S. GEORGE DJORGOVSKI: So that was just the qualitative description, but we'd like to associate some numbers to be really rigorous about this. And so how do we do this? How do we actually quantify distribution of galaxies at large scales? The first thing that was suggested was so-called, a 2-point correlation function, which is accessed probability of finding one galaxy next to another galaxy at some distance r.

Now, if galaxies were all uniformly randomly distributed, say like molecules of gas in this room. Then, there will be constant density, and probability will be exactly proportional to density as uniform. But if galaxies are bunched together, then you find there'll be extra probability at smaller scales.

So you compute what is the actual probability just by counting galaxies in spheres around each other. And then you subtract 1 from that ratio, because it's called the excess probability. So that correlation function for uniformly randomly distributed field of galaxy will be 0. No correlation, that's what it means.

Instead of that, there is clustering and it turns out to be well described by a power wall, with a slope of minus 1.8, and it looks like this. So this is the log of the correlation function, which is actually probably of finding galaxy at some separation versus log of their separation in megaparts 6, which you can't see. And it's almost perfect, but not exactly perfect power law.

At some point, it has to cross 0 and become negative, because if you have excess of galaxies at small separations, that in order to achieve the mean, you have to have deficit and some other separation. And indeed at large separations, this turns to be negative and that's the voids.

So in some sense, you can think of from the uniform distribution, galaxies evacuated from the voids and piled up in the structures that we see, filaments and clusters and what have you. So we can then look what galaxies of different kind do in terms of clustering. And so they're interesting phenomena there. For example, turns out that the bright galaxies cluster more strongly than faint galaxies. High density galaxies cluster more strongly than low density galaxies. Early Hubble types, ellipticals cluster more strongly than spirals and so on.

And so these are all hints about how galaxies form. So in fact, we now understand fairly well, why the galaxies are different kind with different star formation histories cluster large scale differently. And in some sense this is telling you that there is always a connection between large scale environmental galaxies and how they evolve.

Now a more modern way to quantify galaxy clustering is so-called, power spectrum. And if you are familiar with Fourier transforms, which is a standard tool in many different fields, engineering and science. Basically, you decompose the dense field of the universe into whole lot of sound waves in 3D with different amplitudes and random orientation of phases.

And you can then ask the question, how much mass is clumping on what spatial scale? So the Fourier spectrum of sound base will be a very low notes. And there would be a lot of power on low frequencies. And travel will be at low power high frequencies and so on. So you can do the same thing for anything, not just sound, but densities of galaxies in the universe. And you can compute that.

Now, this has the advantage of being easily compared to theoretical models, but its mathematically equivalent to the 2-point correlation function, their so-called, Fourier pair of functions. So one or the other, nowadays this is the more popular way of doing it.

Just to illustrate that, here is some fake distributions. Each corresponds to one almost delta function within narrow distribution of power and particular spatial scale, and depending on what is the location of the spike in the wavelength space. But then tells you what structure you're going to see. If it's at high spatial frequency, there will be a lot of small blobs. If it's a low spatial frequency, there will be few large blobs.

Now, of course, in real life, there is a mixture in all different scales. And when you actually go ahead and measure this, you get something like this. Now this is state of the art measurement or many orders of magnitude. And this is shown in unit is spatial frequency, 1 over wavelength. So it's the large structures on the left, and small structures on the right. And it's a log log plot.

So we can see at very large scales, there is more power at smaller scales. There are few really, really big blobs and more smaller ones. But then it turns over, and cuts off, and somehow that extrapolation is missing. And now we understand why that happens, and this is due to sample cold dark matter. But this is the kind of modeling that's now compared to simulation, observations that one can then try to constrain different models of structured formation.

So all this has been telling you how much mass is clustered at what spatial scale. But it's saying exactly nothing about how is that arranged in space on large scales. And that, in Fourier speak, is phased distribution, which is lost in power spectrum . So you've seen these maps. Now, the real universe looks just like the simulations, sheets, filaments, connections.

And on the right, there is a nice illustration of what it means to ignore the phased distribution. So the image on the top is so-called, Voronoi foam. It's just a fake density distribution. It looks spongy. So we take its Fourier transform, and compute its power spectrum. But take phases of that Fourier transform, scramble them randomly, and then transform back, and you get the image on the bottom. It looks just like a total mess.

So those two density fields have identical power spectrum by construction. One of them has ordered phases. The other one has random phases. And believe it or not, even though so many smart people have worked on this for so many years, decades, we still don't have a reasonably good way of quantifying phase coherent distribution in the universe. And would seem that this is a pretty important part of the total description.

So this is a really good task for you guys. If you figure out how to do something as simple and elegant as 2-point correlation functional or power spectrum. But for the phase distribution, which then tells you about topology of large scale structure. Why does it sponge as opposed to whole bunch of little blobs? That would be a really good thing to do.