8 Structure Formation

Up to this point we have discussed the universe in terms of a homogeneous and isotropic model (which we shall now refer to as the "unperturbed" or the "background" universe). Clearly the universe is today rather inhomogeneous. By *structure formation* we mean the generation and evolution of this inhomogeneity. We are here interested in distance scales from galaxy size to the size of the whole observable universe. The structure is manifested in the existence of luminous galaxies and in their uneven distribution, their *clustering*. This is the obvious inhomogeneity, but we understand it reflects a density inhomogeneity also in other, nonluminous, components of the universe, especially the *cold dark matter*. The structure has formed by gravitational amplification of a small primordial inhomogeneity. There are thus two parts to the theory of structure formation:

1) The generation of this primordial inhomogeneity, "the seeds of galaxies". This is the more speculative part of structure formation theory. We cannot claim that we know how this primordial inhomogeneity came about, but we have a good candidate scenario, *inflation*, whose predictions agree with the present observational data, and can be tested more thoroughly by future observations. In inflation, the structure originates from *quantum fluctuations* of the inflaton field φ near the time the scale in question exits the horizon.

2) The growth of this small inhomogeneity into the present observable structure of the universe. This part is less speculative, since we have a well established theory of gravity, *general relativity*. However, there is uncertainty in this part too, since we do not know the precise nature of the dominant components to the energy density of the universe, the *dark matter* and the *dark energy*. The gravitational growth depends on the equations of state and the streaming lengths (particle mean free path between interactions) of these density components. Besides gravity, the growth is affected by pressure forces.

We shall do the second part first. But before that we discuss statistical measures of inhomogeneity: correlation functions and power spectra.

8.1 Inhomogeneity

We write all our inhomogeneous quantities as a sum of a homogeneous background value, and a perturbation, the deviation from the background value. For example, for energy density and pressure we write

$$\begin{aligned} \rho(t,\mathbf{x}) &= \bar{\rho}(t) + \delta\rho(t,\mathbf{x}) \\ p(t,\mathbf{x}) &= \bar{p}(t) + \delta p(t,\mathbf{x})\,, \end{aligned} \tag{1}$$

where $\bar{\rho}$ and $\bar{p}$ are the background density and pressure, $\mathbf{x}$ is the *comoving* 3D space coordinate, and $\delta\rho$ and δp are the density and pressure perturbations. We further define the relative density perturbation

$$\delta(t,\mathbf{x}) \equiv \frac{\delta\rho(t,\mathbf{x})}{\bar{\rho}(t)}\,. \tag{2}$$

Since $\rho \geq 0$, necessarily $\delta \geq -1$. These quantities can be defined separately for different components to the energy density, e.g., matter, radiation, and dark energy. Perturbations in dark energy are expected to be small, and if it is just vacuum energy, it has no perturbations. When we discuss the later history of the universe, the main interest is in the matter density perturbation,

$$\delta_m(t,\mathbf{x}) \equiv \frac{\delta\rho_m(t,\mathbf{x})}{\bar{\rho}_m(t)}\,, \tag{3}$$

and *then we will often write just* δ *for* δ_m.

We do the split into the background and perturbation so that the background is equal to the mean (volume average) of the full quantity. An important question is, whether the $\bar{\rho}(t)$ and $\bar{p}(t)$ defined this way correspond to a (homogeneous and isotropic) solution of General Relativity, i.e., an FRW universe. We expect the exact answer to be negative, since GR is a nonlinear theory, so that perturbations affect the evolution of the mean. This effect is called *backreaction*.

However, if the perturbations are small, we can make an approximation, where we drop from our equations all those terms which contain a product of two or more perturbations, as these are "higher-order" small. The resulting approximate theory is called *first-order perturbation theory* or *linear* perturbation theory. As the second name implies, the theory is now linear in the perturbations, meaning that the effect of overdensities cancel the effect of underdensities on, e.g., the average expansion rate. In this case the mean values evolve just like they would in the absence of perturbations.

While the perturbations at large scales have remained small, during the later history of the universe the perturbations have grown large at smaller scales. How big is the effect of backreaction, is an open research question in cosmology, since the calculations are difficult, but a common view is that the effect is small compared to the present accuracy of observations. For this course, we adopt this view, and assume that the background universe simultaneously represents a FRW universe ("the universe we would have if we did not have the perturbations") and the mean values of the quantities in the true universe at each time t.

Moreover, in Cosmology II we shall (mostly) assume that the background universe is flat ($K = 0$).

8.1.1 Statistical homogeneity and isotropy

We assume that the origin of the perturbations is some random process in the early universe. Thus over- ($\delta > 0$) and underdensities ($\delta < 0$) occur at randomly determined locations and we cannot expect to theoretically predict the values of $\delta(t, \mathbf{x})$ for particular locations $\mathbf{x}$. Instead, we can expect theory to predict statistical properties of the inhomogeneity field $\delta(t, \mathbf{x})$. The statistical properties are typically defined as averages of some quantities. We will deal with two kind of averages: *volume average* and *ensemble average*; the ensemble average is a theoretical concept, whereas the volume average is more observationally oriented.

We denote the volume average of some quantity $f(\mathbf{x})$ with the overbar, $\bar{f}$, and it is defined as

$$\bar{f} \equiv \frac{1}{V} \int_V d^3x f(\mathbf{x}) \,. \tag{4}$$

The integration volume V in question will depend on the situation.

For the ensemble average we assume that our universe is just one of an *ensemble* of an infinite number of possible universes (*realizations* of the random process) that could have resulted from the random process producing the initial perturbations. To know the random process means to know the probability distribution $\mathrm{Prob}(\gamma)$ of the quantities γ produced by it. (At this stage we use the abstract notation of γ to denote the infinite number of these quantities. They could be the generated initial density perturbations at all locations, $\delta(\mathbf{x})$, or the corresponding Fourier coefficients $\delta_{\mathbf{k}}$. We will be more explicit later.) The ensemble average of a quantity f depending on these quantities γ as $f(\gamma)$ is denoted by $\langle f \rangle$ and defined as the (possibly infinite-dimensional) integral

$$\langle f \rangle \equiv \int d\gamma \mathrm{Prob}(\gamma) f(\gamma) \,. \tag{5}$$

Here f could be, e.g., the value of $\rho(\mathbf{x})$ at some location $\mathbf{x}$. The ensemble average is also called the *expectation value*. Thus the ensemble represents a probability distribution of universes. A

cosmological theory predicts such a probability distribution, but it does not predict in which realization from this distribution we live in. Thus the theoretical properties of the universe we will discuss (e.g., statistical homogeneity and isotropy, and ergodicity, see below) will be properties of this ensemble.

We now make the assumption that, although the universe is inhomogeneous, it is *statistically homogeneous and isotropic*. This is the second version of the *Cosmological Principle*. Statistical homogeneity means that the expectation value $\langle f(\mathbf{x})\rangle$ must be the same at all $\mathbf{x}$, and thus we can write it as $\langle f\rangle$. Statistical isotropy means that for quantities which involve a direction, the statistical properties are independent of the direction. For example, for vector quantities $\mathbf{v}$, all directions must be equally probable. This implies that $\langle\mathbf{v}\rangle = 0$. The assumption of statistical homogeneity and isotropy is justified by inflation: inflation makes the background universe homogeneous and isotropic so that the external conditions for quantum fluctuations are everywhere the same.

If the theoretical properties of the universe are those of an ensemble, and we can only observe one universe from that ensemble, how can we compare theory and observation? It seems reasonable that the statistics we get by comparing different parts of the universe should be similar to the statistics of a given part of the universe over different realizations, i.e., that they provide a *fair sample* of the probability distribution. This is called *ergodicity*. Fields $f(\mathbf{x})$ that satisfy

$$\bar{f} = \langle f\rangle \tag{6}$$

for an infinite volume V (for $\bar{f}$) and an arbitrary location $\mathbf{x}$ (for $\langle f\rangle$) are called *ergodic*. We assume that cosmological perturbations are ergodic. The equality does not hold for a finite volume V; the difference is called *sample variance* or *cosmic variance*. The larger the volume, the smaller is the difference. Since cosmological theory predicts $\langle f\rangle$, whereas observations probe $\bar{f}$ for a limited volume, cosmic variance limits how accurately we can compare theory with observations.[1]

8.1.2 Density autocorrelation function

From ergodicity,

$$\langle\rho\rangle = \bar{\rho} \quad \Rightarrow \quad \langle\delta\rho\rangle = 0 \quad \text{and} \quad \langle\delta\rangle = 0\,. \tag{7}$$

Thus we cannot use $\langle\delta\rangle$ as a measure of the inhomogeneity. Instead we can use the square of δ, which is necessarily nonnegative everywhere, so it cannot average out like δ did. Its expectation value

$$\langle\delta^2\rangle = \frac{\langle\delta\rho^2\rangle}{\bar{\rho}^2} \tag{8}$$

is the *variance* of the density perturbation, and the square root of the variance,

$$\delta_{\rm rms} \equiv \sqrt{\langle\delta^2\rangle} \tag{9}$$

the *root-mean-square* (rms) density perturbation, is a typical expected absolute value of δ at an arbitrary location.[2] It tells us about how strong the inhomogeneity is, but nothing about the shapes or sizes of the inhomogeneities. To get more information, we introduce the correlation function ξ.

We define the *density 2-point autocorrelation function* (often called just *correlation function*) as

$$\xi(\mathbf{x}_1,\mathbf{x}_2) \equiv \langle\delta(\mathbf{x}_1)\delta(\mathbf{x}_2)\rangle\,. \tag{10}$$

[1] Another notation I will use for volume average is $\widehat{f}$, for smaller volumes, e.g., the volume observed in a galaxy survey. I try to reserve $\bar{f}$ for situations where we can assume $\bar{f} = \langle f\rangle$, whereas cosmic variance is the difference between $\widehat{f}$ and $\langle f\rangle$.

[2] In other words, $\delta_{\rm rms}$ is the standard deviation of $\rho/\bar{\rho}$.

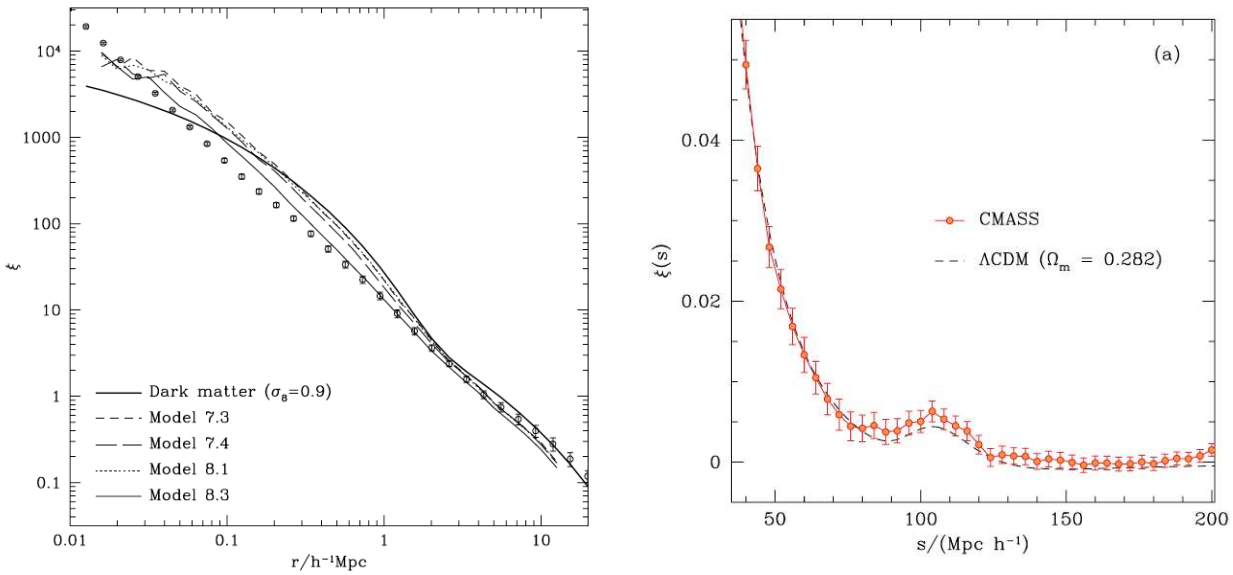


Figure 1: The 2-point correlation function $\xi(r)$ from galaxy surveys. Left: Small scales shown in a log-log plot. The circles with error bars show the observational determination from the APM galaxy survey [1]. The different lines are theoretical predictions by [2] (this is Fig. 9 from [2]). Right: Large scales shown in a linear plot. Red circles with error bars show the observational determination from the CMASS Data Release 9 (DR9) sample of the Baryonic Oscillation Spectroscopic Survey (BOSS). The dashed line is a theoretical prediction from the ΛCDM model. The bump near $100\,h^{-1}$Mpc is the baryon acoustic oscillation (BAO) peak that will be discussed later. This is Fig. 2a from [3].

It is positive if the density perturbation is expected to have the same sign at both $\mathbf{x}_1$ and $\mathbf{x}_2$, and negative for an overdensity at one and underdensity at the other. Thus it probes how density perturbations at different locations are correlated with each other. Due to statistical homogeneity, $\xi(\mathbf{x}_1, \mathbf{x}_2)$ can only depend on the *separation* $\mathbf{r} \equiv \mathbf{x}_2 - \mathbf{x}_1$, so we redefine ξ as

$$\xi(\mathbf{r}) \equiv \langle \delta(\mathbf{x})\delta(\mathbf{x}+\mathbf{r}) \rangle \,. \tag{11}$$

From statistical isotropy, $\xi(\mathbf{r})$ is independent of direction, i.e., spherically symmetric (isotropic),

$$\xi(\mathbf{r}) = \xi(r) \,. \tag{12}$$

We will have use for both the 3D, $\xi(\mathbf{r})$, and 1D, $\xi(r)$, versions. The correlation function is large and positive for r smaller than the size of a typical over- or underdense region, and becomes small for larger separations.

The correlation function at zero separation gives the variance of the density perturbation,

$$\langle \delta^2 \rangle \equiv \langle \delta(\mathbf{x})\delta(\mathbf{x}) \rangle \equiv \xi(0) \,. \tag{13}$$

We can also define a correlation function $\widehat{\xi}(\mathbf{r})$ for a single realization as a volume average,

$$\widehat{\xi}(\mathbf{r}) \equiv \frac{1}{V} \int d^3x\, \delta(\mathbf{x})\delta(\mathbf{x}+\mathbf{r}) \,. \tag{14}$$

Integrating over $\mathbf{r}$ and assuming periodic boundary conditions[3] we get the *integral constraint*

$$\int d^3r\, \widehat{\xi}(\mathbf{r}) \;=\; \frac{1}{V} \int d^3r d^3x\, \delta(\mathbf{x})\delta(\mathbf{x}+\mathbf{r}) \;=\; \frac{1}{V} \int d^3x\, \delta(\mathbf{x}) \int d^3r\, \delta(\mathbf{x}+\mathbf{r}) = 0 \,, \tag{16}$$

[3]The other option is not to use periodic boundary conditions, but to understand the integral in (14) to go over only those $\mathbf{x}$, for which both $\mathbf{x}$ and $\mathbf{x}+\mathbf{r} \in V$. This is what we have to do when V refers to an actual survey.

since the latter integral is $\bar{\delta} = 0$. Since $\xi(\mathbf{r}) = \langle\widehat{\xi}(\mathbf{r})\rangle$ the integral constraint applies to it likewise. Therefore $\xi(r)$ must become negative at some point, so that at such a distance from an overdense region we are more likely to find an underdense region. Going to ever larger separations, ξ as a function of r may oscillate around zero, the oscillation becoming ever smaller in amplitude. Most of the interest in $\xi(r)$ is for the small r within the initial positive region.

8.1.3 Fourier space

The evolution of perturbations is best discussed in Fourier space. Fourier analysis is a method for separating out different distance scales, so that the dependence of the physics on distance scale becomes clear and easy to handle.

For mathematical convenience, we assume the observable part of the universe lies within a fiducial cubic box, volume $V = L^3$, with periodic boundary conditions. This box may be assumed to be much larger than the region of interest, so that these boundary conditions should have no effect. Since the infinite universe is now assumed periodic, the volume average over the infinite universe will be equal to the volume average over the fiducial box. Thus also the ergodicity assumption requires the fiducial volume to be large, so that it can provide a fair sample of the ensemble.

We can now expand any function of space $f(\mathbf{x})$ as a Fourier series

$$f(\mathbf{x}) = \sum_{\mathbf{k}} f_{\mathbf{k}} e^{i\mathbf{k}\cdot\mathbf{x}}\,, \tag{17}$$

where the wave vectors $\mathbf{k} = (k_1, k_2, k_3)$ take values

$$k_i = n_i \frac{2\pi}{L}\,, \qquad n_i = 0, \pm 1, \pm 2, \ldots \tag{18}$$

The Fourier coefficients $f_{\mathbf{k}}$ are obtained as

$$f_{\mathbf{k}} = \frac{1}{V}\int_V f(\mathbf{x}) e^{-i\mathbf{k}\cdot\mathbf{x}} d^3x\,. \tag{19}$$

If $f(\mathbf{x})$ is a perturbation so that its mean value vanishes, then the term $\mathbf{k} = 0$ does not occur.

The Fourier coefficients are complex numbers even though we are dealing with real quantities $f(\mathbf{x})$. From the reality of $f(\mathbf{x})$ follows that

$$f_{-\mathbf{k}} = f_{\mathbf{k}}^*\,. \tag{20}$$

This means that for each pair of terms $f_{\mathbf{k}} e^{i\mathbf{k}\cdot\mathbf{x}} + f_{-\mathbf{k}} e^{-i\mathbf{k}\cdot\mathbf{x}}$ the imaginary parts cancel. The real part of $f_{\mathbf{k}} e^{i\mathbf{k}\cdot\mathbf{x}}$ is

$$\Re\left(f_{\mathbf{k}} e^{i\mathbf{k}\cdot\mathbf{x}}\right) \;=\; \Re f_{\mathbf{k}} \cos\mathbf{k}\cdot\mathbf{x} - \Im f_{\mathbf{k}} \sin\mathbf{k}\cdot\mathbf{x}\,. \tag{21}$$

To visualize a Fourier component $f_{\mathbf{k}} e^{i\mathbf{k}\cdot\mathbf{x}}$ one may thus visualize just this real part, which is a sinusoidal plane wave in the $\mathbf{k}$ direction.

The double integral in (16) then goes over all pairs $(\mathbf{x}_1, \mathbf{x}_2)$ in V and can be written as

$$\frac{1}{V}\int_V d^3x_1\,\delta(\mathbf{x}_1)\int_V d^3x_2\,\delta(\mathbf{x}_2) = 0\cdot 0\,. \tag{15}$$

For an actual survey, V is not really large enough to make cosmic variance negligible. This has two effects: First, the integral does not extend to large enough separations r to capture the full $\xi(r)$, so that typically negative $\xi(r)$ values at large separations r are missed. This would tend to make the integral positive. However, for an actual survey, also the mean density has to be estimated from the survey, so that δ will actually refer to the deviation from the mean density of the survey, which again forces $\int_V d^3x\,\delta(x) = 0$. Thus this forced integral constraint makes the survey typically underestimate the true correlation function.

The Fourier expansion works only if the background universe is flat, although it can be used as an approximation in open and closed universes,[4] if the region of interest is much smaller than the curvature radius.

The separation of neighboring k_i values is $\Delta k_i = 2\pi/L$, so we can write

$$f(\mathbf{x}) \;=\; \sum_{\mathbf{k}} f_{\mathbf{k}} e^{i\mathbf{k}\cdot\mathbf{x}} \left(\frac{L}{2\pi}\right)^3 \Delta k_1 \Delta k_2 \Delta k_3 \;\approx\; \frac{1}{(2\pi)^3}\int f(\mathbf{k}) e^{i\mathbf{k}\cdot\mathbf{x}} d^3k\,, \tag{22}$$

where

$$f(\mathbf{k}) \equiv L^3 f_{\mathbf{k}}\,. \tag{23}$$

replacing the Fourier series with the Fourier integral. The size of the Fourier coefficients depends on the fiducial volume V – increasing V tends to make the $f_{\mathbf{k}}$ smaller to compensate for the denser sampling of $\mathbf{k}$ in Fourier space.

In the limit $V \to \infty$, the approximation in (22) becomes exact, and we have the *Fourier transform* pair

$$\begin{aligned} f(\mathbf{x}) &= \frac{1}{(2\pi)^3}\int f(\mathbf{k}) e^{i\mathbf{k}\cdot\mathbf{x}} d^3k \\ f(\mathbf{k}) &= \int f(\mathbf{x}) e^{-i\mathbf{k}\cdot\mathbf{x}} d^3x\,. \end{aligned} \tag{24}$$

Note that this assumes that the integrals converge, which requires that $f(\mathbf{x}) \to 0$ for $|\mathbf{x}| \to \infty$. Thus we use only the Fourier series for, e.g., $\delta(\mathbf{x})$, but for, e.g., the correlation function $\xi(\mathbf{x})$ the Fourier transform is appropriate.

Even with a finite V we can use the Fourier integral as an approximation. Often it is conceptually simpler to work first with the Fourier series (so that one can, e.g., use the Kronecker delta $\delta_{\mathbf{k}\mathbf{k}'}$ instead of the Dirac delta function $\delta_D(\mathbf{k}-\mathbf{k}')$), replacing it with the integral in the end, when it needs to be calculated. The recipe for going from the series to the integral is

$$\begin{aligned} \left(\frac{2\pi}{L}\right)^3 \sum_{\mathbf{k}} &\;\to\; \int d^3k \\ L^3 f_{\mathbf{k}} &\;\to\; f(\mathbf{k}) \\ \left(\frac{L}{2\pi}\right)^3 \delta_{\mathbf{k}\mathbf{k}'} &\;\to\; \delta_D^3(\mathbf{k}-\mathbf{k}')\,. \end{aligned} \tag{25}$$

so that, e.g.,

$$\sum_{\mathbf{k}} f_{\mathbf{k}}\, e^{i\mathbf{k}\cdot\mathbf{x}} \;\to\; \frac{1}{(2\pi)^3}\int f(\mathbf{k}) e^{i\mathbf{k}\cdot\mathbf{x}} d^3k\,. \tag{26}$$

8.1.4 Power spectrum

We now expand the density perturbation as a Fourier series

$$\delta(\mathbf{x}) = \sum_{\mathbf{k}} \delta_{\mathbf{k}} e^{i\mathbf{k}\cdot\mathbf{x}}\,, \tag{27}$$

with

$$\delta_{\mathbf{k}} = \frac{1}{V}\int_V \delta(\mathbf{x}) e^{-i\mathbf{k}\cdot\mathbf{x}} d^3x \tag{28}$$

[4] An exact treatment in open and closed universes requires expansion in terms of suitable other functions instead of the plane waves $e^{i\mathbf{k}\cdot\mathbf{x}}$.

and $\delta_{-\mathbf{k}} = \delta^*_{\mathbf{k}}$. Note that

$$\langle \delta(\mathbf{x}) \rangle = 0 \quad \Rightarrow \quad \langle \delta_{\mathbf{k}} \rangle = 0\,. \tag{29}$$

In analogy with the correlation function $\xi(\mathbf{x}, \mathbf{x}')$, we may ask what is the corresponding correlation in Fourier space, $\langle \delta^*_{\mathbf{k}} \delta_{\mathbf{k}'} \rangle$. Note that due to the mathematics of complex numbers, correlations of Fourier coefficients are defined with the complex conjugate *. This way the correlation of $\delta_{\mathbf{k}}$ with itself, $\langle \delta^*_{\mathbf{k}} \delta_{\mathbf{k}} \rangle = \langle |\delta_{\mathbf{k}}|^2 \rangle$ is a real (and nonnegative) quantity, the expectation value of the absolute value (modulus) of $\delta_{\mathbf{k}}$ squared, i.e., the variance of $\delta_{\mathbf{k}}$. Calculating

$$\begin{aligned}
\langle \delta^*_{\mathbf{k}} \delta_{\mathbf{k}'} \rangle &= \frac{1}{V^2} \int d^3x e^{i\mathbf{k}\cdot\mathbf{x}} \int d^3x' e^{-i\mathbf{k}'\cdot\mathbf{x}'} \langle \delta(\mathbf{x}) \delta(\mathbf{x}') \rangle \\
&= \frac{1}{V^2} \int d^3x e^{i\mathbf{k}\cdot\mathbf{x}} \int d^3r e^{-i\mathbf{k}'\cdot(\mathbf{x}+\mathbf{r})} \langle \delta(\mathbf{x}) \delta(\mathbf{x}+\mathbf{r}) \rangle \\
&= \frac{1}{V^2} \int d^3r e^{-i\mathbf{k}'\cdot\mathbf{r}} \xi(\mathbf{r}) \int d^3x e^{i(\mathbf{k}-\mathbf{k}')\cdot\mathbf{x}} \\
&= \frac{1}{V} \delta_{\mathbf{k}\mathbf{k}'} \int d^3r e^{-i\mathbf{k}\cdot\mathbf{r}} \xi(\mathbf{r}) \;\equiv\; \frac{1}{V} \delta_{\mathbf{k}\mathbf{k}'} P(\mathbf{k})\,,
\end{aligned} \tag{30}$$

where we used $\langle \delta(\mathbf{x}) \delta(\mathbf{x}+\mathbf{r}) \rangle = \xi(\mathbf{r})$, i.e., independent of $\mathbf{x}$, which results *from statistical homogeneity*, and the orthogonality of plane waves

$$\boxed{\int d^3x e^{i(\mathbf{k}-\mathbf{k}')\cdot\mathbf{x}} \;=\; V \delta_{\mathbf{k}\mathbf{k}'} \;\to\; (2\pi)^3 \delta_D^3(\mathbf{k}-\mathbf{k}')\,.} \tag{31}$$

Note that here $\delta_{\mathbf{k}\mathbf{k}'}$ is the Kronecker delta, 1 for $\mathbf{k} = \mathbf{k}'$, 0 otherwise – nothing to do with the density perturbation! In the limit $V \to \infty$ we get the Dirac delta function $\delta_D^3(\mathbf{k}-\mathbf{k}')$.

Written in terms of $\delta(\mathbf{k}) = V\delta_{\mathbf{k}}$, the result (30) reads as

$$\langle \delta(\mathbf{k})^* \delta(\mathbf{k}') \rangle \;=\; V \delta_{\mathbf{k}\mathbf{k}'} P(\mathbf{k}) \;\to\; (2\pi)^3 \delta_D^3(\mathbf{k}-\mathbf{k}') P(\mathbf{k})\,, \tag{32}$$

Thus, *from statistical homogeneity follows that the Fourier coefficients* $\delta_{\mathbf{k}}$ *are uncorrelated*. The quantity

$$P(\mathbf{k}) \;\equiv\; V \langle |\delta_{\mathbf{k}}|^2 \rangle \;=\; \int d^3r\, e^{-i\mathbf{k}\cdot\mathbf{r}} \xi(\mathbf{r})\,, \tag{33}$$

which gives the *variance* of $\delta_{\mathbf{k}}$, is called the *power spectrum* of $\delta(\mathbf{x})$. Since the correlation function $\to 0$ for large separations, we can replace the integration volume V in (33) with an infinite volume.[5] We see that the power spectrum is the 3D Fourier transform of $\xi(\mathbf{r})$, and therefore also

$$\xi(\mathbf{r}) \;=\; \frac{1}{(2\pi)^3} \int d^3k\, e^{i\mathbf{k}\cdot\mathbf{r}} P(\mathbf{k})\,. \tag{34}$$

Unlike the correlation function, the power spectrum $P(\mathbf{k})$ is positive everywhere. Perturbations at large distance scales are more commonly discussed in terms of $P(\mathbf{k})$ than $\xi(\mathbf{r})$.

From statistical isotropy

$$\xi(\mathbf{r}) = \xi(r) \quad \Rightarrow \quad P(\mathbf{k}) = P(k) \tag{35}$$

(the 3D Fourier transform of a spherically symmetric function is also spherically symmetric), so that the variance of $\delta_{\mathbf{k}}$ depends only on the magnitude k of the wave vector $\mathbf{k}$, i.e., on the corresponding distance scale. Using spherical coordinates and doing the angular integrals

[5] We want to avoid discussing $\xi(r)$ for $r \geq L$, since the artificially assumed periodicity would cause artifacts in the behavior of $\xi(\mathbf{r})$ at such separations. Thus L is assumed so large that ξ is completely negligible at such huge separations.

we obtain (**exercise**) the relation between the 1D correlation function $\xi(r)$ and the 1D power spectrum $P(k)$,

$$\begin{aligned} P(k) &= \int_0^\infty \xi(r)\frac{\sin kr}{kr}4\pi r^2 dr \\ \xi(r) &= \frac{1}{(2\pi)^3}\int_0^\infty P(k)\frac{\sin kr}{kr}4\pi k^2 dk\,, \end{aligned} \tag{36}$$

For the density variance we get

$$\langle\delta^2\rangle \equiv \xi(0) = \frac{1}{(2\pi)^3}\int_0^\infty P(k)4\pi k^2 dk = \frac{1}{2\pi^2}\int_{-\infty}^\infty k^3 P(k) d\ln k \equiv \int_{-\infty}^\infty \mathcal{P}(k) d\ln k\,. \tag{37}$$

where we have defined

$$\boxed{\mathcal{P}(k) \equiv \frac{k^3}{2\pi^2}P(k)\,.} \tag{38}$$

Another common notation for $\mathcal{P}(k)$ is $\Delta^2(k)$. The word "power spectrum" is used to refer to both $P(k)$ and $\mathcal{P}(k)$. Of these two, $\mathcal{P}(k)$ has the more obvious physical meaning: it gives the contribution of a logarithmic interval of scales, i.e., from k to ek, to the density variance. $\mathcal{P}(k)$ is dimensionless, whereas $P(k)$ has the dimension of Mpc^3 (when discussing observed values, it is usually given in units of $h^{-3}\mathrm{Mpc}^3$ as distance determinations are proportional to the Hubble constant).

Exercise: Define $\widehat{\xi}(\mathbf{r})$ as the volume average of $\delta(\mathbf{x})\delta(\mathbf{x}+\mathbf{r})$, i.e., integrate $\mathbf{x}$ over the box V with periodic boundary conditions, and show that

$$\widehat{\xi}(\mathbf{r}) = \frac{V}{(2\pi)^3}\int d^3k\,|\delta_\mathbf{k}|^2 e^{i\mathbf{k}\cdot\mathbf{r}}\,, \tag{39}$$

for a single realization. Note that here we do not need any statistical assumptions (like statistical homogeneity or ergodicity). Contrast this result with (34).

8.1.5 Scales of interest and window functions

In (37) we integrated over all scales, from the infinitely large ($k = 0$ and $\ln k = -\infty$) to the infinitely small ($k = \infty$ and $\ln k = \infty$) to get the density variance.[6] Perhaps this is not really what we want. The average matter density today is $3\times10^{-27}\,\mathrm{kg/m}^3$. The density of the Earth is $5.5\times10^3\,\mathrm{kg/m}^3$ and that of an atomic nucleus $2\times10^{17}\,\mathrm{kg/m}^3$, corresponding to $\delta\approx 2\times10^{30}$ and $\delta\approx10^{44}$. Probing the density of the universe at such small scales finds a huge variance in it, but this is no longer the topic of cosmology – we are not interested here in planetary science or nuclear physics.

Even the study of the structure of individual galaxies is not considered to belong to cosmology, so the smallest (comoving) scale of cosmological interest, at least when we discuss the present universe,[7] is that of a typical separation between neighboring galaxies, of the order of 1 Mpc.

To exclude scales smaller than R ($r < R$ or $k > R^{-1}$) we *filter* the density field with a *window function*. This can be done in **k**-space or **x**-space.

---

[6] Note that large scales correspond to small k and vice versa. To avoid confusion, it is better to use the words *low* and *high* for k, so that *large scales correspond to low k, and small scales correspond to high k*.

[7] In early universe cosmology we may study events, or possible events, related to also smaller comoving scales.

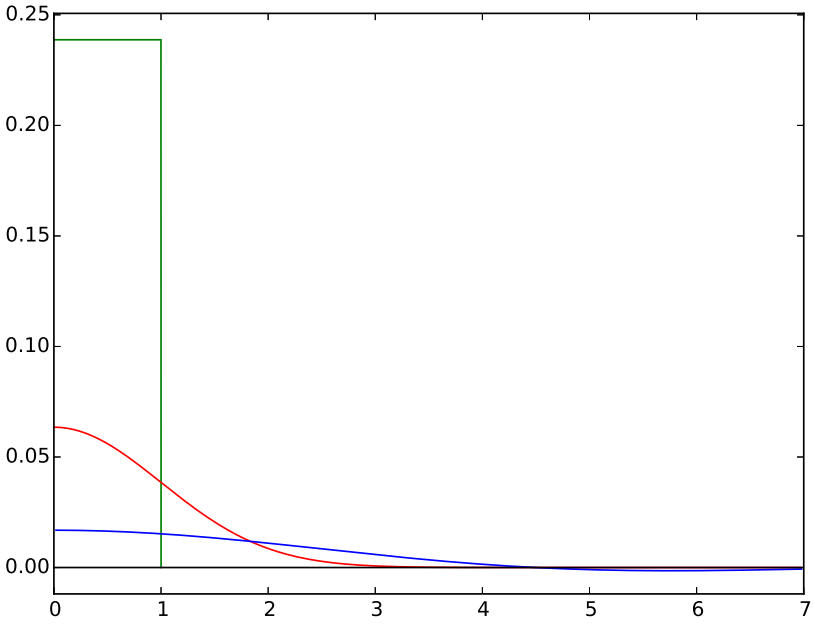


Figure 2: The 3D window functions $W(r)$, top-hat (green), Gaussian (red), and k (blue), for $R = 1$.

