Contents

9 Perturbations after inflation

9.1 Metric perturbations

In the previous chapter we considered the generation of perturbations during inflation. Let us now discuss how these perturbations evolve and form structures. Very non-linear structures such as planets, stars and galaxies have grown from small initial perturbations under the influence of gravity, a process called *structure formation*, and often the distribution of non-linear objects can be treated in terms of linear theory. As in the previous chapter, we will consider linear perturbation theory, so we linearise all equations around a background, and neglect terms higher than first order in the perturbations.

Let us discuss perturbations of the metric. We leave the rigorous development of cosmological perturbation theory to a more advanced course, and just summarise some basic concepts and results. (The interested reader may consult [1, 2] for details; Hannu Kurki-Suonio's lectures may also be useful[1].) We only consider spatially flat backgrounds, as spatial curvature would introduce technical complications, and inflation is expected to make the spatial curvature tiny. The most general linear perturbation around the FLRW metric can be written as

$$\begin{aligned} \mathrm{d}s^2 &= -(1+2\Phi)\mathrm{d}t^2 + 2a(t)(B_{,i} - S_i)\mathrm{d}x^i\mathrm{d}t \\ &\quad + a(t)^2\left[(1-2\Psi)\delta_{ij} + 2E_{,ij} + F_{i,j} + F_{j,i} + 2h_{ij}\right]\mathrm{d}x^i\mathrm{d}x^j \ , \end{aligned} \tag{9.1}$$

where repeated indices are summed over, and a comma stands for derivative with respect to x^i i.e. $f_{,i} \equiv \partial f/\partial x^i$. The perturbations have been written in terms of *irreducible representations* of the group of three-dimensional spatial rotations, in other words in terms of quantities that are closed under rotations. Here Φ, Ψ, B and

[1] https://www.mv.helsinki.fi/home/hkurkisu/cpt/

E are *scalars*, S_i and F_i are *vectors* and h_{ij} is a *tensor*, The vector perturbations are *transverse*, $\delta^{ij}S_{i,j} = \delta^{ij}F_{i,j} = 0$, and the tensor perturbation is transverse and *traceless*, $\delta^{ij}h_{ij} = 0, \delta^{jk}h_{ij,k} = 0$. Physically, tensors correspond to gravity waves, vectors describe rotation and scalars are directly related to the density perturbation, as we will see.

In linear perturbation theory the irreducible scalar, vector and tensor perturbations evolve independently. Vector perturbations are not sourced by inflation (or subsequent evolution) in the linear regime, so we put them to zero, $F_i = S_i = 0$. No tensor perturbations have been detected thus far, but they are an important prediction of inflation.

For the metric perturbation, we have 10 independent functions (vectors included). However, four of them are not physical degrees of freedom, they just correspond to the freedom of choosing the four coordinates. So there are 6 physical degrees of freedom. There are thus different coordinate systems (also called different *gauges*) which describe the same physics. The choice of coordinates is called a choice of gauge[2]. It can be shown that we can choose $E = B = 0$, and that doing so fixes the coordinate system completely. This choice is known as the *longitudinal gauge* and also as *the conformal Newtonian gauge*. We are then left with the metric

$$\mathrm{d}s^2 = -(1+2\Phi)\mathrm{d}t^2 + a(t)^2\left[(1-2\Psi)\delta_{ij} + 2h_{ij}\right]\mathrm{d}x^i\mathrm{d}x^j \ , \tag{9.2}$$

so we have two scalar degrees of freedom and one transverse traceless symmetric tensor, which has two independent degrees of freedom. The metric perturbations $\Phi(t,\boldsymbol{x})$ and $\Psi(t,\boldsymbol{x})$ are called the *Bardeen potentials*[3]. The function Φ is also called the Newtonian potential, since in the Newtonian limit, it becomes equal to the Newtonian potential perturbation, and Ψ is called the Newtonian curvature perturbation, because it determines the curvature of the 3-dimensional $t =$ constant subspaces, which are flat in the unperturbed universe.

9.2 The linear equations of motion

In general relativity, spacetime geometry is described by the metric (expressed in the line element). The evolution of the metric is sourced by matter as described by the *Einstein equation*, the equation of motion of general relativity. We will not go into details of the general description of matter in the theory. We will only consider matter that is an *ideal fluid* or a mixture of several ideal fluids. An ideal fluid has at every point a unique energy density ρ, pressure p, and four-velocity u^α (the index α goes from 0 to 3 and labels spacetime directions). As with the metric, we split the these terms into background plus perturbations,

$$\rho(t,\boldsymbol{x}) = \bar{\rho}(t) + \delta\rho(t,\boldsymbol{x}) \tag{9.3}$$

$$p(t,\boldsymbol{x}) = \bar{p}(t) + \delta p(t,\boldsymbol{x}) \tag{9.4}$$

$$u^\alpha(t,\boldsymbol{x}) = \delta^{\alpha 0} + \delta u^\alpha(t,\boldsymbol{x}) \ , \tag{9.5}$$

[2]More precisely, perturbation theory is formulated in terms of a mapping from the real inhomogeneous and anisotropic spacetime to a background spacetime, and it is the choice of map which is called a "gauge choice". However, the choice of coordinates and choice of mapping are often conflated in cosmological parlance. More simply, change of gauge is a change of coordinates, except that it only affects the perturbations, the background is kept fixed. We will not get into such details.

[3]Warning: sign conventions for Φ and Ψ differ, and the definitions of Ψ and Φ are also sometimes switched with each other.

We will not derive the linearised equation of motion –if you know general relativity, it is easy to do– but just give the result. For the (spatially flat) background we have

$$
\begin{aligned}
3H^2 &= 8\pi G_{\rm N}\bar\rho && (9.6)\\
3(\dot H + H^2) &= -4\pi G_{\rm N}(\bar\rho + 3\bar p) \ , && (9.7)
\end{aligned}
$$

where we have used the relation $\ddot a/a = \dot H + H^2$. For the perturbations, we have

$$
\begin{aligned}
4\pi G_{\rm N}\delta\rho &= \frac{1}{a^2}\nabla^2\Psi - 3H(\dot\Psi + H\Phi) && (9.8)\\
4\pi G_{\rm N}(\bar\rho + \bar p)\delta u_i &= -(\dot\Psi + H\Phi)_{,i} && (9.9)\\
4\pi G_{\rm N}\delta p\delta_{ij} &= \left[(2\dot H + 3H^2)\Phi + H\dot\Phi + \ddot\Psi + 3H\dot\Psi + \frac{1}{2}\frac{1}{a^2}\nabla^2 D\right]\delta_{ij} \\
&\quad -\frac{1}{2}\frac{1}{a^2}D_{,ij} && (9.10)\\
0 &= \ddot h_{ij} + 3H\dot h_{ij} - \frac{1}{a^2}\nabla^2 h_{ij} \ , && (9.11)
\end{aligned}
$$

where $\nabla^2 \equiv \delta^{ij}\partial_i\partial_j$ and $D \equiv \Phi - \Psi$.

From the non-diagonal components of (9.10) we see that $D_{,ij} = 0$ for all $i \neq j$. The general solution of this equation is $D = A(t,x) + B(t,y) + C(t,z)$. In cosmology there are no preferred coordinate axes, so the only physically relevant solution is $D = D(t)$. However, this corresponds to changing the time coordinate, so we can set $D(t) = 0$ without loss of generality. We therefore have $\Phi = \Psi$.[4] To see what the single remaining scalar metric degree of freedom corresponds to, we can manipulate the remaining perturbation equations (9.8)–(9.10). Let us introduce some notation: the *density contrast* is defined as

$$
\delta \equiv \frac{\delta\rho}{\bar\rho} \ . \qquad (9.12)
$$

We also define the background equation of state as $w \equiv \bar p/\bar\rho$, and introduce the variable $v^2 \equiv \delta p/\delta\rho$. We will later see that if $v^2 > 0$, then v corresponds (for certain types of perturbation called *adiabatic*) to the sound speed of the cosmic fluid; if $v^2 < 0$, it is instead related to the instability timescale of the fluid. We can now express the pressure perturbation in terms of v^2 and δ, and write (9.8)–(9.11) as

$$
\begin{aligned}
0 &= \ddot\Phi + H(4 + 3v^2)\dot\Phi - v^2\frac{1}{a^2}\nabla^2\Phi + [2\dot H + (3 + 3v^2)H^2]\Phi && (9.13)\\
\delta &= \frac{2}{3}\frac{1}{(aH)^2}\nabla^2\Phi - 2\frac{1}{H}\dot\Phi - 2\Phi && (9.14)\\
\delta u^i &= \frac{1}{a^2\dot H}\partial_i(\dot\Phi + H\Phi) && (9.15)\\
0 &= \ddot h_{ij} + 3H\dot h_{ij} - \frac{1}{a^2}\nabla^2 h_{ij} \ . && (9.16)
\end{aligned}
$$

[4] In fact, neutrinos develop *anisotropic stress* after neutrino decoupling, so they do not behave like an ideal fluid. Therefore the two Bardeen potentials actually differ from each other by about 10% in the time between neutrino decoupling and matter-radiation equality. After the universe becomes matter-dominated, the neutrinos become unimportant, and Ψ and Φ rapidly approach each other. The same thing happens to photons after photon decoupling, but the universe is then already matter-dominated, so the photons do not cause a significant difference between Ψ and Φ.

From the set of equations (9.13)–(9.15) it follows that the metric perturbation Φ is non-zero only if there is matter. So Φ is generated directly by matter sources, in particular by the density perturbations. In contrast, the tensor perturbation h_{ij} can be non-zero even if the space is empty: they correspond to gravitational waves.

The procedure for solving the perturbed equations is the following.

1) Give the matter model, i.e. give w and v^2.

2) Solve for the evolution of the background and obtain $a(t)$.

3) Solve the perturbation equations.

