3

BEYOND EQUILIBRIUM

The very early universe was hot and dense. As a result, interactions among particles occurred much more frequently than they do today. As an example, a photon today can travel across the observable universe without deflection or capture, so it has a mean free path greater than 10^{28} cm. When the age of the universe was equal to 1 sec, though, the mean free path of a photon was about the size of an atom. Thus in the time it took the universe to expand by a factor of 2, a given photon interacted many, many times. These multiple interactions kept the consitituents in the universe in equilibrium in most cases. Nonetheless, there were times when reactions could not proceed rapidly enough to maintain equilibrium conditions. These times are—perhaps not coincidentally—of the utmost interest to cosmologists today.

Indeed, we will see in this chapter that out-of-equilibrium phenomena played a role in (i) the formation of the light elements during Big Bang nucleosynthesis; (ii) recombination of electrons and protons into neutral hydrogen when the temperature was of order 1/4 eV; and quite possibly in (iii) production of dark matter in the early universe. Each of these three periods, for obvious reasons, is the subject of intense study by many different groups. Often these studies are carried out in parallel, without the link among the three mentioned. I think it is important to understand that all three phenomena are the result of nonequilibrium physics and that all three can be studied with the same formalism: the Boltzmann equation. Section 3.1 introduces the Boltzmann equation and some approximations to it that are common to all three processes. The remaining three sections of the chapter are simply applications of this general formula.

Beyond the intrinsic importance of these nonequilibrium phenomena, this chapter also serves as a bridge between the smooth, homogeneous universe described in Chapter 2 and the inhomogeneous perturbations we will explore in the rest of the book. One way to think of this transition is in terms of phase-space distributions f of Eqs. (2.60) and (2.61). Until now, we have assumed that the chemical potentials are zero and temperatures uniform. In this chapter, we will have to abandon the idea of trivial chemical potentials in order to track abundances of particles losing contact with the plasma. In succeeding chapters, we will move beyond uniformity and explore temperatures which depend on both position and direction of propagation.

3.1 BOLTZMANN EQUATION FOR ANNIHILATION

The Boltzmann equation formalizes the statement that the rate of change in the abundance of a given particle is the difference between the rates for producing and eliminating that species. Suppose that we are interested in the number density n_1 of species 1. For simplicity, let's suppose that the only process affecting the abundance of this species is an annihilation with species 2 producing two particles, imaginatively called 3 and 4. Schematically, $1+2 \leftrightarrow 3+4$; i.e., particle 1 and particle 2 can annihilate producing particles 3 and 4, or the inverse process can produce 1 and 2. The Boltzmann equation for this system in an expanding universe is

$$\begin{aligned} a^{-3}\frac{d\left(n_1 a^3\right)}{dt} &= \int \frac{d^3p_1}{(2\pi)^3 2E_1} \int \frac{d^3p_2}{(2\pi)^3 2E_2} \int \frac{d^3p_3}{(2\pi)^3 2E_3} \int \frac{d^3p_4}{(2\pi)^3 2E_4} \\ &\quad \times (2\pi)^4 \delta^3(p_1+p_2-p_3-p_4)\delta(E_1+E_2-E_3-E_4)\left|\mathcal{M}\right|^2 \\ &\quad \times \{f_3 f_4[1\pm f_1][1\pm f_2] - f_1 f_2[1\pm f_3][1\pm f_4]\}\,. \end{aligned} \tag{3.1}$$

In the absence of interactions, the left-hand side of Eq. (3.1) says that the density times the scale factor cubed is conserved. This reflects the nature of the expanding universe: as the comoving grid expands, the volume of a region containing a fixed number of particles grows as a^3. Therefore, the physical number density of these particles falls off as a^{-3}. Interactions are included in the right-hand side of the Boltzmann equation. Let's consider the interaction term starting from the last line and moving up. Putting aside the $1 \pm f$ terms on the last line, we see that the rate of producing species 1 is proportional to the occupation numbers of species 3 and 4, f_3 and f_4. Similarly the loss term is proportional to $f_1 f_2$. The $1 \pm f$ terms, with plus sign for bosons such as photons and minus sign for fermions such as electrons, represent the phenomena of Bose enhancement and Pauli blocking. If particles of type 1 already exist, a reaction producing more such particles is more likely to occur if 1 is a boson and less likely if a fermion. I have suppressed the momentum dependence of f, but of course all the occupation numbers depend on the corresponding momentum (e.g., $f_1 = f_1(p_1)$). Moving upward, the Dirac delta functions on the second line in Eq. (3.1) enforce energy and momentum conservation; the factors of 2π are the result of moving from discrete Kronecker δ's to the continuous Dirac version. The energies here are related to the momenta via $E = \sqrt{p^2+m^2}$. The amplitude on the second line $\mathcal{M}$ is determined from the fundamental physics in question. For example, if we were interested in the Compton scattering of electrons off of photons, $\mathcal{M}$ would be proportional to α, the fine structure constant. In almost all cases of interest, this amplitude is *reversible*, identical for $1+2 \to 3+4$ and $3+4 \to 1+2$. Indeed, reversibility has been assumed in Eq. (3.1).

The last two lines of Eq. (3.1) depend on the momenta of the particles involved. To find the total number of interactions, we must sum over all momenta. The integrals on the first line do precisely that. As in Figure 2.4, the factors of $(2\pi)^3$ [really $(2\pi\hbar)^3$] represent the volume of one unit of phase space; we want to sum over all such units. Finally, the factors of $2E$ in the denominator arise because, relativistically, the phase space integrals should really be four-dimensional, over the three components of momentum and one of energy. However, these are constrained to lie on the 3-sphere fixed by $E^2 = p^2 + m^2$. In equations,

$$\int d^3p \int_0^\infty dE\ \delta(E^2 - p^2 - m^2) = \int d^3p \int_0^\infty dE\ \frac{\delta\left(E - \sqrt{p^2+m^2}\right)}{2E}. \tag{3.2}$$

Performing the integral over E with the delta function yields the factor of $2E$.

Equation (3.1) is an integrodifferential equation for the phase space distributions. Further, in principle at least, it must be supplemented with similar equations for the other species. In practice, these formidable obstacles can be overcome for many practical cosmological applications. The first, most important realization is that scattering processes typically enforce *kinetic equilibrium*. That is, scattering takes place so rapidly that the distributions of the various species take on the generic Bose–Einstein/Fermi–Dirac forms (Eqs. (2.61) and (2.60)). This form condenses all of the uncertainty in the distribution into a single function of time μ. If annihilations were also in equilibrium, μ would be the chemical potential, and the sum of the chemical potentials in any reaction would have to balance. For example, the reaction $e^+ + e^- \leftrightarrow \gamma + \gamma$ would cause $\mu_{e^+} + \mu_{e^-} = 2\mu_\gamma$. In the out-of-equilibrium cases we will study, the system will not be in *chemical equilibrium* and we will have to solve a differential equation for μ. The great simplifying feature of kinetic equilibrium, though, is that this differential equation will be a single ordinary differential equation, as opposed to the very complicated form of Eq. (3.1).

We will typically be interested in systems at temperatures smaller than $E - \mu$. In this limit, the exponential in the Bose–Einstein or Fermi–Dirac distribution is large and dwarfs the ± 1 in the denominator. Thus, another simplification emerges: we can ignore the complications of quantum statistics. The distributions become

$$f(E) \to e^{\mu/T} e^{-E/T} \tag{3.3}$$

and the Pauli blocking/Bose enhancement factors in the Boltzmann equation can be neglected.

Under these approximations, the last line of Eq. (3.1) becomes

$$\begin{aligned} f_3 f_4[1 \pm f_1][1 \pm f_2] &- f_1 f_2[1 \pm f_3][1 \pm f_4] \\ &\to \quad e^{-(E_1+E_2)/T}\left\{e^{(\mu_3+\mu_4)/T} - e^{(\mu_1+\mu_2)/T}\right\}. \end{aligned} \tag{3.4}$$

Here I have used energy conservation, $E_1 + E_2 = E_3 + E_4$. We will use the number densities themselves as the time-dependent functions to be solved for, instead of μ. The number density of species i is related to μ_i via

Table 3.1. Reactions in This Chapter: $1 + 2 \leftrightarrow 3 + 4$

	1	2	3	4
Neutron-Proton Ratio	n	ν_e or e^+	p	e^- or $\bar{\nu}_e$
Recombination	e	p	H	γ
Dark Matter Production	X	X	l	l

$$n_i = g_i e^{\mu_i/T} \int \frac{d^3p}{(2\pi)^3} e^{-E_i/T} \tag{3.5}$$

where g_i is the degeneracy of the species, e.g., equal to 2 for the two spin states of the photon. It will also be useful to define the species-dependent equilibrium number density as

$$n_i^{(0)} \equiv g_i \int \frac{d^3p}{(2\pi)^3} e^{-E_i/T} = \begin{cases} g_i \left(\frac{m_i T}{2\pi}\right)^{3/2} e^{-m_i/T} & m_i \gg T \\ g_i \frac{T^3}{\pi^2} & m_i \ll T \end{cases} . \tag{3.6}$$