The filtering in $\mathbf{x}$-space is done by convolution. We introduce a (usually spherically symmetric) window function $W(\mathbf{r})$ such that $W(\mathbf{r})$ is relatively large for $|\mathbf{r}| \leq R$ and $W \sim 0$ for $|\mathbf{r}| \gg R$. We use normalization

$$\int d^3r\, W(\mathbf{r}) = 1 \tag{40}$$

and define the filtered density field

$$\delta(\mathbf{x}, R) \;\equiv\; (\delta * W)(\mathbf{x}) \;\equiv\; \int d^3\mathbf{x}'\, \delta(\mathbf{x}') W(\mathbf{x}' - \mathbf{x})\,. \tag{41}$$

The simplest window function is the top-hat window function

$$W_T(\mathbf{r}) \;\equiv\; \left(\frac{4\pi}{3}R^3\right)^{-1} \qquad \text{for} \quad |\mathbf{r}| \leq R \tag{42}$$

and $W_T(\mathbf{r}) = 0$ elsewhere, i.e., $\delta(\mathbf{x})$ is filtered by replacing it with its mean value within the distance R. Mathematically more convenient is the Gaussian window function

$$W_G(\mathbf{r}) \;\equiv\; \frac{1}{(2\pi)^{3/2}R^3} e^{-\frac{1}{2}|\mathbf{r}|^2/R^2}\,. \tag{43}$$

By the convolution theorem, the filtering in Fourier space becomes just multiplication:

$$\delta(\mathbf{k}, R) \;=\; \delta(\mathbf{k}) W(\mathbf{k})\,, \tag{44}$$

where $W(\mathbf{k})$ is the Fourier transform of the window function. For W_T and W_G we have (**exercise**)

$$\begin{aligned} W_T(\mathbf{k}) &= \frac{3(\sin kR - kR\cos kR)}{(kR)^3} \\ W_G(\mathbf{k}) &= e^{-\frac{1}{2}(kR)^2}\,. \end{aligned} \tag{45}$$

We can also define the **k**-space top-hat window function

$$W_k(\mathbf{k}) \;\equiv\; 1 \qquad \text{for} \quad k \le 1/R \tag{46}$$

and $W_k(\mathbf{k}) = 0$ elsewhere. In **x**-space this becomes (**exercise**)

$$W_k(\mathbf{r}) \;=\; \frac{1}{2\pi^2 R^3}\frac{\sin y - y\cos y}{y^3}\,, \qquad \text{where} \quad y \equiv |\mathbf{r}|/R\,. \tag{47}$$

The variance of the filtered density field (**Exercise:** derive the second equalities of both expressions)

$$\begin{aligned}\sigma^2(R) \;&\equiv\; \langle\delta(\mathbf{x},R)^2\rangle \;=\; \frac{1}{(2\pi)^3}\int d^3k\, P(k)|W(\mathbf{k})|^2 \\ \widehat{\sigma}^2(R) \;&\equiv\; \frac{1}{V}\int d^3x\,\delta(\mathbf{x},R)^2 \;=\; \frac{V}{(2\pi)^3}\int d^3k\,|\delta_{\mathbf{k}}|^2|W(\mathbf{k})|^2\,.\end{aligned} \tag{48}$$

is a measure of the inhomogeneity at scale R. For the **k**-space top-hat window this becomes simply

$$\sigma^2(R) \;=\; \frac{1}{(2\pi)^3}\int_0^{R^{-1}} 4\pi k^2 P(k)dk \;=\; \int_{-\infty}^{-\ln R}\mathcal{P}(k)d\ln k\,. \tag{49}$$

One may also ask, whether scales larger than the observed universe (the lower limit $k = 0$ or $\ln k = -\infty$ in the k integrals) are relevant, since we cannot observe the inhomogeneity at such scales. Due to such very-large-scale inhomogeneities, the average density in the observed universe may deviate from the average density of the entire universe. Inhomogeneities at scales somewhat larger than the observed universe could appear as an anisotropy in the observed universe. The importance of such large scales depends on how strong the inhomogeneities at these scales are, i.e., how the power spectrum behaves as $k \to 0$. The present understanding, supported by observations, is that the contribution of such large scales is small.[8]

8.1.6 Power-law spectra

We have observational information and theoretical predictions for $\xi(r)$ and $P(k)$ for a wide range of scales. (We will discuss the theory in detail later.) For certain intervals, they can be approximated by a power-law form,

$$\xi(r) \propto r^{-\gamma} \qquad \text{or} \qquad P(k) \propto k^n\,. \tag{50}$$

When plotted on a log-log scale, such functions appear as straight lines with slope $-\gamma$ and n. The proportionality constant can be given in terms of a reference scale. For $\xi(r)$ we usually choose the scale r_0 where $\xi(r_0) = 1$, so that

$$\xi(r) \;=\; \left(\frac{r}{r_0}\right)^{-\gamma}\,. \tag{51}$$

For $P(k)$ we may write

$$P(k) \;=\; A^2\left(\frac{k}{k_p}\right)^n \qquad \text{or} \qquad \mathcal{P}(k) \;=\; A^2\left(\frac{k}{k_p}\right)^{n+3}\,, \tag{52}$$

[8]The inflation scenario predicts that the universe outside the current horizon is similar to the observable universe up to distances very much larger than the current horizon distance. However at "very very far away", beyond the pre-inflation horizon, the universe may be quite different. For this section, we exclude such "very very far" regions from the concept of "universe". (They are "other universes" in a "multiverse".)

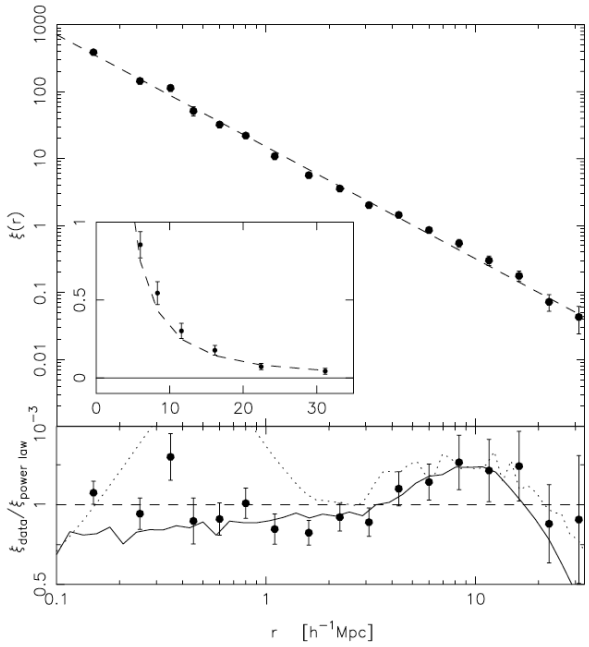


Figure 3: Top panel: The correlation function from the 2dFGRS galaxy survey in log-log scale. The dashed line is the best-fit power law ($r_0 = 5.05\,h^{-1}$Mpc, $\gamma = 1.67$). The inset shows the same in linear scale. Bottom panel: 2dFGRS data (solid circles with error bars) divided by the power-law fit. The solid line is the result from the APM survey and the dotted line from an N-body simulation. This is Fig. 11 from [4].

where k_p is called a *pivot scale* (whose choice depends on the application) and $A \equiv \sqrt{P(k_p)}$ or $\sqrt{\mathcal{P}(k_p)}$ is the amplitude of the power spectrum at the pivot scale.

We define the spectral index $n(k)$ as

$$n(k) \;\equiv\; \frac{d\ln P}{d\ln k}\,. \tag{53}$$

It gives the slope of $P(k)$ on a log-log plot. For a power-law $P(k)$, $n(k) = const = n$. We can study power-law $\xi(r)$ and $P(k)$ as a playground to get a feeling what different values of the spectral index mean, and, e.g., how γ and n are related.[9]

The Fourier transform of a power law is a power law. For the correlation function of (51) we get (**exercise**)

$$\begin{aligned} P(k) &= \frac{4\pi}{k^3}\Gamma(2-\gamma)\sin\frac{(2-\gamma)\pi}{2}\,(kr_0)^\gamma \\ \mathcal{P}(k) &= \frac{2}{\pi}\Gamma(2-\gamma)\sin\frac{(2-\gamma)\pi}{2}\,(kr_0)^\gamma \end{aligned} \tag{54}$$

for $1 < \gamma < 2$ or $2 < \gamma < 3$, and

$$\begin{aligned} P(k) &= \frac{2\pi^2}{k^3}(kr_0)^2 \\ \mathcal{P}(k) &= (kr_0)^2 \end{aligned} \tag{55}$$

for $\gamma = 2$. Thus

$$n \;=\; \gamma - 3 \qquad \text{for } 1 < \gamma < 3, \text{ i.e., } -2 < n < 0\,. \tag{56}$$

[9] In reality the spectral index is very different at small scales than at large scales. Observationally, for small scales, $\gamma \sim 1.8$, and for large scales, $n \sim 1$. We discuss this later.

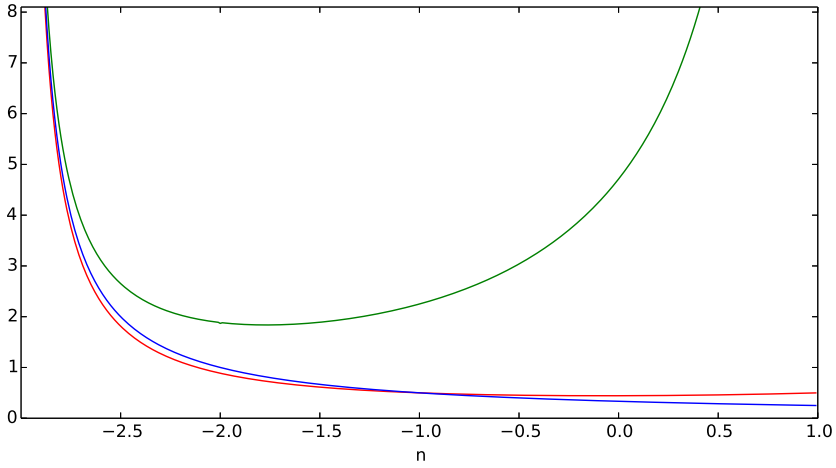


Figure 4: The ratio of $\sigma^2(R)$ to $\mathcal{P}(R^{-1})$ in the case of a power-law spectrum $\mathcal{P}(k) \propto k^{n+3}$ for the three different window functions: top-hat (green), Gaussian (red), and k (blue). They all diverge in the limit $n \to -3$ due to the contributions of ever larger scales ($\ln k \to -\infty$).

The variance

$$\langle \delta^2 \rangle = \xi(0) = \int_0^\infty \mathcal{P}(k) \frac{dk}{k} \propto \int_0^\infty k^{n+2} dk = \frac{1}{n+3} \left[k^{n+3} \right]_0^\infty \qquad \text{for} \quad n \neq -3 \tag{57}$$

diverges at small scales (high k) for $n \geq -3$ and at large scales (low k) for $n \leq -3$. In practice we encounter only the small-scale divergence.

We cure the small-scale divergence with filtering as discussed in Sec. 8.1.5, replacing (57) with (see Eq. 48)

$$\sigma^2(R) \;\equiv\; \langle \delta(\mathbf{x}, R)^2 \rangle \;=\; \int_0^\infty \mathcal{P}(k) |W(k)|^2 \frac{dk}{k} \,. \tag{58}$$

For the three window functions given in Sec. 8.1.5, power-law spectra give

$$\begin{aligned}
\sigma_T^2(R) &= \frac{9}{2^n}(n+1) \sin\frac{n\pi}{2} \, \frac{\Gamma(n-1)}{n-3} \mathcal{P}(R^{-1}) \qquad \text{for } -3 < n < 1 \\
\sigma_G^2(R) &= \frac{1}{2}\Gamma\left(\frac{n+3}{2}\right) \mathcal{P}(R^{-1}) \qquad \text{for } n > -3 \\
\sigma_k^2(R) &= \frac{1}{n+3} \mathcal{P}(R^{-1}) \qquad \text{for } n > -3.
\end{aligned} \tag{59}$$

For $n \geq 1$, the top-hat window is not able to cure the small-scale divergence, since its Fourier transform does not die out fast enough at high k (this is related to the sharp boundary of the window). For integers $-3 < n < 1$ the formula for $\sigma_T^2(R)$ in (59) is not defined, since either $n+1$ or $\sin n\pi/2$ gives 0 and $\Gamma(n-1)$ gives infinity. For these cases

$$\sigma_T^2(R) \;=\; \frac{3\pi}{5}\mathcal{P}(R^{-1}) \,, \quad \frac{9}{4}\mathcal{P}(R^{-1}) \,, \quad \frac{3\pi}{2}\mathcal{P}(R^{-1}) \qquad \text{for } n = -2, \; -1, \; 0 \,. \tag{60}$$

For $n = 1$, $\sigma_G^2(R) = \frac{1}{2}\mathcal{P}(R^{-1})$.

Exercise: Derive these results for $\sigma_G^2(R)$ and $\sigma_k^2(R)$. ($\sigma_T^2(R)$ is more difficult.)

8.1.7 Galaxy 2-point correlation function

The most obvious way to try to measure the cosmological density perturbations is to observe the spatial distribution of galaxies. We treat individual galaxies as mathematical points, so

that each galaxy has a comoving coordinate value $\mathbf{x}$. We define the *galaxy 2-point correlation function* $\xi_g(\mathbf{r})$ as the *excess probability* of finding a galaxy at separation $\mathbf{r}$ from another galaxy:

$$dP \equiv \bar{n}\,[1 + \xi_g(\mathbf{r})]\,dV \tag{61}$$

where $\bar{n}$ is the mean galaxy number density, dV is a volume element that is a separation $\mathbf{r}$ away from a chosen reference galaxy, and dP is the probability that there is a galaxy within dV. (Here dV is assumed so small that there is at most one galaxy in it.)

If the galaxy number density $n(\mathbf{x})$ faithfully traces the underlying matter density, so that

$$\delta_g \equiv \frac{\delta n}{\bar{n}} = \delta \equiv \frac{\delta\rho_m}{\bar{\rho}_m}\,, \tag{62}$$

then ξ_g becomes equal to the matter density autocorrelation function ξ: The probability of finding a galaxy in volume dV_1 at a random location $\mathbf{x}$ is

$$dP_1 \;=\; \langle n(\mathbf{x})\rangle dV_1 \;=\; \langle \bar{n} + \delta n(\mathbf{x})\rangle dV_1 \;=\; \bar{n} dV_1\,. \tag{63}$$

The probability of finding a galaxy pair at $\mathbf{x}$ and $\mathbf{x}+\mathbf{r}$ is

$$\begin{aligned} dP_{12} &= \langle n(\mathbf{x})n(\mathbf{x}+\mathbf{r})\rangle dV_1 dV_2 \;=\; \bar{n}^2\langle[1+\delta(\mathbf{x})][1+\delta(\mathbf{x}+\mathbf{r})]\rangle dV_1 dV_2 \\ &= \bar{n}^2\left[1 + \langle\delta(\mathbf{x})\rangle + \langle\delta(\mathbf{x}+\mathbf{r})\rangle + \langle\delta(\mathbf{x})\delta(\mathbf{x}+\mathbf{r})\rangle\right] dV_1 dV_2 \\ &= \bar{n}^2\left[1 + \langle\delta(\mathbf{x})\delta(\mathbf{x}+\mathbf{r})\rangle\right] dV_1 dV_2\,, \end{aligned} \tag{64}$$

since $\langle\delta(\mathbf{x})\rangle = \langle\delta(\mathbf{x}+\mathbf{r})\rangle = 0$. Dividing dP_{12} with dP_1 we get the probability dP_2 of finding the second galaxy once we have found the first one

$$dP_2 \;=\; \bar{n}\left[1 + \langle\delta(\mathbf{x})\delta(\mathbf{x}+\mathbf{r})\rangle\right] dV_2 \;=\; \bar{n}\,[1 + \xi(\mathbf{r})]\,dV_2\,. \tag{65}$$

Thus $\xi_g = \xi$.

It is probable that the galaxy number density does not trace the matter density faithfully, since galaxy formation is likely to be more efficient in high-density regions. This is called *bias*. Specifically the bias, or *galaxy bias* b_g, is defined as the ratio

$$b_g \equiv \frac{\delta_g}{\delta_m} \quad\Rightarrow\quad \xi_g = b_g^2\xi\,, \tag{66}$$

where the expectation is that $b_g > 1$. In principle the bias could depend on the scale k, the time t (or redshift z), and/or the strength of the density perturbation δ_m. The simplest treatment of bias is to assume b_g is a constant over the observationally relevant ranges of these quantities.

The bias will depend on the type of tracer (all galaxies, specific types of galaxies, galaxy clusters) and is typically larger for more massive objects.

For the galaxy number density field, observationally $\sigma_T(R) \approx 1$ at $R = 8\,h^{-1}$Mpc. This has motivated the definition (will come later) of the quantity σ_8 as a cosmological parameter related to the amplitude of large-scale structure.

8.2 Newtonian perturbation theory

We shall now study the evolution of perturbations during the history of the universe. Initially the perturbations were small and we restrict the quantitative treatment to that part of the evolution when they remained small (for large scales, this extends to the present time and the future). This allows us to use *first-order perturbation theory*, where we drop from our equations all those terms which contain a product of two or more perturbations (as these products are even smaller). The remaining equations will then contain only terms which are either *zeroth order*, i.e., contain only background quantities, or *first order*, i.e., contain exactly one perturbation. If we kept only the zeroth order parts, we would be back to the equations of the homogenous and isotropic universe. Subtracting these from our equations we arrive at the *perturbation equations* where every term is first-order in the perturbation quantities, i.e., it is a *linear equation* for them. This makes the equations easy to handle, we can, e.g., Fourier transform them.

As we discovered in our discussion of inflation, the different cosmological distance scales first exit the horizon during inflation, then enter the horizon during various epochs of the later history. Matter perturbations at subhorizon scales, i.e., after horizon entry, can be treated with *Newtonian perturbation theory*, but scales which are close to horizon size or superhorizon require *relativistic perturbation theory*, which is based on general relativity.

The Newtonian equations for (perfect gas)[10] fluid dynamics with gravity are

$$\frac{\partial \rho}{\partial t'} + \nabla_{\mathbf{r}} \cdot (\rho \mathbf{u}) = 0 \tag{67}$$

$$\frac{\partial \mathbf{u}}{\partial t'} + (\mathbf{u} \cdot \nabla_{\mathbf{r}})\mathbf{u} + \frac{1}{\rho}\nabla_{\mathbf{r}} p + \nabla_{\mathbf{r}} \tilde{\Phi} = 0 \tag{68}$$

$$\nabla^2_{\mathbf{r}} \tilde{\Phi} = 4\pi G \rho \tag{69}$$

Here ρ is the mass density, p is the pressure, and $\mathbf{u}$ is the flow velocity of the fluid. We write $\tilde{\Phi}$ for the Newtonian gravitational potential, since we want to reserve Φ for its perturbation. The subscript $\mathbf{r}$ in $\nabla_{\mathbf{r}}$ emphasizes that the space derivatives are taken with respect to the Newtonian space coordinate $\mathbf{r}$ (instead of a comoving coordinate). Although the Newtonian time coordinate t' is equal to the cosmic time coordinate t, we need to make a distinction between t' and t in partial derivatives as will become clear soon.

The first equation is the law of mass conservation. The second equation is called the *Euler equation*, and it is just "$F = ma$" for a fluid element, whose mass is ρdV. Here the acceleration of a fluid element is not given by $\partial \mathbf{u}/\partial t'$ which just tells how the velocity field changes at a given position, but by $d\mathbf{u}/dt'$, where

$$\frac{d}{dt'} \equiv \frac{\partial}{\partial t'} + (\mathbf{u} \cdot \nabla_{\mathbf{r}}) \tag{70}$$

is the *convective time derivative*, which follows the fluid element as it moves. The two other terms give the forces due to pressure gradient and gravitational field.

We can apply Newtonian physics if:

1) Distance scales considered are $\ll$ the scale of curvature of spacetime (given by the Hubble length in cosmology[11])

2) The fluid flow is nonrelativistic, $u \ll c \equiv 1$.

3) We are considering nonrelativistic matter, $|p| \ll \rho$

[10] perfect gas = no internal friction $\Rightarrow$ pressure is isotropic

[11] As discussed in Chapter 3, the spacetime curvature has two distance scales, the Hubble length H^{-1} and the curvature radius $R_{\rm curv} \equiv a|K|^{-1/2}$. From observations we know that the curvature radius is larger than the Hubble length (at all times of interest), possibly infinite.

The last condition corresponds to particle velocities being nonrelativistic, if the matter is made out of particles. Although the pressure is small compared to mass density, the pressure gradient can be important if the pressure varies at small scales.

Note: Energy density and mass density. In Newtonian gravity, the source of gravity is mass density ρ_m, not energy density ρ. For nonrelativistic matter, the kinetic energies of particles are negligible compared to their masses, and thus so is the energy density compared to mass density, if we don't count the rest energy in it. The Newtonian equations for mass density and energy density are

$$\frac{\partial \rho_m}{\partial t'} + \nabla_{\mathbf{r}} \cdot (\rho_m \mathbf{u}) = 0 \tag{71}$$

$$\frac{\partial \rho_u}{\partial t'} + \nabla_{\mathbf{r}} \cdot (\rho_u \mathbf{u}) + p \nabla_{\mathbf{r}} \cdot \mathbf{u} = 0\,, \tag{72}$$

where $\nabla_{\mathbf{r}} \cdot \mathbf{u}$ gives the rate of change of the volume of the fluid element and $p\nabla_{\mathbf{r}} \cdot \mathbf{u}$ is the work done by pressure. In Newtonian physics, rest energy (mass) is not included in the energy density. Eq. (72) applies whether we include it or not. Define total energy density as

$$\rho \equiv \rho_m + \rho_u\,,$$

where ρ_u is the Newtonian energy density and ρ_m is the mass density. Adding Eqs. (71) and (72) gives

$$\frac{\partial \rho}{\partial t'} + \nabla_{\mathbf{r}} \cdot (\rho \mathbf{u}) + p \nabla_{\mathbf{r}} \cdot \mathbf{u} = 0\,. \tag{73}$$

For nonrelativistic matter $\rho_u \ll \rho_m$ and $p \ll \rho_m$. We can thus drop the last term in (73) and ignore the distinction between mass density and total energy density.

A homogeneously expanding fluid,

$$\rho = \rho(t_0) a^{-3} \tag{74}$$

$$\mathbf{u} = \frac{\dot{a}}{a}\mathbf{r} \tag{75}$$

$$\tilde{\Phi} = \frac{2\pi G}{3}\rho r^2 \tag{76}$$

is a solution to these equations (**exercise**), with a condition to the function $a(t)$ giving the expansion law. It is the Newtonian version of the matter-dominated Friedmann model. Writing $H(t) \equiv \dot{a}/a$ we find that the homogeneous solution satisfies

$$\dot{\rho} + 3H\rho = 0\,, \tag{77}$$

and the condition for $a(t)$ (from the exercise) can be written as

$$\frac{\ddot{a}}{a} = \dot{H} + H^2 = -\frac{4\pi G}{3}\rho\,. \tag{78}$$

You should recognize these equations as the energy-continuity equation and the second Friedmann equation for a matter-dominated FRW universe.[12] The result for $\tilde{\Phi}$, Eq. (76), has no relativistic counterpart, the whole concept of gravitational potential does not exist in relativity (except in special cases; like here in perturbation theory, where we introduce potentials related to perturbations).

[12] The freedom of choosing the initial value of the expansion rate leaves the connection between H and ρ open up to a constant. This constant has the same effect on the time evolution of $a(t)$ as the curvature constant K in the first Friedmann equation, but of course in the Newtonian treatment it is not interpreted as curvature, and it does not otherwise have the same physical effects. We shall (unless otherwise noted) choose this constant so that the background solution matches the *flat* FRW universe. Then we have

$$H^2 = \frac{8\pi G}{3}\rho \qquad \text{or} \qquad 4\pi G\rho = \frac{3}{2}H^2\,. \tag{79}$$

8.2.1 Comoving coordinates

Introduce now a new (comoving) coordinate system $(t, \mathbf{x})$ which is related to the Newtonian coordinate system $(t', \mathbf{r})$ by

$$t' = t \qquad \mathbf{r} = a(t)\mathbf{x}\,. \tag{80}$$

Thus the time coordinate is the same in both coordinate systems, but we need to distinguish between the partial derivatives $\partial/\partial t$ and $\partial/\partial t'$, since in the first $\mathbf{x}$ is kept constant and in the second $\mathbf{r}$ is kept constant. Relate now the partial derivatives:

$$\begin{aligned}
\frac{\partial}{\partial t} &= \frac{\partial t'}{\partial t}\frac{\partial}{\partial t'} + \sum_i \frac{\partial r_i}{\partial t}\frac{\partial}{\partial r_i} = \frac{\partial}{\partial t'} + \sum \dot{a} x_i \frac{\partial}{\partial r_i} = \frac{\partial}{\partial t'} + H\mathbf{r}\cdot\nabla_{\mathbf{r}} \\
\frac{\partial}{\partial x_i} &= \frac{\partial t'}{\partial x_i}\frac{\partial}{\partial t'} + \sum_j \frac{\partial r_j}{\partial x_i}\frac{\partial}{\partial r_j} = \sum_j \delta_{ij} a \frac{\partial}{\partial r_j} = a\frac{\partial}{\partial r_i} \quad \Rightarrow \quad \nabla_{\mathbf{x}} = a\nabla_{\mathbf{r}}\,.
\end{aligned} \tag{81}$$

Thus

$$\frac{\partial}{\partial t'} = \frac{\partial}{\partial t} - H\mathbf{x}\cdot\nabla_{\mathbf{x}} \qquad \text{and} \qquad \nabla_{\mathbf{r}} = \frac{1}{a}\nabla_{\mathbf{x}}\,. \tag{82}$$

(Later we will work exclusively in the comoving coordinates and write just ∇ for $\nabla_{\mathbf{x}}$. The "original" coordinates $\mathbf{r}$ are just an artifact of the Newtonian approach and do not appear in relativistic perturbation theory.)

8.2.2 The perturbation

Now, consider a small perturbation, so that

$$\rho(t', \mathbf{r}) = \bar{\rho}(t') + \delta\rho(t', \mathbf{r}) \tag{83}$$

$$p(t', \mathbf{r}) = \bar{p}(t') + \delta p(t', \mathbf{r}) \tag{84}$$

$$\mathbf{u}(t', \mathbf{r}) = H(t')\mathbf{r} + \mathbf{v}(t', \mathbf{r}) \tag{85}$$

$$\tilde{\Phi}(t', \mathbf{r}) = \frac{2\pi G}{3}\bar{\rho}(t')r^2 + \Phi(t', \mathbf{r})\,, \tag{86}$$

where $\bar{\rho}$, $\bar{p}$, and H denote homogeneous background quantities (solutions of the background, or zeroth-order, equations) and $\delta\rho$, δp, $\mathbf{v}$, Φ are small inhomogeneous perturbations.

Inserting these into the Eqs. (67,68,69) and subtracting the homogeneous equations (76,77,78) we get (**exercise**) the *perturbation equations*

$$\frac{\partial\delta\rho}{\partial t'} + 3H\delta\rho + H\mathbf{r}\cdot\nabla_{\mathbf{r}}\delta\rho + \bar{\rho}\nabla_{\mathbf{r}}\cdot\mathbf{v} = 0 \tag{87}$$

$$\frac{\partial\mathbf{v}}{\partial t'} + H\mathbf{v} + H\mathbf{r}\cdot\nabla_{\mathbf{r}}\mathbf{v} + \frac{1}{\bar{\rho}}\nabla_{\mathbf{r}}\delta p + \nabla_{\mathbf{r}}\Phi = 0 \tag{88}$$

$$\nabla^2_{\mathbf{r}}\Phi = 4\pi G\delta\rho\,. \tag{89}$$

In terms of the comoving coordinates these become (**exercise**):

$$\frac{\partial\delta\rho}{\partial t} + 3H\delta\rho + \frac{\bar{\rho}}{a}\nabla_{\mathbf{x}}\cdot\mathbf{v} = 0 \tag{90}$$

$$\frac{\partial\mathbf{v}}{\partial t} + H\mathbf{v} + \frac{1}{a\bar{\rho}}\nabla_{\mathbf{x}}\delta p + \frac{1}{a}\nabla_{\mathbf{x}}\Phi = 0 \tag{91}$$

$$\nabla^2_{\mathbf{x}}\Phi = 4\pi G a^2\delta\rho\,. \tag{92}$$

In terms of the relative density perturbation $\delta \equiv \delta\rho/\bar{\rho}$ we have $\delta\rho = \bar{\rho}\cdot\delta$ and

$$\frac{\partial\delta\rho}{\partial t} = \dot{\bar{\rho}}\cdot\delta + \bar{\rho}\frac{\partial\delta}{\partial t} \qquad \text{where} \qquad \dot{\bar{\rho}}\cdot\delta = -3H\bar{\rho}\,\delta\,, \tag{93}$$

and we can write

$$\frac{\partial \mathbf{v}}{\partial t} + H\mathbf{v} = \frac{1}{a}\frac{\partial}{\partial t}(a\mathbf{v}) \tag{94}$$

so that the set of perturbation equations becomes

$$\frac{\partial \delta}{\partial t} + \frac{1}{a}\nabla_{\mathbf{x}} \cdot \mathbf{v} = 0 \tag{95}$$

$$\frac{\partial}{\partial t}(a\mathbf{v}) + \frac{1}{\bar{\rho}}\nabla_{\mathbf{x}}\delta p + \nabla_{\mathbf{x}}\Phi = 0 \tag{96}$$

$$\nabla^2_{\mathbf{x}}\Phi = 4\pi G a^2 \bar{\rho}\delta \tag{97}$$

Finally, we Fourier expand the perturbations,

$$\delta(t, \mathbf{x}) = \sum_{\mathbf{k}} \delta_{\mathbf{k}}(t) e^{i\mathbf{k}\cdot\mathbf{x}} \qquad \text{etc.} \tag{98}$$

In Fourier space the perturbation equations become

$$\dot{\delta}_{\mathbf{k}} + \frac{i\mathbf{k}\cdot\mathbf{v}_{\mathbf{k}}}{a} = 0 \tag{99}$$

$$\frac{d}{dt}(a\mathbf{v}_{\mathbf{k}}) + i\mathbf{k}\frac{\delta p_{\mathbf{k}}}{\bar{\rho}} + i\mathbf{k}\Phi_{\mathbf{k}} = 0 \tag{100}$$

$$\Phi_{\mathbf{k}} = -4\pi G\left(\frac{a}{k}\right)^2 \bar{\rho}\delta_{\mathbf{k}}\,. \tag{101}$$

Solving the evolution of the perturbations is a two-step process:

1) Solve the background equations to obtain the functions $a(t)$, $H(t)$, and $\bar{\rho}(t)$. After this, these are *known functions* in the perturbation equations.

2) Solve the perturbation equations.

8.2.3 Vector and scalar perturbations

We now divide the velocity perturbation field $\mathbf{v}(t, \mathbf{r})$ into its rotational (solenoidal, divergence-free) and irrotational (curl-free) parts,

$$\mathbf{v} = \mathbf{v}_\perp + \mathbf{v}_\parallel\,, \tag{102}$$

where $\nabla \cdot \mathbf{v}_\perp = 0$ and $\nabla \times \mathbf{v}_\parallel = 0$. For Fourier components this simply means that $\mathbf{k} \cdot \mathbf{v}_{\perp\mathbf{k}} = 0$ and $\mathbf{k} \times \mathbf{v}_{\parallel\mathbf{k}} = 0$. That is, we divide $\mathbf{v}_\mathbf{k}$ into the components perpendicular and parallel to the wave vector $\mathbf{k}$. The parallel part we can write in terms of a scalar function v, whose Fourier components $v_\mathbf{k}$ are given by

$$\mathbf{v}_{\parallel\mathbf{k}} \equiv v_{\mathbf{k}}\hat{\mathbf{k}}\,, \tag{103}$$

where $\hat{\mathbf{k}}$ denotes the unit vector in the $\mathbf{k}$ direction.

We can now take the perpendicular and parallel parts of Eq. (100),

$$\frac{d}{dt}(a\mathbf{v}_{\perp\mathbf{k}}) = 0 \tag{104}$$

$$\frac{d}{dt}(av_{\mathbf{k}}) + ik\frac{\delta p_{\mathbf{k}}}{\bar{\rho}} + ik\Phi_{\mathbf{k}} = 0\,. \tag{105}$$

We see that the rotational part of the velocity perturbation has a simple time evolution,

$$\mathbf{v}_\perp \propto a^{-1}\,, \tag{106}$$

i.e., it *decays* from whatever initial value it had, inversely proportional to the scale factor.

The other perturbation equations involve only the irrotational part of the velocity perturbation. Thus we can divide the total perturbation into two parts, commonly called the *vector* and *scalar* perturbations, which evolve independent of each other:

1) The vector perturbation: $\mathbf{v}_\perp$.

2) The scalar perturbation: δ, δp, v, Φ, which are all coupled to each other.

The vector perturbations are thus not related to the density perturbations, or the structure of the universe. Also, any primordial vector perturbation should become rather small as the universe expands, at least while first-order perturbation theory applies.[13] They are thus not very important, and we shall have no more to say about them. The rest of our discussion focuses on the scalar perturbations.

8.2.4 The equations for scalar perturbations

We summarize here the Fourier space equations for scalar perturbations:

$$\dot{\delta}_\mathbf{k} + \frac{ikv_\mathbf{k}}{a} = 0 \qquad \Rightarrow \qquad v_\mathbf{k} = i\frac{a}{k}\dot{\delta}_\mathbf{k} \tag{107}$$

$$\frac{d}{dt}(av_\mathbf{k}) + ik\frac{\delta p_\mathbf{k}}{\bar{\rho}} + ik\Phi_\mathbf{k} = 0 \tag{108}$$

$$\Phi_\mathbf{k} = -4\pi G\left(\frac{a}{k}\right)^2 \bar{\rho}\delta_\mathbf{k}\,. \tag{109}$$

Inserting $v_\mathbf{k}$ from (107) and $\Phi_\mathbf{k}$ from (109) into (108) we get

$$\boxed{\ddot{\delta}_\mathbf{k} + 2H\dot{\delta}_\mathbf{k} = -\frac{k^2}{a^2}\frac{\delta p_\mathbf{k}}{\bar{\rho}} + 4\pi G\bar{\rho}\delta_\mathbf{k}\,.} \tag{110}$$

8.2.5 Adiabatic and entropy perturbations

Suppose the equation of state is *barotropic*,

$$p = p(\rho) \tag{111}$$

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

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

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

$$\dot{\bar{p}} = \frac{d\bar{p}}{dt} = \frac{dp}{d\rho}(\bar{\rho})\frac{d\bar{\rho}}{dt} = \frac{dp}{d\rho}\dot{\bar{\rho}}\,.$$

Assuming this derivative $dp/d\rho$ is nonnegative, we call its square root the *speed of sound*

$$c_s \equiv \sqrt{\frac{dp}{d\rho}}\,. \tag{112}$$

[13] Thus we end up with an irrotational velocity field. The rotational motion (e.g., rotation of galaxies) which is common in the present universe at small scales has arisen from higher-order effects from the primordial scalar perturbations, not from the primordial vector perturbations.

