Lecture 6: Dark Matter Halos

Houjun Mo

February 23, 2004
Hierarchical clustering, the matter in the universe condenses to form quasi-equilibrium objects (dark halos) of increasing masses in the passage of time.

The Goal of this lecture is to understand this process and the properties of dark matter halos.
Issues to be addressed

1. The mass function (i.e. the number density of haloes as a function of halo mass) at a given time. The study of this issue may help us to understand the origin of the mass functions of galaxies and of clusters of galaxies.

2. The mass distribution of the progenitors of a given halo. The study of this issue may help us to understand the formation processes of present-day galaxies as well as galaxy populations at high redshifts.

3. The merging rate of haloes. This issue may be related to the understanding of the population of interacting galaxies and the substrutures in clusters of galaxies.

4. Intrinsic properties of haloes, such as density profile, shape and angular momentum. This is an important issue, because it may be related to the nature and morphology of protogalactic collapses.

5. The large-scale environment and spatial clustering of dark haloes. This may help us to understand the relation between galaxy population and the large scale environment.
Halo Mass Function: the Press-Schechter Formalism

Basic idea:

- The cosmic density field is Gaussian
- The collapse of the high density regions can be described by spherical model

Consider a smoothed version of the cosmic density field $\delta(x)$:

$$\delta(x; R) \equiv \int \delta(x') W(x + x'; R) \, d^3x',$$

where $W(x; R)$ is a window function of characteristic radius $R$. 
For a Gaussian density field, $\delta(x,R)$ has also Gaussian distribution:

$$P[\delta(R,t)]d\delta(R,t) = \frac{1}{\sqrt{2\pi\sigma(R,t)}} \exp\left[-\frac{\delta^2(R,t)}{2\sigma^2(R,t)}\right] d\delta(R,t).$$

$$\sigma^2(R,t) = D^2(t)\sigma^2(R); \quad \sigma^2(R) = \equiv \langle \delta^2(x;R) \rangle = \int \Delta^2(k)\tilde{W}^2(kR)\frac{dk}{k},$$

where

$$\Delta^2(k) \equiv \frac{1}{2\pi^2}k^3P(k),$$

and $\tilde{W}(kR)$ is the Fourier transform of the window function.
The forms of the two most commonly used window functions are as follows:

- **Top-hat window:**

  
  \[
  W_{\text{th}}(x; R) = \left( \frac{4\pi}{3} R^3 \right)^{-1} \begin{cases} 
  1 & (\text{for } |x| \leq R) \\
  0 & (\text{for } |x| > R)
  \end{cases}
  \]

  \[
  \tilde{W}_{\text{th}}(kR) = \frac{3 (\sin kR - kR \cos kR)}{(kR)^3};
  \]

- **Gaussian window:**

  \[
  W_{\text{G}}(x; R) = \frac{1}{(2\pi)^{3/2} R^3} \exp \left( -\frac{|x|^2}{2R^2} \right), \quad \tilde{W}_{\text{G}}(kR) = \exp \left[ -\frac{(kR)^2}{2} \right];
  \]

The volumes of windows are related to the smoothing radius \( R \) as

\[
V(R) = \begin{cases} 
4\pi R^3/3 & (\text{for top-hat}) \\
(2\pi)^{3/2} R^3 & (\text{for Gaussian})
\end{cases}
\]
The fraction of points surrounded by a sphere of radius $R$ and mean overdensity exceeds $\delta_c$ is:

$$F_{1/2}(R,t) = \int_{\delta_c}^\infty \frac{d\delta}{\sqrt{2\pi D(t)\sigma(R)}} \exp \left[-\frac{\delta^2}{2D^2(t)\sigma^2(R)}\right] = \frac{1}{2} \text{erfc} \left[\frac{\delta_c/D(t)}{\sqrt{2}\sigma(R)}\right].$$

Press and Schechter suggested that, with $\delta_c$ chosen to be the threshold overdensity for spherical collapse ($\delta_c \approx 1.686$), this fraction be identified with the fraction of particles which are part of collapsed lumps with masses exceeding $M = 4\pi\bar{\rho}a^3R^3/3$.

