this

is the limit of the sequence of complete bipartite graphs with equal parts.

So one needs to be very careful. And so it's not necessarily an intuitive definition. The idea that you just squint your eyes and think about what the image becomes, that works fine for intuition for some examples, but not for others. So we do really need to be careful in giving a precise definition. And here the rearrangement of the rows and columns needs to be taken care of.

So let me be more precise. Starting with a graph  $G$ , I can-- so let me label the vertices by 1 through  $n$ . I can denote by  $W$  sub  $G$  this function, this graphon, obtained by the following procedure.

First, you partition the interval into intervals of length exactly  $1/n$ . And you set  $W$  of  $x$  comma  $y$  to be basically what happened in the procedure above. If  $x$  and  $y$  lie in the box  $I$  sub  $I$  cross  $I$  sub  $J$ , then I put in 1 if  $I$  is adjacent to  $J$  and 0 otherwise-- so this picture, where we obtained by taking the adjacency matrix and transforming it into a pixelated image.

What are some of the things that we would like to do with graph limits or graphs in general?  
Yeah?

**AUDIENCE:** Is the range also 0, 1, squared or 0, 1?

**YUFEI ZHAO:** Thank you. The range is 0, 1. So here are some quantities we are interested in when considering graph limits. So given two graphs,  $G$  and  $H$ , we say that a graph homomorphism-- so a graph homomorphism between from  $G$  to  $H$  is a map of their vertexes such that the edges are preserved. So you have-- and so whenever  $uv$  is an edge of  $H$ , your image vertexes get mapped to an edge of  $G$ .

And we are interested in the number of graph homomorphisms. So often I use uppercase to denote a set of homomorphisms  $G$  to  $H$ -- and lowercase to denote the number. So for example, the number of homomorphisms from a single vertex-- so a single vertex with no edge to a graph  $G$ , that's just the-- what is this quantity? So some number of vertices of  $G$ -- what about homomorphisms from an edge to  $G$ ?

**AUDIENCE:** The number of edges?

**YUFEI ZHAO:** Not quite the number of edges, but twice the number of edges. What about the number of homomorphisms from a triangle to  $G$ ?

**AUDIENCE:** 6 times the number of triangles.

**YUFEI ZHAO:** So yeah, you got the idea-- so the 6 times the number of triangles. So now let me ask a slightly more interesting question. What about the number of homomorphisms from  $H$  to a triangle? What's a different name for this quantity here?

It's the number of proper three colorings. So it's the number of proper three colorings, the number of proper colorings of  $H$  with three labeled colors, red, green, and blue. So think about the three vertices. That's red, green, and blue. And whichever vertex of  $H$  can map to red, color that vertex red.

So you see that there is a one-to-one correspondence between such homomorphisms and proper colorings. So many important graph parameters, graph quantities, can be encoded in terms of graph homomorphisms. And these are the ones that we're going to be looking at most of the time.

When we're thinking about very large graphs, often it's not the number of homomorphisms that concern us, but the density of homomorphisms. And the difference between homomorphisms on one hand and subgraphs is that the homomorphisms are not quite the same as subgraphs, other than this multiplicity, because you might have non-injective homomorphisms. But these non-injective homomorphisms do not end up contributing very much because they only have  $n$  to the number of vertices of  $H$  minus 1 on that border where I think of  $n$  as the number of vertices of  $G$ .  $n$  is supposed to be large. So in terms of graph limits when  $n$  gets large, I don't need to distinguish so much between homomorphisms and subgraphs.

We define the homomorphism density, denoted by the letter  $t$ , from  $H$  to  $G$ , by-- define it to be the fraction of all vertex maps that are homomorphisms. So this is also equivalent to be defined as the probability that a uniform random map from the vertex set of  $H$  to the vertex set of  $G$  is a homomorphism from  $H$  to  $G$ . So it's a graph homomorphism. And this quantity turns out to be quite important. So we're going to be seeing this a lot.