With this defintion, $e^{\mu_i/T}$ can be rewritten as $n_i/n_i^{(0)}$, so the last line of Eq. (3.1) is equal to

$$e^{-(E_1+E_2)/T} \left\{ \frac{n_3 n_4}{n_3^{(0)} n_4^{(0)}} - \frac{n_1 n_2}{n_1^{(0)} n_2^{(0)}} \right\} . \tag{3.7}$$

With these approximations the Boltzmann equation now simplifies enormously. Define the thermally averaged cross section as

$$\langle \sigma v \rangle \equiv \frac{1}{n_1^{(0)} n_2^{(0)}} \int \frac{d^3p_1}{(2\pi)^3 2E_1} \int \frac{d^3p_2}{(2\pi)^3 2E_2} \int \frac{d^3p_3}{(2\pi)^3 2E_3} \int \frac{d^3p_4}{(2\pi)^3 2E_4} e^{-(E_1+E_2)/T}$$
$$\times (2\pi)^4 \delta^3(p_1 + p_2 - p_3 - p_4) \delta(E_1 + E_2 - E_3 - E_4) \left|\mathcal{M}\right|^2 . \tag{3.8}$$

Then, the Boltzmann equation becomes

$$a^{-3} \frac{d\left(n_1 a^3\right)}{dt} = n_1^{(0)} n_2^{(0)} \langle \sigma v \rangle \left\{ \frac{n_3 n_4}{n_3^{(0)} n_4^{(0)}} - \frac{n_1 n_2}{n_1^{(0)} n_2^{(0)}} \right\} . \tag{3.9}$$

We thus have a simple ordinary differential equation for the number density. Although the details will vary from application to application (see Table 3.1), we will in the remainder of the chapter start from this equation when tracking abundances.

One qualitative note about Eq. (3.9). The left-hand side is of order n_1/t, or, since the typical cosmological time is H^{-1}, $n_1 H$. The right-hand side is of order $n_1 n_2 \langle \sigma v \rangle$. Therefore, if the reaction rate $n_2 \langle \sigma v \rangle$ is much larger than the expansion rate, then the terms on the right side will be much larger than the one on the left.

The only way to maintain equality then is for the individual terms on the right to cancel. Thus, when reaction rates are large,

$$\frac{n_3 n_4}{n_3^{(0)} n_4^{(0)}} = \frac{n_1 n_2}{n_1^{(0)} n_2^{(0)}}. \tag{3.10}$$

This equation, which follows virtually by inspection from the Boltzmann equation, goes under different names in different venues. The particle physics community, which first studied the production of heavy relics in the early universe, tends to call it *chemical equilibrium*. In the context of Big Bang nucleosynthesis, it is called *nuclear statistical equilibrium* (NSE), while students of recombination, the process of electrons and protons combining to form neutral hydrogen, use the terminology *Saha equation*.

3.2 BIG BANG NUCLEOSYNTHESIS

As the temperature of the universe cools to 1 MeV, the cosmic plasma consists of:

- **Relativistic particles in equilibrium: photons, electrons and positrons.** These are kept in close contact with each other by electromagnetic interactions such as $e^+e^- \leftrightarrow \gamma\gamma$. Besides a small difference due to fermion/boson statistics, these all have the same abundances.
- **Decoupled relativistic particles: neutrinos.** At temperatures a little above 1 MeV, the rate for processes such as $\nu e \leftrightarrow \nu e$ which keep neutrinos coupled to the rest of the plasma drops beneath the expansion rate. Neutrinos therefore share the same temperature as the other relativistic particles, and hence are roughly as abundant, but they do not couple to them.
- **Nonrelativistic particles: baryons.** If there had been no asymmetry in the initial number of baryons and anti-baryons, then both would be completely depleted by 1 MeV. However, such an asymmetry did exist: $(n_b - n_{\bar{b}})/s \sim 10^{-10}$ initially,[1] and this ratio remains constant throughout the expansion. By the time the temperature is of order 1 MeV, all anti-baryons have annihilated away (Exercise 12) so
$$\eta_b \equiv \frac{n_b}{n_\gamma} = 5.5 \times 10^{-10} \left(\frac{\Omega_b h^2}{0.020}\right). \tag{3.11}$$
There are thus many fewer baryons than relativistic particles when $T \sim$ MeV.

Our task in this section will be to determine how the baryons end up. Were the system to remain in equilibrium throughout, the final state would be dictated solely by energetics, and all baryons would relax to the nuclear state with the lowest energy per baryon, iron (Figure 3.1). However, nuclear reactions, which scale as the second—or higher—power of the density, are too slow to keep the system in equilibrium as the temperature drops. So, in principle, we need to solve the

[1] s is the entropy density which scales as a^{-3}, as we saw in Chapter 2.

equivalent of Eq. (3.9) for all the nuclei, i.e., a set of coupled differential equations. In practice, at least for a qualitative understanding of the result, we can make use of two simplifications that obviate the need to solve the full set of differential equations.

Lightning Introduction to Nuclear Physics

A single proton is a hydrogen nucleus, referred to as ^{1}H or simply p; a proton and a neutron make up deuterium, ^{2}H or D; one proton and two neutrons make tritium, ^{3}H or T. Nuclei with two protons are helium; these can have one neutron (^{3}He) or two (^{4}He). Thus unique elements have a fixed number of protons, and isotopes of a given element have differing numbers of neutrons. The total number of neutrons and protons in the nucleus, the *atomic number*, is a superscript before the name of the element.

The total mass of a nucleus with Z protons and $A - Z$ neutrons differs slightly from the mass of the individual protons and neutrons alone. This difference is called the binding energy, defined as

$$B \equiv Zm_p + (A - Z)m_n - m \tag{3.12}$$

where m is the mass of the nucleus. For example, the mass of deuterium is 1875.62 MeV while the sum of the neutron and proton masses is 1877.84 MeV, so the binding energy of deuterium is 2.22 MeV. Nuclear binding energies are typically in the MeV range, which explains why Big Bang nucleosynthesis occurs at temperatures a bit less than 1 MeV even though nuclear masses are in the GeV range.

Neutrons and protons can interconvert via weak interactions:

$$p + \bar{\nu} \leftrightarrow n + e^+ \quad ; \quad p + e^- \leftrightarrow n + \nu \quad ; \quad n \leftrightarrow p + e^- + \bar{\nu} \tag{3.13}$$

where all the reactions can proceed in either direction. The light elements are built up via electromagnetic interactions. For example, deuterium forms from $p+n \to \mathrm{D}+\gamma$. Then, $\mathrm{D}+\mathrm{D} \to n+{}^3\mathrm{He}$, after which ${}^3\mathrm{He}+\mathrm{D} \to p+{}^4\mathrm{He}$ produces ^{4}He.

The first simplification is that essentially no elements heavier than helium are produced at appreciable levels.[2] So the only nuclei that need to be traced are hydrogen and helium, and their isotopes: deuterium, tritium, and ^{3}He. The second simplification is that, even in the context of this reduced set of elements, the physics splits up neatly into two parts since above $T \simeq 0.1$ MeV, no light nuclei form: only free protons and neutrons exist. Therefore, we first solve for the neutron/proton ratio and then use this abundance as input for the synthesis of helium and isotopes such as deuterium.

[2] An exception is lithium, produced at a part in 10^9–10^{10}, and this trace abundance may be observable today. See, e.g., Pinsonneault *et al.* (2001).

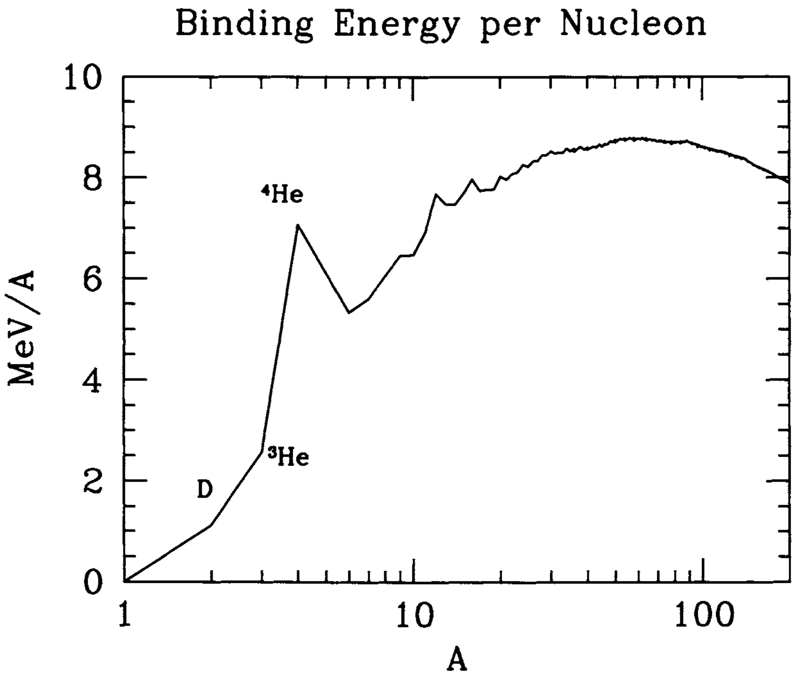

Figure 3.1. Binding energy of nuclei as a function of mass number. Iron has the highest binding energy, but among the light elements, ^{4}He is a crucial local maximum. Nucleosynthesis in the early universe essentially stops at ^{4}He because of the lack of tightly bound isotopes at $A = 5 - 8$. In the high-density environment of stars, three ^{4}He nuclei fuse to form ^{12}C, but the low baryon number precludes this process in the early universe.