A problem: as $M \to 0$, $\sigma(R) \to \infty$ (for power-law spectrum with $n > -3$) and $F_{1/2} \to 1/2$. Only half of the particles are parts of lumps of any mass. Press and Schechter ‘solve’ this, arbitrarily, by multiplying the mass fraction by a factor of 2.
The number density of collapsed lumps with mass in the range $M \rightarrow M + dM$ is then

$$n(M,t) \, dM = -2 \rho \frac{dF_{1/2}}{dR} \frac{dR}{dM} \, dM = -\sqrt{2 \rho} \frac{\delta_c(t)}{M} \frac{d\sigma}{dM} \exp \left[ -\frac{\delta_c^2(t)}{2\sigma^2} \right] \, dM,$$

where $\delta_c(t) \equiv \delta_c/D(t)$ is the critical overdensity linearily extrapolated to the present time.

Note: time enters only through $D(t)$, mass enters through $\sigma(R)$. The mass fraction of the universe in objects with $\sigma^2(R)$ in the range from $\sigma^2 \rightarrow \sigma^2 + d\sigma^2$ is

$$f[\sigma,\delta_c(t)] \, d\sigma^2 = -\frac{1}{\sqrt{2\pi}} \frac{\sigma^3}{\sigma^3} \exp \left[ -\frac{\delta_c^2(t)}{2\sigma^2} \right] \, d\sigma^2.$$
The Press-Schechter formalism provides a useful way to understand how nonlinear structures develop in a hierarchical model. As one can see, haloes with mass $M$ can form in significant number only when $\sigma(M) \gtrsim \delta_c(t)$. If we define a characteristic mass $M_\star(t)$ by

$$\sigma(M_\star) = \frac{\delta_c}{D(t)},$$

then only haloes with $M \ll M_\star$ can form in significant number at time $t$. Since $D(t)$ increases with $t$ and $\sigma(M)$ decreases with $M$ in hierarchical models, we see that haloes with larger and larger masses can form in the passage of time.
Excursion Set Derivation of the Press-Schechter Formula

How to explain the fudge factor of 2 in the Press-Schechter formalism? The derivation by Bond et al. (1991).

Instead with the spherical top-hat filter, smooth $\delta(x)$ a sharp-$k$ filter:

$$\tilde{W}_{Sk}(k; k_c) = \begin{cases} 
1 & \text{(for } |k| \leq k_c) \\
0 & \text{(for } |k| > k_c) 
\end{cases} .$$

The smoothed field is

$$\delta_s(x, t; k_c) = \int d^3k \delta_k(t) \tilde{W}_{Sk}(k; k_c) e^{-i k \cdot x} = \int_{k < k_c} d^3k \delta_k(t) e^{-i k \cdot x} .$$
Consider an increase in $k_c$ by $\Delta k$:

$$
\delta_s(x, t; k_c + \Delta k) = \delta_s(x, t; k_c) + 4\pi\delta_k(t) e^{-ik \cdot x} k^2 \Delta k.
$$

As $k_c$ is increased the value of $\delta_s$ at a given point $x$ executes a random walk.

The advantage of using sharp-$k$ filter:

$$
\langle \Delta \delta_s \delta_s(x; k_c) \rangle = 0
$$

where $\Delta \delta_s = \delta_s(x, t; k_c + \Delta k) - \delta_s(x, t; k_c)$
and so $\Delta \delta_s$ is independent of $\delta_s(x, t; k_c)$. The variance

$$
\langle (\Delta \delta_s)^2 \rangle = \sigma^2(k_c + \Delta k, t) - \sigma^2(k_c, t) = \Delta \sigma^2.
$$

So as we change the window size, the mean density with it execute a random walk governed by a simple Gaussian distribution independent of early steps.
At given time $t$ we can identify all the points, $\{x\}$, where $\delta_s(x, t; K_c) = \delta_c$; and $\delta_s(x, t; k_c) < \delta_c$ for all $k_c < K_c(x)$. $K_c$ is the value of $k_c$ where $\delta_s(x, t; k_c)$ first crosses the barrier at $\delta_s = \delta_c$.