The order of solving the perturbation equations is that (9.13) gives the evolution of Φ, and we then find the corresponding density contrast from (9.14) and the velocity perturbation from (9.15). (We will not be much concerned with the velocity perturbation.) Note an important difference in (9.14) from the classical Poisson equation: there are terms of the metric perturbation without any gradients on the right-hand side. This is a purely general relativistic feature which has very important consequences, as we will see.

9.3 Fourier transformation

Since the equations are linear, they are easily solved in terms of a Fourier transformation. As in the previous chapter, we define

$$\Phi(t,\boldsymbol{x}) = \frac{1}{(2\pi)^{3/2}}\int \mathrm{d}^3k\Phi_{\boldsymbol{k}}(t)e^{i\boldsymbol{k}\cdot\boldsymbol{x}} \ , \tag{9.17}$$

and define $\delta_{\boldsymbol{k}}$, $u^i_{\boldsymbol{k}}$ and $h_{\boldsymbol{k}ij}$ in the same way. Different Fourier modes decouple, and the equations reduce to ordinary second order differential equations for each mode. Inserting (9.17) into (9.13)–(9.16) we get (we drop the velocity equation)

$$0 = \ddot{\Phi}_{\boldsymbol{k}} + H(4+3v^2)\dot{\Phi}_{\boldsymbol{k}} + v^2\frac{k^2}{a^2}\Phi_{\boldsymbol{k}} + [2\dot{H} + (3+3v^2)H^2]\Phi_{\boldsymbol{k}} \tag{9.18}$$

$$\delta_{\boldsymbol{k}} = -\frac{2}{3}\frac{k^2}{(aH)^2}\Phi_{\boldsymbol{k}} - 2\frac{1}{H}\dot{\Phi}_{\boldsymbol{k}} - 2\Phi_{\boldsymbol{k}} \tag{9.19}$$

$$0 = \ddot{h}_{\boldsymbol{k}ij} + 3H\dot{h}_{\boldsymbol{k}ij} + \frac{k^2}{a^2}h_{\boldsymbol{k}ij} \ . \tag{9.20}$$

The above equations have an interesting property. For a fluid for which $v^2 = w$, the last term in (9.18) vanishes due to (9.6) and (9.7). Thus, for long wavelength perturbations, $k \ll aH$, we find that $\Phi_{\boldsymbol{k}} =$ constant is a solution, and (9.19) shows that the density contrast $\delta_{\boldsymbol{k}}$ is then also constant and equal to $-2\Phi_{\boldsymbol{k}}$. The gravitational waves also have a constant solution, regardless of v^2 or the equation of state, as long as $k \ll aH$. So the relativistic equations allow for the possibility that perturbations with wavelengths much larger than the Hubble scale are frozen and remain unaffected by cosmological evolution. Such a feature is not present in Newtonian gravity.

In the first part of the course we saw that the early universe is radiation-dominated until $t = t_{\rm eq} \approx 50\,000$ years, after which the universe is matter-dominated until it becomes (in the ΛCDM model) dominated by the vacuum energy at around

8 billion years. In order to know the evolution of the perturbations, all we need to do is to plug the background evolution we have already calculated into the above equations and solve, keeping in mind that we have to track at least four different components (photons, neutrinos, baryons and dark matter) with different behaviour (i.e. different w and v^2).

The equations (9.18) and (9.19) give the time evolution of the Fourier components, but the spatial dependence (i.e. dependence on $\boldsymbol{k}$) is left unconstrained, and since the equations are linear, all linear combinations of solutions are also solutions. The spatial dependence is fixed by the initial conditions at early times, given by inflation.

9.4 Evolution on super-Hubble scales

In the previous chapter, we calculated the primordial power spectrum of scalar field fluctuations and noted how it is related to the comoving curvature perturbation $\mathcal{R}$, which is conserved on super-Hubble scales[5], $k \ll aH$. Now we want to know how $\mathcal{R}$ is related to Φ, the scalar metric perturbation in the longitudinal gauge, and to the density contrast δ, and how to go from primordial perturbations to the perturbations seen today.

It can be shown that $\mathcal{R}$ is related to Φ as follows:

$$\mathcal{R} \;=\; -\frac{5+3w}{3+3w}\Phi - \frac{2}{3+3w}H^{-1}\dot{\Phi}\;; \tag{9.21}$$

recall that $w \equiv \bar{p}/\bar{\rho}$. Given $\mathcal{R}$, we can read (9.21) as a differential equation from which to solve Φ. During any period when $w =$ constant, the solution is

$$\Phi_{\boldsymbol{k}} \;=\; -\frac{3+3w}{5+3w}\mathcal{R}_{\boldsymbol{k}} + \text{a decaying part}\;. \tag{9.22}$$

Thus, after w has been constant for some time, the Bardeen potential has settled to the constant value

$$\Phi_{\boldsymbol{k}} = -\frac{3+3w}{5+3w}\mathcal{R}_{\boldsymbol{k}}\;. \tag{9.23}$$

In particular, we have

$$\begin{aligned}
\Phi_{\boldsymbol{k}} \;&=\; -\frac{2}{3}\mathcal{R}_{\boldsymbol{k}} \qquad (\text{rad.dom., } w = \tfrac{1}{3}) \\
\Phi_{\boldsymbol{k}} \;&=\; -\frac{3}{5}\mathcal{R}_{\boldsymbol{k}} \qquad (\text{mat.dom., } w = 0)\;.
\end{aligned} \tag{9.24}$$

Using the relation between Φ and δ given in (9.19), we have, for super-Hubble modes and a constant equation of state,

$$\delta_{\boldsymbol{k}} = -2\Phi_{\boldsymbol{k}} = \frac{6+6w}{5+3w}\mathcal{R}_{\boldsymbol{k}}\;. \tag{9.25}$$

We should now find out how the perturbations evolve when they enter the Hubble radius, and how the situation changes as we pass from radiation domination to matter domination to being dominated by vacuum energy. (**Exercise:** According to (9.25), we would get $\delta_{\boldsymbol{k}} = 0$ for $w = -1$. Explain this in physical terms.)

[5]More precisely, $\mathcal{R}$ is conserved when perturbations are *adiabatic*. We will come back to this shortly.

9.5 Hubble entry

When the expansion of the universe decelerates, i.e. after inflation but before the recent period of accelerated expansion, scales are entering the Hubble radius. Short scales enter earlier, large scales enter later. The history of different scales after Hubble radius entry, and thus their present perturbation amplitude, depends on the epoch during which they enter the Hubble radius. Even if the primordial perturbations are scale-free, the perturbations seen today are not, because different scales have been processed differently. The wavelengths of the modes that enter during transitions between epochs are special scales that characterise the present structure of the universe. Particularly important scales are the inverse wavenumber (or wavelength, if we are not careful about factors of 2π) of modes that enter at the moment of matter-radiation equality $t_{\rm eq}$,

$$k_{\rm eq}^{-1} = (a_{\rm eq}H_{\rm eq})^{-1} \approx 13.7\omega_m^{-1}\ \text{Mpc}\ , \tag{9.26}$$

and the inverse wavenumber of the mode that at the time $t_{\rm dec} \approx 380\,000$ yr of photon decoupling,

$$k_{\rm dec}^{-1} = (a_{\rm dec}H_{\rm dec})^{-1} \approx 90\omega_m^{-1/2}\ \text{Mpc}\ . \tag{9.27}$$

A conservative model-independent observational range is $\omega_m = 0.14 \pm 0.01$ [3, 4]. This gives $k_{\rm eq}^{-1} = 98 \pm 7$ Mpc and $k_{\rm dec}^{-1} = 241 \pm 9$ Mpc. For the ΛCDM model the Planck data gives $\omega_m = 0.1430 \pm 0.0011$ [5], which corresponds to $k_{\rm eq}^{-1} = 95.8 \pm 0.7$ Mpc and $k_{\rm dec}^{-1} = 238 \pm 1$ Mpc. The smallest scale that may be considered cosmological is the typical distance between galaxies, about 1 Mpc. [6] This scale entered the Hubble radius during the radiation-dominated epoch well after Big Bang nucleosynthesis.

The present Hubble length is

$$k_0^{-1} = (a_0H_0)^{-1} \approx 3000h^{-1}\ \text{Mpc} \approx 4000\ \text{Mpc}\ , \tag{9.28}$$

where in the last equality we have used $h = 0.7$. If the expansion is accelerating at the moment[7] this scale is actually exiting now, and there are scales, somewhat larger than this, that have briefly entered and then exited again in the recent past. Modes on the largest observable scales $\sim k_0^{-1}$ have essentially remained at their primordial amplitude.

9.6 Composition of the real universe

In the ΛCDM model, the energy density of the universe has five components:

1. cold dark matter (CDM)

[6] In the present universe, structure at smaller scales has undergone a non-linear process of galaxy formation, and it bears little relation to the primordial perturbations. However, observations of the high-redshift universe, especially so-called Lyman-α observations (absorption spectra of high-z quasars, which reveal distant gas clouds along the line of sight), can reveal these structures when they are closer to their primordial state. With such observations, the "cosmological" range of scales can be extended down to ~ 0.1 Mpc. Other observables such as 21 cm radio emission from hydrogen spin flips can in principle take this down even further.

[7] This is the case in the ΛCDM model, but there are also models where the acceleration has transitioned back into deceleration.

2. baryonic matter
3. photons
4. neutrinos
5. vacuum energy .

The existence of baryons, photons, and neutrinos is beyond reasonable doubt, the existence of dark matter is considered established by most cosmologists (however, warm dark matter remains a plausible alternative to cold dark matter), and the existence and nature dark energy is still a subject of debate. As in the first part of the course, we will stick with the ΛCDM model and only consider vacuum energy. We have

$$\rho = \underbrace{\rho_c + \rho_b}_{\rho_m} + \underbrace{\rho_\gamma + \rho_\nu}_{\rho_r} + \rho_\Lambda \ , \tag{9.29}$$

where we have grouped CDM (denoted with c) and baryons together as matter, and photons and neutrinos as radiation. As we have discussed, neutrinos are actually non-relativistic today and so constitute matter. However, for simplicity we will neglect neutrino masses, as we have done before. (Because the contribution of the neutrinos to the total energy density, or the energy density of matter, is small when they become non-relativistic, this approximation is not too bad.)

