Big Bang Thermal Relics

9

- The universe began hot and dense (the big bang), and thereafter expanded and cooled. The early universe underwent a series of phase changes between thermal equilibria—each one a cosmic soup composed of a different mix of particles. This cosmic history left behind telltale thermal distributions of relic particles: light nuclei, neutrinos, and photons.
- The observed abundances of the light nuclear elements (helium, deuterium, etc.) support the view that they are products of big bang nucleosynthesis around a hundred seconds after the big bang.
- When the universe was around 380 000 years old, photons ceased to interact with matter and began to pass freely though the universe. This primordial light remains today as the cosmic microwave background (CMB), with a blackbody spectrum of temperature $T = 2.725\,\mathrm{K}$.
- Similarly, at an earlier cosmic epoch, neutrinos decoupled to form a cosmic neutrino background. Thus the most abundant particles in our universe are relic photons and neutrinos, with a density of about 400 of each per cubic centimeter.
- The CMB is not perfectly uniform. Its dipole anisotropy is primarily determined by our motion in its rest frame (the comoving Robertson–Walker frame); higher multipoles contain much information about the geometry and the matter/energy content of the universe, as well as the initial density perturbation, from which grew the cosmic structure we see today.

During the epochs immediately after the big bang, the universe was much more compact, and the energy associated with the random motions of matter and radiation was much larger. Thus we say that the universe was much hotter. Space was filled with a plasma, in which various particles could be in thermal equilibrium through high-energy interactions. As the universe expanded, it also cooled. As thermal energy lowered, particles (and antiparticles) either disappeared through annihilation, combined into various composites of particles, or decoupled (i.e., ceased to interact) to become free particles. Consequently, there remain today different kinds of thermal relics left behind by the hot big bang.

One approach to studying the universe's history is to start with some initial state that may be guessed based on our knowledge of (or speculation about)

A College Course on Relativity and Cosmology. First Edition. Ta-Pei Cheng.

particle physics. Then we can evolve this proposed universe forward in the hope of ending up with something like the observed universe today. That we can speak of the early universe with any sort of confidence rests with the idea that the universe passed through a series of equilibria. At such stages, its properties were determined, independently of the details of the interactions, by a few parameters such as the temperature, density, and pressure.

The thermodynamical investigation of cosmic history was pioneered by Richard Tolman (1881–1948). This approach of extracting observable consequences from big bang cosmology was vigorously pursued in the late 1940s and early 1950s by Alexander Friedmann's PhD student George Gamow (1904–1968) and his collaborators, Ralph Alpher (1921–2009) and Robert C. Herman (1914–1997).

Here, we shall first give an overview of the thermal history of the universe, in particular the scale dependence of radiation temperature.

9.1 The thermal history of the universe

Once again, it should be pointed out that the calculations carried out in this chapter are rather crude and for illustration only—to give us a flavor of how in principle cosmological predictions can be made. Realistic calculations would typically use the Boltzmann equation,[1] involving many reaction rates with numerous conditions.

9.1.1 Scale dependence of radiation temperature

For the radiation component of the universe, we can neglect particle masses and chemical potentials[2] much smaller than $k_{\rm B}T$ (where $k_{\rm B}$ is Boltzmann's constant). The number density (per unit volume) distributions with respect to the energy E (of the bosons (minus sign in $\pm$) and fermions (plus sign)) in a radiation gas are

$$dn = \frac{g}{2\pi^2(\hbar c)^3}\frac{E^2\,dE}{e^{E/k_{\rm B}T}\pm 1}, \tag{9.1}$$

where $\hbar$ is Planck's constant (with dimensions of energy×length for $\hbar c$) and g is the number of spin states[3] of the particles making up the radiation. We integrate to get the number densities of the bosons and fermions:

$$n_{\rm b} = \frac{4}{3}n_{\rm f} = \zeta(3)\frac{g}{\pi^2}\left(\frac{k_{\rm B}T}{\hbar c}\right)^3, \tag{9.2}$$

where $\zeta(3) = 1.202$ is the Riemann zeta function.[4] We can derive the thermodynamic relation (the Stefan–Boltzmann law) between radiation energy density and temperature by integration, $u = \int E\,dn \sim T^4$:

$$\rho_{\rm R}c^2 \equiv u_{\rm R} = \frac{g^*}{2}a_{\rm SB}T^4, \tag{9.3}$$

[1] The Boltzmann equation relates the number of a particular species of particle to the difference between its production and disappearance rates.

[2] Except for photons, there is no strong theoretical ground to set the chemical potential μ to zero. Nonetheless, since there is nothing requiring a sizable μ, we shall for simplicity assume $|\mu| \ll k_{\rm B}T$.

[3] Particles with mass and spin s have $2s+1$ spin states (e.g., spin-$\frac{1}{2}$ electrons have two spin states), but massless particles (e.g.,spin-1 photons or spin-2 gravitons) have only two spin states. Note that antiparticles are counted separately, so the electron and positron have four spin states between them. Moreover, Standard Model neutrinos have only one spin state, because only left-handed states interact; they come in three flavors, making six states of left-handed neutrinos and right-handed antineutrinos.

[4] The Riemann zeta function is defined as

$$\zeta(n) = \sum_{l=1}^{\infty}\frac{1}{l^n} = \frac{1}{(n-1)!}\int_0^\infty dx\,\frac{x^{n-1}}{e^x-1}.$$

In particular,

$$\zeta(2) = \pi^2/6 \simeq 1.645,$$
$$\zeta(3) \simeq 1.202,$$
$$\zeta(4) = \pi^4/90 \simeq 1.082.$$

where the radiation energy density has been written as $\rho_R c^2$ (cf. Sidenote 31 in Chapter 8) and a_{SB} is the Stefan–Boltzmann constant,

$$a_{SB} = \frac{3!\zeta(4)}{\pi^2}\frac{k_B^4}{(\hbar c)^3} = \frac{\pi^2 k_B^4}{15(\hbar c)^3} = 4.722\ \text{keV m}^{-3}\,\text{K}^{-4}. \tag{9.4}$$

We have summed the energy contributions of all the constituent radiation particles, so that g^* is the effective number of spin states of all the particles making up the radiation:

$$g^* = \sum_i (g_b)_i + \frac{7}{8}\sum_i (g_f)_i, \tag{9.5}$$

where $(g_b)_i$ and $(g_f)_i$ are the spin multiplicities of the ith species of boson and fermion radiation particles, respectively. The factor 7/8 reflects the different integral values for the fermion distribution, with a plus sign in (9.1), vs. the boson distribution, with a minus sign.

From the number and energy densities, we can also compute the average energies $\bar{E} = \rho_R c^2/n$ with $\bar{E}_b = \frac{6}{7}\bar{E}_f$ for bosons and fermions in the radiation. In particular, for photons,

$$\bar{E}_\gamma = \frac{3\zeta(4)k_B T}{\zeta(3)} = \frac{\pi^4 k_B T}{30\zeta(3)} = 2.701 k_B T. \tag{9.6}$$

Scale dependence of temperature Combining the Stefan–Boltzmann law (9.3) $\rho_R \sim T^4$ with our previously derived relation (8.64) for a radiation-dominated system $\rho_R \sim a^{-4}$, we deduce the scaling property for the radiation temperature:

$$T \propto a^{-1}. \tag{9.7}$$

This expresses, in precise scaling terms, our expectation that temperature is high when the universe is compact, so it cools as it expands. Under this temperature-scaling law, the total number distributions $dN = V\,dn$ (for some volume $V \sim a^3$) from (9.1) are unchanged. Because the radiation energy varies inversely with the wavelength, $E \sim \lambda^{-1} \sim a^{-1}$, the combinations $VE^2\,dE$ and $E/k_B T$ are invariant under scale changes. Thus, as the universe expands and the temperature falls, the form of the blackbody spectrum is maintained.