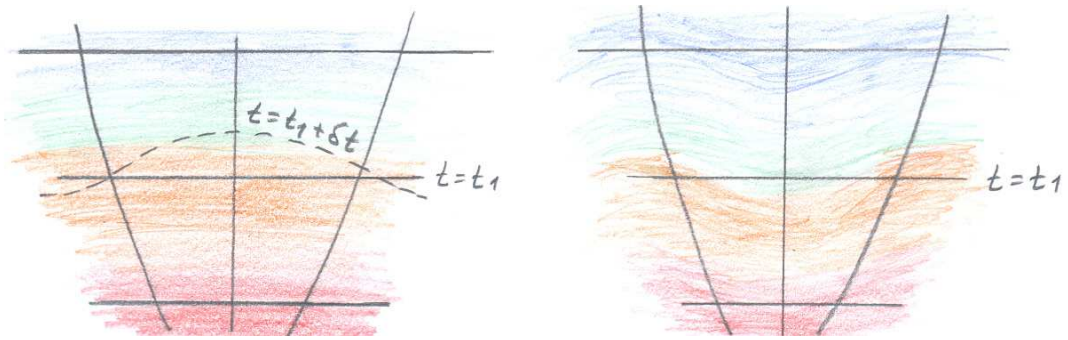


Figure 5: For adiabatic perturbations, the conditions in the perturbed universe (right) at $(t_1, \mathbf{x})$ equal conditions in the (homogeneous) background universe (left) at some time $t_1 + \delta t(\mathbf{x})$.

(We shall indeed find that sound waves propagate at this speed.) We thus have the relation

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

In general, when p may depend on other variables besides ρ, the speed of sound in a fluid is given by

$$c_s^2 = \left(\frac{\partial p}{\partial \rho}\right)_S \tag{113}$$

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 that

$$\frac{\dot{\bar{p}}}{\dot{\bar{\rho}}} = \left(\frac{\partial p}{\partial \rho}\right)_S = c_s^2 \,. \tag{114}$$

Perturbations with the property

$$\frac{\delta p}{\delta \rho} = \frac{\dot{\bar{p}}}{\dot{\bar{\rho}}} \tag{115}$$

are called *adiabatic perturbations* in cosmology.

If $p = p(\rho)$, perturbations are necessarily adiabatic. In the general case the perturbations may or may not be adiabatic. In the latter case, the perturbation can be divided into an adiabatic component and an *entropy perturbation*. An entropy perturbation is a perturbation in the entropy-per-particle ratio.

For adiabatic perturbations we thus have

$$\delta p = c_s^2 \delta \rho = \frac{\dot{\bar{p}}}{\dot{\bar{\rho}}} \delta \rho \,. \tag{116}$$

Adiabatic perturbations have the property that the local state of matter (determined here by the quantities p and ρ) at some spacetime point $(t, \mathbf{x})$ of the perturbed universe is the same as in the background universe at some slightly different time $t + \delta t$, this time difference being different for different locations $\mathbf{x}$. See Fig. 5.

Thus we can view adiabatic perturbations as some parts of the universe being "ahead" and others "behind" in the evolution.

Adiabatic perturbations are the simplest kind of perturbations. Single-field inflation produces adiabatic perturbations, since perturbations in all quantities are proportional to a perturbation $\delta\varphi$ in a single scalar quantity, the inflaton field.

Adiabatic perturbations stay adiabatic while they are outside horizon, but may develop entropy perturbations when they enter the horizon. This happens for many-component fluids (discussed a little later).

Present observational data is consistent with the primordial (i.e., before horizon entry) perturbations being adiabatic.

8.2.6 Adiabatic perturbations in matter

Consider now adiabatic perturbations of a non-relativistic single-component fluid. The equation for the density perturbation is now

$$\boxed{\ddot{\delta}_{\mathbf{k}} + 2H\dot{\delta}_{\mathbf{k}} + \left[\frac{c_s^2 k^2}{a^2} - 4\pi G\bar{\rho}\right]\delta_{\mathbf{k}} = 0\,.} \tag{117}$$

I shall call this the *Jeans equation*[14] (although Jeans considered a static, not an expanding fluid).

This is a second-order differential equation from which we can solve the time evolution of the Fourier amplitudes $\delta_{\mathbf{k}}(t)$ of the perturbation. Before solving this equation we need to first find the background solution which gives the functions $a(t)$, $H(t) = \dot{a}/a$, and $\bar{\rho}(t)$.

The nature of the solution to Eq. (117) depends on the sign of the factor in the brackets. The first term in the brackets is due to pressure gradients. Pressure tries to resist compression, so if this term dominates, we get an oscillating solution, standing density (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,

$$k_J \;=\; \frac{a\sqrt{4\pi G\bar{\rho}}}{c_s} \;=\; \sqrt{\frac{3}{2}}\frac{1}{c_s}\mathcal{H}\,, \tag{118}$$

is called the *Jeans wave number*, and the corresponding wavelength

$$\lambda_J \;=\; \frac{2\pi}{k_J} \;=\; 2\pi c_s\sqrt{\frac{2}{3}}\mathcal{H}^{-1} \tag{119}$$

the *Jeans length*. Here $\mathcal{H} \equiv aH$, the comoving Hubble parameter. In the latter equalities we assumed that the background solution is the flat FRW universe, so that

$$4\pi G\bar{\rho} = \tfrac{3}{2}H^2\,. \tag{120}$$

For nonrelativistic matter $c_s \ll 1$, so that the Jeans length is much smaller than the Hubble length, $k_J \gg \mathcal{H}$. Thus we can apply Newtonian theory for scales both larger and smaller than the Jeans length.

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

$$\ddot{\delta}_{\mathbf{k}} + 2H\dot{\delta}_{\mathbf{k}} + \frac{c_s^2 k^2}{a^2}\delta_{\mathbf{k}} = 0\,. \tag{121}$$

The solutions are oscillating, i.e., we get sound waves. The exact solutions of (121) are Bessel functions, but for small scales we can make a further approximation by first ignoring the middle term (which is smaller than the other two) and the time-dependence of a and c_s to get that $\delta_{\mathbf{k}}(t) \sim e^{\pm i\omega t}$, where $\omega = c_s k/a$. These oscillations are damped by the $2H\dot{\delta}_{\mathbf{k}}$ term, so the amplitude of the oscillations decreases with time. There is no growth of structure for sub-Jeans scales.

[14] In the literature, there is usually no name given to this equation, but the terms *Jeans length* etc. are standard.

Exercise: Sound waves. For short-wavelength modes $k \gg k_J$, density perturbations in the matter-dominated universe satisfy (121). Switch to conformal time, $d\eta = dt/a$, and solve $\delta_{\mathbf{k}}(\eta)$ for the $\Omega_m = 1$, $\Omega_\Lambda = 0$ cosmology, assuming $c_s = const.$ How does the amplitude and frequency of the oscillations change with time and scale factor? (Hint: The solutions are spherical Bessel functions.)

For **scales much longer than the Jeans length** (but still subhorizon), $\mathcal{H} \ll k \ll k_J$, we can approximate the Jeans equation by

$$\ddot{\delta}_{\mathbf{k}} + 2H\dot{\delta}_{\mathbf{k}} - 4\pi G\bar{\rho}\delta_{\mathbf{k}} = 0\,. \tag{122}$$

We dropped the pressure gradient term, which means that this equation applies also to nonadiabatic perturbations for scales where pressure gradients can be ignored. Note that Eq. (122) is the same for all $\mathbf{k}$, i.e., there is no k-dependence in the coefficients. This means that the equation applies also in coordinate space, i.e. for $\delta(\mathbf{x})$, as long as we ignore contributions from scales that do not satisfy $\mathcal{H} \ll k \ll k_J$.

For a matter-dominated universe, the background solution is $a \propto t^{2/3}$, so that

$$H = \frac{\dot{a}}{a} = \frac{2}{3t} \tag{123}$$

and

$$\frac{8\pi G}{3}\bar{\rho} = H^2 = \frac{4}{9t^2} \qquad \Rightarrow \qquad \bar{\rho} = \frac{1}{6\pi G t^2}\,, \tag{124}$$

so the Jeans equation becomes

$$\ddot{\delta}_{\mathbf{k}} + \frac{4}{3t}\dot{\delta}_{\mathbf{k}} - \frac{2}{3t^2}\delta_{\mathbf{k}} = 0\,. \tag{125}$$

The general solution is

$$\delta_{\mathbf{k}}(t) = b_1 t^{2/3} + b_2 t^{-1}\,. \tag{126}$$

The first term is the *growing mode* and the second term the *decaying mode*. After some time the decaying mode has died out, and the perturbation grows

$$\boxed{\delta \propto t^{2/3} \propto a\,.} \tag{127}$$

Thus *density perturbations in matter grow proportional to the scale factor.*

From Eq. (101) we have that

$$\Phi \propto a^2\bar{\rho}\delta \propto a^2 a^{-3} a = \text{const.}$$

The gravitational potential perturbation is constant in time during the matter-dominated era.

8.2.7 Radiation

Since radiation is a relativistic form of energy, we cannot apply the preceding Newtonian discussion to perturbations in radiation. However, the qualitative results are similar.

The equation of state for radiation is $p = \rho/3$, and the speed of sound in a radiation fluid is given by

$$c_s^2 = \frac{dp}{d\rho} = \frac{1}{3}\,.$$

Thus the Jeans length for radiation is comparable to the Hubble length, and the subhorizon scales are also sub-Jeans scales for radiation. Thus for subhorizon radiation perturbations we only get oscillatory solutions. During the radiation-dominated epoch they are not damped by expansion, but the oscillation amplitude stays roughly constant.

Since

$$\Phi \propto a^2 \bar{\rho}\delta \propto a^{-2}\delta\,,$$

the amplitude of the gravitational potential oscillation decays.

Relativistic perturbations in non-expanding space. While the full treatment of relativistic perturbations is beyond the level of this course, we can obtain the limit where we ignore the effect of expansion by combining special relativity and the Newtonian limit of general relativity.

Special relativistic fluid dynamics follows from the energy-momentum continuity equation

$$\frac{\partial T^{\mu\nu}}{\partial x^\nu} \equiv \partial_\nu T^{\mu\nu} \equiv T^{\mu\nu}{}_{,\nu} \;=\; 0\,. \tag{128}$$

For a perfect fluid

$$T^{\mu\nu} = (\rho + p)u^\mu u^\nu + p g^{\mu\nu}\,, \tag{129}$$

where the metric is now that of Minkowski space, $g^{\mu\nu} = \mathrm{diag}(-1,1,1,1)$. The 4-velocity u^μ is related to the 3-velocity $\mathbf{v} = v^i$ by

$$u^\mu = (\gamma, \gamma\mathbf{v})\,, \tag{130}$$

where $\gamma = 1/\sqrt{1-v^2}$.

By contracting the energy tensor $T^{\mu\nu}$ with the 4-velocity u_μ we obtain $u_\nu T^{\mu\nu}{}_{,\nu} = 0$, which gives

$$(\rho u^\mu)_{,\mu} + p u^\mu{}_{,\mu} \;=\; 0\,, \tag{131}$$

the energy continuity equation. Subtracting u^ν times this from (128) we get the special relativistic Euler equation

$$(\rho + p)u^\mu u^\nu{}_{,\mu} + (g^{\mu\nu} + u^\mu u^\nu)p_{,\mu} \;=\; 0\,, \tag{132}$$

where

$$u^\mu u^\nu{}_{,\mu} \equiv a^\nu \tag{133}$$

is the 4-acceleration.

For small velocities, $v \ll 1$, we can approximate $\gamma \approx 1$, so that

$$u^\mu \approx (1, \mathbf{v}) \tag{134}$$

and (131),(132) become

$$\begin{aligned} \frac{\partial \rho}{\partial t} + \nabla\cdot(\rho\mathbf{v}) &= -p\nabla\cdot\mathbf{v} \\ (\rho+p)\left(\frac{\partial}{\partial t} + \mathbf{v}\cdot\nabla\right)\vec{v} &= -\nabla p - \mathbf{v}(\mathbf{v}\cdot\nabla p) \;\approx\; -\nabla p\,. \end{aligned} \tag{135}$$

In the Newtonian limit of general relativity, but without the assumption $p \ll \rho$, the passive gravitational mass density is given by $\rho + p$, so that the gravitational force on a volume element of fluid is given by $-(\rho+p)\nabla\Phi$ and the active gravitational mass density by $\rho + 3p$, so that the gravitational potential is given by

$$\nabla^2\Phi = 4\pi G(\rho + 3p)\,. \tag{136}$$

Thus the Euler equation with gravity becomes

$$(\rho+p)\left(\frac{\partial}{\partial t} + \mathbf{v}\cdot\nabla\right)\mathbf{v} \;=\; -\nabla p - (\rho+p)\nabla\Phi\,. \tag{137}$$

For several fluid components, not interacting with each other except gravitationally, the fluid equations become thus

$$\begin{aligned} \frac{\partial \rho_i}{\partial t} + \nabla\cdot(\rho_i\mathbf{v}_i) &= -p_i\nabla\cdot\mathbf{v}_i \\ (\rho_i+p_i)\left(\frac{\partial}{\partial t} + \mathbf{v}_i\cdot\nabla\right)\mathbf{v}_i &= -\nabla p_i - (\rho_i+p_i)\nabla\Phi \\ \nabla^2\Phi &= 4\pi G\sum_i(\rho_i + 3p_i)\,. \end{aligned} \tag{138}$$

For perturbations $\rho_i = \bar{\rho}_i + \delta\rho_i = \bar{\rho}_i(1+\delta_i)$, $p_i = \bar{p}_i + \delta p_i$, where the background density and pressure are now constant both in space and time, we get to first order in perturbations

$$\begin{aligned} \frac{\partial \delta_i}{\partial t} &= -(1+w_i)\nabla\cdot\mathbf{v}_i \\ (\bar{\rho}_i+\bar{p}_i)\frac{\partial \mathbf{v}_i}{\partial t} &= -\nabla\delta p_i - (\bar{\rho}_i+\bar{p}_i)\nabla\Phi \\ \nabla^2\Phi &= 4\pi G\sum_i(\bar{\rho}_i\delta_i + 3\delta p_i)\,, \end{aligned} \tag{139}$$

where $w_i \equiv \bar{p}_i/\bar{\rho}_i$. For Fourier components this becomes

$$\begin{aligned} \dot{\delta}_{i\mathbf{k}} &= -ik(1+w_i)\mathbf{k}\cdot\mathbf{v}_{i\mathbf{k}} \\ (\bar{\rho}_i+\bar{p}_i)\dot{\mathbf{v}}_{i\mathbf{k}} &= -i\mathbf{k}\delta p_{i\mathbf{k}} - i\mathbf{k}(\bar{\rho}_i+\bar{p}_i)\Phi_{\mathbf{k}} \\ \Phi_{\mathbf{k}} &= \frac{-4\pi G}{k^2}\sum_i(\bar{\rho}_i\delta_{i\mathbf{k}} + 3\delta p_{i\mathbf{k}})\,. \end{aligned} \tag{140}$$

For vector perturbations the second equation gives

$$\dot{\mathbf{v}}_{i\perp\mathbf{k}} = 0 \quad\Rightarrow\quad \mathbf{v}_{i\perp\mathbf{k}} = const\,, \tag{141}$$

and for scalar perturbations the first and second equations become

$$\begin{aligned} \dot{\delta}_{i\mathbf{k}} &= -i(1+w_i)kv_{i\mathbf{k}} \\ \dot{v}_{i\mathbf{k}} &= -ik\frac{\delta p_{i\mathbf{k}}}{\bar{\rho}_i+\bar{p}_i} - ik\Phi_{\mathbf{k}}\,. \end{aligned} \tag{142}$$

If $w_i = const$, so that $\dot{w}_i = 0$, we get the Jeans equation

$$\ddot{\delta}_{i\mathbf{k}} + k^2\frac{\delta p_{i\mathbf{k}}}{\bar{\rho}_i} + k^2(1+w_i)\Phi_{\mathbf{k}} = 0\,. \tag{143}$$

8.2.8 Many fluid components

Assume now that the "cosmic fluid" contains several components i (different types of matter or energy) which *do not interact with each other*, except gravitationally. This means that each component feels only its own pressure[15], and that the components can have different flow velocities. Then the Newtonian equations for each component i are

$$\frac{\partial \rho_i}{\partial t'} + \nabla_{\mathbf{r}}\cdot(\rho_i\mathbf{u}_i) = 0 \tag{144}$$

$$\frac{\partial \mathbf{u}_i}{\partial t'} + (\mathbf{u}_i\cdot\nabla_{\mathbf{r}})\mathbf{u}_i + \frac{1}{\rho_i}\nabla_{\mathbf{r}}p_i + \nabla_{\mathbf{r}}\tilde{\Phi} = 0 \tag{145}$$

$$\nabla^2_{\mathbf{r}}\tilde{\Phi} = 4\pi G\rho\,, \tag{146}$$

where $\rho = \sum\rho_i$. Note that there is only one gravitational potential $\tilde{\Phi}$, due to the total density, and this way the different components do interact gravitationally.

We again have the homogeneous solution, where now each component has to satisfy

$$\dot{\rho}_i + 3H\rho_i = 0\,, \tag{147}$$

[15] In standard cosmology, we actually have just one component, the baryon-photon fluid, which feels its own pressure, and the other components do not feel even *their own* pressure (neutrinos after decoupling) or do not even *have* pressure (cold dark matter). But we shall first do this general treatment, and do the application to standard cosmology later.

and the expansion law

$$\dot{H} + H^2 = -\frac{4\pi G}{3}\rho\,, \tag{148}$$

is determined by the total density.

We can now introduce the density, pressure, and velocity perturbations for each component separately,

$$\begin{aligned}
\rho_i(t', \mathbf{r}) &= \bar{\rho}_i(t) + \delta\rho_i(t', \mathbf{r}) && (149)\\
p_i(t', \mathbf{r}) &= \bar{p}_i(t) + \delta p_i(t', \mathbf{r}) && (150)\\
\mathbf{u}_i(t', \mathbf{r}) &= H(t)\mathbf{r} + \mathbf{v}_i(t', \mathbf{r})\,, && (151)
\end{aligned}$$

but there is only one gravitational potential perturbation,

$$\tilde{\Phi}(t', \mathbf{r}) = \frac{2\pi G}{3}\bar{\rho}r^2 + \Phi(t', \mathbf{r})\,. \tag{152}$$

Following the earlier procedure, we obtain the perturbation equations for the fluid components,

$$\begin{aligned}
\frac{\partial}{\partial t'}\delta\rho_i + 3H\delta\rho_i + H\mathbf{r}\cdot\nabla_{\mathbf{r}}\delta\rho_i + \bar{\rho}_i\nabla\cdot\mathbf{v}_i &= 0 && (153)\\
\frac{\partial}{\partial t'}\mathbf{v}_i + H\mathbf{v}_i + H\mathbf{r}\cdot\nabla_{\mathbf{r}}\mathbf{v}_i + \frac{1}{\bar{\rho}_i}\nabla_{\mathbf{r}}\delta p_i + \nabla_{\mathbf{r}}\Phi &= 0 && (154)\\
\nabla^2_{\mathbf{r}}\Phi &= 4\pi G\delta\rho && (155)
\end{aligned}$$

in Newtonian coordinate space, and

$$\begin{aligned}
\dot{\delta}_{i\mathbf{k}} + \frac{i\mathbf{k}\cdot\mathbf{v}_{i\mathbf{k}}}{a} &= 0 && (156)\\
\frac{d}{dt}\left(a\mathbf{v}_{i\mathbf{k}}\right) + i\mathbf{k}\frac{\delta p_{i\mathbf{k}}}{\bar{\rho}_i} + i\mathbf{k}\Phi_{\mathbf{k}} &= 0 && (157)\\
\Phi_{\mathbf{k}} &= -4\pi G\frac{a^2}{k^2}\sum\bar{\rho}_i\delta_{i\mathbf{k}} && (158)
\end{aligned}$$

in comoving Fourier space. Here $\delta\rho = \sum\delta\rho_i$ and

$$\delta_i \equiv \frac{\delta\rho_i}{\bar{\rho}_i}\,. \tag{159}$$

Separating out the scalar perturbations we finally get

$$\ddot{\delta}_{i\mathbf{k}} + 2H\dot{\delta}_{i\mathbf{k}} = -\frac{k^2}{a^2}\frac{\delta p_{i\mathbf{k}}}{\bar{\rho}_i} + 4\pi G\delta\rho_{\mathbf{k}}\,, \tag{160}$$

where

$$\delta\rho_{\mathbf{k}} = \sum_j \bar{\rho}_j\delta_{j\mathbf{k}}\,. \tag{161}$$

8.2.9 Adiabatic and entropy perturbations again

The simplest inflation models predict that the primordial perturbations are adiabatic. This means that locally the perturbed universe at some $(t, \mathbf{x})$ looks like the background universe at some time $t + \delta t(\mathbf{x})$. See Sec. 8.2.5.

$$\left.\begin{aligned}\delta\rho_i(\mathbf{x}) &= \dot{\bar{\rho}}_i\delta t(\mathbf{x})\\ \delta p_i(\mathbf{x}) &= \dot{\bar{p}}_i\delta t(\mathbf{x})\end{aligned}\right\} \quad\Rightarrow\quad \begin{cases}\dfrac{\delta p_i}{\delta\rho_i} = \dfrac{\dot{\bar{p}}_i}{\dot{\bar{\rho}}_i}\\[2ex] \dfrac{\delta\rho_i}{\delta\rho_j} = \dfrac{\dot{\bar{\rho}}_i}{\dot{\bar{\rho}}_j}\end{cases} \quad\Rightarrow\quad \frac{\delta_i}{\delta_j} = \frac{\dot{\bar{\rho}}_i}{\bar{\rho}_i}\frac{\bar{\rho}_j}{\dot{\bar{\rho}}_j} \tag{162}$$

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

$$\dot{\bar{\rho}}_i \;=\; -3H(\bar{\rho}_i + \bar{p}_i) \;\equiv\; -3H(1+w_i)\bar{\rho}_i\,, \tag{163}$$

where $w_i \equiv \bar{p}_i/\bar{\rho}_i$ is the *equation-of-state parameter* of fluid component i. Thus for adiabatic perturbations,

$$\frac{\delta_i}{1+w_i} = \frac{\delta_j}{1+w_j} \tag{164}$$

(which is thus related to $\bar{\rho}_i \propto a^{-3(1+w_i)}$). For matter components $w_i \approx 0$, and for radiation components $w_i = \frac{1}{3}$. Thus, for adiabatic perturbations, all matter components have the same perturbation

$$\delta_i = \delta_m$$

and all radiation perturbations have likewise

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

We can define a *relative entropy perturbation*[16] between two components

$$S_{ij} \;\equiv\; -3H\left(\frac{\delta\rho_i}{\dot{\bar{\rho}}_i} - \frac{\delta\rho_j}{\dot{\bar{\rho}}_j}\right) \;=\; \frac{\delta_i}{1+w_i} - \frac{\delta_j}{1+w_j} \tag{165}$$

to describe a deviation from the adiabatic case. The relative entropy perturbation is a perturbation in the ratio of the number densities of the two species. For a nonrelativistic species

$$\rho_i = m_i n_i \quad\Rightarrow\quad \delta\rho_i = m_i\delta n_i \qquad \text{and} \qquad \delta_i \equiv \frac{\delta\rho_i}{\bar{\rho}_i} = \frac{\delta n_i}{\bar{n}_i}\,, \tag{166}$$

whereas for an ultrarelativistic species ($\mu \ll T$ and $m \ll T$)

$$\begin{aligned} \rho_i \;&\propto\; T_i^4 \quad\Rightarrow\quad \delta\rho_i = \bar{\rho}_i\cdot 4\frac{\delta T_i}{T_i} \\ n_i \;&\propto\; T_i^3 \quad\Rightarrow\quad \delta n_i = \bar{n}_i\cdot 3\frac{\delta T_i}{T_i} \\ &\Rightarrow\quad \delta_i \equiv \frac{\delta\rho_i}{\bar{\rho}_i} = \frac{4}{3}\frac{\delta n_i}{\bar{n}_i}\,. \end{aligned} \tag{167}$$

For both cases

$$\delta_i = (1+w_i)\frac{\delta n_i}{\bar{n}_i}\,. \tag{168}$$

Thus

$$S_{ij} \;=\; \frac{\delta n_i}{\bar{n}_i} - \frac{\delta n_j}{\bar{n}_j} \;=\; \frac{\delta(n_i/n_j)}{\bar{n}_i/\bar{n}_j}\,. \tag{169}$$

Even if perturbations are initially adiabatic, relative entropy perturbation may develop inside the horizon. We shall encounter such a case in Sec. 8.3.4.

[16] There is a connection to entropy/particle of the different components, but we need not concern ourselves with it now. It is not central to this concept, and it is perhaps somewhat unfortunate that it has become customary, for historical reasons, to use the word "entropy" for these perturbations.

8.2.10 The effect of a homogeneous component

The energy density of the real universe consists of several components. In many cases it is reasonable to ignore the perturbations in some components (since they are relatively small in the scales of interest). We call such components *smooth* and we can add them together into a single smooth component $\rho_s = \bar{\rho}_s$.

Consider the case where we have perturbations in a nonrelativistic ("matter") component ρ_m, and the other components are smooth. Then

$$\rho = \rho_m + \rho_s \tag{170}$$

but

$$\delta\rho = \delta\rho_m \equiv \bar{\rho}_m \delta\,. \tag{171}$$

We write just δ for $\delta\rho_m/\bar{\rho}_m$, since there is no other density perturbation, but note that now $\delta \neq \delta\rho/\bar{\rho}$ (beware of this trap!).

Assuming adiabatic perturbations, we have then from Eq. (160) that

$$\ddot{\delta}_{\mathbf{k}} + 2H\dot{\delta}_{\mathbf{k}} + \left[\frac{c_s^2 k^2}{a^2} - 4\pi G\bar{\rho}_m\right]\delta_{\mathbf{k}} = 0\,. \tag{172}$$

The difference from Eq. (117) is that now the background energy density in the "gravity" term still contains only the matter component $\bar{\rho}_m$, but the expansion law, $a(t)$ and $H(t)$ comes from the full background energy density $\bar{\rho} = \bar{\rho}_m + \bar{\rho}_s$.

Newtonian perturbation theory can be applied even with the presence of relativistic energy components, like radiation and dark energy, as long as they can be considered as smooth components and their perturbations can be ignored. Then they contribute only to the background solution. In this case we have to calculate the background solution using general relativity, i.e., the background solution is a FRW universe, but the perturbation equations are the Newtonian perturbation equations. We can also consider a non-flat (open or closed) FRW universe, as long as we only apply perturbation theory to scales much shorter than the curvature radius (and the Hubble length). Thus the background quantities are to be solved from the Friedmann and energy continuity equations

$$H^2 + \frac{K}{a^2} = \frac{8\pi G}{3}\rho \tag{173}$$

$$\dot{H} + H^2 = -\frac{4\pi G}{3}(\rho + 3p) \tag{174}$$

$$\dot{\rho} = -3H(\rho + p)\,. \tag{175}$$

Example: Matter perturbations in flat vacuum-dominated universe. Consider the case where $\rho_s = \rho_{\rm vac} \gg \rho_m$ and matter is approximated as pressureless (we do not then have to make a separate adiabaticity assumption, since the pressure term does not appear). Then the Jeans equation becomes

$$\ddot{\delta}_{\mathbf{k}} + 2H\dot{\delta}_{\mathbf{k}} - 4\pi G\bar{\rho}_m\delta_{\mathbf{k}} = 0\,. \tag{176}$$

To estimate the relative order of magnitude of the three terms it is better to divide the equation by H^2,

$$H^{-2}\ddot{\delta}_{\mathbf{k}} + 2H^{-1}\dot{\delta}_{\mathbf{k}} - \frac{4\pi G\bar{\rho}_m}{H^2}\delta_{\mathbf{k}} = 0\,, \tag{177}$$

so that the Hubble time H^{-1} provides the time scale for the time derivatives. Now

$$H^2 = \frac{8\pi G}{3}\rho_{\rm cr} \approx \frac{8\pi G}{3}\rho_{\rm vac} = const \tag{178}$$

and in the last term $\delta_{\mathbf{k}}$ is multiplied with $\frac{3}{2}\bar{\rho}_m/\rho_{\rm vac} \ll 1$, so that we can drop the last term and approximate the Jeans equation by

$$\ddot{\delta}_{\mathbf{k}} + 2H\dot{\delta}_{\mathbf{k}} \;=\; 0\,. \tag{179}$$

We see immediately that $\delta_{\mathbf{k}} = const$ is a solution. For the other solution, solve first $\dot{\delta}_{\mathbf{k}}$:

$$\frac{d\dot{\delta}_{\mathbf{k}}}{dt} = -2H\dot{\delta}_{\mathbf{k}} \quad \Rightarrow \quad \frac{d\dot{\delta}_{\mathbf{k}}}{\dot{\delta}_{\mathbf{k}}} = -2H dt\,, \tag{180}$$

whose solution is $\ln\dot{\delta}_{\mathbf{k}} = -2Ht + const$ or $\dot{\delta}_{\mathbf{k}} = Ce^{-2Ht}$. Integrating this gives

$$\delta_{\mathbf{k}} \;=\; Ae^{-2Ht} + B\,, \tag{181}$$

with a constant term and an exponentially decaying term. Thus in a vacuum-dominated universe matter perturbations stay constant (after the decaying term has died out); or to be more precise and referring to the original equation (177), the relative change in $\delta_{\mathbf{k}}$ in a Hubble time is of order $\bar{\rho}_m/\rho_{\rm vac} \ll 1$

We shall do this calculation more accurately later, including the transition from matter domination to vacuum domination. The main lesson now is that the increased expansion rate due to the presence of a smooth component slows down the growth of perturbations.

Exercise: Find the solution for the Jeans equation for pressureless matter perturbations when a) the energy density is dominated by a smooth radiation component b) when there is no other energy component, but the universe has the open geometry ($K < 0$) and is curvature dominated, considering only scales $\ll$ curvature radius.

8.2.11 Meszaros equation

Consider now a flat universe with just cold dark matter (CDM) and radiation, and ignore perturbations in radiation so that radiation can be taken as a smooth component. This approximation may be motivated by noting that subhorizon radiation perturbations do not grow. The CDM is pressureless, and thus the CDM sound speed is zero, and so is the CDM Jeans length. Thus, for CDM, all scales are larger than the Jeans scale, and we don't get an oscillatory behavior. Instead, perturbations grow at all scales.

We get the equation for the CDM perturbation from Eq. (172) by setting $c_s = 0$ (or rather, $\delta p = 0$; we need not invoke the assumption of adiabaticity, since CDM is pressureless),

$$\ddot{\delta}_{\mathbf{k}} + 2H\dot{\delta}_{\mathbf{k}} - 4\pi G\bar{\rho}_m\delta_{\mathbf{k}} = 0\,. \tag{182}$$

Note that the equation is the same for all $\mathbf{k}$ and therefore it applies also in the coordinate space, i.e., for $\delta(\mathbf{x})$. To simplify notation, we drop the subscript $_{\mathbf{k}}$.

The Friedmann equation is

$$H^2 = \left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3}\bar{\rho}\,,$$

where $\bar{\rho} = \bar{\rho}_m + \bar{\rho}_r$ and $\bar{\rho}_m \propto a^{-3}$ and $\bar{\rho}_r \propto a^{-4}$.

A useful trick is to study this as a function of a instead of t or η. We define a new time coordinate,

$$y \equiv \frac{a}{a_{\rm eq}} = \frac{\bar{\rho}_m}{\bar{\rho}_r}\,. \tag{183}$$

($y = 1$ at $t = t_{\rm eq}$.) Now

$$4\pi G\bar{\rho}_m \;=\; 4\pi G\frac{y}{y+1}\bar{\rho} \;=\; \frac{3}{2}\frac{y}{y+1}H^2 \tag{184}$$

and Eq. (182) becomes

$$\ddot{\delta} + 2H\dot{\delta} - \frac{3}{2}\frac{y}{y+1}H^2 = 0\,. \tag{185}$$

Performing the change of variables from t to y (**Exercise**; you may need the 2nd Friedmann equation), we arrive at the equation (where $' \equiv d/dy$)

$$\boxed{\delta'' + \frac{2+3y}{2y(1+y)}\delta' - \frac{3}{2y(1+y)}\delta \;=\; 0\,,} \tag{186}$$

known as the *Meszaros equation*.

It has two solutions, one growing, the other one decaying. The growing solution is

$$\boxed{\delta \;=\; \delta_{\rm prim}\left(1+\frac{3y}{2}\right) \;=\; \delta_{\rm prim}\left(1+\frac{3}{2}\frac{a}{a_{\rm eq}}\right)\,.} \tag{187}$$

We see that the perturbation remains frozen to its primordial value, $\delta \approx \delta_{\rm prim}$, during the radiation-dominated period. By $t = t_{\rm eq}$, it has grown to $\delta = \frac{5}{2}\delta_{\rm prim}$.

During the matter-dominated period, $y \gg 1$, the CDM perturbation grows proportional to the scale factor,

$$\delta \propto y \propto a \propto t^{2/3}\,. \tag{188}$$

8.3 Perturbations at subhorizon scales in the real universe

8.3.1 Horizon entry

Newtonian perturbation theory is valid only at subhorizon scales, $k \gg \mathcal{H}$, or $k^{-1} \ll \mathcal{H}^{-1}$. During "normal", decelerating expansion, i.e., after inflation but before the recent onset of dark energy domination, scales are entering the horizon. Short scales enter first, large scales enter later. We have not yet studied what happens to perturbations outside the horizon (for that we need (general) relativistic perturbation theory, to be discussed In Sec. 8.4). So, for the present discussion, whatever values the perturbation amplitudes $\delta_{\mathbf{k}}$ have soon after horizon entry, are to be taken as an initial condition, the *primordial perturbation*[17]. Observations actually suggest that different scales enter the horizon with approximately equal perturbation amplitude, whose magnitude is characterized by the number[18] few $\times\, 10^{-5}$.