Different points on the random walk can be divided into three categories:

1. Points with $\delta_s > \delta_c$ for $k_c = K_c$;

2. Points with $\delta_s < \delta_c$ for $k_c = K_c$ but $\delta_s > \delta_c$ for some $k_c < K_c$;

3. Points with $\delta_s < \delta_c$ for all $k_c \leq K_c$.

We want the fraction of mass elements in class (iii) since this is the fraction of mass elements with first upcrossing at $k_c > K_c$. 
Note that for every random walk leading to an element with $\delta_s = \delta_0 > \delta_c$ in class (i) there is an *equally probable* walk leading to an element in class (ii) with $\delta_s = \delta_c - (\delta_0 - \delta_c) = 2\delta_c - \delta_0$. Hence the distribution of $\delta_s$ for points with first upcrossing at $k_c > K_c$ is

$$P_{FU}(\delta_s) \, d\delta_s = \frac{1}{\sqrt{2\pi} D(t) \sigma(M_c)} \left\{ \exp \left[ -\frac{\delta_s^2}{2D^2(t)\sigma^2(M_c)} \right] - \exp \left[ -\frac{(2\delta_c - \delta_s)^2}{2D^2(t)\sigma^2(M_c)} \right] \right\}$$

and the fraction of mass elements with first upcrossing at $k_c > K_c$ is

$$F(> K_c) = \int_{-\infty}^{\nu_c} P_{FU} \, d\delta_s = \int_{-\infty}^{\nu_c} \frac{dx}{\sqrt{2\pi}} e^{-x^2/2} - \int_{\nu_c}^{\infty} \frac{dx}{\sqrt{2\pi}} e^{-x^2/2},$$

where $\nu_c \equiv \delta_c(t)/\sigma(M_c)$. 
Thus,

\[ f[\sigma(M_c), \delta_c(t)] \, dS = \frac{\partial F}{\partial S} \, dS = \frac{1}{\sqrt{2\pi} [S(M_c)]^{3/2}} \exp \left[ -\frac{\delta_c^2(t)}{2S(M_c)} \right] \, dS. \]

The form of \( f \) is exactly the same as given by the Press-Schechter formalism, except that the variance \( S \) are calculated in two different types of windows. The origin of the famous factor of 2 in the Press-Schechter formalism is now clear. The original treatment included only the first of the two terms in \( P_{FU} \); an equal contribution should come from the second term.
Interpretation of the Excursion Set

According to spherical model, a region will collapse at a time when its overdensity reaches the critical overdensity $\delta_c$. One might therefore think that all the mass elements which have first upcrossing at mass scale in the range corresponding to $(S, S + dS)$ are parts of collapsed haloes with this mass at time $t$, and so $f$ in the above equation could be interpreted as the fraction of mass elements which are part of such collapsed objects.

The first part of the inference is not correct while the second part is.

The simplest way to understand this is to consider a uniform spherical region of mass $M$ and radius $R$, with critical overdensity at some time $t$, which is embedded in a larger region with lower overdensity. According to spherical model, all mass elements in the spherical region will be part of a collapsed halo of mass $M$ at time $t$. Now consider an interior point $x$ which is at a distance $r$ from the surface of the spherical region, and calculate the mean overdensities within spheres centred on this point. Clearly, the overdensity reaches the critical density only for spheres
with radii smaller than \( r \). Thus, the mass element at \( x \) has its first upcrossing at a mass near \( M(r/R)^3 \) which is smaller than \( M \) if \( x \) is not at the centre of \( M \). From this simple example, we see that the mass at the first upcrossing of a mass element only gives the lower limit of the collapsed object which the mass element is part of. It then follows that the integration of \( f \) over mass,

\[
F(M,t) = \int_{M}^{\infty} f[S(M'), \delta_c(t)] \frac{dS(M')}{dM'} \, dM',
\]