Both of these simplifications — no heavy elements at all and only n/p above 0.01 MeV — rely on the physical fact that, at high temperatures, comparable to nuclear binding energies, any time a nucleus is produced in a reaction, it is destroyed by a high-energy photon. This fact is reflected in the fundamental equilibrium equation (3.10). To see how, let's consider this equation applied to deuterium production, $n + p \leftrightarrow \mathrm{D} + \gamma$. Since photons have $n_\gamma = n_\gamma^{(0)}$, the equilibrium condition becomes

$$\frac{n_D}{n_n n_p} = \frac{n_D^{(0)}}{n_n^{(0)} n_p^{(0)}}. \tag{3.14}$$

The integrals on the right, as given in Eq. (3.6), lead to

$$\frac{n_D}{n_n n_p} = \frac{3}{4} \left(\frac{2\pi m_D}{m_n m_p T} \right)^{3/2} e^{[m_n + m_p - m_D]/T}, \tag{3.15}$$

the factor of 3/4 being due to the number of spin states (3 for D and 2 each for p and n). In the prefactor, m_D can be set to $2m_n = 2m_p$, but in the exponential the small difference between $m_n + m_p$ and m_D is important: indeed the argument of the exponential is by defintion equal to the binding energy of deuterium, $B_D = 2.22$ MeV. Therefore, as long as equilibrium holds,

$$\frac{n_D}{n_n n_p} = \frac{3}{4}\left(\frac{4\pi}{m_p T}\right)^{3/2} e^{B_D/T}. \tag{3.16}$$

Both the neutron and proton density are proportional to the baryon density, so roughly,

$$\frac{n_D}{n_b} \sim \eta_b \left(\frac{T}{m_p}\right)^{3/2} e^{B_D/T}. \tag{3.17}$$

As long as B_D/T is not too large, the prefactor dominates this expression. And the prefactor is very small because of the smallness of the baryon-to-photon ratio, Eq. (3.11).

The small baryon to photon ratio thus inhibits nuclei production until the temperature drops well beneath the nuclear binding energy. At temperatures above 0.1 MeV, then, virtually all baryons are in the form of neutrons and protons. Around this time, deuterium and helium are produced, but the reaction rates are by now too low to produce any heavier elements. We could have anticipated this by considering Figure 3.1. The lack of a stable isotope with mass number 5 implies that heavier elements cannot be produced via $^4\text{He}+p \to X$. In stars, the triple alpha process $^4\text{He}+^4\text{He}+^4\text{He}\to {}^{12}\text{C}$ produces heavier elements, but in the early universe, densities are far too low to allow three nuclei to find one another on relevant time scales.

3.2.1 Neutron Abundance

We begin by solving for the neutron–proton ratio. Protons can be converted into neutrons via weak interactions, $p+e^- \to n+\nu_e$ for example. As we will see, reactions of this sort keep neutrons and protons in equilibrium until $T \sim$ MeV. Thereafter, one must solve the rate equation (3.9) to track the neutron abundance.

From Eq. (3.6), the proton/neutron equilibrium ratio in the nonrelativistic limit ($E = m + p^2/2m$) is

$$\frac{n_p^{(0)}}{n_n^{(0)}} = \frac{e^{-m_p/T}\int dp\, p^2\, e^{-p^2/2m_pT}}{e^{-m_n/T}\int dp\, p^2\, e^{-p^2/2m_nT}}. \tag{3.18}$$

The integrals here are proportional to $m^{3/2}$, but the resulting ratio $(m_p/m_n)^{3/2}$ is sufficiently close to unity that we can neglect the mass difference. However, in the exponential the mass difference is very important, and we are left with

$$\frac{n_p^{(0)}}{n_n^{(0)}} = e^{\mathcal{Q}/T} \tag{3.19}$$

with $\mathcal{Q} \equiv m_n - m_p = 1.293$ MeV. Therefore, at high temperatures, there are as many neutrons as protons. As the temperature drops beneath 1 MeV, the neutron fraction goes down. If weak interactions operated efficiently enough to maintain equilibrium indefinitely, then it would drop to zero. The main task of this section is to find out what happens in the real world where weak interactions are not so efficient.

It is convenient to define

$$X_n \equiv \frac{n_n}{n_n + n_p}, \tag{3.20}$$

that is, X_n is the ratio of neutrons to total nuclei. In equilibrium,

$$X_n \to X_{n,\mathrm{EQ}} \equiv \frac{1}{1 + \left(n_p^{(0)}/n_n^{(0)}\right)}. \tag{3.21}$$

To track the evolution of X_n, let's start from Eq. (3.9), with 1 = neutron, 3 = proton, and 2, 4 = leptons in complete equilibrium ($n_l = n_l^{(0)}$). Then,

$$a^{-3}\frac{d\left(n_n a^3\right)}{dt} = n_l^{(0)}\langle\sigma v\rangle \left\{ \frac{n_p n_n^{(0)}}{n_p^{(0)}} - n_n \right\}. \tag{3.22}$$

We have already determined the ratio $n_n^{(0)}/n_p^{(0)} = e^{-Q/T}$ and we can identify $n_l^{(0)}\langle\sigma v\rangle$ as λ_{np}, the rate for neutron $\to$ proton conversion since it multiplies n_n in the loss term. Also if we rewrite n_n on the left as $(n_n + n_p)X_n$, then the total density times a^3 can be taken outside the derivative, leaving

$$\frac{dX_n}{dt} = \lambda_{np}\left\{(1 - X_n)e^{-Q/T} - X_n\right\}. \tag{3.23}$$

Equation (3.23) is a differential equation for X_n as a function of time, but it contains the temperature T and the reaction rate λ_{np}, both of which have complicated time dependences. It is simplest therefore to recast the equation using as the evolution variable

$$x \equiv \frac{Q}{T}. \tag{3.24}$$

The left-hand side of Eq. (3.23) then becomes $\dot{x}\, dX_n/dx$, so we need an expression for $dx/dt = -x\dot{T}/T$. Since $T \propto a^{-1}$,

$$\frac{1}{T}\frac{dT}{dt} = -H = -\sqrt{\frac{8\pi G\rho}{3}}, \tag{3.25}$$

the second equality following from Eq. (2.39). Nucleosynthesis occurs in the radiation-dominated era, so the main contribution to the energy density ρ comes from relativistic particles. Recall from Chapter 2 that the contribution to the energy density from relativistic particles is

$$\rho = \frac{\pi^2}{30}T^4 \left[\sum_{i=\mathrm{bosons}} g_i + \frac{7}{8} \sum_{i=\mathrm{fermions}} g_i \right] \qquad (i \text{ relativistic})$$

$$\equiv g_* \frac{\pi^2}{30} T^4. \tag{3.26}$$

The effective numbers of relativistic degrees of freedom, g_*, is a function of temperature. At temperatures of order 1 MeV, the contributing species are: photons ($g_\gamma = 2$), neutrinos ($g_\nu = 6$), and electrons and positrons ($g_{e^+} = g_{e^-} = 2$). Adding up leads to $g_* \simeq 10.75$, roughly constant throughout the regime of interest. Then, Eq. (3.23) becomes

$$\frac{dX_n}{dx} = \frac{x\lambda_{np}}{H(x=1)} \left\{ e^{-x} - X_n \left(1 + e^{-x} \right) \right\} \tag{3.27}$$

with

$$H(x=1) = \sqrt{\frac{4\pi^3 G Q^4}{45}} \times \sqrt{10.75} = 1.13 \text{ sec}^{-1}. \tag{3.28}$$

Finally, we need an expression for the neutron–proton conversion rate, λ_{np}. Under the approximations we are using, the rate is (Bernstein 1988 or Exercise 3)

$$\lambda_{np} = \frac{255}{\tau_n x^5} \left(12 + 6x + x^2 \right), \tag{3.29}$$

with the neutron lifetime $\tau_n = 886.7$ sec. Thus, when $T = Q$ (i.e., when $x = 1$), the conversion rate is 5.5 sec^{-1}, somewhat larger than the expansion rate. As the temperature drops beneath 1 MeV, though, the rate rapidly falls below the expansion rate, so conversions become inefficient.

We can now integrate Eq. (3.27) numerically to track the neutron abundance (Exercise 4). Figure 3.2 shows the results of this integration. Note that the result agrees extremely well at temperatures above ~ 0.1 MeV with the exact solution which includes proper statistics, nonzero electron mass, and changing g_*. The neutron fraction X_n does indeed fall out of equlibrium once the temperature drops below 1 MeV: it freezes out at 0.15 once the temperature drops below 0.5 MeV.