And because of this remark over here, in the limit, this quantity of graph homomorphism densities in the limit as the number of vertices  $G$  goes to infinity and  $H$  fixed, the homomorphism densities approaches the same limit as subgraph densities. So you should regard these two quantities as basically the same thing. Any questions so far?

So all of these quantities so far defined are for-- so everything is defined so far for graphs, so what happens between graphs and graphs. So what about for graphons? I'll give you this limit object, this analytic object. I can still define densities by integrals now. So suppose I start with a symmetric measurable function.

So tell me, for example, a graphon. But I can let my range be even more generous. Starting with such a function, I define the graph homomorphism density from a fixed graph  $H$  to this graphon or kernel, more generally, to be the following integral, where I'm-- before writing down the full form, let me first give you an example. I think it will be more helpful.

So if I'm looking at a triangle going to  $W$ , what I would like is the integral that captures the triangle density. So this quantity here, if I let  $x$ ,  $y$ , and  $z$  vary over  $0$  and  $1$ ,  $0$  through  $1$ , independently and uniformly, then this quantity here captures the triangle density in  $W$ . In fact, and I'll state this more precisely in a second-- if you look at the translation from graph to graphon and combine that translation with this definition here, you recover the triangle density. More generally, for  $H$  instead of a triangle, the  $H$  density in a graphon is defined to be the integral of-- instead of this product here, I take a product corresponding to the graph structure of  $H$  with one factor for each edge of  $H$ . And the variables go over the vertex set of  $H$ .

So this is the definition of homomorphism densities, not for graphs, but for symmetric measurable functions, in particular, for graphons. And we define it this way because-- and we use the same symbols because these two definitions agree. If you start with a graph and look at the  $H$  density in  $G$ , then this quantity here is equal to the  $H$  density in the graphon associated to the graph  $G$  constructed as we did just now. So make sure you understand why this is true and why we defined the densities this way. Any questions so far?

So we've given the definition of graph homomorphism density. And we've defined these objects, these graphons. And I mentioned even something about the idea of a limit. But in what sense can we have a limit of graphs?

So here is an important definition on the convergence of graphs. So in what sense can we say that a sequence of graphs converge? So we say that a sequence of graphs  $G_n$ -- graphs or graphons, so these two definitions are interchangeable for what I'm about to say regarding limits for graphons, in which case I'm going to denote them by  $W_n$ . So we say the sequence is convergent if the sequence of subgraph densities-- of course, if you are looking at graphons, then you should look at the graphon, the subgraph density in-- homomorphism

density in graphons if this sequence converges as  $n$  goes to infinity for every graph  $H$ .

So that's the definition of what it means for a sequence of graphs to converge, which so far looks actually quite different from what we discussed intuitively. But I will state some theorems towards the end of this lecture explaining what the connections are. So intuitively what I said earlier is that you have a sequence of graphs that are convergent if you have some vague notion of one image morphing into a sequence of images morphing into this final image. Still hold that thought in your mind. But that's not a rigorous definition yet.

The definition we will use for convergence is if all the subgraph-- all the homomorphism densities were equivalently subgraph densities, they converge. So this is the definition. It's not required. So this is basically rigorous as stated.

Just as a remark, it's not required that the number of vertices goes to infinity, although you really should think that that is the case. So just to put it out there-- so I can have a sequence of constant graphs and they will still be convergent. And that's still OK. But you should think of the number of vertices going to infinity. Yeah?

**AUDIENCE:** What is  $F$  in the definition?

**YUFEI ZHAO:**  $F$  is  $H$ . Thank you. Any other questions?

So there are some questions that we'd like to discuss. And this will occupy the next few lectures in terms of proving the following statements. One is do you always have graph limits? If you have a convergent sequence of graphs, do they always approach a limit?

And just because something is convergent doesn't mean you can represent the limit necessarily. So it turns out the answer is yes. It turns out that-- and this makes it a good theory, a good, useful theory, and an easy theory to use, that there is always a limit object whenever you have convergence.

And the other question is while we have described intuitively one notion of convergence and also defined more rigorously another definition of convergence, are these two notions compatible? And what does this even mean, this idea of image becoming closer and closer to a final image? What does that even mean? So these are some of the questions that I would like to address.

