Lecture 4

Jarle Brinchmann

25/02/2014

1 Introduction

Last week we covered the final part of the linear growth of structure and looked at how the fact that we have dark matter helps baryons perturbations form and grow after decoupling. In this lecture we will towards the end take a step away from linear perturbations and move into the non-linear regime, which after all is where galaxies form.

2 The power spectrum

Last week we saw that the correlation function, $\xi(r)$, is the Fourier transform of the power spectrum, in the sense that

$$\xi(r) = \int \frac{d^3\vec{k}}{(2\pi)^3} P(k) e^{-i\vec{k}\cdot\vec{x}}. \tag{1}$$

The power spectrum can be defined, or taken to be

$$P(k) \propto \left\langle |\delta_k|^2 \right\rangle, \tag{2}$$

and we also defined the power per decade, $\Delta^2(k)$ to be

$$\Delta^2(k) = \frac{1}{2\pi^2} k^3 P(k). \tag{3}$$

The primordial power-spectrum we will assume as given — inflationary theories typically predict that $P(k) \propto k^n$ at some early time for instance. We are then concerned with how this evolve with time and encapsulate this process with the *transfer function*, $T(k)$, which we typically write as

$$P(k;t) \propto P(k;t_i) T^2(k). \tag{4}$$

We can determine the form of this function from the growth of perturbations which we covered last week.

We first make the assumption that the primordial power spectrum at time t_i was a power-law

$$P(k; t_i) = Ak^n, \tag{5}$$

which is a common assumption as we have no accepted theory for the initial perturbations. In this case the growth can be written as:

$$P(k; t) \propto A \left(\frac{a(t)}{a_{\rm eq}}\right)^2 \begin{cases} 0 & r < r_{\rm FS} \\ k^{n-4} & r_{\rm FS} < r < r_{\rm eq} \\ k^n & r > r_{\rm eq} \end{cases}, \tag{6}$$

where the most important aspect of this equation is the scale dependence and I will suppress the time-dependence in general in the following. From this we can observe that on scales larger than the horizon at matter-radiation equality, $r > r_{\rm eq} \approx 100$Mpc, we expect to see the shape of the primordial power spectrum.

Observationally $n \approx 1$ and $n = 1$ is known as the Harrison-Zeldovich (H-Z) spectrum. The choice of a H-Z spectrum is natural for a number of reasons. One reason is that it can be shown that $k^3 P(k)$ at the time of horizon entry is scale-free if $n = 1$, implying that no particular scale is picked out since the power will be equal in each logarithmic interval. Another reason has to do with variations in the gravitational potential. If we write the Poisson equation in proper coordinates for a perturbation δ giving rise to a perturbation of the potential $\delta\Phi$, we have

$$\nabla^2 \delta\Phi = 4\pi G \rho_b \delta \tag{7}$$

which if we Fourier transform it gives

$$\delta\hat{\Phi}_k = -4\pi G \rho_b \frac{\hat{\delta}_k}{k^2} \tag{8}$$

which means that

$$\Delta^2_{\delta\hat{\Phi}_k} \propto k^3 \langle|\delta\hat{\Phi}_k|^2\rangle \propto k^3 \langle|\hat{\delta}(k)|^2\rangle \frac{1}{k^4} \propto k^{n-1}, \tag{9}$$

which says that if $n = 1$ no scale will contribute excessively to the overall variation in Φ. This is merely a consistency argument but $n = 1$ does have a range of such nice properties thus it is a favoured theoretical assumption.

3 Window functions

A density estimate at a point is rather meaningless (at least from an operational point of view), because it requires a volume. While one can mathematically define the density at a point with a limiting process this is not particularly useful for us. What we instead want to do is to apply some sort of filter to the density field and measure the density within this filter. To do this we introduce the concept of **window functions**.

