Dark Matter Halos

Dark matter halos form through gravitational instability. 

**Density perturbations grow linearly until they reach a critical density, after which they turn around from the expansion of the Universe and collapse to form virialized dark matter halos.**

These halos continue to grow in mass (and size), either by accreting material from their neighborhood or by merging with other halos.

In this module we will talk about:

- The mass function of dark matter halos (i.e. the number density of halos as a function of halo mass), its dependence on cosmology, and its evolution with redshift.
- The mass distribution of the progenitors of individual halos.
- The merger rate of dark matter halos.
- Intrinsic properties of dark matter halos, such as density profile, shape and angular momentum.

They will have a large impact on galaxies
Dark Matter Halos

In the peak formalism, it is \textit{assumed} that the material which will collapse to form non-linear objects (i.e. dark matter halos) of given mass can be \textit{identified} in the \textit{initial} density field by first smoothing it with a filter of the appropriate scale and then locating all peaks above some threshold.

The properties of peaks in a Gaussian random field can be analysed in a mathematically rigorous way, but the formalism nevertheless has significant limitations:

- it cannot be used to obtain the mass function of nonlinear objects (dark halos)
- it does not provide a model for how dark matter halos grow with time

Alternative formalism: Press-Schechter 1974

- mass function, the mass distribution of progenitors, merger rate and clustering properties
Density Peaks

In the linear regime the overdensity field grows as \( \delta(x,t) \propto D(t) \), with \( D(t) \) the linear growth rate:

for a given cosmology [which governs \( D(t) \)] we can determine the properties of the density field at any given time from those at some early time \( t_i \).

**Can we extend this to the non-linear regime** in order, for example, to predict the distribution and masses of collapsed objects from \( \delta(x,t_i) \) without following the non-linear dynamics in detail?

**An object of mass \( M \) forms from an overdense region with volume \( V = M/\rho \) in the initial density field.**

It is tempting to assume that such a region will correspond to a peak in the density field after smoothing with window of characteristic scale \( R \propto M^{1/3} \).

This suggests that a study of the density peaks of the smoothed density field can shed some light on the number density and spatial distribution of collapsed dark matter halos.
Shortcomings of the Peak Formalism

It is tempting to interpret the number density of peaks in terms of a number density of collapsed objects of mass \( M \propto \rho R^3 \). However, there is a serious problem with this identification, because a mass element which is associated with a peak of \( \delta_1(x) = \delta_s(x; R_1) \) can also be associated with a peak of \( \delta_2(x) = \delta_s(x; R_2) \), where \( R_2 > R_1 \).

Should such a mass element be considered part of an object of mass \( M_1 \) or of mass \( M_2 \)?

If \( \delta_2 < \delta_1 \) the mass element can be considered part of both. The overdensity will first reach (at time \( t_1 \)) the critical value for collapse on scale \( R_1 \), and then (at time \( t_2 > t_1 \)) on scale \( R_2 \). \( M_1 \) is the mass of a collapsed object at \( t_1 \) which merges to form a bigger object of mass \( M_2 \) at \( t_2 \).

In the opposite case, where \( \delta_2 > \delta_1 \), the mass element apparently can never be part of a collapsed object of mass \( M_1 \) but rather must be incorporated directly into a larger system of mass \( M_2 \). Such peaks of the field \( \delta_1 \) should therefore be excluded when calculating the number density of \( M_1 \) objects.

This difficulty is known as the “cloud-in-cloud” problem.

What is required to predict the mass function of collapsed objects is a method to partition the linear density field \( \delta(x) \) at some early time into a set of disjoint regions (patches) each of which will form a single collapsed object at some later time.
Press-Schechter Formalism

Consider the overdensity field \( \delta(x,t) \), which in the linear regime evolves as \( \delta(x,t) = \delta_0(x)D(t) \), where \( \delta_0(x) \) is the overdensity field \textit{linearly extrapolated} to the present time and \( D(t) \) is the linear growth rate normalized to unity at the present.

According to the spherical collapse model described, regions with \( \delta(x,t) > \delta_c \approx 1.69 \), or equivalently, with:

\[
\delta_0(x) > \frac{\delta_c}{D(t)} \equiv \delta_c(t),
\]

will have collapsed to form virialized objects.

In order to assign masses to these collapsed regions, Press & Schechter (1974) considered the smoothed density field

\[
\delta_s(x; R) \equiv \int \delta_0(x') W(x + x'; R) \, d^3x'
\]

where \( W(x; R) \) is a window function of characteristic radius \( R \) corresponding to a mass \( M = \gamma f \rho R^3 \), with \( \gamma f \) a filter-dependent shape parameter (i.e. \( \gamma f = 4\pi/3 \) for a spherical top-hat filter).

