# Chapter 10 Structure formation and the cosmic web

We have seen from the CMB that the early Universe had a very uniform density distribution, with only very small anisotropies ($\sim 1$ part in $10^5$). The CMB anisotropies visible today are thought to originate in quantum fluctuations in the distribution of matter and radiation that occurred at the time of inflation, which led to the acoustic oscillations of baryons and photons discussed in Chapter 7. The CMB anisotropies follow an angular power spectrum with distinct peaks that correspond to physical scales imprinted on the early Universe, and relate to the size of the horizon (the distance light can have travelled in the time so far elapsed) at that time.

In this chapter, and the one that follows, you will consider how the density perturbations present at the epoch of last scattering subsequently evolved to produce large-scale structure in the Universe. You will explore how gravity enabled the small density inhomogeneities of the early Universe to grow with time to produce galaxies, stars, planets and – ultimately – us.

## Objectives

Working through this chapter will enable you to:

- explain the principles and key equations describing the gravitational collapse of matter to form stable dark-matter halos in virial equilibrium
- mathematically resolve tensions between local gravitational collapse and the overall expansion of the Universe
- describe qualitatively how realistic 3D collapse processes evolve and how they differ from simple spherical models
- discuss how numerical simulations are used to model structure formation and how they can be compared with observations.

# 10.1 Growth of density perturbations

## 10.1.1 Hubble flow

● Consider two galaxies isolated in space. Will they move towards or away from each other?

○ The answer is that it depends on the circumstances. Gravity will try to pull them together; expansion of the Universe will tend to separate them. They will move apart unless they are close enough and massive enough for gravity to dominate.

Locally, we can observe the infall of galaxies towards us or their motion away from us by measuring their velocities. Galaxies falling towards us will appear to be blueshifted (i.e. they have negative velocities); the light from galaxies that are moving away from us (i.e. with positive velocities) will appear redshifted.

The radial velocities of galaxies near to the Milky Way are shown in Figure 10.1, which can be compared with the Hubble diagrams in Figures 1.2 and 1.3. In Figure 10.1 we can identify two distinct regions. First, within about 1.3 Mpc, galaxies have random radial velocities with respect to us, which can be either positive or negative. These galaxies incorporate all of the Milky Way's satellites, as well as the Andromeda Galaxy (Messier 31) and its satellites (including Messier 32 and Messier 33). Any of these galaxies could merge with ours in the future; taken together, they define the Local Group of galaxies.

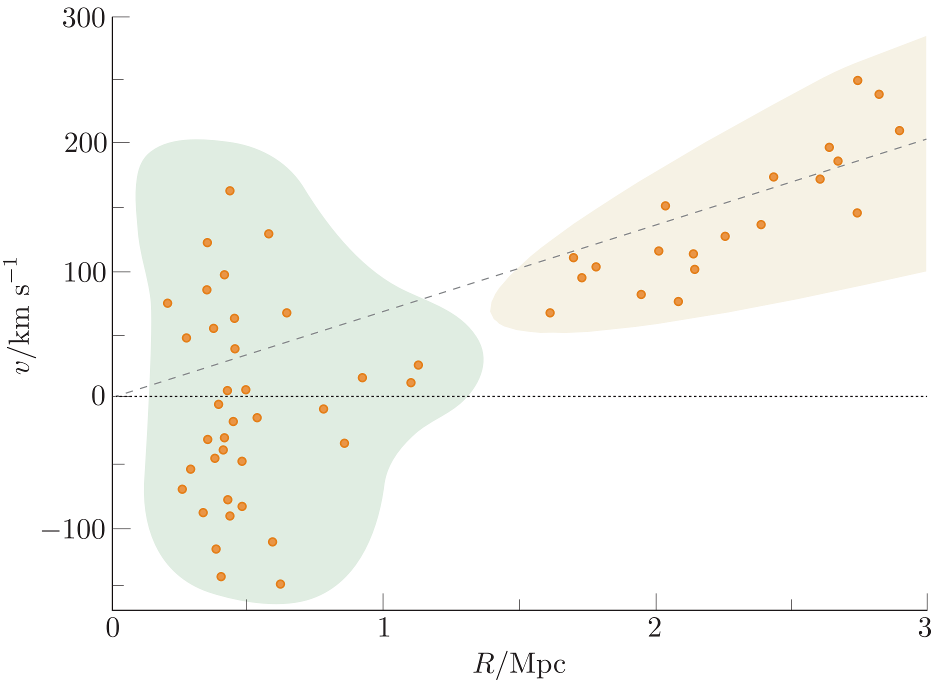



**Figure 10.1** The observed relationship between distance and velocity for galaxies within 3 Mpc of Earth. The dashed line shows the Hubble–Lemaître relationship. The green region shows galaxies that appear bound to the Local Group; the orange-region galaxies have velocities that follow the Hubble flow.

The second region in Figure 10.1, outside the Local Group, shows galaxies whose radial velocity, $V$, is always positive, and becomes well correlated with distance from us, $R$. This is a sign that these objects are being carried away by the Universe's expansion. Although the galaxies in this region are still gravitationally attracted to us, the distance is too great and expansion too fast for them to ever merge with us, assuming the currently favoured model for the evolution of the scale factor, $a(t)$. These objects are described as being entrained in the Hubble flow.

### Exercise 10.1

A galaxy lies 3 Mpc from the Milky Way and is receding from us at $200\,\mathrm{km\,s^{-1}}$. Ignoring continued expansion of the Universe and the presence of any other galaxies, how long would it take gravity to slow the Milky Way and the other galaxy to rest? Assume that each has a mass of $10^{12}\,\mathrm{M_\odot}$.

## 10.1.2 Collapse or expansion?

We have previously discussed the interplay of gravitational attraction and gas pressure in the context of acoustic oscillations of the photon–baryon fluid at the time of CMB production (Chapter 7). Gravity draws matter (and radiation) inwards towards overdense regions, while gas pressure opposes this infall and causes matter to expand. This contest between gravity and pressure applies to many astrophysical situations – gravitational collapse of gas causes conversion of gravitational potential energy to thermal energy, increasing temperature and therefore pressure, and so potentially causing expansion. Stable stars form when these two forces are able to achieve balance. But in order for collapsed structures to form in the first place, the scales need initially to tip in favour of gravity.