Remark 9.1 *In the context of the Newtonian interpretation of the cosmological (Friedmann) equations, we can understand energy conservation in an expanding universe as follows: while the total number of radiation particles $N = nV$ does not change during expansion, the total radiation energy ($\sim Nk_B T$) scales as a^{-1}. This loss of radiation energy as the scale a increases is balanced by the increase in the gravitational energy of the universe. The gravitational potential energy is also inversely proportional to the scale, but is negative. Thus, it increases (becomes less negative) with an increase in a.*

Example 9.1 Relation between radiation temperature and cosmic time

The early universe was dominated by radiation, so it grew like $a \propto t^{1/2}$, cf. (8.71). We can drop the curvature (k) term in the Friedmann equation (8.42), and replace $\dot{a}/a$ by $(2t)^{-1}$, so that the radiation energy density is related to cosmic time by

$$\rho_{\rm R} c^2 = \frac{3}{32\pi} \frac{c^2}{G_{\rm N}} t^{-2}. \tag{9.8}$$

The left-hand side can be expressed in terms of the thermal energy by the Stefan–Boltzmann law (9.3):

$$\rho_{\rm R} c^2 = \frac{g^* \pi^2}{30} \frac{(k_{\rm B} T)^4}{(\hbar c)^3}. \tag{9.9}$$

Thus, in a radiation-dominated universe, time is related to temperature by

$$\begin{aligned} t &= \left(\frac{45 \hbar^3 c^5}{16\pi^3 g^* G_{\rm N} k_{\rm B}^4 T^4} \right)^{1/2} \\ &= \frac{0.3012}{\sqrt{g^*}} \left(\frac{T_{\rm Pl}}{T} \right)^2 t_{\rm Pl} = \frac{3.26 \times 10^{20}\,{\rm K}^2 \cdot {\rm s}}{\sqrt{g^*} T^2}, \end{aligned} \tag{9.10}$$

where we have used the Planck time and temperature from (7.28). For an effective multiplicity[5] of $g^* = 10\frac{3}{4}$, this gives an easy-to-remember numerical relation between the cosmic time in seconds and temperature in kelvins:

$$t\,(s) \simeq \frac{10^{20}}{[T(K)]^2}. \tag{9.11}$$

From this estimate we can see that the big bang nucleosynthesis at temperature $T_{\rm bbn} \simeq 10^9$ K (cf. (9.21)) took place about a hundred seconds after the big bang: $t_{\rm bbn} = O(10^2)$ s.

[5] While electrons and positions contribute to the radiation, (9.5) gives $g^* = 2(\text{photons}) + \frac{7}{8} \times 4(\,e^+ \,\&\, e^-) + \frac{7}{8} \times 6(\nu \,\&\, \bar{\nu}) = 10\frac{3}{4}$; (9.11) is good to $\simeq 1\%$. When the electrons and positrons mostly vanish at reheating when $k_{\rm B} T \simeq m_e c^2 \simeq 0.5\,{\rm MeV}$, g^* drops to $7\frac{1}{4}$. Moreover, reheating (Exercise 9.1) increases subsequent temperatures by a factor of $(11/4)^{1/3}$ from what they would be, so $t(T)$ must go up by a factor of $(11/4)^{1/3} \simeq 1.96$. Thus (9.11) is low at later epochs: $t(s) \simeq 2.37 \times 10^{20}/[T({\rm K})]^2$.

9.1.2 Different thermal equilibrium stages

After the big bang (the inflationary epoch to be discussed in Chapter 10), the cooling of the universe allowed the existence of different mixtures of particles in the equilibrium plasma. Composites such as nucleons, nuclei, and atoms were able to form and survive. When the falling thermal energy $k_{\rm B} T$ could no

longer produce various types of particle–antiparticle pairs, the antiparticles annihilated with particles and disappeared.[6] Particles eventually ceased to interact with (decoupled from) the cosmic soup.

Determining what reactions took place to maintain each thermal equilibrium involves dynamical calculations, taking into account reaction rates in an expanding and cooling medium. The basic requirement for a given particle interaction to be significant at a given epoch is that the time interval between particle scatterings be much shorter than the age of the cosmos. This can be expressed as the Gamow condition that the reaction rate Γ must be faster than the expansion rate of the universe as measured by Hubble's constant:

$$\Gamma \geq H. \tag{9.12}$$

The reaction rate Γ of a particle is the product of the number density n of the particles with which it interacts, their relative velocity v, and the reaction cross section σ:[7]

$$\Gamma = nv\sigma. \tag{9.13}$$

Particle velocity enters because nv is the flux of the interacting particles (the number of particles passing through unit area in unit time). The velocity distribution is determined by the thermal energy of the system. The cross section can be measured in a laboratory or predicted by theory. The condition for a transition to a new equilibrium phase is that $\Gamma = H$. Since the cosmic age $\sim H^{-1}$, one can think of this condition as requiring on average one interaction since the beginning of the universe. This condition can be used to solve for the thermal energy and the redshift value at which a new equilibrium stage started. These different thermal stages (ordered by their time since the beginning of the universe and their average thermal energies) can be summarized as follows:

A chronology of the universe

- $\lesssim$ **10^{-43} s** ($k_B T \gtrsim 10^{19}$ GeV) Planck epoch. Quantum gravity is expected to be relevant at such high energy scales.
- $\simeq$ **10^{-35} s** ($k_B T \simeq 10^{16}$ GeV) Inflation. The big bang is described as an exponential expansion in which the scale factor of space grew by something like 30 orders of magnitude owing to vacuum energy during a phase transition likely associated with the grand unification of particle physics. This enormous expansion left a homogeneous, flat universe with the seeds for subsequent formation of cosmic structure.
- $\simeq$ **10^{-35}–10^{-14} s** ($k_B T \simeq 10^{16}$–10^{2} GeV) Early stages of a radiation-dominated universe with all the fundamental Standard Model species present. Any phase transitions or nontrivial dynamics would be due to physics beyond the Standard Model. It is often speculated that this early

[6] Why there was an excess of particles over antiparticles is currently an area of active research in particle physics and cosmology.

[7] You might imagine the cross section σ as the cross-sectional area of one of the interacting particles if the other particle were a point. Properly, the reaction rate should be thermally averaged: $\Gamma = n\overline{\sigma v}$. Since the product σv is nearly constant (except at a resonance or cutoff threshold), this average for our purpose can be trivially done. **Exercise:** check the dimensions of reaction rate as given here.

universe was also populated with yet undiscovered particles such as the supersymmetric particles.

- $\simeq$ **10^{-14}–10^{-10} s** ($k_B T \simeq 100\,\mathrm{GeV}$) Electroweak symmetry breaking of the Standard Model. Up to this time, the gauge bosons of the weak interaction were massless, just like the photons of electromagnetism. After this phase transition, W and Z bosons, as well as all quarks and leptons, gained their masses.
- $\simeq$ **10^{-5} s** ($k_B T \simeq 200\,\mathrm{MeV}$) Quark/gluon confinement. Strongly interacting particles (previously deconfined in a plasma) were bound into hadrons. Unstable hadrons (e.g., pions) shortly fell out of equilibrium, leaving only nucleons.
- $\simeq$ **1 s** ($k_B T \simeq 1\,\mathrm{MeV}$) Neutron–proton freeze-out and neutrino decoupling. Nucleons ceased interconverting into each other through the weak interaction involving neutrinos, thus fixing their ratio (although neutrons would very slowly decay into protons until bound later into stable nuclei). Another consequence of ending weak interactions is that neutrinos decoupled, becoming free-streaming to form the cosmic neutrino background.
- $\simeq$ **4–8 s** ($k_B T \simeq 0.5\,\mathrm{MeV}$) Positron disappearance and photon reheating. The early universe contained comparable numbers of electrons and positrons, with a slight excess ($\simeq 10^{-9}$) of electrons. When the thermal radiation energy fell below the rest energies, they could no longer be produced; they annihilated with one another, leaving only the few leftover electrons present in matter today. This annihilation also boosted the temperature of the cosmic photons (compared with the previously decoupled cosmic neutrinos).
- $\simeq$ **200 s** ($k_B T \simeq 0.1\,\mathrm{MeV}$) Nucleosynthesis. Protons and neutrons combined into charged ions of light nuclei: helium, deuterium, etc.
- $\simeq$ **70 kyr** ($k_B T \simeq 1\,\mathrm{eV}$) Radiation–matter equality. The radiation energy density fell below that of nonrelativistic matter (which falls more slowly), thereby changing the dynamics (power-law expansion) of the universe.
- $\simeq$ **380 kyr** ($k_B T \simeq 0.3\,\mathrm{eV}$) Photon decoupling. The thermal energy of radiation dropped too low to ionize the just-formed neutral atoms. Photons could free-stream to form the cosmic microwave background observed today.
- $\simeq$ **100 Myr** Formation of stars and galaxies.
- $\simeq$ **8 Gyr** Transition to an accelerating universe; $\simeq$ **10 Gyr** Equality of matter and dark energy. The matter energy fell below the ever-constant dark energy density. As the dark energy came to dominate, the expansion of the universe stopped slowing down and began to accelerate.
- $\simeq$ **14 Gyr** The present epoch.