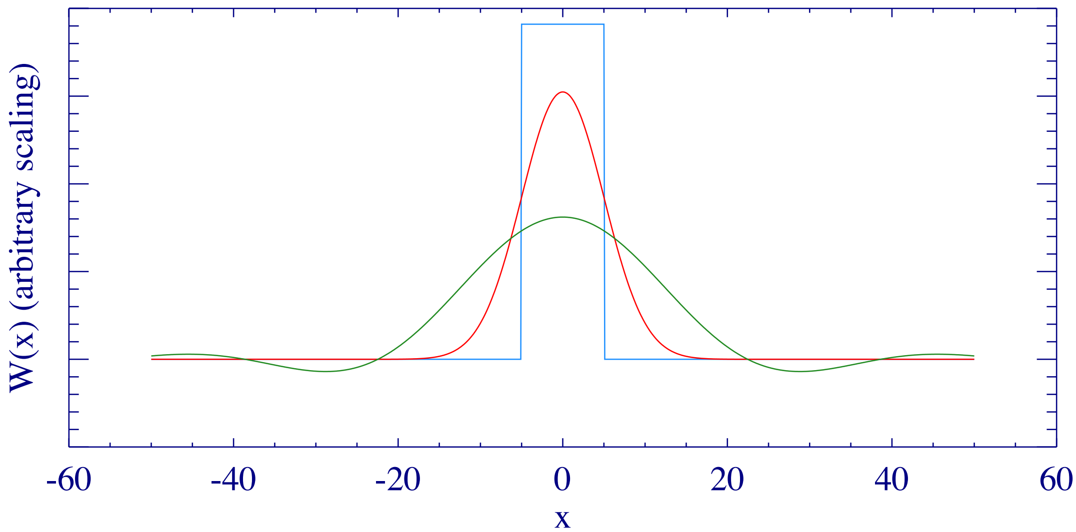

Figure 1: A comparison of the three window functions discussed in the text. The blue shows the top-hat, the red the Gaussian and the green the sharp k-space window function. I set $R = 5$ for this particular plot.

With a window function, $W(\vec{x}; R)$ we can define a smoothed density field, $\delta(\vec{x}; R)$ through:

$$\delta(\vec{x}; R) = \int \delta(\vec{x'}) W(\vec{x} - \vec{x'}; R) d\vec{x'} \tag{10}$$

Since this is a convolution, we know from the convolution theorem in Fourier analysis that in Fourier space this is a simple multiplication of Fourier transforms — this can be very useful in practice:

$$\hat{\delta}(\vec{k}; R) = \hat{\delta}(\vec{k}) \hat{W}(\vec{k}; R). \tag{11}$$

In principle any function could be a window function (as long as they are reasonably integrable), but in cosmology there are three main window functions that are mostly used. They are illustrated in Figure 1 and mathematically they are:

The spherical top-hat This is a function that is defined through

$$W_{\rm TH}(\vec{x}; R) = \frac{1}{V_{\rm TH}} \begin{cases} 1 & |x| \leq R \\ 0 & |x| > R \end{cases}, \tag{12}$$

where $V_{\rm TH} = 4\pi R^3/3$. Its Fourier transform is

$$\hat{W}_{\rm TH}(\vec{k}; R) = \frac{3(\sin kR - kR \cos kR)}{(kR)^3} \tag{13}$$

The Gaussian window is defined through

$$W_G(\vec{x}; R) = \frac{1}{V_G} e^{-|x|^2/2R^2}, \tag{14}$$

which has a volume, $V_G = (2\pi)^{3/2} R^3$, and its Fourier transform is:

$$\hat{W}_G(\vec{k}; R) = e^{-(kR)^2/2}. \tag{15}$$

Sharp k-space filter This is a filter that is particularly useful for theoretical arguments. It is the equivalent to the top-hat filter in Fourier space. So it is defined in Fourier space as

$$\hat{W}_{\mathrm{kTH}}(\vec{k}; R) = \begin{cases} 1 & |kR| \leq 1 \\ 0 & |kR| > 1 \end{cases} . \tag{16}$$

In real space this takes the form:

$$W_{\mathrm{kTH}}(\vec{x}; R) = \frac{3}{V_{\mathrm{kTH}}} \left| \frac{\vec{x}}{R} \right|^{-3} \left(\sin \frac{|x|}{R} - \frac{|x|}{R} \cos \frac{|x|}{R} \right), \tag{17}$$

where the volume is defined to be $V_{\mathrm{kTH}} = 6\pi^2 R^3$, which is a convenient definition because it satisfies $W_{\mathrm{kTH}}(0; R)V(R) = 1$, but the integral of W_{kTH} over all space (which would normally define the volume) diverges logarithmically so is not defined.

For each of these window functions we can then define a mass through:

$$M(R) = \bar{\rho} V_W(R), \tag{18}$$

where R is the co-moving radius under consideration and $\bar{\rho}$ is the mean density of the Universe.

4 Statistical properties of the filtered density field

Looking back at equation 10 we see that $\delta(\vec{x}; R)$ is a linear combination of $\delta(\vec{x})$. In the linear regime we know that δ is distributed as a Gaussian. It is then a standard theorem that says that a linear combination of Gaussian distributions is again a Gaussian. Thus to characterise the field $\delta(\vec{x}; R)$, we need the mean and the variance of the field.