We will start by ignoring dark matter, and considering the general case of a diffuse gas cloud, which, depending on its mass, might eventually form a galaxy or a cluster of stars. Such a cloud will be able to collapse if its mass exceeds a specific mass. This condition is known as the Jeans criterion.

### The Jeans criterion

The **Jeans criterion** states that a cloud of gas will only collapse to form a stable, bound object if it is sufficiently massive for the gravitational attraction to exceed the initial pressure that supports it. The **Jeans mass**, $M_\mathrm{J}$, is the minimum mass that a particular cloud of gas – of particle number density $n$ and temperature $T$ – must exceed in order to collapse:

$$M_\mathrm{J} = \frac{9}{4} \times \sqrt{\frac{1}{2\pi n}} \times \frac{1}{\langle m \rangle^2} \times \left(\frac{k_\mathrm{B}T}{G}\right)^{3/2} \tag{10.1}$$

where $\langle m \rangle$ is the average particle mass. This can also be written in terms of a **Jeans length**, $\lambda_\mathrm{J}$, which is the size of a region of density $\rho$ that contains the Jeans mass:

$$M_\mathrm{J} = \frac{4\pi}{3}\rho\lambda_\mathrm{J}^3 \tag{10.2}$$

where

$$\lambda_\mathrm{J} \approx \sqrt{\frac{k_\mathrm{B}T}{G\rho\,\langle m \rangle}}$$

In summary, a gas cloud of a given density and temperature needs to be more massive than $M_J$, or larger than $\lambda_J$, in order to collapse.

Another way to describe the situation is to ask whether an overdense region will grow with time if its density is greater than the mean background density, $\langle \rho \rangle$, by a factor $\delta$, so that

$$\delta = \frac{\rho - \langle \rho \rangle}{\langle \rho \rangle} \tag{10.3}$$

In other words, will the region become even more overdense? For a baryonic gas (which is assumed to behave as an ideal gas) this would be true if the mass in the overdensity exceeds $M_J$.

The concept of Jeans mass is helpful to understand the factors influencing gravitational collapse, but there are several reasons why the formulae above, which assume a baryonic gas, are too simplified to describe the growth of structure in the early Universe. The next section considers what additional factors must be included for a more accurate model.

### 10.1.3 Limitations of the Jeans criterion

The roles played by radiation and dark matter are two of the most important complications in how ordinary matter collapses to form galaxies.

● Recalling the discussion of acoustic oscillations in Chapter 7, how would you expect radiation to affect the growth of baryon overdensities in the early Universe?

○ Radiation dominated the energy density of the Universe at early times. From general relativity – and discussions in Chapter 7 – we expect the energy density of radiation to contribute to the gravitation driving collapse. Radiation also dominated the *pressure* of the photon–baryon fluid prior to recombination and decoupling, which would increase the opposing pressure compared to a baryon-only scenario.

You will see in later sections that the loss of radiation pressure when the photon–baryon fluid decoupled at the time the CMB was produced is what enabled galaxy-mass overdensities of baryons to begin to collapse rather than oscillate. After the period of radiation domination, which ended after nucleosynthesis took place, and prior to recombination, the dominant form of matter according to the favoured cosmological model was dark matter.

● How would you expect the dark-matter content of the Universe to affect the growth of baryon overdensities in the early Universe?

○ Dark matter is assumed to dominate the matter content of the Universe at all times. As discussed in Chapter 7, around the time of decoupling, baryons (together with photons) oscillated within dark-matter overdensities. Therefore, once matter dominates radiation, the collapse of ordinary matter to form galaxies needs to account for the dominant gravitational influence of dark matter.

The Jeans analysis cannot be applied directly to dark-matter collapse. Particle dark matter does not interact electromagnetically. Collapsing dark matter can be considered to have an effective pressure, related to the increasing kinetic energy of the dark-matter particles, but its pressure cannot be defined in the same way as for an ordinary gas.

A final important consideration is that gravitational collapse takes place in an expanding universe, which changes the density and pressure of all of its components. Collapse to form a stable, bound object must happen on a timescale that is faster than expansion can separate the collapsing material.

# 10.2 Collapse of dark-matter halos

The remainder of this chapter looks at how dark-matter overdensities collapse, beginning with a uniform sphere and gradually making it more realistic. The next chapter then discusses how the baryonic gas later collapsed within these dark-matter overdensities to form galaxies. This order of discussion traces the process as it evolved through time, and also charts how simulations of the early Universe have had to evolve, to take into account the progressively more detailed physics of the real Universe.

**Online resources: Python demonstrations**

A set of Python demonstrations and short exercises based on solving the equations presented in this section is provided in the online module resources. These resources are designed to be relatively simple tasks that reinforce your understanding of these topics, so we strongly encourage you to complete these as you work through this section. This will let you see how changing the initial conditions of the Universe would change the structure that evolves.

## 10.2.1 Collapse of a spherical overdensity

We start by modelling a spherically symmetric cloud of material, which can be modelled as a set of concentric spherical shells of matter. (See Section 9.3.1 for discussion of a similar geometry of concentric shells.) The collapse of such a cloud is dictated by the shell theorems of Newton, briefly mentioned when deriving the Friedmann equations in Section 4.1.1, which describe the gravitational forces acting on a given spherical shell of matter.

**Newton's shell theorems for a spherically symmetric body**

- A body located at a radius smaller than that of a shell of matter experiences no net gravitational acceleration from that shell.
- A body outside a shell of matter experiences gravitational acceleration of the same magnitude and direction as if all the matter in that shell were concentrated at its centre.

Let's first simplify the Universe into a uniform medium of density $\langle\rho\rangle$, comprised entirely of pressureless matter (e.g. dark matter or diffuse baryons). We place into this Universe a single overdensity of $\rho = \langle\rho\rangle(1+\delta)$ and radius $R$, and split this sphere up into infinitesimally thin shells. Because we are performing this calculation in only one dimension, each shell can also be considered to represent the radial trajectory of an individual dark-matter particle. Take a moment to think how this sphere might collapse: which parts of the sphere will collapse in which order?