At temperatures below 0.1 MeV, two reactions we have not included yet become important: neutron decay ($n \to p + e^- + \bar{\nu}$) and deuterium production ($n + p \to \text{D} + \gamma$). Decays can be added trivially by adding in a factor of e^{-t/τ_n} to the results of Figure 3.2. By the time decays become important, electrons and positrons have annihilated, so g_* in Eq. (3.26) is 3.36 and the time–temperature relation is (Exercise 5):

$$t = 132 \,\text{sec} \left(\frac{0.1\text{MeV}}{T} \right)^2. \tag{3.30}$$

We will see shortly that production of deuterium, and other light elements, begins in earnest at $T \sim 0.07$ MeV. By then, decays have depleted the neutron fraction by a factor of $\exp[-(132/886.7)(0.1/0.07)^2] = 0.74$. So the neutron abundance at the onset of nucleosynthesis is 0.15×0.74, or

$$X_n(T_{\text{nuc}}) = 0.11. \tag{3.31}$$

We now turn to light element formation to understand the ramifications of this number.

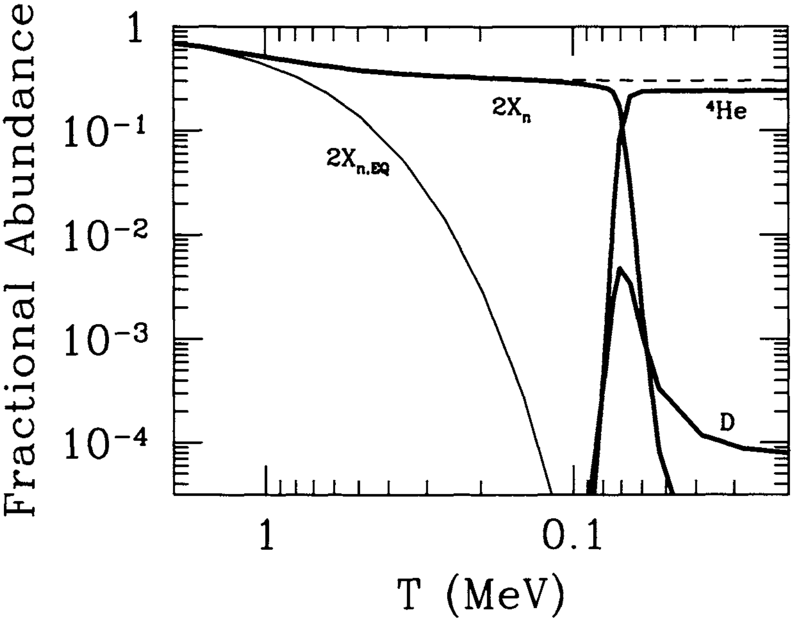

Figure 3.2. Evolution of light element abundances in the early universe. Heavy solid curves are results from Wagoner (1973) code; dashed curve is from integration of Eq. (3.27); light solid curve is twice the neutron equilibrium abundance. Note the good agreement of Eq. (3.27) and the exact result until the onset of neutron decay. Also note that the neutron abundance falls out of equilibrium at $T \sim$MeV.

3.2.2 Light Element Abundances

A useful way to approximate light element production is that it occurs instantaneously at a temperature $T_{\rm nuc}$ when the energetics compensates for the small baryon to photon ratio. Let's consider deuterium production as an example, with Eq. (3.17) as our guide. The equilibrium deuterium abundance is of order the baryon abundance (i.e. if the universe stayed in equilibrium, all neutrons and protons would form deuterium) when Eq. (3.17) is of order unity, or

$$\ln(\eta_b) + \frac{3}{2}\ln(T_{\rm nuc}/m_p) \sim -\frac{B_D}{T_{\rm nuc}}. \tag{3.32}$$

Equation (3.32) suggests that deuterium production takes place at $T_{\rm nuc} \sim 0.07$ MeV, with a weak logarithmic dependence on η_b.

Since the binding energy of helium is larger than that of deuterium, the exponential factor $e^{B/T}$ favors helium over deuterium. Indeed, Figure 3.2 illustrates that helium is produced almost immediately after deuterium. Virtually all remaining neutrons at $T \sim T_{\rm nuc}$ then are processed into ^{4}He. Since two neutrons go into ^{4}He, the final ^{4}He abundance is equal to half the neutron abundance at $T_{\rm nuc}$. Often, results are quoted in terms of mass fraction; then,

$$X_4 \equiv \frac{4n_{^4He}}{n_b} = 2X_n(T_{\rm nuc}). \tag{3.33}$$

Figure 3.2 shows that this relation holds. Indeed, to find the final helium mass fraction, we need only take the neutron fraction at $T_{\rm nuc}$, Eq. (3.31), and multiply by 2, so the final helium mass fraction is 0.22. This rough estimate, obtained by solving a single differential equation, is in remarkable agreement with the exact solution, which can be fit via (Olive, 2000; Kolb and Turner, 1990)

$$Y_p = 0.2262 + 0.0135 \ln(\eta_b/10^{-10}). \tag{3.34}$$

One important feature of this result is that it depends only logarithmically on the baryon fraction. We saw in Eq. (3.32) that $T_{\rm nuc}$ has this dependence. You might think that the exponential sensitivity to $T_{\rm nuc}$ in the decay fraction would turn this into linear dependence. However, $T_{\rm nuc}$ is sufficiently early that only a small fraction of neutrons have decayed: the exponential in this regime is linear in the time. Therefore, the final helium abundance maintains only logarithmic dependence on the baryon density.

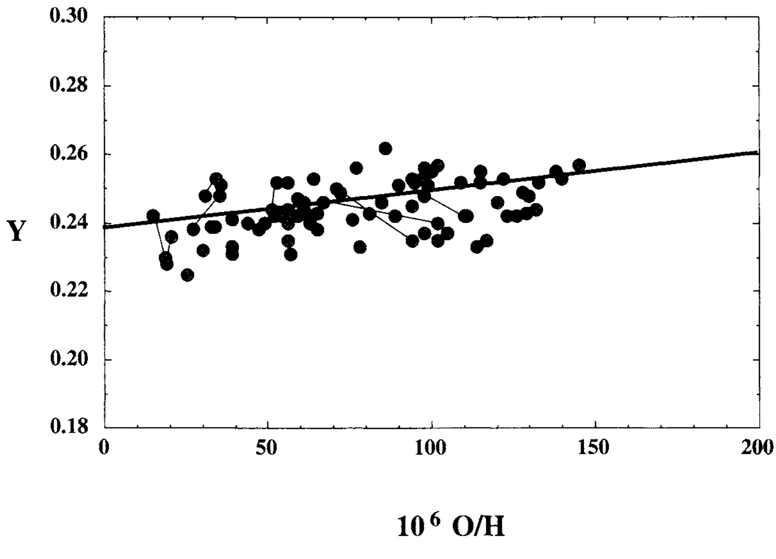

Figure 3.3. Helium abundance ($Y \equiv Y_p$) as a function of oxygen/hydrogen ratio. Lower oxygen systems have undergone less processing, so the helium abundance in those systems is closer to primordial. Line, and extrapolation to $Y_p = 0.238$, from Olive (2000). Data from Pagel *et al.* (1992), Skillman and Kennicutt (1993), Skillman *et al.* (1994), and Izotov and Thuan (1998). Short lines connect the same region observed by different groups.

The prediction agrees well with the observations, as indicated in Figure 3.3. The best indication of the primordial helium abundance comes from the most unprocessed systems, typically identified by low metallicities. As Figure 3.3 indicates, the primordial abundance almost certainly lies between 0.22 and 0.25. Although there have been claims of discord in the past, the agreement remains one of the pillars of observational cosmology.

Figure 3.2 shows that not all of the deuterium gets processed into helium. A trace amount remains unburned, simply because the reaction which eliminates it (D $+p \rightarrow {}^3$He $+ \gamma$) is not completely efficient. Figure 3.2 shows that after $T_{\rm nuc}$, deuterium is depleted via these reactions, eventually freezing out at a level of order 10^{-5}–10^{-4}. If the baryon density is low, then the reactions proceed more slowly, and the depletion is not as effective. Therefore, low baryon density inevitably results in more deuterium; the sensitivity is quite stark, as illustrated in Figure 1.8. As a result, deuterium is a powerful probe of the baryon density. Complementing this sensitivity is the possiblity of measuring deuterium in gas clouds as $z \sim 3$ by looking for absorbtion in the spectra of distant QSOs. For example, O'Meara *et al.* (2001) combine the measurements of primordial deuterium in four systems to obtain, D/H$= 3.0 \pm 0.4 \times 10^{-5}$, corresponding to $\Omega_b h^2 = 0.0205 \pm 0.0018$.

3.3 RECOMBINATION

As the temperature drops to ~ 1 eV, photons remain tightly coupled to electrons via Compton scattering and electrons to protons via Coulomb scattering. It will come as no surprise that at these temperatures, there is very little neutral hydrogen. Energetics of course favors the production of neutral hydrogen with a binding energy of $\epsilon_0 = 13.6$ eV, but the high photon/baryon ratio ensures that any hydrogen atom produced will be instantaneously ionized. This phenomenon is identical to the delay in the production of light nuclei we saw above, replayed on the atomic scale.