Until the decoupling of photons and matter at $t = t_{\rm dec}$, baryons and photons are tightly coupled, so for $t < t_{\rm dec}$ it is useful to treat them as a single component,

$$\rho_{b\gamma} \equiv \rho_b + \rho_\gamma \ . \tag{9.30}$$

We treat the other components as non-interacting (except via gravity). The description of matter as an ideal fluid (i.e. one with a unique density and velocity at every point in space) applies to components whose particle mean free paths are smaller than the scales of interest. After decoupling, photons *free-stream*, i.e. they move almost without scattering, and cannot be discussed as an ideal fluid. On the other hand, the density contrast in the photon component does not grow after decoupling, so we can neglect the effect of photon perturbations compared to perturbations in the matter after decoupling.[8] We make the same approximation for the neutrinos, treating them as an ideal fluid of radiation. We also consider CDM as an ideal fluid: after non-linear structure formation, the ideal fluid assumption of a unique velocity at every point will not be valid any more, but in the linear regime the approximation is reasonable. If the dark energy is vacuum energy, it is perfectly smooth, with no perturbations. (In more complicated dark energy models, perturbations of dark energy are typically not important on small scales, but may have an effect on large scales.)

[8] The CMB perturbations carry important information, and will be the focus of our attention in the next section. However, their influence on the evolution of the total density perturbation is small.

9.7 Multifluid matter

Let us now discuss the general case when the matter consists of several components, which individually can be treated as ideal fluids and which interact with each other only gravitationally. This means that each component sees only its own pressure, and the components can have different flow velocities. Labelling the components with the subscript i, we introduce separate density, pressure, and velocity perturbations for each,

$$\begin{aligned} \rho_i(t,\boldsymbol{x}) &= \bar\rho_i(t) + \delta\rho_i(t,\boldsymbol{x}) && (9.31)\\ p_i(t,\boldsymbol{x}) &= \bar p_i(t) + \delta p_i(t,\boldsymbol{x}) && (9.32)\\ u_i^\alpha(t,\boldsymbol{x}) &= \delta^{\alpha 0} + \delta u_i^\alpha(t,\boldsymbol{x}) \;, && (9.33) \end{aligned}$$

and the total quantities are

$$\begin{aligned} \bar\rho &= \sum_i \bar\rho_i \;, \quad \bar p = \sum_i \bar p_i && (9.34)\\ \delta\rho &= \sum_i \delta\rho_i \;, \quad \delta p = \sum_i \delta p_i \;. && (9.35) \end{aligned}$$

The individual density contrasts are

$$\delta_i \equiv \frac{\delta\rho_i}{\bar\rho_i} \;, \tag{9.36}$$

and the total density contrast is

$$\delta = \frac{\delta\rho}{\bar\rho} = \frac{\sum_i \delta\rho_i}{\sum_j \bar\rho_j} \;. \tag{9.37}$$

Note that the total density contrast is not the sum of the individual density contrasts. Instead, the density contrasts are weighted by the mean densities,

$$\delta = \sum_i \delta_i \frac{\bar\rho_i}{\bar\rho} \;. \tag{9.38}$$

9.8 Adiabatic and isocurvature perturbations

Before going to the evolution of the different components, let us discuss perturbations in the multifluid case. Suppose that the equation of state is *barotropic*

$$p = p(\rho) \;, \tag{9.39}$$

i.e. the pressure is uniquely determined by the energy density. Then the perturbations δp and $\delta\rho$ are necessarily related by the derivative $\mathrm{d}p/\mathrm{d}\rho$ of the function $p(\rho)$,

$$p = \bar p + \delta p = \bar p(\bar\rho) + \frac{\mathrm{d}p}{\mathrm{d}\rho}(\bar\rho)\delta\rho \quad\Rightarrow\quad \delta p = \frac{\mathrm{d}p}{\mathrm{d}\rho}\delta\rho \;.$$

The time derivatives of the background quantities $\bar p$ and $\bar\rho$ are related by the same derivative,

$$\dot{\bar p} = \frac{\mathrm{d}\bar p}{\mathrm{d}t} = \frac{\mathrm{d}p}{\mathrm{d}\rho}(\bar\rho)\frac{d\bar\rho}{\mathrm{d}t} = \frac{\mathrm{d}p}{\mathrm{d}\rho}\dot{\bar\rho} \;.$$

Assuming the derivative $\mathrm{d}p/\mathrm{d}\rho$ is non-negative, its square root is the *sound speed*

$$c_s \equiv \sqrt{\frac{\mathrm{d}p}{\mathrm{d}\rho}} \ . \tag{9.40}$$

We thus have, for barotropic equation of state, the relation

$$v^2 \equiv \frac{\delta p}{\delta \rho} = \frac{\dot{\bar{p}}}{\dot{\bar{\rho}}} = c_s^2 \ .$$

In general, p may depend on other variables besides ρ. The sound speed is then given by

$$c_s^2 = \left(\frac{\partial p}{\partial \rho}\right)_S \tag{9.41}$$

where the subscript S indicates that the derivative is taken so that the entropy of the fluid element is kept constant. Since the background universe expands adiabatically (meaning that there is no entropy production), we have

$$\frac{\dot{\bar{p}}}{\dot{\bar{\rho}}} = \left(\frac{\partial p}{\partial \rho}\right)_S = c_s^2 \ . \tag{9.42}$$

Perturbations with the property

$$\frac{\delta p}{\delta \rho} = \frac{\dot{\bar{p}}}{\dot{\bar{\rho}}} \tag{9.43}$$

are called *adiabatic perturbations*. If $p = p(\rho)$, perturbations are necessarily adiabatic. In the general case, perturbations may or may not be adiabatic. If they are not, the perturbations can be divided into adiabatic perturbations and *isocurvature perturbations*. An adiabatic perturbation corresponds to a change in the total energy density, whereas isocurvature perturbations correspond to perturbations between the different components. For adiabatic perturbations we have

$$\delta p = c_s^2 \delta \rho = \frac{\dot{\bar{p}}}{\dot{\bar{\rho}}} \delta \rho \ . \tag{9.44}$$

Adiabatic perturbations are the simplest kind of perturbations. It is an important prediction of single-field inflation that the perturbations are adiabatic, since all scalar perturbations in all quantities are proportional to the scalar field perturbation $\delta\varphi$.

Adiabatic perturbations have the property that the local state of matter (determined here by the quantities p and ρ) at some spacetime point $(t, \boldsymbol{x})$ of the perturbed universe is the same as in the background universe at some slightly different time $t + \delta t(t, \boldsymbol{x})$, with a different time difference for different locations $\boldsymbol{x}$. We can thus think of adiabatic perturbations in terms of some parts of the universe being ahead and others behind in the evolution, as visualised in figure 1.

For different components i and j we have

$$\left.\begin{aligned} \delta\rho_i(\boldsymbol{x}) &= \dot{\bar{\rho}}_i \delta t(\boldsymbol{x}) \\ \delta p_i(\boldsymbol{x}) &= \dot{\bar{p}}_i \delta t(\boldsymbol{x}) \end{aligned}\right\} \quad \Rightarrow \quad \begin{cases} \dfrac{\delta p_i}{\delta \rho_i} = \dfrac{\dot{\bar{p}}_i}{\dot{\bar{\rho}}_i} \\ \dfrac{\delta \rho_i}{\delta \rho_j} = \dfrac{\dot{\bar{\rho}}_i}{\dot{\bar{\rho}}_j} \end{cases} \tag{9.45}$$

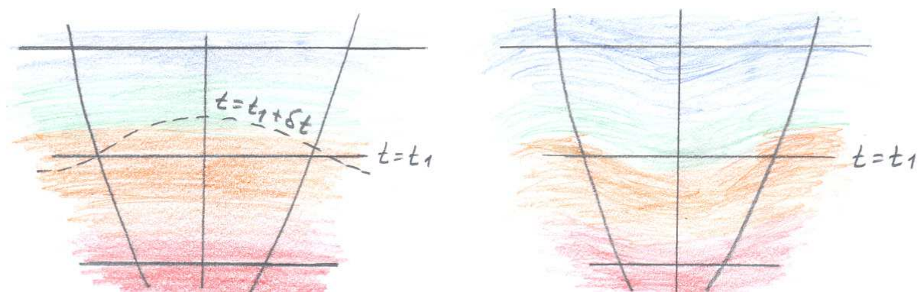

Figure 1: For adiabatic perturbations, the conditions in the perturbed universe (right) at $(t_1, \boldsymbol{x})$ equal conditions in the (homogeneous) background universe (left) at some time $t_1 + \delta t(\boldsymbol{x})$.

If there is no energy transfer between the fluid components at the background level, the energy continuity equation is satisfied by each one separately,

$$\dot{\bar{\rho}}_i = -3H(\bar{\rho}_i + \bar{p}_i) = -3H(1 + w_i)\bar{\rho}_i\,, \tag{9.46}$$

Thus for adiabatic perturbations we have

$$\frac{\delta_i}{1 + w_i} = \frac{\delta_j}{1 + w_j}\,. \tag{9.47}$$

For matter components $w_i = 0$, and for radiation components $w_i = \frac{1}{3}$. Thus, for adiabatic perturbations, all matter components have the same perturbation

$$\delta_i = \delta_m \tag{9.48}$$

and we likewise have for all radiation perturbations

$$\delta_i = \delta_r = \frac{4}{3}\delta_m\,. \tag{9.49}$$

The *isocurvature perturbation* between two components is defined as

$$S_{ij} \equiv -3H\left(\frac{\delta\rho_i}{\dot{\bar{\rho}}_i} - \frac{\delta\rho_j}{\dot{\bar{\rho}}_j}\right) = \frac{\delta_i}{1 + w_i} - \frac{\delta_j}{1 + w_j}\,, \tag{9.50}$$

and it describes deviation from the adiabatic case.