Thermal relics of light nuclei, neutrinos, and CMB

In the following sections, we shall discuss two particular epochs in the history of the universe that left observable relics in our present-day cosmos. In Section 9.2, we study the epoch of big bang nucleosynthesis at $t_{\rm bbn} \simeq 200\,{\rm s}$. In Section 9.3, we study the epoch at $t_\gamma \simeq 380\,000$ years when the photons decoupled to form the CMB radiation.

In Box 9.2 in Section 9.3.1, we shall also briefly comment on epochs prior to the two mentioned above. About a second after the big bang, neutrinos decoupled to form the first half of the main thermal relic particles of the universe. This was followed by reheating (Exercise 9.1): the disappearance of positrons (and most of the electrons) from the cosmic soup. Because this reheating took place after neutrino decoupling, the relic neutrinos (as yet undetected) should have a cooler thermal distribution than the CMB photons, which decoupled at a later epoch.

9.2 Primordial nucleosynthesis

When we look around our universe, we see mostly hydrogen but very little of the heavy elements. The observed amounts of the heavy elements can all be satisfactorily accounted for by the known nuclear reactions taking place inside stars and supernovae. On the other hand, everywhere we look, besides hydrogen, we also see a significant amount of helium. (The helium abundance can be deduced from measurements of the intensities of spectral lines of helium-4 (^{4}He) in stars, planetary nebulae, and galactic as well as extragalactic HII regions.) The data indicate a ^{4}He mass fraction close to a quarter:

$$y \equiv \left(\frac{^4{\rm He}}{{\rm H} + {}^4{\rm He}}\right)_{\rm mass}, \qquad \text{with} \quad y_{\rm obs} \simeq 0.24. \tag{9.14}$$

Similarly, we observe much smaller uniform abundances of other light elements: deuterium (D), helium-3 (^{3}He), and lithium-7 (^{7}Li). These light nuclear elements are theorized to have been synthesized in the early universe by the path described below.

Proton–neutron equilibrium and freeze

By the time of $10^{-5}\,{\rm s}$ after the big bang (when the thermal energy was about 200 MeV), quarks had coalesced into nucleons, and unstable particles had vanished from the rapidly cooling universe. The cosmic soup was then composed of protons (p^+), neutrons (n), electrons (e^-), positrons (e^+), three flavors of neutrinos (ν) and antineutrinos ($\bar\nu$), and photons. Up till the first second after the big bang, neutrons and protons could interconvert by weak interactions such as

$$n + e^+ \rightleftarrows p^+ + \bar\nu \quad \text{and} \quad p^+ + e \rightleftarrows n + \nu. \tag{9.15}$$

Neutrons and protons were in thermal equilibrium, so their number ratio $\lambda = n_n/n_p$ was governed by the Boltzmann distribution $\exp(-E/k_B T)$:

$$\lambda = \exp\left[-\left(\frac{E_n - E_p}{k_B T}\right)\right] \simeq \exp\left[-\left(\frac{m_n - m_p}{k_B T}\right)c^2\right] \simeq \exp\left(\frac{-1.3\,\text{MeV}}{k_B T}\right) \tag{9.16}$$

for nonrelativistic neutrons and protons with a difference in rest energy of 1.3 MeV.

When thermal energy fell below $k_B T_{fr} \simeq 1$ MeV ($T_{fr} \simeq 10^{10}$ K at time $t_{fr} \simeq 1$ s per (9.11)), the interconversion rate Γ fell below the rate of relative expansion H. We call this the freeze-out time, because when the neutron and protons left thermal equilibrium, their ratio was frozen at $\lambda_{fr} \simeq 1/6$. In fact, their ratio would subsequently drop owing to neutron beta decay, $n \longrightarrow p + e + \bar{\nu}$, but not too rapidly, since the mean lifetime of a free neutron, $\tau_n \simeq 880$ s, is considerably longer than the then age of the universe.

Epoch of primordial nucleosynthesis

Protons and neutrons tend to join (through strong interactions) into bound nuclear states.[8]. However, during this epoch, as soon as they were formed, they were blasted apart by energetic photons (photodissociation):

$$p^+ + n \rightleftarrows D^+ + \gamma. \tag{9.17}$$

As the universe cooled, there were fewer photons energetic enough to photodissociate the deuteron (the reaction proceeding from right to left), so deuterons accumulated. The following nucleon-capture reactions could then build up heavier elements:

$$D^+ + n \rightleftarrows {}^3H^+ + \gamma, \quad D^+ + p^+ \rightleftarrows {}^3He^{2+} + \gamma, \tag{9.18}$$

and

$${}^3H^+ + p^+ \rightleftarrows {}^4He^{2+} + \gamma, \quad {}^3He^{2+} + n \rightleftarrows {}^4He^{2+} + \gamma. \tag{9.19}$$

These reversible nuclear reactions would not go further to form heavier nuclei; because helium-4 is particularly tightly bound, the formation of nuclei with five nucleons is not energetically favored. Lacking a stable $A = 5$ nucleus, the synthesis of lithium with mass numbers six or seven from stable helium-4 requires the much less abundant deuterons or tritium:

$${}^4He^{2+} + D^+ \rightleftarrows {}^6Li^{3+} + \gamma \quad \text{and} \quad {}^4He^{2+} + {}^3H^+ \rightleftarrows {}^7Li^{3+} + \gamma. \tag{9.20}$$

Big bang nucleosynthesis could not progress further to produce even heavier elements ($A > 7$), because there is no stable $A = 8$ element. Beryllium-8 almost

[8] A nucleus is composed of Z (the atomic number) protons and N neutrons, giving it the mass number $A = Z + N$. Since chemical properties are determined by the number of protons, we can identify Z from the name of the element; e.g., hydrogen has $Z = 1$ and helium $Z = 2$. Nuclei having the same Z but different numbers of neutrons are isotopes. From the mass number, usually denoted by a superscript on the left side of the nucleus symbol, we can figure out the number of neutrons. The most abundant helium isotope is helium-4 (^{4}He) with two neutrons, followed by helium-3 (^{3}He) with one neutron. Hydrogen isotopes have specific names: the deuteron has one proton and one neutron ($^2\text{H} \equiv \text{D}$), and the tritium nucleus (^{3}H) has two neutrons. Superscript + to the right of the nucleus symbols denotes positively charged ions.

immediately disintegrates into a pair of helium-4's. (Only by packing very large concentrations of helium-4 at high temperatures can mature stars proceed to form stable carbon-12.)