So in the next few things that I would like to discuss, first, I want to give you a definition of a

distance between two graphons or two graphs. If I give you two graphs, how similar or dissimilar are they-- so that we have this metric. And then we can talk about convergence in metric spaces. So let's take a quick break.

So given this notion of convergence, I would like to define the notion of distance between graphs so that convergence corresponds to convergence in the metric space sense of distance going to 0. So how can we define distance? First, let me tell you that there's a trivial way. And so there's a way in which you look at that definition and produce a distance out. And here's what you can do.

I can convert that definition to a metric by setting the distance between two graphs  $G$  and  $G'$  to be the following quantity, obtained by-- what would I like to do? I would like to say the distance goes to 0 if and only if the homomorphism densities, they are all close to each other. And so I can sum up all the homomorphism densities and look at their differences between  $G$  and  $G'$ . And I simply enumerate the list of all possible graphs.

I want to be just slightly more careful with this definition here because I want something which-- so when I write this, this number might be infinite for all pairs  $G$  and  $G'$ . So if I just add a scaling factor here, then-- and this is some distance. So this is some distance. And you see that it matches the definition up there.

But it's completely useless. It might as well-- might as well not have said anything because it's tautologically the same as what happened up there. And if I give you two graphs, it doesn't really tell you all that much information except to encapsulate that definition into a single number. Great. So I'm just-- the point of this is just to tell you that there is always a trivial way to define distance.

But we want some more interesting ways. So what can we do? So here is an attempt, which is that of an edit distance. So we have seen this before when we discussed removal lemmas.

The edit distance is the number of edges you need to change to go from one graph to the other graph. And this seems like a pretty reasonable thing to do. And it is an important quantity for many applications, but turns out not the right one for all application. And here is the reason. So this is why the edit distance is-- by edit distance, I mean 1 over the number of vertex squared times the number of edge changes needed.

So there's normalization so that the distance is always between 0 and 1. But this is not a very

good notion for the following reason. If I take two copies of the Erdos-Reyni random graph  $G_{n, 1/2}$ , what do you think is the edit distance between two such random graphs? How many edges? Yeah?

**AUDIENCE:** Isn't it roughly one-half of the number of edges because there's like a one-half probably that won't be there or not be there [INAUDIBLE]?

**YUFEI ZHAO:** So yeah, so let me try to rephrase what you are saying. So suppose I have this  $G$  and  $G$  prime both sitting on top of the vertex set  $n$ . So if I'm not allowed to rearrange the vertices, how many edge changes do I need to go from one to the other? I need about  $1/2$ .

So one-half the time, I'm going to have a wrong edge there. Now you can make this number just slightly smaller by permuting the vertices. But actually you will not improve that much. It is still going to be roughly that edit distance, which is quite large. This is almost as large as you can possibly get between two arbitrary graphs.

So if we want to say that random graphs, they approach a limit, a single limit, then this is not a very good notion because they are quite far apart for every  $n$ . So this is the reason why the more obvious suggestion of an edit distance might not be such a great idea. So what should we use instead? So we should take inspiration from what we discussed in quasirandomness. You have a question.

**AUDIENCE:** Is the edit distance only for two graphs of the same vertex set?

**YUFEI ZHAO:** So the question is, is the edit distance only for two graphs with the same vertex set? Let's say yes. So we'll see later on, you can also compare graphs with different number of vertices. So hold onto that thought.

So I would like to come up with a notion of distance between graphs that is inspired by our discussion of quasirandomness earlier. So think about the discussion of quasirandomness or quasirandom graphs. In what sense can  $G$  be close to a constant, let's say  $p$ ? And so this was the Chung-Graham-Wilson theorem that we proved a few lectures ago. So in what sense can  $G$  be close to  $p$ ?

And one of those definitions was discrepancy. And discrepancy says that if the following quantity is small for all subsets  $x$  and  $y$ , which are subsets of vertices of  $G$ -- so you remember, all of you remember, this part, the discrepancy hypothesis for quasirandomness. And this is a kind of definition that we would like to describe when two graphs are similar to each other,

when they are close in this discrepancy sense.

