1

The Homogeneous and Isotropic Universe

Notation

In this book we denote the derivative with respect to physical time by a prime and the derivative with respect to conformal time by a dot,

$$\tau = \text{physical (cosmic) time} \qquad \frac{dX}{d\tau} \equiv X', \tag{1.1}$$

$$t = \text{conformal time} \qquad \frac{dX}{dt} \equiv \dot{X}. \tag{1.2}$$

Spatial 3-vectors are denoted by a boldface symbol such as $\mathbf{k}$ or $\mathbf{x}$ whereas four-dimensional spacetime vectors are denoted as $x = (x^\mu)$.

We use the metric signature $(-, +, +, +)$ throughout the book.

The Fourier transform is defined by

$$f(\mathbf{k}) = \int d^3x \, f(\mathbf{x}) \, e^{i\mathbf{k}\cdot\mathbf{x}}, \tag{1.3}$$

so that

$$f(\mathbf{x}) = \frac{1}{(2\pi)^3} \int d^3k \, f(\mathbf{k}) \, e^{-i\mathbf{k}\cdot\mathbf{x}}. \tag{1.4}$$

We use the same letter for $f(\mathbf{x})$ and for its Fourier transform $f(\mathbf{k})$. The spectrum $P_f(k)$ of a statistically homogeneous and isotropic random variable f is given by

$$\langle f(\mathbf{k}) f^*(\mathbf{k}') \rangle = (2\pi)^3 \, \delta(\mathbf{k} - \mathbf{k}') P_f(k). \tag{1.5}$$

Since it is isotropic, $P_f(k)$ is a function only of the modulus $k = |\mathbf{k}|$.

Throughout this book we use units where the speed of light, c; Planck's constant, $\hbar$; and Boltzmann's constant, k_B, are unity: $c = \hbar = k_B = 1$. Length and time therefore have the same units and energy, mass, and momentum also have the same units, which are inverse to the unit of length. Temperature has the same units

https://doi.org/10.1017/9781316471524.002 Published online by Cambridge University Press

as energy. We may use cm^{-1} to measure energy, mass, and temperature, or eV^{-1} to measure distances or times. We shall use whatever unit is convenient to discuss a given problem. Conversion factors can be found in Appendix 1.

1.1 Homogeneity and Isotropy

Modern cosmology is based on the hypothesis that our Universe is to a good approximation homogeneous and isotropic on sufficiently large scales. This relatively bold assumption is often called the "cosmological principle." It is an extension of the Copernican principle stating that not only should our place in the Solar System not be a special one, but also that the position of the Milky Way in the Universe should be in no way statistically distinguishable from the position of other galaxies. Furthermore, no direction should be distinguished. The Universe looks statistically the same in all directions. This, together with the hypothesis that the matter density and geometry of the Universe are smooth functions of the position, implies homogeneity and isotropy on sufficiently large scales. Isotropy around each point together with analyticity actually already implies homogeneity of the Universe.[1] A formal proof of this quite intuitive result can be found in Straumann (1974).

But which scale is "sufficiently large"? Certainly not the Solar System or our Galaxy. But also not the size of galaxy clusters. [In cosmology, distances are usually measured in Mpc (Megaparsec). 1 Mpc $= 3.2615 \times 10^6$ light years $= 3.0856 \times 10^{24}$ cm is a typical distance between galaxies; the distance between our neighbor Andromeda and the Milky Way is about 0.7 Mpc. These and other connections between frequently used units can be found in Appendix 1.]

It turns out that the scale at which the *galaxy distribution* becomes homogeneous is difficult to determine. From the analysis of the Sloan Digital Sky Survey (SDSS) it has been concluded that the irregularities in the galaxy density are still on the level of a few percent on scales of 100 Mpc (Hogg *et al.*, 2005). Fortunately, we know that the *geometry* of the Universe shows only small deviations from the homogeneous and isotropic background, already on scales of a few Mpc. The geometry of the Universe can be tested with the peculiar motion of galaxies, with lensing, and in particular with the cosmic microwave background (CMB).

The small deviations from homogeneity and isotropy in the CMB are of utmost importance, since, most probably, they represent the "seeds," that, via gravitational instability, have led to the formation of large-scale structure, galaxies, and eventually solar systems with planets that support life in the Universe.

[1] If "analyticity" is not assumed, the matter distribution could also be fractal and still statistically isotropic around each point. For a detailed elaboration of this idea and its comparison with observations see Sylos Labini *et al.* (1998).

https://doi.org/10.1017/9781316471524.002 Published online by Cambridge University Press

Furthermore, we suppose that the initial fluctuations needed to trigger the process of gravitational instability stem from tiny quantum fluctuations that have been amplified during a period of inflationary expansion of the Universe. I consider this connection of the microscopic quantum world with the largest scales of the Universe to be of breathtaking philosophical beauty.

In this chapter we investigate the background Universe. We shall first discuss the geometry of a homogeneous and isotropic spacetime. Then we investigate two important events in the thermal history of the Universe. Finally, we study the paradigm of inflation. This chapter lays the basis for the following ones where we shall investigate *fluctuations* on the background, most of which can be treated in first-order perturbation theory.

1.2 The Background Geometry of the Universe

1.2.1 The Friedmann Equations

In this section we assume a basic knowledge of general relativity. The notation and sign convention for the curvature tensor that we adopt are specified in Appendix 2, Section A2.1.

Our Universe is described by a four-dimensional spacetime $(\mathcal{M}, g)$ given by a pseudo-Riemannian manifold $\mathcal{M}$ with metric g. A homogeneous and isotropic spacetime is one that admits a slicing into homogeneous and isotropic, that is, maximally symmetric, 3-spaces. There is a preferred geodesic time coordinate τ, called "cosmic time," such that the 3-spaces of constant time, $\Sigma_\tau = \{\mathbf{x} | (\tau, \mathbf{x}) \in \mathcal{M}\}$, are maximally symmetric spaces, hence spaces of constant curvature. The metric g is therefore of the form

$$ds^2 = g_{\mu\nu}\, dx^\mu\, dx^\nu = -d\tau^2 + a^2(\tau)\gamma_{ij}\, dx^i\, dx^j. \tag{1.6}$$

The function $a(\tau)$ is called the scale factor and γ_{ij} is the metric of a 3-space of constant curvature K. Depending on the sign of K this space is locally isometric to a 3-sphere ($K > 0$); a three-dimensional pseudo-sphere ($K < 0$); or flat, Euclidean space ($K = 0$). In later chapters of this book we shall mainly use "conformal time" t defined by $a\, dt = d\tau$, so that

$$ds^2 = g_{\mu\nu}\, dx^\mu\, dx^\nu = a^2(t)\left(-dt^2 + \gamma_{ij}\, dx^i\, dx^j\right). \tag{1.7}$$

The geometry and physics of homogeneous and isotropic solutions to Einstein's equations were first investigated mathematically in the early 1920s by Friedmann (1922, 1924) and physically as a description of the observed expanding Universe

https://doi.org/10.1017/9781316471524.002 Published online by Cambridge University Press

in 1927 by Lemaître.[2] Later, Robertson (1936), Walker (1936), and others rediscovered the Friedmann metric and studied several additional aspects. However, since we consider the contributions by Friedmann and Lemaître to be far more fundamental than the subsequent work, we shall call a homogeneous and isotropic solution to Einstein's equations a "Friedmann–Lemaître universe" (FL universe) in this book.

It is interesting to note that the Friedmann solution breaks Lorentz invariance. Friedmann universes are not invariant under boosts; there is a preferred cosmic time τ, the proper time of an observer who sees a spatially homogeneous and isotropic universe. Like so often in physics, the Lagrangian and therefore also the field equations of general relativity are invariant under Lorentz transformations, but a specific solution in general is not. In that sense we are back to Newton's vision of an absolute time. But on small scales, for example, the scale of a laboratory, this violation of Lorentz symmetry is, of course, negligible.

The topology is not determined by the metric and hence by Einstein's equations. There are many compact spaces of negative or vanishing curvature (e.g., the torus), but there are no infinite spaces with positive curvature. A beautiful treatment of the fascinating, but difficult, subject of the topology of spaces with constant curvature and their classification is given in Wolf (1974). Its applications to cosmology are found in Lachieze-Rey and Luminet (1995).

Forms of the metric γ, which we shall often use, are

$$\gamma_{ij}\,dx^i\,dx^j = \frac{\delta_{ij}\,dx^i\,dx^j}{(1+\frac{1}{4}K\rho^2)^2}, \tag{1.8}$$

$$\gamma_{ij}\,dx^i\,dx^j = dr^2 + \chi^2(r)\left(d\theta^2 + \sin^2(\theta)\,d\varphi^2\right), \tag{1.9}$$

$$\gamma_{ij}\,dx^i\,dx^j = \frac{dR^2}{1-KR^2} + R^2\left(d\theta^2 + \sin^2(\theta)\,d\varphi^2\right), \tag{1.10}$$

where in Eq. (1.8)

$$\rho^2 = \sum_{i,\,j=1}^{3} \delta_{ij}x^i x^j, \quad \text{and} \quad \delta_{ij} = \begin{cases} 1 & \text{if } i=j, \\ 0 & \text{else}, \end{cases} \tag{1.11}$$

and in Eq. (1.9);

$$\chi(r) = \begin{cases} r & \text{in the Euclidean case,} \quad K=0, \\ \frac{1}{\sqrt{K}}\sin(\sqrt{K}r) & \text{in the spherical case,} \quad K>0, \\ \frac{1}{\sqrt{|K|}}\sinh(\sqrt{|K|}r) & \text{in the hyperbolic case,} \quad K<0. \end{cases} \tag{1.12}$$

[2] In the English translation of (Lemaître, 1927) from 1931 Lemaître's somewhat premature but pioneering arguments that the observed Universe is actually expanding have been omitted.

https://doi.org/10.1017/9781316471524.002 Published online by Cambridge University Press

Often one normalizes the scale factor such that $K = \pm 1$ whenever $K \neq 0$. One has, however, to keep in mind that in this case r and K become dimensionless and the scale factor a has the dimension of length. If $K = 0$ we can normalize a arbitrarily. We shall usually normalize the scale factor such that $a_0 = 1$ and the curvature is not dimensionless. The coordinate transformations that relate these coordinates are determined in Exercise 1.1.

Owing to the symmetry of spacetime, the energy–momentum tensor can only be of the form

$$(T_{\mu\nu}) = \begin{pmatrix} -\rho g_{00} & \mathbf{0} \\ \mathbf{0} & P g_{ij} \end{pmatrix}. \tag{1.13}$$

There is no additional assumption going into this ansatz, such as the matter content of the Universe being an ideal fluid. It is a simple consequence of homogeneity and isotropy and is also verified for scalar field matter, a viscous fluid, or free-streaming particles in a FL universe. As usual, the energy density ρ and the pressure P are defined as the time- and space-like eigenvalues of (T^{μ}_{ν}).

The Einstein tensor can be calculated from the definition (A2.12) and Eqs. (A2.32)–(A2.39),

$$G_{00} = 3\left[\left(\frac{a'}{a}\right)^2 + \frac{K}{a^2}\right] \qquad \text{(cosmic time)}, \tag{1.14}$$

$$G_{ij} = -\left(2a''a + a'^2 + K\right)\gamma_{ij} \qquad \text{(cosmic time)}, \tag{1.15}$$

$$G_{00} = 3\left[\left(\frac{\dot{a}}{a}\right)^2 + K\right] \qquad \text{(conformal time)}, \tag{1.16}$$

$$G_{ij} = -\left(2\left(\frac{\dot{a}}{a}\right)^{\bullet} + \left(\frac{\dot{a}}{a}\right)^2 + K\right)\gamma_{ij} \qquad \text{(conformal time)}. \tag{1.17}$$

The Einstein equations relate the Einstein tensor to the energy–momentum content of the Universe via $G_{\mu\nu} = 8\pi G T_{\mu\nu} - g_{\mu\nu}\Lambda$. Here Λ is the so-called cosmological constant. In an FL universe the Einstein equations become

$$\left(\frac{a'}{a}\right)^2 + \frac{K}{a^2} = \frac{8\pi G}{3}\rho + \frac{\Lambda}{3} \qquad \text{(cosmic time)}, \tag{1.18}$$

$$2\frac{a''}{a} + \frac{(a')^2}{a^2} + \frac{K}{a^2} = -8\pi G P + \Lambda \qquad \text{(cosmic time)}, \tag{1.19}$$

$$\left(\frac{\dot{a}}{a}\right)^2 + K = \frac{8\pi G}{3}a^2\rho + \frac{a^2\Lambda}{3} \qquad \text{(conformal time)}, \tag{1.20}$$

$$2\left(\frac{\dot{a}}{a}\right)^{\bullet} + \left(\frac{\dot{a}}{a}\right)^2 + K = -8\pi G a^2 P + a^2\Lambda \qquad \text{(conformal time)}. \tag{1.21}$$

https://doi.org/10.1017/9781316471524.002 Published online by Cambridge University Press

Energy "conservation," $T^{\mu\nu}_{;\mu} = 0$, yields

$$\dot{\rho} = -3(\rho + P)\left(\frac{\dot{a}}{a}\right) \quad \text{or, equivalently} \quad \rho' = -3(\rho + P)\left(\frac{a'}{a}\right). \tag{1.22}$$

This equation can also be obtained by differentiating Eq. (1.18) or (1.20) and inserting (1.19) or (1.21); it is a consequence of the contracted Bianchi identities (see Appendix 2, Section A2.1). Equations (1.18)–(1.21) are the Friedmann equations. The quantity

$$H(\tau) \equiv \frac{a'}{a} = \frac{\dot{a}}{a^2} \equiv \mathcal{H}a^{-1}, \tag{1.23}$$

is called the Hubble rate or the Hubble parameter, where $\mathcal{H}$ is the comoving Hubble parameter. At present, the Universe is expanding, so that $H_0 > 0$. We parameterize it by

$$H_0 = 100\, h\ \mathrm{km\,s^{-1}\,Mpc^{-1}} \simeq 3.241 \times 10^{-18}\, h\ \mathrm{s^{-1}} \simeq 0.3336 \times 10^{-3}\ \mathrm{h\,Mpc^{-1}}.$$

Observations show (Freedman *et al.*, 2001) that $h \simeq 0.72 \pm 0.1$. Equation (1.22) is easily solved in the case $w = P/\rho =$ constant. Then one finds

$$\rho = \rho_0 (a_0/a)^{3(1+w)}, \tag{1.24}$$

where ρ_0 and a_0 denote the value of the energy density and the scale factor at present time, τ_0. In this book cosmological quantities indexed by a "0" are evaluated today, $X_0 = X(\tau_0)$. For nonrelativistic matter, $P_m = 0$, we therefore have $\rho_m \propto a^{-3}$ while for radiation (or any kind of massless particles) $P_r = \rho_r/3$ and hence $\rho_r \propto a^{-4}$. A cosmological constant corresponds to $P_\Lambda = -\rho_\Lambda$ and we obtain, as expected, $\rho_\Lambda =$ constant. If the curvature K can be neglected and the energy density is dominated by one component with $w =$ constant, inserting Eq. (1.24) into the Friedmann equations yields the solutions

$$a \propto \tau^{2/3(1+w)} \propto t^{2/(1+3w)} \qquad w = \text{constant} \neq -1, \tag{1.25}$$

$$a \propto \tau^{2/3} \propto t^2 \qquad w = 0, \quad \text{(dust)}, \tag{1.26}$$

$$a \propto \tau^{1/2} \propto t \qquad w = 1/3, \quad \text{(radiation)}, \tag{1.27}$$

$$a \propto \exp(H\tau) \propto 1/|t| \qquad w = -1, \quad \text{(cosmol. const.)}. \tag{1.28}$$

It is interesting to note that if $w < -1$, so-called phantom matter, we have to choose $\tau < 0$ to obtain an expanding universe and the scale factor diverges in finite time, at $\tau = 0$. This is the so-called big rip. Phantom matter has many problems but it is discussed in connection with the supernova type 1a (SN1a) data, which are compatible with an equation of state with $w < -1$ or with an ordinary cosmological constant (Caldwell *et al.*, 2003). For $w < -\frac{1}{3}$ the time coordinate t

https://doi.org/10.1017/9781316471524.002 Published online by Cambridge University Press

has to be chosen as negative for the Universe to expand and spacetime cannot be continued beyond $t = 0$. But $t = 0$ corresponds to a cosmic time, the proper time of a static observer, $\tau = \infty$; this is not a singularity. (The geodesics can be continued until affine parameter ∞.)

We also introduce the adiabatic sound speed c_s determined by

$$c_s^2 = \frac{P'}{\rho'} = \frac{\dot{P}}{\dot{\rho}}. \tag{1.29}$$

From this definition and Eq. (1.22) it is easy to see that

$$\dot{w} = 3\mathcal{H}(1+w)\big(w - c_s^2\big). \tag{1.30}$$

Hence $w =$ constant if and only if $w = c_s^2$ or $w = -1$. Note that already in a simple mixture of matter and radiation $w \neq c_s^2 \neq$ constant (see Exercise 1.3).

Equation (1.18) implies that for a critical value of the energy density given by

$$\rho(\tau) = \rho_c(\tau) = \frac{3H^2}{8\pi G} \tag{1.31}$$

the curvature and the cosmological constant vanish. The value ρ_c is called the critical density. The ratio $\Omega_X = \rho_X/\rho_c$ is the "density parameter" of the component X. It indicates the fraction that the component X contributes to the expansion of the Universe. We shall make use especially of

$$\Omega_r \equiv \Omega_r(\tau_0) = \frac{\rho_r(\tau_0)}{\rho_c(\tau_0)}, \tag{1.32}$$

$$\Omega_m \equiv \Omega_m(\tau_0) = \frac{\rho_m(\tau_0)}{\rho_c(\tau_0)}, \tag{1.33}$$

$$\Omega_K \equiv \Omega_K(\tau_0) = \frac{-K}{a_0^2 H_0^2}, \tag{1.34}$$

$$\Omega_\Lambda \equiv \Omega_\Lambda(\tau_0) = \frac{\Lambda}{3H_0^2}. \tag{1.35}$$

1.2.2 The "Big Bang" and "Big Crunch" Singularities

We can absorb the cosmological constant into the energy density and pressure by redefining

$$\rho_{\text{eff}} = \rho + \frac{\Lambda}{8\pi G}, \qquad P_{\text{eff}} = P - \frac{\Lambda}{8\pi G}.$$

https://doi.org/10.1017/9781316471524.002 Published online by Cambridge University Press

Since Λ is a constant and $\rho_{\text{eff}} + P_{\text{eff}} = \rho + P$, the conservation equation (1.22) still holds. A first interesting consequence of the Friedmann equations is obtained when subtracting Eq. (1.18) from (1.19). This yields

$$\frac{a''}{a} = -\frac{4\pi G}{3}(\rho_{\text{eff}} + 3P_{\text{eff}}). \tag{1.36}$$

Hence if $\rho_{\text{eff}} + 3P_{\text{eff}} > 0$, the Universe is decelerating. Furthermore, Eqs. (1.22) and (1.36) then imply that in an expanding and decelerating universe

$$\frac{\rho'_{\text{eff}}}{\rho_{\text{eff}}} < -2\frac{a'}{a},$$

so that ρ decays faster than $1/a^2$. If the curvature is positive, $K > 0$, this implies that at some time in the future, $\tau_{\max}$, the density has dropped down to the value of the curvature term, $K/a^2(\tau_{\max}) = 8\pi G\rho_{\text{eff}}(\tau_{\max})$. Then the Universe stops expanding and recollapses. Furthermore, this is independent of curvature; as a' decreases the curve $a(\tau)$ is concave and thus cuts the $a = 0$ line at some finite time in the past. This moment of time is called the "big bang." The spatial metric vanishes at this value of τ, which we usually choose to be $\tau = 0$; and spacetime cannot be continued to earlier times. This is not a coordinate singularity. From the Ricci tensor given in Eqs. (A2.32) and (A2.33) one obtains the Riemann scalar

$$R = 6\left[\frac{a''}{a} + \left(\frac{a'}{a}\right)^2 + \frac{K}{a^2}\right],$$

which also diverges if $a \to 0$. Also the energy density, which grows faster than $1/a^2$ as $a \to 0$, diverges at the big bang.

If the curvature K is positive, the Universe contracts after $\tau = \tau_{\max}$ and, since the graph $a(\tau)$ is convex, reaches $a = 0$ at some finite time τ_c, the time of the "big crunch." The big crunch is also a physical singularity beyond which spacetime cannot be continued.

It is important to note that this behavior of the scale factor can be implied only if the so-called strong energy condition holds, $\rho_{\text{eff}} + 3P_{\text{eff}} > 0$. This is illustrated in Fig. 1.1.