Newton's shell theorems mean we can treat the overdensity as equivalent to a point mass at radius $R$ and ignore anything outside it. The overdensity has mass

$$M = \frac{4\pi R^3}{3}\langle\rho\rangle(1+\delta) \tag{10.4}$$

Newtonian gravity dictates that a particle will accelerate towards the centre of the overdensity at a rate of

$$\begin{aligned} g &= \ddot{R} \\ &= -\frac{GM}{R^2} \\ &= -\frac{G}{R^2}\left[\frac{4\pi R^3}{3}\langle\rho\rangle(1+\delta)\right] \\ &= -\frac{4\pi G}{3}R\langle\rho\rangle(1+\delta) \end{aligned} \tag{10.5}$$

By modelling the evolution of $R$ over time (or by starting from Kepler's third law), it can be shown that the time it takes the overdensity to collapse to a central point is given by the **free-fall timescale**:

$$t_{\text{ff}} = \sqrt{\frac{3\pi}{32G\langle\rho\rangle(1+\delta)}} \tag{10.6}$$

Converting back from $\langle\rho\rangle(1+\delta)$ to $\rho$, and evaluating the numerical constants, leads to the simpler expression

$$t_{\text{ff}} \approx 18.45 \text{ hours} \left(\frac{\rho}{1\,\text{kg}\,\text{m}^{-3}}\right)^{-1/2} \tag{10.7}$$

## Exercise 10.2

Use information in Table 10.1 to calculate (i) the mean total density and (ii) the free-fall timescale for both the Milky Way and the Local Group.

**Table 10.1** Masses and radii for the Milky Way and the Local Group.

| System | Mass/$M_\odot$ | Radius/kpc |
|---|---|---|
| Milky Way | $1.5 \times 10^{12}$ | 26 |
| Local Group | $3.7 \times 10^{12}$ | $\sim$1000 |

- Note that Equation 10.6 depends only on the matter density. How quickly will shells near the centre of our overdensity collapse compared to those further out?
- Because we have been considering a homogeneous overdensity, all shells have the same density and so their free-fall timescales are identical. All particles within the sphere will reach $R = 0$ at the same time.

When $R = 0$, the entire sphere has collapsed to a point. The entire mass $M$ is contained in this point, giving it effectively an infinite density ($\delta \to \infty$). If we were considering collapse of a baryonic gas, then this could be prevented by the fact that collapse would heat up the gas, leading to increasing pressure, which would create a pressure gradient and (at some point) halt the collapse. However, the situation for non-interacting dark matter is different. We therefore need a way to modify our model dark-matter Universe to stop it collapsing into (very massive) black holes.

### Online resources: free-fall collapse demo

The online module resources include a Jupyter Notebook that models free-fall collapse by directly modelling how particles evolve under Newtonian gravity. Use this notebook to explore free-fall collapse and confirm the timescale calculations made in Exercise 10.2.

## 10.2.2 Incorporating expansion of the Universe

One simple modification we can make to our model is to include the expansion of the Universe. As you saw in earlier chapters, this expansion is usually characterised by the behaviour of the scale factor, $a(t)$, which evolves according to the overall matter and energy density. We will explore how expansion alters gravitational collapse by considering a spatially flat ($k = 0$), matter-dominated ($\Lambda = 0$) universe, which is an appropriate model for the early stages of gravitational collapse.*

We can take Equation 10.5, divide through by $R$, and write it as the sum of two terms:

$$\frac{\ddot{R}}{R} = -\frac{4\pi G}{3}\langle\rho\rangle - \frac{4\pi G}{3}\langle\rho\rangle\delta \tag{10.8}$$

Because $\langle\rho\rangle$ is the mean density of the Universe (at the time being considered), the first term represents the self-gravity of any typical region of the Universe of radius $R$. In our $\Lambda = 0$ model, this corresponds to the global deceleration of the Universe's expansion owing to the self-gravity of the matter within it, which occurs via changes to the scale factor, $a$. If we set $\delta = 0$ we obtain

$$\frac{\ddot{R}}{R} = -\frac{4\pi G}{3}\langle\rho\rangle \tag{10.9}$$

*The derivation in this section follows Ryden (2017), pp. 213–216.

If we consider a region of size $R = a$, then the left-hand side becomes $\ddot{a}/a$, and Equation 10.9 is equivalent to the cosmological acceleration equation (Equation 4.35) for the $k = 0$ matter-only ($\Lambda = 0$) case.

The second term in Equation 10.8 represents the local, additional self-gravity of the particular overdensity being considered. To compute this, we can first recognise that the mass of the overdensity is constant, so following Equation 10.4,

$$R^3 \propto \frac{1}{\langle \rho \rangle (1+\delta)}$$

However, because $\langle \rho \rangle \propto a^{-3}$, this becomes

$$R \propto a(1+\delta)^{-1/3} \tag{10.10}$$

We are interested in the rate of change of $\delta$ with time, because this describes how the overdensity grows. Taking the second time derivative of Equation 10.10 and using the chain rule for derivatives of $(1+\delta)$ gives, after some manipulation:

$$\frac{\ddot{R}}{R} = \frac{\ddot{a}}{a} - \frac{1}{3}\ddot{\delta}(1+\delta)^{-1} - \frac{2}{3}\frac{\dot{a}}{a}\dot{\delta}(1+\delta)^{-1} + \frac{4}{9}\dot{\delta}^2(1+\delta)^{-2} \tag{10.11}$$

When $|\delta| \ll 1$, this can be simplified† to

$$\frac{\ddot{R}}{R} \approx \frac{\ddot{a}}{a} - \frac{1}{3}\ddot{\delta} - \frac{2}{3}\frac{\dot{a}}{a}\dot{\delta} \tag{10.12}$$

We can now substitute for the left-hand side using Equation 10.8 to find

$$-\frac{4\pi G}{3}\langle \rho \rangle - \frac{4\pi G}{3}\langle \rho \rangle \delta \approx \frac{\ddot{a}}{a} - \frac{1}{3}\ddot{\delta} - \frac{2}{3}\frac{\dot{a}}{a}\dot{\delta} \tag{10.13}$$

We can relate the first term on the left-hand side to $\ddot{a}/a$ via Equation 10.9 where a region of size $a$ is considered. This cancels the first term on each side to leave

$$-\frac{1}{3}\ddot{\delta} - \frac{2}{3}\frac{\dot{a}}{a}\dot{\delta} = -\frac{4\pi G}{3}\langle \rho \rangle \delta \tag{10.14}$$

Remembering that the Hubble parameter $H = \dot{a}/a$, and rearranging slightly gives

$$\ddot{\delta} + 2H\dot{\delta} = 4\pi G \langle \rho \rangle \delta \tag{10.15}$$

We now have what looks like a relatively straightforward differential equation that describes how a density perturbation changes with time. In order to solve this for $\delta(t)$ and find out how the perturbation grows or shrinks, we need to recall that both $H$ and $\langle \rho \rangle$ change with time as well.