According to a detailed rate calculation, below the thermal energy

$$k_{\mathrm{B}} T_{\mathrm{bbn}} \simeq O(0.1)\,\mathrm{MeV}, \tag{9.21}$$

photons were no longer energetic enough to photodissociate the bound nuclei. This corresponds to a temperature on the order of $T_{\mathrm{bbn}} \simeq 10^9\,\mathrm{K}$ and a cosmic age (cf. (9.11)) of $t_{\mathrm{bbn}} \simeq O(10^2)\,\mathrm{s}$. The net effect of the above reactions from (9.17) to (9.19) was to bind almost all the neutrons into helium-4 nuclei, because there were more protons than neutrons:

$$2\mathrm{n} + 2\mathrm{p}^+ \longrightarrow {}^4\mathrm{He}^{2+} + \gamma. \tag{9.22}$$

We can then conclude that the resultant number density n_{He} for helium-4 must equal half of the neutron density n_{n}. The number density of hydrogen n_{H} (the protons leftover after all the others were bound with neutrons into helium ions) should equal the proton number density minus that of the neutrons. The helium mass m_{He} is about four times the nucleon mass m_{N}. This yields a helium mass fraction of

$$y \equiv \left(\frac{{}^4\mathrm{He}}{\mathrm{H} + {}^4\mathrm{He}}\right)_{\mathrm{mass}} = \frac{n_{\mathrm{He}} m_{\mathrm{He}}}{n_{\mathrm{H}} m_{\mathrm{H}} + n_{\mathrm{He}} m_{\mathrm{He}}}$$

$$= \frac{(n_{\mathrm{n}}/2) \cdot 4 m_{\mathrm{N}}}{(n_{\mathrm{p}} - n_{\mathrm{n}}) m_{\mathrm{N}} + (n_{\mathrm{n}}/2) \cdot 4 m_{\mathrm{N}}} = \frac{2\lambda}{1+\lambda}, \tag{9.23}$$

where λ is the neutron-to-proton ratio, $n_{\mathrm{n}}/n_{\mathrm{p}}$. Between freeze-out $t_{\mathrm{fr}} \simeq 1\,\mathrm{s}$ and nucleosynthesis $t_{\mathrm{bbn}} \simeq O(100)\,\mathrm{s}$, the ratio dropped owing to free-neutron decay[9] from $\lambda_{\mathrm{fr}} \simeq 1/6$ to its ultimate value $\lambda \simeq 1/7$. This yields a primordial helium-4 mass fraction very close to the observed ratio of 0.24:

$$y = \frac{2\lambda}{1+\lambda} \simeq \frac{2/7}{8/7} = \frac{1}{4}. \tag{9.24}$$

In summary, once deuterium was formed by the fusion of protons and neutrons, this chain of fusion reactions proceeded rapidly, so that by about 180 s after the big bang, nearly all the neutrons were bound into helium. Since these reactions were not perfectly efficient, trace amounts of deuterium and helium-3 were left over. (Any leftover tritium would also decay into helium-3.) Formation of nuclei beyond helium progressed slowly; only small amounts of lithium-6 and -7 were synthesized in the big bang. Again, we must keep in mind that the crude calculations presented here are for illustrative purposes only. They are meant to give us a simple picture of the physics involved in such cosmological deductions. Realistic calculations often include many simultaneous reactions. The detailed computations leading to theoretical predictions such as (9.24) must also consider the following:

[9] Neutrons are stable once bound into nuclei.

1. The Standard Model assumes three flavors of light neutrinos (ν_e, ν_μ, and ν_τ), which we include in the effective degrees of freedom g^* in (9.5) for the radiation energy density. Additional neutrinos or other light particles would increase the expansion rate, making freeze-out earlier and hotter, thereby increasing the neutron/proton ration and the helium-4 mass fraction. The observed abundance thus constrain new theories that include more neutrinos or any light exotic particles beyond the Standard Model.
2. The baryon mass density ρ_B affects the cooling rate of the universe. Deuterium is particularly sensitive to ρ_B. Thus we can use the observed abundance of deuterium (one in every 300 000 hydrogens!) to constrain the baryon density.[10] The best fit, as shown in plots like Fig. 9.1, is at $\rho_B \simeq 0.5 \times 10^{-30}\,\mathrm{g/cm^3}$, or, as a fraction of the critical density:

$$\Omega_B = 0.0475 \pm 0.0006. \tag{9.25}$$

As we already pointed out in Section 8.1.3, when compared with the total mass density $\Omega_M = \Omega_B + \Omega_{DM} \simeq 0.31$, this shows that ordinary atomic matter (baryons) is only a small part of the matter in the universe. Furthermore, we can obtain an estimate of baryon number density n_B by dividing the baryon energy density $\rho_B c^2 = \Omega_B \cdot \rho_c c^2 \simeq 220\,\mathrm{MeV/m^3}$ by the energy

[10] As can be seen from (9.17)–(9.19), the production of all light elements passed through deuterium. The remnant abundance observed today of such an intermediate state is very sensitive to the reaction and cooling rates. This Ω_B value deduced from nucleosynthesis is confirmed by other cosmological observations, notably the CMB anisotropy. In fact, the quoted value in (9.25) is from the CMB summary by the Planck Collaboration (2014).

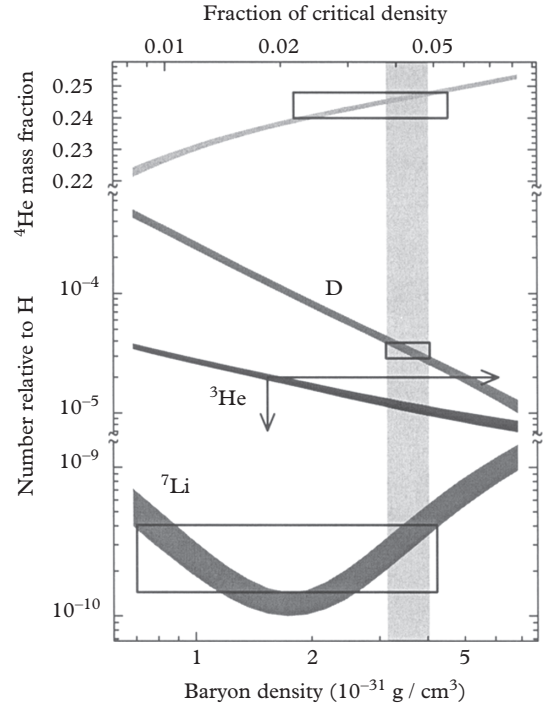

Figure 9.1 *The abundances of light nuclear elements vs. the baryon mass density ρ_B of the universe. The curves are big bang nucleosynthesis predictions, and the boxes are observational results: the vertical heights represent uncertainties in observation, and the horizontal widths the ranges of ρ_B for which theory can accommodate observation. The shaded vertical column represents the range of ρ_B for which theory and observation agree for all four elements. Its uncertainty (the width of the column) is basically determined by the deuterium abundance, which is well measured and strongly dependent on ρ_B. The graph is reproduced with permission from (Burles, Nollett, and Turner 2001). ©2001 American Astronomical Society.*

of each nucleon, whose average can be taken to be the rest energy of a nucleon (939 MeV), because these particles are nonrelativistic:

$$n_{\mathrm{B}} \simeq 0.23/\mathrm{m}^3. \tag{9.26}$$

9.3 Photon decoupling and cosmic microwave background

The early universe after nucleosynthesis contained a plasma of photons and charged matter coupled by their mutual electromagnetic interactions. As the universe expanded and cooled, baryonic matter (ions and electrons) congealed into neutral atoms, so the cosmic soup lost its ability to entrap the photons. These free thermal photons survived as the CMB radiation we see today. The nearly uniformly distributed relic photons obey a blackbody spectrum. Their discovery gave strong support to the hot big bang theory of the beginning of our universe, as it is difficult to think of any other alternative to account for the existence of such a physical phenomenon on the cosmic scale. Furthermore, its slight temperature fluctuation, the CMB anisotropy, is a picture of the early universe. Careful study of this anisotropy has furnished and will continue to provide us with detailed information about the history and composition of the universe. This is a major tool for quantitative cosmology.

9.3.1 Universe became transparent to photons

The epoch when charged nuclear ions and electrons combined into neutral atoms is called the photon-decoupling time[11] t_γ. This took place when the thermal energy of the photons dropped below the threshold required to ionize the newly formed atoms. Namely, the dominant reversible reaction during the age of ions,

$$\mathrm{e}^- + \mathrm{p}^+ \rightleftarrows \mathrm{H} + \gamma, \tag{9.27}$$

ceased to proceed from right to left when the photon energy fell below the ionization energy. All the charged electrons and ions were swept up and bound themselves into stable neutral atoms.

One would naturally expect the thermal energy at the decoupling time, $k_{\mathrm{B}} T_\gamma$, to have been comparable to the typical atomic binding energy $O(\mathrm{eV})$. In fact, a detailed calculation yields the thermal energy and redshift of this cosmic epoch respectively as

$$k_{\mathrm{B}} T_\gamma \simeq 0.26\ \mathrm{eV} \quad \text{and} \quad z_\gamma \simeq 1100. \tag{9.28}$$

Dividing out Boltzmann's constant k_{B}, this energy corresponds to a photon temperature of

$$T_\gamma \simeq 3000\,\mathrm{K}. \tag{9.29}$$

[11] The photon-decoupling time is also referred to in the literature as the recombination time. We do not use this terminology often, as up to this time ions and electrons had never been combined. The name has been used because of the analogous situation in the interstellar plasma, where such atomic formation is indeed a recombination.