So now, instead of a graph and a number, what if now I have two graphs? I'll give you two graphs of  $G$  and  $G$  prime. And what I would like to say is that, if for now, so if they have the same vertex set, I want to say that they are close if I have that the number of edges between  $x$  and  $y$  in  $G$  is very close to the number of edges between  $x$  and  $y$  in  $G$  prime. And I normalize by the number of vertices squared, so  $n$  this number of vertices.

And I would like to find out the worst possible scenario, so overall,  $x$  and  $y$  subsets of the vertex set. If this quantity is small, then I would like to say that  $G$  and  $G$  prime are close to each other. So this is inspired by this discrepancy notion. Can you see anything wrong with this definition here? Yeah?

**AUDIENCE:** [INAUDIBLE]

**YUFEI ZHAO:** So permutations are vertices. So just like in the checkerboard example we saw earlier, you have two graphs. And if they are indeed labeled graphs in the same labeled vertex set, then this is the definition more or less what we used. I will define it more precisely in a second. But if they are unlabeled vertices, we need to possibly optimize permutations over rearrangements of vertices, which actually turns out to be quite subtle.

So I'm going to give precise definitions in a second. But this one here, so think about permuting vertices. But it's actually a bit more subtle than that.

So here are some actual definitions. I'm going to define this quantity called a cut norm. So this chapter is all going to be somewhat functional analytic in nature. So get used to the analytic language.

So the cut norm of  $W$  is defined to be the following quantity denoted by this norm with a box in the subscript, which is defined to be-- if I look at this  $W$ , and I integrate it over a box, and I would like to maximize this quantity here over choices of boxes  $S$  and  $T$ , they are subsets of the interval measurable subsets. So choose your-- so over all possible choices of measurable subsets  $S$  and  $T$ , if I integrate  $W$  over  $S$  cross  $T$ , what is the furthest I can get from 0? So this is the definition of cut norm. And you can already see that it has some relations to what were discussed up there.

But while we're talking about norms, let me just mention a few other norms that might come up

later on when we discuss graph limits. So there will be a lot of norms throughout. So in particular, the  $l_p$  norm is going to play a frequent role. So  $l_p$  norm is defined by looking at the peak norm of the absolute value, integrated and then raised to  $1/p$ . And so the infinity norm-- so this is almost, but not quite the same as the sup-- so almost the same as the supremum, but not quite because I need to ignore subsets of measure 0.

So I can write down a formal definition in a second. But I need to-- if I change  $W$  on the subset of the measure 0, I shouldn't change any of these norms. And so the one way to define this essential supremum-- it's called an essential sup-- is that it is the largest-- so it is the smallest  $\lambda$  such that-- so the smallest number  $m$  such that the measure of the set taking value bigger than  $m$  this set has measure 0. So it's the threshold above which you-- this, it has measure 0.

And the  $l_2$  norm will play a particularly special role. And for the  $l_2$  norm, you're really in the Hilbert space, in which case we are going to have inner products. And we denote inner products using the square-- using these brackets.

So everything is real. I don't have to worry about complex conjugates. So comparing with the discussion up there, we see that a sequence of-- so sequence  $G_n$  of quasirandom graphs has a property that the associated graphons converge to  $p$  in the cut norm.

For quasirandom graphs, there is no issue having to do with permutations because the target is invariant upon permutations. But if I give you two different graphs, then I need to think about their permutations. And to study permutations of vertices, the right way to do this is to consider measure-preserving transformations.

So we say that  $\phi$  from the interval to the interval is measure-preserving because first of all, it has to be a measurable map. And everything I'm going to talk about are measurable. So sometimes I will even omit mentioning it. So it is measure-preserving if, for all measurable subsets  $A$  of this interval, one has that the pullback of  $A$  has the same measure as  $A$  itself.

Let me give you an example. So you have to be also slightly careful with this definition if you think about the pushforward that's false. It has to be the pullback.