As long as the reaction[3] $e^- + p \leftrightarrow \mathrm{H} + \gamma$ remains in equilibrium, the condition in Eq. (3.10) (with $1 = e, 2 = p, 3 =$H) ensures that

$$\frac{n_e n_p}{n_H} = \frac{n_e^{(0)} n_p^{(0)}}{n_H^{(0)}}. \tag{3.35}$$

We can go further here by recognizing that the neutrality of the universe ensures that $n_e = n_p$. Let's define the free electron fraction

$$X_e \equiv \frac{n_e}{n_e + n_H} = \frac{n_p}{n_p + n_H}, \tag{3.36}$$

the denominator equal to the total number of hydrogen nuclei. Carrying out the integrals on the right of Eq. (3.35) leads then to

$$\frac{X_e^2}{1 - X_e} = \frac{1}{n_e + n_H} \left[\left(\frac{m_e T}{2\pi} \right)^{3/2} e^{-[m_e + m_p - m_H]/T} \right] \tag{3.37}$$

where we have made the familiar approximation of neglecting the small mass difference of H and p in the prefactor. The argument of the exponential is $-\epsilon_0/T$.

[3]Here p stands for free protons and H for neutral hydrogen, i.e., a proton with an electron attached.

Neglecting the relatively small number of helium atoms, the denominator $n_e + n_H$ (or $n_p + n_H$) is equal to the baryon density, $\eta_b n_\gamma \sim 10^{-9}T^3$. So when the temperature is of order ϵ_0, the right-hand side is of order $10^9(m_e/T)^{3/2} \simeq 10^{15}$. In that case, Eq. (3.37) can be satisfied only if the denominator on the left is very small, that is if X_e is very close to 1: all hydrogen is ionized. Only when the temperature drops far below ϵ_0 does appreciable recombination take place. As X_e falls, the rate for recombination also falls, so that equilibrium becomes more difficult to maintain. Thus, in order to follow the free electron fraction accurately, we need to solve the Boltzmann equation, just as we did for the neutron–proton ratio.

In this case, Eq. (3.9) for the electron density becomes

$$\begin{aligned} a^{-3}\frac{d\left(n_e a^3\right)}{dt} &= n_e^{(0)} n_p^{(0)} \langle\sigma v\rangle \left\{ \frac{n_H}{n_H^{(0)}} - \frac{n_e^2}{n_e^{(0)} n_p^{(0)}} \right\} \\ &= n_b \langle\sigma v\rangle \left\{ (1 - X_e)\left(\frac{m_e T}{2\pi}\right)^{3/2} e^{-\epsilon_0/T} - X_e^2 n_b \right\} \end{aligned} \tag{3.38}$$

where the last line follows since the ratio $n_e^{(0)} n_p^{(0)}/n_H^{(0)}$ is equal to the term in square brackets in Eq. (3.37). Meanwhile, since $n_b a^3$ is constant it can be passed through the derivative on the left after expressing n_e as $n_b X_e$, so that

$$\frac{dX_e}{dt} = \left\{ (1 - X_e)\beta - X_e^2 n_b \alpha^{(2)} \right\} \tag{3.39}$$

where the ionization rate is typically denoted

$$\beta \equiv \langle\sigma v\rangle \left(\frac{m_e T}{2\pi}\right)^{3/2} e^{-\epsilon_0/T} \tag{3.40}$$

and the recombination rate

$$\alpha^{(2)} \equiv \langle\sigma v\rangle. \tag{3.41}$$

The recombination rate has superscript $^{(2)}$ because recombination to the ground ($n = 1$) state is not relevant. Ground-state recombinations lead to production of an ionizing photon, and this photon immediately ionizes a neutral atom. The net effect of such a recombination is zero: no new neutral atoms are formed this way. The only way for recombination to proceed is via capture to one of the excited states of hydrogen; to a good approximation, this rate is

$$\alpha^{(2)} = 9.78 \frac{\alpha^2}{m_e^2} \left(\frac{\epsilon_0}{T}\right)^{1/2} \ln\left(\frac{\epsilon_0}{T}\right). \tag{3.42}$$

The Saha approximation, Eq. (3.37), does a good job predicting the redshift of recombination, but fails as the electron fraction drops and the system goes out of equilibirium. Therefore, the detailed evolution of X_e must be obtained by a numerical integration of Eq. (3.39) (Exercise 8). Results from numerical integration of Eq. (3.39) are shown in Figure 3.4.

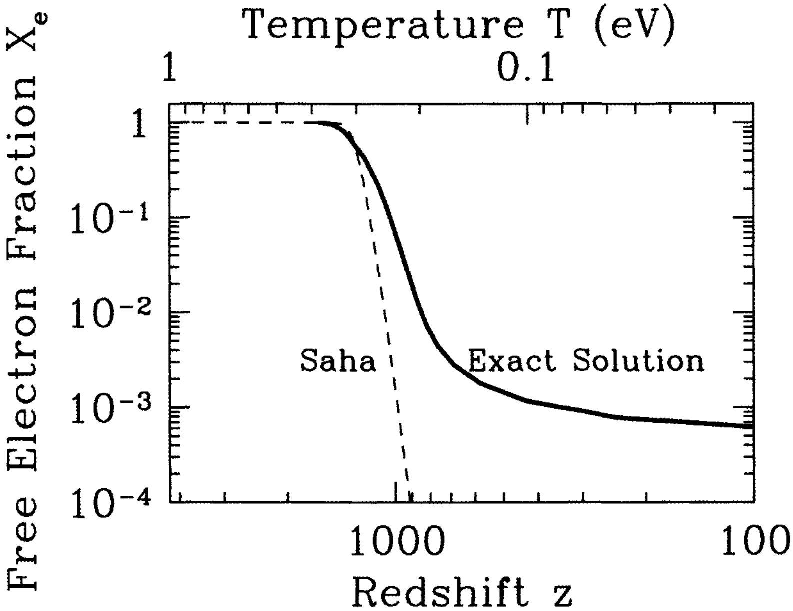

Figure 3.4. Free electron fraction as a function of redshift. Recombination takes place suddenly at $z \sim 1000$ corresponding to $T \sim 1/4$ eV. The Saha approximation, Eq. (3.37), holds in equilibrium and correctly identifies the redshift of recombination, but not the detailed evolution of X_e. Here $\Omega_b = 0.06, \Omega_m = 1, h = 0.5$.

The computation of the neutron/proton ratio affects the abundance of light elements today. Similarly, the evolution of the free electron abundance has major ramifications for observational cosmology. Recombination at $z_* \sim 1000$ is directly tied to the *decoupling* of photons from matter.[4] This decoupling, in turn, directly affects the pattern of anisotropies in the CMB that we observe today.

Decoupling occurs roughly when the rate for photons to Compton scatter off electrons becomes smaller than the expansion rate.[5] The scattering rate is

$$n_e \sigma_T = X_e n_b \sigma_T \tag{3.43}$$

where $\sigma_T = 0.665 \times 10^{-24}$ cm^2 is the Thomson cross section, and I continue to ignore helium, thereby assuming that the total number of hydrogen nuclei (free protons + hydrogen atoms) is equal to the total baryon number. Since the ratio of the

[4] Notice from Figure 1.4 that even though photons stop scattering off electrons at $z \sim 1000$, electrons do scatter many times off photons until much later. This is not a contradiction: there are many more photons than baryons. In any event, many cosmologists shy away from the word *decoupling* to describe what happens at $z \sim 1000$ for this reason.

[5] In Chapter 8 we will define a more precise measure of decoupling, making use of the *visibility function*, the probability that a photon last scattered at a given redshift. Using the visibility function, we will show that a CMB photon today most likely last scattered at a slightly higher redshift than inferred by the simple $n_e \sigma_T = H$ criterion.

baryon density to the critical density is $m_p n_b/\rho_{cr} = \Omega_b a^{-3}$, n_b can be eliminated in Eq. (3.43) in favor of Ω_b:

$$n_e \sigma_T = 7.477 \times 10^{-30} \text{cm}^{-1} X_e \Omega_b h^2 a^{-3}. \tag{3.44}$$

Dividing by the expansion rate leads to

$$\frac{n_e \sigma_T}{H} = 0.0692 a^{-3} X_e \Omega_b h \frac{H_0}{H}. \tag{3.45}$$

The ratio on the right depends on the Hubble rate, which is given in Eq. (1.2). From that equation or from Figure 1.3, we see that at early times, the main contribution comes from either matter or radiation, so $H/H_0 = \Omega_m^{1/2} a^{-3/2}[1 + a_{\text{eq}}/a]^{1/2}$. Then,

$$\frac{n_e \sigma_T}{H} = 113 X_e \left(\frac{\Omega_b h^2}{0.02}\right) \left(\frac{.15}{\Omega_m h^2}\right)^{1/2} \left(\frac{1+z}{1000}\right)^{3/2} \left[1 + \frac{1+z}{3600}\frac{0.15}{\Omega_m h^2}\right]^{-1/2}. \tag{3.46}$$

Here I have normalized with "best fit" values for the baryon and matter densities. When the free electron fraction X_e drops below $\sim 10^{-2}$, photons decouple. From Figure 3.4, we see that X_e drops very quickly from unity to 10^{-3}. Therefore, decoupling takes place during recombination.