is the fraction of mass elements which are part of collapsed objects of mass exceeding \( M \), and \( f \) itself is just the fraction of mass elements which are part of collapsed objects of the mass in the range corresponding to \((S, S + dS)\). Thus, the excursion-set approach predicts how much mass ends up in collapsed objects of certain mass at a given time in a statistical sense, but it does not predict directly the host object of a particular mass element.
Halo bias: halo number density is modulated in space
\[ N[M_1|\delta_0(R)]dM_1 = \frac{M_0}{M_1} f(1|0) \frac{dS_1}{dM_1} dM_1, \quad f(1|0) = \frac{\delta_1 - \delta_0}{\sqrt{2\pi(S_1 - S_0)^{3/2}}} \exp \left[ -\frac{(\delta_1 - \delta_0)^2}{2(S_1 - S_0)} \right] \]

where \( S_1 = \sigma^2(M_1), S_0 = \sigma^2(M_0). \)

\[ \delta_h(1|0) = b_h(M_1, \delta_1) \delta_m, \quad \delta_h = \Delta n_h/\bar{n}_h; \quad \delta_m = \Delta \rho_m/\rho_m \]

Halo bias (Mo & White 1996):

\[ b_h = 1 + \frac{D(t_1)}{D(t_0)} \left( \frac{v_1^2 - 1}{\delta_c} \right), \quad v_1 = \delta_c(t_1)/\sigma(M_1). \]

Linear bias of the halo-mass, halo-halo correlations:

\[ \xi_{hm}(r; M) = b_h \xi_m(r) \quad \xi_{hh}(r; M) = b_h^2 \xi_m(r). \]
Tests by numerical simulations
An improvement of the PS formalism

- **Basic idea (Sheth, Mo, Tormen 2001):**
  - Collapse of dark halos depends not only on overdensity, but also on the shape of the shear field: so ellipsoidal collapse instead of spherical collapse
  - In a Gaussian density field the ellipticity of the shear field depends on the mass of the region in consideration
  - The threshold over-density for collapse depends on the mass scale in consideration

\[
\frac{\delta_{ec}(M)}{\delta_{sc}} = 1 + \beta \left[ \frac{\sigma^2(M)}{\delta_{sc}^2} \right]^\gamma
\]

where \( \beta \approx 0.5, \gamma = 0.6 \).
• Mass function

\[ n(M, z) = A \left(1 + \frac{1}{v'^2q}\right) \sqrt{\frac{2}{\pi M dM}} \exp \left(-\frac{v'^2}{2}\right) \]

where \( v' = \sqrt{av} \), \( v = \delta_c/\sigma \), \( a = 0.707 \), \( q = 0.3 \), \( A = 0.322 \).

• Bias factor

\[ b = 1 + \frac{1}{\delta_c} \left[v'^2 + b v'^2(1-c) - \frac{v'^2c/\sqrt{a}}{v'^2c + b(1-c)(1-c/2)}\right], \quad \text{(for halos)} \]

\[ b_0 = 1 + \frac{D(z)}{\delta_c} \left[v'^2 + b v'^2(1-c) - \frac{v'^2c/\sqrt{a}}{v'^2c + b(1-c)(1-c/2)}\right], \quad \text{(for descendants)} \]

where \( b = 0.5 \) and \( c = 0.6 \), \( D(z) \) is the linear growth factor.
Mass functions and bias factors:
Some applications

The correlation of clusters at $z = 0$ (Mo, Jing & White 96)

The correlation of Lyman-break galaxies at $z \sim 3$ (Mo, Mao & White 99)
Some interesting predictions (Mo & White 2002)
Halo Progenitor Distribution

An advantage of the excursion set approach is that it provides a neat way to calculate the properties of the progenitors which give rise to any given class of collapsed objects. For example one can calculate the mass function at $z = 5$ of those clumps (progenitors) which are today part of rich clusters of mass $10^{15} M_\odot$. This is important, because it may give a model to describe how a collapsed object was build up by the accretion and merger of small objects.