The \textit{ansatz} of the Press-Schechter (PS) formalism is that the probability that \( \delta_s > \delta_c(t) \) is \textit{the same as the fraction of mass elements that at time} \( t \) \textit{are contained in halos with mass greater than} \( M \).

The probability that an overdense region of radius \( R \) has collapsed gives the fraction of mass in objects above that mass.
Tests of PS

Many crude assumptions in Press-Schechter (PS) and extended PS (EPS) - important to test HMF with simulations, which follow the growth and collapse of structures directly by solving the equations of motion for dark matter particles.

The PS mass function over-predicts the abundance of sub-$M^*$ halos, and under-predicts that of halos with $M > M^*$. The PS mass function for ellipsoidal collapse, on the other hand, fits simulations much better.

![Graph showing PS mass functions vs PS (spherical) and PS (ellipsoidal) mass functions](image)

However, comparisons depend on halos identification in the simulations. For example, (FOF) algorithm defines halos as structures whose particles are separated by distances less than the linking length, $b$, times the mean interparticle distance. A heuristic argument, based on spherical collapse, suggests that one should use $b \approx 0.2$, for which the mean overdensity of a halo is $\sim 180$. Different values for $b$ result in different mass functions.

More stringent test to e-PS is to predict the formation sites and masses of individual halos, and compare with simulations. The excursion-set formalism is not expected to work on an object-by-object basis. However, the mass at first upcrossing of a mass element should give a lower limit to the mass of the collapsed object to which the mass element belongs. For an ellipsoidal collapse barrier, this prediction was found to be consistent with $N$-body simulations.
Merger Trees

The conditional mass function derived above can be used to describe the statistical properties of the merger histories of dark matter halos.

The merger history of a given dark matter halo can be represented pictorially by a merger tree such as that shown in Fig. 7.4. In this figure, time increases from top to bottom, and the widths of the branches of the tree represent the masses of the individual progenitors.

If we start from an early time, the mass which ends up in a halo at the present time $t_0$ (the trunk of the tree) is distributed over many small branches; pairs of small branches merge into bigger ones as time goes on.

Fig. 7.4. Illustration of a merger tree, depicting the growth of a dark matter 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 halos. A horizontal slice through the tree, such as that at $t = t_f$ gives the distribution of the masses of the progenitor halos at a given time. [Adapted from Lacey & Cole (1993)]
Main Progenitor

The full merger history of any individual dark matter halo is a complex structure containing a lot of information. A useful subset of this information is provided by the main progenitor history, sometimes called the mass accretion history or mass assembly history, which restricts attention to the main “trunk” of the merger tree. This main trunk is defined by following the branching of a merger tree back in time, and selecting at each branching point the most massive (main) progenitor.

The left-hand panel of Fig. 7.5 shows 25 examples of main-progenitor histories as function of redshift, $M_{\text{main}}(z)$, normalized by the present-day halo mass, $M_0$.

All histories correspond to a halo of $M_0 = 5 \times 10^{11} h^{-1} M_\odot$ in an EdS cosmology. Note that these main progenitor histories reveal a large amount of scatter.

The right-hand panel of Fig. 7.5 shows the average main progenitor histories, $\langle M_{\text{main}}(z)/M_0 \rangle$, obtained by averaging 1000 random realizations for a given $M_0$. Results are shown for five values of $M_0$, as indicated. As is evident, on average halos that are more massive assemble later. As we demonstrate below, this is a characteristic of hierarchical structure formation.
Internal Structure of Dark Matter Halos

To a first approximation, we can model a dark matter halo as a spherical object. In this case, the internal mass distribution is fully described by a density profile, $\rho(r)$.

As discussed above, different halos have different formation histories, so we may expect a significant halo-to-halo variation in density profile.

But dark matter halos are highly non-linear, and it may be that information regarding their formation histories has largely been erased by their non-linear collapse. In the latter case, density profiles may be more closely related to the violent relaxation process than to initial conditions.

In the similarity model of spherical collapse, if we start with an initial perturbation $\delta_i(r) \propto r^{-3\varepsilon}$, then the final profile will be $\rho(r) \propto r^{-\gamma}$, where $\gamma = 1$ for $\varepsilon \leq 2/3$, and $\gamma = 3/(1+3\varepsilon)$ for $\varepsilon > 2/3$.