The history of the different scales after horizon entry, and thus their present perturbation amplitude, depends on at what epoch they enter. The scales which enter during transitions between epochs are thus special scales which should characterize the present structure of the universe. Such important scales are the scale (**exercise**)

$$k_{\rm eq}^{-1} \;=\; (\mathcal{H}_{\rm eq})^{-1} \;\sim\; 13.7\,\Omega_m^{-1} h^{-2}\,{\rm Mpc} \;\equiv\; 13.7\,\omega_m^{-1}\,{\rm Mpc}\,, \tag{189}$$

which enters at the time $t_{\rm eq}$ ($1 + z_{\rm eq} = 23\,902\,\omega_m$) of matter-radiation equality, and the scale

$$\begin{aligned} k_{\rm dec}^{-1} \;&=\; (\mathcal{H}_{\rm dec})^{-1} \;\sim\; 91\,\Omega_m^{-1/2}\left[1 + \frac{\Omega_r}{\Omega_m}(1+z_{\rm dec})\right]^{-1/2} h^{-1}{\rm Mpc} \\ &\equiv\; 91\,\omega_m^{-1/2}\left[1 + \frac{\omega_r}{\omega_m}(1+z_{\rm dec})\right]^{-1/2} {\rm Mpc}\,, \end{aligned} \tag{190}$$

which enters at the time $t_{\rm dec}$ ($z_{\rm dec} = 1090$) of photon decoupling. Here $\omega_r = 4.184 \times 10^{-5}$ includes relativistic neutrinos, since the result above only requires them to be relativistic at $t_{\rm dec}$. For $\Omega_\Lambda = 0.7$, $\Omega_m = 0.3$, $h = 0.7$, these scales are

$$\begin{aligned} k_{\rm eq}^{-1} \;&=\; 65\,h^{-1}{\rm Mpc} \;=\; 93\,{\rm Mpc} \\ k_{\rm dec}^{-1} \;&=\; 145\,h^{-1}{\rm Mpc} \;=\; 207\,{\rm Mpc}\,. \end{aligned} \tag{191}$$

The smallest "cosmological" scale is that corresponding to a typical distance between galaxies, about 1 Mpc.[19] This scale entered during the radiation-dominated epoch (well after Big Bang nucleosynthesis).

The scale corresponding to the present "horizon" (i.e. Hubble length) is

$$k_0^{-1} \;=\; (\mathcal{H}_0)^{-1} \;=\; 2998\,h^{-1}{\rm Mpc} \;\sim\; 4300\,{\rm Mpc}\,. \tag{192}$$

Because of the acceleration due to dark energy, 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. The horizon entry is not to be taken as an instantaneous process, so these scales were

[17] We shall later redefine *primordial perturbation* to refer to the perturbation at the epoch when all cosmologically interesting scales were well outside the horizon, which is the standard meaning of this concept in cosmology.

[18] Although in coordinate space the relative density perturbation $\delta(\mathbf{x})$ is a dimensionless number, the Fourier quantity $\delta_{\mathbf{k}}$ is not. The size of $\delta_{\mathbf{k}}$ is characterized by the dimensionless value $\mathcal{P}(k)^{1/2}$.

[19] In the present universe, structure at smaller scales has been messed up by galaxy formation, so that it bears little relation to the primordial perturbations at these scales. 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.

never really subhorizon enough for the Newtonian theory to apply to them. Thus we shall just consider scales $k^{-1} < k_0^{-1}$. The largest observable scales, of the order of k_0^{-1}, are essentially at their "primordial" amplitude now.

We shall now discuss the evolution of the perturbations at these scales ($k^{-1} < k_0^{-1}$) after horizon entry, using the Newtonian perturbation theory presented in the previous section.

8.3.2 Composition of the real universe

The present understanding is that there are five components to the energy density of the universe,

1. cold dark matter (c)
2. baryonic matter (b)
3. photons (γ)
4. neutrinos (ν)
5. dark energy (d)

(during the time of interest for this section, i.e., from some time after BBN until the present). Thus

$$\rho = \underbrace{\rho_c + \rho_b}_{\rho_m} + \underbrace{\rho_\gamma + \rho_\nu}_{\rho_r} + \rho_d \,. \tag{193}$$

(Note that ρ_c here is the CDM density, not the critical density, for which we write $\rho_{\rm cr}$.)

Baryons and photons interact with each other until $t = t_{\rm dec}$, so for $t < t_{\rm dec}$ they have to be discussed as a single component,

$$\rho_{b\gamma} = \rho_b + \rho_\gamma \,. \tag{194}$$

The other components do not interact with each other, except gravitationally, during the time of interest. The fluid description of Sec. 8.2 can only be applied to components whose particle mean free paths are shorter than the scales of interest. After decoupling, photons "free stream" and cannot be discussed as a fluid. On the other hand, the photon component becomes then rather homogeneous quite soon, so we can approximate it as a "smooth" component[20]. The same applies to neutrinos for the whole time since the BBN epoch, until the neutrinos become nonrelativistic. This will happen to at least two of the three neutrino species, and then they should be treated as matter (hot dark matter), not radiation. According to observations, the neutrino masses are small enough, not to have a major impact on structure formation. (However, for accurate work this must be taken into account and this effect on structure formation provides the tightest cosmological limits to neutrino masses.) Thus we shall here approximate neutrinos as a smooth radiation component. Dark energy is believed to be relatively smooth. If it is a cosmological constant (vacuum energy) then it is perfectly homogeneous.

The discussion in Sec. 8.2 applies to the case, where ρ can be divided into two components,

$$\rho = \rho_m + \rho_s \,, \tag{195}$$

where the perturbation is only in the matter component ρ_m and $\rho_s = \bar{\rho}_s$ is homogeneous. For perturbations in radiation components and dark energy the Newtonian treatment is not

[20] As long as we are interested in density perturbations only. When we are interested in the CMB anisotropy, the momentum distribution of these photons becomes the focus of our attention.

enough. Unfortunately, we do not have quite this two-component case here. Based on the above discussion, a reasonable approximation is given by a separation into three components:

$$t < t_{\rm dec}: \quad \rho = \rho_c + \rho_{b\gamma} + \rho_s \qquad (\rho_s = \rho_\nu + \rho_d) \tag{196}$$

$$t > t_{\rm dec}: \quad \rho = \rho_c + \rho_b + \rho_s \qquad (\rho_s = \rho_\gamma + \rho_\nu + \rho_d)\,. \tag{197}$$

After decoupling, both ρ_c and ρ_b are matter-like ($p \ll \rho$) and we'll discuss in Sec. 8.3.4 how this case is handled. Before decoupling, the situation is more difficult, since $\rho_{b\gamma}$ is not matter-like, as the pressure provided by photons is large. Here we shall be satisfied with a crude approximation for this period.

The most difficult period is that close to decoupling, where the photon mean free path λ_γ is growing rapidly. The fluid description, which we are here using for the perturbations, applies only to scales $\gg \lambda_\gamma$, whereas the photons are smooth only for scales $\ll \lambda_\gamma$. Thus this period can be treated properly only with large numerical "Boltzmann" codes, such as CMBFAST or CAMB.

8.3.3 CDM density perturbations

Cold dark matter is the dominant structure-forming component in the universe (dark energy dominates the energy density at late times, but does not form structure, or, if it does, these structures are very weak, not far from homogeneous). Observations indicate that $\rho_b \sim 0.2\rho_c$. Thus we get a first approximation to the behavior of the CDM perturbations by ignoring the baryon component and equating

$$\rho_m \approx \rho_c\,.$$

The CDM is pressureless, and thus the CDM sound speed is zero, and so is the CDM Jeans length. Thus, for CDM, all scales are larger than the Jeans scale, and we don't get an oscillatory behavior. Instead, perturbations grow at all scales. On the other hand, as we shall discuss in Sec. 8.3.4, perturbations in $\rho_{b\gamma}$ oscillate before decoupling. Therefore the perturbations in $\rho_{b\gamma}$ will be smaller than those in ρ_c, and we can make a (crude) approximation where we treat $\rho_{b\gamma}$ as a homogeneous component before decoupling. This is important, since although $\rho_b \ll \rho_c$, this is not true for $\rho_{b\gamma}$ at earlier, radiation-dominated, times. At decoupling $\rho_b < \rho_\gamma < \rho_c$. Before matter-radiation equality, there is an epoch when $\rho_c < \rho_\gamma$, but $\delta\rho_c > \delta\rho_{b\gamma}$. For simplicity, we now approximate

$$\rho = \rho_m + \rho_r + \rho_d \tag{198}$$

where $\rho_m = \rho_c$ and $\rho_r = \rho_\gamma + \rho_\nu$ is a smooth component (ρ_ν truly smooth, ρ_γ truly smooth after decoupling, and (crudely) approximated as smooth before decoupling). We have ignored baryons, since they are a subdominant part of $\rho_{b\gamma}$ before decoupling, and a subdominant matter component after decoupling. Likewise, ρ_d is also smooth, and becomes important only close to present times.

We can now study the growth of CDM perturbations even during the radiation-dominated period, as the radiation-component is taken as smooth and affects only the expansion rate. We can study it all the way from horizon entry to the present time, or until the perturbations become nonlinear ($\delta_c = \delta\rho_c/\bar\rho_c \sim 1$).

For the radiation- and matter-dominated epochs, including the transition in between, we then have the case of Sec. 8.2.11, and the matter perturbation grows as (187),

$$\delta \;=\; \delta_{\rm prim}\left(1 + \frac{3y}{2}\right) \;=\; \delta_{\rm prim}\left(1 + \frac{3}{2}\frac{a}{a_{\rm eq}}\right)\,.$$

The perturbation remains frozen to its primordial value, $\delta \approx \delta_{\rm prim}$, during the radiation-dominated period. By $t = t_{\rm eq}$, it has grown to $\delta = \frac{5}{2}\delta_{\rm prim}$. During the matter-dominated

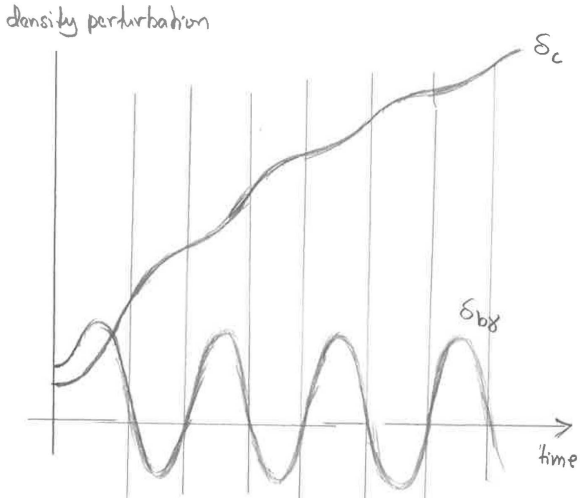


Figure 6: Growth of CDM perturbation during radiation-dominated epoch for the case of adiabatic primordial perturbations (qualitative). The time axis represents conformal time.

period, the CDM perturbation grows proportional to the scale factor,

$$\delta \propto y \propto a \propto t^{2/3} .$$

However, in the case of adiabatic primordial perturbations, the above approximation misses an important effect: an additional logarithmic growth factor $\sim \ln(k/k_{\rm eq})$ the CDM perturbations get from the gravitational effect (ignored in the above) of the oscillating radiation perturbation during the radiation-dominated epoch. To get this boost the CDM perturbation must initially be in the same direction (positive or negative) as the radiation perturbation, which is the case for adiabatic primordial perturbations.

For adiabatic primordial perturbations, the baryon, CDM, and radiation perturbations are related at horizon entry as $\delta_c = \delta_b = \frac{3}{4}\delta_\gamma$. Consider scales that enter during the radiation-dominated epoch ($t < t_{\rm eq} < t_{\rm dec}$). The gravitational effect is dominated initially by the radiation perturbation, which begins to oscillate after horizon entry; the baryon perturbation will oscillate with it until $t_{\rm dec}$. CDM, on the other hand, does not feel the radiation pressure responsible for the oscillation, it sees only the gravitational effect of the baryon-photon fluid. In the first phase of the oscillation period δ_c is of the same sign as $\delta_{b\gamma}$, so $\delta_{b\gamma}$ adds to the gravitational pull to increase δ_c. Since at first $\delta\rho_{b\gamma} > \delta\rho_c$, this additional pull is larger than that of CDM itself, leading to a much faster growth of δ_c (which otherwise would grow very little during the radiation domination). The flow of CDM is accelerated toward CDM overdensities. In the next phase of the oscillation, the sign of $\delta_{b\gamma}$ reverses, and now the pull of $\delta\rho_{b\gamma}$ on CDM is in the opposite direction, and will slow down the flow of CDM toward overdensities. But this is not enough to reverse the CDM flow before the sign of $\delta_{b\gamma}$ changes again and begins to accelerate CDM again toward CDM overdensities.

Thus the effect of the radiation oscillations is to increase δ_c stepwise, one step for each oscillation period. See Fig. 6. As the $\bar\rho_\gamma/\bar\rho_c$ ratio decreases the relative increases per step decrease; but this effect keeps adding steps until $t_{\rm dec}$. The smaller the scale (the higher the k) the more steps there are between horizon entry (t_k) and $t_{\rm dec}$, and the larger the first steps. An analytic calculation (too complicated for this course, but it is done in [9] and I do it in Cosmological Perturbation Theory) of this effect, in the small-scale limit $k \gg k_{\rm eq}$ and still ignoring baryons, so that the oscillating radiation perturbation is just photons, gives that it

leads to a boost by a factor $\sim 7.5\ln(0.17k/k_{\rm eq})$, so that (187) is modified to

$$\delta_c \;\approx\; \delta_{\rm prim}\left(1+\frac{3}{2}\frac{a}{a_{\rm eq}}\right)7.5\ln\left(0.17\frac{k}{k_{\rm eq}}\right) \qquad \text{for } k\gg 6k_{\rm eq} \text{ and } t>t_{\rm dec} \tag{199}$$

(for $k \lesssim 6k_{\rm eq}$ the logarithm is negative; this approximate result does not apply for such large scales). There is more discussion of this result in Sec. 8.4.4, where we compare this approximate analytical result to a more accurate numerical result from CAMB.

8.3.4 Baryon density perturbations

Although CDM is the dominant matter component in the universe, we cannot directly see it. The main method to observe the density perturbation today is to study the distribution of galaxies. But the part of galaxies that we can see is baryonic. Thus to compare the theory of structure formation to observations, we need to study how the perturbation in the baryonic component evolves.

Baryon Jeans length and speed of sound. We define the baryon Jeans length as $\lambda_J = 2\pi k_J^{-1}$, where

$$k_J^{-1} = \frac{c_s}{a\sqrt{4\pi G\bar\rho_b}}\,, \tag{200}$$

and c_s is the speed of sound for baryons (i.e., in the baryon-photon fluid before decoupling, and in the baryon fluid after decoupling). This definition compares the pressure felt by baryons to baryon gravity, so it addresses the question whether the baryon density perturbation can grow under its own gravity. This is not the question we face in reality, since at early times the gravity was dominated by the radiation perturbation and later by the CDM perturbation. The baryon Jeans length can still be used for order-of-magnitude estimates of at what scales the baryon perturbation can grow, and for the argument that we cannot match observations without CDM.

In general,

$$c_s^2 = \left(\frac{\partial p}{\partial\rho}\right)_\sigma\,, \tag{201}$$

where σ refers to constant entropy per baryon. Since in our case the entropy is completely dominated by photons,

$$s_{b\gamma} \sim s_\gamma = \frac{4\pi^2}{45}T^3 = \frac{2\pi^4}{45\zeta(3)}n_\gamma\,, \tag{202}$$

we have

$$\sigma \equiv \frac{s_{b\gamma}}{n_b} \sim \frac{s_\gamma}{n_b} = \frac{2\pi^4}{45\zeta(3)}\frac{n_\gamma}{n_b} \approx 3.6016\frac{1}{\eta}\,, \tag{203}$$

where η is the baryon-to-photon ratio.

We find the speed of sound by varying $\rho_{b\gamma}$ and $p_{b\gamma}$ *adiabatically*, (i.e., keeping σ, the entropy/baryon constant), which in this case means keeping η constant. Now

$$\begin{aligned}
&\rho_b = mn_b = m\eta n_\gamma = m\eta\frac{2\zeta(3)}{\pi^2}T^3 \quad\Rightarrow\quad \delta\rho_b = \bar\rho_b\cdot 3\frac{\delta T}{T}\\
&\rho_\gamma = \frac{\pi^2}{15}T^4 \quad\Rightarrow\quad \delta\rho_\gamma = \bar\rho_\gamma\cdot 4\frac{\delta T}{T}\\
&p_\gamma = \frac{\pi^2}{45}T^4 \quad\Rightarrow\quad \delta p_\gamma = \bar p_\gamma\cdot 4\frac{\delta T}{T} = \bar\rho_\gamma\cdot\frac{4}{3}\frac{\delta T}{T}\,.
\end{aligned}$$

Since $p_b \ll p_\gamma \quad\Rightarrow\quad \delta p_b \ll \delta p_\gamma$, we get

$$c_s^2 = \frac{\delta p}{\delta\rho} = \frac{\delta p_\gamma}{\delta\rho_\gamma+\delta\rho_b} = \frac{\frac{4}{3}\bar\rho_\gamma}{4\bar\rho_\gamma+3\bar\rho_b} = \frac{1}{3}\frac{1}{1+\frac{3}{4}\frac{\bar\rho_b}{\bar\rho_\gamma}}\,. \tag{204}$$

This was a *calculation* of the speed of sound, which one gets by varying the pressure and density adiabatically. It is independent of whether the actual perturbations we study are adiabatic or not.

This result, Eq. (204), applies before decoupling. As we go back in time, $\bar{\rho}_b/\bar{\rho}_\gamma \to 0$ and $c_s^2 \to 1/3$. As we approach decoupling, $\bar{\rho}_b$ becomes comparable to (but still smaller than) $\bar{\rho}_\gamma$ and the speed of sound falls, but not by a large factor.

Newtonian perturbation theory applies only to subhorizon scales. The ratio of the (comoving) baryon Jeans length

$$\lambda_J = \frac{2\pi c_s}{a\sqrt{4\pi G\bar{\rho}_b}}$$

to the comoving Hubble length

$$\mathcal{H}^{-1} = \frac{1}{a\sqrt{\frac{8\pi G}{3}\bar{\rho}}}$$

is

$$\frac{\lambda_J}{\mathcal{H}^{-1}} = \mathcal{H}\lambda_J = 2\pi\sqrt{\frac{2\bar{\rho}}{3\bar{\rho}_b}}c_s\,.$$

Thus we see that before decoupling the baryon Jeans length is comparable to the Hubble length, and thus all scales for which our present discussion applies are sub-Jeans. Therefore, if baryon perturbations are adiabatic[21], they oscillate before decoupling[22].

After decoupling, the baryon component sees just its own pressure. This component is now a gas of hydrogen and helium. This gas is monatomic for the epoch we are now interested in. Hydrogen forms molecules only later. For a non-relativistic monatomic gas,

$$c_s^2 = \frac{5T_b}{3m}\,, \tag{205}$$

where we can take $m \approx 1\,\mathrm{GeV}$, since hydrogen dominates. Down to $z \sim 100$, residual free electrons maintain enough interaction between the baryon and photon components to keep $T_b \approx T_\gamma$. After that the baryon temperature falls faster,

$$T_b \propto (1+z)^2 \qquad \text{whereas} \qquad T_\gamma \propto 1+z \tag{206}$$

(as shown in an exercise in Chapter 4). For example, at $1+z = 1000$, soon after decoupling, $T_b = 2725\,\mathrm{K} = 0.2348\,\mathrm{eV}$ and the speed of sound is $c_s = 5930\,\mathrm{m/s}$. The baryon density is $\bar{\rho}_b = \Omega_b(1+z)^3\rho_{\rm cr} = \omega_b(1+z)^3\,1.88\times10^{-26}\,\mathrm{kg/m^3}$, and we get for the Jeans length

$$\lambda_J = (1+z)\frac{\sqrt{\pi}c_s}{\sqrt{G\bar{\rho}_b}} \tag{207}$$

that soon after decoupling

$$\lambda_J(1+z=1000) \;=\; \omega_b^{-1/2}\,0.96\times10^3\,\mathrm{pc} \;=\; \eta_{10}\,0.016\,\mathrm{Mpc} \;\sim\; 0.095\,\mathrm{Mpc}\,, \tag{208}$$

where $\eta_{10} \equiv 10^{10}\eta = 274\,\omega_b$ or $\omega_b = 0.00365\,\eta_{10}$, and the last number is for $\eta_{10} \sim 6$.

We define the baryon Jeans mass

$$M_J \equiv \bar{\rho}_{b0}\frac{\pi}{6}\lambda_J^3 \tag{209}$$

as the mass of baryonic matter within a sphere whose diameter is λ_J. Note that since λ_J is defined as a comoving distance, we must use here the present (mean) baryon density $\bar{\rho}_{b0}$. At $1+z = 1000$, the baryon Jeans mass is $\omega_b^{-1/2}\,1.3\times10^5\,\mathrm{M}_\odot = \eta_{10}^{-1/2}\,2.1\times10^6\,\mathrm{M}_\odot \sim 9\times10^5\,\mathrm{M}_\odot$ for $\eta_{10} \sim 6$. This corresponds to

[21] If there is an initial baryon entropy perturbation, i.e., a perturbation in baryon density without an accompanying 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 is generated, and the baryon perturbation begins to oscillate around its initial value (instead of oscillating around zero).

[22] We have not calculated this exactly, since all our calculations have been idealized, i.e., we have used perturbation theory which applies only to matter-dominated perturbations, and here we have ignored the CDM component. But this qualitative feature will hold also in the exact calculation, and this will be enough for us now.

the mass of a globular cluster and is much less than the mass of a galaxy. Thus, for our purposes, the baryonic component is pressureless after decoupling, i.e., baryon pressure can be ignored in the evolution of perturbations at cosmological scales (greater than $\sim 1\,\mathrm{Mpc}$). (The pressure cannot be ignored for smaller scale physics like the formation of individual galaxies.)

The baryon Jeans length after decoupling is $\ll$ Mpc. It would be relevant if we were interested in the process of the formation of individual galaxies, but here we are interested in the larger scales reflected in perturbations in the galaxy number density. Thus for our purposes, the baryonic component is pressureless after decoupling.

After decoupling, the evolution of the baryon density perturbation is governed by the gravitational effect of the dominant matter component, the CDM.

We now have the situation of Sec. 8.2.10, except that we have two matter components,

$$\rho = \rho_c + \rho_b + \rho_s \,, \tag{210}$$

where we approximate $\rho_s = \rho_\gamma + \rho_\nu + \rho_d$ as homogeneous. With the help of Sec. 8.2.8, the discussion is easy to generalize for the present case.

We can ignore the pressure of both ρ_b and ρ_c. Therefore their perturbation equations are

$$\ddot{\delta}_c + 2H\dot{\delta}_c = 4\pi G\bar{\rho}_m\delta \tag{211}$$

$$\ddot{\delta}_b + 2H\dot{\delta}_b = 4\pi G\bar{\rho}_m\delta \tag{212}$$

where $\bar{\rho}_m = \bar{\rho}_c + \bar{\rho}_b$ is the total background matter density and

$$\delta = \frac{\delta\rho_c + \delta\rho_b}{\bar{\rho}_c + \bar{\rho}_b} \tag{213}$$

is the total matter density perturbation.

We can now define the *baryon-CDM entropy perturbation*,

$$S_{cb} \equiv \delta_c - \delta_b \,, \tag{214}$$

which expresses how the perturbations in the two components deviate from each other. Subtracting Eq. (212) from (211) we get an equation for this entropy perturbation,

$$\ddot{S}_{cb} + 2H\dot{S}_{cb} = 0 \,. \tag{215}$$

We assume that the primordial perturbations were adiabatic, so that we had $\delta_b = \delta_c$, i.e, $S_{cb} = 0$ at horizon entry. For large scales, which enter the horizon after decoupling, an S_{cb} never develops, so the evolution of the baryon perturbations is the same as CDM perturbations.

But for scales which enter before decoupling, an S_{cb} develops because the baryon perturbation is then coupled to the photon perturbation, whereas the CDM perturbation is not. After decoupling, $\delta_b \ll \delta_c$, since δ_c has been growing, while δ_b has been oscillating. The initial condition for Eqs. (211,212,215) is then $S_{cb} \sim \delta_c$ ("initial" time here being the time of decoupling $t_{\rm dec}$). During the matter-dominated epoch, when $a \propto t^{2/3}$, so that $H = 2/3t$, the solution for S_{cb} is

$$S_{cb} = A + Bt^{-1/3} \,, \tag{216}$$

whereas for δ_c it is, neglecting the effect of baryons on it, from Eq. (126),

$$\delta_c = Ct^{2/3} + Dt^{-1} \sim Ct^{2/3} \,. \tag{217}$$

We call the first term the "growing" and the second term the "decaying" mode (although for S_{cb} the "growing" mode is actually just constant). For δ_c the growing and decaying modes have been growing and decaying since horizon entry, so we can now drop the decaying part of δ_c.

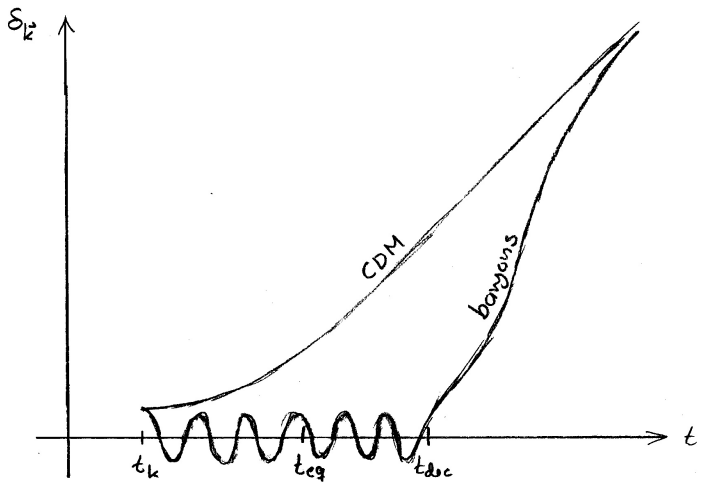


Figure 7: Evolution of the CDM and baryon density perturbations after horizon entry (at $t = t_k$). The figure is just schematic; the upper part is to be understood as having a $\sim$ logarithmic scale; the difference $\delta_c - \delta_b$ stays roughly constant, but the fractional difference becomes negligible as both δ_c and δ_b grow by a large factor.

To work out the precise initial conditions, we would need to work out the behavior of S_{cb} during decoupling. However, we really only need to assume that initially there is no strong cancellation between the growing and decaying modes in (216), so that $S_{cb} = \delta_c - \delta_b$ either shrinks or stays roughly constant near the initial value of δ_c. While δ_c grows by a large factor, δ_b must follow it to keep the difference close to the initial small value of δ_c, so that $\delta_b/\delta_c \to 1$.

Thus the baryon density contrast δ_b grows to match the CDM density contrast δ_c (see Fig. 7), and we have eventually $\delta_b = \delta_c = \delta$ to high accuracy.

The baryon density perturbation begins to grow only after $t_{\rm dec}$. Before decoupling the radiation pressure prevents it. Without CDM it would grow only as $\delta_b \propto a \propto t^{2/3}$ after decoupling (during the matter-dominated period; the growth stops when the universe becomes dark energy dominated). Thus it would have grown at most by the factor $a_0/a_{\rm dec} = 1 + z_{\rm dec} \sim 1100$ after decoupling. In the anisotropy of the CMB we observe the baryon density perturbations at $t = t_{\rm dec}$. They are too small (about 10^{-4}) for a growth factor of 1100 to give the present observed large scale structure[23].

With CDM this problem was solved. The CDM perturbations begin to grow earlier, at $t \sim t_{\rm eq}$, and by $t = t_{\rm dec}$ they are much larger than the baryon perturbations. After decoupling the baryons have lost the support from photon pressure and fall into the CDM gravitational potential wells, catching up with the CDM perturbations.

This allows the baryon perturbations to be small at $t = t_{\rm dec}$ and to grow after that by much more than the factor 10^3, matching observations. This is one of the reasons we are convinced that CDM exists.[24]

The whole subhorizon evolution history of all the different cosmological scales of perturbations is summarized by Fig. 8.

[23] This assumes adiabatic primordial perturbations, since we are seeing δ_γ, not δ_b. For a time, primordial baryon entropy perturbations $S_{b\gamma} = \delta_b - \frac{3}{4}\delta_\gamma$ were considered a possible explanation, but more accurate observations have ruled this model out.

[24] Historically, the above situation became clear in the 1980's when the upper limits to CMB anisotropy (which was finally discovered by COBE in 1992) became tighter and tighter. By today we have accurate detailed measurements of the structure of the CMB anisotropy which are compared to detailed calculations including the CDM so the argument is raised to a different level—instead of comparing just two numbers we are now comparing entire power spectra (to be discussed later).

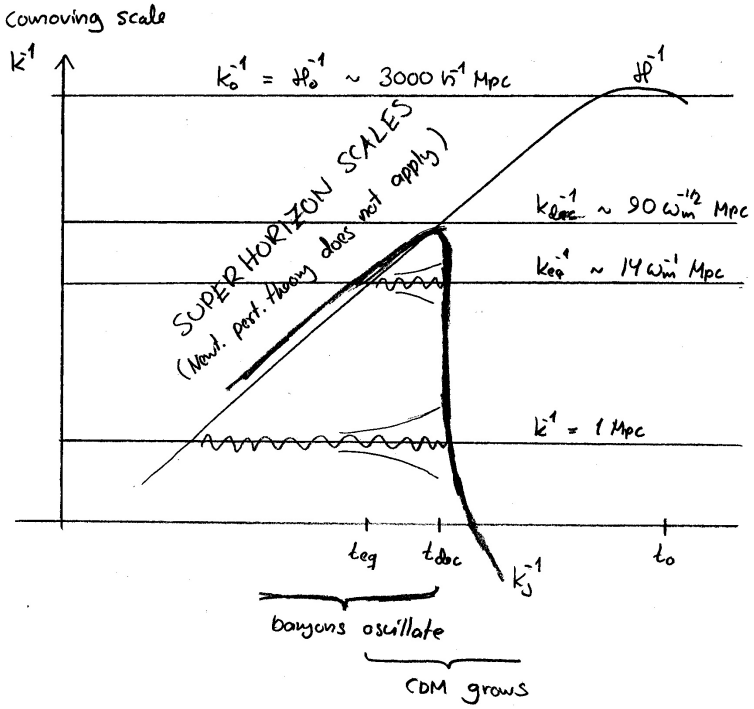


Figure 8: A figure summarizing the evolution of perturbations at different subhorizon scales. The baryon Jeans length k_J^{-1} drops precipitously at decoupling so that all cosmological scales became super-Jeans after decoupling, whereas all subhorizon scales were sub-Jeans before decoupling. The wavy lines symbolize the oscillation of baryon perturbations before decoupling, and the opening pair of lines around them symbolize the $\propto a$ growth of CDM perturbations after $t_{\rm eq}$. There is also an additional weaker (logarithmic) growth of CDM perturbations between horizon entry and $t_{\rm eq}$.

8.3.5 Late-time growth in the ΛCDM model

At late times, dark energy begins to accelerate the expansion, which will slow down the growth of the density perturbation. In the ΛCDM model dark energy is just a constant vacuum energy, so it has no perturbations and thus affects just the background. The perturbations are in CDM and baryons, and we can ignore the pressure term in the Jeans equation, since at such small scales where baryon pressure gradients would be important, first-order perturbation theory is not valid anyway at late times. Thus we are facing a similar calculation as we did in Sec. 8.2.11, the solution of Eq. (182),

$$\ddot{\delta}_{\mathbf{k}} + 2H\dot{\delta}_{\mathbf{k}} - 4\pi G\bar{\rho}_m\delta_{\mathbf{k}} = 0\,, \tag{218}$$

where $4\pi G\bar{\rho}_m = \frac{3}{2}\Omega_m H_0^2 a^{-3}$, with $\delta_b = \delta_c = \delta$, but instead of radiation we have now vacuum energy contributing to the background solution, which is the Concordance Model discussed in Cosmology I (Chapter 3):

$$a(t) \;=\; \left(\frac{\Omega_m}{\Omega_\Lambda}\right)^{1/3} \sinh^{2/3}\left(\tfrac{3}{2}\sqrt{\Omega_\Lambda}H_0 t\right)\,. \tag{219}$$

The Hubble parameter is given by

$$H = H_0\sqrt{\Omega_m a^{-3} + \Omega_\Lambda}\,. \tag{220}$$

Again, it is better to use the scale factor as time coordinate. The difference in the power of a in the behavior of the two density components is now 3 instead of 1, which makes the calculation more difficult. We follow here Dodelson[9]. After the change of variable from t to a,

(218) becomes (**exercise**)

$$\delta'' + \left(\frac{H'}{H} + \frac{3}{a}\right)\delta' - \frac{3\Omega_m}{2a^5}\left(\frac{H_0}{H}\right)^2 \delta = 0\,, \tag{221}$$

where $' \equiv d/da$. The decaying solution is

$$\delta \;\propto\; H \;\propto\; \sqrt{\Omega_m a^{-3} + \Omega_\Lambda} \tag{222}$$

and the growing solution is

$$\delta \;\propto\; H\int^a \frac{dx}{H^3x^3} \;\propto\; \sqrt{\Omega_m a^{-3} + \Omega_\Lambda}\int^a \frac{x^{3/2}dx}{\left(1+\frac{\Omega_\Lambda}{\Omega_m}x^3\right)^{3/2}} \tag{223}$$

The effect of changing the lower limit of integration can be incorporated in the decaying solution; so we can set the lower limit to 0. (Equation (221) is valid in general for matter perturbations with an additional smooth background component. The first forms of the solutions (222) and (223) are valid when the smooth component is vacuum energy or negative curvature.)

In the limit $a \ll 1$, or rather, $\Omega_\Lambda \ll \Omega_m a^{-3}$, the decaying solution becomes

$$\delta \propto a^{-3/2} \propto t^{-1} \tag{224}$$

and the growing solution becomes

$$\delta \propto a \propto t^{2/3} \tag{225}$$

the familiar results for the matter-dominated universe from Sec. 8.2.6. We can ignore the decaying mode, since it has become completely negligible when the vacuum energy begins to have an effect.

