# Lecture 5

Jarle Brinchmann

5/03/2013

## 1 Introduction

Last week we showed that the correlation function can be written as the Fourier transform of the power spectrum,

$$\xi(|\vec{x}_1 - \vec{x}_2|) = \xi(r) = \int \frac{d^3\vec{k}}{(2\pi)^3} P(k) e^{-i\vec{k}\cdot\vec{r}}, \tag{1}$$

and in the case of isotropy we can write

$$\xi(r) = \int \Delta^2(k) \frac{\sin kr}{kr} d\ln k. \tag{2}$$

We also looked at the practical way of measuring densities using a window function and defined the smoothed density field as

$$\delta(\vec{x}; R) = \int \delta(\vec{x'}) W(\vec{x} - \vec{x'}; R) d\vec{x'} \tag{3}$$

in real space, and

$$\hat{\delta}(\vec{k}; R) = \hat{\delta}(\vec{k}) \hat{W}(\vec{k}; R). \tag{4}$$

in Fourier space where $\hat{W}$ is the Fourier transform of the window function. For each smoothing length, $R$, we can then associate a mass

$$M = \gamma_W \rho_b R^3, \tag{5}$$

where $\gamma_W$ is a filter-dependent number and $\rho_b$ is the mean density of the Universe.

This then leads us to the expressions for the mass variance

$$\sigma^2(R) = \sigma^2(M) = \sigma_M^2 = \frac{1}{2\pi^2} \int k^3 P(k) \hat{W}(k; R)^2 \frac{dk}{k}, \tag{6}$$

which for a power-law power-spectra, $P(k) \propto k^n$ gives us

$$\sigma^2(M) \propto k^{n+3} \propto M^{-(n+3)/3} \tag{7}$$

We also discussed the spherical collapse model which allows us to connect linear theory with the non-linear evolution of perturbations:

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

$$\delta_{\rm NL} = 1 \qquad\qquad \delta_L \approx 0.57, \tag{8}$$

where $\delta$ corresponds to the exact solution and $\delta_L$ to the overdensity we would find if linear theory was correct.

2. The density at the time of turn-around, when the perturbation stops expanding is given by

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

where $\rho_b$ 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. Finally the object is said to virialise, or collapse, at $t_{\rm coll} = 2t_{\rm TA}$ when the spherical model would predict infinite density. In that case the density of the region is given by

$$\rho_{\rm coll} = (1 + \delta_{\rm coll}) = 18\pi^2\rho_b \approx 178\rho_b \qquad\qquad \delta_L \approx 1.69 = \delta_c. \tag{10}$$

where we define a linear critical overdensity for collapse as $\delta_c = 1.69$ and this is a quantity we will use repeatedly later. Note also that the radius of the collapsed object is $r_{\rm coll} = r_{\rm TA}/2$ according to the virial theorem.

The exact results do depend on the exact cosmology chosen but the $\delta_c$ 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.

# 2 Redshift space versus real space

*MvdBW section 6.3; C&L section 18.5; Pdm3 section 7.7; Peacock section 16.5.*

We tend to implicitly assume that we know the distances to objects in our analysis. However except for the very most nearby systems, distances are usually impossible to determine precisely. Instead we rely on redshifts. At low redshift, the distance is then related to the redshift by $d = H_0^{-1}cz$. This makes the assumption that galaxies are test particles following the Hubble expansion, but in reality of course galaxies have peculiar velocities.