So for example, the map which sends-- so an easy example, the map which sends  $x$  to  $x$  plus  $1/2$ -- so think about a circle as your space. And here I am just rotating the circle by one-half rotation. So it's obviously measure-preserving. I am not changing any measures.

Slightly more interesting example, quite a bit more interesting example is setting  $x$  to  $2x$ . This is also measure-preserving. And you might be puzzled for a second why it's measure-preserving because it sounds like it's dilating everything by a factor of 2.

But if you look at the definition-- and so here is again mod 1. If you look at the definition, if you look at, let's say, a subset  $A$ , which is-- so what should I think? For example, so if that is my  $A$ , so what's the inverse of  $A$ ? So it's this set.

So the measure is preserved upon this pullback. And so if you pushforward, then you might dilate by a factor of 2. But when you pullback, the measure gets preserved.

So these measure-preserving transformations are going to play role of permutations of vertices. So it turns out that these things are actually-- they are quite subtle technically. And I am going to, as much as I can, ignore some of the measure theoretic technicalities. But they are quite subtle.

So for example, so now let me give you a definition for the distance between two graphons. I write, starting with a symmetric measurable function  $W$ , so I write  $W^\phi$  to denote the function obtained as follows. So I think of this as relabeling the vertices of a graph. And now I define this distance. So this is going to be called the cut distance between two symmetric measurable functions,  $U$  and  $W$ , to be the infimum over all measure-preserving bijections.

So this is the definition for the distance between two graphons. To take the optimal-- and my question does it-- I am looking at  $\inf$ . So I haven't told you yet whether you can take a single one. And it turns out that's a subtle issue. And generally it doesn't exist.

But I look over all measure-preserving bijections  $\phi$ . And I look at the distance between  $W$  and  $W^\phi$ , optimized over the best possible measure-preserving bijection. So this  $\inf$  is really an  $\inf$ . It's not always obtained. And actually, this example here is a great example for-- you can create an example for why  $\inf$  is not always obtained from the discussion over here.

For example, if  $U$  is the function  $x$  times  $y$ , this is a graphon  $xy$  and  $W$  is  $U \circ v$ , where  $v$  is the map distance  $x$  to  $2x$ , then in your mind, you should think of these two as really the same graphons. You are applying the measure-preserving transformation. It's like doing a permutation. But because  $\phi$  is not bijective, you cannot putting  $\phi$  here to get these two things to be the same.

So there are some subtleties. So this is really an example just to highlight there's some subtleties here, which I am going to try to ignore as much as possible. But I will always give you correct definitions. Any questions? Yeah? Yeah?

**AUDIENCE:** So can we expect the cut distance between these two sets to be 0 [INAUDIBLE]?

**YUFEI ZHAO:** So the question, do we expect the cut distance between these two to be 0? And the answer is yes. So we do expect them to be 0. And they are 0. They are equal to 0.

And let me just tell you one something that is new. And this is one of those statements that has a lot of measure theoretic technicalities. For all graphons  $U$  and  $W$ , it turns out that there exist measure-preserving maps-- so not necessarily bijections, but measure-preserving maps from  $[0, 1]$  interval to itself, such that the distance between  $U$  and  $W$ , the cut distance, is obtained by the cut norm difference between-- the difference between  $U \circ \phi$  and  $W \circ \psi$ . So don't worry about it. So far, we have defined this notion of a cut distance between two graphons. But now I'll give you two graphs. So what do you do for two graphs? Or I can-- yeah?

**AUDIENCE:** You can take the graphon associated it.

**YUFEI ZHAO:** Great. So take the graphon associated with these graphs and consider their cut distance. So for graphs  $G$  and  $G'$ , and potentially even a different number of vertices, I can define the distance, the cut distance between these two graphs to be the distance between the associated graphons. And similarly, if I have a graph and a graphon, I can also compare their distance.

So what does this actually mean? So if I give you two graphs, even with the same number of vertices, it's not quite the same thing as a permutation of vertices. It's a bit more subtle. Now why is it more subtle than just permuting the vertices?