Let's forget all we just learned and ask what would happen if the universe remained ionzied throughout its history. In that hypothetical case, $X_e = 1$, and Eq. (3.46) can be trivially solved to find the redshift of decoupling. Setting the right hand side to 1 leads to

$$1 + z_{\text{decouple}} = 43 \left(\frac{0.02}{\Omega_b h^2}\right)^{2/3} \left(\frac{\Omega_m h^2}{0.15}\right)^{1/3} \qquad \text{(no recombination)}. \tag{3.47}$$

Equation (3.47) tells us that even if the electrons remained ionized throughout the history of the universe, eventually the photons decoupled simply because expansion made it more difficult to find the increasingly dilute electrons. In theory, we do not expect the electrons to remain ionized throughout, so this calculation would appear academic. However, Eq. (3.47) is relevant for a more general reason. We do expect that at some late time, the electrons were reionized. We expect this because the universe we observe back to redshift $z \sim 6$ appears to be ionized. If the universe was reionized at very late times, much after the z_{decouple} of Eq. (3.47), there would not be a huge change in the CMB anisotropy pattern. However, if the universe was reionized earlier than this redshift, multiple scattering of the photons would dramatically alter the primordial anisotropy pattern set up at $z \sim 1000$. Observations of the most distant quasars (Becker *et al.*, 2001; Fan *et al.*, 2002) suggest that reionization took place at $z \simeq 6$, so the alteration is expected to be slight.

3.4 DARK MATTER

There is strong evidence for nonbaryonic dark matter in the universe, with $\Omega_{\text{dm}} \simeq 0.3$. Perhaps the most plausible candidate for dark matter is a weakly interacting massive particle (WIMP), which was in close contact with the rest of the cosmic plasma at high temperatures, but then experienced *freeze-out* as the temperature dropped below its mass. Freeze-out is the inability of annihilations to keep the particle in equlibrium. Indeed, were it kept in equilibrium indefinitely, its abundance would be suppressed by $e^{-m/T}$: there would be no such particles in the observable universe. The purpose of this section, then, is to solve the Boltzmann equation for such a particle, determining the epoch of freeze-out and its relic abundance. The hope is that, by fixing its relic abundance so that $\Omega_{\rm dm} \simeq 0.3$, we will learn something about the fundamental properties of the particle, such as its mass and cross section. We then might use this knowledge to detect the particles in a laboratory.

In the generic WIMP scenario, two heavy particles X can annihilate producing two light (essentially massless) particles l. The light particles are assumed to be very tightly coupled to the cosmic plasma, so they are in complete equilibrium (chemical as well as kinetic), with $n_l = n_l^{(0)}$. There is then only one unknown, n_X, the abundance of the heavy particle. We can use Eq. (3.9) to solve for this abundance:

$$a^{-3}\frac{d\left(n_X a^3\right)}{dt} = \langle\sigma v\rangle \left\{(n_X^{(0)})^2 - n_X^2\right\}. \tag{3.48}$$

To go further, recall that the temperature typically scales as a^{-1}, so if we multiply and divide the factor of $n_X a^3$ inside the parentheses on the left by T^3, we can remove $(aT)^3$ outside the derivative, leaving $T^3 d(n_X/T^3)/dt$. Let's define then

$$Y \equiv \frac{n_X}{T^3}. \tag{3.49}$$

The differential equation for Y becomes

$$\frac{dY}{dt} = T^3\langle\sigma v\rangle \left\{Y_{\rm EQ}^2 - Y^2\right\}, \tag{3.50}$$

with $Y_{\rm EQ} \equiv n_X^{(0)}/T^3$.

To go further, as in the neutron–proton case, it is convenient to introduce a new time variable,

$$x \equiv m/T \tag{3.51}$$

where m, the mass of the heavy particle, sets a rough scale for the temperature during the region of interest. Very high temperature corresponds to $x \ll 1$, in which case reactions proceed rapidly so $Y \simeq Y_{\rm EQ}$. Since the X particles are relativistic at these epochs, the $m \ll T$ limit of Eq. (3.6) implies that $Y \simeq 1$. For high x, the equilibrium abundance $Y_{\rm EQ}$ becomes exponentially suppressed (e^{-x}). Ultimately, X particles will become so rare because of this suppression that they will not be able to find each other fast enough to maintain the equilibrium abundance. This is the onset of freeze-out. To change from t to x, we need the Jacobian $dx/dt = Hx$. Dark matter production typically occurs deep in the radiation era where the energy density scales as T^4, so $H = H(m)/x^2$. Then the evolution equation becomes

$$\frac{dY}{dx} = -\frac{\lambda}{x^2}\left\{Y^2 - Y_{\rm EQ}^2\right\}, \tag{3.52}$$

where the ratio of the annihilation rate to the expansion rate is parameterized by

$$\lambda \equiv \frac{m^3\langle\sigma v\rangle}{H(m)}. \tag{3.53}$$

In many theories λ is a constant, but in some, the thermally averaged cross section is temperature dependent; this leads to slight numerical changes in the following but unchanged qualitative solutions.

Equation (3.52) is a form of the Riccati equation, for which in general there are no analytic solutions. In this case, though, we can make use of our understanding of the freeze-out process to get an analytic expression for the final freeze-out abundance $Y_\infty \equiv Y(x = \infty)$. Let's review this understanding in the context of Eq. (3.52). The left-hand side is of order Y (for $x \sim 1$) while the right is of order $Y^2\lambda$. We will see that λ is typically quite large, so as long as Y is not too small, the right-hand side must zero itself by setting $Y = Y_{\rm EQ}$. At late times, as $Y_{\rm EQ}$ drops precipitously, the terms on the right-hand side will no longer be much larger than the one on the left. In fact, well after freeze-out, Y will be much larger than $Y_{\rm EQ}$: the X particles will not be able to annihilate fast enough to maintain equilibrium. Thus at late times,

$$\frac{dY}{dx} \simeq -\frac{\lambda Y^2}{x^2} \qquad (x \gg 1). \tag{3.54}$$

Integrate this analytically from the epoch of freeze-out x_f until very late times $x = \infty$ to get

$$\frac{1}{Y_\infty} - \frac{1}{Y_f} = \frac{\lambda}{x_f}. \tag{3.55}$$

Typically Y at freeze-out Y_f is significantly larger than Y_∞, so a simple analytic approximation is

$$Y_\infty \simeq \frac{x_f}{\lambda}. \tag{3.56}$$

This approximation is incomplete, in that it depends on the freeze-out temperature, which we have not determined. Although more precise determinations are possible (Exercise 10), a simple order-of-magnitude estimate for the dark matter problem is $x_f \sim 10$.

Figure 3.5 shows the numerical solution to Eq. (3.52) for several different values of λ. The abundances do track the equilibrium abundances until $m/T \sim 10$, after which they level off to a constant. The rough estimate $Y_\infty \sim 10/\lambda$ is seen to be a reasonable approximation for the relic abundance. Note that particles with larger cross sections (e.g. in the figure, $\lambda = 10^{10}$) freeze out later, and this later freeze-out carries along with it a lower relic abundance. Also note from the inset in Figure 3.5 that the distinction between Bose–Einstein, Fermi–Dirac, and Boltzmann statistics is important only at temperatures above the particle's mass. For temperatures relevant to the freeze-out process, our use of Boltzmann statistics is completely warranted.

There is one more piece of physics needed in order to determine the present-day abundance of these heavy particle relics. After freeze-out, the heavy particle density

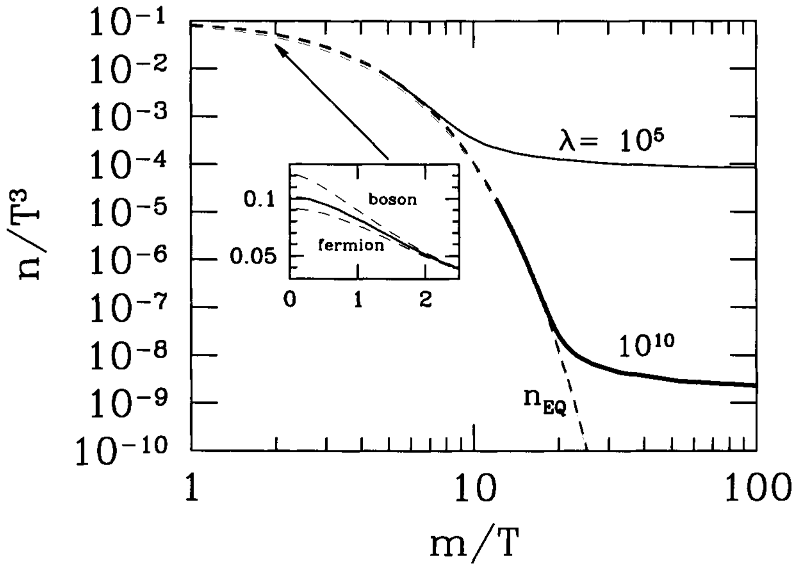

Figure 3.5. Abundance of heavy stable particle as the temperature drops beneath its mass. Dashed line is equilibrium abundance. Two different solid curves show heavy particle abundance for two different values of λ, the ratio of the annihilation rate to the Hubble rate. Inset shows that the difference between quantum statistics and Boltzmann statistics is important only at temperatures larger than the mass.

simply falls off as a^{-3}. So its energy density today is equal to $m(a_1/a_0)^3$ times its number density where a_1 corresponds to a time sufficiently late that Y has reached its aymptotic value, Y_∞. The number density at that time is $Y_\infty T_1^3$, so

$$\rho_X = mY_\infty T_0^3 \left(\frac{a_1 T_1}{a_0 T_0}\right)^3 \simeq \frac{mY_\infty T_0^3}{30}. \tag{3.57}$$