Adiabatic perturbations remain adiabatic while they are outside the Hubble radius and are frozen. However, adiabatic perturbations can (and in general do) evolve into a mixture of adiabatic and isocurvature perturbations once they enter the Hubble radius, because different components evolve differently. In more complicated models of inflation and reheating there can also be primordial isocurvature perturbations (in which case neither isocurvature nor adiabatic perturbations are conserved even on super-Hubble scales). For example, if there are two fields during inflation, and one decays to CDM and the other produces the rest of the matter, there would be primordial isocurvature perturbations between CDM and other forms of matter. Present observational data is consistent with the primordial perturbations being purely adiabatic, and any isocurvature contribution is constrained to be at most a few %, with the precise bound depending on the type of isocurvature perturbation [6].

9.8.1 Multifluid evolution

The background evolution is given by

$$
\begin{aligned}
3H^2 &= 8\pi G_\mathrm{N}\bar\rho && (9.51)\\
3\frac{\ddot a}{a} &= -4\pi G_\mathrm{N}(\bar\rho + 3\bar p) && (9.52)\\
0 &= \dot{\bar\rho}_i + 3H(\bar\rho_i + \bar p_i)\ . && (9.53)
\end{aligned}
$$

If we had energy transfer between components, the left-hand side of (9.53) would be non-zero for the individual components (but still zero for the total energy density and pressure).

Just like the background expansion is sourced by the total energy density and pressure, the metric perturbations are sourced by the perturbations in the total energy density and pressure, so we have, from section 9.3,

$$
\begin{aligned}
0 &= \ddot\Phi_{\boldsymbol k} + H(4+3v^2)\dot\Phi_{\boldsymbol k} + v^2\frac{k^2}{a^2}\Phi_{\boldsymbol k} + [2\dot H + (3+3v^2)H^2]\Phi_{\boldsymbol k} && (9.54)\\
\delta_{\boldsymbol k} &= -\frac{2}{3}\frac{k^2}{(aH)^2}\Phi_{\boldsymbol k} - 2\frac{1}{H}\dot\Phi_{\boldsymbol k} - 2\Phi_{\boldsymbol k}\ . && (9.55)
\end{aligned}
$$

9.9 The radiation-dominated era

After reheating (or, more accurately, preheating: the matter does not need to have a thermal distribution for the energy density to scale like radiation) the universe is dominated by radiation. As late as at BBN, matter contributes only about a fraction of 10^{-6} to the total energy density. So let us first see how the density perturbations evolve in the radiation-dominated universe. In this era, we have to good accuracy for the background energy density $\rho \approx \rho_r \propto a^{-4}$, and spatial curvature and vacuum energy are negligible. We therefore get from (9.6) $a \propto t^{1/2}$, $H = 1/(2t)$. We assume that the perturbations are adiabatic, so we have from (9.49) $\delta_m = \frac{3}{4}\delta_r$ on super-Hubble scales. As long as the growth of matter perturbations on sub-Hubble scales is not too strong (we discuss this below), we then have $\delta\rho_r \gg \delta\rho_m$, so $\delta p/\delta\rho = \delta p_r/\delta\rho \approx \delta p_r/\delta\rho_r = 1/3$ and $\delta \approx \delta_r$ (see (9.37)) to good accuracy. Hence $v^2 = c_s^2 = \frac{1}{3}$.

The general solution of (9.54) and (9.55) is then

$$
\begin{aligned}
\Phi_{\boldsymbol k}(t) &= [y\cos y - \sin y]\, a^{-3}A_{1\boldsymbol k} + [y\sin y + \cos y]\, a^{-3}A_{2\boldsymbol k} && (9.56)\\
\delta_{\boldsymbol k}(t) &= 4\left[(y^2-1)\sin y + y\left(1-\frac{1}{2}y^2\right)\cos y\right] a^{-3}A_{1\boldsymbol k}\\
&\quad +4\left[(1-y^2)\cos y + y\left(1-\frac{1}{2}y^2\right)\sin y\right] a^{-3}A_{2\boldsymbol k}\ , && (9.57)
\end{aligned}
$$

where the behaviour has been conveniently expressed in terms of the variable $y \equiv k/(\sqrt{3}aH) \propto a \propto t^{1/2}$. There are two limiting regimes, perturbations much larger than the Hubble radius ($y \ll 1$) and perturbations deep inside the Hubble radius ($y \gg 1$).

For $y \ll 1$, the mode proportional to $A_{2\boldsymbol k}$ in $\Phi_{\boldsymbol k}$ decays as a^{-3}, while the amplitude of the $A_{1\boldsymbol k}$ mode stays constant, and likewise for $\delta_{\boldsymbol k}$. The non-decaying mode behaviour in the long-wavelength limit is

$$\begin{aligned}\Phi_{\boldsymbol{k}}(t) &= -\frac{1}{9\sqrt{3}}\left(\frac{k}{H_0}\right)^3 A_{1\boldsymbol{k}} = \text{constant} \\ \delta_{\boldsymbol{k}}(t) &= \frac{2}{9\sqrt{3}}\left(\frac{k}{H_0}\right)^3 A_{1\boldsymbol{k}} = \text{constant} \ . \end{aligned} \tag{9.58}$$

On sub-Hubble scales, $y \gg 1$, we have (again dropping the decaying mode)

$$\begin{aligned}\Phi_{\boldsymbol{k}}(t) &= \frac{1}{\sqrt{3}}\left(\frac{k}{H_0}\right) a^{-2}\cos y A_{1\boldsymbol{k}} \propto a^{-2}\cos y \\ \delta_{\boldsymbol{k}}(t) &= -\frac{2}{3\sqrt{3}}\left(\frac{k}{H_0}\right)^3 \cos y A_{1\boldsymbol{k}} \propto \cos y \ . \end{aligned} \tag{9.59}$$

So the gravitational potential decays, while the density perturbation oscillates around a constant amplitude.

Though the physical wavelength of the mode grows like $\propto a$, the Hubble radius stretches faster, $H^{-1} \propto a^2$. (Viewed in comoving terms, the wavelength stays constant, while $aH \propto a^{-1}$ drops.) For super-Hubble modes, the decaying mode becomes negligible, while the non-decaying mode remains constant. Once the wavelength of the mode becomes smaller than the Hubble radius, the density contrast starts to oscillate, and the gravitational potential decays. In both cases, the perturbations remain small.

What about perturbations in the matter during the radiation-dominated era? Baryons are tightly coupled to radiation until $z \approx 1090$, so they have the same perturbations as the radiation fluid. (We will later come back to what happens when baryons and photons decouple; that occurs in the matter dominated era.) However, dark matter decouples from the thermal bath earlier than the baryons, since it interacts weakly. We assume here that dark matter is cold, so its pressure is negligible. After the decoupling of dark matter, its energy density and pressure satisfy the continuity equation individually. Since the dark matter contributes negligibly to the background and to the gravitational potential, we can take (9.56) as a given and see how the dark matter perturbation evolves in this gravitational potential. The derivation for the equations the dark matter density contrast is not complicated, but it requires a bit more general relativity than we have on this course, so we just give the result. For a general FRW background and general metric perturbation Φ, we have

$$\ddot{\delta}_{c\boldsymbol{k}} + 2H\dot{\delta}_{c\boldsymbol{k}} = 3\ddot{\Phi}_{\boldsymbol{k}} + 6H\dot{\Phi}_{\boldsymbol{k}} - \frac{k^2}{a^2}\Phi_{\boldsymbol{k}} \ . \tag{9.60}$$

It is clear that the solution for super-Hubble modes $k \ll aH$ is $\delta_{c\boldsymbol{k}} = \text{constant}$, given that $\Phi_{\boldsymbol{k}} = \text{constant}$. In the opposite limit $k \gg aH$ we get, by inputting $a \propto t^{1/2}$ and (9.59), the solution

$$\delta_{c\boldsymbol{k}} = \tilde{A}_{1\boldsymbol{k}} + \tilde{A}_{2\boldsymbol{k}} \ln y \ , \tag{9.61}$$

where the coefficients $\tilde{A}_{1\boldsymbol{k}}$ and $\tilde{A}_{2\boldsymbol{k}}$ can be written in terms of $A_{1\boldsymbol{k}}$ and $A_{2\boldsymbol{k}}$. (**Exercise.** Calculate $\tilde{A}_{1\boldsymbol{k}}$ and $\tilde{A}_{2\boldsymbol{k}}$ in terms of $A_{1\boldsymbol{k}}$ and $A_{2\boldsymbol{k}}$.) (Recall that if we assume adiabatic initial conditions, we have $\delta_m = \frac{3}{4}\delta_r \approx \delta$.) So, in contrast to baryons, the density contrast of cold dark matter grows logarithmically during the radiation dominated era. The dark matter perturbations thus have a head start on perturbations in baryonic matter, which is tightly coupled to the photons.