It is easy to show the fraction, among all of the mass elements of all $(M_2, \delta_2)$ patches, which was in collapsed objects of mass $M_1$ at the earlier time $t_1$:

$$f(S_1, D_1|S_2, D_2) \, dS_1 = \frac{1}{\sqrt{2\pi} (S_1 - S_2)^{3/2}} \exp \left[ -\frac{(\delta_1 - \delta_2)^2}{2(S_1 - S_2)} \right] \, dS_1,$$

where $S_1 \equiv \sigma^2(M_1)$, $S_2 \equiv \sigma^2(M_2)$, $\delta_1 \equiv \delta_c/D(t_1)$, $\delta_2 \equiv \delta_c/D(t_2)$. 

\[
\text{\textcopyright 2023}
\]
The mass distribution of the progenitors of objects of mass $M_2$ is therefore

$$n(M_1,t_1|M_2,t_2)\,dM_1 = \frac{M_2}{M_1} f(S_1,D_1|S_2,D_2) \frac{dS_1}{dM_1} \, dM_1.$$  

Note that $D$, $t$ and $\delta = \delta_c/D(t)$ are equivalent variables, so are $S$ and $M$. From this we can calculate, for example, the fraction of the material in present-day rich clusters, which at $z = 3$ was in haloes with mass exceeding $10^{12} M_\odot$. Since a mass of this order must be assembled to make a bright galaxies, this clearly limits when the bright galaxies currently observed in clusters have formed.
Other derived quantities

Straightforward manipulations using the calculus of probabilities allow us to construct models for the distributions of many interesting properties of dark haloes.

Halo Formation Time

A convenient operational definition for the formation time of a halo of mass $M$ is the time when the largest of its progenitors first reaches mass $M/2$.

$$P(t_f < t_1; M_2, t_2) = \int_{M_2/2}^{M_2} \frac{M_2}{M_1} f(S_1, D_1|S_2, D_2) \frac{dS_1}{dM_1} dM_1.$$

The differential probability distribution for $t_f$ is then given by

$$\frac{dP}{dt_f}(t_f; M_2, t_2) = \int_{M_2/2}^{M_2} \frac{M_2}{M_1} \left| \frac{\partial}{\partial t_f} [f(S_1, D_f|S_2, D_2)] \right| \frac{dS_1}{dM_1} dM_1.$$
Merging Trees

The growth of a halo as the result of a series of mergers. Time increases from top to bottom, and the width of the tree branches represent the masses of the individual parent haloes. A slice throught the tree horizontally gives the distribution of masses in the parent haloes at a given time. The present time $t_0$ and the formation time $t_f$ are marked by horizontal lines, where the formation time is defined as the time at which a parent halo containing in excess of half of the mass of the final halo was first created. [after Lacey & Cole (1993)]
How to construct a merger tree?

To construct a merger tree of a halo, start from the trunk and work upwards.

Consider a halo with mass $M_2$ at a time $t_2$.

At a slightly earlier time $t_1 = t_2 - \Delta t$, the progenitor distribution is given by the conditional mass function. If $\Delta t$ is chosen small enough, the halo will either remain to be one halo or split into two haloes. The probability of there being a merger at $t_1$, with the smaller progenitor having mass $M_1$ (the larger one has $M'_1 \equiv M_2 - M_1$) can be taken to be

$$dP = n(M_1, t_1 | M_2, t_2) \, dM_1,$$

where $n(M_1, t_1 | M_2, t_2)$ is the conditional mass function.
Internal Structures of Dark Matter Haloes

Halo Density Profiles

From high-resolution simulations:

$$\rho(r) = \rho_{\text{crit}} \frac{\delta_{\text{ch}}}{(r/r_s)(1 + r/r_s)^2},$$

where $r_s$ is a scale radius.

[NFW profile (Navarro et al. 1997)]. Concentration $c \equiv r_{\text{vir}}/r_s$.