The typical density profile around each particle in a linear density field with scale-free power spectrum, $P(k) \propto k^n$, is $\delta_i(r) \sim \xi(r) \propto r^{-(n+3)}$. The typical halo profile will then be:

$$\rho(r) \propto \begin{cases} r^{-2} & \text{(for } n \leq -1) \\ r^{-(3n+9)/(n+4)} & \text{(for } n > -1) \end{cases} \quad \text{(7.134)}$$

Over the range of interest, $-3 < n < 0$, this model thus predicts that virialized halos resemble isothermal spheres.
The simplest plausible model is therefore to assume that dark matter halos are truncated singular isothermal spheres:

\[ \rho(r) \propto r^{-2} \quad \text{for} \quad r \leq r_h. \]  

(7.135)

We define the limiting radius of a dark matter halo, \( r_h \), to be the radius within which the mean matter density is

\[ \rho_h = \Delta_h \bar{\rho} = \Delta_h \rho_{\text{crit}} \Omega_m, \]  

(7.136)

where \( \rho \) is the mean matter density of the Universe at the time in question, and \( \rho_{\text{crit}} \) is the corresponding critical density for closure. Delta is the density contrast. Under this assumption, we have

\[ \rho(r) = \frac{V_h^2}{4\pi G r^2}, \quad r_h = \sqrt{\frac{200}{\Delta_h \Omega_m}} \frac{V_h}{10H(t)}, \quad V_h = \sqrt{\frac{GM_h}{r_h}}, \]  

(7.137)

where \( M_h \) is the mass of the halo (i.e. the mass within \( r_h \)), \( V_c \) is its circular velocity at \( r_h \), and \( H(t) \) is the Hubble constant at time \( t \). It is common practice to refer to \( r_h \) and \( V_c \) as the virial radius and virial velocity.

A physically motivated choice would be to take \( \Delta_h = \Delta_{\text{vir}} \). However, since the criterion for virialization is not strict, other definitions are also in use in the literature. For example, some studies adopt \( \rho_h = 200 \times \rho \), so that \( \Delta_h = 200 \), while others use \( \rho_h = 200 \times \rho_{\text{crit}} \), so that \( \Delta_h = 200/\Omega_m \). Note that different definitions imply different relations among \( M_h, r_h \) and \( V_h \).
The isothermal model is, at best, an approximation. Many effects may cause deviations from the profile predicted by the simple similarity model.

(i) collapse may never reach an equilibrium state in the outer region of a dark matter halo,
(ii) non-radial motion may be important
(iii) mergers associated with the (hierarchical) formation of a halo may render the spherical-collapse model invalid.
Can only be addressed numerically.

Using high resolution $N$-body simulations of structure formation in a CDM cosmogony, Navarro et al. (1996) showed that the density profiles of the simulated dark matter halos are shallower than $r^{-2}$ at small radii and steeper at large radii. Navarro, Frenk & White (NFW) profile:

$$
\rho(r) = \rho_{\text{crit}} \frac{\delta_{\text{char}}}{(r/r_s)(1+r/r_s)^2} \quad \text{(NFW profile).}
$$

Here $r_s$ is a scale radius, and $\delta_{\text{char}}$ is a characteristic overdensity. The logarithmic slope of the NFW profile changes gradually from $-1$ near the center to $-3$ at large radii, and only resembles that of an isothermal sphere at radii $r \sim r_s$.

In a follow-up paper, Navarro et al. (1997) found Eq. (7.138) to be a good representation of the equilibrium density profiles of dark matter halos of all masses and in all CDM-like cosmogonies. Thus, halos formed by dissipationless hierarchical clustering seem to have a universal density profile. The enclosed mass of the NFW profile is given by

$$
M(r) = 4\pi \rho \delta_{\text{char}} r_s^3 \left[ \ln(1 + cx) - \frac{cx}{1 + cx} \right],
$$

where $x \equiv r/r_h$, and

$$
c \equiv \frac{r_h}{r_s} \quad \text{(7.140)}
$$
is the halo concentration parameter.
Internal Structure of Dark Matter Halos

Note that the total mass, $M_h$, of the halo is given by Eq. (7.139) with $x = 1$, and that this mass depends on the definition chosen for $r_h$ (different definitions for the bounding radius of a halo are in common use - different mass and concentration). Using Eqs. (7.138) – (7.140), we obtain a relation between characteristic overdensity and concentration parameter,

$$
\delta_{\text{char}} = \frac{\Delta_h}{3 \ln(1+c) - c/(1+c)}.
$$

-for a given cosmology, the NFW profile is completely characterized by its mass, $M$, and its concentration parameter, $c$, or equivalently by $r$ and $\delta_{\text{char}}$.