Also, we note that the average photon energy (cf. (9.6)) at decoupling was

$$\bar{E}_\gamma = 2.7 \times 0.26\,\text{eV} \simeq 0.70\,\text{eV}. \tag{9.30}$$

From (8.35), this redshift factor in (9.28) also tells us that the universe was about a thousand times smaller in linear dimension, and by (8.64) a trillion times denser, on average, than it is today.

After the cosmic time t_γ, the decoupled photons could travel freely through the universe, but they kept the blackbody spectrum, whose shape was unchanged as the universe expanded. These relic photons cooled according to the scaling law $T \propto a^{-1}$. Thus, the big bang cosmology predicts that everywhere in the present universe there should be a sea of primordial photons following a blackbody spectrum.

What should the photon temperature be now? From the estimates of $T_\gamma \simeq 3000$ K and $z_\gamma \simeq 1100$, we can use (9.7) and (8.35) to deduce

$$T_{\gamma,0} = a_\gamma T_\gamma = \frac{T_\gamma}{1 + z_\gamma} \simeq 2.7\,\text{K}. \tag{9.31}$$

A blackbody spectrum of this temperature $T_{\gamma,0}$ has its maximal intensity (power per unit area per unit wavelength)[12] at the wavelength $\lambda_{\max}$ such that $\lambda_{\max} T_{\gamma,0} \simeq 0.290\,\text{cm}\cdot\text{K}$ (known as the Wien displacement constant). Thus $T_{\gamma,0} \simeq 2.7\,\text{K}$ implies a thermal spectrum with the maximal energy density at $\lambda_{\max}$ on the order of a millimeter—namely, a relic background radiation in the microwave range.[13]

[12] The reader should be cautioned not to confuse the maximal energy density per unit wavelength, discussed here, with the maximal energy density per unit frequency, as shown for example in Fig. 9.2. They are related but not the same.

[13] While this electromagnetic radiation is outside the visible range, we can still "see" it, because such a microwave noise constitutes a percentage of the (cathode-ray) television snow between channels.

Example 9.2 An estimate of the recombination time

The redshift can be translated into the cosmic age with the following estimate. To the extent that we can ignore the repulsive effect of dark energy, the universe has been matter-dominated since the photon-decoupling time,[14] so the time dependence of the scale factor is $a \propto t^{2/3}$, cf. (8.71):

$$\frac{a_0}{a_\gamma} = \left(\frac{t_0}{t_\gamma}\right)^{2/3} = \frac{1 + z_\gamma}{1 + z_0}. \tag{9.32}$$

As it turns out, the age of the universe is fairly close to the Hubble time $t_0 \simeq 14$ Gyr. We get

$$t_\gamma = (1100)^{-3/2} \times 14\,\text{Gyr} \simeq 3.8 \times 10^5\,\text{years} \tag{9.33}$$

[14] As we shall discuss below, the universe ceased to be radiation-dominated way before the recombination time. Also, the dark energy became significant only recently (on the cosmic timescale)—see the "cosmological coincidence problem" to be discussed in Section 10.2.2 in particular (10.32).

In summary, photons in the early universe were tightly coupled to ionized matter (especially electrons) through Thomson scattering. Such interactions stopped

around a redshift of $z_\gamma \simeq 1100$, when the universe had cooled sufficiently to form neutral atoms (mainly hydrogen). Ever since this last-scattering time, the photons have traveled freely through the universe, and have redshifted to microwave frequencies as the universe expanded. This primordial light should appear today as the CMB thermal radiation with a temperature of about 3 K.

Box 9.1 The discovery of CMB radiation

The observational discovery of the CMB radiation was one of the great scientific events of the modern era. It made the big bang cosmology much more credible, as it is difficult to see how else such thermal radiation could have been produced. The discovery and its interpretation also constitute an interesting story. Gamow (1946, 1948), Alpher, and Herman (1948) first predicted that a direct consequence of the big bang model is the presence of a relic background of radiation with a temperature of a few degrees. However, their contribution was not widely appreciated, and no effort was mounted to detect such a microwave background. Only in 1964 did Robert Dicke (1916–1997) at Princeton University rediscover this result and lead a research group (including James Peebles, Peter Roll, and David Wilkinson) to detect this CMB. While they were constructing their apparatus, Dicke was contacted by Arno Penzias (1933–) and Robert W. Wilson (1936–) at the nearby Bell Laboratories. Penzias and Wilson had used a horn-shaped microwave antenna over the preceding year to do astronomical observations. This Dicke radiometer had originally been used in a trial satellite communication experiment and was known to have some excess noise. Not content to ignore it, they made a careful measurement of this background radiation, finding it to be independent of direction, time of day, and season of the year. While puzzling over the cause of such radiation, they were informed by one of their colleagues of the Princeton group's interest in the detection of a cosmic background radiation. (Peebles had given a colloquium on this subject at another university.) This resulted in the simultaneous publication of two papers: one in which Penzias and Wilson (1965) announced their discovery, the other by the Princeton group (Dicke et al. 1965) explaining the cosmological significance of the discovery.

Because of microwave absorption by water molecules in the atmosphere, it is desirable to carry out CMB measurements at locations having low humidity and/or at high altitude. Thus some of the subsequent observations were done with balloon-borne instruments launched in Antarctica (low temperature, low humidity, and high altitude)—or, even better, above the atmosphere from a satellite. Satellite observations were first accomplished in the early 1990s by the Cosmic Background Explorer (COBE) satellite observatory,

continued

Box 9.1 *continued*

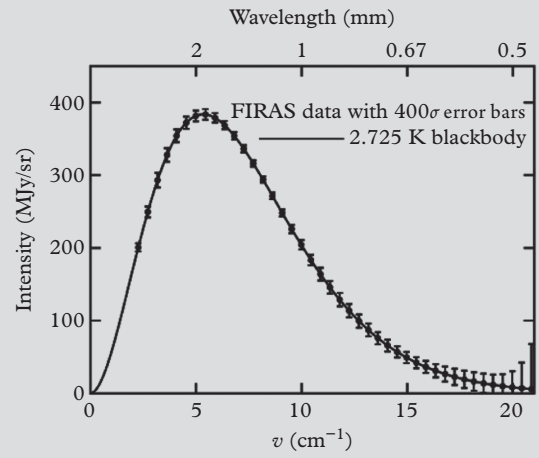

Figure 9.2 *The cosmic background radiation spectrum measured by the FIRAS instrument on the COBE satellite, showing a perfect fit to the blackbody distribution. The horizontal axis is the frequency ∼ 1/(wavelength in cm). The vertical axis is the power per unit area per unit frequency per unit solid angle in megajanskies per steradian. In order to make the error bars visible, these estimated uncertainties have been multiplied by a factor of 400. The graph is based on data from (Fixsen et al. 1996). ©1996 American Astronomical Society.*

led by John Mather (1946–) and George Smoot (1945–), which obtained results (Fig. 9.2) showing that the CMB radiation followed a perfect blackbody spectrum with a temperature

$$T_{\gamma,0} = 2.725 \pm 0.002\,\mathrm{K}. \tag{9.34}$$

The COBE observations not only confirmed that the thermal nature of the cosmic radiation was very uniform (the same temperature in every direction), but also discovered a minute anisotropy at the microkelvin level. This has been interpreted as resulting from the matter density perturbations, that, through subsequent gravitational clumping, gave rise to the cosmic structures we see today: galaxies, clusters of galaxies, voids, etc.[15]

[15] This will be discussed further in Section 9.3.2.