9.10 The matter-dominated era

9.10.1 CDM density perturbations

For cold dark matter, it is simple to determine how the perturbations evolve. In the matter-dominated era, we have $\rho \approx \rho_m \propto a^{-3}$, so we get $a \propto t^{2/3}$, $H = 2/(3t)$. Assuming that the initial radiation density contrast is not much larger than that of CDM, we can neglect perturbations in the radiation fluid in the matter-dominated era (as $\delta\rho_r = \delta_r \bar{\rho}_r$). This is always true for adiabatic perturbations. We therefore have $v^2 \approx c_s^2 \approx 0$. The general solution of (9.54) and (9.55) is then

$$\begin{aligned} \Phi_{\boldsymbol{k}}(t) &= B_{1\boldsymbol{k}} + a^{-5/2} B_{2\boldsymbol{k}} \\ \delta_{\boldsymbol{k}}(t) &= -(2y^2+2)B_{1\boldsymbol{k}} - (2y^2-3)a^{-5/2}B_{2\boldsymbol{k}} \ , \end{aligned} \tag{9.62}$$

where $y \equiv k/(\sqrt{3}aH) \propto a^{1/2} \propto t^{1/3}$. Note that with $c_s^2 = 0$, the equation (9.54) for the gravitational potential contains no spatial derivatives, so there are no oscillating solutions. (This is physically obvious: with zero sound speed, there are no sound waves.) For super-Hubble modes, $k \ll aH$, the behaviour is qualitatively the same as in the radiation-dominated era: the decaying mode becomes negligible, and the amplitude of the non-decaying mode remains constant, both for the gravitational potential and the density contrast. However, the short wavelength behaviour is quite different. The gravitational potential is constant, and the density contrast grows like $(aH)^{-2} \propto a \propto t^{2/3}$. It is also noteworthy that (neglecting the decaying mode), the metric perturbation during the matter-dominated era is constant on all scales, not just on super-Hubble wavelengths.

As the universe changes from radiation domination to matter domination, the coefficient $B_{1\boldsymbol{k}}$ is determined in terms of the radiation era coefficient $A_{1\boldsymbol{k}}$: more precisely, the full solution describes a smooth interpolation between the two eras.

9.10.2 Baryon density perturbations

Falling into CDM potential wells. Although CDM is the dominant matter component in the universe, observations are of (light emitted by) baryonic matter. The main method to observe the density perturbations today is to study the distribution of galaxies. To compare the theory of structure formation to observations, it is crucial to know how perturbations in the baryonic component evolve. The issue is complicated by the coupling between baryons and photons.

Before decoupling, baryons evolve as part of the tightly coupled baryon-photon fluid. After decoupling, they are an independent fluid, and the evolution of the baryon density perturbation is driven by the gravitational effect by the total matter density contrast, which includes both baryons and CDM, and is dominated by the latter. On large scales, we can ignore the pressure of the baryonic component, and then δ_b has the same evolution equation as δ_c, namely (9.60). According to (9.50), the baryon-CDM isocurvature perturbation is then

$$S_{cb} = \delta_c - \delta_b \ , \tag{9.63}$$

and it expresses how perturbations in the two components deviate from each other. For both δ_c and δ_b, the right-hand side of (9.60) is the same, so subtracting the equations we get an equation for S_{cb}:

$$\ddot{S}_{cb} + 2H\dot{S}_{cb} = 0 \ . \tag{9.64}$$

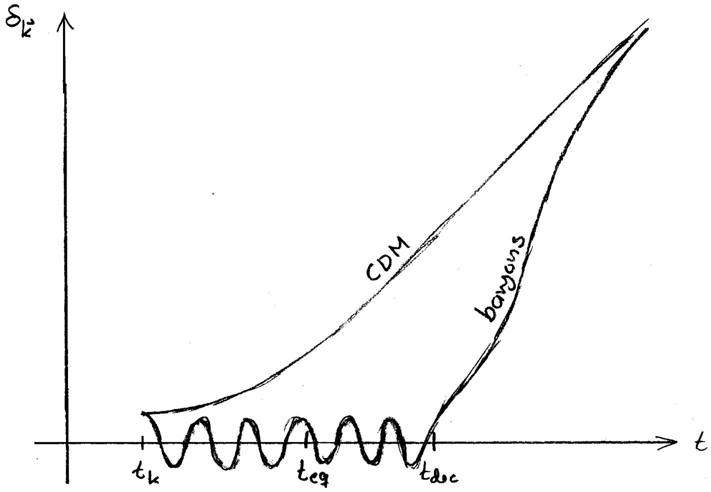

Figure 2: Evolution of the CDM and baryon density perturbations after horizon entry (at $t = t_k$). The figure is just schematic; the upper part is to be understood as having a $\sim$ logarithmic scale; the difference $\delta_c - \delta_b$ stays roughly constant, but the fractional difference becomes negligible as both δ_c and δ_b grow by a large factor.

We assume that the primordial perturbations were adiabatic, so that we originally had $\delta_b = \delta_c$, i.e. $S_{cb} = 0$ at Hubble entry. For large scales, which enter the Hubble radius after decoupling, a non-zero S_{cb} does not develop, so the evolution of the baryon perturbations is the same as CDM perturbations. (This is for linear scales: when perturbations become non-linear, baryons and CDM behave differently.)

But for scales which enter the Hubble radius before decoupling, a non-zero S_{cb} develops, because baryon perturbations are coupled to photon perturbations, but CDM perturbations are not. After decoupling, $\delta_c \gg \delta_b$, since δ_c grows, and δ_b oscillates. During the matter-dominated epoch, the solution for S_{cb} is

$$S_{cb} = A + Bt^{-1/3} \ , \tag{9.65}$$

so if we drop the decaying mode, we have $\delta_b = A + \delta_c$. During matter domination, $\Phi_{\boldsymbol{k}}$ is constant according to (9.62), and from (9.60) we find that the growing mode behaves like $\delta_c \propto t^{2/3}$. Thus the constant A (related to the initial density contrasts) quickly becomes irrelevant, and the baryon density contrast δ_b grows to match the CDM density contrast δ_c (see figure 2), and we eventually have $\delta_b = \delta_c = \delta$ to high accuracy.

The baryon density perturbation begins to grow only after t_{dec}, because before decoupling the radiation pressure prevents growth. Without CDM, the density contrast would grow only as $\delta_b \propto a \propto t^{2/3}$ after decoupling (during the matter-dominated period, and the growth stops when the universe becomes dark energy dominated). Thus it would have grown at most by the factor $a_0/a_{\text{dec}} = 1 + z_{\text{dec}} \approx 1090$ after decoupling. In the anisotropy of the CMB we observe the baryon density perturbations at $t = t_{\text{dec}}$. They are too small (about 10^{-5}) for a growth factor of 1090 to give the present observed large scale structure[9].

[9]This assumes adiabatic primordial perturbations, since we see δ_γ, not δ_b. For a time, primordial

CDM solves this problem. CDM perturbations begin to grow earlier, logarithmically in a during the radiation-dominated era and linearly from $t \sim t_{\rm eq}$ onwards, and by $t = t_{\rm dec}$ they are much larger than baryon perturbations. After decoupling the baryons lose support from photon pressure and fall into the CDM gravitational potential wells, catching up with the CDM perturbations. This allows the small baryon perturbations at $t = t_{\rm dec}$ to grow by much more than the factor 10^3 until today. Thus, smallness of the CMB anisotropy is one of the strongest pieces of evidence for dark matter.

The above situation became clear in the 1980s when the upper limits to CMB anisotropy (which was finally discovered by COBE in 1992) became tighter and tighter. Today we have precise measurements of the structure of the CMB anisotropy which are compared to detailed calculations that include CDM, and the argument is raised to a different level – instead of comparing just two numbers we now look at entire power spectra, as we will discuss in the next chapter.

The Jeans equation. Before decoupling, baryons see the photon pressure and their own pressure, while after decoupling, they just see their own pressure. Baryon pressure is much smaller than photon pressure, but it is important on small scales. At the background level, the baryon pressure can be taken to be zero, $\bar{p}_b = 0$, but the perturbation is non-zero, $\delta p_b \neq 0$. After decoupling, baryonic matter is a gas of hydrogen and helium. If we ignore the formation of molecules in the gas and neglect the contribution of helium, so that the gas is monoatomic, we have

$$v^2 = \frac{\delta p_b}{\delta \rho_b} \approx T_b \frac{\delta n_b}{\delta \rho_b} = \frac{T_b}{m_N} , \tag{9.66}$$

where we have taken into account that the temperature is very uniform, and $m_N \approx$ 1 GeV is the nucleon mass. Note that in this case $v^2 = c_s^2 = \partial p_b / \partial \rho_b$. Down until $z \sim 100$, residual free electrons maintain enough interaction between the baryon and photon components to keep $T_b \approx T_\gamma$. During this period, we thus have $c_s^2 \approx 10^{-13}(1+z) \propto 1/a$, using the fact that $T_\gamma = T_0(1+z)$ After that the baryon temperature falls faster than the photon temperature,

$$T_b \propto a^{-2} \qquad \text{whereas} \qquad T_\gamma \propto a^{-1}$$

(as shown in an exercise in chapter 4).

However, even a tiny pressure can be important on small scales. If we take the analogue of (9.60) for the baryonic component, which includes a tiny pressure contribution (we skip the derivation), we get the *Jeans equation*[10], valid on sub-Hubble scales,

$$\ddot{\delta}_{b\boldsymbol{k}} + 2H\dot{\delta}_{b\boldsymbol{k}} + \left(c_s^2 \frac{k^2}{a^2} - 4\pi G_{\rm N} \bar{\rho} \right) \delta_{b\boldsymbol{k}} = 0 \ . \tag{9.67}$$

We have assumed that the universe is spatially flat, so we can also can write this as

$$\ddot{\delta}_{b\boldsymbol{k}} + 2H\dot{\delta}_{b\boldsymbol{k}} + \left(c_s^2 \frac{k^2}{a^2} - \frac{3}{2} H^2 \right) \delta_{b\boldsymbol{k}} = 0 \ . \tag{9.68}$$

baryon entropy perturbations $S_{b\gamma} = \delta_b - \frac{3}{4}\delta_\gamma$ were considered a possible way out, but more precise observations have ruled this out.

[10] Often the Jeans equation is derived starting from the equations of Newtonian gravity, in which context it was originally presented.