To fix the proportionality coefficient in the growing mode, we write it as

$$\delta = A\left(\Omega_m a^{-3} + \Omega_\Lambda\right)^{1/2}\int_0^a \frac{x^{3/2}dx}{\left(1+\frac{\Omega_\Lambda}{\Omega_m}x^3\right)^{3/2}} \tag{226}$$

and note that in the limit $\Omega_\Lambda \ll \Omega_m a^{-3}$ it becomes

$$\delta \;\approx\; A\Omega_m^{1/2}a^{-3/2}\int_0^a x^{3/2}dx \;=\; \frac{2}{5}\Omega_m^{1/2}Aa\,. \tag{227}$$

At $a = a_0 = 1$ this would give

$$\frac{2}{5}\Omega_m^{1/2}A \equiv \tilde{\delta} \quad\Rightarrow\quad A = \frac{5}{2}\Omega_m^{-1/2}\tilde{\delta}\,, \tag{228}$$

where we have defined $\tilde{\delta}$ as the value δ would have "now"[25] if there were no vacuum energy, i.e., the universe had stayed matter dominated.

[25]Note that we defined "now" as $a = a_0 = 1$, or in more physical terms as $T = T_0 = 2.7255\,\mathrm{K}$; not as $t = t_0$. The comparison situation ($\tilde{}$) we have in mind is that the early universe (where vacuum energy has no effect) is the same as in the ΛCDM model, but there is no vacuum energy to accelerate the expansion at late times, so that by "now" the expansion rate, i.e., H_0, is smaller than we observe in reality. The present matter density $\rho_{m0} = (3/8\pi G)\Omega_m H_0^2$ is the same as in the ΛCDM model, but $\tilde{\Omega}_m = 1$, so $\tilde{H}_0 = \Omega_m^{1/2}H_0$. The age of the universe is $\tilde{t}_0 = \frac{2}{3}\tilde{H}_0^{-1} = \frac{2}{3}\Omega_m^{-1/2}H_0^{-1}$, which for $h = 0.7$ and $\Omega_m = 0.3$ gives $\tilde{t}_0 = 17.0\times10^9$ years, instead of the $t_0 = 13.5\times10^9$ years of the ΛCDM model.

Thus we write (226) as

$$\boxed{\delta \;=\; \tilde{\delta}\,\frac{5}{2}\left(a^{-3}+\frac{\Omega_\Lambda}{\Omega_m}\right)^{1/2}\int_0^a \frac{x^{3/2}dx}{\left(1+\frac{\Omega_\Lambda}{\Omega_m}x^3\right)^{3/2}}\,.} \tag{229}$$

Unfortunately, the integral in (229) does not give an elementary function. (I think it is a so-called hypergeometric function, which does not give much useful information compared to just integrating (229) numerically.) We can see that at late (future) times, when $a \gg 1$, there is very little growth, since the factor outside the integral approaches a constant and for any $a_1 \gg 1$ and $a_2 \gg 1$, the contribution to the integral,

$$\int_{a_1}^{a_2} \frac{x^{3/2}dx}{\left(1+\frac{\Omega_\Lambda}{\Omega_m}x^3\right)^{3/2}} \;\approx\; \left(\frac{\Omega_m}{\Omega_\Lambda}\right)^{3/2}\int_{a_1}^{a_2} x^{-3}\,dx \;=\; \frac{1}{2}\left(\frac{\Omega_m}{\Omega_\Lambda}\right)^{3/2}\left(a_1^{-2}-a_2^{-2}\right) \tag{230}$$

is very small.

It turns out that the integral can be done if we extend it to the infinite future (**exercise**) : As $a \to \infty$,

$$\delta \;\to\; \delta(\infty) \;\equiv\; \tfrac{5}{2}\tilde{\delta}\left(a^{-3}+\frac{\Omega_\Lambda}{\Omega_m}\right)^{1/2}\int_0^\infty \frac{x^{3/2}dx}{\left(1+\frac{\Omega_\Lambda}{\Omega_m}x^3\right)^{3/2}} \;=\; \tfrac{5}{6}\tilde{\delta}\left(\frac{\Omega_m}{\Omega_\Lambda}\right)^{1/3} B(\tfrac{5}{6},\tfrac{2}{3})\,, \tag{231}$$

where

$$B(p,q) \;\equiv\; \int_0^1 t^{p-1}(1-t)^{q-1}dt \;=\; \frac{\Gamma(p)\Gamma(q)}{\Gamma(p+q)} \tag{232}$$

is the beta function and

$$B\left(\tfrac{5}{6},\tfrac{2}{3}\right) \approx 1.725\,. \tag{233}$$

Thus the perturbations "freeze", i.e., approach a final value

$$\delta(\infty) = 1.437\left(\frac{\Omega_m}{\Omega_\Lambda}\right)^{1/3}\tilde{\delta}\,. \tag{234}$$

which for $\Omega_m = 0.3$, $\Omega_\Lambda = 0.7$ gives

$$\delta(\infty) = 1.084\,\tilde{\delta}\,, \tag{235}$$

i.e., the perturbations will never become much stronger than what they in the matter-dominated model would be already "now". To get the present density perturbation $\delta(a=1)$ one has to do (229) numerically. This is done in Fig. 9, from which one can read that $\delta(a=1) \approx 0.78\,\tilde{\delta}$.

For perturbations that entered horizon well before matter-radiation equality $t_{\rm eq}$, we have from (199) that

$$\tilde{\delta} \approx \delta_{\rm prim}\left(1+\frac{3}{2a_{\rm eq}}\right)7.5\ln\left(0.17\frac{k}{k_{\rm eq}}\right)\,, \tag{236}$$

assuming that this is still $\ll 1$ so that first-order perturbation theory remains valid.

For $\Omega_m = 0.3$, $\Omega_\Lambda = 0.7$, $h = 0.7$, we have $k_{\rm eq}^{-1} = 65\,h^{-1}$Mpc and $a_{\rm eq} = 1/3514$. Equation (236) gives then for the scale $k^{-1} = 8\,h^{-1}$Mpc,

$$\tilde{\delta} \approx 13\,000\,\delta_{\rm prim} \qquad \text{and} \qquad \delta(a=1) \approx 10\,000\,\delta_{\rm prim}. \tag{237}$$

The contribution of the boost from radiation oscillation is a factor $7.5\ln(0.17k/k_{\rm eq}) \approx 2.4$ at this scale (and for smaller scales it is more). Actually, this scale is still too large (too low

k) for (236) to apply; in reality the factor is somewhat larger.[26] We chose the reference scale $k^{-1} = 8\,h^{-1}$Mpc, since observationally, the variance σ_T^2 of the top-hat-filtered density field of the galaxy distribution today is ≈ 1 at this scale. Because of the galaxy bias b_g, the corresponding variance for the matter distribution is less by a factor b_g^{-2}, but still not far from 1, meaning that the linear perturbation theory approximation is beginning to break. Because of nonlinear effects, the perturbation today is somewhat larger than our prediction from linear theory. We noted earlier that the perturbations entered the horizon with amplitude $\mathcal{P}(k)^{1/2} \approx \text{few} \times 10^{-5}$. Thus today the amplitude at $k = 1/8\,h^{-1}$Mpc should be somewhat more than few$\times 10^{-1}$. From Fig. 4 we see that σ_T^2 is typically a bit more that 2 times $\mathcal{P}$ at the same scale (depending on the shape of $\mathcal{P}(k)$). So indeed we get a prediction that it should be somewhat less than 1. We will do this comparison more quantitatively later.

8.3.6 Growth function

Inside the horizon, after photon decoupling the linear growth of matter perturbations is independent of scale (once the decaying mode has died out and ignoring the subcosmological scales where pressure gradients have a role). Thus it can be described by a function that depends on time (or scale factor, or redshift) only, called the *growth function*,

$$D(a) \equiv \frac{\delta(a)}{\delta_{\rm ref}} \tag{238}$$

where $\delta(a)$ is the density perturbation ($\delta_{\mathbf{k}}$ or $\delta(\mathbf{x})$; $D(a)$ is the same function for any $\mathbf{k}$ or $\mathbf{x}$) when scale factor is a and $\delta_{\rm ref}$ is it at some reference time. The choice of reference time fixes the normalization of D. During matter domination, $D(a) \propto a$ and a common normalization is to normalize so that $D(a) = a$ during matter domination. This corresponds to setting $\delta_{\rm ref} = \tilde{\delta}$. So that in the ΛCDM model

$$D(a) \;=\; \frac{5}{2}\left(a^{-3} + \frac{\Omega_\Lambda}{\Omega_m}\right)^{1/2} \int_0^a \frac{x^{3/2}dx}{\left(1 + \frac{\Omega_\Lambda}{\Omega_m}x^3\right)^{3/2}}\,, \tag{239}$$

(from the onset of matter domination).

We define the *growth rate*

$$f \equiv \frac{d\ln D}{d\ln a} = \frac{d\ln\delta}{d\ln a} = \frac{a}{\delta}\frac{d\delta}{da}\,, \tag{240}$$

which is independent of this normalization.

For the ΛCDM model of Sec. 8.3.5, we get from (229) (**exercise**)

$$f(a) \;=\; \frac{1}{1+\frac{\Omega_\Lambda}{\Omega_m}a^3}\left(\tfrac{5}{2}a\frac{\tilde{\delta}}{\delta} - \tfrac{3}{2}\right) \;=\; \frac{1}{1+\frac{\Omega_\Lambda}{\Omega_m}a^3}\left[\frac{a}{\left(a^{-3}+\frac{\Omega_\Lambda}{\Omega_m}\right)^{1/2}\int_0^a \frac{x^{3/2}dx}{\left(1+\frac{\Omega_\Lambda}{\Omega_m}x^3\right)^{3/2}}} - \frac{3}{2}\right]. \tag{241}$$

It turns out that a good approximation to (241) is

$$f(a) \approx \Omega_m(a)^\gamma\,, \qquad \text{where } \gamma = 0.55\,, \tag{242}$$

where γ is called the *growth index*. (We have assumed General Relativity, and the measurement of the growth index from galaxy surveys is a way of testing gravity theory.) We plot D, f, and the approximation (242) for ΛCDM in Fig. 9.

[26] One should instead use the BBKS transfer function (which still ignores effect of baryons) here, which gives a larger factor. On the other hand, baryonic effects decrease the result somewhat; but the net effect is a larger factor than 2.4. These corrections will be discussed in Sec. 8.4.4.

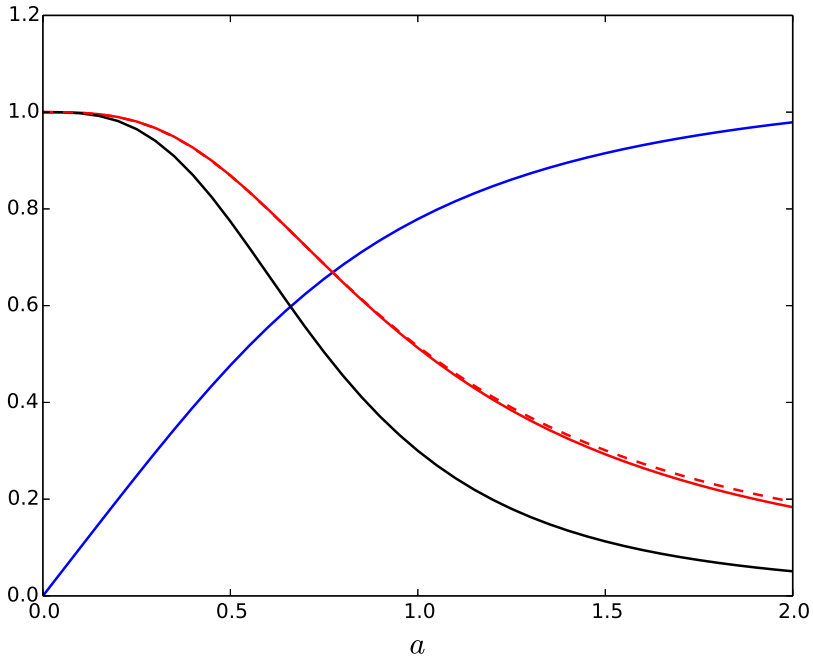


Figure 9: The growth function $D(a)$ (blue, with normalization $\delta_{\rm ref} = \tilde{\delta}$), matter density parameter $\Omega_m(a)$ (black), growth rate $f(a)$ (red), and the approximation (242) (red, dashed) for ΛCDM with $\Omega_m = 0.3$.

8.4 Relativistic perturbation theory

For scales comparable to, or larger than the Hubble scale, Newtonian perturbation theory does not apply, because we can no more ignore the curvature of spacetime. Therefore we need to use (general) relativistic perturbation theory. Instead of the Newtonian equations of gravity and fluid mechanics, the fundamental equation is now the *Einstein equation* of general relativity (GR). We assume a background solution, which is homogeneous and isotropic, i.e., a solution of the Friedmann equations, and study small perturbations around it. This particular choice of the background solution means that we are doing a particular version of relativistic perturbation theory, called *cosmological perturbation theory*.

The evolution of the perturbations while they are well outside the horizon is simple, but the mathematical machinery needed for its description is complicated. This is due to the coordinate freedom of general relativity. For the background solution we had a special coordinate system (time slicing) of choice, the one where the $t = const$ slices are homogeneous. The perturbed universe is no more homogeneous, it is just "close to homogeneous", and therefore we no more have a unique choice for the coordinate system. We should now choose a coordinate system where the universe is close to homogeneous on the time slices, but there are many different possibilities for such slicing. This freedom of choosing the coordinate system in the perturbed universe is called *gauge freedom*, and a particular choice is called a *gauge*.[27] The most important part of the choice of gauge is the choice of the time coordinate, because it determines the slicing of the spacetime into $t = const$ slices, "universe at time t". Sometimes the term 'gauge' is used to refer only to this slicing.

[27] If you are familiar with *gauge field theories*, like electrodynamics, the concept of 'gauge' may look different here. The mathematical similarity appears when the perturbation equations are developed. In relativistic perturbation theory gauge has this geometric origin (this is where the use of the word "gauge" comes from), unlike in electrodynamics.

Because the perturbations are defined in terms of the chosen coordinate system, they look different in different gauges. We can, for example, choose the gauge so that the perturbation in one scalar quantity, e.g., proper energy density, disappears, by choosing the $\rho = const$ 3-surfaces as the time slices (this is called "the uniform energy density gauge").

The true nature of gravitation is spacetime curvature, so perturbations should be described in terms of curvature.

We leave the actual development of cosmological perturbation theory to a more advanced course (Cosmological Perturbation Theory, lectured in spring 2020), and just summarize here some basic concepts and results.

In the Newtonian theory gravity was represented by a single function, the gravitational potential Φ. In GR, gravity is manifested in the geometry of spacetime, described in terms of the metric. Thus in addition to the density, pressure, and velocity perturbations, we have a perturbation in the metric. The perturbed metric tensor is

$$g_{\mu\nu} = \bar{g}_{\mu\nu} + \delta g_{\mu\nu} \, . \tag{243}$$

For the background metric, $\bar{g}_{\mu\nu}$, we choose that of the flat Friedmann-Robertson-Walker universe,

$$ds^2 = \bar{g}_{\mu\nu} dx^\mu dx^\nu = -dt^2 + a(t)^2 (dx^2 + dy^2 + dz^2) \tag{244}$$

The restriction to the flat case is an important simplification, because it allows us to Fourier expand our perturbations in terms of plane waves.[28] Fortunately the real universe appears to be flat, or at least close to it. And earlier it was even flatter. Inflation predicts a very flat universe.

For the metric perturbation, we have now 10 functions $\delta g_{\mu\nu}(t, \mathbf{x})$. So there appears to be ten degrees of freedom. Four of them are not physical degrees of freedom, since they just correspond to our freedom in choosing the four coordinates. So there are 6 real degrees of freedom.

Two of these metric degrees of freedom couple to density and pressure perturbations and the irrotational velocity perturbation. These are the *scalar perturbations*. Two couple to the rotational velocity perturbation to make up the *vector perturbations*. The remaining two are not coupled to the cosmic fluid at all[29], and are called *tensor perturbations*. They are gravitational waves, which do not exist in Newtonian theory.

The vector perturbations decay in time, and are not produced by inflation, so they are the least interesting.

Although the tensor perturbations also are not related to growth of structure, they are produced in inflation and affect the cosmic microwave background (CMB) anisotropy and polarization. Different inflation models produce tensor perturbations with different amplitudes and spectral indices (to be explained later), so they are an important diagnostic of inflation. No tensor perturbations have been detected in the CMB so far, but they could be detected in the future with more sensitive instruments if their amplitude is large enough.[30]

The three kinds of perturbations evolve independently of each other in linear perturbation theory, so they can be studied separately. We shall first concentrate on the scalar perturbations, returning to the tensor perturbations later.

[28] In the Newtonian case this restriction was not necessary, and we could apply it to any Friedmann model, as there is no curvature of spacetime in the Newtonian view, and only the expansion law $a(t)$ of the Friedmann model is used. The Newtonian theory of course is only valid for small scales were the curvature can indeed be ignored.

[29] This is true in first-order perturbation theory in the perfect fluid approximation, but not in general.

[30] Typically, large-field inflation models produce tensor perturbations with much larger amplitude than small-field inflation models. In the latter case they are likely to be too small to be detectable.

8.4.1 Gauges for scalar perturbations

Consider now scalar perturbations. The gauges discussed in the following assume scalar perturbations.

Perturbations appear different in different gauges. When needed, we use superscripts to indicate in which gauge the quantity is defined: C for the comoving gauge and N for the Newtonian gauge. Some other gauges are the synchronous gauge (S), spatially flat gauge (Q), and the uniform energy density gauge (U).

There are two common ways to specify a gauge, i.e., the choice of coordinate system in the perturbed universe:

- A statement about the relation of the coordinate system to the fluid perturbation. This will lead to some condition on the metric perturbation.
- A statement about the metric perturbation. This will then lead to some condition on the coordinate system.

The two gauges (C and N) we shall refer to in the following, give an example of each.

The *comoving gauge* is defined so that the space coordinate lines $\mathbf{x} = const$ follow fluid flow lines, and the time slice, the $t = const$ hypersurface, is orthogonal to them. Thus the velocity perturbation is zero in this gauge,

$$v^C = 0\,. \tag{245}$$

The *conformal-Newtonian* gauge, also called the longitudinal gauge, or the zero-shear gauge, and sometimes, for short, just the Newtonian gauge, is defined by requiring the metric to be of the form

$$ds^2 = -(1+2\Phi)dt^2 + a^2(1-2\Psi)(dx^2+dy^2+dz^2)\,. \tag{246}$$

This means that we require

$$\delta g_{0i} = 0, \quad \delta g_{11} = \delta g_{22} = \delta g_{33}, \quad \text{and} \quad g_{ij} = 0 \quad \text{for} \quad i \neq j\,. \tag{247}$$

(This is possible for scalar perturbations). The two metric perturbations, $\Phi(t,\mathbf{x})$ and $\Psi(t,\mathbf{x})$ are called *Bardeen potentials*.[31] Φ is also called the Newtonian potential, since in the Newtonian limit ($k \gg \mathcal{H}$ and $p \ll \rho$), it becomes equal to the Newtonian gravitational potential perturbation. Thus we can use the same symbol for it. Ψ is also called the Newtonian curvature perturbation, because it determines the curvature of the 3-dimensional $t = const$ subspaces, which are flat in the unperturbed universe (since it is the flat FRW universe).

It turns out that the difference $\Phi - \Psi$ is caused only by anisotropic stress (or anisotropic pressure). We shall here consider only the case of a perfect fluid. For a perfect fluid the pressure (or stress) is necessary isotropic. Thus we have only a single metric perturbation[32]

$$\Psi = \Phi \tag{248}$$

The density perturbations in these two gauges become equal in the limit $k \gg \mathcal{H}$ (inside horizon), and we can then identify them with the "usual" density perturbation δ of Newtonian theory.

[31] Warning: The sign conventions for Ψ differ, and many authors call them Ψ and Φ instead.

[32] In reality, neutrinos develop anisotropic pressure after neutrino decoupling. Therefore the two Bardeen potentials actually differ from each other by about 10 % between the times of 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 happens to photons after photon decoupling, but the universe is then already matter-dominated, so they do not cause a significant $\Psi - \Phi$ difference.

8.4.2 Evolution at superhorizon scales

When the perturbations are outside the horizon (meaning that the wavelength of the Fourier mode we are considering is much longer than the Hubble length), very little happens to them, and we can find quantities which remain constant for superhorizon scales. Such a quantity is the (comoving gauge) curvature perturbation $\mathcal{R}(\mathbf{x})$, which describes how curved is the $t = const$ slice in the comoving gauge.[33] *For adiabatic perturbations, the curvature perturbation $\mathcal{R}$ stays constant in time outside the horizon.*

Using gauge transformation equations $\mathcal{R}$ can be related to the metric in the Newtonian gauge. The result is

$$\mathcal{R} \;=\; -\frac{5+3w}{3+3w}\Phi - \frac{2}{3+3w}H^{-1}\dot{\Phi}\,, \tag{251}$$

where $w \equiv \bar{p}/\bar{\rho}$.

Because $\mathcal{R}_\mathbf{k}$ stays constant while $k \ll \mathcal{H}$, it is a very useful quantity for "carrying" the perturbations from their generation at horizon exit during inflation to horizon entry at later times. We now define *the primordial perturbation* to refer to the perturbation at the epoch when it is well outside the horizon. For adiabatic perturbations, the primordial perturbation is completely characterized by the set of these constant values $\mathcal{R}_\mathbf{k}$. We shall later discuss how the primordial perturbation is generated by inflation, and how these superhorizon values $\mathcal{R}_\mathbf{k}$ are determined by it.

However, we would like to describe the perturbation in more "familiar" terms, the gravitational potential perturbation Φ and the density perturbation δ. While $\mathcal{R}_k$ remains constant this turns out to be easy. Eq. (251) can be written as a differential equation for $\Phi_\mathbf{k}$,

$$\frac{2}{3}H^{-1}\dot{\Phi}_\mathbf{k} + \frac{5+3w}{3}\Phi_\mathbf{k} \;=\; -(1+w)\mathcal{R}_\mathbf{k}\,. \tag{252}$$

During any period, when also $w = const$, the solution of this equation is

$$\Phi_\mathbf{k} \;=\; -\frac{3+3w}{5+3w}\mathcal{R}_\mathbf{k} + \text{a decaying part}\,. \tag{253}$$

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

$$\boxed{\Phi_\mathbf{k} = -\frac{3+3w}{5+3w}\mathcal{R}_\mathbf{k}} \qquad (w = const)\,. \tag{254}$$

In particular, we have the relations

$$\Phi_\mathbf{k} \;=\; -\frac{2}{3}\mathcal{R}_\mathbf{k} \qquad (\text{rad.dom},\ w = \tfrac{1}{3}) \tag{255}$$

$$\Phi_\mathbf{k} \;=\; -\frac{3}{5}\mathcal{R}_\mathbf{k} \qquad (\text{mat.dom},\ w = 0)\,. \tag{256}$$

[33] Technically, $\mathcal{R}$ is defined in terms of the trace of the space part of the comoving gauge metric perturbation ($-\Psi$ is the corresponding quantity in the Newtonian gauge), and it is related to the scalar curvature ${}^{(3)}R^C$ of the *comoving gauge* time slice (the ${}^{(3)}$ reminds us that we are considering a 3-dimensional subspace, and the C refers to the comoving gauge) so that

$${}^{(3)}R^C = -4a^{-2}\nabla^2\mathcal{R}\,. \tag{249}$$

For Fourier components we have then that

$$\mathcal{R}_\mathbf{k} \equiv \frac{1}{4}\left(\frac{a}{k}\right)^2 {}^{(3)}R^C_\mathbf{k}\,. \tag{250}$$

Another similar quantity is the (uniform-density-gauge) curvature perturbation ζ that is defined the same way, but for the uniform-density-gauge time slice. For superhorizon scales they are equal, $\mathcal{R} = \zeta$ (in the limit $k \ll \mathcal{H}$).

8.4.3 From outside to inside horizon

After the perturbation has entered horizon, we can use the Newtonian perturbation theory result, Eq. (109), which gives the density perturbation as

$$\delta_{\mathbf{k}} \;=\; -\left(\frac{k}{a}\right)^2 \frac{\Phi_{\mathbf{k}}}{4\pi G\bar{\rho}} \;=\; -\frac{2}{3}\left(\frac{k}{aH}\right)^2 \Phi_{\mathbf{k}} \;=\; -\frac{2}{3}\left(\frac{k}{\mathcal{H}}\right)^2 \Phi_{\mathbf{k}}\,, \tag{257}$$

where we used the background relation

$$H^2 = \frac{8\pi G}{3}\bar{\rho} \quad \Rightarrow \quad \boxed{4\pi G\bar{\rho} = \frac{3}{2}H^2}\,. \tag{258}$$

The problem is to get $\Phi_{\mathbf{k}}$ from the superhorizon epoch where it is constant (as long as $w = const$) through the horizon entry to the subhorizon epoch where it evolves according to Newtonian theory. We do this for the two cases, large ($k \ll k_{\rm eq}$) and small ($k \gg k_{\rm eq}$) scales, below.

Large scales. For scales k which enter while the universe is matter dominated, this is easy, since in this case Φ_k stays constant the whole time (until dark energy becomes important). Thus we can relate these constant values of $\Phi_{\mathbf{k}}$, and the corresponding subhorizon density perturbations $\delta_{\mathbf{k}}$ during the matter-dominated epoch to the primordial perturbations $\mathcal{R}_{\mathbf{k}}$ by

$$\begin{aligned} \Phi_{\mathbf{k}} &= -\frac{3}{5}\mathcal{R}_{\mathbf{k}} \qquad \text{(mat.dom)} \\ \delta_{\mathbf{k}} &= -\frac{2}{3}\left(\frac{k}{\mathcal{H}}\right)^2 \Phi_{\mathbf{k}} \;=\; \frac{2}{5}\left(\frac{k}{\mathcal{H}}\right)^2 \mathcal{R}_{\mathbf{k}} \;\propto\; \frac{1}{(aH)^2} \propto t^{2/3} \propto a \end{aligned} \tag{259}$$

Note that by $\mathcal{R}_{\mathbf{k}}$ we refer always to the constant primordial value, when we use it in equations, like (259), that give other quantities at later times.

Small scales. For perturbations which enter during the radiation-dominated epoch, the potential $\Phi_{\mathbf{k}}$ does not stay constant. We learned earlier, that in this case the radiation density perturbation oscillates with roughly constant amplitude, which means that the amplitude for the potential Φ must decay $\propto a^2\bar{\rho} \propto a^{-2}$. This oscillation applies to the baryon-photon fluid, whereas the CDM density perturbation grows slowly. After the universe becomes matter dominated, it is these CDM perturbations that matter.

We shall now make a crude estimate how the amplitudes of these smaller-scale perturbations during the matter-dominated epoch are related to the primordial perturbations (in particular, we ignore the slow groth of the CDM perturbation during tehradiation-dominated epoch). These perturbations enter during the radiation-dominated epoch. Assume that the relation $\Phi_{\mathbf{k}} = -\frac{2}{3}\mathcal{R}_{\mathbf{k}}$ holds all the way to horizon entry ($k = \mathcal{H}$). Assume then that the Newtonian relation (257) holds already. Then

$$\delta_{\mathbf{k}} \;\approx\; -\frac{2}{3}\left(\frac{k}{\mathcal{H}}\right)^2 \Phi_{\mathbf{k}} \;=\; -\frac{2}{3}\Phi_k \approx \frac{4}{9}\mathcal{R}_{\mathbf{k}} \tag{260}$$

at horizon entry. The universe is now radiation-dominated, and therefore $\delta_{r\mathbf{k}} = \delta_{\mathbf{k}}$. We are assuming primordial adiabatic perturbations and therefore the adiabatic relations $\delta_c = \frac{3}{4}\delta_r$, $\delta_\gamma = \delta_r$ hold at superhorizon scales. Assume that these relations hold until horizon entry. After that $\delta_{\gamma\mathbf{k}}$ begins to oscillate, whereas $\delta_{c\mathbf{k}}$ grows slowly. Thus we have that at horizon entry

$$\delta_{c\mathbf{k}} \approx \frac{3}{4}\delta_{\mathbf{k}} \approx \frac{1}{3}\mathcal{R}_{\mathbf{k}}\,. \tag{261}$$

Ignoring the slow growth of δ_c we get that $\delta_{c\mathbf{k}}$ stays at this value until the universe becomes matter-dominated at $t = t_{\rm eq}$, after which we can approximate $\delta_{\mathbf{k}} \approx \delta_{c\mathbf{k}}$ and $\delta_{\mathbf{k}}$ begins to grow according to the matter-dominated law, $\propto a \propto 1/\mathcal{H}^2$.

Thus

$$\delta_{\mathbf{k}}(t_{\rm eq}) \approx \frac{1}{3}\mathcal{R}_{\mathbf{k}} \tag{262}$$

and

$$\delta_{\mathbf{k}}(t) \approx \frac{1}{3}\mathcal{R}_{\mathbf{k}} \left(\frac{\mathcal{H}_{\rm eq}}{\mathcal{H}}\right)^2 = \frac{1}{3}\mathcal{R}_{\mathbf{k}} \left(\frac{k_{\rm eq}}{\mathcal{H}}\right)^2 \qquad \text{for } t > t_{\rm eq}\,, \tag{263}$$

as long as the universe stays matter dominated.

8.4.4 Transfer function

For large scales ($k \ll k_{\rm eq}$) which enter the horizon during the matter-dominated epoch, we got

$$\delta_{\mathbf{k}}(t) = \frac{2}{5}\left(\frac{k}{\mathcal{H}}\right)^2 \mathcal{R}_{\mathbf{k}} \qquad (k \ll k_{\rm eq})\,, \tag{264}$$

for as long as the universe stays matter dominated.

This is a simple result, and we use this as a reference for the more complicated result at smaller scales. That is, we define a *transfer function* $T(k,t)$ so that

$$\delta_{\mathbf{k}}(t) = \frac{2}{5}\left(\frac{k}{\mathcal{H}}\right)^2 T(k,t)\mathcal{R}_{\mathbf{k}} \tag{265}$$

where $\mathcal{R}_{\mathbf{k}}$ refers to the primordial perturbation. Thus by definition $T(k,t) = 1$ for $k \ll k_{\rm eq}$.[34]

Using the rough estimate from the previous subsection we get, comparing (263) to (259), that

$$T(k,t) \approx \frac{5}{6}\left(\frac{k_{\rm eq}}{k}\right)^2 \tag{266}$$

during the matter-dominated epoch, where we can drop the factor $\frac{5}{6}$, since this is anyway just a rough estimate.

Once we are well into the matter-dominated era, perturbations at all scales grow $\propto a \propto 1/(aH)^2$ and the transfer function becomes independent of time,[35]

$$\begin{aligned} T(k) &= 1 && k \ll k_{\rm eq} \\ T(k) &\sim \left(\frac{k_{\rm eq}}{k}\right)^2 && k \gg k_{\rm eq} \end{aligned} \tag{267}$$

A more accurate calculation, including the gravitational effect of radiation perturbation oscillations on the CDM perturbation (see Sec. 8.3.3), assuming adiabatic primordial perturbations and still ignoring baryons, adds a logarithmic growth factor, the ratio between (199) and the $\delta = \delta_{\rm prim}(a/a_{\rm eq})$ of (263), and gives[36]

$$T(k) \;\approx\; 11.3\left(\frac{k_{\rm eq}}{k}\right)^2 \ln\left(\frac{0.17k}{k_{\rm eq}}\right) \qquad k \gg k_{\rm eq}\,, \tag{269}$$

[34] With the given definition for $T(k,t)$, this holds for $t \ll t_0$, i.e, before we entered the present dark-energy-dominated epoch.

[35] We shall later define other transfer functions, but this is *the* transfer function $T(k)$ of structure formation theory. It relates the perturbations inside the horizon during the matter-dominated epoch to the primordial perturbations, and it is independent of time.

[36] This calculation is presented in Dodelson[9] (Sections 7.3 and 7.4). Dodelson (7.69) gives the result as

$$T(k) \;=\; \frac{12k_{\rm eq}^2}{k^2}\ln\left(\frac{k}{8k_{\rm eq}}\right)\,. \tag{268}$$

For some reason I get the somewhat different numerical factors in (269) when I do the same calculation.

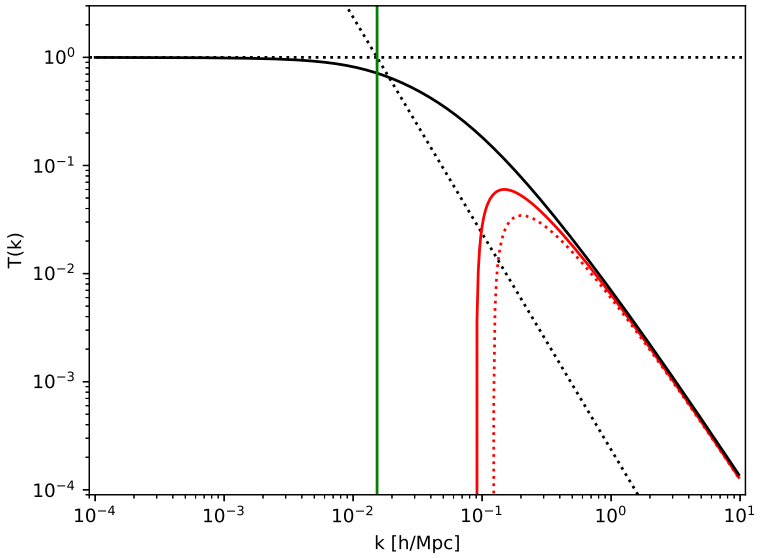


Figure 10: Transfer function $T(k)$ for CDM, adiabatic primordial fluctuations. The black curve is the BBKS transfer function (270), the red curve is the small-scale approximate analytical result (269) (the dotted red curve is the Dodelson version (268)), the two black dotted lines correspond to (267), and the green vertical line gives $k = k_{\rm eq}$. The k scale is for our reference model, $\Omega_m = 0.3$, $h = 0.7$, for which $k_{\rm eq} = 0.0153\,h/{\rm Mpc} = 1/(65\,h^{-1}{\rm Mpc})$.

where we approximated $1 + a/a_{\rm eq} \approx a/a_{\rm eq}$ (appropriate for application at late times, $a \gg a_{\rm eq}$). Note that that logarithm is negative for $k \lesssim 6k_{\rm eq}$; the equation is not supposed to apply yet for this low k.