The photon and baryon numbers

Knowing the CMB photon temperature $T_{\gamma,0} = 2.725\,\mathrm{K}$, we can calculate the relic photon number density via (9.2):

$$n_{\gamma,0} = \frac{2.404}{\pi^2}\left(\frac{k_\mathrm{B}T_{\gamma,0}}{\hbar c}\right)^3 \simeq 410/\mathrm{cm}^3. \tag{9.35}$$

That is, there are now in the universe, on average, 410 photons for every cubic centimeter. Clearly this density is much higher than the baryon number density obtained in (9.26). The baryon-to-photon number ratio, with more precise inputs, comes out to be

$$\frac{n_\mathrm{B}}{n_\gamma} = (6.05 \pm 0.07) \times 10^{-10}. \tag{9.36}$$

For every proton or neutron, there are almost two billion photons. This explains why the average energy at decoupling (0.70 eV from (9.30)) fell so far below the hydrogen ionization energy of 13.6 eV before ionization stopped. There was such a high density of photons that even though the average photon energy was only 0.70 eV, there remained up to that time a sufficient number of high-energy photons (at the tail end of the distribution) to hold off the transition to a new equilibrium phase.

This ratio (9.36) should hold all the way back to the photon-decoupling time, because not only was the number of free photons unchanged, but also the baryon number, since all the interactions in this low-energy range (in fact, all the Standard Model interactions we have ever observed) respect the law of baryon number conservation. In Exercise 9.2, this density ratio will be used to estimate the cosmic time at which the universe switched from being radiation-dominated to matter-dominated.

Box 9.2 The cosmic neutrino background

We have already discussed the cosmic microwave background formed from decoupled photons. There is another cosmic background of comparable abundance, formed by decoupled neutrinos. Neutrinos have only weak interaction. Their collision cross section with other particles (e.g., electrons) is small. This cross section has a strong energy dependence. In the early universe, high-energy neutrinos interacted strongly enough with the cosmic particle soup to be in thermal equilibrium. They played a role in keeping neutrons and protons in equilibrium (9.15) until freeze-out at $k_\mathrm{B}T_\mathrm{fr} \simeq 1\,\mathrm{MeV}$. Shortly thereafter, when the temperature fell to about $k_\mathrm{B}T_\nu \simeq 0.3\,\mathrm{MeV}$, they decoupled and become free-streaming. Thus the cosmic neutrino background was formed considerably earlier than the CMB (at 0.26 eV), prior even to primordial nucleosynthesis (at 0.1 MeV). While neutrinos are about as numerous as photons, the present numbers of all other particles, including dark matter particles, are very much smaller—about one per every few billion photons and neutrinos, cf. (9.36).

continued

Box 9.2 *continued*

Cosmic neutrino temperature

According to (9.7), the radiation temperature scales as a^{-1}, regardless of whether or not the particles are coupled. Thus one would expect the temperature of the cosmic neutrino background to be the same as that of the CMB. This is not the case, because in between the neutrino and photon decoupling times, the photon temperature got a boost from electron–positron annihilation (into photons). As a result, the photon temperature is somewhat higher than the neutrino temperature (see Exercise 9.1):

$$T_\gamma = \left(\frac{11}{4}\right)^{1/3} T_\nu. \tag{9.37}$$

Given that the CMB has at the present epoch a temperature $T_\gamma = 2.7\,\mathrm{K}$, the cosmic neutrino background should have a temperature of 1.9 K.

Neutrino number density

From the neutrinos' temperature, one can fix their number density via (9.2). Because neutrinos are fermions and photons are bosons, we have

$$\frac{n_\nu}{n_\gamma} = \frac{3}{4}\left(\frac{T_\nu}{T_\gamma}\right)^3 \frac{g_\nu}{g_\gamma}. \tag{9.38}$$

As mentioned in Sidenote 3, there are two photon states and six neutrino states (left-handed neutrinos and right-handed antineutrinos for each of the three lepton flavors). Plugging in these multiplicities and the temperature ratio from (9.37) yields

$$\frac{n_\nu}{n_\gamma} = \frac{9}{11}. \tag{9.39}$$

Neutrino contribution to the radiation energy density

From the Stefan–Boltzmann law for the radiation energy density (9.3), $u_\mathrm{R} \propto g^* T^4$, with g^* being the effective spin degrees of freedom, as shown in (9.5), $g^*_\gamma = 2$ and $g^*_\nu = \frac{7}{8} \times 6$. Therefore,

$$\frac{\rho_\nu}{\rho_\gamma} = \frac{u_\nu}{u_\gamma} = \frac{g^*_\nu}{g^*_\gamma}\left(\frac{T_\nu}{T_\gamma}\right)^4 = \frac{21}{8}\left(\frac{4}{11}\right)^{4/3} = 0.68, \tag{9.40}$$

where to reach the numerical result we have again used the temperature ratio of (9.37).

Exercise 9.1 Photon temperature boost by e^+e^- annihilation

Since neutrinos and photons were once coupled and in thermal equilibrium, their temperatures were the same: $T'_\nu = T'_\gamma$. *The reaction* $e^+ + e^- \leftrightarrows \gamma + \gamma$ *ceased to proceed from right to left when the photon energy fell below* 0.5 MeV.[16] *The disappearance of positrons increased the photons' number and hence their temperature. This temperature boost can be calculated through the entropy conservation condition. Entropy* S *is related to energy* U *as* $dS = (1/T)\,dU = (V/T)\,du$, *where* V *and* u *are respectively the volume and energy density. Given* $u \propto g^* T^4$ *from (9.3), the key entropy dependences can be identified as* $S \propto g^* V T^3$. *By comparing the volume and photon temperature change as required by the entropy conservation condition* $S'_{e^+} + S'_{e^-} = S_\gamma - S'_\gamma$ *in the annihilation reaction and the corresponding volume and neutrino temperature change for the uncoupled neutrinos* $S'_\nu = S_\nu$, *show that the final photon and neutrino temperatures are related by (9.37):* $T_\gamma = (11/4)^{1/3}\, T_\nu$.

[16] Recall that the rest energy of an electron or positron is about 0.5 MeV.

Exercise 9.2 The radiation–matter equality time

The early universe was radiation-dominated; it then gave way to a matter-dominated system. The radiation–matter equality time t_{RM} *is defined to be the cosmic time at which the energy densities of radiation and matter were equal:*

$$1 = \frac{\rho_R(t_{RM})}{\rho_M(t_{RM})} = \frac{\Omega_R(t_{RM})}{\Omega_M(t_{RM})}. \tag{9.41}$$

Calculate t_{RM} *by the following steps:*

(a) From the scaling behavior of the radiation and matter densities, relate the scale factor a_{RM} *at the radiation–matter equality time to the matter-to-radiation density ratio now,* $\Omega_M(t_0)/\Omega_R(t_0)$.

(b) This density ratio can be calculated to have the value of 3300 *with the following inputs: radiation is composed of photons and neutrinos, so the total radiation energy density* $\Omega_R(t_0)$ *is related to the photon density by (9.40); the matter content* $\Omega_M(t_0)$ *can be deduced from the baryon fraction of matter* $\Omega_B(t_0)/\Omega_M(t_0) \simeq 0.05/0.32$ *and the photon-to-baryon number ratio* n_γ/n_B. *The energy of baryonic matter* $E_B(t_0)$ *can be calculated by adding up the nucleon rest energies, while the photon energy* $\bar{E}_\gamma(t_0)$ *can be deduced from its value of* 0.7 eV *at redshift* $z_\gamma \simeq 1100$.

(c) Follow the worked Example 9.2, using the result for a_{RM} *from parts (a) and (b) and a cosmic age* $t_0 = 14$ Gyr *to show that the radiation–matter dominance transition happened approximately* 73 000 *years after the big bang.*

9.3.2 CMB anisotropy as a baby picture of the universe

The CMB shows a high degree of isotropy. After subtracting off the Milky Way foreground radiation, one obtains in every direction the same blackbody temperature. However, the isotropy is not perfect. The great achievement of COBE (improved by the better angular resolution of NASA's Wilkinson Microwave Anisotropy Probe (WMAP), and now of the European Space Agency's Planck satellite)[17] was the first detection of slight variations of temperature: first at the 10^{-3} level associated with the motion of our Local Group of galaxies, then at the 10^{-5} level, which, as we shall explain, holds the key to our understanding the origin of structure in the universe—how the primordial plasma evolved into stars, galaxies, and clusters of galaxies. Furthermore, the CMB fluctuation provides us with another means to measure the matter/energy content of the universe, as well as many cosmological parameters. The free-streaming photons of the CMB were created around $t_\gamma \simeq 3.8 \times 10^5$ years, and the universe has an approximate age of $t_0 \simeq 1.4 \times 10^{10}$ years. If one were to regard the universe as a one-hundred-year-old person, the CMB anisotropy image would be like her one-day-old baby picture.