So here we are using measure-preserving transformations, which doesn't see your atomic vertices. So we might split up your vertices. So you might take a vertex and chop it in half and send one half somewhere and another half somewhere else because these guys, they don't care about your vertices anymore. So it's not quite the same as permuting vertices.

But it's some kind of-- so you allow some kind of splitting and rearrangement and overlays. So you can write out this distance in this format, find out the best way to split and overlay and to rearrange that way. But it's much cleaner to define it in terms of graphons. Yes?

**AUDIENCE:** Is this why we take bijections up there [INAUDIBLE]?

**YUFEI ZHAO:** The question is, is that why we take bijections up there? And no, so up there, if I wrote instead measure-preserving maps, it's still a correct definition and it's the same definition. And the fact that these two are equivalent goes to some measure theory, which I will not-- do not want to indulge you Great. But the moral of the story is you take two graphons and rearrange the vertices in some way, in the best way, overlay them on top of each other and take the difference and look at the cut norm. And so that's the distance.

So I want to finish by stating the main theorems that form graph limit theory. And these address the questions I mentioned right before the break. So do there exist limits? And do these two different notions of one having to do with distance and another having to do with homomorphism densities, how do they relate to each other? Are they consistent?

So the first theorem, Theorem 1, has to do with the equivalence of the convergence, namely, that if you have a sequence of graphs or graphons, the sequence is convergent in the sense, up there, if and only if they are convergent in the sense of-- in this metric space. So remember what convergence means in the metric space is that of a Cauchy sequence-- so if and only if it is a Cauchy sequence with respect to this cut distance. So it's just-- maybe for many of you, it's been a while since you took 18-100. So let remind you a Cauchy sequence, in this case, it means that, if I look at the distance between two graphs, if I look far enough out, then I can contain the rest of the sequence in an arbitrarily small ball. So the sup positive  $\epsilon$  of this guy here, goes to 0 as  $n$  goes to infinity.

But because we don't know yet whether the limit exists, so I can't talk about them getting closer and closer to a limit. But they mutually get closer to each other. So Theorem 1 tells us that these two notions, one having to do with homomorphism densities, is consistent and in fact equivalent to the appropriate notion in the metric space.

So let's use a symbol. So we say that  $G_n$  converges to  $W$ , or in the case of a sequence of graphons. So we can do that as well.

So here we say that  $G_n$  converges with  $W$ , if whenever you look at the  $f$  density in  $G_n$ , this sequence converges to the corresponding  $f$  density in  $W$  for every  $f$ , and similarly, if you have a graphon instead of a graph. So that definition was just whether a sequence is convergent. Here it converges to this graphon  $W$ .

And the question is, if you give me a convergent sequence, is there a limit? Does it converge to some limit? And the answer is yes. And that's the second theorem, which tells us the existence of a limit, of the limit object. So the statement is that every convergent sequence of graph or graphons has a limit graphon.

So now I want you to imagine this space of graphons. So we'll have this space containing all the graphons. And let me denote this space by this curly  $W_0$ . So this, the 0 is-- don't worry about it. It's more just convention.

But let me also put a tilde on top for the following reason. Let this be the space of graphons where we identify graphons with distance 0. So then the space combined with this metric is a metric space. It is the space of graphons.

And so the third theorem is that it's the compactness of the space of graphons, namely, that this space is compact. Because we're in the metric space, compactness in the usual sense of every open cover has a finite subcover is equivalent to the slightly more intuitive notion of sequential compactness-- every sequence has a convergent subsequence. And then it's also, if you have a limit, so it converges to some limit.

So how should you think of Theorem 3? So it's about compactness and some tautological notion. But intuitively, you should think of compactness as saying-- and the English word, the English meaning of the word compact is small.

You should think of this space as being quite small, which is rather counterintuitive because we're looking at the space of graphons, certainly at least as large as the space of graphs, but really all functions from the square to the interval. This seems like a pretty large space. But this theorem here says that, in fact, that space is quite small. And where have we also seen that philosophy before?