We noted earlier that we are considering a matter-dominated, $\Lambda = 0$ Universe, which is an appropriate model for the early stages of structure

†This is because terms in $1+\delta$ reduce to 1, and the final $\dot{\delta}^2$ term can be neglected because the square of a quantity $\ll 1$ must be much smaller than terms involving the unsquared quantity.

formation. In Chapter 4 you saw that in this model $a(t) \propto t^{2/3}$, which means that $H(t) \propto t^{-1}$, and $\langle \rho \rangle = \rho_{\rm m} \propto t^{-2}$.

The following example uses this information to consider the possible solutions to Equation 10.15.

---

### Example 10.1

With the appropriate proportionality constants (from the Friedmann equation), the proportionalities for $H$ and $\langle \rho \rangle$ can be written in full as

$$H(t) = \frac{2}{3}t^{-1} \qquad \text{and} \qquad \langle \rho \rangle = \rho_{\rm m}(t) = \frac{3}{8\pi G}H(t)^2 = \frac{3}{8\pi G}\frac{4}{9}t^{-2}$$

Use this information to show that

$$\delta(t) = k_1 t^{2/3}$$

where $k_1$ is a constant, is a valid solution to Equation 10.15, and therefore that small perturbations are expected to grow in size in the early matter-dominated Universe, even when expansion of the Universe is accounted for.

### Solution

To show that the provided expression is a valid solution to Equation 10.15, we start by differentiating it (twice) to find the following expressions for $\dot{\delta}$ and $\ddot{\delta}$:

$$\dot{\delta} = \frac{2}{3}k_1 t^{-1/3} \qquad \text{and} \qquad \ddot{\delta} = -\frac{2}{9}k_1 t^{-4/3}$$

We now need to substitute these expressions into Equation 10.15, along with the expressions provided for the time-dependence of $H$ and $\langle \rho \rangle$. Substituting all four expressions in gives a left-hand side of

$$-\frac{2}{9}k_1 t^{-4/3} + 2\frac{2}{3}t^{-1}\frac{2}{3}k_1 t^{-1/3} = \frac{6}{9}k_1 t^{-4/3}$$

and a right-hand side of

$$4\pi G\frac{3}{8\pi G}\frac{4}{9}t^{-2}k_1 t^{2/3} = \frac{6}{9}k_1 t^{-4/3}$$

Therefore the two sides of the equation match, and so we have shown that $\delta(t) = k_1 t^{2/3}$ is a valid solution for the evolution of $\delta$ with time. Because the exponent is positive, this tells us that – in the early, matter-dominated situation being considered – overdense regions will grow with time.

---

Example 10.1 has shown that structure formation can proceed successfully in the early, matter-dominated Universe, despite the counteracting effect of the expansion of spacetime. We can also describe this growth as a function of the scale factor and therefore of the observable quantity (redshift) as follows:

$$\delta(t) \propto t^{2/3} \propto a(t) \propto \frac{1}{1+z} \tag{10.16}$$

It is important, however, to remember that our analysis assumed that $\delta \ll 1$, and so this relation becomes increasingly inaccurate as $\delta \to 1$.

The next exercise considers the implications for what we observe in the real Universe.

**Exercise 10.3**

The CMB anisotropies (e.g. Figure 6.9) correspond to density deviations of $\delta \approx 10^{-5}$ at the time of decoupling ($z \approx 1090$). At the time of writing, the galaxies with the highest redshift that have been reliably observed have $z \approx 13$. Use Equation 10.16 to show that the overdensities implied by the CMB could not have evolved into galaxies at $z \approx 13$. Based on what you read in earlier chapters, how might this inconsistency be resolved?

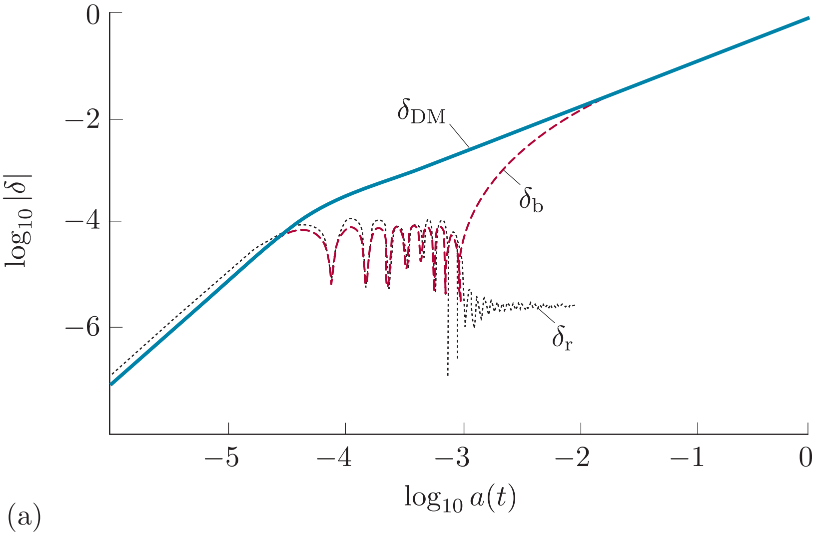
In this section you have seen how the expansion of the Universe can be incorporated into a basic model of gravitational collapse. However, this has not solved the problem of how dark matter is prevented from collapsing to an infinitely dense point. In order to address this, we need to think further about the nature of dark matter and how it must behave during gravitational collapse.