1.2.3 Cosmological Distance Measures

It is notoriously difficult to measure distances in the Universe. The position of an object in the sky gives us its angular coordinates, but how far away is the object from us? This problem had plagued cosmology for centuries. It took until 1915–1920 when Hubble discovered that the "spiral nebulae" are actually not situated inside our own galaxy but much further away. This then led to the discovery of the expansion of the Universe.

https://doi.org/10.1017/9781316471524.002 Published online by Cambridge University Press

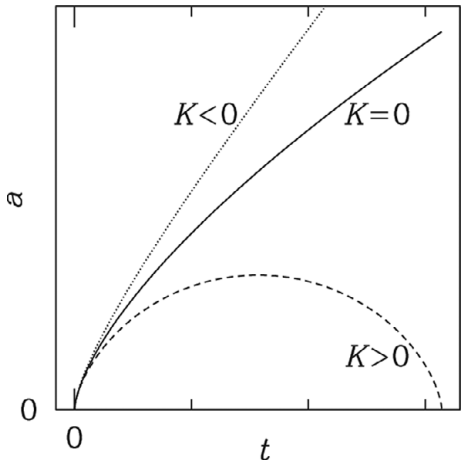

Fig. 1.1 The kinematics of the scale factor in a Friedmann–Lemaître universe that satisfies the strong energy condition, $\rho_{\rm eff} + 3P_{\rm eff} > 0$.

For cosmologically distant objects, a third coordinate, which is today relatively easy to obtain, is the redshift z experienced by the photons emitted from the object. A given spectral line with intrinsic wavelength λ is redshifted due to the expansion of the Universe. If it is emitted at some time τ, it reaches us today with wavelength $\lambda_0 = \lambda a_0/a(\tau) = (1+z)\lambda$. This leads to the definition of the cosmic redshift

$$z(\tau) + 1 = \frac{a_0}{a(\tau)}. \tag{1.37}$$

On the other hand, an object at physical distance $d = a_0 r$ away from us, at redshift $z \ll 1$, recedes with speed $v = H_0 d$. To the lowest order in z, we have $\tau_0 - \tau \approx d$ and $a_0 \approx a(\tau) + a'(\tau_0 - \tau)$, so that

$$1 + z \approx 1 + \frac{a'}{a}(\tau_0 - \tau) \approx 1 + H_0 d.$$

For objects that are sufficiently close, $z \ll 1$. We therefore have $v \approx z$ and hence $H_0 = z/d$. This is the method usually applied to measure the Hubble constant.

There are different ways to measure distances in cosmology, all of which give the same result in a Minkowski universe but differ in an expanding universe. They are, however, simply related, as we shall see.

One possibility is to define the distance d_A to a certain object of given physical size Δ seen at redshift z_1 such that the angle subtended by the object is given by

$$\vartheta = \Delta/d_A, \qquad d_A = \Delta/\vartheta. \tag{1.38}$$

This is the angular diameter distance; see Fig. 1.2.

https://doi.org/10.1017/9781316471524.002 Published online by Cambridge University Press

Fig. 1.2 The two ends of the object emit a flash simultaneously from A and B at z_1 which reaches us today. The angular diameter distance to A (or B) is defined by $d_A = \Delta/\vartheta$.

We now derive the expression

$$d_A(z) = \frac{1}{\sqrt{|\Omega_K|}H_0(1+z)}\chi\left(\sqrt{|\Omega_K|}H_0\int_0^z \frac{dz'}{H(z')}\right), \tag{1.39}$$

for the angular diameter distance to redshift z. In a given cosmological model, this allows us to express the angular diameter distance for a given redshift as a function of the cosmological parameters.

To derive Eq. (1.39) we use the coordinates introduced in Eq. (1.9). Without loss of generality we set $r = 0$ at our position. We consider an object of physical size Δ at redshift z_1 simultaneously emitting a flash at both of its ends, A and B. Hence $r = r_1 = t_0 - t_1$ at the position of the flashes, A and B at redshift z_1. If Δ denotes the physical arc length between A and B we have $\Delta = a(t_1)\chi(r_1)\vartheta = a(t_1)\chi(t_0 - t_1)\vartheta$, that is,

$$\vartheta = \frac{\Delta}{a(t_1)\chi(t_0 - t_1)}. \tag{1.40}$$

According to Eq. (1.38) the angular diameter distance to t_1 or z_1 is therefore given by

$$a(t_1)\chi(t_0 - t_1) \equiv d_A(z_1). \tag{1.41}$$

To obtain an expression for $d_A(z)$ in terms of the cosmic density parameters and the redshift, we have to calculate $(t_0 - t_1)(z_1)$.

Note that in the case $K = 0$ we can normalize the scale factor a as we want, and it is convenient to choose $a_0 = 1$, so that comoving scales become physical scales today. However, for $K \neq 0$, we have already normalized a such that $K = \pm 1$ and $\chi(r) = \sin r$ or $\sinh r$. In this case, we have no normalization constant left and a_0 has the dimension of a length. The present spatial curvature of the Universe then is $\pm 1/a_0^2$.

The Friedmann equation Eq. (1.20) reads

$$\dot{a}^2 = \frac{8\pi G}{3}a^4\rho + \frac{1}{3}\Lambda a^4 - Ka^2, \tag{1.42}$$

https://doi.org/10.1017/9781316471524.002 Published online by Cambridge University Press

where $\dot{a} = da/dt$. To be specific, we assume that ρ is a combination of dust, cold, nonrelativistic "matter" of $P_m = 0$ and radiation of $P_r = \rho_r/3$.

Since $\rho_r \propto a^{-4}$ and $\rho_m \propto a^{-3}$, we can express the terms on the right-hand side of Eq. (1.42) as

$$\frac{8\pi G}{3} a^4 \rho = H_0^2 \left(a_0^4 \Omega_r + \Omega_m a a_0^3 \right), \tag{1.43}$$

$$\frac{1}{3} \Lambda a^4 = H_0^2 \Omega_\Lambda a^4, \tag{1.44}$$

$$-K a^2 = H_0^2 \Omega_K a^2 a_0^2. \tag{1.45}$$

The Friedmann equation then implies

$$\frac{da}{dt} = H_0 a_0^2 \left(\Omega_r + \frac{a}{a_0} \Omega_m + \frac{a^4}{a_0^4} \Omega_\Lambda + \frac{a^2}{a_0^2} \Omega_K \right)^{1/2}, \tag{1.46}$$

so that

$$\begin{aligned} r(z_1) &= t_0 - t_1 \\ &= \frac{1}{H_0 a_0} \int_0^{z_1} \frac{dz}{\left[\Omega_r (z+1)^4 + \Omega_m (z+1)^3 + \Omega_\Lambda + \Omega_K (z+1)^2 \right]^{1/2}} \\ &= \frac{1}{a_0} \int_0^{z_1} \frac{dz}{H(z)}. \end{aligned} \tag{1.47}$$

Here we have used $z + 1 = a_0/a$ so that $da = -dz a_0/(1+z)^2$.

In principle, we could of course also add other matter components such as, for example, "quintessence" (Caldwell and Steinhardt, 1998), which would lead to a somewhat different form of the integral (1.47), but for definiteness, we remain with matter, radiation, and a cosmological constant.

From $-K/H_0^2 a_0^2 = \Omega_K$ we obtain $H_0 a_0 = 1/\sqrt{|\Omega_K|}$ for $\Omega_K \neq 0$. The expression for the angular diameter distance thus becomes

$$d_A(z) = \begin{cases} \frac{1}{\sqrt{|\Omega_K|} H_0 (z+1)} \chi \left(\sqrt{|\Omega_K|} \int_0^z \frac{dz'}{\left[\Omega_r (z'+1)^4 + \Omega_m (z'+1)^3 + \Omega_\Lambda + \Omega_K (z'+1)^2 \right]^{1/2}} \right) \\ \qquad \text{if } K \neq 0 \\ \frac{1}{H_0 (z+1)} \int_0^z \frac{dz'}{\left[\Omega_r (z'+1)^4 + \Omega_m (z'+1)^3 + \Omega_\Lambda \right]^{1/2}} \\ \qquad \text{if } K = 0. \end{cases} \tag{1.48}$$

Using the Friedmann equation, this formula can also be written in the more general form of Eq. (1.39).

https://doi.org/10.1017/9781316471524.002 Published online by Cambridge University Press

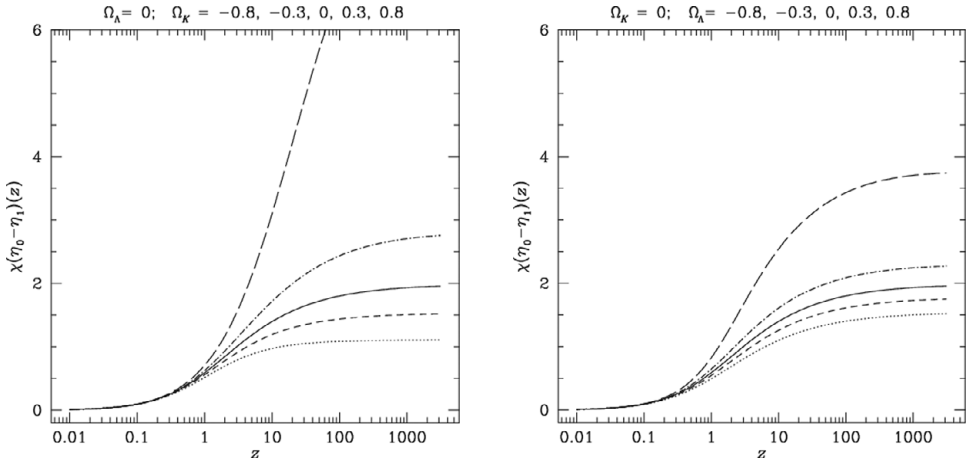

Fig. 1.3 The function $\chi(t_0 - t_1)$ as a function of the redshift z for different values of the cosmological parameters Ω_K (left, with $\Omega_\Lambda = 0$) and Ω_Λ (right, with $\Omega_K = 0$), namely -0.8 (dotted), -0.3 (short-dashed), 0 (solid), 0.3 (dot-dashed), 0.8 (long-dashed).

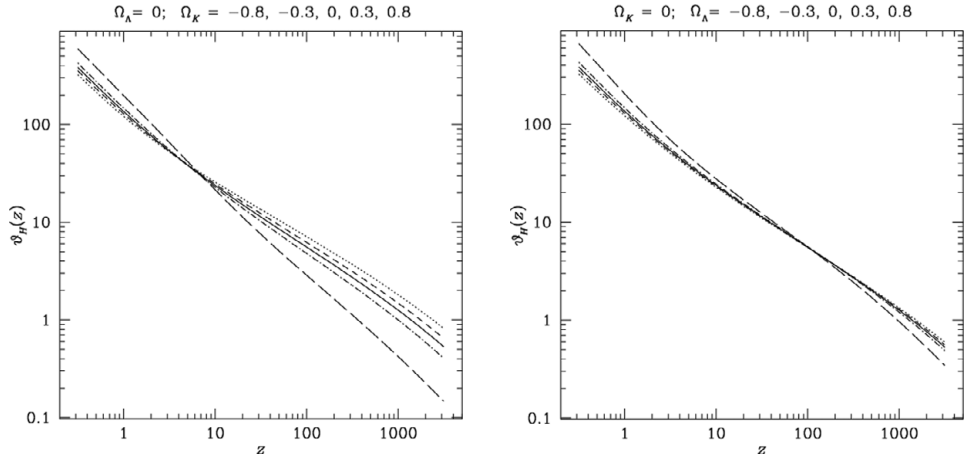

Fig. 1.4 $\vartheta_H(z_1)$ (in degrees) for different values of the cosmological parameters Ω_K and Ω_Λ. The line styles are as in Fig. 1.3.

In general, the integral in Eq. (1.48) has to be solved numerically. It determines the angle $\vartheta(\Delta, z) = \Delta / d_A(z)$ under which an object of size Δ placed at redshift z is seen (see Figs. 1.3 and 1.4).

If we are able to measure the redshifts and the angular extensions of a certain class of objects at different redshifts, of which we know the intrinsic size Δ, comparing with Eq. (1.48) allows us, in principle, to determine the parameters Ω_m, Ω_Λ, Ω_K, and H_0.

https://doi.org/10.1017/9781316471524.002 Published online by Cambridge University Press

Observationally we know for certain that $10^{-5} < \Omega_r \leq 10^{-4}$ as well as $0.1 \leq \Omega_m \lesssim 1$, $|\Omega_\Lambda| \lesssim 1$, and $|\Omega_K| \lesssim 1$.

If we are interested in small redshifts, $z_1 \lesssim 10$, we may therefore safely neglect Ω_r. In this region, Eq. (1.48) is very sensitive to Ω_Λ and provides an excellent mean to constrain the cosmological constant.

At high redshift, $z_1 \gtrsim 1000$, neglecting radiation is no longer a good approximation.

We shall later also need the opening angle of the *horizon* distance,

$$\vartheta_H(z_1) = \frac{t_1}{\chi(t_0 - t_1)}, \tag{1.49}$$

$$t_1 = \frac{1}{H_0 a_0} \int_{z_1}^{\infty} \frac{dz}{\left[\Omega_r (z+1)^4 + \Omega_m (z+1)^3 + \Omega_\Lambda + \Omega_K (z+1)^2\right]^{1/2}}. \tag{1.50}$$

(Clearly this integral diverges if $\Omega_r = \Omega_m = 0$. This is exactly what happens during an inflationary period and leads there to the solution of the horizon problem; see Section 1.5.)

Neglecting Ω_r, for $\Omega_\Lambda = 0$ and small curvature, $0 < |\Omega_K| < \Omega_m z_1$ at high enough redshift, $z_1 \geq 10$, one has $t_0 - t_1 \simeq 2\sqrt{|\Omega_K|/\Omega_m} = 2/(H_0 a_0 \sqrt{\Omega_m})$. With $\chi(x) \simeq x$, which is valid for small curvature, this yields $\vartheta(\Delta, z_1) \simeq \sqrt{\Omega_m} H_0 a_0 \Delta/(2a_1) = \frac{1}{2}\sqrt{\Omega_m} H_0 \Delta/(z_1+1)$ (see also Exercise 1.10).

Another important distance measure in cosmology is the luminosity distance. It is defined as follows. Let L be the luminosity (energy emitted per second) of a source at redshift z_1 and F its flux (energy received per second per square centimeter) arriving at the observer position. We define the luminosity distance to the source by

$$d_L(z_1) \equiv \left(\frac{L}{4\pi F}\right)^{1/2}. \tag{1.51}$$

We now want to show that $d_L(z_1) = (1+z_1)^2 d_A(z_1)$.

In a proper time interval of the emitter, $d\tau_1 = a(t_1)\,dt$, the source emits the energy $La(t_1)\,dt$. This energy is redshifted by a factor of $(1+z_1)^{-1} = a(t_1)/a(t_0)$. It is then distributed over a sphere with radius $a(t_0)\chi(t_0 - t_1)$. So that the flux per proper time of the observer $d\tau_0 = a(t_0)\,dt$ becomes

$$F = \frac{La^2(t_1)}{4\pi a^4(t_0)\chi^2(t_0 - t_1)},$$

https://doi.org/10.1017/9781316471524.002 Published online by Cambridge University Press

leading to

$$d_L(z_1) = \frac{a(t_0)^2}{a(t_1)}\chi(t_0 - t_1) = (1 + z_1)^2 d_A(z_1). \tag{1.52}$$

The luminosity distance hence contains two additional factors $(1 + z)$ compared to the angular diameter distance. One of them is due to the "redshift" of proper time and the other is due to the redshift of photon energy.

1.3 Recombination and Decoupling

We assume that, at sufficiently early times, reaction rates for particle interactions are much faster than the expansion rate, so that the cosmic fluid is in thermal equilibrium. During its expansion, the Universe then cools adiabatically. At early times, it is dominated by a relativistic radiation background with

$$\rho = C/a^4 = \frac{g_{\rm eff}}{2} a_{\rm SB} T^4. \tag{1.53}$$

This behavior implies that $T \propto a^{-1}$. Here $g_{\rm eff}$ is the effective number of degrees of freedom, which we define below and $a_{\rm SB}$ is the Stefan–Boltzmann constant, $a_{\rm SB} = \pi^2/15$ in our units. For massless (or extremely relativistic) fermions and bosons in thermal equilibrium at temperature T with N_b respectively N_f spin degrees of freedom we have (remember that we use units such that $\hbar = k_B = c = 1$)

$$\begin{aligned}\rho_b &= \frac{N_b 4\pi}{(2\pi)^3}\int_0^\infty \frac{p^3\,dp}{\exp(p/T) - 1} = \frac{N_b T^4}{2\pi^2}\int_0^\infty \frac{x^3\,dx}{\exp(x) - 1} \\ &= \frac{N_b T^4}{2\pi^2}\Gamma(4)\zeta(4) = \frac{N_b T^4 \pi^2}{30},\end{aligned} \tag{1.54}$$

$$\begin{aligned}\rho_f &= \frac{N_f 4\pi}{(2\pi)^3}\int_0^\infty \frac{p^3\,dp}{\exp(p/T) + 1} = \frac{N_f T^4}{2\pi^2}\int_0^\infty \frac{x^3\,dx}{\exp(x) + 1} \\ &= \frac{N_f T^4}{2\pi^2}\Gamma(4)\zeta(4)\frac{7}{8} = \frac{7}{8}\frac{N_f T^4 \pi^2}{30},\end{aligned} \tag{1.55}$$

where Γ denotes the Gamma-function and ζ is the Riemann zeta-function and we make use of the integrals (Gradshteyn and Ryzhik, 2000)

$$I_b(\alpha) = \int_0^\infty \frac{x^\alpha\,dx}{\exp(x) - 1} = \Gamma(\alpha + 1)\zeta(\alpha + 1), \tag{1.56}$$

$$I_f(\alpha) = \int_0^\infty \frac{x^\alpha\,dx}{\exp(x) + 1} = \left[1 - \left(\frac{1}{2}\right)^\alpha\right]\Gamma(\alpha + 1)\zeta(\alpha + 1). \tag{1.57}$$

https://doi.org/10.1017/9781316471524.002 Published online by Cambridge University Press

Furthermore, $\zeta(2) = \pi^2/6$, $\zeta(4) = \pi^4/90$, and $\Gamma(n) = (n-1)!$ for $n \in \mathbb{N}$; see Abramowitz and Stegun (1970).

Hence $\rho = \rho_b + \rho_f = \frac{g_{\rm eff}}{2} a_{SB} T^4$ for $a_{SB} = \pi^2 k_B^4/(15\,\hbar^3 c^2) = \pi^2/15$ and $g_{\rm eff} = N_b + 7/8 N_f$, if all the particles are at the same temperature T. If the temperatures are different, such as, for example, the neutrino temperature after electron–positron annihilation, this has to be taken into account with a factor $(T_\nu/T_\gamma)^4$ multiplying N_ν in $g_{\rm eff}$.

At temperatures below the electron mass, at $T < m_e \simeq 0.511$ MeV, only neutrinos and photons are still relativistic. Very recently, $T \lesssim 0.06$ eV at least some of the neutrinos also become nonrelativistic so that the density parameter of relativistic particles today is probably given only by the photon density,[3]

$$\Omega_{\rm rel} = \Omega_\gamma = \frac{8\pi G}{3H_0^2} a_{\rm SB} T_0^4 = 2.49 \times 10^{-5} h^{-2}. \tag{1.58}$$

Here we have set $T_0 = 2.725$ K. The present CMB temperature is the most precisely measured number in cosmology. Its value is (Fixsen, 2009)

$$T_0 = 2.72548 \pm 0.00057 \text{K}. \tag{1.59}$$

The pressure of relativistic particles is given by $P = T_i^i/3 = \rho/3$. The thermodynamic relation $dE = T\,dS - P\,dV$ therefore gives for the entropy density $s = dS/dV$

$$s = \frac{dS}{dV} = \frac{1}{T}\left(\frac{dE}{dV} + P\right) = \frac{\rho + P}{T} = \frac{4\rho}{3T}. \tag{1.60}$$

Using the expression for the energy density (1.54) and (1.55) this gives for each particle species X

$$s_X = \begin{cases} \frac{2\pi^2}{45} N_X T^3 & \text{for bosons,} \\ \frac{7\pi^2}{180} N_X T^3 & \text{for fermions.} \end{cases} \tag{1.61}$$

The particle density for relativistic particles is given by

$$n_X = \frac{N_X}{2\pi^2} \int \frac{p^2}{\exp(p/T) \pm 1}\, dp = \begin{cases} T^3 \frac{N_X}{\pi^2} \zeta(3) & \text{for bosons,} \\ T^3 \frac{N_X}{\pi^2} \zeta(3) \frac{3}{4} & \text{for fermions.} \end{cases} \tag{1.62}$$

[3] At present only neutrino mass differences are known from oscillation experiments. The lowest neutrino mass could still be zero, or at least lower than T_0. From oscillation experiments, however, we know that the heaviest neutrino mass is at least 0.05eV (see Olive *et al.*, 2014).

https://doi.org/10.1017/9781316471524.002 Published online by Cambridge University Press

The particle and entropy densities both scale like T^3. Using $\zeta(3) \simeq 1.202\,057$ we obtain

$$s_X \simeq \begin{cases} 3.6 \cdot n_X & \text{for bosons,} \\ 4.2 \cdot n_X & \text{for fermions.} \end{cases} \tag{1.63}$$

The photons obey a Planck distribution ($\epsilon = ap =$ the photon energy),

$$f(\epsilon) = \frac{1}{e^{\epsilon/T} - 1}. \tag{1.64}$$

At a temperature of about $T \sim 4000\,\mathrm{K} \sim 0.4$ eV, the number density of photons with energies above the hydrogen ionization energy ($= \Delta = 1\,\mathrm{Ry} = 13.6$ eV) drops below the baryon density of the Universe, and the protons begin to (re)combine to neutral hydrogen. Even though electrons and protons were not combined to neutral hydrogen before, this process is called "recombination" rather than "combination."

