Lecture 12: Random fields: correlation function, power spectrum, and formation of halos.

L: Some unknown random process is responsible for the initial fluctuations in cosmology. The favorite hypothesis at present for the origin of these fluctuations is quantum fluctuations during the epoch of inflation. Whatever they are, the origin and nature of the fluctuations is a fundamental question in cosmology.

L: The density fluctuation is a *random field*. We have a random number at each point in space, $\delta(\vec{r})$, where the mass density is:

$$\rho(\vec{r}) = \bar{\rho}[1 + \delta(\vec{r})]. \tag{1}$$

However, these random numbers form a sort of continuous function, so there is obviously a correlation among the values of δ at nearby points. We define the correlation function:

$$\xi(\vec{r}) = <\delta(\vec{x}) \cdot \delta(\vec{r} + \vec{x})> \quad . \tag{2}$$

The average here is understood as an *ensemble* average: we have a random realization out of a statistical distribution of random fields. We average over a very large number of realizations. The ergodic theorem tells us that we can replace the average over random realizations by an average over a very large volume in one realization. Of course, in practice we have only one Universe we can observe. Actually, we have only our observable horizon to observe, so we never really have an arbitrarily large volume. Cosmologists simply hope that the region of our Universe we can observe within our present horizon is representative of a theory describing all the Universe: we cannot do any better.

Normally, we expect the correlation function $\xi(\vec{r})$ to be homogeneous and isotropic (to depend only on the modulus r, and not on $\vec{x}$), unless the Universe violates the cosmological principle and is inhomogeneous or anisotropic in the statiscal properties of the fluctuation correlations. So, there are primordial small amplitude fluctuations, but we assume homogeneity and isotropy holds *statistically* on the properties of these fluctuations, on all scales.

Q: How do we measure the correlation function? In practice, we often measure only from observations of the galaxy distribution, defining a galaxy density $n_g(\vec{x}) = \bar{n}_g[1 + \delta_g(\vec{x})]$. We then measure the galaxy correlation:

$$\bar{n}_g \xi_g(\vec{r}) = \frac{1}{\bar{n}_g} <n_g(\vec{x}) \cdot n_g(\vec{r} + \vec{x})> \quad . \tag{3}$$

Note that $\bar{n}_g \xi_g(\vec{r})$ is also the average number of galaxies around any random galaxy. In general, $\bar{\xi}(\vec{r})$ is the average overdensity around a random mass particle, as a function of the distance $\vec{r}$.

So, we can use galaxy surveys to measure their correlation function, but what is the relation between galaxy and mass correlation functions? In other words, how do galaxies trace mass? A common assumption that is made is *linear bias*:

$$\delta_g(\vec{x}) = b_g \delta(\vec{x}) \ . \tag{4}$$

One can show this linear bias is valid in the linear regime, when we average the galaxy distribution over regions large enough to have small amplitude density fluctuations at present, and if the process of galaxy formation depends only on local properties (the density field in a small region around the galaxy compared to the scale in which we measure the correlation).

Q: Can we observe the mass density fluctuations directly? Yes, for example using weak gravitational lensing, and also through the CMB fluctuations using the theory of how CMB intensity is determined by the underlying evolution in the mass and radiation density fluctuations.

L: Now, we consider the Fourier transform of the density field:

$$\delta(\vec{k}) = \int d^3x \, \delta(\vec{x}) e^{i\vec{k}\cdot\vec{x}} \ . \tag{5}$$

Q: What is the correlation of Fourier modes? We can calculate this under the assumption of homogeneity:

$$<\delta(\vec{k}) \cdot \delta^*(\vec{k}')> = \int d^3x\, d^3x'\, e^{i(\vec{k}\cdot\vec{x}-\vec{k}'\cdot\vec{x}')} <\delta(\vec{x}) \cdot \delta^(\vec{x}')> = \tag{6}$$

$$= \int d^3\left(\frac{\vec{x}+\vec{x}'}{2}\right) d^3\left(\vec{x}-\vec{x}'\right) e^{i(\vec{x}+\vec{x}')(\vec{k}-\vec{k}')/2+(\vec{x}-\vec{x}')(\vec{k}+\vec{k}')/2}\, \xi(\vec{x}-\vec{x}') = \tag{7}$$