We see that the small pressure term c_s^2 is enhanced on small scales by the term k^2. If k is sufficiently large, this term dominates, no matter how small c_s^2 is. The nature of the solution to the Jeans equation depends on the sign of the factor in brackets. Pressure resists compression, so if the first term dominates, we get an oscillating solution, i.e. sound waves. The second term in the brackets is due to gravity. If this term dominates, the perturbations grow. The wavenumber for which the terms are equal,

$$k_J = a\frac{\sqrt{4\pi G_{\mathrm{N}}\bar{\rho}}}{c_s} = \sqrt{\frac{3}{2}}\frac{aH}{c_s} \; , \tag{9.69}$$

is called the *Jeans wavenumber*, and the corresponding wavelength

$$\lambda_J = \frac{2\pi}{k_J} \tag{9.70}$$

is *Jeans length*.

For **scales much smaller than the Jeans length**, $k \gg k_J$, we can approximate the Jeans equation by

$$\ddot{\delta}_{b\boldsymbol{k}} + 2H\dot{\delta}_{b\boldsymbol{k}} + c_s^2\frac{k^2}{a^2}\delta_{b\boldsymbol{k}} = 0 \; . \tag{9.71}$$

The solutions oscillate with angular frequency $\omega = 3c_s k/a$ (assuming that c_s is constant, or changes slowly – this is not really quite true, as we have seen). The oscillations are damped by the $2H\dot{\delta}_{\boldsymbol{k}}$ term, thus the amplitude of the oscillations decreases with time. There is no growth of structure on sub-Jeans scales.

For **scales much longer than the Jeans length** (but still sub-Hubble), $aH \ll k \ll k_J$, we have

$$\ddot{\delta}_{b\boldsymbol{k}} + 2H\dot{\delta}_{b\boldsymbol{k}} - \frac{3}{2}H^2\delta_{b\boldsymbol{k}} = 0 \; . \tag{9.72}$$

In the matter-dominated era we have $a \propto t^{2/3}$, and the general solution is

$$\delta_{b\boldsymbol{k}}(t) = C_{1\boldsymbol{k}}t^{2/3} + C_{2\boldsymbol{k}}t^{-1} \; , \tag{9.73}$$

So baryon perturbations on scales larger than the Jeans length but smaller than the Hubble length grow just like CDM perturbations, as we discussed earlier.

The ratio of the Jeans length to the Hubble length is, from (9.69)

$$\frac{\lambda_J}{(aH)^{-1}} = 2\pi\sqrt{\frac{2}{3}}c_s \; . \tag{9.74}$$

Before decoupling, the baryons see the photon pressure, and $c_s^2 \sim \frac{1}{3}$. From (9.74) we would then conclude that before decoupling the baryonic Jeans length is comparable to the Hubble length, so that all sub-Hubble modes are sub-Jeans. Therefore, all sub-Hubble baryon modes oscillate before decoupling. However, this argument is not really correct, because the Jeans equation is not valid when c_s^2 is large. Also, in the period close to decoupling the photon mean free path λ_γ grows rapidly. The fluid description, which we are using for the perturbations, applies only on scales $\gg \lambda_\gamma$, whereas the photon gas is smooth only on scales $\ll \lambda_\gamma$. The

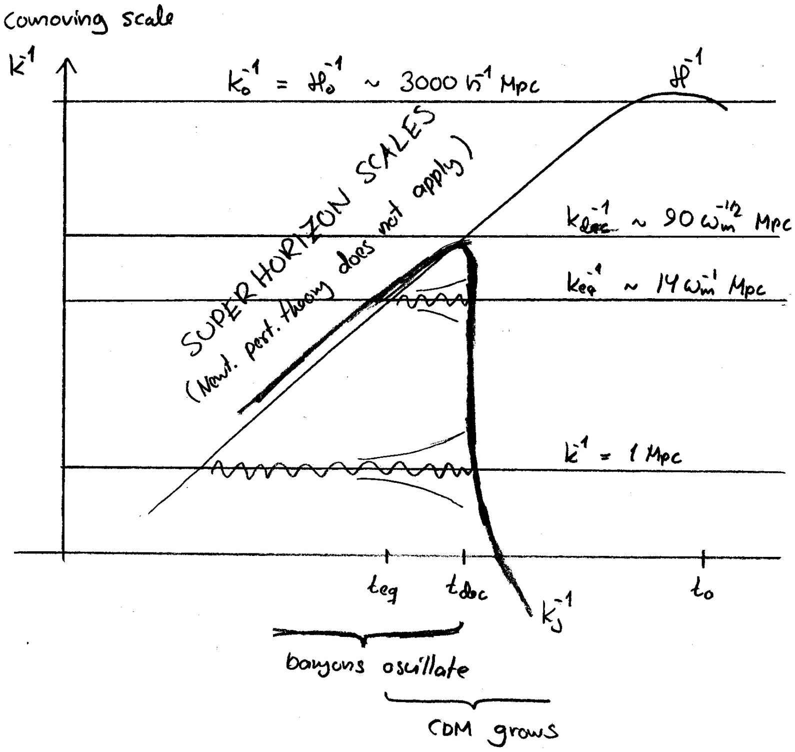

Figure 3: The evolution of perturbations on different subhorizon scales. The baryon Jeans length k_J^{-1} drops precipitously at decoupling so that all cosmological scales became super-Jeans after decoupling, whereas all subhorizon scales were also sub-Jeans before decoupling. The wavy lines symbolise the oscillation of baryon perturbations before decoupling, and the opening pair of lines around them symbolise the $\propto a$ growth of CDM perturbations after $t_{\rm eq}$. There is also logarithmic growth of CDM perturbations between horizon entry and $t_{\rm eq}$.

behaviour during this period can be treated properly only with numerical codes, such as COSMOMC. Nevertheless, the conclusion that all baryonic sub-Hubble modes oscillate before decoupling is correct, at least when perturbations are adiabatic.[11]

After decoupling, the Jeans length grows. However, at all times until today, it is $\ll$ Mpc. It would be relevant if we were interested in the process of the formation of individual galaxies, but here we are looking at larger scales reflected in perturbations of the galaxy number density. Thus for our purposes, the baryonic component is pressureless after decoupling.

The sub-Hubble evolution history of the different cosmological scales of perturbations is summarised in figure 3.

[11]If there is an initial baryon isocurvature perturbation, i.e. a perturbation in baryon density without the corresponding radiation perturbation, it will initially begin to grow in the same manner as a CDM perturbation, since the pressure perturbation provided by the photons is missing. (Such a baryon entropy perturbation corresponds to a perturbation in the baryon-photon ratio η.) But as the movement of baryons drags the photons with them, a radiation perturbation will be generated, and the baryon perturbation will begin to oscillate around its initial value (instead of oscillating around zero).

9.11 The transfer function

Let us summarise the evolution of the linearly perturbed universe. The universe expands as $a \propto t^{1/2}$ in the radiation-dominated era, and then as $a \propto t^{2/3}$ in the matter-dominated era, with a smooth transition around redshift $z_{\rm eq} = 3500$ at 50 000 years. During both eras, perturbations with wavelengths larger than the Hubble radius remain frozen. This means that the properties of the super-Hubble perturbations (i.e. the growing mode amplitudes $A_{1\boldsymbol{k}}$) are preserved from the inflationary era.[12]

As perturbations enter the Hubble radius during the radiation-dominated era, the gravitational potential decays, while the density contrast of photons and baryons oscillates. The density contrast of dark matter grows logarithmically. As the universe becomes matter-dominated, the density contrast of sub-Hubble modes starts to grow $\propto a$, and the gravitational potential stays constant. When the universe becomes dominated by dark energy, perturbations stop growing. (**Exercise:** Show this.)

These effects modify the primordial value of the perturbations, and this is encoded in the *transfer function*. We also express the relation between the primordial curvature perturbation and $\mathcal{R}_{\boldsymbol{k}}$ and any other quantity we are interested in via a transfer function. Since we have only one source of perturbations and perturbations are assumed to be small, the value of any perturbation g at time t is related to the primordial perturbation $\mathcal{R}_{\boldsymbol{k}}$ linearly:

$$g_{\boldsymbol{k}}(t) = T_g(t,k)\mathcal{R}_{\boldsymbol{k}} \ , \tag{9.75}$$

where $T_g(t,k)$ is the transfer function for perturbation g. The transfer function depends only on the magnitude k and not on the direction of $\boldsymbol{k}$, because perturbations evolve in a homogeneous and isotropic background. Often the transfer function separates, $T_g(t,k) = f_g(t)F_g(k)$. In particular, this is the case for cold dark matter if the decaying mode can be neglected. The transfer function incorporates all the physics that determines how structure evolves in the linear regime. The power spectrum of g is

$$\mathcal{P}_g(t,k) = T_g(t,k)^2\mathcal{P}_{\mathcal{R}}(k) \ . \tag{9.76}$$

On scales $k^{-1} \gg 10$ Mpc, perturbations are still small today, and one does not have to go beyond the linear regime transfer function. For smaller scales, corresponding to galaxies and galaxy clusters, the density perturbations have become large at late times, and the physics of structure growth has become nonlinear. As the perturbations become non-linear, modes with different wavenumber become coupled. This nonlinear evolution is typically studied using large numerical simulations, which mostly use Newtonian gravity, although in the past 10 years there are increasingly sophisticated cosmological simulations using general relativity. There are also many analytical results, most (but by no means all) of them in Newtonian gravity.

On scales that are still super-Hubble today, the relation between the density contrast and the primordial perturbations is simple. We have $\delta_m \approx \delta = -2\Phi = \frac{6}{5}\mathcal{R}$, where we have used (9.24). So for $k \gg a_0 H_0$, we simply have $T_\delta(t,k) = \frac{6}{5}$.

[12] In fact, as the equation of state is not barotropic during the transition from radiation domination to matter domination, the amplitude of perturbations undergoes a small change even on super-Hubble scales.