Peculiar velocities complicates the relationship between distance and redshift and this effect is scale-dependent. On large scales, we will see the effect of collapse (in the co-moving frame) which will lead to overdensities to be slightly compressed along the line-of-sight in redshift space as indicated in the top line on the left in Figure 1. Turn-around takes place on slightly smaller scale and at turn-around the infall velocity equals the local Hubble

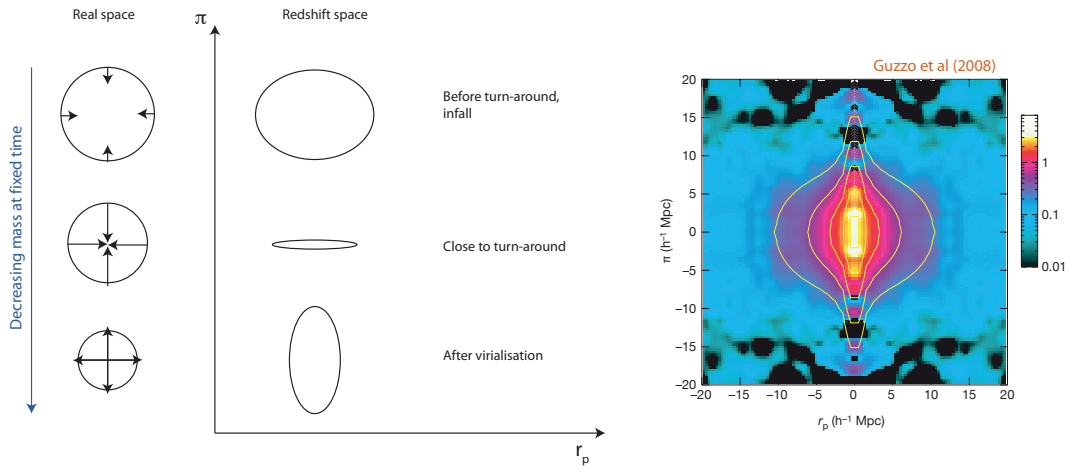


Figure 1: On the left an illustration of the effect of redshift space distortions. It is qualitatively trying to capture the situation at a fixed age. From the top downwards we proceed to smaller masses, from the slowly collapsing stages (in co-moving coordinates) before turn-around through turn-around when the velocities are equal to the expansion velocity and the structure is completely collapsed along one axis. Finally on the smallest scales virialisation has taken place at which point peculiar velocities exceed the Hubble velocities and you see significant elongation, "finger-of-god", along the line of sight. The right-band plot shows an observational determination of the correlation function (which shows the same overall behaviour as the density) from the VIMOS Very Deep Survey (Guzzo et al 2008; Nature). Here you can see clearly the small-scale elongation and the large-scale flattening implied by the illustration on the left.

expansion so a structure is very flattened in redshift space (middle line in the figure) and finally where virialization has happened — on smaller scales again, velocities are large relative to the local expansion velocity and structures are elongated along the line of sight in redshift space, the so-called "finger-of-god" effect.

## 2.1 Linear redshift distortions

Consider a small volume, $d^3\vec{r}$ with a number density $n(\vec{r})$. The number here must match what we get in the same volume in redshift space. To firm up our notation, let us write the distance to an object in redshift space as $\vec{s}$, while the proper distance in real space is $\vec{r}$ (breaking my habit of using $r$ to designate comoving coordinates — but this is the notation generally used in this area).

We can then relate $\vec{s}$ and $\vec{r}$ through

$$\vec{s} = \vec{r} + \frac{v_r}{H}\hat{\vec{r}}, \tag{11}$$

where $H$ is the Hubble parameter and this is used to convert the velocity along the line of sight, $v_r$ into distance units. Note that derivations in this field often use equation 11 but with distances in velocity units without pointing this out.

Using this we can first write down the conservation of numbers

$$n(\vec{r})d^3\vec{r} = n^{(s)}(\vec{s})d^3\vec{s}, \tag{12}$$

where I here and in the following use a subscript $s$ to indicate redshift space quantities. Inserting the definition of the overdensity, $\delta$, we get

$$\left[1 + \delta^{(s)}(\vec{s})\right] d^3\vec{s} = \left[1 + \delta(\vec{r})\right] d^3\vec{r}. \tag{13}$$

I have tacitly assumed that the mean densities in redshift space and real space are the same — this is in fact not obvious and the literature contains expressions incorporating variations in this ratio using selection functions.