Helium has already recombined earlier. The binding energy of the first electron to the He nucleus is $4\Delta = 54.4$eV. Using the Saha equation derived in the next section for the transition $\mathrm{He}^{+2} \rightarrow \mathrm{He}^{+}$, one finds that the recombination of the first electron transition takes place at $T_{2\to1} \simeq 1.4 \times 10^4$K. The binding energy of the second electron to the He nucleus is 24.6 eV and, again using the Saha equation, one finds that the transition $\mathrm{He}^{+} \rightarrow$ He takes place at $T_{1\to0} \simeq 0.5 \times 10^4$K (see Exercise 1.5).

Before (re)combination photons and baryons are tightly coupled by Thomson scattering of electrons. During recombination the free electron density drops sharply and the mean free path of the photons grows larger than the Hubble scale. At the temperature $T_{\mathrm{dec}} \sim 3000\,\mathrm{K}$ (corresponding to the redshift $z_{\mathrm{dec}} \simeq 1100$ and the physical time $\tau_{\mathrm{dec}} \simeq a_{\mathrm{dec}} t_{\mathrm{dec}} \simeq 10^5$ yr) photons decouple from the electrons and the Universe becomes transparent. We now want to study this process in somewhat more detail.

1.3.1 The Physics of Recombination

From Eq. (1.63) with $N_\gamma = 2$ we obtain that the photon entropy is given by

$$s_\gamma = \frac{4\pi^2}{45} T^3 \simeq 3.6 n_\gamma.$$

The conserved baryon number n_B satisfies $a^3 n_B =$ constant; hence $n_B \propto a^{-3} \propto T^3$. The entropy per baryon is therefore a constant,

$$\sigma = s_\gamma / n_B = \frac{\frac{4\pi^2}{45} T_0^3}{\Omega_B \rho_c(\tau_0)/m_p} = 1.4 \times 10^8 \frac{T_{2.7}^3}{\Omega_B h^2}. \tag{1.65}$$

https://doi.org/10.1017/9781316471524.002 Published online by Cambridge University Press

Here we have used (see Appendix 1)

$$\rho_c(\tau_0) = 1.88h^2 \times 10^{-29}\,\mathrm{g\,cm^{-3}} = 8.1h^2 \times 10^{-11}\,(\mathrm{eV})^4,$$
$$m_p = 9.38 \times 10^8\ \mathrm{eV}, \qquad \text{(proton mass)},$$
$$T(\tau_0) = 2.3T_{2.7} \times 10^{-4}\ \mathrm{eV}, \quad T_{2.7} = T(\tau_0)/2.7\,\mathrm{K}.$$

As we shall see in the next section, the baryon density is approximately $\Omega_B h^2 \simeq 2.2 \times 10^{-2}$ so that $\sigma \simeq 10^{10}$. Correspondingly the ratio between the baryon and photon density is

$$\eta_B = n_B/n_\gamma = 2.7 \times 10^{-8}\left(\frac{\Omega_B h^2}{T_{2.7}^3}\right) \simeq 6 \times 10^{-10}. \tag{1.66}$$

As long as hydrogen is ionized, the timescale of interaction between photons and electrons (Thomson scattering) and between electrons and protons (Rutherford scattering) is much faster than expansion and we may therefore consider the latter as adiabatic. At every moment, the electron, proton, and photon plasma is in thermal equilibrium. As long as the temperature is above the ionization energy of neutral hydrogen, $T > 1\,\mathrm{Ry} = \Delta = \alpha^2 m_e/2 = 13.6$ eV, all hydrogen atoms that form are rapidly dissociated. Most electrons and protons are free and the neutral hydrogen density is very low. At some sufficiently low temperature, however, there will no longer be sufficiently many energetic photons around to disrupt neutral hydrogen and the latter becomes more and more abundant. To determine the temperature at which this transition, called "recombination,"[4] happens, we apply the standard rules of equilibrium statistical mechanics to the reaction

$$e^- + p \longleftrightarrow \mathrm{H} + \gamma\ (13.6\,\mathrm{eV}). \tag{1.67}$$

Supposing that pressure and temperature are fixed and only the number of free electrons, N_e; free protons, N_p; hydrogen atoms, N_H; and photons, N_γ, can change, the second law of thermodynamics implies that the Gibbs potential G is constant,

$$0 = dG = \mu_p\, dN_p + \mu_e\, dN_e + \mu_H\, dN_H + \mu_\gamma\, dN_\gamma,$$

Here μ_X denotes the chemical potential of species X. The different dN_X are not independent. Particle number conservation implies

$$dN_p + dN_H = dN_e + dN_H = 0. \tag{1.68}$$

As there is no conservation of photons, the chemical potential of photons vanishes, $\mu_\gamma = 0$. With this and Eq. (1.68) the Gibbs equation, $dG = 0$, implies

$$\mu_e + \mu_p - \mu_H = 0. \tag{1.69}$$

[4] The expression "combination" would be more adequate, since this is the first time that neutral hydrogen forms, but it is difficult to change historical misnamings. . . .

https://doi.org/10.1017/9781316471524.002 Published online by Cambridge University Press

In principle, this result is valid only in full thermal equilibrium. But Thomson scattering between electrons and photons does not change the photon energy and the Rydberg photons are not readily thermalized. They actually have time to ionize another hydrogen atom before they lose energy. Therefore, neglecting recombination into excited states of the hydrogen atom is a bad approximation, but it leads roughly to the right recombination temperature.

In this discussion, where we are more interested in the basic concepts than in accuracy, we also neglect helium that has recombined earlier. We set $n_p + n_H = n_B$, which induces an error of about 25%. For an accurate calculation of the final ionization fraction, one would have to take into account both the recombination of helium and the recombination into excited states of hydrogen. We briefly discuss this in Section 1.3.3. Despite these complications, a discussion of recombination into the ground state gives the correct orders of magnitude for the recombination and decoupling redshifts which we now derive.

In thermal equilibrium, electrons, protons and hydrogen atoms obey a Maxwell–Boltzmann distribution. Their number densities are given by (see Exercise 1.7)

$$n_e = \frac{2}{(2\pi)^3}(2\pi m_e T)^{3/2} \exp\left(-\frac{m_e - \mu_e}{T}\right), \tag{1.70}$$

$$n_p = \frac{2}{(2\pi)^3}(2\pi m_p T)^{3/2} \exp\left(-\frac{m_p - \mu_p}{T}\right), \tag{1.71}$$

$$n_H = \frac{4}{(2\pi)^3}(2\pi m_H T)^{3/2} \exp\left(-\frac{m_H - \mu_H}{T}\right). \tag{1.72}$$

We now make use of the fact that the Universe is globally neutral, $n_e = n_p$. Furthermore, the binding energy of hydrogen $\Delta = \alpha^2 m_e/2$ (here $\alpha \simeq 1/137$ is the fine structure constant) is given by $\Delta = m_e + m_p - m_H$. With this we obtain

$$\frac{n_e^2}{n_H} = \frac{n_e n_p}{n_H} = \left(\frac{m_e T}{2\pi}\right)^{3/2} e^{-\Delta/T}. \tag{1.73}$$

Here we have neglected the small difference between the hydrogen and proton mass in the second factor of Eqs. (1.71) and (1.72) but not in the exponential. This is the *Saha equation*. The corresponding equation for helium, setting $n_{\mathrm{He}^+} = n_{\mathrm{He}^{2+}}$, yields the $\mathrm{He}^{2+} \to \mathrm{He}^+$ transition temperature and accordingly the $\mathrm{He}^+ \to \mathrm{He}$ transition temperature (see Exercice 1.5).

We now define the ionization fraction x_e by $x_e \equiv n_e/(n_e + n_H)$. In Section 1.4, we shall find that about 25% of all baryons in the Universe are bound in the form of He^4 so that $n_p + n_H = n_e + n_H = 0.75 n_B$. Equation (1.73) then leads to

https://doi.org/10.1017/9781316471524.002 Published online by Cambridge University Press

$$\frac{x_e^2}{1-x_e} = \frac{n_e^2}{n_H(n_p+n_H)} = \frac{1}{0.75 n_B}\left(\frac{m_e T}{2\pi}\right)^{3/2} e^{-\Delta/T}. \tag{1.74}$$

Inserting the entropy per baryon, $\sigma = (4\pi^2/45)T^3/n_{\rm B}$, in this equation yields

$$\frac{x_e^2}{1-x_e} = \frac{45\sigma}{0.75\times 4\pi^2}\left(\frac{m_e}{2\pi T}\right)^{3/2} e^{-\Delta/T}. \tag{1.75}$$

At very high temperatures, $T \gg \Delta$, the ionization fraction x_e is close to 1. Recombination happens roughly when $\sigma \exp(-\Delta/T)$ is of the order of unity. If $\sigma \sim 1$ this corresponds to $T \sim \Delta$. The fact that the entropy per baryon is very large, $\sigma = 1.4\times 10^8 (\Omega_B h^2)^{-1} \sim 10^{10}$, delays recombination significantly. Since there are so many more photons than baryons in the Universe, even at a temperature much below $\Delta = 13.6$ eV there are still enough photons in the high-energy tail of the Planck distribution to keep the Universe ionized.

To be more specific, we define the recombination temperature $T_{\rm rec}$ as the temperature when $x_e = 0.5$ (as we shall see, the precise value is of little importance). Equation (1.75) then leads to

$$\left(\frac{T_{\rm rec}}{1\ {\rm eV}}\right)^{-3/2} e^{-\Delta/T_{\rm rec}} = 0.97\times 10^{-16}\ \Omega_B h^2. \tag{1.76}$$

For $\Omega_B h^2 \simeq 0.022$ we obtain

$$T_{\rm rec} = 3722\,{\rm K} = 0.321\,{\rm eV},\ \ z_{\rm rec} = 1353.$$

The function $x_e(T)$ is shown in Fig. 1.5. Clearly, this function grows very steeply from $x_e \sim 0$ to $x_e \sim 1$ at $T \sim 3700$ K and $T_{\rm rec}$ depends only weakly on the value chosen for $x_e(T_{\rm rec})$.

Interestingly, at temperature $T_{\rm rec}$ the baryon and photon densities are of the same order, $\rho_\gamma(T_{\rm rec}) \simeq \rho_B(T_{\rm rec})$. This seems to be a complete coincidence. More precisely, the ratio of these two densities is given by

$$\begin{aligned}\frac{\rho_\gamma}{\rho_B} &= \frac{(\pi^2/15)T^4}{n_B m_p} = \frac{\pi^2 T_0^4}{15 n_B(t_0) m_p}(z+1)\\ &\simeq 2\times 10^{-5}\left(\Omega_B h^2\right)^{-1}(z+1).\end{aligned} \tag{1.77}$$

This ratio is equal to 1 at redshift $z_{\gamma b}$ given by

$$(1+z_{\gamma b}) = 10^3\left(\frac{\Omega_B h^2}{2\times 10^{-2}}\right) \simeq 10^3 \sim 1+z_{\rm rec}. \tag{1.78}$$

https://doi.org/10.1017/9781316471524.002 Published online by Cambridge University Press

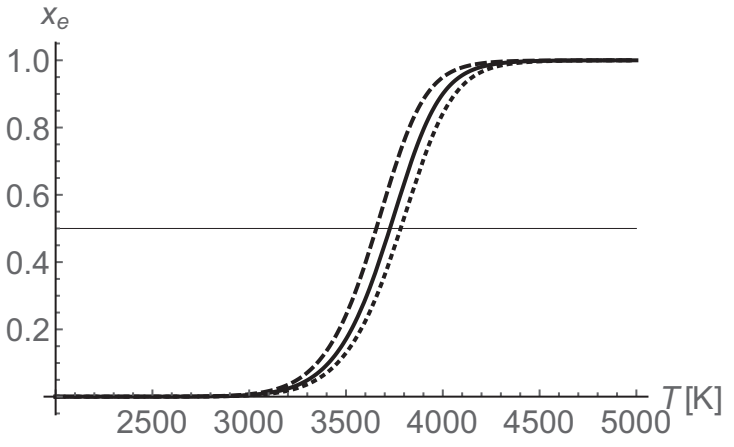

Fig. 1.5 The ionization fraction x_e as a function of the temperature is obtained via the Saha equation for $\Omega_B h^2 = 0.022$ (solid curve), for $\Omega_B h^2 = 0.01$ (dashed curve), and for $\Omega_B h^2 = 0.04$ (dotted curve). Our definition of recombination, $x_{\rm rec} = 0.5$, is indicated. Note that x decays from $x_e \simeq 1$ to $\simeq 0$ between $T = 4000$ and 3400 K. Below about $x_e \sim 0.9$ the shape of the true ionization fraction significantly differs from this Saha-equation result and levels off at the final ionization fraction computed in the text that follows.

1.3.2 Final Ionization and Photon Decoupling

We have determined the temperature at which electrons and protons recombine to neutral hydrogen. The Saha equation predicts an exponentially falling fraction of free electrons. But this is correct only as long as thermal equilibrium is established. As the free electron fraction drops, the interaction rate between electrons and protons decreases, and at some point the remaining free electrons and protons are too sparse to find each other, thermal equilibrium is lost, and the number of free electrons remains constant. But also the photon–electron interaction rate decreases. Whenever an interaction rate Γ drops below the expansion rate of the Universe,

$$\Gamma < H,$$

one considers the corresponding reaction as "frozen." It becomes negligible. The temperature at which $\Gamma = H$ is called the "freeze out" temperature of the reaction with rate Γ.

When the recombination rate drops below the expansion rate, recombination freezes out and the ionization fraction remains constant. When the scattering rate of photons on electrons falls below the expansion rate of the Universe, photons become free to propagate without further scattering. We want to calculate both the final ionization fraction, x_R, and the redshift, $z_{\rm dec}$, of the decoupling of photons.

https://doi.org/10.1017/9781316471524.002 Published online by Cambridge University Press

Let us first determine the temperature T_g at which the process of recombination freezes out. The cross section of the reaction $p^+ + e^- \rightarrow \mathrm{H} + \gamma$ is (see, e.g., Rybicki and Lightman, 1979)

$$\langle \sigma_R v \rangle \simeq 4.7 \times 10^{-24} \left(\frac{T}{1\,\mathrm{eV}} \right)^{-1/2} \mathrm{cm}^2. \tag{1.79}$$

Here v is the thermal electron velocity and we have used the fact that $3T = m_e v^2$. The reaction rate is therefore

$$\begin{aligned} \Gamma_R &= n_p \langle \sigma_R v \rangle = x_e \left(\frac{0.75 n_B}{n_\gamma} \right) n_\gamma \langle \sigma_R v \rangle \\ &\simeq 2.1 \times 10^{-10}\,\mathrm{cm}^{-1} \left(\frac{T}{1\,\mathrm{eV}} \right)^{7/4} \exp(-\Delta/2T)(\Omega_B h^2)^{1/2}, \end{aligned}$$

where we have inserted the Saha equation, assuming that the ionization fraction is much smaller than 1, that is,

$$x_e \simeq (\sqrt{45\sigma/0.75}/2\pi)(m_e/2\pi T)^{3/4} \exp(-\Delta/2T) \ll 1.$$

We have also used Eq. (1.66).

To determine the expansion rate $H(T)$, we neglect curvature or a possible cosmological constant, which is certainly a good approximation for all redshifts larger than, say, 5. We also assume that the Universe is matter dominated at freeze-out, which induces an error of about 15% in H. The Friedmann equation (1.18) then gives

$$\begin{aligned} H^2 &\simeq \frac{8\pi G}{3} \rho \simeq \frac{8\pi G}{3} \rho_0 (a_0/a)^3 \\ &= \frac{8\pi G}{3} \Omega_m \rho_c(t_0)(T/T_0)^3, \end{aligned}$$

so that

$$H \simeq 3 \times 10^{-23}\,\mathrm{cm}^{-1} (\Omega_m h^2)^{1/2} \left(\frac{T}{1\,\mathrm{eV}} \right)^{3/2}. \tag{1.80}$$

Equation (1.80) is a very useful formula, valid whenever the Universe is dominated by nonrelativistic matter or dust, $P \ll \rho$, and curvature or a cosmological constant are negligible.

The temperature T_g is defined by $\Gamma_R(T_g) = H(T_g)$, which finally leads to

$$\left(\frac{T_g}{1\,\mathrm{eV}} \right)^{1/4} e^{-\Delta/2T_g} = 1.4 \times 10^{-13} \left(\frac{\Omega_m}{\Omega_B} \right)^{1/2}. \tag{1.81}$$

https://doi.org/10.1017/9781316471524.002 Published online by Cambridge University Press

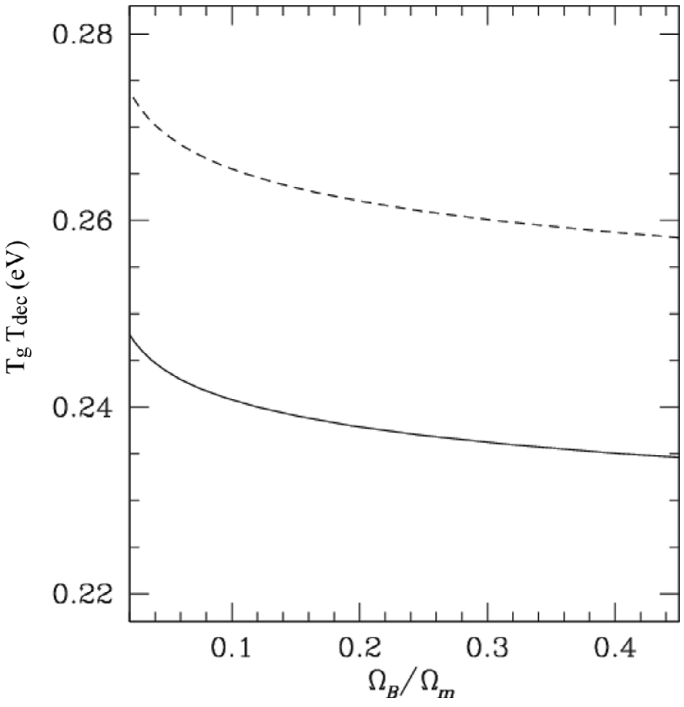

Fig. 1.6 The freeze-out temperatures of recombination, T_g (solid curve), and of Thomson scattering, $T_{\rm dec}$ (dashed curve), as functions of Ω_B/Ω_m.

This result is independent of h. For $\Omega_m \simeq 6.4\Omega_B$ (the value inferred from observations; see Planck Coll. XIII, 2016), we obtain $T_g \simeq 0.24$ eV and $z_g \simeq 1010$ (see Fig. 1.6). T_g depends only weakly on the ratio Ω_B/Ω_m.

The final ionization fraction is given by

$$x_R \simeq x_e(T_g) \simeq 7.3 \times 10^{-6} \left(\frac{T_g}{1\,\text{eV}}\right)^{-1} \Omega_m^{1/2}/(\Omega_B h) \simeq 3 \times 10^{-5}\Omega_m^{1/2}/(\Omega_B h). \tag{1.82}$$

A more detailed numerical analysis, taking into account the contribution of radiation to the expansion rate and, especially, the recombination into excited states of the hydrogen atoms and the presence of helium, gives $x_R \sim 1.2 \times 10^{-5}\Omega_m^{1/2}/(\Omega_B h)$ (Peebles, 1993; Mukhanov, 2005; Weinberg, 2008). We use this result to calculate the optical depth τ to Thomson scattering of photons by free electrons up to a redshift $z < z_g$ in a recombined universe. The optical depth to z is the scattering probability of a photon integrated from z until today. With the Thomson cross section

$$\sigma_T = \frac{8\pi}{3}\alpha^2 m_e^{-2} \simeq 6.65 \times 10^{-25}\,\text{cm}^2, \tag{1.83}$$

https://doi.org/10.1017/9781316471524.002 Published online by Cambridge University Press

one finds

$$\tau(z) \equiv \int_{t(z)}^{t_0} \sigma_T n_e a dt \simeq 0.046 x_R (1+z)^{3/2} \Omega_B \Omega_m^{-1/2} h. \tag{1.84}$$

With the residual ionization computed in Eq. (1.82) we obtain $\tau(z = 800) \simeq 0.01$. As we shall see in Section 9.3, the Universe is reionized at low redshift $z \sim 7.5$, which increases the optical depth by about a factor of 6. This rescattering of CMB photons is relevant for the evolution of fluctuations, as we shall discuss in Section 9.3.