### 10.2.3 Collapsing dark matter and baryons

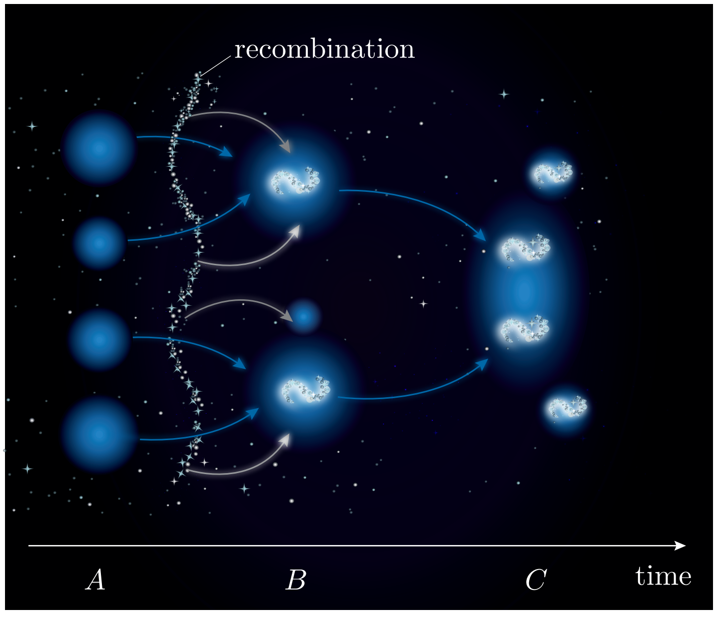
Models of the Universe with overdensities based on the $1:100\,000$ fractional temperature fluctuations of the CMB do not accurately reproduce observed galaxies: we require overdensities at the time of recombination to be at least 100 times greater. This tells us an important fact about the nature of dark matter.

The density of baryonic matter must be strongly coupled to the density of radiation at the time of recombination, so the overdensities cannot be enhanced using any kind of known, baryonic material. Instead, we need some kind of gas- or fluid-like substance that has mass, and so can create its own overdensities, but interacts weakly with the radiation and matter around it. This requirement strongly suggests that dark matter comprises some sort of particle beyond the Standard Model of particle physics, one that does not interact electromagnetically with radiation and ordinary matter particles.

Figure 10.2 illustrates the growth of density perturbations arising from dark matter, and the differing behaviours of baryons and radiation. Panel (a) shows that the dark-matter overdensity, $\delta_{\mathrm{DM}}$, grows smoothly with time, a process that starts at around the time of matter–radiation equality ($z \approx 3400$). Panel (b) shows a more schematic view of the differing behaviours of dark matter (blue regions) and baryonic gas (white regions).




(a)




(b)

**Figure 10.2** (a) The growth of density enhancements in radiation ($\delta_{\rm r}$), dark matter ($\delta_{\rm DM}$) and baryonic matter ($\delta_{\rm b}$); (b) a graphical illustration of this effect at three different times, $A$, $B$ and $C$.

At time $A$, before recombination, overdensities of dark matter develop. These overdensities start clumping together into larger structures (blue arrows). At time $B$, after recombination, baryonic gas (white arrows) is decoupled from radiation, and can start pooling in these overdensities, forming the first galaxies. By time $C$, the overdensities have merged into largely isolated galaxy clusters, each containing many galaxies, which are themselves in the process of merging together.

### 10.2.4 Virialisation

In order to solve the problem of dark-matter collapse to infinite density, we need to assume a more realistic overdensity in which density increases smoothly towards the centre of the overdensity. In this situation the central, high-density parts have a shorter $t_{ff}$ so collapse first; the outer parts take longer to collapse.

This has the consequence that smaller objects (with higher mean density) collapse faster than bigger ones; for example, halos on the scale of dwarf galaxies should collapse first, then begin to coalesce into larger objects like the dark-matter halos that host spiral and elliptical galaxies today.

If we consider a single large overdensity, with shells at different radii, then we can revisit how the collapsing shells evolve. Although each shell passes through $R = 0$, thereby theoretically creating infinite density, each infinitesimally thin shell has infinitesimal mass.

- What happens to a shell's energy as it passes through $R = 0$?
- Gravitational potential energy has been converted into kinetic energy, so the shell continues to expand again. However, the Universe has continued to expand and dilute around it, so instead it remains an overdensity, held together by gravity: it has become a **bound halo**.

Shells rebound at different times and so they will cross one another. This means that each shell no longer has the same mass inside it, and the gravitational potential at each radius changes rapidly. The repeated crossings of shells eventually stabilise to a constant, centrally concentrated mass profile. In the resulting stable configuration, dark matter particles at each radius are sustained by the balance of gravity and kinetic energy and obey the virial theorem (see Section 9.3.1).

It's worth considering how this kinetic energy manifests itself. An initially homogeneous thin shell of particles has become thoroughly mixed into a much more dynamic system of particles repeatedly passing one another and moving with respect to the centre of mass – in other words, behaving like a typical gas. Dark matter appears to be collisionless and non-baryonic, so we *cannot* equate it to a monatomic ideal gas with $\frac{1}{2}mv^2 = \frac{3}{2}k_{\mathrm{B}}T$. Nevertheless, we can think of the distribution of kinetic energies in dark matter as being semi-equivalent to a kinetic temperature. In other words, it is the kinetic temperature of the dark matter that prevents its further collapse. Structure formation places requirements on this microphysical behaviour of dark matter, which helps to constrain models for candidate dark-matter particles. The collapse of baryonic gas within these dark-matter halos will be considered in the next chapter.

# 10.3 Collapse in three dimensions

## 10.3.1 Collapse of an ellipsoid

So far, we have only considered spherical collapse; however, it is more realistic to consider the collapse of a non-spherical object.

### Online resources: ellipsoidal collapse demo

Another demonstration found in the online module resources leads you through the collapse and virialisation of a 3D ellipsoid, so you can visualise how this works in a more realistic setting.

Let's take our sphere and squash two of its dimensions until it becomes a cigar-shaped ellipsoid, as in Figure 10.3. Each axis of the ellipsoid will now collapse and virialise at a different rate, starting with the smallest. When the first axis collapses to $R = 0$, the ellipsoid turns into a flat ellipse, known as a **Zeldovich pancake**, after the physicist Yakov Zeldovich who first developed this theory. Collapsing and virialising the second axis forms a long, cigar-like **filament**; collapsing and virialising the third axis allows matter to flow along the filament into a central dark-matter halo.