On scales that are sub-Hubble today, the situation is a bit more involved. Let us make a crude estimate of the transfer function on those scales. Let us first look at scales that enter before matter-radiation equality, $k^{-1} < k_{\rm eq}^{-1} \approx 13.7\omega_m^{-1}$ Mpc $\approx$ 100 Mpc. We make the approximation that the relation (9.24) $\Phi_{\boldsymbol{k}} = -\frac{2}{3}\mathcal{R}_{\boldsymbol{k}}$ holds all the way to Hubble entry ($k = aH$), though it is strictly only valid for $k \ll aH$. From (9.56) and (9.57) we have at Hubble entry ($k = aH$, or $y = 1/\sqrt{3}$) $\delta_{\boldsymbol{k}} \approx -\frac{5}{2}\Phi_{\boldsymbol{k}} = \frac{5}{3}\mathcal{R}_{\boldsymbol{k}}$. With adiabatic initial conditions, we have $\delta_m = \frac{3}{4}\delta_r \approx \frac{3}{4}\delta$. We thus get

$$\delta_{c\boldsymbol{k}} \approx \frac{3}{4}\delta_{\boldsymbol{k}} \approx \frac{5}{4}\mathcal{R}_{\boldsymbol{k}} \ . \tag{9.77}$$

at Hubble entry. If we neglect the logarithmic growth of the CDM density perturbations, their amplitude stays at this level until the universe becomes matter-dominated at $t = t_{\rm eq}$, after which we can approximate $\delta_{\boldsymbol{k}} \approx \delta_{c\boldsymbol{k}}$ and $\delta_{\boldsymbol{k}}$ begins to grow according to the matter-dominated law, $\propto 1/(aH)^2 \propto a$. Putting in the logarithmic growth from Hubble entry to matter-radiation equality, the perturbations are in addition enhanced by a factor $\ln(a_{\rm eq}/a_{\rm entry}) = \ln[a_{\rm entry}H_{\rm entry}/(a_{\rm eq}H_{\rm eq})] = 2\ln(k/k_{\rm eq})$, where the subscript entry refers to Hubble entry. So all in all we have, for $k \gg k_{\rm eq}$ in the matter-dominated era

$$\begin{aligned}\delta_{\boldsymbol{k}}(t) &\approx \frac{5}{4}\left(\frac{a_{\rm eq}H_{\rm eq}}{aH}\right)^2 \ln\frac{k}{k_{\rm eq}}\mathcal{R}_{\boldsymbol{k}} \\ &= \frac{5}{4}\left(\frac{k_{\rm eq}}{aH}\right)^2 \ln\frac{k}{k_{\rm eq}}\mathcal{R}_{\boldsymbol{k}} \ . \end{aligned}\tag{9.78}$$

In contrast, for perturbations that enter the Hubble radius during matter domination $k \ll k_{\rm eq}$, we have

$$\begin{aligned}\delta_{\boldsymbol{k}}(t) &= -\frac{2}{3}\left(\frac{k}{aH}\right)^2\Phi_{\boldsymbol{k}} \\ &= \frac{2}{5}\left(\frac{k}{aH}\right)^2\mathcal{R}_{\boldsymbol{k}} \ , \end{aligned}\tag{9.79}$$

where we have used the relation given by (9.24), $\Phi = -\frac{3}{5}\mathcal{R}$.

For a scale-invariant spectrum of primordial comoving curvature perturbations, the amplitude of the density perturbations grows on small scales like k^2. All modes enter ($k = aH$) with approximately the same amplitude, but their amplitude then grows when they are sub-Hubble. However, the modes which entered during the radiation-dominated era have not grown during that era, so their growth is damped by the extra term $(k_{\rm eq}/k)^2$ (modulo the logarithmic growth). This behaviour can be parametrised by introducing a new transfer function $T(k)$, which is defined as

$$\delta_{\boldsymbol{k}} = \frac{2}{5}\left(\frac{k}{aH}\right)^2\mathcal{R}_{\boldsymbol{k}}T(k) \ . \tag{9.80}$$

Putting the above results together, we have

$$\begin{aligned}T(k) &= 1 && k \ll k_{\rm eq} \\ T(k) &\approx \left(\frac{k_{\rm eq}}{k}\right)^2 \ln\frac{k}{k_{\rm eq}} && k \gg k_{\rm eq} \ , \end{aligned}\tag{9.81}$$

where we have dropped factors of order unity from the case $k \gg k_{\rm eq}$, since the calculation is anyway approximate. If we wanted a transfer function which is continuous, we could replace $\ln(k/k_{\rm eq})$ with $\ln(e + k/k_{\rm eq})$. However, our calculation is rather crude, and we should take into account the transition from radiation to matter domination in more detail. An analytical fit to a numerical calculation gives [7]

$$T(k) = \frac{\ln(1+2.34q)}{2.34q\left[1+3.89q+(16.1q)^2+(5.46q)^3+(6.71q)^4\right]^{1/4}} , \qquad (9.82)$$

where $q \approx k e^{f_b}/(14k_{\rm eq})$, and the *baryon fraction* $f_b \equiv \omega_b/\omega_m$ takes into account interactions between baryons and photons, which dampen the matter perturbations. The form (9.82) is called the BBKS transfer function after Bardeen, Bond, Kaiser and Szalay. For the realistic value $f_b = 0.15$, it has an error of around 30% around the turning value $k_{\rm eq}$, while it is accurate for high and low values of k. In detailed calculations, numerical solutions of the baryon-photon-dark matter system are used to derive the transfer function. There are publicly available computer programs for doing this, such as COSMOMC. One of the main effects missing from both (9.81) and (9.82) is *baryon acoustic oscillations* in the regime $k > k_{\rm eq}$. These are remnants of the oscillations of the baryon-photon fluid before decoupling, which are imprinted on the pattern of density fluctuations (and thus the the distribution of galaxies) today. Since there is much more dark matter than baryons, the oscillations are only a small feature in the overall power spectrum, but they carry important cosmological information, much like the CMB anisotropies we discuss in the next chapter. Further discussion of the baryon acoustic oscillations is beyond the scope of this course.

In the ΛCDM model, the universe becomes dark energy dominated as we approach the present time. The equation of state parameter w becomes negative and Φ begins to decay, so the growth of the density perturbations is damped. This effect is not very big up until today (and we shall not calculate it now), since the universe has expanded by less than a factor of 2 after the onset of dark energy domination, but it is important for detailed comparison of observation and theory.

We have calculated everything using linear perturbation theory. It breaks down when the perturbations become large (it's also said that perturbations become nonlinear), $|\delta(\boldsymbol{x})| \sim 1$. This has happened for scales $k^{-1} \lesssim 10$ Mpc by now. When the perturbation becomes nonlinear, i.e. an overdense region becomes about twice as dense as the average density of the universe, it collapses rapidly, and forms a gravitationally bound structure, such as a galaxy or a cluster of galaxies. Further collapse is prevented by the angular momentum of the structure. Stars and gas and CDM particles in a galaxy orbit around the centre of mass of the bound structure, and galaxies in galaxy groups and clusters have more complicated orbits around each other. Underdense regions start to depart from the linear behaviour when they are roughly half as dense as the background. Such regions become ever emptier, as they expand faster than the background.

9.12 The meaning of scale invariance

Inflation predicts and observations give evidence for an almost scale invariant primordial power spectrum. Let us forget the "almost" for a moment and discuss what it means for the primordial power spectrum to be scale-invariant.

The primordial spectrum is something we have at super-Hubble scales, where we have discussed it in terms of the comoving curvature perturbation $\mathcal{R}$. Recall that the perturbation spectrum is called scale invariant when

$$\mathcal{P}_{\mathcal{R}}(k) = A^2 = \text{const.} \;, \tag{9.83}$$

where in the real universe $A \approx 4.6 \times 10^{-5}$.

In terms of the other definition of the power spectrum, $P(k) \equiv (2\pi^2/k^3)\mathcal{P}(k)$ we have

$$\begin{aligned} P_{\mathcal{R}}(k) &\propto k^{-3}\mathcal{P}_{\mathcal{R}} \propto k^{-3} \\ P_{\delta}(k) &\propto k^{-3}\mathcal{P}_{\delta} \propto k\mathcal{P}_{\mathcal{R}} \propto k \;, \end{aligned} \tag{9.84}$$

For $\mathcal{P}_{\mathcal{R}}(k) \propto k^{n-1}$ we have $P_\delta(k) \propto k^n$. This is the reason for the -1 in the definition of the spectral index in terms of $\mathcal{P}_{\mathcal{R}}$ – it was originally defined in terms of P_δ.

We might ask why inflation generates a scale-invariant spectrum – not the mathematical reason (we calculated that in the previous chapter) but the physical idea. During inflation the universe is close to a de Sitter universe, with the metric

$$ds^2 = -dt^2 + e^{2Ht}(dx^2 + dy^2 + dz^2) \;.$$

with $H =$ constant. The de Sitter universe is an example of a *maximally symmetric spacetime*. In addition to being homogeneous (in the space directions), it also looks the same at all times. (This is not obvious from the metric, just like spatial homogeneity is not obvious from the metric for FRW universes with non-zero spatial curvature.) Therefore, modes of different wavelength get the same perturbations imprinted on them regardless of when they leave the Hubble radius.

We would now like to see how the scale-invariance relates to the density perturbation. The power spectrum of density perturbations is

$$\mathcal{P}_\delta(k) = \frac{4}{25}\left(\frac{k}{aH}\right)^4 T(k)^2\mathcal{P}_{\mathcal{R}}(k) \;, \tag{9.85}$$

and for the gravitational potential we have

$$\mathcal{P}_\Phi(k) = \frac{9}{25}\mathcal{P}_{\mathcal{R}}(k)T(k)^2 = \text{constant for } k < k_{\text{eq}} \;. \tag{9.86}$$

We see that perturbations in the gravitational potential are scale invariant (apart from the transfer function), but perturbations in density are not. Instead the density perturbation spectrum is steeply rising on small scales, meaning that there is more structure at small scales than at large scales. Thus the scale invariance refers to the metric perturbations. The density perturbation then turns at $\sim k_{\text{eq}}$ to become almost flat (growing $\sim \ln k$) at small scales, due to the inhibition of the growth of density perturbations during the radiation-dominated era. We can also say that the scale-invariance refers to the density perturbations as they enter the Hubble radius, i.e. density perturbations on all scales enter the Hubble radius with the same amplitude $\frac{2}{5}A \approx 2 \times 10^{-5}$.