To include the intermediate scales, which enter close matter-radiation equality, requires numerical computation. For $\omega_b \ll \omega_c$ (i.e., still essentially ignoring baryons), Bardeen, Bond, Kaiser, and Szalay [10] gave a fitting formula, the BBKS transfer function

$$T(k) \;=\; \frac{\ln(1+2.34q)}{2.34q}\,\frac{1}{[1+3.89q+(16.1q)^2+(5.64q)^3+(6.71q)^4]^{1/4}}\,, \qquad (270)$$

where $q = 0.073(k/k_{\rm eq})$, to such numerical results. See Fig. 10 for these results. The slope of the BBKS transfer function is

$$\frac{d\ln T}{d\ln q} \;=\; \frac{2.34q}{(1+2.34q)\ln(1+2.34q)} - \frac{1}{4}\frac{3.89q+2(16.1q)^2+3(5.64q)^3+4(6.71q)^4}{1+3.89q+(16.1q)^2+(5.64q)^3+(6.71q)^4} - 1\,. \qquad (271)$$

For later reference, we note that for our reference model, $\Omega_m = 0.3$, $h = 0.7$, this gives $d\ln T/d\ln q = -1.184$ at $k = 1/(8\,h^{-1}{\rm Mpc})$.

According to present understanding, the universe becomes dark energy dominated as we approach the present time. The equation-of-state parameter w begins to decrease (becomes negative) and therefore Φ begins to change again. The growth of the density perturbations is slowed down as we saw in Secs. 8.3.5 and 8.3.6. Since this affect all scales the same way, we can model this with the growth function $D(a)$, and keep the transfer function $T(k)$ unaffected.

We are still missing the effect of baryons. There are publicly available computer programs (such as CMBFAST, CAMB[37], and CLASS[38]; you give your favorite values for the cosmological

[37] `https://camb.info/`
[38] `http://class-code.net/`

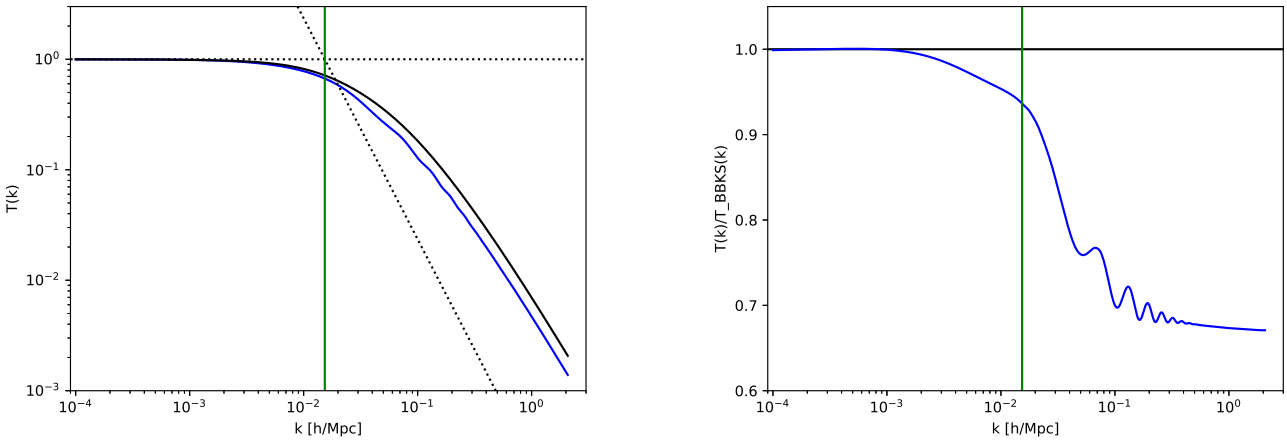


Figure 11: Left: Transfer function $T(k)$ calculated with CAMB (blue curve) for adiabatic primordial fluctuations in the flat ΛCDM model with $\omega_b = 0.023$, $\omega_c = 0.124$, $h = 0.7$ (so that $\Omega_m = 0.3$), and massless neutrinos (a neutrino mass 0.06 eV for one neutrino species changes the transfer function by less than the width of the curve). The black curve is the BBKS transfer function (270), the two black dotted lines correspond to (267), and the green vertical line gives $k = k_{\rm eq}$, as in Fig. 10. The main difference from BBKS is due to baryons. Right: The ratio (blue) of the $T(k)$ from CAMB to the BBKS transfer function.

parameters as input) that include it and other small physical effects we have ignored. They represent the current state of the art. The exact result can be given in form of the transfer function $T(k)$ we defined above. We show in Fig. 11 a transfer function calculated with CAMB. The effect of *baryon acoustic oscillations* (i.e., the oscillations of $\delta_{b\gamma}$ before decoupling, which leave a trace in δ_b) shows up as a small-amplitude wavy pattern in the $k > k_{\rm eq}$ part of the transfer function, since different modes k were at a different phase of the oscillation when that ended around $t_{\rm dec}$.

Everything has still been calculated using linear perturbation theory. In linear perturbation theory one gets the present-day density power spectrum by multiplying the primordial power spectrum by (the square of) the transfer function and growth function.

Linear perturbation theory breaks down when the perturbations become large, $\delta(\mathbf{x}) \sim 1$. We say that the perturbation becomes nonlinear. This has happened for the smaller scales, $k^{-1} < 10$ Mpc by now. Nonlinear effects speed up the growth of density perturbations. They cannot be captured in a transfer function and a growth function, since now the ratio between the present-day and primordial power spectra depends on the primordial power spectrum. CAMB can also calculate nonlinear effects but within a more restricted set of cosmological models, because this is based on results from N-body simulations.

When the perturbation becomes sufficiently nonlinear, i.e., an overdense region becomes significantly denser (say, twice as dense) as the average density of the universe, it collapses and forms a gravitationally bound structure, e.g. a galaxy or a cluster of galaxies. Further collapse is prevented by the angular momentum of the structure. Galaxies in a cluster and stars (and CDM particles) in a galaxy orbit around the center of mass of the bound structure.

8.4.5 Tensor perturbations

In addition to scalar and vector perturbations, in general relativistic perturbation theory we have tensor perturbations. They have the nice property that we do not have to worry about different gauges, since they are gauge invariant in the sense that, if we first do a gauge transformation and then separate out the scalar, vector, and tensor parts, the tensor part has remained unchanged.

These are perturbations of the metric that for one Fourier mode take the form

$$\begin{aligned} ds^2 &= -dt^2 + a(t)^2\left[(1+h)dx^2 + (1-h)dy^2 + dz^2\right] \\ &= a(\eta)^2\left[-d\eta^2 + (1+h)dx^2 + (1-h)dy^2 + dz^2\right] \end{aligned} \tag{272}$$

where

$$h = h_{\mathbf{k}}(t)e^{ikz} \tag{273}$$

is the perturbation and η is conformal time. In (272) we have chosen the z axis in the direction of the wave vector, so that $\mathbf{k} = k\hat{k}$ and $\mathbf{k}\cdot\mathbf{x} = kz$. Since the metric is a real quantity, in (272) and (278) h should be interpreted as the real part of h; like one should always do when one makes physical interpretations for a single Fourier mode. Remember that when one sums over Fourier components the imaginary parts of $h_{\mathbf{k}}(t)e^{ikz} + h_{-\mathbf{k}}(t)e^{-ikz}$ cancel since $h_{-\mathbf{k}} = h^*_{\mathbf{k}}$, and thus the imaginary parts have no physical significance, they are just a mathematical convenience.

The effect of the tensor perturbation is to stretch space in one direction (here x if h is positive) and compress it in the other direction (here y) orthogonal to the wave vector of the Fourier mode. In (272) we also chose the orientation of the x and y axes so that they correspond to these stretch/compress directions. But of course the perturbation could be oriented differently. We get the other possibilities by rotating the pattern around the wave vector $\mathbf{k}$ by some angle φ, which is mathematically equivalent to rotating the coordinate system by angle $-\varphi$.

In matrix form the metric is

$$[g_{\mu\nu}] = a^2\begin{bmatrix} -1 & & & \\ & 1+h & & \\ & & 1-h & \\ & & & 1 \end{bmatrix} \tag{274}$$

After rotation by φ around the z axis it becomes

$$[g_{\mu\nu}] = a^2\begin{bmatrix} 1 & & & \\ & \cos\varphi & -\sin\varphi & \\ & \sin\varphi & \cos\varphi & \\ & & & 1 \end{bmatrix}\begin{bmatrix} -1 & & & \\ & 1+h & & \\ & & 1-h & \\ & & & 1 \end{bmatrix}\begin{bmatrix} 1 & & & \\ & \cos\varphi & \sin\varphi & \\ & -\sin\varphi & \cos\varphi & \\ & & & 1 \end{bmatrix} \tag{275}$$

Rotation by 45°, i.e., $\cos\varphi = \sin\varphi = 1/\sqrt{2}$, gives

$$[g_{\mu\nu}] = a^2\begin{bmatrix} -1 & & & \\ & 1 & h & \\ & h & 1 & \\ & & & 1 \end{bmatrix} \tag{276}$$

We call (274) the + mode and (276) the × mode. An arbitrary orientation of the stretch/compress pattern can be obtained as a linear combination of these two modes, so that the general form of the tensor perturbation is

$$[g_{\mu\nu}] = a^2\begin{bmatrix} -1 & & & \\ & 1+h_+ & h_\times & \\ & h_\times & 1-h_+ & \\ & & & 1 \end{bmatrix} \tag{277}$$

or

$$\begin{aligned} ds^2 &= -dt^2 + a(t)^2 \left[(1+h_+)dx^2 + 2h_\times dxdy + (1-h_+)dy^2 + dz^2\right] \\ &= a(\eta)^2 \left[-d\eta^2 + (1+h_+)dx^2 + 2h_\times dxdy + (1-h_+)dy^2 + dz^2\right] \end{aligned} \tag{278}$$

for a Fourier mode in the z direction. Thus we have two Fourier amplitudes $h_{+\mathbf{k}}(t)$ and $h_{\times\mathbf{k}}(t)$ for each wave vector $\mathbf{k}$. In the following we mostly write just $h(t)$ to represent an arbitrary such mode.

The evolution equation for $h(t)$,

$$\ddot{h} + 3H\dot{h} + \left(\frac{k}{a}\right)^2 h = 0 \qquad \Leftrightarrow \qquad H^{-2}\ddot{h} + 3H^{-1}\dot{h} + (k/\mathcal{H})^2\, h = 0\,, \tag{279}$$

can be obtained from the Einstein equation. This derivation is beyond the level of this course, but the equation has a simple and plausible form: it is the wave equation with a damping term $3H\dot{h}$; the wave velocity is the speed of light = 1.

For superhorizon scales we can ignore the last term, and we get $h = const$ as a solution and another solution where $\dot{h} \equiv dh/dt \propto a^{-3}$ so it also approaches a constant. Thus tensor perturbations remain essentially constant outside the horizon.

For evolution inside the horizon we get oscillatory solutions and then it is better to work with conformal time. The $h(\eta)$ evolution equation is

$$h'' + 2\mathcal{H}h' + k^2 h = 0 \qquad \Leftrightarrow \qquad \mathcal{H}^{-2}h'' + 2\mathcal{H}^{-1}h' + (k/\mathcal{H})^2\, h = 0\,, \tag{280}$$

where $' \equiv d/d\eta$. If we first ignore the middle term, we get solutions of the form $h \propto e^{\pm ik\eta}$, where $-$ represents a wave moving in the $\mathbf{k}$ direction and $+$ in the $-\mathbf{k}$ direction. These are *gravitational waves*. They propagate at the speed of light and they are transverse waves. During one half-period of the wave oscillation, space is stretched in one direction orthogonal to the direction of propagation, and compressed in the other orthogonal direction. During the next half-period the opposite happens. The amplitude of the stretching is given by h, meaning that the maximum stretching is by factor $1 + |h|$ and the maximum compression is by factor $1 - |h|$.

The middle term in (280) represents the damping of gravitational terms due to the expansion of the universe. Write

$$h(\eta) = A(\eta)e^{-ik\eta} \tag{281}$$

and insert this into (280) to get

$$A'' + 2\mathcal{H}A' - 2ik(A' + \mathcal{H}A) = 0\,. \tag{282}$$

For $k \gg \mathcal{H}$, the part $2ik(A' + \mathcal{H}A)$ dominates the left-hand side, and we get

$$A' + \mathcal{H}A \;=\; A' + \frac{a'}{a}A \;=\; \frac{1}{a}(aA)' \;=\; 0 \quad \Rightarrow \quad aA = const \quad \Rightarrow \quad A \propto a^{-1}\,. \tag{283}$$

Thus gravitational waves are damped inside the horizon as a^{-1} independent of the expansion law.

For simple expansion laws one can also solve Eq. (280) exactly, covering also horizon entry/exit. These solutions are Bessel functions.

8.5 Nonlinear growth

When δ grows the evolution becomes nonlinear, requiring a more complicated discussion. One can get further with higher-order perturbation theory or something called the Zeldovich approximation, but eventually one has to resort to numerical simulations. We shall not discuss these in this course. The spherically symmetric special case can be done analytically by basing it on solutions for FRW universes with different densities. We do it below for an overdensity in a flat matter-dominated background universe.

8.5.1 Closed Friedmann model

In Cosmology I we derived the expansion law for the closed ($\Omega > 1$) matter-dominated FRW universe. It cannot be given in closed form as $a(t)$, but can be given in terms of an auxiliary variable, the *development angle* ψ, as

$$\begin{aligned} a(\psi) &= a_i \frac{\Omega_i}{2(\Omega_i - 1)}(1 - \cos\psi) = a(\psi)\frac{\Omega(\psi)}{2[\Omega(\psi) - 1]}(1 - \cos\psi) \\ t(\psi) &= H_i^{-1}\frac{\Omega_i}{2(\Omega_i - 1)^{3/2}}(\psi - \sin\psi) = H(\psi)^{-1}\frac{\Omega(\psi)}{2[\Omega(\psi) - 1]^{3/2}}(\psi - \sin\psi)\,, \end{aligned} \tag{284}$$

where a_i, Ω_i, and H_i are the scale factor, density parameter, and Hubble parameter at some reference time t_i (usually chosen as the present time t_0, but below we will instead choose t_i to be some early time, when Ω is still very close to 1). In the second forms we took advantage of the fact that we can choose t_i to be any time during the development and replaced it with the "current" time. See Fig. 12 for the shape of $a(t)$. This curve is called a *cycloid*. (It is the path made by a point at the rim of a wheel.) From (284) we solve

$$\Omega(\psi) = \frac{2}{1 + \cos\psi}\,. \tag{285}$$

Calculating $da/dt = da/d\psi \times d\psi/dt$ we find (**exercise**)

$$H(\psi) = 2H_i\frac{(\Omega_i - 1)^{3/2}}{\Omega_i}\frac{\sin\psi}{(1 - \cos\psi)^2}\,. \tag{286}$$

The matter density is given by

$$\rho(\psi) = \rho_i\left(\frac{a_i}{a(\psi)}\right)^3 = 8\rho_i\frac{(\Omega_i - 1)^3}{\Omega_i^3(1 - \cos\psi)^3}\,. \tag{287}$$

The scale factor reaches a maximum $a_{\rm ta}$ (and the density a minimum) at the "turnaround" time $t_{\rm ta}$, when $\psi = \pi$, so that

$$a_{\rm ta} = a_i\frac{\Omega_i}{\Omega_i - 1}\,, \qquad t_{\rm ta} = \frac{\pi}{2}H_i^{-1}\frac{\Omega_i}{(\Omega_i - 1)^{3/2}}\,, \qquad \text{and} \qquad \rho(t_{\rm ta}) = \rho_i\frac{(\Omega_i - 1)^3}{\Omega_i^3}\,. \tag{288}$$

At this point $H = 0$ and then the universe begins to shrink. Since

$$\rho_i = \frac{3\Omega_i H_i^2}{8\pi G} \qquad \text{we have} \qquad \rho(t_{\rm ta}) = \frac{3\pi}{32Gt_{\rm ta}^2}\,. \tag{289}$$

The universe collapses at $t_{\rm coll} = 2t_{\rm ta}$, when $\psi = 2\pi$ and $a = 0$ again.

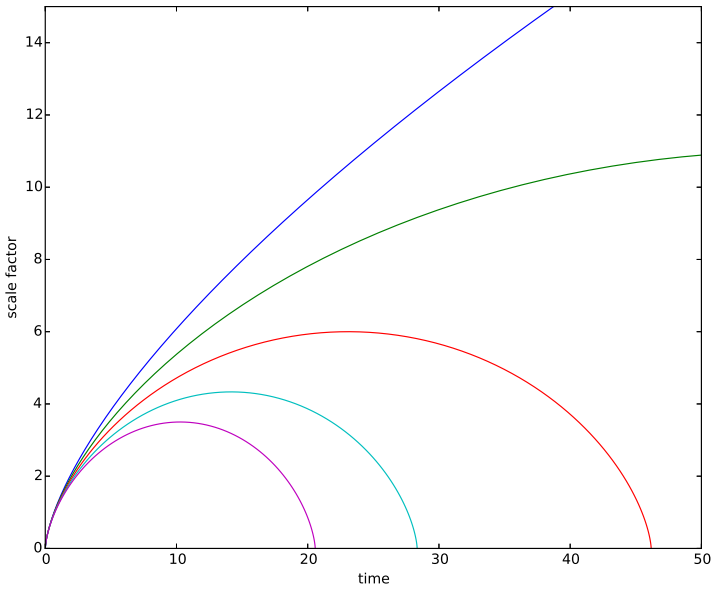


Figure 12: The expansion law for the flat matter-dominated universe (blue) and for closed matter-dominated universes with different initial values $\Omega_i > 1$ for the density parameter. Both axes are linear, the units are arbitrary.

8.5.2 Spherical collapse

The expansion law (284) will hold also for a spherically symmetric overdense region within a flat ($\Omega = 1$) matter-dominated FRW universe. Denote the quantities for this flat background universe by $\bar{a}, \bar{H}, \bar{\rho}$. (Time t is the same for both solutions and $\bar{\Omega} = 1$, so we don't need notations for them.) The background universe has

$$\bar{H}^2 = \frac{8\pi G}{3}\bar{\rho} = \left(\frac{2}{3t}\right)^2 \quad \Rightarrow \quad \bar{\rho} = \frac{1}{6\pi G t^2} \tag{290}$$

Thus we see that at $t_{\rm ta}$, the density of the overdense region is

$$\rho(t_{\rm ta}) = \frac{9\pi^2}{16}\bar{\rho}(t_{\rm ta}) \approx 5.5517\bar{\rho}(t_{\rm ta})\,, \tag{291}$$

i.e., at the turnaround time the density contrast has the value

$$\delta_{\rm ta} = \frac{9\pi^2}{16} - 1 \approx 4.5517\,. \tag{292}$$

Until then the overdense region has been expanding, although slower than the surrounding background universe. At turnaround the overdense region begins to shrink (in terms of proper distance).

The preceding applies both for an overdense region with homogeneous density and for one with a spherically symmetric density profile. In the latter case, we have to apply it separately for each spherical shell, and the density ρ refers, not to the density of the shell, but to the mean density within the shell, as it is the total mass within the shell that is responsible for the gravity affecting the expansion or contraction of the shell. To avoid shell crossing the density profile

has to decrease outward, so that outer shells do not collapse before inner shells.[39]

In linear perturbation theory, which applies when $\delta \ll 1$, density perturbations in the flat matter-dominated universe grow as

$$\delta^{\rm lin} \propto a \propto t^{2/3} \,. \tag{293}$$

When the density contrast δ becomes large it begins to grow faster. Compare now the linear growth law to the above result for δ at turnaround.

The initial density contrast δ_i is given by $\rho_i = (1+\delta_i)\bar\rho_i$. On the other hand

$$\bar H_i^2 = \frac{8\pi G}{3}\bar\rho_i \qquad \text{and} \qquad H_i^2 = \frac{8\pi G}{3}\Omega_i\rho_i \tag{294}$$

so that

$$1+\delta_i = \Omega_i \frac{H_i^2}{\bar H_i^2} \qquad \text{or at any time} \qquad 1+\delta = \Omega\frac{H^2}{\bar H^2} \,. \tag{295}$$

Thus the density contrast is not simply given by $\Omega - \bar\Omega = \Omega - 1$, since also the Hubble parameters are different for the two solutions. We can sort out the separate contributions from $\Omega_i - 1$ and $(H_i/\bar H_i)^2$ at an early time when $\Omega - 1 \ll 1$ and $\psi \ll 1$, by expanding Ω, H and $\bar H$ from (285), (286) and (290&284) in terms of ψ (**exercise**) to get

$$\Omega_i \approx 1+\tfrac{1}{4}\psi_i^2 \qquad \text{and} \qquad \frac{H_i^2}{\bar H_i^2} \approx 1-\tfrac{1}{10}\psi^2 \quad \Rightarrow \quad 1+\delta_i \approx 1+\tfrac{3}{20}\psi^2 \quad \Rightarrow \quad \delta_i \approx \tfrac{3}{5}(\Omega_i - 1) \,. \tag{296}$$

We can now give the linear prediction for the density contrast at turnaround time[40]:

$$\delta_{\rm ta}^{\rm lin} \;=\; \frac{\bar a_{\rm ta}}{\bar a_i}\delta_i \;=\; \left(\frac{t_{\rm ta}}{t_i}\right)^{2/3}\delta_i \;\approx\; \left(\frac{3\pi}{4}\right)^{2/3}\frac{\delta_i}{\Omega_i - 1} \;\approx\; \frac{3}{5}\left(\frac{3\pi}{4}\right)^{2/3} \;\approx\; 1.0624 \,, \tag{297}$$

where we approximated

$$t_{\rm ta} \;\approx\; \frac{\pi}{2}\bar H_i^{-1}\frac{1}{(\Omega_i - 1)^{3/2}} \qquad \text{and} \qquad t_i = \tfrac{2}{3}\bar H_1^{-1} \,. \tag{298}$$

Thus we conclude that density perturbations begin to collapse when the linear prediction is $\delta \sim 1$, at which time the true density perturbation is already over 4 times stronger.

The collapse is completed at $t_{\rm coll} = 2t_{\rm ta}$, when the linear prediction gives

$$\delta_{\rm coll}^{\rm lin} \;=\; 2^{2/3}\delta_{\rm ta}^{\rm lin} \;\approx\; 1.6865 \,. \tag{299}$$

The above special case can be extended to the situation where the background universe is a closed or open Friedmann model (i.e., a matter-dominated FRW universe), and to the ΛCDM model, with more complicated math.

8.5.3 Without spherical symmetry

I suppose these idealized cases would lead to a supermassive black hole at the center of symmetry (for perturbations at cosmological scales, for a smaller scale perturbation we might end up with a star). In reality overdensities are never exactly spherically symmetric. The deviation from spherical symmetry increases as the collapse progresses. For an ellipsoidal overdensity the flattest direction collapses first leading first to a "Zeldovich pancake", and the second flattest

[39] More precisely, the density of an outer shell must not be more than the mean density inside it. We should also include in our model an underdense region around our overdense region so that their combined mean density equals that of the background universe, so as not to affect the evolution of the surroundings.

[40] Note that Kolb&Turner[11], p. 328, misses the factor 3/5 .

next leading then to an elongated structure. In the situation where the density refers to a number density of galaxies instead of a smooth continuous density, the galaxies will pass the center point at various distances (instead of colliding at the center as in the perfectly spherically symmetric case), after which they will move away from the center and will be decelerated, eventually falling back in and ending up orbiting the center, forming a cluster of galaxies.

For the real universe the different distance scales are in a different stage of the collapse. The largest distance scales are still "falling in", leading to flattened structures at the largest scales and elongated structures, "filaments", at somewhat smaller scales. These structures surround rounder underdense regions, "voids". Smaller scales have already collapsed into galaxy clusters.

8.6 Perturbations during inflation

So far we have developed perturbation theory describing the substance filling the universe in fluid terms, i.e., giving the perturbations in terms of $\delta\rho$ and δp. During inflation the universe is dominated by a scalar field, the inflaton φ, so it is better to give the perturbation directly as a perturbation in the inflaton field,

$$\varphi(t,\mathbf{x}) = \bar{\varphi}(t) + \delta\varphi(t,\mathbf{x})\,. \tag{300}$$

8.6.1 Evolution of inflaton perturbations

In Minkowski space the field equation for a scalar field is

$$\ddot{\varphi} - \nabla^2\varphi + V'(\varphi) = 0\,. \tag{301}$$

In the flat Friedmann-Robertson-Walker universe (the background universe) the field equation is

$$\ddot{\varphi} + 3H\dot{\varphi} - a^{-2}\nabla^2\varphi + V'(\varphi) = 0\,. \tag{302}$$

(Here $\nabla = \nabla_{\mathbf{x}}$, i.e., with respect to the comoving coordinates $\mathbf{x}$, and therefore the factor $1/a$ appears in front of it.)

We ignore for the moment the perturbation in the spacetime metric and just insert (300) into Eq. (302),

$$(\bar{\varphi}+\delta\varphi)^{\cdot\cdot} + 3H(\bar{\varphi}+\delta\varphi)^{\cdot} - a^{-2}\nabla^2(\bar{\varphi}+\delta\varphi) + V'(\bar{\varphi}+\delta\varphi) = 0\,. \tag{303}$$

Here $V'(\bar{\varphi}+\delta\varphi) = V'(\bar{\varphi}) + V''(\bar{\varphi})\delta\varphi$ and $\bar{\varphi}(t)$ is the homogeneous background solution from our earlier discussion of inflation. Thus $\nabla^2\bar{\varphi} = 0$, and $\bar{\varphi}$ satisfies the background equation

$$\ddot{\bar{\varphi}} + 3H\dot{\bar{\varphi}} + V'(\bar{\varphi}) = 0\,. \tag{304}$$

Subtracting the background equation from the full equation (303) we get the perturbation equation

$$\delta\ddot{\varphi} + 3H\delta\dot{\varphi} - a^{-2}\nabla^2\delta\varphi + V''(\bar{\varphi})\delta\varphi = 0 \tag{305}$$

In Fourier space we have

$$\delta\ddot{\varphi}_{\mathbf{k}} + 3H\delta\dot{\varphi}_{\mathbf{k}} + \left[\left(\frac{k}{a}\right)^2 + m^2(\bar{\varphi})\right]\delta\varphi_{\mathbf{k}} = 0\,, \tag{306}$$

or

$$H^{-2}\delta\ddot{\varphi}_{\mathbf{k}} + 3H^{-1}\delta\dot{\varphi}_{\mathbf{k}} + \left[\left(\frac{k}{aH}\right)^2 + \frac{m^2}{H^2}\right]\delta\varphi_{\mathbf{k}} = 0\,, \tag{307}$$

where

$$m^2(\bar{\varphi}) \equiv V''(\bar{\varphi})\,. \tag{308}$$

During inflation, H and m^2 change slowly. Thus we make now an approximation where we treat them as constants. If the slow-roll approximation is valid, $m^2 \ll H^2$, since

$$\frac{m^2}{H^2} = 3M_{\rm Pl}^2\frac{V''}{V} = 3\eta \ll 1\,. \tag{309}$$

Thus we can ignore the m^2/H^2 in Eq. (307)[41]. The general solution becomes then

$$\delta\varphi_{\mathbf{k}}(t) = A_{\mathbf{k}}w_k(t) + B_{\mathbf{k}}w_k^*(t)\,, \tag{312}$$

[41] The general solution to (306), when H and m^2 are constants, is

$$\delta\varphi_{\mathbf{k}}(t) = a^{-3/2}\left[A_{\mathbf{k}}J_{-\nu}\left(\frac{k}{aH}\right) + B_{\mathbf{k}}J_{\nu}\left(\frac{k}{aH}\right)\right]\,, \tag{310}$$

where

$$w_k(t) = \left(i + \frac{k}{aH}\right) \exp\left(\frac{ik}{aH}\right) . \tag{313}$$

(**Exercise:** Show that this is a solution of (306) when $H = const$ and $m^2 = 0$.) The time dependence of (312) is in

$$a = a(t) \propto e^{Ht} . \tag{314}$$

Well before horizon exit, $k \gg aH$, the argument of the exponent is large. As $a(t)$ increases the solution oscillates rapidly and its amplitude is damped. *After horizon exit*, $k \ll aH$, the solution stops oscillating and *approaches the constant value* $i(A_{\mathbf{k}} - B_{\mathbf{k}})$.

We have cheated by ignoring the metric perturbation. We should use GR and write the curved-spacetime field equation using the perturbed metric. Perturbations in a scalar field couple only to scalar perturbations, so we need to consider scalar perturbations only. For example, in the conformal-Newtonian gauge the correct perturbation equation is

$$\delta\ddot{\varphi}^N_{\mathbf{k}} + 3H\delta\dot{\varphi}^N_{\mathbf{k}} + \left[\left(\frac{k}{a}\right)^2 + V''(\bar{\varphi})\right]\delta\varphi^N_{\mathbf{k}} \;=\; -2\Phi_{\mathbf{k}}V(\bar{\varphi}) + \left(\dot{\Phi}_{\mathbf{k}} + 3\dot{\Psi}_{\mathbf{k}}\right)\dot{\bar{\varphi}} . \tag{315}$$

That is, there are additional terms which are first order in the metric and zeroth order (background) in the scalar field φ.

Fortunately, it is possible to choose the gauge so that the terms with the metric perturbations are negligible *during inflation*[42], and the previous calculation applies in such a gauge. The comoving gauge is not such a gauge, so a gauge transformation is required to obtain the comoving gauge curvature perturbation $\mathcal{R}$. Gauge transformations are beyond the scope of these lectures, but the result is

$$\mathcal{R} = -H\frac{\delta\varphi}{\dot{\bar{\varphi}}} . \tag{316}$$

Thus it is clear what we want from inflation. We want to find the inflaton perturbations $\delta\varphi_{\mathbf{k}}$ some time after horizon exit. We can use the constant value the solution (312) approaches after horizon exit. Then Eq. (316) gives us $\mathcal{R}_k$, which remains constant while the scale k is outside the horizon, and is indeed the primordial $\mathcal{R}_{\mathbf{k}}$ discussed in the previous section. And then we can use the results of Sec. 8.4 to get $\delta_{\mathbf{k}}$.

We are still missing the initial conditions for the solution (312). These are determined by quantum fluctuations, which we shall discuss in Sec. 8.6.3. Quantum fluctuations produce the initial conditions in a random manner, so that we can predict only their statistical properties. It turns out that the quantum fluctuations are a *Gaussian process*, a term which specifies certain statistical properties, which we shall discuss next before returning to the application to inflaton fluctuations.

8.6.2 Statistical properties of Gaussian perturbations

The statistical (Gaussian) nature of the inflaton perturbations $\delta\varphi(\mathbf{x})$ are inherited later by other perturbations, which depend linearly on them. Let us therefore discuss a generic Gaussian

where J_ν is the Bessel function of order ν and

$$\nu = \sqrt{\frac{9}{4} - \frac{m^2}{H^2}} . \tag{311}$$

With $m^2 = 0$, $\nu = \frac{3}{2}$. Bessel functions of half-integer order are spherical Bessel functions which can be expressed in terms of trigonometric functions, or $e^{\pm ikx}$.

[42] One such gauge is the *spatially flat gauge* Q. For scalar perturbations it is possible to choose the time coordinate so that the time slices have Euclidean geometry. This leads to the spatially flat gauge. (There are still perturbations in the spacetime curvature; they show up when one considers the time direction).

perturbation

$$g(\mathbf{x}) = \sum_{\mathbf{k}} g_{\mathbf{k}} e^{i\mathbf{k}\cdot\mathbf{x}}\,, \tag{317}$$

where the set of Fourier coefficients $\{g_{\mathbf{k}}\}$ is a result of a *statistically homogeneous and isotropic Gaussian random process*. We assume $g(\mathbf{x})$ is real, so that $g_{-\mathbf{k}} = g^*_{\mathbf{k}}$. We write $g_{\mathbf{k}}$ in terms of its real and imaginary part,

$$g_{\mathbf{k}} = \alpha_{\mathbf{k}} + i\beta_{\mathbf{k}}\,. \tag{318}$$

For Fourier analysis of statistically homogeneous and isotropic random perturbations, see sections (8.1.1, 8.1.3, 8.1.4), where the probability distribution was treated as unknown. The new ingredient (in addition to the assumption that the perturbations are small, allowing the use of first-order perturbation theory, which we introduced in Sec. 8.2), is that the probability distribution is known to be Gaussian. This means that

$$\begin{aligned}\mathrm{Prob}(g_{\mathbf{k}}) &= \frac{1}{2\pi s^2_{\mathbf{k}}} \exp\left(-\frac{1}{2}\frac{|g_{\mathbf{k}}|^2}{s^2_{\mathbf{k}}}\right) \\ &= \frac{1}{\sqrt{2\pi} s_{\mathbf{k}}} \exp\left(-\frac{1}{2}\frac{\alpha^2_{\mathbf{k}}}{s^2_{\mathbf{k}}}\right) \times \frac{1}{\sqrt{2\pi} s_{\mathbf{k}}} \exp\left(-\frac{1}{2}\frac{\beta^2_{\mathbf{k}}}{s^2_{\mathbf{k}}}\right)\,,\end{aligned} \tag{319}$$

i.e., the real and imaginary parts are independent Gaussian random variables[43] with equal variance $s^2_{\mathbf{k}}$.

The *expectation value* of a quantity which depends on $g_{\mathbf{k}}$ as $f(g_{\mathbf{k}})$ is given by

$$\langle f(g_{\mathbf{k}})\rangle \equiv \int f(g_{\mathbf{k}})\mathrm{Prob}(g_{\mathbf{k}}) d\alpha_{\mathbf{k}} d\beta_{\mathbf{k}}\,, \tag{320}$$

where the integral is over the complex plane, i.e.,

$$\int_{-\infty}^{\infty} d\alpha_{\mathbf{k}} \int_{-\infty}^{\infty} d\beta_{\mathbf{k}}\,.$$

We immediately get (**exercise**) the *mean*

$$\langle g_{\mathbf{k}}\rangle = 0 \tag{321}$$

and *variance*

$$\langle |g_{\mathbf{k}}|^2\rangle = 2s^2_{\mathbf{k}} = \langle \alpha^2_{\mathbf{k}} + \beta^2_{\mathbf{k}}\rangle \tag{322}$$

of $g_{\mathbf{k}}$.