To make progress we need to know the Jacobian $d^3\vec{s}/d^3\vec{r}$ and carrying out the derivatives and calculating the determinant of the variable change using equation 11 we find that

$$d^3\vec{s} = \left(1 + \frac{v_r}{Hr}\right)^2 \left(1 + \frac{1}{H}\frac{\partial v_r}{\partial r}\right) d^3\vec{r}. \tag{14}$$

Now for an observer that is far away, the $v_r/r$ term can be neglected — although note that in careful analysis when the data cover a wide region of sky this might not be appropriate. But what about the $dv_r/dr$ term? In order to estimate this, we need to take a step back and consider the the continuity equation

$$\frac{\partial\delta}{\partial t} + \frac{1}{a}\nabla_{\vec{x}}(1+\delta)\vec{v} = 0. \tag{15}$$

If we linearise this equation we get

$$\frac{\partial \delta}{\partial t} = \delta H \frac{d \ln \delta}{d \ln a} = -\frac{1}{a} \nabla_{\vec{x}} \vec{v}. \tag{16}$$

The reason for writing the equation in this form is that $d \ln \delta / d \ln a \approx \Omega_m^{0.6}$ to a good approximation, and is usually denoted $f(\Omega_m)$. Armed with this we finally have

$$\delta(\vec{x}) = -\frac{\nabla_{\vec{x}} \cdot \vec{v}}{aHf(\Omega_m)}. \tag{17}$$

This immediately tells us that in the linear regime small perturbations create velocities and in principle, at least, one might expect that velocities can be used to determine cosmological parameters. For our use here, however, the key observation here is that as long as $\delta \ll 1$ we will have that $dv_r/dr \ll 1$ so we can combine equation 13 with equation 14 to get

$$\delta^{(s)}(\vec{s}) \approx \delta(\vec{r}) - \frac{1}{H} \frac{\partial v_r}{\partial r}. \tag{18}$$

in the linear case.

If we now go to Fourier space, we can write equation 17

$$\vec{v}_k = f(\Omega_m) H \frac{ia\vec{k}}{k^2} \delta_{\vec{k}}, \tag{19}$$

and now projecting this along the line of sight ($\vec{r}$) we have

$$v_{r,k} = f(\Omega_m) H \mu \frac{ia}{k} \delta_k, \tag{20}$$

where $\mu$ is the cosine of the angle between $\vec{k}$ and the line of sight. In order to put equation **??** in Fourier space, we also need an expression for $\frac{\partial}{\partial r}$ and this is

$$\frac{\partial}{\partial r} = \frac{1}{a} \hat{\vec{r}} \frac{\partial}{\partial \vec{x}} \rightarrow \frac{ik\mu}{a}, \tag{21}$$

where the arrow indicates transformation to Fourier space (consider operating on $e^{i\vec{k}\cdot\vec{x}}$).

Taken together this puts equation 18 in the form

$$\delta^{(s)}_{\vec{k}} = \delta_{\vec{k}} + f(\Omega_m) \mu^2 \delta_{\vec{k}}. \tag{22}$$

From which we can conclude that the perturbations in redshift space are boosted by a factor $1 + f(\Omega_m)\mu^2$ relative to real space — a factor that will also come into the power spectrum in redshift space. Furthermore, we see that the power spectrum, and thus the correlation function, in redshift space will be anisotropic because of the dependence on $\mu^2 = \cos^2 \theta$.

In practice we observe galaxies, while the derivation above focuses on mass. If we assume that the density of galaxies is linearly related to the true mass density through a linear constant $b$, called the bias parameter, the pre-factor $f(\Omega_m)$ is changed to

$$\beta = \frac{f(\Omega_M)}{b} \approx \frac{\Omega_M^{0.6}}{b}, \tag{23}$$

and this has been exploited in a number of studies to constrain cosmology. A key study at higher redshift is that of Guzzo et al (2008, Nature, 451, 541). These studies are usually phrased in terms of the correlation function instead of the density directly and the resulting map from Guzzo et al is shown on the right in Figure 1. Guzzo et al find $\beta \approx 0.7$ and studies in the nearby Universe has found a range of $\beta = 0.5$ to $\beta = 0.8$, with the latter being appropriate for IRAS galaxies.