The relation between density and gravitational potential perturbations reflects the nature of gravity: a 1% overdense region 100 Mpc across generates a much deeper potential well than a 1% overdense region 10 Mpc across, since the former has 1000 times more mass. Therefore we need much stronger density perturbations at smaller scales to have an equal contribution to Φ.

Thus the perturbations get rapidly stronger on smaller scales, down to $k_{\text{eq}}^{-1} \sim$ 100 Mpc. The $\sim$ 100 Mpc scale appears indeed quite prominent in large scale structure surveys. Towards smaller scales the structures get stronger, but quite slowly. On sufficiently small scales perturbations are so large that first order perturbation theory begins to fail: this limit is crossed at around $k^{-1} \sim 10$ Mpc. Nonlinear effects cause the density power spectrum to rise more steeply than calculated by perturbation theory on scales smaller than this.

The present-day density power spectrum $\mathcal{P}_\delta(k)$ can be determined observationally from the distribution of galaxies. The quantity plotted is usually $P_\delta(k) \equiv (2\pi^2/k^3)\mathcal{P}_\delta(k)$. It goes as

$$\begin{aligned} P_\delta(k) &\propto k^n && \text{for} \quad k \ll k_{\text{eq}} \\ P_\delta(k) &\propto k^{n-4}\ln k && \text{for} \quad k \gg k_{\text{eq}} \ . \end{aligned} \tag{9.87}$$

9.13 Towards the non-linear regime

We earlier presented a simple argument for why dark matter is needed, based on the 10^{-5} amplitude of the observed CMB anisotropies. Because baryons are tightly coupled with photons at the time of last scattering, their density contrast δ_b is also $\sim 10^{-5}$, and since density perturbations grow only linearly with the scale factor, an expansion factor of ~ 1000 is not enough to produce non-linear perturbations. However, the density contrast of dark matter, which is not coupled to the baryons, grows logarithmically during the radiation-dominated era, and so factor of one thousand amplification is enough to give non-linear structures today.

With the more detailed look above, we note that even without the transfer function, the amplitude of the density perturbation, unlike the gravitational potential, depends on the scale. The conclusion that non-linear baryonic structures on the presently observed scales could not have formed without dark matter is correct, but the argument is a bit more subtle. Perturbations on comoving length scale R become non-linear when their density contrast becomes of order unity. The density contrast smoothed on a ball of radius R around the point $\boldsymbol{x}$ is

$$\delta(\boldsymbol{x}, R) \equiv \frac{1}{V}\int W\left(\frac{|\boldsymbol{x}' - \boldsymbol{x}|}{R}\right)\delta(\boldsymbol{x}')d^3x' \ , \tag{9.88}$$

where $W(y)$, the *window function*, is some function which falls off rapidly as $y > 1$, i.e. $|\boldsymbol{x} - \boldsymbol{x}'| > R$, and $V \equiv \int d^3x W(x/R)$ is the *volume* of W. A typical choice of W is a Gaussian, $W(x/R) = \exp[-x^2/(2R^2)]$.

We are not interested in any specific point $\boldsymbol{x}$, but in the typical value of $|\delta(\boldsymbol{x}, R)|$ (the average of $\delta(\boldsymbol{x}, R)$ is zero), so we consider the mean square density contrast

$$\sigma^2(R) \equiv \langle \delta(\boldsymbol{x}, R)^2 \rangle \ . \tag{9.89}$$

where $\langle\rangle$ stands for the spatial average. As we are considering the linear density field, this is just the average over the background space, $\langle \delta(\boldsymbol{x}, R)^2 \rangle = (\int d^3x)^{-1} \int d^3x \delta(\boldsymbol{x}, R)^2$. Structures start forming on comoving scale R when $\sigma(R)$, which grows linearly with the scale factor, reaches unity. Doing a Fourier transform, we can write the mean square density contrast as

$$\sigma^2(R) = \int_0^\infty \frac{\mathrm{d}k}{k}\mathcal{P}_\delta(k,t)W(kR)^2 \ , \tag{9.90}$$

where for a Gaussian window function we have $W(kR) = e^{-\frac{1}{2}k^2R^2}$. For a power law spectrum of density perturbations, $\mathcal{P}(k) = Ak^{n+3}$, we have (**Exercise:** Show this.)

$$\sigma^2(R) = \frac{1}{2}\Gamma\left(\frac{n+3}{2}\right)\mathcal{P}_\delta(R^{-1}) \ . \tag{9.91}$$

So the mean square density contrast on a given comoving scale R is roughly given by the value of the power spectrum at $k = R^{-1}$. The real power spectrum is more complicated because of the transfer function, but it's still the case that the amplitude of density perturbations on a given scale is roughly given by the power spectrum on that scale.

If the transfer function were to continue to have the $k^2 \ln k$ behaviour for very large k without limit, we would have $\mathcal{P}_\delta(k) \sim k^{n-1}[\ln(k/k_{\rm eq})]^2$. So if $n \geq 1$, the power spectrum would reach non-linear values at all times, on sufficiently small scales. So we would always have non-linear structures, albeit on very small scales! However, the radiation-dominated era after inflation has a finite duration, so the amount of logarithmic growth is limited. There is also another effect which wipes out structure on small scales, namely the motion of the dark matter particles, called *free-streaming.*

Even CDM has a finite temperature, which means that dark matter particles have thermal motions, and this smooths density perturbations below some scale, as particles from overdense and underdense regions mix and balance the density perturbations out. For CDM, the transfer function is modified by the term $e^{-k^2/k_{\rm fs}^2}$ for $k \gg k_{\rm fs}$, where $k_{\rm fs}$ is the free-streaming scale, related to the distance the dark matter particles have moved since decoupling. For $k < k_{\rm fs}$, structure formation is unaffected, but on small scales, perturbations are highly suppressed. The smallest scale on which structures form is given by the free-streaming length, which for a WIMP is approximately [8]

$$k_{\rm fs} \approx \left(\frac{m}{100\ \mathrm{GeV}}\right)^{1/2}\left(\frac{T_D}{30\ \mathrm{MeV}}\right)^{1/2} \mathrm{pc}^{-1} \ , \tag{9.92}$$

where m is the mass of the dark matter particle and T_D is its decoupling temperature. The smallest structures for a typical WIMP are therefore of comoving length 1 pc. They form around a redshift of $z = 40 \ldots 60$.

For warm dark matter, the free-streaming scale is larger, so structures on larger scales are wiped out. For example, for light sterile neutrinos (sterile neutrinos are neutrinos that don't have any Standard Model interactions, but they mix with the ordinary neutrinos via neutrino oscillations; they are one prominent warm dark matter candidate), the transfer function is instead modified approximately with the term $[1 + (k/k_{\rm fs})^2]^{-5}$, with [9]

$$k_{\rm fs} \approx \left(\frac{m}{500\ \mathrm{eV}}\right) \ \mathrm{Mpc}^{-1} \ . \tag{9.93}$$

If the sterile neutrino mass were 500 eV, all structures on comoving scales smaller than a Mpc would have been suppressed, in drastic conflict with observations. However, for a mass of say 5 keV, galaxies still form, but smaller structures are suppressed. This was proposed as an explanation for why there seemed to be fewer observed satellites of the Milky Way than predicted in CDM models, but more recent observations have shown that there is no conflict between predictions of CDM and observations as regards the abundance of dwarf galaxies, narrowing the room for warm dark matter.[13] Viewed from another perspective, observations of structures can be used to constrain particle physics dark matter models.

References

[1] V. F. Mukhanov, H. A. Feldman and R. H. Brandenberger, *Theory of cosmological perturbations. Part 1. Classical perturbations. Part 2. Quantum theory of perturbations. Part 3. Extensions*, *Phys. Rept.* **215** (1992) 203–333.

[2] V. Mukhanov, *Physical Foundations of Cosmology*. Cambridge University Press, Oxford, 2005, 10.1017/CBO9780511790553.

[3] M. Vonlanthen, S. Räsänen and R. Durrer, *Model-independent cosmological constraints from the CMB*, *JCAP* **08** (2010) 023, [1003.0810].

[4] B. Audren, J. Lesgourgues, K. Benabed and S. Prunet, *Conservative Constraints on Early Cosmology: an illustration of the Monte Python cosmological parameter inference code*, *JCAP* **02** (2013) 001, [1210.7183].

[5] Planck collaboration, N. Aghanim et al., *Planck 2018 results. VI. Cosmological parameters*, *Astron. Astrophys.* **641** (2020) A6, [1807.06209].

[6] Planck collaboration, P. A. R. Ade et al., *Planck 2015 results. XX. Constraints on inflation*, *Astron. Astrophys.* **594** (2016) A20, [1502.02114].

[7] J. M. Bardeen, J. R. Bond, N. Kaiser and A. S. Szalay, *The Statistics of Peaks of Gaussian Random Fields*, *Astrophys. J.* **304** (1986) 15–61.

[8] A. M. Green, S. Hofmann and D. J. Schwarz, *The power spectrum of SUSY - CDM on sub-galactic scales*, *Mon. Not. Roy. Astron. Soc.* **353** (2004) L23, [astro-ph/0309621].

[9] S. H. Hansen, J. Lesgourgues, S. Pastor and J. Silk, *Constraining the window on sterile neutrinos as warm dark matter*, *Mon. Not. Roy. Astron. Soc.* **333** (2002) 544–546, [astro-ph/0106108].

[13] Note that $k_{\rm fs}^{-1}$ is the comoving scale of the linear perturbation from which the structure formed. The corresponding actual size of the structure today is smaller, because structures contract and then stop expanding when they form, whereas in linear theory they would have been stretched linearly with the scale factor.