Navarro et al. (1997) showed that characteristic overdensity is closely related to formation time. Halos which form earlier are more concentrated. For a definition of formation time similar to that of Eq. (7.91), they found that the ‘natural’ relation $\delta_{\text{char}} \propto \Omega_m(1 + z_f)^3$ describes how the overdensity of halos varies with their formation redshift.

Halos that have experienced a recent major merger typically have low concentrations, $c \sim 4$, while halos which have experienced a longer phase of relatively quiescent growth have larger concentrations (Zhao et al., 2003a, 2008). Physically, the central structure of a dark matter halo seems to be established through violent relaxation during a phase of major mergers, which leads to a ‘universal’ NFW profile with $c \sim 4$. Later accretion does not add much material to its inner regions, thus increasing $r_h$ while leaving $r_s$ almost unchanged.

More massive halos assemble later, and thus had their last major merger more recently - lower concentration.
Although the NFW profile is widely used, numerical simulations of high resolution (Navarro et al., 2004; Hayashi & White, 2008; Gao et al., 2008; Springel et al., 2008) have shown that density profiles of dark matter halos show small deviations from NFW form and are more accurately described by an Einasto (1965) profile,

\[ \rho(r) = \rho_2 \exp \left( \frac{-2}{\alpha} \left( \frac{r}{r_2} \right)^{-\alpha} - 1 \right) \]  

(7.143)

with \( r_2 \) the radius at which the logarithmic slope of the density distribution is equal to \(-2\) and \( \rho_2 = \rho(r_2) \). The best-fit values for the index \( \alpha \) typically span the range \( 0.12 < \alpha < 0.25 \) and increase systematically with increasing mass. Note that the systematic variation of \( \alpha \) with halo mass demonstrates a small but significant deviation of mean density profiles from any “universal” shape.

A characteristic property of the Einasto profile is that its logarithmic slope is a power-law in radius:

\[ \frac{\text{dln}\rho}{\text{dln}r} = -2 \left( \frac{r}{r_2} \right)^\alpha. \]  

(7.144)

Thus, contrary to an NFW profile, which has a central \( r^{-1} \) cusp, the logarithmic slope of the Einasto profile continues to become shallower as \( r \rightarrow 0 \) (however it only becomes \(<-1\) at \( 0.01r_2 \)).

Why near-universal profiles? Different initial conditions give similar results, so they must result from relaxation processes in very general circumstances. The formation of a dark matter halo is in general clumpy and chaotic meaning that violent relaxation must play an important role (however cold collapse from asymmetric initial conditions produces objects with near-NFW density profiles, so hierarchical growth of structure is apparently not required).
Halo Shapes

The collapse of overdensities in the cosmic density field is generically aspherical so halos should not be spherical. Halo equidensity surfaces can be described by ellipsoids, each of which is characterized by the lengths of its axes \((a_1 \geq a_2 \geq a_3)\). These axes can be used to specify the dimensionless shape parameters,

\[
s = \frac{a_3}{a_1}, \quad q = \frac{a_2}{a_1}, \quad p = \frac{a_3}{a_2},
\]

(7.145)

and/or the triaxiality parameter:

\[
T = \frac{a_1^2 - a_2^2}{a_1^2 - a_3^2} = \frac{1 - q^2}{1 - s^2},
\]

(7.146)

Note that oblate and prolate shapes correspond to \(T = 0\) and \(T = 1\), respectively. The majority of CDM halos in numerical simulations have \(0.5 < T < 0.85\), and \(0.5 < s < 0.75\). In particular, less massive halos are more spherical, and halos of a given mass become flatter with increasing redshift.

Shape is tightly correlated with its merger history: halos that assembled earlier are more spherical. Halos that experienced a recent major merger have a tendency to be close to prolate, with the major axis reflecting the direction along which the last merger event occurred. There is a strong tendency for the minor axes of halos to lie perpendicular to large-scale filaments. This alignment is found to be stronger for more massive halos, and shows that the shapes of dark matter halos reflect the large-scale tidal field in which they are embedded.
Halo Substructure

When a small halo merges with a significantly larger halo it becomes a sub-halo orbiting within the potential well of its host. As it orbits, it is subjected to strong tidal forces from the host, which cause it to lose mass. In addition, the orbit itself evolves, as the sub-halo is subjected to dynamical friction which causes it to lose energy and angular momentum to the dark matter particles of its host. Whether a sub-halo survives as a self-bound entity depends on its mass (relative to that of its host), its density profile (which is related to its formation redshift) and its orbit.
Halo Substructure

The mass loss of sub haloes can be written as:

\[
\frac{dm}{dt} = \frac{dm}{dr_{\text{tid}}} \frac{dr_{\text{tid}}}{dt}
\]  