As long as the temperature is larger than T_g, the reaction $p + e \longleftrightarrow \mathrm{H} + \gamma$ is in thermal equilibrium. When the temperature drops below T_g, the recombination process freezes out and the degree of ionization remains nearly constant.

Let us also note that in deriving the Saha equation (1.73), we used the fact that the process of recombination is in thermal equilibrium, which we have verified only now since freeze-out happens after recombination, $T_g < T_{\mathrm{rec}}$.

We finally calculate the redshift of the decoupling of photons. The process that remains effective longest is elastic Thomson scattering. Its rate is given by

$$\begin{aligned} \Gamma_T &= \sigma_T n_e = \sigma_T x_e \left(\frac{0.75 n_B}{n_\gamma}\right) n_\gamma \\ &\simeq 2.6 \times 10^{-11}\,\mathrm{cm}^{-1} (\Omega_B h^2)^{1/2} \left(\frac{T}{1\,\mathrm{eV}}\right)^{9/4} \exp(-\Delta/2T). \end{aligned} \tag{1.85}$$

Comparing it to the expansion rate, we find T_{dec}, which is defined by $H(T_{\mathrm{dec}}) = \Gamma_T(T_{\mathrm{dec}})$. A rough estimate gives $T_{\mathrm{dec}} \sim 0.26$ eV (see Fig. 1.6), which corresponds to $z_{\mathrm{dec}} \sim 1100$. Again we have assumed $x_e \ll 1$ in Eq. (1.85), which is justified since $T_{\mathrm{dec}} \sim 3000$ K (see Fig. 1.5).

Even though after z_{dec} photons decouple from electrons, the latter are still coupled to photons. The scattering rate of electrons, given by $\Gamma_e = \sigma_T x_e n_\gamma = \sigma_T x_R n_\gamma$ at low redshifts, is sufficient to keep the remaining electrons and with them baryonic matter in thermal equilibrium with the photons until about $z \sim 100$. Therefore, even after recombination the matter temperature is equal to the temperature of the CMB and does not decay like $1/a^2$, as would be expected from a pure thermal gas of massive particles (see Section 1.3.4). This is an example of two species, electrons and photons, where the former is in thermal equilibrium with the latter but not vice versa.

1.3.3 An Accurate Treatment of Recombination

So far, we have given an approximate treatment of the process of recombination and photon decoupling. This yields the correct orders of magnitude, but to determine

https://doi.org/10.1017/9781316471524.002 Published online by Cambridge University Press

especially the anisotropies of the CMB and the polarization, in Chapters 4 and 5, with good precision, this is not sufficient. For precise results it is necessary to treat recombination of hydrogen into higher levels, especially 2S, but also Raman scattering by which the electrons of a hydrogen atoms are scattered into a higher energy level and then decay into a lower level by the emission of a photon of different energy.

It is actually interesting to note that recombination into the ground state (1S) is not efficient at all because the ionization cross section is very high for the resonant Rydberg photons so that most of these just ionize another hydrogen atom before being redshifted out of the resonance, leading to no net recombination. The same is true for recombination into the 2P excited state. The Lyα photons from the 2P$\rightarrow$1S transition are quickly absorbed and excite another hydrogen atom, which is then reionized via a 2P-ionization photon. The single most efficient channel is the capture of electrons into the 2S level, from which they can decay into the ground state via the emission of two photons. By angular momentum conservation, the emission of a single photon is not possible. The inverse process, excitation from 1S to 2S, is a three-body process and therefore highly unlikely. Even though the rate of the transition $(e, p) \rightarrow \mathrm{H(2S)} \rightarrow \mathrm{H(1S)}$ is relatively low, it wins against direct recombination into the ground state and subsequent cosmological redshifting of the photon before the next ionization can take place. Since the binding energy of the 2S state is lower, this delays recombination somewhat. A semianalytic, dynamical treatment including recombination into the 2S and 2P states can be found in Peebles (1993), Mukhanov (2005), and Weinberg (2008).

For accurate results of helium and hydrogen recombination, as they are required to accurately study the CMB anisotropies discussed in the next chapters, a numerical computation is needed that takes into account many (up to 300) excited states and their decay. The most popular publicly available code for this is "RECFAST" (Seager *et al.*, 1999). The latest work (Shaw and Chluba, 2011), including even more details, has still found changes by up to 3% in the free electron fraction throughout the recombination process from $z \sim 2200$ (helium recombination) to $z \simeq 800$.

Interestingly, recombination also leads to lines and other distortions in the CMB frequency spectrum that might be observable with a future satellite mission measuring the CMB spectrum with high accuracy; see Rubino-Martin *et al.* (2006) and Wong *et al.* (2006).

1.3.4 Propagation of Free Photons and the CMB

After t_{dec}, photons cease any interaction with the cosmic fluid and propagate freely. It is straightforward to estimate that the cross section for Rayleigh scattering with hydrogen atoms is much too weak to be relevant (see Exercise 1.6).

https://doi.org/10.1017/9781316471524.002 Published online by Cambridge University Press

The free propagation of photons after decoupling is described with the Liouville equation for the photon distribution function, which we now develop. Since photons do not interact anymore, they simply move along geodesics. The Liouville equation translates this to a differential equation for the 1-particle distribution function f of the photons. The function f describes the particle density in the phase space P_0, the photon mass-shell, given by

$$P_0 = \{(x,p) \in T\mathcal{M} \mid g_{\mu\nu}(x)p^\mu p^\nu = 0\}, \quad f : P_0 \to \mathbb{R}.$$

The distribution function f gives the number of particles per phase space volume $|g|\, d^3x\, d^3p$ at fixed time t. In some general geometry a specific space-like hypersurface Σ has to be chosen and one then has to show that f does not depend on this choice [more details are found in Ehlers (1971) and Stewart (1971)]. In cosmology, due to the symmetries present, we simply use the hypersurfaces of constant time, $\Sigma = \Sigma_t$.

We choose the coordinates (x^μ, p^i) on the seven-dimensional mass-shell ($0 \le \mu \le 3$ and $1 \le i \le 3$). The energy p^0 is then determined by the mass-shell condition $g_{\mu\nu}(x)p^\mu p^\nu = 0$. Liouville's equation now says that the 1-particle distribution remains unchanged if we follow the geodesic motion of the particles, that is,

$$0 = \frac{df}{dt} = \dot{x}^\mu \partial_\mu f + \dot{p}^i \frac{\partial f}{\partial p^i},$$
$$0 = p^\mu \partial_\mu f - \Gamma^i_{\mu\nu} p^\mu p^\nu \frac{\partial f}{\partial p^i} \equiv L_{X_g} f\ . \tag{1.86}$$

A particle distribution obeying this equation is often also called a geodesic spray (see Abraham and Marsden, 1982). If the particles are not free, but collisions are so rare that an equilibrium description is not adequate, one uses the Boltzmann equation,

$$L_{X_g} f = C[f], \tag{1.87}$$

where $C[f]$ is the so-called collision integral, which depends on the details of the interactions.

It may be disturbing to some readers that we take over these concepts from non-relativistic physics so smoothly to the relativistic case. In cosmology, this does not cause any problems. But in general, it is true that the collision integral is not always well defined and certain conditions have to be posed to the nature of the spacetime and of the interaction. This problem has been studied in detail by Ehlers (1971).

Since the photons are massless, $|\mathbf{p}|^2 = \gamma_{ij} p^i p^i = (p^0)^2$. Here p^0 is the 0-component of the momentum 4-vector in *conformal* time so that $\epsilon = ap^0$ is

https://doi.org/10.1017/9781316471524.002 Published online by Cambridge University Press

the physical photon energy. Isotropy of the distribution implies that f depends on p^i only via $p \equiv |\mathbf{p}| = p^0$, and so

$$\frac{\partial f}{\partial p^i} = \frac{\partial p}{\partial p^i}\frac{\partial f}{\partial p} = \frac{p_i}{p}\frac{\partial f}{\partial p}. \tag{1.88}$$

Furthermore, f depends on x^i only through $p = \sqrt{\gamma_{ij}p^i p^i}$. Spatial derivatives are therefore given by

$$\begin{aligned} p^i \partial_i f &= \frac{1}{2} p^i \gamma_{lm,i}\frac{p^l p^m}{p}\frac{\partial f}{\partial p} = \frac{1}{2} p_j \gamma^{ij}\gamma_{lm,i}\frac{p^l p^m}{p}\frac{\partial f}{\partial p} \\ &= \frac{1}{2}\gamma^{ij}\left(\gamma_{li,m} + \gamma_{mi,l} - \gamma_{lm,i}\right)\frac{p_j p^l p^m}{p}\frac{\partial f}{\partial p} \\ &= \Gamma^j_{lm}\frac{p^l p^m p_j}{p}\frac{\partial f}{\partial p}. \end{aligned}$$

This leads to

$$p^i \partial_i f - \Gamma^i_{\mu\nu}\frac{p^\mu p^\nu p_i}{p}\frac{\partial f}{\partial p} = -\left(\Gamma^i_{j0} + \Gamma^i_{0j}\right)\frac{p^j p p_i}{p}\frac{\partial f}{\partial p} = -2p^2\frac{\dot{a}}{a}\frac{\partial f}{\partial p},$$

where we have used the expressions in Appendix 2, Section A2.3 for $\Gamma^i_{\mu\nu}$ and $p = p^0$. Inserting this result into (1.86) we obtain, with Eq. (1.88),

$$\partial_t f - 2p\frac{\dot{a}}{a}\frac{\partial f}{\partial p} = 0, \tag{1.89}$$

which is satisfied by an arbitrary function $f = f(pa^2) = f(a\epsilon)$. Hence the distribution of free-streaming photons changes only by redshifting the *physical energy* $\epsilon = ap^0$ or the *physical momentum* $a|\mathbf{p}| = \epsilon$. Therefore, setting $T \propto a^{-1}$ even after recombination, the blackbody shape of the photon distribution remains unchanged. This radiation of free photons with a perfect blackbody spectrum is the CMB. Its physics, especially its fluctuation and polarization, are the main topic of this book.

The same result is also obtained for massive particles,

$$\partial_t f - 2p\frac{\dot{a}}{a}\frac{\partial f}{\partial p} = 0, \tag{1.90}$$

where $p = |\mathbf{p}|$; hence the momentum is simply redshifted. Therefore, massive particles that decouple when they are still relativistic keep their extremely relativistic Fermi–Dirac (or Bose–Einstein) distribution, $f = (\exp(ap/T) \pm 1)^{-1}$, with a temperature that simply scales as $T \propto 1/a$. This is especially important for the cosmic neutrinos, which probably have masses in the range of a $0.1\text{eV} > m_\nu \gtrsim 0.01$ eV. But, as we shall see in the next section, they decouple at $T \sim 1.4$ MeV. We therefore expect them to be distributed according to an extremely

https://doi.org/10.1017/9781316471524.002 Published online by Cambridge University Press

relativistic Fermi–Dirac distribution, which is not a thermal distribution for nonrelativistic neutrinos. By the same argument, particles that decouple once they are nonrelativistic keep their Maxwell Boltzmann distribution, $f \propto \exp\left[(ap)^2/(mT)\right]$, if we assume the temperature to scale as $T \propto a^{-2}$, which is also the scaling in thermal equilibrium for massive particles [see discussion after Eq. (1.93)].

Note, however, that after decoupling the particles are no longer in thermal equilibrium and the T in their distribution function is not a temperature in the thermodynamical sense but merely a parameter, representing a measure of the mean kinetic energy.

The situation is different for the electron–proton–hydrogen plasma. As we have seen, the free electrons still scatter with photons and keep the same temperature as the latter. In other words: even though most photons are no longer interacting with the electrons, the latter are still interacting with the photons. (To have one collision with all the remaining electrons, only a fraction of about 10^{-14} of the photons have to be involved!)

Soon after recombination, the baryon energy density exceeds the photon energy density and one might expect that this would change the evolution of the temperature. To investigate this we use the energy conservation equation of the baryon–photon system. We neglect the tiny number of free electrons. The energy density and pressure are then given by

$$\rho = n_B m_B + (3/2) n_B T + \frac{\pi^2}{15} T^4, \tag{1.91}$$

$$p = n_B T + \frac{\pi^2}{45} T^4. \tag{1.92}$$

The energy conservation equation, $d\rho/da = -3(\rho + p)/a$, now gives

$$\frac{a}{T}\frac{dT}{da} = -\frac{3n_B + \frac{4\pi^2}{15}T^3}{(3/2)n_B + \frac{4\pi^2}{15}T^3} = -\frac{\sigma + 1}{\sigma + 1/2}. \tag{1.93}$$

Since $\sigma \gg 1$, the photons are so much more numerous than the baryons that the latter have no influence on the temperature, which keeps evolving as $1/a$. Note, however, that in the absence of photons, the temperature of a monoatomic gas would decrease like $1/a^2$ as mentioned earlier (just consider the limit $\sigma \to 0$).

The blackbody spectrum of the CMB photons is extremely well verified observationally (see Fig. 1.7 and Chapter 10). The limits on deviations are often parameterized in terms of three parameters: the chemical potential μ, the Compton-y parameter (which quantifies a well-defined change in the spectrum arising from interactions with a nonrelativistic electron gas at a different temperature; we

https://doi.org/10.1017/9781316471524.002 Published online by Cambridge University Press

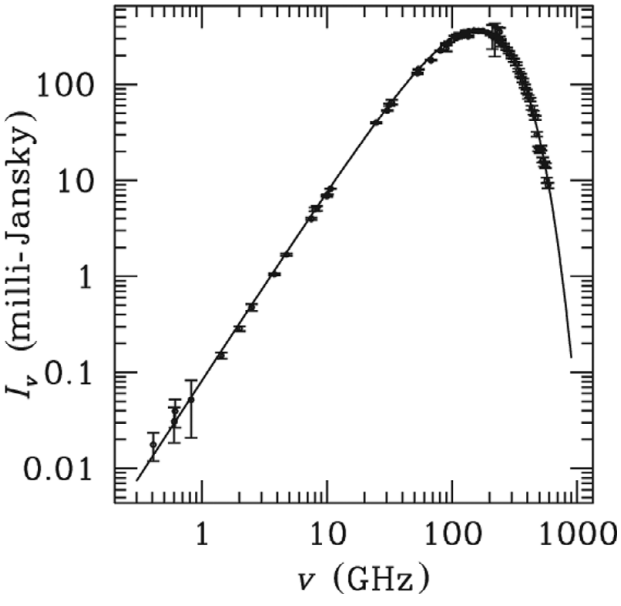

Fig. 1.7 The spectrum of the cosmic background radiation. I_ν is the energy flux per frequency. The data are from many different measurements that are all compiled in Kogut *et al.* (2007). The points around the top are the measurements from the FIRAS experiment on COBE (Fixsen *et al.*, 1996). The line traces a blackbody spectrum at a temperature of 2.728 K (the data are courtesy of Susan Staggs).

shall discuss this in detail in Chapter 10), and $Y_{\rm ff}$ (describing a contamination by free–free emission).

The present 95% confidence limits on these parameters are (Particle Data Group, 2006)

$$|\mu| < 9 \times 10^{-5}, \quad |y| < 1.2 \times 10^{-5}, \quad |Y_{\rm ff}| < 1.9 \times 10^{-5}. \tag{1.94}$$

These limits are mainly derived from the COBE satellite data, which had been taken more than 25 years ago. It would be very interesting to have newer data and better limits on these spectral distortions, as we will discuss in Chapter 10.

The CMB photons not only have a very thermal spectrum, but they are also distributed very isotropically, apart from a dipole that is (most probably) mainly due to our motion relative to the surface of last scattering.

Indeed, an observer moving with velocity $\mathbf{v}$ relative to a source in direction $\mathbf{n}$ emitting a photon with proper momentum $\mathbf{p} = -\epsilon\mathbf{n}$ sees this photon redshifted with frequency

$$\epsilon' = \gamma\epsilon\,(1 - \mathbf{nv})\,, \tag{1.95}$$

https://doi.org/10.1017/9781316471524.002 Published online by Cambridge University Press

where $\gamma = 1/\sqrt{1-v^2}$ is the relativistic γ-factor. For an isotropic emission of photons coming from all directions **n** this leads to a dipole anisotropy to first order in **v**. This dipole anisotropy, which is of the order of

$$\left(\frac{\Delta T}{T}\right)_{\text{dipole}} \simeq 1.2 \times 10^{-3},$$

had already been discovered in the 1970s (Conklin, 1969; Henry, 1971). Interpreting it as due to our motion with respect to the last scattering surface implies a velocity for the Solar System barycenter of $v = 371 \pm 0.5\,\text{km}\,\text{s}^{-1}$ at 68% CL (Particle Data Group, 2006).

In addition to the dipole, the COBE[5] DMR experiment (differential microwave radiometer) has found fluctuations of order

$$\sqrt{\left\langle\left(\frac{\Delta T}{T}\right)^2\right\rangle} \sim (\text{a few}) \times 10^{-5}, \tag{1.96}$$

on all angular scales $\theta \geq 7^\circ$ (Smoot *et al.*, 1992). On smaller angular scales many experiments found fluctuations (we shall describe the experimental results in more detail later) all of which satisfy $|\Delta T/T| \lesssim 10^{-4}$.

As we shall see in Chapter 2, the CMB fluctuations on large scales provide a measure for the deviation of the geometry from the Friedmann–Lemaître one. The geometry perturbations are thus small, and we may calculate their effects by *linear perturbation theory*. On smaller scales, $\Delta T/T$ reflects the fluctuations in the energy density in the baryon/radiation plasma prior to recombination. Their amplitude is just about right to allow the formation of the presently observed nonlinear structures (such as galaxies, clusters, etc.) by gravitational instability.

These findings strongly support our hypothesis that the large-scale structure (i.e., the galaxy distribution) observed in the Universe has been formed by gravitational instability from small ($\sim 10^{-4}$) initial fluctuations. As we shall see in Chapters 2, 4, and 5, such initial fluctuations leave an interesting "fingerprint" on the cosmic microwave background.

1.4 Nucleosynthesis

1.4.1 Expansion Dynamics at $T \sim$ a Few MeV

At high temperatures, $T > 30$ MeV, none of the light nuclei (deuterium, ^{2}H, helium-4, ^{4}He, helium-3, ^{3}He or lithium, ^{7}Li) are stable. At these temperatures, we expect the baryons to form a simple mixture of protons and neutrons in thermal

[5] Cosmic Background Explorer, NASA satellite launched 1990.

https://doi.org/10.1017/9781316471524.002 Published online by Cambridge University Press

equilibrium with each other and with electrons, photons, and neutrinos. The highest binding energy is the one of ^{4}He, which is about 28 MeV. Nevertheless, ^{4}He cannot form at this temperature because the baryon density of the Universe is not high enough for three- or even four-body interactions to occur in thermal equilibrium. Therefore, before any nucleosynthesis can occur, the temperature has to drop below the binding energy of deuterium, which is about 2.2 MeV. But even at this temperature there are still far too many high-energy photons around for deuterium to be stable. This is due to the very low baryon to photon ratio, $\eta_B \simeq 10^{-10}$. Just as recombination is delayed from the naively expected temperature $T = 13.7$ eV to about $T_{\rm rec} \sim 0.3$ eV, nucleosynthesis does not happen at $T \sim 2.2$ MeV but around $T_{\rm nuc} \sim 0.1$ MeV. Most of the neutrons present at that temperature are converted into ^{4}He. Only small traces remain as deuterium or are burned into ^{3}He and ^{7}Li.