The mean of the field is 0, because when you take an average of δ over space you must get zero because the density on average must by definition be equal to the mean density of the Universe. The variance of the field can also be easily calculated (consult the calculation of the link between the correlation function and the power spectrum if you are uncertain):

$$\sigma^2(R) \stackrel{\text{def}}{=} \langle \delta(\vec{x}; R)^2 \rangle \tag{19}$$

$$= \frac{1}{2\pi^2} \int k^3 P(k) \hat{W}(k; R)^2 \frac{dk}{k} \tag{20}$$

$$= \int \Delta^2(k) \hat{W}(k; R)^2 \frac{dk}{k}, \tag{21}$$

where I have assumed that $P(k)$ and the window function is isotropic.

With this we can then write

$$P\left(\delta(\vec{x};R)\right)d\delta = \frac{1}{\sqrt{2\pi}\sigma(R)}e^{-\delta^2/2\sigma^2(R)}d\delta, \tag{22}$$

which is a very useful identify for use when calculating characteristics of the density field.

5 Mass variance

The preceding gave a theoretical exposition. In practice one might more reasonably look at the possibility of calculating the variance of mass contained within a volume with radius R as you sample different parts of the Universe. In this case we can define, for each volume, a mass:

$$M(\vec{x};R) = \rho(\vec{x};R)V(R), \tag{23}$$

where $V(R)$ is the volume of the window function considered (given above). We can then define the dimensionless mass variance as

$$\sigma^2(M) = \frac{1}{N}\sum\left(\frac{M(\vec{x},R)-\bar{M}}{\bar{M}}\right)^2 = \left\langle\left(\frac{M(\vec{x},R)-\bar{M}}{\bar{M}}\right)^2\right\rangle_{\vec{x}}, \tag{24}$$

where a bar denotes average values, and where the sum is over the N regions sampled. This is a spatial average, emphasised by the subscript $\vec{x}$ in the second equality.

At the same time we can write the mass entires as:

$$M(\vec{x};R) = [1+\delta(\vec{x};R)]\bar{\rho}V(R) \tag{25}$$

$$\bar{M} = \bar{\rho}V(R), \tag{26}$$

which means that we have

$$\frac{M(\vec{x};R)-\bar{M}}{\bar{M}} = \frac{\delta M}{M} = \delta(\vec{x};R). \tag{27}$$

The second equality defines the commonly used notational shortcut, $\delta M/M$.

If we assume ergodicity, then

$$\sigma^2(M) = \sigma^2(R), \tag{28}$$

and we will make use of this throughout. Written out this gives:

$$\sigma^2(M) = \left\langle\left(\frac{M(\vec{x},R)-\bar{M}}{\bar{M}}\right)^2\right\rangle = \int \Delta^2(k)\hat{W}(k;R)^2\frac{dk}{k}, \tag{29}$$

which in the case of a sharp k-space filter with radius k_F, can be approximated by:

$$\sigma^2(M) \approx \Delta^2(k_F) = k_F^3 P(k_F). \tag{30}$$

In the case of other window functions the same scaling holds approximately. We can then introduce a power-law power spectrum, $P(k) \propto k^n$ and find:

$$\sigma^2(M) \propto k^{n+3} \propto M^{-(n+3)/3}. \tag{31}$$

To ensure that $\sigma^2(M)$ does not diverge on large scales, this requires that $n > -3$. Note that for this power spectrum the scaling is easily derived for all window functions (see problem set).

6 Non-linear scaling laws

We can use the formalism in the preceding to derive non-linear scaling laws. Ie. to estimate at what time/redshift, structures go non-linear. To do this, it is useful to recall the following simple scalings, which are valid in the linear regime in the matter dominated era:

$$\delta \propto a \propto t^{2/3} \tag{32}$$

$$\xi \propto P(k) \propto \delta^2 \propto a^2 \propto t^{4/3} \tag{33}$$

$$\sigma_M \propto a \propto t^{2/3}. \tag{34}$$

We now set the requirement for a perturbation to go non-linear as $\delta \approx 1$, or equivalently $\sigma(M) \sim 1$. Observationally we have that $\sigma(M) \sim 1$ on scales of about $8h^{-1}$Mpc. From a theory point of view we can then write:

$$\sigma_M(t_{\rm NL}) = \sigma_M(t_i)\left(\frac{t_{\rm NL}}{t_i}\right)^{2/3} = 1 \tag{35}$$

$$\Downarrow$$

$$t_{\rm NL} \propto \sigma_M(t_i)^{-3/2} \tag{36}$$

$$t_{\rm NL} \propto M^{(n+3)/4}. \tag{37}$$

So for $n > -3$ we have that smaller scales go non-linear earlier, while for $n = -3$ all scales collapse simultaneously. For CDM we can calculate this from the power spectrum and find that $n \sim -2$ on the scale of galaxies (see Figure 2). So in CDM structures collapse on small scales first.