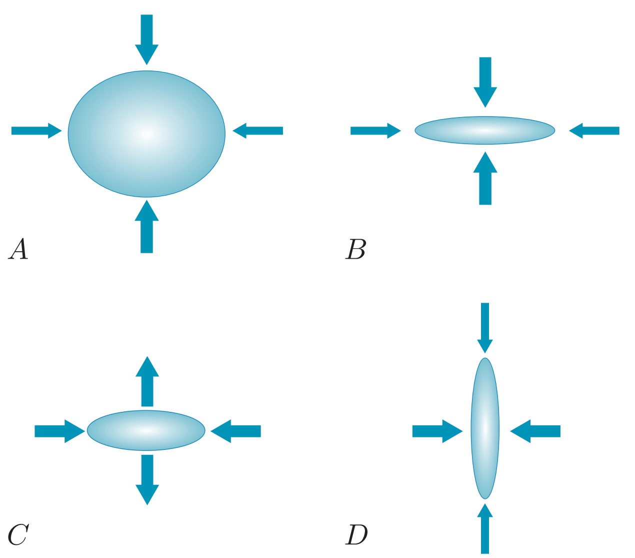



**Figure 10.3** The collapse of an ellipsoid structure into a Zeldovich pancake, shown over four snapshots starting from time $A$ to time $D$.

The slight ellipticity at time $A$ means that the force along the shortest axis is stronger, so this axis collapses first, forming a flattened structure (time $B$). Remember that compression also works in the third dimension, which is not easily shown here. The vertical axis rebounds (time $C$) while the horizontal axis is still collapsing. Finally, the horizontal axis collapses (time $D$), by which point the vertical axis may already have rebounded to its maximum extent. Note that the enhanced elongation (i.e. the pancake shape in three dimensions) is generally retained after the initial collapse.

Now let's consider many overlapping ellipsoids with voids between them. Matter will be pulled out of the voids into these collapsing ellipsoids, which will flatten into sheets (Zeldovich pancakes) joined together by filaments, at the junctions of which will be individual dark-matter halos. We're left with the foam-like structure of the cosmic web, as observed in the real Universe (e.g. as shown in Figure 1.1).

## 10.3.2 Numerical simulations

The real Universe follows the principles discussed in the previous sections, but is more complex than can be fully captured by this kind of intuitive reasoning and cartoon diagrams. Instead, it is necessary to turn to computer simulations like those of the online demonstrations for this chapter, but considerably more complex.

### Online resources: numerical simulation videos

Visualising the complexities of structure formation in the Universe is difficult. The online module resources include links to videos of numerical simulations, where you can watch these collapse processes happen in physically realistic scenarios.

Cosmological simulations start with a near-homogeneous universe, populated with small overdensities and underdensities. The statistical properties of these density variations and their initial expanding motions are generated from an underlying model of cosmology, containing parameters calibrated against our observations of our Universe. Physics is encoded into the model to let the different components of the Universe (radiation, dark matter, baryonic matter and the cosmological constant) interact with themselves and each other according to the laws of physics.

Given these initial physical conditions, the simulation is begun and the model universe is stepped forward in small time steps. The gravitational physics incorporated into the model acts to enhance the initial overdensities so that they grow into the dark-matter halos in which galaxies and galaxy clusters form.

### Types of simulation

Simulations must be programmed efficiently so that they run in the shortest time possible. This allows the Universe to be simulated in the smallest possible time steps, and with the highest possible spatial resolution, to ensure that physics happening on short timescales and in small regions is not ignored.

Astrophysical simulations tend to use one of a number of established modelling approaches. The choice of method depends on the physical situation the programmers want to study. For example, this chapter discusses the collapse of dark-matter halos on large scales, and studies focusing on this collapse would typically use a different approach from the detailed studies of galaxy evolution discussed in the next chapter. Commonly used methods include the following.

- **$N$-body simulations**: These split the Universe up into point masses, each with its own properties (e.g. mass). This makes them useful for looking at situations where objects behave mostly like particles, such as tracking the movements of star particles in galaxy collisions (note that these are not individual stars, but particles that each represent many stars). However, they are not very useful for simulating diffuse material, such as gas or dark matter. The Python demonstrations in this chapter are simple $N$-body simulations.
- **Smoothed-particle hydrodynamic (SPH) simulations**: These are essentially $N$-body simulations in which each particle is 'smoothed', meaning it covers a finite volume. These are useful in situations where there is complex motion, such as in the turbulent flows of galaxy and star formation.
- **Grid simulations**: These take the opposite approach to SPHs, by splitting the Universe up into cubes on a regular 3D grid and applying hydrodynamical equations to evolve the fluid within each cell, with grid cells assigned properties such as density, temperature, etc. Grid simulations are useful for the treatment of gas physics, but they are very inefficient if most of the simulation is empty space, and they can miss important details if the size of the grid is too large. Some grid simulations solve this with **nested sampling**, in which important grid squares are sampled by a sub-grid and can be incremented in smaller time steps than in the rest of the simulation. Grid simulations are also less useful if material moves around within the simulation, because they then have to track what goes through each face of each grid cube.
- **Moving-mesh simulations**: Moving or dynamic meshes solve the problems of empty space and moving material by splitting the Universe up into less-regular regions, tied together by a set of mesh points. These mesh points can be concentrated in the most dynamic regions of space and move with the material they are tracing. However, these still have limitations, for example difficulty in allowing media, such as stars and gas, to pass through each other, because material from both fluids is physically bound to the same mesh nodes.

### Online resources: $N$-body simulation demo

A final demonstration for this chapter, also available in the online module resources, leads you through the first steps of building an $N$-body simulation. This basic set-up can be modified to simulate real systems in the Universe.

Figure 10.4 shows a set of snapshots from a sophisticated moving-mesh cosmological simulation called Illustris. The four rows show the evolution over four redshifts of dark-matter density, baryonic gas density, temperature and metallicity.