The second equality here is nontrivial. You might expect that aT remains constant through the evolution of the universe, so that the ratio a_1T_1/a_0T_0 would be unity. It is not, for the same reason that the CMB and neutrinos have different temperatures. We saw in Chapter 2 that photons are heated by $e^\pm$ annihilation, while neutrinos which have already decoupled are not. Similarly, as the universe expands, photons are heated by the annihilation of the zoo of particles with masses between 1 MeV and 100 GeV, so T does not fall simply as a^{-1}. You can show in Exercise 11 that as a result $(a_1T_1/a_0T_0)^3 \simeq 1/30$. Finally, to find the fraction of critical density today contributed by X, insert our expression for Y_∞ and divide by $\rho_{\rm cr}$:

$$\begin{aligned}\Omega_X &= \frac{x_f}{\lambda}\frac{mT_0^3}{30\rho_{\rm cr}} \\ &= \frac{H(m)x_fT_0^3}{30m^2\langle\sigma v\rangle\rho_{\rm cr}}.\end{aligned} \tag{3.58}$$

To find the present density of heavy particles, then, we need to compute the Hubble rate when the temperature was equal to the X mass, $H(m)$, for which we need the energy density when the temperature was equal to m. The energy density in the radiation era is given by Eq. (3.26) with g_* a function of temperature. Therefore,

$$\Omega_X = \left[\frac{4\pi^3 G g_*(m)}{45}\right]^{1/2} \frac{x_f T_0^3}{30\langle\sigma v\rangle \rho_{\rm cr}}. \tag{3.59}$$

We see that Ω_X does not explicitly depend on the mass of the X particle.[6] So it is mainly the cross section which determines the relic abundance.

Let's now see what order of magnitude is needed to get dark matter today, i.e., to get $\Omega_X = \Omega_{\rm dm} \simeq 0.3$. At the temperatures of interest for dark matter production, $T \sim 100$ GeV, $g_*(m)$ includes contributions from all the particles in the standard model (three generations of quarks and leptons, photons, gluons, weak bosons, and perhaps even the Higgs boson) and so is of order 100. Normalizing $g_*(m)$ and x_f by their nominal values leads to

$$\Omega_X = 0.3h^{-2}\left(\frac{x_f}{10}\right)\left(\frac{g_*(m)}{100}\right)^{1/2} \frac{10^{-39}{\rm cm}^2}{\langle\sigma v\rangle}. \tag{3.60}$$

The fact that this estimate is of order unity for cross sections of order 10^{-39} cm^2 is taken as a good sign: there are several theories which predict the existence of particles with cross sections this small.

Perhaps the most notable of these theories is supersymmetry, the theory which predicts that every particle has a partner with opposite statistics. For example, the supersymmetric partner of the spin zero Higgs boson is the spin-1/2 *Higgsino* (fermion). Initially it was hoped that the observed fermions could be the partners of the observed bosons, but this hope is not realized in nature. Instead, supersymmetry must be broken and all the supersymmetric partners of the known particles must be so massive that they have not yet been observed even in accelerators: they must have masses greater than 10 to 100 GeV. Which of the supersymmetric partners is the best candidate for the dark matter today? The particle must be neutral since the evidence points to dark matter that is truly dark, i.e., does not interact much with the known particles and especially does not emit photons. The particle must also be stable: if it could decay to lighter particles, then decays would have kept it in equilibrium throughout the early universe and there would be none left today. The first of these criteria restricts the dark matter to be the partner of one of the neutral particles, such as the Higgs or the photon.[7] The second requires the particle to be the lightest (LSP for lightest supersymmetric partner) of these, for any heavier particles could decay into the lightest one plus some ordinary particles.

[6] There is a slight implicit dependence on mass in the freeze-out temperature x_f and in g_*, which is to evaluated when $T = m$.

[7] The other neutral boson in the Standard Model is the Z vector boson. The lightest supersymmetric partner is likely a linear combination of the partners of all of these.

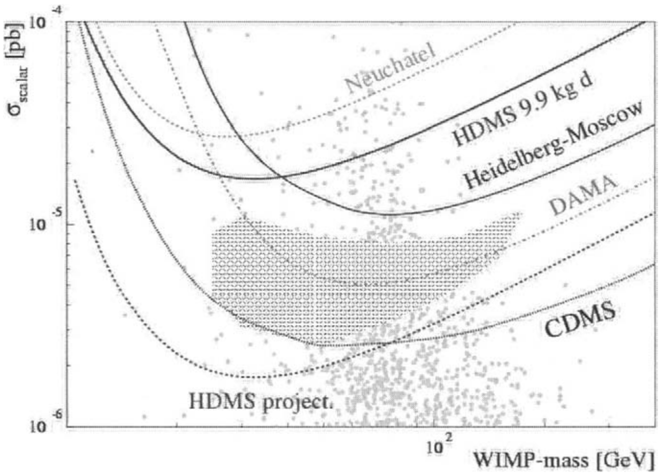

Figure 3.6. Constraints on supersymmetric dark matter (Baudis *et al.*, 2000). Region above the solid curves is excluded, while filled region is reported detection by DAMA (Bernabei *et al.*, 1999). Lowest curve (labled *HDMS project*) is only a projected limit based on one future experiment. The CDMS experiment (Abusaidi *et al.*, 2000) appears to rule out the DAMA detection. Points scattered throughout correspond to different parameter choices in a class of supersymmetric models. Note the limits on the cross section are in units of picobarns (1 picobarn = 10^{-36} cm^2).

Not only would weakly interacting particles such as LSP's annihilate in the early universe, but if they were around today they would scatter off ordinary matter. Although it is difficult to detect these reactions because the rate is so low, a number of experiments have been performed searching for dark matter particles. Figure 3.6 shows the limits on the masses and cross sections of dark matter from these experiments (note that the scattering cross section, while related to the annihilation cross section, is not identically equal to it). Apart from a tantalizing detection from the DAMA experiment, so far we have only upper limits. However, as indicated in the figure, in the coming years the experiments are expected to pierce into the region predicted by supersymmetric theories.

3.5 SUMMARY

The light elements in the universe formed when the temperature of the cosmic plasma was of order 0.1 MeV. Roughly a quarter of the mass of the baryons is in the form of ^{4}He, the remaining in the form of free protons with only trace amounts of deuterium, ^{3}He, and lithium.

These elements remain ionized until the temperature of the universe drops well below the ionization energy of hydrogen. The epoch of recombination — at which time electrons and protons combine to form neutral hydrogen — is at redshift $z \sim 1000$ corresponding to a temperature $T \sim 0.25$ eV. Before recombination, photons and electrons and protons are tightly coupled with one another because of Compton and Coulomb scattering. After this time, photons travel freely through the universe without interacting, so the photons in the CMB we observe today offer an excellent snapshot of the universe at $z \sim 1000$. The importance of this snapshot cannot be overstated.

The details of both nucleosynthesis and recombination are heavily influenced by the fact that the reactions involved eventually become too slow to keep up with the expansion rate. This feature may also be responsible for the production of dark matter in the universe. We explored the popular scenario wherein a massive, neutral stable particle stops annihilating when the temperature drops significantly beneath its mass. The present-day abundance of such a particle can be deterimined in terms of its annihilation cross section, as in Eq. (3.60). Larger cross sections correspond to more efficient annhilation and therefore a lower abundance today. Roughly cross sections of order 10^{-40} are needed to get the dark matter abundance observed today. Such cross sections and the requisite stable, neutral particles emerge fairly natrually in extensions of the Standard Model of particle physics, such as supersymmetry, and may well be tested by accelerator experiments in the near future.

SUGGESTED READING

The Early Universe (Kolb and Turner) contains especially clear treatments of the heavy particle freeze-out problem and Big Bang nucleosynthesis, also based on the Boltzmann equation. *Kinetic Theory in an Expanding Universe* (Bernstein) offers some semianalytic solutions to these problems, as well as the requisite Boltzmann-ology.

Work on nucleosynthesis in the early universe dates back to Gamow and collaborators, summarized in Alpher, Follin, and Herman (1953). The first post-CMB papers were Peebles (1966) and Wagoner, Fowler, and Hoyle (1967); these got the basics right: 25% helium and roughly the correct amount of deuterium. Yang *et al.* (1984) helped many people of my generation understand that the baryonic density could be constrained with observations of the light elements. A nice review article on nucleosynthesis is Olive, Steigman, and Walker (2000).

As I tried to indicate in the text, the process of recombination is very rich; it involves some subtle physics. The original paper which worked through all the details was by Peebles (1968). Ma and Bertschinger (1996) however managed to describe the physics succintly in just one page in their Section 5.8. Seager, Sasselov, and Scott (1999) have presented a more accurate treatment (although, as they emphasize, Peebles' more intuitive work holds up remarkably well), including many small effects previously neglected.