The typical size of a give perturbation that has gone non-linear is

$$M \propto t^{4/(n+3)}, \tag{38}$$

while their typical proper size is:

$$L \propto Ra(t_{\rm NL}) \propto R t_{\rm NL}^{2/3} \propto M^{(n+5)/6}. \tag{39}$$

The density of an object forming at $t = t_{\rm NL}$ is some constant times the mean density at the time, which means that the density of the object is:

$$\rho \propto \rho_b(t_{\rm NL}) \propto t_{\rm NL}^{-2} \propto M^{-(n+3)/2}. \tag{40}$$

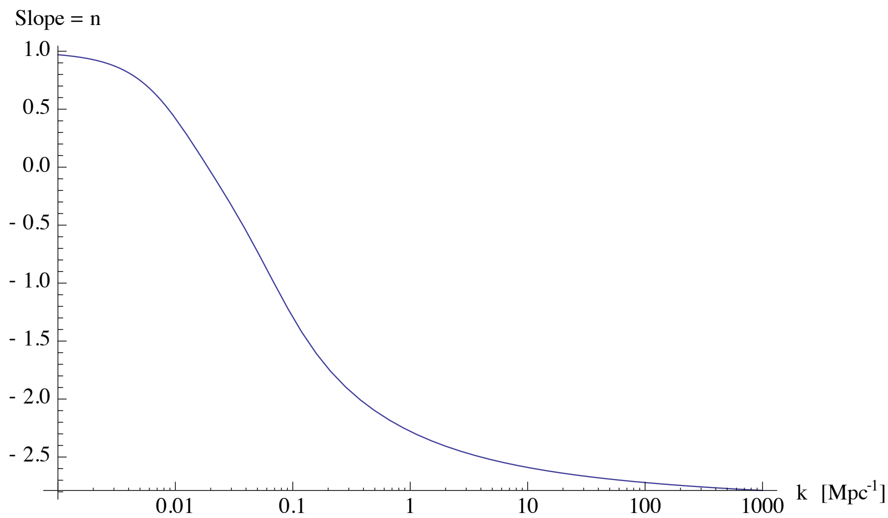

Figure 2: The slope of the power spectrum, $P(k)$, for CDM using the Bardeen et al (1986) transfer function, as a function of k in 1/Mpc. On galaxy scales (for the linear considerations) we have k $\sim$ 0.2–1.5 and a slope around -2.

Thus we see that for $n > -3$, bigger objects are less dense.

We can also write down the virial theorem:

$$\frac{M^2}{L} \sim Mv^2, \tag{41}$$

which gives us

$$v^2 \propto MM^{-(n+5)/6} \propto M^{(1-n)/6}, \tag{42}$$

with v^2 being the velocity dispersion of the system.

On galaxy scales, as mentioned above, the slope of the power spectrum is ~ -2, so if we insert that in the equation above we have

$$v \propto M^{(1-n)/12} \tag{43}$$

$$v \propto M^{1/4}, \tag{44}$$

and if galaxy luminosity, L, traces dark matter halo mass, M, then we get $L \propto v^4$, which is very close to the Faber-Jackson relation.