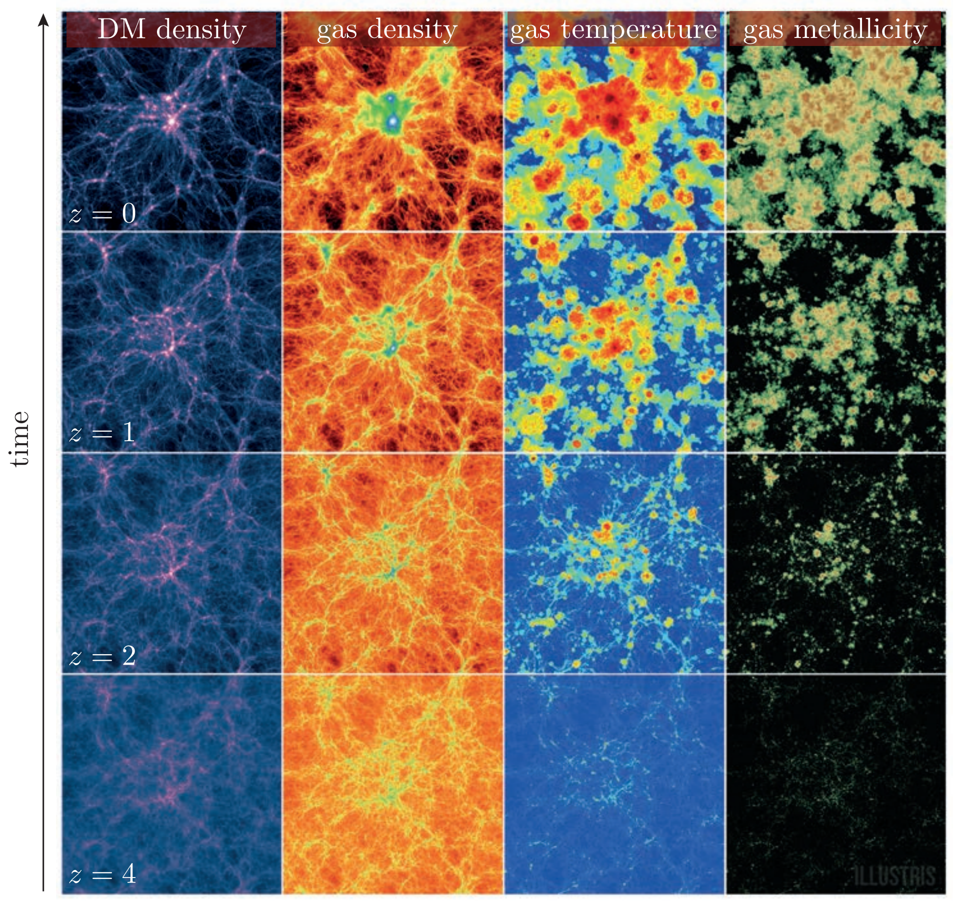



**Figure 10.4** The growth of structure in the Illustris cosmological simulation, shown via snapshots of the distribution of four modelled physical properties (dark matter density, gas density, temperature and metallicity) at four different redshifts ($z$).

In the next chapter you will explore in more detail how baryonic gas evolves within the landscape of the dark-matter structures whose evolution we have examined in this chapter. By incorporating additional physical processes, simulations like the one shown in Figure 10.4 can trace when, where and how individual galaxies form within the cosmic web of dark matter and gas.

### 10.3.3 Testing simulation predictions

The results of cosmological simulations such as the one shown in Figure 10.4 are constantly being compared with a wide variety of observations of galaxies in the real Universe. These comparisons are extremely successful in many cases, but can also reveal interesting disagreements that are used to identify areas where our theories are incomplete or perhaps entirely wrong.

Modern simulations are thought to model the formation of dark-matter structure extremely well. The dark-matter structure in simulations, and in the real Universe, can be characterised by the **halo mass function**, which is the distribution of dark-matter halos according to mass. Figure 10.5 shows the number of halos, $N\,(>M)$, in a fixed volume of a large cosmological simulation.

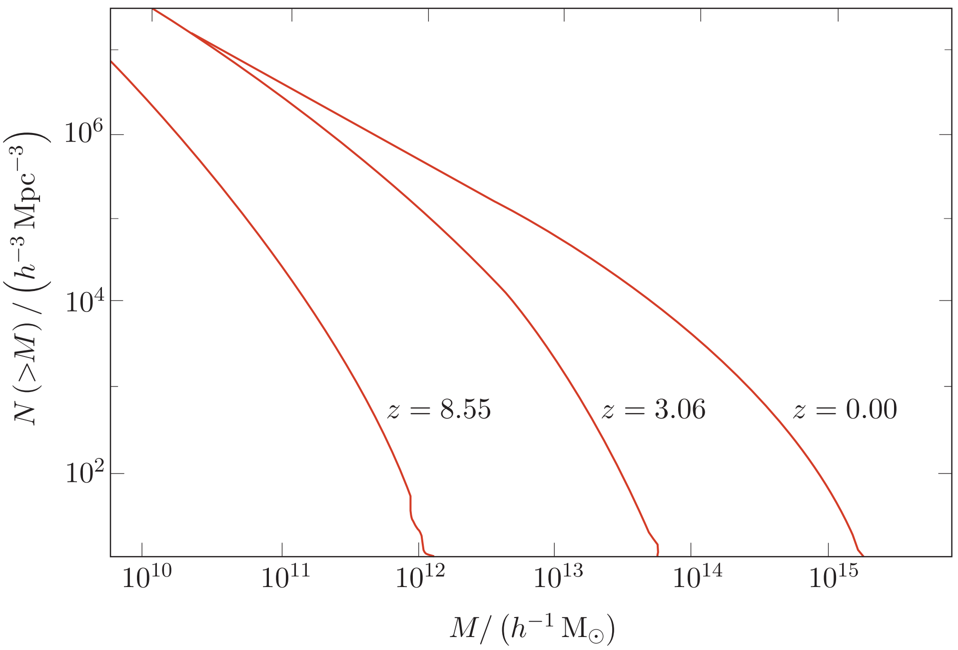



**Figure 10.5** The halo mass function at three redshifts from the Millennium Simulation, showing the number of dark-matter halos ($N$) above a given halo mass ($M$).

- What can you conclude from Figure 10.5 about the relative numbers of low- and high-mass halos and how these change with time?

- At all three redshifts there are many more low-mass halos than high-mass ones. As the age of the Universe increases (redshift decreases), halos become increasingly massive. Halos of the size of galaxy clusters ($10^{14}$ to $10^{15}\,\mathrm{M}_\odot$) only form at late times (after $z = 3$).