Jing et al.
Asymmetry and Substructure
Angular Momentum

The angular momentum of a dark halo is represented by its spin parameter:

$$\lambda = \frac{|E|^{1/2}J}{GM^{5/2}},$$

where $J$, $E$ and $M$ are the total angular momentum, energy and mass. For an isolated system, these quantities are all conserved during dissipationless gravitational evolution, and so is $\lambda$ itself.

$$\lambda = \text{rotational energy/total energy}$$
Numerical simulations show that \( \lambda \) has approximately a lognormal distribution

\[
p(\lambda) \, d\lambda = \frac{1}{\sqrt{2\pi \sigma_\lambda}} \exp \left[ -\frac{\ln^2(\lambda/\overline{\lambda})}{2\sigma^2_\lambda} \right] \frac{d\lambda}{\lambda},
\]

with \( \overline{\lambda} \approx 0.05 \) and \( \sigma_\lambda = 0.5 \).

Detail simulations show that the \( \overline{\lambda} \) declines weakly with increasing halo mass and that the distribution is similar for haloes in different large-scale environments and in different cosmogonic models.

Other issues:

(1) The distribution of angular momentum in individual halos

(2) Physical reasons: torque by large-scale structure
Observational Tests of Massive halos

Individual galaxies:

Groups and clusters:
Halo Model of Mass Clustering in the Universe

Ingredient:

• halo mass function

• halo correlation functions

• halo density profiles

We can obtain the correlation functions of mass (original idea due to Neyman and Scott (1952))
An example: the power spectrum (Ma & Fry 2000)

\[ P(k) = P_{1h}(k) + P_{2h}(k) \]

\[ P_{1h}(k) = \frac{1}{\bar{\rho}^2} \sum_M \tilde{\rho}^2(M, k), \]

\[ P_{2h}(k) = \frac{1}{\bar{\rho}^2} \sum_{M_1, M_2} \tilde{\rho}(M_1, k)\tilde{\rho}(M_2, k)P_{hh}(k; M_1, M_2). \]

Note that \( \tilde{\rho} \) can either be the mass density profile of dark halos, or the galaxy number density profile in halos.
For simplicity, we write
\[ \rho(x) = mu(x|m), \]  
(1)
x is the position relative to the halo center, and \( \int u(x|m)d^3x = 1. \)

Dividing the space into many small cells, \( \Delta V_i \) \((i = 1, 2, \cdots)\), so small that none of them contains more than one halo center. The occupation number, \( N_i \), is either 1 or 0, so \( N_i = N_i^2 = N_i^3 = \cdots \)

\[ \rho(x) = \sum_i N_i m_i u(x - x_i|m_i), \]  
(2)

\( m_i \) is the mass of the halo in \( \Delta V_i \). Note that

\[ \langle N_i m_i u(x - x_i|m_i) \rangle = \int dm n(m) m \Delta V_i u(x - x_i|m) \]

where \( n(m) \) is the halo mass function.
It is more convenient to work in Fourier space:

\[
\rho(k) = \sum_i N_i m_i \tilde{u}(k|m_i) e^{i k \cdot x_i},
\]

(3)

\(\tilde{u}(k|m_i)\) is the Fourier transform of the density profile.

Using the properties of \(N_i\),

\[
P(k) \equiv V_u \langle |\delta_k|^2 \rangle = P^{1h}(k) + P^{2h}(k),
\]

(4)

\[
P^{1h}(k) = \frac{1}{\bar{\rho}^2} \int \! dm m^2 n(m) |u(k|m)|^2,
\]

(5)

\[
P^{2h}(k) = \frac{1}{\bar{\rho}^2} \int \! dm_1 m_1 n(m_1) u(k|m_1) \int \! dm_2 m_2 n(m_2) u(k|m_2) P_{hh}(k|m_1,m_2),
\]

(6)

with \(P_{hh}(k|m_1,m_2)\) the power spectrum of halos of mass \(m_1\) and \(m_2\).
Models can be constructed for all low-order statistics of clustering:

- Two-point correlation function and power spectrum
- Higher-order correlation functions
- Redshift-space correlation functions
- Redshift-distortions and velocity statistics