(7.152)

where \( r_{\text{tid}} \) is its instantaneous tidal radius. We assume the latter to be the radius where the original (unstripped) density of the subhalo equals the density of its host’s halo at its current orbital radius, \( r_{\text{orb}} \): i.e. \( \rho_{\text{sub}}(r_{\text{tid}}) = \rho_{\text{host}}(r_{\text{orb}}) \). If we assume that both \( \rho_{\text{sub}}(r) \) and \( \rho_{\text{host}}(r) \) are singular isothermal spheres, then

\[
\frac{1}{r_{\text{tid}}} \frac{dr_{\text{tid}}}{dt} = \frac{1}{r_{\text{orb}}} \frac{dr_{\text{orb}}}{dt}
\]

(7.153)

so that both \( m \) and \( r_{\text{tid}} \) decrease in direct proportion to \( r_{\text{orb}} \). The evolution of \( r_{\text{orb}} \) is governed by dynamical friction, and, assuming a circular orbit, is given by

\[
\frac{dr_{\text{orb}}}{dt} = -0.428 \frac{Gm}{V_c r_{\text{orb}}} \ln \Lambda,
\]

(7.154)

with \( V_c \) the circular velocity of the host halo. To leading order, the Coulomb logarithm \( \ln \Lambda \) is just a function of the mass ratio \( m/M \), which is a constant (equal to \( v_c^2/V_c^2 \) where \( v_c \) is the circular velocity of the subhalo) under the physically reasonable assumption that \( M \) should be taken as the mass interior to the satellite’s orbit. In this very simple model the satellite’s mass and the radius of its (circular) orbit thus decrease linearly with time:

\[
\frac{m}{m_i} = \frac{r_{\text{orb}}}{r_{\text{orb,i}}} = 1 - 0.428 \ln \Lambda \frac{v_c^3 t}{V_c^2 r_{\text{orb,i}}},
\]

(7.155)

where \( m_i \) and \( r_{\text{orb,i}} \) are the initial values of \( m \) and \( r_{\text{orb}} \) respectively.
Halo Substructure

Since we expect $\ln \Lambda$ to be of the order of a few and $r_{\text{orb},i} \sim r_{\text{vir}}$, this predicts that newly accreted subhalos should merge completely with their host after of order $0.1(V/V_c)^3 \sim 0.1 m_i/M$ times the initial orbital periods, where $M$ is the total mass of the host. Thus infalling subhalos with mass greater than a few percent that of the main object are predicted to merge rapidly, whereas substantially lower mass objects will survive for long times with little stripping.

High resolution numerical simulations show that while this conclusion is qualitatively correct, Eq. (7.155) is a poor description of the typical mass loss behavior (Diemand et al., 2007). Most accreted halos fall in on highly elongated orbits and lose a large fraction of their mass at first pericenter.

- The most massive and lowest concentration objects continue to lose mass thereafter and are often completely disrupted.
- Less massive and more concentrated objects often stabilize on their new orbit at their reduced mass. Thereafter they evolve rather little either in mass or in orbit.

Regardless of this, host halos assembling earlier will end up with less mass in subhalos, simply because there has been more time for mass loss to operate. In a CDM cosmogony, halos that are more massive assemble later, and are thus expected to have a larger subhalo mass fraction, $f_{\text{0}}$, than lower mass objects.
Another important property of a dark matter halo is its angular momentum. As originally pointed by Hoyle (1949) and first demonstrated explicitly using numerical simulation by Efstathiou & Jones (1979), asymmetric collapse in an expanding universe produces objects with significant angular momentum. This is traditionally parametrized through a dimensionless spin parameter,

$$\lambda = \frac{J}{GM^{5/2}},$$  \hspace{1cm} (7.156)

where $J$, $E$ and $M$ are the total angular momentum, energy and mass of the halo, respectively. For an isolated system, all these quantities are conserved during dissipationless gravitational evolution, and so, therefore, is $\lambda$ itself. The spin parameter thus defined is roughly the square root of the ratio between the rotational and the total energy of the system, and so characterizes the overall importance of angular momentum relative to random motion.

The angular momentum growth stops once a proto-galaxy separates from the overall expansion and starts to collapse. As an approximation, the final angular momentum of the proto-galaxy may therefore be estimated as the value of $J$ predicted at the time $t_f$ when $D(t_f)\delta = 1$. According to the Zel’dovich approximation, this is approximately the time when the object collapses to form a pancake.

This angular momentum may not correspond to the final angular momentum of a dark matter halo, because a dark matter halo may acquire significant amounts of angular momentum during the late stages of non-linear collapse and due to mergers with other halos.