The dipole anisotropy Although each point on the sky has a blackbody spectrum, in one half of the sky the spectrum corresponds to a slightly higher temperature, while the other half is slightly cooler with respect to the average background temperature: $\delta T/T \simeq 1.237 \times 10^{-3} \cos\theta$, where θ is measured from the hottest spot on the sky. The dipole distortion is a simple Doppler shift, caused by the net motion of our own galaxy due to the gravitational attraction resulting from the uneven distribution of masses in our cosmic neighborhood. Namely, it shows directly that the Local Group is traveling toward the Virgo Cluster of galaxies at about 600 km/s. This peculiar motion[18] is measured with respect to the frame in which the CMB is isotropic.

The existence of such a CMB rest frame does not contradict special relativity. SR only says that no internal physical measurements can detect absolute motion. Namely, physics laws must be covariant; they may not single out an absolute rest frame. Covariance does not mean that we cannot compare motion relative to a cosmic structure such as the microwave background. Space may not be a thing, but the CMB certainly is. Nor does the CMB rest frame violate the isotropy assumption underlying Robertson–Walker spacetime. Each point in an isotropic spacetime must have an observer (defined to be at rest in the comoving frame) to whom the universe appears the same in every direction. But there is no reason to expect the earth (or the sun, the Milky Way, or the Local Group) to be at rest. There is certainly no expectation that every observer sees the same thing every way he looks. At the same time, there is no definitive explanation why the CMB rest frame defines the inertial frames for us. However, Mach's principle would certainly suggest that this is not a coincidence.[19]

[17] WMAP was preceded by other groups such as the Boomerang and Maxima high-altitude balloon observations in the late 1990s. See the further discussion in Section 10.1, where we discuss the CMB evidence for a flat universe.

[18] The quoted number represents the observational result after subtracting out the orbital motion of COBE around the earth (∼8 km/s) and the seasonal motion of the earth around the sun (∼30 km/s). The measured value is the vector sum of the orbital motion of the solar system around the galactic center (∼220 km/s), the motion of the Milky Way around the center of mass of the Local Group of galaxies (∼80 km/s), and the motion of the Local Group (630 ± 20 km/s) in the general direction of the constellation Hydra. The last, the peculiar motion of our small galaxy cluster toward the large mass concentration in the neighboring part of the universe, reflects the gravitational attraction by the very massive Virgo Cluster at the center of our Local Supercluster, which is in turn accelerating toward the Hydra–Centaurus Supercluster.

[19] This paragraph ties together several fundamental principles motivating GR and cosmology. Review Section 1.1 for a discussion of covariance and Mach's principle (in particular Sidenote 2 in Chapter 1 and Sidenote 1 in Chapter 8), Section 4.1.2 for Einstein's conviction that "space is not a thing," and Section 8.2.1 for the assumption of an isotropic universe.

Exercise 9.3 Temperature dipole anisotropy as Doppler effect

By converting temperature variation to that of light frequency, show that the Doppler effect implies that an observer moving with a nonrelativistic velocity $\mathbf{v}$ *through an isotropic CMB would see a temperature dipole anisotropy* $\delta T/T = (v/c)\cos\theta$*, where the angle* θ *is measured from the direction of the motion.*

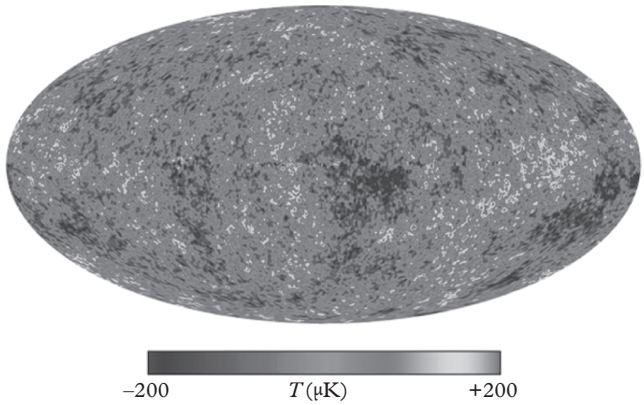

Figure 9.3 *The temperature fluctuation of the CMB is a snapshot of the baby universe at the photon decoupling time. The foreground emission of the Milky Way has been subtracted out. The data are from nine-year observations by the WMAP Collaboration, reproduced with permission from (Bennett et al. 2013). ©2013 American Astronomical Society.*

Physical origin of the temperature inhomogeneity Aside from this 10^{-3}-level dipole anisotropy, the background radiation is seen to be quite isotropic. The CMB is a snapshot of the early universe, so its observed isotropy is direct evidence that our working hypothesis of a homogeneous and isotropic universe is essentially valid as far back as the photon decoupling time. Nevertheless, this isotropy should not be perfect. The observed universe has all sorts of structure; some of the superclusters of galaxies and the largest voids are as large as 100 Mpc across. Such a basic feature of our universe must be reflected in the CMB in the form of small temperature anisotropies. There must have been some matter density nonuniformity at t_γ, which would have brought about temperature inhomogeneity; photons traveling from denser regions would be gravitationally redshifted and therefore arrive cooler, while photons from less-dense regions would do less work and arrive warmer. Such small temperature variations ($\simeq 30\,\mu$K) coming from different directions were finally detected, providing evidence for a primordial density nonuniformity that, under gravitational attraction, grew into the structures of stars, galaxies, and clusters of galaxies that we observe today.

Dark matter forms the cosmic scaffolding of matter distribution The temperature variation $\delta T/T = O(10^{-5})$ is smaller than expected based on the observed structure of baryonic matter. But this discrepancy can be resolved by

the existence of dark matter. Dark matter has no electromagnetic interaction; its density inhomogeneity does not directly produce temperature inhomogeneities in the cosmic fluid that can be seen in the CMB anisotropy. In fact, we expect that the gravitational clumping of the dark matter in forming structure took place first. The corresponding clumping of the baryonic matter (having electromagnetic interaction) was countered by radiation pressure until after photons were decoupled. Once the baryonic structure was formed, it tended to fall into the gravitational potential wells of the already-formed dark-matter scaffolding. This extra early growth of density perturbation for the nonbaryonic dark matter means that less baryonic inhomogeneity at t_γ is needed to produce the structure seen today. Thus a lesser $\delta T/T$ in the baryonic matter at t_γ (and hence in the CMB) is needed if the bulk of the matter has no electromagnetic interaction.

Cosmic inflation, primordial gravitational waves, and CMB polarization Where did the primordial density perturbations come from? As we shall discuss in Chapter 10, the favored theory of the big bang is that at the cosmic time $O(10^{-35}\,\text{s})$, the fluctuation of some (scalar) quantum field led to a state having a large cosmological constant, which drove an exponential expansion of the universe. This stretched the quantum fluctuations of the density to macroscopic sizes, seeding subsequent structure formation, and brought about the CMB temperature anisotropy discussed above. Furthermore, this anisotropic radiation would in turn lead, through Compton scattering, to CMB polarization (to be discussed in Box 10.3). Of particular interest is the theoretical prediction that such an inflationary epoch would generate tensor perturbations of the spacetime metric (i.e., gravitational waves), which would give rise to a unique pattern of polarization (called *B*-mode polarization). Its definitive detection would provide us with direct evidence of gravitational waves and of the inflationary theory of the big bang.