Let us study this in some more detail. At the time of recombination, the relativistic particle species are the photon and, probably, three types of neutrinos. As we shall see in the next paragraph, the neutrino temperature is actually a factor of $(4/11)^{1/3}$ lower than the temperature of the photons. With Eqs. (1.54) and (1.55), the energy density of these particles while they are relativistic is given by

$$\rho_{\rm rel}(t) = \left[\rho_\gamma(t) + \rho_\nu(t)\right] = \left[1 + 3\frac{7}{8}(4/11)^{4/3}\right]\frac{\pi^2}{15}T^4, \tag{1.97}$$

$$\simeq 10^{-33}\,{\rm g\,cm^{-3}}\left(\frac{T}{T_0}\right)^4, \tag{1.98}$$

$$\simeq \rho_c(t_0)\Omega_{\rm rel}h^2(1+z)^4, \text{ where}$$

$$\Omega_{\rm rel}h^2 \simeq 4.4\times 10^{-5}\,. \tag{1.99}$$

Note that at temperatures below the highest neutrino mass, this is no longer the energy density of relativistic particles; therefore $\Omega_{\rm rel}$ is not the density parameter of relativistic particles today. Above the neutrino mass threshold and below the electron mass threshold we have

$$\frac{\rho_{\rm rel}}{\rho_m} = \frac{\Omega_{\rm rel}}{\Omega_m}(1+z) \simeq 4.4\times 10^{-5}\left(\frac{1}{\Omega_m h^2}\right)(1+z), \tag{1.100}$$

Since $\Omega_m h^2 \simeq 0.14$, the redshift $z_{\rm eq}$ above which the Universe is dominated by relativistic particles is about

$$z_{\rm eq} \simeq 3.2\times 10^3, \quad T_{\rm eq} \simeq 1\,{\rm eV}. \tag{1.101}$$

At temperatures significantly above $T_{\rm eq}$, we can also neglect a possible contribution from curvature or a cosmological constant to the expansion of the Universe, so that for

$$z \gg z_{\rm eq} \qquad P = \frac{1}{3}\rho, \quad a \propto \tau^{1/2} \propto t. \tag{1.102}$$

https://doi.org/10.1017/9781316471524.002 Published online by Cambridge University Press

At these high temperatures the energy density of the Universe is given by

$$\rho = g_{\rm eff}\frac{\pi^2}{30}T^4 \quad \text{where} \quad g_{\rm eff} = N_B(T) + \frac{7}{8}N_F(T). \tag{1.103}$$

Here, N_B and N_F denote the number of bosonic and fermionic degrees of freedom of relativistic particles (i.e., particles with mass $m < T$) that are in thermal equilibrium at temperature T.

To discuss the physical processes at work at some temperature T, we need to know the spectrum of relativistic particles and their interactions at this temperature. Here, we shall study the Universe at $10\,{\rm keV} < T < 100\,{\rm MeV}$, where the physics is well known. The only relativistic particles present at these temperatures are electrons, positrons, photons, and three types of neutrinos. (The muons have a mass of $m_\mu \simeq 105.66$ MeV.) Even if the individual neutrino masses are not very well constrained, the oscillation experiments (Particle Data Group, 2004) imply that their masses are below 1 eV if there is no degeneracy. As we shall see later, also CMB data estimate masses below this value. Therefore, we may neglect the neutrino masses in our treatment. The baryon number is well conserved at these temperatures, so that we may set η_B equal to its present value, $\eta_B = n_B/n_\gamma \simeq 2.7 \times 10^{-8}\Omega_B h^2 = $ constant. We neglect the small contribution from muon/anti-muon pairs that decay exponentially $\propto \exp(-m_\mu/T)$ via the reaction

$$\mu + \bar{\nu}_\mu \to e + \bar{\nu}_e.$$

Thermal equilibrium between photons and electron/positrons is maintained mainly via the process $e^- + e^+ \longleftrightarrow 2\gamma$ (or $3\gamma\ldots$). The conservation of the chemical potential during this reaction implies

$$\mu_e + \bar{\mu}_e = 2\mu_\gamma = 0. \tag{1.104}$$

The last equals sign comes from the fact that photons can be generated and destroyed, their number is not conserved, and hence their chemical potential vanishes in thermal equilibrium. Here we use the notation $e^+ = \bar{e}$ and $\mu_{\bar{e}} = \bar{\mu}_e$. The difference in the density of electrons and positrons is therefore

$$n_e - \bar{n}_e = \frac{1}{\pi^2}\int p^2\,dp\left[\frac{1}{\exp\left(\frac{E-\mu_e}{T}\right)+1} - \frac{1}{\exp\left(\frac{E+\mu_e}{T}\right)+1}\right]. \tag{1.105}$$

At low temperatures this number is dictated by the neutrality of the Universe, and $n_e - \bar{n}_e \sim n_B$ is much smaller than $n_e + \bar{n}_e \sim n_\gamma$. Therefore, the chemical potential is much smaller than the electron mass, $\mu_e \ll m_e$. At high temperatures, $T \gg m_e$,

https://doi.org/10.1017/9781316471524.002 Published online by Cambridge University Press

we may therefore expand the electron number density in the small parameter μ_e/T. At first order this yields

$$n_e - \bar{n}_e \simeq \frac{2\mu_e}{\pi^2 T} \int p^2\, dp \frac{\exp(p/T)}{\left[\exp(p/T)+1\right]^2} = \frac{2\mu_e T^2}{\pi^2}\zeta(2). \tag{1.106}$$

With $n_\gamma = 2T^3\zeta(3)/\pi^2$ this yields

$$\frac{n_e - \bar{n}_e}{n_\gamma} \simeq 1.4\frac{\mu_e}{T} \sim \frac{n_B}{n_\gamma} \simeq 2.7 \times 10^{-8}\Omega_B h^2. \tag{1.107}$$

We can therefore neglect the small chemical potential of the electrons and positrons. The interaction $e + \bar{e} \longleftrightarrow \nu + \bar{\nu}$ also implies that $\mu_\nu = -\bar{\mu}_\nu$. But unfortunately, the number $n_\nu - \bar{n}_\nu$ that determines, together with $n_e - \bar{n}_e$, the lepton number of the Universe, is not known from observations. We suppose that the lepton number, like the baryon number, is small and that we may also neglect the chemical potential of the neutrinos. Comparing our results with observations, we can check this hypothesis later.

At $T \lesssim 100$ MeV photons, electron/positrons, and neutrinos are still relativistic, so that $N_B = 2$ and $N_F = 4 + 6$; hence

$$g_{\text{eff}}(T \sim 100\,\text{MeV}) = \frac{43}{4} = 10.75. \tag{1.108}$$

The Hubble parameter is given by

$$\left(\frac{a'}{a}\right)^2 = H^2 = \frac{1}{4\tau^2} = \frac{8\pi G}{3}\rho = \frac{8\pi^3 G}{90} g_{\text{eff}} T^4.$$

With the Planck mass, m_P, defined by $G = 1/m_P^2 = 1/(1.22 \times 10^{19}\,\text{GeV})^2$, we find

$$H^2(T) \simeq 2.76 g_{\text{eff}}(T) \left(\frac{T^2}{m_P}\right)^2, \tag{1.109}$$

$$H \simeq 0.21\sqrt{g_{\text{eff}}} \left(\frac{T}{1\,\text{MeV}}\right)^2 \text{s}^{-1}, \tag{1.110}$$

$$\tau = \frac{1}{2H} \simeq 0.3 g_{\text{eff}}(T)^{-1/2} \left(\frac{m_P}{T^2}\right) \simeq 2.4\,\text{s} \left(\frac{1\,\text{MeV}}{T}\right)^2 g_{\text{eff}}^{-1/2}. \tag{1.111}$$

Here we have used the formulas in Appendix 1 to convert MeV's to seconds, $1\ \text{MeV} = 1.5192 \times 10^{21}\text{s}^{-1}$. The temperature of $T \sim 100$ MeV corresponds thus to an age of $\tau \sim 7 \times 10^{-5}$ s, and $T = 1$ MeV corresponds to $\tau \sim 0.7$ s. The relations (1.110) and (1.111) can be applied as long as the Universe is dominated by relativistic particles.

https://doi.org/10.1017/9781316471524.002 Published online by Cambridge University Press

1.4.2 Neutrino Decoupling

Neutrinos are kept in thermal equilibrium via the exchange of a W-boson, $e + \bar{\nu} \longleftrightarrow e + \bar{\nu}$ and $\nu + \bar{e} \longleftrightarrow \nu + \bar{e}$, or a Z-boson, $e + \bar{e} \longleftrightarrow \nu + \bar{\nu}$. At low energies, $E \ll m_{Z,W} \sim 100$ GeV, we can determine the cross sections within the 4-fermion theory of weak interaction. Within this approximation, the effective interaction Langrangian is given by

$$\begin{aligned} L_{\text{int}} &= \frac{G_F}{\sqrt{2}} J_\mu^\dagger J^\mu + \text{hermitean conjugate} \\ &= \frac{G_F}{\sqrt{2}} \left(u_e^* \gamma_\mu \frac{1}{2}(1-\gamma^5) u_\nu \right) \left(u_\nu^* \gamma^\mu \frac{1}{2}(1-\gamma^5) u_e \right) + \text{h.c.}, \end{aligned} \tag{1.112}$$

where the coupling parameter, G_F, is the Fermi constant, and γ^μ are Dirac's gamma-matrices, $\gamma_5 = i\gamma^0\gamma^1\gamma^2\gamma^3$.

$$G_F = 1.166 \times 10^{-5}\,\text{GeV}^{-2} = (293\,\text{GeV})^{-2}. \tag{1.113}$$

The fermion $V - A$ current J_μ is expressed in terms of the electron and neutrino spinors $u_{e,\nu}$ and the Dirac γ-matrices.

The cross section of the W- and Z-boson exchange processes are identical within this approximation and they are given by

$$\sigma_F \simeq G_F^2 E^2 \sim G_F^2 T^2,$$

The involved particle density is $n_F(T) = g_F(T)\zeta(3)T^3/\pi^2 \sim 1.3T^3$, where we have set $g_F(T) = 3/4 N_F(T) = 30/4$ for the three types of left-handed neutrinos and the $e^\pm$ s. Since the particles are relativistic, we can set $v \sim 1$ so that we obtain an interaction rate of

$$\Gamma_F = \langle \sigma_F v \rangle n_F \simeq 1.3 G_F^2 T^5.$$

Comparing this with the expansion rate H obtained in (1.109), we find

$$\frac{\Gamma_F}{H} \simeq 0.24 T^3 m_P G_F^2 \simeq \left(\frac{T}{1.4\,\text{MeV}} \right)^3. \tag{1.114}$$

At temperatures below $T_F \sim 1.4$ MeV the mean number of interactions of a neutrino within one Hubble time, H^{-1}, becomes less than unity and the neutrinos effectively decouple. The plasma becomes transparent to neutrinos that are no longer in thermal equilibrium with electrons and positrons and hence photons and baryons.

https://doi.org/10.1017/9781316471524.002 Published online by Cambridge University Press

As we have discussed in the previous section, even at temperatures far below their mass $m_\nu \gtrsim 0.05$ eV, their particle distribution remains an extremely relativistic Fermi–Dirac distribution with temperature

$$T_\nu = T_F \frac{a_F}{a},$$

since they are no longer in thermal equilibrium and their distribution is affected solely by redshifting of the momenta.

As long as the photon/electron/baryon temperature also scales like $1/a$, the neutrinos conserve the same temperature as the thermal plasma, but when the number of degrees of freedom, g_{eff}, changes, the plasma temperature decays for a brief period of time less rapidly than $1/a$ and therefore remains higher than the neutrino temperature. This is exactly what happens at the electron–positron mass threshold, $T = m_e \simeq 0.5$ MeV. Below that temperature, only the process $e + \bar{e} \to 2\gamma$ remains in equilibrium while $2\gamma \to e + \bar{e}$ is exponentially suppressed. We calculate the reheating of the photons gas by electron–positron annihilation, assuming that the process takes place in thermal equilibrium and that the entropy remains unchanged. This is well justified because the cross section of this process is very high. Denoting the entropy inside a volume of size Va^3 before and after electron–positron annihilation by S_i and S_f, we therefore have $S_i = S_f$. With Eq. (1.60) this yields

$$S_i = \frac{2}{3} a_{\text{SB}}\, g_{\text{eff},i} (Ta)_i^3 V, \quad S_f = \frac{2}{3} a_{\text{SB}}\, g_{\text{eff},f} (Ta)_f^3 V.$$

The electron–positron degrees of freedom disappear in this process so that $g_{\text{eff},f} = 2$ while $g_{\text{eff},i} = 2 + 4(\frac{7}{8}) = 11/2$. From $S_i = S_f$ we therefore conclude

$$(Ta)_f = (Ta)_i \left(\frac{11}{4}\right)^{1/3}.$$

The neutrino temperature is not affected by $e^\pm$ annihilation, so that $(T_\nu a)_f = (T_\nu a)_i = (Ta)_i$. For the last equals sign we have used that the neutrino and photon temperatures are equal before $e^\pm$ annihilation. At temperatures $T \ll m_e$ we therefore have

$$T = \left(\frac{11}{4}\right)^{1/3} T_\nu. \tag{1.115}$$

Since there are no further annihilation processes, this relation remains valid until today and the present Universe not only contains a thermal distribution of photons, but also a background of cosmic neutrinos that have an extremely relativistic Fermi–Dirac distribution with temperature

$$T_\nu(\tau_0) = (4/11)^{1/3} T_0 = 1.95\,\text{K}. \tag{1.116}$$

https://doi.org/10.1017/9781316471524.002 Published online by Cambridge University Press

We set

$$g_0 = 2 + \frac{7}{8}6\left(\frac{4}{11}\right)^{4/3} \simeq 3.36, \quad \text{and} \tag{1.117}$$

$$g_{0S} = 2 + \frac{7}{8}6\left(\frac{4}{11}\right) \simeq 3.91. \tag{1.118}$$

These are respectively the effective degrees of freedom of the energy and entropy densities as long as all the neutrinos are relativistic. Until then we therefore have

$$\rho_{\rm rel}(T) = \frac{\pi^2}{30} g_0 T^4 \simeq 8.1 \times 10^{-34}\,{\rm g\,cm^{-3}} \left(\frac{T}{T_0}\right)^4, \tag{1.119}$$

$$s(T) = \frac{2\pi^2}{45} g_{0S} T^3 \simeq 3 \times 10^3\,{\rm cm^{-3}} \left(\frac{T}{T_0}\right)^3. \tag{1.120}$$

The neutrino cross section at low energies is extremely weak, and so far the neutrino background has not been observed directly (see Exercise 1.9).

1.4.3 The Helium Abundance

The observed abundance of helium is universally about

$$\frac{n_{He}\, m_{He}}{n_H\, m_H} \equiv Y \simeq 0.24. \tag{1.121}$$

It is well known that this amount of helium cannot have been produced in stars. We now want to investigate how much helium is produced in the primordial Universe. At temperatures of a few MeV nuclei and baryons are non-relativistic and the equilibrium distribution for a nucleus with atomic mass (i.e., number of protons and neutrons) A and proton number Z is given by

$$n_A = N_A \left(\frac{m_A T}{2\pi}\right)^{3/2} \exp\left(-\frac{m_A - \mu_A}{T}\right). \tag{1.122}$$

The proton density is given in Eq. (1.71). The neutron density is correspondingly

$$n_n = 2\left(\frac{m_B T}{2\pi}\right)^{3/2} \exp\left(-\frac{m_n - \mu_n}{T}\right). \tag{1.123}$$

Here, we neglect the small difference $Q = m_n - m_p = 1.293$ MeV in the prefactor, setting $m_n \sim m_p \sim m_B$. The conservation of the chemical potentials in nuclear reactions implies

$$\mu_A = Z\mu_p + (A - Z)\mu_n,$$

https://doi.org/10.1017/9781316471524.002 Published online by Cambridge University Press

so that

$$\exp\left(-\frac{m_A-\mu_A}{T}\right)=\left(e^{\mu_p/T}\right)^Z\left(e^{\mu_n/T}\right)^{(A-Z)}e^{-m_A/T},$$
$$=\frac{1}{2^A}\left(\frac{2\pi}{m_BT}\right)^{3A/2}\exp(B_A/T)n_p^Zn_n^{A-Z}.$$

Here, $B_A=Zm_p+(A-Z)m_n-m_A$ is the binding energy of the nucleus (A,Z). In thermal equilibrium, the density of this ion is then given by

$$n_A=\frac{N_A}{2^A}A^{3/2}\left(\frac{2\pi}{m_BT}\right)^{3(A-1)/2}n_p^Zn_n^{A-Z}\exp(B_A/T). \tag{1.124}$$

Here we have again neglected the nucleon mass difference Q and the binding energy B_A in the prefactor by setting $m_A\sim Am_B$, but not in the exponential.

We define the various mass abundances by

$$Y_A\equiv\frac{An_A}{n_B}=\frac{An_A}{\eta_Bn_\gamma},$$
$$Y_p\equiv\frac{n_p}{n_B}=\frac{n_p}{\eta_Bn_\gamma},$$
$$Y_n\equiv\frac{n_n}{n_B}=\frac{n_n}{\eta_Bn_\gamma}.$$

Hence the thermal abundance of the nucleus (A,Z) is given by

$$Y_A=F(A)\left(\frac{T}{m_B}\right)^{3(A-1)/2}\eta_B^{A-1}Y_p^ZY_n^{A-Z}e^{B_A/T}, \tag{1.125}$$

$$\text{where}\quad F(A)=N_AA^{5/2}\zeta(3)^{A-1}\pi^{-(A-1)/2}2^{(3A-5)/2}. \tag{1.126}$$

This equation shows nicely the influence of the radiation entropy on nucleosynthesis. If we had $\eta_B\sim 1$, the nucleus (A,Z) would become stable and relatively abundant at $T\sim B_A$. At this temperature the formation of (A,Z) [controlled by the factor $\exp(B_A/T)$] is sufficiently important to counterbalance photodissociation (controlled by the factor η_B^{A-1}). In equilibrium, the exponential $\exp(B_A/T)$ is then of the order of $\eta_B^{1-A}\sim 1$ and the ratio Y_A then approaches the value $Y_A\sim Y_p^ZY_n^{A-Z}$. However, if η_B is very small, the equilibrium between production of (A,Z) and photodissociation is delayed until $\exp(-B_A/T)\sim\eta_B^{A-1}\ll 1$, that is, to much lower temperatures. Neglecting the numerical factor $F(A)$, the temperature T_A, defined by $Y_A(T_A)\sim Y_p(T_A)^ZY_n(T_A)^{A-Z}$, is

$$T_A\sim\frac{B_A}{(A-1)\left[\ln(\eta_B^{-1})+3/2\ln(m_B/T_A)\right]}.$$

https://doi.org/10.1017/9781316471524.002 Published online by Cambridge University Press

For the deuteron with binding energy $B_2 = 2.22$ MeV we find

$$T_2 \sim 0.085\,\text{MeV}. \tag{1.127}$$

The reaction rate Γ_{np} of the process $n + p \longleftrightarrow {}^2\text{H} + \gamma$ is given by

$$\Gamma_{np} = \langle \sigma_{np} v \rangle n_p \simeq 1.8 \times 10^{-17} (T/T_0)^3 \eta_B \,\text{s}^{-1} \simeq 10^{12} \eta_B \left(\frac{T}{\text{MeV}}\right)^3 \text{s}^{-1},$$

where we have used $\langle \sigma_{np} v \rangle = \text{constant} = 4.55 \times 10^{-20}\,\text{cm}^3\,\text{s}^{-1}$ at temperatures $1\,\text{keV} \leq T \leq 10\,\text{MeV}$, and $n_p = \eta_B n_\gamma \simeq 420 \eta_B (T/T_0)^3\,\text{cm}^{-3}$. Using $H \simeq 0.4\,(T/\text{MeV})^2\,\text{s}^{-1}$, we conclude that this interaction remains in thermal equilibrium as long as $T \gtrsim 0.004$ MeV. So the assumption of a thermal deuterium abundance is justified. As already mentioned, three-body interactions are not in thermal equilibrium; their reaction rate contains an additional factor $n_B/n_\gamma = \eta_B \ll 1$.

Therefore, at temperature T_2 only deuterium can form and subsequently virtually all the neutrons present are burned into ${}^4\text{He}$. To determine the helium abundance, we have to determine the neutron density at this temperature. Let us first determine the temperature at which β and inverse β processes drop out of equilibrium,

$$\nu + n \longleftrightarrow p + e, \;\; \bar{e} + n \longleftrightarrow p + \bar{\nu}, \;\; n \to p + e + \bar{\nu}.$$

On one hand, particle conservation imposes

$$\mu_n - \mu_p = \mu_e - \mu_\nu.$$

On the other hand, the neutrality of the Universe requires $n_p = n_e$. Since $m_e \ll m_p$, Eqs. (1.70) and (1.71) imply $\mu_e \ll \mu_p$. Finally, setting $\mu_\nu \sim 0$, the chemical potentials of the neutron and the proton are approximately equal, that is, $\mu_n \simeq \mu_p$. The ratio of their densities is thus simply given by the mass difference $Q = m_n - m_p$,

$$\frac{n_n}{n_p} = \frac{Y_n}{Y_p} = \exp(-Q/T).$$

This ratio remains constant as long as the reactions $n \longleftrightarrow p$ are sufficiently rapid. At the decoupling temperature of these reactions,

$$\Gamma(T_D) = H(T_D) \simeq 3\frac{T_D^2}{m_P},$$

the ratio (n_n/n_p) is hence given by

$$\left(\frac{n_n}{n_p}\right)(T_D) = \exp(-Q/T_D).$$

https://doi.org/10.1017/9781316471524.002 Published online by Cambridge University Press

Afterwards, the neutron density decays exponentially by β-decay, $n \to p + e + \bar{\nu}$,

$$n_n(\tau) = n_n(\tau_D) \exp\left(-\frac{\tau - \tau_D}{\tau_n}\right) \quad \text{for } \tau > \tau_D, \tag{1.128}$$

where $\tau_n \simeq 886$ s is the neutron lifetime.