# 3 The topology of the large-scale structure

*Note I here do not refer to topology in the mathematical sense. This is also a very interesting topic for large-scale structure studies but fall outside the scope of these lectures. Section 5.6 in MvdBW contains some discussion on the topic of this section.*

A gravitational perturbation will typically be tri-axial and will therefore collapse along its shortest axis first. This leads to a flattened structure and this will subsequently collapse along *its* shortest axis leading to a filamentary structure and if collapse can continue, the final shape will be point like, in reality, because of virialisation this will be a halo.

This then leads to the basic topology of the large-scale structure of the Universe. We have

**1D collapse** This leads to so-called pancakes, sheet-like structures.

**2D collapse** The sheets themselves will collapse along their shortest axes and this will lead to filamentary structures.

**3D collapse** Finally we get collapse to a virialised halo. In simulations we find halos normally at the intersections of filaments.

# 4 The peak formalism & biased galaxy formation

*Section 7.1 in MvdBW; Peacock 16.3 also has a discussion and Bardeen et al (1986, ApJ, 304, 15) is the main original reference for much of this.*

We focus our attention in this course mostly on what we would call collapsed objects by identifying regions with overdensity above a certain threshold with collapsed objects. This identification has its problems but it provides some results in good agreement with simulations.

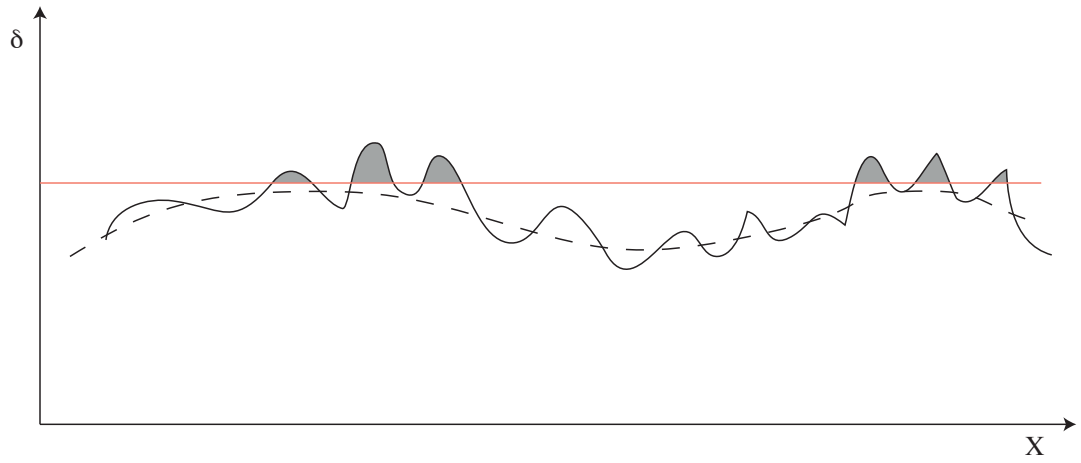


Figure 2: An cartoon illustration of why high peaks in the density field are expected to be highly clustered. The dashed line shows the large-scale density field and the solid line represents the density field smoothed on some small scale. If the red line represents the density threshold for collapse, we see that only very high peaks lie above it at this time and these cluster strongly because they lie themselves on peaks in the large-scale density field.

An alternative approach is to focus on peaks in the (smoothed) density field. It is not unreasonable to hypothesise that objects will form close to peaks in the density field.

The advantage of this is that the statistics of peaks in the density field can be worked out analytically for Gaussian fields (Bardeen et al 1986, ApJ, 304, 15). The details of this are not important for us here, but the cartoon plot of the density field in Figure 2 illustrates the main salient points of interest for us. The red line in that figure shows line above which we suppose that all peaks correspond to objects. Since the location of peaks is modulated by the large-scale density field (dashed line), the chances of having another peak close is high for the high peaks and they cluster strongly.