Box 9.3 The statistical study of CMB anisotropy

Cosmological theories can be checked by a statistical study of the CMB temperature $T(\theta,\phi)$ across the celestial sphere. With an average

$$\langle T\rangle = \frac{1}{4\pi}\int T(\theta,\phi)\sin\theta\, d\theta\, d\phi = 2.725\,\text{K}, \tag{9.42}$$

the temperature fluctuation

$$\frac{\delta T}{T}(\theta,\phi) \equiv \frac{T(\theta,\phi)-\langle T\rangle}{\langle T\rangle} \tag{9.43}$$

has a root-mean-square value of $\sqrt{\langle(\delta T/T)^2\rangle} = 1.1\times10^{-5}$. How do we describe such a temperature variation and connect it to the underlying cosmological theory? Recall that for a function of one variable, a useful approach

is Fourier expansion of the function in a series of sine waves with frequencies that are integral multiples of the fundamental frequency (of the wave with the largest wavelength). Similarly, we expand the temperature fluctuation in terms of spherical harmonics (think of them as vibration modes on the surface of an elastic sphere):

$$\frac{\delta T}{T}(\theta,\phi) = \sum_{l=0}^{\infty}\sum_{m=-l}^{l} a_{lm} Y_l^m(\theta,\phi). \tag{9.44}$$

The multipole number l represents the number of nodes (locations of zero amplitude) between equator and poles, while m is the longitudinal node number. For a given l, there are $2l+1$ values for m: $-l, -l+1, \ldots, l-1, l$. The series can of course be inverted so that the expansion coefficients a_{lm} are expressed in terms of the temperature fluctuation.

Cosmological theories predict statistical information about CMB temperature fluctuation. The most useful statistic is the 2-point correlation. Consider two points on a unit sphere at $\hat{\mathbf{n}}_1$ and $\hat{\mathbf{n}}_2$, separated by θ. We define the correlation function

$$C(\theta) \equiv \left\langle \frac{\delta T}{T}(\hat{\mathbf{n}}_1)\frac{\delta T}{T}(\hat{\mathbf{n}}_2)\right\rangle_{\hat{\mathbf{n}}_1\cdot\hat{\mathbf{n}}_2=\cos\theta}, \tag{9.45}$$

where the angle brackets denote the averaging over an ensemble of realizations of the fluctuation.[20] The inflationary cosmology predicts that the fluctuation is Gaussian[21] (i.e., maximally random) and is thus independent of the a_{lm}. Namely, the multipoles a_{lm} are uncorrelated. For different values of l and m,

$$\langle a_{lm}\rangle = 0, \qquad \langle a_{lm}^* a_{l'm'}\rangle = C_l \delta_{ll'}\delta_{mm'}, \tag{9.46}$$

which defines the power spectrum C_l as a measure of the relative strength of the spherical harmonics in the decomposition of the temperature fluctuations. Namely, it measures the typical size of the temperature irregularity on a given angular scale. The lack of m dependence reflects the (azimuthal) rotational symmetry of the underlying cosmological model. When we plug (9.44) into (9.45) with the conditions (9.46), the expansion is simplified[22] to

$$C(\theta) = \frac{1}{4\pi}\sum_{l=0}^{\infty}(2l+1)C_l P_l(\cos\theta), \tag{9.47}$$

where $P_l(\cos\theta)$ is the Legendre polynomial. Namely, the information carried by $C(\theta)$ in the angular space can be represented by C_l in the space of multipole number l. For the large multipole, the Legendre polynomial has the

continued

[20] In principle, this means averaging over many universes. Since we have only one universe, this ensemble averaging is carried out by averaging over multiple moments with different m, which in theory should be equal because of spherical symmetry. But for small l, there are fewer m available, which makes the average more uncertain. Figure 9.4 exhibits this cosmic variance for low l.

[21] If the temperature fluctuation were not Gaussian, higher-order correlations would contain additional information.

[22] We also need to use the addition theorem of spherical harmonics:

$$\sum_m Y_l^{*m}(\hat{\mathbf{n}}_1) Y_l^m(\hat{\mathbf{n}}_2) = \frac{2l+1}{4\pi} P_l(\cos\theta_{12}).$$

Box 9.3 *continued*

asymptotic form[23] $P_l(\cos\theta) \simeq \cos l\theta$. At peak locations in this regime, there is a correspondence between angular separation and multipole number as

$$l \simeq \frac{\pi}{\theta}. \tag{9.48}$$

Large l's correspond to small angular scales, with $l \simeq 10^2$ corresponding to degree-scale separation; cf. Box 10.2.

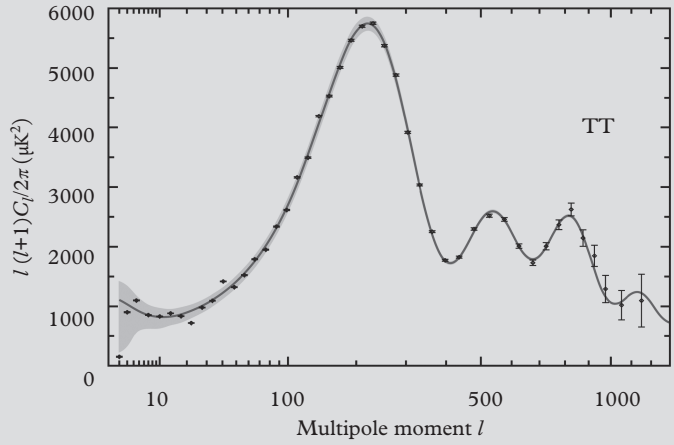

Figure 9.4 *The angular power spectrum of the CMB temperature anisotropy. The theoretical curve follows from the ΛCDM (dark energy Λ–cold dark matter) model. The first major peak at multipole number $l \simeq 200$ is evidence for a flat universe. The theoretical uncertainty for low multipoles is due to cosmic variance (cf. Sidenote 20). The data are from nine-year observations by the WMAP Collaboration, reproduced with permission from (Bennett et al. 2013). ©2013 American Astronomical Society.*

The power spectrum C_l can be measured by observations and compared with theoretical predictions. An example is displayed in Fig. 9.4. We note the prominent features:

(i) a fairly flat region of lower l values, the Sachs–Wolfe plateau, due to the variation of gravitational potential on the last-scattering surface;

(ii) a series of oscillations, the acoustic peaks, coming from the complex motions of the cosmic fluid when the CMB was first created;

(iii) the damping tail in the large-l region, reflecting the fact that the last-scattering surface has finite thickness.

Theoretical predictions are made by selecting cosmological parameters that yield the best match with the observed power spectrum. The cosmological parameters used in our presentation are mainly from the CMB power spectra obtained by the WMAP and Planck satellites.

[23] The exact limit form is

$$P_l(\cos\theta) = \sqrt{\frac{2}{\pi l \sin\theta}}\cos\left[\left(l+\frac{1}{2}\right)\theta - \frac{\pi}{4}\right].$$

In other words, when one looks at the anisotropy distribution shown in Fig. 9.3, one sees spots of many different sizes. This pattern can be translated into the power spectrum of Fig. 9.4, which encodes the key cosmological information that can be compared with theoretical predictions.

Review questions

1. Give an argument for the scaling behavior of the radiation temperature: $T \sim a^{-1}$. Show that under such a scaling law, the shape of the blackbody radiation spectrum is unchanged as the universe expands; i.e., a redshifted blackbody spectrum is simply a colder blackbody spectrum.
2. What is the condition (called the Gamow condition) for any particular set of interacting particles to be in thermal equilibrium during a given epoch of the expanding universe?
3. Cosmic helium synthesis combines two protons and two neutrons into a helium nucleus. The Boltzmann distribution at a thermal energy of the order of MeV yields a neutron-to-proton number density ratio $n_n/n_p \simeq 1/7$. From this, how would you estimate the cosmic helium mass fraction?
4. How can one use the theory of big bang nucleosynthesis and the observed abundance of light elements such as deuterium and helium to deduce the baryon number density Ω_B, and that the number of neutrino flavors should be three? (The reader is not asked why there are three neutrino flavors, but how the astrophysical observation is only compatible with three flavors of light neutrinos.)
5. What physical process took place around the photon-decoupling time t_γ? What are the average thermal energy and temperature at t_γ? Given the redshift $z_\gamma \simeq 10^3$, calculate the expected photon temperature now.
6. What was the cosmic time at which the universe made the transition from a radiation-dominated to a matter-dominated system. How does it compare with the nucleosynthesis and photon-decoupling times?
7. The abundance of cosmic background neutrinos should be comparable to the CMB. How come we have not detected them?
8. Why would the peculiar motion of our galaxy show up as a CMB dipole anisotropy?
9. Besides the dipole anisotropy, how does the CMB temperature anisotropy reflect the origin of cosmic structure?
10. Why would the presence of a significant amount of dark matter reduce the baryonic matter inhomogeneity (and hence the CMB temperature inhomogeneity) required to account for the observed structure in the universe?