We now want to determine the temperature T_D. We can again use Fermi theory to determine the different cross sections. For nucleons, the pure $V - A$ current, $\psi^* \gamma_\mu (1 - \gamma_5)\psi$, is replaced by $\psi^* \gamma_\mu (g_V + g_A \gamma_5)\psi$, which takes into account the internal structure of the nucleons. In the Born approximation the cross section becomes (see, e.g., Maggiore, 2005)

$$\sigma(\nu + n \to p + e) = \frac{G_F^2}{\pi}(g_V^2 + 3g_A^2) v_e E_e^2.$$

The constants g_V and g_A are determined experimentally (e.g., by measuring the neutron lifetime), $g_V \simeq 1.00$ and $g_A \simeq 1.25$. The interaction rate per neutron is obtained by multiplying the preceding result with $v_\nu n_\nu$,

$$\Gamma(\nu + n \to p + e) = \langle \sigma v_\nu \rangle n_\nu = \frac{1}{2\pi^2} \int \frac{p_\nu^2 \, dp_\nu}{e^{p_\nu / T_\nu} + 1} v_\nu \sigma \left(1 - \frac{1}{e^{E_e/T} + 1}\right).$$

The factor $1 - 1/[\exp(E_e/T) + 1]$ is the probability that the electron state with energy E_e is free (it implements the Pauli principle). To simplify the integral we first use energy conservation, $E_\nu + E_n = E_p + E_e$. Since all the energies involved are of the order of MeV, we can set $E_n - E_p \sim m_n - m_p = Q = 1.293\,\text{MeV}$ and $E_e = p_\nu + Q$. Furthermore, $E_e = m_e \gamma = m_e / \sqrt{1 - v_e^2}$, which implies $v_e = \sqrt{(p_\nu + Q)^2 - m_e^2}/E_e$. Inserting these simplifications, we obtain finally

$$\Gamma(\nu + n \to p + e) = \frac{G_F^2 (g_V^2 + 3g_A^2) m_e^5}{2\pi^3} \times \int_0^\infty \frac{e^{\alpha(x+q)} x^2 (x+q) \sqrt{(x+q)^2 - 1}}{(1 + e^{\alpha(x+q)})(1 + e^{\beta x})} \, dx, \tag{1.129}$$

where we have set $x = p_\nu / m_e$, $\alpha = m_e / T_\gamma$, $\beta = m_e / T_\nu$, and $q = Q/m_e \simeq 2.5$. To compute the other processes we note that the matrix element $\mathcal{M}(p_\nu, p_n, p_p, p_e)$ that appears in the amplitude for $\nu + n \longleftrightarrow p + e$ is invariant under the transformations $(\mathbf{p}_\nu, p_n, p_p, \mathbf{p}_e) \to (-\mathbf{p}_\nu, p_n, p_p, -\mathbf{p}_e)$, and $(\mathbf{p}_\nu, p_n, p_p, \mathbf{p}_e) \to (-\mathbf{p}_\nu, p_n, p_p, \mathbf{p}_e)$, where p_ν, p_n, p_p and p_e are the momenta of the neutrino, neutron, proton and electron respectively,

$$\mathcal{M}(p_\nu, p_n, p_p, p_e) = \mathcal{M}(-p_\nu, p_n, p_p, -p_e),$$
$$\mathcal{M}(p_\nu, p_n, p_p, p_e) = \mathcal{M}(-p_\nu, p_n, p_p, p_e).$$

https://doi.org/10.1017/9781316471524.002 Published online by Cambridge University Press

This observation allows us immediately to determine the reaction rates of the other processes. We simply have to take into account the different phase space constraints. With $x = E_e/m_e$ (the other parameters as earlier), we obtain

$$\Gamma(e + p \to n + \nu) = \frac{G_F^2(g_V^2 + 3g_A^2)m_e^5}{2\pi^3} \times \int_q^\infty \frac{e^{\beta(x-q)}x(x-q)^2\sqrt{x^2-1}\,dx}{(1+e^{\beta(x-q)})(1+e^{\alpha x})}, \tag{1.130}$$

and

$$\Gamma(n \to p + e + \bar{\nu}) \simeq \frac{G_F^2(g_V^2 + 3g_A^2)m_e^5}{2\pi^3} \times \int_1^q \frac{e^{\alpha x}e^{\beta(q-x)}(x-q)^2x\sqrt{x^2-1}\,dx}{(1+e^{\beta(q-x)})(1+e^{\alpha x})}, \tag{1.131}$$

$$\Gamma(n \to p + e + \bar{\nu})|_{T \ll m_e} \simeq 1.6\frac{G_F^2}{2\pi^3}(g_V^2 + 3g_A^2)m_e^5 = \tau_n^{-1} \tag{1.132}$$

$$\tau_n^{-1} = \frac{1}{886\ \mathrm{s}}$$

for the β-decay of the neutron at low temperature.

The products $\tau_n\Gamma$ are functions of the temperature T. When $T \gg Q$, the kinetic energy in the system $e + \bar{\nu}$ is much higher than the electron mass. Hence $x \pm q \simeq x$ at the positions that contribute most to the foregoing integrals and the reaction rates go like

$$\left.\begin{matrix}\tau_n\Gamma(n \to p)\\ \tau_n\Gamma(p \to n)\end{matrix}\right\} \propto T^5, \quad \text{for } T \gg Q.$$

In the regime $0.1\,\mathrm{MeV} \le T \le 1\,\mathrm{MeV}$, the product $\tau_n\Gamma(n \to p)$ is roughly proportional to $T^{4.4}$. The same is true for $\tau_n\Gamma(p \to n)$. But the phase space for β-decay is larger than for the reaction $p \to n$, so that $\tau_n\Gamma(n \to p) > \tau_n\Gamma(p \to n)$. Once the temperature drops below about 0.1 MeV, $\tau_n\Gamma(p \to n)$ decays exponentially while $\tau_n\Gamma(n \to p)$ converges to 1 [see Fig. 1.8, where $\tau_n\Gamma(n \to p)$, $\tau_n\Gamma(p \to n)$, and the expansion rate $\tau_n H$ are shown as functions of the temperature].

According to Fig. 1.8, the line $\tau_n H$ intersects the lines $\tau_n\Gamma(n \to p)$ and $\tau_n\Gamma(p \to n)$ around $T = 0.8$ MeV. A more detailed analysis gives a decoupling temperature of $T_D \simeq 0.7$ MeV, below which the three reactions are no longer in thermal equilibrium.

https://doi.org/10.1017/9781316471524.002 Published online by Cambridge University Press

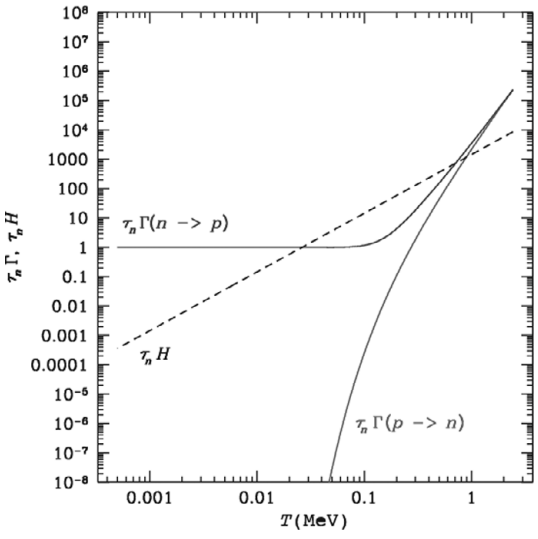

Fig. 1.8 The weak interaction rates, $\tau_n\Gamma(p \to n)$ and $\tau_n\Gamma(n \to p)$, are shown as functions of the temperature. The expansion rate, $\tau_n H$, is also indicated.

Another way to see this dropping out of the thermal equilibrium of weak interaction is to compare the true neutron abundance, Y_n, with the one obtained in thermal equilibrium. A semianalytical calculation gives (see Bernstein *et al.*, 1989) the behavior plotted in Fig. 1.9.

At decoupling, the ratio of the neutron to proton density is

$$\left(\frac{n_n}{n_p}\right)(T_D) = \exp(-Q/T_D) \simeq 1/6, \tag{1.133}$$

so that

$$Y_n = 1/7 \quad \text{and} \quad Y_p = 6/7. \tag{1.134}$$

Since T_2, the temperature of deuterium formation, is lower than T_D, in the interval $T_D > T > T_2$, neutrons simply β-decay. At τ_2 given by $T_2 = T(\tau_2) = 0.085\,\text{MeV}$ their density is

$$\left(\frac{n_n}{n_p}\right)(T_2) = e^{-Q/T_D}\exp(-\tau_2/\tau_n) \simeq 0.8/6 \simeq 1/7, \tag{1.135}$$

and therefore

$$Y_n = 1/8 \quad \text{and} \quad Y_p = 7/8. \tag{1.136}$$

https://doi.org/10.1017/9781316471524.002 Published online by Cambridge University Press

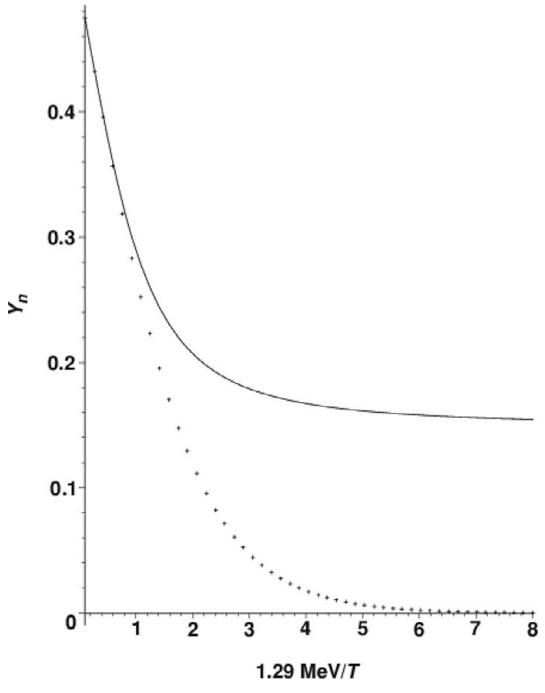

Fig. 1.9 The true neutron abundance as a function of $\Delta m/T$ (solid line) is compared with the equilibrium abundance (dotted line). Clearly, weak interaction freezes out around $T \sim 0.6 \times \Delta m \sim 0.7$ MeV.

For this we have used $\tau_2 \simeq 1.3\,\mathrm{s}\,(1/0.085)^2 \simeq 180$ s. Once deuterium is formed, helium-4 is very rapidly synthesized via the reactions

$$
\begin{aligned}
{}^2\mathrm{H} + {}^2\mathrm{H} &\longrightarrow n + {}^3\mathrm{He} \\
{}^3\mathrm{He} + {}^2\mathrm{H} &\longrightarrow p + {}^4\mathrm{He}
\end{aligned}
$$

$$
\begin{aligned}
{}^2\mathrm{H} + {}^2\mathrm{H} &\longrightarrow \mathrm{p} + {}^3\mathrm{H} \\
{}^3\mathrm{H} + {}^2\mathrm{H} &\longrightarrow n + {}^4\mathrm{He}
\end{aligned}
$$

$$
{}^2\mathrm{H} + {}^2\mathrm{H} \longrightarrow \gamma + {}^4\mathrm{He}
$$

and essentially all deuterium is transformed in ^{4}He. The helium abundance is thus in good approximation, given by half the neutron abundance at temperature $T_2 \simeq 0.085$ MeV. With this approximation we obtain a helium-4 abundance of

$$
Y_{^4He} = \frac{4(n_n/2)}{n_n + n_p} = \frac{2(n_n/n_p)}{n_n/n_p + 1} \simeq \frac{1}{4}. \tag{1.137}
$$

https://doi.org/10.1017/9781316471524.002 Published online by Cambridge University Press

In this expression we have used the neutron abundance from Eq. (1.136). Considering that τ_2 scales like $\sqrt{\log \eta_B}$ while T_D depends strongly on the expansion rate H, which is proportional to $\sqrt{g_{\rm eff}} \propto \sqrt{N_\nu (4/11)^{4/3} + 1}$, we conclude that the helium-4 abundance is very sensitive on the number of neutrino families, but does not change very rapidly with η_B. Historically, the cosmological helium-4 abundance has been the first experimental data to determine the number of (light) neutrino families in the range $N_\nu = 3.24 \pm 1.2$, when allowing for very generous error bars in the measurements (Fields and Sarkar, 2006). Presently, the Z-boson decay width, which has been measured very accurately with the LEP accelerator at CERN, gives the tightest value (see Particle Data Group, 2006), $N_\nu = 3.07 \pm 0.12$ at 95% confidence.

1.4.4 Deuterium, Helium-3 and Lithium-7

Nucleosynthesis starts at $T \sim 0.1$ MeV, corresponding to $\tau \sim 130$ s and terminates after a few minutes. Apart from ^{4}He very small amounts of all other elements up to lithium-7 are formed (some deuterium, tritium, and helium-3 remain unprocessed). All these elements except deuterium, helium-3, and lithium-7 decay radioactively and their primordial abundance can no longer be observed today.

The amount of deuterium and helium-3 that is not burned into helium-4 is a steep function of the baryon abundance in the Universe. The higher the baryon density, the more efficient is the conversion of deuterium and helium-3 into helium-4 (see Fig. 1.10). This can be used to determine the baryon density in the Universe very accurately. Measuring the primordial deuterium abundance is an art by itself on which we shall not dwell here. Most recent results are obtained by measuring it from the absorption lines in hydrogen (Ly-α) clouds intervening in the line of sight between us and quasars. Within generous error bars one obtains $2 \times 10^{-5} < Y_{^2H}/Y_p < 2 \times 10^{-4}$. This gives $4.7 \times 10^{-10} < \eta_B < 6.5 \times 10^{-10}$ [for more details see Olive *et al.* (2000), Burles *et al.* (2001), and Particle Data Group (2006)].

As one sees in Fig. 1.10, the lithium abundance is not a monotone function of η_B. This is so since, depending on the value of η_B, two different processes lead to lithium formation. If the baryon density is small, $\eta_B < 3 \times 10^{-10}$, lithium abundance is determined by the competition between the production process $^4\mathrm{He} + {}^3\mathrm{H} \to {}^7\mathrm{Li} + \gamma$ and the destruction process $^7\mathrm{Li} + p \to {}^4\mathrm{He} + {}^4\mathrm{He}$. In this regime, the abundance decays with growing η_B. For $\eta_B > 3 \times 10^{-10}$, the dominant channel goes over beryllium production $^4\mathrm{He} + {}^3\mathrm{He} \to {}^7\mathrm{Be} + \gamma$, which is then converted into lithium-7 via the reaction $^7\mathrm{Be} + e \to {}^7\mathrm{Li} + \gamma$. The destruction process is the same as at low density. Since the conversion of beryllium into lithium increases with increasing baryon density, lithium abundance grows with η_B, for $\eta_B > 3 \times 10^{-10}$. The lithium abundance has a minimum around $\eta_B \simeq 3 \times 10^{-10}$. Inference of the

https://doi.org/10.1017/9781316471524.002 Published online by Cambridge University Press

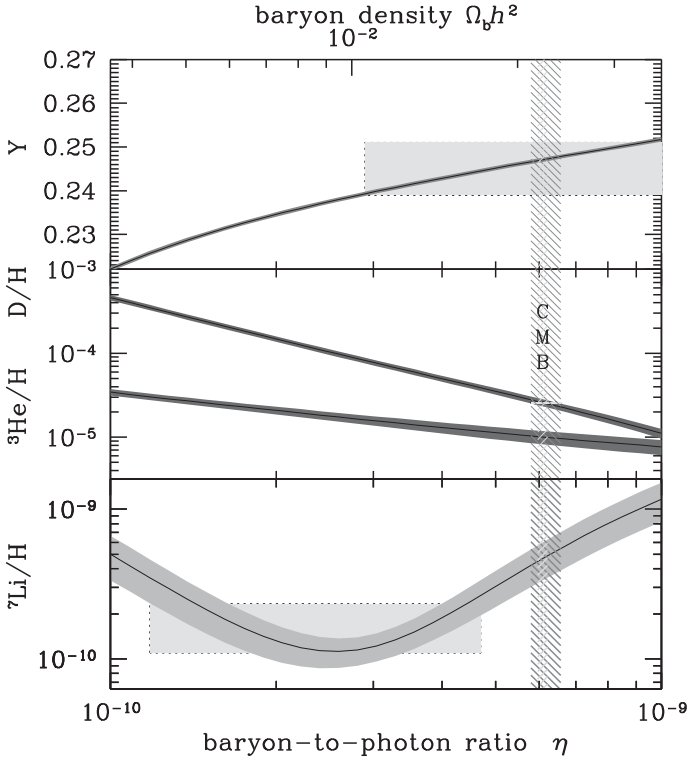

Fig. 1.10 The primordial element abundance as a function of the parameter $\eta_B = n_B/n_\gamma$. The bands compatible with the observations of the different nuclei are indicated. The horizontal band shows the range of η_B (or equivalently $\Omega_B h^2$) compatible with the nucleosynthesis data while the narrow vertical range is compatible with CMB anisotropies (see Chapter 9). It agrees very well with deuterium and helium abundances from nucleosynthesis but not so well with the lithium abundance. Figure from Tanabashi *et al.* (2019)

primordial lithium abundance is still a matter of considerable debate. It nevertheless allows us to constrain $10^{-10} < \eta_B < 10^{-9}$.

Finally, in the regime $10^{-10} < \eta_B < 10^{-9}$ the helium-4 abundance is well approximated by the formula

$$Y_{^4He} = 0.23 + 0.011 \ln(\eta_{10}) + 0.013(N_\nu - 3), \tag{1.138}$$

where we have introduced $\eta_{10} = \eta_B/10^{-10}$. All the present observations of light elements taken together limit $4.7 < \eta_{10} < 6.5$, leading to $0.017 < \Omega_B h^2 < 0.024$ [a constantly updated review can be found in Tanabashi *et al.* (2019)]. It is remarkable that this value is in very good agreement with the result obtained from measurements of the fluctuations in the CMB, which are based on completely different physics (see Chapter 9).

https://doi.org/10.1017/9781316471524.002 Published online by Cambridge University Press

This value is much larger than the density of luminous baryons that make up the stars and gas in the galaxies, and that lead only to $\Omega_L h^2 \simeq 0.004$. Hence most baryons in the Universe are not luminous. On the other hand, dynamical measurements and, more accurately, the anisotropies in the CMB, require an energy density of nonrelativistic matter today of about $\Omega_m h^2 \simeq 0.13$. We discuss constraints on cosmological parameters from CMB data in detail in Chapter 9. To satisfy both constraints, the matter density of the Universe has to be dominated to about 80% by nonbaryonic, so-called dark matter (dark in this context means that this matter does not interact with photons). So far, this dark matter has not been observed directly, but many experiments are underway and are starting to reach promising sensitivities. There are several candidates for dark matter particles. Most notably, the lightest supersymmetric particle, but also the gravitino, axion, or primordial black holes are viable candidates.

The good agreement of N_ν and $\Omega_B h^2$ obtained from the study of primordial nucleosynthesis with other experiments, confirms that the Universe has been in a thermal state expanding adiabatically back to temperatures of the order of $T \sim 1$ MeV. For earlier times we have no experimental evidence. However, if the Universe has been in a thermal state at a temperature of $T \sim 200$ MeV, $\tau \sim 0.1$ s, it has then undergone a confinement transition leading from a quark gluon plasma at higher temperatures to baryons (such as the proton and neutron) and mesons (such as pions). If it has also been in thermal equilibrium at temperatures of up to $T \sim 200$ GeV, $\tau \sim 0.001$ s, it has then undergone the electroweak transition giving masses to the $W^\pm$ and Z bosons. At even higher temperatures we have no experimentally confirmed theory of fundamental interactions. Maybe, at $T \sim$ a few TeV the Universe becomes supersymmetric. Maybe, at $T \sim 10^{16}$ GeV a phase transition from a previous grand unified symmetry to the (supersymmetric) standard model symmetries took place. At this or higher energies the Universe may also have gone through (or emerged from) a superstring phase. To date such questions remain entirely speculative. Their quantitative investigation, especially possible observable signatures of a superstring phase, is an active field of research.

1.5 Inflation

1.5.1 Cosmological Problems

We first discuss the motivation for, and some consequences of a so-called inflationary phase. We then exemplify the idea with a cosmology dominated by a scalar field. It is, however, clear that this realization has to be regarded as a toy model because the actual physical degrees of freedom relevant in the very early Universe, where such a period has most probably to be situated (see Chapter 3), are not

https://doi.org/10.1017/9781316471524.002 Published online by Cambridge University Press

known. In that sense this section is on a different level from the previous ones. We do not have any direct evidence that an inflationary phase has taken place in our Universe. Such a period just addresses several otherwise mysterious initial conditions of the observed Universe. The most significant observed "prediction" of inflation is a nearly scale-invariant spectrum of initial fluctuations that we shall discuss in Chapter 3. What is more serious is that we have no "direct" experimental evidence of the existence of an "inflaton field."