If we write $\nu_s = \delta_S/\sigma_S$, where $S$ means that the density field is smoothed on a scale $M$, it turns out (see the references given above) that for $\nu_S \gg 1$ we can write

$$\delta_{\rm peak}(\nu_S|\nu_b) \approx \frac{\nu_S}{\sigma_S}\delta_b = b_{\rm peak}\delta_b, \tag{24}$$

where $\nu_b$ and $\delta_b$ corresponds to the significance and density contrast of the background field respectively. We can then read the last equality as saying that the density field of peaks is biased relative to that of the mass.

Since the correlation function is related to the density contrast squared, it should be clear that a similar relation holds for the correlation function, and indeed

$$\xi_{\rm peak}(r) = b^2_{\rm peak}\xi(r), \tag{25}$$

holds in this limit as well.

Both from the illustration, and equation (24) we can make two important observations valid for Cold Dark Matter models:

1. At fixed mass, clustering was stronger at high-$z$ — because the threshold for collapse would be higher in the linearly extrapolated density field.

2. At fixed time, more massive haloes cluster more strongly. This is perhaps most easily seen from the equation because $\sigma_S$ is smaller for more more massive haloes.

# 5 Mass functions — the Press-Schechter formalism

*This is discussed in section 7.2 in MvdBW; section 7.4 in Pdm3; a brief discussion is also in 14.5 C&L. The original paper on the Press-Schechter mass function is Press & Schechter (1974, ApJ, 187, 425)*

What is a mass function? It is a function that tells you how many objects there is in a narrow range in mass. Normally it is normalised to give the number of unit volume or some similar normalisation. In galaxy formation the most widely used theoretical mass function is the very influential Press-Schechter (PS) mass function.

The PS mass function originates from linear theory so it is useful to recall that the time-dependence of the density field is:

$$\delta(\vec{x}, t) = \delta_0(\vec{x}) D(t), \tag{26}$$

where subscript 0 refers to the value of the field today, and $D(t)$ is the growth factor, normalised to unity today. For all scales of interest here (galaxy to cluster masses), the evolution of this is well approximated by $D(t) = 1/(1+z)$, at least at early times and unless otherwise noted this is what we will assume this. There are fitting functions that can be used for cases where $\Omega_m \neq 1$, see for instance Carroll et al (1992, ARA&A 30, 499) or equation 4.76 in MvdBW.

By referring to the spherical collapse model, we say that a perturbation has collapsed when

$$\delta(\vec{x}, t) > \delta_c \approx 1.69 \quad \Leftrightarrow \quad \delta_0(\vec{x}) > \frac{\delta_c}{D(t)} \stackrel{\text{def}}{=} \delta_c(t). \tag{27}$$

Of course in reality this statement only makes sense if we adopt a window function and calculate the smoothed field on a scale $R$, with associated mass, $M$, as given above.

To get to the PS mass function, we make use of the PS ansatz: The fraction of mass elements in halos with mass $> M$ is equal to the probability of having $\delta_S > \delta_c(t)$, $P(\delta_S > \delta_c(t))$, where $\delta_S$ is the density field smoothed on scale $M$.

While this ansatz is not supported by simulations it turns out to be a fruitful approximation to make although it has weaknesses that we will return to in a short while. To

make use of this ansatz we first calculate the probability:

$$\begin{aligned} P(\delta_S > \delta_c(t)) &= \frac{1}{\sqrt{2\pi}\sigma(M)} \int_{\delta_c(t)}^{\infty} e^{-\delta_S^2/2\sigma^2(M)} d\delta \\ &= \frac{1}{2} \operatorname{erfc}\left[\frac{\delta_c(t)}{\sqrt{2}\sigma(M)}\right], \end{aligned} \tag{28}$$

where erfc is the complementary error function ($\operatorname{erfc}(x) = 2/\sqrt{\pi}\int_x^\infty e^{-t^2}dt$).