Jungman, Kamionkowski, and Griest (1996) is a comprehensive review of all aspects of supersymmetric dark matter. Many papers have explored limits on supersymmetric dark matter candidates from cosmology and accelerators. To mention just several: Roszkowski (1991) showed that the Higgsino is likely *not* the lightest supersymmetric partner; Nath and Arnowitt (1992) and Kane *et al.* (1994) showed that the *bino*, the partner of the initial gauge eigenstate B, is the most likely — both from the point of view of physics and cosmology — LSP; Ellis *et al.* (1997) combined accelerator constraints with those from cosmology to place a lower limit on the mass of the LSP; Edsjo and Gondolo (1997) included some subtle effects in the relic abundance calculation which affects the limits if the LSP is composed primarily of the partner of the Higgs boson; and Bottino *et al.* (2001) explored the consequences if the signal seen in the DAMA signal is due to dark matter particles.

EXERCISES

Exercise 1. Compute the equilibrium number density (i.e., zero chemical potential) of a species with mass m and degeneracy $g = 2$ in the limits of large and small m/T. Take these limits for all three types of statistics: Boltzmann, Bose–Einstein, and Fermi–Dirac. You will find Eqs. (C.26) and (C.27) helpful for the high-T Bose–Einstein and Fermi–Dirac limits.

Exercise 2. In the text, we treated $e^{\pm}$ as relevant to the energy density at temperatures above m_e, but irrelevant afterwards (Eq. (3.30) and the discussion leading up to it). Track the $e^{\pm}$ density through annihilation assuming $n_{e^{\pm}} = n_{e^{\pm}}^{(0)}$. This equality holds during the BBN epoch because electromagnetic interactions (e.g., $e^+ + e^- \leftrightarrow \gamma + \gamma$) are so strong. When does the density fall to 1% of the photon energy density? If $\eta_b \simeq 6 \times 10^{-10}$, at what temperature do you expect n_{e^-} to depart from $n_{e^-}^{(0)}$?

Exercise 3. Compute the rate for neutron-to-proton conversion, λ_{np}. Show that it is equal to Eq. (3.29). There are two processes which contribute to λ_{np}: $n + \nu_e \to p + e^-$ and $n + e^+ \to p + \bar{\nu}_e$. Assume that all particles can be described by Boltzmann statistics and neglect the mass of the electron. With these approximations the two rates are identical.
(a) Use Eq. (3.8) to write down the rate for $n + \nu_e \to p + e^-$. Perform the integrals over heavy particle momenta to get

$$\lambda_{np} = n_{\nu_e}^{(0)} \langle \sigma v \rangle = \frac{\pi}{4m^2} \int \frac{d^3 p_\nu}{(2\pi)^3 2p_\nu} \, e^{-p_\nu/T} \times \int \frac{d^3 p_e}{(2\pi)^3 2p_e} \delta(Q + p_\nu - p_e) \left|\mathcal{M}\right|^2 . \tag{3.61}$$

(b) The amplitude squared is equal to $|\mathcal{M}|^2 = 32 G_F^2 (1 + 3g_A^2) m_p^2 p_\nu p_e$, where g_A is the axial-vector coupling of the nucleon. The present best measurement of g_A is via the neutron lifetime, $\tau_n = \lambda_0 G_F^2 (1 + 3g_A^2) m_e^5 / (2\pi^3)$, where the phase space integral

$$\lambda_0 \equiv \int_1^{Q/m_e} dx x (x - Q/m_e)^2 (x^2 - 1)^{1/2} = 1.636. \tag{3.62}$$

Carry out the integrals in Eq. (3.61) to get the rate, λ_{np} in terms of τ_n. Don't forget to multiply by 2 for the two different reactions.

Exercise 4. Solve the rate equation (3.27) numerically to determine the neutron fraction as a function of temperature. Ignore decays. There are (at least) two ways to perform this computation. The first is to treat it as a simple ordinary differential equation and solve numerically. The second is to proceed analytically and reduce the problem to an evaluation of a single numerical integral. This second method, which I'll lead you through here, is based on a numerical coincidence noted by Bernstein, Brown, and Feinberg (1988).
(a) Using standard differential equation techniques, show that a formal solution to Eq. (3.27) is

$$X_n(x) = \int_{x_i}^{x} dx' \, \frac{\lambda_{np}(x') e^{-x'}}{x' H(x')} \, e^{\mu(x') - \mu(x)} \tag{3.63}$$

where x_i is some initial, very high temperature, and

$$\mu(x) \equiv \int_{x_i}^{x} \frac{dx'}{x' H(x')} \, \lambda_{np}(x') \left[1 + e^{-x'}\right]. \tag{3.64}$$

(b) Use Eqs. (3.29) and (3.26) to compute the integrating factor μ analytically. Show that it is equal to

$$\mu = -\frac{255}{\tau_n Q}\left[\frac{4\pi^3 G Q^2 g_*}{45}\right]^{-1/2} \times \left[\left(\frac{4}{x^3}+\frac{3}{x^2}+\frac{1}{x}\right)+\left(\frac{4}{x^3}+\frac{1}{x^2}\right)e^{-x}\right]\Bigg|_{x_i}^{x}. \tag{3.65}$$

The simple form for μ is the result of numerical coincidence alone.
(c) With the results of part **(b)**, do the single numerical integral in **(a)** numerically. Compare the asymptotic result at $x=\infty$ with the result in the text, $X_n(x=\infty)=0.15$.

Exercise 5. Integrate the Friedmann equation (1.2) to verify the time–temperature relation in Eq. (3.30) in the epoch after $e^{\pm}$ annihilation, but before matter domination.

Exercise 6. Determine η_b in terms of $\Omega_b h^2$. Show that it is given by Eq. (3.11).

Exercise 7. An important parameter for CMB anisotropies is the sound speed at decoupling. This is determined by the ratio of baryons to photons.
(a) Find

$$R \equiv \frac{3\rho_b}{4\rho_\gamma}$$

as a function of a. Evaluate it at decoupling. Your answer should depend on $\Omega_b h^2$.
(b) We will see in Chapter 8 that the sound speed of the combined photon/baryon fluid is

$$c_s = \sqrt{\frac{1}{3(1+R)}}. \tag{3.66}$$

Use your answer from (a) to plot the sound speed at decoupling as a function of $\Omega_b h^2$.

Exercise 8. Solve for the evolution of the free electron fraction. Do not compare your results with Figure 3.4 until you finish part **(d)**. Throughout, take parameters $\Omega_m = 1, \Omega_b = 0.06, h = 0.5$.
(a) Use as an evolution variable $x \equiv \epsilon_0/T$ instead of time in Eq. (3.39). Rewrite the equation in terms of x and the Hubble rate at $T=\epsilon_0$.
(b) Using the methods of Section 3.4, find the final freeze-out abundance of the free electron fraction, $X_e(x=\infty)$.
(c) Numerically integrate the equation from **(a)** from $x=1$ down to $x=1000$. What is the final frozen-out X_e?
(d) Peebles (1968) argued that even captures to excited states would not be important except for the fraction of times that the $n=2$ state decays into two photons or

expansion redshifts the Lyman alpha photon so that it cannot pump up a ground-state atom. Quantitatively, he multiplied the right-hand side of Eq. (3.39) by the correction factor,

$$C = \frac{\Lambda_\alpha + \Lambda_{2\gamma}}{\Lambda_\alpha + \Lambda_{2\gamma} + \beta^{(2)}} \tag{3.67}$$

where the two-photon decay rate is $\Lambda_{2\gamma} = 8.227\text{sec}^{-1}$; Lyman alpha production is $\beta^{(2)} = \beta e^{3\epsilon_0/4T}$; and

$$\Lambda_\alpha = \frac{H(3\epsilon_0)^3}{(8\pi)^2}. \tag{3.68}$$

Do this and show how it changes your final answer. Now compare the freeze-out abundance with the result of **(c)** and the evolution with Figure 3.4.

Exercise 9. Find the redshift of decoupling as a function of Ω_b. If you do not have the evolution code of Exercise 8, use the Saha equation to determine X_e.

Exercise 10. Find an approximation to the freeze-out temperature of annihilating heavy particles by setting x_f such that $n^{(0)}(x_f)\langle\sigma v\rangle = H(x_f)$.

Exercise 11. Typically the temperature of the cosmic plasma cools as a^{-1} with the expansion. However, when particles annihilate, they deposit energy into the plasma, thereby slowing the cooling. (Scherrer and Turner, 1986, showed that the annihilations do not actually *heat* the universe: T never increases, it simply decreases more slowly than a^{-1}.) Use the fact that the entropy density (Eq. (2.66)) scales as a^{-3} to compute the ratio of $(aT)^3$ at $T = 10$ GeV (roughly the time when WIMPs decouple) to its present value today.

Exercise 12. Suppose that there were no baryon asymmetry so that the number density of baryons exactly equaled that of anti-baryons. Determine the final relic density of (baryons+anti-baryons). At what temperature is this asymptotic value reached?

Exercise 13. There is a fundamental limitation on the annihilation cross section of a particle with mass m. Because of unitarity, $\langle\sigma v\rangle$ must be less than or equal to $1/m^2$, give or take a factor of order unity. Determine Ω_X for a particle which saturates this bound, i.e., for a particle with $\langle\sigma v\rangle = 1/m^2$. For what value of m is Ω_X equal to 1? (Keep x_f and g_* equal to the nominal values given in Eq. (3.60).) Note that if m is greater than this critical value, $\Omega_X > 1$, which is ruled out. This is a strong argument against stable particles (and therefore dark matter candidates) with masses above this critical value.