We include a possible cosmological constant into the energy density and the pressure, so that Eqs. (1.20) and (1.21) reduce to

$$\mathcal{H}^2 = \frac{8\pi G}{3} a^2 \rho - K, \tag{1.139}$$

$$\dot{\mathcal{H}} = -\frac{4\pi G}{3} a^2 (\rho + 3P) = \left(\frac{\ddot{a}}{a}\right) - \mathcal{H}^2. \tag{1.140}$$

If $\rho + 3P > 0$ at all times, the homogeneous and isotropic cosmological model has several important problems.

First, as we have discussed in Section 1.2.2, there is the big bang singularity in the finite past, $t = 0$. At this time $a = 0$ and the curvature diverges.

Furthermore, the causal horizon at (conformal) time t, that is, the distance a photon has traveled from $t = \tau = 0$ until time t, is given by $a(t)t = a(t) \int_0^{\tau(t)} a^{-1}\, d\tau$. Since for $\rho + 3P > 0$, a grows slower than linear in τ, this integral converges, is finite. As we have seen (in Eq. (1.25)), $a(\tau) \propto \tau^{\frac{2}{3(1+w)}}$ if $w = P/\rho$ is constant.

For example, the size of the causal horizon at recombination is seen today under the angle of about 1°, if the Universe was radiation ($w = 1/3$) and matter ($w = 0$) dominated up to recombination see Exercise 1.10. It is therefore very mysterious that we see the same microwave background temperature on patches separated by much more than 1°, which had never been in causal contact before the microwave photons had been emitted. This is the "**horizon problem**."

Another problem is the following: the Friedmann equations, (1.139) and (1.140), allow us to derive an evolution equation for $\Omega(t) \equiv 8\pi G\rho a^4/3\dot{a}^2 \equiv 1 + K/\mathcal{H}^2$,

$$\frac{d}{dt}(\Omega(t) - 1) = (\Omega(t) - 1)\frac{8\pi G a^2}{3}\left(\frac{\rho + 3P}{\mathcal{H}}\right). \tag{1.141}$$

This shows that, in an expanding universe with $\rho + 3P > 0$, $\Omega = 1$ is an unstable fixed point of evolution: if $\Omega(t) > 1$, the derivative is positive and $\Omega(t)$ increases while for $\Omega(t) < 1$, the derivative is negative and $\Omega(t)$ decreases. For a present value of $0.1 < \Omega_0 < 2$ we need $|\Omega(\eta_{\rm nuc}) - 1| \sim (z_{\rm eq}/z_{\rm nuc}^2)|\Omega_0 - 1| \leq 10^{-15}$ at nucleosynthesis, or $|\Omega(t_P) - 1| \leq 10^{-60}$ at the Planck time, $\tau_P = \sqrt{\hbar G/c^5} \simeq 5.4 \times 10^{-44}$ s. Why is $\Omega(t)$ still of order unity so long after the only timescale in the problem that is τ_P?

https://doi.org/10.1017/9781316471524.002 Published online by Cambridge University Press

This "**flatness problem**" can also be formulated as an "**entropy problem**." The entropy inside the curvature radius is already of the order of $S_K \geq 10^{88}$ at the Planck time.

Another problem is the "**monopole problem**" or more generically the problem of unwanted "relics." Most particle physics models produce some stable "relics" at very high temperatures, which are not observed in the present Universe. A very rapid phase of expansion can help to dilute such relics.

To resolve these problems one introduces an "inflationary phase." Inflation is a phase during which the strong energy condition, $\rho + 3P > 0$, is violated and expansion can therefore be much more rapid than linear in τ.

1.5.2 Scalar Field Inflation

We now study the most common solution of the aforementioned problems, namely the introduction of a period in which the dynamics of the Universe is dominated by a scalar field, ϕ which is usually called the "inflaton." The scalar field Lagrangian is given by

$$\mathcal{L}_\phi = -\frac{1}{2}\,\partial_\mu\phi\,\partial^\mu\phi - W(\phi). \tag{1.142}$$

The sign of the kinetic term in the foregoing Lagrangian may differ from what you are used to from quantum field theory. This comes from the fact that we use the metric signature $(-,+,+,+)$.

The field ϕ can, in principle, interact with other fields such as fermions, gauge bosons, and so forth, but we assume that this interaction can be neglected during inflation, and that energy and pressure are dominated by the contribution from the inflaton. The energy–momentum tensor of ϕ is given by

$$T_{\mu\nu} = \frac{-2}{\sqrt{-g}}\frac{\partial}{\partial g^{\mu\nu}}(\sqrt{-g}\mathcal{L}_\phi),$$

where $g = \det(g_{\mu\nu})$. This yields

$$\begin{aligned} T_{\mu\nu} &= \partial_\mu\phi\,\partial_\nu\phi + g_{\mu\nu}\mathcal{L}_\phi \\ &= \partial_\mu\phi\,\partial_\nu\phi - \frac{1}{2}g_{\mu\nu}\,\partial_\lambda\phi\,\partial^\lambda\phi - g_{\mu\nu}W(\phi). \end{aligned}$$

Here we have used that the derivatives of the determinant A of an arbitrary matrix A_{ab} with respect to the elements of its inverse, A^{ab}, are given by $\partial A/\partial A^{ab} = AA_{ab}$.

For the energy density and pressure we thus obtain

$$\rho_\phi = -T^0_0 = \frac{1}{2a^2}\dot{\phi}^2 + \frac{1}{2a^2}(\nabla\phi)^2 + W(\phi), \tag{1.143}$$

https://doi.org/10.1017/9781316471524.002 Published online by Cambridge University Press

and

$$P_\phi = \frac{1}{3}T_i^i = \frac{1}{2a^2}\dot{\phi}^2 - \frac{1}{6a^2}(\nabla\phi)^2 - W(\phi). \tag{1.144}$$

We now assume that there exists some region of space within which we may neglect the spatial derivatives of ϕ, at some initial time τ_i, and the temporal derivative is much smaller than the potential,

$$\nabla\phi(\mathbf{x},\tau_i) \ll \dot{\phi}(\mathbf{x},\tau_i) \ll W(\phi). \tag{1.145}$$

Furthermore, we assume that the potential is positive,

$$W(\phi(\mathbf{x},\tau_i)) > 0. \tag{1.146}$$

We then have

$$\frac{3H^2}{8\pi G} = \rho = \rho_\phi = \frac{1}{2a^2}\dot{\phi}^2 + W(\phi) \simeq W(\phi), \tag{1.147}$$

$$P = P_\phi = \frac{1}{2a^2}\dot{\phi}^2 - W(\phi) \simeq -W(\phi), \tag{1.148}$$

so that $P_\phi \simeq -\rho_\phi$ and $\rho_\phi + 3P_\phi \simeq -2W(\phi) < 0$. (We have neglected a possible curvature term. Qualitatively nothing changes if we include it, since it soon becomes subdominant.)

This is the basic idea of inflation: at some early time, in some sufficiently large patch, the Universe is dominated by the potential of a slowly varying (slow rolling) scalar field, and hence it is in an inflationary phase. During inflation this patch expands rapidly and the causal horizon becomes very large and $\Omega(t)$ tends to 1, so that the curvature term is soon negligible. As time goes on, the scalar field starts evolving faster and inflation eventually comes to an end when the time derivative ϕ'^2 grows to the order of W. The scalar field then soon reaches the minimum of the potential and starts to oscillate. We suppose that at large values of $a^{-1}\dot{\phi}$, the coupling of the inflaton to other fields becomes significant so that it decays into a thermal mix of elementary particles, leading to a radiation-dominated universe. There are many detailed realizations of this basic picture that can be found in the literature; see, for example, Liddle and Lyth (2000). It is, however very difficult to deduce them from a serious high-energy physics theory such as string theory.

Let us study slow roll inflation in somewhat more detail. When neglecting spatial derivatives, the equation of motion of the scalar field becomes ($W_{,\phi} \equiv dW/d\phi$)

$$\ddot{\phi} + 2\left(\frac{\dot{a}}{a}\right)\dot{\phi} + a^2 W_{,\phi} = 0, \tag{1.149}$$

$$\phi'' + 3\left(\frac{a'}{a}\right)\phi' + W_{,\phi} = 0, \tag{1.150}$$

https://doi.org/10.1017/9781316471524.002 Published online by Cambridge University Press

in conformal time, Eq. (1.149), and in cosmic time, Eq. (1.150). During slow rolling, the first term of this equation is negligible with respect to the two others, so that

$$3\left(\frac{a'}{a}\right)\phi' \simeq -W_{,\phi}. \tag{1.151}$$

The slow roll conditions are therefore

$$\frac{1}{2}\phi'^2 \ll W \quad \text{and} \quad |\phi''| \ll 3H|\phi'|. \tag{1.152}$$

With $H = a'/a$, slow rolling also implies that $H' \ll H^2$. Taking the time derivative of Eq. (1.147) and replacing ϕ' by (1.151) yields the slow roll conditions

$$\epsilon_1 \equiv -\frac{H'}{H^2} = \frac{\mathcal{H}^2 - \dot{\mathcal{H}}}{\mathcal{H}^2} \approx \frac{m_P^2}{16\pi}\left(\frac{W_{,\phi}}{W}\right)^2 \simeq \frac{3}{2}\frac{\phi'^2}{W} \ll 1. \tag{1.153}$$

The second condition of Eq. (1.152) gives

$$\left|\frac{\phi''}{3H\phi'}\right| \ll 1.$$

We now set

$$\epsilon_2 \equiv -\frac{m_P^2}{24\pi}\left(\frac{W_{,\phi\phi}}{W}\right) \quad \text{and require} \quad |\epsilon_2| \ll 1. \tag{1.154}$$

Note that ϵ_1 is always positive while ϵ_2 can have either sign. With $H^2 \simeq 8\pi W/(3m_P^2)$, and the derivative of $\phi' = -W_{,\phi}/(3H)$, one finds that the inequalities (1.153) and (1.154) are equivalent to the slow roll conditions (1.152). The parameters ϵ_1 and ϵ_2 are the slow roll parameters. Inflation terminates when ϵ_1 approaches unity. In the literature one often uses the notation $\epsilon \equiv \epsilon_1$ and $\delta \equiv -\epsilon_2/3$.

Taking the derivative (w.r.t. t) of Eq. (1.153) in the last equals sign, one obtains

$$\dot{\epsilon}_1 = 2\epsilon_1\,(3\epsilon_2 + 2\epsilon_1)\,\mathcal{H}, \qquad \frac{\dot{\epsilon}_1}{\mathcal{H}\epsilon_1} = 6\epsilon_2 + 4\epsilon_1 \equiv \eta. \tag{1.155}$$

The last equation can also be used as a definition of ϵ_2 (or, more consistently, η). The advantage of this definition is its independence of the realization of slow roll inflation by means of a scalar field. A more systematic procedure along these lines is to define $\tilde{\epsilon}_1 \equiv \epsilon_1$ and $\tilde{\epsilon}_2 = (\dot{\tilde{\epsilon}}_1/\tilde{\epsilon}_1)\mathcal{H}^{-1} = \eta$, $\tilde{\epsilon}_3 = (\dot{\tilde{\epsilon}}_2/\tilde{\epsilon}_2)\mathcal{H}^{-1}$, and so forth. Our parameter ϵ_2 is related to $\tilde{\epsilon}_2 \equiv \eta$ via

$$\epsilon_2 = -\frac{2}{3}\epsilon_1 + \frac{1}{6}\eta. \tag{1.156}$$

While ϵ_2 is usually of the same order of magnitude as ϵ_1, we expect η to be significantly smaller.

https://doi.org/10.1017/9781316471524.002 Published online by Cambridge University Press

As an example we consider power law expansion, $a \propto t^q$. In this case we have

$$\mathcal{H} = \frac{q}{t}, \quad \epsilon_1 = 1 + \frac{1}{q}, \quad \epsilon_2 = -\frac{2}{3}\epsilon_1, \quad \tilde{\epsilon}_2 = \tilde{\epsilon}_n = 0. \tag{1.157}$$

During slow roll inflation, $q \sim -1$, the parameters ϵ_1 and ϵ_2 are small. Also note that $\epsilon_2 = -(2/3)\epsilon_1$ during power law expansion. The parameters $\tilde{\epsilon}_i$, $i > 1$ describe the deviation from power law expansion. They have been used in the literature to derive a systematic slow roll expansion to higher orders (Schwarz *et al.*, 2001). In this book we shall not go beyond the first order and we use the standard parameters ϵ_1 and ϵ_2 to make contact with the standard literature.

There are two principally different possibilities for slow roll inflation.

(1) We first consider a potential that is simply $\propto \phi^n$, so that $W_{,\phi\phi}/W \sim (W_{,\phi}/W)^2 \sim \phi^{-2}$. The slow roll conditions then require $\phi \gg m_P$ and inflation stops when the inflaton becomes of order the Planck mass. These models are termed **large-field inflation**. Setting $W = (\lambda/n)m_P^4(\phi/m_P)^n$, during the inflationary phase, Eq. (1.151) together with Eq. (1.147) implies

$$\sqrt{\frac{24\pi\lambda}{n}} m_P \, (\phi/m_P)^{n/2} \, \phi' = -\lambda m_P^3 \left(\frac{\phi}{m_P}\right)^{n-1}. \tag{1.158}$$

Dividing by ϕ^{n-1}, if $n \neq 4$ the left-hand side becomes the derivative of $(\phi/m_P)^{2-n/2}$, which hence is a constant. If $n = 4$, the left-hand side is $\propto 1/\phi$, that is, the derivative of $\log(\phi/m_P)$. The general solution is therefore given by

$$\phi(\tau)^{(4-n)/2} = \phi_i^{(4-n)/2} - \frac{4-n}{2}\sqrt{n\lambda/24\pi}\, m_P(\tau - \tau_i) \quad \text{if } n \neq 4, \tag{1.159}$$

$$\phi(\tau) = \phi_i \exp\left(-\sqrt{\frac{\lambda}{6\pi}} m_P(\tau - \tau_i)\right) \qquad \text{if } n = 4. \tag{1.160}$$

Inserting now $\phi' = -\sqrt{\lambda n/24\pi}\, m_P^2 (\phi/m_P)^{n/2-1}$ in the Friedmann equation,

$$(\log(a))' = \sqrt{\frac{8\pi\lambda}{3n}} m_P (\phi/m_P)^{n/2},$$

we obtain

$$\frac{d\log(a)}{d\phi} = -\frac{8\pi}{n}\frac{\phi}{m_P^2},$$

with solution

$$a(\tau) = a_i \exp\left(\frac{4\pi}{n m_P^2}(\phi_i^2 - \phi^2)\right). \tag{1.161}$$

This case is illustrated in Fig. 1.11.

https://doi.org/10.1017/9781316471524.002 Published online by Cambridge University Press

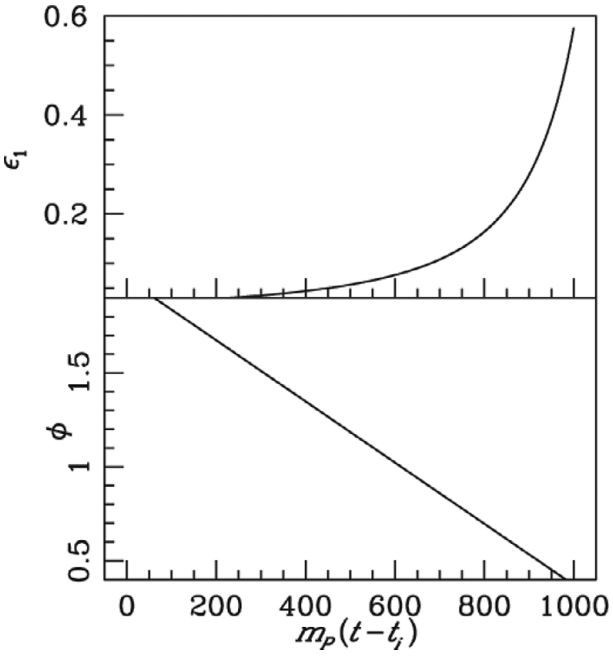

Fig. 1.11 Large-field inflation for $W = \lambda m_P^2 \phi^2/2$. The bottom panel shows the inflaton ϕ in units of m_P rolling linearly in time. In the upper panel the evolution of the slow roll parameter, $\epsilon_1(t)$, is indicated. As long as $\phi > m_P$, $\epsilon_1 = -2\epsilon_2$ stays small. At $\phi \sim m_P$, ϵ_1 starts to grow and inflation stops.

(2) If the potential is more complicated and has a very flat regime in the vicinity of its maximum $\phi = \sigma \ll m_P$, like, for example, the Coleman–Weinberg potential (Kolb and Turner, 1990),

$$W(\phi) = \frac{1}{2}\sigma^4 + \phi^4\left[\ln\left(\frac{\phi^2}{\sigma^2}\right) - \frac{1}{2}\right],$$

we speak of **small-field inflation**. This potential passes through 0 at $\phi = \sigma$. In this case, the slow roll conditions are satisfied for field values $|\phi| \lesssim \sigma$, which are much smaller than the Planck mass.

During a potential dominated phase where $\rho \sim -P \sim W \sim$ constant, the solutions of the Friedmann equations are

$$a = a_0 \exp(\tau H) = \frac{1}{H|t|} \quad (-\infty < t < 0, \quad -\infty < \tau < \infty), \tag{1.162}$$

$$H^2 = \frac{8\pi G}{3} W = \text{constant}, \tag{1.163}$$

$$\mathcal{H} = aH = \frac{1}{|t|}. \tag{1.164}$$

https://doi.org/10.1017/9781316471524.002 Published online by Cambridge University Press

The limit $\tau \to \infty$ corresponds to $t \to 0$. The foregoing solution is a portion of de Sitter spacetime.[6]

Denoting by indices i and f the beginning and the end of inflation, the number of e-foldings of expansion during inflation is given by

$$N(\phi_f, \phi_i) = \ln\left(\frac{a(\tau_f)}{a(\tau_i)}\right).$$

Using

$$N(\phi_f, \phi_i) = \ln a_f - \ln a_i = \int_{a_i}^{a_f} \frac{da}{a},$$

we obtain

$$N(\phi_f, \phi_i) = \int_{a_i}^{a_f} \frac{1}{a}\, da = \int_{\tau_i}^{\tau_f} \frac{a'}{a}\, d\tau = \int_{\tau_i}^{\tau_f} H\, d\tau. \tag{1.165}$$

With Eq. (1.151) we can write

$$H\, d\tau = H\, \frac{d\tau}{d\phi}\, d\phi = H \frac{d\phi}{\phi'} = -\frac{3H^2\, d\phi}{W_{,\phi}}.$$

The number of e-foldings is hence given by

$$N(\phi_f, \phi_i) = -3 \int_{\phi_i}^{\phi_f} \frac{H^2}{W_{,\phi}}\, d\phi \simeq -\frac{8\pi}{m_P^2} \int_{\phi_i}^{\phi_f} \frac{W}{W_{,\phi}}\, d\phi = -2\sqrt{\pi} \int_{\phi_i}^{\phi_f} \frac{1}{\sqrt{\epsilon_1}} \frac{d\phi}{m_P}$$
$$\sim \frac{8\pi}{n} \frac{\phi_i^2}{m_P^2}. \tag{1.166}$$

The last $\sim$ sign is valid only for large-field inflation, where $W \propto \phi^n$ and we suppose

$$\phi_f \sim m_P \ll \phi_i.$$

The slow roll conditions imply

$$N_{\rm tot} = N(\phi_f, \phi_i) \gg 1. \tag{1.167}$$

For $w = P/\rho =$ constant we have

$$|\Omega(\tau) - 1| = \frac{3|K|}{8\pi G a^2 \rho} \propto a^{1+3w}.$$

During an inflationary phase, $w = -1$, $|\Omega(\tau) - 1|$ decreases like $1/a^2$. To reduce it from a value of order unity down to $\sim 10^{-60}$ we therefore need about $30 \ln(10) \sim 70$ e-foldings of inflation.

[6] de Sitter spacetime is the solution to the Einstein equation $G_{\mu\nu} = \Lambda g_{\mu\nu}$ with $\Lambda > 0$. The solution with $\Lambda < 0$ is called anti-de Sitter; see Hawking and Ellis (1973).

https://doi.org/10.1017/9781316471524.002 Published online by Cambridge University Press

1.5.3 Preheating and Reheating

When inflation ends, ϕ decays rapidly and starts oscillating about its minimum. The details of this process depend on the couplings of the inflaton to other degrees of freedom, which eventually decay into the degrees of freedom of the standard model. For this discussion we consider a simple toy model with $\mathcal{L}_\phi = -\frac{1}{2}\partial_\mu\phi\partial^\mu\phi - \frac{1}{2}m_\phi^2\phi^2$. At the end of inflation the inflaton oscillates as

$$\phi = \phi_0(\tau)\cos(m_\phi\tau)$$

with a slowly varying amplitude $\phi_0(\tau) \simeq m_P$. The inflatons have vanishing momentum and their number density is

$$n_\phi = \frac{\rho_\phi}{m_\phi} = \frac{1}{2m_\phi}((\phi')^2 + m_\phi^2\phi^2) \simeq \frac{1}{2}m_\phi m_P^2. \tag{1.168}$$

For example, for $m_\phi = 10^{15}$ GeV this amounts to the huge number density of $n_\phi \sim 10^{95}\ \mathrm{cm}^{-3}$.