The halo mass functions from simulations can be directly compared with observations. Figure 10.6 shows a comparison of the galaxy cluster mass function measured from X-ray observations with a model obtained by fitting a function to cosmological simulation results similar to those in Figure 10.5.

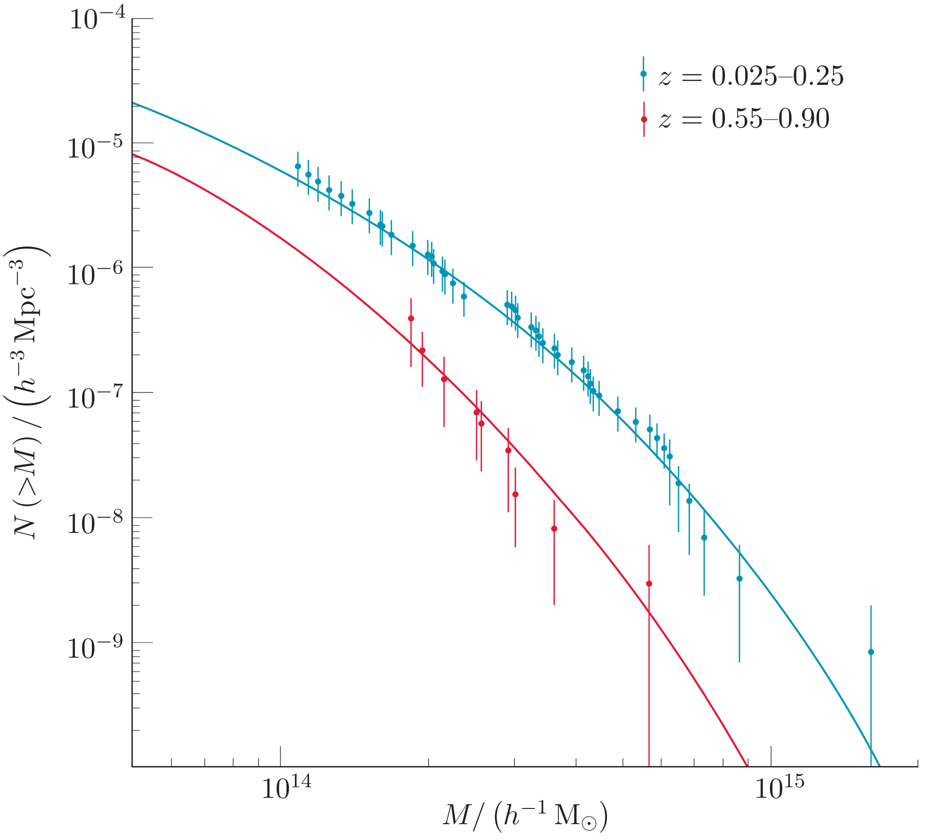



**Figure 10.6** A comparison of the observed cluster mass distribution at two redshifts (blue and red data points), with solid lines indicating model fits based on simulation results for the same redshifts.

It is an impressive success both of modern cosmology and developments in numerical simulation methods that it is possible to predict the distribution of halo masses in the present-day Universe on the basis of evolution under gravity of tiny density perturbations created at very early times. The next chapter will take the story further by exploring how galaxies form within the structures we have explored here.

# 10.4 Summary of Chapter 10

- The expansion of space (Hubble flow) causes galaxies to recede from each other, except on scales of a few megaparsecs, where gravity is strong enough to counteract this expansion.
- The **Local Group** is our gravitationally bound set of galaxies, comprising the galaxies within about 1.3 Mpc of us. It exists as part of larger structures that have formed from the anisotropies present before the formation of the CMB.

- The simplest model for gravitational collapse of structure is based on consideration of the **Jeans criterion**. In this model, objects with masses greater than the Jeans mass can collapse:

$$M_\mathrm{J} = \frac{9}{4} \times \sqrt{\frac{1}{2\pi n}} \times \frac{1}{\langle m \rangle^2} \times \left(\frac{k_\mathrm{B}T}{G}\right)^{3/2} \qquad \text{(Eqn 10.1)}$$

- Overdense regions (halos) will start to collapse on the **free-fall timescale**

$$t_\mathrm{ff} \approx 18.45\ \text{hours}\ \left(\frac{\rho}{1\,\mathrm{kg\,m^{-3}}}\right)^{-1/2} \qquad \text{(Eqn 10.7)}$$

but the collapse timescale will be increased by the expansion of the Universe, which acts to slow collapse.

- Accounting for both gravitational collapse and expansion of the Universe, the growth of an overdensity, $\delta = (\rho - \langle\rho\rangle)/\langle\rho\rangle$ (Equation 10.3), can be modelled to obtain the following equation:

$$\ddot{\delta} + 2H\dot{\delta} = 4\pi G\langle\rho\rangle\delta \qquad \text{(Eqn 10.15)}$$

- Overdensities initially grow at a rate of

$$\delta(t) \propto t^{2/3} \propto a(t) \propto \frac{1}{1+z} \qquad \text{(Eqn 10.16)}$$

- Overdensities in the CMB, originating from the oscillating photon–baryon fluid, aren't large enough to grow the observed galaxy structure. This motivates the need for (non-baryonic) dark matter, of a form that is not coupled to radiation, to drive the collapse of structure.
- Dark-matter halos begin to grow at matter–radiation equality. Baryonic gas can start to collapse into the halos after recombination, when radiation pressure no longer opposes its collapse.
- A collapsing gravitational halo will rebound and virialise to form a semi-stable **bound halo**.
- The most realistic model considers collapse in three dimensions. Collapse first progresses as sheets (i.e. **Zeldovich pancakes**), then **filaments**, and then halos. This creates a foam-like structure, consistent with the observed cosmic web.
- Observed structure in the Universe can be recreated using computer simulations. These simulations take an initial model universe and use physical laws to move that universe forward in time.
- The simulations use different computational techniques to keep track of parcels of material. These include ***N*-body simulations**, hydrodynamical **grid simulations** (fixed or **moving-mesh**, with or without **nested sampling**) and **smoothed-particle hydrodynamic (SPH) simulations**.
- Simulations predict many observable quantities that enable simulation results to be tested. The **halo mass function** is one important, testable prediction of cosmological simulations.