The distribution has one free parameter, the real positive number $s_{\mathbf{k}}$ which gives the width (determines the variance) of the distribution. From statistical isotropy and homogeneity follows that $s_{\mathbf{k}} = s(k)$ and

$$\langle g^*_{\mathbf{k}} g_{\mathbf{k}'}\rangle = 0 \qquad \text{for} \quad \mathbf{k} \neq \mathbf{k}'\,. \tag{323}$$

We can combine Eqs. (322) and (323) into a single equation,

$$\langle g^*_{\mathbf{k}} g_{\mathbf{k}'}\rangle \;=\; 2\delta_{\mathbf{k}\mathbf{k}'} s^2_{\mathbf{k}} \;=\; \delta_{\mathbf{k}\mathbf{k}'}\langle |g_{\mathbf{k}}|^2\rangle \;\equiv\; \frac{\delta_{\mathbf{k}\mathbf{k}'}}{V} P(k) \;\equiv\; \frac{2\pi^2\delta_{\mathbf{k}\mathbf{k}'}}{Vk^3}\mathcal{P}(k)\,, \tag{324}$$

where

$$\mathcal{P}(k) \;\equiv\; \left(\frac{L}{2\pi}\right)^3 4\pi k^3 \langle |g_{\mathbf{k}}|^2\rangle \;=\; \frac{V}{2\pi^2} k^3 \langle |g_{\mathbf{k}}|^2\rangle\,, \tag{325}$$

[43] $g_{\mathbf{k}}$ is a complex Gaussian random variable and $\alpha_{\mathbf{k}}$ and $\beta_{\mathbf{k}}$ are real Gaussian random variables.

which gives the dependence of the variance of $g_{\mathbf{k}}$ on the wave number k, is the *power spectrum* of g.

Going back to coordinate space, we find

$$\langle g(\mathbf{x})\rangle = \left\langle \sum_{\mathbf{k}} g_{\mathbf{k}} e^{i\mathbf{k}\cdot\mathbf{x}} \right\rangle = \sum_{\mathbf{k}} \langle g_{\mathbf{k}}\rangle e^{i\mathbf{k}\cdot\mathbf{x}} = 0 \tag{326}$$

The square of the perturbation can be written as

$$g(\mathbf{x})^2 = \sum_{\mathbf{k}} g^*_{\mathbf{k}} e^{-i\mathbf{k}\cdot\mathbf{x}} \sum_{\mathbf{k}'} g_{\mathbf{k}'} e^{i\mathbf{k}'\cdot\mathbf{x}} \tag{327}$$

since $g(\mathbf{x})$ is real. The typical amplitude of the perturbation is described by the variance, the expectation value of this square,

$$\begin{aligned} \langle g(\mathbf{x})^2\rangle &= \sum_{\mathbf{k}\mathbf{k}'} \langle g^*_{\mathbf{k}} g_{\mathbf{k}'}\rangle e^{i(\mathbf{k}'-\mathbf{k})\cdot\mathbf{x}} = \sum_{\mathbf{k}} \langle |g_{\mathbf{k}}|^2\rangle = 2\sum_{\mathbf{k}} s^2_{\mathbf{k}}\,. = \left(\frac{2\pi}{L}\right)^3 \sum_{\mathbf{k}} \frac{1}{4\pi k^3}\mathcal{P}(k) \\ &\to \frac{1}{4\pi}\int \frac{d^3k}{k^3}\mathcal{P}(k) = \int_0^\infty \frac{dk}{k}\mathcal{P}(k) = \int_{-\infty}^\infty \mathcal{P}(k) d\ln k\,. \end{aligned} \tag{328}$$

Note that there is no $\mathbf{x}$-dependence in the result, since this is an expectation value. $g(\mathbf{x})^2$ of course varies from place to place, but its expectation value from the random process is the same everywhere—the perturbed universe is statistically homogeneous. Thus the power spectrum of g gives the contribution of a logarithmic scale interval to the variance of $g(\mathbf{x})$.

An alternative definition for the power spectrum is

$$P(k) \equiv V\langle |g_{\mathbf{k}}|^2\rangle \tag{329}$$

While this definition is simpler, the result for the variance of $g(\mathbf{x})$ in terms of it and thus the interpretation is more complicated. Because of the common use of this latter definition, we shall make reference to both power spectra, and distinguish them by the different typeface. They are related by

$$P(k) = \frac{2\pi^2}{k^3}\mathcal{P}(k)\,. \tag{330}$$

For Gaussian perturbations, the power spectrum gives a complete statistical description. All statistical quantities can be calculated from it. In particular, (**exercise**)

$$\langle |g_{\mathbf{k}}|^4\rangle = 2\langle |g_{\mathbf{k}}|^2\rangle^2\,. \tag{331}$$

Exercise: Show that, if α is a real Gaussian random variable, with $\langle\alpha\rangle = 0$, then

$$\langle |\alpha|^4\rangle = 3\langle |\alpha|^2\rangle^2,$$

and that, if g is a complex Gaussian random variable (real and imaginary parts independent of each other), with $\langle g\rangle = 0$, then

$$\langle |g|^4\rangle = 2\langle |g|^2\rangle^2.$$

For a single realization, define $\hat{P}(\mathbf{k}) \equiv V|g_{\mathbf{k}}|^2$. From (331)

$$\langle \hat{P}(\mathbf{k})^2\rangle = 2P(k)^2\,. \tag{332}$$

The typical deviation of $\hat{P}(\mathbf{k})$ from its expectation value is given by the square root of the variance

$$\langle |\hat{P}(\mathbf{k}) - P(k)|^2\rangle \equiv \langle \hat{P}(\mathbf{k})^2\rangle - P(k)^2 = P(k)^2\,, \tag{333}$$

which (the last equality) only holds for Gaussian perturbations. For a single mode $\mathbf{k}$, this variance is large, but we can define

$$\hat{P}(k) \equiv \frac{V}{N_k} \sum_{\mathbf{k}} |g_{\mathbf{k}}|^2 \,, \tag{334}$$

where the sum is over all $\mathbf{k}$ for which $k - \frac{1}{2}\Delta k < |\mathbf{k}| \leq k + \frac{1}{2}\Delta k$, where Δk is the width of a k-bin (a shell in $\mathbf{k}$-space) over which we average, and N_k is the number of Fourier modes $\mathbf{k}$ in the bin. We then get (**exercise**)

$$\langle |\hat{P}(k) - P(k)|^2 \rangle \;=\; \frac{2}{N_k} P(k)^2 \,. \tag{335}$$

This is an example of the cosmic variance discussed in Sec. 8.1.1: the power spectrum $\hat{P}(k)$ measured from a finite volume V deviates from its expectation value $P(k)$, and thus we can measure $P(k)$ only with finite accuracy. The estimate $\hat{P}(k)$ is an average over N_k modes and its variance is reduced by the factor $N_k/2$ (since $g_{-\mathbf{k}} = g_{\mathbf{k}}^*$ we have $N_k/2$ independent modes, i.e., $N_k/2$ independent complex random variables or N_k independent real random variables.) The density of $\mathbf{k}$ modes in $\mathbf{k}$-space goes as $1/V$, so the larger the survey volume V, the more modes we have in a $\mathbf{k}$-bin (the larger os N_k). Also, for higher k, there are more modes in a k-bin: a given survey volume samples small scales better than large scales.

Exercise: Derive Eq. (335). Note that the $g_{\mathbf{k}}$ are independent variables so that $\langle |g_{\mathbf{k}}|^2 |g_{\mathbf{k}'}|^2 \rangle = \langle |g_{\mathbf{k}}|^2 \rangle \langle |g_{\mathbf{k}'}|^2 \rangle$ for $\mathbf{k}' \neq \pm\mathbf{k}$, but since $g_{-\mathbf{k}} = g_{\mathbf{k}}^*$, $\langle |g_{\mathbf{k}}|^2 |g_{\mathbf{k}'}|^2 \rangle = \langle |g_{\mathbf{k}}|^4 \rangle$.

It can be shown (under weak assumptions about the power spectrum), that statistically homogeneous and isotropic Gaussian perturbations are ergodic, so that we do not need to make a separate assumption of ergodicity.[44]

8.6.3 Generation of primordial perturbations from inflation

Subhorizon scales during inflation are microscopic[45] and therefore quantum effects are important. Thus we should study the inflaton field using quantum field theory.

This goes beyond the level of this course, so we have relegated the discussion into an appendix. The basic idea is that for scales that are inside horizon there are quantum fluctuations, called *vacuum fluctuations*, in the inflaton field. For a homogeneous inflaton field, the Fourier amplitudes $\delta\varphi_{\mathbf{k}}$ of its perturbations would be identically zero, but analogous to a quantum harmonic oscillator, it is not possible for them to stay there, but instead they fluctuate around this value.

We saw in Sec. 8.6.1 that the classical solutions to the evolution of $\delta\varphi_{\mathbf{k}}$ reach a constant value after horizon exit (in the approximation $H = const$ during horizon exit). The quantum treatment gives that at this stage we can neglect further quantum fluctuations and treat $\delta\varphi_{\mathbf{k}}$ classically—the fluctuations "freeze".

The final result is that well after horizon exit, $k \ll \mathcal{H}$, the Fourier amplitudes $\delta\varphi_{\mathbf{k}}$ have acquired a power spectrum

$$\mathcal{P}_\varphi(k) \;\equiv\; V \frac{k^3}{2\pi^2} \langle |\delta\varphi_{\mathbf{k}}|^2 \rangle \;=\; \left(\frac{H}{2\pi}\right)^2 . \tag{336}$$

[44] Liddle & Lyth [6], in Sec. 4.3.3, make this claim but do not provide a proof.

[45] We later give an upper limit to the inflation energy scale, i.e., $V(\varphi)$ at the time cosmological scales exited the horizon, $V^{1/4} < 1.9 \times 10^{16}$ GeV. From $H^2 = V(\varphi)/3M_{\rm Pl}^2$ we have $H < 9 \times 10^{13}$ GeV or for the Hubble length $H^{-1} > 2.3 \times 10^{-30}$ m. This is a lower limit to the horizon size, but it is not expected to be very many orders of magnitude larger.

After this we can ignore further quantum effects and treat the later evolution of the inflaton field, both the background and the perturbation, classically. The effect of the vacuum fluctuations was to produce "out of nothing" the perturbations $\delta\varphi_{\mathbf{k}}$. We can't predict their individual values; their production from quantum fluctuations is a random process. We can only calculate their statistical properties. Closer investigation reveals that this is a Gaussian random process. All $\delta\varphi_{\mathbf{k}}$ acquire their values as independent random variables (except for the reality condition $\delta\varphi_{-\mathbf{k}} = \delta\varphi_{\mathbf{k}}^*$) with a Gaussian probability distribution. Thus all statistical information is contained in the power spectrum $\mathcal{P}_\varphi(k)$.

The result (336) was obtained treating H as a constant. However, over long time scales, H does change. The main purpose of the preceding discussion was to follow the inflaton perturbations through the horizon exit. After the perturbation is well outside the horizon, we switch to other variables, namely the curvature perturbation $\mathcal{R}_{\mathbf{k}}$, which, unlike $\delta\varphi_{\mathbf{k}}$, remains constant outside the horizon, even though H changes. Therefore we have to use for each scale k a value of H which is representative for the evolution of that particular scale through the horizon. That is, we choose the value of H at horizon exit,[46] so that $aH = k$. Thus we write our power spectrum result as

$$\boxed{\mathcal{P}_\varphi(k) \;=\; V\frac{k^3}{2\pi^2}\langle|\delta\varphi_{\mathbf{k}}|^2\rangle \;=\; \left(\frac{H}{2\pi}\right)^2_{aH=k}} \,, \tag{337}$$

to signify that the value of H for each k is to be taken at horizon exit of that particular scale. Equation (337) is our main result from inflaton fluctuations.

8.6.4 Transfer functions

Since the inflaton fluctuations are assumed to be the origin of structure, all later perturbations are related to the inflaton perturbations $\delta\varphi_{\mathbf{k}}$. As long as all inhomogeneities are small ("perturbations"), the relationship is linear. We can express these linear relationships as *transfer functions* $T(t,k)$, e.g.,

$$g_{\mathbf{k}}(t) = T_{g\varphi}(t,k)\delta\varphi_{\mathbf{k}}(t_k)\,. \tag{338}$$

The linearity implies several things:

1. The Fourier coefficient $g_{\mathbf{k}}$ depends only on the Fourier coefficient of $\delta\varphi$ corresponding to the same wave vector $\mathbf{k}$, not on any other $\mathbf{k}'$.

2. The relationship is linear, so that if $\delta\varphi_{\mathbf{k}}$ were, e.g., twice as big, then so would $g_{\mathbf{k}}$ be.

3. The perturbations of g inherit the Gaussian statistics of $\delta\varphi$.

We could also define transfer functions relating perturbations at any two different times, t and t', and call them $T(t,t',k)$, but here we are referring to the inflaton perturbations at the time of horizon exit, t_k, which is different for different k. Actually, by $\delta\varphi_{\mathbf{k}}(t_k)$ we mean the constant value the perturbation approaches after horizon exit in the $H = const = H_k$ approximation.

That the transfer function depends only on the magnitude k results from the fact that physical laws are isotropic. The transfer function of Eq. (338) will then relate the power spectra of $\{g_{\mathbf{k}}(t)\}$ and $\{\delta\varphi_{\mathbf{k}}(t_k)\}$ as

$$\mathcal{P}_g(t,k) = T_{g\varphi}(t,k)^2\mathcal{P}_\varphi(k)\,. \tag{339}$$

The transfer functions thus incorporate all the physics that determines how structure evolves.

[46] One can do a more precise calculation, where one takes into account the evolution of $H(t)$. The result is that one gets a correction to the amplitude of $\mathcal{P}_\mathcal{R}(k)$, which is first order in slow-roll parameters, and a correction to its spectral index n which is second order in the slow-roll parameters.

For the largest scales, $k^{-1} \gg 10h^{-1}\,\mathrm{Mpc}$, the perturbations are still small today, and one needs not go beyond the transfer function. For smaller scales, corresponding to galaxies and galaxy clusters, the inhomogeneities have become large at late times, and the physics of structure growth has become nonlinear. This nonlinear evolution is typically studied using large numerical simulations. Fortunately, the relevant scales are small enough that Newtonian physics is usually sufficient.

We are now in position to put together all the results we obtained. From Eq. (316)

$$\mathcal{R}_{\mathbf{k}} = -H\frac{\delta\varphi_{\mathbf{k}}}{\dot{\bar{\varphi}}}\,, \tag{340}$$

so that

$$T_{\mathcal{R}\varphi}(k) = -\frac{H_k}{\dot{\bar{\varphi}}(t_k)} \tag{341}$$

and

$$\mathcal{P}_{\mathcal{R}}(k) = \left(\frac{H}{\dot{\bar{\varphi}}}\right)^2 \mathcal{P}_{\varphi}(k) = \left[\left(\frac{H}{\dot{\bar{\varphi}}}\right)\left(\frac{H}{2\pi}\right)\right]^2_{\mathcal{H}=k}\,, \tag{342}$$

where we used the result (337).

This primordial spectrum is the starting point for calculating structure formation (discussed already) and the CMB anisotropy (Chapter 9). Thus CMB and large-scale structure observations can be used to constrain $\mathcal{P}_{\mathcal{R}}$ together with other cosmological parameters.

8.6.5 Generation of primordial gravitational waves

The quantum fluctuations at subhorizon scales during inflation apply also to the spacetime itself. We do not yet have a complete theory of quantum gravity, so we do not know how spacetime behaves in the Planck era. At lower energy scales the spacetime fluctuations are smaller and for small perturbations around a FRW universe we can use the linearized equations for metric perturbations, for which quantization is straightforward. In fact, the proper treatment of the generation of inflaton perturbations, where we include the scalar metric perturbations in the inflaton perturbation equation (see Eq. 315), contains also the quantum treatment of scalar metric perturbations.

Likewise, we have quantum fluctuations of tensor metric perturbations during inflation. These do not couple to density perturbations, but they become classical gravitational waves after horizon exit. These *primordial gravitational waves* have an effect on CMB anisotropy and polarization.

In the quantum treatment, $(M_{\mathrm{Pl}}/\sqrt{2})h$ fluctuates like a scalar field, so that in inflation the gravitational wave amplitudes h acquire a spectrum

$$\mathcal{P}_h(k) \;\equiv\; 4\frac{V}{2\pi^2}k^3\langle|h_{\mathbf{k}}|^2\rangle \;=\; 4\frac{2}{M_{\mathrm{Pl}}^2}\left(\frac{H}{2\pi}\right)^2_{\mathcal{H}=k} \;=\; \frac{8}{M_{\mathrm{Pl}}^2}\left(\frac{H}{2\pi}\right)^2_{\mathcal{H}=k} \tag{343}$$

(the factor 4 in this customary definition is related to the way h appears in several places in the metric and to there being two modes for each $\mathbf{k}$).

The *tensor-to-scalar ratio* is the ratio of the two primordial spectra (343) and (342),

$$r \;\equiv\; \frac{\mathcal{P}_h(k)}{\mathcal{P}_{\mathcal{R}}(k)} \;=\; \frac{8}{M_{\mathrm{Pl}}^2}\left(\frac{\dot{\bar{\varphi}}}{H}\right)^2_{\mathcal{H}=k}\,. \tag{344}$$

8.7 The primordial spectrum

8.7.1 Primordial spectrum from slow-roll inflation

The final result of the previous section is thus that inflation generates primordial scalar perturbations $\mathcal{R}_{\mathbf{k}}$ with the power spectrum

$$\mathcal{P}_{\mathcal{R}}(k) \;=\; \left[\left(\frac{H}{\dot\varphi}\right)\left(\frac{H}{2\pi}\right)\right]^2_{\mathcal{H}=ak} \;=\; \frac{1}{4\pi^2}\left(\frac{H^2}{\dot\varphi}\right)^2_{t=t_k}. \tag{345}$$

and primordial tensor perturbations with the power spectrum

$$\mathcal{P}_h(k) \;=\; \frac{8}{M_{\rm Pl}^2}\left(\frac{H}{2\pi}\right)^2_{t=t_k}. \tag{346}$$

In this section φ and $\dot\varphi$ refer to the background values.

Applying the slow-roll equations

$$H^2 = \frac{V}{3M_{\rm Pl}^2} \quad \text{and} \quad 3H\dot\varphi = -V' \quad \Rightarrow \quad \frac{\dot\varphi}{H} = -M_{\rm Pl}^2\frac{V'}{V}$$

these become

$$\begin{aligned}\mathcal{P}_{\mathcal{R}}(k) &= \frac{1}{12\pi^2}\frac{1}{M_{\rm Pl}^6}\frac{V^3}{V'^2} \;=\; \frac{1}{24\pi^2}\frac{1}{M_{\rm Pl}^4}\frac{V}{\varepsilon}\\ \mathcal{P}_h(k) &= \frac{2}{3\pi^2}\frac{V}{M_{\rm Pl}^4},\end{aligned} \tag{347}$$

where ε is the slow-roll parameter. The tensor-to-scalar ratio is thus

$$r \;\equiv\; \frac{\mathcal{P}_h(k)}{\mathcal{P}_{\mathcal{R}}(k)} \;=\; 16\varepsilon\,. \tag{348}$$

According to present observational CMB and large-scale structure data, the amplitude of the primordial power spectrum is about

$$\mathcal{P}_{\mathcal{R}}(k)^{1/2} \approx 5\times 10^{-5} \tag{349}$$

at cosmological scales. This gives a constraint on inflation

$$\left(\frac{V}{\varepsilon}\right)^{1/4} \;\approx\; 24^{1/4}\sqrt{\pi}\sqrt{5\times 10^{-5}}M_{\rm Pl} \;\approx\; 0.028M_{\rm Pl} \;=\; 6.8\times 10^{16}\,{\rm GeV}\,. \tag{350}$$

The best chance of detecting primordial gravitational waves is based on their effect on CMB. They have not been observed so far and the present upper limit is about [8]

$$r < 0.07 \quad \Rightarrow \quad \mathcal{P}_h(k)^{1/2} < 1.3\times 10^{-5} \quad \text{and} \quad \varepsilon < 0.004\,. \tag{351}$$

This implies an upper limit to the inflation energy scale[47]

$$V^{1/4} \;\approx\; \varepsilon^{1/4}0.028M_{\rm Pl} \;<\; 0.007M_{\rm Pl} \;=\; 1.7\times 10^{16}\,{\rm GeV}\,. \tag{352}$$

[47] For the epoch when perturbations at observable cosmological scales were generated. During earlier phases of inflation the energy scale was higher than in that epoch.

Since during inflation, V and V' change slowly while a wide range of scales k exit the horizon, $\mathcal{P}_\mathcal{R}(k)$ and $\mathcal{P}_h(k)$ should be slowly varying functions of k. We define the *spectral indices* n_s and n_t of the primordial spectra as

$$\begin{aligned} n_s(k) - 1 &\equiv \frac{d\ln\mathcal{P}_\mathcal{R}}{d\ln k} \\ n_t(k) &\equiv \frac{d\ln\mathcal{P}_h}{d\ln k}\,. \end{aligned} \tag{353}$$

(The -1 is in the definition of n_s for historical reasons, to match with the definition in terms of density perturbations, see Sec. 8.7.2.) If the spectral index is independent of k, we say that the spectrum is *scale free*. In this case the primordial spectra have the *power-law* form

$$\mathcal{P}_\mathcal{R}(k) = A_s^2\left(\frac{k}{k_p}\right)^{n_s-1} \qquad \text{and} \qquad \mathcal{P}_h(k) = A_t^2\left(\frac{k}{k_p}\right)^{n_t}\,, \tag{354}$$

where k_p is some chosen reference scale, "pivot scale", and A_s and A_t are the amplitudes at this pivot scale.

If the power spectrum is constant,

$$\mathcal{P} = \text{const.}\,, \tag{355}$$

corresponding to $n_s = 1$ and $n_t = 0$, we say that the spectrum is *scale invariant*. A scale-invariant scalar spectrum is also called the *Harrison–Zeldovich* spectrum.

If $n_s \neq 1$ or $n_t \neq 0$, the spectrum is called *tilted*. A tilted spectrum is called *red*, if $n_s < 1$ (more structure at large scales), and *blue* if $n_s > 1$ (more structure at small scales).

Using Eqs. (347) and (353) we can calculate the spectral index for slow-roll inflation.

Since $\mathcal{P}(k)$ is evaluated from Eqs. (345) and (346) or (347) when $k = aH$,

$$\frac{d\ln k}{dt} = \frac{d\ln(aH)}{dt} = \frac{\dot a}{a} + \frac{\dot H}{H} = (1-\varepsilon)H \approx H\,,$$

where we used $\dot H = -\varepsilon H^2$ (in the slow-roll approximation) in the last step. Thus

$$\frac{d}{d\ln k} = \frac{1}{1-\varepsilon}\frac{1}{H}\frac{d}{dt} = \frac{1}{1-\varepsilon}\frac{\dot\varphi}{H}\frac{d}{d\varphi} = -\frac{M_\mathrm{Pl}^2}{1-\varepsilon}\frac{V'}{V}\frac{d}{d\varphi} \approx -M_\mathrm{Pl}^2\frac{V'}{V}\frac{d}{d\varphi}\,. \tag{356}$$

Let us first calculate the scale dependence of the slow-roll parameters:

$$\frac{d\varepsilon}{d\ln k} = -M_\mathrm{Pl}^2\frac{V'}{V}\frac{d}{d\varphi}\left[\frac{M_\mathrm{Pl}^2}{2}\left(\frac{V'}{V}\right)^2\right] = M_\mathrm{Pl}^4\left[\left(\frac{V'}{V}\right)^4 - \left(\frac{V'}{V}\right)^2\frac{V''}{V}\right] = 4\varepsilon^2 - 2\varepsilon\eta \tag{357}$$

and, in a similar manner (**exercise**),

$$\frac{d\eta}{d\ln k} = \ldots = 2\varepsilon\eta - \xi\,, \tag{358}$$

where we have defined a third slow-roll parameter

$$\xi \equiv M_\mathrm{Pl}^4\frac{V'}{V^2}V'''\,. \tag{359}$$

The parameter ξ is typically second-order small in the sense that $\sqrt{|\xi|}$ is of the same order of magnitude as ε and η. (Therefore it is sometimes written as ξ^2, although nothing forces it to be positive.)

We are now ready to calculate the spectral indices:

$$\begin{aligned} n_s - 1 &= \frac{1}{\mathcal{P}_\mathcal{R}}\frac{d\mathcal{P}_\mathcal{R}}{d\ln k} = \frac{\varepsilon}{V}\frac{d}{d\ln k}\left(\frac{V}{\varepsilon}\right) = \frac{1}{V}\frac{dV}{d\ln k} - \frac{1}{\varepsilon}\frac{d\varepsilon}{d\ln k} \\ &= -M_{\rm Pl}^2\frac{V'}{V}\cdot\frac{1}{V}\frac{dV}{d\varphi} - 4\varepsilon + 2\eta = -6\varepsilon + 2\eta \\ n_t &= \frac{1}{\mathcal{P}_h}\frac{d\mathcal{P}_h}{d\ln k} = -M_{\rm Pl}^2\frac{V'}{V}\frac{1}{V}\frac{dV}{d\varphi} = -M_{\rm Pl}^2\left(\frac{V'}{V}\right)^2 = -2\varepsilon\,. \end{aligned} \tag{360}$$

Since $\varepsilon > 0$, the tensor spectrum is necessarily red. (This follows already from (346), since H is decreasing, or from (347) since V is decreasing.) Slow-roll requires $\varepsilon \ll 1$ and $|\eta| \ll 1$, so *both spectra are close to scale invariant*. For scalar perturbations this is verified by observation. Based on CMB anisotropy data from the Planck satellite, the Planck Collaboration [8] finds

$$n_s = 0.965 \pm 0.004\,. \tag{361}$$

If one were able to measure all three values n_s, r, and n_t from observations, one could solve from them the slow-roll parameters ε and η and moreover, check the *consistency condition*

$$n_t = -\frac{r}{8} \tag{362}$$

for single-field slow-roll inflation. This consistency condition is the only truly quantitative prediction of the inflation scenario (as opposed to some specific inflation model) – all the other predictions (Ω_k very small, n_s close to 1 and n_t close to 0, primordial perturbations Gaussian) are of qualitative nature, not a specific number not equal to 0 or 1.

Unfortunately, the existing upper limit to r already means that it will be difficult to ever determine the spectral index n_t with sufficient accuracy to distinguish between $n_t = -r/8$ and $n_t = 0$. The most sensitive probe to primordial gravitational waves is provided by polarization of CMB on which they will imprint a characteristic pattern (discussed briefly in the next chapter). The theoretical limit to detection is $r \sim 10^{-4}$ and there are proposals[48] for future CMB satellite missions that could reach $r \sim 10^{-3}$. If r is significantly larger than these detection limits, after detection one could still measure n_t accurately enough to distinguish, say, $n_t \approx -1$, $n_t \approx 0$ (which includes the case $n_t = -r/8$), and $n_t \approx 1$ from each other. There have been other proposals (other than inflation) for very-early-universe physics, which predict primordial tensor perturbations that deviate from scale invariance this much or more.

The Japanese space agency ISAS/JAXA selected in May 2019 the 3-year LiteBIRD mission[49] to be launched in the late 2020s. LiteBIRD is expected to measure r with accuracy $\Delta r \sim 10^{-3}$ ($1\ \sigma$). A significant American and European participation in LiteBIRD is expected.

Detection of primordial gravitational waves, i.e., measurement of r, would be enough to determine ε and η and thus the inflation energy scale from Eq. (350).

One can also calculate the scale-dependence of the spectral index (**exercise**):

$$\frac{dn_s}{d\ln k} = 16\varepsilon\eta - 24\varepsilon^2 - 2\xi\,. \tag{363}$$

It is second order in slow-roll parameters, so it's expected to be even smaller than the deviation from scale invariance, $n_s - 1$. Planck Collaboration finds it consistent with zero to accuracy $\mathcal{O}(10^{-2})$, as expected.

Cosmologically observable scales have a range of about $\Delta\ln k \sim 10$. Planck measured the CMB anisotropy over a range $\Delta\ln k \sim 6$ (missing the shortest scales, where the CMB is expected

[48] See, e.g., `http://www.core-mission.org/`

[49] `http://litebird.jp/eng/`

to have negligible anisotropy). Some inflation models have $|n_s - 1|$, r, and $|dn_s/d\ln k|$ larger than the Planck results, while others do not. These observations already ruled out many inflation models.

Example: Consider the simple inflation model

$$V(\varphi) = \frac{1}{2}m^2\varphi^2 \,.$$

In Chapter 7 we already calculated the slow-roll parameters for this model:

$$\varepsilon = \eta = 2\frac{M_{\rm Pl}^2}{\varphi^2}$$

and we immediately see that $\xi = 0$. Thus

$$\begin{aligned} n_s &= 1 - 6\varepsilon + 2\eta \;=\; 1 - 8\left(\frac{M_{\rm Pl}}{\varphi}\right)^2 \\ \frac{dn_s}{d\ln k} &= 16\varepsilon\eta - 24\varepsilon^2 - 2\xi \;=\; -32\left(\frac{M_{\rm Pl}}{\varphi}\right)^4 \\ r &= 16\varepsilon \;=\; 32\left(\frac{M_{\rm Pl}}{\varphi}\right)^2 \\ n_t &= -2\varepsilon \;=\; -4\left(\frac{M_{\rm Pl}}{\varphi}\right)^2 \end{aligned}$$

To get the numbers out, we need the values of φ when the relevant cosmological scales exited the horizon. The number of inflation e-foldings after that should be about $N \sim 50$. We have

$$N(\varphi) \;=\; \frac{1}{M_{\rm Pl}^2}\int_{\varphi_{\rm end}}^{\varphi}\frac{V}{V'}d\varphi \;=\; \frac{1}{M_{\rm Pl}^2}\int\frac{\varphi}{2} \;=\; \frac{1}{4M_{\rm Pl}^2}\left(\varphi^2 - {\varphi_{\rm end}}^2\right) ,$$

and we estimate $\varphi_{\rm end}$ from $\varepsilon(\varphi_{\rm end}) = 2M_{\rm Pl}^2/{\varphi_{\rm end}}^2 = 1 \quad \Rightarrow \quad \varphi_{\rm end} = \sqrt{2}M_{\rm Pl}$ to get

$$\varphi^2 = {\varphi_{\rm end}}^2 + 4M_{\rm Pl}^2 N = 2M_{\rm Pl}^2 + 4M_{\rm Pl}^2 N \approx 4M_{\rm Pl}^2 N \,.$$

Thus

$$\left(\frac{M_{\rm Pl}}{\varphi}\right)^2 = \frac{1}{4N}$$

and

$$\begin{aligned} n_s &= 1 - \frac{2}{N} \;\approx\; 0.96 \\ \frac{dn_s}{d\ln k} &= -\frac{2}{N^2} \;\approx\; -0.0008 \\ r &= \frac{8}{N} \;\approx\; 0.16 \\ n_t &= -\frac{1}{N} \;\approx\; -0.02 \end{aligned}$$

We see that this model is ruled out by the observed upper limit $r < 0.1$.[50]

---

[50] There was enormous excitement in early 2014, when the BICEP2 collaboration[12] claimed to have detected the effect of primordial gravitational waves with $r = 0.20^{+0.07}_{-0.05}$ in CMB polarization, consistent with this inflation model. However, it turned out that their data was contaminated by polarized emission from dust in our own galaxy.[13]

8.7.2 Scale invariance of the primordial power spectrum

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 spectrum to be scale invariant.

The primordial spectrum is something we have at superhorizon scales, where we have discussed it in terms of the comoving curvature perturbation $\mathcal{R}$, and we are calling it scale invariant, when

$$\mathcal{P}_{\mathcal{R}}(k) = A_s^2 = \text{const.} \tag{364}$$

We would like the spectrum in terms of more familiar concepts like the density perturbation, but at superhorizon scales the density perturbation is gauge dependent.

For small scales the perturbation spectrum gets modified when the scales enter the horizon, but for large scales $k \ll k_{\rm eq}$ the spectrum maintains its primordial shape, at least as long as the universe stays matter dominated. This allows the discussion of the primordial spectrum at subhorizon scales, where we can talk about the density perturbations without specifying a gauge.

From Eq. (259), the gravitational potential and density perturbation are related to the curvature perturbation as

$$\begin{aligned} \Phi_{\mathbf{k}} &= -\frac{3}{5}\mathcal{R}_{\mathbf{k}} \qquad \text{(mat.dom)} \\ \delta_{\mathbf{k}} &= -\frac{2}{3}\left(\frac{k}{\mathcal{H}}\right)^2 \Phi_{\mathbf{k}} \;=\; \frac{2}{5}\left(\frac{k}{\mathcal{H}}\right)^2 \mathcal{R}_{\mathbf{k}}\,, \end{aligned} \tag{365}$$

giving

$$\mathcal{P}_\Phi(k) = \frac{9}{25}\mathcal{P}_{\mathcal{R}}(k) = \frac{9}{25}A_s^2 = const \tag{366}$$

$$\begin{aligned} \mathcal{P}_\delta(t,k) &= \frac{4}{9}\left(\frac{k}{\mathcal{H}}\right)^4 \mathcal{P}_\Phi(k) = \frac{4}{25}\left(\frac{k}{\mathcal{H}}\right)^4 \mathcal{P}_{\mathcal{R}}(k) \\ &= \frac{4}{25}\left(\frac{k}{\mathcal{H}}\right)^4 A_s^2 \propto t^{4/3}k^4 \end{aligned} \tag{367}$$

Thus perturbations in the gravitational potential are scale invariant, but perturbations in density are not. Instead the density perturbation spectrum is steeply rising, meaning that there is much more structure at small scales than at large scales. Thus the scale invariance refers to the gravitational aspect of perturbations, which in the Newtonian treatment is described by the gravitational potential, and in the GR treatment by spacetime curvature.

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.