So in Szemerédi's Graph Regularity Lemma, the underlying philosophy there is that, even though the possibilities, the space of possibilities for graph is quite large, once you apply Szemerédi's Regularity Lemma, and once you are OK with some epsilon approximations, there is only a small description, this bounded description, of a graph. And you can work with that description. And these two philosophies, it's no coincidence that they are consistent with each other because we will use Szemerédi's Regularity Lemma to prove this compactness.

In fact, we will use a slightly weaker version of Szemerédi's Regularity Lemma to prove

compactness. And then you will see that, from the compactness, one can use properties of the compactness to boost to a stronger version of regularity. But the underlying philosophy here is that this compactness is in some sense a quantity. It's a qualitative reformulation, analytic reformulation of Szemerédi's Graph Regularity Lemma. OK, so--

So this topic, this graph limits, which we'll explore for the next few lectures, including giving a proof of all three of these main theorems, nicely encapsulates the past couple of topics we have done. So on one hand, Szemerédi's Regularity Lemma, or some version of that, will be used in proving the existence of the limit and also the compactness. And also it's philosophically and in some sense related and very much equivalent in some sense and related to these notions. It is also related to quasirandomness-- in particular, quasirandom graphs that we did a few lectures ago, where in quasirandom graphs, we are really looking at the constant graphon in this language.

And now we expand our horizons. And instead of just looking at the constant graphon, we can now consider arbitrary graphons. They are also this model for a very large graph. Any questions? Yeah?

**AUDIENCE:** Can we prove the theorem analytically and then deduce the Regularity Lemma with it?

**YUFEI ZHAO:** The question is, can we prove Theorem 3 analytically and deduce the Regularity Lemma? So you will see once you see the proof. It depends on what you mean. But roughly, the answer is yes.

But there's a very important caveat. It's that, because we are using compactness, any argument involving compactness gives no quantitative bounds. So you will have a proof of the Szemerédi Regularity Lemma that tells you there is a bound for each epsilon. But it doesn't tell you what the bound is. Yeah?

**AUDIENCE:** Doesn't Theorem 3 imply Theorem 1 because of the [INAUDIBLE]?

**YUFEI ZHAO:** Does Code Theorem 3 imply Theorem 1? And the answer is no because in Theorem 1, the notion of convergence is about homomorphism densities. So Theorem 1 is about these two different notions of convergence and that they are equivalent to each other. Theorem 3 is just about the metric. It's about the cut metric.

And so Theorem 1 is-- the point of Theorem 1 is that you have these two-- you have these two notions of convergence, one having to do with subgraph densities and the other having to do

with a cut distance. And in fact, they are equivalent notions. So all great questions-- any others?

**AUDIENCE:** And for that  $F$ , is  $F$  a graphon because the [INAUDIBLE]? Is  $F$  a graphon or a graph?

**YUFEI ZHAO:** The question is, is  $F$  a graph or a graphon?  $F$  is always a graph. So in  $t F, W$ , I do not define this quantity for graphon  $F$ . So this quantity here, I have only allowed the second argument to be a graphon. The first argument is not allowed to be a graphon. It doesn't make sense. Yeah?

**AUDIENCE:** Doesn't Theorem 1 and 2 together imply Theorem 3?

**YUFEI ZHAO:** The question is, doesn't Theorem 1 and Theorem 2 together imply Theorem 3? So first of all, Theorem 1 is really-- it's not about compactness. So it's really about the equivalence of two different notions of convergence. It's like you have two different metrics. I am showing that these two metrics are equivalent to each other.

Theorem 2 and Theorem 3 are quite intimately related. So Theorem 2 is about-- Theorem 2, so they are quite related. But they're not quite the same.

So let me just give you the real line analogy, going back to what we said in the beginning. So Theorem 2 is kind of like saying that the real numbers is complete. Every convergent sequence has a limit, whereas Theorem 3 is more than that.

It's also bounded in some sense. But here, there is no notion of bounded. It's compact. But the main-- you should think of these two are very much related to each other. But here it's-- but they are not equivalent.

Anything else? Great. So that's all for today.