$$= \int d^3s\, e^{i\vec{s}\cdot(\vec{k}-\vec{k}')} \int d^3r\, e^{i\vec{r}\cdot(\vec{k}+\vec{k}')/2}\, \xi(\vec{r}) = \tag{8}$$

$$= (2\pi)^3\delta^D(\vec{k}-\vec{k}') \int d^3r\, e^{i\vec{k}\cdot\vec{r}}\, \xi(\vec{r}) = (2\pi)^3\delta^D(\vec{k}-\vec{k}')\, P(\vec{k}) \ . \tag{9}$$

In this equation we have defined the power spectrum, which represents the characteristic amplitude of Fourier modes of wave number $\vec{k}$. Just like the correlation function, if the fluctuations are isotropic in their statistical properties, the power spectrum should be isotropic, and is defined by:

$$P(k) = \int d^3r\, e^{i\vec{k}\cdot\vec{r}}\xi(r) \ . \tag{10}$$

Because of the Dirac delta function, we find that homogeneity in the correlation function implies that there is no 2-point correlation in Fourier modes of different wave number: two modes are always independent, although in general there may still be correlations when you look at 3 or more modes. The correlation function is given by the inverse Fourier transform:

$$\xi(r) = \int \frac{d^3k}{(2\pi)^3}\, e^{-i\vec{k}\cdot\vec{r}} P(k) \ . \tag{11}$$

It is also useful to define the characteristic amplitude of fluctuations on the Fourier scale k, which is contributed by all Fourier modes on a sphere in Fourier space of radius k, so:

$$\Delta^2(k) d\log k = \frac{4\pi k^2 dk}{(2\pi)^3} P(k) \ ; \qquad \Delta^2(k) = \frac{k^3 P(k)}{2\pi^2} \ . \tag{12}$$

Q: Example: if $P(k) \propto k^n$, how does the typical relative mass fluctuation depend on scale?

$$\left(<\delta^2(R)>\right)^{1/2} \propto \Delta^2(k \sim \pi/R) \propto R^{-(3+n)/2} \ . \tag{13}$$

Q: In the matter-dominated epoch, how does the characteristic epoch of collapse of halos scale with their mass?

$$a_{\rm col} \propto \left(<\delta^2(R)>\right)^{-1/2} \propto R^{(3+n)/2} \propto M^{(3+n)/6} \ , \tag{14}$$

where the last equality is because halos form from the mass contained in a region of comoving size R, $M \sim R^3$.

Q: In the matter-dominated universe, how does the characteristic velocity dispersion of halos scale with their mass?

$$\sigma^2 \sim GM/(a_{\rm col}R) \propto M^{2/3}/a_{\rm col} \sim\propto M^{(1-n)/6} \ . \tag{15}$$

So, this tells us how the characteristic velocity dispersion of halos formed during the formation of structure will vary with the halo mass depending on the slope of the power spectrum.

Q: Example: what happens if the initial fluctuations are simply derived from a process of randomly distributed seeds or galaxies, or a Poisson process? The correlation function is then a delta function,

$$\xi(x) = \bar{n}_g^{-1} \delta^D(\vec{x}) \ , \tag{16}$$

where the normalization is determined by requiring that $\bar{n}_g \int d^3x \xi(x) = 1$. So the power spectrum is:

$$P(k) = \int d^3r \, e^{i\vec{k}\cdot\vec{r}} \xi(r) = \frac{1}{\bar{n}_g} \ . \tag{17}$$

The power spectrum is a constant. As you expect in a Poisson process, the fluctuation scales as the inverse square root of the total number of galaxies (or mass) contained, $\left(<\delta^2(R)>\right)^{1/2} \propto R^{-3/2} \propto M^{-1/2}$.

L: Gaussian fields: These are special random fields where the correlation function, or its Fourier transform the power spectrum, fully defines all the statistical properties of the random realizations of the field. Each Fourier mode has a Gaussian-distributed amplitude regulated by the power spectrum, and a random phase, and the Fourier modes have no further correlations among them at the 3-point or any n-point level: they are totally independent.