As we go to smaller masses, we would expect this to give the total fraction of mass in the Universe that is locked up in bound objects. We recall from last lecture that $\sigma^2(M) \propto M^{-(n+3)/3}$, so for $n > -3$, which is the case for CDM, we have that as we go to small masses, $\sigma \to \infty$ and hence the argument of erfc goes to 0, and as $\operatorname{erfc}(0) = 1$, we find that the probability is 50%. In the PS ansatz this would then mean that only 50% of the mass of the Universe resides in bound structures. This only counts overdense systems, and ignores systems that are underdense but might be embedded in a much larger overdense region. This is unsatisfactory, so PS then proposed (with somewhat unsatisfactory justification) that one should multiply the probability by 2 to correct for this. We will see next lecture that it is possible to do a more rigorous derivation and hence get the factor of 2 naturally. For now, let just assume that:

$$F(> M) = 2P[\delta_S > \delta_c(t)]. \tag{29}$$

We can then use this to get the mass function. This is

$$n(M,t)dM = \frac{\bar{\rho}}{M}\frac{\partial F(> M)}{\partial M} dM, \tag{30}$$

where the first term, $\bar{\rho}/M$, is the number density of objects with mass $M$ and the second term gives the fraction of these to count in this mass interval. Writing this out in full we get

$$n(M,t)dM = \frac{\bar{\rho}}{M} 2 \frac{\partial}{\partial M}\left\{\frac{1}{2}\operatorname{erfc}\left[\frac{\delta_c(t)}{\sqrt{2}\sigma(M)}\right]\right\} dM \tag{31}$$

$$= \sqrt{\frac{2}{\pi}}\frac{\bar{\rho}}{M^2}\frac{\delta_c(t)}{\sigma(M)} e^{-\delta_c^2(t)/2\sigma^2(M)} \left|\frac{d\ln\sigma(M)}{d\ln M}\right| dM. \tag{32}$$

The last line here defines the PS mass function. We can write it in a slightly different form by introducing the significance of a peak:

$$\nu \stackrel{\text{def}}{=} \frac{\delta_c(t)}{\sigma(M)}, \tag{33}$$

which measures how many sigma a fluctuation is. Using this we can recast the PS formula in the form

$$n(M,t)dM = \frac{\bar{\rho}}{M^2} f_{\rm PS}(\nu) \left|\frac{d\ln\nu}{d\ln M}\right| dM \tag{34}$$

, where

$$f_{\rm PS}(\nu) = \sqrt{\frac{2}{\pi}}\nu e^{-\nu^2/2}. \tag{35}$$

Note that the mass function so calculated does depend on the window function chosen.

### 5.1 Time-dependence

The various quantities, $\sigma(M)$, that go into the PS mass function are evaluated at $z = 0$, thus the time-dependence is encapsulated in the time-evolution of $\delta_c(t)$. We have in general that

$$\delta(z=0) = \frac{\delta_c}{D(z)} \approx \frac{1.69}{D(z)} \tag{36}$$

and inserting this into the PS equation in the form of equation (32), then we get:

$$n(M,z)dM = \sqrt{\frac{2}{\pi}}\frac{\bar{\rho}}{M^2}\frac{\delta_c}{D(z)\sigma}e^{-(\delta_c/D(z))^2/2\sigma^2}\left|\frac{d\ln\sigma}{d\ln M}\right|dM, \tag{37}$$

where $\sigma \equiv \sigma(M)$.

It is worth noting that if a perturbation of a given mass was a $3\sigma$ fluctuation at $z = 2$ another perturbation with the same mass today would only be a $1\sigma$ fluctuation. Halos that were rare as collapsed objects at that redshift, are much more common today. This is the reason why you would expect to see very few massive galaxy clusters at high redshift for instance, because they would correspond to very rare perturbations.