However, if we extrapolate Eq. (367) back to horizon entry, $k = \mathcal{H}$, we get

$$\delta_H^2(k) \equiv \text{“}\mathcal{P}_\delta(k,t_k)\text{”} \equiv \frac{4}{25}\mathcal{P}_{\mathcal{R}}(k) = \left(\frac{2}{5}A_s\right)^2 = const \tag{368}$$

Thus for scale-invariant primordial perturbations, *density perturbations of all scales enter the horizon with the same amplitude*, $\delta_H = (2/5)A_s \sim 2\times 10^{-5}$. Since the density perturbation at the horizon entry is actually a gauge-dependent quantity, and our extension of the above Newtonian relation up to the horizon scale is not really allowed, this statement should be taken just qualitatively (hence the quotation marks around the $\mathcal{P}_\delta$). As such, it applies also to the

smaller scales which enter during the radiation-dominated epoch, since the perturbations only begin to evolve after horizon entry.

What is the deep reason that inflation generates (almost) scale invariant perturbations? During inflation the universe is almost a de Sitter universe, which has the metric

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

with $H = const.$ In GR we learn that it 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. Therefore, as different scales exit at different times they all obtain the same kind of perturbations.

In terms of the other definition of the power spectrum, $P(k) \equiv (2\pi^2/k^3)\mathcal{P}(k)$, the relations (367) for scale-invariant perturbations give

$$\begin{aligned} P_{\mathcal{R}}(k) &\propto k^{-3}\mathcal{P}_R \propto k^{-3} \\ P_\delta(k) &\propto k^{-3}\mathcal{P}_\delta \propto k^4 P_{\mathcal{R}} \propto k\mathcal{P}_R \propto k \end{aligned} \tag{369}$$

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.

8.8 The power spectrum today

8.8.1 Density perturbations

From Eq. (265), the density perturbation spectrum at late times is

$$\mathcal{P}_\delta(k) = \frac{4}{25}\left(\frac{k}{\mathcal{H}}\right)^4 T(k,t)^2 \mathcal{P}_{\mathcal{R}}(k) \tag{370}$$

where, from Eq. (267)

$$\begin{aligned} T(k) &= 1 && \text{for} \quad k \ll k_{\rm eq} \\ T(k) &\sim \left(\frac{k_{\rm eq}}{k}\right)^2 \ln k && \text{for} \quad k \gg k_{\rm eq}\,. \end{aligned}$$

(For the more precise form of $T(k)$ see Sec. 8.4.4, Eq. (270) and Fig. 11.) Thus the present-day density power spectrum rises steeply $\propto k^4$ at large scales, but turns at $\sim k_{\rm eq}$ to become less steep (growing $\sim \ln k$) at small scales. This is because the growth of density perturbations was inhibited while the perturbations were inside the horizon during the radiation-dominated epoch. The $\sim \ln k$ factor comes from the slow growth of CDM perturbations during this time.

Thus the structure in the universe appears stronger at smaller scales, down to $k_{\rm eq}^{-1} \sim 100\,$Mpc. The $\sim 100\,$Mpc scale is indeed quite prominent in large scale structure surveys, like the 2dF-GRS and SDSS galaxy distribution surveys. Towards smaller scales the structure keeps getting stronger, but now more slowly. However, perturbations are now so large that first-order perturbation theory begins to fail, and that limit is crossed at around $k^{-1} \sim k_{\rm nl}^{-1} \sim 10\,$Mpc. Nonlinear effects cause the density power spectrum to rise more steeply than calculated by perturbation theory at scales smaller than this.

The present-day density power spectrum $\mathcal{P}_\delta(k)$ can be determined observationally from the distribution of galaxies (Fig. 14). The quantity plotted is usually $P_\delta(k)$. It should go as

$$\begin{aligned} P_\delta(k) &\propto k^n && \text{for} \quad k \ll k_{\rm eq} \\ P_\delta(k) &\propto k^{n-4}\ln k && \text{for} \quad k \gg k_{\rm eq}\,. \end{aligned} \tag{371}$$

See Fig. 15.

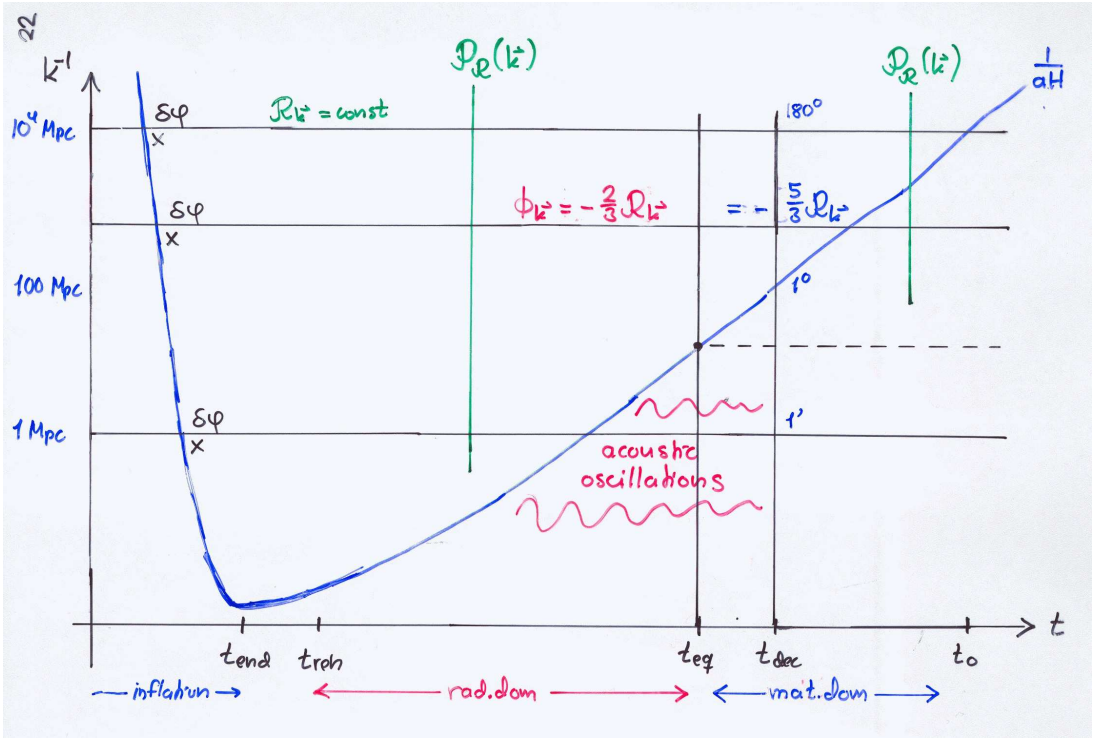


Figure 13: The whole picture of structure formation theory from quantum fluctuations during inflation to the present-day power spectrum at t_0.

Example: Errors on $P(k)$ estimation. The accuracy of power spectrum estimation from a galaxy survey is affected by *sample variance* and *shot noise*. Sample (or cosmic) variance was discussed earlier. Shot noise comes from the fact that instead of observing a continuous matter distribution, we are sampling it with a finite number density n of galaxies. To reduce sample variance one should increase the survey volume V_s. To reduce shot noise, one should increase n, by observing also smaller and fainter galaxies. We skip the math for shot noise (discussed in Galaxy Survey Cosmology), and give just the result that for Gaussian perturbations the combined effect of sample variance and shot noise is

$$\langle |\hat{P}(k) - P(k)|^2 \rangle \;=\; \frac{2}{N_k}\left[P(k) + \frac{1}{\bar{n}}\right]^2 , \tag{372}$$

where $\bar{n}$ is the mean number density of galaxies in the survey. (This is Eq. (335) modified by adding the effect of shot noise.)

Let us see how well this corresponds to the $P(k)$ estimate obtained in [15] for the SDSS survey (Fig. 15). Table I of [15] gives their estimate $\hat{P}(k)$ and their estimated uncertainty $\Delta P(k)$ for each of their 20 k-bins in numbers. We copy these numbers into our Table 1. The number of LRG galaxies in the survey was $N = 53\,860$. The effective sky area of the survey was $\Omega_s = 4259\,\mathrm{deg}^2 = 1.297\,\mathrm{sr}$, and the surveyed redshift range was $z = 0.155$–0.474. From Figs. 1 and 2 of [15] this corresponds to a comoving distance range $r = r_{\min} - r_{\max} = 450 - 1300\,h^{-1}\mathrm{Mpc}$, from which we get a survey volume and number density

$$\begin{aligned} V_s &= \Omega_s \left(\frac{r_{\max}^3}{3} - \frac{r_{\min}^3}{3} \right) = 0.9107\,h^{-3}\mathrm{Gpc}^3 , \\ \bar{n} &= \frac{N}{V_s} = 6.408 \times 10^{-5}\,h^3\mathrm{Mpc}^{-3} . \end{aligned} \tag{373}$$

If the survey volume were a cube, $V_s = L^3$, the side of the cube would be $L = 969.3\,h^{-1}\mathrm{Mpc}$, which means the components of $\mathbf{k}$ are integer multiples of $k_f = 2\pi/L = 0.006482\,h/\mathrm{Mpc}$. The volume of a k-bin

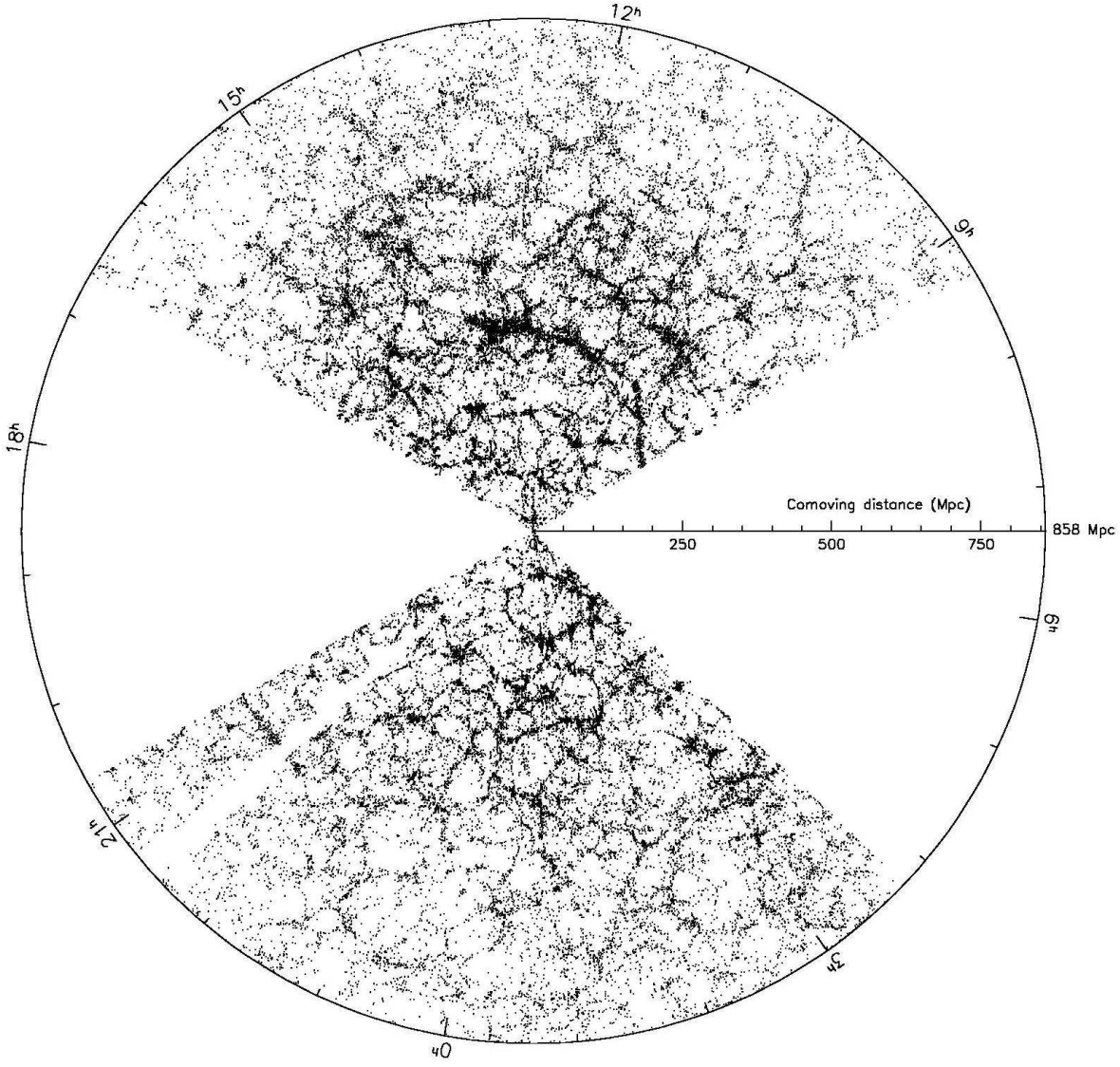


Figure 14: Distribution of galaxies according to the Sloan Digital Sky Survey (SDSS). This figure shows galaxies that are within 2° of the equator and closer than 858 Mpc (assuming $H_0 = 71$ km/s/Mpc). Figure from astro-ph/0310571[14].

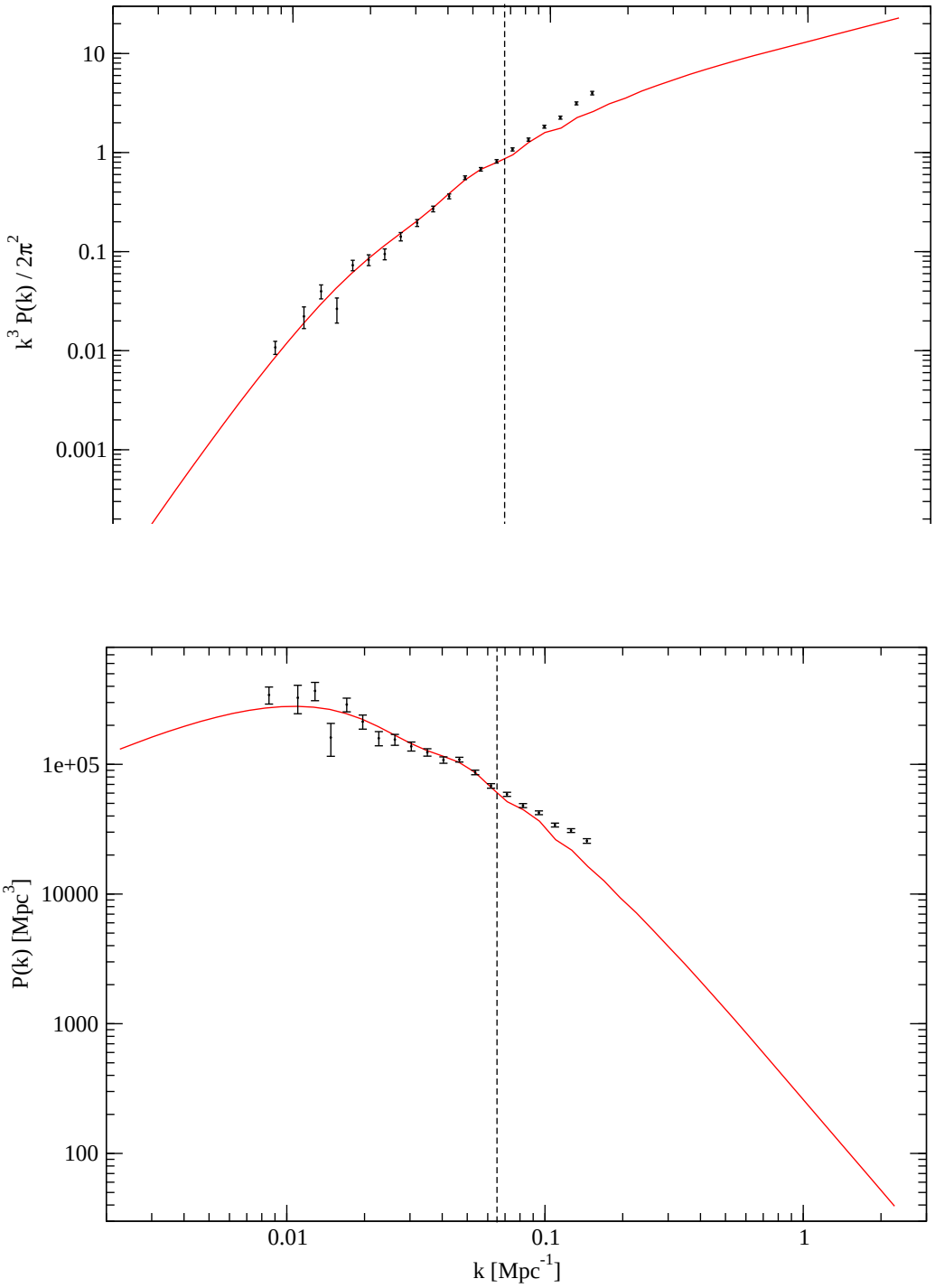


Figure 15: The power spectrum from the SDSS obtained using luminous red galaxies (LRG) [15]. Note that LRG are more strongly biased, $b \approx 1.9$, than galaxies on average, so to get the $P(k)$ of matter, one should divide by $b^2 \approx 3.6$. The top figure shows $\mathcal{P}_\delta(k)$ and the bottom figure $P_\delta(k)$. A Hubble constant value $H_0 = 71.4$ km/s/Mpc has been assumed for this figure. (These galaxy surveys only obtain the scales up to the Hubble constant, and therefore the observed $P_\delta(k)$ is usually shown in units of $h^{-3}\,\mathrm{Mpc}^3$ as a function of h Mpc^{-1}, so that no value for H_0 need to be assumed.) The black bars are the observations and the red curve is a theoretical fit, from linear perturbation theory, to the data. The bend in $P(k)$ at $k_{\mathrm{eq}} \sim 0.01$ Mpc^{-1} is clearly visible in the bottom figure. Linear perturbation theory fails when $\mathcal{P}(k) \gtrsim 1$, and therefore the data points do not follow the theoretical curve to the right of the dashed line (representing an estimate on how far linear theory can be trusted). Figure by R. Keskitalo.

is

$$V_k \;=\; \frac{4\pi}{3}\left(k_{\rm max}^3 - k_{\rm min}^3\right)\,. \tag{374}$$

and the number of $\mathbf{k}$ modes in the bin is $N_k = V_k/k_f^3$, which we have added as the third column in Table 1.

Using the estimated $\hat{P}(k)$ in place of $P(k)$ on the rhs of Eq. (372), we obtain from it our estimate for the relative uncertainty as

$$\frac{\sqrt{\langle|\hat{P}(k)-P(k)|^2\rangle}}{P(k)} \;\approx\; \frac{\sqrt{\frac{2}{N_k}}\left[\hat{P}(k)+\frac{1}{\bar{n}}\right]}{\hat{P}(k)}\,, \tag{375}$$

which we have added as the last column of Table 1.

Compared to the actual SDSS analysis the above is extremely naive. The facts that the survey volume is not a cube and that because of selection effects the mean number density of galaxies decreases as a function of z required a more sophisticated analysis. Also the k-bins SDSS used where not sharp $[k_{\rm min}, k_{\rm max}]$, but instead for each bin they used a smooth window function (of k) giving more weight to the modes near the center of the bin and some weight even to modes outside $[k_{\rm min}, k_{\rm max}]$. In addition, nonlinear effects cause deviations from Gaussianity, which typically increase the sample variance, and there are other effects contributing to the estimate uncertainty; so that we should expect (375) to be an underestimate. Nevertheless, comparing the last column of Table 1 (our naive estimate) to the second-to-last column (the SDSS uncertainty estimate) we see that we are on the right track.

SDSS LRG Power spectrum

$k_{\rm min}$	$k_{\rm max}$	N_k	$\hat{P}(k)$	$\Delta P(k)$	$\Delta P(k)/\hat{P}(k)$	Eq. (375)
0.008	0.017	68	124884	18775	0.150	0.193
0.013	0.018	56	118814	29400	0.247	0.214
0.016	0.022	101	134291	21638	0.161	0.157
0.018	0.025	151	58644	16647	0.284	0.146
0.021	0.028	195	105253	12736	0.121	0.116
0.025	0.033	312	77699	9666	0.124	0.096
0.029	0.037	404	57870	7264	0.126	0.089
0.033	0.043	670	56516	5466	0.097	0.070
0.037	0.051	1261	50125	3991	0.080	0.052
0.042	0.057	1708	45076	2956	0.066	0.046
0.050	0.066	2499	39339	2214	0.056	0.040
0.057	0.075	3640	39609	1679	0.042	0.033
0.066	0.086	5360	31566	1284	0.041	0.029
0.076	0.099	8171	24837	991	0.040	0.025
0.088	0.113	11710	21390	778	0.036	0.023
0.101	0.128	16407	17507	629	0.036	0.021
0.108	0.145	27511	15421	516	0.033	0.017
0.126	0.165	38320	12399	430	0.035	0.016
0.159	0.190	43665	11237	382	0.034	0.016
0.181	0.218	68135	9345	384	0.041	0.014

Table 1: The first and second column give roughly the extent of each k-bin used in units of $h\,\mathrm{Mpc}^{-1}$. (Note that the bins overlap.) N_k is the corresponding number of k modes in the bin. $\hat{P}(k)$ and $\Delta P(k)$ are the estimates from [15] in units of $(\,h^{-1}\mathrm{Mpc})^3$, and in the last two columns we compare their ratio to the uncertainty estimate we get from the survey characteristics.

8.8.2 σ_8

Instead of A_s, the amplitude of the primordial power spectrum at a pivot scale, astronomers like to use the parameter σ_8 to describe the amplitude of cosmological perturbations of a cosmological

model. σ_8 is defined as the top-hat window matter density variance $\sigma_T(R)$ at scale $R = 8\,h^{-1}$Mpc predicted by linear perturbation theory in the cosmological model.

We mentioned earlier (Sec. 8.1.7) that according to observations, $\sigma_{T,g}(R = 8\,h^{-1}\text{Mpc}) \approx 1$. This means that nonlinear effects are becoming important at this scale. The difference between $\sigma_{T,g}(R = 8\,h^{-1}\text{Mpc})$ and σ_8 is due to 1) galaxy bias and 2) nonlinear effects. Both should work in the direction of making $\sigma_{T,g}(R = 8\,h^{-1}\text{Mpc})$ larger than σ_8, so we expect that the correct cosmological model has $\sigma_8 < 1$, but not necessarily by very much.

To calculate σ_8 in a cosmological model specified by, say, A_s, n_s, Ω_m, Ω_Λ, ω_m, and ω_b, is not simple and requires numerical computation. This is done, e.g., by the CAMB code. The computation involves the transfer function $T(k)$, the growth function $D(a)$, and doing the integral (48a) for $\sigma_T^2(R)$.

For the Planck 2018 best-fit ΛCDM model, $\sigma_8 = 0.8210$.

Planck 2018 best-fit model σ_8. Let's see how well the results of this chapter allow us to calculate the Planck 2018 best-fit ΛCDM model σ_8. The model has parameter values (see Table I, `Plik` best-fit column in [8])

$$\begin{aligned} \ln(10^{10}A_s^2) &= 3.0448 \quad \Rightarrow \quad A_s^2 = 21.006\times 10^{-10} \\ n_s &= 0.96605 \\ H_0 &= 0.8120 \\ \Omega_m &= 0.3158 \\ \omega_b &= 0.022383\,, \end{aligned} \tag{376}$$

where A_s is for the pivot scale $k_p = 0.05\,\text{Mpc}^{-1}$.

The primordial power spectrum is

$$\mathcal{P}_\mathcal{R}(k) = A_s^2\left(\frac{k}{k_p}\right)^{n_s-1}. \tag{377}$$

We want to find the present-day linear power spectrum $\mathcal{P}_\delta(k)$. If the universe would have stayed matter dominated until today, $\delta_\mathbf{k}$ and $\mathcal{R}_\mathbf{k}$ would be related by

$$\tilde{\delta}_\mathbf{k} = \frac{2}{5}\left(\frac{k}{\tilde{H}_0}\right)^2 T(k)\mathcal{R}_\mathbf{k}\,, \tag{378}$$

where $\tilde{}$ denotes matter-dominated-model quantities, and "today" is defined by $a = 1$. As we noted in the footnote in Sec. 8.3.5, the comparison is to a matter-dominated model with the same matter density today, so it has a smaller total density and thus a smaller Hubble constant, $\tilde{H}_0 = \Omega_m^{1/2}H_0 = 37.83\,\text{km/s/Mpc}$. The true linear $\delta_\mathbf{k}$ differs from (378) due to the different growth function. For the matter-dominated model $D_m(a) = a$, so $D_m(1) = 1$, and

$$\begin{aligned} \delta_\mathbf{k}(t) &= \frac{2}{5}\frac{D(a)}{a}\left(\frac{k}{\tilde{\mathcal{H}}(a)}\right)^2 T(k)\mathcal{R}_\mathbf{k} \\ \delta_\mathbf{k}(t_0) &= \frac{2}{5}D(1)\left(\frac{k}{\tilde{H}_0}\right)^2 T(k)\mathcal{R}_\mathbf{k}\,, \end{aligned} \tag{379}$$

where $D(a)$ is the growth function of the ΛCDM model, given by Eq. (239). Thus the power spectra are related by

$$\mathcal{P}_\delta(k) = \frac{4}{25}D(1)^2\left(\frac{k}{\tilde{H}_0}\right)^4 T(k)^2\mathcal{P}_\mathcal{R}(k)\,. \tag{380}$$

We now calculate $\mathcal{P}_\delta(k)$ for $k = 1/(8\,h^{-1}\text{Mpc})$. For this k,

$$\begin{aligned} \frac{k}{k_p} &= \frac{1}{0.05\,\text{Mpc}^{-1}8\,h^{-1}\text{Mpc}} = \frac{h}{0.4} = 1.683 \\ \frac{k}{\tilde{H}_0} &= 666.9\,, \end{aligned} \tag{381}$$

so that

$$\mathcal{P}_\delta(k) \;=\; 65.30 D(1)^2 T(k)^2 \,. \tag{382}$$

The BBKS transfer function, using $k_{\rm eq}^{-1} = 13.7 \Omega_m^{-1} h^{-2}\,{\rm Mpc} = 64.44\, h^{-1}{\rm Mpc}$, gives for this k

$$T_{\rm BBKS}(k) = 0.1435\,, \qquad \text{with slope} \quad -1.17933\,, \tag{383}$$

so taht

$$\mathcal{P}_\delta(k) \;=\; 1.346 D(1)^2 \left(\frac{T(k)}{T_{\rm BBKS}(k)} \right)^2 \,. \tag{384}$$

We should now calculate $D(1)$ for the Planck best-fit ΛCDM model, and run CAMB to get the true $T(k)$ fro this model, but I am lazy here, and use the results for our reference model ($\Omega_m = 0.3$, $h = 0.7$) from Sec. 8.3.5 and Fig. 11: $D(1) = 0.78$ and $T(k)/T_{\rm BBKS}(k) \approx 0.7$ to give

$$\mathcal{P}_\delta(k) \approx 0.40\,. \tag{385}$$

Finally we should calculate σ_T^2 for $R = 8\, h^{-1}{\rm Mpc}$, which is an integral of $\mathcal{P}_\delta(k)$ over k, but we try to get away with just using the value and slope at $k = 1/R$, i.e., we approximate $\mathcal{P}_\delta(k)$ with a power-law function. The slope is $n = 0.96605 - 2 \times 1.17933 = -1.39261$ (modifying the slope of $\mathcal{P}_\mathcal{R}(k)$ with the slope of $T_{\rm BBKS}(k)$) and, from (59),

$$\sigma_8^2 \;\approx\; \frac{9}{2^n}(n+1) \sin\frac{n\pi}{2}\; \frac{\Gamma(n-1)}{n-3} \mathcal{P}_\delta(k) \;=\; 1.939\, \mathcal{P}_\delta(k) \;\approx\; 0.776\,, \tag{386}$$

giving

$$\sigma_8 \approx 0.881 > 0.8210\,. \tag{387}$$

Using the power-law approximation should give an overestimate, since the true $\mathcal{P}_\delta(k)$ bends down compared to it on both sides of the chosen k value, and indeed we overestimated σ_8, but not badly.

8.8.3 Primordial gravitational waves

We found that outside the horizon tensor perturbations remain constant,

$$h_{\mathbf{k}}(t) = h_{\mathbf{k},{\rm prim}} = const\,, \tag{388}$$

whereas inside the horizon they become gravitational waves whose amplitude decays

$$|h_{\mathbf{k}}(t)| \propto a^{-1}\,. \tag{389}$$

Define the transfer function for gravitational waves

$$T_h(k) \equiv \frac{|h_{\mathbf{k}}(t_0)|}{h_{\mathbf{k},{\rm prim}}}\,, \tag{390}$$

so that the present-day power spectrum of primordial gravitational waves is

$$\mathcal{P}_{\rm grav}(k, t_0) = T_h(k)^2 \mathcal{P}_h(k)\,. \tag{391}$$

Make the approximation that the transition from (388) to (389) is instantaneous at horizon entry defined as

$$k = \mathcal{H} = aH\,. \tag{392}$$

Denote these values of a, H, and $\mathcal{H}$ by a_k, H_k, and $\mathcal{H}_k$. Then

$$T_h(k) \;=\; \frac{a_k}{a_0} \;=\; a_k\,. \tag{393}$$

The shape of the transfer function is determined by the rate at which different comoving scales k enter horizon as the universe expands. This is determined by the evolution of the comoving Hubble distance $\mathcal{H}^{-1}$.

In the matter-dominated universe

$$a \propto t^{2/3} \quad \text{and} \quad H = \frac{2}{3t} \propto a^{-3/2} \quad \Rightarrow \quad \mathcal{H} \propto a^{-1/2} \,. \tag{394}$$

Make first the approximation that the universe is still matter dominated. Then

$$T_h(k) \;=\; \frac{a_k}{a_0} \;=\; \left(\frac{\mathcal{H}_k}{\mathcal{H}_0}\right)^{-2} \;=\; \left(\frac{k}{a_0 H_0}\right)^{-2} \qquad (H_0 < k < k_{\rm eq}) \tag{395}$$

for scales that entered during the matter-dominated epoch.

To correct this result for the effect of dark energy at late times, we note that because of dark energy, the comoving Hubble distance $\mathcal{H}^{-1} = (aH)^{-1}$ stopped growing and began to shrink, so that the scale $k = H_0$ is actually exiting now, and it entered at an earlier time t_1 when the expansion was still (barely) matter dominated. Thus the above result for $T_h(k)$ should apply (roughly) at that earlier time:

$$T_h(t_1, k) \;=\; \left(\frac{k}{a_1 H_1}\right)^{-2} \;=\; \left(\frac{k}{a_0 H_0}\right)^{-2} \qquad (H_0 < k < k_{\rm eq}) \tag{396}$$

While the scale $k = H_0$ was inside the horizon, the universe expanded by about a factor of two, so the correct transfer function is about half of (395).

Exercise: Extend the result (395) to scales $k > k_{\rm eq}$. You can make the approximation where the transition from radiation-dominated expansion law to matter-dominated expansion law is instantaneous at $t_{\rm eq}$. (This approximation actually underestimates $T_h(k > k_{\rm eq})$ by a factor that roughly compensates the overestimation in (395) from ignoring dark energy at late times.)

Gravitational waves were detected for the first time on September 14, 2015 at the LIGO observatory. These were not primordial gravitational waves; they were caused by a collision of two black holes about 400 Mpc from here, and they were observed only for about 0.2 seconds. The peak amplitude was $h \approx 10^{-21}$. LIGO is sensitive to frequencies near 100 Hz, and with further refinements it is expected to reach a sensitivity of $h = 10^{-22}$. Assume the primordial tensor perturbations had amplitude $h = 10^{-5}$ (close to the upper limit from CMB observations). What is their amplitude today at the 100 Hz frequency?

ESA is planning to launch a space gravitational wave observatory (LISA) in 2034. It would have similar sensitivity as LIGO, but for frequencies lower by a factor 10^{-4}. What do you conclude about the prospect for observing primordial gravitational waves this way?

References

[1] C.M. Baugh, *The real-space correlation function measured from the APM Galaxy Survey*, Mon. Not. R. Astron. Soc. **280**, 267 (1996), astro-ph/9512011

[2] A.J. Benson, R.G. Bower, C.S. Frenk, C.G. Lacey, C.M. Baugh, and S. Cole, *What shapes the luminosity function of galaxies?*, Astrophys. J. **599**, 38 (2003), astro-ph/0302450

[3] A.G. Sánchez et al., *The clustering of galaxies in the SDSS-III Baryon Oscillation Spectroscopic Survey: cosmological implications of the large-scale two-point correlation function*, Mon. Not. R. Astron. Soc. **425**, 415 (2012), arXiv:1203.6616

[4] E. Hawkins et al., *The 2dF Galaxy Redshift Survey: correlation functions, peculiar velocities and the matter density of the Universe*, Mon. Not. R. Astron. Soc. **346**, 78 (2003),astro-ph/0212375

[5] J.A. Peacock: Cosmological Physics (Cambridge University Press 1999), Chapter 16

[6] A.R. Liddle and D.H. Lyth: Cosmological Inflation and Large-Scale Structure (Cambridge University Press 2000)

[7] Planck Collaboration, Astronomy & Astrophysics **594**, A13 (2016), arXiv:1502.01589

[8] Planck Collaboration, *Planck 2018 results. VI. Cosmological parameters*, arXiv:1807.06209

[9] S. Dodelson: Modern Cosmology (Academic Press 2003), Chapter 7

[10] J.M. Bardeen, J.R. Bond, N. Kaiser, A.S. Szalay, *The statistics of peaks in Gaussian randomn fields*, Astrophys. J. **304**, 15 (1986), Appendix G

[11] E.W. Kolb and M.S. Turner: The Early Universe (Addison-Wesley 1990)

[12] P.A.R. Ade et al., Phys. Rev. Lett. **112**, 241101 (2014), arXiv:1403.3985

[13] Planck Collaboration, Astronomy & Astrophysics **586**, A133 (2016), arXiv:1409.5738

[14] J. Richard Gott III et al., *A Map of the Universe*, Astrophys. J. **624**, 463 (2005), astro-ph/0310571

[15] M. Tegmark et al., *Cosmological Constraints from the SDSS Luminous Red Galaxies*, Phys. Rev. **D74**, 123507 (2006), astro-ph/0608632