7 Spherical collapse

See e.g. section 5.1 in MvdBW; Section 5.10 in Pdm3; 14.1 in C&L

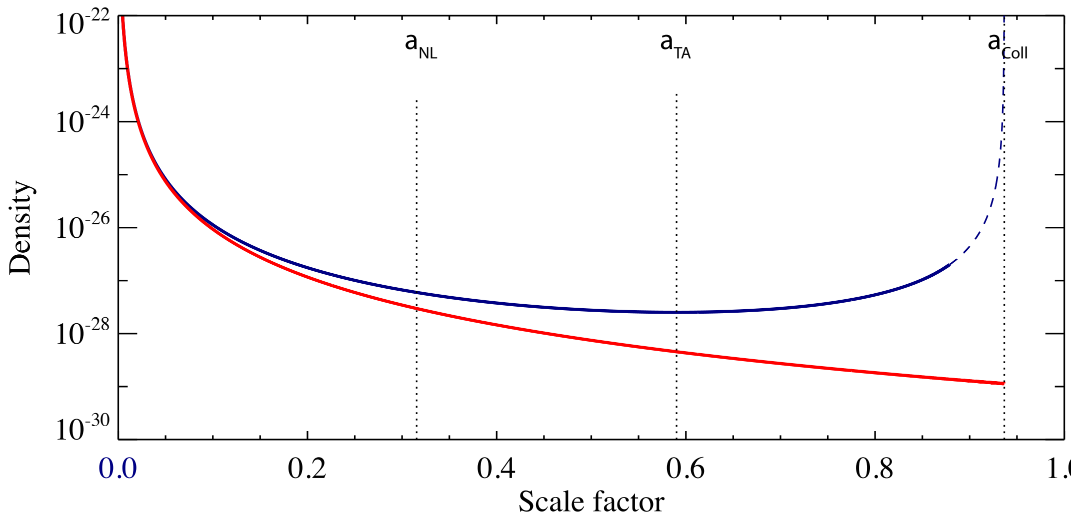

Figure 3: A calculation of the density evolution for a spherical top-hat perturbation as a function of a. The red line shows the evolution of the background Universe, for which the density declines $\propto a^{-3}$. The dark blue line shows the same for the spherical perturbation. The dashed line shows the evolution approximately after virialisation. The time of turn-around, non-linearity and collapse are also indicated.

In the preceding we focused on linear evolution, but to make progress we really have to get some understanding of how non-linear structures forms since that is what real galaxies are. A simple, but very useful, model for this is the spherical collapse model. The idea of this model is to consider a spherical overdensity in a flat, $\Omega_m = 1$, Universe. We can solve the evolution of the spherical overdensity exactly and it will evolve like a closed FRW Universe. In this case the solution is usually given in parametric form as

$$r = A(1 - \cos\theta) \tag{45}$$

$$t = B(\theta - \sin\theta), \tag{46}$$

where r is the radius of the overdensity and t the time-variable. From this we can immediately conclude that r will start small and expand until $\theta = \pi$. At this point, turn-around, $t_{\mathrm{TA}} = \pi B$. The second feature is that r becomes zero at $\theta = 2\pi$, at which point we can reasonably say that the object has collapsed. At that time, $t = 2\pi B = 2t_{\mathrm{TA}}$.

Using this model we can also estimate the overdensity of the non-linear structure and at the same time compare this to what linear theory would have predicted at the same point. The reason for this is that if we can find a good link between linear and non-linear theory, we can use linear theory (which is much easier to work with) to make predictions for the non-linear Universe. This gives a series of stages (see also Figure 3):

1. A scale will go non-linear at scale factor a_{NL}, at which point we have

$$\delta_{\mathrm{NL}} = 1 \qquad \delta_L \approx 0.57, \tag{47}$$

where δ corresponds to the exact solution and δ_L to the overdensity we would find if linear theory was correct.

2. As explained above, initially the perturbation will expand with the surrounding Universe, but at some point it will stop expanding and start to contract. The point at which expansion stops, we refer to as the time of turn-around, when the scale factor is $a_{\rm TA}$ and the time $t_{\rm TA}$. This happens when

$$\rho_{\rm TA} = (1+\delta_{\rm TA}) = \frac{9\pi^2}{16}\rho_b(t_{\rm TA}) \qquad\qquad \delta_L \approx 1.067, \tag{48}$$

   where ρ_b is the background Universe. So the overdensity at this time is already 4 times the surrounding Universe and clearly in the non-linear regime. The extrapolated linear overdensity would be 1.067 at this stage.

3. In a real object, deviations from spherical symmetry will ensure that it does not collapse to a point. Instead it will virialise. It is conventional to refer to the time of collapse as twice the time of turn-around, $t_{\rm coll} = 2t_{\rm TA}$, at which point we have (from the virial theorem) that the radius is $r_{\rm coll} = r_{\rm TA}/2$. The densities at that point are:

$$\rho_{\rm coll} = (1+\delta_{\rm coll}) = 18\pi^2\rho_b \approx 178\rho_b \qquad\qquad \delta_L \approx 1.69 = \delta_c. \tag{49}$$

   where we define a linear critical overdensity for collapse as $\delta_c = 1.69$ and this is a quantity we will use repeatedly later.

The exact results do depend on the exact cosmology chosen but the δ_c threshold is fairly robust to changes in the cosmology. In a short while we will use these results to derive the mass function of dark matter halos and look at how that compares with galaxies in the Universe.