Independent of the detailed form of the potential, to lowest order, ϕ is a harmonic oscillator with frequency $m_\phi^2 \simeq W_{,\phi\phi}(\phi_0)$ (as long as the quadratic term in the potential does not vanish). For a harmonic oscillator, when averaging over one period we have

$$\langle W\rangle = \frac{1}{2a^2}\langle\dot\phi^2\rangle,$$

so that

$$\langle p_\phi\rangle = \left\langle \frac{1}{2a^2}\dot\phi^2 - W\right\rangle = 0, \quad \text{and hence } \langle\rho_\phi\rangle \propto a^{-3}.$$

We assume that during these oscillations, the coupling of ϕ to other degrees of freedom becomes relevant and the inflaton finally decays into a mix of elementary particles. In a first approximation we can describe the coupling with the other degrees of freedom by means of a term of dissipation of the form $\Gamma\dot\phi$ in the equation of motion for ϕ,

$$\phi'' + 3H\phi' + \Gamma\phi' = -W_{,\phi}(\phi). \tag{1.169}$$

As long as $H \gg \Gamma$ (during inflation), particle production is negligible. When $H \simeq \Gamma$, reheating takes place and the inflaton energy is rapidly dissipated into other particles that couple to the inflaton.

In order to discuss the decay of the inflaton in somewhat more detail, we consider a toy model in which the interaction is dominated by the coupling of ϕ to a scalar field χ with Lagrangian

$$\mathcal{L}_\chi = -\frac{1}{2}\,\partial_\mu\chi\,\partial^\mu\chi - \frac{1}{2}m_\chi^2\chi^2. \tag{1.170}$$

https://doi.org/10.1017/9781316471524.002 Published online by Cambridge University Press

The interaction between the inflaton ϕ and the matter field χ is supposed to be of the form

$$\mathcal{L}_{\text{int}} = -\frac{1}{2} g \phi \chi^2, \tag{1.171}$$

where g is a coupling constant with the dimension of mass. The full Lagrangian is then given by

$$\mathcal{L} = \mathcal{L}_\phi + \mathcal{L}_{\text{int}} + \mathcal{L}_\chi. \tag{1.172}$$

The decay rate of the ϕ particles in Born approximation is

$$\Gamma_\phi \sim \frac{g^2}{m_\phi}.$$

However, inserting this into Eq. (1.169) is a good approximation only as long as the mean number of χ particles already present in a given momentum mode k is small so that we may neglect stimulated emission. The effective mass of χ-particles is $m_{\text{eff}} = \sqrt{m_\chi^2 + g\phi(t)}$ so that their momentum is

$$k = \left(\frac{m_\phi^2}{4} - m_{\text{eff}}^2 \right)^{1/2}.$$

Here we have taken into account that each inflaton decays into two χ-particles. Now, $\phi(t) \in [-m_P, m_P]$. Hence, if $m_\phi^2 \gg m_\chi^2 + g m_P$, the band of possible momenta is given by $k \in [k_0 - \Delta k, k_0 + \Delta k]$ with

$$k_0 = \sqrt{\frac{m_\phi^2}{4} - m_\chi^2} \simeq \frac{m_\phi}{2} \quad \text{and} \quad \Delta k \simeq \frac{g m_P}{m_\phi} \ll k_0.$$

Because $\Delta k \ll k_0$ this situation is called "narrow band preheating." As we shall see in the text that follows, this process leads to resonant amplification.

The number of χ-particles with momentum k is roughly given by the total number of χ-particles divided by the number of "elementary phase space volumes," $(2\pi)^3$, in the allowed volume of phase space, $4\pi k_0^2 (2\Delta k)$. This yields

$$N_k \simeq \frac{4\pi^2 n_\chi}{g m_\phi m_P} \simeq \frac{2\pi^2 m_P n_\chi}{g n_\phi}.$$

For the second $\simeq$ sign we made use of Eq. (1.168). This occupation number exceeds unity as soon as a fraction g/m_P of ϕ-particles is converted into χ-particles. After that moment, stimulated emission can no longer be neglected and Eq. (1.169) becomes a bad approximation. Since g/m_P typically is very small, this is usually the case very soon. As we shall now see, when this happens, stimulated emission leads to resonant production of χ-particles in certain k-bands.

https://doi.org/10.1017/9781316471524.002 Published online by Cambridge University Press

To calculate the generation of χ-particles in more detail we vary the Lagrangian with respect to χ to obtain the χ-equation of motion,

$$\chi'' + 3H\chi' - a^{-2}\nabla^2\chi + (m_\chi^2 + g\phi_0\cos(m_\phi\tau))\chi = 0.$$

To study qualitatively the decay of the ϕ-particles into χ, we neglect expansion by setting $H = 0$, $a = 1$ and $\phi_0 =$ constant. Fourier transforming the above equation, we then obtain for the mode χ_k

$$\chi_k'' + \left[\omega_k^2 + 2\mu\cos(m_\phi\tau)\right]\chi_k = 0, \quad \mu = \frac{g\phi_0}{2}, \quad \omega_k^2 = k^2 + m_\chi^2 \ .$$

This equation is known as the *Mathieu equation*. Its solutions are characterized by resonance bands of widths $\Delta\omega_k^{(n)}$ centered at the frequencies

$$\omega_k^{(n)} = \frac{n}{2}m_\phi.$$

The widths are of the order of

$$\frac{\Delta\omega_k^{(n)}}{\omega_k^{(n)}} \simeq \left(\frac{2\mu}{\omega_k^{(n)2}}\right)^n = \frac{\Delta k^{(n)}}{k^{(n)}} \simeq \left(\frac{4gm_P}{n^2m_\phi^2}\right)^n \propto \left(\frac{g}{n^2}\right)^n.$$

For frequencies within these bands, χ_k is amplified exponentially fast (for more details on the Mathieu equation and resonant amplification, see Arnold, 1978). Since the width of the nth resonance is proportional to g^n, it appears only at nth order in perturbation theory. For small couplings only the first resonance $\omega_k^{(1)} = m_\phi/2$ with $\Delta\omega_k^{(1)} = \Delta k$ is relevant.

When we take into account the expansion of the Universe, the frequency ω_k becomes time dependent. A given frequency therefore spends only a finite time in the resonance band and the energy transfer from ϕ into χ remains perfectly finite. Nevertheless, this parametric resonance is much more efficient than the decay obtained by some effective damping rate Γ.

After parametric resonance, χ is not yet in a thermal state. This period is therefore called "preheating." After preheating, the coupling of χ to other degrees of freedom leads to thermalization; this process is called *reheating*. The importance of preheating lies in its efficiency in transferring energy. If the χ-field couples strongly to the standard-model particles, reheating and thermalization can proceed much faster over resonant decay than over the necessarily weak average coupling of the inflaton to other particles.

If the condition $m_\phi^2 > m_\chi^2 + g\phi(t)$ is not satisfied, $\Delta\omega_k^{(1)} = \Delta k$ is not small and we have "broad-band" resonance. In this case, the effective mass of the χ-particles can be larger than the mass of the ϕ-particles and only the coherent decay of several inflatons can lead to χ-production. For a discussion of the main physical processes in this case see Mukhanov (2005). One of the most interesting consequences of

https://doi.org/10.1017/9781316471524.002 Published online by Cambridge University Press

broad-band resonance is that it can lead to the production of particles that are heavier than the inflaton.

Changing the coupling to $L_{\rm int} = -\frac{1}{2}\tilde{g}\phi^2\chi^2$ does not affect the generic behavior of preheating. We can again obtain a Mathieu equation but with resonant frequency $\omega^{(n)} = nm_\phi$ and width $\Delta\omega^{(n)}/\omega^{(n)} = (2\tilde{g}\phi_0^2/\omega^{(n)2})^n$; see Exercise 1.11. Due to the Pauli exclusion principle, couplings of the inflaton to fermions cannot give raise to parametric resonance. The details of the reheating process and the temperature at the end of reheating depend on the particle physics model describing the coupling of ϕ and χ to other particles, especially to the standard model particles. The reheating temperature can go from 1 TeV $< T <$ 10^{13} GeV.

1.5.4 Resolution of the "Cosmological Problems"

At the end of the reheating process, $\tau = \tau_{\rm rh}$, all the energy is supposed to be thermalized and the Universe is dominated by relativistic particles, satisfying $P = \rho/3$ such that

$$\rho \propto a^{-4}.$$

To determine the duration of inflation necessary in order to solve the horizon problem, we consider the entropy, S_H, contained in a volume that corresponds to one Hubble scale, H_i^{-1}, at the beginning of inflation. Since expansion is adiabatic after inflation, the entropy inside a given physical volume remains constant. The requirement that the present Hubble scale, H_0^{-1}, be smaller than the size of the causal horizon is therefore equivalent to $S_H > S_{H_0}$, where S_{H_0} denotes the entropy inside the volume H_0^{-3}. The entropy inside a causal volume, $H_i^{-3}(a/a_i)^3$, is given by its value

$$S_H \simeq H_i^{-3}\left(\frac{a_{\rm rh}}{a_i}\right)^3 T_{\rm rh}^3,$$

after reheating. The Hubble parameter at the beginning of inflation is

$$H_i^2 \simeq \frac{8\pi}{3m_P^2} W(\phi_i),$$

so that

$$S_H \simeq H_i^{-3}\left(\frac{a_f}{a_i}\right)^3\left(\frac{a_{\rm rh}}{a_f}\right)^3 T_{\rm rh}^3 \simeq \frac{m_P^3}{W_i^{3/2}} e^{3N_{\rm tot}}\frac{\rho_f}{T_{\rm rh}}.$$

For the last $\simeq$ sign we have assumed that the Universe was roughly matter dominated from the end of inflation until the end of reheating, $\rho \propto a^{-3}$ and $\rho_{\rm rh} \sim T_{\rm rh}^4$. With $\rho_f \sim W_f$, this yields

https://doi.org/10.1017/9781316471524.002 Published online by Cambridge University Press

$$S_H \simeq \frac{m_P^3 W_f}{T_{\rm rh} W_i^{3/2}} e^{3N_{\rm tot}}.$$

In order to solve the entropy problem, we require that this entropy is at least as large at the entropy in the present Hubble horizon, $S_H > S_{H_0} \simeq T_0^3 H_0^{-3} \simeq 10^{88}$. This now results in

$$N_{\rm tot} \geq N_{\rm min} = \frac{88}{3}\ln(10) + \ln\left(\frac{T_{\rm rh}^{1/3} W_i^{1/2}}{m_P W_f^{1/3}}\right). \tag{1.173}$$

For example, in a model with $W = \frac{1}{2}m_\phi^2\phi^2$, we have large-field inflation that stops roughly when $\phi = \phi_f \simeq m_P$ so that $W_f = \frac{1}{2}(m_\phi m_P)^2$ and $W_i = \frac{1}{2}(m_\phi \phi_i)^2$. Hence

$$N_{\rm min} = \frac{88}{3}\ln(10) + \frac{1}{3}\ln\left(\frac{T_{\rm rh} m_\phi}{m_P^2}\right) + \ln\left(\frac{\phi_i}{m_P}\right).$$

If $N_{\rm tot} \geq N_{\rm min}$ the horizon problem is also solved. Indeed, since the entropy inside a comoving volume is conserved after inflation, the present volume of radius H_0^{-1} has grown out of a radius that was smaller than H_i^{-1} at the beginning of inflation, and therefore was already in causal contact before the beginning of inflation.

To solve the flatness problem we must enlarge the curvature scale to $R_K(\tau_0) \geq H_0^{-1}$. This is equivalent to $S_K(\tau_0) \geq S_H(\tau_0) \simeq 10^{88}$. With

$$R_K^3(\tau_{\rm rh}) = R_K^3(\tau_i)\left(\frac{a_f}{a_i}\right)^3 = \frac{H_i^{-3}}{|\Omega_i - 1|^{3/2}}\left(\frac{a_f}{a_i}\right)^3,$$

this leads to

$$N_{\rm tot} \geq N_{\rm min} + \frac{1}{2}\log|\Omega_i - 1|. \tag{1.174}$$

Comparing $N_{\rm min}$ with Eq. (1.166), we find that successful inflation with a simple $\frac{1}{2}m_\phi^2\phi^2$ potential requires $\phi_i \gtrsim$ a few times m_P. After an inflationary period that is sufficiently long, so that the conditions (1.173) and (1.174) are satisfied, both the horizon and flatness problems are resolved. During such an inflationary phase also all unwanted relics are diluted by a factor of $\exp(3N_{\rm tot})$.

Finally, it is important to note that we do not require a perfectly homogeneous and isotropic universe, or even thermal equilibrium prior to inflation. We just need a small "patch" in an otherwise arbitrary, chaotic, universe, within which the gradient and kinetic energy are much smaller than the potential energy, so that the slow roll conditions are satisfied. This patch then inflates to encompass the entire present Hubble volume. This idea of "chaotic inflation" goes back to Linde (1989) and it is of course much more satisfactory than a model in which the Universe has to start out with homogeneous and isotropic spatial sections before inflation.

https://doi.org/10.1017/9781316471524.002 Published online by Cambridge University Press

When discussing inflation, one of the most mysterious problems of gravity becomes apparent: while adding a constant to the potential W of the scalar field does not affect any of the other interactions, it severely alters gravity. It modifies cosmic expansion in the same way as adding a cosmological constant. What determines the correct level of a potential? This question is equivalent to the problem of the cosmological constant. Why is the present cosmological constant so small, $\Lambda/(8\pi G) \simeq (2 \times 10^{-3}\,\mathrm{eV})^4$, much smaller than any fundamental energy scale? The problem is even more serious when we remember that in quantum field theory we use the freedom to add or subtract a constant from the potential by absorbing the infinite zero-point energy into it. Furthermore, at each phase transition this zero-point energy changes by a finite, calculable amount. Before the discovery of the accelerated expansion of the Universe, which is most simply interpreted as a cosmological constant, $\Lambda/(8\pi G) \simeq (2 \times 10^{-3}\,\mathrm{eV})^4 \neq 0$, it was justifiable to assume that the freedom of the cosmological constant has to be used in order to annihilate any vacuum energy contribution from quantum field theory, so that the effective cosmological constant would vanish, $\Lambda_{\mathrm{eff}} = \Lambda + 8\pi G W_0 = 0$. Present observations, however, indicate that this compensation takes place only approximately, leaving a small but nonvanishing effective cosmological constant, $\Lambda_{\mathrm{eff}} \neq 0$, which starts to dominate the expansion of the Universe just at present time, when there are sufficiently developed intelligent beings in the Universe that wonder about it. In all the cosmic past, this cosmological constant was completely negligible, and in all the cosmic future, it will be the only relevant contribution to expansion. Only at present it is comparable with the mean mass density of the Universe. Apart from the bizarre value of Λ_{eff}, we thus also have a strange coincidence problem.

This is presently one of the deepest problems of physics. Ordinary quantum field theory does not determine the vacuum energy of quantum fields, but only changes that may happen depending on the external conditions. We may hope that a quantum theory of gravity addresses the cosmological constant problem. The cosmological constant may even represent our first observational data related to quantum gravity.

Exercises

(The exercises marked with an asterisk are solved in Appendix 11 which is not in this printed book but can be found online.)

1.1 **Coordinates**
Find the coordinate transformation leading from the coordinates used in Eq. (1.9) to those of Eq. (1.10) and finally of Eq. (1.8).

https://doi.org/10.1017/9781316471524.002 Published online by Cambridge University Press

1.2 **FL universes are conformally flat**

Show that FL universes are conformally flat (also when the curvature does not vanish) and find the coordinate transformation $(\tau, r) \to (\sigma, \rho)$ such that

$$-d\tau^2 + a^2(\tau)\gamma_{ij}\, dx^i\, dx^j = A^2(\sigma,\rho)\eta_{\mu\nu}\, dX^\mu\, dX^\nu, \quad (1.175)$$

with $\sigma = X^0$ and $\rho^2 = \sum_{i=1}^3 (X^i)^2$.

1.3 **Matter and radiation mixture**

Consider an FL universe containing a mixture of nonrelativistic matter (dust) and radiation with vanishing curvature. The respective densities and pressures are ρ_m, ρ_r, and $P_m = 0$, $P_r = \rho_r/3$. We denote the ratio of radiation to matter by $R = \rho_r/\rho_m$.

(a) Determine w and c_s^2 as functions of R. What is the time dependence of R?
(b) For a given redshift $z_{\rm eq} \gg 1$ of matter and radiation equality determine the scale factor as a function of conformal and of physical time; normalize the scale factor to 1 at equality, $a_{\rm eq} = 1$.
(c) Determine $t_{\rm eq}$ and $\tau_{\rm eq}$ as functions of $z_{\rm eq}$, and H_0.

1.4 **Cosmological constant***

Investigate the dynamics of an FL universe with matter ($P = 0$) and a cosmological constant Λ.

(1) Show for a sufficiently small cosmological constant and positive curvature that the Universe recollapses in a "big crunch," while for a larger cosmological constant or nonpositive curvature, the Universe expands forever.
(b) Show furthermore that for an even higher cosmological constant there are solutions that have no big bang in the past, but issue from a previous contracting phase. The transition from the contracting to an expanding phase is called the "bounce."
(c) Make a plot in the plane $(\Omega_m, \Omega_\Lambda)$ distinguishing the regimes determined earlier.
(d) For case (b), determine (numerically) the redshift of the bounce as a function of Ω_Λ for fixed $\Omega_m = 0.1$. Discuss.

1.5 **Helium recombination**

Write the Saha equation Eq. (1.73) for the two helium recombination processes and use it to determine the helium recombination temperatures, $T_{2\to1}$ and $T_{1\to0}$.

https://doi.org/10.1017/9781316471524.002 Published online by Cambridge University Press

Hint: As a simplifying assumption neglect the fact that helium uses up some of the electrons and simply set $n_e \simeq n_B$.

1.6 **Rayleigh scattering**
The Rayleigh scattering cross section for an atom is $\sigma_{\rm Ra} \simeq \alpha^2/\lambda^4$, where α is the polarizability and λ the photon wavelength. For hydrogen atoms $\alpha \simeq 3.8 \times 10^{-24}{\rm cm}^3$. Show that after recombination the Rayleigh scattering rate of CMB photons on hydrogen atoms is much smaller than the expansion rate H.

1.7 **Distribution functions**
Show that in the nonrelativistic limit, $m \gg T$ both, the Fermi–Dirac and the Bose–Einstein distributions reduce to a Maxwell–Boltzmann distribution and the number and energy density are given by

$$n = \frac{2}{(2\pi)^3} \exp(-(m-\mu)/T)(2\pi m T)^{3/2}, \quad \rho = mn, \quad (1.176)$$

where μ is the chemical potential.

1.8 **Liouville equation**
Using that, in an FL universe the distribution function f depends only on (conformal) time t and $p = \sqrt{\gamma_{ij} p^i p^j}$, derive Eq. (1.88).

1.9 **The neutrino background**
Determine the neutrino cross section for the reaction $e^- + \bar{\nu} \to e^- + \bar{\nu}$ at energy $E_\nu = T_\nu(t_0)$. Compare it with the cross section of the neutrinos detected in the super-Kamiokande experiment. Keeping the efficiency of super-Kamiokande, how large a water tank would you need to detect neutrinos from the cosmic background?

1.10 **Angular diameter distance**
Determine the angular diameter distance to the last scattering surface under the assumptions $K = \Lambda = 0$. Under which angle do we presently see the causal horizon of this time, $a(t_{\rm rec})t_{\rm rec}$? How does this result change if one admits a cosmological constant so that $\Omega_m = 0.3$ and $\Omega_\Lambda = 0.7$?

1.11 **Resonant amplification**
Consider an inflaton coupled to a scalar field with $\mathcal{L}_\chi = -\frac{1}{2}\,\partial_\mu\chi\,\partial^\mu\chi - \frac{1}{2}\,m_\chi^2\chi^2$ and $\mathcal{L}_{\rm int} = -\frac{1}{2}\tilde{g}\phi^2\chi^2$. Consider the equation of motion of χ in the classical background solution $\phi(t) = \phi_0 \cos(m_\phi \tau)$. Neglect the cosmic expansion. Show that the equation for each Fourier mode can be written as a Mathieu equation and discuss the resonance frequencies and widths.

https://doi.org/10.1017/9781316471524.002 Published online by Cambridge University Press