## 6 Example use of the PS mass function

If we define a characteristic mass, $M$, such that it corresponds to a $\nu = 1$ fluctuation. From equation (33) we then have that

$$D(z)\sigma(M) = \delta_c, \tag{38}$$

which we can use to rewrite the PS mass function in the form

$$n(M,z)dM = \sqrt{\frac{2}{\pi}}\frac{\bar{\rho}}{M^2}\left|\frac{d\ln\sigma(M)}{d\ln M}\right|\frac{\sigma(M_*)}{\sigma(M)}e^{-\frac{1}{2}\left(\frac{\sigma(M_*)}{\sigma(M)}\right)^2}dM. \tag{39}$$

We can gain more insight into this equation by approximating $P(k)$ as a power-law, $P(k) \propto k^n$. In this case we have

$$\sigma(M) \propto M^{-(n+3)/6} \Rightarrow \frac{\sigma(M_*)}{\sigma(M)} = \left(\frac{M}{M_*}\right)^{(n+3)/6}, \tag{40}$$

giving $d\ln\sigma/d\ln M = -(n+3)/6$. In turn this gives the PS mass function as

$$n(M,z)dM = \frac{\bar{\rho}}{M^2}\frac{1}{\sqrt{2\pi}}(1+\frac{n}{3})\left(\frac{M}{M_*}\right)^{(n+3)/6}e^{-(M/M_*)^{(n+3)/3}/2}. \tag{41}$$

If we contrast this to observations, we know that the galaxy stellar mass function or the luminosity function can be written as Schechter functions:

$$\phi(X)dX = \phi_* \left(\frac{X}{X_*}\right)^{\alpha} e^{-X/X_*} \frac{dX}{X_*}. \tag{42}$$

So both this and the PS mass function has an exponential cut-off and a power-law faint end (low mass) behaviour. That looks promising but the devil is in the details. Observationally the faint end slope is $\sim -1.3$ although it does depend on what redshift/galaxy population you use.

For the PS mass function we can use that the CDM power spectrum has a slope $\sim -2$ on galaxy scales. Inserting this into equation (41) we get, that at $M \ll M_*$:

$$n(M, z) \propto M^{-11/6}, \tag{43}$$

and hence a faint end slope of about $-2$. This is considerably steeper than the mass/luminosity functions and is a clear indication that we need to suppress galaxy formation in small halos.

Likewise at the massive end, while there is an exponential cut-off, it is not $e^{-M}$ but rather $e^{-M^{1/3}}$, which declines much less quickly. Again this leads to an excess of large halos relative to the galaxy population. I illustrate this difference in Figure 3.

Understanding the difference between these two functions is a key challenge of galaxy formation theory.

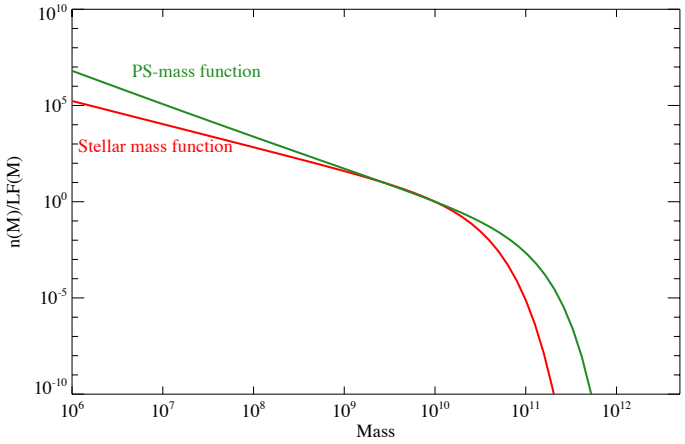


Figure 3: A comparison of the Press-Schechter mass function at $z = 0$ for a CDM cosmology with WMPA7 cosmological parameters with a generic galaxy mass function with a faint-end slope of $-1.2$. The galaxy mass function has a characteristic mass of $10^{10} M_{\odot}$ and arbitrary normalisation. For display purposes, the halo mass function has been shifted by a factor $\sim 3000$ in mass to match the galaxy mass function. Note the clear excess of low-mass and high-mass halos.